

BioSerenity-E1 : un modèle de fondation EEG pour des applications cliniques cliniques multiples

G. Jubien¹, R. G. Bettinardi¹, M. Rahmouni¹, A. Bussalb¹, F. Le Gac¹, P. Di Piazza¹, C. Flipo¹, U. Gimenez¹

¹ BioSerenity, Paris, France

CONTEXTE & MOTIVATION

Le défi de l'IA appliquée à l'EEG

Le développement d'algorithmes IA pour l'EEG clinique se heurte à plusieurs obstacles fondamentaux : la rareté des données annotées, la complexité des signaux neurophysiologiques et la diversité des configurations matérielles. Les approches classiques nécessitent de larges jeux de données labellisés pour chaque tâche, rendant le développement de nouveaux cas d'usage long et coûteux.

Le paradigme des modèles de fondation

À l'instar des grands modèles de langage (LLMs), les modèles de fondation apprennent des représentations généralisables à partir de vastes corpus non étiquetés, puis s'adaptent rapidement à de nouvelles tâches cliniques par fine-tuning. BioSerenity a développé BioSerenity-E1 [1], premier modèle de fondation EEG de la société, pré-entraîné de façon auto-supervisée sur plus de 4 000 heures d'EEG cliniques et publics.

MÉTHODES — ARCHITECTURE DU MODÈLE DE FONDATION

Pré-entraînement auto-supervisé

Corpus : > 4 000 h d'EEG (données BioSerenity + bases publiques)

Architecture : Transformer avec mécanismes d'attention intra- et inter-canaux, capturant les dépendances spatio-temporelles du signal EEG.

Stratégie : Tokenisation spectrale + masked token prediction (70 % de masquage). Le modèle reconstruit les tokens masqués, apprenant ainsi la structure intrinsèque du signal.

Downstream tasks

Pour évaluer la capacité de généralisation et performance de BioSerenity-E1, plusieurs tâches ont été développées. La diversité des applications retenues permet de s'assurer que le modèle est performant dans des contextes variés, et non optimisé pour un seul cas d'usage.

