

Apprentissage profond et superpixels pour la segmentation d'images

Applications à la segmentation d'organes en imagerie abdominale

Pierre-Henri Conze

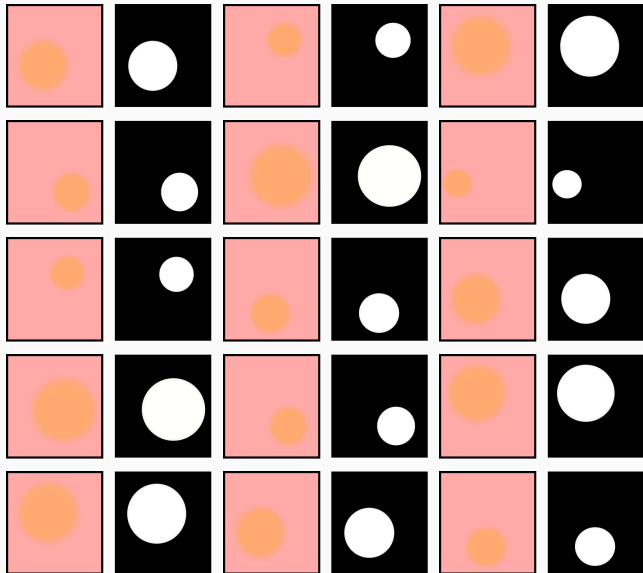
pierre-henri.conze@imt-atlantique.fr

IMT Atlantique, LaTIM UMR 1101, Inserm

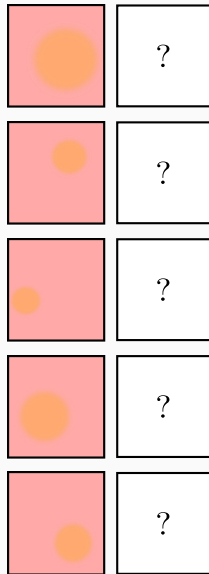
Introduction

Segmentation d'images par apprentissage

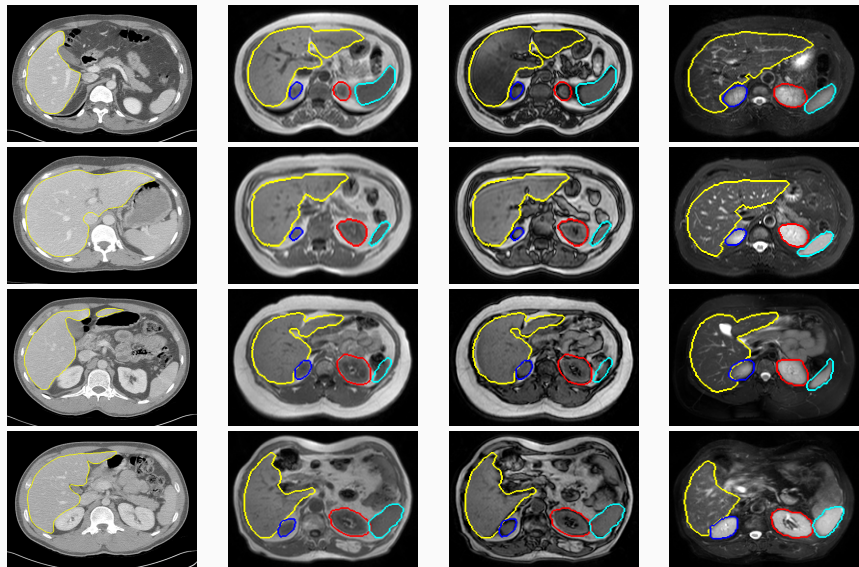
Apprentissage



Test



Challenge CHAOS



CT

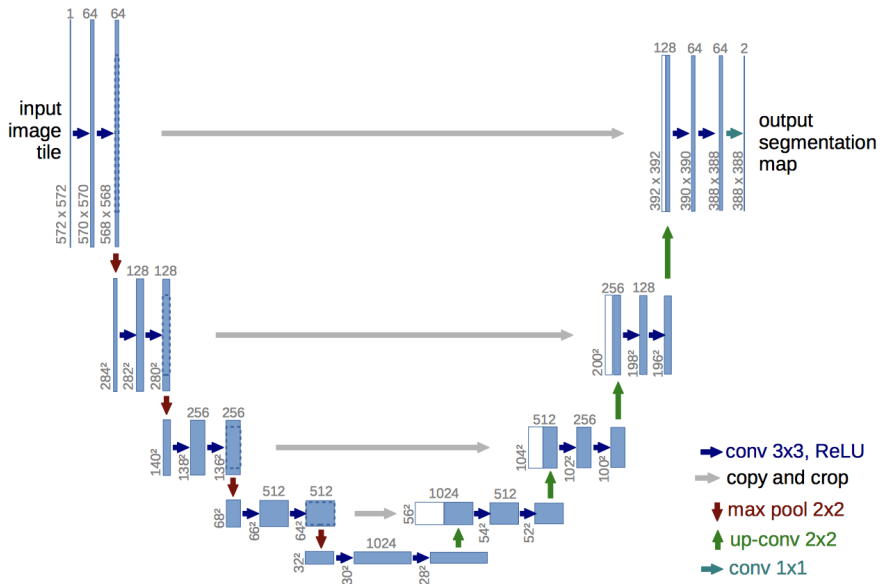
T1DUALin

T1DUALout

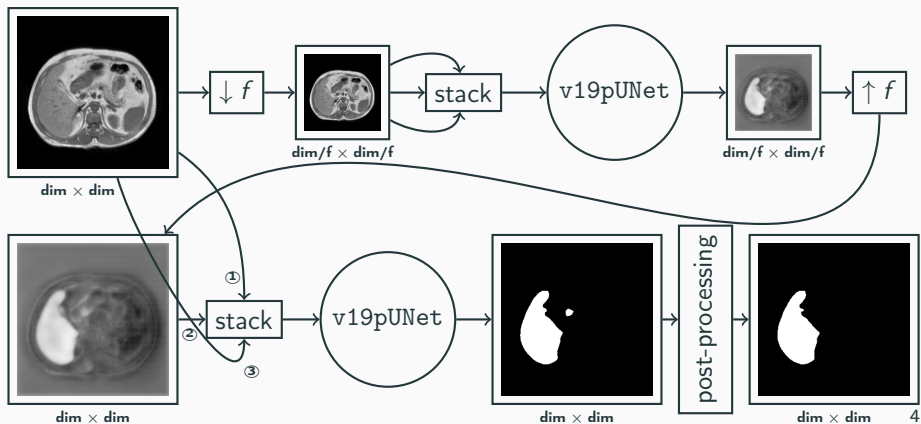
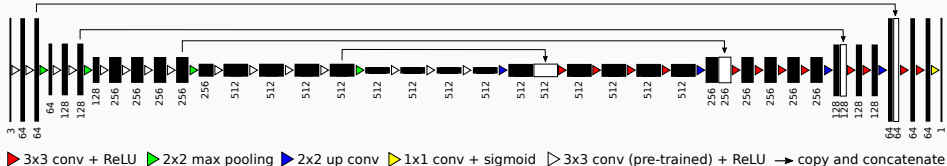
T2SPIR

<https://chaos.grand-challenge.org>

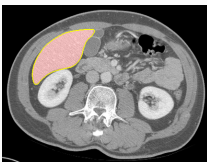
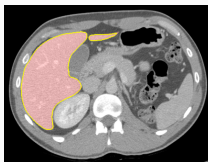
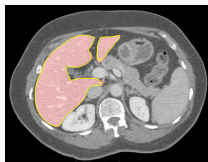
U-Net [Ronneberger et al., 2015]



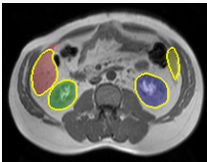
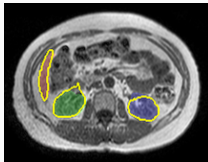
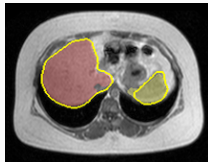
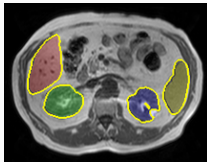
Architectures dérivées de U-Net



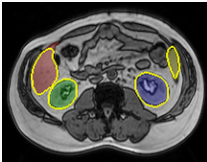
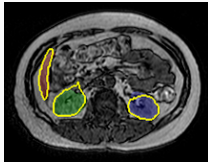
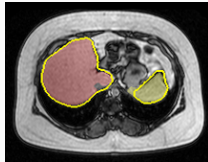
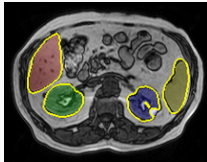
CT



