



Imagerie de sources électrophysiologiques par apprentissage supervisé

Sarah Reynaud, Adrien Merlini, Douraied Ben Salem, François Rousseau

► To cite this version:

Sarah Reynaud, Adrien Merlini, Douraied Ben Salem, François Rousseau. Imagerie de sources électrophysiologiques par apprentissage supervisé. ORASIS 2023 : 19ième édition du colloque d'ORASIS, journées francophones des jeunes chercheurs en vision par ordinateur, Laboratoire LIS, UMR 7020, May 2023, Carqueiranne, France. hal-04219527

HAL Id: hal-04219527

<https://hal.science/hal-04219527v1>

Submitted on 27 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Imagerie de sources électrophysiologiques par apprentissage supervisé (*Electrophysiological source imaging by supervised learning*)

S. Reynaud¹

A. Merlini²

D. Ben Salem³

F. Rousseau¹

¹ IMT Atlantique, LaTIM U1101 INSERM, Brest

² IMT Atlantique, Lab-STICC UMR CNRS 6285, Brest

³ CHU, UBO, LaTIM U1101 INSERM, Brest

sarah.reynaud@imt-atlantique.fr

Résumé

L'estimation de l'activité à l'origine des mesures d'électroencéphalographie (EEG) est un problème inverse mal posé. Les méthodes d'inversion nécessitent la définition d'un a priori sur la solution. Les méthodes d'apprentissage automatique supervisé permettent d'estimer la fonction inverse directement à partir d'un ensemble de données. Les travaux présentés dans cet article reposent sur l'utilisation de simulations réalistes de données de sources étendues, unique ou multiples, s'appuyant sur l'outil de simulation SEREEGA [13]. Plus précisément, nous étudions deux modèles (un réseau convolutif à une dimension et un réseau récurrent long short term memory (LSTM)), comparés aux méthodes variationnelles minimum norm energy [12] (MNE) et standardized low resolution brain electromagnetic tomography [14] (sLORETA). Les résultats sur données simulées montrent l'apport potentiel des méthodes par apprentissage supervisé au problème de l'imagerie de source électrophysiologique spatio-temporelle, mais doivent être complétés par de plus nombreuses expériences pour évaluer la capacité du réseau à généraliser sur des données réelles.

Mots Clef

EEG, apprentissage profond, imagerie cérébrale, problème inverse

Abstract

Estimating the activity at the origin of electroencephalography (EEG) measurements is an ill-posed inverse problem. Many methods allow to solve this problem by imposing different prior on the solution. Machine learning could allow to learn the inverse function directly from the data and thus make the complex choice of one of the multiple solutions of the inverse problem more reliable. This work is based on simulations of electrophysiological data containing single or multiple extended sources, using the SEREEGA simulation toolbox [13], to train a one-dimensional convolutional network and to compare the results of this learning approach to those obtained by a re-

current long short term memory (LSTM) network from the literature and by minimum norm energy [12] (MNE) and standardized low resolution brain electromagnetic tomography [14] (sLORETA) methods. These results on simulated data are encouraging about the potential contribution of learning-based methods to the problem of spatiotemporal electrophysiological source imaging. More experiments still need to be done in order to also evaluate the network's ability to generalize to real data.

Keywords

EEG, deep learning, neuro-imaging, inverse problem

1 Introduction

L'électroencéphalographie (EEG) permet de mesurer de manière non-invasive et avec une résolution temporelle de l'ordre de la milliseconde l'activité électrique du cerveau, à l'aide d'électrodes placées sur le scalp. Le modèle direct en EEG permet de faire le lien entre une activité cérébrale connue, modélisée par une distribution de courants, et les mesures EEG par le biais de la matrice de *leadfield*. L'imagerie de source électrophysiologique correspond à l'estimation de la distribution de courants à partir des données EEG. C'est un problème inverse mal posé, pour le résoudre il est nécessaire d'ajouter un a priori sur l'activité à estimer [4]. Le choix de l'a priori sur les données influence énormément la solution finale et peut-être sujet à de nombreux questionnements. Certaines activités cérébrales comme celles présentant des sources étendues, ou multiples restent des défis pour l'imagerie de source électrophysiologique (*electrophysiological source imaging*, ESI). Dans ce contexte, les approches par apprentissage automatique et notamment l'utilisation de réseaux de neurones pourraient permettre d'apprendre la fonction inverse à partir de données réalistes.

