6/19
One-shot Recognition

Kevin Tang

Project Review

- Learn a linear/non-linear projection to a lower dimension
- "Learn how to learn" by training on multiple subsets of the training data
  - One subset = 1 pseudo training example, 1 pseudo testing example
- Evaluate with one-shot recognition (one training example)
  - Using K-Nearest Neighbors

This Week's Progress

- Fixed a lot of bugs in implementation
- New corrected results for previous setups
- Optimized code with efficient data structures
- Implemented different baseline systems for comparison
- Further extended system
  - Different distance metrics
  - Multi-layer for nonlinearity

Experimental Setup

- Setup for experiments:
  - 1000 different sets of:
  - Training classes – subset = 1 training example, 10 testing examples
  - Testing classes – subset = 1 training example, 10 testing examples
  - Reduction to 10 dimensions

Corrected Results

- PCA (baseline)
  - ~17.30% mean accuracy
- Linear Projection
  - ~12.85% mean accuracy
  - Not fully optimized
- Logistic Projection
  - ~18.23% mean accuracy

New Extensions

- Additional baseline measurements
  - No dimensionality reduction
  - A = Identity matrix
- Different distance metrics
  - log(1 + norm(O(x) – O(y))^2)
  - sum(log(1 + (O(x) – O(y)).^2))
- Higher dimensionality reduction
  - Increase to 50 dimensions
  - Lampert has 85 attributes (dimensions)
- Run on training classes
- Multi-layer dimensionality reduction
  - Provides more non-linearity

Results

- Same experimental setup as before
- A = Identity matrix (baseline)
  - ~14.87% mean accuracy
  - PCA: ~17.30% mean accuracy
- Log distance, logistic projection
  - ~18.16% mean accuracy
  - Logistic Euclidean: ~18.23% mean accuracy
- Log distance summed, linear projection
  - ~19.81% mean accuracy
  - Linear Euclidean: ~12.85% mean accuracy

More Results

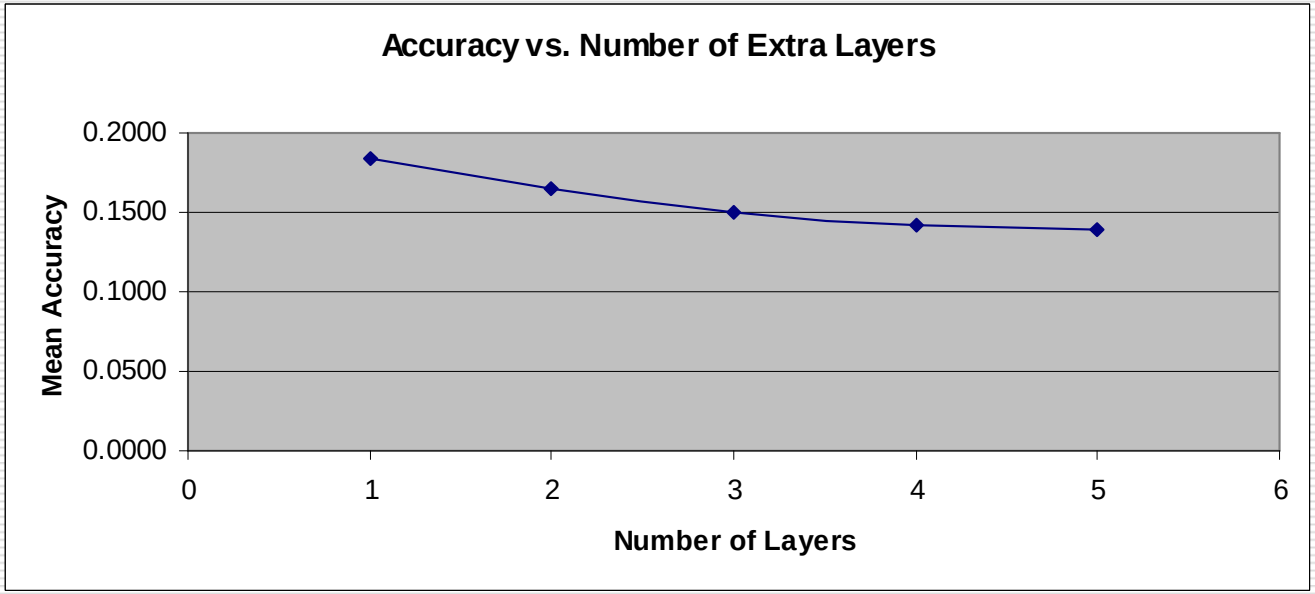
- Higher dimensionality reduction
  - Logistic projection with 50 dimension reduction
  - ~20.22% mean accuracy
  - Best results so far

More Results

- Evaluate on training classes
- Logistic projection (50 dimensions)
  - ~19.94% mean accuracy
  - Chance accuracy: 2.5%
- PCA (50 dimensions) (baseline)
  - ~4.63% mean accuracy
- A = Identity matrix (baseline)
  - ~4.03% mean accuracy

More Results

- Cascaded multiple logistic projections to make multi-layer, more non-linearity
- Not helping, as we are over-fitting
  - Training classes accuracy for 5-layer: ~51.41%



Next Steps

- 6 sets of features, right now only using rgsift
  - Incorporate all feature sets together
  - Extend multi-layer idea here
- Implement more sophisticated form of Stochastic Gradient Descent
- Possible extension to fast object recognition
- Obtain actual images from Lampert