RECOMBINER: Robust and Enhanced Compression with Bayesian Implicit Neural Representations

Jiajun He*[,1], Gergely Flamich*[,1], Zongyu Guo[2], José Miguel Hernández Lobato[1]
*equal contribution, [1]University of Cambridge, [2]University of Science and Technology of China

Paper

TL;DR: We improve the compression performance of Bayesian INRs by using reparameterized weights, learned positional embeddings and hierarchical weight priors.

Motivation

COMBINER:

- Represent data as NN $\hat{\mathbf{y}} = g(\mathbf{x} \mid \mathbf{w})$
- Overfit posterior $q_{\mathbf{w}}$ to data $\mathcal{D}$ using

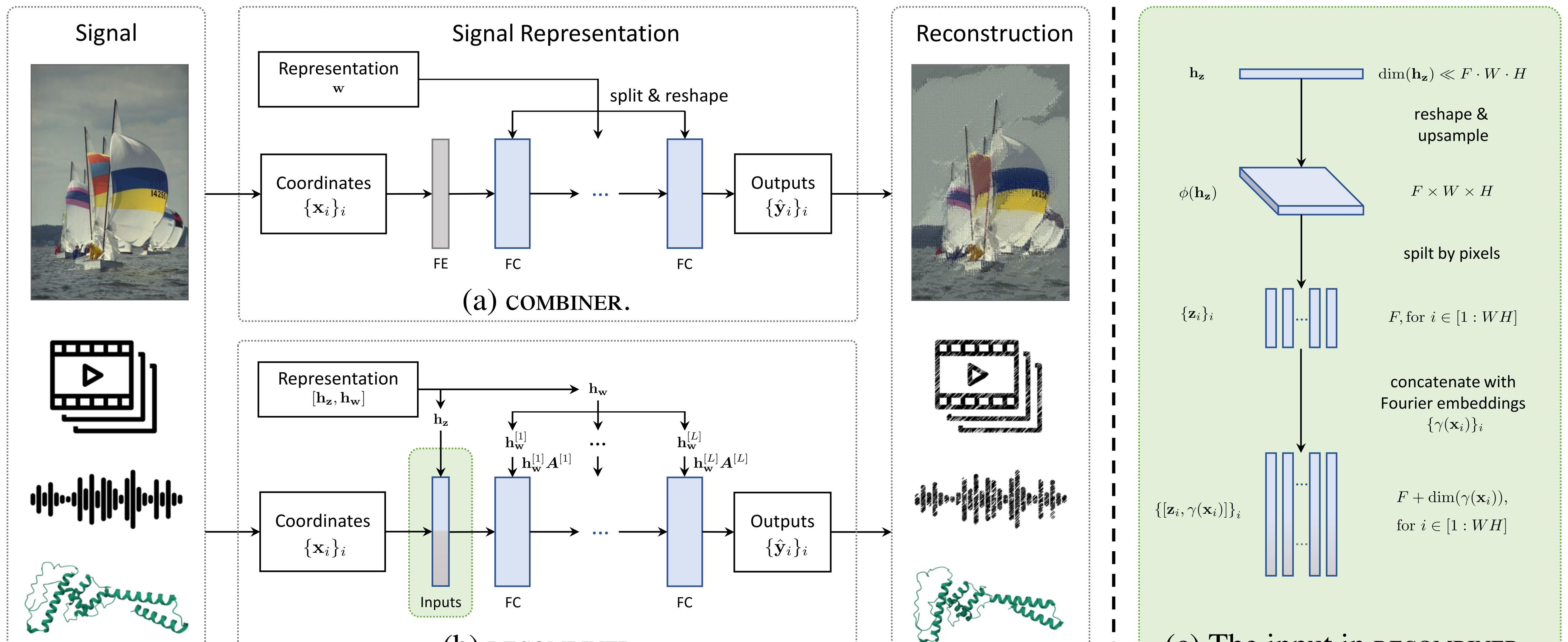
$$\beta D_{\mathrm{KL}}[q_{\mathbf{w}} \| p_{\mathbf{w}}] + \frac{1}{D}\sum_{i=1}^{D} \mathbb{E}\left[\Delta(\mathbf{y}_i, \hat{\mathbf{y}}_i)\right]$$

