

**DOSSIER DE CANDIDATURE
POUR THESE EN COTUTELLE
POUR LA RENTREE 2025
FINANCEMENT : BOURSE**

Dossier complété et revêtu des signatures à transmettre impérativement pour le :

1^{er} avril 2025 au plus tard,

A la Direction de la Recherche et Valorisation

secretariat.recherche@univ-littoral.fr

Rappel : les doctorants bénéficiant des financements de cotutelle n'ont pas de contrat de travail et ne perçoivent pas de salaire. Ils ont le statut étudiant et bénéficient d'une bourse.

Montant maximum cofinancé par l'ULCO : 1 600€/mois (exemple : si le doctorant bénéficie d'une bourse de 1000€ pendant 36 mois par le pays partenaire, l'ULCO complètera avec 600€ de bourse pendant les 18 mois passés à l'ULCO)

Titre de la thèse : Apprentissage par Transformers pour l'analyse des signaux EEG. Application à la détection de l'épilepsie

Laboratoire d'accueil ULCO : LISIC

Directeur de thèse ULCO : F. Dufrenois

Directeur de thèse Partenaire : Yasser Mohanna

e-mail : yamoha@ul.edu.lb

Merci de fournir une attestation de financement du partenaire (celle-ci peut être conditionnée au cofinancement ULCO)

Partenariat/thématique : (Merci de cocher le nom du partenaire et de sélectionner la thématique de la thèse)

Si nouveau partenariat, compléter la rubrique AUTRES et nous indiquer l'université d'origine, le directeur de thèse, le laboratoire, ainsi que les conditions de financement du doctorant dans le pays d'origine

• **LIBAN – Université Libanaise**

- Thématique :

- (1) La qualité de l'air
- (2) Le milieu aquatique
- (3) L'obésité, la nutrition et les activités sportives,
- (4) Les énergies propres et renouvelables
- (5) La gestion et le traitement des déchets
- (6) L'urbanisme
- (7) **santé**

○ **LIBAN – CNRS Libanais**

- Thématique :

- (1) La qualité de l'air
- (2) Le milieu aquatique
- (3) L'obésité, la nutrition et les activités sportives

- (4) Les énergies propres et renouvelables
- (5) La gestion et le traitement des déchets
- (6) L'urbanisme

○ **MAROC - Université Mohammed VI Polytechnique**

- Thématique :

- (1) Mathématiques Appliquées, Sciences de Données
- (2) Sciences de l'Ingénieur (informatique, traitement du signal et de l'industrie 4.0)
- (3) Ingénierie des systèmes complexes

○ **MAROC - Université Hassan II**

- Thématique :

- (1) Economie Gestion
- (2) Sciences et Technologies

○ **MAROC - Université Cadi Ayyad (inclus dans la convention de partenariat A2U / CNRST du Maroc)**

- Thématique :

- (1) Sciences marines
- (2) Intelligence artificielle
- (3) Energie renouvelables et énergie alternatives
- (4) Environnement et ressources naturelles
- (5) Santé
- (6) SHS

Le candidat identifié bénéficie-t-il d'une bourse Doctorants-Moniteurs PASS du CNRST : oui-non

○ **MAROC - Université Ibn Zohr (inclus dans la convention de partenariat A2U / CNRST du Maroc)**

- Thématique :

- (1) Sciences marines
- (2) Intelligence artificielle
- (3) Energie renouvelables et énergie alternatives
- (4) Environnement et ressources naturelles
- (5) Santé
- (6) SHS

Le candidat identifié bénéficie-t-il d'une bourse Doctorants-Moniteurs PASS du CNRST : oui-non

○ **MAROC - Université Mohammed V (inclus dans la convention de partenariat A2U / CNRST du Maroc)**

- Thématique :

- (1) Sciences marines
- (2) Intelligence artificielle
- (3) Energie renouvelables et énergie alternatives
- (4) Environnement et ressources naturelles
- (5) Santé
- (6) SHS

Le candidat identifié bénéficie-t-il d'une bourse Doctorants-Moniteurs PASS du CNRST : oui-non

○ **MAROC - Université Mohammed Premier (inclus dans la convention de partenariat A2U / CNRST du Maroc)**

