

Machine Learning & Art

Bart.Lamiroy@univ-reims.fr



UNIVERSITÉ
DE REIMS
CHAMPAGNE-ARDENNE

Plan du cours

- Motivation première :
La Muse (E. Pottier)
- Technologie et fondements scientifiques :
Réseaux convolutionnels récurrents (RCNN) profonds et réseaux antagonistes génératifs (GAN)
- Exemples et analyse des limites
- Références et code



Motivation

<https://lamuse.univ-reims.fr> : La muse

Recherche sur le processus de création artistique

- l'angoisse de la page blanche,
- le rôle de la contrainte dans le processus créatif,
- le choix des thèmes pour la création d'une œuvre.

La muse: Leveraging Artificial Intelligence for Sparking Inspiration

[Bart Lamiroy](#), [Emmanuelle Potier](#)

[CRESTIC - Centre de Recherche en Sciences et Technologies de l'Information et de la Communication - EA 3804](#)

Question :

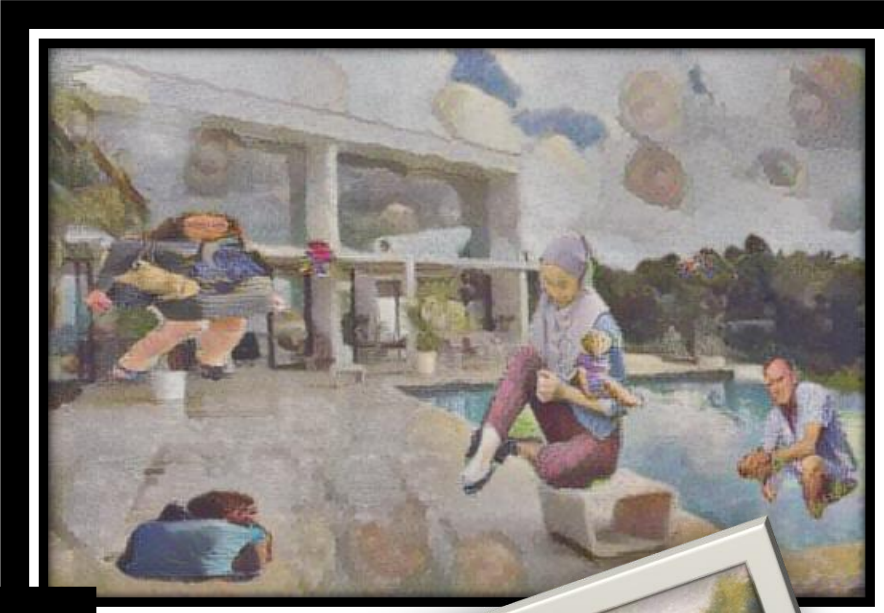
Comment créer des cas d'études *pertinentes* pour un artiste peintre ?

Comment mesurer la pertinence

- Règles de composition
- Cohérence de style pictural
- Adéquation du thème, époque, contexte ...



Générer des réinterprétations d'œuvres connues ...



Démarche suivie

- Utilisation de Mask-RCNN pour l'analyse de la composition
- Substitution d'éléments en maintenant la cohérence de la composition
- Transfert de style en utilisant des CNN
- Upsampling



Technologie et fondements scientifiques

1. Mask-RCNN

Mask R-CNN, K. He, G. Gkioxari, P. Dollar, R. Girshick ; Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2961-2969. (<https://arxiv.org/abs/1703.06870>)

facebook.

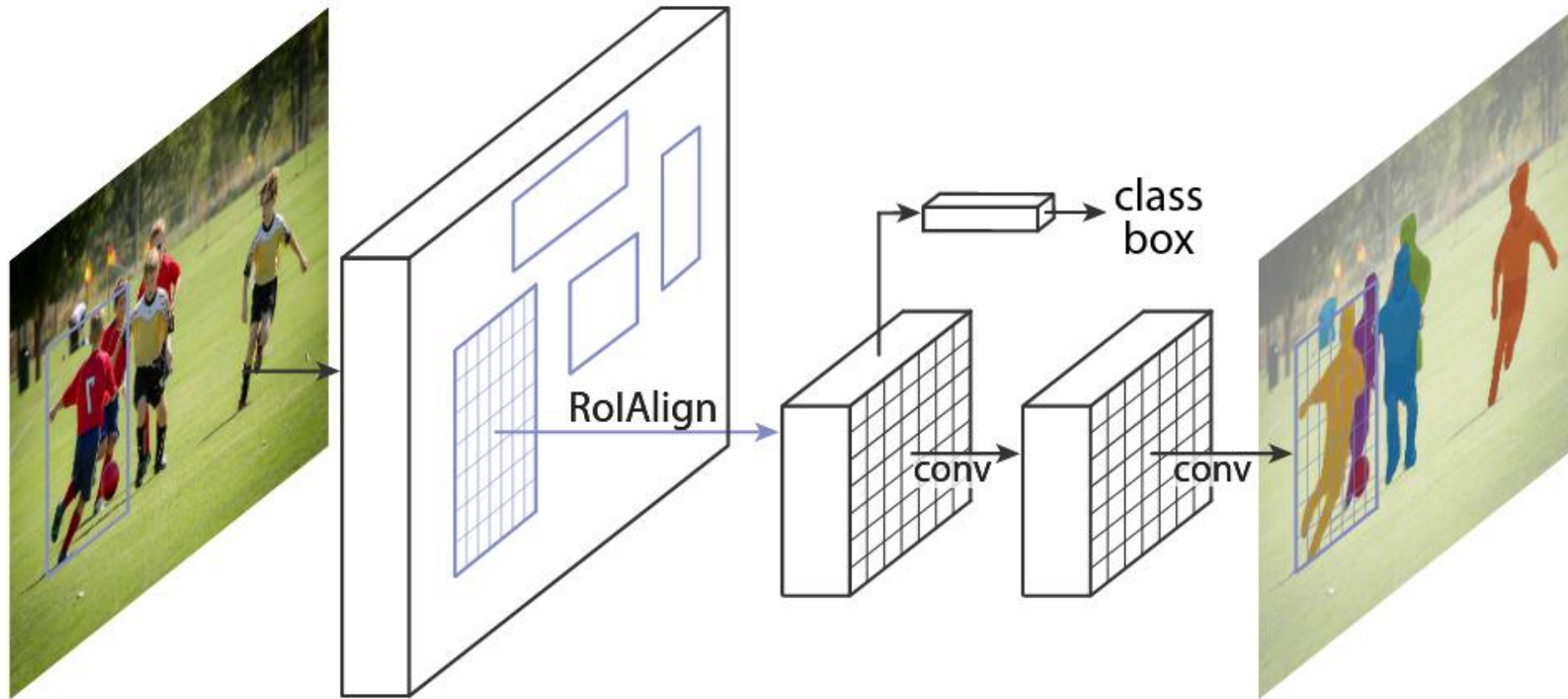


M-RCNN : principes de base

- Cœur du réseau : classification via des CNN
<https://cs231n.github.io/convolutional-networks/>
- Inconvénient des CNN de base : détection/classification d'objets multiples dans une même image
 - **R-CNN** subdivise l'image en régions et fait de la classification par région
 - **Fast R-CNN** optimise la sélection des régions via RoIPool (Region of Interest Pooling)
 - **Faster R-CNN** améliore encore la sélection par l'introduction d'une couche supplémentaire (*Region Proposal Network*)
- M-RCNN ajoute par ailleurs la segmentation par masque

