

Application de techniques d'analyses d'image à l'identification d'organismes marins

Villon Sébastien

Encadré par : Marc Chaumont, Jérôme Pasquet, Gérard Subsol
Thomas Claverie, David Mouillot, Sébastien Villéger
Université De Montpellier

11 Juin 2015



**UNIVERSITÉ
DE MONTPELLIER**



- 1 Contexte
- 2 Classification et Deep learning
 - Généralités
 - Réseaux convolutionnels
- 3 Analyses
 - Structure du réseau
 - Structure des données
 - Bilan des analyses
- 4 Conclusion et perspectives

Sommaire

- 1 Contexte
- 2 Classification et Deep learning
- 3 Analyses
- 4 Conclusion et perspectives

Contexte écologique

- Une mesure de l'impact de l'activité humaine
- Des enjeux environnementaux
- Chercher à gagner en temps, et ne pas être restreint par une ressource humaine limitée

Figure 1: Obtenir une méthode de comptage et de reconnaissance reproductible



Contexte technique

Figure 2: Les techniques d'étude actuelles ne permettent pas d'utiliser les quantités gigantesques de données que nous stockons



- Un essor des techniques d'étude par vidéo
- Une qualité d'image croissante
- Un volume de données de plus en plus important

Données

Figure 3: lifeCLEF : une initiative internationale pour comparer les techniques d'annotations et de classification automatique d'images



- Utilisation du jeu de données fishCLEF ¹
- 15 espèces, 22000 images, 76 vidéos de 15 secondes



¹<http://www.imageclef.org/lifeclef/2015>

Sommaire

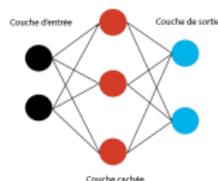
- 1 Contexte
- 2 Classification et Deep learning
 - Généralités
 - Réseaux convolutionnels
- 3 Analyses
- 4 Conclusion et perspectives

Introduction, le machine learning

- Création d'un modèle par apprentissage automatique
- Quelques différences : Apprentissage supervisé (SVM, support Vector Machine [Freund and Schapire, 1999, Cortes and Vapnik, 1995]), non supervisé (K plus proches voisins, distances euclidiennes [Pornpanomchai et al., 2013])

Les réseaux neuronaux

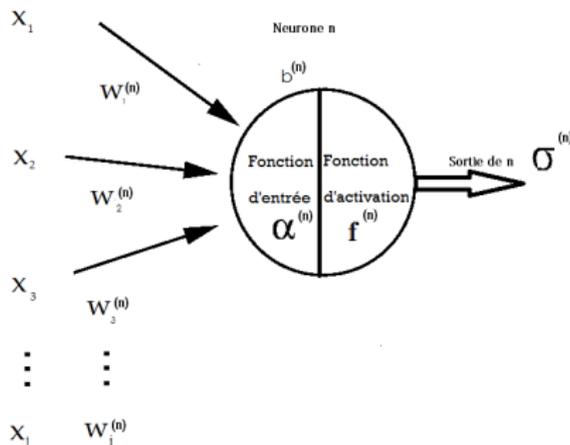
Figure 4: Réseau neuronal simple ou "Artificial Neural Networks" (ANN) [Atkinson and Tatnall, 1997]



- Une alternative aux méthodes de classification
- Peu d'utilisation dans le domaine de la reconnaissance d'espèces sous-marine

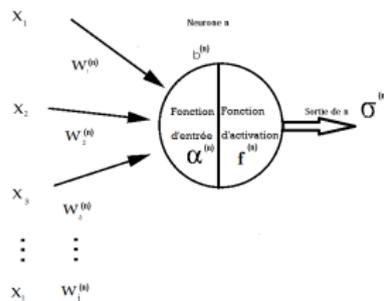
Vue détaillée d'un neurone

Figure 5: Détail d'un neurone n avec une fonction d'entrée $\alpha^{(n)}$, une fonction d'activation $f^{(n)}$, une pondération $w^{(n)} \in \mathbb{R}^c$ avec c le nombre de connexions en entrée, et un biais b^n