Les premiers travaux utilisant de l'apprentissage automatique pour l'ESI se sont concentrés sur l'estimation de la position de quelques dipôles équivalents à partir d'un ou plusieurs instants du signal EEG en utilisant des perceptrons multicouches (MLP) [19] ou des réseaux récurrents

long short term memory (LSTM) [10]. Une autre famille d’approches consiste à estimer l’amplitude de sources placées sur le maillage du cerveau qui constitue une grille prédéfinie dans l’espace. Deux grandes catégories de méthodes existent alors, la première se concentre principalement sur la distribution spatiale des sources en estimant la distribution des sources pour un seul instant [9], [18], [17], tandis que la seconde vise à estimer la distribution spatio-temporelle c’est-à-dire à reconstruire la distribution des sources pour plusieurs instants consécutifs. Dans ce cas, l’entrée type d’un réseau est une fenêtre de T échantillons temporels de données EEG $\mathbf{M} \in \mathbb{R}^{N_e \times T}$, pour N_e électrodes et la sortie est la fenêtre de l’activité de sources correspondante $\mathbf{D} \in \mathbb{R}^{N_s \times T}$, pour N_s sources. Ce travail se concentre sur la dernière catégorie d’approches : l’imagerie de source électrophysiologique, spatio-temporelle.

L’un des aspects importants dans la conception d’un algorithme d’apprentissage profond est de représenter au mieux les données EEG afin d’en extraire les informations spatiales et temporelles. En considérant les données EEG comme une série temporelle multivariée, un réseau LSTM bidirectionnel peut être utilisé pour estimer l’activité de sources étendues dans l’espace [8]. Un auto-encodeur utilisant des convolutions sur les différentes dimensions de manière séquentielle a aussi été proposé [11] ainsi qu’un réseau basé sur un module spatial (ResNet) et un module temporel (LSTM) pour extraire les informations spatio-temporelles des données [16].

Dans ce travail, des données simulées (décrites dans la Section 2.1) sont utilisées pour entraîner des réseaux de neurones pour résoudre le problème inverse en EEG. Les réseaux étudiés (CNN-1D et LSTM) ainsi que des méthodes variationnelles (MNE, sLORETA) sont présentées dans les Sections 2.2 et 2.3 respectivement. Les métriques présentées en Section 2.4 visent à évaluer les différentes solutions sur les aspects spatiaux, temporels et d’amplitude. Ces métriques sont utilisées pour comparer les performances des différentes méthodes dans la Section 3.

2 Méthodes

2.1 Simulation des données

La mesure directe de l’activité électrophysiologique du cerveau est un processus invasif, il n’existe généralement pas de données « vérité-terrain » de l’activité cérébrale au niveau des sources correspondant aux enregistrements EEG. Afin d’entraîner des réseaux de neurones de manière supervisée pour la résolution du problème inverse mais aussi afin d’évaluer les résultats des différentes méthodes d’ESI, des données avec une vérité-terrain sont nécessaires.

L’activité mesurée par EEG à la surface du crâne correspond à l’activité de groupes de neurones actifs de manière synchrone, qui peut-être modélisée par un dipôle équivalent de courant [6]. Le lien entre l’activité de chaque dipôle équivalent, ou source, à une position donnée dans le cerveau et les mesures de potentiel réalisées à l’extérieur du crâne est représenté par la matrice de *leadfield* selon le

modèle direct $\tilde{\mathbf{M}} = \mathbf{GD}$, où $\tilde{\mathbf{M}}$ est la distribution de potentiels non bruitée à des positions données (électrodes) sur le scalp, $\mathbf{D}$ est la distribution de source et $\mathbf{G}$ est la matrice de *leadfield*. La matrice de *leadfield* rassemble des informations anatomiques et biophysiques sur le cerveau et est calculée en résolvant le problème direct, c’est-à-dire en résolvant l’équation de Poisson sur le domaine de la tête.

Ainsi pour générer des paires synthétiques de données EEG et de données d’activité cérébrale, une distribution de sources est générée puis projetée sur les électrodes à l’aide du modèle direct. La simulation d’une distribution de sources peut être divisée en trois parties : la définition du modèle de tête qui regroupe les informations sur la géométrie et la biophysique du cerveau, la définition du motif spatial et du motif temporel de l’activité de la source. Dans ce travail, nous avons utilisé mne-python [3] pour créer le modèle de tête et SEREEGA [13] pour simuler les sources et les activités EEG résultantes. Deux ensembles de données ont été créés : un ensemble contenant des données pour lesquelles une source étendue unique est active et un ensemble avec des données pour lesquelles plusieurs sources étendues sont actives. Le même modèle de tête est utilisé pour les deux ensembles de données.

Modèle de tête. Le modèle de tête utilisé est défini à partir du *template* IRM fsaverage (mne-python, FreeSurfer). La matrice de *leadfield* est calculée à l’aide d’une méthode aux éléments de frontière (*Boundary Element Method*, BEM) avec trois couches : cerveau (cortex), crâne et peau avec des valeurs de conductivité respectives de 0.3 S/m^2 , 0.006 S/m^2 et 0.3 S/m^2 . Le maillage cortical est sous-échantillonné pour créer un espace source avec $N_s = 1284$ sources corticales ayant une orientation fixe, normales à la surface. Le montage d’électrodes est un montage standardisé de $N_e = 90$ électrodes (*standard_1020*).

