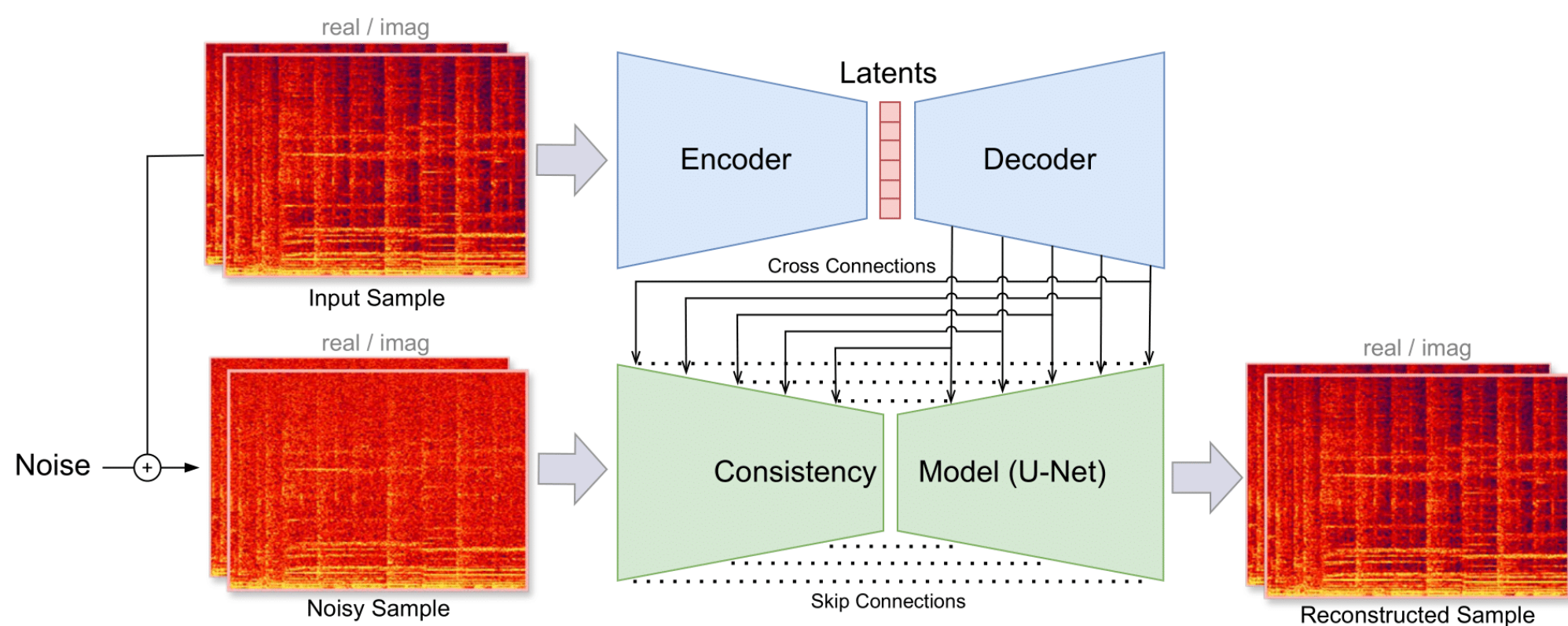


Music2Latent: Consistency Autoencoders for Latent Audio Compression

Marco Pasini

2023



Finding

Music2Latent is a consistency autoencoder that encodes samples into a continuous latent space with a high compression ratio. It is trained fully end-to-end using a single loss function. Its representations perform competitively on downstream tasks, and can be used to train generative models. It provides more accurate reconstructions than comparable models.

Question

We improve on the original consistency models training recipe by introducing continuous noise levels with a novel exponential schedule. We also adopt frequency wise self-attention and an adaptive frequency scaling mechanism to deal with the unbalanced magnitude distribution over different frequency bins. We validate the effectiveness via ablation studies.