Source : <https://viso.ai/deep-learning/mask-r-cnn/>

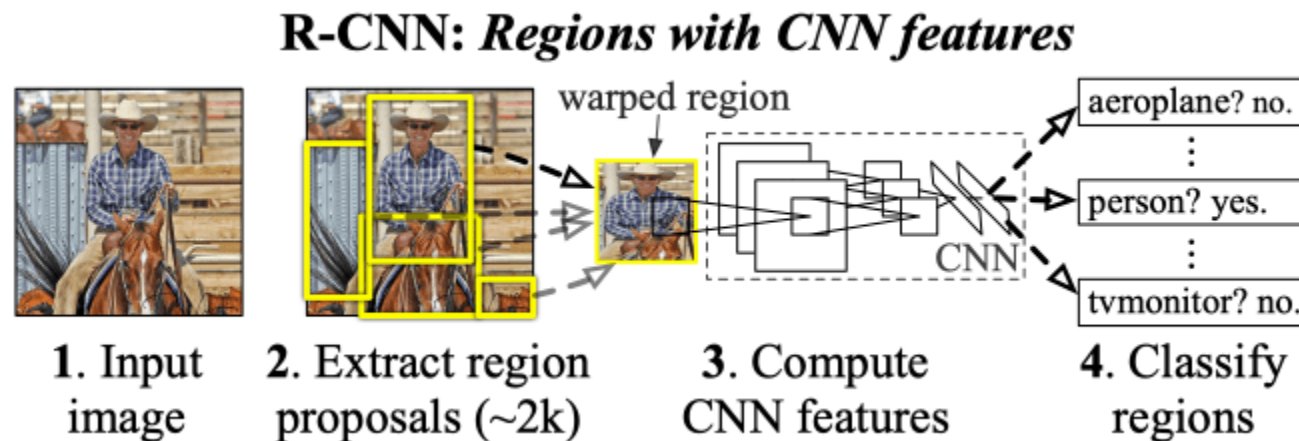
M-RCNN : architecture



Source : <https://arxiv.org/abs/1703.06870>

Etape 1 : R-CNN

« *Rich feature hierarchies for accurate object detection and semantic segmentation* », Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik (<https://arxiv.org/abs/1311.2524>)



La sélection des régions initiales est faite par « Selective Search » Uijlings, K. van de Sande, T. Gevers, and A. Smeulders, IJCV, 2013 ([10.1007/s11263-013-0620-5](https://arxiv.org/abs/1311.2524))

<https://learnopencv.com/selective-search-for-object-detection-cpp-python/>

Etape 1 : R-CNN

Inconvénients :

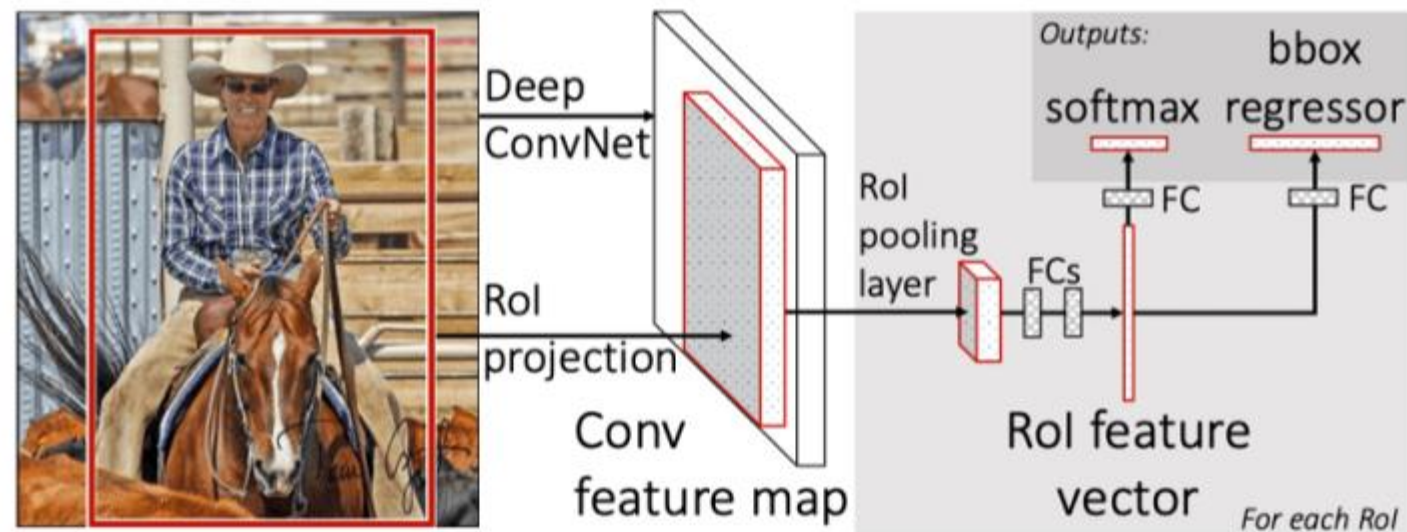
- CNN indépendant pour chaque région
- Entraînement et calcul des descripteurs (filtres convolutionnels) redondant
- Classification binaire pour chaque classe

Etape 2 : Fast R-CNN

« Fast R-CNN », R. Girshick, 2015, <https://arxiv.org/abs/1504.08083>

Avantages :

- Unique calcul des convolutions (redimensionnement via RoI Pooling)
- Classification multiclasse via softmax



Etape 3 : Faster RCNN

Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, S. Ren, K. He, R. Girshick and J. Sun in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017.