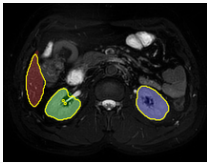
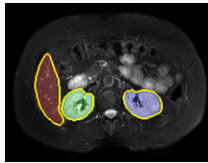
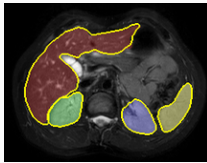
T1-DUALin



T1-DUALout

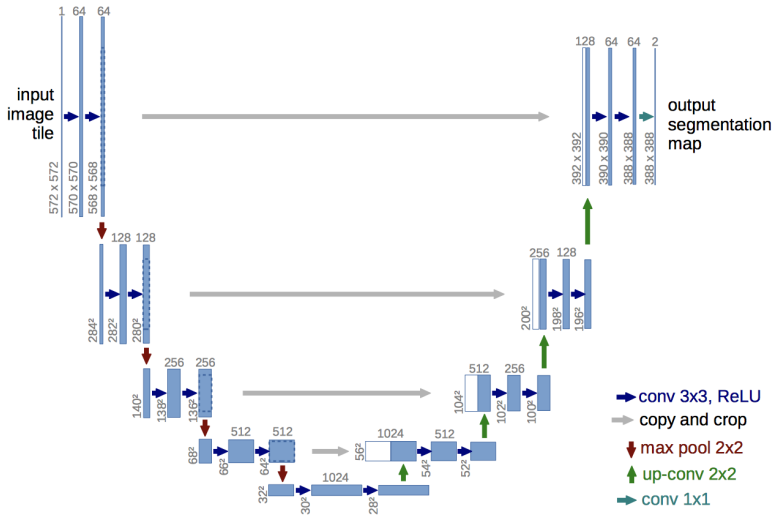


T2-SPiR



— prédiction ■ foie ■ rein droit ■ rein gauche ■ rate

Limitations



descripteurs calculés avec supports fixes → erreurs de délimitations (contours d'objets)
absence d'information de contexte large échelle → relations inter-objets sous-estimées

Décomposition en superpixels

Définition

superpixels : primitives visuelles générées par agrégation de pixels voisins partageant des caractéristiques similaires

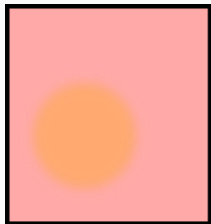
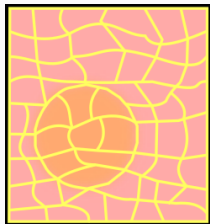


image source I



superpixels f_i

$\mathcal{F} = \{f_i\}_{i \in \{1, \dots, |\mathcal{F}|\}}$ ensemble de $|\mathcal{F}|$ superpixels décomposant l'image I

Algorithmes [Stutz et al., 2018]

watershed-based : waterpixels

[Machairas et al., 2015]

density-based : quickshift

[Vedaldi and Soatto, 2008]

graph-based : Felzenszwalb

[Felzenszwalb and Huttenlocher, 2004]

contour evolution

patch-based

clustering-based : SLIC

[Achanta et al., 2012]

energy optimization

wavelet-based

...

