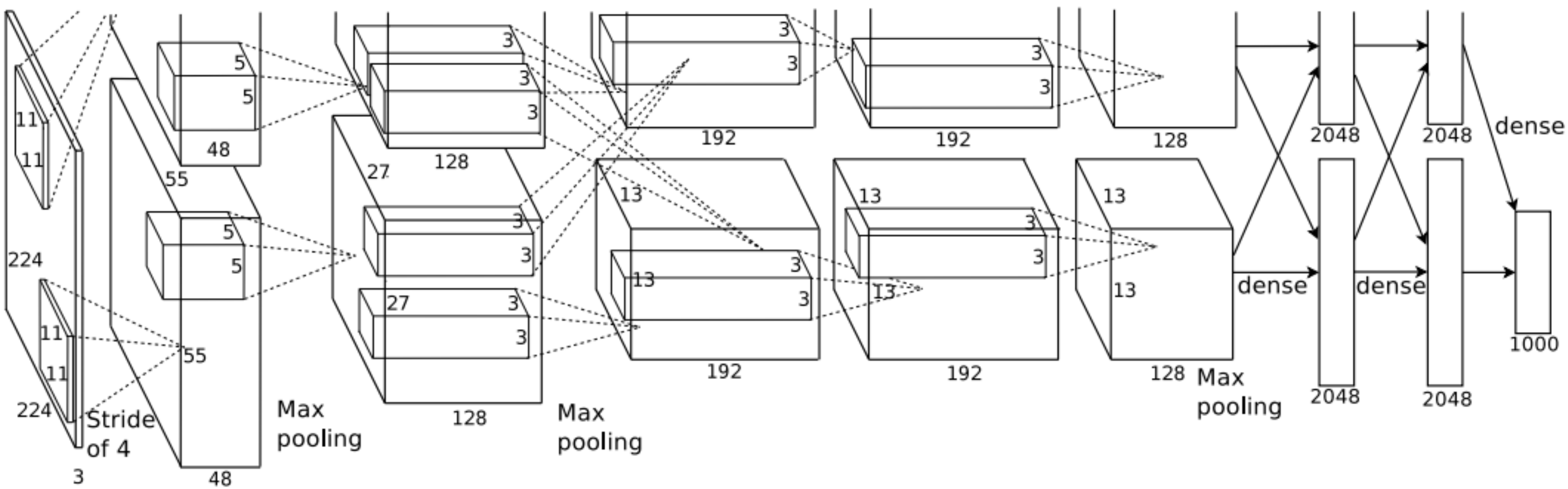
Génération d'images par apprentissage profond



GIF-4105/7105 Photographie Algorithmique
Jean-François Lalonde

Merci à Alexei Efros, James Hays, Philip Isola, Andrew Owens, Andrea Vedaldi, Derek Hoiem

140 milliard d'images
6 milliard ajoutée à chaque mois

6 milliard d'mages

1 milliard d'images
accédées par jour

72 heures téléversées à
chaque minute

3.5 trillion photographs

90% du traffic sur Internet sera des données *visuelles*

P. Isola

30% des vidéos sur Youtube ont moins de 10 vues

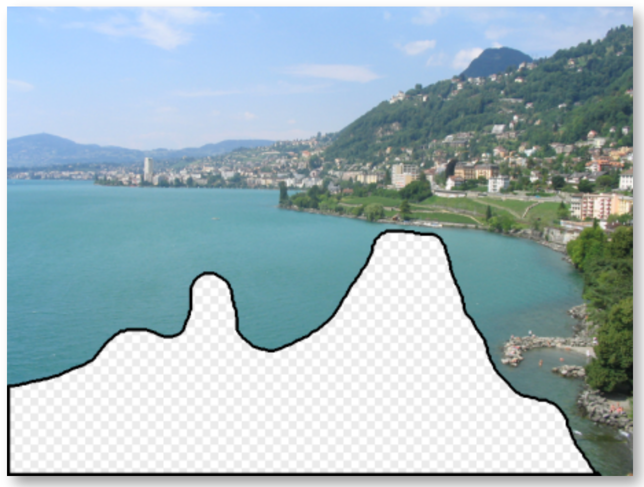
«Digital Dark Matter»

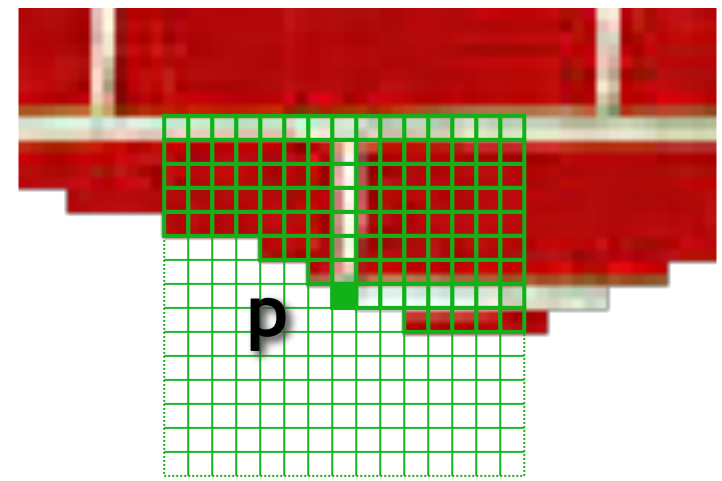
[Perona 2010]

Défi principal

- Comment utiliser toutes ces données?
- Avec l'apprentissage profond!
- Mais tout d'abord, utilisons un exemple de système (sans «deep learning») qui utilise des données massives

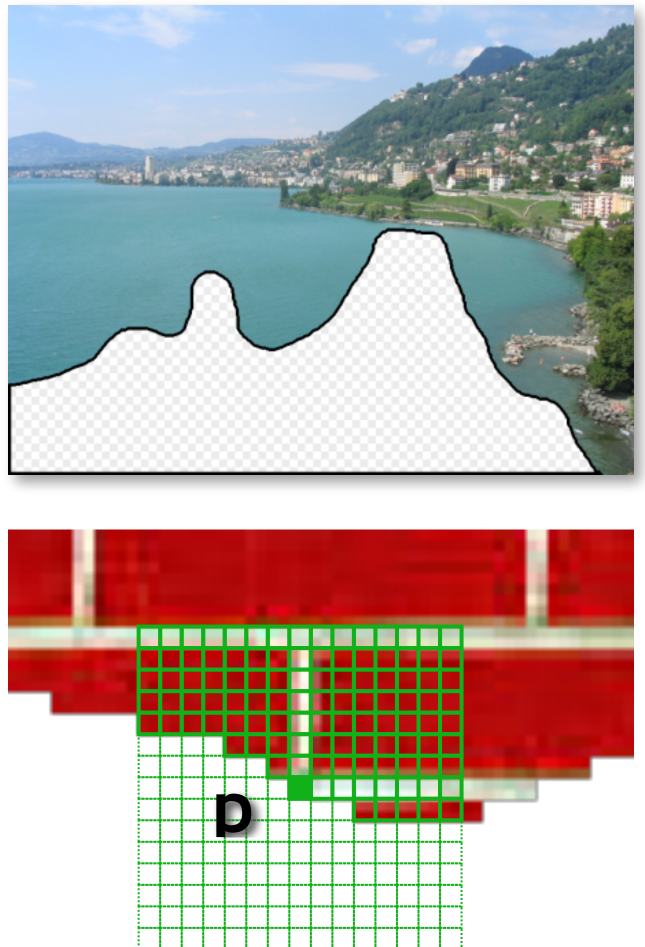
Rappel: limites





- Lent
- Fonction d'appariement définie manuellement (peut sembler arbitraire)
- Chaque méthode fonctionne seulement pour leur domaine particulier (e.g. scènes «typiques» à l'extérieur, textures semi-régulières, etc.)

P. Isola

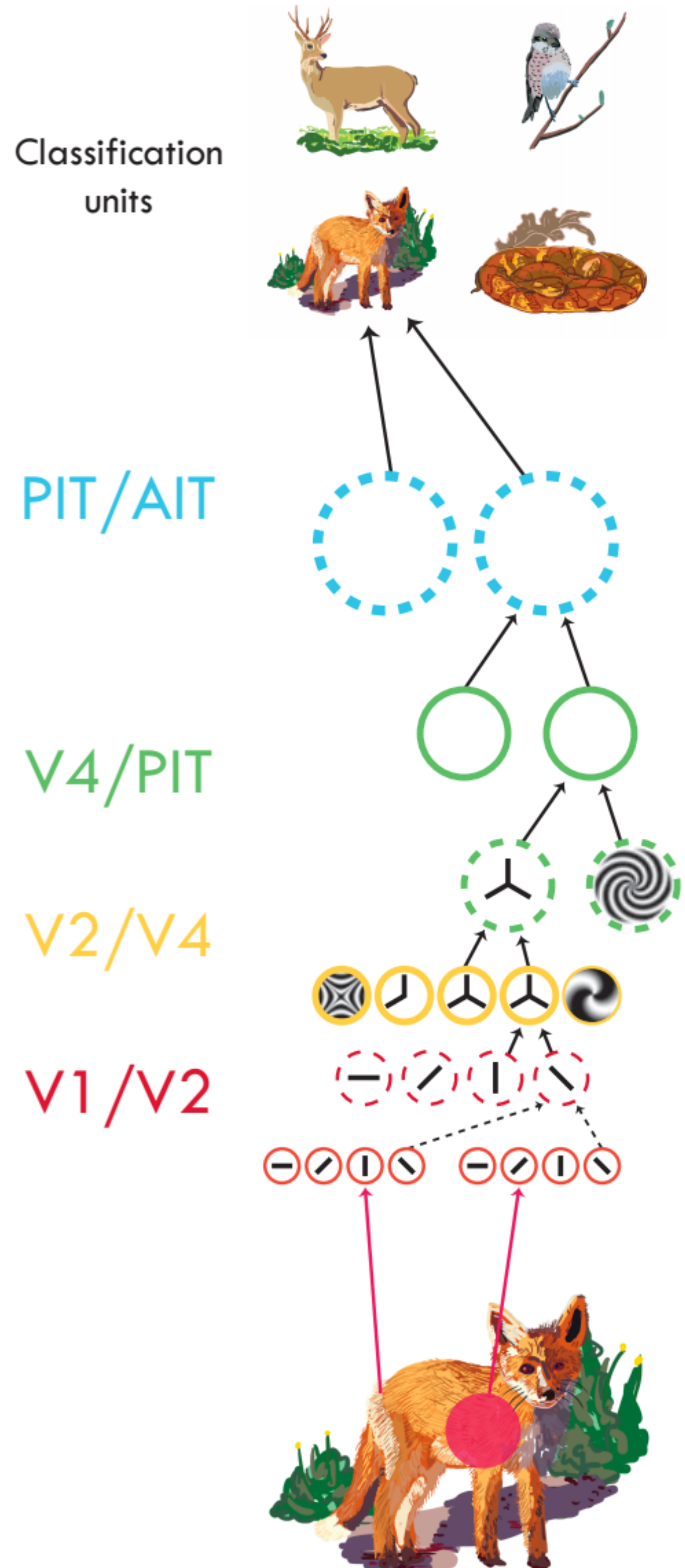


Stocke toute l'information
dans une mémoire

«lookup table»

Assume que le monde
possède une structure simple

Apprend une représentation
qui capture cette structure



Serre, 2014

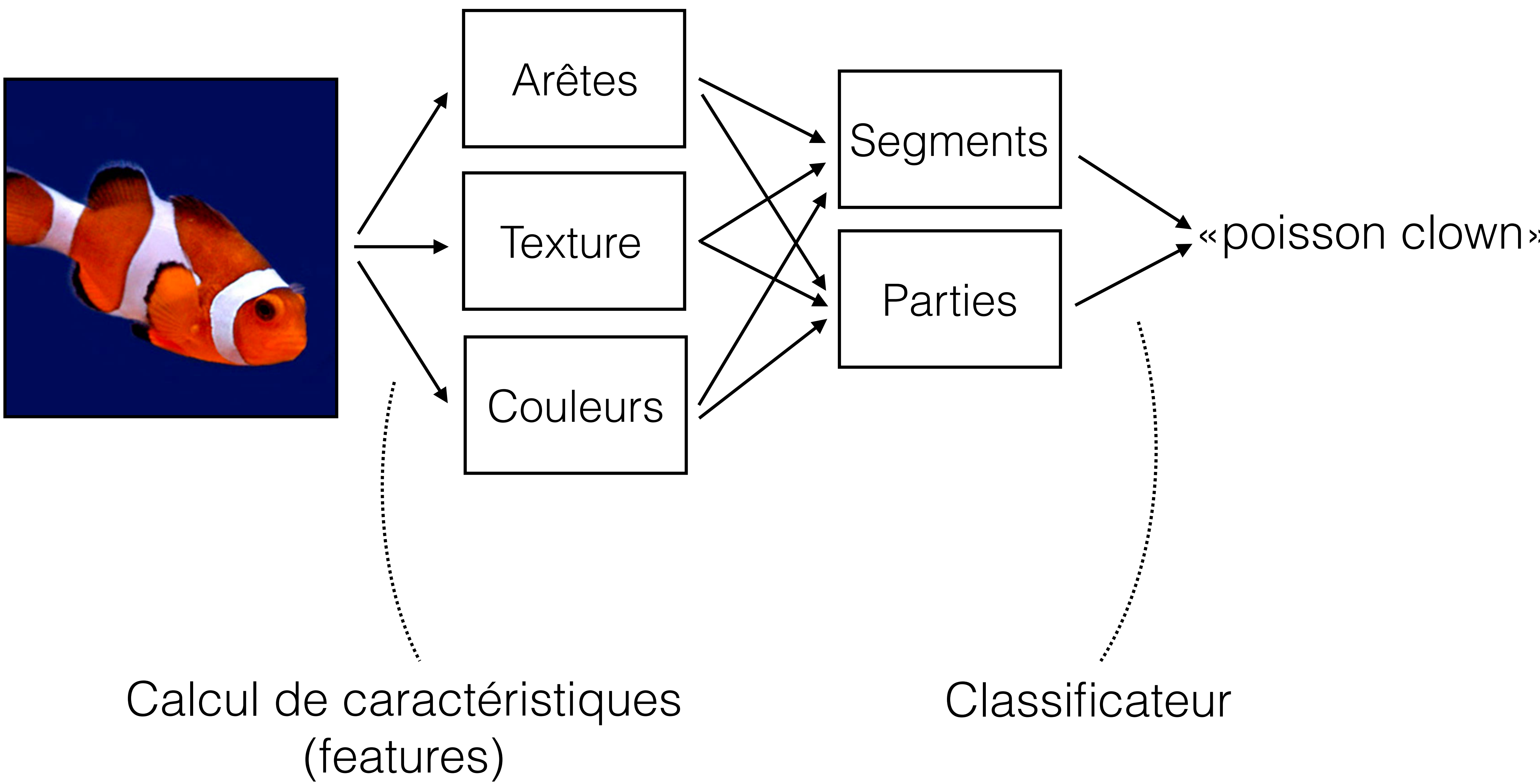
Idée de base

→ Cerveau/Machine → «poisson clown»

Hiérarchie d'unités *simples*

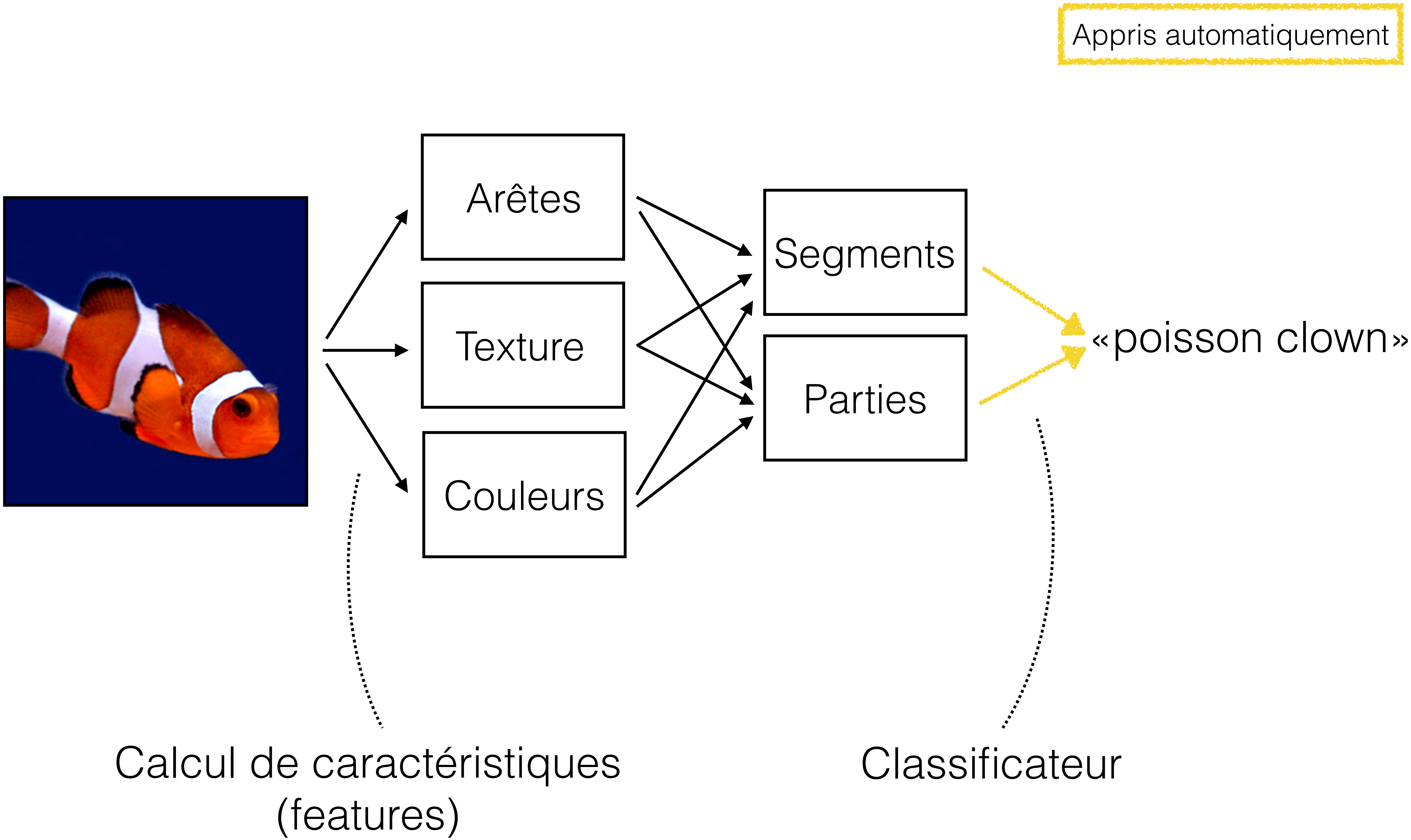
P. Isola

Reconnaissance d’objets: approche «traditionnelle»



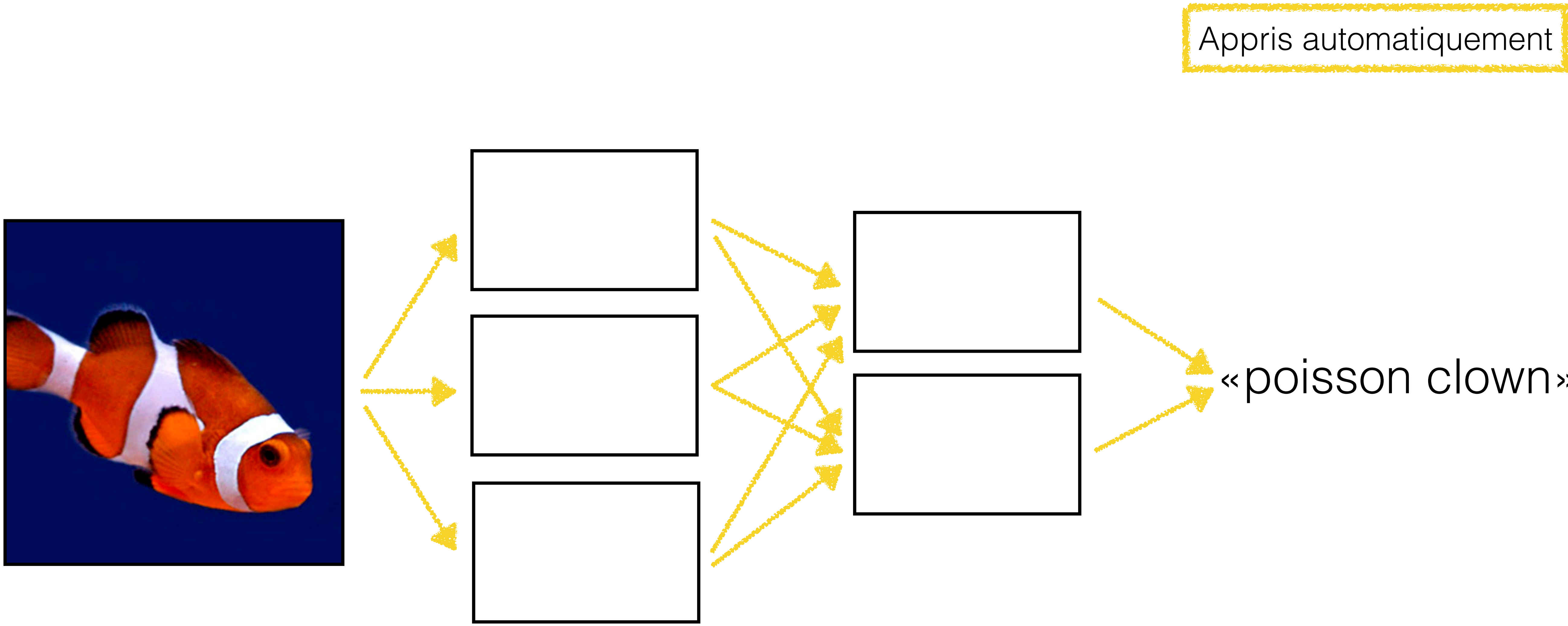
P. Isola

Reconnaissance d’objets: approche «traditionnelle»



P. Isola

Réseau de neurones



P. Isola

Réseau de neurones

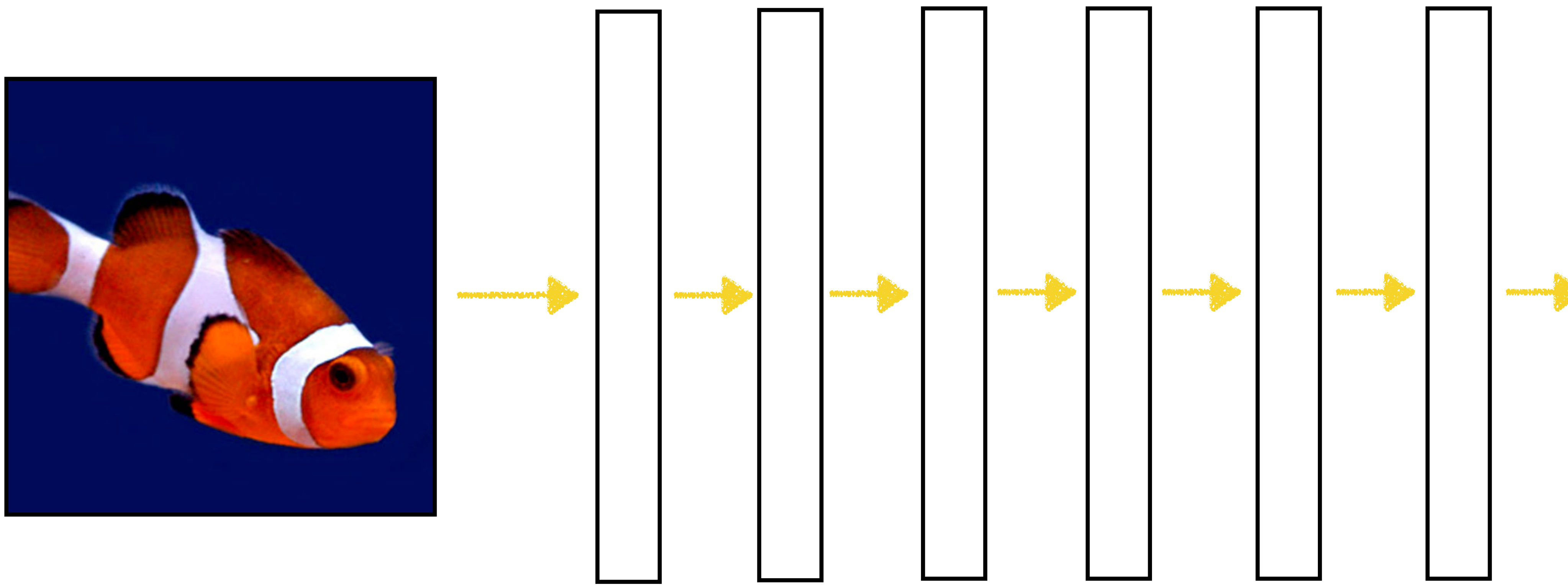
Appris automatiquement

«poisson clown»

P. Isola

Réseau de neurones *profond*

Appris automatiquement



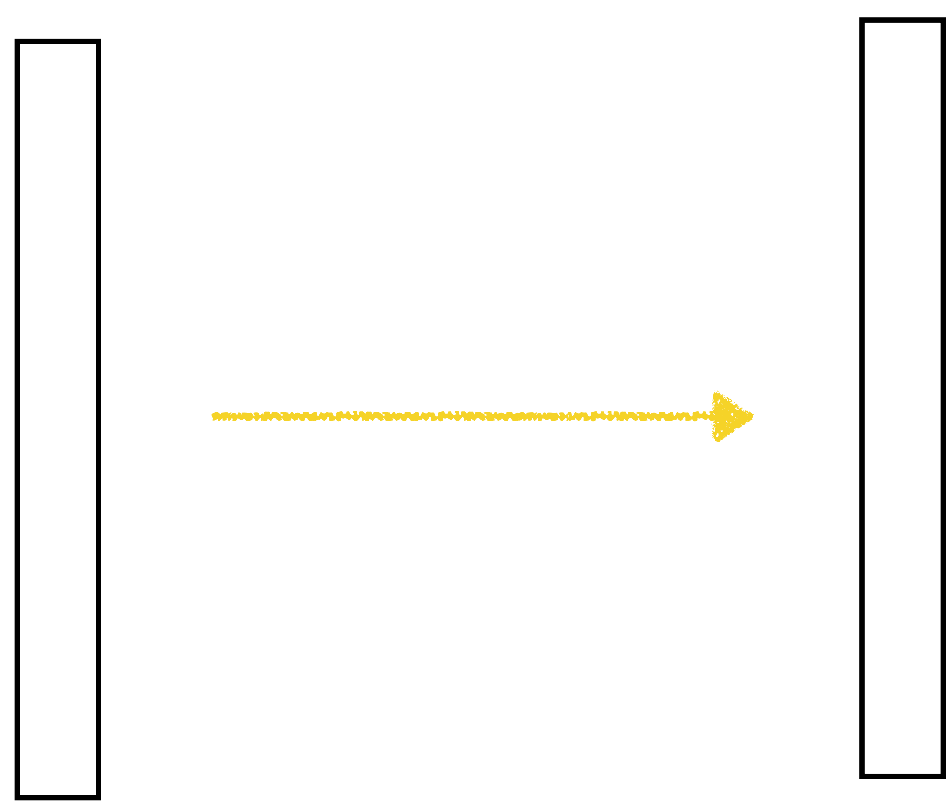
«poisson clown»

P. Isola

Calculs dans un réseau de neurones

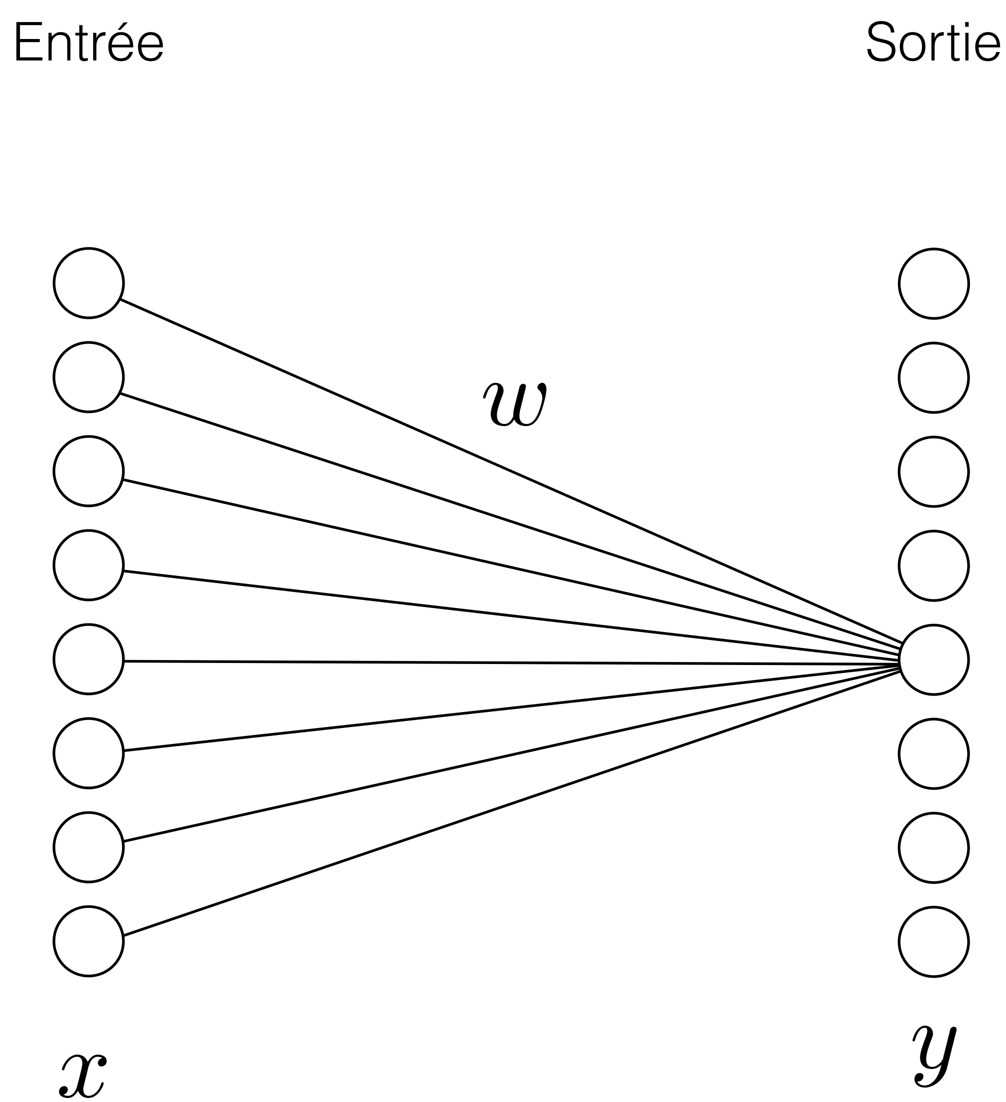
Que font ces flèches?
Quels calculs sont effectués?

