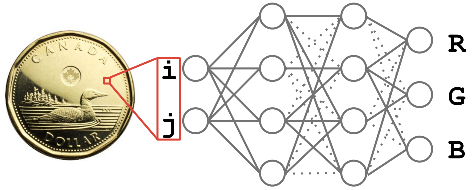
L$_0$onie: Compressing COINs with L0-constraints

Juan Ramirez, Jose Gallego-Posada

Mila Université de Montréal

Background

Implicit Neural Representations (INRs) train neural networks mapping coordinates to features, aiming to **approximate** a given object.



COIN: Compress an image by training an INR that maps from pixel locations to RGB values.

$$\min_{\boldsymbol{\theta}} \mathcal{L}_{\mathbf{I}}(\boldsymbol{\theta}) \triangleq \sum_{p\in\mathcal{P}} ||f_{\boldsymbol{\theta}}(\boldsymbol{x}_p) - \boldsymbol{y}_p||_2^2.$$

Challenges of COIN

Model size trade-off: larger models are more expressive but lead to worse compression rates

- ▲ Must solve a costly optimization problem to compress an image.
- ▲ Selecting the model architecture: tuning for low distortion entails architecture search.

Framework

Key idea: Train larger *"overparameterized"* COIN model, and sparsify during training to match desired BPP.

Main advantages:

- 🥳**Better reconstructions, faster!**
- 🥳**No sacrifice in compression rate.**

Use differentiable L0-reparametrization: $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}} \odot \boldsymbol{z}$.

$\tilde{\boldsymbol{\theta}}$: parameter magnitudes and $\boldsymbol{z}$: stochastic gates with params $\boldsymbol{\phi}$.

We consider a constrained optimization problem:

Quantize to `dtype`

Budget BPP

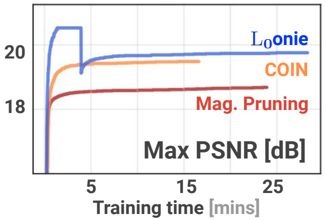
$$\min_{\tilde{\boldsymbol{\theta}},\boldsymbol{\phi}} \mathcal{L}_{\mathbf{I}}(\tilde{\boldsymbol{\theta}} \odot \hat{\boldsymbol{z}}(\boldsymbol{\phi})) \text{ s.t. } \text{BPP}(\texttt{cast}(\tilde{\boldsymbol{\theta}} \odot \hat{\boldsymbol{z}}(\boldsymbol{\phi})) \leq \tau_{\text{BPP}}$$

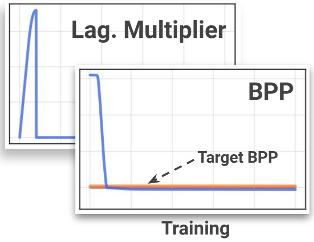
Medians of stochastic gates

Training dynamics of Lagrangian

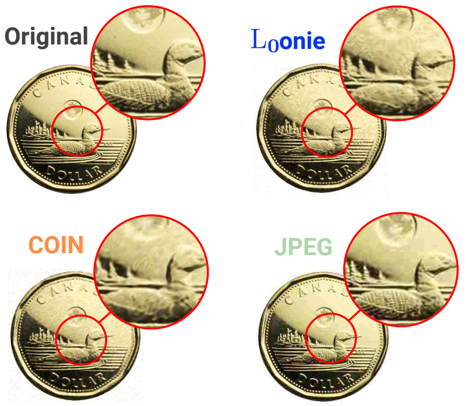
$$\tilde{\boldsymbol{\theta}}^*, \boldsymbol{\phi}^*, \lambda_{\text{co}}^* = \underset{\tilde{\boldsymbol{\theta}},\boldsymbol{\phi}}{\text{argmin}} \underset{\lambda_{\text{co}}\geq 0}{\text{argmax}} \mathfrak{L}(\tilde{\boldsymbol{\theta}}, \boldsymbol{\phi}, \lambda_{\text{co}})$$

$$\mathfrak{L}(\tilde{\boldsymbol{\theta}}, \boldsymbol{\phi}, \lambda_{\text{co}}) \triangleq \mathcal{L}_{\mathbf{I}}(\tilde{\boldsymbol{\theta}} \odot \hat{\boldsymbol{z}}(\boldsymbol{\phi})) + \lambda_{\text{co}} \left(\text{BPP}(\texttt{cast}(\tilde{\boldsymbol{\theta}} \odot \hat{\boldsymbol{z}}(\boldsymbol{\phi})) - \tau_{\text{BPP}}\right)$$





Results



Reconstructions at 0.6 BPP