Split-Brain Autoencoders: Unsupervised Learning by Cross-Channel Prediction
Richard Zhang  Phillip Isola  Alexei A. Efros
University of California, Berkeley

PROBLEM STATEMENT
Learn a “useful” representation in the unsupervised regime

Our contributions
1) Propose split-brain autoencoders, a method for unsupervised learning which uses raw data as its own supervisory signal
2) Demonstrate state-of-the-art performance
3) Ablation studies by studying cross-channel prediction problems and loss functions

METHODS

Traditional Autoencoder
Learn by reconstructing raw data from itself

Disadvantage: simple identity function can be learned, relies on special constraints (e.g., bottleneck) to force abstraction

Cross-Channel Encoder
Learn by predicting subset of channels from another
- E.g., predict color (ab) from grayscale (L channel) [Zhang et al. ECCV 2016]
- Straightforward; little engineering effort

Disadvantages: Only extracts features from a subset of data

Split-Brain Autoencoder
Aggregate complementary cross-channel prediction tasks
- Demonstrates general validity of cross-channel encoding
- Achieve state-of-the-art on various representation learning benchmarks
- Straightforward; little engineering effort (simply add split in network)

SPLIT BRAIN AUTOENCODERS

Image Representation Evaluation (Pascal VOC fine-tuning)

Image Representation Evaluation (Linear Classifiers)

Image B-D Evaluation (NYU-D Detection)

Supervised Pretraining
Unsupervised Pretraining

Ours (Split-Brain)