The Contributions of Deep Learning to Computer Vision: Application to Medical Images

Nicolas Thome
Professor at Conservatoire National des Arts et Métiers (Cnam)
Machine Learning Team (MSDMA) - Computer Science Department
Prenom.Nom@cnam.fr - http://cedric.cnam.fr/~thomen/

April 4, 2018
**Context: Big Data**

- Superabundance of visual data: images, videos, *etc*
  
  **BBC:** 2.4M videos  
  **Social media,**  
  *e.g.* Facebook: 1B each day  
  **100M monitoring cameras**

- Obvious need for **Visual Recognition**

- Huge number of applications: mobile visual search, medical imaging, robotics, autonomous driving, augmented reality, *etc*
Challenge: filling the semantic gap

What we perceive vs What a computer sees

- Illumination variations
- View-point variations
- Deformable objects
- intra-class variance
- etc

⇒ How to design "good" intermediate representations?
Before DL: **handcrafted intermediate representations** for each domain

Since DL: **Representation Learning**

---

**Handcrafted**

- Before DL:
  - **Local Features** (SIFT/ HoG)
  - Coding / Pooling
  - Classifier

- Since DL:
  - Learned
  - Learned
  - Learned

**Image**
Outline

1. Convolutionnal Neural Networks (ConvNets)

2. Deep learning for Medical Images
Neural Networks

- The formal neuron

\[ y = f(w^T x + b) \]

- Stacking several formal neurons ⇒ Perceptron
The Multi-Layer Perceptron (MLP)

- Perceptron: limited to linear decision boundaries
- Stacking layers of neural networks ⇒ more complex and rich functions

- Basis of the “deep learning” field
- All parameters trained with backpropagation with class labels
Convolutional Neural Networks (ConvNets)

- Scalability issue with Fully Connected Networks (MLP) + no local information!

ConvNets: sparse connections, shared weights = compact + local features
Convolutional Neural Networks (ConvNets)

- Convolution on tensors, i.e. multidimensional arrays: $T$ of size $W \times H \times D$
  - Convolution: $C[T] = T'$, $T'$ tensor of size $W' \times H' \times K$
  - Each filter locally connected with shared weights ($K$ number of filters)

- **Elementary block:** Convolution + Non linearity (e.g. ReLU) + pooling

- Stacking several Blocks: intuitive hierarchical information extraction
Deep Learning History

- 80’s: training Convolutional Neural Networks (CNN) with back-propagation $\Rightarrow$ postal code reading [LeCun et al., 1989]

- 90’s: golden age of kernel methods, NN = black box

- 2000’s: BoW + SVM : state-of-the-art CV
Deep Learning History


- Two main practical reasons:
  1. Huge number of labeled images (10^6 images)
  2. GPU implementation for training
Outline

1. Convolutional Neural Networks (ConvNets)
2. Deep learning for Medical Images
Deep Learning (DL) for Medical Image Diagnostic

- Deep ConvNets require large-scale annotated datasets
- **BUT:** Transferring Representations learned from ImageNet
  Extract layer (fixed-size vector) $\Rightarrow$ "Deep Features" (DF)

- Now state-of-the-art for any visual recognition task [Azizpour et al., 2016]
- DF very robust to domain shifts, e.g. medical images
Deep Learning (DL) for Medical Image Diagnostic

- DL & ConvNets: performance boost for classification in medical images
- Transfer & fine-tuning (ImageNet), e.g. Polyp Detection [Tajbakhsh et al., 2016]
- ConvNets trained from scratch, e.g. Mammography Classification [Kooi et al., 2017]
- ConvNets: winners of recent challenges based on deep learning: Mammography, Melanoma Detection, etc
- **Semantic segmentation**: assigning a label to each image pixel
Deep Learning segmentation: classifying image regions around each pixel
- Standard computer vision models based on Fully Convolutional Networks (FCN)

- FCN base models for many state-of-the-art methods segmentation methods, e.g. leading approach in Liver Tumor Segmentation (LiTS’17) challenge [Li et al., 2017]
Deep Learning for Medical Images

- Successful exportation of DL solutions boost performances... **BUT**
- ... Medical images very different from natural images:
  - Discriminative pattern often tiny, *e.g.* Mammography 0.5% – 1.2% cancer pixel [Akselrod-Ballin et al., 2017] vs > 50% ImageNet or > 30% VOC
  - ⇒ **Strong imbalance** between $\oplus$ and $\ominus$ (background) classes

Calcification(0.5%)  Mass(1.2%)
Successful exportation of DL solutions boost performances... **BUT**

... Medical images very different from natural images:

- **3D volumes vs 2D Images**
- Hierarchical / nested detection or organs, e.g. tumor inside liver
Resolution loss through the network

- **Introduction of skip connections** in U-Net [Ronneberger et al., 2015]
Specific Deep Learning Architectures for Medical Images

Representation Learning with 3D Inputs?

- Use 3D convolution, e.g. V-Net [Milletari et al., 2016], 3D U-Net [Çiçek et al., 2016] or [Lu et al., 2017]
Specific Training Schemes for Medical Images

Class imbalance

Use a specific loss function, e.g.

- Weighted cross entropy, U-Net [Ronneberger et al., 2015]
- Dice score, V-Net [Milletari et al., 2016] or [Fidon et al., 2017, Sudre et al., 2017]

\[
S = \frac{2|T \cap P|}{|T| + |P|}
\]
Specific Training Schemes for Medical Images

Exploit *prior* knowledge between organs, e.g. tumors only in liver

- Cascaded FCNNs for liver-tumor segmentation [Christ et al., 2016]
Conclusion

- Deep Learning & ConvNets: state-of-the-art solutions for medical image analysis
  - **Representation learning** ⇒ better visual features
- Exporting solutions from computer vision: transfer for classification, RPN for localization, FCN for segmentation, etc
  - Some adaptation required: spatial resolution, class imbalance, 3D data, etc
- Other crucial steps for deploying DL solution in Healthcare: uncertainty estimate and explainability ⇒ vanilla DL models poor at these tasks
Thank you for your attention!

Questions?


Large scale deep learning for computer aided detection of mammographic lesions.

Imagenet classification with deep convolutional neural networks.

Backpropagation applied to handwritten zip code recognition.

H-denseunet: Hybrid densely connected unet for liver and liver tumor segmentation from CT volumes.
*CoRR*, abs/1709.07330.

Automatic 3d liver location and segmentation via convolutional neural network and graph cut.


U-Net: Convolutional Networks for Biomedical Image Segmentation.

Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations.
In *DLMIA/ML-CDS@MICCAI*. 
Convolutional Neural Networks for Medical Image Analysis: Fine Tuning or Full Training?
*IEEE Transactions on Medical Imaging, PP(99):1–1.*