Intern, REQ24-944
New Image Representations based on invertible Neural Networks (NN)

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.

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

Summary

Implicit neural representations (INR) [1,2,3,4] have recently emerged as a new deep learning-based tool to model and represent various types of signals, including images and videos. For image and video compression, these implicit representations can pave the way towards a radically new paradigm: a neural network overfitted on an image to be encoded, transmitting the image amounts to transmitting the weights of the network, and the decoding is a straightforward evaluation of the neural network for each decoded pixel. Hybrid INR models [3,5,6] are a variant of INR where the pixel coordinates are first mapped to a latent code (or a feature vector). These latent codes are then further used as input for the neural network.

Invertible Neural Networks (INNs) are a class of neural networks designed to be bijective, meaning they have both a forward and an inverse mapping. This property allows the input to be exactly reconstructed from the output. INNs are particularly useful in applications such as probabilistic modeling, generative modeling, and solving inverse problems [7]. They achieve invertibility through specific architectural designs, such as normalizing flows and invertible residual networks [8].

The Internship proposes to investigate the use of invertible networks for INR-based video compression.

References:

[1] 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.
[2] H. Chen, B. He, H. Wang, Y. Ren, S. N. Lim, and A. Shrivastava, "NeRV: Neural Representations for Videos," in *Neurips* 2021
[3] 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.
[4] Kwan, H.M., Gao, G., Zhang, F., Gower, A. and Bull, D., 2024. Hinerv: Video compression with hierarchical encoding-based neural representation. Advances in Neural Information Processing Systems, 36.
[5] 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.
[6] T. Leguay, T. Ladune, P. Philippe, O. Déforges (2024) Cool-chic video: Learned video coding with 800 parameters. In: 2024 Data Compression Conference (DCC). IEEE, pp 23–32
[7] L. Ardizzone, J. Kruse, S. Wirkert, D. Rahner, E. W. Pellegrini, R. S. Klessen, L. Maier-Hein, C. Rother, and

U. Köthe, "Analyzing Inverse Problems with Invertible Neural Networks," arXiv:1808.04730 [cs.LG], Feb. 2019.
[8] J.-J. Huang and P. L. Dragotti, "Invertible Neural Networks and their Applications," Imperial College London, 2022.

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
- Contribution to paper / patent writing

Qualifications

- Education: M2 Research
- Skills: machine/deep learning, python
- Some experience in pytorch would be appreciated.

Keywords:

- Machine learning
- Deep learning
- Image/video coding
- Implicit neural representation

Expected Outcomes:

- Competitive performance of the developed solution with the state of the art
- If results and time allow paper submission / patent filing.

Location: Rennes, France

Mentors:, François Schnitzer, Olivier Le Meur