

# Explorer des corpus d'images, l'IA au service du patrimoine

Le Deep Learning pour l'histoire de l'art

Nicolas Gonthier

Université Paris Saclay - Télécom ParisTech - LTCI

28 mars 2018



université  
PARIS-SACLAY



# Motivation

- 1 Motivation
- 2 Le Deep Learning
- 3 La classification et la détection d'objets dans les peintures

# Présentation :

## Mon projet de thèse

Transfert d'apprentissage pour l'analyse iconographique d'œuvres d'art graphiques

## Directions

- Yann Gousseau et Saïd Ladjal (Télécom ParisTech)
- Olivier Bonfait (Université de Bourgogne)



# Motivation

Objectif : Indexation **iconographique** automatique



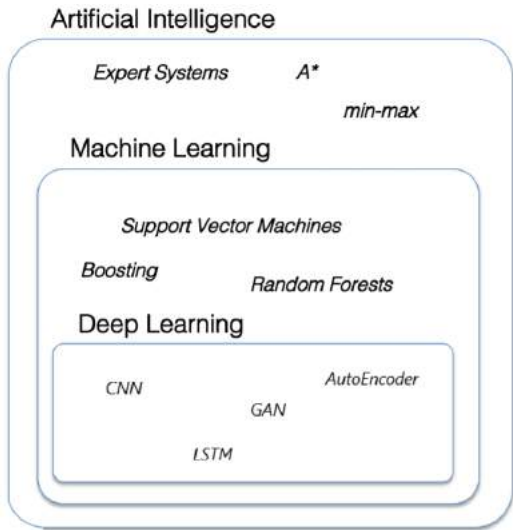
*La Mort de Cléopâtre, Bohn Germann-August von (1812-1899), Nantes, musée d'Arts*

**Mots-clé :**  
Cléopâtre VII,  
femme,  
nudité,  
suicide,  
vipère,  
lit,  
table,  
panier,  
palmier,  
colonne,  
montagne,  
crépuscule,  
etc

# Le Deep Learning

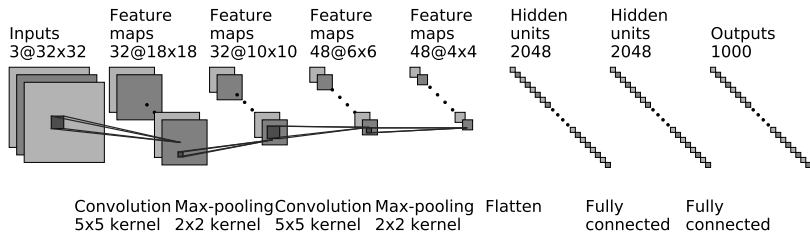
- 1 Motivation
- 2 Le Deep Learning
- 3 La classification et la détection d'objets dans les peintures

# Le Deep Learning, une sous branche de l'IA



Algorithme basé  
sur l'exemple

## Les réseaux de neurones profonds - Deep Learning



## Renouveau du Deep Learning en 2012



Résultats d'AlexNet sur ILSVRG-2010, [KSH12]

- 1.2 million d'images d'entraînement
- 1000 classes
- 60 million de paramètres
- 150000 images de test : "top-5 test error rate of 15.3%"
- Remarque en 2016 : 3.08% [SIVA16]

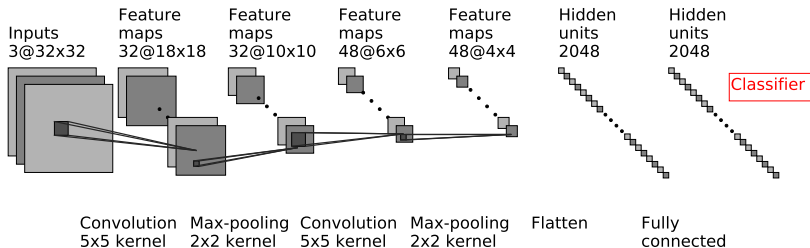


# La classification et la détection d'objets dans les peintures

- 1 Motivation
- 2 Le Deep Learning
- 3 La classification et la détection d'objets dans les peintures**

# Principe du transfert d'apprentissage

**Idée** : utiliser les descripteurs haut niveau d'un réseau entraîné pour une autre tâche et uniquement entraîné un classifieur<sup>1</sup> pour la nouvelle classe :



## Principe du transfert d'apprentissage

On utilise un réseau entraîné pour réaliser de la classification de photos pour calculer des descripteurs pour réaliser la classification des peintures.

