Sandwiched Video Compression: Efficiently Extending the Reach of Standard Codecs with Neural Wrappers

Berivan Isik,¹ Onur G. Guleryuz,² Danhang Tang,² Jonathan Taylor,² Philip A. Chou²
¹Stanford University, ²Google

Neural Video Compression

Pros:
- Flexible – can learn non-linear transforms

Cons:
- Very complex networks (How to predict flows, how to warp the previous reconstructions with the predicted flows, how to do residual compensation)
- Computationally costly
- Slow training and inference
- Memory inefficient

Sandwiched Video Compression Architecture

Loss = Σ_i^T D_i(t) + λ Σ_i^T R_i(t)

Sandwiched Video Compression

● Keeps pros of neural video compression and avoids cons.
● Adds conveniences of a standard codec while repurposing it scenarios it is not designed for.
● Incorporates important parts of a video codec into a carefully-designed differentiable codec proxy.
● Works by message-passing/data-embedding between pre-post with temporally consistent modulation patterns.
● LR-HR scenario: 6.5dB improvements.
● LPIPS scenario: 30% improvements in rate at same quality.
● Slim 57K parameter model obtains 99% reduction in parameters over baseline model with similar results.

The codebase will be open-sourced soon!!!

Also meaningful compression improvements (%5 in rate at same PSNR).

LPIPS akin to robustified E-LPIPS. LPIPS-trained models also obtain %10 rate gains VMAF.

Please ask the presenter to show you clips for all scenarios.