Entrée Sortie



P. Isola

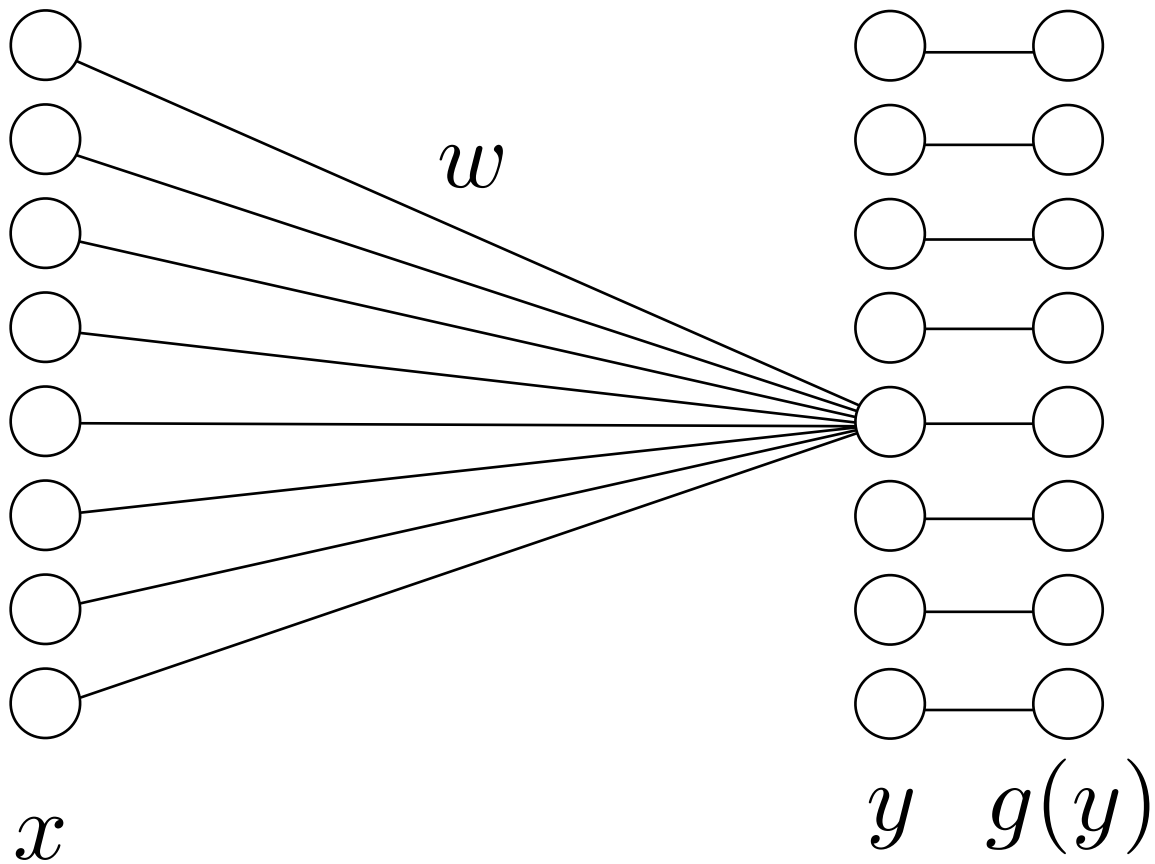
Calculs dans un réseau de neurones



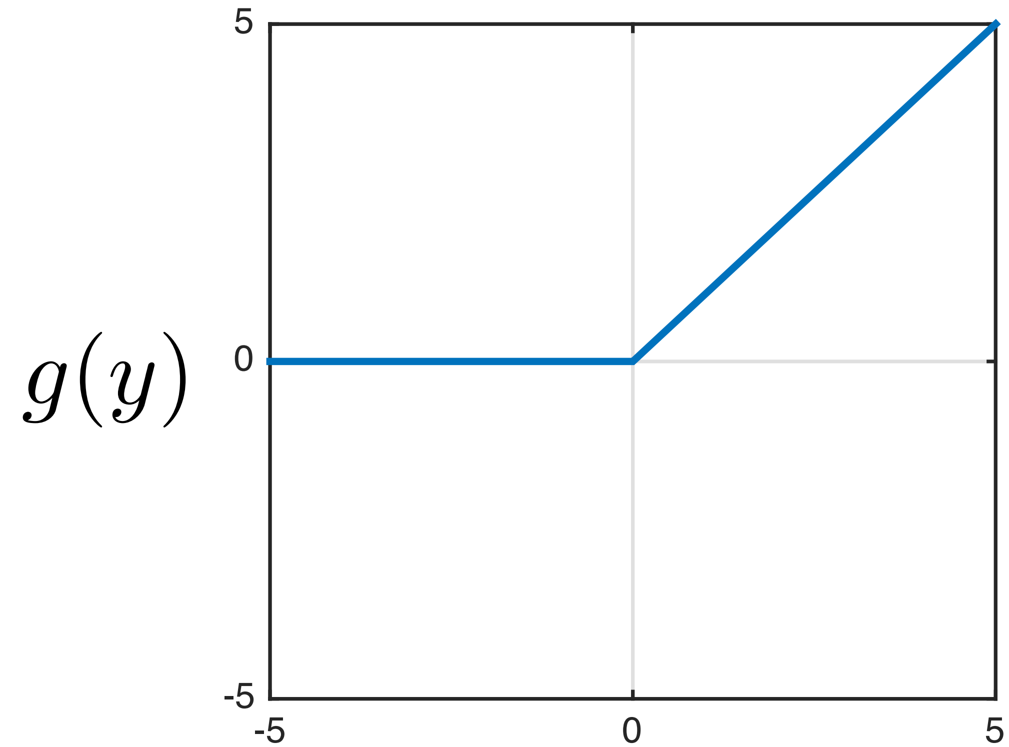
$$y_j = \sum_i w_{ij} x_i$$

P. Isola

Calculs dans un réseau de neurones



«Rectified linear unit» (ReLU)



$$g(y) = \max(0, y)$$

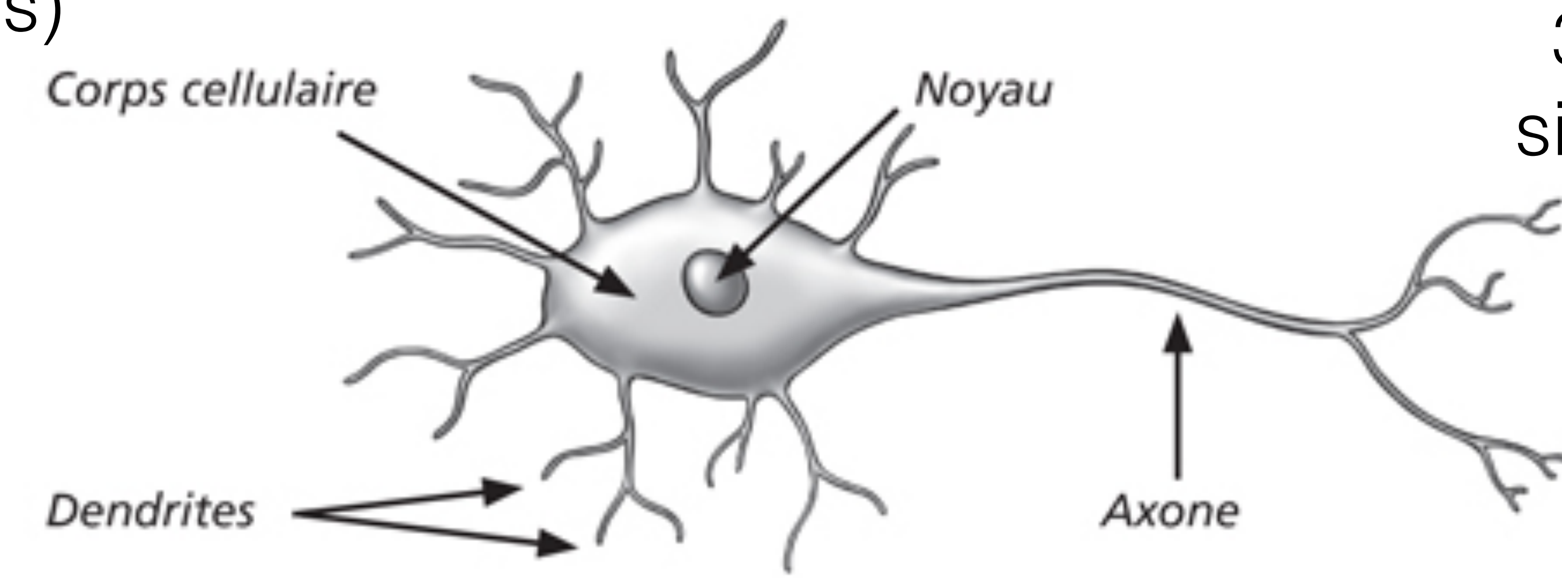
P. Isola

Inspiration: neurone biologique

$$g = f\left(\sum_i w_i x_i\right)$$

1. signaux provenant d'autres neurones (entrées)

3. signal transmis si potentiel suffisant (activation)

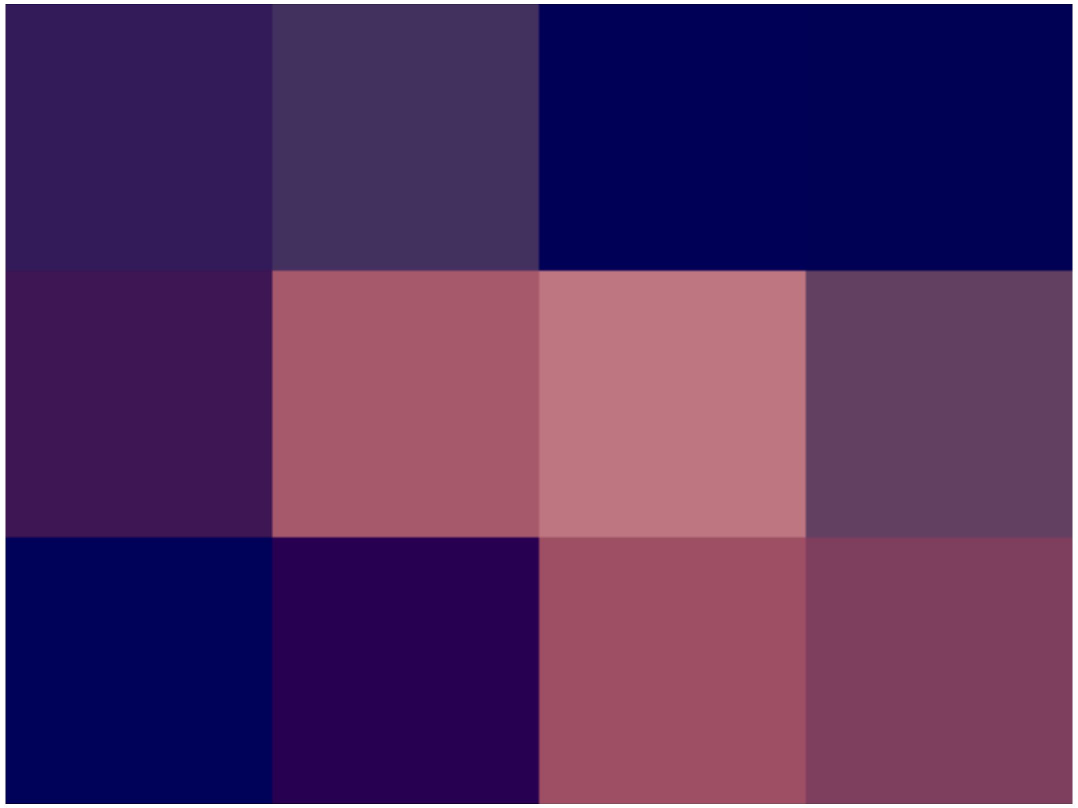


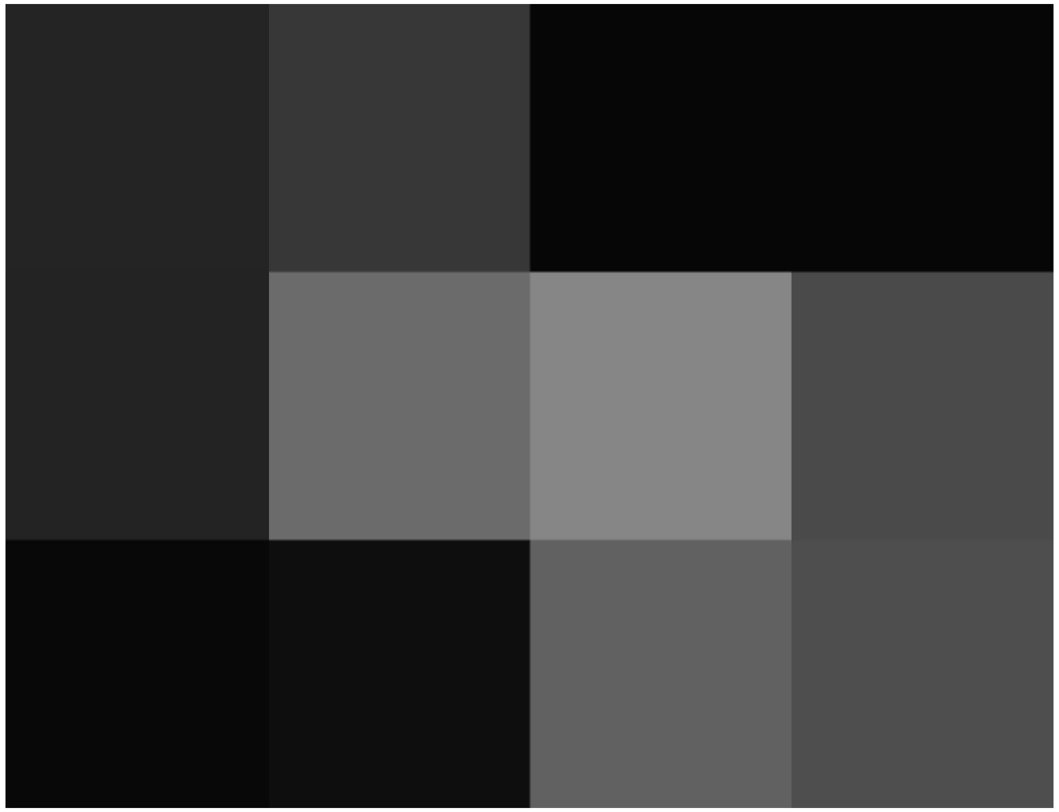
2. accumulation (ou soustraction) de potentiel électrique (somme pondérée)

Image: Éditions Thierry Souccar

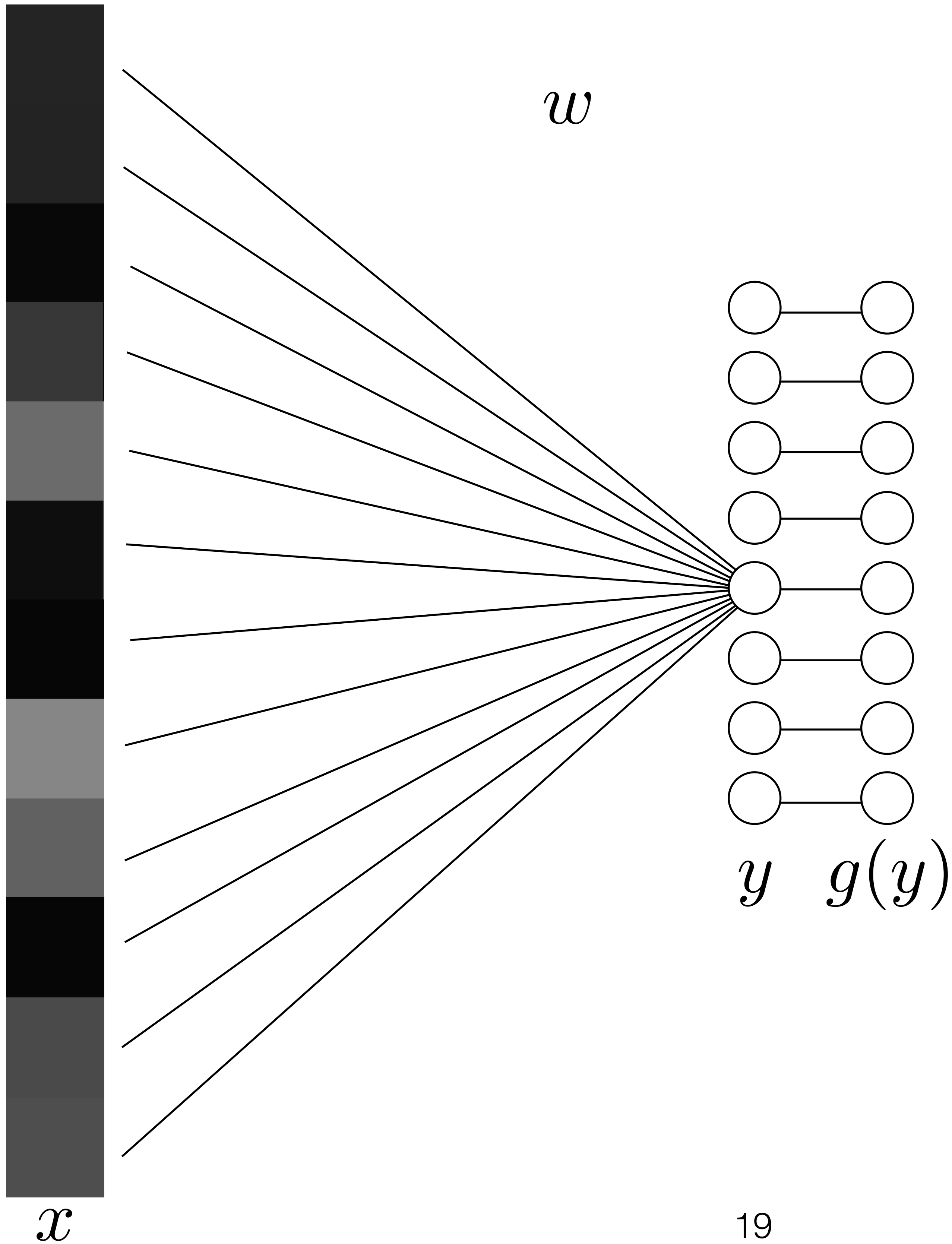
Comment traiter une image?

Ex: résolution 3x4

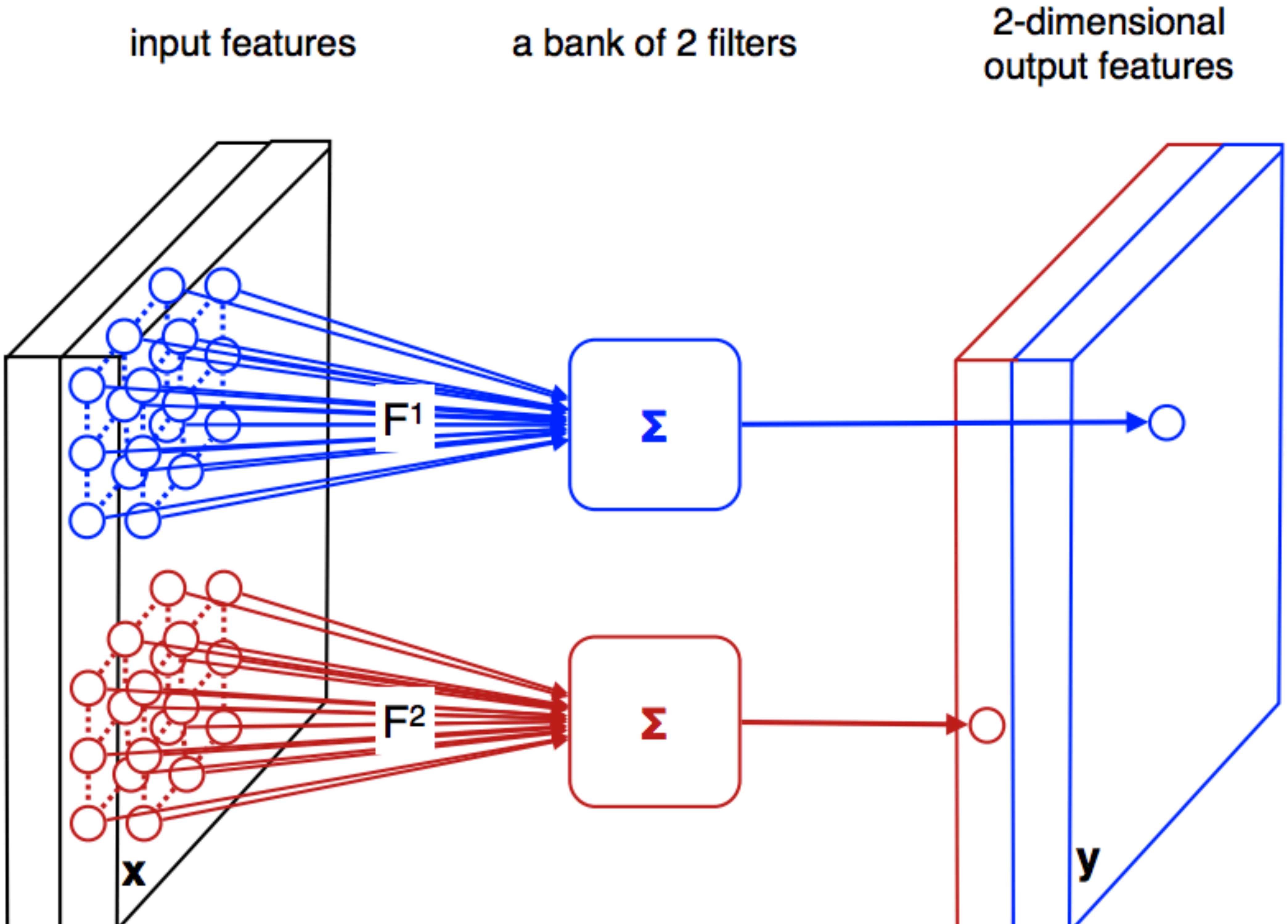




Comment traiter une image?



Réseaux de neurones *à convolution* (CNN)

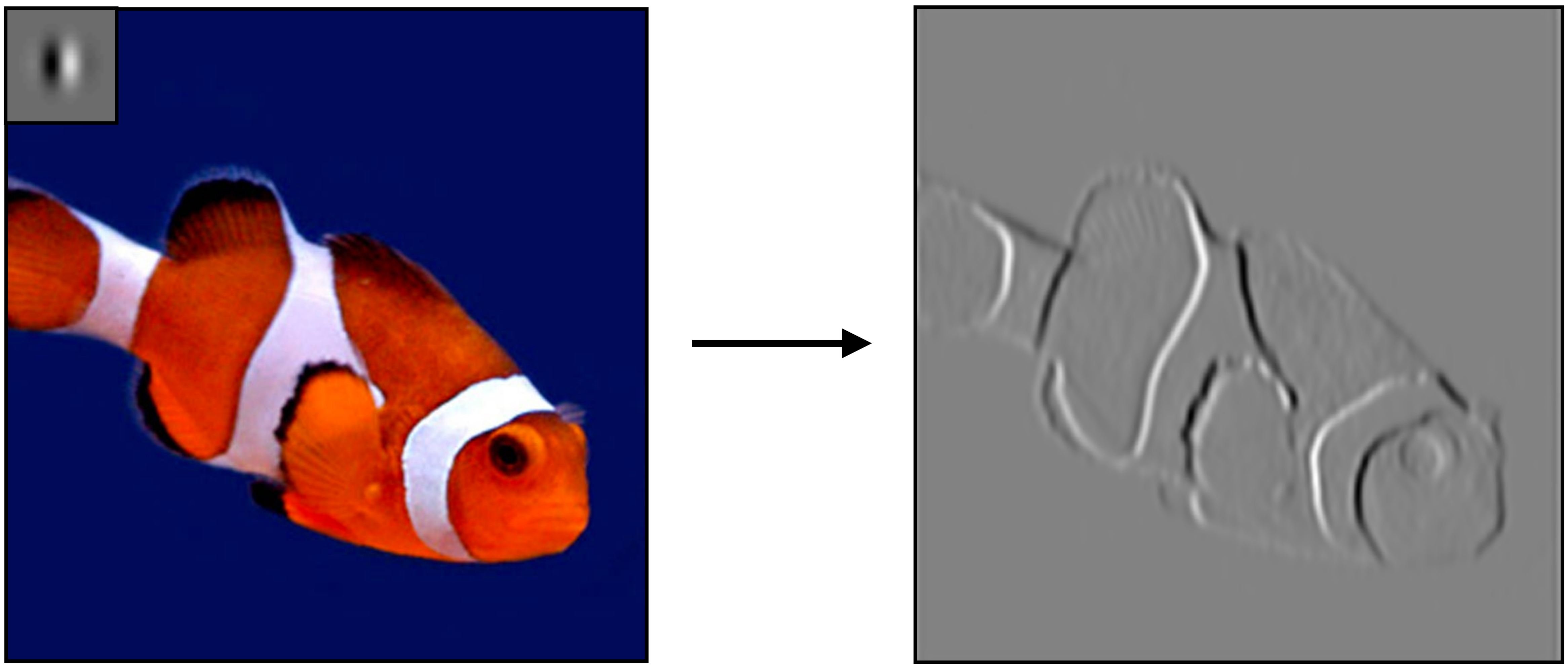


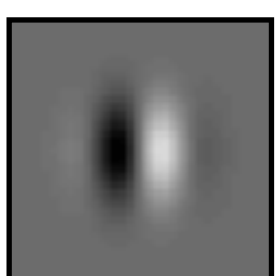
A. Vedaldi

Réseaux de neurones à convolution (CNN)

Convolution

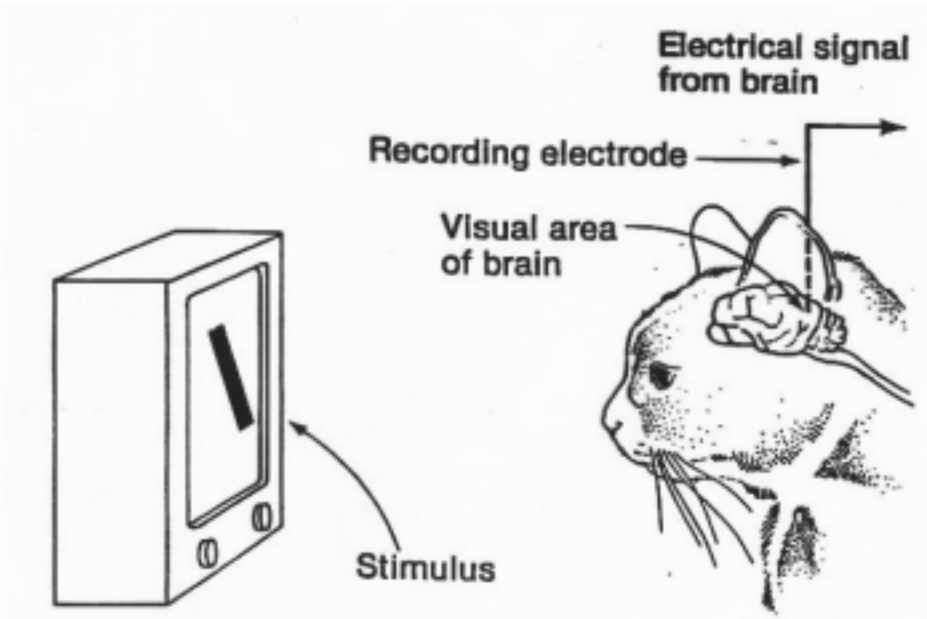
Avantage: les paramètres sont *partagés*
(toutes les fenêtres de l'image sont traitées de la même façon)

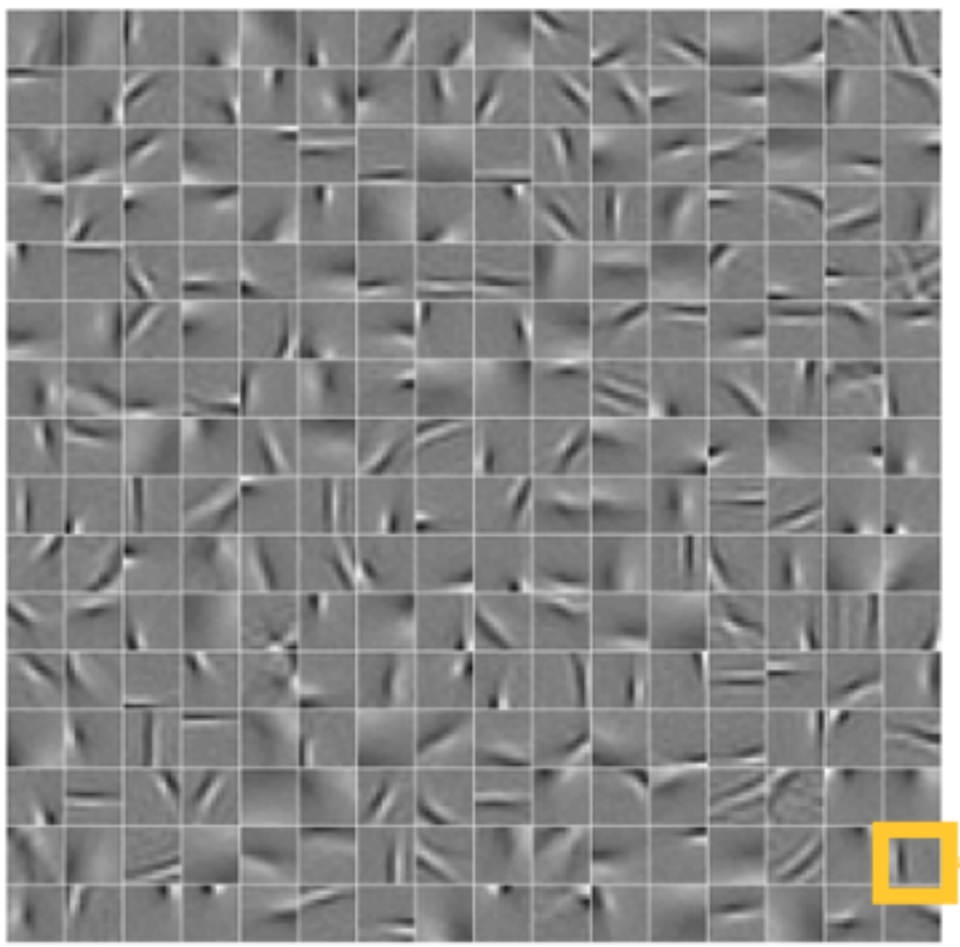


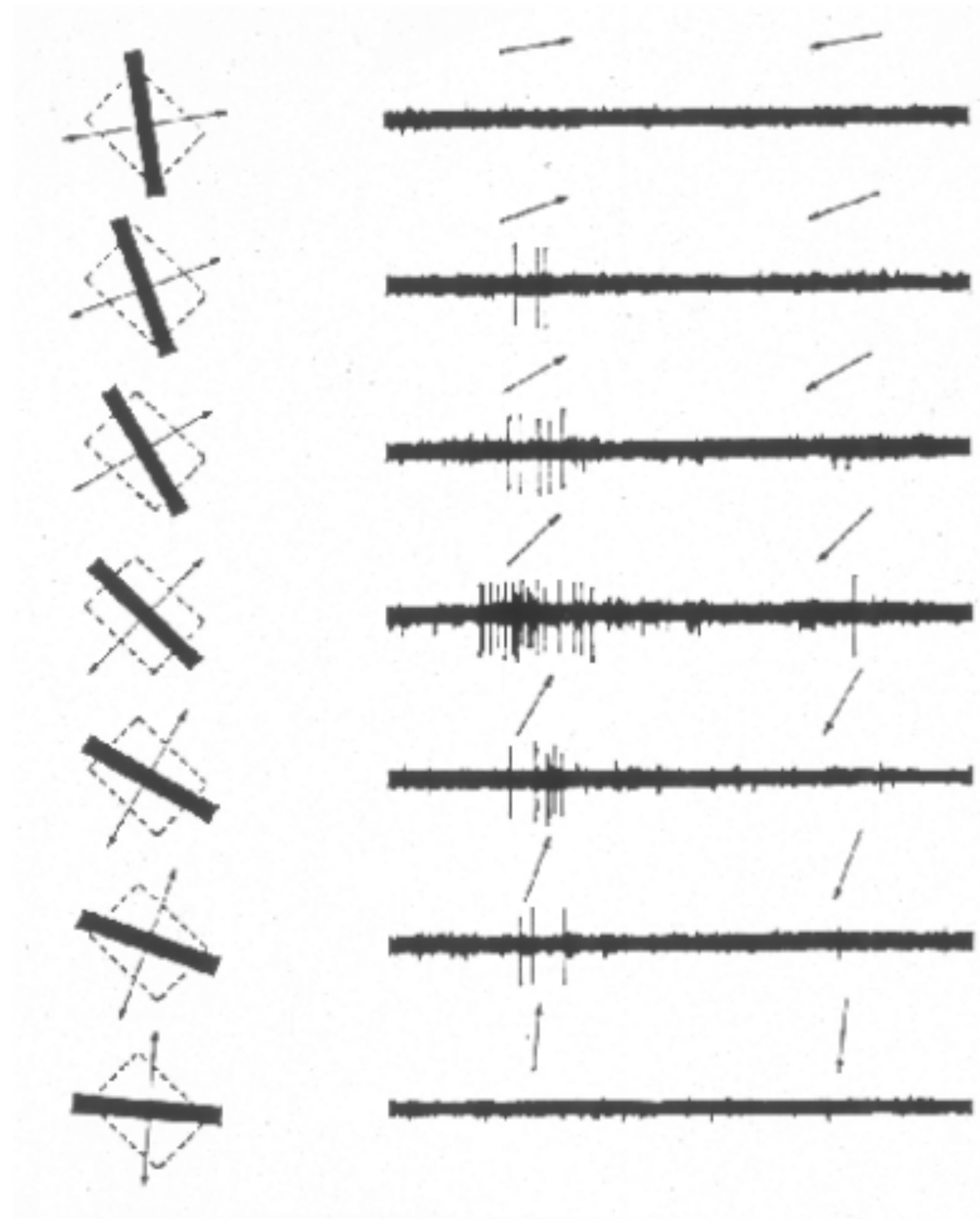
filtre 

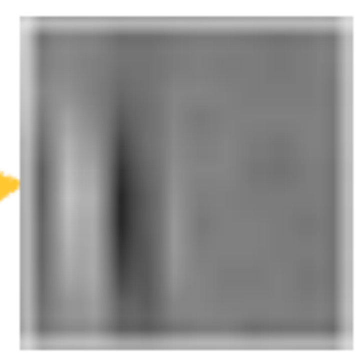
P. Isola

[Hubel and Wiesel 59]





