

Détection de crise d'épilepsie en temps réel au moyen des signaux de mouvements et de la reconnaissance d'activité

F. Peigné (1)(2), L. Dumont (1), C. Le Couedic (1), D. Rigal (2), A. Larbi(2), P. Latour, J. Petit (3)
 (1) Association Aura - Paris (France), (2) L3S, (3) Institut la Teppe - Tain L'Hermitage (France)



INTRODUCTION

Le suivi des crises d'épilepsie est crucial pour la surveillance de l'épilepsie, l'évaluation du traitement et la possibilité d'évaluer si la survenue des crises d'épilepsie correspond à un chronotype particulier.

Ce dernier point permettrait la réalisation d'un agenda de crise pour les personnes épileptiques. Actuellement, les études montrent que les agendas de crises d'épilepsie reposant sur les seules personnes épileptiques ou sur les proches aidants ne sont pas totalement fiables, avec en particulier de nombreux oublis des crises nocturnes.

Les dispositifs portables peuvent être mieux tolérés et mieux adaptés à une évaluation précise de la fréquence et la temporalité de crises d'épilepsie à long terme.

Nous proposons dans ce document de partager la problématique et les pistes d'amélioration de la détection de crise d'épilepsie à l'aide de modèles de reconnaissance d'activité par le mouvement.

Cette étude fait partie d'un projet global (DetecTeppe) dont l'ambition est d'obtenir un dispositif offrant :

- Un système de détection de crise en temps réel.
- Basée sur une analyse multimodale biométrique (i.e. fréquence cardiaque et du mouvement) des crises d'épilepsies.
- Interactif pour le patient.

La double compétence médicale et technique entre l'Institut la Teppe, ses partenaires et l'association Aura a été mise à profit pour réaliser cette étude, un ensemble d'outils open-source permettant de travailler sur ces données est mis à disposition.

MÉTHODE

Objectif

Établir un agenda personnalisé de suivi des crises d'épilepsie pour améliorer la prise en charge des personnes épileptiques par les professionnels de santé.

Baseline et entraînement des modèles

Pour détecter ces crises d'épilepsie, nous avons développé des modèles par apprentissage statistiques qui reconnaissent des motifs de mouvement spécifiques à la crise.

En 2010 Guillaume Becq a publié un modèle Baseline de détection de crise Tonic Cloniques entraîné sur un sous ensemble de 30 crises, et s'appuyant sur une feature unique (Entropie) [1]. Sensibilité : 0.80, Spécificité temporelle : 0.95 Bénéficiant de la démocratisation des outils d'intelligence artificielle et de l'augmentation de la puissance de calcul CPU, nous avons poursuivi ces travaux en entraînant 4 nouveaux modèles de Machine Learning sur un ensemble de features plus large.

Traitement du signal

Compte tenu de l'absence d'information sur l'orientation du capteur nous avons choisi de travailler sur la norme euclidienne du signal. MSG et MSD désignent les membres supérieurs gauche et droit. Le capteur tête a été retiré en raison de la trop grande variance des données (manque de fiabilité).

Nous avons découpé le signal en fenêtres de 15 secondes avec overlap de 50%.

Nous avons calculé et utilisé les features suivantes :

Nom feature	Interprétabilité
Std_MSG	Ecart type signal
Max_MSG	Max signal
Mean_absolute_MSG	Moyenne de la Valeur absolue
MSG_F_5-9, 9-13, 13-23	Spectre entre 5-9Hz etc.
std range crossing rate	Nombre d'aller retour en dehors du range [écart-type,écart-type]
Ent_MSG	Entropie pour MSG
Freq_max_MSG	Fréquence du pic d'activité MSG

Train/test split

On écarte à chaque fois du jeu d'apprentissage la personne que l'on cherche à prédire et on ne garde que les crises concernées par le label pour l'apprentissage ("Leave one patient out"). Cela pose parfois le problème du très faible nombre de crises pour l'apprentissage (labels N et A).