Vue détaillée d'un neurone

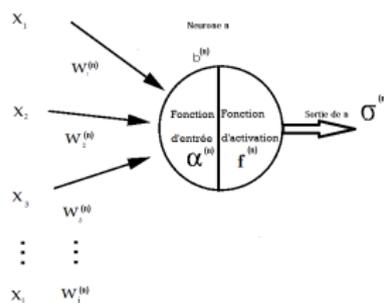
Figure 6: Détail d'un neurone



$$\sigma^{(n)}(\mathbf{x}^{(n)}) = f^{(n)}(\mathbf{w}^{(n)}\mathbf{x}^{(n)} + b^{(n)})$$

Vue détaillée d'un neurone

Figure 7: Détail d'un neurone

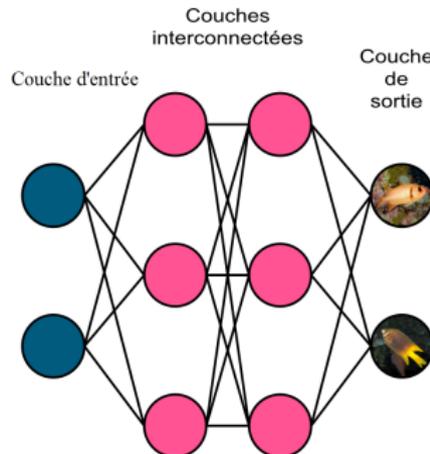


Fonction ReLU : $f(r) = \max(0, r)$

Fonction Sigmoïde : $f(r) = \frac{1}{1+e^{-r}}$

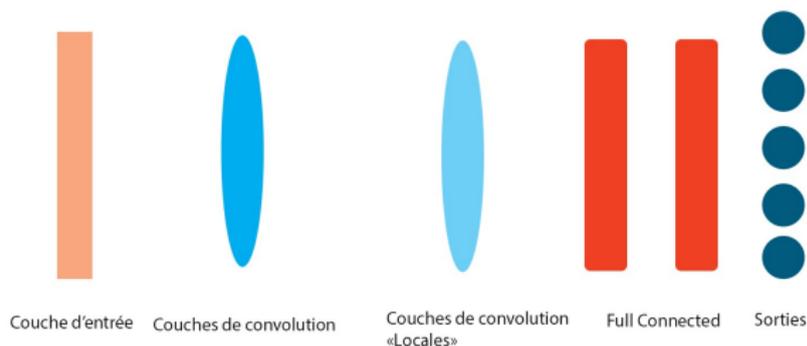
Réseaux simples et Deep-learning

Figure 8: Un réseau de neurones "Deep-learning" : davantage de couches cachées ou *full-connected* [Bengio et al., 2013], et la création de filtres intelligentes



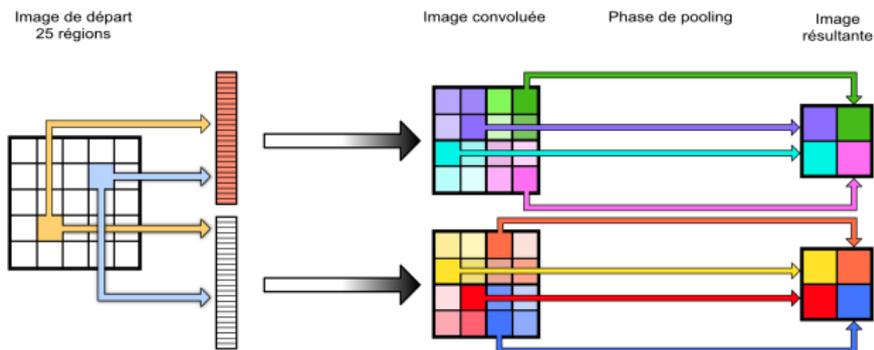
Introduction

Figure 9: Les réseaux convolutionnels ajoutent de l'information pour la classification grâce à l'introduction de couche(s) de convolution. Les filtres dynamiques s'adaptent intelligemment pour s'adapter au modèle [Schmidhuber, 2015].