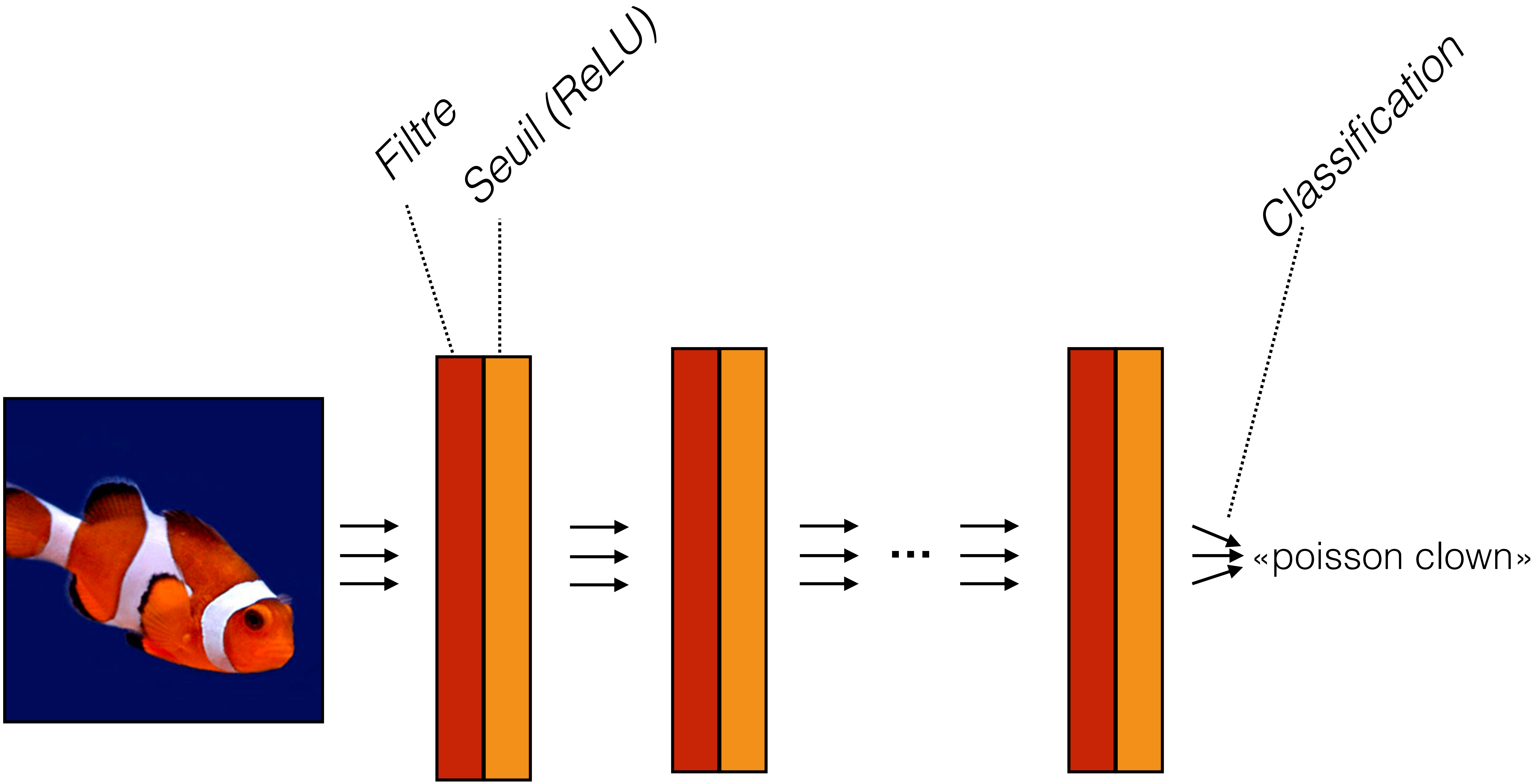




oriented filter

A. Vedaldi

Calculs dans un réseau de neurones

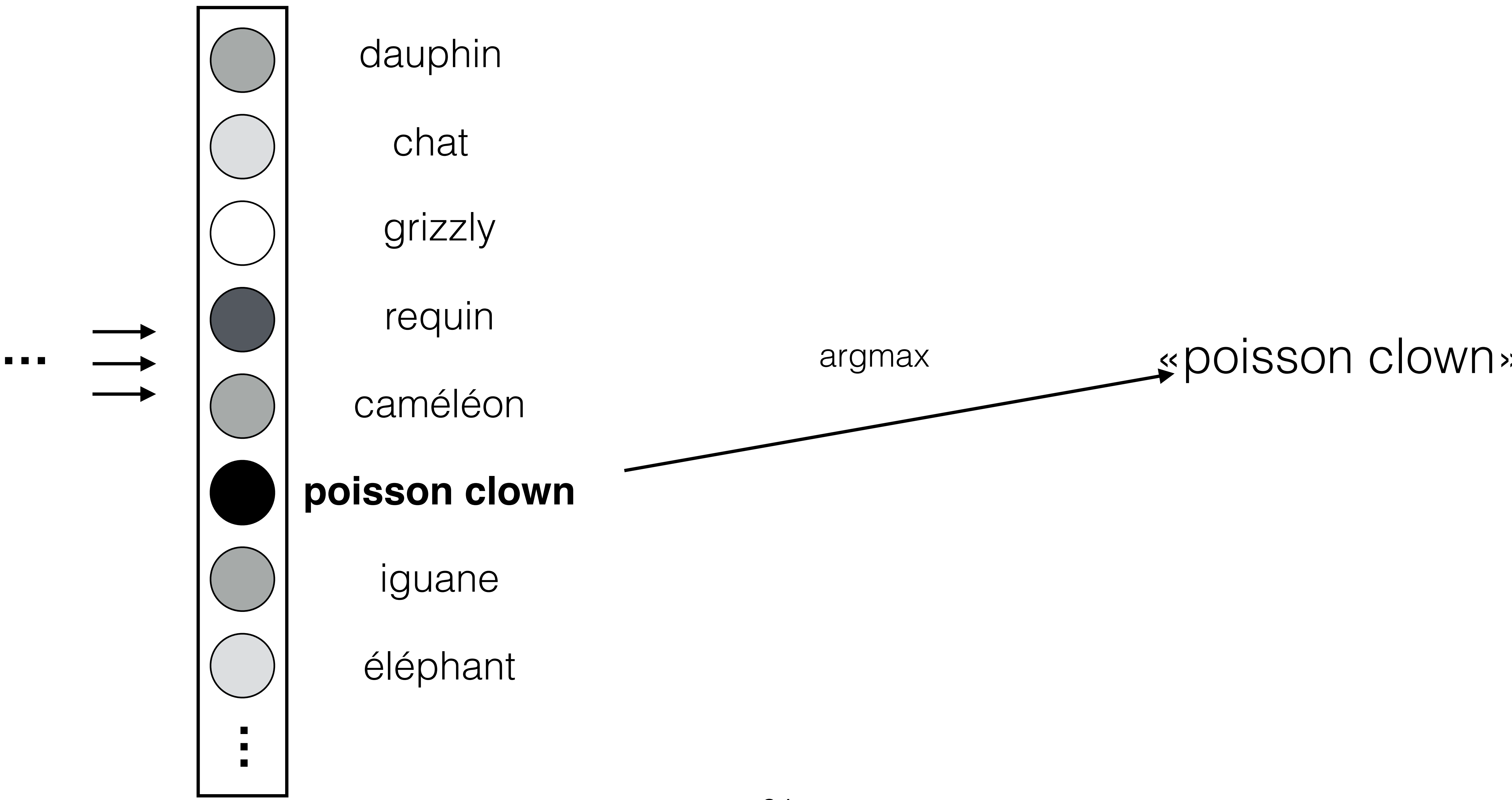


$$f(\mathbf{x}) = f_L(\dots f_2(f_1(\mathbf{x})))$$

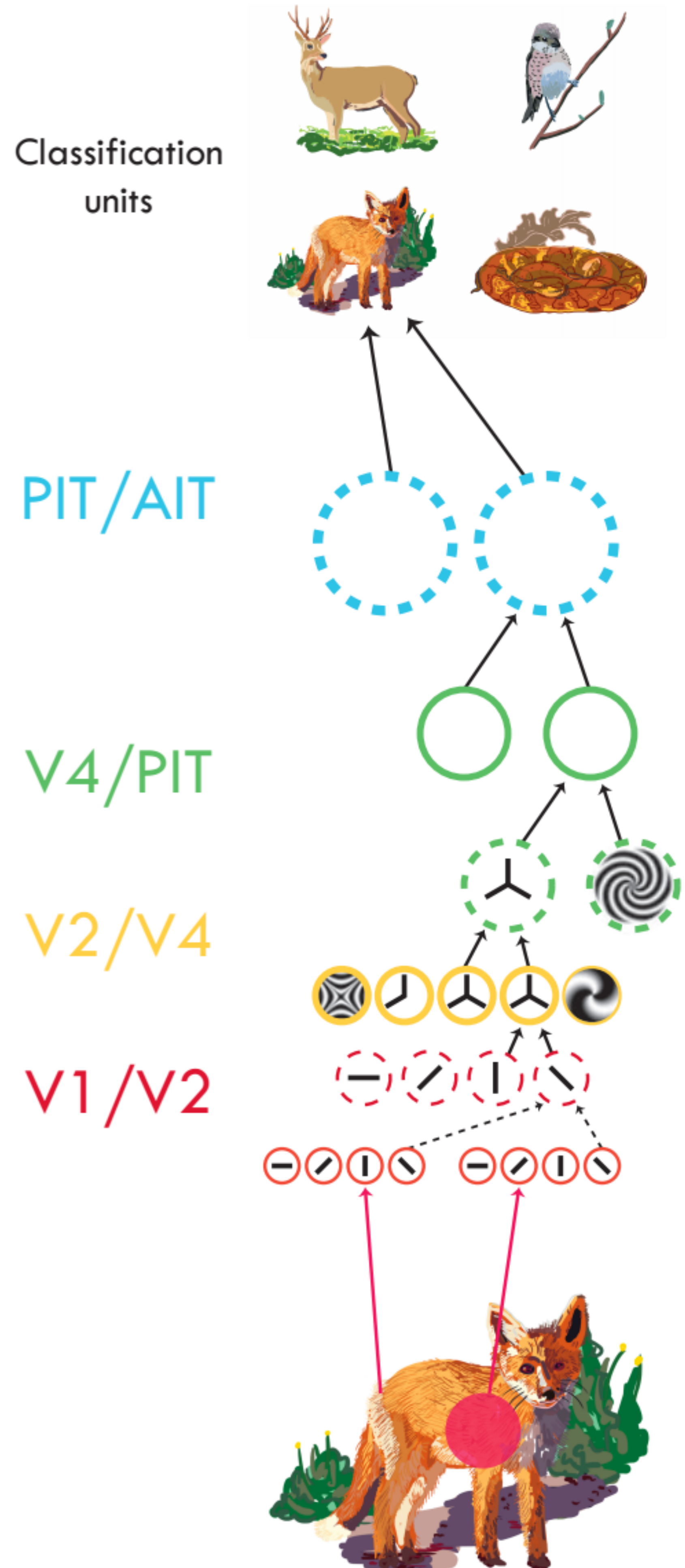
P. Isola

Calculs dans un réseau de neurones

Dernière couche
(classification)

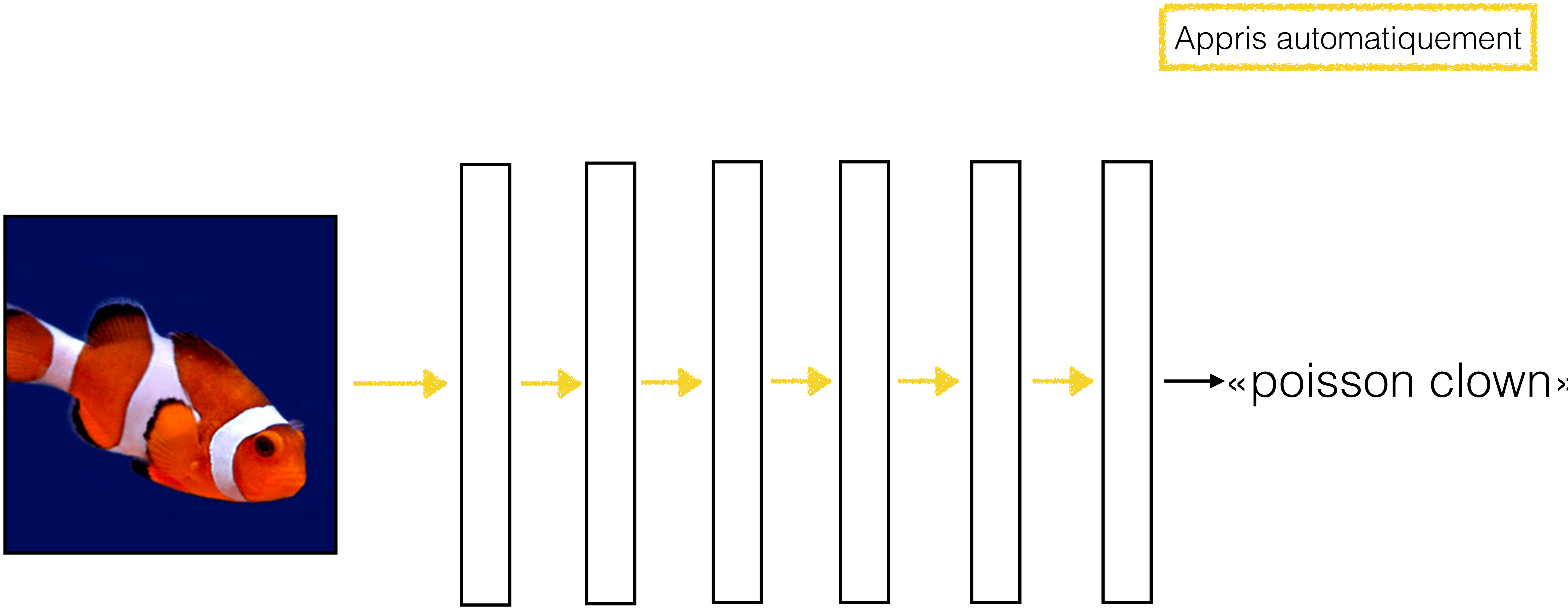


P. Isola



Serre, 2014

Apprentissage par réseaux profonds



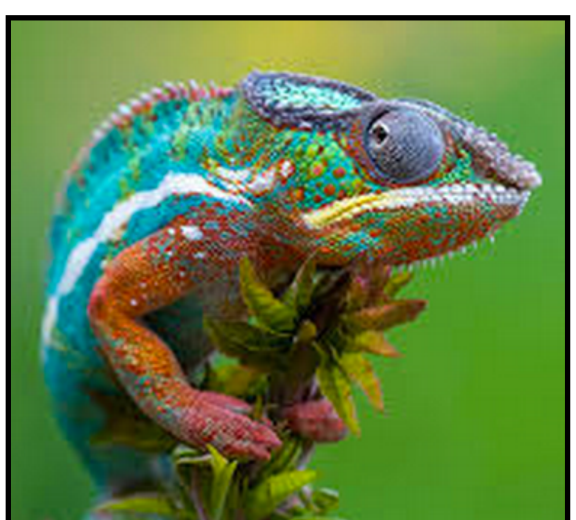
Idée: on modifie les poids (dans un CNN, les filtres!) jusqu'à temps que la sortie corresponde à ce qu'on veut

P. Isola

Apprentissage par réseaux profonds

→ «poisson clown»

→ «grizzly»



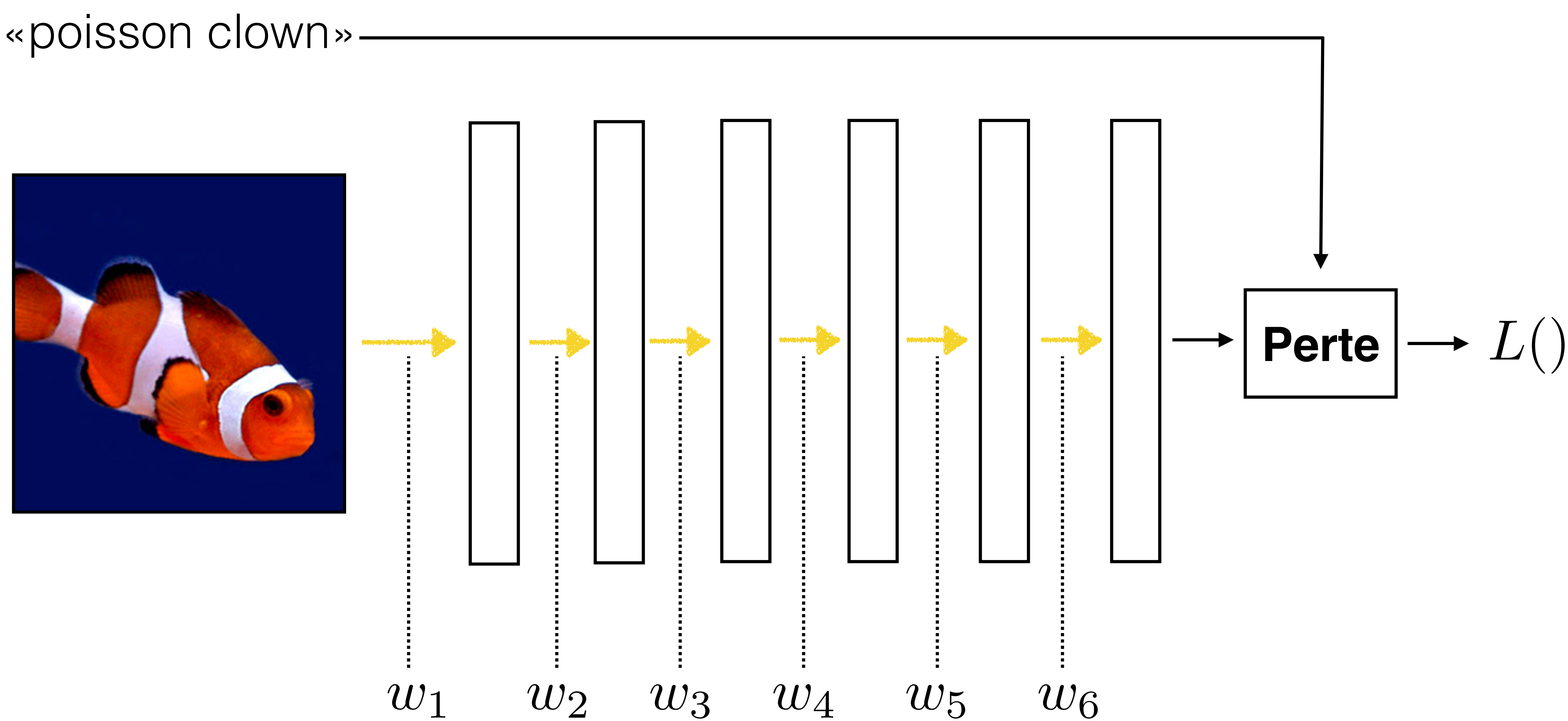
→ «caméléon»

Nous voudrions entraîner le réseau à associer chaque image à la bonne étiquette

P. Isola

Apprentissage par réseaux profonds

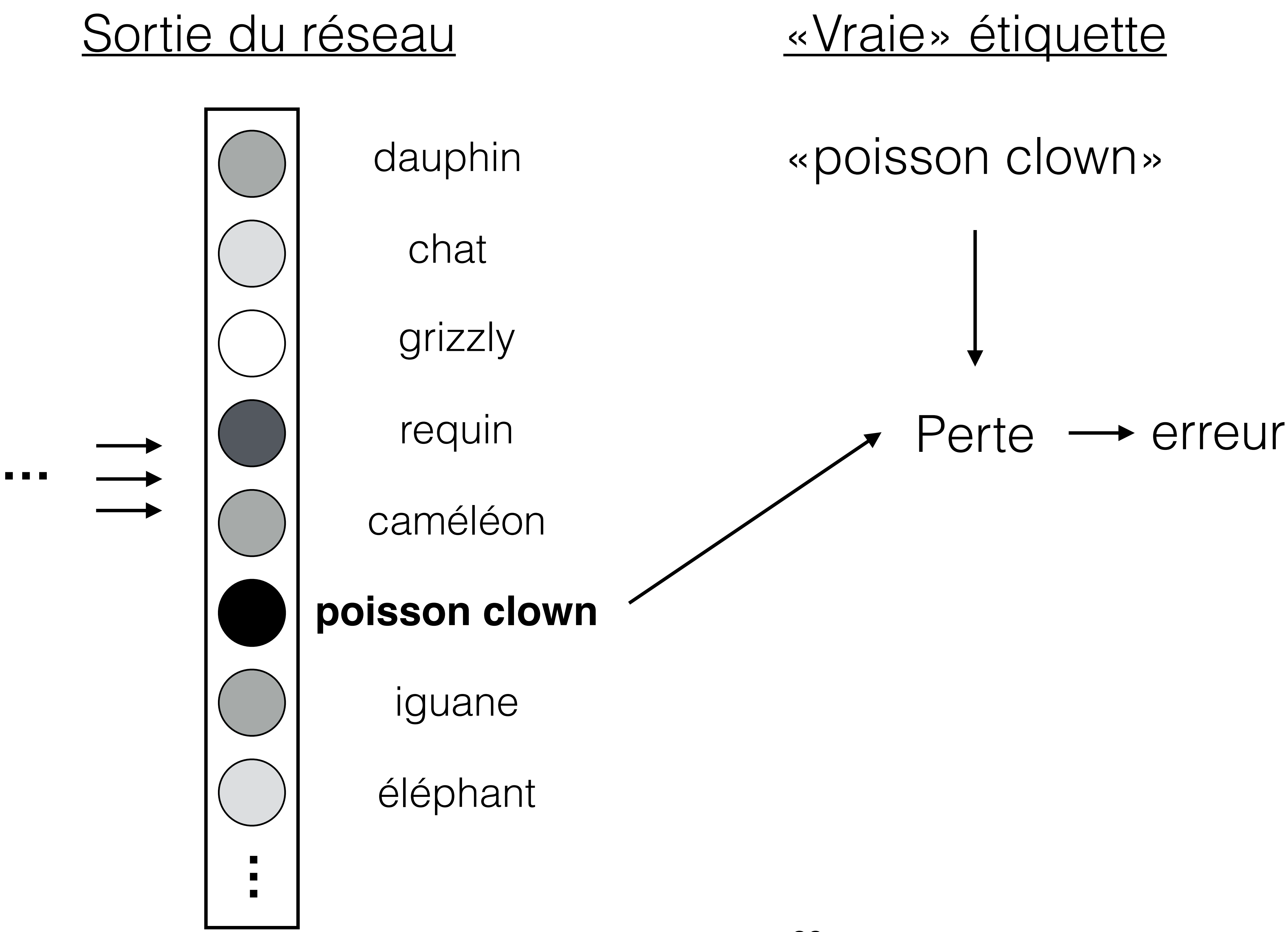
Appris automatiquement



$$\boxed{\underset{\mathbf{w}}{\operatorname{argmin}} \; L(w_1, \ldots, w_6)}$$

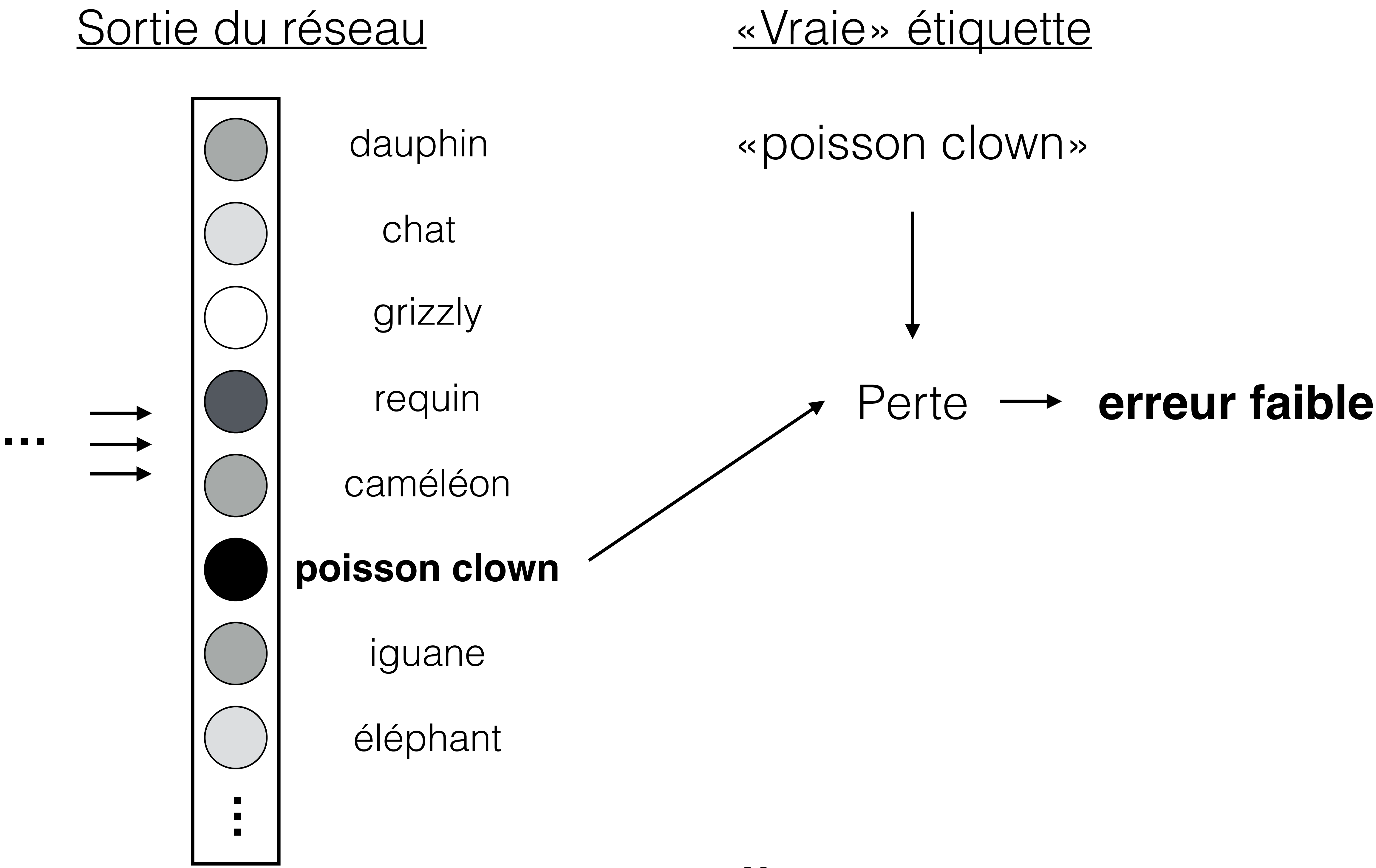
P. Isola

Fonction de perte (*loss function*)



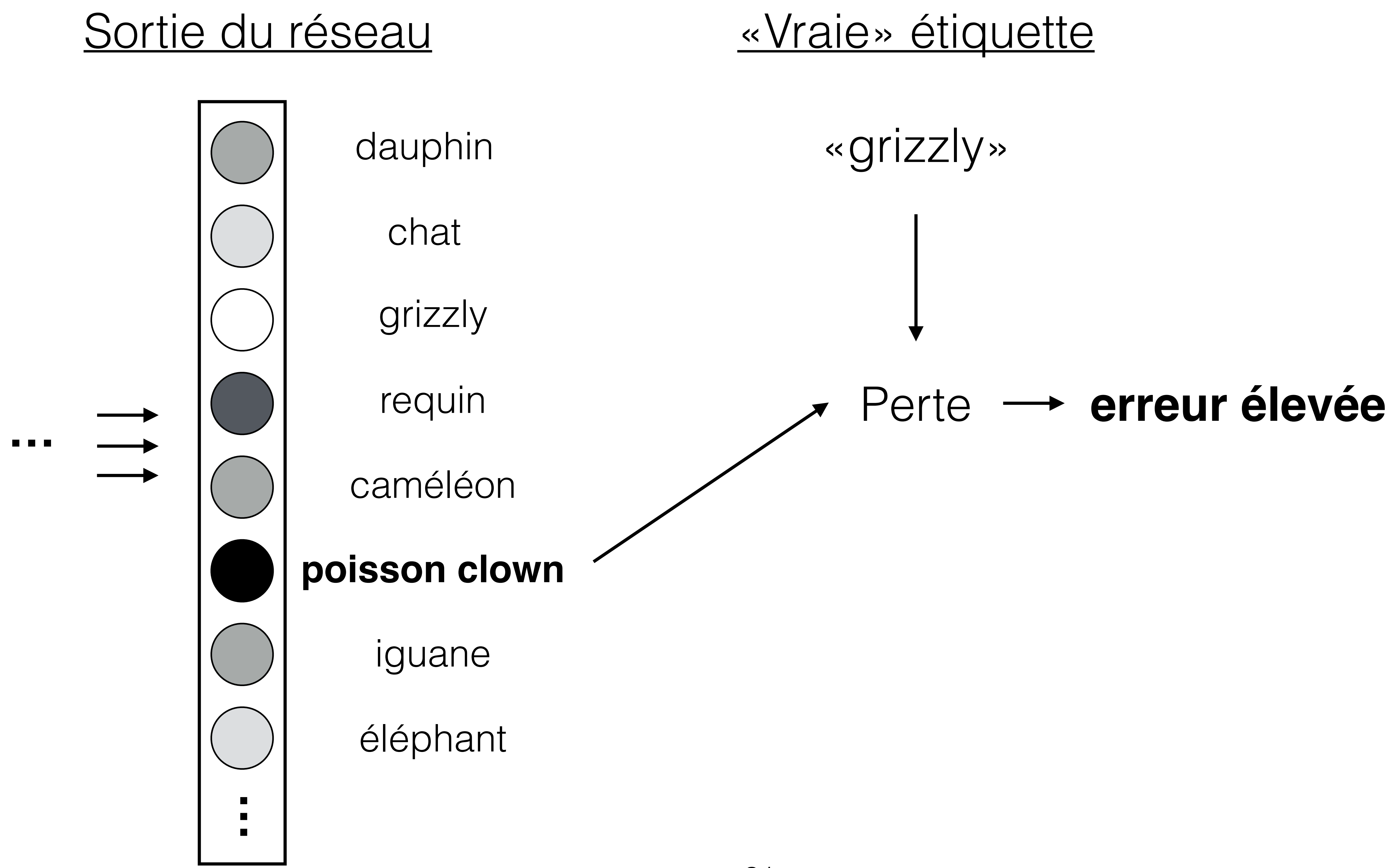
P. Isola

Fonction de perte (*loss function*)



P. Isola

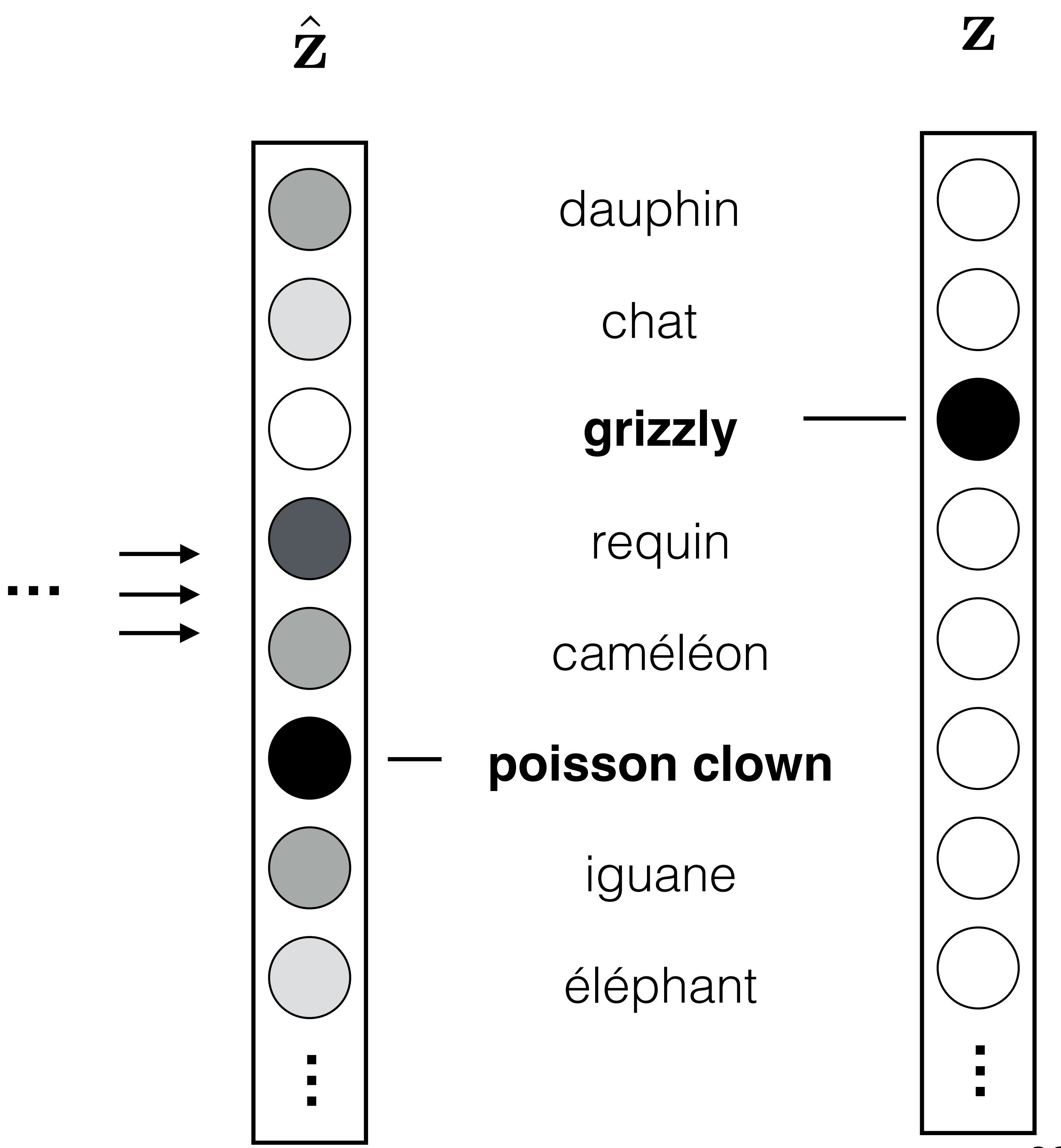
Fonction de perte (*loss function*)



P. Isola

Fonction de perte pour classification

<u>Sortie du réseau</u> <u>«Vraie» étiquette</u>



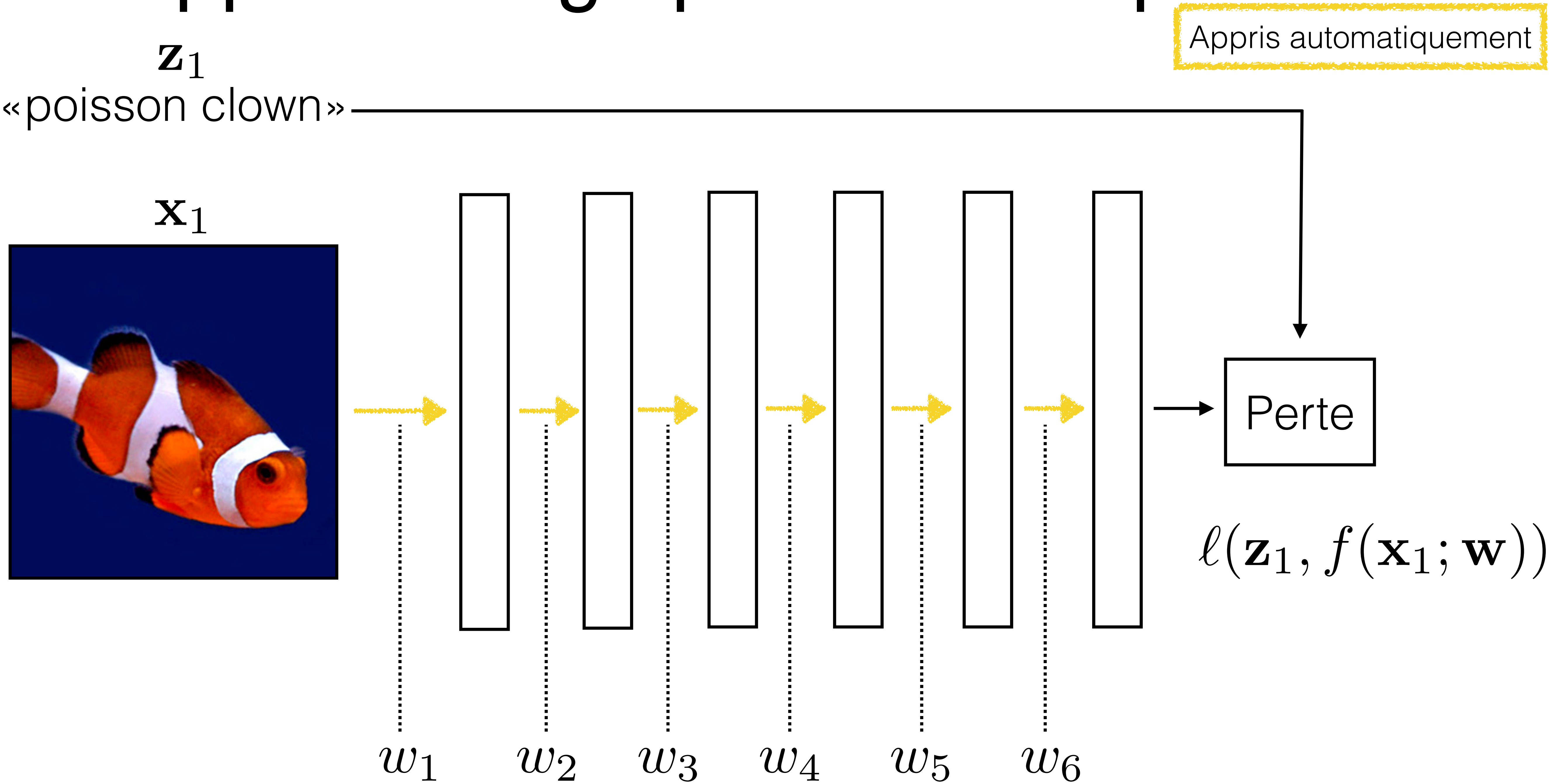
Probabilité des données observées sous le modèle

$$H(\hat{\mathbf{z}}, \mathbf{z}) = -\sum_c \hat{\mathbf{z}}_c \log \mathbf{z}_c$$