1. un SVM linéaire par exemple

# Recherche dans une base d'images d'art

Classification de peintures issues de **Your paintings Art UK** [CZ14, CZ16]

Split	Aero	Bird	Boat	Chair	Cow	Din	Dog	Horse	Sheep	Train	Total
Train	74	319	862	493	255	485	483	656	270	130	3463
Val	13	72	222	140	52	130	113	127	76	35	865
Test	113	414	1059	569	318	586	549	710	405	164	4301

*Répartition des classes*



## Reproduction et extension des résultats de [CZ16]

Il s'agit ici d'une recherche par mot-clé et non par requête visuelle :

Net	Augmentation	mAP	mAP [CZ16]
VGG16, 4096D,relu7 [SZ14]	none	69.8	68.7
	Stretch	72.0	70.8
VGG19, 4096D, relu7 [SZ14]	none	69.6	•
	Stretch	72.4	•
ResNet-152 2048D [HZRS15]	none	73.3	72.5
	Stretch	76.0	75.0
InceptionResNet v2 1536D [SIVA16]	none	74.9	•
	Stretch	77.6	•

Images d'entrée des réseaux  $224 \times 224$  ou  $299 \times 299$

**Conclusion :** Il est possible de classifier des peintures à l'aide de descripteurs appris sur des photos, avec deux cents exemples.

# Oxford Painting Search

Démonstration en ligne sur les peintures de ArtUk :  
<http://zeus.robots.ox.ac.uk/artsearch/>



Search results page 3 of 200 (34,800 results)



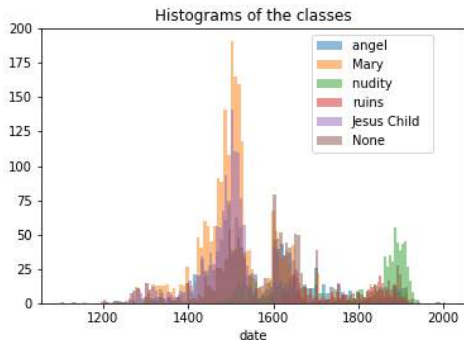
Source: [CZ14]

# Création d'une base d'apprentissage nouvelle

5 classes différentes avec des images issues de WikiCommons

Classe	Angel	Mary	nudity	ruins	Jesus_Child	None	Total
#	1269	2336	890	459	1605	1392	5473
%	23%	43%	16%	8%	19%	25%	●

Séparer en 40% pour l'entraînement et 60% pour les tests.



*Histogramme des dates de création des œuvres ou bien une date de l'artiste*

# Exemples d'apprentissage pour Ange



angel Mary nudity



angel Mary Jesus\_Child

# Exemples d'apprentissage pour Ange



angel Mary



angel



# Résultats de classification :

Net	angel	Mary	nudity	ruins	Jesus Child	mAP
VGG19 relu7	0.55	0.82	0.77	0.91	0.75	0.76
Strech	0.584	0.846	0.793	0.920	0.785	0.786
ResNet152 2048D	0.58	0.85	0.81	0.90	0.8	0.79
Strech	0.595	0.877	0.835	0.916	0.830	0.811
InceptionResNetv2 1536D	0.51	0.84	0.79	0.91	0.77	0.76
Strech	0.546	0.862	0.817	0.917	0.800	0.789

