Intern, REQ-946
New Image Representations based on 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

Kolmogorov-Arnold network (KAN) [2] have recently emerged as new type of neural network architecture for machine learning and computer vision tasks, which is an alternative to the multilayer perceptron (MLP). In particular, KAN learns the B-spline activation functions directly instead of the learning the linear weights in the MLP's and offers the competitive performance compared to MLP. The KAN could be also employed to parametrize the implicit functions and offers alternative to the MLP based implicit functions [2,3]. Implicit neural representations (INR) have recently emerged as a new tool to model and represent various types of signals, including images [1,4] and videos [5]. In particular for image and video compression, these implicit representations can pave the way towards a radically new paradigm: once 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.

This internship proposes to study the representations of image and videos using KAN based implicit representations, and to explore how these new paradigm helps to improve these implicit representations for image and video compression. More specifically, as the KAN learns the activation functions, and in order to encode the learnable parameters in the bit-stream, it needs to be quantized. The impact due to the quantization error in KAN based INR might be more severe than MLP based INR. Thus, one of the directions of this internship is also to tackle this problem through quantization aware training and other techniques. The internship will also explore entropic coding and sparsification of the activation functions.

References:

[1] Sitzmann, V., Martel, J., Bergman, A., Lindell, D., & Wetzstein, G. (2020). Implicit neural representations with periodic activation functions. Advances in Neural Information Processing Systems, 33, 7462-7473.
[2] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T.Y. and Tegmark, M., 2024. Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*.
[3] Mehrabian, A., Adi, P.M., Heidari, M. and Hacihaliloglu, I., 2024. Implicit Neural Representations with Fourier Kolmogorov-Arnold Networks. *arXiv preprint arXiv:2409.09323*.
[4] Strümpler, Y., Postels, J., Yang, R., Van Gool, L., & Tombari, F. (2022). Implicit neural representations for image compression. European Conference on Computer Vision, 2022.
[5] 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*.

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:

- 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: Bharath Bhushan Damodaran, François Schnitzer

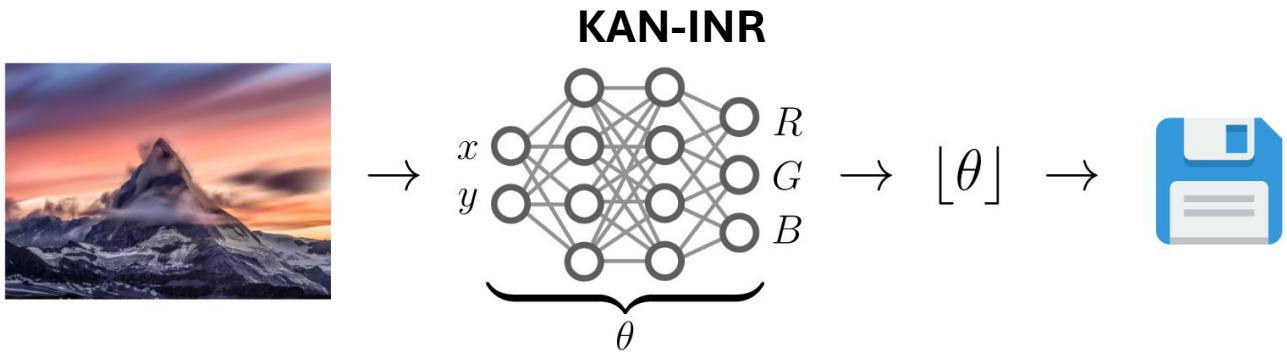

Image from E. Dupont, A. Goliński, M. Alizadeh, Y. W. Teh, and A. Doucet, "COIN: COmpression with Implicit Neural representations," *arXiv:2103.03123 [cs, eess]*, Apr. 2021.

InterDigital is an equal employment opportunity employer. InterDigital will not engage in or tolerate unlawful discrimination with regard to any employment decision, policy or practice based on a person's sex, gender, pregnancy (including childbirth, breastfeeding and related medical conditions), age, race, color, religion, creed, national origin, ancestry, citizenship, military status, veteran status, mental or physical disability, medical condition, genetic information, sexual orientation, gender identity or expression, or any other factor protected by applicable federal, state or local law. This policy applies to all terms and conditions of employment, including, but not limited to, recruiting, hiring, compensation, benefits, training, assignments, evaluations, coaching, promotion, discipline, discharge and layoff.