Formulation du problème

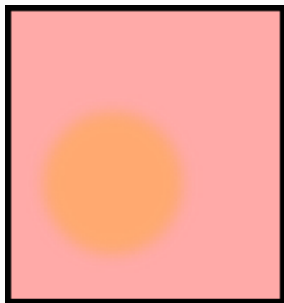
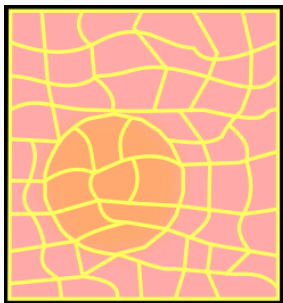
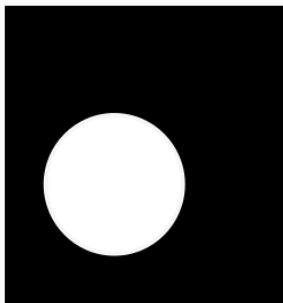


image source



superpixels

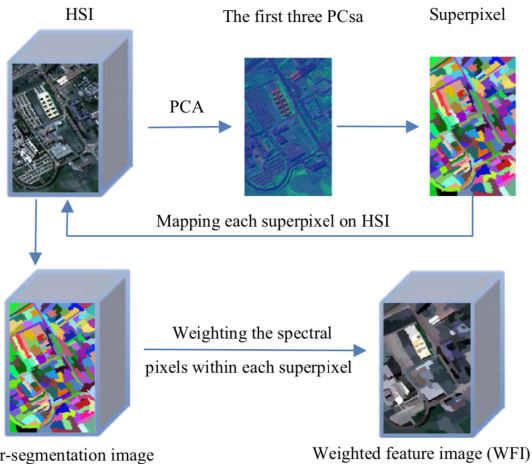


segmentation

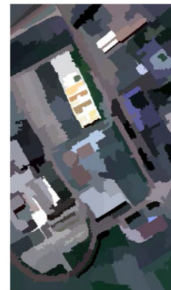
Apprentissage profond et superpixel pour la segmentation d'images ?

Travaux existants

Superpixel-based deep neural networks [Shi and Pun, 2018]

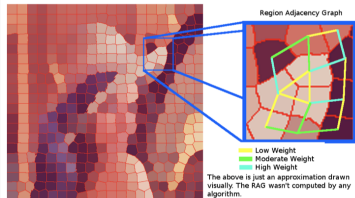
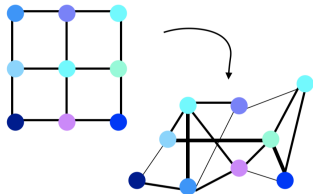
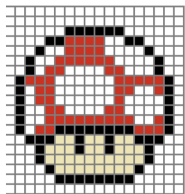


False color HSI
(bands 30,40,50)



False color WFI
(bands 30,40,50)

Convolutions sur domaines non-réguliers



Domaine

Données régulières

Grille régulière des pixels

Même nombre de voisins

Ordonnancement intrinsèque

Descripteurs

Intensité du pixel

Tâches

Classification d'images

Segmentation d'images

Données irrégulières

Structure de graphe

Nombre variable de voisins par noeud

Par d'ordonnancement

Vecteur associé à chaque noeud

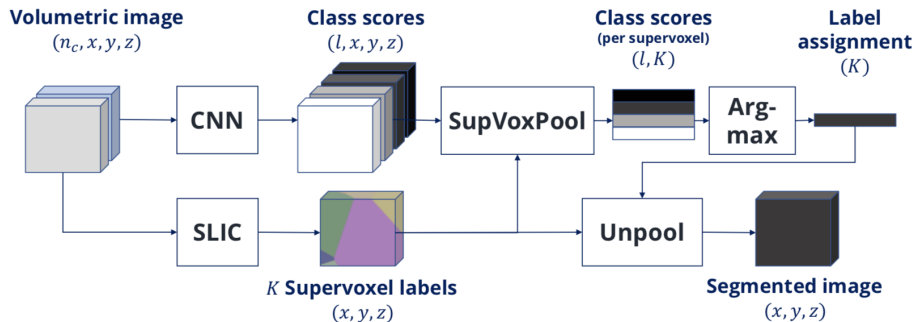
Classification du graphe entier

Classification de chaque noeud

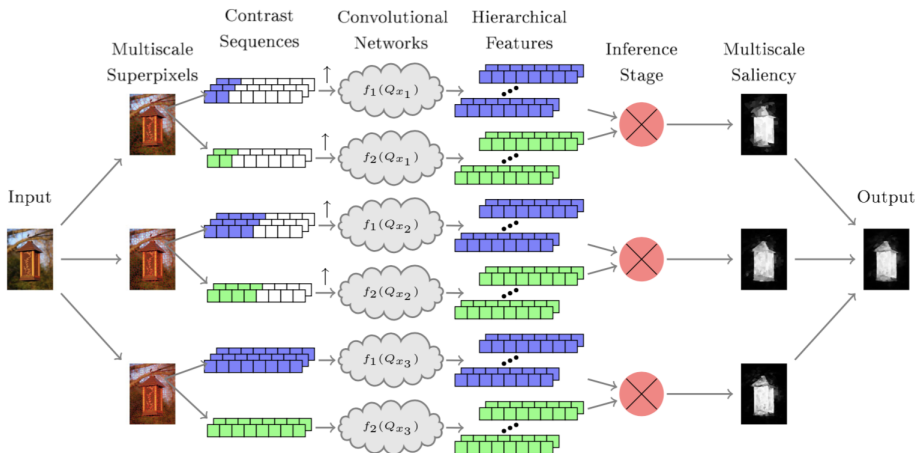
S. Parisot, *Deep learning on graphs : convolutional neural networks on irregular domains*

<https://biomedic.doc.ic.ac.uk/miccai17-grail/talks/parisot.pdf>

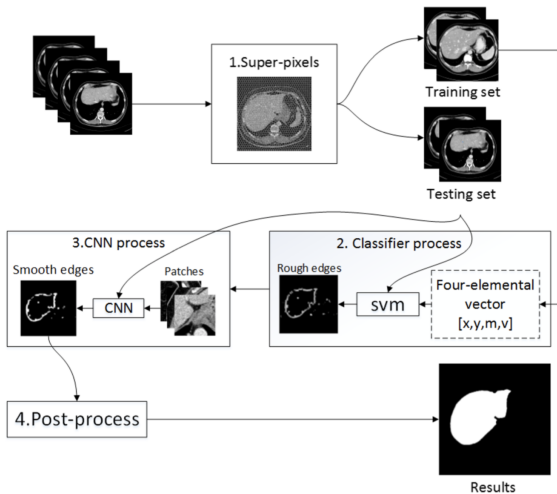
Supervoxel pooling [Schuermans et al., 2018]