- Principe du FPN (*Feature Pyramid Network*)

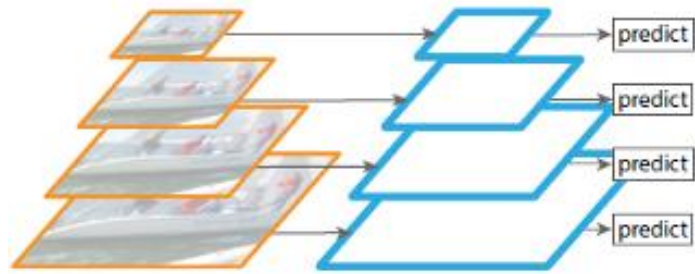
Feature Pyramid Networks for Object Detection, TY. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, S. Belongie; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2117-2125

- Basé sur des réseaux convolutionnels classiques (ResNet, VGG, Inception) pour l'extraction de descripteurs
- Reprojection de carte des descripteurs dans l'image d'origine.

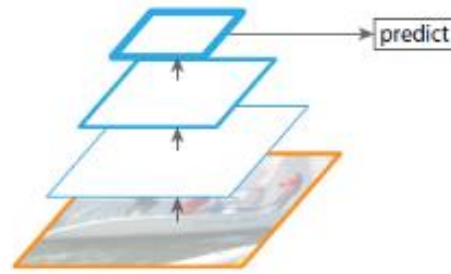
Ce type de réseau permet de maintenir une cohérence sémantique sur différents niveaux de résolution.

- Combiné avec un RPN (*Region Proposal Network*)

