Innovation 1: Signal Compression Deep Generative Models

DEVCOM ARMY RESEARCH LABORATORY   IoBT REIGN

IoBT nodes will sense and transmit high-dimensional signals such as images and videos for remote decision-making. IoBTs will have limited bandwidth and hence, signals need to be compressed before transmission over network such that the remote decision-making can make correct decisions by recovering original signal from compressed representation.

Contribution: We devised a new approach to compress high-dimensional signals using deep generative models for compression.

- Compressed Sensing: Recover signal $x \in R^n$ from (possibly noisy) measurements $M(x) = y \in R^m$ where $m \ll n$. Traditionally, M is the measurement matrix-a random matrix, for example, a Gaussian or Bernoulli matrix, which meets the Restricted Isometry Property (RIP) with a large probability. Classical approaches enable such recovery by imposing structures on signals such as l-sparseness in some known basis. For instance, images are sparse in wavelet basis, images are sparse in Fourier basis. Recover signal by

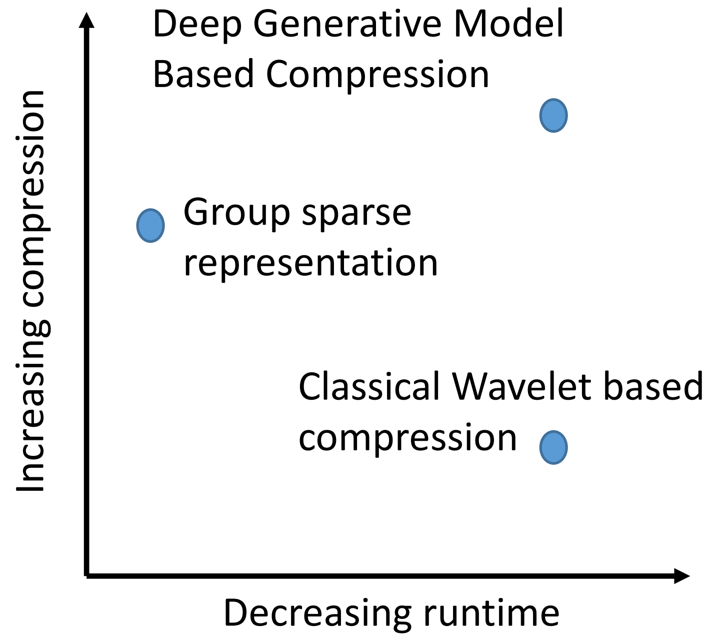
$$\min_x ||x||_0 \quad s.t. Mx = y$$

$$\underbrace{\boxed{\text{m x n}}}_{M} \times \underbrace{\boxed{\text{n x 1}}}_{x} = \underbrace{\boxed{\text{mx1}}}_{y}$$

- Deep generative model based compression: Deep learning can be used to *learn structure* from the data. From data, we can learn a generator $G: R^k \rightarrow R^n$ mapping latent space to signal space such that signal $x \approx G(z)$. Recover signal by

$$z^* = \min_z || MG(z) - y ||^2$$

Reconstructed signal is $x^* = G(z^*)$. Measurement error is $|| MG(z^*) - y ||^2$.Reconstruction error is $|| G(z^*) - x^* ||^2$. No longer a convex optimization problem but amenable to SGD.



Notational plot showing the utility of deep generative models for compression

Innovation 1: Deep learning based task-aware compression

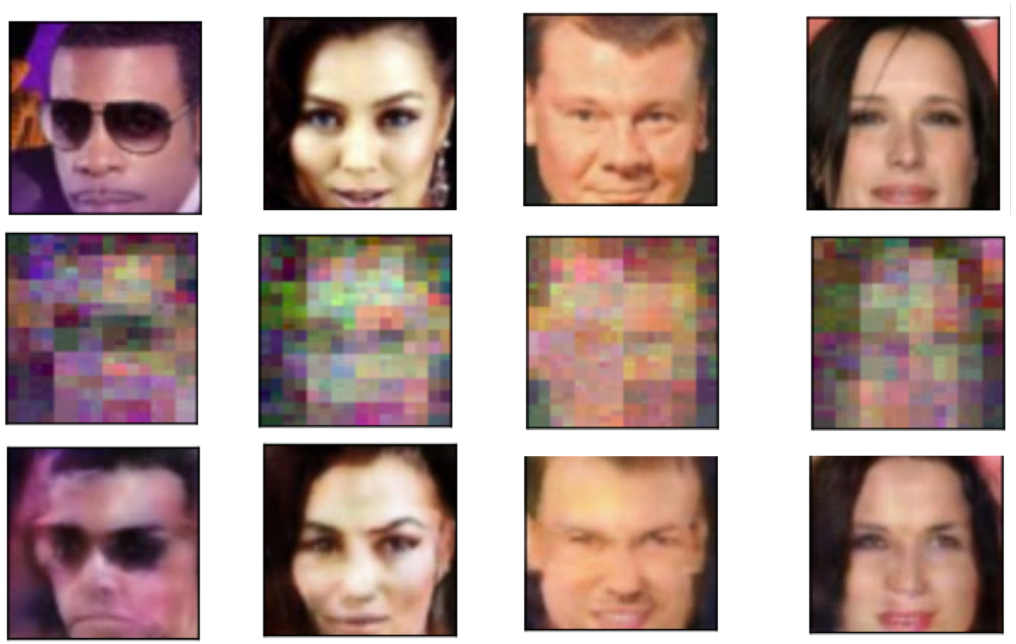
DEVCOM ARMY RESEARCH LABORATORY

IoBT REIGN

IoBT nodes will sense and transmit high-dimensional signals such as images and videos for remote decision-making. IoBTs will have limited bandwidth and hence, signals need to be compressed before transmission over network such that the remote decision-making can make correct decisions by recovering original signal from compressed representation.

Experiments and Results:

Reconstruction results on celebA with m = 500 measurements (of n = 12288 dimensional vector). Original images (top row), and reconstructions by Lasso with wavelet basis (second row), and deep generator based method (last row).



Downstream ML Task: Gender prediction accuracy after reconstruction at 91% accuracy

Summary and Impact:

- Replacing traditional methods of identifying sparse basis for each domain manually to the use of deep learning models to discover structure in the input signals enables more efficient compression and opens up opportunity for multimodal compression important for heterogeneous sensing capability in IoBTs.

Contact: Susmit Jha (SRI)

$$\min_{x} ||m - Ax||^2 . s.t. P(x) = 0$$

$$x^{k+1} = (A^T A + \rho A)^\wedge - 1 \quad (A^T m + \rho(p^k - u^k))$$

$$p^{k+1} .= \mathrm{argmin}_{\mathrm{y}}\, \lambda\, P(y) + \frac{\rho}{2} \left|\left|x^{k+1} - y + u^k\right|\right|^2$$

$$u^{k+1} = u^k + x^{k+1} - y^{k+1}$$