ECE 599/692 – Deep Learning

Lecture 4 – CNN: Practical Issues

Hairong Qi, Gonzalez Family Professor
Electrical Engineering and Computer Science
University of Tennessee, Knoxville
http://www.eecs.utk.edu/faculty/qi
Email: hqi@utk.edu

Outline

- Lecture 3: Core ideas of CNN
  - Receptive field
  - Pooling
  - Shared weight
  - Derivation of BP in CNN
- Lecture 4: Practical issues
  - Normalized input and initialization of hyperparameters
  - Cross validation
  - Momentum
  - Learning rate
  - Activation functions
  - Pooling strategies
  - Regularization
- Lecture 5: Variants of CNN
  - From LeNet to AlexNet to GoogleNet to VGG to ResNet
- Lecture 6: Implementation
- Lecture 7: Applications of CNN – Binary hashing
- Lecture 8: Applications of CNN – Person re-identification
Cost functions

- Least square quadratic
  - The learning slowdown problem
- Cross-entropy
  - How does it solve the slowdown problem?

Activation functions

- Sigmoid
- Softmax
- Tanh
- ReLU
Regularization methods

• The problem of overfitting
• Weight decay or L2 regularization
• L1 regularization
• Dropout
• Artificial expansion of the training data

Learning rates

• Always use smaller rates

From [Duda&Hart:2001]
Momentum

• Taking into account of previous changes

\[
\omega_{st}^{k+1} = \omega_{st}^{k} - c_{k} \frac{\partial E^{k}}{\partial \omega_{st}^{k}} \\
\omega_{st}^{k+1} = \omega_{st}^{k} + (1 - c_{k}) \Delta \omega_{bp}^{k} + c_{k} (\omega_{st}^{k} - \omega_{st}^{k-1})
\]

From [Duda&Hart:2001]

Weight initialization

• Gaussian distribution with different std

![Graph showing classification accuracy over epochs with two lines representing old and new approaches to weight initialization.](image)