Métriques d'évaluation

Nous avons utilisé les métriques de référence requises par la FDA pour la certification médicale des systèmes de crise d'épilepsie :

- Taux de fausse alertes sur 24 heures (< 2 / jours)
- Sensibilité de détection des crises > 70 %
- Spécificité sur des fenêtres temporelles

PRÉSENTATION DE LA BASE DE DONNÉE CAPAMETRIM

Basée sur l'étude réalisée par l'hôpital de Grenoble entre 2005 et l'année 2010 sous la direction de Guillaume Becq dans le cadre de ses recherches sur l'épilepsie au terme du projet « Epimouv », rassemblant 3 bases de données différentes.

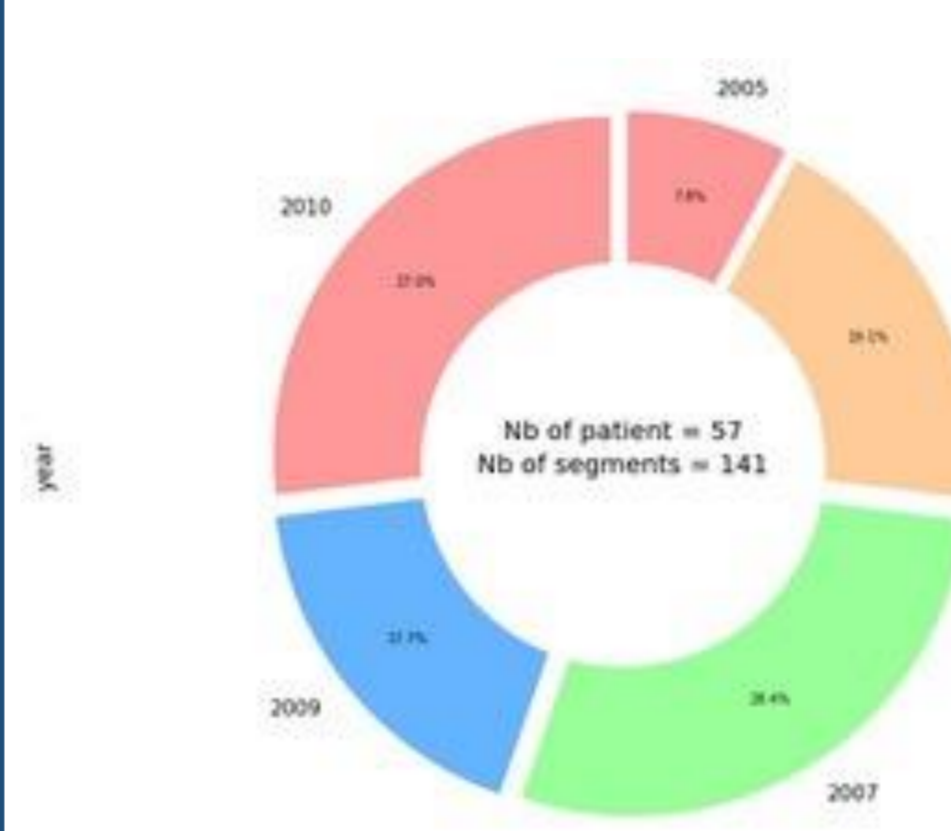


Fig 1: Provenance du signal

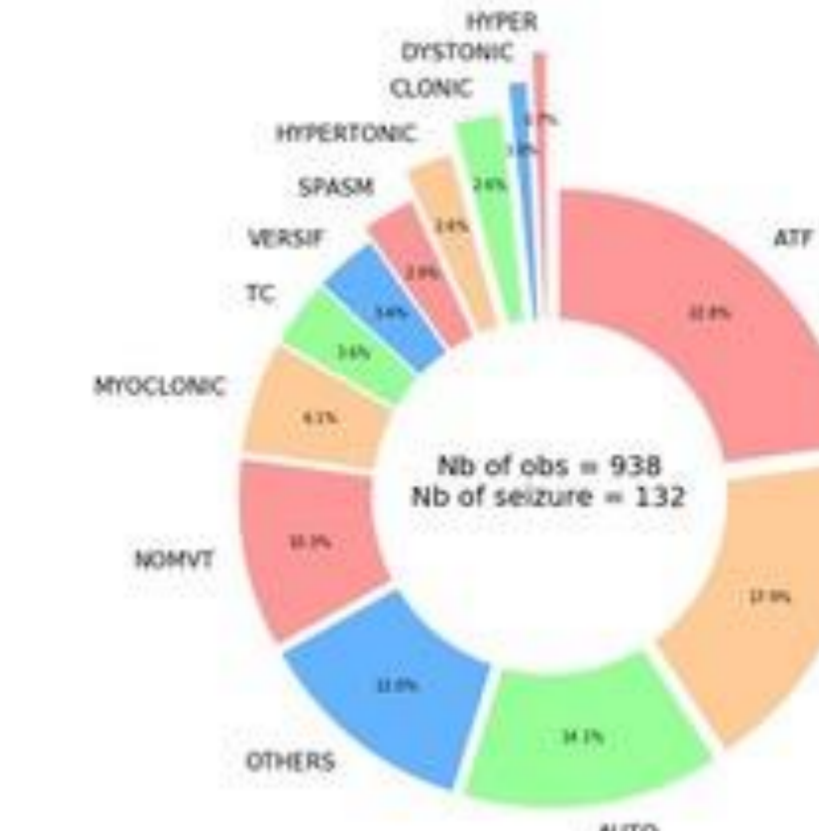


Fig 2: Répartition des crises et labels

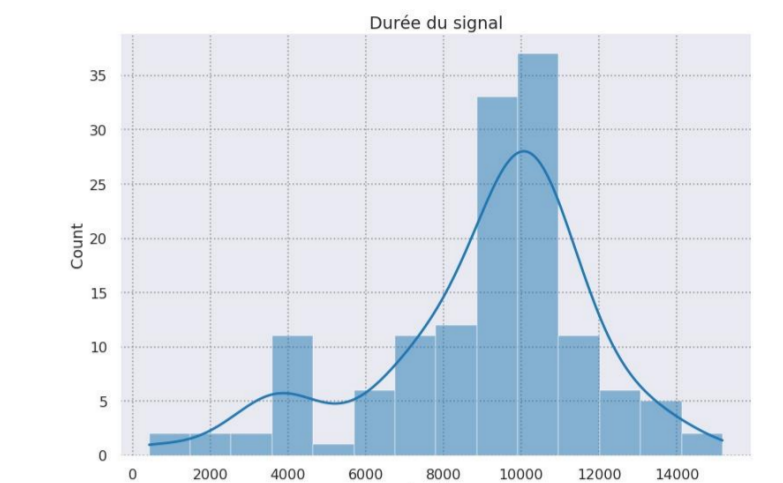


Fig 3: Durée par examen (secondes)



Fig 4: Durée par crise (secondes)



Acquisition / Capteurs

Acquisition faite avec trois capteurs positionnés sur respectivement :

1. poignet gauche (MSG)
2. poignet droit (MSD)
3. tête (T)

Portés par des patients allongés et sous vidéosurveillance.

En raison d'un manque de fiabilité des données, le capteur tête n'a pas été exploité. Chaque capteur mesure la gravité selon 3 axes x,y,z.



Nettoyage des données

Les signaux sont écartés si :

- Leur norme euclidienne moyenne ne se situe pas entre 0.97g et 1.03g.
- Inconsistance dans les labels temporels
- Datas manquantes
- Suppression du capteur tête (inconsistance du signal)

Sur 57 patients et 141 crises, 25 patients contenant 60 crises sont finalement retenus après ces filtres de qualité.



Labelling des données

Les données de crises ont été analysées par le Dr Latour.

L'analyse a consisté à identifier la séméiologie la plus longue entre t0 et t2 (respectivement début et fin d'observation de la crise) sur la base des annotations réalisées dans l'étude EPIMOUV. Les labels finaux utilisés sont N, A, T, TC, Auto.