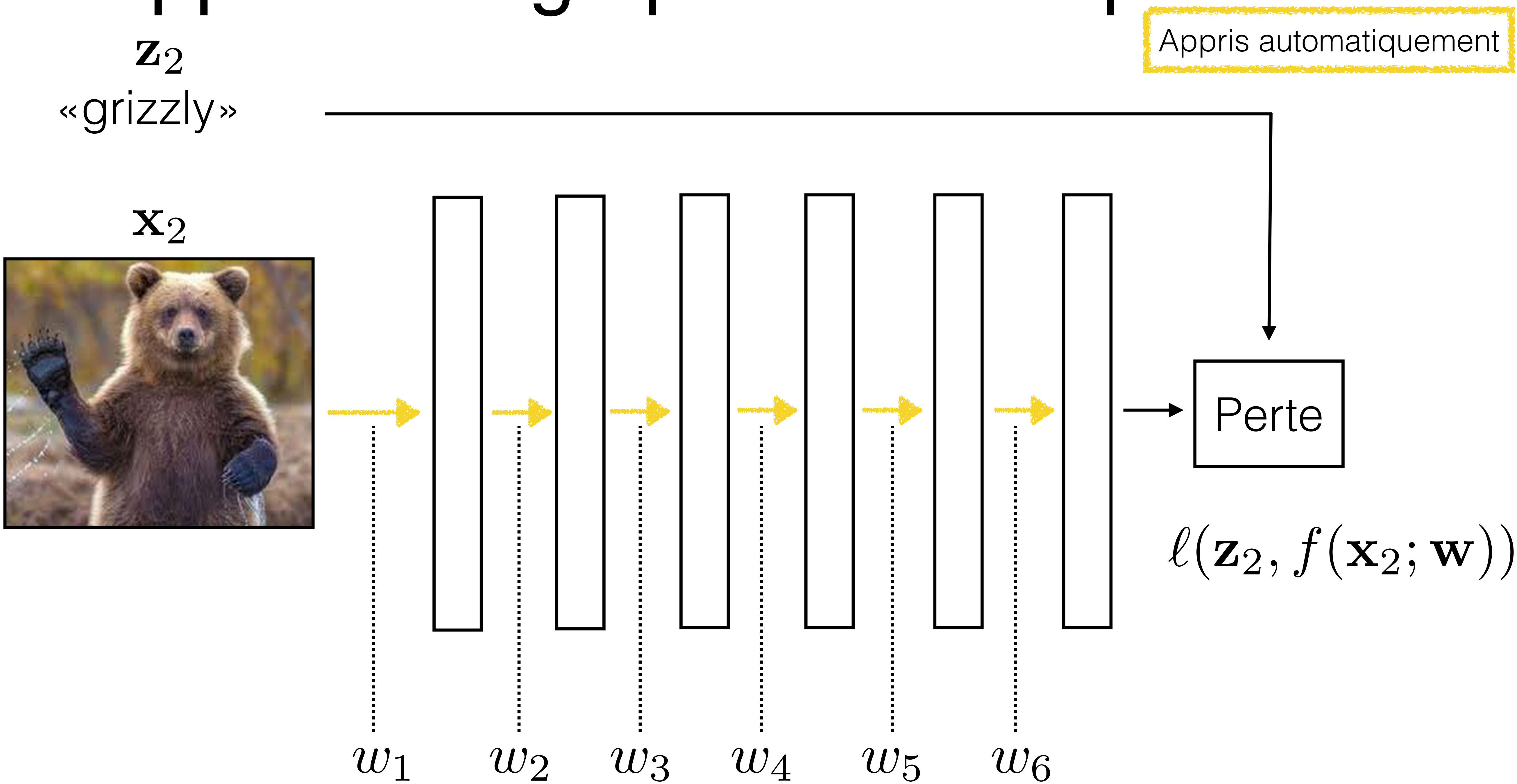
Apprend un modèle de type $p(c|\mathbf{x})$

P. Isola

Apprentissage par réseaux profonds



Apprentissage par réseaux profonds



P. Isola

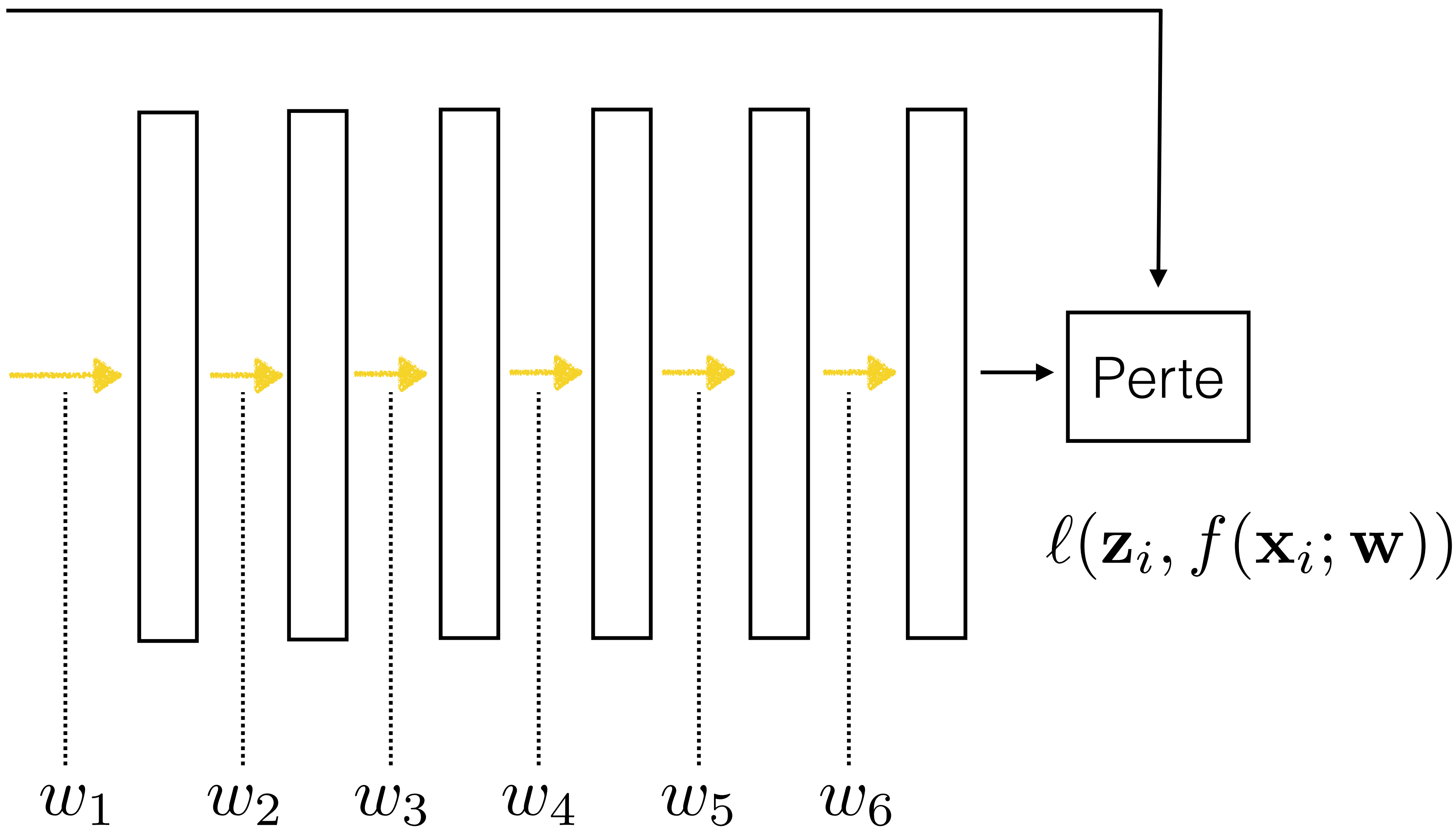
Apprentissage par réseaux profonds

Appris automatiquement

$\mathbf{z}_i$

«caméléon»

$\mathbf{x}_i$



$$\operatorname*{argmin}_{\mathbf{w}} \sum_i \ell(\mathbf{z}_i, f(\mathbf{x}_i; \mathbf{w}))$$

P. Isola

Descente du gradient

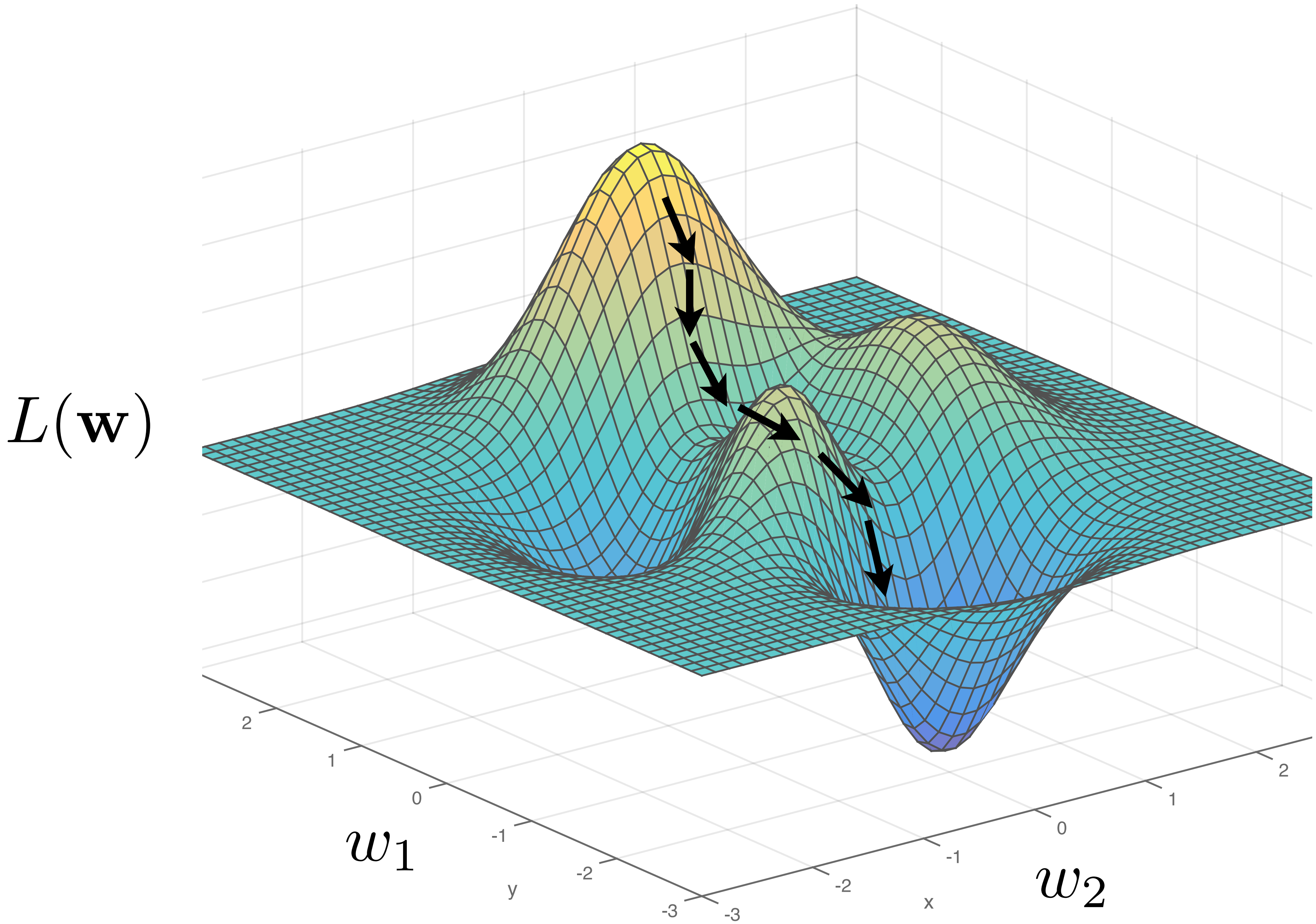
$$\boxed{\underset{\mathbf{w}}{\arg\min} \sum_i \ell(\mathbf{z}_i, f(\mathbf{x}_i; \mathbf{w})) = L(\mathbf{w})}$$

Une itération de descente du gradient

$$\mathbf{w}^{t+1} = \mathbf{w}^t - \eta_t \frac{\partial L(\mathbf{w^t})}{\partial \mathbf{w}}$$

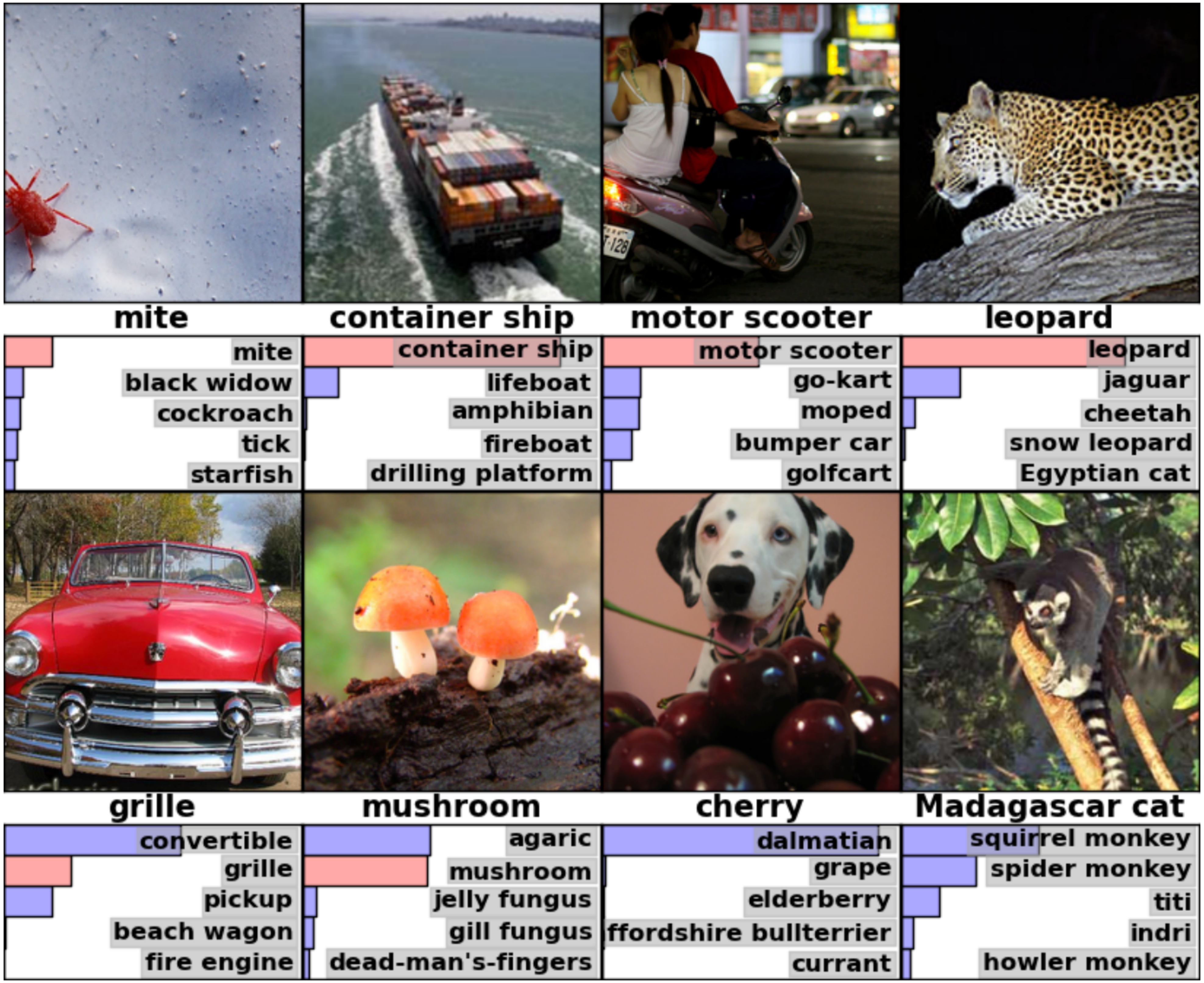
taux d'apprentissage

Descente du gradient



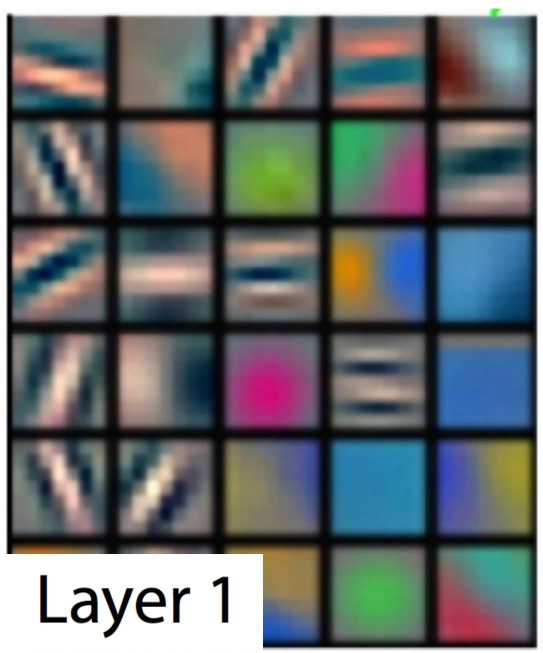
P. Isola

$$p(c|\mathbf{x})$$



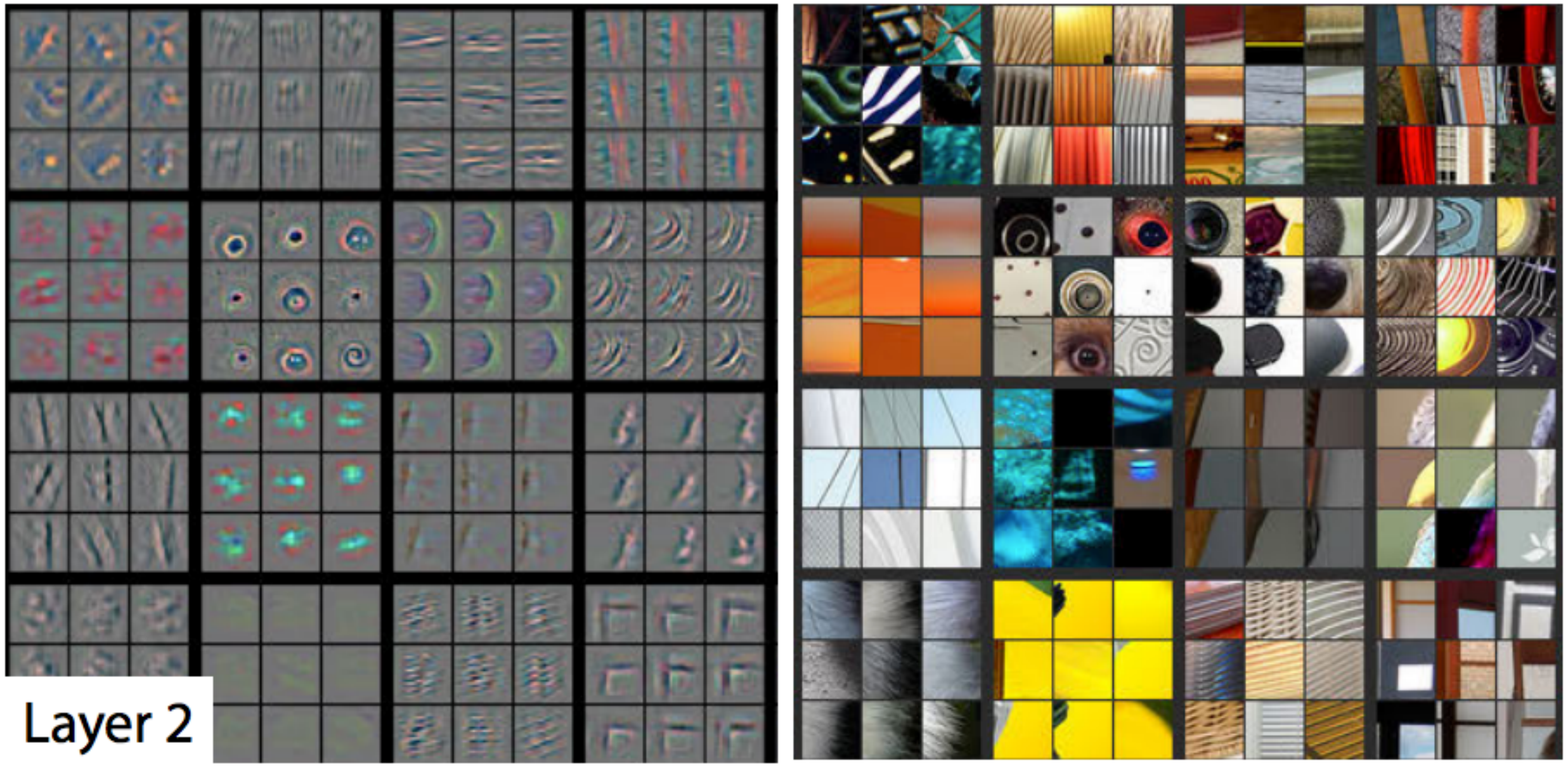
[Krizhevsky et al. NIPS 2012]

Qu'est-ce que le réseau apprend?



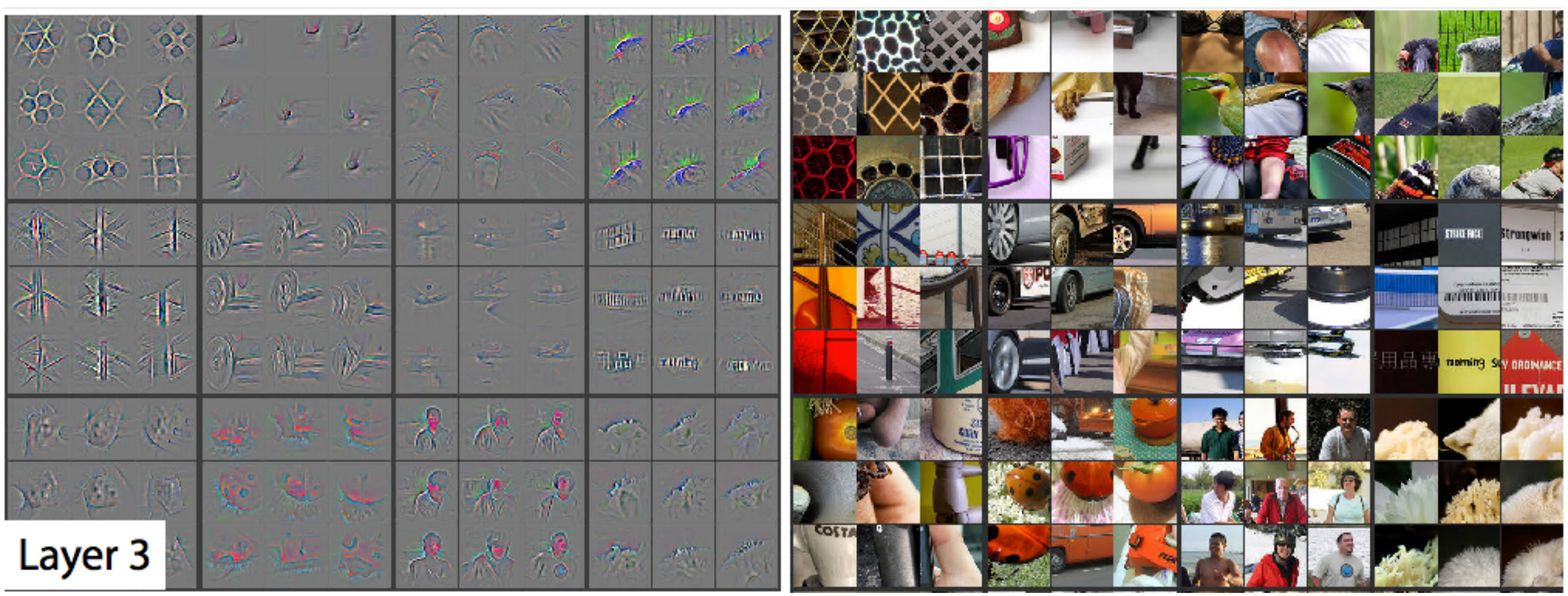
Credit: R. Fergus

Qu’est-ce que le réseau apprend?



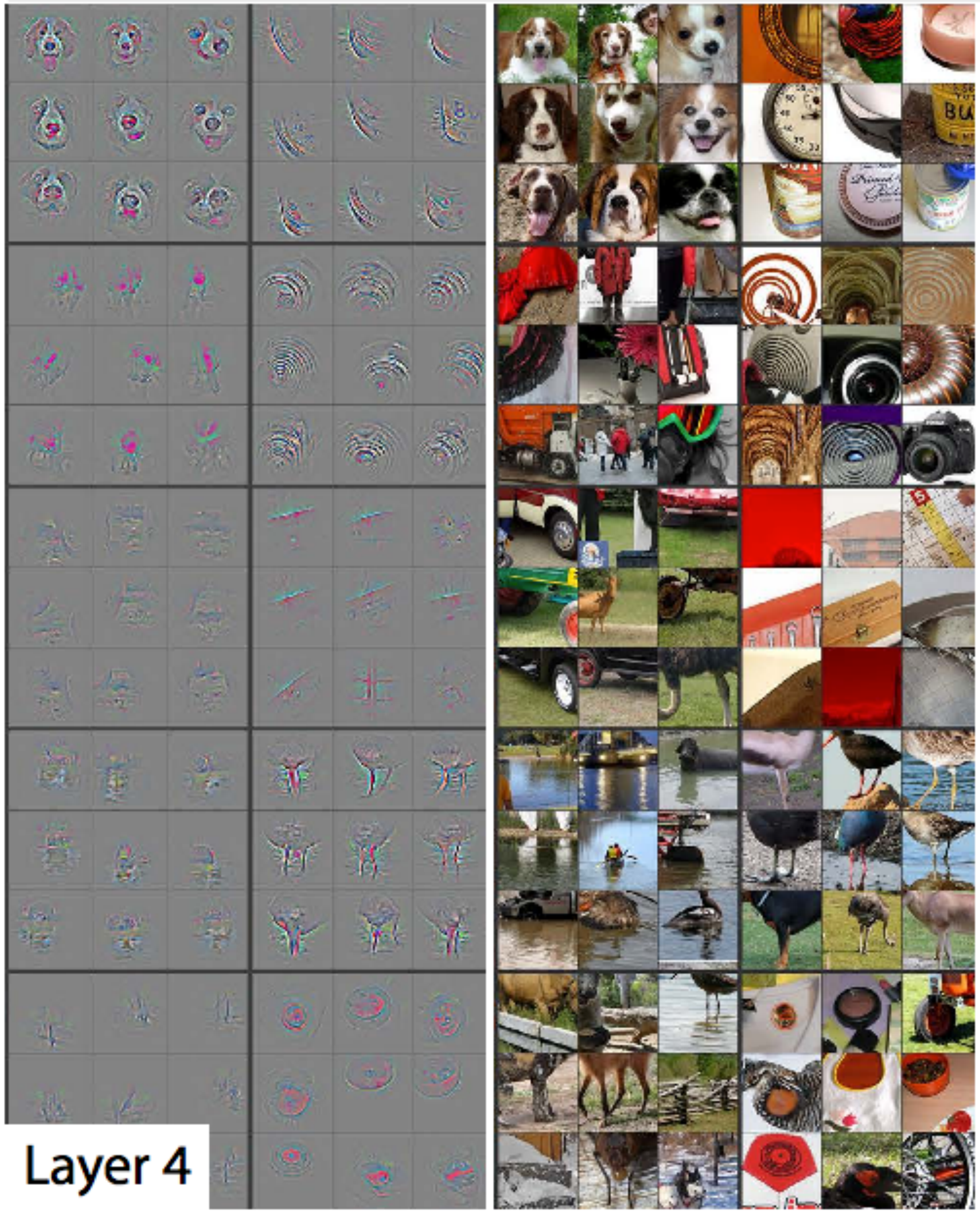
Credit: R. Fergus

Qu'est-ce que le réseau apprend?



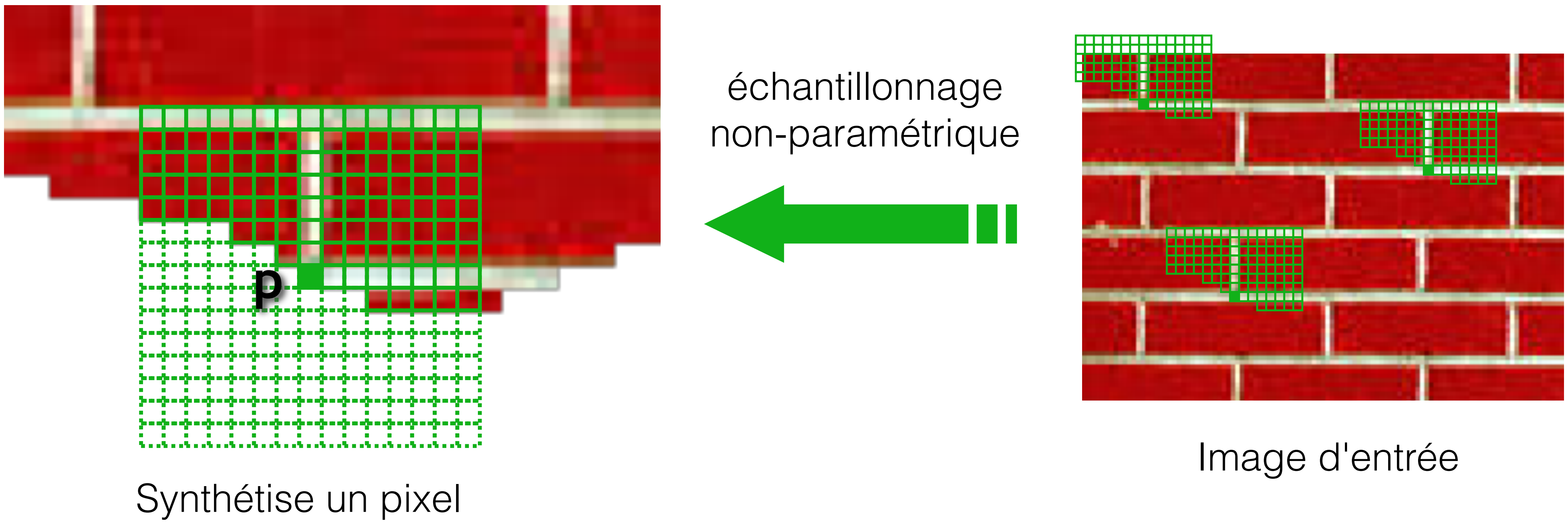
Credit: R. Fergus

Qu’est-ce que le réseau apprend?