- Encode a sample $\mathbf{w} \sim q_{\mathbf{w}}$ using REC

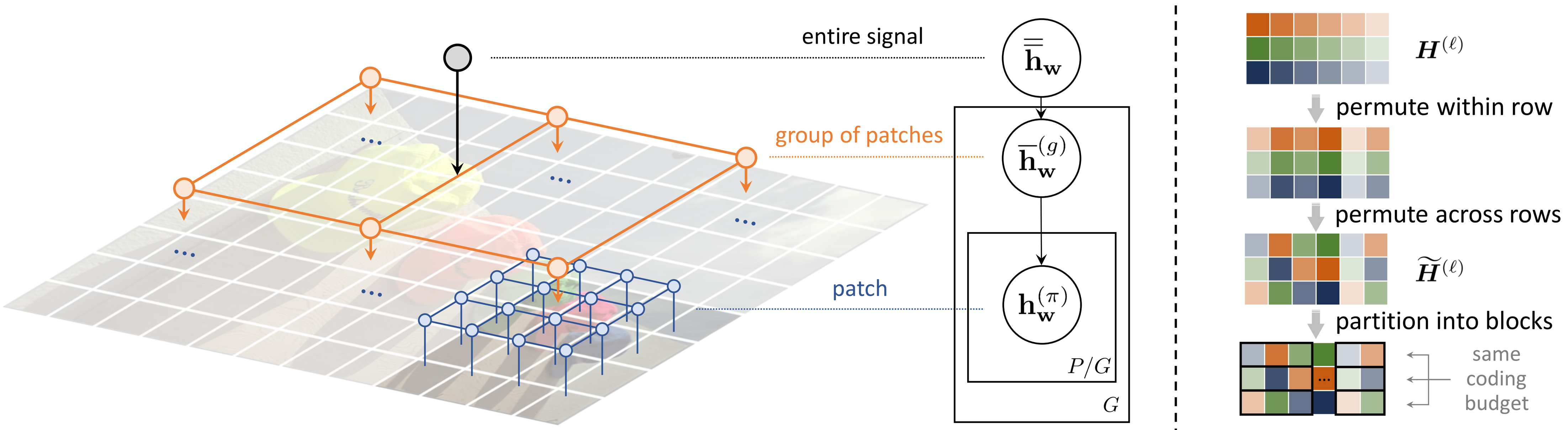
Challenges:

- Overfitting: COMBINER uses fully factorized Gaussian variational posterior $q_{\mathbf{w}}$.
- Stable optimization
- Scaling to high-res data

RECOMBINER

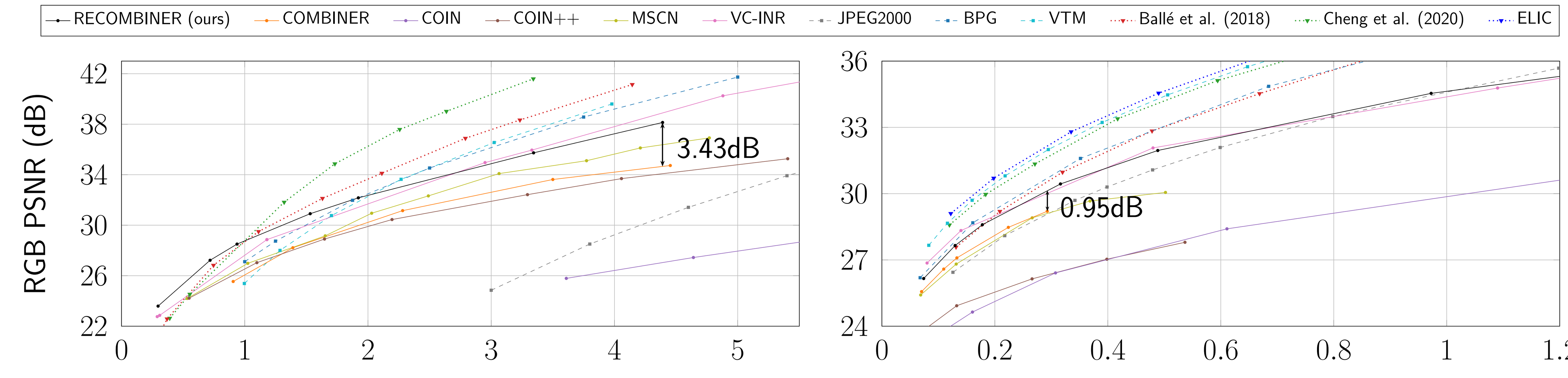


(a) COMBINER. (b) RECOMBINER. (c) The input in RECOMBINER.

