Deep Learning in Medical Imaging Analysis

Dinggang Shen, Ph.D.

IDEA Lab, Department of Radiology, and Biomedical Research Imaging Center (BRIC)
University of North Carolina at Chapel Hill
Deep Learning
For imaging-based diagnosis, manual labeling is expensive, and hence ground-truth data are very limited.
## Unsupervised Learning

<table>
<thead>
<tr>
<th>Method</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• PCA</td>
<td>✓ Linear</td>
</tr>
<tr>
<td></td>
<td>✓ Not optimal for non-Gaussian data</td>
</tr>
<tr>
<td>• Gaussian Mixture Models</td>
<td>✓ Require knowledge for the number of clusters</td>
</tr>
<tr>
<td>• K-Means</td>
<td>✓ Challenging when applied to high-dimensional data</td>
</tr>
<tr>
<td>• ICA</td>
<td>✓ Linear model</td>
</tr>
<tr>
<td>• Sparse Coding</td>
<td>✓ Shallow model (e.g., single-layer representation)</td>
</tr>
<tr>
<td>• Non-Linear Embedding</td>
<td></td>
</tr>
</tbody>
</table>

- All these methods involve just **one step** of mapping!
  - Mapping is shallow, **not deep**!
  - Thus, **not** able to represent the **complex** mapping!
Deep Learning – Why hot?

• Deep mapping and representation

- Deeper representations → abstractions → disentangling

- Manifolds are expanded and flattened

The following 5 slides edited from Dr. Yoshua Bengio's tutorial
Deep Learning – Why hot?

- Deep mapping and representation
  - Each level transforms the data into a representation, which can be easily modeled
  - Unfolding it more will map the original data to a factorized (uniform-like) distribution

Performance increase with layers
Deep Learning – Why hot?

- Successive model layers learn deeper intermediate representations.

Prior: Underlying factors and concepts compactly expressed without multiple levels of abstraction.
Neural Network – Why not working

• Issues with previous neural network (NN)
  – Gradient-based method ➔ propagate errors from the last layer to the previous layers
  – Last layer represents high nonlinear function (i.e., a jump function in binary classification) ➔ unstable and large gradient in small range, but zero in most places
Neural Network – Why not working

- Effect of initial conditions in Deep Nets

No two training trajectories end up in the same place → huge number of effective local minima

Pre-training: Transfer knowledge from previous learning (representation and explanatory factors) → cases with few examples → shared underlying explanatory factors, between $P(X)$ and $P(Y|X)$
Deep Learning – Why working now

- Three main reasons
  - New layer-wise training algorithm [Science 2006]
    - Each time, train on simple task
  - Big data, compared to 20 years ago
  - Powerful computers
    - Previous algorithms may be theoretically working, but practically not converged to good local minima with the previous less-powerful computers
Deep Learning

$P(v, h^1, h^2, \ldots, h^l) = P(v|h^1) P(h^1|h^2) \cdots P(h^{l-2}|h^{l-1}) P(h^{l-1}, h^l)$
Deep Learning – Greedy Training

- First step
  - Construct an RBM with an input layer $v$ and a hidden layer $h$
  - Train the RBM
Deep Learning – Greedy Training

• Second step
  – Stack another hidden layer on top of the RBM to form a new RBM
  – Fix $W^1$, sample $h^1$ from $Q(h^1|v)$ as input
  – Train $W^2$ as RBM
Deep Learning – Greedy Training

- Third step
  - Continue to stack layers on top of the network, and train it as previous step, with samples sampled from $Q(h^2|h^1)$
- And so on…
Deep Learning – Stacked Auto-Encoder

Pretraining | Unrolling | Fine-tuning

Reducing the Dimensionality of Data with Neural Networks

High-dimensional data can be converted to low-dimensional codes by training a multilayer neural network with a small central layer to reconstruct high-dimensional input vectors. Gradient descent can be used for fine-tuning the weights in such "autoencoder" networks, but this works well only if the initial weights are close to a good solution. We describe an effective way of initialization the weights that allow deep autoencoder networks to learn low-dimensional codes that work much better than principal components analysis as a tool to reduce the dimensionality of data.
Deep Learning – Stacked Auto-Encoder

