

B. Sc. level internship - 2021 for 2 months

Subject	GPU optimisation of Light Gated Recurrent Units in the context of End-to-End Automatic Speech Recognition
Advisor	Titouan Parcollet, Assoc. Prof.
Location	Avignon Université, LIA, France
Context	This internship is strongly linked to SpeechBrain. An international and collaborative project initiated by the LIA and the Mila (CA).
Details	<p>Light Gated Recurrent Units (LiGRU)[1] have been designed to carefully address the task of Automatic Speech Recognition (ASR) within the DNN-HMM framework. In this very specific context, LiGRU outperformed traditional RNN such as GRU and LSTM by an important margin both in terms of concepts error rate (i.e. words, phonemes, characters) over various datasets [2] and computation efficiency. Indeed, a LiGRU cell relies on the combination of a single memory state combined with ReLU activation functions and batch normalisation leading to a decrease of the number of neural parameters by a factor of 3 in comparison with a LSTM cell. More recently, LiGRU have also reached state-of-the-art performance with end-to-end ASR [3].</p> <p>Unfortunately, the spread of LiGRU across the research community remains limited due to a very practical limitation : LSTM and GRU are commonly implemented in CuDNN within deep learning frameworks (e.g. PyTorch and Tensorflow) leading to a drastic training speed improvement. Therefore, and despite being more efficient in terms of FLOPs, the current implementations of the LiGRU (PyTorch-Kaldi and SpeechBrain) remain slower to use than other architectures. During this internship, the candidate is expected to provide a solution to this issue by investigating various implementations of the LiGRU within SpeechBrain [4] with GPU-accelerated technologies : Numba, CuDNN ...</p> <p>On a practical note, all the benchmarks to compare the solution found by the intern to the state-of-the-art across various tasks already exist and will be used, potentially leading to a publication and to an integration as a collaborator of the SpeechBrain project.</p> <p>Bibliography: [1] Ravanelli, M., Brakel, P., Omologo, M., & Bengio, Y. (2018). Light gated recurrent units for speech recognition. <i>IEEE Transactions on Emerging Topics in Computational Intelligence</i>. [2]Ravanelli, M., Parcollet, T., & Bengio, Y. (2019, May). The pytorch-kaldi speech recognition toolkit. In <i>ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)</i> (pp. 6465-6469). IEEE. [3]Parcollet, T., Morchid, M., & Linares, G. (2020, May). E2E-SINCNET: Toward fully end-to-end speech recognition. In <i>ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)</i> (pp. 7714-7718). IEEE. [4] Website : http://www.darnault-parcollet.fr/Parcollet/hiddennohare/speechbrain.github.io/</p>
Keywords	Artificial Intelligence, Recurrent Neural Networks, Speech Recognition.

Calendrier prévisionnel

La période de stage est un intervalle continu. Il est important de mentionner que dès la date de fermeture de l'université, le stagiaire sera tenu de travailler à distance. Les conditions de présence du stagiaire au sein du LIA dépendront du contexte sanitaire avec la mise en place du télétravail si nécessaire.