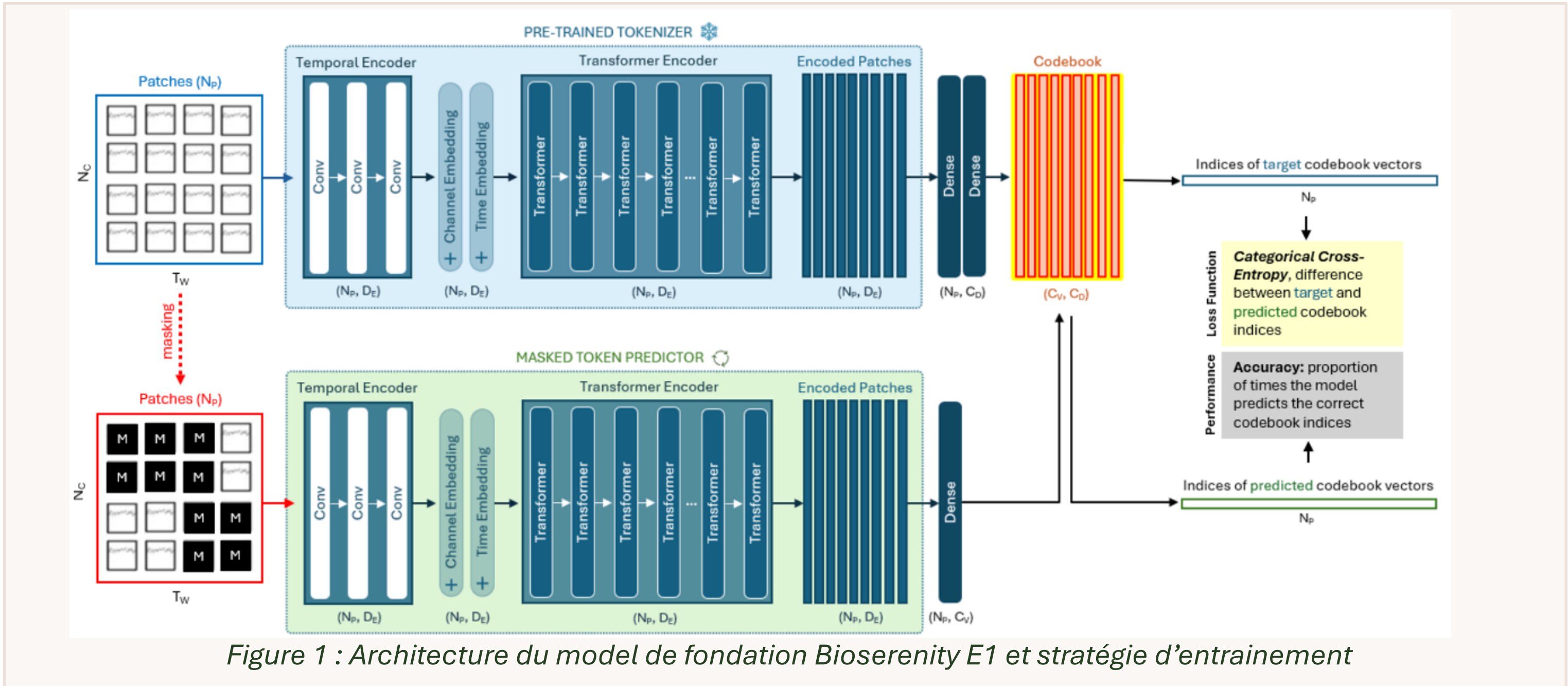


Figure 1 : Architecture du model de fondation Bioserenity E1 et stratégie d'entrainement

MÉTHODES — 4 CAS D'USAGE CLINIQUES

01 Normal / Anormal [2]	02 Crises d'épilepsie	03 Sommeil pédiatrique	04 Maladie d'Alzheimer
Objectif Classifier des EEG de routine (~20 min, adultes ≥ 18 ans) en normal ou anormal.	Objectif Identification automatique des événements ictaux en contexte ambulatoire et hospitalier.	Objectif Classification automatisée des stades de sommeil (critères AASM) chez l'enfant de 6 à 18 ans.	Objectif Détection binaire (MA vs. contrôle) à partir d'EEG de repos de 20 min.
Données d'entraînement 2 500 EEG multicentriques équilibrés (Neurophy, France, 2020–2023).	Données d'entraînement & de test Entraînement sur TUH EEG Seizure Corpus (n=5255 en entraînement / 1572 en test) & TUH EEG Artifact Corpus (n=300 en entraînement) [3]	<i>Ce modèle a été entraîné avec la version single channel du LBM</i>	Données d'entraînement & de test 1 646 sujets (779 MA, 867 contrôles appariés âge/comorbidités). Split sujet-indépendant 60/20/20.
Données de test Dataset A interne (n = 4 480, 1 expert) Dataset B multi-annoté (n = 198, 3 experts , consensus au 2/3 de la majorité) TUH Abnormal EEG Corpus benchmark public (n = 276).	Données de test Évaluation sur dataset de Siena (n=39)	Données d'entraînement & de test Entraînement sur 410 PSG Adultes. Finetuning sur PSG pédiatriques (n=209 en entraînement / 53 en test),	Données de test Validation externe sur cohorte indépendante (n = 65) [4].
Architecture Encodeur gelé + tête de prédiction (convolutions, batch norm, dropout p = 0,4, global average pooling, couches linéaires).	Architecture N fenêtres de 16 s + deux branches parallèles : encodeur BioSerenity-E1 & CNN spectrogramme +Fusion des vecteurs + réseau récurrent	Données de test Validation externe (HFME, Lyon) : 46 PSG sains <i>En cours de validation sur 100 SAS et 100 narcolepsies.</i>	Architecture Encodeur gelé + tête de classification avec mécanismes d'attention intra- et inter-canaux ; agrégation au niveau de l'enregistrement pour une prédiction sujet-niveau.
Etude Clinique 395 EEG de routine multicentriques, 3 experts, consensus au 2/3 de la majorité	Etude Clinique <i>En cours d'étude clinique</i>	Architecture 6 EEG + 2 EOG + 1 EMG, epochs de 30 s. Architecture canal-agnostique : encodeur gelé + GRU bidirectionnel + mécanisme d'attention finale.	

