

Abstract

Deep Learning (DL) models make it possible to automate bio-signal analysis. When dealing with Electroencephalography (EEG), a widely used tool to monitor brain activity, this translates into accurate detection of neurological events and scalable systems suitable for real-time applications. Within this framework, the present PhD work aims to design and optimize DL architectures tailored to patient-specific seizure detection, specifically addressing two main aspects of the problem. On the one hand, the proper management of the spatial complexity of EEG signals, which requires extracting meaningful information from multichannel data distributed across different brain regions; on the other hand, the reduction of the computational complexity of DL algorithms, in order to make their use feasible in a wearable context. This latter aspect is highly relevant, as it aligns with the growing interest in *personalized medicine* and remote healthcare, where the analysis of brain activity through continuous monitoring can help clinicians track the progression of neurological conditions more accurately.

The research activity is divided into two lines of investigation and is evaluated using two public datasets, i.e., the Children’s Hospital Boston-Massachusetts Institute of Technology (CHB-MIT) database, widely adopted in the literature, and the Siena Scalp EEG dataset. The first study investigates how to effectively reduce the spatial dimensionality of EEG signals and exploit the resulting reduced input in a lightweight DL model trained on raw EEG data for supervised detection of epileptic seizures. To this end, we propose a data-driven selection algorithm that identifies the most informative scalp regions involved in seizure detection and a computationally efficient Convolutional Neural Network (CNN) that processes data from only two

EEG channels in a patient-specific setting, where both training and testing are performed on data from the same subject.

The two selected channels are representative of the two main phases of epileptic activity, i.e. ictal activity corresponding to labeled seizure events and no-seizure activity corresponding to the absence of ictal activity. The results show high performance compared to the state of the art. On the CHB-MIT dataset, the model achieves a balanced accuracy of 83% averaged across patients and an overall accuracy of 99%, correctly detecting 152 of 181 epileptic events and outperforming the same architecture applied to the four-channel fixed configuration commonly used in wearable EEG applications. Even on the Siena Scalp, the model achieves a balanced accuracy of 83% (with 31 of 36 epileptic events) compared to 78% obtained in the case of the fixed four-channel configuration.

Based on these results, the proposed method is further evaluated in a single-channel framework, in order to show the impact of choosing informative versus non-informative locations, and across various implementation strategies that address the patient-independent scenario, assessing how the model could generalize across patients. In the first evaluation, each of the two selected channels is tested individually, together with a non-informative EEG channel not highlighted by the channel selection procedure. The analysis demonstrates that using a non-informative channel leads to a 22% decrease in balanced accuracy on both datasets, confirming that not all channels carry the information required for seizure detection. Moreover, although the dual-channel configuration provides higher performance by combining complementary information from both ictal and no-seizure representative channels, the single-channel setup still offers a compelling trade-off between performance and simplicity. On the CHB-MIT dataset, the ictal-representative single channel achieves a balanced accuracy of 77% and detects 136 out of 181 epileptic events. Under the same training conditions, the no-seizure single channel reaches a balanced accuracy of 69% and detects 98 out of 173 events. Similar results are replicated on the Siena dataset, with a balanced accuracy of 81% for the ictal-representative channel and 77% for the no-seizure channel.

In the second evaluation, we investigate different implementation strategies for the patient-independent scenario. Specifically, we compare a fully personalized configuration, in which both the channel subset and the lightweight DL model are customized using only data acquired from a single patient; a fully generalized configuration, where both channel selection and model training are derived from a multi-patient dataset, and two hybrid approaches. In the first hybrid strategy, channel selection is performed on a large multi-patient dataset, while the model is trained on data from a specific patient. In the second hybrid strategy, the model is trained on a multi-patient dataset, whereas EEG channels are selected based on data from a specific patient. This analysis suggests that, although the fully customized approach achieves the best performance, an intermediate strategy, which consists of fixing the electrode layout while customizing the detection model, offers an effective balance between simplicity and accuracy, with a small decrease in balanced accuracy (80% vs. 83% on CHB-MIT and 73% vs. 83% on Siena Scalp).