Super-CNN [He et al., 2015]



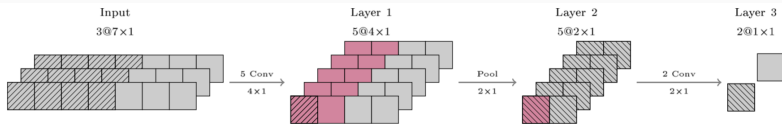
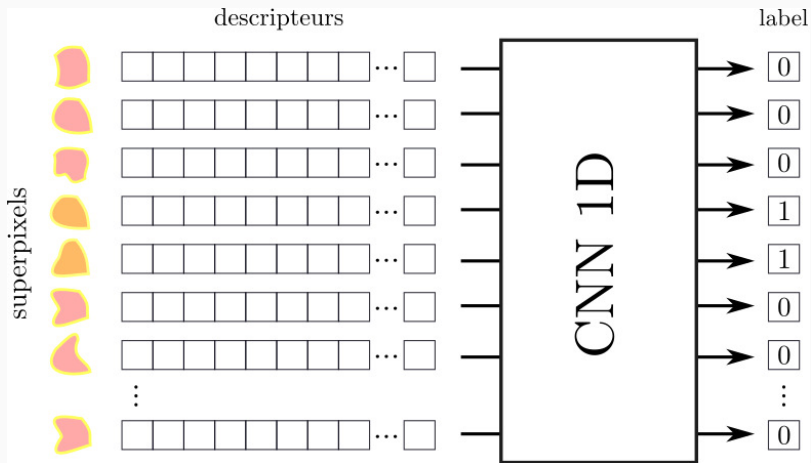
Superpixel signature SVM classification + CNN [Liu et al., 2018]



1. signature associée à chaque superpixel : positions, moyenne/variance intensités
2. classification SVM : all-organ area (1), mixed area (0), non-organ area (-1)
3. CNN basé patch pour raffiner les contours
4. opérations morphologiques

Méthode

Classification de superpixels

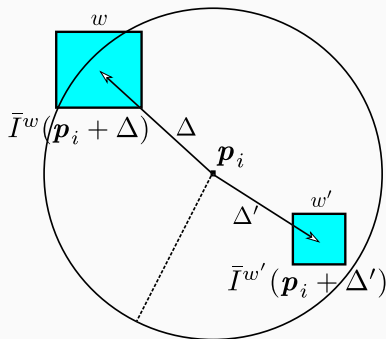


Descripteurs de contexte basés pixels [Criminisi et al., 2009]

$\bar{I}^w(\mathbf{p}_i)$: intensité moyenne dans une fenêtre de taille w centré autour du pixel \mathbf{p}_i

Pour chaque pixel $\mathbf{p}_i = (x_i, y_i)$, N descripteurs de contexte peuvent être

calculés comme suit : $\phi_n(\mathbf{p}_i) = \bar{I}^w(\mathbf{p}_i + \Delta) - \beta \times \bar{I}^{w'}(\mathbf{p}_i + \Delta')$



Une fois générés aléatoirement, les jeux de paramètres $\{w, w', \Delta, \Delta', \beta \in \{0, 1\}\}$ restent inchangés quelque soient le pixel \mathbf{p}_i et l'image I

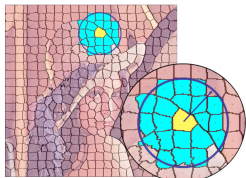
Descripteurs de contexte basés superpixels

Soit f_i un superpixel de \mathcal{F} de barycentre $c(f_i)$ et d'intensité moyenne $\bar{I}(f_i)$

Superpatch [Giraud et al., 2017]

$P_r(f_i)$ de rayon r agrège les supervoxels $f_l \in \mathcal{F}$ tel que $\|c(f_l) - c(f_i)\|_2 \leq r$

- correspond aux superpixels dont le barycentre est situé dans un cercle de rayon r centrée en $c(f_i)$
- contient au moins f_i



Intensité moyenne d'un superpatch

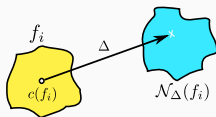
Moyenne des intensités moyennes estimées pour chaque

$$f_l \in P_r(f_i) : \bar{I}[P_r(f_i)] = \frac{1}{|P_r(f_i)|} \sum_{f_l \in P_r(f_i)} \bar{I}(f_l)$$

Supervoxels distants

$\mathcal{N}_\Delta(f_i)$ correspond au superpixel du voisinage étendue de f_i obtenu en appliquant le déplacement Δ à partir de $c(f_i)$

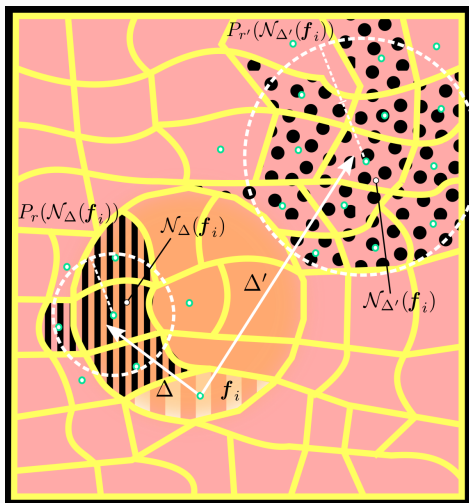
→ $\mathcal{N}_\Delta(f_i)$ est donc le superpixel contenant la position $c(f_i) + \Delta$



Descripteurs de contexte basés superpixels

N descripteurs de contexte sont calculés pour chaque superpixel \mathbf{f}_i comme

$$\text{suit : } \phi_n(\mathbf{f}_i) = \bar{I}[P_r(\mathcal{N}_\Delta(\mathbf{f}_i))] - \beta \times \bar{I}[P_{r'}(\mathcal{N}_{\Delta'}(\mathbf{f}_i))]$$



Les paramètres $\{r, r', \Delta, \Delta', \beta \in \{0, 1\}\}$ restent inchangés $\forall \mathbf{f}_i$ et $\forall I$