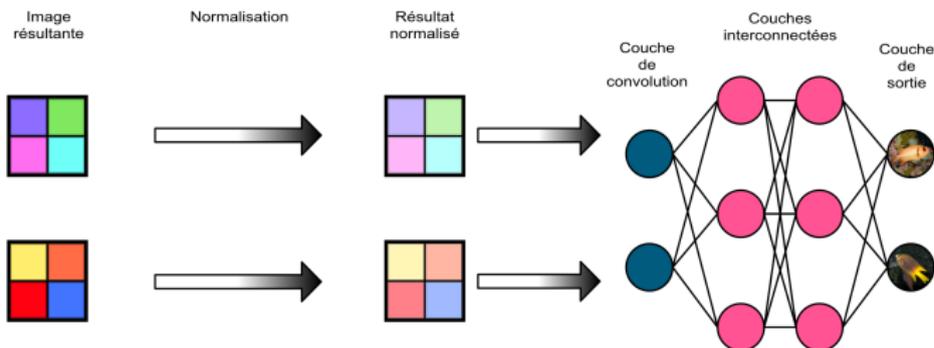
Détail d'une couche de convolution

Figure 10: Convolution et Pooling



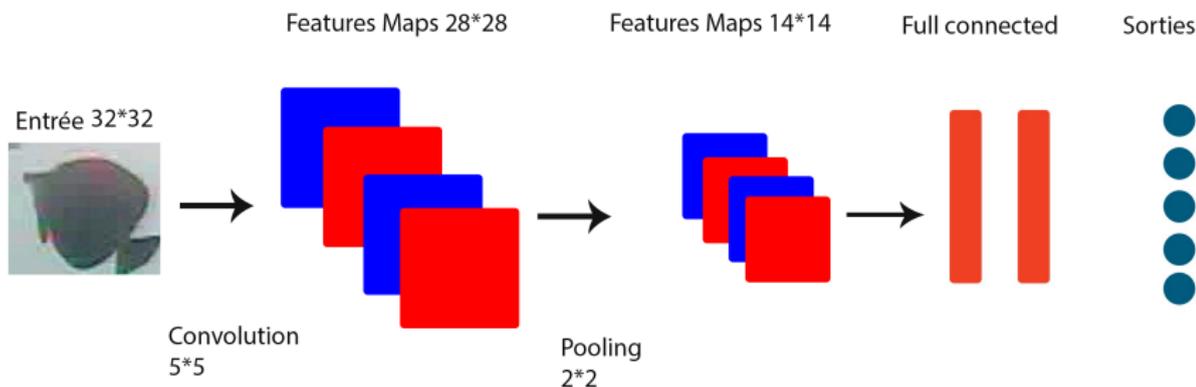
Détail d'une couche de convolution

Figure 11: Normalisation et connexion à un ANN



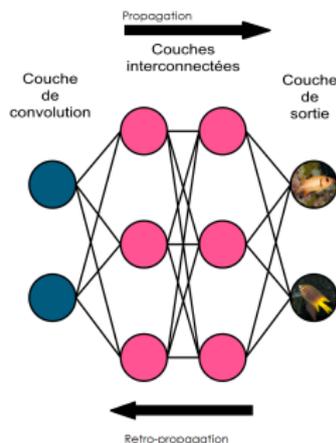
Détail d'un réseau de convolution

Figure 12: Évolution d'une image pendant une phase de convolution



Apprentissage, la mise à jour des poids

Figure 13: Principe de la rétro-propagation : Propagation de l'exemple, calcul de l'erreur, puis modification des poids



Sommaire

- 1 Contexte
- 2 Classification et Deep learning
- 3 Analyses**
 - Structure du réseau
 - Structure des données
 - Bilan des analyses
- 4 Conclusion et perspectives

Méthodologie

Figure 14: Images d'entrées



- Utilisation d'images pour l'apprentissage
- Utilisation de différents ensembles pour l'apprentissage et les tests
- Apprentissage, validation et test

Données de départ

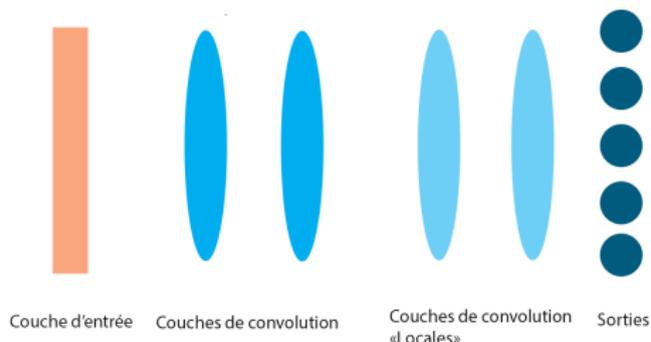
Figure 15: Base de données FishCLEF



- 15 espèces, 22000 images
- Des données inégales : de 24 à 3000 images
- Première étape de traitement : Création d'un corpus C1 robuste au changement de taille grâce au redimensionnement

