Intern, REQ24-947
Fast training for Hybrid Implicit Neural Networks (INR)

About InterDigital

InterDigital is a global research and development company focused primarily on wireless, video, artificial intelligence ("AI"), and related technologies. We design and develop foundational technologies that enable connected, immersive experiences in a broad range of communications and entertainment products and services. We license our innovations worldwide to companies providing such products and services, including makers of wireless communications devices, consumer electronics, IoT devices, cars and other motor vehicles, and providers of cloud-based services such as video streaming. As a leader in wireless technology, our engineers have designed and developed a wide range of innovations that are used in wireless products and networks, from the earliest digital cellular systems to 5G and today's most advanced Wi-Fi technologies. We are also a leader in video processing and video encoding/decoding technology, with a significant AI research effort that intersects with both wireless and video technologies. Founded in 1972, InterDigital is listed on Nasdaq.

InterDigital is a registered trademark of InterDigital, Inc.

For more information, visit: www.interdigital.com.

Summary

INR, standing for Implicit Neural Representations, has recently been applied to image and video coding. INR are neural networks (e.g. MLPs) that estimate a function representing a given signal continuously. The training, based on discretely represented samples of the signal, consists in overfitting the neural network. Hybrid INR networks, extend INR by mapping latent variables (or a feature vector) to data coordinates. These latent variables are then further used as input for the neural network. In this internship, the objective is to propose a fast-encoding scheme to reduce as much as possible the encoding time (i.e. the training). A particular focus will be to constrain latent variables to be as compact as possible.

References:
- E. Dupont, A. Goliński, M. Alizadeh, Y. W. Teh, and A. Doucet, "COIN: COmpression with Implicit Neural representations," *arXiv:2103.03123 [cs, eess]*, Apr. 2021.
- H. Chen, B. He, H. Wang, Y. Ren, S. N. Lim, and A. Shrivastava, "NeRV: Neural Representations for Videos," in *Neurips* 2021
- T. Müller, A. Evans, C. Schied, and A. Keller, "Instant Neural Graphics Primitives with a Multiresolution Hash Encoding," *ACM Trans. Graph.*, vol. 41, no. 4, pp. 1–15, Jul. 2022.
- T. Ladune, P. Philippe, F. Henry, and G. Clare, "COOL-CHIC: Coordinate-based Low Complexity Hierarchical Image Codec," Dec. 11, 2022, *arXiv*: arXiv:2212.05458.

Duration: 5-6 months, starting January-April 2025

Responsibilities

- State-of-the-art and analysis of advantages/problems
- Implementation and documentation of a possible solution
- Evaluation and reporting of results

Qualifications

- Education: M2 Research

- Skills: machine/deep learning/AI, computer vision, python
- Some experience in pytorch would be appreciated.

Keywords:

- Implicit neural representation
- Computer vision
- Image/video coding

Expected Outcomes:
- Implementation and results of a fast-encoding scheme
- If results and time allow paper submission / patent filing.

Location: Rennes, France

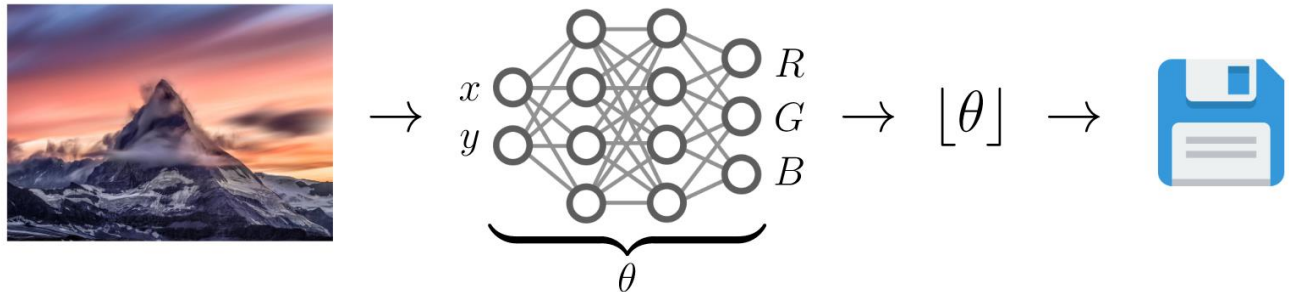**Mentors:** Anne Lambert, François Schnitzer

Image from E. Dupont, A. Goliński, M. Alizadeh, Y. W. Teh, and A. Doucet, "COIN: COmpression with Implicit Neural representations," *arXiv:2103.03123 [cs, eess]*, Apr. 2021.