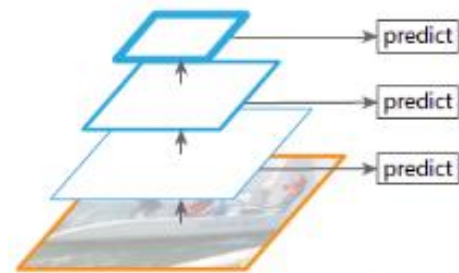
Principes et précurseurs



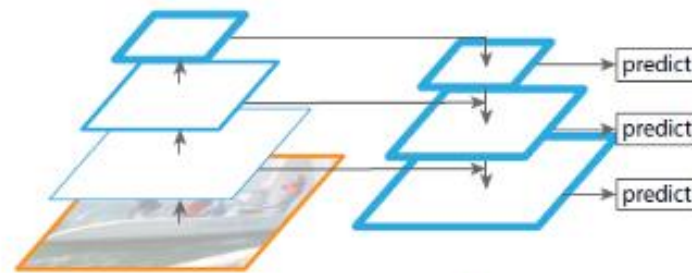
(a) Featurized image pyramid



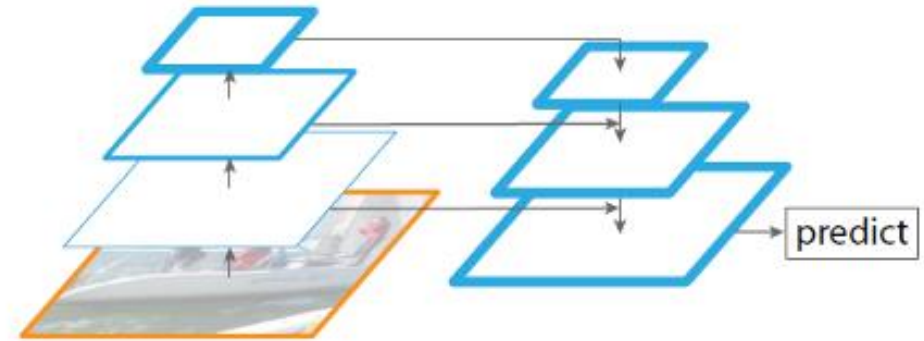
(b) Single feature map



(c) Pyramidal feature hierarchy

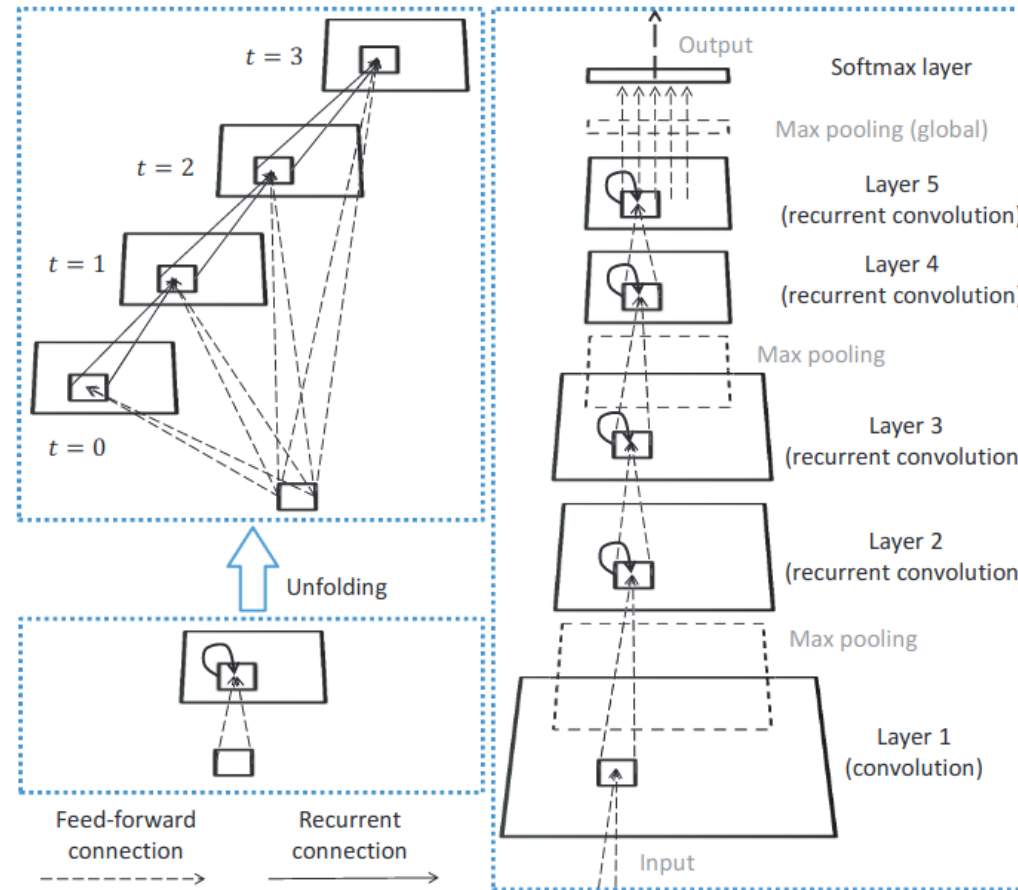


(d) Feature Pyramid Network



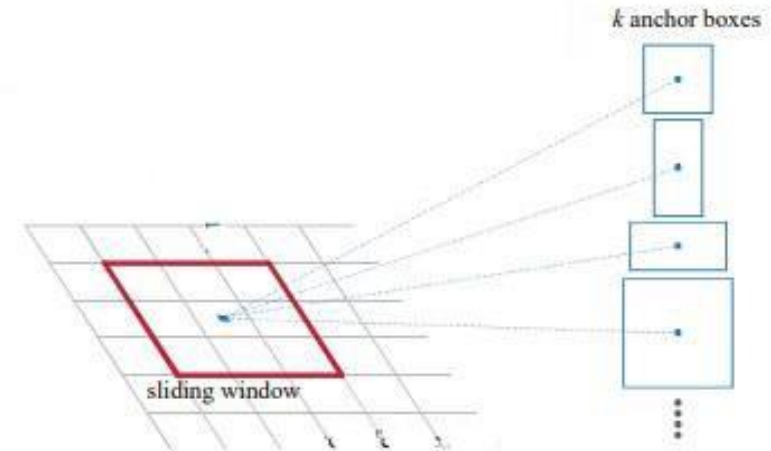
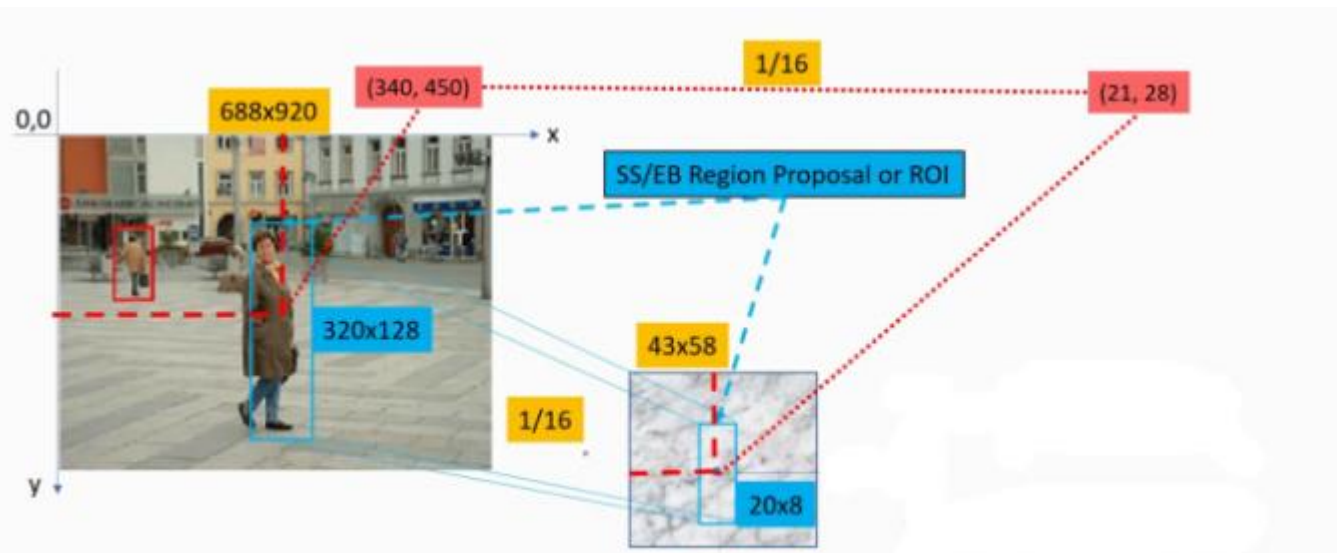
(e) Similar Structure with (d)

Recurrent Convolutional Networks



Source : [Recurrent Convolutional Neural Network for Object Recognition](#), M. Liang, X. Hu, CVPR 2015

Region Proposal Network



i = anchor index in minibatch

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)$$

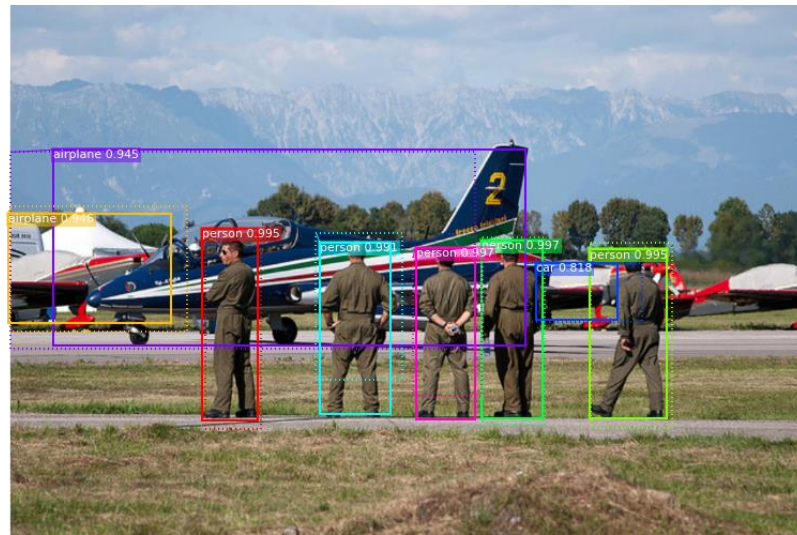
L_{cls} is associated with **Log loss**.
 L_{reg} is associated with **Smooth L1 loss**.
 p_i is the **Predicted probability of being an object for anchor i** .
 t_i is the **Coordinates of the predicted bounding box for anchor i** .
 p_i^* is the **Ground truth objectness label**.
 t_i^* is the **True box coordinates**.

In practice $\lambda = 10$, so that both terms are roughly equally balanced

N_{cls} = Number of anchors in minibatch (~ 256)
 N_{reg} = Number of anchor locations (~ 2400)

M-RCNN vs. F-RCNN

Extension des régions à des masques au niveau pixel, plutôt que des boîtes englobantes.



Pour jouer avec : https://github.com/matterport/Mask_RCNN

Petit rappel du projet initial ...