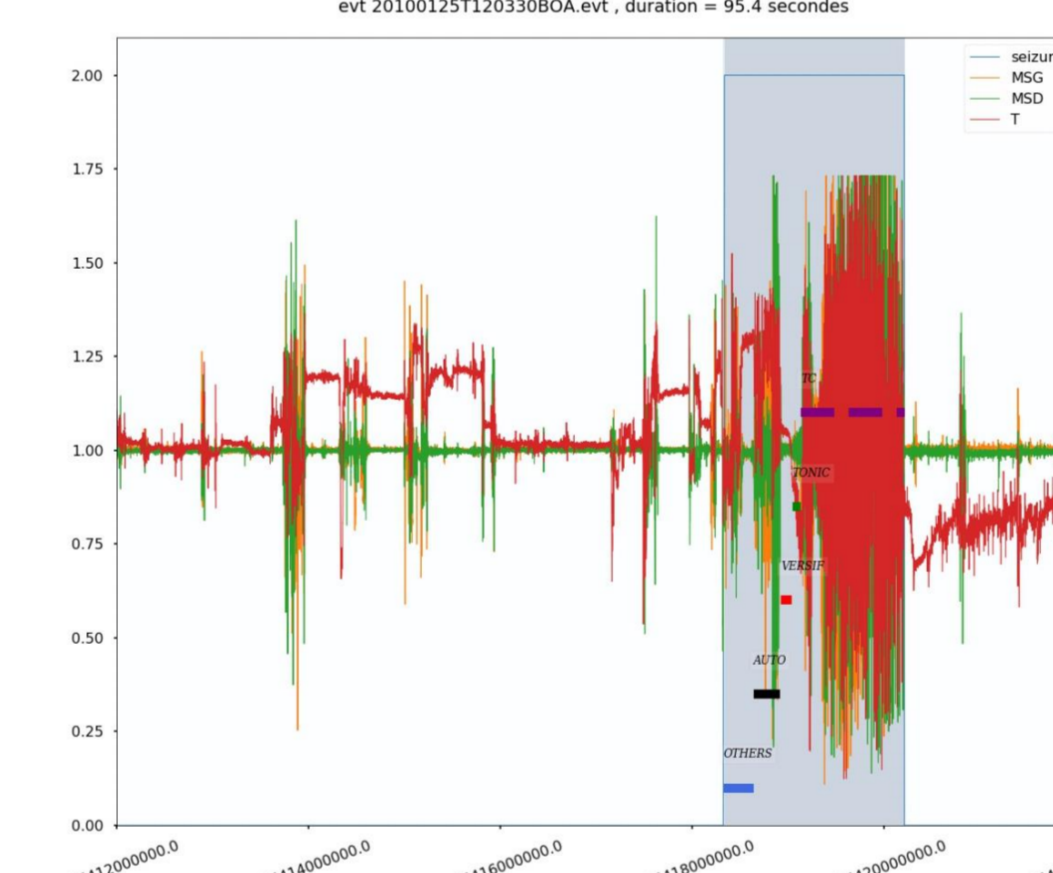


Fig 5: Exemple de crise

Classe utilis	Sous-classe	Manifestations
N = NOMVT		Pas de mouvement
A = AUTO		Automatismes
T = TONIC	V VERSIF HT	Manifestations toniques; les mouvements versifs et hypertoniques sont inclus dans cette classe
TC = TONICO CLONIC	C CLONIC	Les mouvements cloniques sont inclus dans cette classe
H, HM = HYPER		Mouvements hypermoteurs
O = OTHERS	ATF	Autres manifestations motrices non pathologiques ou complexes; les artéfacts sont inclus dans cette classe
NaN		Non scorés

