INSA Rouen Normandie ⁕ Normandie Université ⁕ Henri-Becquerel Center, Rouen
LITIS laboratory ✧ Learning Team
S. Belharbi, C. Chatelain, R.Hérault, S. Adam, S. Thureau, M. Chastan, R. Modzelewski.

Spotting L3 slice in CT scans using deep convolutional network and transfer learning

⁕ Medical application ⁕

Soufiane Belharbi

soufiane.belharbi@insa-rouen.fr

sbelharbi.github.io

INSA Rouen Normandie

July 8, 2018

Problem setup: L3 slice localization in CT scans

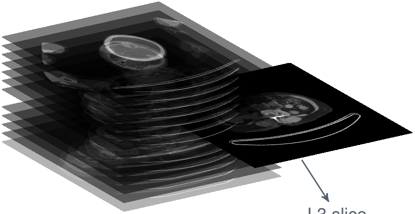
Context: Collaboration with Henri-Becquerel center at Rouen (cancer).

Main goal: Estimate the sarcopenia[1] level from a computerized tomography (CT) scan based only on the third lumbar vertebra (L3).

☞ A CT scan is stack of N slices (2D images). ☞ N is variable.

☞ In a CT scan, a specific slice is selected to represent the L3.

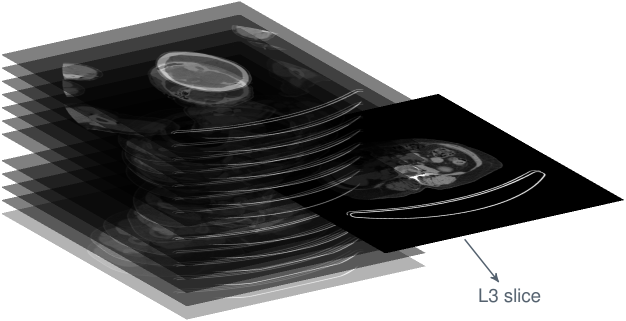
$\Rightarrow$ Need to locate the slice representing the third lumbar vertebra.



Find the L3 slice within a whole CT scan.

1. Sarcopenia: loss of skeletal muscle mass.

Problem setup: L3 slice localization in CT scans

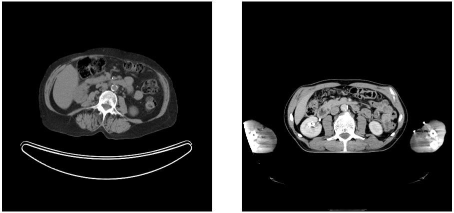


Finding the L3 slice within a whole CT scan.

- **L3CT1**:
  a dataset composed of 642 CT scans provided by Henri-Becquerel center.
- **Available annotation**:
  the position of the 3^{rd} lumber vertebra. (i.e., the number of the correct slice in the CT scan)

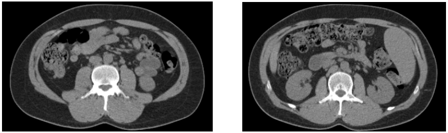
Problems:

⚠ Inter-patients variability.



L3 slices from two different patients: [Left] Patient A. [Right] Patient B.

⚠ Visual similarity of the vertebrae slices of the same patient.



Two slices from the same patient: [Left] an L3 slice. [Right] a non L3 slice.

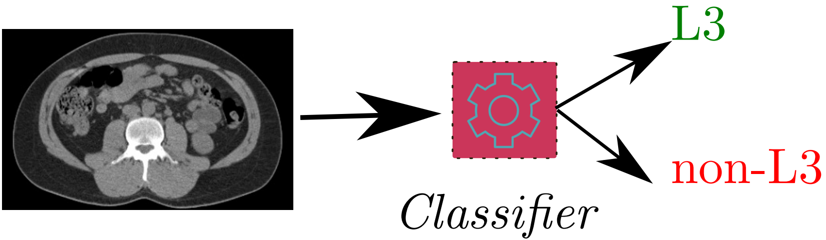
☞ The need to use the context to localize the L3 slice.

☞ **Machine Learning**!

Classification (discrete value) [X]

Classify each slice for: “L3” or “Not L3”:

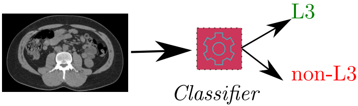
- Simple. ☺
- ⚠ No context. ☹

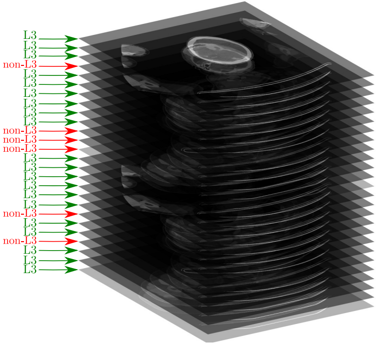


Classification (discrete value) [X]

Classify each slice for: “L3” or “Not L3”:

- Simple.
- ⚠ No context.