A. The codes produced by a 500-250-125-2 Auto-Encoder

B. The codes produced by 2D LSA

C. The fraction of retrieved documents
Application 1

Segmentation

- Hippocampus Segmentation using 7T MRIs
- Infant Brain Segmentation
Hippocampus Segmentation

Hippocampus Segmentation Using 7T MR Images
Hippocampus Segmentation

- Challenges in hippocampus segmentation using 1.5T/3T and 7T
  - Low imaging resolution
  - Low contrast
  - Much richer structural information
  - Less partial volume effect
  - But, severe intensity inhomogeneity problem

Hand-Crafted Features

- Limited discriminative power of hand-crafted features

Extracting patches from a 7T MR image

Responses of Haar filters for the image patches
Hierarchical Feature Extraction via Unsupervised Deep Learning

- **Stacked two-layer convolutional ISA (Independent Subspace Analysis)**

![Diagram showing the process of hierarchical feature extraction via unsupervised deep learning.]

- Basis filters $W$ in 1st layer
- Activations $P$ in 1st layer
- Image patches $X$
- Dimension-reduced activations from 1st layer
- Basis filters $W'$ in 2nd layer
- Activations $P'$ in 2nd layer
- PCA

Learned basis filters by the 1st ISA
Multi-Atlas-based Segmentation using Deep Learning Features

Training Stage:
- Aligned training images in each atlas space, 1…N
- Image patches
- 2-layer ISA
- Learned features
- Classifier sequence 1
- ACM

Testing Stage:
- Image patches
- 2-layer ISA
- Learned features
- Classifier sequence 1
- ACM
- Adaptively weighted fusion
- Level set
- Probability map
- Segmentation result

Subject image space
Results

Comparison Results Using 20 Leave-One-Out Cases

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>R</th>
<th>RO</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand-Crafted Haar + Texture Features</td>
<td>0.843</td>
<td>0.847</td>
<td>0.772</td>
<td>0.865</td>
</tr>
<tr>
<td>Hierarchical Patch Representations</td>
<td>0.883</td>
<td>0.881</td>
<td>0.819</td>
<td>0.894</td>
</tr>
</tbody>
</table>

P = Precision; R = Recall; RO = Relative overlap; SI = Similarity index
Infant Brain Segmentation

Multi-modality Isointense Infant Brain Image Segmentation
Infant Brain Segmentation

- Challenges in infant brain segmentation
  - Low tissue contrast
  - Low spatial resolution

WM and GM exhibit almost the same level of intensity in both T1 and T2 MR images.
Deep Convolutional Neural Network (CNN)

T1  T2  FA

Multi-modality images

convolution  convolution  local response normalization (across maps)

full connection  soft-max  true label

cross entropy loss

CSF  GM  WM

Tissue segmentation
### Results

Segmentation performance in terms of **Dice ratio** achieved by the CNN, RF, SVM, CLS, MV