Credit: R. Fergus

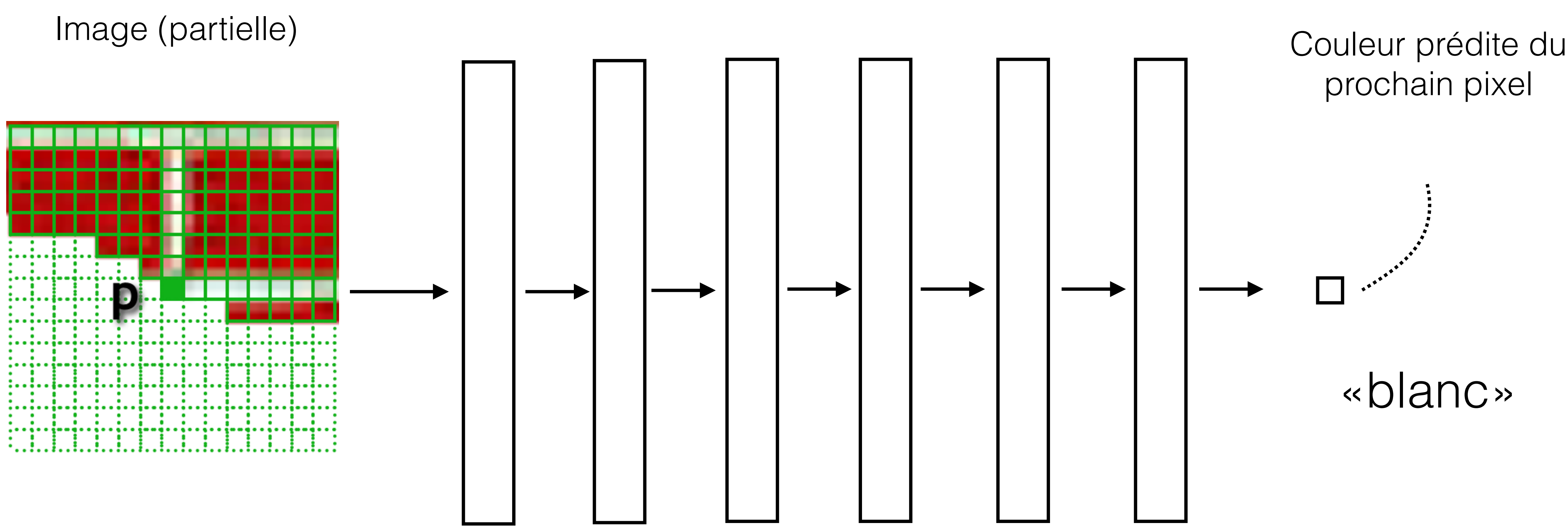
Synthèse de texture: rappel



Modélise $P(p|N(p))$

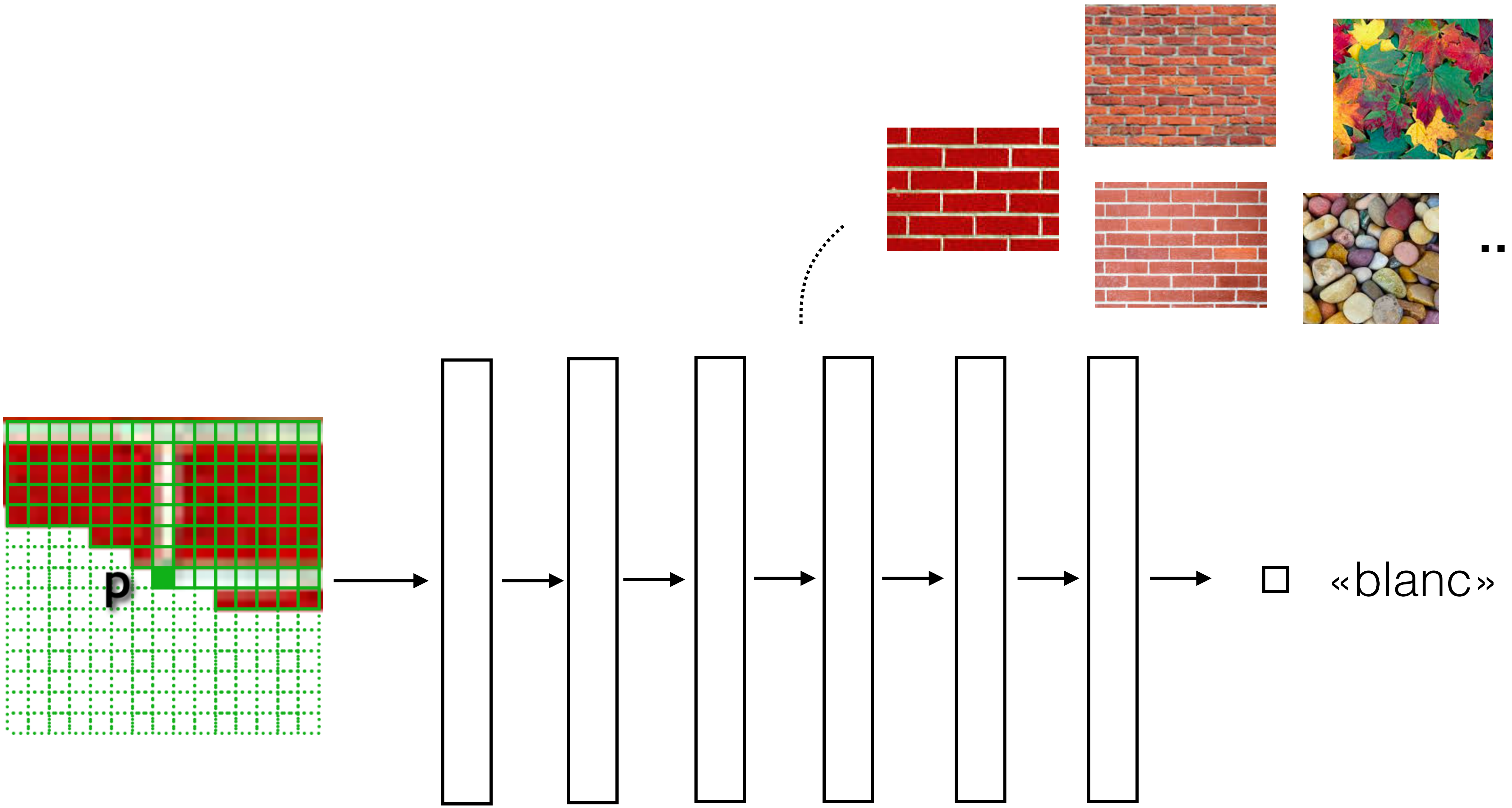
[Efros & Leung 1999]

Synthèse de texture par réseau profond



[van der Oord et al. 2016]

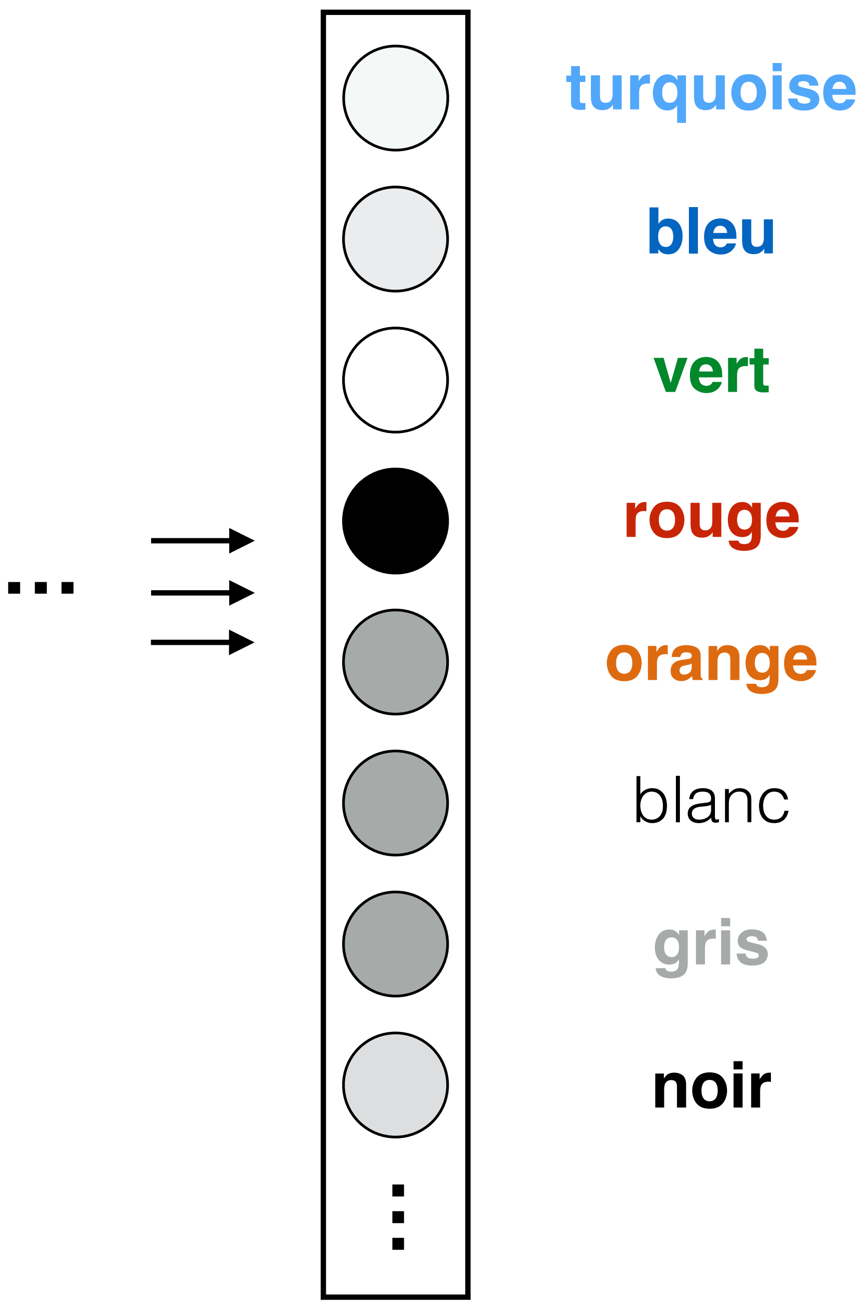
Synthèse de texture par réseau profond



[van der Oord et al. 2016]

Échantillonnage

Sortie du réseau

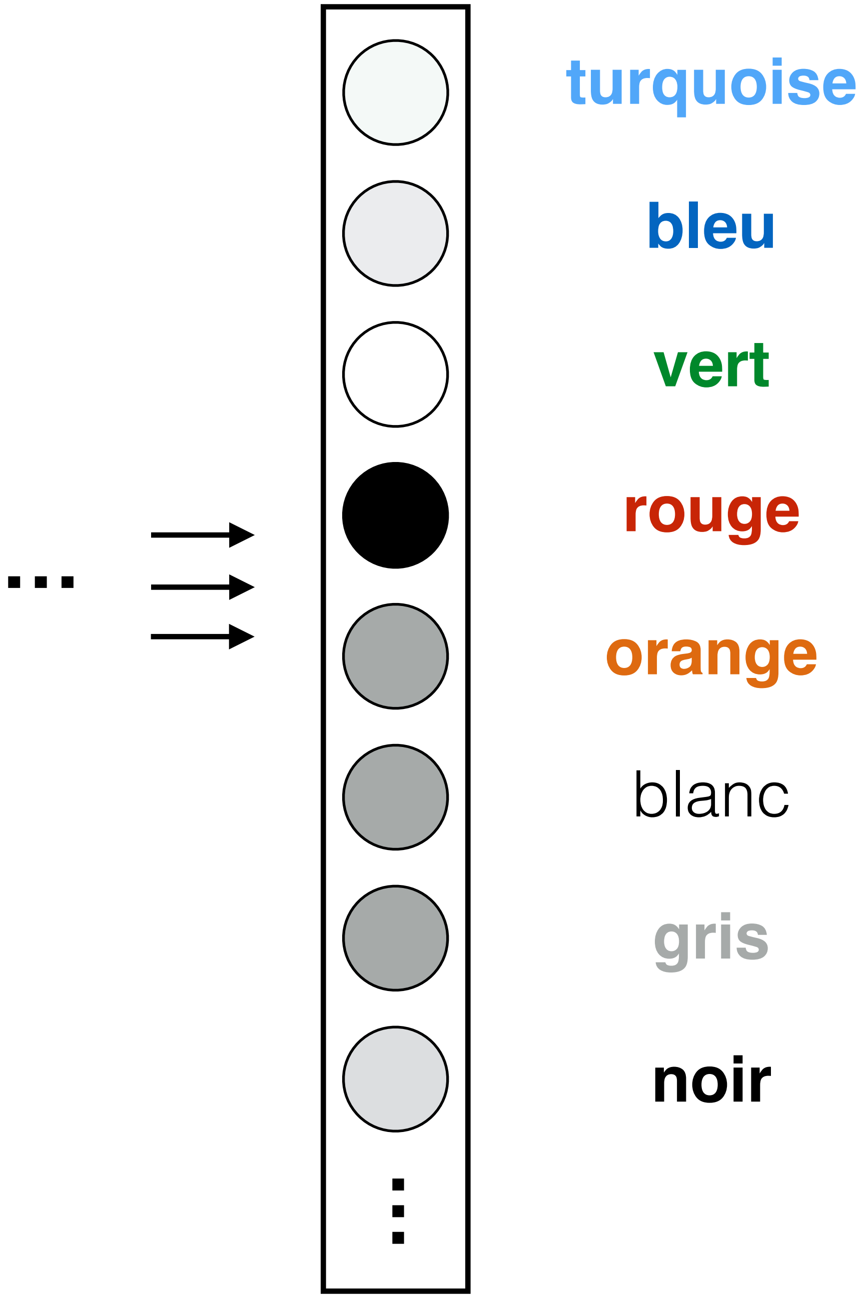


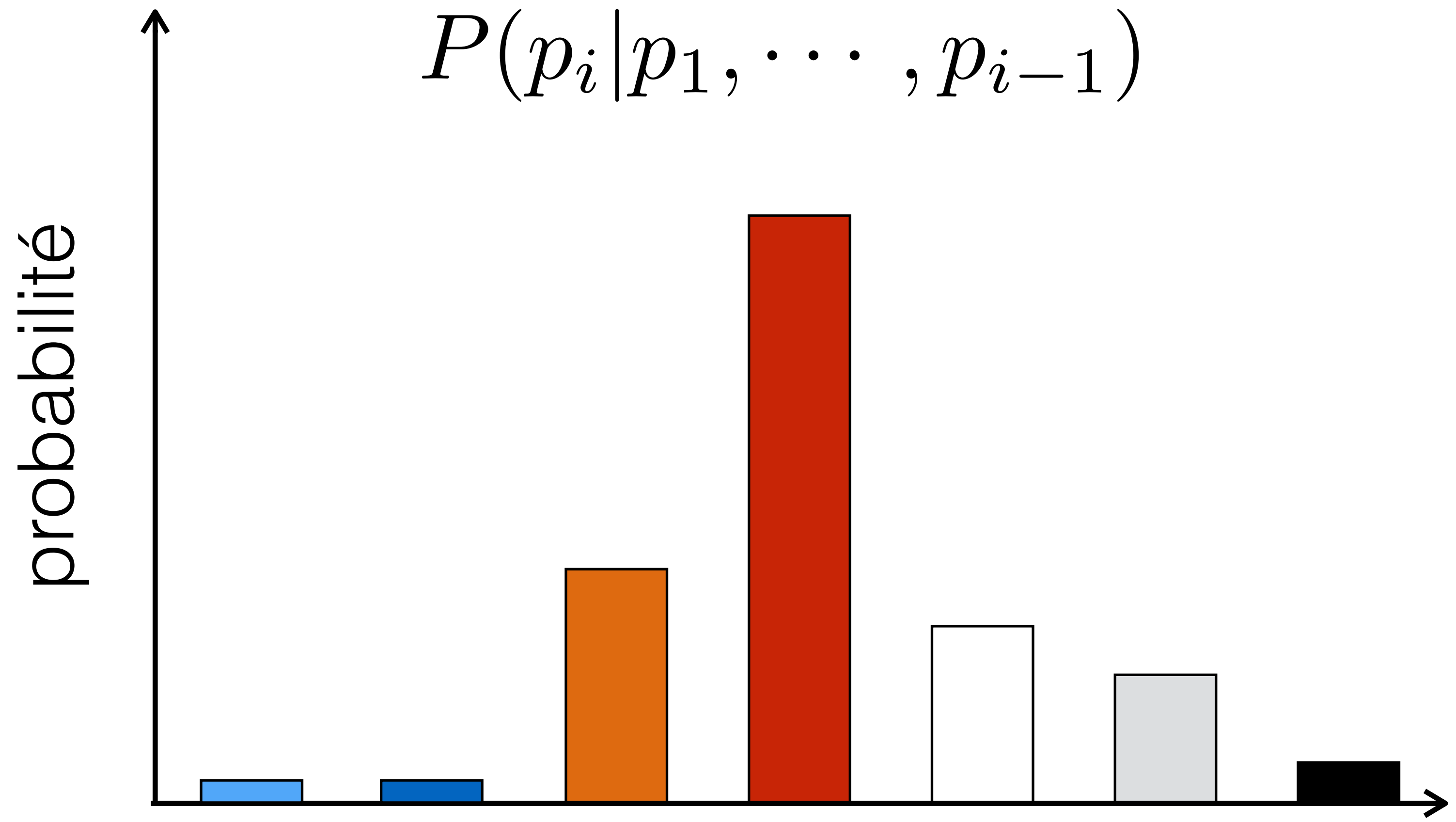
$$P(p_i | p_1, \cdots, p_{i-1})$$

[van der Oord et al. 2016]

Échantillonnage

<u>Sortie du réseau</u>

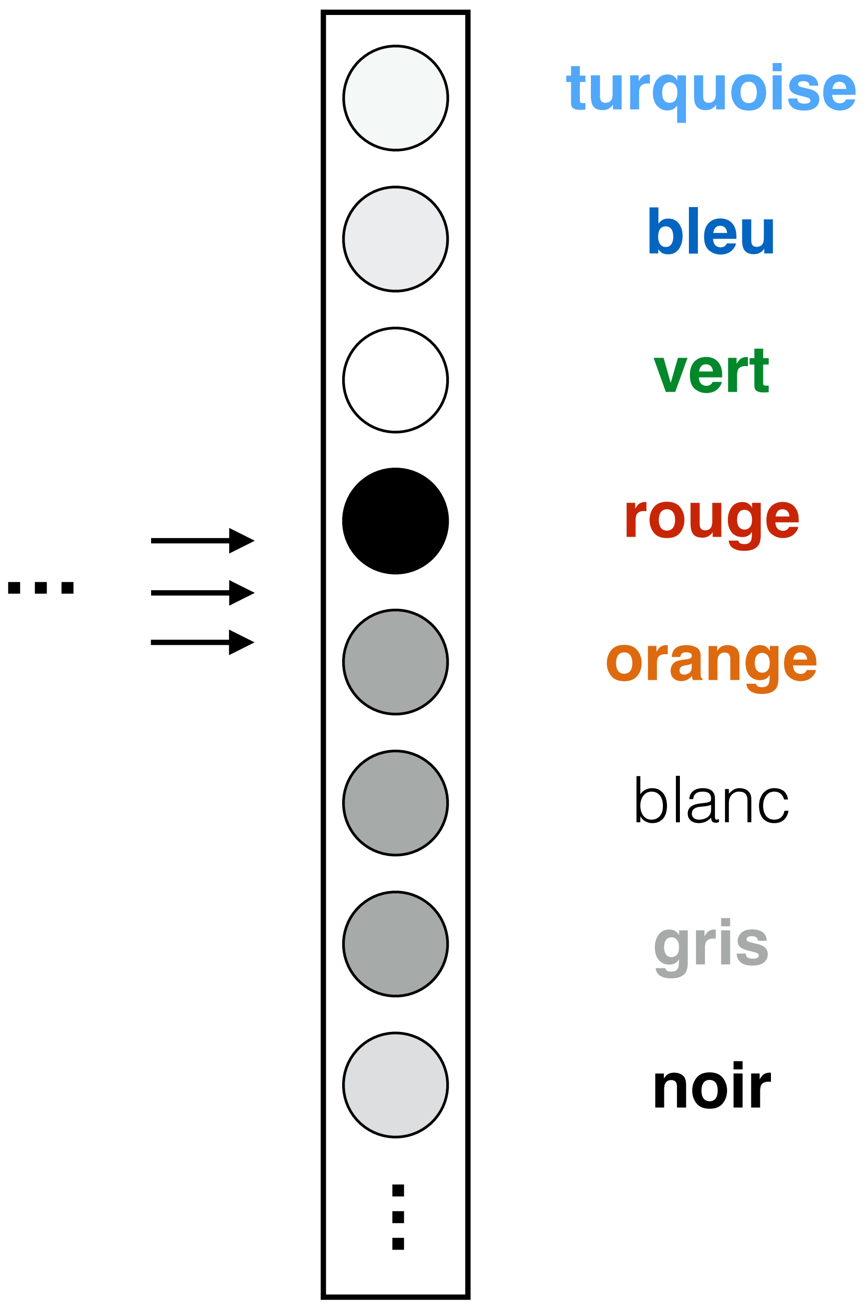


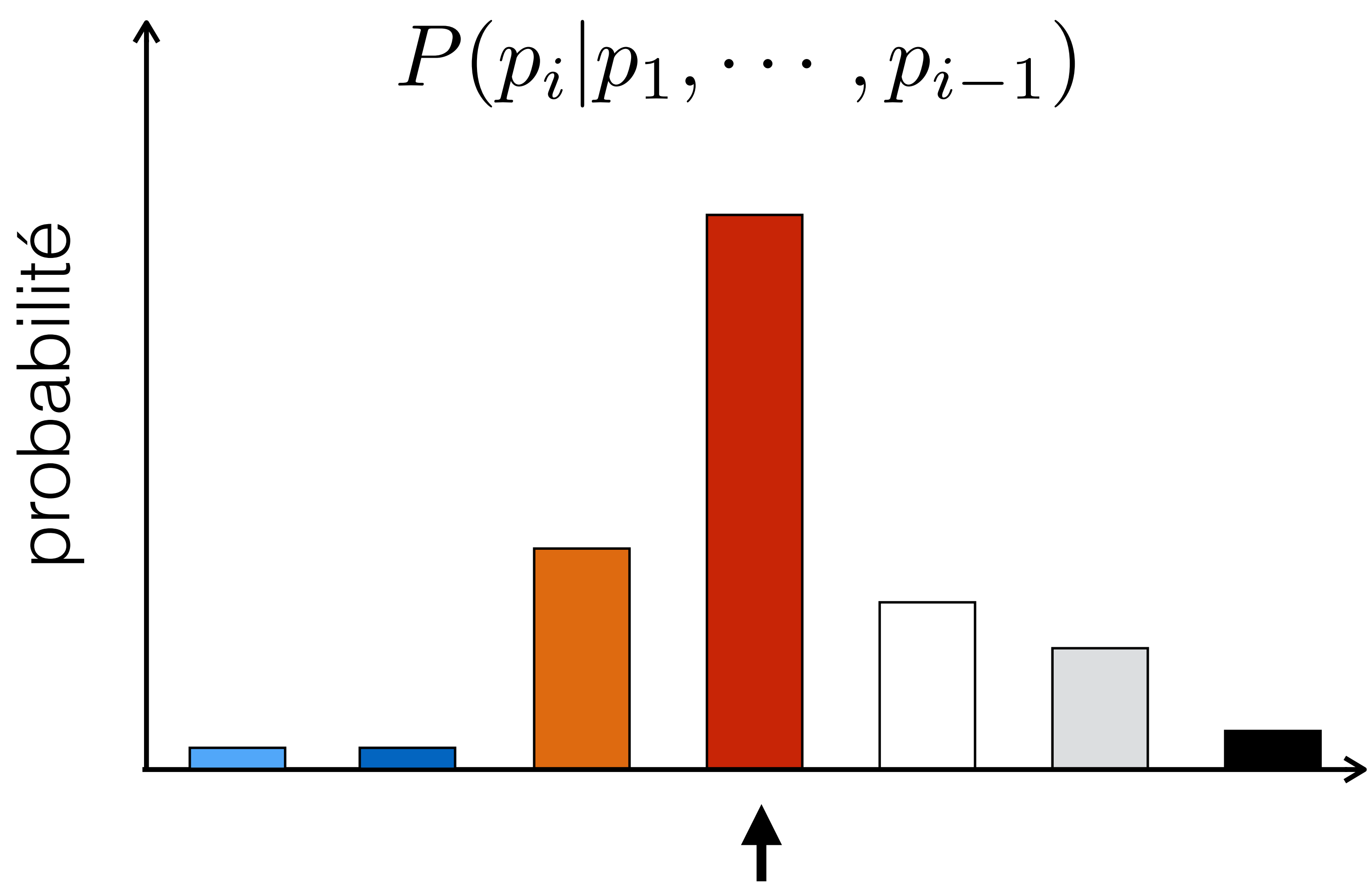


[van der Oord et al. 2016]

Échantillonnage

Sortie du réseau

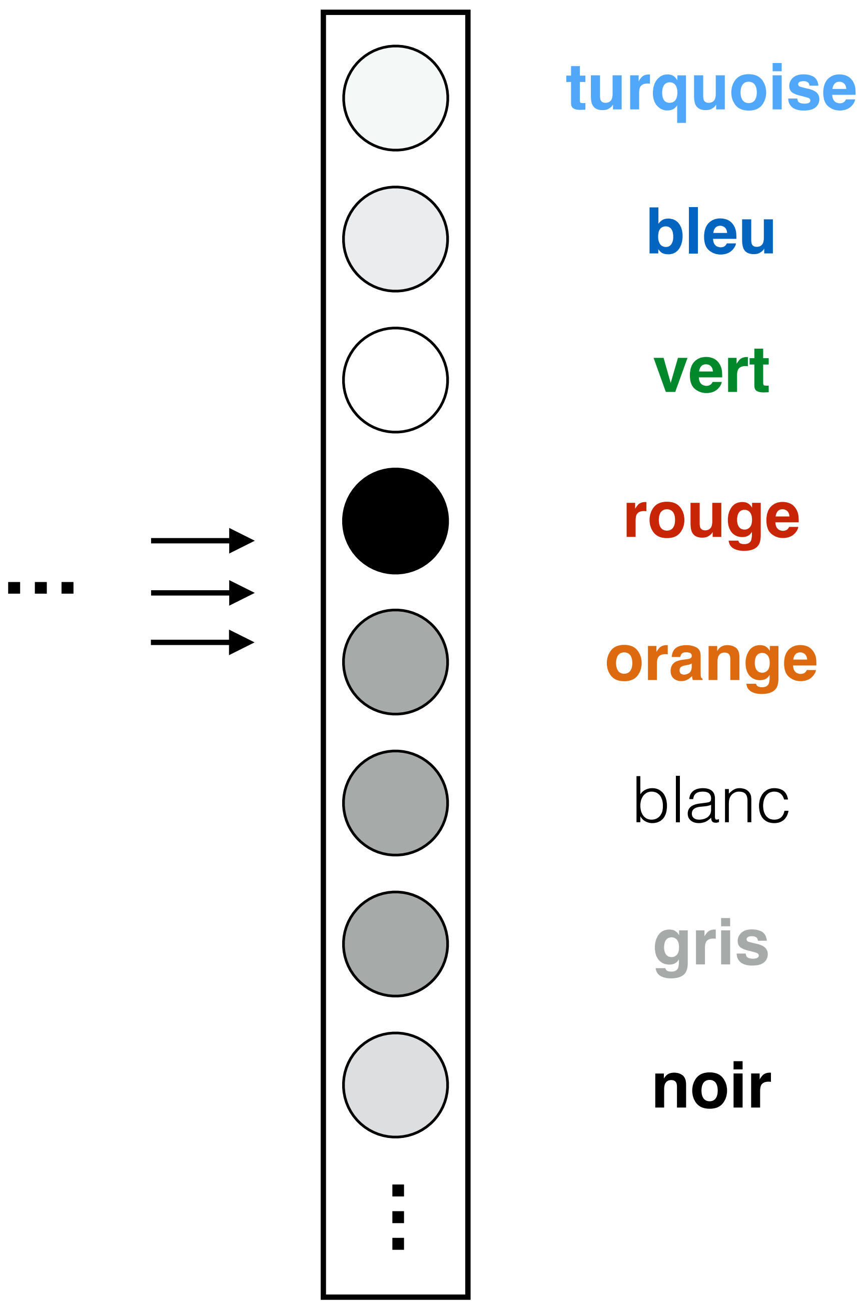


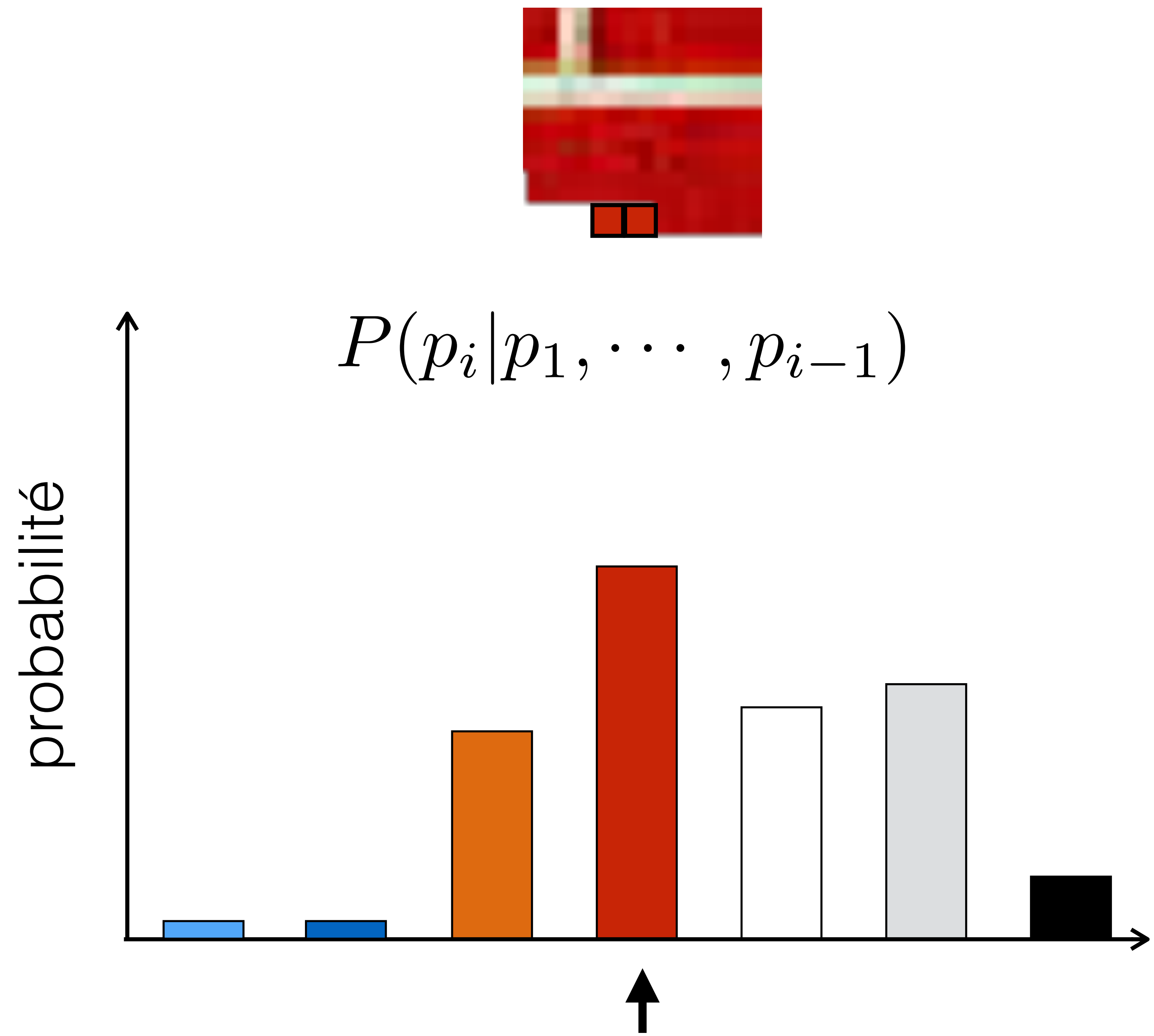


[van der Oord et al. 2016]

Échantillonnage

Sortie du réseau

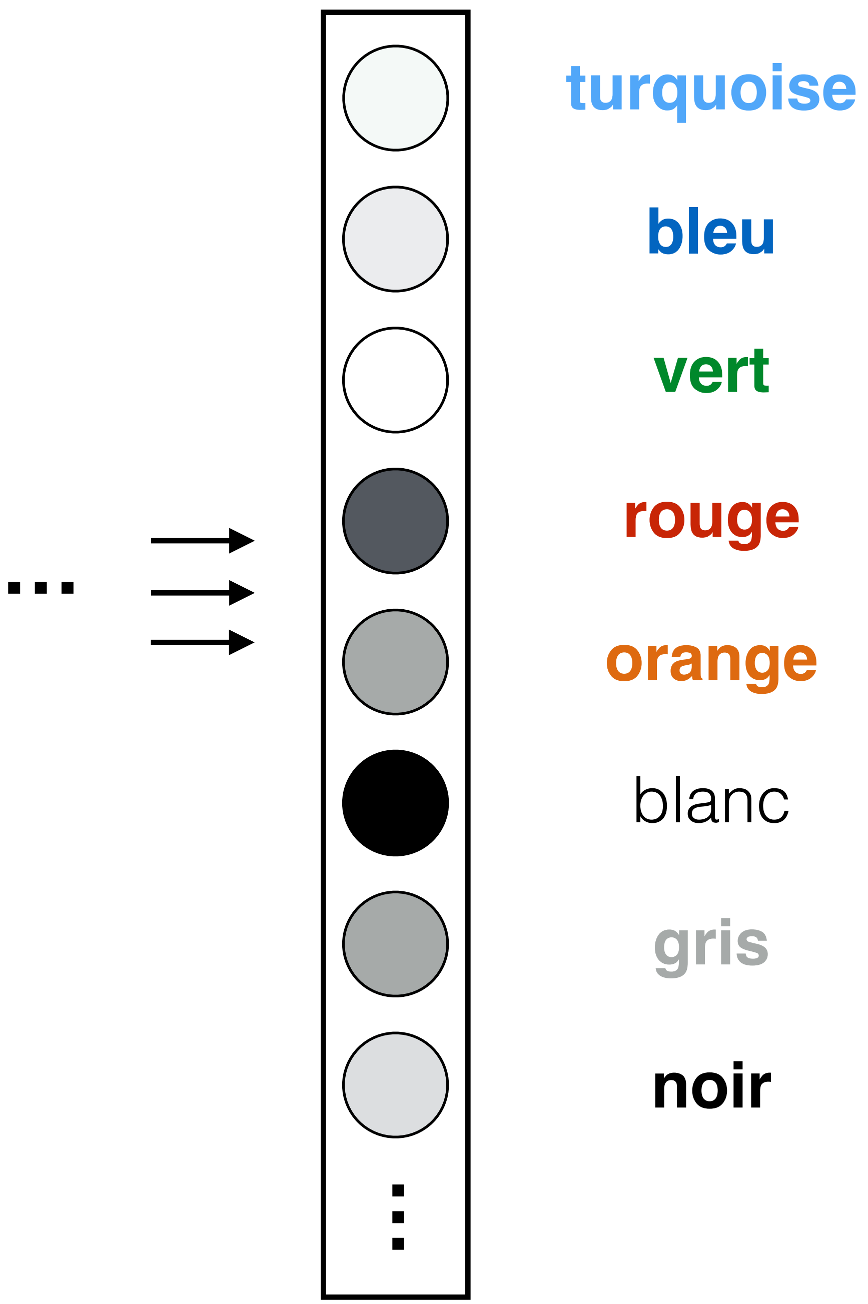


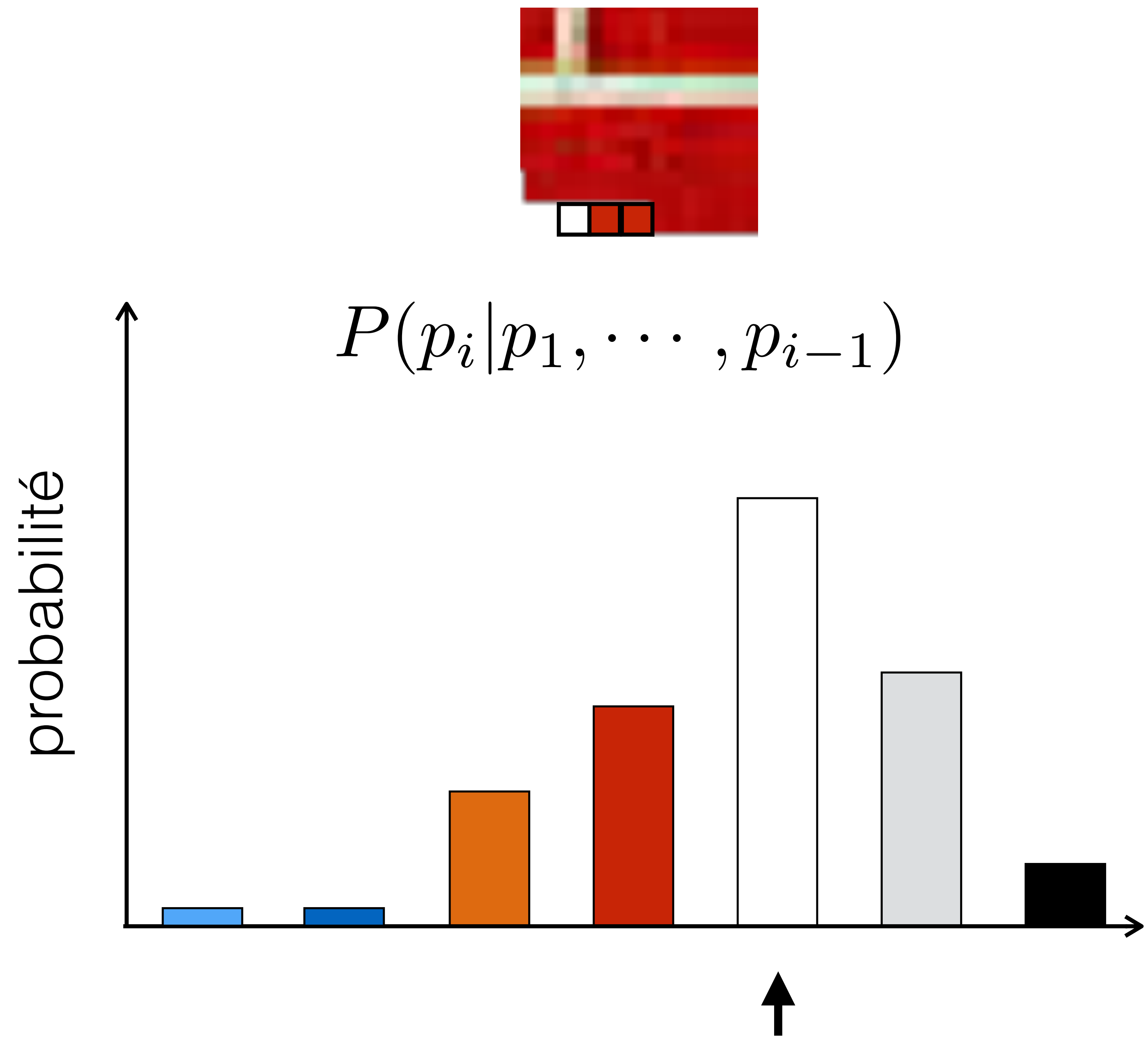


[van der Oord et al. 2016]