Source unique étendue. Une source étendue est constituée d’une zone de dipôles voisins qui sont actifs de manière similaire. Dans cette situation l’un des défis de l’ESI est d’estimer non seulement le centre de la source mais aussi son extension correcte.

Motif spatial Le motif spatial fait référence à l’ensemble des sources actives. Une source étendue est une région de sources actives voisines créée en sélectionnant une source noyau et un ordre d’extension, et en ajoutant récursivement les sources voisines de la source noyau à la région comme expliqué dans [5]. Plus formellement : notons $\mathcal{S}$ l’ensemble des sources de l’espace source. Soit $s \in \mathcal{S}$ une source noyau sélectionnée de manière aléatoire. Le voisinage de premier ordre de s , noté $\mathcal{V}^{(1)}(s)$, est l’ensemble des sources partageant une arête avec s dans le maillage cortical sous-échantillonné. Pour créer une région d’ordre o on calcule récursivement :

$$\mathcal{V}^{(k)} = \left\{ \mathcal{V}^{(k)}(s) \cup \{ \mathcal{V}^{(k-1)}(s') ; s' \in \mathcal{V}^{(k-1)}(s) \} \right\}$$

pour $k \in \llbracket 2, o \rrbracket$.

Paramètre	Intervalle
Ordre d'extension	$\llbracket 1; 5 \rrbracket$
Amplitude [nAm]	$[0.5; 1.5]$
Centre [ms]	$[125; 375]$
Largeur [ms]	$[48; 51]$

TABLE 1 – Paramètres de la simulation source unique étendue, simulation avec une unique source active qui correspond à une zone étendue d'activité. La zone active est définie par une position et une extension variable et le signal assigné aux sources actives possède une amplitude, un centre et une largeur variables.

Motif temporel Le motif temporel est la forme temporelle du signal assigné à chacune des sources actives. Pour le jeu de données de source unique étendue le motif temporel est similaire à celui d'un potentiel évoqué (*event related potential*, ERP). Il est créé en utilisant la fonction de la *toolbox* SEREEGA correspondant à la simulation d'ERP. Le signal est une fonction gaussienne en temps de la forme

$$f(t) = ae^{-\frac{1}{2}(\frac{6(t-c)}{w})^2}, \quad (1)$$

définie par son amplitude a , son centre c et sa largeur $w = 6\sigma$. Ces paramètres sont choisis aléatoirement dans des intervalles donnés pour chaque échantillon du jeu de données. Les paramètres utilisés pour cette simulation se trouvent dans le tableau 1. Le même signal temporel est assigné à toutes les sources d'une région avec une amplitude décroissante selon une Gaussienne fonction de la distance entre la source considérée et la source noyau de la région. Chaque exemple a une durée de 500 ms avec une fréquence d'échantillonnage de 512 Hz soit 256 échantillons temporels : $\mathbf{D} \in \mathbb{R}^{N_s \times 256}$.

Une donnée EEG non bruitée $\tilde{\mathbf{M}}$ est ensuite générée en projetant une distribution de source $\mathbf{D}$ sur l'espace des électrodes par le biais du modèle direct.

Un exemple de distribution des sources et de l'EEG correspondant est présenté Figure 1.

Sources multiples étendues. Un autre défi de l'ESI est d'estimer l'activité de sources multiples (étendues) dans le cerveau. Nous avons donc considéré ce cas en simulant un ensemble de données avec une ou plusieurs sources étendues actives.

Motif spatial Le même principe que pour le jeu de données source unique est utilisé pour sélectionner les sources d'une région. Dans ce jeu de données, plusieurs noyaux et ordres d'extension sont choisis aléatoirement dans des intervalles donnés et la distribution des sources est composée de plusieurs régions qui ne se chevauchent pas.

Motif temporel Le même signal gaussien de type ERP que pour l'ensemble de données de source unique étendue est attribué aux sources actives avec le même principe de décroissance des amplitudes des différentes sources actives au sein d'une région. De la variabilité entre les régions d'une même distribution de sources est ajoutée, ce

Paramètre	Intervalle entre exemples	Intervalle entre régions
Nombre de région	$[1; 5]$	$\emptyset$
Extension ordre	$[1; 5]$	$\emptyset$
Amplitude [nAm]	$[0.5; 1.5]$	$[0.15; 1.7]$
Centre [ms]	$[125; 375]$	$[87; 487]$
Largeur [ms]	$[48; 51]$	$[43; 56]$

TABLE 2 – Paramètres de simulation et intervalles pour le jeu de données multi-sources, variabilité entre différents exemples du jeu de données et entre différentes régions d'un même exemple.

qui signifie que différentes régions d'un même échantillon de distribution de sources auront des formes d'ondes différentes. Chaque échantillon a une durée de 500 ms et une fréquence d'échantillonnage de 512 Hz. Les paramètres et intervalles utilisés pour cette simulation sont indiqués dans le tableau 2.