<table>
<thead>
<tr>
<th></th>
<th>Subj. 1</th>
<th>Subj. 2</th>
<th>Subj. 3</th>
<th>Subj. 4</th>
<th>Subj. 5</th>
<th>Subj. 6</th>
<th>Subj. 7</th>
<th>Subj. 8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CSF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNN</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.84</td>
<td>0.85</td>
<td>0.82</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>RF</td>
<td>0.82</td>
<td>0.81</td>
<td>0.83</td>
<td>0.81</td>
<td>0.83</td>
<td>0.85</td>
<td>0.79</td>
<td>0.80</td>
</tr>
<tr>
<td>SVM</td>
<td>0.74</td>
<td>0.77</td>
<td>0.77</td>
<td>0.74</td>
<td>0.70</td>
<td>0.78</td>
<td>0.72</td>
<td>0.73</td>
</tr>
<tr>
<td>CLS</td>
<td>0.81</td>
<td>0.82</td>
<td>0.73</td>
<td>0.86</td>
<td>0.84</td>
<td>0.82</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td>MV</td>
<td>0.71</td>
<td>0.69</td>
<td>0.68</td>
<td>0.63</td>
<td>0.61</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>GM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNN</td>
<td>0.85</td>
<td>0.86</td>
<td>0.88</td>
<td>0.82</td>
<td>0.81</td>
<td>0.87</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>RF</td>
<td>0.83</td>
<td>0.85</td>
<td>0.88</td>
<td>0.81</td>
<td>0.80</td>
<td>0.85</td>
<td>0.85</td>
<td>0.84</td>
</tr>
<tr>
<td>SVM</td>
<td>0.79</td>
<td>0.80</td>
<td>0.83</td>
<td>0.75</td>
<td>0.74</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>CLS</td>
<td>0.83</td>
<td>0.84</td>
<td>0.85</td>
<td>0.83</td>
<td>0.81</td>
<td>0.87</td>
<td>0.86</td>
<td>0.84</td>
</tr>
<tr>
<td>MV</td>
<td>0.85</td>
<td>0.84</td>
<td>0.85</td>
<td>0.80</td>
<td>0.78</td>
<td>0.80</td>
<td>0.84</td>
<td>0.83</td>
</tr>
<tr>
<td><strong>WM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNN</td>
<td>0.88</td>
<td>0.81</td>
<td>0.88</td>
<td>0.85</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
<td>0.88</td>
</tr>
<tr>
<td>RF</td>
<td>0.86</td>
<td>0.78</td>
<td>0.87</td>
<td>0.84</td>
<td>0.85</td>
<td>0.86</td>
<td>0.84</td>
<td>0.84</td>
</tr>
<tr>
<td>SVM</td>
<td>0.82</td>
<td>0.74</td>
<td>0.76</td>
<td>0.80</td>
<td>0.80</td>
<td>0.79</td>
<td>0.71</td>
<td>0.76</td>
</tr>
<tr>
<td>CLS</td>
<td>0.84</td>
<td>0.81</td>
<td>0.80</td>
<td>0.82</td>
<td>0.84</td>
<td>0.82</td>
<td>0.83</td>
<td>0.81</td>
</tr>
<tr>
<td>MV</td>
<td>0.86</td>
<td>0.80</td>
<td>0.85</td>
<td>0.82</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
</tr>
</tbody>
</table>

CNN: Convolutional Neural Network  
RF: Random Forest  
SVM: Support Vector Machine  
CLS: Coupled Level Sets  
MV: Majority Voting
Results

- Original multi-modality data (T1, T2 and FA)
- Manual segmentations (CSF, GM, and WM)
- Segmentation results by CNN
- Segmentation results by RF
Application 2

Registration

- Brain MRI Registration
Brain MRI Registration

Feature-based Symmetric Registration of Brain MR Images
Feature-based Symmetric Image Registration

- S-HAMMER (Symmetric HAMMER)

Stacked Auto-Encoder (SAE)

Input image patches

Morphological signatures for image registration

Features learned in the first layer
Deep Learning based Intrinsic Features

- By local patches
- By SIFT
- By unsupervised deep learning
## Results

### Dice ratios of WM, GM, and VN on ADNI dataset (%)

<table>
<thead>
<tr>
<th>Methods</th>
<th>VN</th>
<th>GM</th>
<th>WM</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demons</td>
<td>90.2</td>
<td>76.0</td>
<td>85.7</td>
<td>84.0</td>
</tr>
<tr>
<td>M+PCA</td>
<td>90.5</td>
<td>76.6</td>
<td>85.5</td>
<td>84.1</td>
</tr>
<tr>
<td>M+DP</td>
<td>90.9</td>
<td>76.5</td>
<td>85.8</td>
<td>84.4</td>
</tr>
<tr>
<td>HAMMER</td>
<td>91.5</td>
<td>75.5</td>
<td>85.4</td>
<td>84.1</td>
</tr>
<tr>
<td>H+PCA</td>
<td>91.7</td>
<td>76.9</td>
<td>86.5</td>
<td>85.0</td>
</tr>
<tr>
<td><strong>H+DP</strong></td>
<td><strong>95.0</strong></td>
<td><strong>78.6</strong></td>
<td><strong>88.1</strong></td>
<td><strong>86.6</strong></td>
</tr>
</tbody>
</table>