Tab. 1 : Liste des classe et sous classe de crises

RÉSULTATS / CONCLUSION

Models / seizure type	Hyperparameters	All crises (60)				TC : 11				T : 23				O : 15				N : 5				A : 5			
		sens	spec	sz detc	fp 24h	sens	spec	sz detc	FP24	sens	spec	sz detc	FP24	sens	spec	sz detc	FP24	sens	spec	sz detc	FP24	sens	spec	sz detc	FP24
Forêt aléatoire	max_depth = 4, n_estimators = 30, max_features = 10	0,55	0,993	33	40	0,73	0,996	8	24	0,26	0,995	6	29	0,87	0,975	13	145	0,0	0,959	0	236	0,80	0,936	4	371
Régression logistique		0,67	0,986	40	80	0,73	0,993	8	39	0,35	0,981	8	107	0,80	0,970	12	171	0,0	0,938	0	358	0,80	0,986	4	81
Kmeans	2 clusters	1,00	0,470	60	3048	1,00	0,690	11	1786	0,83	0,468	19	3067	0,93	0,430	14	3282	0,8	0,695	4	1755	1,00	0,336	5	3823
Xgboost	booster=gbtree, max_depth=5, eta=0.2	0,77	0,960	46	230	0,82	0,997	9	17	0,39	0,980	9	115	0,87	0,974	13	150	0,0	0,980	0	115	0,80	0,989	4	63

Tab. 2 : Evaluation de la performance de 4 algorithmes de ML selon les types de crise. sens (sensibilité temporelle sur fenêtre de 15 seconds), spec (spécificité temporelle sur fenêtre de 15 seconds), sz detc (nb de crise détectée sur au moins une fenêtre), FP24 (Taux de fausse alerte sur 24 heures)

Conclusion

- Gain en spécificité avec perte de sensibilité sur les crises Tonic Clonique comparé au modèle baseline.
- L'augmentation du nombre de feature rend le modèle plus robuste
- Le découpage train/test par type de crises rend les modèles moins performants qu'un apprentissage sur tous le jeu de données en raison du trop faible nombre de données pour chaque type de crises.
- Très grande variance dans la prédictibilité selon le « label » des crises.
- La détection presque systématique des crises tonico-clonique et même d'autres type de crise à définir à l'aide de signaux d'accélérométrie et d'ECG semble un objectif raisonnable, le problème principal est d'éviter un trop grand nombre de faux positifs. De plus les modèles permettront à minima une « surveillance statistique ».

limites

- Les signaux, acquis sur des personnes allongées, ne correspondent pas à ce que l'on peut observer en situation réelle. Pour robustifier ces modèles dans une utilisation courante, il sera important d'utiliser des données de mouvement.
- Limites du matériel d'acquisition d'alors (plafond à 1G), l'utilisation de matériel plus récent permettra d'affiner le signal et donc les prédictions => amélioration du matériel d'acquisition lors de la prochaine étude clinique.

3 modèles généralement performants sur la classification binaire de séries temporelles ont été entraînés : forêt aléatoire, régression logistique et Xgboost. Kmeans avait pour objet d'observer si un "cluster de crise" se forme spontanément. Les 3 modèles donnent des résultats assez similaires le meilleur étant Xgboost. Les crises donnant les meilleurs résultats de prédiction sont les crises labellisées tonico-clonique mais ces résultats sont insuffisants à ce jour pour prétendre au statut d'instrument médical, ils produisent trop de faux positifs.

Analyses des features



- Le feature le plus discriminant du meilleur modèle est std range crossing rate, c'est une mesure synthétique de la fréquence du tremblement.
- On observe une variance importante de l'utilisation des features par la forêt aléatoire, c'est compréhensible car l'écart-type, la norme L1, l'entropie, mesurent tous un niveau global d'activité.
- Le spectre apporte un véritable surcroît d'information au modèle.

Pistes d'amélioration

- Le mouvement ne traduisant pas totalement la présence de crise, coupler ce signal au signal ECG qui donne déjà individuellement de bons résultats permettra d'augmenter la qualité des prédictions.
- La superposition d'un modèle de reconnaissance d'activité devrait permettre d'écarter les trop nombreux faux positifs en identifiant les activités concernées.
- Une labélisation ad hoc des crises « prédictibles » est envisagée.
- L'utilisation du deep learning pourrait permettre de générer des features plus pertinentes. Ces approches demandent cependant beaucoup de données.
- Pour assurer la reproductibilité de nos résultats, l'ensemble de la chaîne de traitement et entraînement est disponible sur notre GitHub en open source : [3]