Chaîne de traitement

- 1 - collection d'images (training, test)
- 2 - décomposition en superpixels (training, test) : **maskSLIC**
- 3 - étiquetage des superpixel de training
- 4 - estimation de **descripteurs de contexte** basés superpixels (training, test)
- 5 - équilibrage inter-classe (training) : **algorithme SMOTE**
- 6 - **[optionnel]** réduction de dimension
- 7 - classification des superpixels de test : **CNN-1D, combinaison de CNN-1D + majority voting,...**
- 8 - reconstruction des cartes de segmentation 3D
- 9 - post-traitement : + grande composante connectée
- 10 - évaluation

Résultats préliminaires

Résultats préliminaires obtenus avec :

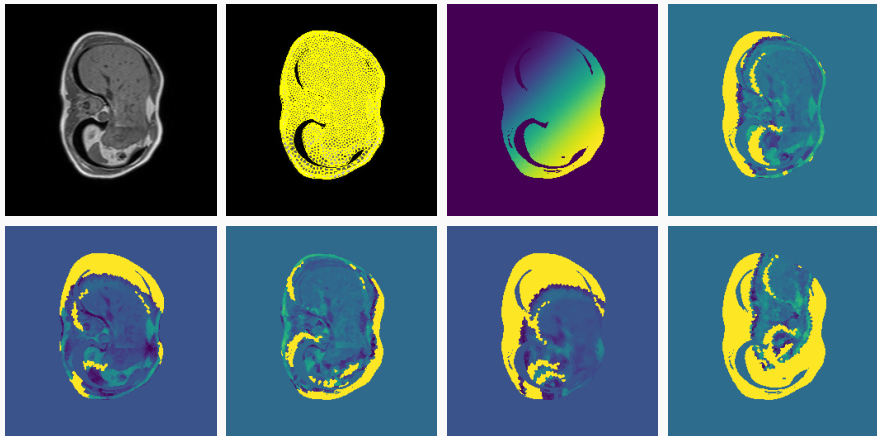
10 examens IRM T1 d'apprentissage, 10 de test

2000 superpixels par coupe axiale (uniquement zone utile)

400 descripteurs de contexte par superpixel

60% seuil de l'étiquetage des superpixels d'apprentissage

20 epochs (50 pour U-Net)

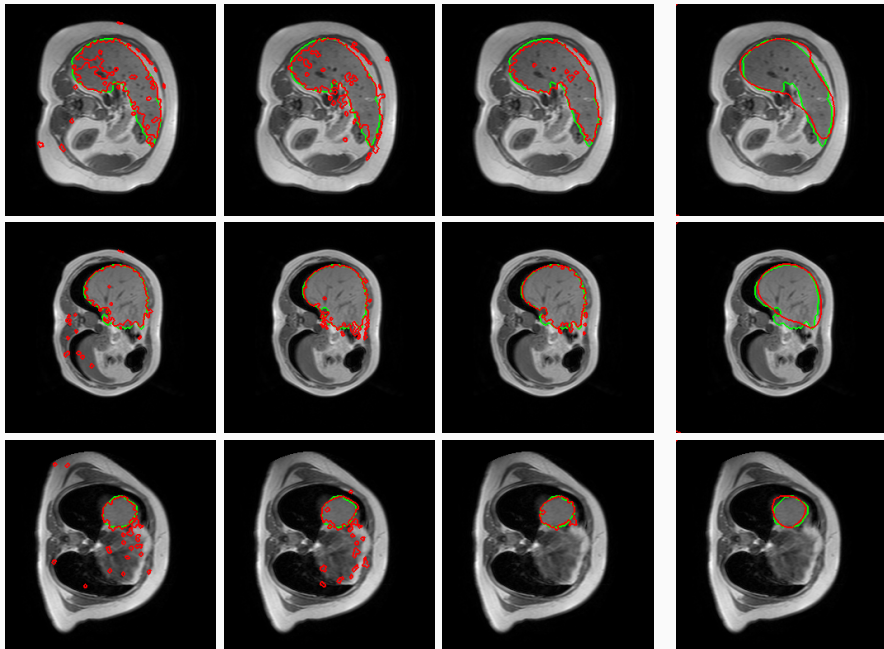


Forêts aléatoires

CNN-1D 5 conv + 2 FC

5 × CNN-1D + MV

U-Net sans data augment.



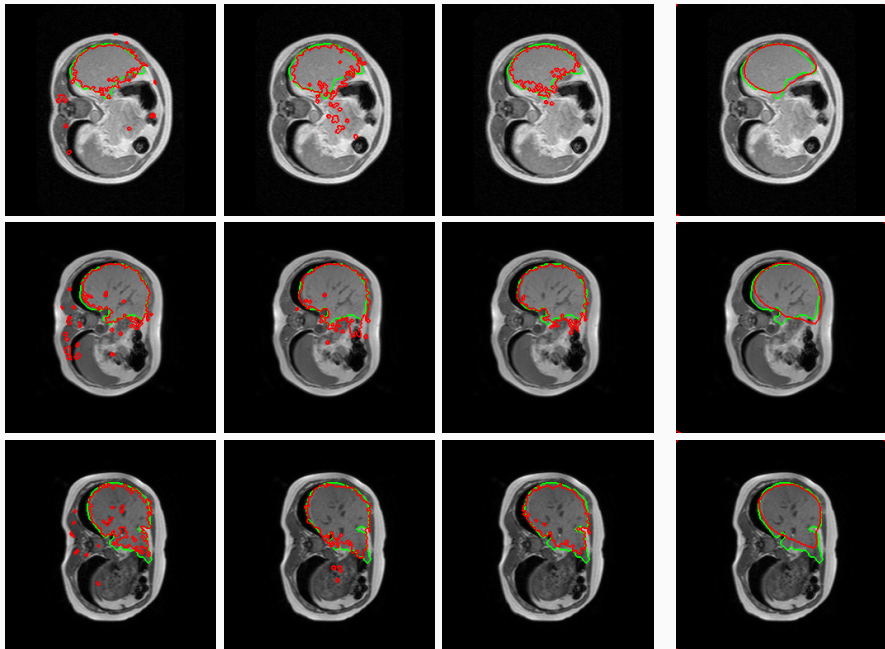
— prédiction — vérité-terrain

Forêts aléatoires

CNN-1D 5 conv + 2 FC

5 × CNN-1D + MV

U-Net sans data augment.