Bruit. Du bruit blanc gaussien est ajouté aux données EEG à un rapport signal à bruit (SNR) donné. Si l'on note $\mathbf{M} \in \mathbb{R}^{N_e \times T}$ la matrice des données EEG sans bruit de dimension nombre d'électrodes N_e par nombre d'échantillons de temps T , et $\epsilon \in \mathbb{R}^{N_e \times T}$, $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ un bruit blanc gaussien, le SNR peut être ajusté en calculant

$$\mathbf{M} = \tilde{\mathbf{M}} + \frac{\epsilon}{\|\epsilon\|_F} \frac{\|\mathbf{M}\|_F}{\sqrt{\text{SNR}}}. \quad (2)$$

Des données EEG avec des SNR de 30 dB, 20 dB et 10 dB sont simulées comme différents ensembles de données pour évaluer la robustesse au bruit d'une méthode.

2.2 Méthodes de résolution du problème inverse par apprentissage supervisé

Dans ce travail, deux méthodes sont utilisées pour étudier l'apport de l'apprentissage automatique à l'ESI : un réseau *long short term memory* bidirectionnel (réseau LSTM) de 7 102 464 paramètres, adapté au plus proche de celui proposé par Hecker *et al.* [8] ainsi qu'un réseau convolutif à une dimension (*convolutional neural network* 1D, CNN 1D) de 513 400 paramètres. Le réseau LSTM est composé de deux couches LSTM bidirectionnelles et d'une couche dense qui permet d'obtenir en sortie le nombre de sources de l'espace source considéré à partir des caractéristiques en sortie des couches récurrentes.

Le CNN 1D est composé d'une couche convolutive à une dimension et d'une couche dense. La dimension des noyaux de convolution est $N_e \times 5$ afin de prendre en compte tous les canaux EEG et un voisinage temporel. Les deux réseaux sont entraînés avec une similarité cosinus pour fonction de coût.

La similarité cosinus ne prend pas en compte l'amplitude des données. Pour calculer l'amplitude des sources estimées, une normalisation est réalisée afin que les signaux

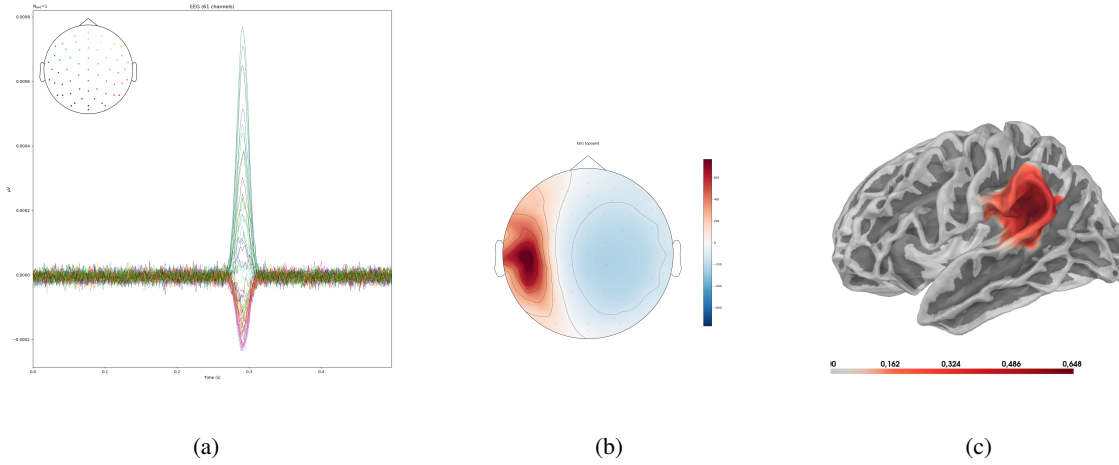


FIGURE 1 – Exemple d’une paire de données EEG et source : (a) *butterfly plot* d’un signal EEG avec un SNR de 10 dB. (b) tracé topographique à l’instant d’activité maximale du signal EEG, (c) distribution des sources visualisée sur le cortex au moment de l’activité maximale (hémisphère gauche).

EEG simulés à partir des sources estimées aient la même puissance globale à chaque instant que les données EEG appliquées en entrée du réseau (*global field power GFP*, méthode adaptée de [8]).

2.3 Méthodes de résolution du problème inverse non basées sur l’apprentissage

Deux méthodes de résolution du problème inverse en EEG, basées sur un a priori de minimisation de l’énergie de la source estimée, sont utilisées ici comme algorithmes de comparaison avec les méthodes par apprentissage automatique : *minimum norm estimates* (MNE) [12] et *standardized low-resolution brain electromagnetic tomography* (sLORETA) [14]. Les deux méthodes sont implémentées dans la librairie mne-python [3], et nécessitent le calcul de la covariance du bruit des données. Pour calculer cette matrice de covariance les points du signal simulé qui correspondent à du bruit et non pas à du signal actif sont agrégés pour obtenir un signal contenant seulement des échantillons de bruit.

