Unsupervised Learning

JESSICA THOMPSON

UNIVERSITY OF ALABAMA

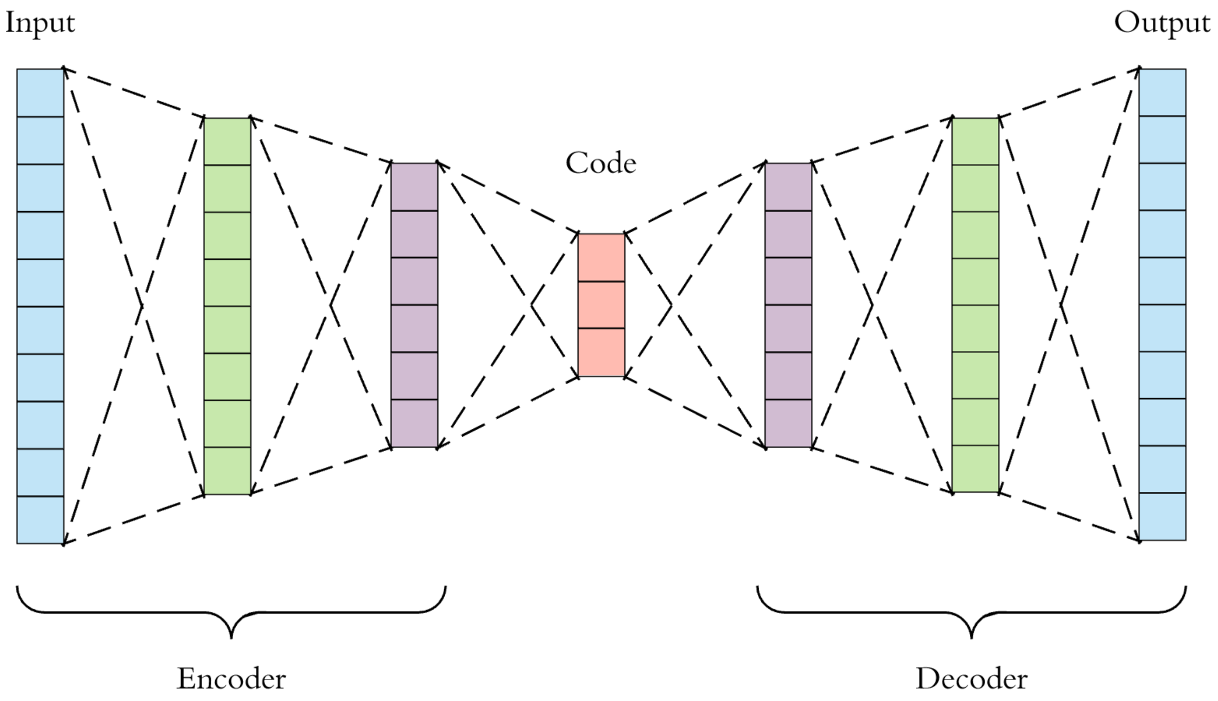
OCTOBER 24, 2018

Unsupervised Learning

- Attempt to find patterns in data without explicit labels
- Useful for images which are difficult or time-consuming to label
- Could use to train a network to find pairs of etch pits

Autoencoder

- Autoencoder finds representation of data in low dimensional space
- The compressed data may correlate to parameters we care about (such as number of etch pits)

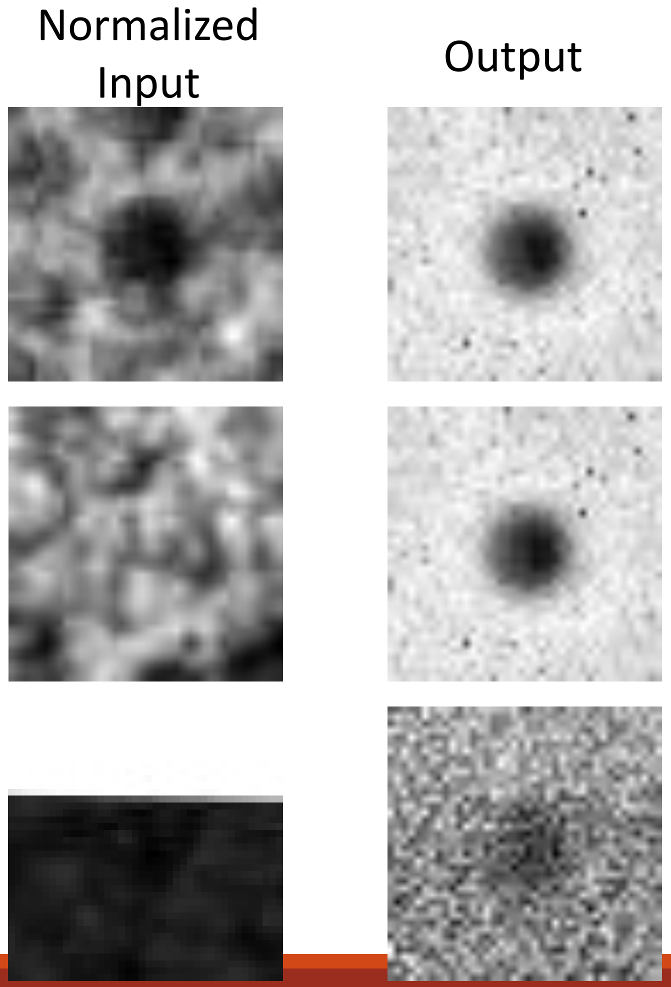


Current Work

- First attempt: train autoencoder on set of 500 images
  - This is the dataset we were able to classify with ~80% accuracy through supervised learning last year
- The autoencoder architecture compresses images to a 12-dimensional representation
- Loss is calculated as Euclidean distance between original image and decoded image + L2 regularization

Results

- Autoencoder generates a generic image as the output for every input image
- Almost every output image looks identical, but they do have subtle pixel differences
- These subtle differences evidently enable the network to reach a lower loss value than if every image was exactly the same
- Conclusion: network is differentiating images to some extent, though not in the way expected



Results

- I have varied network architecture
  - Number of dense layers used
  - Whether or not convolutional layers are used
  - Number of “dimensions” in encoded representation

- In every case, the output images all appear to be identical, with one “etch pit” in the center of image

Next Steps

- Check if encoded representation correlates to mean pixel intensity of image, etc.
- Add a Kullback-Leibler Divergence term to loss function
  - This should force the encoded representation to retain information about the image, rather than staying mostly static
- Train an autoencoder on newer dataset
  - In the old dataset, the “generic image” is a decent representation of nearly 60% of the images