Technologie et fondements scientifiques

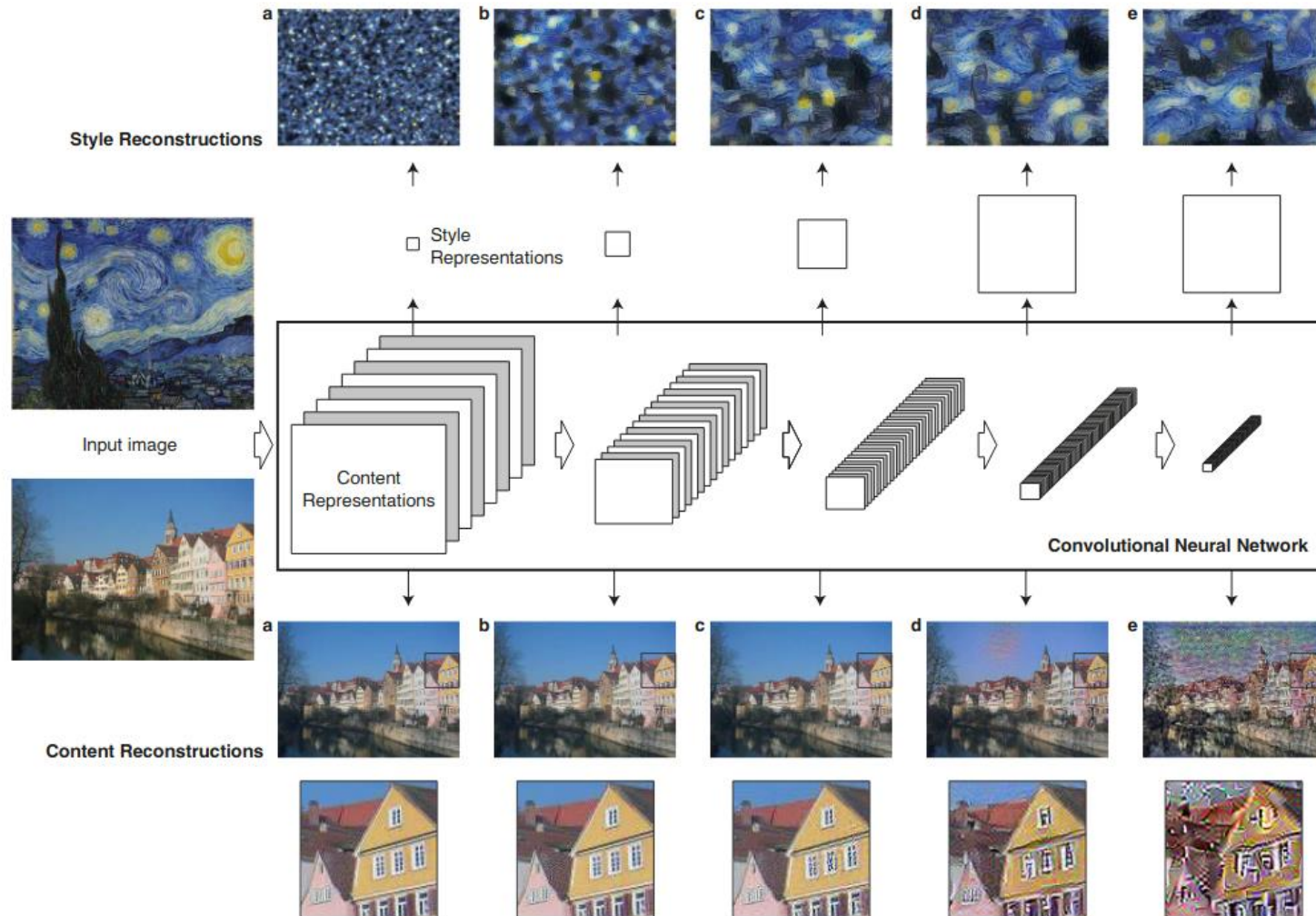
2. CNN

[1] ***Image Style Transfer Using Convolutional Neural Networks***, L. A. Gatys, A. S. Ecker and M. Bethge, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 2414-2423 (https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Gatys_Image_Style_Transfer_CVPR_2016_paper.pdf).

[2] ***Exploring the Structure of a Real-time, Arbitrary Neural Artistic Stylization Network***, G. Ghiasi, H. Lee, M. Kudlur, V. Dumoulin, J. Shlens ; Proceedings of the British Machine Vision Conference (BMVC), 2017. (<https://arxiv.org/abs/1705.06830>)