La méthode MNE correspond à l’utilisation d’un terme de régularisation en norme L_2 (régularisation de Tikhonov), de la forme

$$\hat{\mathbf{d}} = \underset{\mathbf{d}}{\operatorname{argmin}} \{ \|\mathbf{m} - \mathbf{G}\mathbf{d}\|^2 + \lambda \|\mathbf{d}\|_2^2 \}, \quad (3)$$

où $\mathbf{d}$ est un vecteur de sources à un instant t , $\mathbf{m}$ le vecteur EEG correspondant, $\mathbf{G}$ la matrice de *leadfield* et λ le coefficient de régularisation. sLORETA part de la solution de MNE et ajoute une étape de normalisation pour mieux estimer les sources profondes notamment.

2.4 Métriques d’évaluation

Le but des différentes métriques d’évaluation en ESI est d’évaluer les aspects spatiaux et temporels de la solution

par rapport à la vérité-terrain. Pour un instant t_0 la distribution de sources est un vecteur d’amplitudes, et pour chacun des éléments indexés du vecteur (chaque source) une position correspondante dans l’espace est connue. Une solution idéale permettrait d’estimer parfaitement le vecteur d’amplitude, et donc de localiser les sources actives à leur position réelle.

Dans le cas d’une distribution avec une source unique étendue, l’objectif est d’estimer correctement la source avec une activité maximale, l’extension de la région et l’amplitude de l’activité [7, 15]. Toutes les mesures, à l’exception du PSNR, sont calculées au moment de l’activité maximale de la source noyau de la vérité-terrain. Le PSNR est calculé sur l’ensemble de la distribution de sources afin de fournir également une information sur l’estimation globale de la série temporelle.

L’**erreur de localisation** (*localisation error*, **LE**) est définie comme

$$LE(\mathbf{D}_{:,t_0}, \hat{\mathbf{D}}_{:,t_0}) = \operatorname{dist}(\mathbf{r}_s, \mathbf{r}_{\hat{s}}), \quad (4)$$

où $\operatorname{dist}(\cdot, \cdot)$ est la distance euclidienne entre la vraie source maximale s et la source maximale estimée $\hat{s}$. Pour une solution idéale, $LE = 0$.

Aire sous la courbe ROC (*receiver operating characteristic*) (**AUC**) l’AUC est généralement utilisée pour mesurer la précision d’un classifieur [2]. Pour le problème d’estimation de la distribution des sources la valeur absolue de la distribution estimée est mise à l’échelle entre 0 et 1 puis seuillées pour se ramener à un cas de classification binaire dont les classes sont les sources actives et inactives. L’AUC caractérise la capacité d’une méthode à estimer correctement l’extension de la source et à ne pas créer de valeurs aberrantes.

Métrique	Aspect de la distribution de source qui est évalué
Erreur de localisation (LE)	Localisation
MSE normalisée (nMSE)	Amplitude
AUC	Extension
Erreur temporelle (TE)	Temporel
PSNR (sur plusieurs instants)	Amplitude

TABLE 3 – Aspect évalué par les différentes métriques d'évaluation. L'erreur de localisation, la MSE normalisée et l'AUC sont calculées pour un seul instant tandis que le PSNR est calculé sur l'ensemble de la série temporelle.

Erreur quadratique moyenne normalisée (nMSE) la nMSE est l'erreur quadratique moyenne (MSE) entre les distributions de sources normalisées, c'est-à-dire la distribution de sources divisée par sa valeur d'amplitude maximale. La normalisation de la distribution de sources permet une meilleure comparaison des valeurs des métriques entre les différentes méthodes. Ainsi elle prend en compte le motif global estimé plus que la valeur précise des amplitudes. Elle est également plus informative puisque "l'amplitude" estimée par différentes méthodes n'a pas toujours la même signification.

La MSE ou la nMSE ne fournissent pas une forte information sur l'erreur de localisation : une activité parasite aura le même effet sur la MSE si elle est localisée très près du patch de sources actives que si elle en est éloignée.

Erreur temporelle (time error, TE) L'erreur temporelle donne une information sur la capacité d'une méthode à estimer correctement la forme d'onde des sources (activité temporelle). L'erreur temporelle est la valeur absolue de la différence entre l'instant d'activité maximale de la vérité-terrain et l'instant d'activité maximal pour la source estimée (instant du maximum d'amplitude de la source noyau estimée). $TE = |t_{\max} - \hat{t}_{\max}|$.

Peak Signal to Noise Ratio (PSNR) Le PSNR est calculé entre les signaux normalisés (comme pour la nMSE). Le PSNR est calculé sur l'ensemble de la distribution de sources (plusieurs échantillons temporels) pour donner une information sur l'estimation de l'ensemble de l'activité temporelle.