The second study analyzes how to best exploit the spatial relationships between EEG electrodes in scenarios where the goal is to use information from all channels, rather than selecting only a few. Traditional CNNs, which assume input organized on a regular grid, are not ideal for this purpose, as EEG electrodes are arranged on a curved and irregular surface (the scalp) and do not follow a structured two-dimensional grid layout. For this reason, the use of Graph Neural Networks (GNNs) represents a more natural and effective approach to modeling EEG data, in which each channel is described as a node in the graph and the edges encode the functional connection (FC) between the electrodes. To identify the FC measure that best highlights differences between seizure phases while keeping the graph as sparse as possible, thereby enabling a lighter graph-based network fed by more selective and informative inputs, a data-driven statistical comparison of two widely used measures, i.e., phase locking value (PLV) and Pearson correlation coefficient (PCC), is carried out. As a result, PLV emerges as the most effective FC measure, and the developed architecture, which combines Graph Convolutional Networks (GCN)

for spatial feature extraction and a Long Short-Term Memory (LSTM) module for representing temporal dynamics, achieves a balanced accuracy of 97.7%, comparable to the state of the art while using more than 50% fewer parameters. Using PCC, performance is reduced to a balanced accuracy of 93.13% with the same proposed GNN model. Similar results are observed on the Siena Scalp-EEG dataset, where a balanced accuracy of 96.32% is obtained using PLV and 93.68% using PCC.

In conclusion, even if the main result of the entire PhD work is based on the implementation of lightweight DL frameworks for the extraction of meaningful descriptors of brain functioning, particular attention is paid to result visualization, which is essential for the clinical adoption of the models and for effective collaboration between human experts and Artificial Intelligence (AI) systems during the diagnostic process. Indeed, one of the main limitations to the adoption of AI in healthcare is the "black box" nature of state-of-the-art models. In this perspective, both proposed approaches generate intermediate outputs. The first returns EEG topographic maps, which allow the spatial distribution of the most relevant activity in distinguishing between the phases of the epileptic seizure to be visualized. The second, through the use of GNNExplainer, an interpretability tool that identifies the most influential nodes for model predictions, obtains relevance scores at the node level, maintaining a clear focus on the regions of the brain involved. These intermediate readings show strong consistency between the two approaches: graph-based explanation tools highlight regions of interest in accordance with the channels selected in the first study, reinforcing the reliability and cross-validation of the methodological paths adopted.

Abstract in italiano

I modelli di Deep Learning (DL) consentono di automatizzare l'analisi dei biosegnali. Nel caso dell'elettroencefalografia (EEG), uno strumento ampiamente utilizzato per monitorare l'attività cerebrale, ciò si traduce in un rilevamento accurato degli eventi neurologici e in sistemi scalabili adatti ad applicazioni in tempo reale. In questo contesto si inserisce il presente lavoro di dottorato che mira a progettare e ottimizzare architetture DL su misura per il rilevamento delle crisi epilettiche specifiche per ciascun paziente, affrontando in particolare due aspetti principali del problema. Da un lato, la corretta gestione della complessità spaziale dei segnali EEG, che richiede l'estrazione di informazioni significative da dati multicanale distribuiti in diverse regioni del cervello; dall'altro, la riduzione della complessità computazionale degli algoritmi DL, al fine di renderne possibile l'uso in un contesto indossabile. Quest'ultimo aspetto è molto rilevante, in quanto in linea con il crescente interesse per la *medicina personalizzata* ed il monitoraggio a distanza, dove l'analisi dell'attività cerebrale attraverso il monitoraggio continuo può aiutare i medici a seguire con maggiore precisione la progressione delle condizioni neurologiche.

