IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Post-Doc positions at IMT Atlantique : Learning and Data Reconstruction over Coded Data

2 post-doc positions for 12 months, with possible extension to 18 months, at IMT Atlantique, Brest, France. Expected starting date: spring 2023

Contacts:
- Elsa Dupraz (elsa.dupraz@imt-atlantique.fr),
- François-Xavier Socheleau (fx.socheleau@imt-atlantique.fr)

Motivation

Every minute, 500 hours of video are uploaded on Youtube, and 240,000 images are added on Facebook. There is an absolute need to rely on advanced Machine Learning methods so as to sort, organize, and recommend the content to users. But as a crucial step, the data should first be transmitted from users to a server for further storage and processing. The conventional communication framework assumes that the data should be completely reconstructed, even with some distortions, by the server. Instead, **this project aims to develop a novel communication paradigm by considering the learning performance as one of the key criterion for the design of the communication system**.

We consider a setup in which a coded bitstream is transmitted to the server through a noisy channel. From the received bitstream, the server may either reconstruct the data with a certain distortion, or perform learning (image classification, content recognition, etc.) without prior data reconstruction. **The design objective for the source/channel coding scheme is to reduce as much as possible the coding rate in order to address both distortion and learning performance criterion**. Several recent works show that there is a fundamental tradeoff between data reconstruction and learning in terms of coding rate [1].

Challenges

Two main approaches have been investigated in the literature of learning over coded data. **The first approach considers conventional source-channel coding techniques dedicated to data reconstruction**, and aims to apply the learning task directly onto the coded data, possibly after some partial decoding [2]. This approach is especially relevant in our context since it allows to perform both data reconstruction and learning from the same coded data. However, the coding scheme is optimized for data reconstruction only, which may lower the learning performance.

T**he second approach consists of replacing encoders and decoders by full end-to-end Deep Learning techniques** such as Variational Auto-Encoders (VAE), specifically trained for the considered learning task [3]. In our context, this would require to train the VAE so that one or the other task (data reconstruction or learning) can be applied onto the same received data. As a main drawback of this approach, in some applications, it can be difficult to obtain some relevant data to train the system offline prior to implementation.

Candidates Profile

To address these issues, we are seeking for two post-docs:
1) One post-doc to work on conventional Source/Channel coding approaches, in which learning is only applied at the server side. For this position, the candidate is expected to have a background on source and/or channel coding.
2) One post-doc to work on full end-to-end Deep-Learning-based approaches for this problem. For this position, the candidate is expected to have a background on Deep-Learning

In both cases, the developed solutions will be evaluated first onto standard Machine Learning datasets, and then on data issued from underwater acoustic sensors. The later application is challenging because of the few amount of data available for training, and because of its unconventional channel model.

Both positions are for 12 months, and can be extended to 18 months upon common interest. The retained candidates will be expected to collaborate together and with the other project members.

How to apply

Application should contain: a full CV including publication record and contact details of one or two referees (former PhD advisor, etc.).
No need to include a motivation letter, but please explain in a few words in the e-mail why you apply for the position, and why you think your background is relevant.

Please send your application to Elsa Dupraz (elsa.dupraz@imt-atlantique.fr) and François-Xavier Socheleau (fx.socheleau@imt-atlantique.fr).

References

[1] Y. Blau and T. Michaeli. Rethinking lossy compression: The rate-distortion-perception tradeoff. In International Conference on Machine Learning (PMLR), pp. 675-685, 2019.
[2] M. Ehrlich and L. S. Davis, “Deep residual learning in the jpeg transform domain,” in Proceedings of the IEEE International Conference on Computer Vision, pp. 3484–3493, 2019.
[3] M. Jankowski, D. Gündüz, and K. Mikolajczyk, “Wireless image retrieval at the edge,” IEEE Journal on Selected Areas in Communications, vol. 39, no. 1, pp. 89–100, 2020.