— prédiction — vérité-terrain

	dice	sens	spec	jacc	verr
Forêts aléatoires	72.90	70.43	97.76	57.78	216.9
CNN-1D _{2 conv} + 2 FC	77.80	79.18	97.70	64.31	131.3
CNN-1D _{5 conv} + 2 FC	78.71	<u>81.58</u>	97.56	65.71	156.3
5 CNN-1D + MV	81.44	74.92	99.20	69.22	257.0
CNN-1D _{5 conv} + 2 FC + PP	<u>83.28</u>	81.50	98.63	<u>71.72</u>	<u>81.2</u>
5 CNN-1D + MV + PP	83.12	74.75	<u>99.54</u>	71.43	321.0
U-Net	84.52	84.62	98.54	73.55	183.3

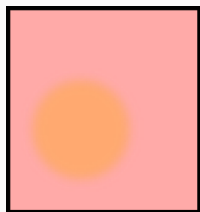
Perspectives

Augmentation de données

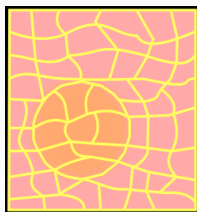
- 1 - collection d'images (training, test) + data augmentation
- 2 - décomposition en superpixels (training, test)
 - modification des paramètres de calcul des superpixels (compactness...)
 - utilisation de plusieurs méthodes de décomposition en superpixels
- 3 - étiquetage des superpixel de training
- 4 - estimation de descripteurs de contexte basés superpixels (training, test)
- 5 - équilibrage inter-classe (training) : algorithme SMOTE
- 6 - [optionnel] réduction de dimension
- 7 - classification des superpixels de test : CNN-1D, combinaison de CNN-1D + majority voting,...
- 8 - reconstruction des cartes de segmentation 3D
- 9 - post-traitement : + grande composante connectée
- 10 - évaluation

Décomposition en superpixels multi-échelles

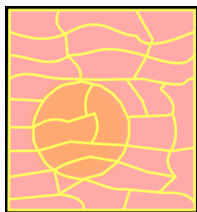
Décomposons $I: \Omega \subset \mathbb{N}^3 \rightarrow \mathbb{N}$ en $M+1$ partitions $\mathcal{P} = \{\mathcal{P}_0, \dots, \mathcal{P}_M\}$ où \mathcal{P}_m est un ensemble de superpixels $\{\mathbf{f}_i^m\}$ tels que $\mathbf{f}_i^m \cap \mathbf{f}_{j \neq i}^m = \emptyset$ et $\bigcup_i \mathbf{f}_i^m = \Omega$



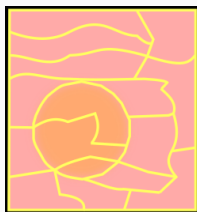
I



\mathcal{P}_6



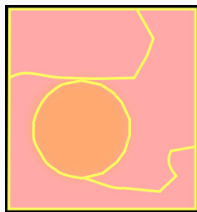
\mathcal{P}_5



\mathcal{P}_4



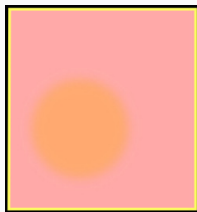
\mathcal{P}_3



\mathcal{P}_2



\mathcal{P}_1

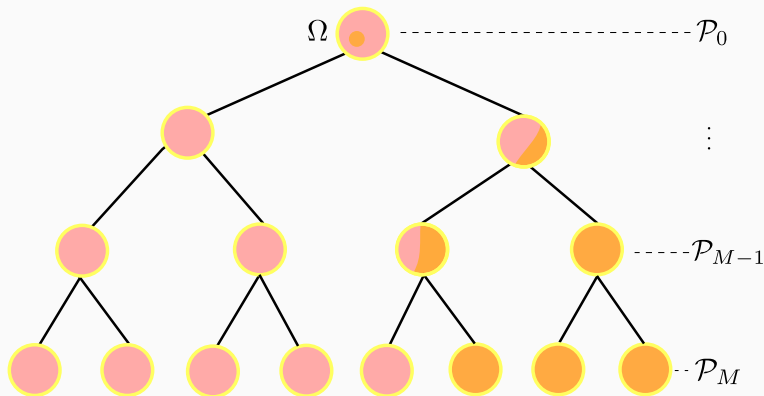


$\mathcal{P}_0 = \Omega$

Représentation multi-échelles

Une fois \mathcal{P}_M obtenue par décomposition en superpixels, les partitions \mathcal{P}_m avec $0 \leq m < M$ sont obtenues par fusion itérative des superpixels (de \mathcal{P}_M vers \mathcal{P}_0).

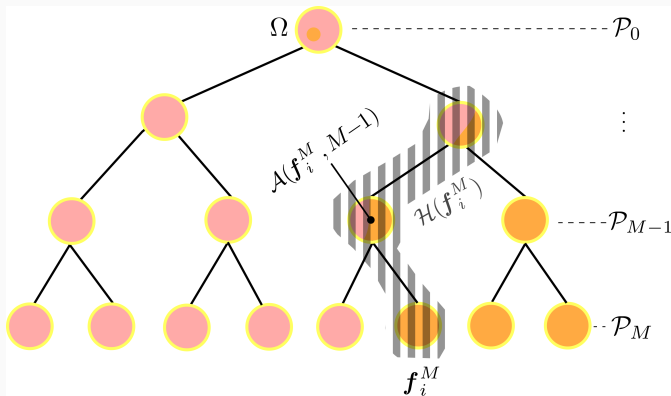
La séquence de partitions $\mathcal{P} = \{\mathcal{P}_m\}$ est encodée au sein des couches d'un arbre multi-échelles $\mathcal{Q} = \{\mathcal{Q}_m\}$. La couche \mathcal{Q}_m associe à chaque superpixel $\mathbf{f}_i^m \in \mathcal{P}_m$ un ensemble de L superpixels enfants $\{\mathbf{f}_j^{m+1}\} \subset \mathcal{P}_{m+1}$ tel que $\mathbf{f}_i^m = \bigcup_j \mathbf{f}_j^{m+1}$.