### Averaged Dice ratios of 54 ROIs on LONI LPBA40 dataset (%)

<table>
<thead>
<tr>
<th>Methods</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demons</td>
<td>68.9</td>
</tr>
<tr>
<td>M+PCA</td>
<td>68.9</td>
</tr>
<tr>
<td>M+DP</td>
<td>69.2</td>
</tr>
<tr>
<td>HAMMER</td>
<td>70.2</td>
</tr>
<tr>
<td>H+PCA</td>
<td>70.6</td>
</tr>
<tr>
<td><strong>H+DP</strong></td>
<td><strong>72.7</strong></td>
</tr>
</tbody>
</table>

Demons: Diffeomorphic Demons  
M+PCA: Multi-channel Demons + PCA  
M+DP: Multi-channel Demons + Deep learning  
H+PCA: HAMMER + PCA  
H+DP: HAMMER + Deep learning
Results

The overlap ratio for hippocampus: 68.5% → 78.4%

Registration results on 7T MRI brain images
Application 3

Disease Diagnosis

- AD/MCI Diagnosis
- CADx for Lung Nodules and Breast Lesions
AD/MCI Diagnosis

Classification of Alzheimer’s Disease and Mild Cognitive Impairment
Alzheimer’s Disease (AD)

- The most common form of dementia
  - An irreversible neurodegenerative disease that causes disruptions in memory, cognition, and eventually death
  - A growing epidemic: Worldwide, nearly 44 million people are living with AD
  - Cannot delay or halt the progression of AD

- Prodromal stage of AD: Mild Cognitive Impairment (MCI)
Computer-Aided Diagnosis for AD

- Neuroimaging modalities for diagnosis
  - MRI, PET, fMRI, ...

- Previous works: simple low-level features
  - MRI: gray matter tissue volumes
  - PET: mean signal intensities
  - CSF: biomarker measures

Vulnerable to noises and/or artifacts
Latent Feature Representation

• Hidden or latent high-level information
  – Deep architecture can be efficiently used to discover latent or hidden representation in self-taught learning
  – Overcome the vulnerability to noise/artifacts in the data by encoding in a hierarchical feature space

• Unsupervised greedy training
  – Allows us to benefit from the target-unrelated samples to discover general latent feature representations
  – Leverages for enhancement of the accuracy

Latent Feature Representation with SAE

- Feature extraction
  - MRI
  - PET
  - CSF
- Pre-training
- Fine-tuning
- Feature representation with Stacked Auto-Encoder
- Label prediction
- Clinical scores regression
- Multi-task learning
- Multi-kernel SVM learning
- AD/MCI diagnosis
- Label
- MMSE ADAS-Cog
- MRI kernel
- PET kernel
- CSF kernel
- Stacked Auto-Encoder
- Augmented feature vector

Multi-modality fusion

Feature selection

Latent feature representation

Simple

Complex
Results

Classification results using ADNI dataset (51 AD, 52 HC, 43 MCI-C, 56 MCI-NC)
Multi-Modal Fusion

- Fusing complementary information from multiple modalities helps enhance diagnostic accuracy

- Previous approaches
  - Independent steps of feature extraction and modality fusion

- Deep Boltzmann Machine (DBM)
  - High-level feature representation via deep learning

- Multi-Modal DBM (MM-DBM)
  - Inherent relations between modalities of MRI and PET
Multi-Modal Fusion