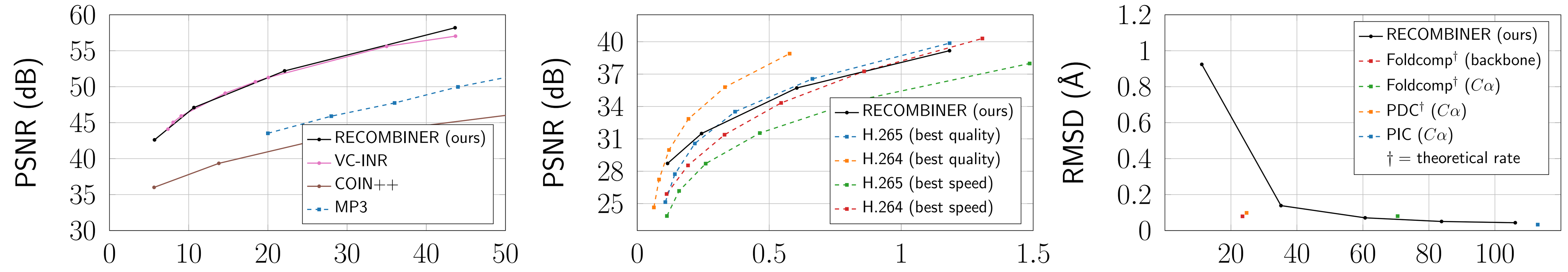


(a) Three-level hierarchical model and the corresponding graphical model. (b) Permutation.

Experimental Results

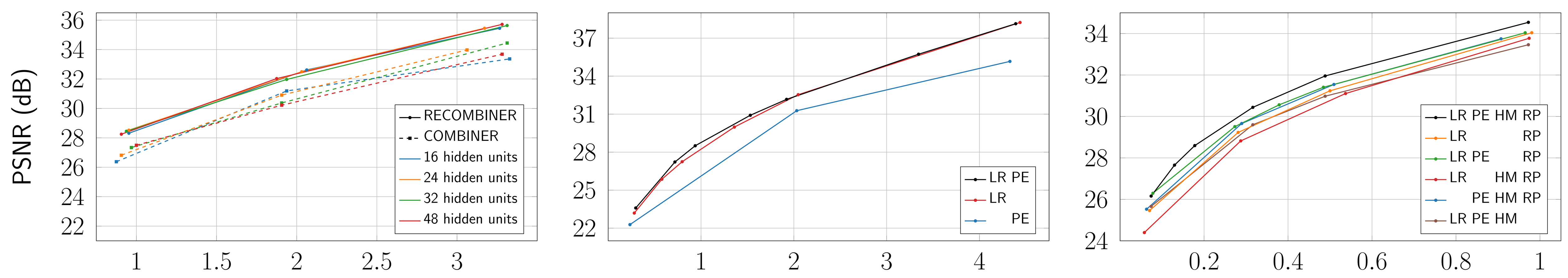


(a) RD curve on CIFAR-10 (image) (b) RD curve on Kodak (image)

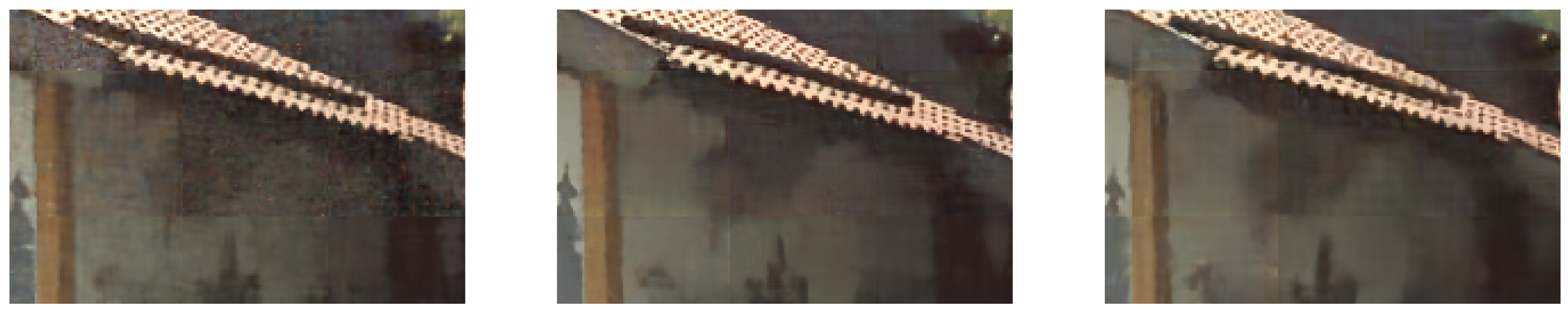


(b) RD curve on LibriSpeech (audio). (c) RD curve on UCF-101 (video). (d) RD curve on protein data.

Ablation Study



(a) Robustness to model size (b) Ablation study on CIFAR-10. (c) Ablation study on Kodak.



(a) w/o positional encodings; bitrate 0.287 bpp; PSNR 25.62 dB. (b) with positional encodings; bitrate 0.316 bpp; PSNR 26.85 dB. (c) with positional encodings; bitrate 0.178 bpp; PSNR 25.05 dB.