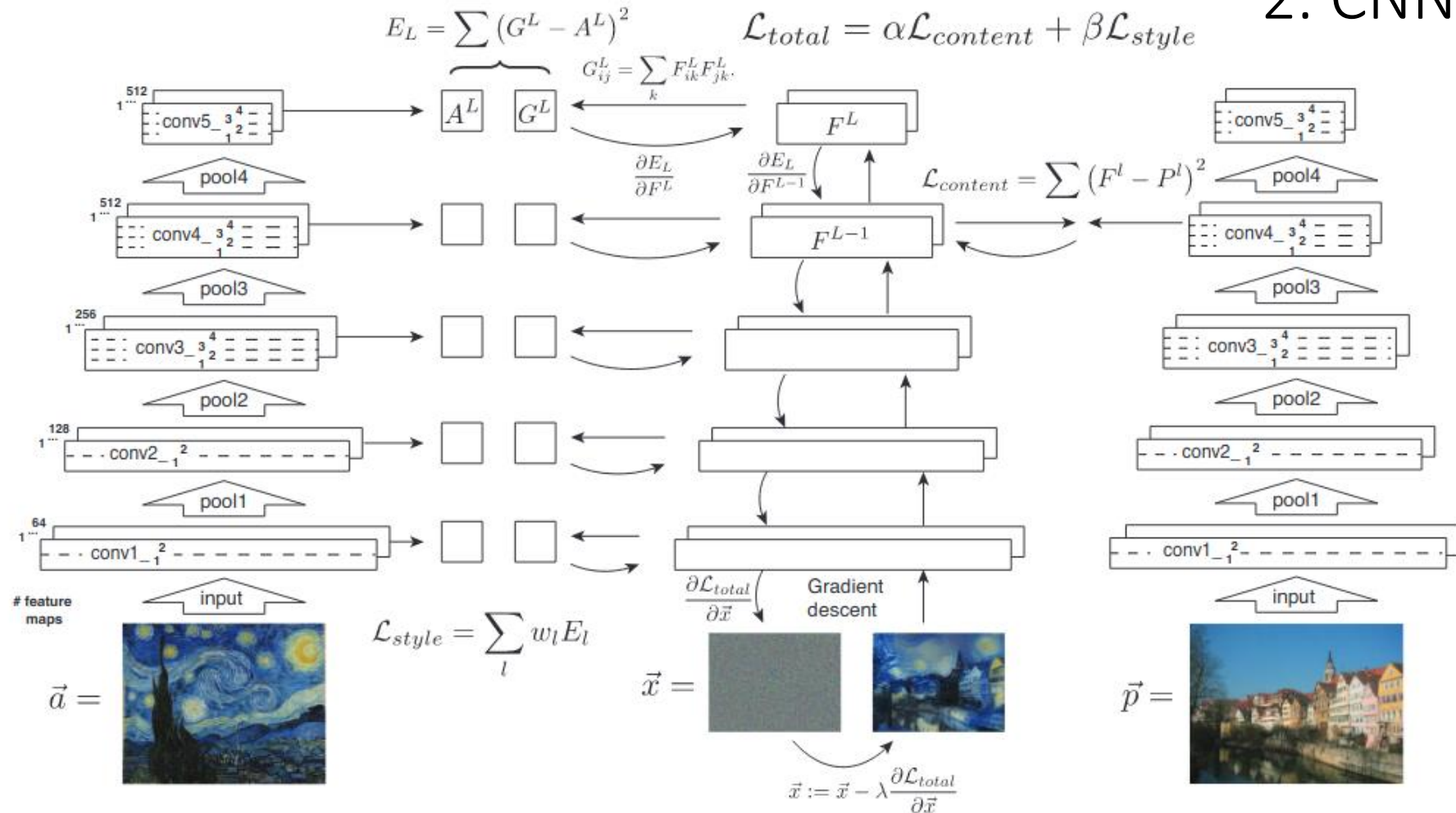
Technologie et fondements scientifiques

2. CNN



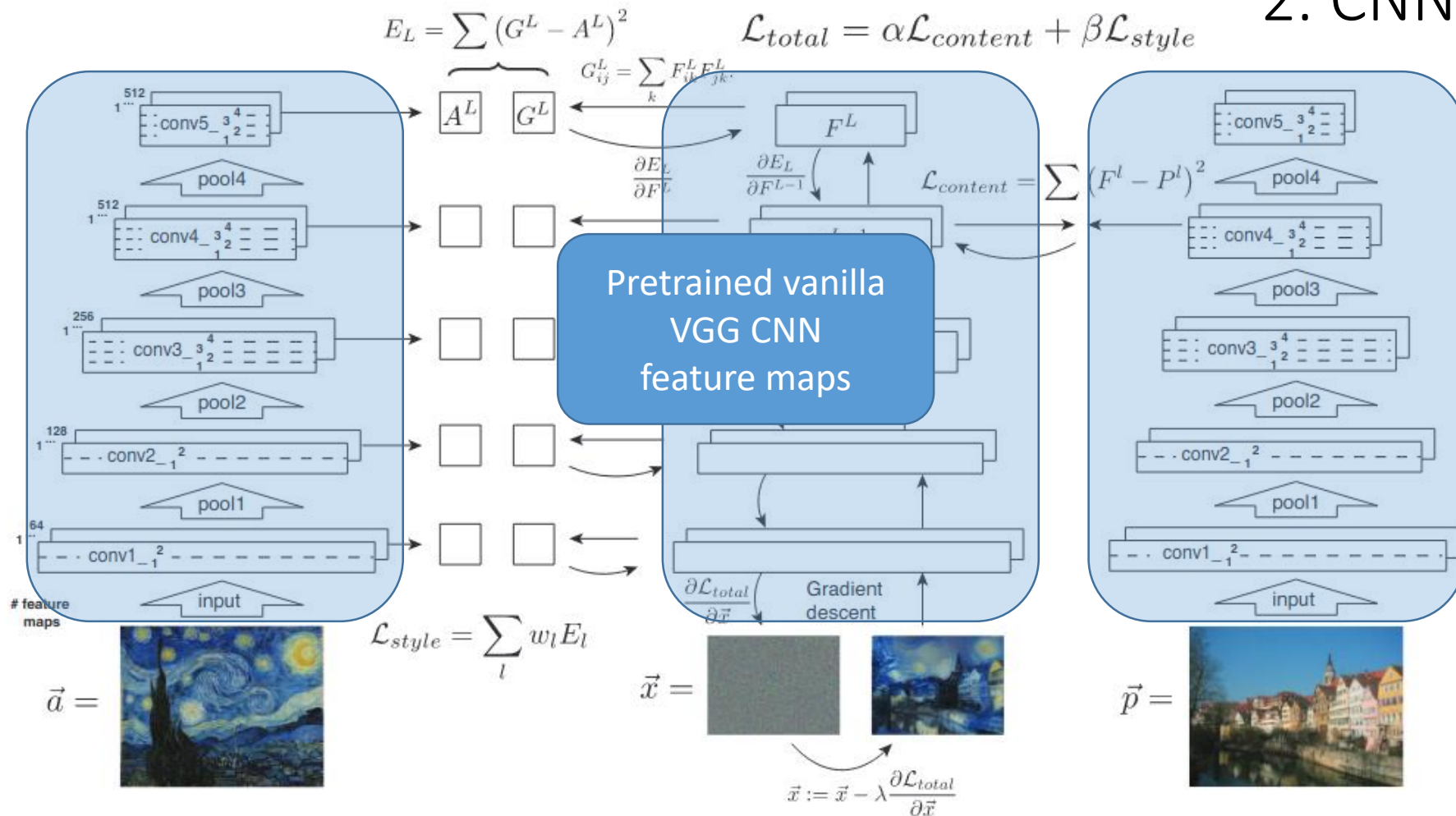
Technologie et fondements scientifiques

2. CNN



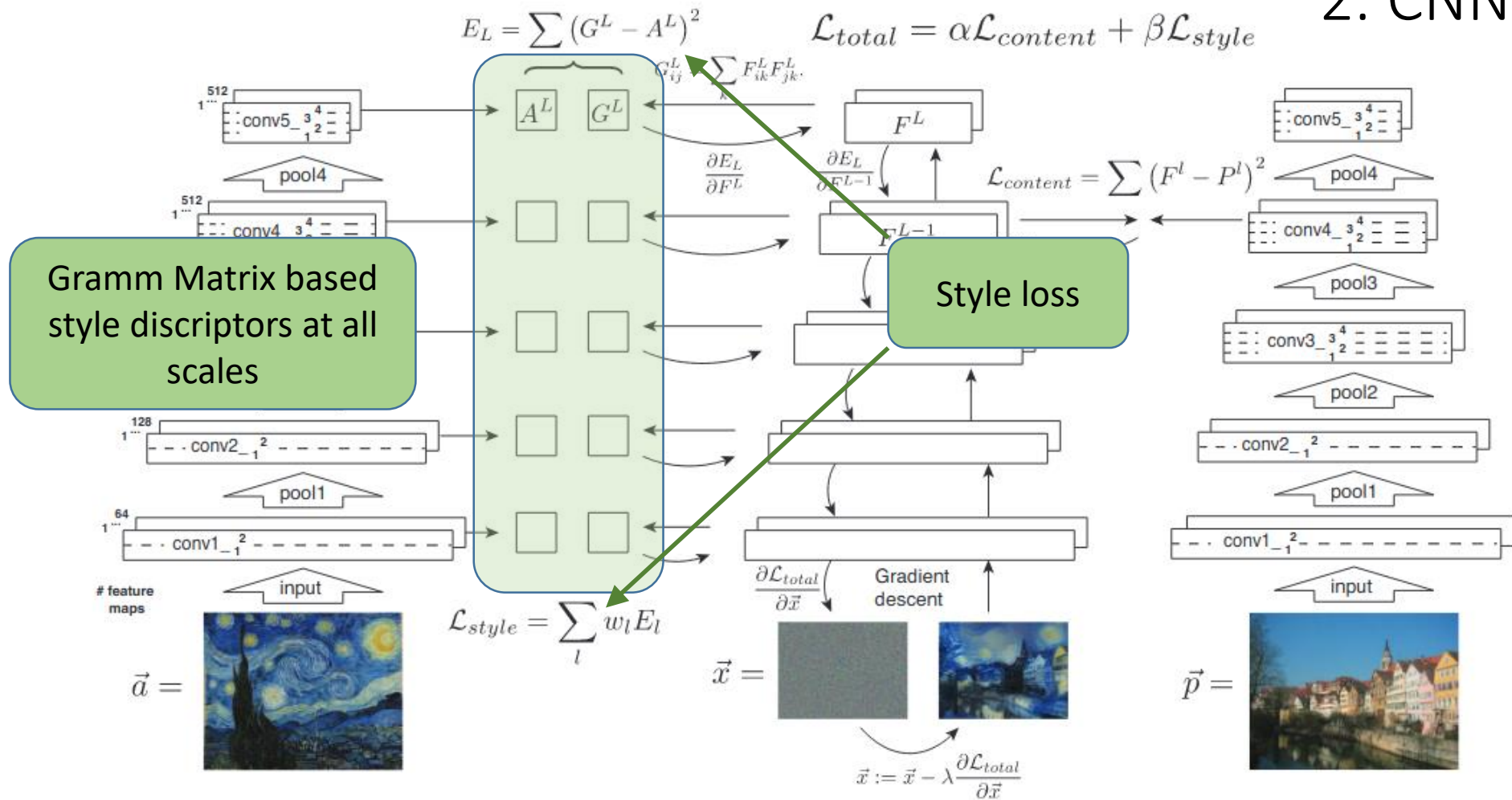
Technologie et fondements scientifiques

2. CNN



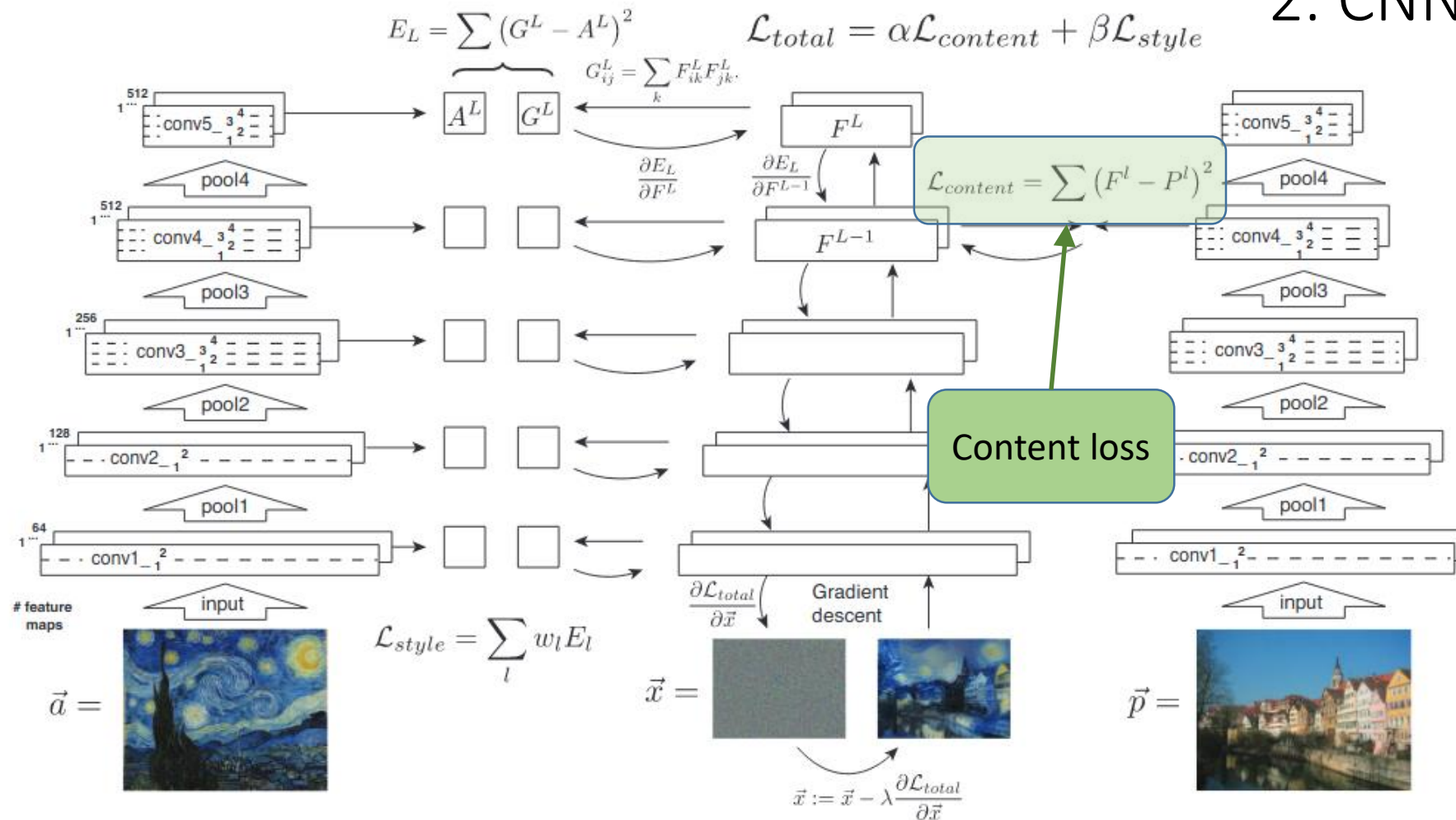
Technologie et fondements scientifiques

2. CNN



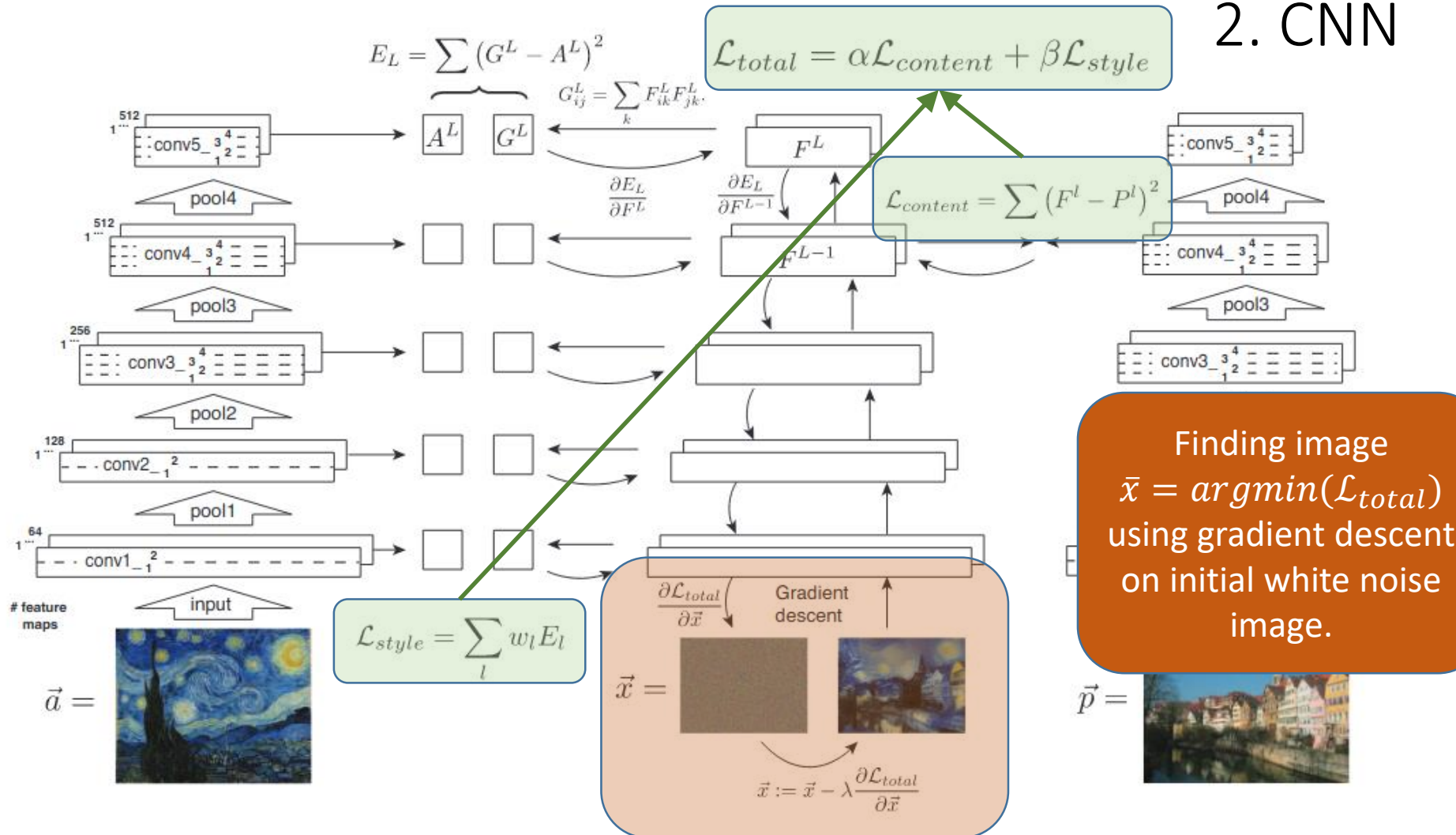