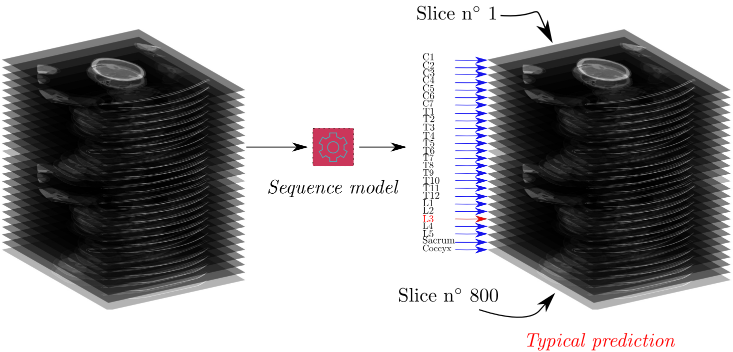


Typical prediction (no context)

Sequence labeling [X]

Label all the slices (vertebrae): L1, L2, L3, . . . :

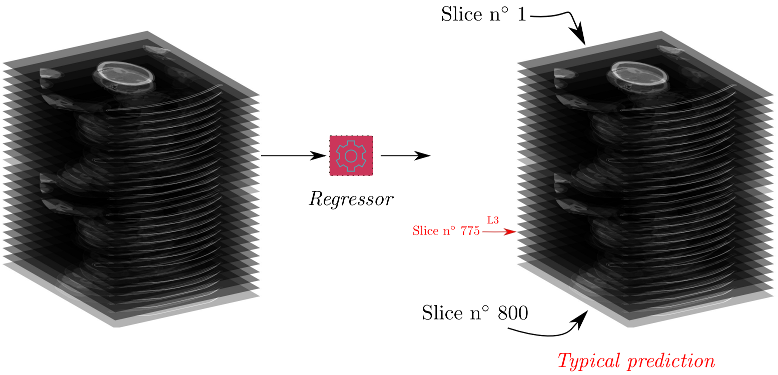
- Global analysis: context. ☺
- Existing work with promising results. ☺
- ⚠ Requires labeling more than one slice. ☹



Regression (real value) [✓]

Predict the height (position) of the L3 slice inside the CT scan:

- Global analysis: context. ☺
- Requires labeling only the L3 slice position. ☺



Which model for regression?

- State of the art in computer vision: Deep learning, convolutional neural network (CNN).
- ⚠ Requires fixed input size (when using dense layers).
- ⚠ Needs a large number of training samples.

Issues

- ⚠ High dimension input: $\underbrace{1 \text{ scan} = N \times 512 \times 512}_{\text{Problem 1: large input space}}$, with $400 < N < 1200$.
- ⚠ Implies: $\underbrace{\textit{Variability}}_{\text{Problem 2: Different input size}}$ of the height of each scan (depends on N).
- ⚠ Dataset with annotated L3 position: $\underbrace{642 \text{ patients}}_{\text{Problem 3: few training data}}$. (L3CT1 dataset)

Which model for regression?

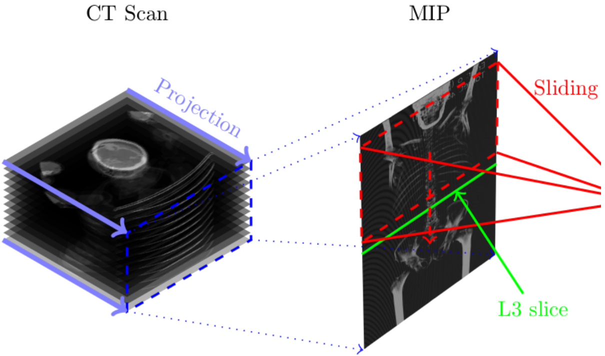
- State of the art in computer vision: Deep learning, convolutional neural network (CNN).
- ⚠ Requires fixed input size (when using dense layers).
- ⚠ Needs a large number of training samples.

Issues

- ⚠ High dimension input: $\underbrace{1 \text{ scan} = N \times 512 \times 512}_{\text{Problem 1: large input space}}$, with $400 < N < 1200$.
- ⚠ Implies: $\underbrace{\textit{Variability}}_{\text{Problem 2: Different input size}}$ of the height of each scan (depends on N).
- ⚠ Dataset with annotated L3 position: $\underbrace{642 \text{ patients}}_{\text{Problem 3: few training data}}$. (L3CT1 dataset)

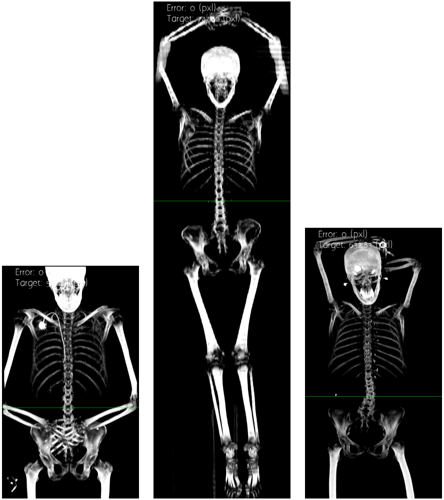
Problem 1: High dimension input

- $131M$ inputs for one example (large input dimension):
  - Frontal or lateral **Maximum Intensity Projection (MIP)**.
- $512 \times 512 \times N \Longrightarrow 512 \times N$.
- Preserves pertinent information (skeletal structure).



