
Segmentation multi-labels de l'EEG néonatale pour l'assistance à l'interprétation

Vivien Kraus^{*1}, Guillaume Dollé², Alexandra Givernaud³, Nathalie Kamgnia³, Gauthier Loron^{1,3}, Jonathan Beck³, Nathalie Bednarek^{1,3}, François Rousseau⁴, and Nicolas Passat¹

¹Université de Reims Champagne Ardenne, CRESTIC, Reims, France – Université de Reims Champagne-Ardenne – France

²Université de Reims Champagne Ardenne, CNRS, LMR, UMR 9008, Reims, France – Université de Reims Champagne-Ardenne, Centre National de la Recherche Scientifique – France

³Service de médecine néonatale et réanimation pédiatrique, CHU de Reims, France – CHU Reims – France

⁴IMT Atlantique, LaTIM U1101 INSERM, UBL, Brest, France – IMT Atlantique – France

Résumé

La recherche clinique concernant le diagnostic de l'encéphalopathie anoxo-ischémique peut s'appuyer sur l'interprétation de l'électroencéphalographie (EEG) néonatale. La présence et la fréquence d'apparition de figures physiologiques ou pathologiques dans le signal joue un rôle déterminant, aussi bien pour le diagnostic que pour le suivi du patient. La recherche de ces figures dans le signal constitue une part importante du temps consacré à son interprétation. En effet, cela revient à chercher des motifs de l'ordre de la seconde dans un signal de plusieurs heures.

La description de certaines de ces figures est connue dans la littérature, en termes d'amplitude, localisation, durée, ou fréquence des ondes à rechercher. Cependant, la précision de cette description ne peut pas remplacer l'expertise d'un électrophysiologiste formé. S'il est possible pour un tel expert de détecter la présence de ces figures, il n'est en revanche pas envisageable d'avoir une information quantitative de leur fréquence d'apparition tout au long du suivi, qui recouvre les premiers jours de vie.

La détection automatique de ces figures constituerait une assistance à l'interprétation du signal EEG, aussi bien pour le diagnostic que pour le suivi. L'apprentissage automatique supervisé est un cadre utile pour la reconnaissance de motifs connus. Il nécessite cependant des données reliant les variations du signal à la détection des figures d'intérêt.

Les données d'EEG de Lytonepal, une cohorte française multi-centres qui regroupe environ 800 patients diagnostiqués comme souffrant d'une encéphalopathie anoxo-ischémique sévère à modérée, sont disponibles pour cette étude ancillaire. Elle contient en particulier 81 tracés enregistrés dans les 6 premières heures de vie. Pour pallier le manque de données quant à la détection de figures, un outil d'annotation a été développé. Il a été utilisé par 3 experts pour faire une annotation de trois figures reconnues importantes dans l'interprétation de l'EEG néonatale dans les 6 premières heures de vie.

*Intervenant

Nous proposons de compléter ces annotations par des données synthétiques simulant les cinq figures les plus précisément décrites dans la littérature de l'EEG néonatal à terme. Les annotations synthétiques sont introduites indépendamment sur les fenêtres ayant la plus grande activité, et déformées légèrement pour augmenter la difficulté de détection automatique. La vraisemblance physiologique n'est pas recherchée.

Nous avons conçu un réseau de neurones profond pour l'apprentissage simultané de la segmentation de ces 8 types de figures. L'utilisation de figures synthétiques renforce l'importance de la forme, l'amplitude, la durée, fréquence et localisation de ces figures, afin que la détection automatique des figures se fonde sur des principes similaires à la lecture humaine. La segmentation consiste en la prédiction, pour chaque fenêtre de 10 secondes extraite du signal, pour chaque point, dérivation, et type de figure, d'un score de détection. L'optimisation se fait grâce à Pytorch, pour un CNN U-Net 1-D entraîné sur un accélérateur graphique. Le volume des données nécessite l'utilisation d'un supercalculateur. Une première phase d'optimisation des hyperparamètres cherche à maximiser la performance de détection des figures synthétiques, puis l'effet de ces données générées est étudié en diminuant la fréquence d'introduction par rapport aux données réelles.