- **Multi-modal input images**
  - MRI
  - PET

- **Patch extraction**
  - $2^{K \times [w \times w \times w]}$

- **Patch-level feature learning**
  - $K \times F_S$

- **Image-level classifier learning**
  - Patch-level SVM learning
  - Spatially distributed “mega-patch” construction
  - Weighted ensemble SVM classifier learning

$I$: image size, $w$: patch size, $K$: # of selected patches, $m$: modality index, $F_S$: # of hidden units in the top-layer of multi-modal Deep Boltzmann Machine (DBM)

Multi-Modal Fusion

Hidden Layer 1

MRI

PET

Learned weights
## Results

<table>
<thead>
<tr>
<th>Methods</th>
<th>Dataset (AD/MCI/NC)</th>
<th>AD vs. NC (%)</th>
<th>MCI vs. NC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kohannim et al., 2010</td>
<td>MRI+PET+CSF (40/83/43)</td>
<td>90.7</td>
<td>75.8</td>
</tr>
<tr>
<td>Walhovd et al., 2010</td>
<td>MRI+CSF (38/73/42)</td>
<td>88.8</td>
<td>79.1</td>
</tr>
<tr>
<td>Hinrichs et al., 2011</td>
<td>MRI+PET (48/119/66)</td>
<td>92.4</td>
<td>n/a</td>
</tr>
<tr>
<td>Westman et al., 2012</td>
<td>MRI+CSF (96/162/111)</td>
<td>91.8</td>
<td>77.6</td>
</tr>
<tr>
<td>Zhang and Shen, 2012</td>
<td>MRI+PET+CSF (51/99/52)</td>
<td>93.3</td>
<td>83.2</td>
</tr>
<tr>
<td><strong>Proposed method</strong></td>
<td>MRI+PET (93/204/101)</td>
<td><strong>93.5</strong></td>
<td><strong>85.2</strong></td>
</tr>
</tbody>
</table>
Missing Data

- Incomplete multi-modal neuroimaging data
  - Not all subjects have all data modalities
  - The accuracy of disease diagnosis can be improved if the missing data could be estimated
  - The relationship between different data modalities is complicated and nonlinear

More than 50% of the subjects in ADNI dataset do not have PET data

Deep Learning for Multi-modality Data Completion

Input: 3D MRI patch 1@15x15x15

Hidden layers:
- 3D feature maps 10@9x9x9
- 3D feature maps 10@3x3x3

Full connection

Output: 3D PET patch 1@3x3x3

Full connection

3D CNN architecture for imaging data completion
Deep Learning for Multi-modality Data Completion

Predicted PET

Ground truth PET

AD
## Results

### Performance comparison of classification tasks

(308 subjects with both MRI and PET)

<table>
<thead>
<tr>
<th></th>
<th>MCI vs. NC</th>
<th>pMCI vs. sMCI</th>
<th>AD vs. NC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PET</td>
<td>True data</td>
<td>0.70±0.02</td>
<td>0.68±0.02</td>
</tr>
<tr>
<td></td>
<td>3D CNN</td>
<td>0.69±0.03</td>
<td>0.68±0.02</td>
</tr>
<tr>
<td></td>
<td>KNN</td>
<td>0.63±0.02</td>
<td>0.63±0.03</td>
</tr>
<tr>
<td></td>
<td>Zero</td>
<td>0.62±0.02</td>
<td>0.61±0.02</td>
</tr>
</tbody>
</table>

### Performance comparison of classification tasks

(830 subjects with MRI and true/estimated PET)

<table>
<thead>
<tr>
<th></th>
<th>MCI vs. NC</th>
<th>pMCI vs. sMCI</th>
<th>AD vs. NC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI</td>
<td>0.75±0.03</td>
<td>0.72±0.03</td>
<td>0.92±0.02</td>
</tr>
<tr>
<td>PET</td>
<td>3D CNN</td>
<td>0.73±0.03</td>
<td>0.70±0.02</td>
</tr>
<tr>
<td></td>
<td>KNN</td>
<td>0.64±0.02</td>
<td>0.61±0.03</td>
</tr>
<tr>
<td></td>
<td>Zero</td>
<td>0.61±0.02</td>
<td>0.59±0.03</td>
</tr>
<tr>
<td>MRI+PET</td>
<td>3D CNN</td>
<td>0.76±0.02</td>
<td>0.72±0.02</td>
</tr>
<tr>
<td></td>
<td>KNN</td>
<td>0.72±0.02</td>
<td>0.68±0.03</td>
</tr>
<tr>
<td></td>
<td>Zero</td>
<td>0.72±0.03</td>
<td>0.63±0.03</td>
</tr>
</tbody>
</table>
Landmark-based Deep Feature Learning for AD Diagnosis