Dans le cas d'activités contenant des sources multiples, l'instant d'activité maximale de chaque source de la vérité-terrain est d'abord estimé. Ensuite pour chaque paire (noyau s , instant d'activité maximale t_0), la source estimée est identifiée comme étant la source ayant une amplitude maximale dans un voisinage d'ordre 7 du noyau de la vérité-terrain à l'instant t_0 . Les métriques sont calculées noyau par noyau de la même manière que pour l'activité d'une source unique et la valeur finale est la moyenne des métriques associées à chaque noyau. Le PSNR est calculé de la même manière que pour l'activité d'une source unique.

3 Résultats

3.1 Paramètres d'entraînement

Chacun des jeux de données, avec source unique et sources multiples et pour chaque valeur de SNR, contient 10 000 exemples. Pour chaque expérience, 8000 exemples sont utilisés pour l'entraînement et 2000 pour la validation. L'optimiseur utilisé est ADAM, avec un taux d'apprentissage de 0.001. Pour le LSTM, la taille de batch est de 8 et les gradients sont bornés à une norme maximale de 1 pour être en adéquation avec les paramètres d'entraînement de l'article original [8]. Pour le CNN 1D, la taille de batch est de 16 et les gradients ne sont pas bornés.

3.2 Imagerie de source électrophysiologique de données avec une source étendue

Les résultats sont présentés dans le Tableau 4 et un exemple de visualisation de données sources estimées par les différentes méthodes est présenté Figure 2a.

Les résultats obtenus avec les deux réseaux de neurones sont comparables et sont pour toutes les métriques évaluées supérieurs aux résultats obtenus pour MNE et sLORETA. La différence est particulièrement notable en matière d'erreur de localisation, pour laquelle les réseaux ont une erreur 2 fois plus faible que sLORETA. La Figure 2a montre également la meilleure capacité des réseaux de neurones à estimer la bonne extension de la source, ce qui est aussi exprimé par l'AUC. La forme temporelle est globalement bien reconstruite par toutes les méthodes, mais l'erreur d'amplitude est moindre pour les réseaux, surtout par rapport à sLORETA, qui ne retourne pas une amplitude au sens physique de l'amplitude des sources.

3.3 Imagerie de source électrophysiologique de données avec plusieurs sources étendues

Les résultats pour plusieurs sources étendues sont présentés dans le Tableau 5. De plus, un exemple d'estimation spatiale pour un échantillon contenant deux sources visibles dans l'hémisphère gauche est présenté Figure 2b. Là encore les deux réseaux utilisés ont des performances supérieures que les méthodes MNE et sLORETA. De même, visuellement les deux sources sont bien reconstruites. L'extension est estimée plus précisément par le CNN 1D et le LSTM que par les deux méthodes à minimisation de l'énergie, qui peinent à reconstruire deux sources distinctes.

4 Conclusion

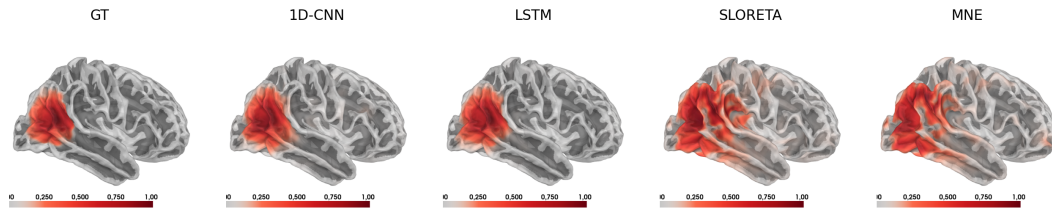
Pour résoudre le problème inverse mal posé que constitue l'ESI, de nombreux algorithmes utilisent un a priori sur l'activité cérébrale à estimer. Cependant, le choix d'un a priori réaliste n'est pas évident et influence la solution obtenue. L'étude de source étendue, unique ou multiples, est un exemple d'activité cérébrale qui reste un défi pour beaucoup de méthodes. Les méthodes d'apprentissage profond appliquées à l'ESI permettent d'apprendre la relation

SNR	Methods Metrics	CNN1d	LSTM	MNE	sLORETA
30dB	AUC $\uparrow$	0.9860	0.9878	0.7862	0.7844
	LE [mm] $\downarrow$	2.94	2.45	9.87	6.25
	nMSE $\downarrow$	0.0039	0.0024	0.0231	0.0328
	PSNR $\uparrow$	39.5436	41.8593	32.0799	29.7051
	time error [ms] $\downarrow$	0.05	0.45	1.25	1.07
20dB	AUC $\uparrow$	0.9859	0.9874	0.7847	0.7847
	LE [mm] $\downarrow$	3.31	2.48	10.07	6.26
	nMSE $\downarrow$	0.0046	0.0024	0.0233	0.0354
	PSNR $\uparrow$	38.7850	41.6129	29.4784	27.5368
	time error [ms] $\downarrow$	0.11	0.46	2.62	2.00
10dB	AUC $\uparrow$	0.9864	0.9892	0.7840	0.7872
	LE [mm] $\downarrow$	3.31	2.52	10.29	6.37
	nMSE $\downarrow$	0.0051	0.0024	0.0267	0.0417
	PSNR $\uparrow$	37.8922	39.7914	26.2729	24.3173
	time error [ms] $\downarrow$	0.32	0.51	5.22	3.82