Notations

Soient :

- $\iota(\mathbf{f}_i) = \{0, 1\}$ la classe associée au superpixel \mathbf{f}_i (segmentation binaire)
- $\mathcal{A}(\mathbf{f}_i^m, s)$ l'ancêtre de \mathbf{f}_i^m d'échelle s avec $m > 1$ et $1 \leq s < m$ ($\mathcal{P}_0 = \Omega$)
- la hiérarchie $\mathcal{H}(\mathbf{f}_i^M)$, l'ensemble des ancêtres de \mathbf{f}_i^M dans \mathcal{Q}
 - $\mathcal{H}(\mathbf{f}_i^M) = \{\mathbf{f}_i^M, \mathcal{A}(\mathbf{f}_i^M, M-1), \mathcal{A}(\mathbf{f}_i^M, M-2), \dots, \mathcal{A}(\mathbf{f}_i^M, 1)\}$



Simple-échelle

① Descripteurs de contexte simple-échelle : $\Phi(\mathbf{f}_i) = \{\phi_n(\mathbf{f}_i)\}$ avec

$$\phi_n(\mathbf{f}_i) = \bar{I}[P_r(\mathcal{N}_\Delta(\mathbf{f}_i))] - \beta \times \bar{I}[P_{r'}(\mathcal{N}_{\Delta'}(\mathbf{f}_i))]$$

② Classification de superpixels : $\text{CNN1D}(\Phi(\mathbf{f}_i)) = P(\iota(\mathbf{f}_i) = 1)$

Objectif commun aux stratégies multi-échelles

→ classification des superpixels $\mathbf{f}_i^M \in \mathcal{P}_M$: $P(\iota(\mathbf{f}_i^M) = 1) \forall i$

Multi-échelles : stratégie 1 [Conze et al., 2017a]

① Descripteurs de contexte simple-échelle calculés à chaque échelle :

$$\Phi(\mathbf{f}_i^m) = \{\phi_n(\mathbf{f}_i^m)\}$$

② Concatenation des descripteurs calculés pour les superpixels de la hiérarchie $\mathcal{H}(\mathbf{f}_i^M)$: $\Phi(\mathcal{H}(\mathbf{f}_i^M)) = [\Phi(\mathbf{f}_i^M), \Phi(\mathbf{f}_i^{M-1}), \dots, \Phi(\mathbf{f}_i^1)]$

③ Classification de superpixels : $\text{CNN1D}(\Phi(\mathcal{H}(\mathbf{f}_i^M))) = P(\iota(\mathbf{f}_i^M) = 1)$

Stratégies d'extension multi-échelles

Multi-échelles : stratégie 2 inspiré de [Conze et al., 2017b]

① Descripteurs de contexte simple-échelle calculés à chaque échelle :

$$\Phi(\mathbf{f}_i^m) = \{\phi_n(\mathbf{f}_i^m)\}$$

② Classification des superpixels de chaque échelle indépendamment :

$$\text{CNN1D}(\Phi(\mathbf{f}_i^m)) = \text{P}(\iota(\mathbf{f}_i^m) = 1) \quad \forall m$$

③ Fusion de résultats par hiérarchie :

$$\text{P}^*(\iota(\mathbf{f}_i^M) = 1) = \prod_{\mathbf{f}_i^m \in \mathcal{H}(\mathbf{f}_i^M)} \text{P}(\iota(\mathbf{f}_i^m) = 1)$$

Multi-échelles : stratégie 3

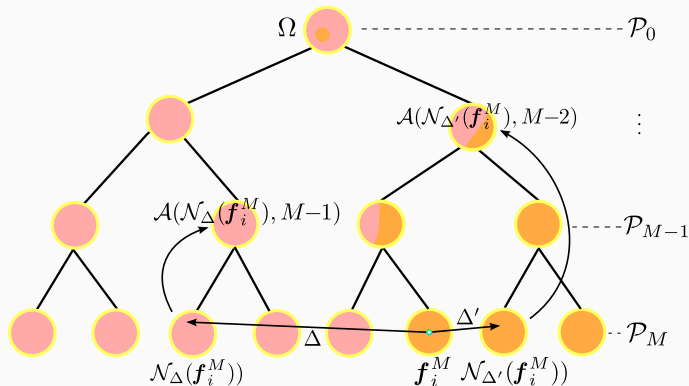
① Descripteurs de contexte multi-échelles :

$\phi_n(\mathbf{f}_i^M) = \bar{l}[\mathcal{A}(\mathcal{N}_\Delta(\mathbf{f}_i^M), s)] - \beta \times \bar{l}[\mathcal{A}(\mathcal{N}_{\Delta'}(\mathbf{f}_i^M), s')]$ avec les paramètres $\{s, s', \Delta, \Delta', \beta\}$ restant inchangés $\forall \mathbf{f}_i^M \in \mathcal{P}_M$ et $\forall l$

② Classification de superpixels : $\text{CNN1D}(\Phi(\mathbf{f}_i^M)) = \text{P}(\iota(\mathbf{f}_i^M) = 1)$

Descripteurs de contexte basés superpixels multi-échelles

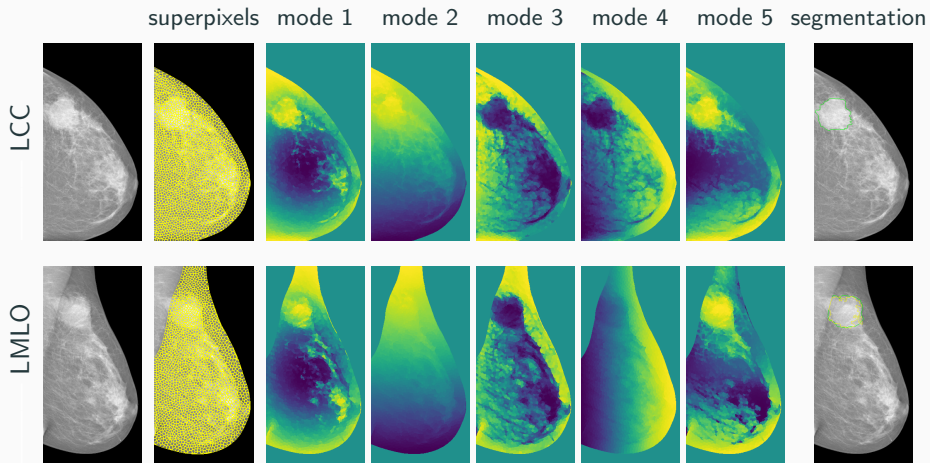
N descripteurs de contexte sont calculés pour chaque superpixel \mathbf{f}_i^M comme suit :

$$\phi_n(\mathbf{f}_i^M) = \bar{I}[\mathcal{A}(\mathcal{N}_\Delta(\mathbf{f}_i^M), s)] - \beta \times \bar{I}[\mathcal{A}(\mathcal{N}_{\Delta'}(\mathbf{f}_i^M), s')]$$