MR Image Pre-processing

Training data

Landmark Detection

Fast Landmark Detection

Patch Extraction

Patch-based Deep Feature Learning

CNN 1 → Deep Feature 1

... → ...

CNN P → Deep Feature P

Deep Features

Disease Diagnosis

Image Retrieval
Landmark-based Deep Feature Learning for AD Diagnosis

Structure of the 3DCNN model used in this study
Landmark-based Deep Feature Learning for AD Diagnosis

Manifold visualization of AD and NC subjects in the ADNI-2 dataset, by t-SNE projection in learned 3DCNN layers.
Results

Comparison of AD/NC classification results:

- ROIF: ROI based features
- VBF: Voxel based features
- LMF: Landmark based morphological features (concatenation/ensemble)
- 3DCNN_FC9/FC10/FC11: Features from the FC9/FC10/FC11 layer in 3DCNN (concatenation/ensemble)
- 3DCNN_Label_ens: Ensemble of labels obtained from 3DCNN
Results

Comparison of AD/NC classification results

MAP: Mean Average Precision
MCC: Matthews Correlation Coefficient
#Correct@K: Number of correct results in top K returned results
Computer-aided Diagnosis
Applications to *Pulmonary Nodules in CT Scans* and *Breast Lesions in Ultrasound Images*
Diagnosis for Lung Nodules and Breast Lesions

### Ultrasound Breast Lesions

- **Benign**
- **Malignant**

### CT Lung Nodules

#### Radiologist’s Diagnosis
- Inter-observer Variation
- Intra-observer variation
- Dependence on Experience
- Human Error

#### Computer-aided Diagnosis
- Decision Support
- Resolve Intra-observer Variation
- Avoid Unnecessary Biopsy
Deep Learning CADx vs. Conventional CADx

• Deep learning CADx
  – Automatic feature extraction and selection
  – Free of intermediate image processing steps (e.g., image segmentation)

SDAE Architecture

Pre-training
- Output layer
- Hidden layers
- Input layer

Supervised Training
- Soft-max classification
- Hidden layers
- Input layer

Resized ROI

Resizing

Aspect Ratio, Resizing Factors

Raw ROIs
Results

Comparison of classification accuracy

RANK: Ranklet Transform + Grey Level Co-occurrence Matrix (GLCM) Features (Yang et al., 2013)
CURVE: Curvelet Transform + GLCM Features (Sun et al., 2013)
MORPH: Clinical Size and Diameter

Ultrasound Breast Lesions

CT Lung Nodules

SINGLE: Use ROI of median slice for each nodule
ALL: Use all ROIs for each nodule and vote ROIs at testing
Outcome Prediction

Brain Tumor Patient’s Survival Time Prediction
Brain Tumor Patient’s Survival Time Prediction

- High-grade gliomas
  - One of most deadly tumors with fast grow rate and poor prognosis
  - Pre-operative outcome prediction with high accuracy is critical for better treatment planning

Prediction based on Tumor Tissue Appearance

T1 MRI

DTI

λ₁

λ₂

λ₃

FA

MD

RD

B₀

fMRI

0.01-0.027 Hz

0.027-0.073 Hz

0.073-0.198 Hz

0.198-0.25 Hz

0-0.25 Hz

power

freq.
Multi-modal/Multi-channel Deep Learning using 3D-CNN

T1
- Conv + Pooling Layers
- Fully Connected Layers
- Softmax Loss

DTI
- Conv + Pooling Layers
- Fusion Layer
- Fully Connected Layers
- Softmax Loss

fMRI
- Conv + Pooling Layers
- Fusion Layer
- Fully Connected Layers
- Softmax Loss