Technologie et fondements scientifiques

2. CNN

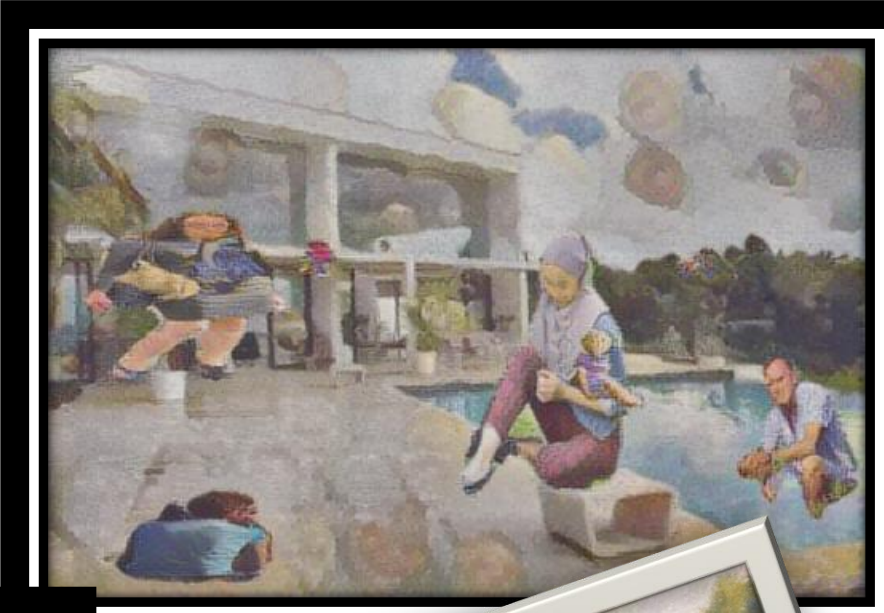


Technologie et fondements scientifiques

2. CNN



Générer des réinterprétations d'œuvres connues ...



Analyse des limites



<https://lamuse.univ-reims.fr>

Références et Code

[1] ***Mask R-CNN***, K. He, G. Gkioxari, P. Dollar, R. Girshick ; Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2961-2969. (<https://arxiv.org/abs/1703.06870>)

[2] ***Exploring the Structure of a Real-time, Arbitrary Neural Artistic Stylization Network***, G. Ghiasi, H. Lee, M. Kudlur, V. Dumoulin, J. Shlens ; Proceedings of the British Machine Vision Conference (BMVC), 2017. (<https://arxiv.org/abs/1705.06830>)

- https://github.com/matterport/Mask_RCNN
- <https://tfhub.dev/google/magenta/arbitrary-image-stylization-v1-256/2>