- Thématique :

- (1) Sciences marines
- (2) Intelligence artificielle
- (3) Energie renouvelables et énergie alternatives

- (4) Environnement et ressources naturelles
- (5) Santé
- (6) SHS

○ **BELGIQUE - UCLouvain - Université Catholique de Louvain**

- Thématique :

- (1) Economie Gestion

○ **CANADA - UQAR : Université du Québec à Rimouski**

- Thématique :

- (1) Sciences de la Mer

○ **AUTRES : Nouveau partenariat :**

- Université d'origine :
- Laboratoire :
- Directeur de thèse :
- Conditions de financement :
- Thématique :

***LABORATOIRE D'ACCUEIL**

Nom du laboratoire d'accueil : LISIC

Nombre de HDR dans le laboratoire : 20

Nombre de thèses encadrées dans le laboratoire (rentrée 2024) : 25

Cotutelles en cours au sein du laboratoire : 9

Durée moyenne des thèses soutenues dans le laboratoire, sur la période 2020-2024 : 3 ans 3 mois

ENCADREMENT

Nom, Prénom du directeur de laboratoire : Verel Sebastien

Nom, Prénom du directeur de thèse (si différent du directeur de laboratoire) : Dufrenois Franck

Nombre de doctorats en préparation sous la direction du directeur de thèse : 1

Avis détaillé du directeur de thèse : Ce projet de thèse s'inscrit pleinement dans mes activités de recherche sur l'apprentissage supervisé mais également dans le cadre de l'alliance A2U. Il aborde une problématique de recherche innovante basée sur une nouvelle architecture de réseaux de neurones profonds : les transformers. Construit autour du principe de l'attention, il a révolutionné le monde du deep learning en particulier le traitement du langage naturel. L'atout de cette méthodologie est de prendre en compte à la fois le contexte et la composante temporelle dans les données. L'objectif de cette thèse est de construire un modèle de transformers adapté pour traiter des séries temporelles multidimensionnelles telles que les signaux

EEG. L'application visée est la détection des crises d'épilepsie. Il s'inscrit donc dans la problématique santé, thématique défendue à la fois par l'ULCO et l'université Libanaise. Au delà de l'application, l'analyse des signaux EEG et leur décodage soulève encore de nombreuses difficultés liées à leur nature non-stationnaire, bruitée et multimodale. Cette thèse a pour ambition d'apporter une réponse innovante à l'ensemble de ces défis à travers un modèle d'IA prometteur.

Signature du directeur de thèse



Avis détaillé du directeur de laboratoire :

Le projet de thèse s'inscrit dans une collaboration de longue date entre l'université libanaise et le LISIC. Le sujet de thèse s'est donc construit entre le LISIC (F. Dufrenois) et le Liban (Yasser Mohanna) de manière naturelle.

Le sujet porte sur l'une des priorités du LISIC en apprentissage automatique (IA) sur des techniques avancées en apprentissage profond (réseaux de neurones avec architecture de type transformer) pour une problématique en santé (EEG) où les données médicales sont disponibles de manière ouverte. A ce titre, le sujet rejoint l'une des problématiques de l'EUR MAIA en santé.

Pour toutes ces raisons, je donne un avis favorable à ce projet de thèse.

Signature du directeur de laboratoire



Intitulé du projet de thèse

Apprentissage par Transformers pour l'analyse des signaux EEG.
Application à la détection de l'épilepsie