Feature Fusion
## Results

### Comparison of prediction accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>ACC</th>
<th>SEN</th>
<th>SPE</th>
<th>PPR</th>
<th>NPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical features</td>
<td>62.96</td>
<td>66.39</td>
<td>58.53</td>
<td>63.18</td>
<td>65.28</td>
</tr>
<tr>
<td>Haar</td>
<td>69.17</td>
<td>72.81</td>
<td>65.84</td>
<td>65.56</td>
<td>73.14</td>
</tr>
<tr>
<td>SIFT</td>
<td>80.56</td>
<td>85.71</td>
<td>77.27</td>
<td>70.59</td>
<td>89.47</td>
</tr>
<tr>
<td>2D-CNN</td>
<td>81.25</td>
<td>80.95</td>
<td>81.82</td>
<td>88.35</td>
<td>74.23</td>
</tr>
<tr>
<td><strong>Proposed</strong></td>
<td><strong>89.95</strong></td>
<td><strong>96.87</strong></td>
<td><strong>83.90</strong></td>
<td><strong>84.94</strong></td>
<td><strong>93.93</strong></td>
</tr>
</tbody>
</table>

ACC=Accuracy; SEN=Sensitivity; SPE=Specificity; PPR=positive predictive rate; NPR=negative predictive rate
Application 4

Image Synthesis

- Estimating CT from MRI
- 7T MRI Construction
Estimating CT from MRI

Estimating CT Images from MRI by Fully Convolutional Networks
Estimated CT from MRI

- CT images
  (+) Dose planning
  (+) PET attenuation correction
  (−) Radiation

- Challenge in estimating CT from MRI
  – Hard to train the networks (3D)

A Basic FCN (Fully Convolutional Network) Architecture

Deeply Supervised Nets (DS-FCN)

\[ J = \gamma R(w_1) + \lambda_1 L_1(w_1; X_{MR}, Y_{CT}) \]

- Regularizer
- Loss function at level 1

Crop
Predicted CT Patch

Deeply Supervised Nets (DS-FCN)

\[ J = \gamma R(w_1) + \lambda_1 L_1(w_1; X_{MR}, Y_{CT}) + \gamma R(w_2) + \lambda_2 L_2(w_2; X_{MR}, Y_{CT}) \]

Deeply Supervised Nets (DS-FCN)

\[ J = \gamma R(w_1) + \lambda_1 L_1(w_1; X_{MR}, Y_{CT}) + \gamma R(w_2) + \lambda_2 L_2(w_2; X_{MR}, Y_{CT}) + \gamma R(w_3) + \lambda_3 L_3(w_3; X_{MR}, Y_{CT}) \]

DS-FCN with Auto-Context Model (ACM)

MRI

DS-FCN

Predicted CT

DS-FCN

Predicted CT

DS-FCN

Predicted CT

...
Results

Comparison of estimated CT images

Atlas: Multi-atlas-based method
SRF: Structure Random Forest
SRF+: Structure Random Forest w/ Auto-context Model Refinement
## Results

### Performance comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>MAE</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean(std)</td>
<td>Med.</td>
</tr>
<tr>
<td>Atlas</td>
<td>64.6(6.6)</td>
<td>65.9</td>
</tr>
<tr>
<td>SR</td>
<td>54.3(10.0)</td>
<td>54.0</td>
</tr>
<tr>
<td>SRF+</td>
<td>48.1(4.6)</td>
<td>48.3</td>
</tr>
<tr>
<td>DS-FCN</td>
<td>42.4(5.1)</td>
<td>42.6</td>
</tr>
<tr>
<td>DS-FCN+ACM</td>
<td><strong>38.7(4.6)</strong></td>
<td><strong>38.9</strong></td>
</tr>
</tbody>
</table>