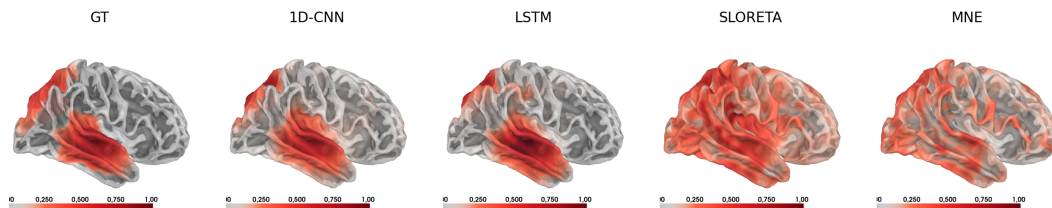
TABLE 4 – Évaluation de méthodes de résolution du problème inverse par apprentissage automatique vs méthodes linéaires, dans le cas de données simulées avec une unique source étendue active. Problème inverse avec $N_e=90$ électrodes, $N_s = 1284$ sources. Validation sur 2000 exemples.

SNR	Methods Metrics	CNN1d	LSTM	MNE	sLORETA
30dB	AUC $\uparrow$	0.7818	0.7782	0.6600	0.6558
	LE [mm] $\downarrow$	4.42	5.22	11.05	7.87
	nMSE $\downarrow$	0.0053	0.0045	0.0232	0.0375
	PSNR $\uparrow$	37.6688	38.3320	30.3987	27.6743
	time error [ms] $\downarrow$	1.91	3.23	7.42	6.61
20dB	AUC $\uparrow$	0.7807	0.7763	0.6528	0.6510
	LE [mm] $\downarrow$	4.76	5.48	11.87	8.63
	nMSE $\downarrow$	0.0062	0.0046	0.0257	0.0440
	PSNR $\uparrow$	36.8713	37.6045	27.6764	25.3797
	time error [ms] $\downarrow$	2.65	5.76	15.14	10.89
10dB	AUC $\uparrow$	0.7786	0.7761	0.6459	0.6461
	LE [mm] $\downarrow$	5.11	5.77	12.93	9.84
	nMSE $\downarrow$	0.0082	0.0050	0.0309	0.0550
	PSNR $\uparrow$	33.9728	34.4751	24.7509	22.3054
	time error [ms] $\downarrow$	8.09	13.45	25.86	18.28

TABLE 5 – Évaluation de méthodes de résolution du problème inverse par apprentissage automatique vs deux méthodes linéaires dans le cas de données avec plusieurs sources étendues.



(a) Résultats dans le cas d'une activité avec une source unique étendue.



(b) Résultats pour le cas d'une activité contenant plusieurs sources étendues.

FIGURE 2 – Visualisation de l'activité estimée par les différentes méthodes. Les résultats sont seuillés de sorte à ce que les amplitudes inférieures à 25% de l'activité maximale sont considérées inactives. Pour le cas des sources multiples l'activité visualisée est la moyenne des activités aux différents maxima de chaque source noyau. Dans l'exemple en (b), il y a trois régions actives, donc trois noyaux : l'activité visualisée est la moyenne des activités de ces trois sources noyaux. GT = vérité terrain.

entre les données EEG en entrée et l'activité cérébrale en sortie sans l'ajout externe d'un a priori. Pour pouvoir entraîner un réseau à cette tâche, des données simulées sont utilisées. Dans cet article, les performances d'un réseau récurrent LSTM et d'un réseau convolutif à une dimension sont comparées aux méthodes de l'état de l'art MNE et sLORETA, reposant sur une minimisation de l'énergie de la solution. Les résultats sur données réalistes simulées de sources étendues, unique ou multiples, sont encourageants quand à l'utilisation des réseaux de neurones.

Dans le cadre des approches par apprentissage, le choix de l'architecture est important et doit être bien adapté aux données : les deux réseaux ont des performances similaires mais le réseau LSTM de Hecker *et al.* possède environ 15 fois moins de paramètres que le CNN 1D. Pour valider de manière plus exhaustive l'efficacité de ces méthodes, des évaluations sur données réelles doivent également être réalisées pour étudier la capacité des réseaux à généraliser, capacité primordiale pour leur utilisation clinique. Enfin, l'approche étudiée dans cet article consiste à apprendre de manière supervisée l'opérateur inverse. Une future direction de recherche pourra consister à développer une approche variationnelle plus flexible [1], notamment en apprenant le terme de régularisation.