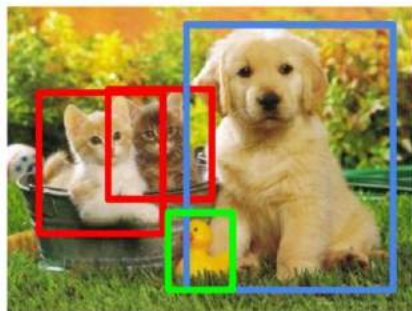
# Algorithme de détection

## Classification

## Object Detection

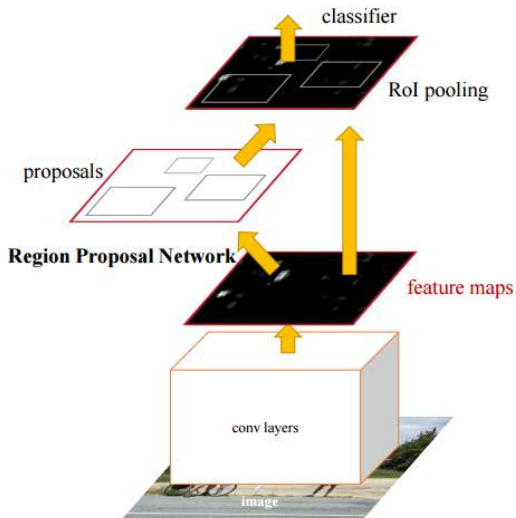


CAT



CAT, DOG, DUCK

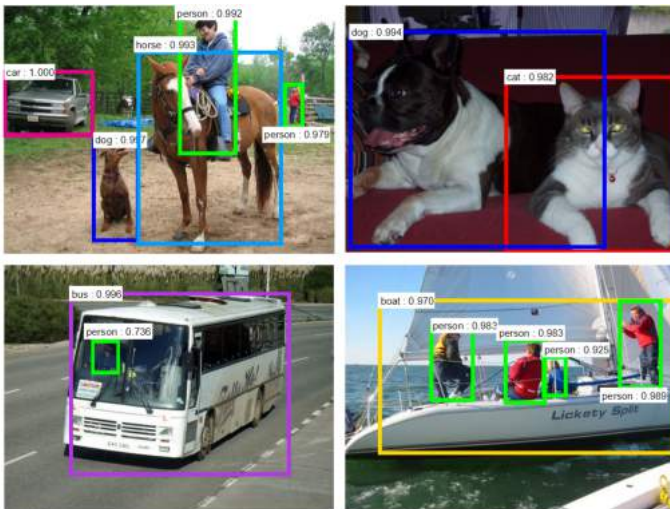
# Algorithme de détection : Faster RCNN [RHGS15]



Le réseau produit un ensemble de détection avec estimation de la localisation des objets (après une suppression des éléments non maximales à l'aide d'un seuillage).

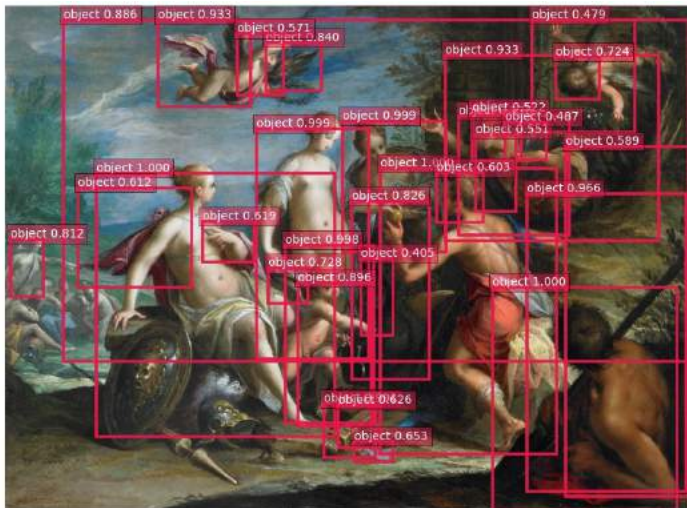
Les images considérées sont de l'ordre de  $600 \times 600$  pixels.