Domaine scientifique

Machine Learning- IA

Résumé

Les troubles neurologiques sont l'une des principales causes d'invalidité signalées dans le monde. Il s'agit d'un grave problème de santé qui attend des solutions efficaces et des plans de traitement. Par exemple, les crises d'épilepsie font souffrir périodiquement 50 millions de personnes affectant tous les aspects de la vie. Ces troubles neurologiques sont liés à une connectivité fonctionnelle complexe et peuvent être diagnostiqués par des signaux d'électroencéphalographie (EEG) du cuir chevelu. Le signal EEG est caractérisé par des bandes de fréquences spécifiques. Il contient de riches informations spatio-temporelles sur les activités cérébrales pour découvrir la connectivité fonctionnelle dynamique entre les différentes régions du cerveau, identifier les modèles de troubles cérébraux et localiser les lésions telles que les foyers épileptiques. Dans cette thèse, nous proposons d'identifier les principales crises d'épilepsie à partir d'un modèle d'apprentissage profond nouveau appelé « transformers ». Introduit initialement pour le traitement du langage naturel, puis pour la classification d'images, il s'est imposé comme un concurrent sérieux aux modèles standards tels que les réseaux de neurones convolutionnels. Basé sur le principe de l'attention, il semble être un outil adapté pour intégrer la dépendance à la fois temporel et spatial des signaux EEG. Cette thèse aura pour but de vérifier cette hypothèse et montrer les capacités du transformers à décoder ces signaux. Notre étude sera basée sur la base de données européenne EPILESTIAE correspondant à 6488 heures d'enregistrement EEG avec plus de 25 crises d'épilepsie

Projet de thèse

a) Le sujet de recherche choisi et son contexte scientifique

Les signaux EEG enregistrent l'activité électrique du cerveau et constituent de ce fait la porte d'entrée pour analyser et comprendre son fonctionnement. La technologie EEG trouve de nombreuses applications par exemple, dans le décodage de tâches cognitives, le décodage affectif ou émotionnel, la détection de

troubles ou de maladies neurologiques (épilepsie, Alzheimer, trouble du sommeil), pour en citer quelques unes.

Le décodage s'effectue généralement en trois étapes distinctes: une phase de prétraitement permettant d'éliminer ou de réduire le bruit et certains artefacts, une phase d'extraction de caractéristiques et enfin une phase de classification basée sur des modèles d'apprentissage le plus souvent supervisés.

Dans le cadre cette thèse, nous proposons d'étudier le problème de la classification des signaux EEG à l'aide d'une technique d'apprentissage récente basée sur le concept d'attention: « les Transformers » [3]. Conçus initialement pour le traitement du langage naturel et ensuite pour résoudre des problèmes de classification en vision, ils sont devenus des concurrents sérieux aux méthodes standards telles que les réseaux de neurones convolutionnels (CNN) et les réseaux de neurones récurrents (RNN). L'utilisation des Transformers dans le cadre de l'analyse des signaux EEG reste encore assez confidentiel et ceci constitue une nouvelle opportunité [4]. L'un des points forts de ce type d'approche est sa capacité à prendre en compte les dépendances à long terme. Cette propriété peut donc constituer un atout essentiel pour prendre en compte la corrélation naturelle existante dans les signaux EEG. Cette thèse aura pour objectif de vérifier cette hypothèse et de comparer les résultats obtenus avec des techniques d'apprentissage profonds. Cette thèse sera appliquée à la détection de l'épilepsie.

L'étudiant en thèse aura de bonnes bases en programmation (python, matlab) ainsi que des notions de traitement du signal.

b) L'état du sujet dans le laboratoire et l'équipe d'accueil

Ce projet repose sur une longue coopération avec l'Université Libanaise (UL) au travers d'encadrement de stages de Master et de cotutelles de thèses. Cette collaboration a d'abord été initiée par D. Hamad à Tripoli (Faculté de Génie 1) et à Beyrouth, (facultés des sciences, branches 2 et 3), puis poursuivie par A. Bigand à Beyrouth (UL 1) à partir de 2011. A partir de cette date, la collaboration avec l'UL s'est d'abord déroulée sous forme d'un encadrement de thèse en co-tutelle (Hala Hijazi, co-dirigée par le Prof. O.Bazzi, UL, soutenance effectuée le 19/12/13, [1]). En parallèle, le Prof. O.Bazzi et A. Bigand avons mis en place une collaboration pédagogique qui a débouché sur la mise en place d'un co-diplôme entre les master TSI (ULCO) et STIP (UL). La première étape de cette collaboration repose sur l'échange d'étudiants pour le stage de master 2 et la définition d'un cours à l'UL (Pattern Recognition and Machine Learning). La deuxième étape de cette collaboration a été constituée par la montée en puissance des

objectifs précédents (stages de master 2, échange de cours) et la poursuite de nos travaux de recherche communs sous forme d'une nouvelle thèse (A.Darwich), avec le professeur Y. Mohanna. Nous espérons poursuivre maintenant cet effort de collaboration par le biais d'un nouveau projet reposant sur un axe prometteur du Machine Learning, les « Transformers ».