Échantillonnage

Sortie du réseau



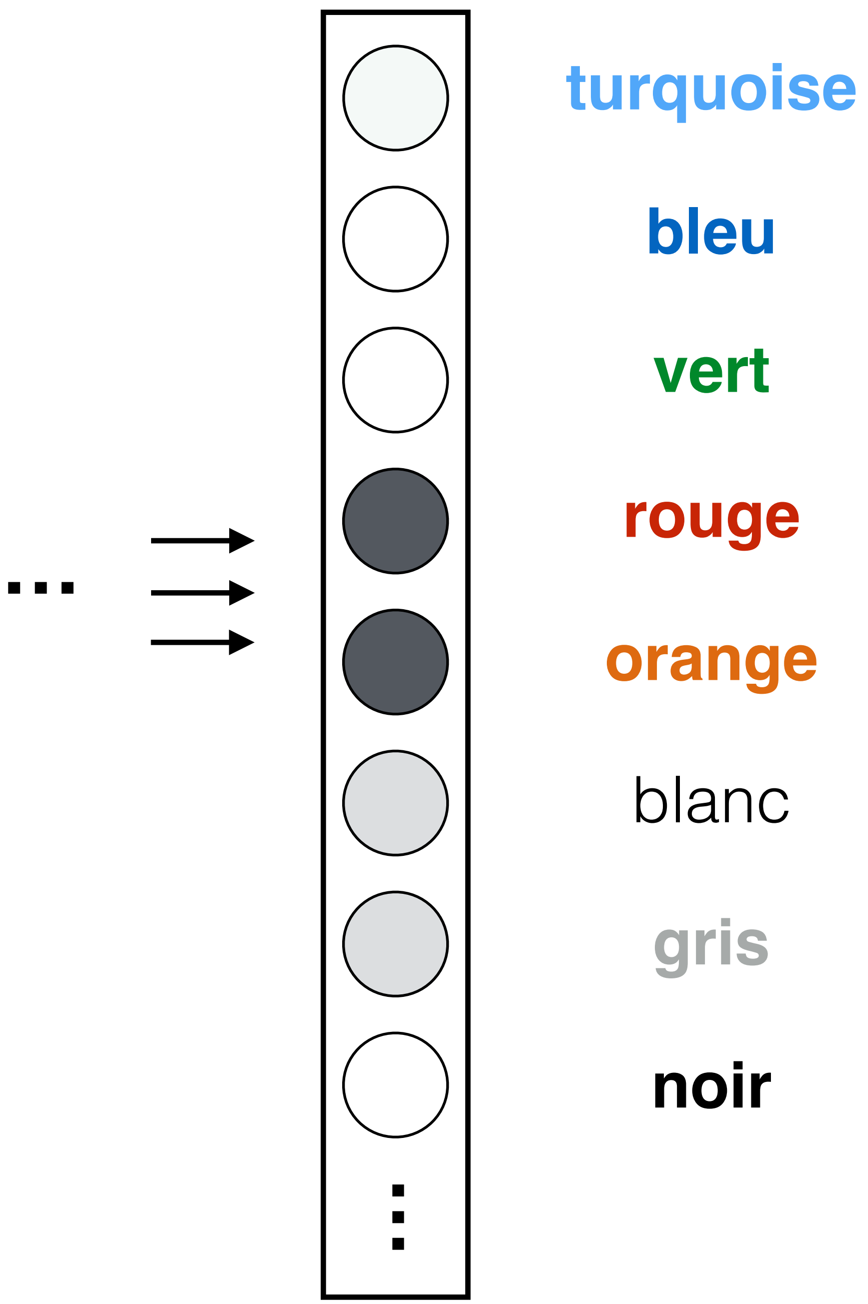


$$P(p_i | p_1, \cdots, p_{i-1})$$

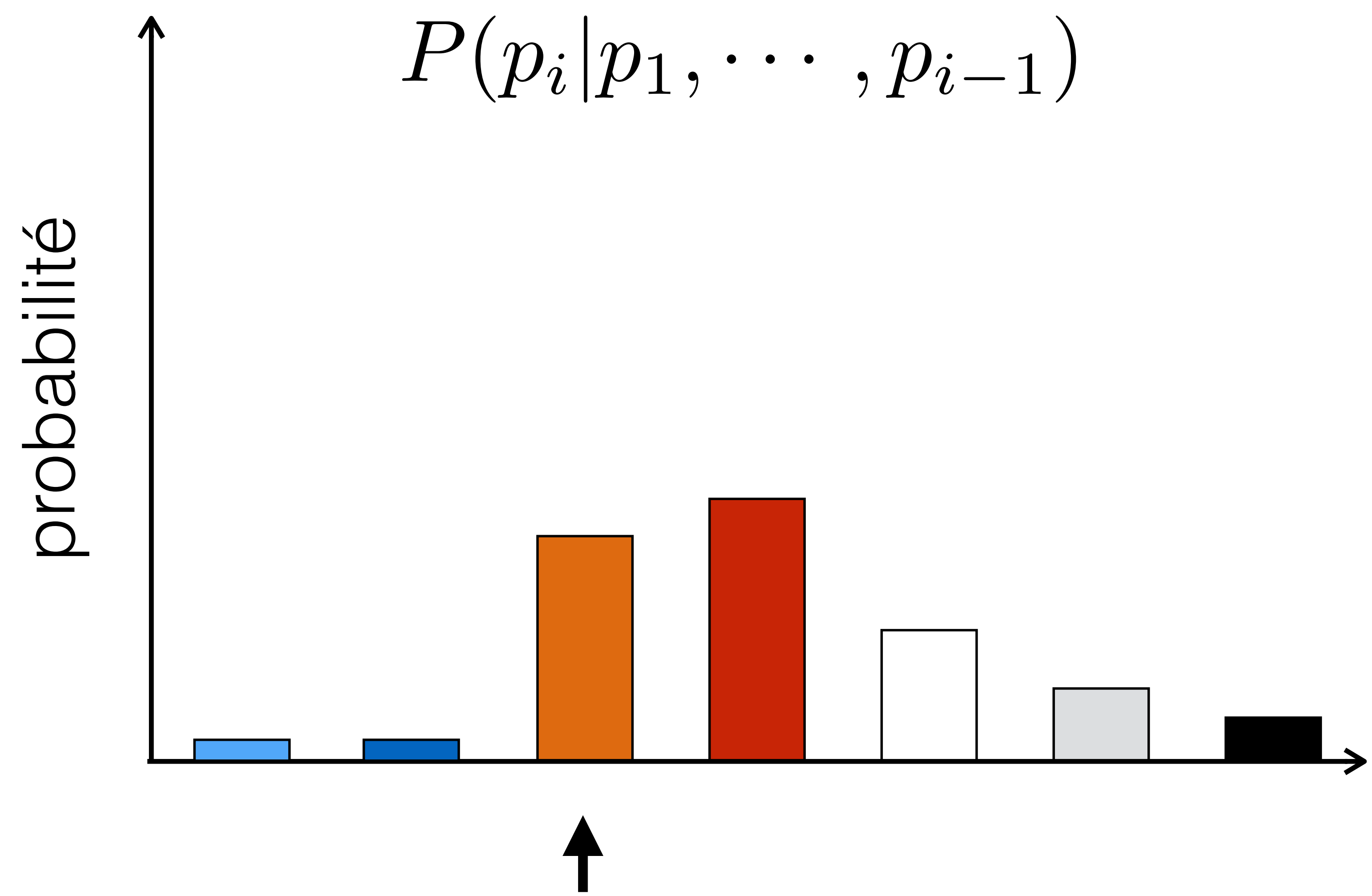
[van der Oord et al. 2016]

Échantillonnage

Sortie du réseau

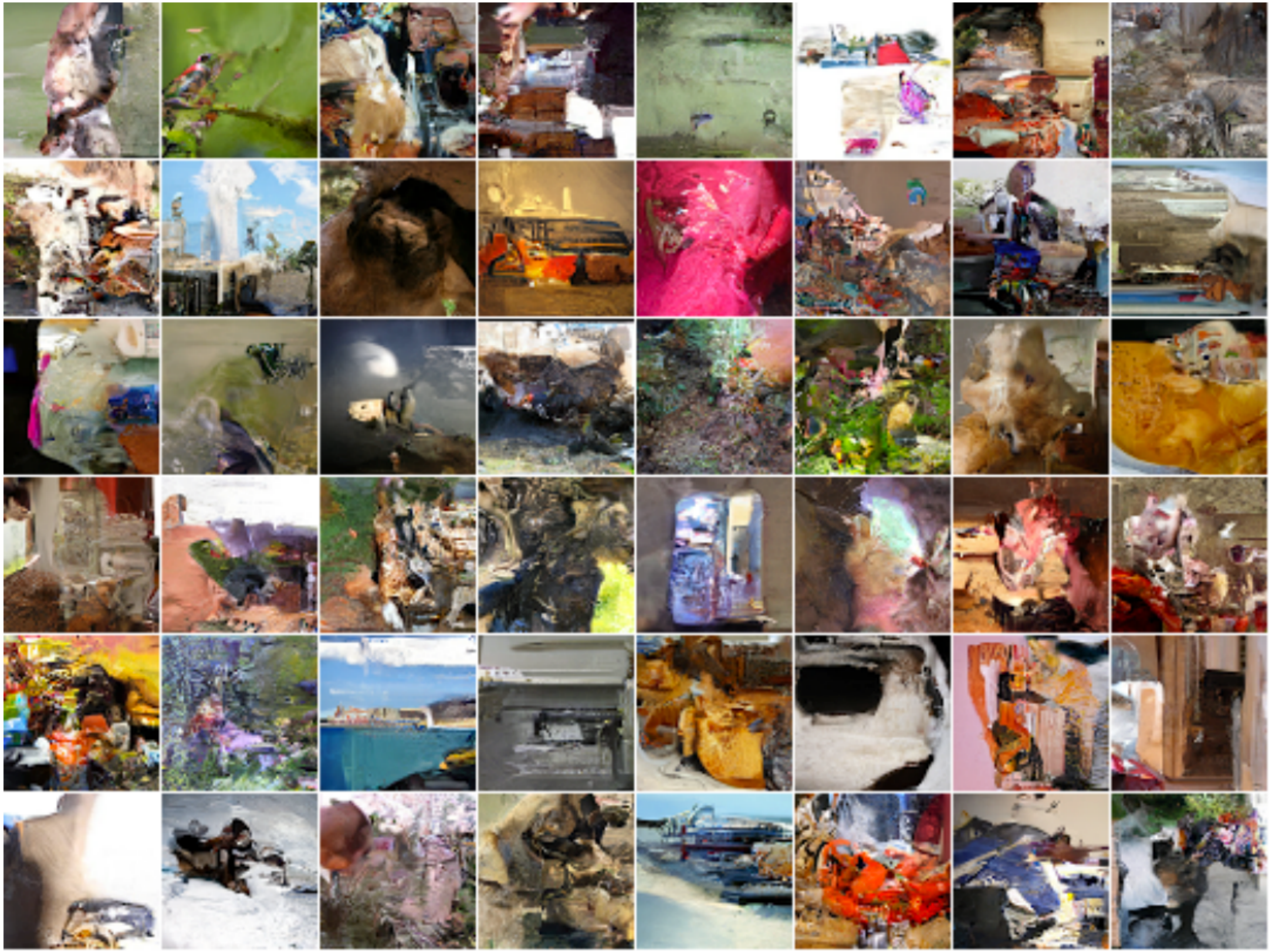


$$P(p_i | p_1, \cdots, p_{i-1})$$



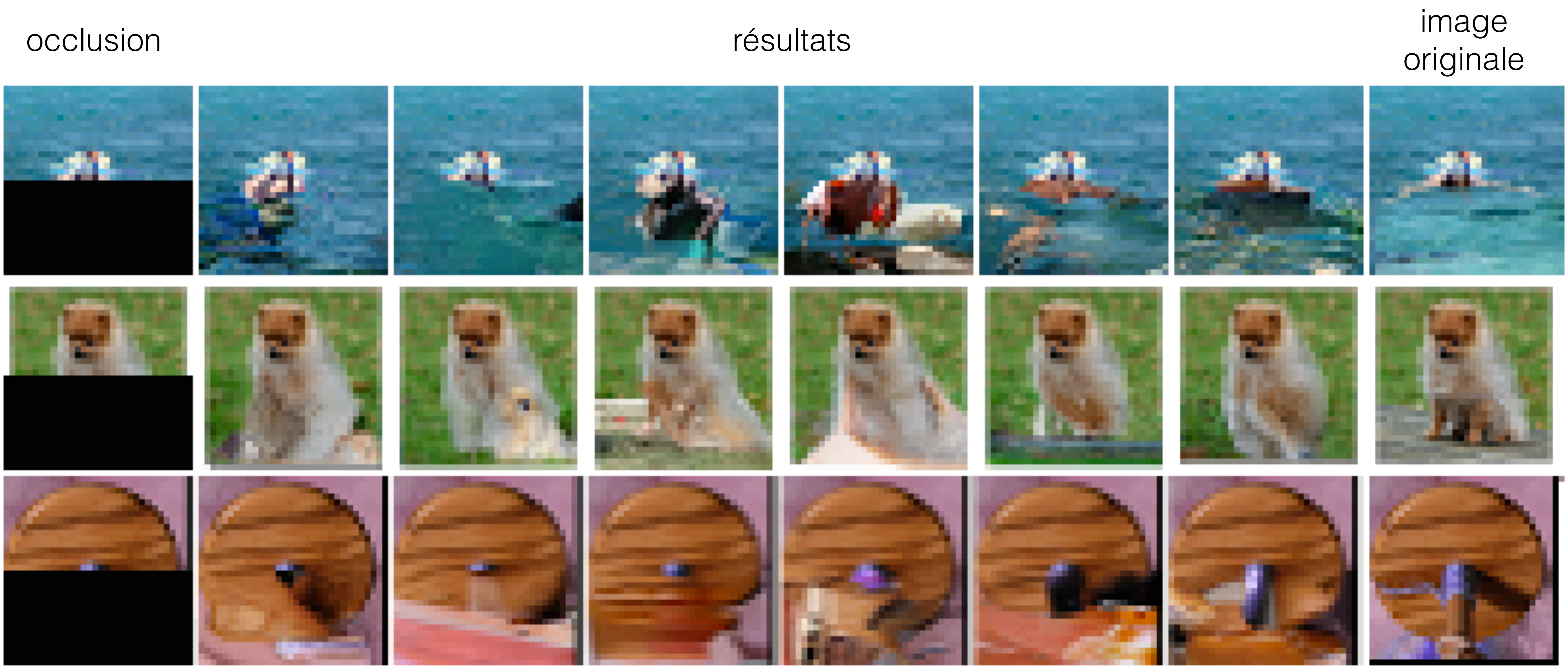
[van der Oord et al. 2016]

Générer des images entières



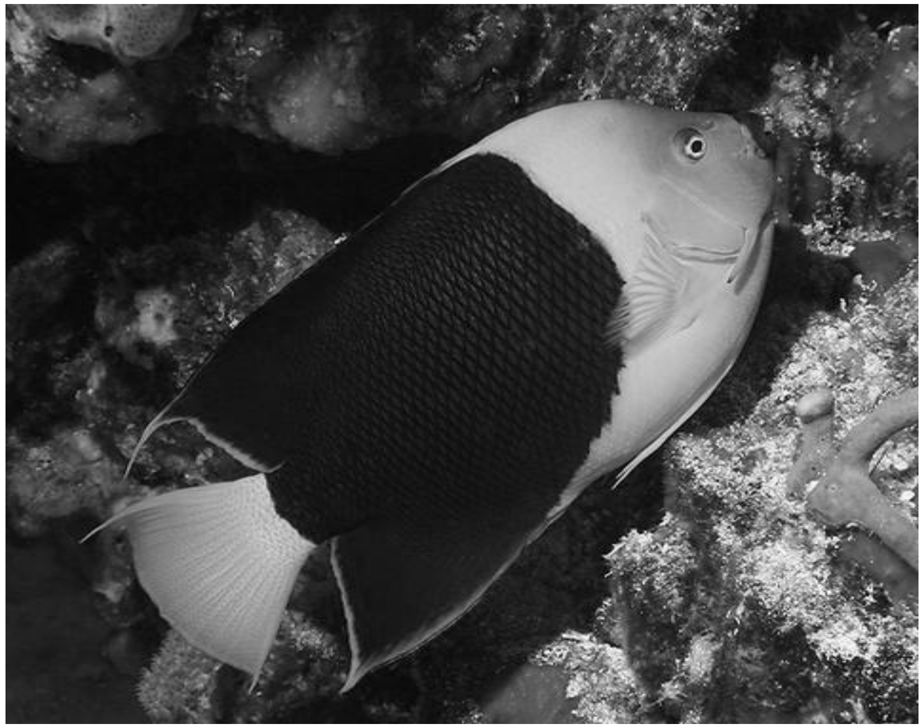
[van der Oord et al. 2016]

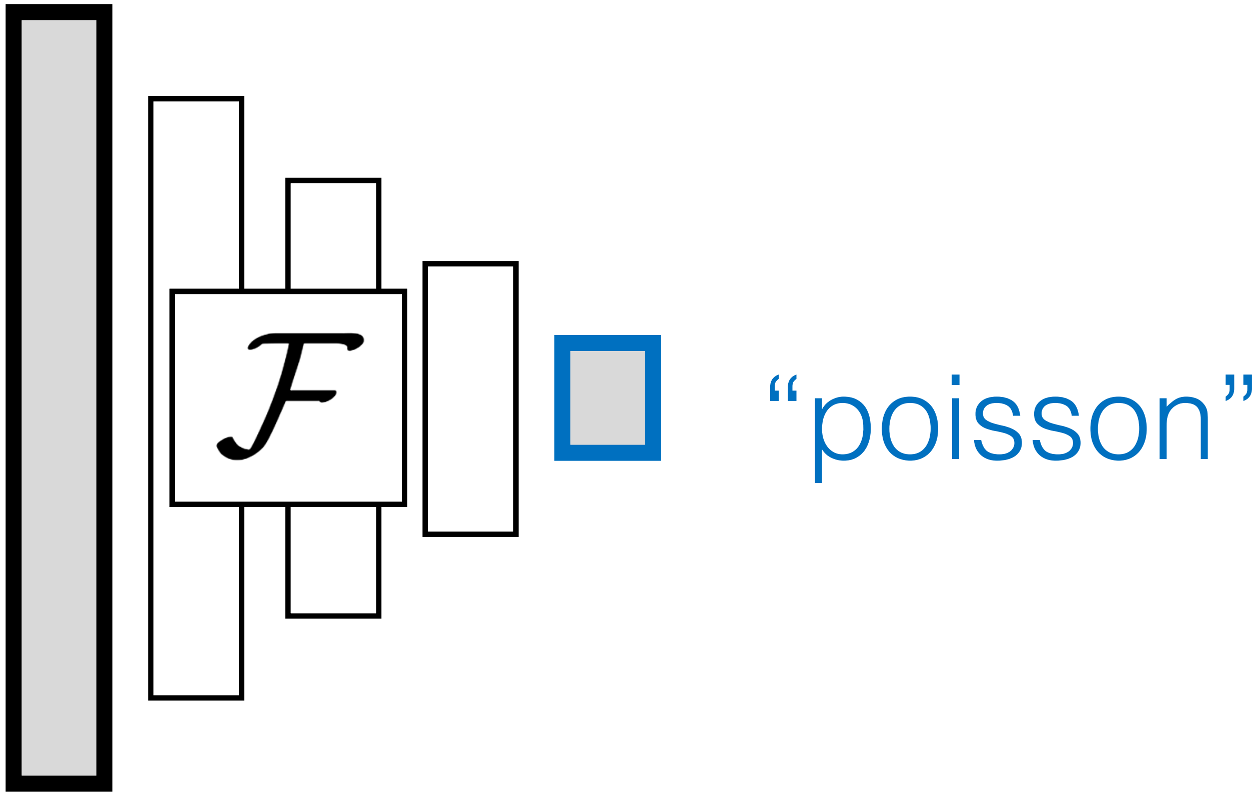
Compléter une image



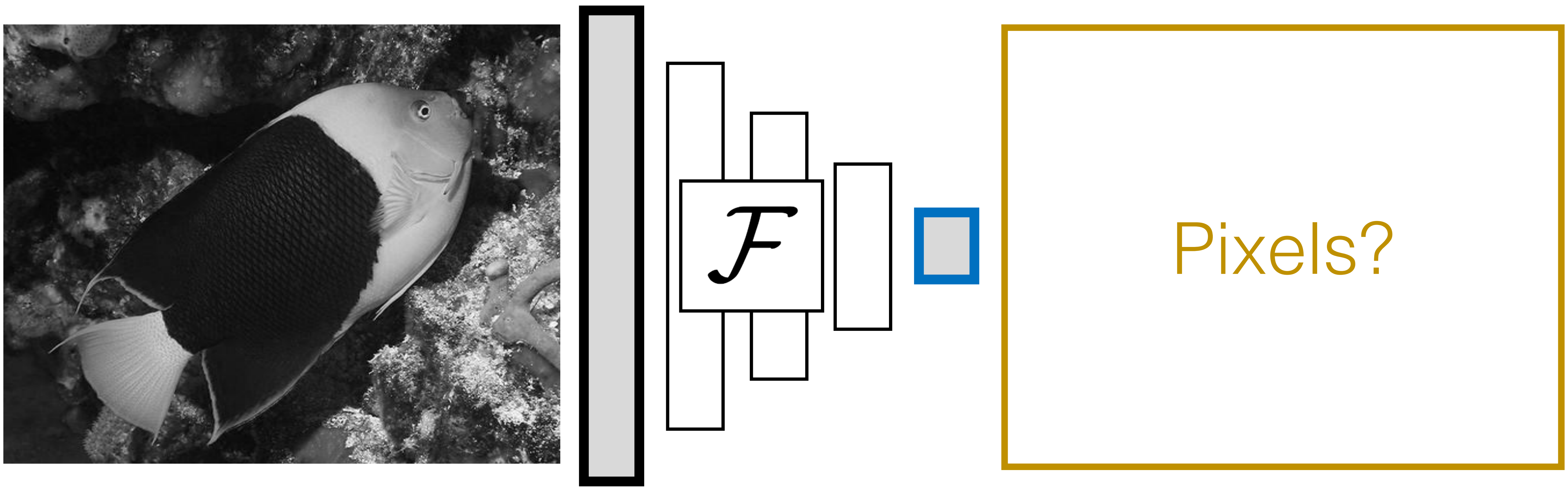
[van der Oord et al. 2016]

Réseaux profonds discriminatifs

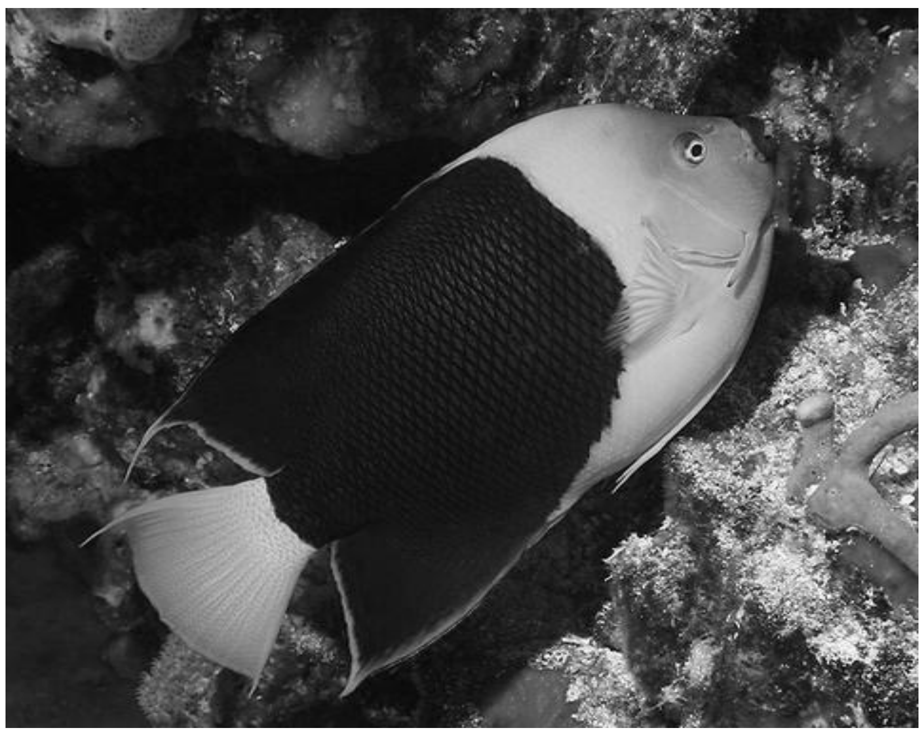


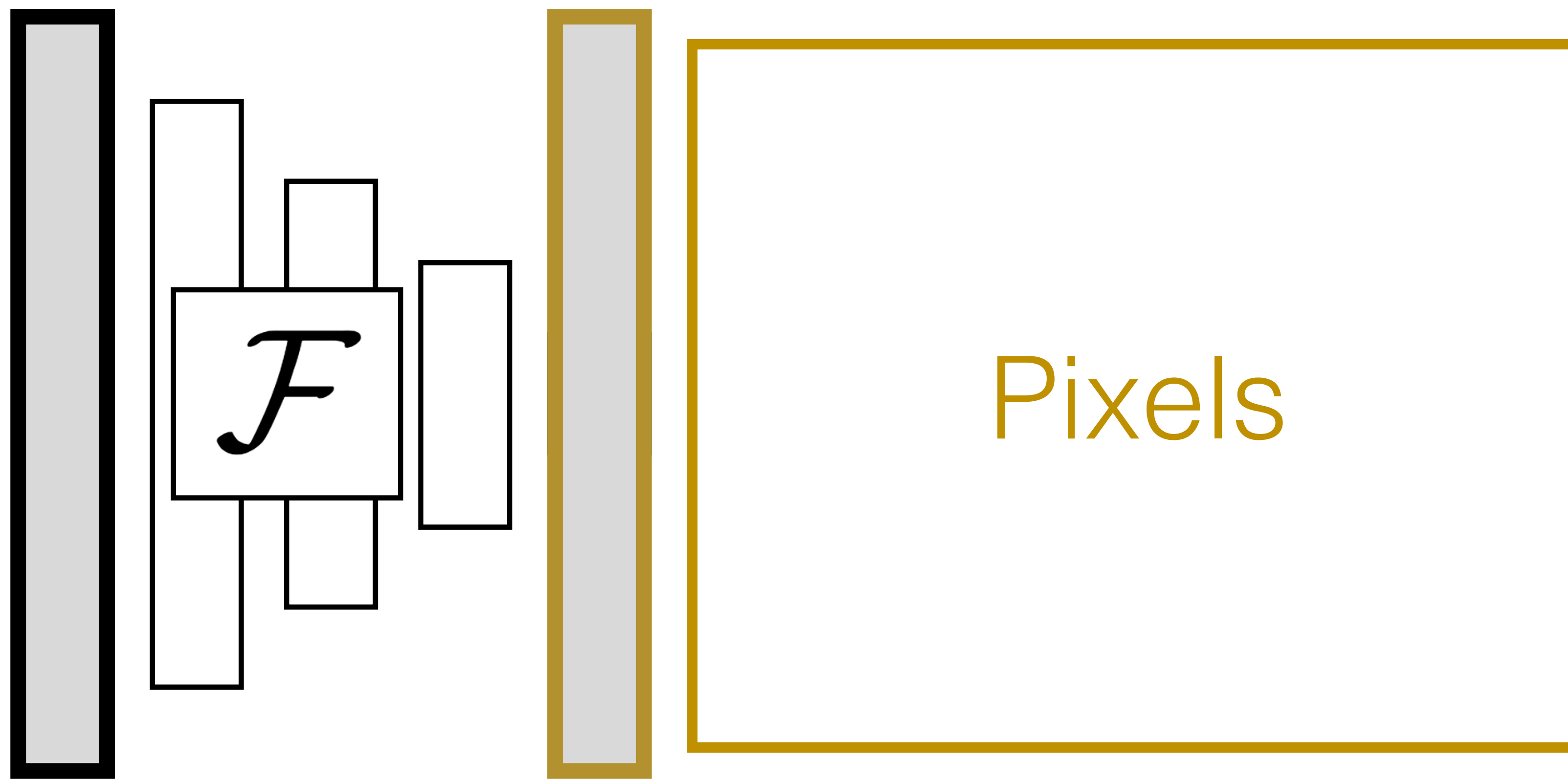


Réseaux profonds discriminatifs



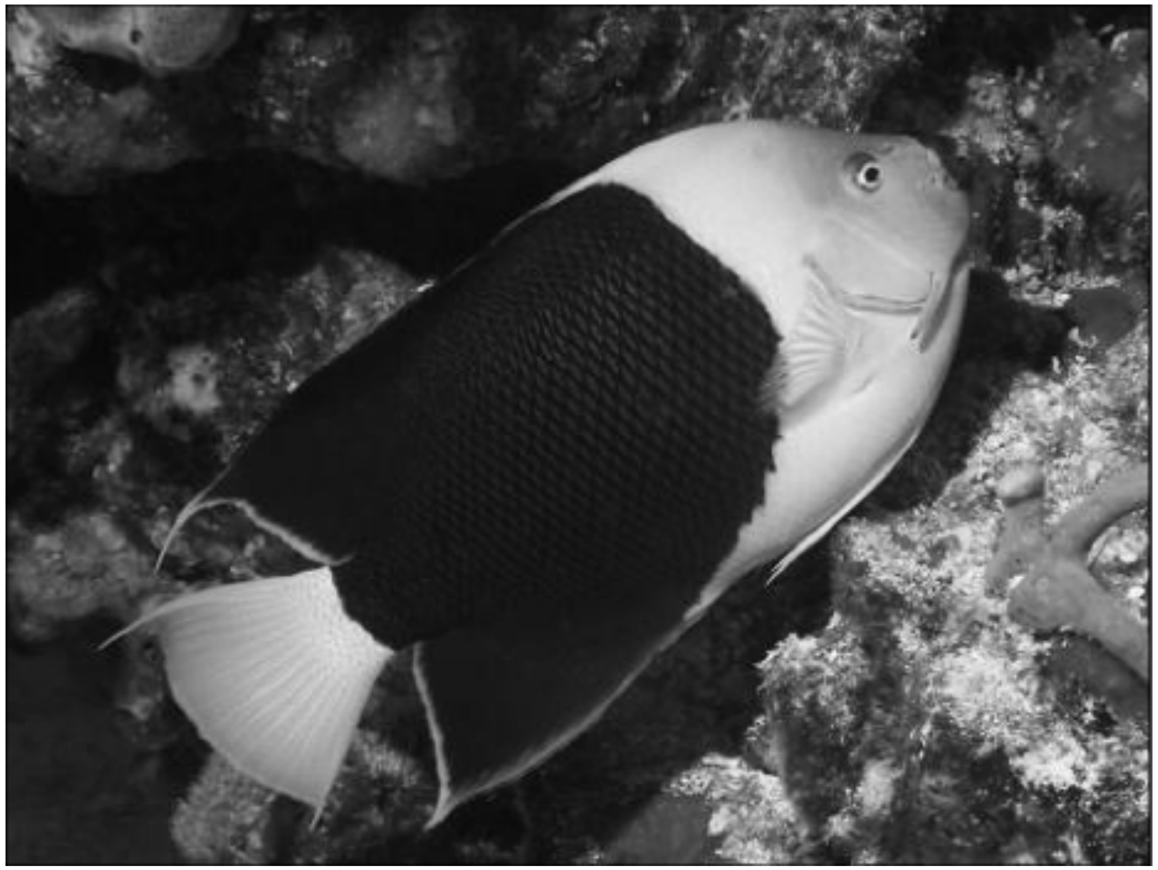
Réseaux profonds *génératifs*





Ansel Adams, Yosemite Valley Bridge

[Zhang et al. 2016]