Premiers résultats

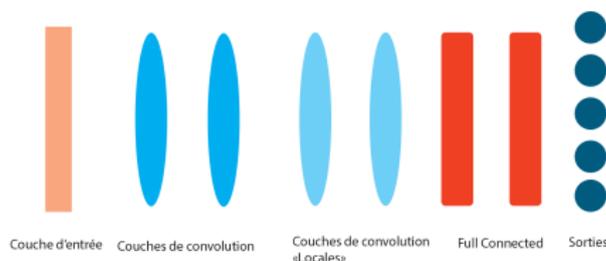
Figure 16: Réseau Cifar [Krizhevsky et al., 2012] - 2 couches de convolution de 32 filtres, une fonction de max pooling. Fonction d'activation des neurones ReLU, couche finale softmax



12% de bonne classification sur 10 jeux de données

Amélioration n°1

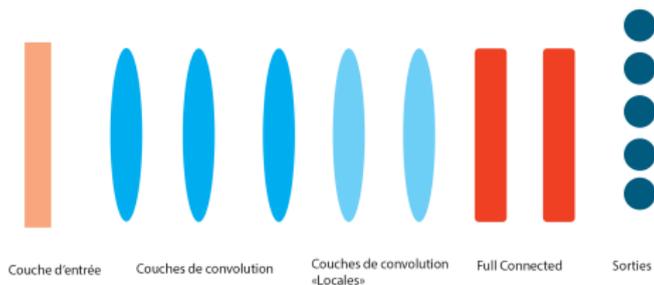
Figure 17: Réseau 2 - 2 couches cachées de 1000 neurones, augmentation de la taille des noyaux de convolution



40% de bonne classification sur 10 jeux de données
Double le temps d'apprentissage

Amélioration n°2

Figure 18: Réseau 3,4,5 - 3 couches de convolution, et deux couches cachées de 4000 neurones. Fonction average pooling pour les réseaux 4 et 5, fonction sigmoïde pour les réseaux 3 et 4.



47% de bonne classification, triple le temps d'apprentissage

Création des différents corpus

- C2 : Robuste à la rotation, et taille de la base augmentée : 86 000 images
- C3 : Classes équilibrées, et taille de la base augmentée : 220 000 images
- C4 : Possibilité de détecter le fond marin , et taille de la base augmentée: 300 000 images

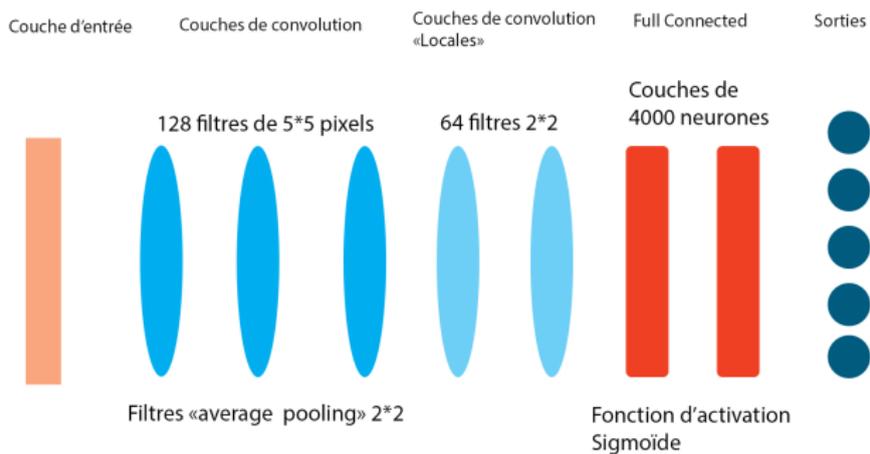
Analyse des résultats

Utilisation des réseaux 3,4,5

- C2 : 70% de bonne classification
- C3 : 28% de bonne classification
- C4 : 31% de bonne classification
- Un nouveau corpus C5 : 80% de bonne classification avec le réseau final

Réseau final

Figure 19: Réseau final - 3 couches de convolution, 2 couches de convolution locales et 2 couches cachées.