D'autre part, ce projet s'inscrit aussi dans le projet A2U IA. H. Hijazi [1] a mis au point des outils d'interprétation et de visualisation augmentée de données, par l'intermédiaire de méthodes de réduction de la dimension pour l'analyse exploratoire de données multidimensionnelles. Ces méthodes ont été étendues à l'apprentissage semi-supervisé (thèse de co-tutelle de M. A. Darwich soutenue en mars 2018, [2]) et devraient être appliquées aux données multi-modales sans problème par le biais de ce projet.

Le travail se fera donc en étroite collaboration avec l'UL. Il s'appuiera sur une collaboration active avec des collègues de l'Université Libanaise par une application des méthodes obtenues par le doctorant aux problèmes de *santé au Liban* et des publications communes.

Ce projet de thèse s'inscrit dans une thématique nouvelle de l'équipe IMAP, ***l'apprentissage automatique*** et un projet phare de l'université : la **santé**.

c) Le programme et l'échéancier de travail

Le travail de thèse devrait se décomposer de la façon suivante :

- Année 1 : Bibliographie, analyses des données et prise en main du modèle de « Transformer ».
- Année 2 : Développement et simulation du modèle sur des bases de données spécifiques. Par exemple, « The European Epileptic Dataset » qui fait partie du projet « EPILESTIAE » financé par l'union européenne. Ce Jeu de données est volumineux contenant des signaux EEG enregistrés sur 300 sujets âgés entre 13 et 67 ans et représentant un large spectre de symptômes épileptiques.
- Année 3 : Rédaction du manuscrit. Valorisation et communication.

Le candidat mènera alternativement ses travaux de recherche dans chacun des établissements partenaires. Celui-ci effectuera des séjours de recherche auprès de chacun des établissements. La durée des périodes de séjour sera annuellement : 6 mois à l'UL et 6 mois à l'ULCO (voir le tableau ci-dessous), en accord

avec les deux établissements, pendant l'application de la présente convention :

	ULCO	Université Libanaise
2025-2026	1 Janvier 2026 – 30 Juin 2026	1 Septembre 2025 – 30 Décembre 2025
2026-2027	1 Janvier 2027 – 30 Juin 2027	1 Septembre 2026– 30 Décembre 2026
2027-2028	1 Janvier 2028 – 30 Juin 2028	1 Septembre 2027– 30 Décembre 2027

d) Les collaborations prévues

Ces travaux seront effectués en étroite collaboration avec l'Université du Liban mais également avec l'université d'Ottawa, en particulier avec le Pr. Cousineau avec qui nous avons initié une collaboration à travers un projet du Bonus Qualité International (BQI) en 2025. Ce projet est basé sur la l'apprentissage et la classification des signaux EEG à partir de stimuli visuels et entre directement dans le cadre de cette thèse. Le Pr. Cousineau sera associé à la démarche afin d'envisager le transfert des méthodologies développées dans cette thèse à leurs problématiques.

e) Publications

- [1] H.Hijazi, O.Bazzi, A.Bigand : " A new nonlinear discriminant analysis algorithm using a combined version of LDA and LLE", Congrès "ICPV2011 ", pp. 106 à 109, 2011, Las Vegas, USA.
- [2] Ali Darwich, *Pierre-Alexandre Hébert*, André Bigand, Yasser Mohanna : "*Background Subtraction Based on a New Fuzzy Mixture of Gaussians for Moving Object Detection*". J. Imaging 4(7): 92, 2018
- [3] Vaswani, Ashish; Shazeer, Noam; Parmar, Niki; Uszkoreit, Jakob; Jones, Llion; Gomez, Aidan N; Kaiser, Łukasz; Polosukhin, Illia . "Attention is All you Need" . *Advances in Neural Information Processing Systems*. **30**, 2025.
- [4] Vafaei, Elnaz and Mohammad Saleh Khajeh Hosseini. "Transformers in EEG Analysis: A Review of Architectures and Applications in Motor Imagery, Seizure, and Emotion Classification." In *Sensors* 25, 2025

