Application des réseaux neuronaux convolutifs à la quantification des lésions pulmonaires dans les pneumopathies infiltratives idiopathiques

Catalin FETITA - Telecom SudParis

Résumé : Les pneumopathies interstitielles idiopathiques sont une sous-classe des pathologies interstitielles pulmonaires dont la cause est souvent inconnue et qui se manifestent comme un processus de dégradation continu et irréversible de la fonction pulmonaire. Leur suivi quantitatif dans le temps par imagerie tomodensitométrique nécessite le développement d'outils d'aide au diagnostic automatiques, capables de discriminer des patterns texturaux parfois très proches entre le parenchyme normal et les diverses atteintes des tissus (fibrose, verre dépoli, emphysème, ...).

Dans ce contexte de classification de texture pulmonaire, les nouvelles approches par apprentissage profond utilisant des réseaux neuronaux convolutionnels (RNC) semblent très prometteuses. Cet exposé en présente le principe ainsi que des résultats comparatifs obtenus avec deux architectures RNC différentes.

Références :
Application of convolutional neural networks to the quantification of lung lesions in idiopathic interstitial pneumonia

Catalin FETITA
Sebastian TARANDO

ARTEMIS Department, TELECOM SudParis
Institut Mines-Telecom, Evry, France
Presentation overview

- Clinical problem, objective of the study and challenges: target on *interstitial lung diseases (ILDs)*
- Materials and methods: a CNN-based approach
- Test on two CNN architectures
- Results and discussion
- Conclusion and future work
Clinical context

Background: social impact of lung diseases

- Very rare
- More than 300 different ILDs
  - IIP (~55% of ILDs, >60 years)
  - Sarcoidosis
  - PICE
  - HCL
  - ...

In 2012, 115043 people died from lung disease (20% of all deaths), among which 5292 from IIP (4.6% of deaths from lung disease)

New drugs become available and need evaluation

UK deaths from IPF compared with other lung diseases (2012).
Source: https://statistics.blf.org.uk/pulmonary-fibrosis
Clinical context

ILDs diagnosis and follow-up

- CT as reference modality for ILDs diagnosis and follow-up:
  - High spatial resolution
  - Contrast healthy vs. pathological tissue

- IIP CT scans biomarkers investigation (consensus ATS 2011)
  - Reticulation
  - Traction Bronchiectasis
  - Honeycombing
  - Ground Glass
  - Emphysema (tobacco)
Clinical context

CT examples

- Examples of CT biomarkers for infiltrative lung diseases (ILDs)

ground glass (IIP)  
emphysema
Clinical context

CT examples

- Examples of CT biomarkers for infiltrative lung diseases (ILDs)

Examples of CT biomarkers for infiltrative lung diseases (ILDs):

- Fibrosis (IIP)
- Honeycombing
Objective

Automated computer-aided diagnosis system

Develop an automated system for a **quantitative follow-up** of pulmonary diseases (ILDs) using volumetric CT

- accurate and robust for inclusion as **second reader** in radiological workflow
- process the entire CT lung volume
- provide the lung region partitioning into defined phenotypes (classes)
Objective

Challenges of ILD classification

- Ontological overlap between classes

![Images of lung scans and corresponding tissue samples showing normal, emphysema, fibrosis, and ground glass]

- Low overlap
- High overlap

![Venn diagram illustrating the overlap between normal, ground glass, emphysema, fibrosis, and AEF in IIP]

CNN-based quantification of IIP by C. Fetita & S. Tarando
Objective

Challenges of ILD classification

- **Image texture variability:**
  - Inter-patient
  
  ![Lung CT images showing variability](image)

  Lack of CT acquisition normalization

  - Intra-patient

  ![Lung CT images showing variability](image)

  - Emphysema
  - G.Glass

antero-posterior gravitational gradient
Materials and Methods

A Machine Learning Approach

- Classification method: oriented towards machine learning

ML with image data as input and deep architectures

(Deep Learning = SoA in the field)

Artificial Neural Networks

→ CNNs

high performance in computer vision
Structuring the information

Size of input data: decisions are based on limited local information => patch size capture the “context”

5x5 patch

11x11 patch
Structuring the information

- Large input patches are problematic
  - Using each voxel as an input feature rapidly increases the number of weights
  - More training data required to generalize well

5x5 input patch

11x11 input patch
Structuring the information

- Consider spatial invariance
  - Structures and objects can appear at various places in the image ⇒ very difficult to learn for MLPs

- Ordering the information: from low- to high-level
  - Similar low-level features can be present at different places in the image ⇒ smarter to first combine spatially close information

Image values
  \[ \downarrow \]
  Gradients
  \[ \downarrow \]
  Simple shapes
  \[ \downarrow \]
  ...