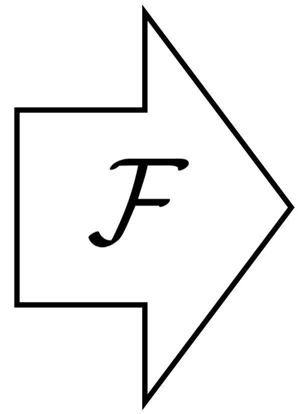
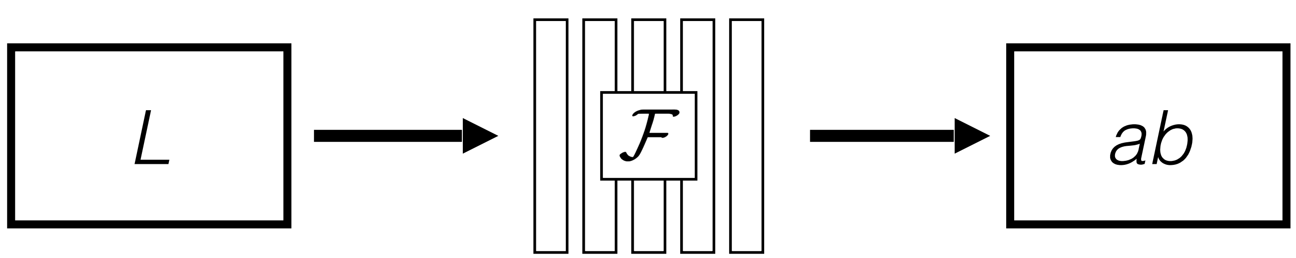


Image d’entrée: canal L seulement

Information en couleurs: canaux ab (de Lab)

$$\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$$

$$\widehat{\mathbf{Y}} \in \mathbb{R}^{H \times W \times 2}$$



P. Isola

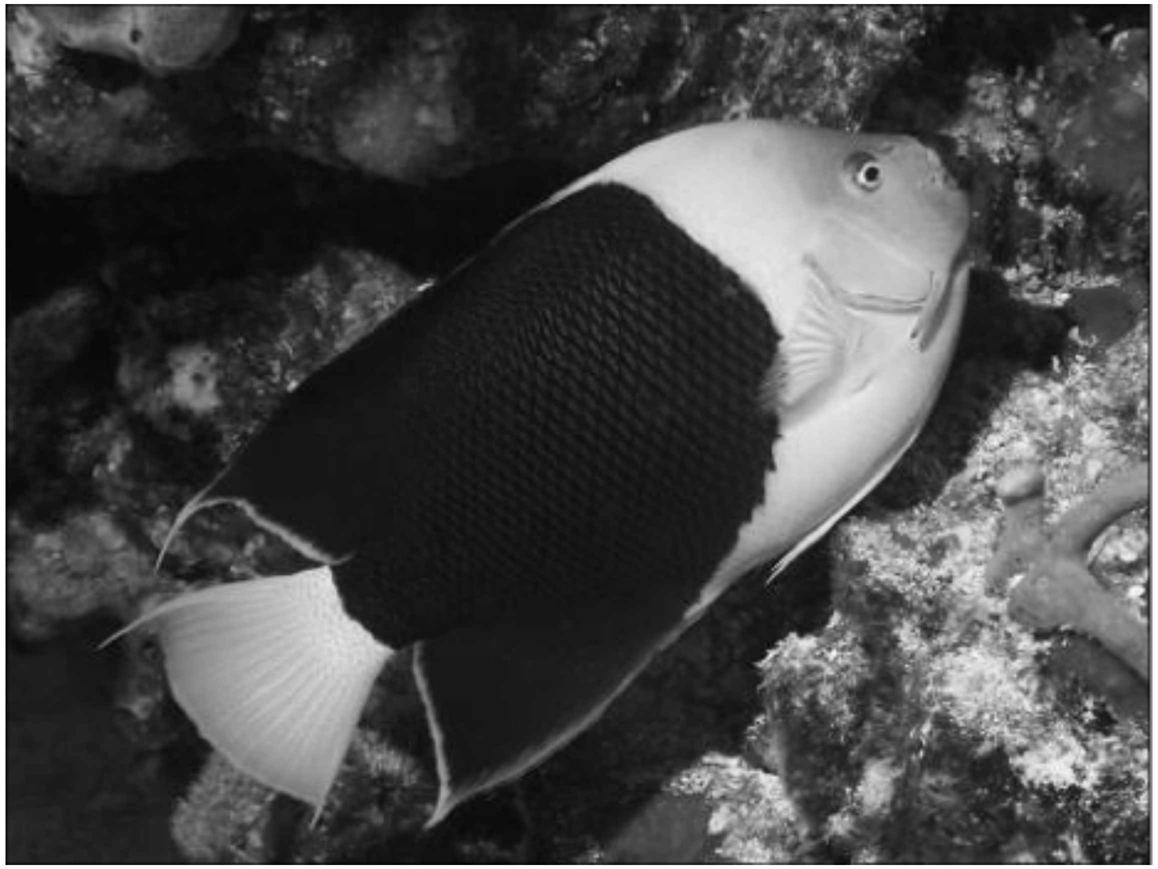
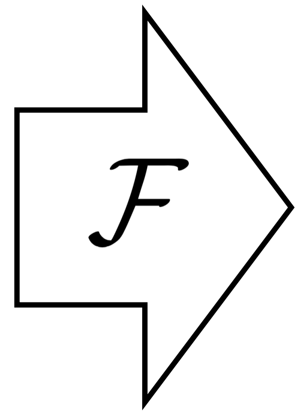


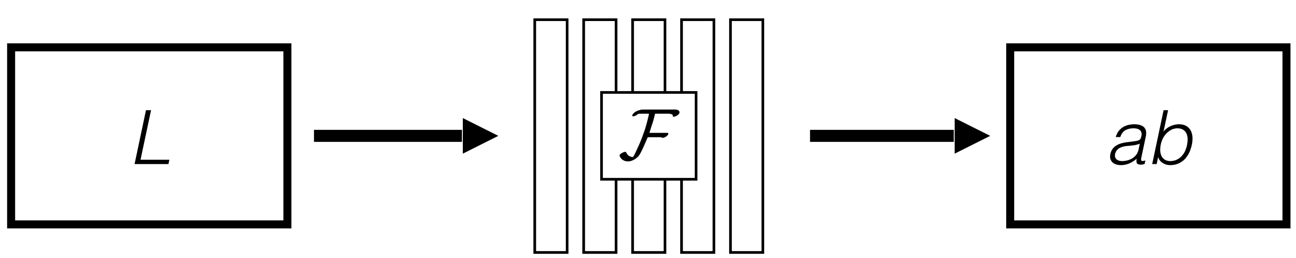
Image: canal L (de Lab)

$$\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$$



Concaténation (L,ab), conversion vers RGB

$$(\mathbf{X}, \widehat{\mathbf{Y}})$$



P. Isola

Malheureusement, minimiser la somme des différences au carré ne fonctionne pas! ☹

Entrée

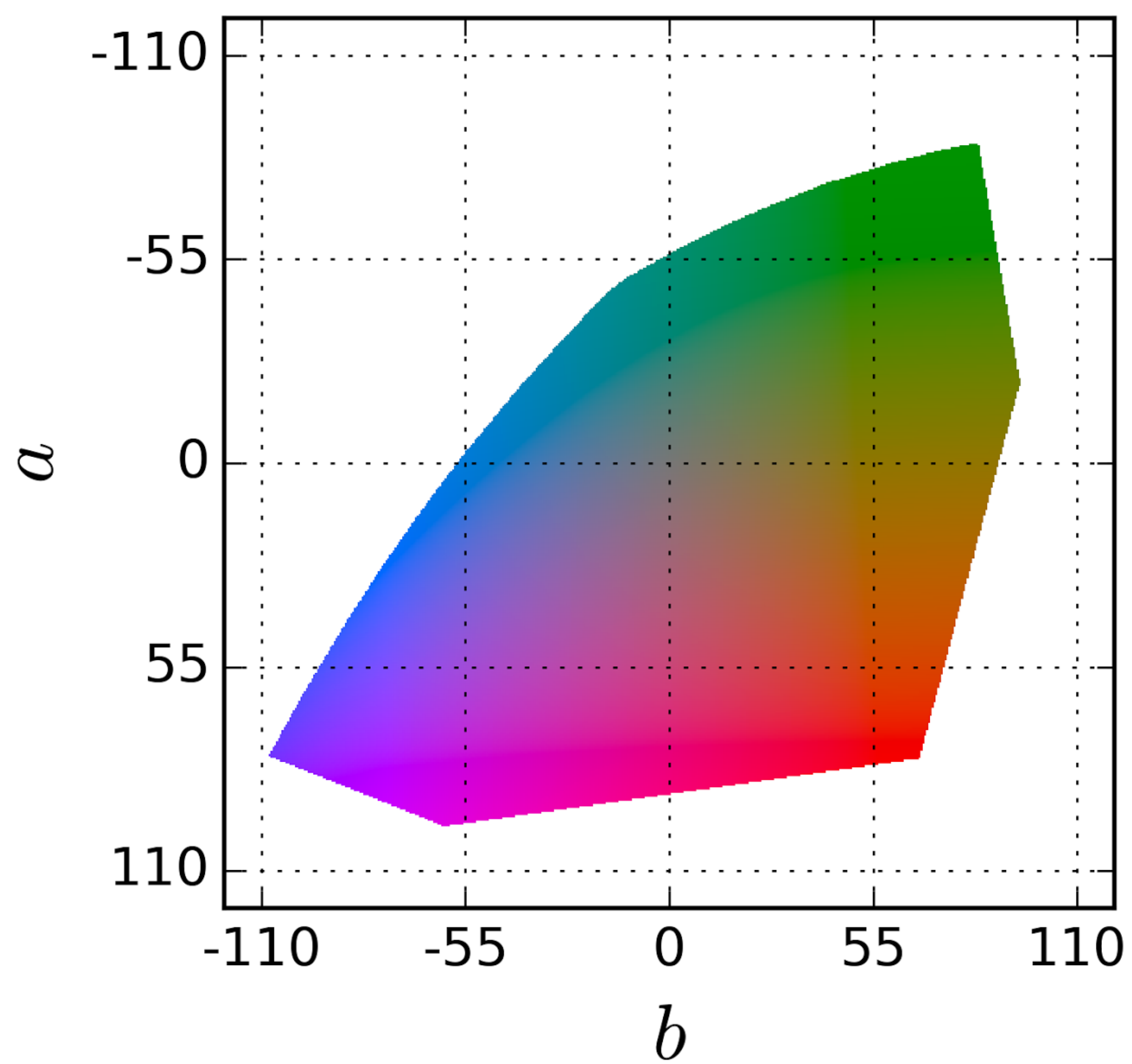
Sortie

Vérité

$$L_2(\widehat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2} \sum_{h,w} \| \mathbf{Y}_{h,w} - \widehat{\mathbf{Y}}_{h,w} \|_2^2$$

Meilleure fonction de perte

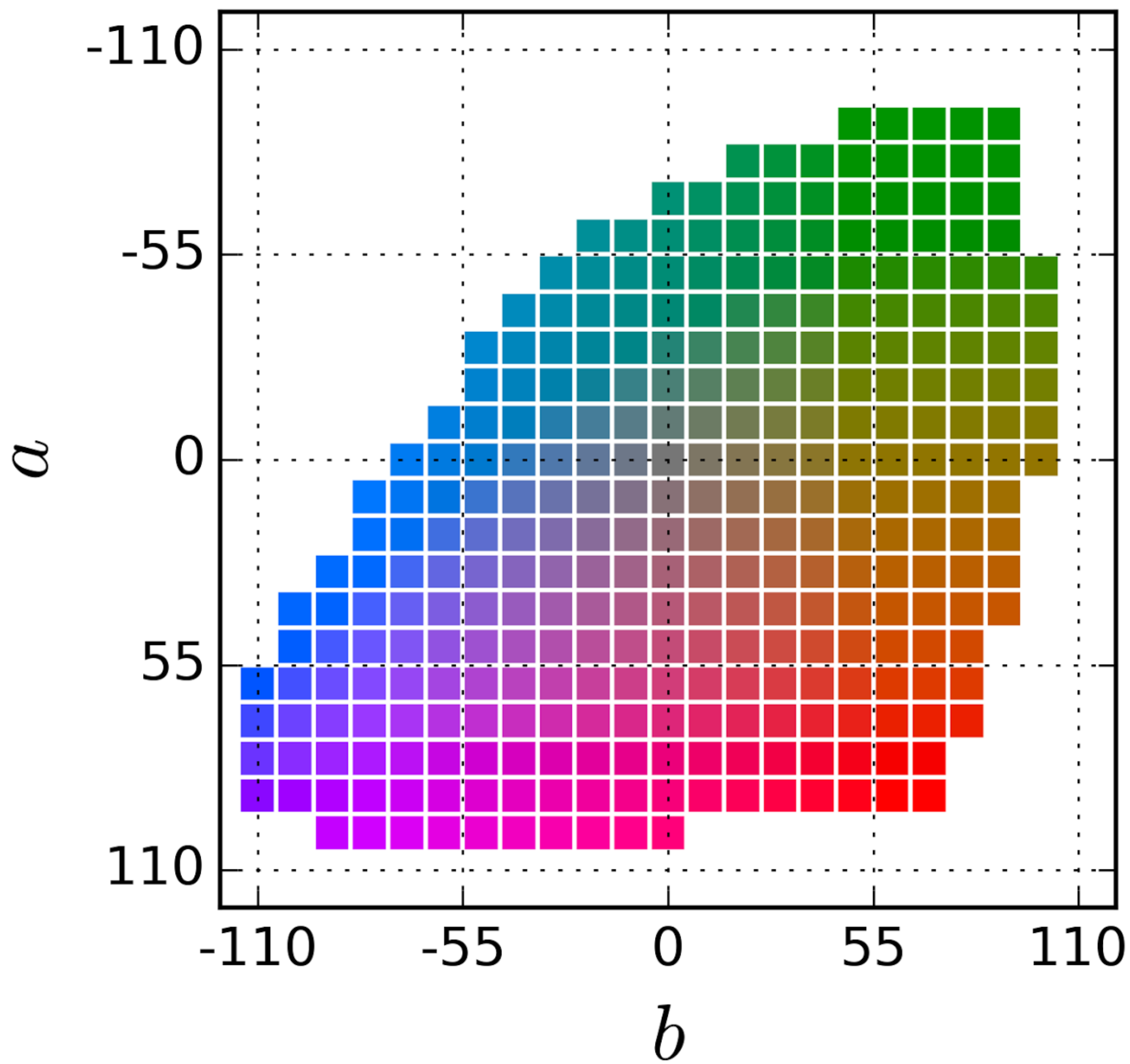
Espace (continu) des couleurs ab



P. Isola

Meilleure fonction de perte

Discrétisation de l'espace de couleurs ab

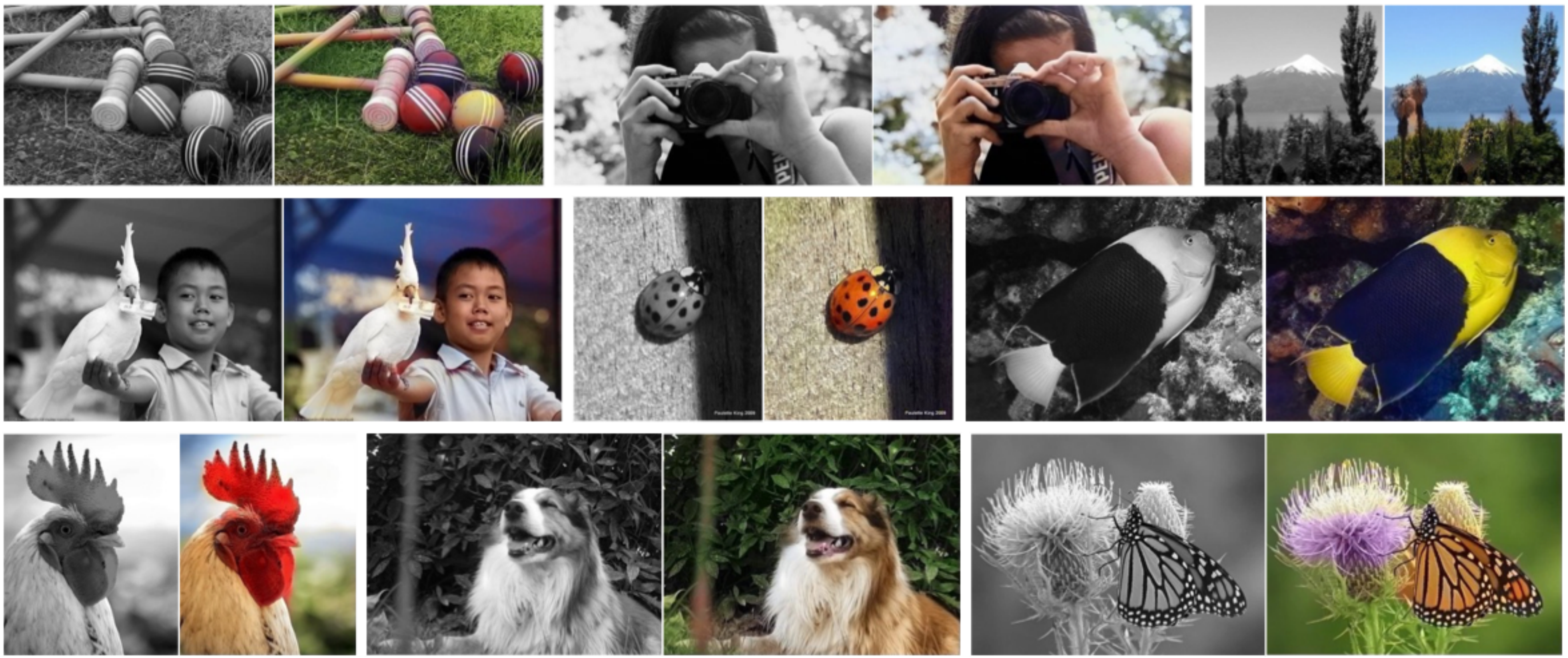


Prédire la *distribution* sur l'espace ab!

Permet de modéliser les distributions multimodales (plusieurs couleurs sont plausibles)

P. Isola

Bons résultats

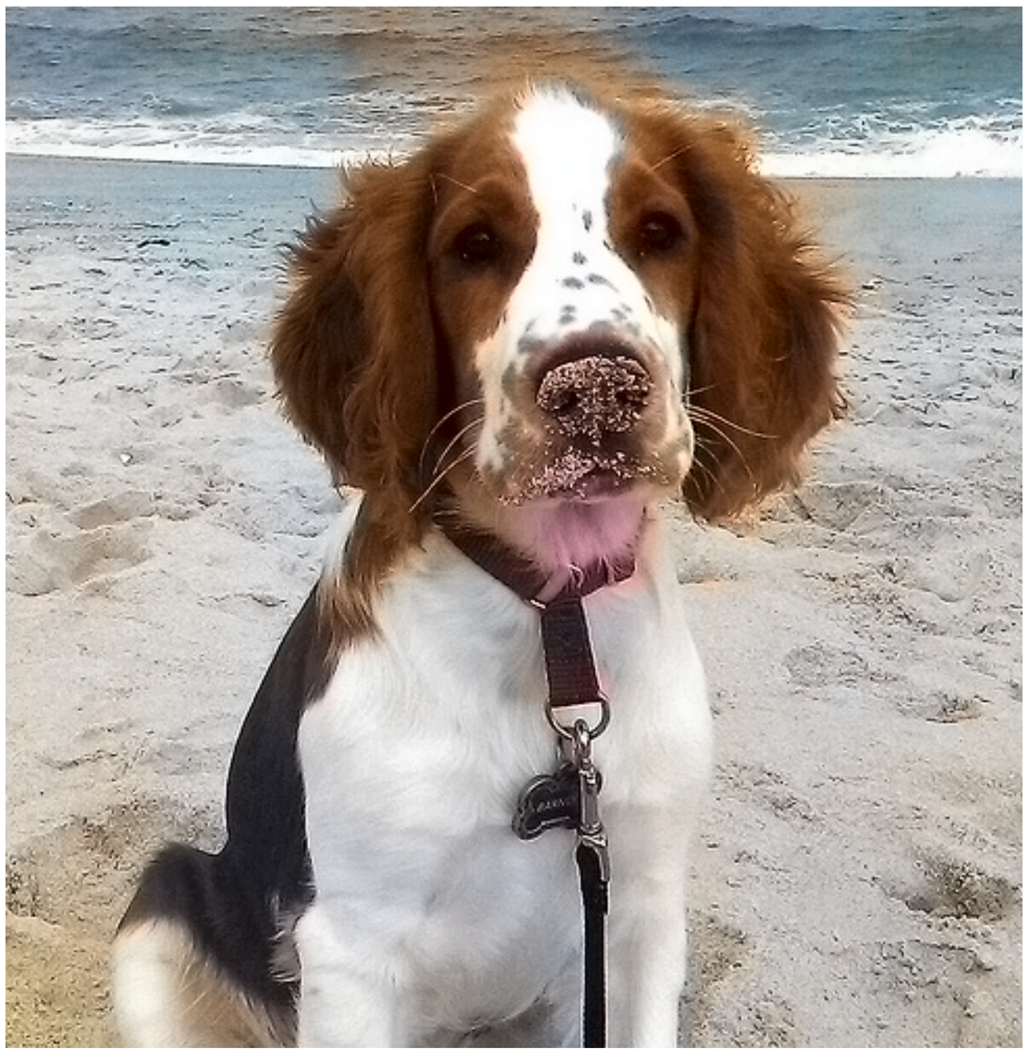


P. Isola

Moins bons résultats

P. Isola

Biais

P. Isola

Reddit /u/SherySantucci

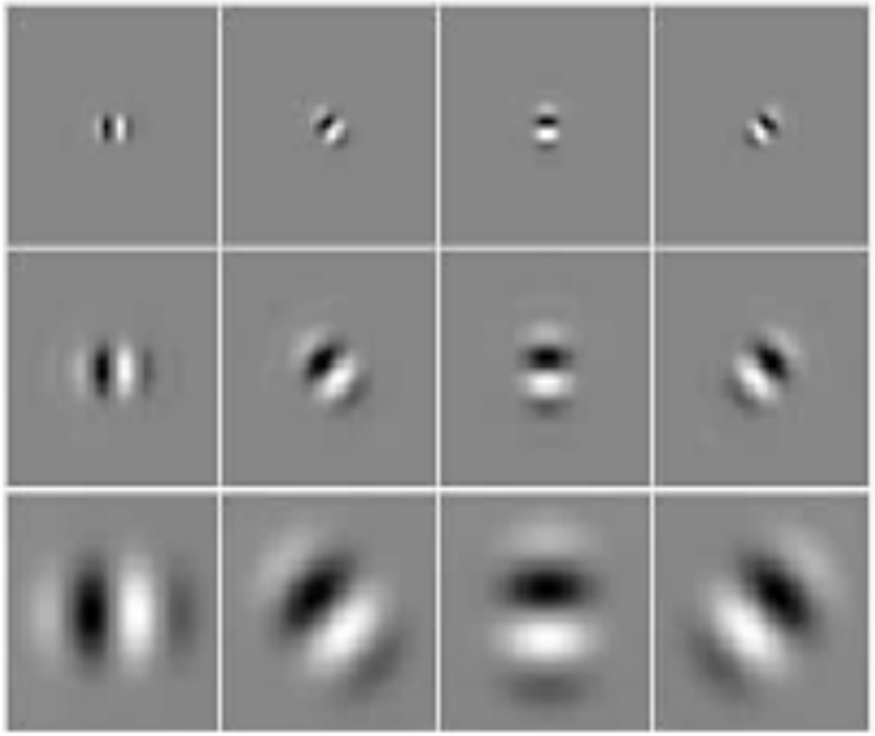
Reddit ColorizeBot

Photo de Reddit /u/Timteroo,
Murale de Eduardo Kobra

Reddit ColorizeBot

P. Isola

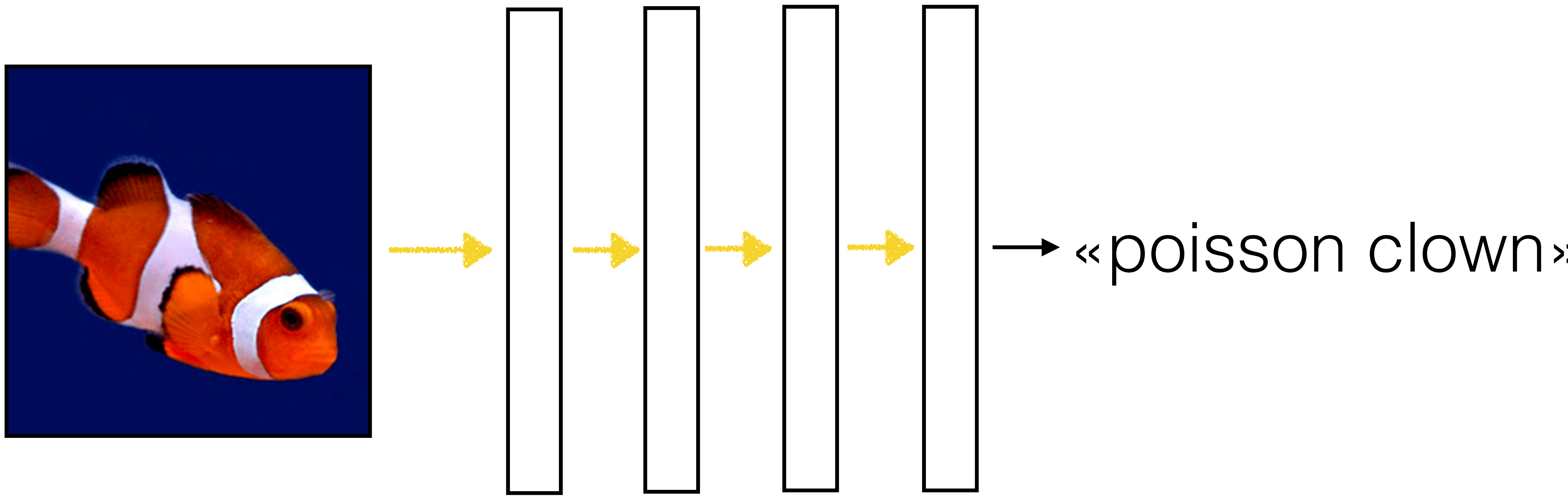
Représenter les textures



Pourquoi *ces* caractéristiques exactement?

A. Owens

Choix de la représentation

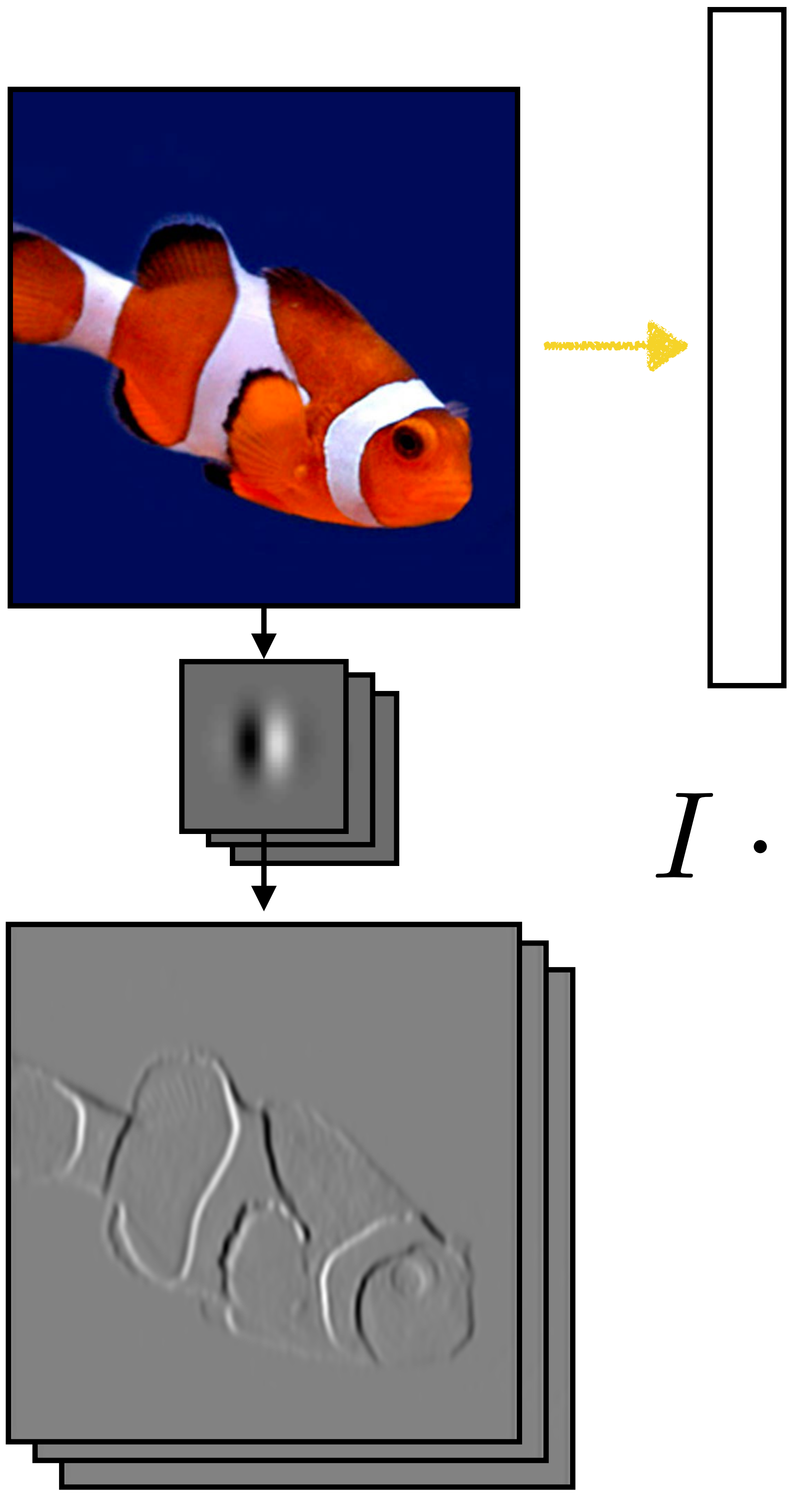


Au lieu de banques de filtres pré-déterminés, utilisons les caractéristiques apprises par un réseau de neurones!

Un réseau de neurone entraîné à reconnaître les objets devient un calculateur de caractéristiques très robustes.