Atlas: Multi-atlas-based method  
SRF: Structure Random Forest  
SRF+: Structure Random Forest w/ Auto-context Model Refinement
7T MRI Reconstruction
Convolutional Neural Network For Reconstruction of 7T-like Images from 3T MRI
**7T vs. 3T MRI**

- **7T MRI**
  - Higher spatial resolution and better tissue contrast, compared to 3T MRI
  - SNR of 7T MRI $\approx 2.3 \times$ SNR of 3T MRI

- Lower availability and higher price of 7T MRI scanners
  - 20,000 3T scanners vs. 40 7T scanners in the world

---

7T MRI Reconstruction

CNN framework for 7T MRI reconstruction
Results

LIS: Local Image Similarity
M-CCA: Multi-level CCA
Results

Our results are more reliable than others when the quality of training & testing images are different.
Thank you!

For more details, please visit:
http://bric.unc.edu/ideagroup
Or google: unc idea
Application 4

Prostate Labeling

• MRI Prostate Segmentation for Cancer Diagnosis and Treatment
MRI Prostate Segmentation

- MRI prostate images provide good soft tissue contrast
  - MRI-guided transperineal prostate core biopsy
  - MRI-guided radiotherapy planning
  - Quantitative analysis using MR images

Challenges

- Large inter-subject anatomical appearance variability
- Inhomogeneity
- Large inter-subject shape variability

Discover abstract latent feature representations from the target-unrelated samples through greedy learning

**Pre-training:** Unsupervised Stacked Auto-Encoder


Optimize the deep-learned feature representation for a certain task to enhance accuracy

**Fine-tuning:** Supervised Stacked Auto-Encoder
Latent Feature Representation

Feature Difference Maps
Proposed Framework

**Learning Stage**
- Deep Learning with Stacked Auto-Encoder
- Pre-training
- Fine-tuning
- Sparse Shape Model

**Segmenting Stage**
-_likelihood Map
- Deformable Model
- Sparse Shape Constraint
- Target Image

**Training Images (Atlases)**
- Sparse Shape Model

**Deep Learning based Feature Representation**
- Sparse Patch Matching

BRIC Research Lab IDEA
## Results

### Comparison of segmentation performance

<table>
<thead>
<tr>
<th>Method</th>
<th>Intensity</th>
<th>Haar</th>
<th>HOG</th>
<th>LBP</th>
<th>Handcraft</th>
<th>Unsupervised SSAE</th>
<th>Supervised SSAE</th>
<th>Supervised SSAE w/ DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dice (%)</td>
<td>85.3±6.2</td>
<td>85.6±4.9</td>
<td>85.7±4.9</td>
<td>85.5±4.3</td>
<td>85.9±4.5</td>
<td>86.7±4.4</td>
<td>87.1±4.2</td>
<td>87.8±4.0</td>
</tr>
<tr>
<td>Precision (%)</td>
<td>85.1±8.3</td>
<td>85.9±8.5</td>
<td>85.3±8.7</td>
<td>83.7±7.7</td>
<td>87.3±7.4</td>
<td>87.3±7.3</td>
<td>87.1±7.3</td>
<td>91.6±6.5</td>
</tr>
<tr>
<td>Hausdorff</td>
<td>8.68±4.24</td>
<td>8.50±2.86</td>
<td>8.51±2.69</td>
<td>8.59±2.38</td>
<td>8.55±2.91</td>
<td>8.65±2.69</td>
<td>8.12±2.89</td>
<td>7.43±2.82</td>
</tr>
<tr>
<td>ASD (mm)</td>
<td>1.87±0.93</td>
<td>1.76±0.52</td>
<td>1.74±0.50</td>
<td>1.75±0.44</td>
<td>1.77±0.54</td>
<td>1.68±0.49</td>
<td>1.66±0.49</td>
<td>1.59±0.51</td>
</tr>
</tbody>
</table>

Dice: Dice ratio; Hausdorff: Hausdorff distance; ASD: Average surface distance