An In-Depth Look at Feature Transformation Ability of CNN

A Case Study on MNIST Dataset

Huangshi Tian

October 15, 2018

Department of Computer Science and Engineering

Group Members (in Alphabetical Order)

Beijing Fang	Department of Civil and Environmental Engineering
Huangshi Tian	Department of Computer Science and Engineering
Yunfei Yang	Department of Mathematics

Outline

1. Feature extraction with ScatNet and ResNet.
2. Feature visualization with various methods.
3. Feature testing with SVM and random forests.
4. Discussion and conclusion.

Feature Extraction

Scattering Network

- Package: ScatNet in MATLAB
- Parameters:
  - maximum scattering order M = 2
  - number of scale J = 3
  - number of orientations L = 6
- Postprocessing: take spatial averages of scattering coefficients

$$\text{features of image } x = \sum_u S_j x(u)$$

Residual Network

- Model: an 18-layer ResNet pre-trained over ImageNet dataset, with its final FC layer removed
- Preprocessing: rescale MNIST images from 28×28 to 224×224 with bi-linear interpolation and triple it to emulate RGB channels

Feature Visualization

Methodology

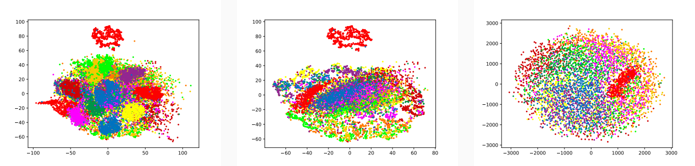
Global Level

- **Multidimensional Scaling (MDS)** preserving pairwise distances
- **Principal Component Analysis (PCA)** preserving global variation
- **Isometric Feature Mapping (Isomap)** preserving pairwise geodesic distances
- **Spectral Embedding (SE)** preserving connectivity

Local Level

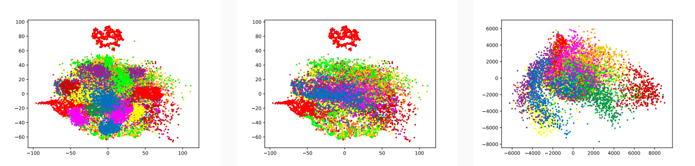
- **Locally Linear Embedding (LLE)** preserving local distances among neighbors
- **t-distributed Stochastic Neighbor Embedding (t-SNE)** preserving local distribution

Results of Gloabl Methods (1)



(a) ResNet (b) ScatNet (c) Raw

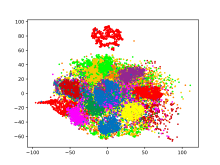
Figure 1: Visualization of features generated with Multidimensional Scaling (MDS).



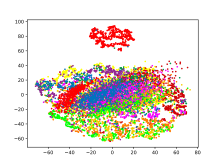
(a) ResNet (b) ScatNet (c) Raw

Figure 2: Visualization of features generated with Isometric Feature Mapping (Isomap).

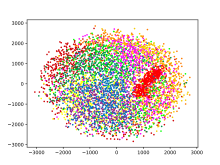
Results of Gloabl Methods (2)



(a) ResNet

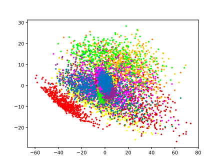


(b) ScatNet

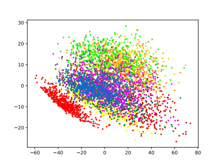


(c) Raw

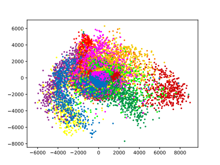
Figure 3: Visualization of features generated with Spectral Embedding (SE).



(a) ResNet



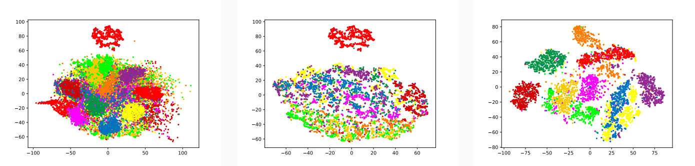
(b) ScatNet



(c) Raw

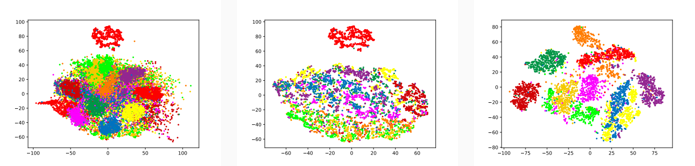
Figure 4: Visualization of features generated with Principal Component Analysis (PCA).

Results of Local Methods



(a) ResNet (b) ScatNet (c) Raw

Figure 5: Visualization of features generated with Locally Linear Embedding (LLE).



(a) ResNet (b) ScatNet (c) Raw

Figure 6: Visualization of features generated with t-distributed Stochastic Neighbor Embedding (t-SNE).

Feature Testing

Methodology

- Preprocessing: normalization by standard deviation and dropping constants
- Comparison Targets:
  - Algorithms: Support Vector Machine (SVM), Random Forest (RF)
  - Features: raw, ScatNet, ResNet
- Procedure:
  - tune parameters with five-fold cross validation and grid search
  - train models on datasets of different size

Results

Training Size	Raw Features (717)		ScatNet Features (127)		ResNet Features (512)		Average (Models)	
	SVM	RF	SVM	RF	SVM	RF	SVM	RF
300	80.22%	78.26%	82.15%	71.53%	85.06%	78.60%	82.48%	76.13%
1000	87.95%	89.09%	91.05%	80.95%	92.71%	87.69%	90.57%	85.91%
2000	90.05%	92.08%	92.67%	84.22%	93.52%	89.57%	92.08%	88.62%
5000	93.21%	94.35%	95.12%	87.31%	95.25%	90.65%	94.53%	90.77%
10000	94.91%	95.37%	96.09%	88.96%	96.17%	91.32%	95.72%	91.88%
20000	96.06%	96.24%	97.01%	90.73%	96.88%	92.17%	96.65%	93.05%
40000	96.99%	96.85%	97.48%	91.85%	97.49%	92.86%	97.32%	93.85%
60000	97.17%	97.22%	97.72%	92.57%	97.82%	93.00%	97.57%	94.26%
Average (training size)	92.07%	92.43%	93.66%	86.02%	94.36%	89.48%	**93.36%**	**89.31%**
Average (features)	92.25%		89.84%		91.92%			

Conclusion

Conclusion (Conjecture)

- On MNIST dataset, raw images could generally serve as better features than those transformed by CNN.
- When training samples are scarce, ScatNet could better capture structural difference than ResNet.
- CNN could contract global distance (metrics) among data samples, but may perturb the local structure at the meantime.
- When data samples are ample, ResNet could better cluster data samples by generating closer features.