RÉSULTATS

NORMAL / ANORMAL	CRISES D'ÉPILEPSIE	SOMMEIL PÉDIATRIQUE	MALADIE D'ALZHEIMER
87.4 % - 98.8 %	71.8 % - 0.2	0.79	0.74 – 0.78
Sensitivité – Spécificité / Etude Clinique	Sensibilité - FPh / Dataset Siena	Macro-F1 / Dataset externe HFME	Balanced accuracy (cohorte externe)
Résultats complémentaires :	Résultats complémentaires :	Résultats complémentaires :	Résultats complémentaires :
Dataset A : 86.46% [84.94-87.99] - 91.93% [90.83-93.00] Dataset B : 91.39 [86.71-95.65] - 97.87 [93.33-100] TUAB (validation externe) : 65.87 [56.91-73.28] - 96.00 [92.62-98.74]	TUSZ : 72.3 % - 2.1 <i>Résultats préliminaires prometteurs pour la surveillance continue en conditions réelles.</i>	Précision par classe : 0.91 pour Eveil, 0.59 pour N1, 0.8 pour N2, 0.87 pour N3, 0.91 pour REM. Dataset de test interne : 0.75 de macro-F1	Spécificité très élevée (0.97 – 1.00) & sensibilité modérée (0.47 – 0.75) sur la cohorte externe indépendante (n = 65). Balanced accuracy (0.69 – 0.70) sur le jeu de test interne (n = 155).

DISCUSSION

Robustesse et Généralisabilité

BioSerenity-E1 démontre une bonne capacité de généralisation à travers des tâches cliniques hétérogènes, avec des performances maintenues de manière cohérente sur les jeux de données internes et les cohortes externes indépendantes.

Facilitation du développement

Les résultats obtenus sur la tâche Normal/Anormal montrent qu'à performances équivalentes, le fine-tuning de BioSerenity-E1 nécessite environ 10 fois moins de données labellisées qu'une approche from scratch, facilitant le développement de nouvelles applications cliniques, notamment pour les pathologies rares [2].

Limitations & perspectives

L'architecture actuelle reste contrainte à un nombre fixe de canaux d'entrée. Les prochains développements visent une entrée flexible aux montages hétérogènes, une montée en échelle du modèle, et l'évaluation sur un spectre plus large de cas d'usages.

Conclusion

BioSerenity-E1 constitue une infrastructure IA pérenne pour la neurophysiologie clinique. Sa capacité à s'adapter rapidement à des tâches variées (classification de l'EEG, pathologies neurodégénératives, sommeil pédiatrique, crises d'épilepsie) avec un coût d'annotation réduit en fait une plateforme différenciante.

L'algorithme Normal/Anormal est d'ores et déjà intégré dans un dispositif médical certifié CE via Neurophy. L'objectif est d'étendre progressivement de nouveaux algorithmes, actuellement en évaluation dans des études cliniques, vers la production.

Références :

- [1] Bettinardi R.G., Rahmouni M., Gimenez U. BioSerenity-E1: a self-supervised EEG model for medical applications. arXiv:2503.10362 (2025) .
- [2] Bussalb A. et al. Automatic detection of abnormal clinical EEG: comparison of a finetuned foundation model with two deep learning models. arXiv:2505.21507 (2025) .
- [3] Shah V. et al. The Temple University Hospital Seizure Detection Corpus. *Front. Neuroinform.* 12:83 (2018) .
- [4] Miltiadous A. et al. A Dataset of Scalp EEG Recordings of Alzheimer's Disease, Frontotemporal Dementia and Healthy Subjects. *Data* 8(6):95 (2023) .