Proposed approach: Regression for L3 localization

Issue 2: Different input size > Solution: Sliding window

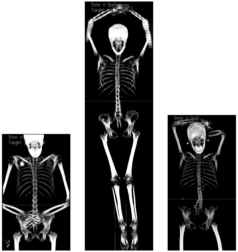


Examples of normalized frontal MIP images with the L3 slice position.

Problem 2: Different input size

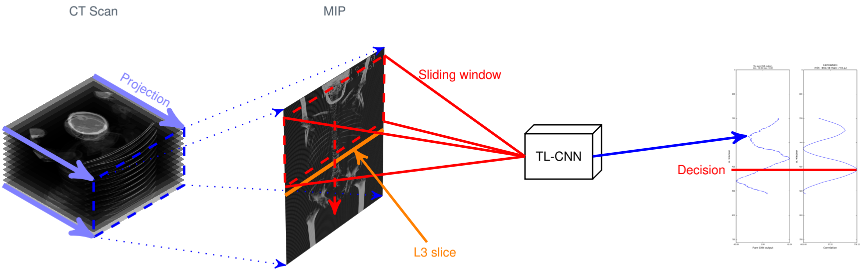
Classical problem in computer vision

- Sliding window technique
- Post-processing



Examples of normalized frontal MIP images with the L3 slice position.

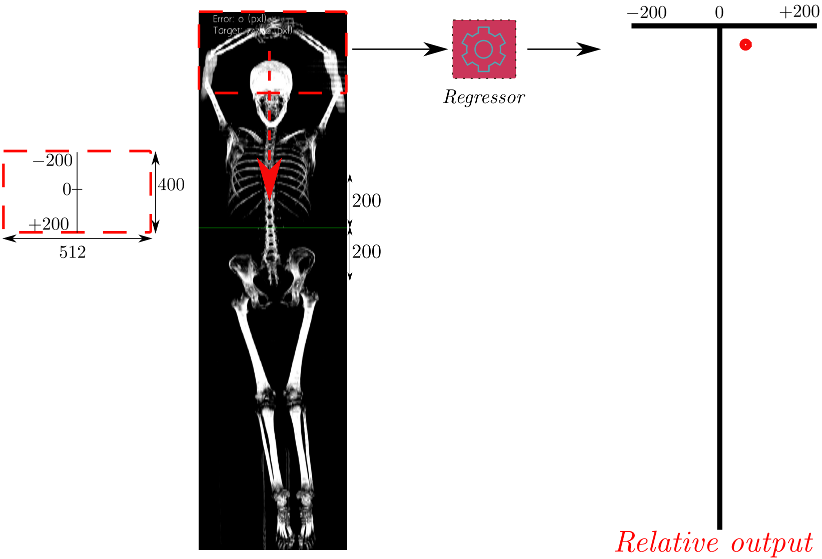
Issue 2: Different input size > Solution: Sliding window