L'attività di ricerca è suddivisa in due linee di indagine ed è valutata utilizzando due set di dati pubblici, ovvero il database del Children's Hospital Boston-Massachusetts Institute of Technology (CHB-MIT), ampiamente adottato in letteratura, ed il dataset Siena Scalp. Il primo studio indaga come ridurre efficacemente la dimensionalità spaziale dei segnali EEG e sfruttare l'input ridotto risultante in un modello DL leggero addestrato su dati EEG grezzi per il rilevamento supervisionato delle crisi epilettiche. A tal fine, si propone un algoritmo di selezione basato sui dati che identifica le regioni dello scal-

po più informative coinvolte nel rilevamento delle crisi e una rete neurale convoluzionale (CNN), efficiente dal punto di vista computazionale, che elabora i dati provenienti da soli due canali EEG in un contesto specifico per ciascun paziente, in cui sia l'addestramento che il test vengono eseguiti su dati provenienti dallo stesso soggetto. I due canali selezionati sono rappresentativi delle due fasi principali dell'attività epilettica, ovvero la fase ictale corrispondente agli eventi epilettici etichettati e la fase no-seizure corrispondente all'assenza di attività epilettica. I risultati mostrano prestazioni elevate rispetto allo stato dell'arte. Sul set di dati CHB-MIT, il modello raggiunge un'accuratezza bilanciata dell'83% in media tra i pazienti e un'accuratezza complessiva del 99%, rilevando correttamente 152 dei 181 eventi epilettici e superando la stessa architettura applicata alla configurazione fissa a quattro canali proposta in letteratura per le applicazioni EEG indossabili. Anche sul Siena Scalp, il modello raggiunge un'accuratezza bilanciata dell'83% (con 31 eventi epilettici su 36 rilevati) rispetto al 78% ottenuto nel caso della configurazione fissa a quattro canali.

Sulla base di questi risultati, il metodo proposto è stato ulteriormente valutato in un framework a canale singolo, al fine di mostrare l'impatto della scelta di posizioni informative rispetto a quelle non informative, e attraverso varie strategie di implementazione che affrontano lo scenario indipendente dal paziente, valutando come il modello potrebbe generalizzare tra i pazienti. Nella prima valutazione, ciascuno dei due canali selezionati viene testato individualmente, insieme a un canale EEG non informativo non emerso dalla procedura di selezione dei canali. L'analisi dimostra che l'utilizzo di un canale non informativo porta a una diminuzione del 22% dell'accuratezza bilanciata su entrambi i dataset, confermando che non tutti i canali trasportano l'informazione necessaria alla rilevazione dell'attacco. Inoltre, sebbene la configurazione a doppio canale fornisca prestazioni più elevate combinando informazioni complementari provenienti da canali rappresentativi sia ictali che no-seizure, la configurazione a canale singolo offre comunque un compromesso interessante tra prestazioni e semplicità. Sul dataset CHB-MIT, il canale singolo rappresentativo della fase ictale raggiunge un'accuratezza bilanciata del 77% e rileva

136 dei 181 eventi epilettici. Nelle stesse condizioni di addestramento, il canale singolo non epilettico raggiunge un'accuratezza bilanciata del 69% e rileva 98 dei 173 eventi. Risultati simili sono stati replicati sul dataset Siena Scalp, con un'accuratezza bilanciata del 81% per il canale rappresentativo della fase ictale e del 77% per il canale no-seizure.