Materials and Methods

Convolutional (CNN) vs Artificial Neural Networks (ANN / MLP)
Materials and Methods

Convolutional (CNN) vs Artificial Neural Networks (ANN / MLP)

- Structuring the information
  - Consider spatial invariance
    
    **CNNs adopt the strategy of sharing weights:**
    - similar structures at different locations produce the same output
    - reduces the number of weights to learn

  - Ordering the information: from low- to high-level
    - CNNs stack several convolutional layers to extract features from low- to high-level via successive convolutions
    - Achieves larger “receptive” fields at high-level
Convolutional Neural Networks

A convolutional layer convolves an image with multiple smaller kernels

- The kernels are not predefined, but learned
- Each convolution kernel generates a feature map
- Stacking convolutional layers leads to more complex features
  - play the role of feature computation and selection
- CNN ends with a fully-connected layer (combines convolutions and MLP)
  - play the role of classifier
Strategies for large receptive fields

- Large receptive fields: more possibilities to combine features (higher expressiveness)

- Can also be achieved by using small kernels and deeper architectures, but
  - very deep networks hard to train
  - more likely to overfit

**Solution: pooling**
Strategies for large receptive fields: **pooling**

- Reduce the size of the feature map by integrating spatial neighborhood
- Non-overlapping ROIs => select maximum, average, ... in each ROI

Pooling => advantage for learning deep CNNs

Trainable convolutional parameters: **8200**
Strategies for large receptive fields: pooling

- Reduce the size of the feature map by integrating spatial neighborhood
- Non-overlapping ROIs => select maximum, average, ... in each ROI

Pool (pooling) => advantage for learning deep CNNs

Trainable convolutional parameters: 1800
Materials and Methods

I. CNN: choice of the architecture

- Sliding patch window to classify its central pixel
I. CNN: choice of the architecture

- Two CNN architectures in competition (sliding window approach)

**Variant of Le-Net**
(Lecun et al, 1998)

*Handwritten character recognition*

**Extension (deeper) of T-CNN**  (Anthimopoulos et al, 2016)

*Texture analysis*
Materials and Methods

II. Experimental setup – Construction of the database

- Controlled acquisition protocol (Avicenne Hospital)
  - similar convolution kernel (same image quality)
  - pixel size in the range of 0.4 – 0.9 mm
  - image resolution 512x512 or 768x768
  - slice thickness in the range of 0.625 – 1.25 mm
  - different manufacturers (Philips, Siemens, GE)

- 1 expert radiologist and 3 junior radiologist to perform/validate annotations (~10 images per case with GT annotation)

- 130 cases with GT available, split at patient level in training (100), validation (10), testing (20) sets
Materials and Methods

II. Experimental setup – Construction of the database

- Examples of annotated GT in the database

- Emphysema-GGlass-Fibrosis
- Ground Glass
- Fibrosis

Legend:
- N: Normal
- F: Fibrosis
- E: Emphysema
- GDG: Ground Glass
III. Data preprocessing

- Bias introduced by normal high densities of vascular tree
  - Removal of vascular tree opacities (local connected filtering)

- Remove ambiguities normal vs. pathological tissue
  - Images rescaled to match 0.4 mm/pixel spacing

Tarando et al, SPIE MI 2018
Materials and Methods

III. Data preprocessing

- Bias introduced by normal high densities of vascular tree
  - Removal of vascular tree opacities (local connected filtering)

Remove ambiguities normal vs. pathological tissue

- Images rescaled to match 0.4 mm/pixel spacing

Tarando et al, SPIE MI 2018
Creating training/validation datasets:

- 130 sets of MDCT of Avicenne Hospital (Bobigny, France) database
- From the training CT dataset 16x16 pixels half-overlapping patches were extracted (falling 80% inside the marked ROIs)
Creating training/validation datasets:

• 130 sets of MDCT of Avicenne Hospital (Bobigny, France) database

• From the training CT dataset 16x16 pixels half-overlapping patches were extracted (falling 80% inside the marked ROIs)

• Horizontal flipping and rotations are applied in order to artificially increase the number of samples and avoid over-fitting the neural network

<table>
<thead>
<tr>
<th></th>
<th>Training set (# of patches)</th>
<th>Validation set (# of patches)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>186160</td>
<td>2000</td>
</tr>
<tr>
<td>Emphysema</td>
<td>203312</td>
<td>2000</td>
</tr>
<tr>
<td>Fibrosis/HC</td>
<td>113508</td>
<td>1300</td>
</tr>
<tr>
<td>Ground glass</td>
<td>54451</td>
<td>500</td>
</tr>
</tbody>
</table>
IV. Training