1. MIP transformation
2. CNN prediction
3. Post processing (Correlation)

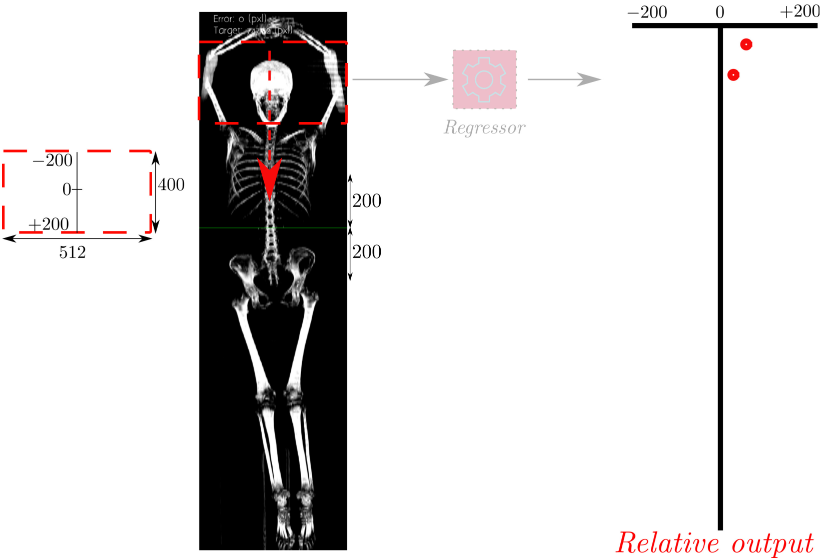
Sliding window

Issue 2: Different input size > Solution: Sliding window

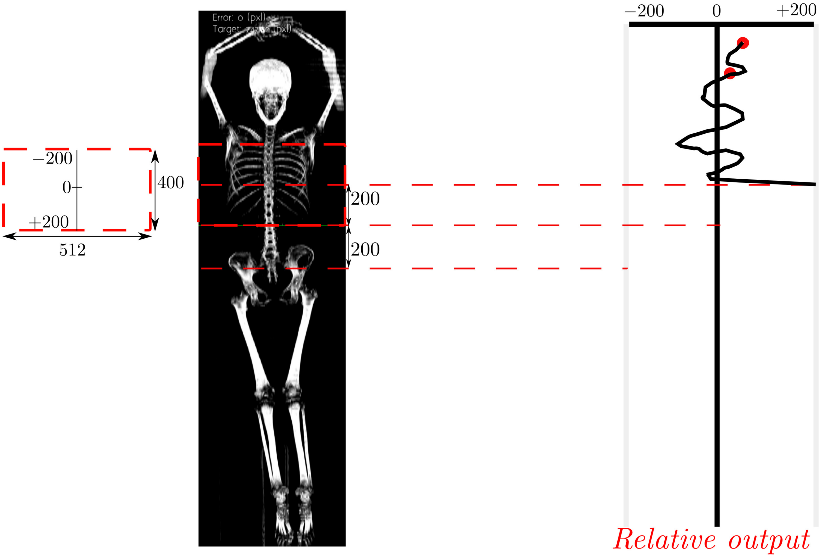


Proposed approach: Regression for L3 localization

Issue 2: Different input size > Solution: Sliding window

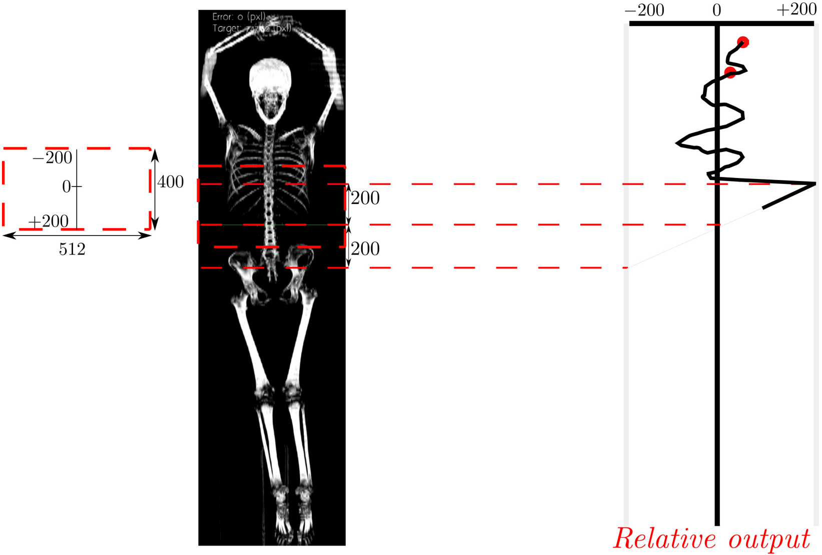


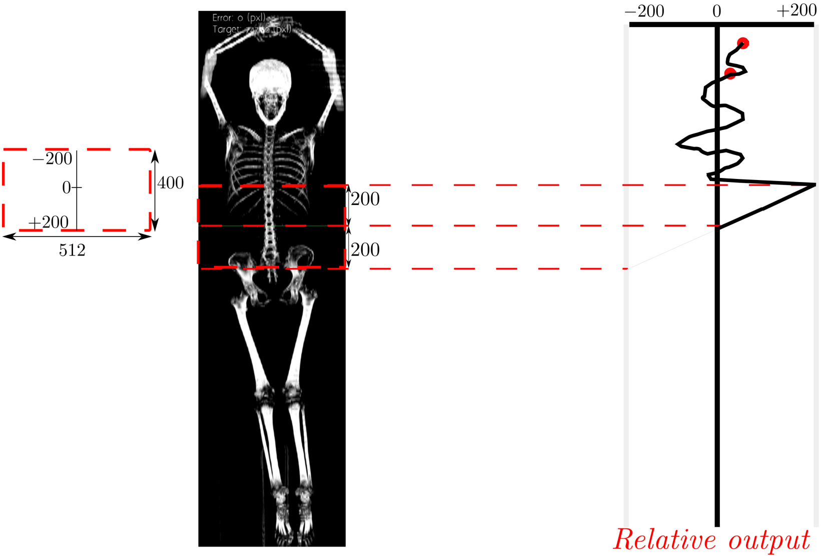
Issue 2: Different input size > Solution: Sliding window