Nella seconda valutazione, sono esaminate diverse strategie di implementazione per lo scenario indipendente dal paziente. Nello specifico, si confronta una configurazione completamente personalizzata, in cui sia la selezione dei canali che il modello DL leggero sono personalizzati utilizzando solo i dati acquisiti da un singolo paziente; una configurazione completamente generalizzata, in cui sia la selezione dei canali che l'addestramento del modello derivano da un set di dati multi-paziente; e due approcci ibridi. Nella prima strategia ibrida, la selezione dei canali viene eseguita su un ampio set di dati multi-paziente, mentre il modello viene addestrato sui dati di un paziente specifico. Nella seconda strategia ibrida, il modello viene addestrato su un set di dati multi-paziente, mentre i canali EEG vengono selezionati per un paziente specifico. Questa analisi suggerisce che, sebbene l'approccio completamente personalizzato raggiunga le migliori prestazioni, una strategia intermedia, che consiste nel fissare la disposizione degli elettrodi personalizzando il modello di rilevamento, offre un equilibrio efficace tra semplicità e accuratezza, con una leggera diminuzione dell'accuratezza bilanciata (80% contro 83% sul CHB-MIT e 73% contro 83% sul Siena Scalp).

Il secondo studio analizza come sfruttare al meglio le relazioni spaziali tra gli elettrodi EEG in scenari in cui l'obiettivo è quello di utilizzare le informazioni provenienti da tutti i canali, piuttosto che selezionarne solo alcuni. Le CNN tradizionali, che presuppongono un input organizzato su una griglia regolare, non sono ideali per questo scopo, poiché gli elettrodi EEG sono disposti su una superficie curva ed irregolare (lo scalp) e non seguono una disposizione strutturata su una griglia bidimensionale. Per questo motivo, l'uso delle reti neurali grafiche (GNN) si presta come un approccio più naturale ed efficace per modellare i dati EEG, in cui ogni canale è descritto come un nodo nel grafo ed i bordi codificano la connessione funzionale

(FC) tra gli elettrodi. Per identificare la misura FC che meglio evidenzia le differenze tra le fasi delle crisi epilettiche mantenendo il grafo il più sparso possibile, consentendo così una rete basata su grafici più leggera alimentata da input più selettivi ed informativi, si effettua un confronto statistico basato sui dati di due misure ampiamente utilizzate, ovvero il valore di Phase locking value (PLV) ed il Coefficiente di Correlazione di Pearson (PCC). Per quanto riguarda i risultati, il PLV emerge come la misura FC più efficace e l'architettura sviluppata, che combina reti convoluzionali a grafo (GCN) per l'estrazione di caratteristiche spaziali ed un modulo di Long Short-Term Memory per rappresentare le dinamiche temporali, raggiunge un'accuratezza bilanciata del 97.7%, paragonabile allo stato dell'arte, utilizzando oltre il 50% in meno di parametri. Utilizzando il PCC, le prestazioni scendono ad un'accuratezza bilanciata del 93.13% con lo stesso modello GNN. Risultati simili si osservano sul dataset Siena Scalp, dove si ottiene un'accuratezza bilanciata del 96.32% con PLV e del 93.68% con PCC.

In conclusione, sebbene il risultato principale del lavoro di dottorato si basi sull'implementazione di framework DL leggeri per l'estrazione di descrittori significativi del funzionamento del cervello, particolare attenzione è rivolta alla visualizzazione dei risultati, essenziale per l'adozione clinica dei modelli e per un'efficace collaborazione tra esperti e sistemi di Intelligenza Artificiale (AI) nel processo diagnostico. Infatti, uno dei principali limiti all'adozione dell'IA in ambito sanitario è la natura "scatola nera" dei modelli. In questa prospettiva, entrambi gli approcci proposti generano output intermedi. Il primo restituisce mappe topografiche che consentono di visualizzare la distribuzione spaziale dell'attività più rilevante nel distinguere le fasi della crisi. Il secondo, attraverso l'uso di GNNE explainer, uno strumento di interpretabilità che identifica i nodi più influenti per le previsioni del modello, ottiene punteggi di rilevanza a livello di nodo, mantenendo una chiara attenzione sulle aree cerebrali coinvolte. Queste letture intermedie mostrano una forte coerenza tra i due approcci: gli strumenti di spiegazione basati su grafo evidenziano regioni di interesse in accordo con i canali evidenziati nel primo studio, rafforzando l'affidabilità e la validazione incrociata dei percorsi metodologici adottati.