## Faster RCNN [RHGS15] exemple de sortie



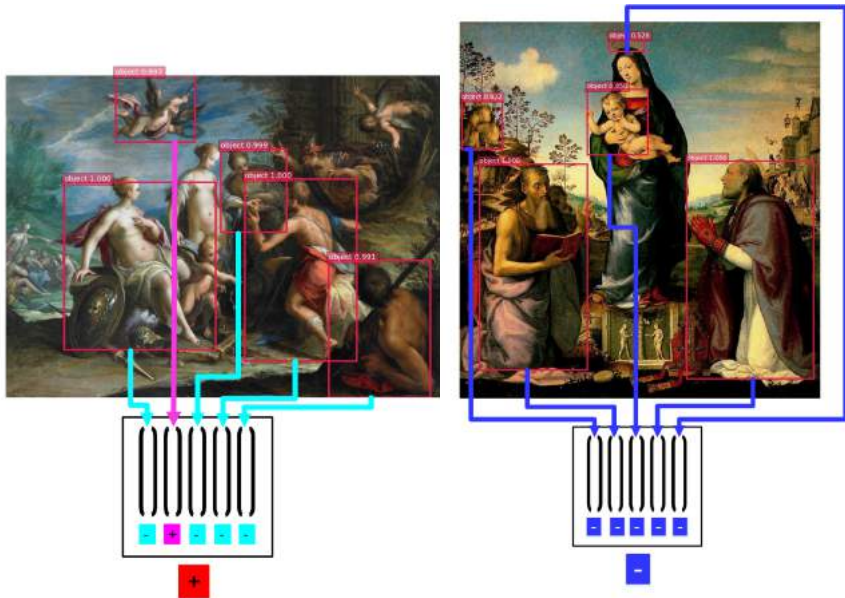
**Question :** Comment faire du transfert d'apprentissage avec un tel réseau ?

# Exemple de 30 zones par image



Quelle proposition contient l'objet qui nous intéresse ?

# Problème d'apprentissage multi-instances



# Phase de sélection : résultats

On ne garde que la zone la plus positive pour les exemples positifs :



Choix correct

Cadres rouges = objet avec le plus haut score, cadre vert = choisi lors de notre phase de sélection.



Choix partiel

# Résultats : Détection d'anges

Les cadres verts sont les détections avec la phase de sélection et ceux en rouge sans cette phase :  
Amélioration





# Résultats : Détection d'anges

Les cadres verts sont les détections avec la phase de sélection et ceux en rouge sans cette phase :



# Résultats : Détection d'anges

Les cadres verts sont les détections avec la phase de sélection et ceux en rouge sans cette phase :

Erreur



# Détection d'autres éléments : amélioration possible



# Conclusion

Limites :

- Attention aux biais dans la base d'apprentissage
- Encore besoin de beaucoup d'exemples

Possibilité :

- Annoter automatiquement des images avec des détections multiples
- Réaliser de la recherche pleine image en quelque sorte

# Pistes futures

- Augmentation de la base
- Amélioration de la brique de sélection par apprentissage multi-instance
- Travailler sur l'invariance au style

# Références I

- ▶ [CZ14] Elliot J. Crowley and Andrew Zisserman.  
In search of art.  
In *Workshop at the European Conference on Computer Vision*, pages 54–70.  
Springer, 2014.
- ▶ [CZ16] Elliot J. Crowley and Andrew Zisserman.  
The Art of Detection.  
In *European Conference on Computer Vision*, pages 721–737. Springer, 2016.
- ▶ [HZRS15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun.  
Deep Residual Learning for Image Recognition.  
*arXiv :1512.03385 [cs]*, December 2015.
- ▶ [KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton.  
Imagenet classification with deep convolutional neural networks.  
In *Advances in Neural Information Processing Systems*, pages 1097–1105, 2012.
- ▶ [RHGS15] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun.  
Faster R-CNN : Towards Real-Time Object Detection with Region Proposal Networks.  
*arXiv :1506.01497 [cs]*, June 2015.

# Références II

- ▶ [SIVA16] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alex Alemi. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *arXiv :1602.07261 [cs]*, February 2016.
- ▶ [SZ14] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv :1409.1556 [cs]*, September 2014.