Proposed approach: Regression for L3 localization

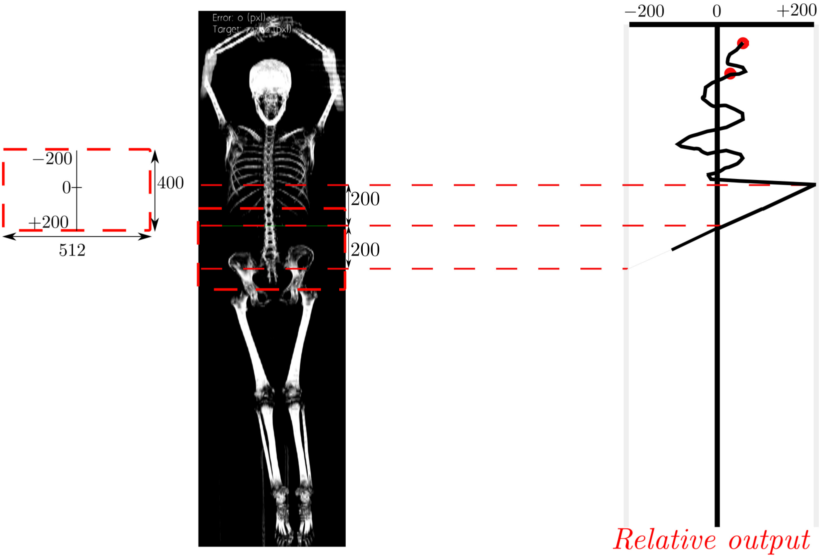
Issue 2: Different input size > Solution: Sliding window



