Contribution ID: **36** Type: **Oral Presentation**

Machine Learning Pipelines on Medical Imaging

Tuesday, September 24, 2019 2:30 PM (15 minutes)

The use of Artificial Intelligence (AI) over medical data allows the extraction of features associated to the disease from medical images using data-characterisation and modelling algorithms. The use of advanced machine learning algorithms is changing the way image processing is performed, evolving from analytic solutions to models built up with supervised training techniques working in complex Convolutional Neural Network (CNN) architectures. However, advanced AI techniques require a deep understanding of the behaviour of the model and non-trivial programming skills. This limits the application of AI to researchers who have a deep understanding of the medical problem but lack from those specific technical skills.
In this work, we will compare an application for the automatic diagnosis of Rheumatic Heart Disease (RHD) from echocardio videos on children, implemented using Keras with an equivalent application deployed using a machine learning workflow system (LEMONADE).
The processing pipeline requires 7 steps: frame splitting, which splits a video into frames; automatic classification into doppler and anatomical images by color inspection (only doppler images are used during the rest of the pipeline); color-based segmentation through k-means clustering; image preprocessing and view classification by using a CNN; first- and second-order texture analysis and blood-flow velocity calculation; z-score features normalization; and classification of extracted features through machine learning techniques into RHD positive or healthy studies.
The implementation in Keras uses the pre-trained models for the classification of the views within the estimation of the RHD. All the components are delivered as containers, facilitating their distribution and the integration of new components in LEMONADE.
The processing backend is a Kubernetes cluster provided of GPU nodes attached through PCI passthrough to the Virtual Machines and the containers. This way there is no penalty on the usage of the GPUs from the applications. Data are stored directly on a persistent storage object exported through an SSH server. As communications are encrypted, data access is measured separately.

Primary authors: Dr DOS SANTOS, Walter (Universidade Federal de Minas Gerais); Prof. MEIRA JR., Wagner (Universidade Federal de Minas Gerais); Mr CAMACHO-RAMOS, Eduardo (QUIBIM); Ms JIMÉNEZ-PASTOR, Ana (QUIBIM); Dr ALBERICH-BAYARRI, Ángel (QUIBIM); BLANQUER ESPERT, Ignacio (Universitat Politècnica de València)

Presenter: BLANQUER ESPERT, Ignacio (Universitat Politècnica de València)

Session Classification: IBERGRID Contributions

Track Classification: Development of Innovative Software Services