L'utilisation des superpatches n'est ici pas nécessaire dans la mesure où exploiter les ancêtres permet une bonne prise en compte du contexte spatial étendu.

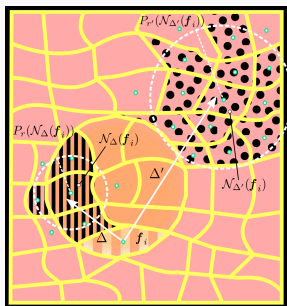
Les paramètres $\{s, s', \Delta, \Delta', \beta \in \{0, 1\}\}$ restent inchangés $\forall \mathbf{f}_i^M \in \mathcal{P}_M$ et $\forall l$

Fusion d'informations multi-vues [Tilquin et al., 2018]



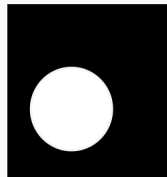
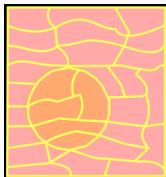
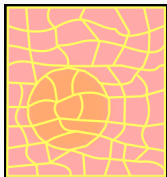
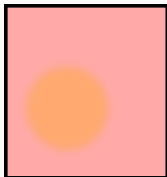
Conclusion

Conclusion et perspectives

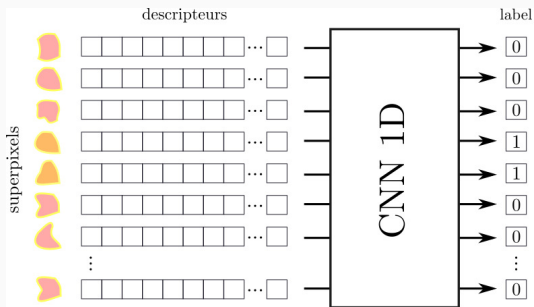


descripteurs de contexte
basés superpixels

extension multi-échelles



fusion d'informations (multi-vues, longitudinales...)



classification de superpixels par deep learning
augmentation données

-  Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Susstrunk, S. (2012). **SLIC superpixels compared to state-of-the-art superpixel methods.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11) :2274–2282.
-  Conze, P.-H., Noblet, V., Rousseau, F., Heitz, F., De Blasi, V., Memeo, R., and Pessaux, P. (2017a). **Scale-adaptive supervoxel-based random forests for liver tumor segmentation in dynamic contrast-enhanced ct scans.** *International journal of computer assisted radiology and surgery*, 12(2) :223–233.
-  Conze, P.-H., Tilquin, F., Noblet, V., Rousseau, F., Heitz, F., and Pessaux, P. (2017b). **Hierarchical multi-scale supervoxel matching using RF for automatic semi-dense abdominal image registration.** In *International Symposium on Biomedical Imaging*.
-  Criminisi, A., Shotton, J., and Bucciarelli, S. (2009). **Decision forests with long-range spatial context for organ localization in ct volumes.** In *Medical Image Computing and Computer-Assisted Intervention*, pages 69–80.
-  Felzenszwalb, P. F. and Huttenlocher, D. P. (2004). **Efficient graph-based image segmentation.** *International journal of computer vision*, 59(2) :167–181.
-  Giraud, R., Ta, V. T., Bugeau, A., Coupé, P., and Papadakis, N. (2017). **Superpatchmatch : An algorithm for robust correspondences using superpixel patches.** *IEEE Transactions on Image Processing*, 26(8) :4068–4078.
-  He, S., Lau, R. W., Liu, W., Huang, Z., and Yang, Q. (2015). **Supercnn : A superpixelwise convolutional neural network for salient object detection.** *International journal of computer vision*, 115(3) :330–344.
-  Liu, X., Guo, S., Yang, B., Ma, S., Zhang, H., Li, J., Sun, C., Jin, L., Li, X., Yang, Q., et al. (2018). **Automatic organ segmentation for ct scans based on super-pixel and convolutional neural networks.** *Journal of digital imaging*, 31(5) :748–760.
-  Machairas, V., Faessel, M., Cárdenas-Peña, D., Chabardes, T., Walter, T., and Decencière, E. (2015). **Waterpixels.** *IEEE Transactions on Image Processing*, 24(11) :3707–3716.
-  Ronneberger, O., Fischer, P., and Brox, T. (2015). **U-net : Convolutional networks for biomedical image segmentation.** In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.
-  Schuurmans, M., Berman, M., and Blaschko, M. B. (2018). **Efficient semantic image segmentation with superpixel pooling.** *arXiv preprint arXiv :1806.02705*.
-  Shi, C. and Pun, C.-M. (2018). **Superpixel-based 3d deep neural networks for hyperspectral image classification.** *Pattern Recognition*, 74 :600–616.
-  Stutz, D., Hermans, A., and Leibe, B. (2018). **Superpixels : An evaluation of the state-of-the-art.** *Computer Vision and Image Understanding*, 166 :1–27.
-  Tilquin, F., Conze, P.-H., Pessaux, P., Lamard, M., Queller, G., Noblet, V., and Heitz, F. (2018). **Robust supervoxel matching combining mid-level spectral and context-rich features.** In *International Workshop on Patch-based Techniques in Medical Imaging*, pages 39–47.
-  Vedaldi, A. and Soatto, S. (2008). **Quick shift and kernel methods for mode seeking.** In *European conference on computer vision*, pages 705–718. Springer.

★ Merci ★

Pour toute information : pierre-henri.conze@imt-atlantique.fr