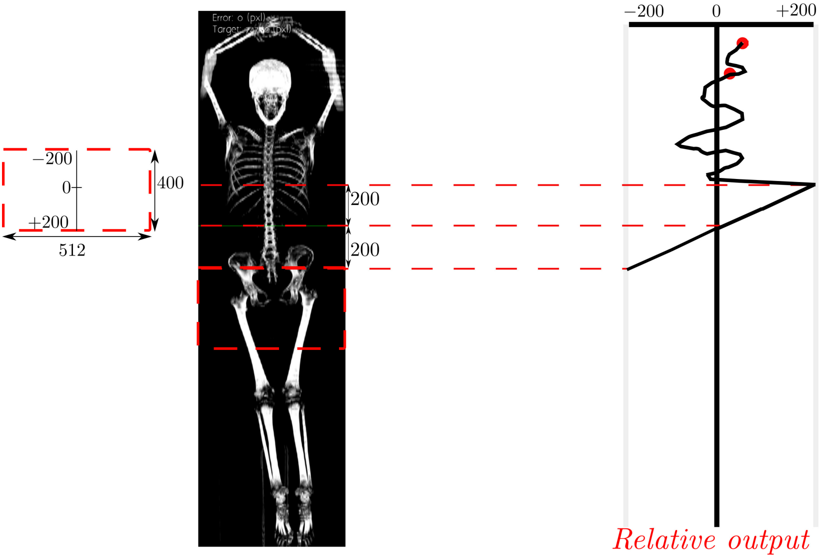

−200
0
+200
400
512
Error: 0 (pxl)
200
200
−200
0
+200
Relative output

Proposed approach: Regression for L3 localization

Issue 2: Different input size > Solution: Sliding window

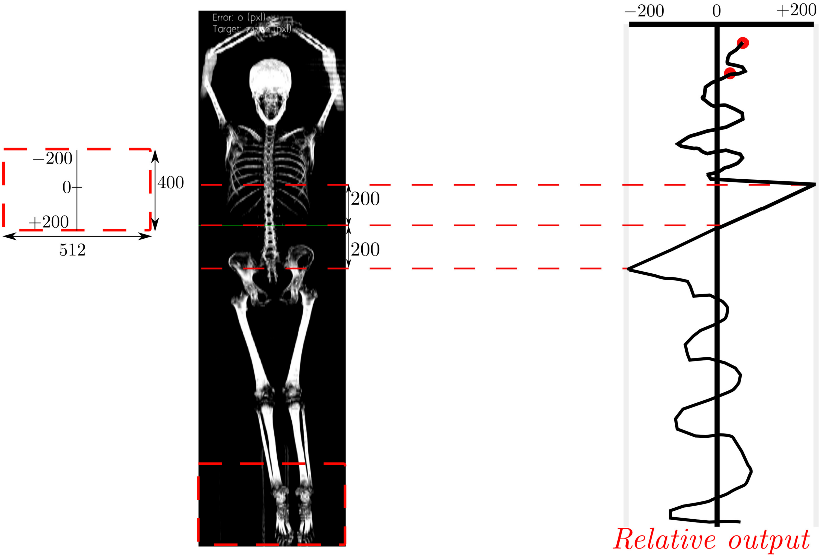


Issue 2: Different input size > Solution: Sliding window



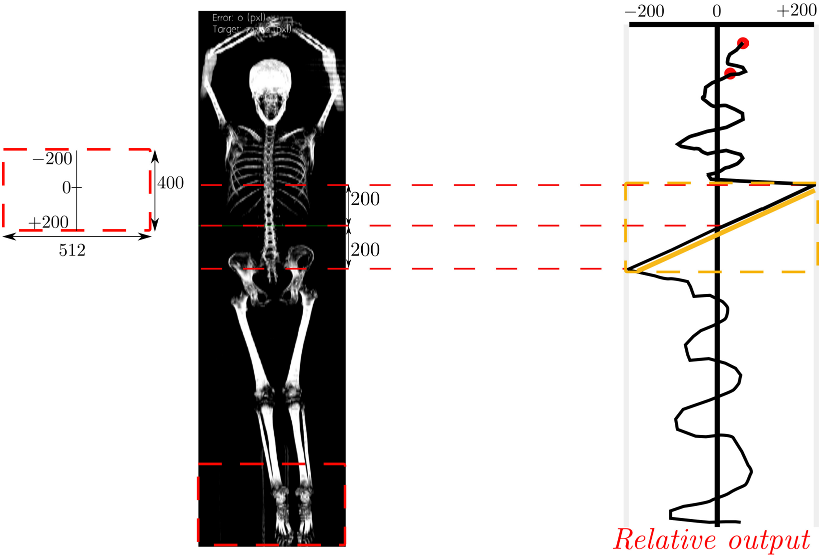
Proposed approach: Regression for L3 localization

Issue 2: Different input size > Solution: Sliding window

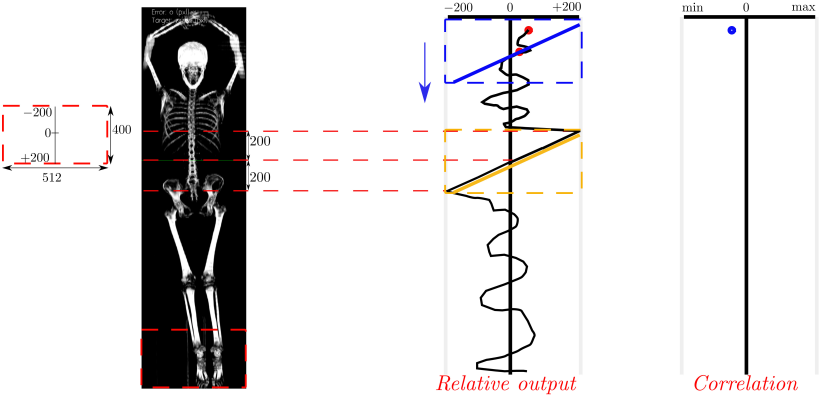


Proposed approach: Regression for L3 localization

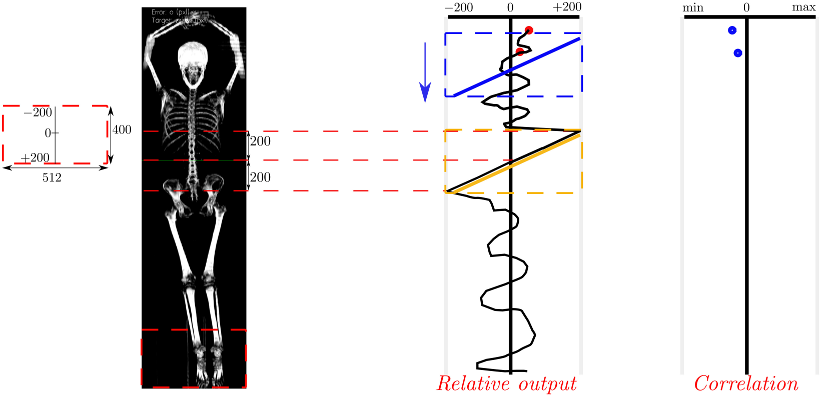
Issue 2: Different input size > Solution: Sliding window



