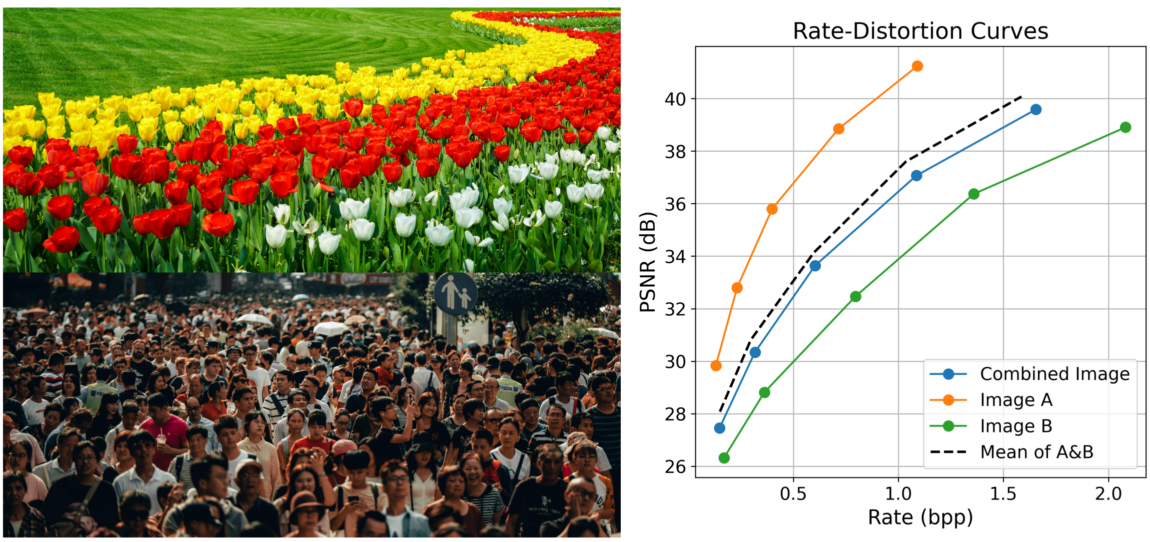
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

Prof. R. Wattenhofer

Cool-Chic MoE: Fast and Expressive Image Decoding

Learned image compression continues to advance by leveraging deep learning models. Although these models improve performance, they are often computationally expensive. Running something like a diffusion model is significantly more costly than using traditional codecs. Cool-Chic [2] addresses this by using compact decoders that bring the computational cost closer to that of conventional codecs. A small decoder is lightweight enough to be sent alongside the latent representation, allowing it to overfit to the distribution of a specific image. It is designed to decode only that image.
While decoding becomes efficient, encoding remains expensive because each image requires fitting both a latent and a dedicated decoder. Recent work [1] attempts to balance this by training an encoder-decoder architecture where the decoder is kept as small as the Cool-Chic decoder. This speeds up encoding significantly, but now the decoder must generalize to the entire natural image distribution.
If we have a general decoder that can work across diverse images, it would be desirable to avoid executing the entire decoder for every image. Instead, we could design a model in which different regions of the image utilize different parts of the decoder. The system would automatically select which components to run based on the image content.



Concatenated input images

Rate-distortion curves

Figure 1: Left: Concatenated images A and B. Right: Rate-distortion (RD) curves for compressing the combined image, compressing only image A or B, and the mean of A and B's RD curves.

Figure 1 illustrates this idea. It shows two images that can either be compressed together using a shared decoder or separately with their own decoders. The average of the individual rate-distortion curves (A and B) is higher than that of the shared decoder, though using separate decoders incurs twice the cost in transmitting decoder parameters. By designing a model that dynamically chooses which parts of a modular decoder to use, we aim to achieve a good trade-off between compression efficiency and flexibility.

Requirements:

Strong programming skills in languages such as Python, and a keen interest in learned compression.

Weekly meetings will be scheduled to address questions, discuss progress, and brainstorm future ideas.

Contact

In a few short sentences, please describe your interest in this project and any relevant coding experience or background (e.g., projects or coursework).

- Till Aczel: `taczel@ethz.ch`, ETZ G60.1

References

[1] Théophile Blard et al. "Overfitted image coding at reduced complexity". In: *2024 32nd European Signal Processing Conference (EUSIPCO)*. IEEE. 2024, pp. 927–931.

[2] Théo Ladune et al. "Cool-chic: Coordinate-based low complexity hierarchical image codec". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023, pp. 13515–13522.