

Titre : Apprentissage faiblement supervisé pour la caractérisation des emboles cérébraux à l'aide d'un dispositif de Doppler transcrânien (TCD).

Résumé :

Cette thèse porte sur la classification et la caractérisation des signaux transitoires de haute intensité (HITS) provenant d'appareils portables de Doppler transcrânien (TCD). L'objectif principal est d'aider les cliniciens à différencier les emboles solides (ES) et gazeux (EG) des artefacts (Art.) générés pendant les sessions de surveillance TCD. Les emboles sont des particules qui peuvent circuler dans les artères cérébrales, parfois les obstruer et provoquer un accident vasculaire cérébral. L'identification et la classification des HITS entre ES, EG et Art. ne sont pas évidentes et requièrent d'importantes connaissances spécialisées. De ce fait, même les cliniciens peuvent avoir des difficultés à différencier les HITS, ce qui complique le traitement des patients. Par conséquent, sa détection, sa classification et sa caractérisation sont des facteurs clés pour améliorer la prise en charge des patients dans les centres de santé.

Dans ce travail, nous proposons des modèles d'apprentissage profond capables d'effectuer une classification précise entre les ES, EG et Art., avec une mémoire et une consommation d'énergie limitées. Notre approche combine l'annotation semi-automatique des données, l'apprentissage multi-représentation et des techniques de compression de modèles. Nous évaluons nos différentes contributions en utilisant plusieurs ensembles de données médicales in vivo, en plus de la classification des HITS.

Notre méthode s'est avérée efficace, avec d'excellents résultats de classification et une faible consommation de mémoire et d'énergie pour plusieurs tâches de classification de signaux médicaux. En particulier, pour la classification des HITS, à notre connaissance, notre travail est le seul à proposer une classification in vivo des HITS issus de TCD portables entre ES, EG et Art. Nos principales contributions sont les suivantes. Premièrement, nous avons proposé une méthode d'annotation semi-automatique des données basée sur des mesures de qualité locales avec une erreur d'annotation contrôlée, permettant d'étiqueter rapidement un grand ensemble de données en utilisant un petit nombre d'échantillons étiquetés. Deuxièmement, nous proposons un modèle de classification multi-représentation guidé et régularisé permettant de classer avec précision les HITS, en tirant simultanément parti du signal Doppler et de sa représentation temps-fréquence. Enfin, nous avons proposé de nouvelles techniques de compression de modèles basées sur l'élagage et la quantification extrême, permettant de réduire les besoins en mémoire et énergie des modèles entraînés.

Enfin, en travaillant en collaboration avec Atys Medical, nous avons pu incorporer nos modèles dans leur logiciel de gestion des données. Nous espérons que les méthodes développées dans ce travail pourront aider les cliniciens dans la prise en charge de leurs patients.

Title: Weakly-supervised learning for emboli characterization with Transcranial Doppler (TCD) monitoring.

Abstract:

This thesis focuses on the classification and characterization of high intensity transient signals coming from portable transcranial Doppler (TCD) ultrasound devices. The main objective is to help clinicians identify solid and gaseous emboli from artifacts generated during TCD monitoring sessions. In fact, emboli are solid or gaseous particles that can circulate in the cerebral arteries, sometimes blocking them and causing ischemic stroke. However, the identification and classification of HITS between solid embolus, gaseous embolus, and artifacts is not evident and require important expert knowledge. Because of this, even clinicians can have some trouble differentiating these different types of HITS, which makes the treatment of patients difficult. Therefore, its detection, classification and characterization are key factors to improve patient management in healthcare centers.

In this work, we propose deep learning models capable of doing an accurate classification between solid embolus, gaseous embolus, and artifacts, with limited memory and energy consumption. Our approach combines semi-automatic data annotation, with multi-feature learning and model compression techniques. We evaluate the different components of our approach using several medical in vivo datasets, besides HITS classification.

Our method proved to be effective, with great classification results, and low memory and energy consumption on several medical signal classification tasks. More precisely, for HITS classification, to our knowledge, our work is the only one proposing an in vivo classification of portable TCD HITS between solid embolus, gaseous embolus, and artifacts.

The main contributions of this work are the following. Firstly, we proposed a semi-automatic data annotation method based on local quality metrics with controlled annotation error, allowing to quickly label a large dataset, using a small number of labeled samples. Secondly, we propose a hybrid guided and regularized multi-feature classification model allowing to accurately classify HITS, simultaneously taking advantage of the raw Doppler signal, and its time-frequency representation.

Finally, we proposed new model compression techniques based on pruning and extreme quantization, allowing to reduce the memory requirements of the trained models, as well as the energy consumption. Finally, as we worked in close cooperation with Atys Medical, manufacturer of portable TCD devices, we were able to incorporate our developed models into their data management software. Even though validation is still needed, we hope that the models and methods developed in this work can help clinicians with their patient management.