Issue 2: Different input size > Solution: Sliding window

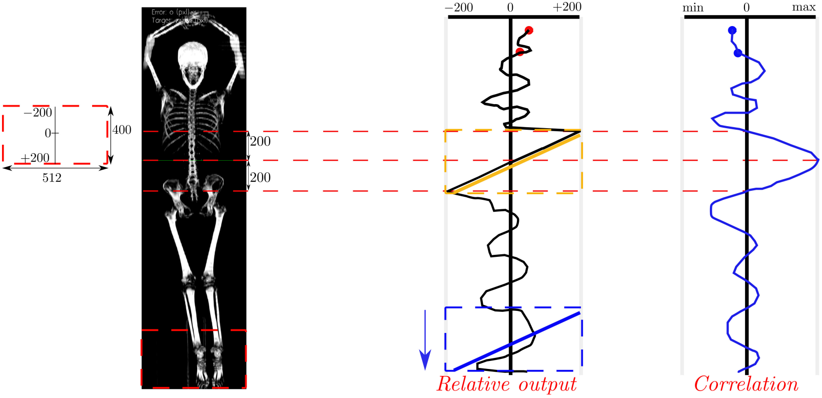


Issue 2: Different input size > Solution: Sliding window

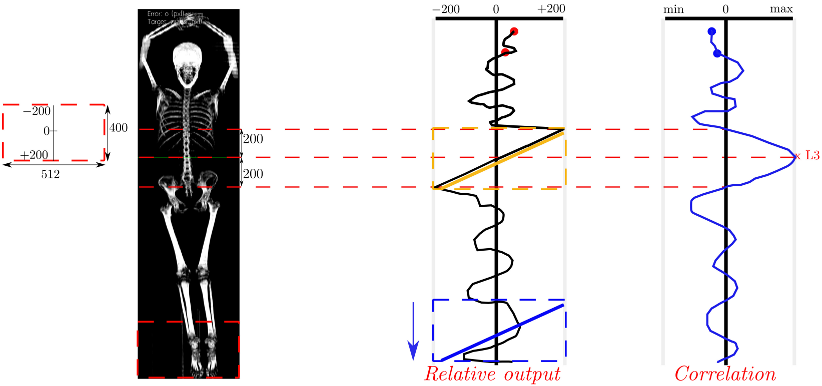


Proposed approach: Regression for L3 localization

Issue 2: Different input size > Solution: Sliding window

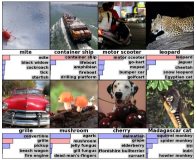


Issue 2: Different input size > Solution: Sliding window



Problem 2: Few data (642 patients)

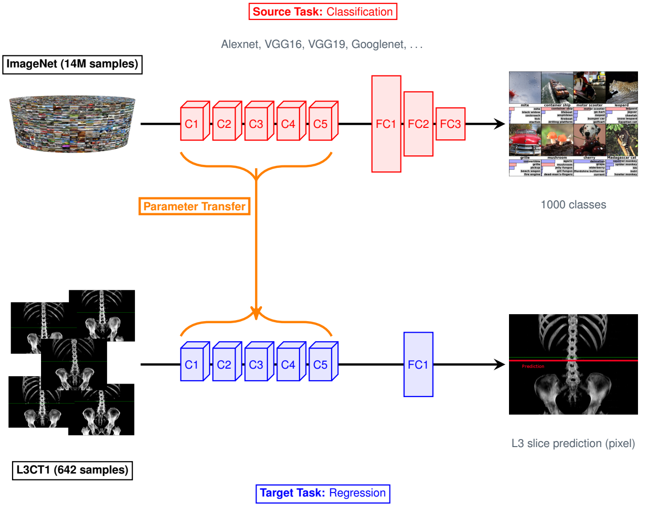
- Use pre-trained CNNs over **large datasets**
  - Alexnet, GoogleNet, VGG16, VGG19, . . . for **classification**
  - Pre-trained models over ImageNet: 14 millions of natural images [Fei-Fei and Russakovsky 2013].



Source task with abundant data.

Proposed approach: Regression for L3 localization

Issue 3: Lack of data > Solution: Transfer learning



System training using transfer learning.

Cross-validation:

			Pre-trained			
	RF500	CNN4	Alexnet	VGG16	VGG19	Googlenet
Average cross-validation error (5 folds) (slice)	10.50 ± 10.80	2.78 ± 2.48	2.45 ± 2.42	$\mathbf{1.82 \pm 2.32}$	1.83 ± 1.83	2.54 ± 4.22
Number of parameters	–	55 K	2 M	14 M	20 M	6[1] M
Average processing time (second/CT scan) (K40)	–	04.46	06.37	13.28	16.02	17.75[1]

RF500 (random forest with 500 decision trees), CNN4 (Homemade model), and Alexnet/VGG16/VGG19/GoogleNet (Pre-trained models).

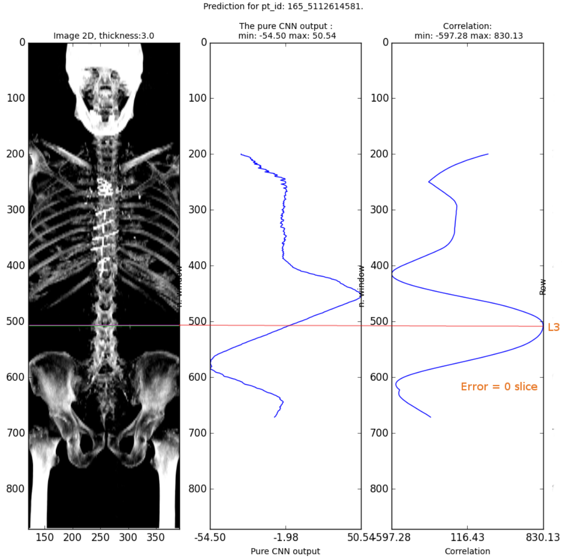
Possible speedup: reduce the number of sampled windows $\Rightarrow$ Increase stride.