- LeNet vs DT-CNN training curves: loss (mean for all classes expressed in %) vs. number of epoch (an entire pass for all training samples):

**LeNet**

**DT-CNN**

optimState (ADAM) =
{
  learningRate : 8e-05
  beta2 : 0.999
  beta1 : 0.9
}

optimState (SGD) =
{
  learningRateDecay : 0
  weightDecay : 0
  learningRate : 0.08
  momentum : 0.95
}
Materials and Methods

IV. Training

- LeNet vs DT-CNN training curves: accuracy (mean for all classes expressed in %) vs. number of epoch (an entire pass for all training samples):

<table>
<thead>
<tr>
<th></th>
<th>Train acc</th>
<th>Validation acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeNet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>80.292%</td>
<td>46.127%</td>
</tr>
<tr>
<td>IIP</td>
<td>76.694%</td>
<td>67.753%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Train acc</th>
<th>Validation acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT-CNN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>87.875%</td>
<td>53.462%</td>
</tr>
<tr>
<td>IIP</td>
<td>82.915%</td>
<td>67.863%</td>
</tr>
</tbody>
</table>

N: Number of samples
Results

Testing the classification system

- LeNet vs DT-CNN qualitative comparison on test dataset:

![LeNet](image1)

![DT-CNN](image2)

**PER**
Results

Testing the classification system

- LeNet vs DT-CNN qualitative comparison on test dataset:
Results

Testing the classification system

- LeNet vs DT-CNN qualitative comparison on test dataset:

![Image of CT scan with classifications](image)

**LeNet**

- Green: N
- Blue: E
- Pink: IIP

**DT-CNN**

**LED**
Results

Testing the classification system

- LeNet vs DT-CNN qualitative comparison on test dataset:

- **LED**
- **DT-CNN**
Focus: DL in ILD

V. Testing for 3-class problem (N, E, IIP); DB2

- Qualitative analysis

![CT image](image1)

**MAU**

![Classification results](image2)

LeNet

DT-CNN

**N**

**E**

**IIP**
Focus: DL in ILD

V. Testing for 3-class problem (N, E, IIP); DB2

- Qualitative analysis

**MAU**

![MAU Image]

**LeNet**

![LeNet Image]

**DT-CNN**

![DT-CNN Image]
Focus: DL in ILD

V. Testing for 3-class problem (N, E, IIP); DB2

- Qualitative analysis
Focus: DL in ILD

V. Testing for 3-class problem (N, E, IIP); DB2

- Qualitative analysis

![Image of LAU and LeNet/DT-CNN outputs with color labels for N, E, F, GG, and IIP]
Results

Testing the classification system

DT-CNN quantitative evaluation on test dataset (20 patients) after noise removal:

- sensitivity (True Positive Rate, TPR)

\[
TPR = \frac{[TP]}{[TP] + [FN]}
\]

- specificity (True Negative Rate, TNR)

\[
TNR = \frac{[TN]}{[TN] + [FP]}
\]

- accuracy (ACC)

\[
ACC = \frac{[TP] + [TN]}{N}
\]

<table>
<thead>
<tr>
<th></th>
<th>Sens</th>
<th>Spec</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>81.3%</td>
<td>77%</td>
<td>80.3%</td>
</tr>
<tr>
<td>E</td>
<td>55%</td>
<td>97%</td>
<td>96%</td>
</tr>
<tr>
<td>IIP</td>
<td>76.3%</td>
<td>85%</td>
<td>83%</td>
</tr>
</tbody>
</table>
Discussion

Lessons learned

- Similar qualitative response between LeNet et DT-CNN

- DT-CNN: more accurate boundaries between structures due to the deeper architecture and absence of pooling

- Training data is crucial for accurate results:
  - representativeness of classes searched for
  - well-balanced samples between classes (data augmentation is good but not sufficient)
  - accurate annotation of labels
Conclusion

- CAD system for lung texture classification issued from a deep T-CNN architecture combined with a pre-filtering stage for vessel removal

- The DT-CNN performance achieved on average for all three classes (N, E, IIP) is 70.86% sensitivity, 86.3% specificity, 86.4% accuracy

- Problems of representativeness of the training database and ontological class overlap were not completely solved. Possible directions:
  - Increase clinical database (currently 130 cases)
  - Considering other clinical parameters to be included in the decision making, especially to discriminate between fibrosis and ground glass (4-class problem)
  - Other CNN architectures to investigate such clinical problem
Application of convolutional neural networks to the quantification of lung lesions in idiopathic infiltrative pneumonia

catalin.fetita@telecom-sudparis.eu