5 Remerciements

Les recherches qui ont conduit à ces résultats ont été financées par le Labex CominLabs dans le cadre du projet CYCLE et l'ANR (AI4CHILD ANR-19-CHIA-0015-01).

Références

- [1] R. Fablet, B. Chapron, L. Drumetz, E. Mémin, O. Pannekoucke, and F. Rousseau. Learning Variational Data Assimilation Models and Solvers. *Journal of Advances in Modeling Earth Systems*, 13(10), October 2021.
- [2] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8) :861–874, June 2006.
- [3] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, and Matti S. Hämäläinen. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7(267) :1–13, 2013.
- [4] Roberta Grech, Tracey Cassar, Joseph Muscat, Kenneth P Camilleri, Simon G Fabri, Michalis Zervakis, Petros Xanthopoulos, Vangelis Sakkalis, and Bart Vanrumste. Review on solving the inverse problem in EEG source analysis. *Journal of NeuroEngineering and Rehabilitation*, 5(1) :25, December 2008.
- [5] C. Grova, J. Daunizeau, J.-M. Lina, C.G. Bénar, H. Benali, and J. Gotman. Evaluation of EEG localization methods using realistic simulations of interictal spikes. *NeuroImage*, 29(3) :734–753, February 2006.
- [6] Becker Hanna. *Denoising, separation and localization of EEG sources in the context of epilepsy*. PhD thesis, Nice Sophia Antipolis, 2014.

- [7] Olaf Hauk, Matti Stenroos, and Matthias Treder. Towards an Objective Evaluation of EEG/MEG Source Estimation Methods : The Linear Tool Kit. preprint, Neuroscience, June 2019.
- [8] Lukas Hecker, Rebekka Rupprecht, Ludger Tebartz van Elst, and Jürgen Kornmeier. Long-Short Term Memory Networks for Electric Source Imaging with Distributed Dipole Models. Technical report, bioRxiv, April 2022.
- [9] Lukas Hecker, Rebekka Rupprecht, Ludger Tebartz Van Elst, and Jürgen Kornmeier. ConvDip : A Convolutional Neural Network for Better EEG Source Imaging. *Frontiers in Neuroscience*, 15 :569918, June 2021.
- [10] Gert Van Hoey, Jeremy De Clercq, Bart Vanrumste, Rik Van de Walle, Ignace Lemahieu, Michel D’Havé, and Paul Boon. EEG dipole source localization using artificial neural networks. *Physics in Medicine and Biology*, 45(4) :997–1011, April 2000.
- [11] Gexin Huang, Zhu Liang Yu, Wei Wu, Ke Liu, ZhengHui Gu, Feifei Qi, YuanQing Li, and Jiawen Liang. Electromagnetic Source Imaging via a Data-Synthesis-Based Denoising Autoencoder, July 2021.
- [12] M. S. Hämmäläinen and R. J. Ilmoniemi. Interpreting magnetic fields of the brain : minimum norm estimates. *Medical & Biological Engineering & Computing*, 32(1) :35–42, January 1994.
- [13] Laurens R. Krol, Juliane Pawlitzki, Fabien Lotte, Klaus Gramann, and Thorsten O. Zander. Sereega : Simulating event-related eeg activity. *Journal of Neuroscience Methods*, 309 :13–24, 2018.
- [14] R. D. Pascual-Marqui. Standardized low-resolution brain electromagnetic tomography (sLORETA) : technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, 24 Suppl D :5–12, 2002.
- [15] John G. Samuelsson, Noam Peled, Fahimeh Mamasli, Jyrki Ahveninen, and Matti S. Hämmäläinen. Spatial fidelity of MEG/EEG source estimates : A general evaluation approach. *NeuroImage*, 224 :117430, January 2021.
- [16] Rui Sun, Abbas Sohrabpour, Gregory A. Worrell, and Bin He. Deep neural networks constrained by neural mass models improve electrophysiological source imaging of spatiotemporal brain dynamics. *Proceedings of the National Academy of Sciences*, 119(31) :e2201128119, August 2022.
- [17] Rui Sun, Abbas Sohrabpour, Shuai Ye, and Bin He. SIFNet : Electromagnetic Source Imaging Framework Using Deep Neural Networks. preprint, Bioengineering, May 2020.
- [18] Chen Wei, Kexin Lou, Zhengyang Wang, Mingqi Zhao, Dante Mantini, and Quanying Liu. Edge Sparse Basis Network : A Deep Learning Framework for EEG Source Localization, June 2021.
- [19] Q. Zhang, M. Yuasa, H. Nagashino, and Y. Kinouchi. Single dipole source localization from conventional EEG using BP neural networks. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286)*, volume 4, pages 2163–2166, Hong Kong, China, 1998. IEEE.