Example VGG16:

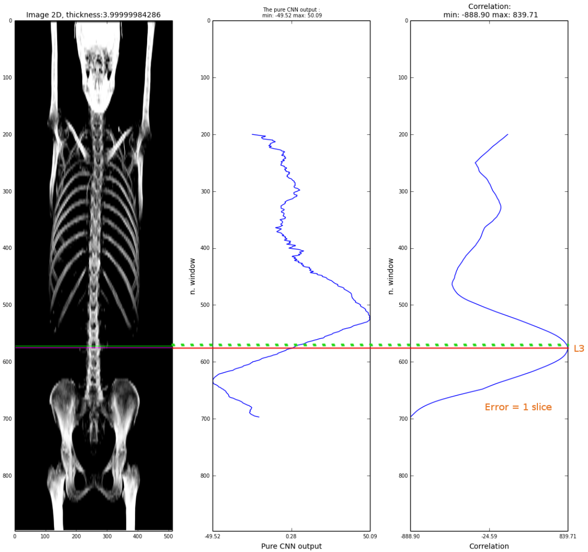
- **stride=1**: $\sim$ $\mathbf{13}$ ***seconds/CT scan*** with a an error of $\mathbf{1.82 \pm 2.32}$.
- **stride=4**: $\sim$ $\mathbf{02}$ ***seconds/CT scan*** with a an error of $\mathbf{1.91 \pm 2.69}$.

1. Due to implementation.

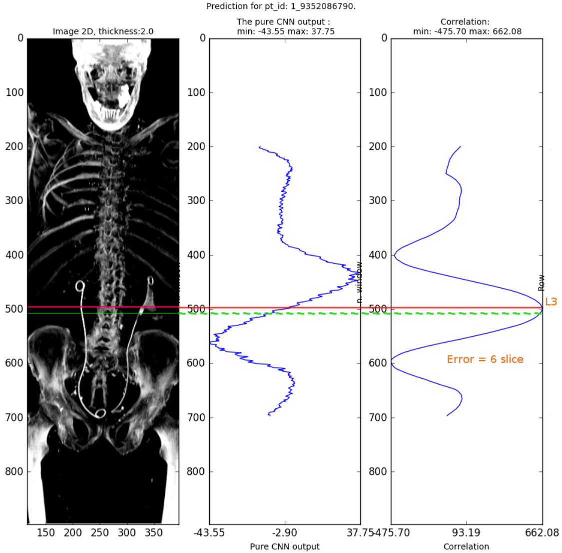
Experiments: Qualitative results



Experiments: Qualitative results



Experiments: Qualitative results



Setup: Intra-annotator variability

- New evaluation set: 43 CT scans annotated by the same reference radiologist (who annotated the L3CT1 dataset).
- Ask 3 other radiologists to localize the L3 slice.
- Perform this experiment twice: t_1, t_2.

Errors (slices) / operator	Ragiologist $\#1$	Radiologist $\#2$	Radiologist $\#3$
t_1	0.81 ± 0.97	0.72 ± 1.51	0.51 ± 0.62
t_2	0.77 ± 0.68	0.95 ± 1.61	0.86 ± 1.30

Intra-annotator variability.

Setup: Intra-annotator variability

- New evaluation set: 43 CT scans annotated by the same reference radiologist (who annotated the L3CT1 dataset).
- Ask 3 other radiologists to localize the L3 slice.
- Perform this experiment twice: t_1, t_2.

Errors (slices) / operator	Ragiologist $\#1$	Radiologist $\#2$	Radiologist $\#3$	CNN4	VGG16
t_1	0.81 ± 0.97	0.72 ± 1.51	0.51 ± 0.62	2.37 ± 2.30	1.70 ± 1.65
t_2	0.77 ± 0.68	0.95 ± 1.61	0.86 ± 1.30	2.53 ± 2.27	1.58 ± 1.83

Performance radiologists vs. automatic systems.

Conclusion

- Adapted pipeline for L3 localization: pre-processing, CNN, post-processing.
- Obtained average error: 1.82 slice ($< 5mm$) (maximum error: 9 slices).
  - Average thickness of a vertebra $\approx 2.5cm \Rightarrow$ Still within the L3 vertebra.
- Learn context: sliding window (double checked using correlation: context over multiple windows.)
- Generic framework: can be easily adapted for detecting other subjects given the required annotation.
- Use of transfer learning alleviates the lack of training data.

Perspectives: ⚠ Running time of VGG16 over CPUs is time consuming.
- Possible solution: Prune unnecessary convolution filters.

Valorization:
- Integrate this work with the software of the projet "**BodyComp.AI**" (diffused to European centers for cancer treatment).
- "**BodyComp.AI**" has won one of the 2017 French Innovative Unicancer Prize.

Publications:

- S. Belharbia, C. Chatelain, R. Hérault, S. Adam, S. Thureau, M. Chastan, and R. Modzelewski. ***Spotting L3 slice in CT scans using deep convolutional network and transfer learning***, Computers in Biology and Medicine, vol. 87, pp. 95-103, 2017.

Thank you for your attention!

Questions?

soufiane.belharbi@insa-rouen.fr
sbelharbi.github.io

Computation resource

UNIVERSITÉ DE ROUEN NORMANDIE

INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES ROUEN NORMANDIE

UFR Sciences et Techniques's data center

INSA Rouen Normandie