Importance de la structure du réseau

- Réussite de la couche de convolution locale diminuant le temps de convergence par un facteur 3
- Impact des fonctions d'activation
- Importance de l'adaptation de la taille filtres de pooling et de convolution

Importance de la structure des données

- Importance de la qualité et de la taille des images
- Influence de la quantité d'exemples pour l'apprentissage
- Plus de diversité, plus de qualité

Sommaire

- 1 Contexte
- 2 Classification et Deep learning
- 3 Analyses
- 4 Conclusion et perspectives

Contributions

- Mise en place d'un système d'identification
- Une capacité de reconnaissance des images fournies au réseau de 80%
- Analyse de l'influence des corpus de données
- Influence des structures et des optimisations des réseaux de neurones profonds
- Analyse des techniques d'utilisation des réseaux convolutionnels

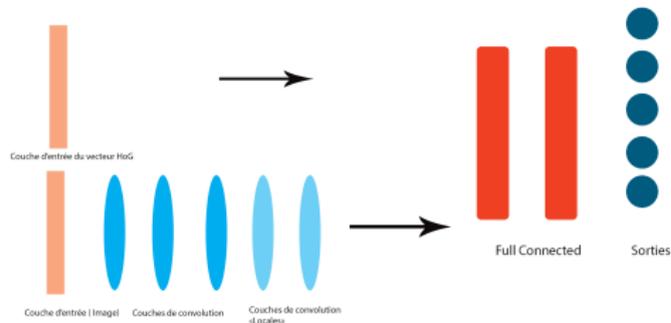
Perspectives et améliorations

- Utilisation du tracking pour ajouter de l'information
- Approche par multi-résolution
- Extraction des matrices de confusions pour analyser a posteriori les résultats
- Utilisation de connaissances à priori, comme le comportement social des poissons
- Amélioration des données utilisées par le réseau
- Continuation de mes recherches et de mon travail au cours d'une année d'ingénierie en grâce au labex NUMEV

Perspectives

- Intégration d'informations caractéristiques supplémentaires

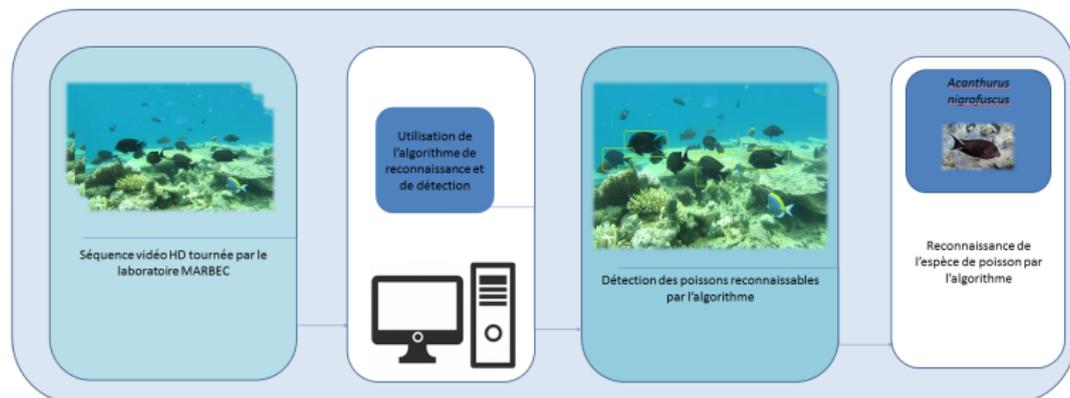
Figure 20: Comparaison des qualités d'images



Perspectives

- Détection de poissons dans des vidéos sous-marines

Figure 21: Reconnaissance et détection



References I

-  Atkinson, P. M. and Tatnall, A. (1997).
Introduction neural networks in remote sensing.
International Journal of remote sensing, 18(4):699–709.
-  Bengio, Y., Boulanger-Lewandowski, N., and Pascanu, R.
(2013).
Advances in optimizing recurrent networks.
In *Acoustics, Speech and Signal Processing (ICASSP), 2013
IEEE International Conference on*, pages 8624–8628. IEEE.
-  Cortes, C. and Vapnik, V. (1995).
Support-vector networks.
Machine learning, 20(3):273–297.

References II

-  Freund, Y. and Schapire, R. E. (1999).
Large margin classification using the perceptron algorithm.
Machine learning, 37(3):277–296.
-  Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012).
Imagenet classification with deep convolutional neural networks.
In Advances in neural information processing systems, pages 1097–1105.
-  Pornpanomchai, C., Lurstwut, B., Leerasakultham, P., and Kitiyanan, W. (2013).
Shape- and texture-based fish image recognition system.
In Kasetsart J.

References II



Schmidhuber, J. (2015).

Deep learning in neural networks: An overview.

Neural Networks, 61:85–117.