A. Owens

Extraire les caractéristiques d’une image



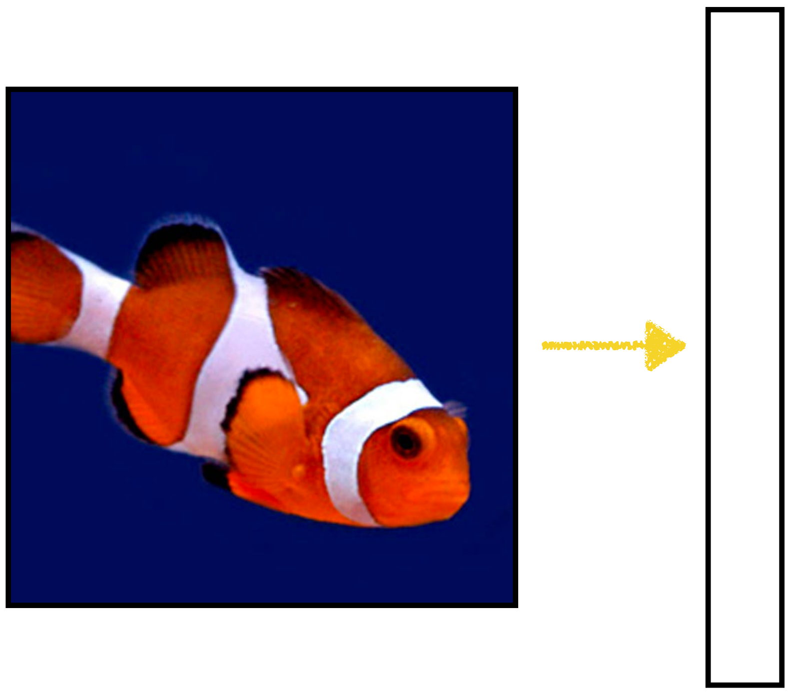
$$I \cdot W_1$$

convolution
(= multiplication matricielle)

A. Owens

Extraire les caractéristiques d'une image

«conv1»



$$\lfloor I \cdot W_1 \rfloor$$

convolution suivie d'une non-linéarité

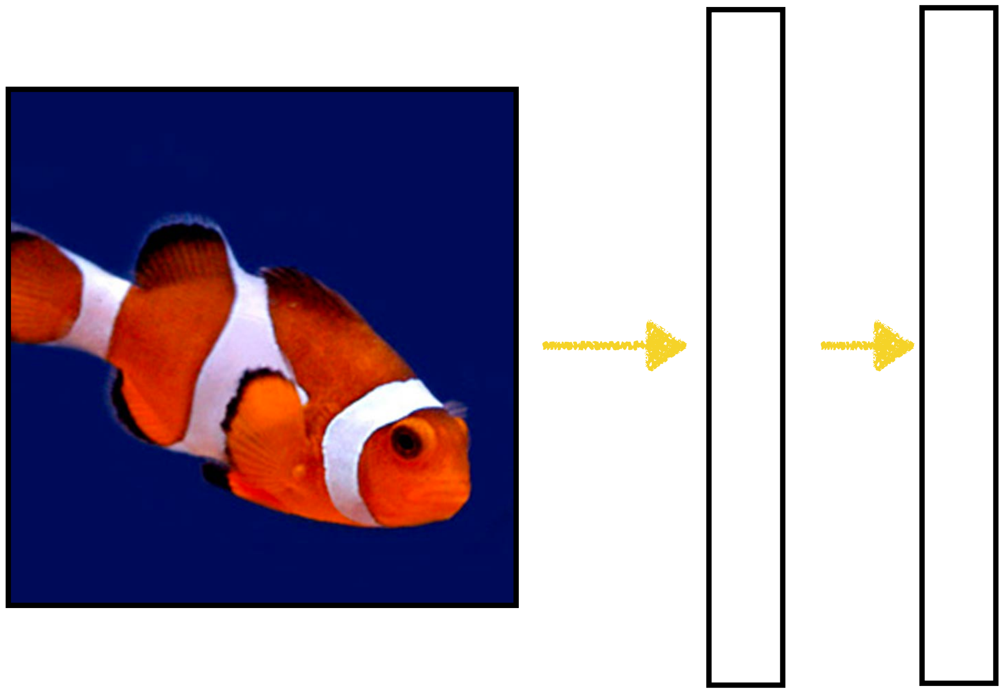
où $$\lfloor \vec{x} \rfloor_i = \max(x_i, 0)$$

i.e. valeurs négatives = 0

A. Owens

Extraire les caractéristiques d'une image

«conv2»

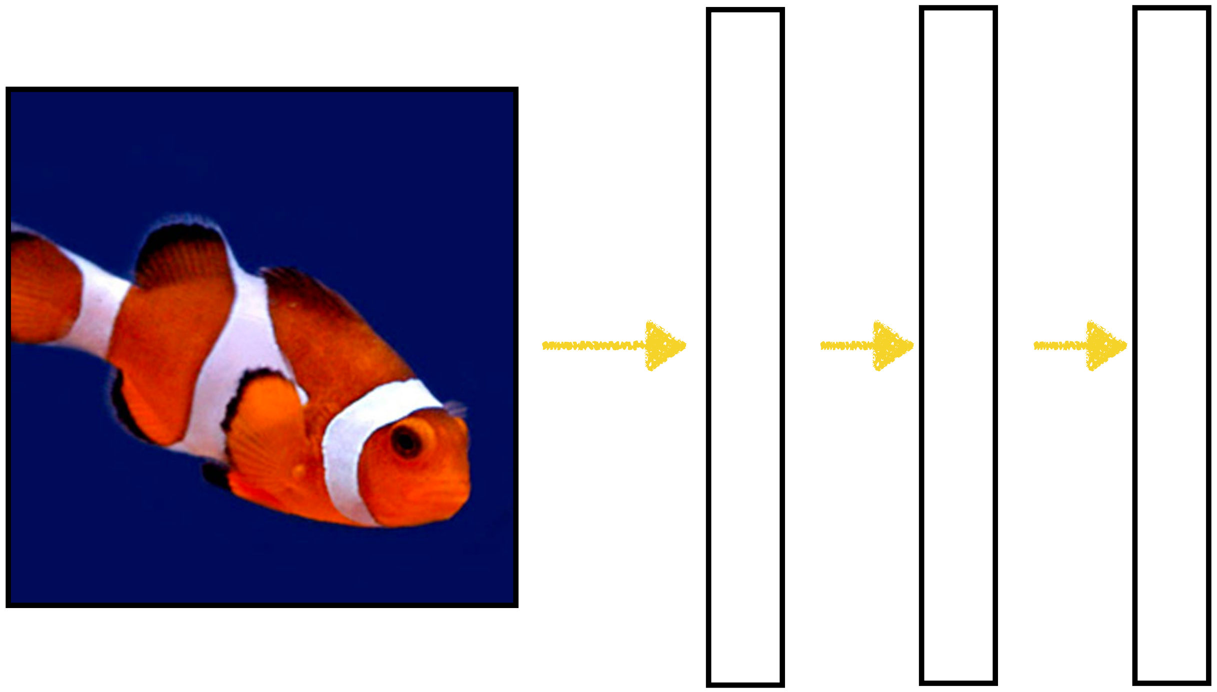


$$\lfloor \lfloor I \cdot W_1 \rfloor \cdot W_2 \rfloor$$

A. Owens

Extraire les caractéristiques d’une image

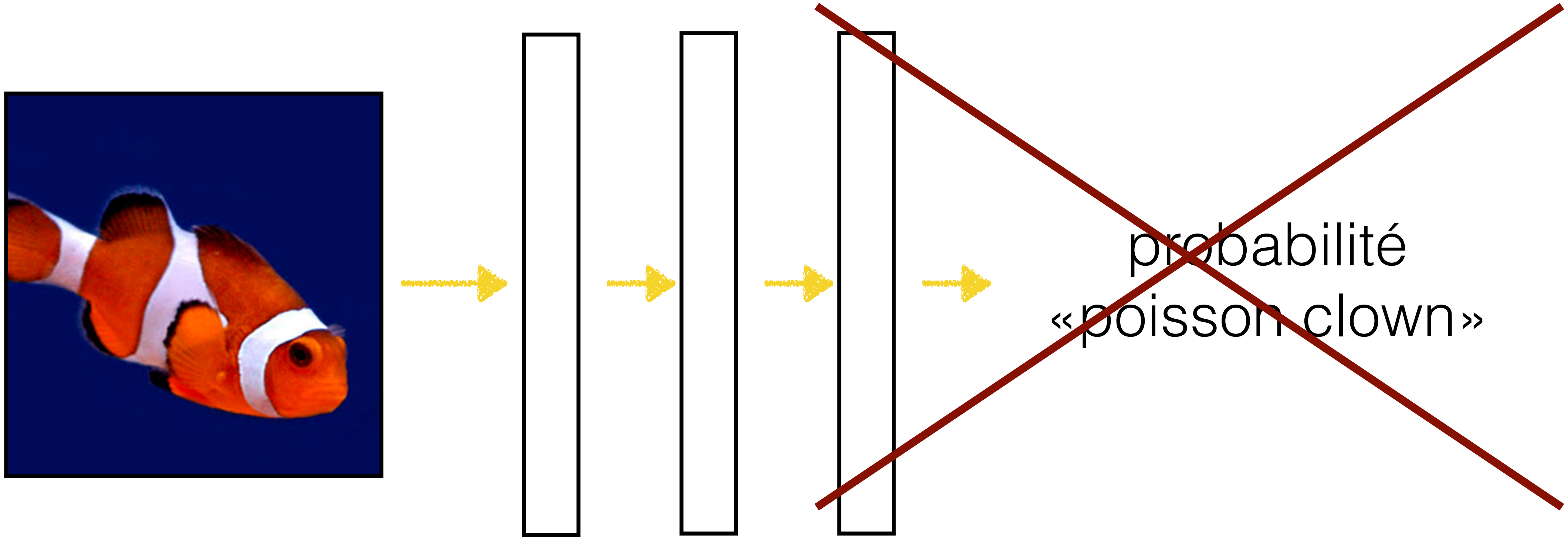
«conv3»



$$\lfloor \lfloor \lfloor I \cdot W_1 \rfloor \cdot W_2 \rfloor \cdot W_3 \rfloor$$

A. Owens

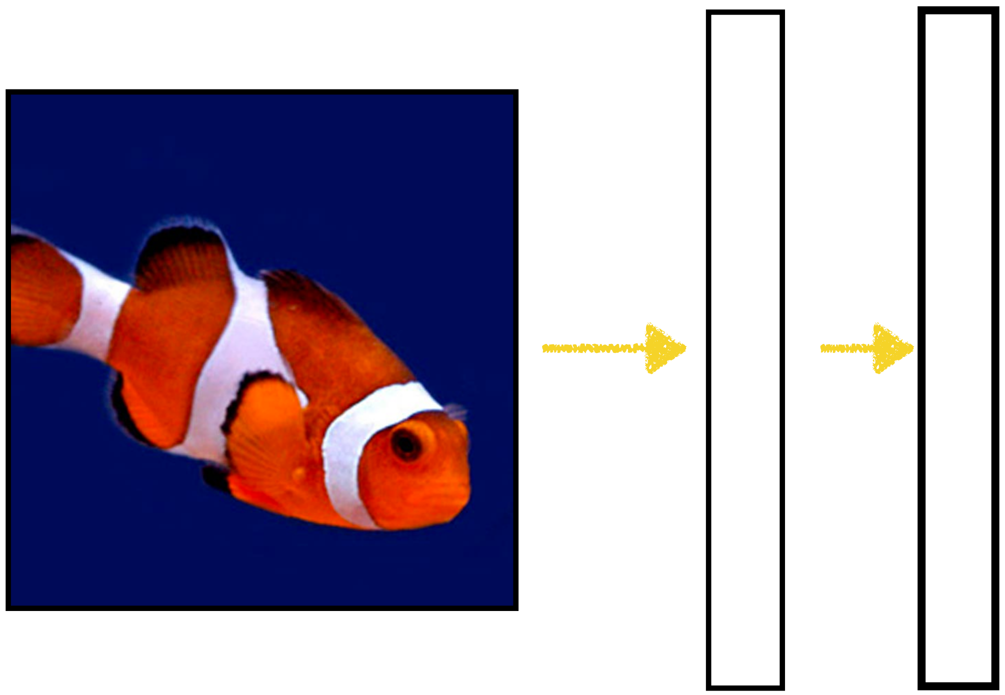
Extraire les caractéristiques d’une image



$$\lfloor\lfloor\lfloor I \cdot W_1 \rfloor \cdot W_2 \rfloor \cdot W_3 \rfloor \cdot w_{\text{“clown fish”}}$$

A. Owens

Extraire les caractéristiques d'une image

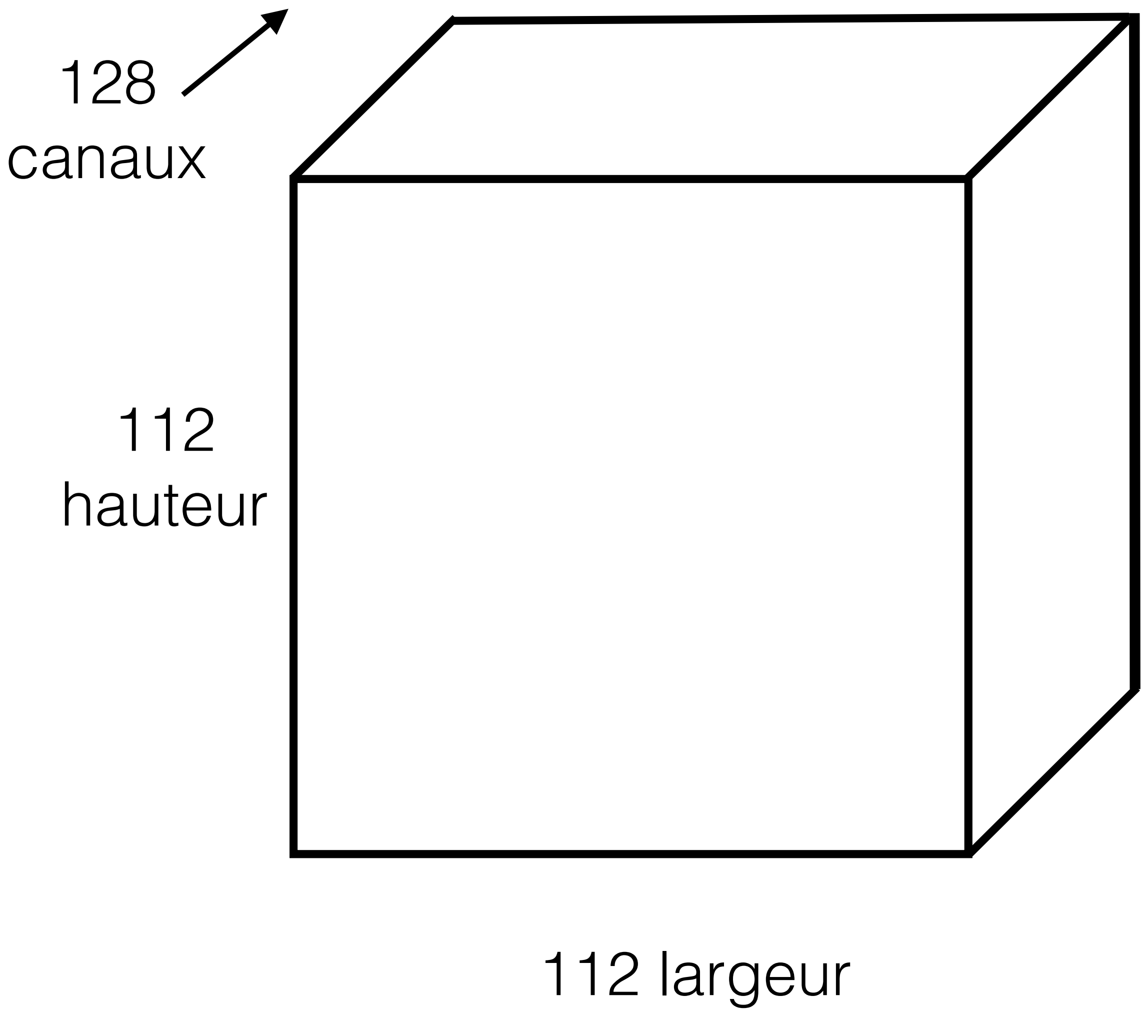


«activations» de la couche «conv2»

$$\lfloor \lfloor I \cdot W_1 \rfloor \cdot W_2 \rfloor$$

A. Owens

Extraire les caractéristiques d'une image



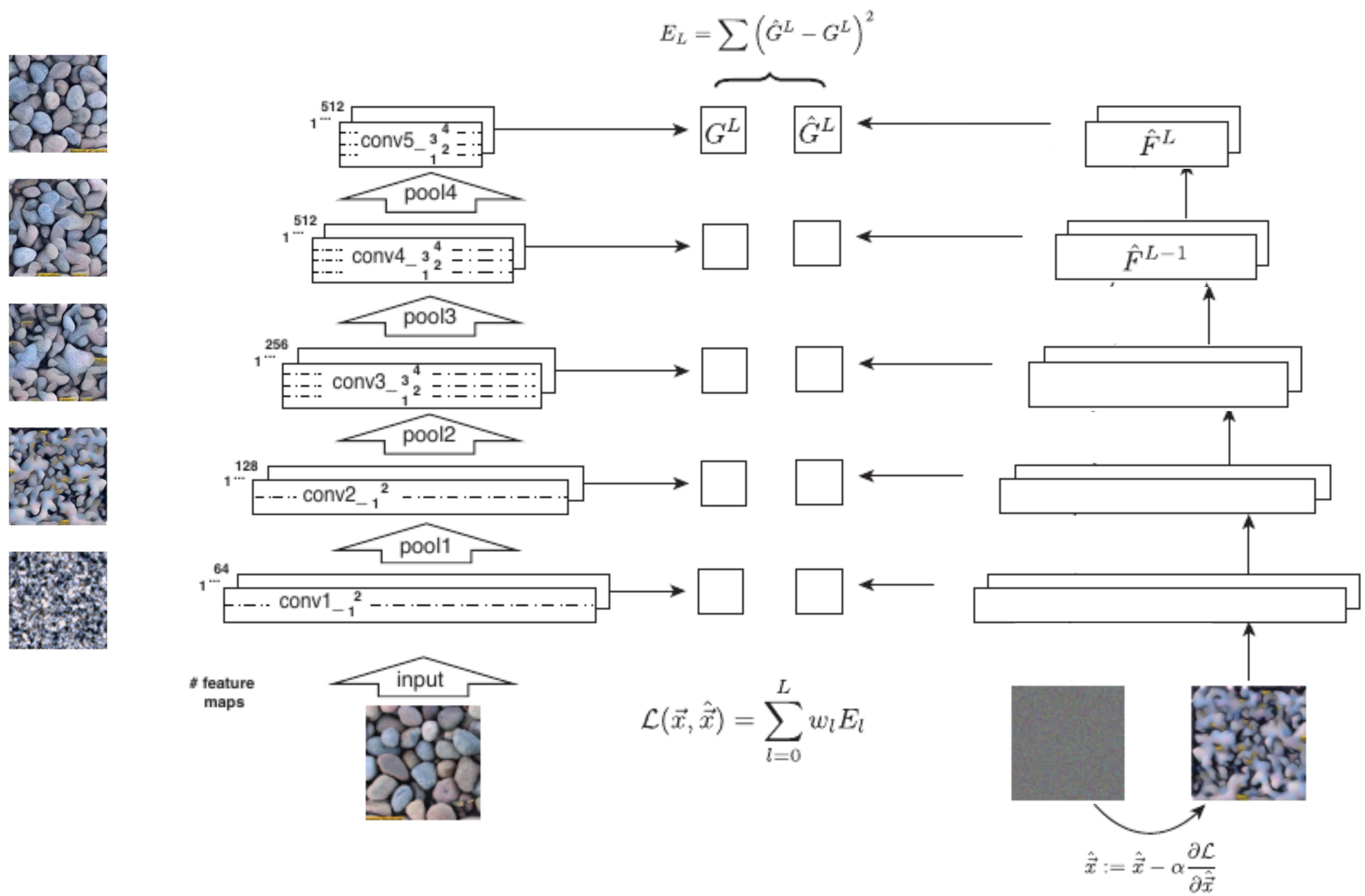
activations de «conv2»

e.g. (112 x 112) x 128 pour conv2

d'un modèle populaire (VGG19)

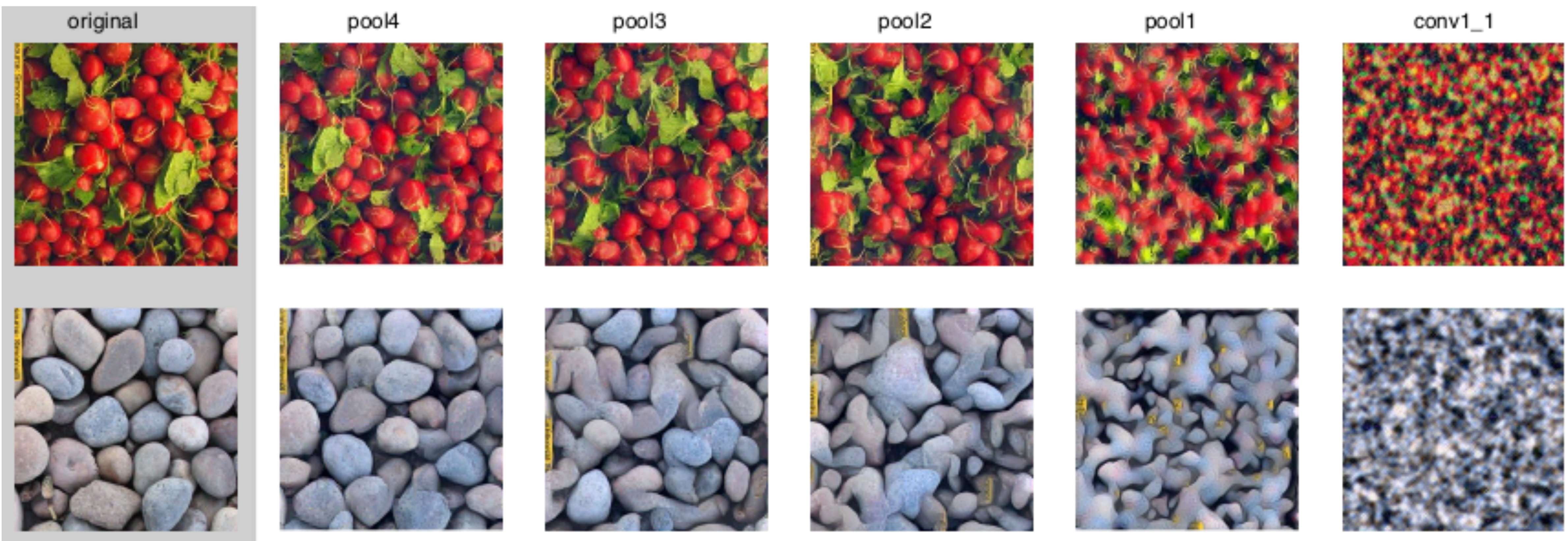
A. Owens

Synthèse de textures par CNN



J. Zhang

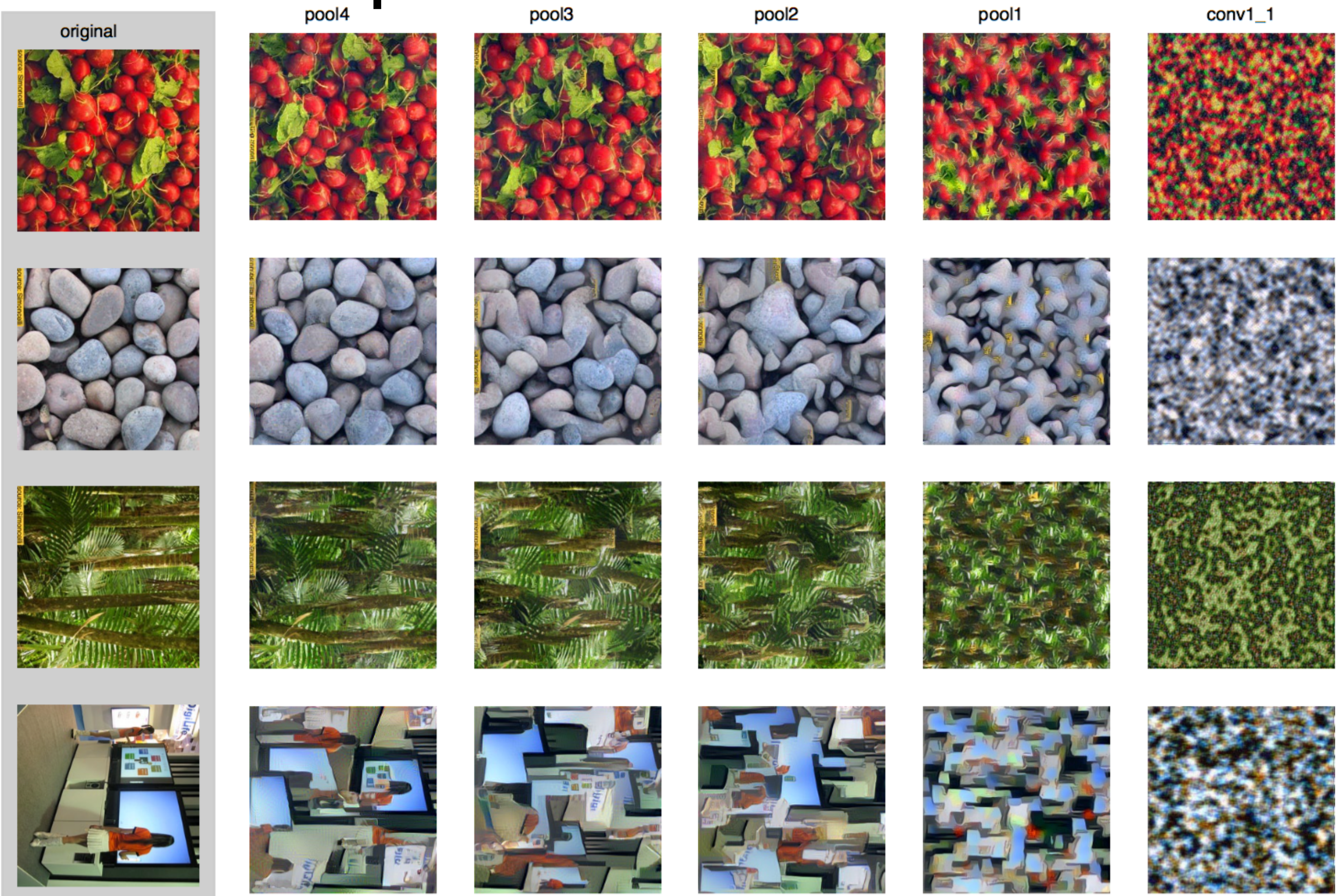
Synthèse de textures par CNN



[Gatys et al. 2005]

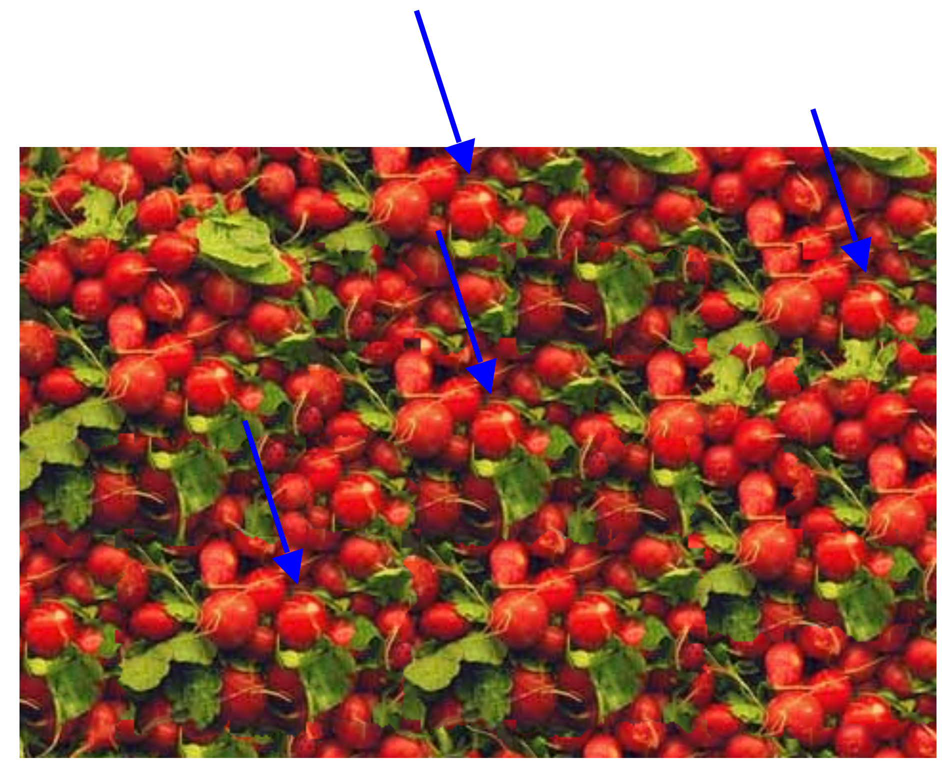
Haute → Basse

Statistiques de textures



Synthèse de textures par CNN

Échantillonner la texture originale



[Gatys et al. 2005]

Synthèse de textures par CNN

Synthesis

Source

[Gatys et al. 2005]

Synthèse de textures par CNN

Synthesis Source

[Gatys et al. 2005]

Capturer le style artistique

Comment transférer le style d'une peinture vers une photo?

[Gatys et al. 2016]

Capturer le style artistique

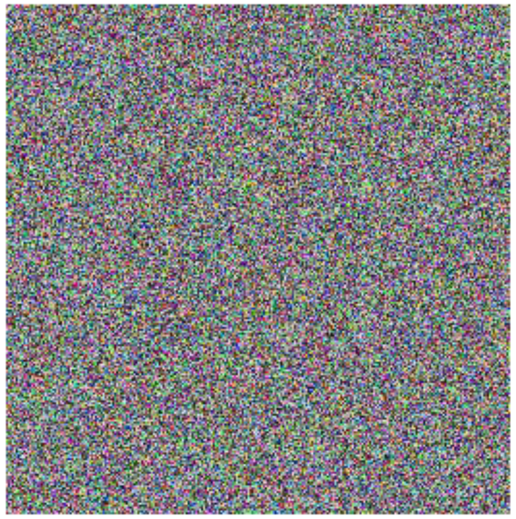
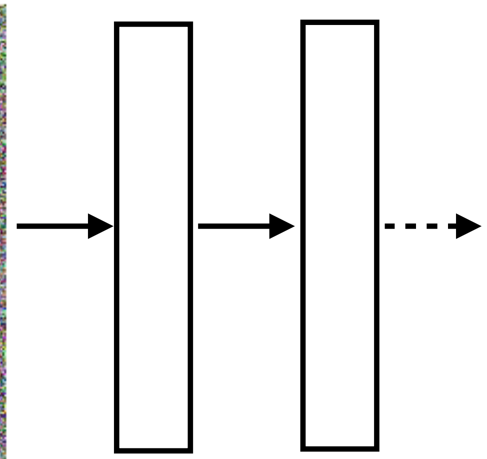


Image de synthèse



$\hat{G}$

$$\sum_{i,j} (\hat{G}_{ij} - G_{ij})^2$$

Rajouter le contenu de la photo

$$\sum_i \sum_{x,y} (c_i(x,y) - \hat{c}_i(x,y))^2$$

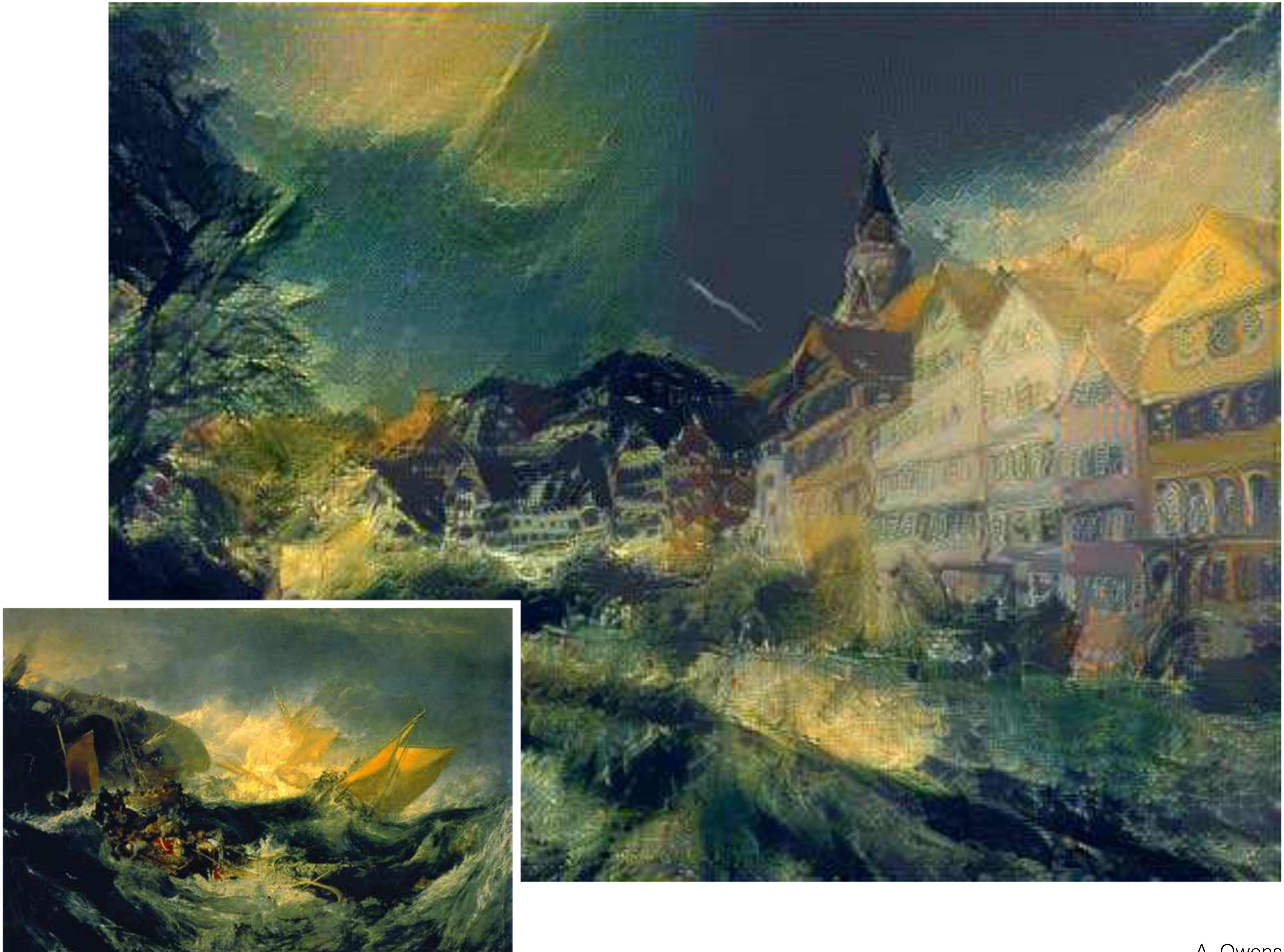
$c_i(x,y)$

$\hat{c}_i(x,y)$

A. Owens

A. Owens

A. Owens



A. Owens

A. Owens

A. Owens

Londres le jour

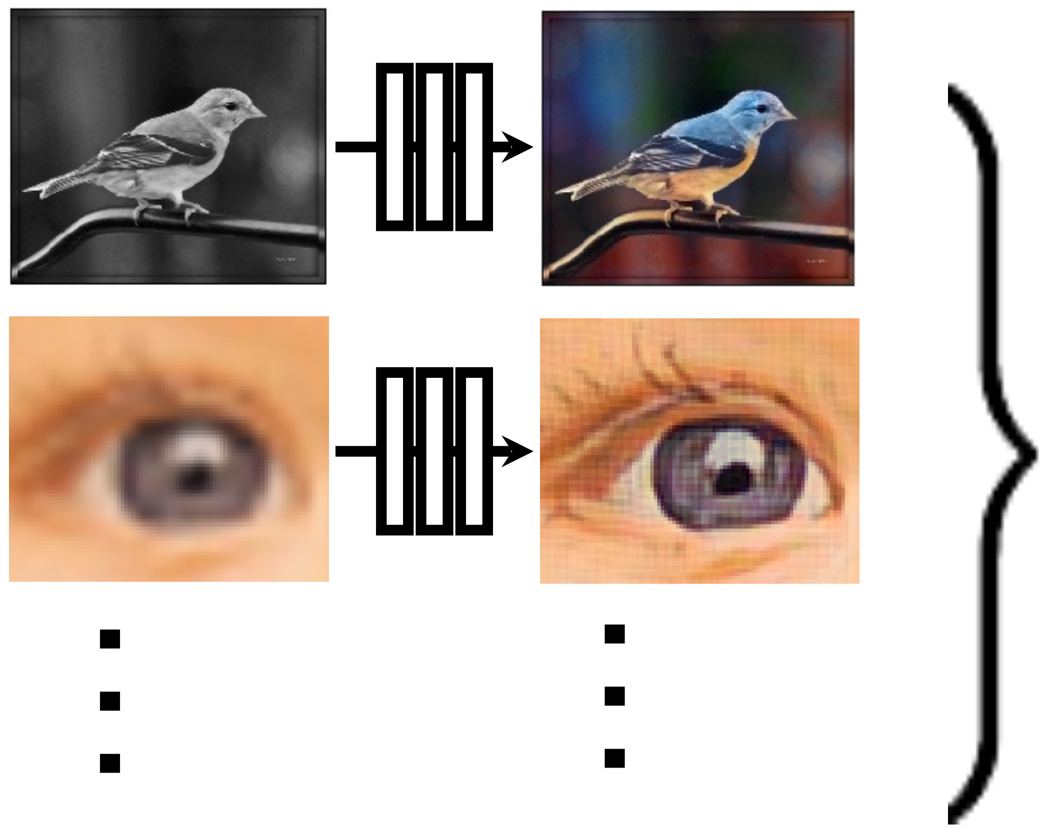
New York le soir

A. Owens

Londres le soir?

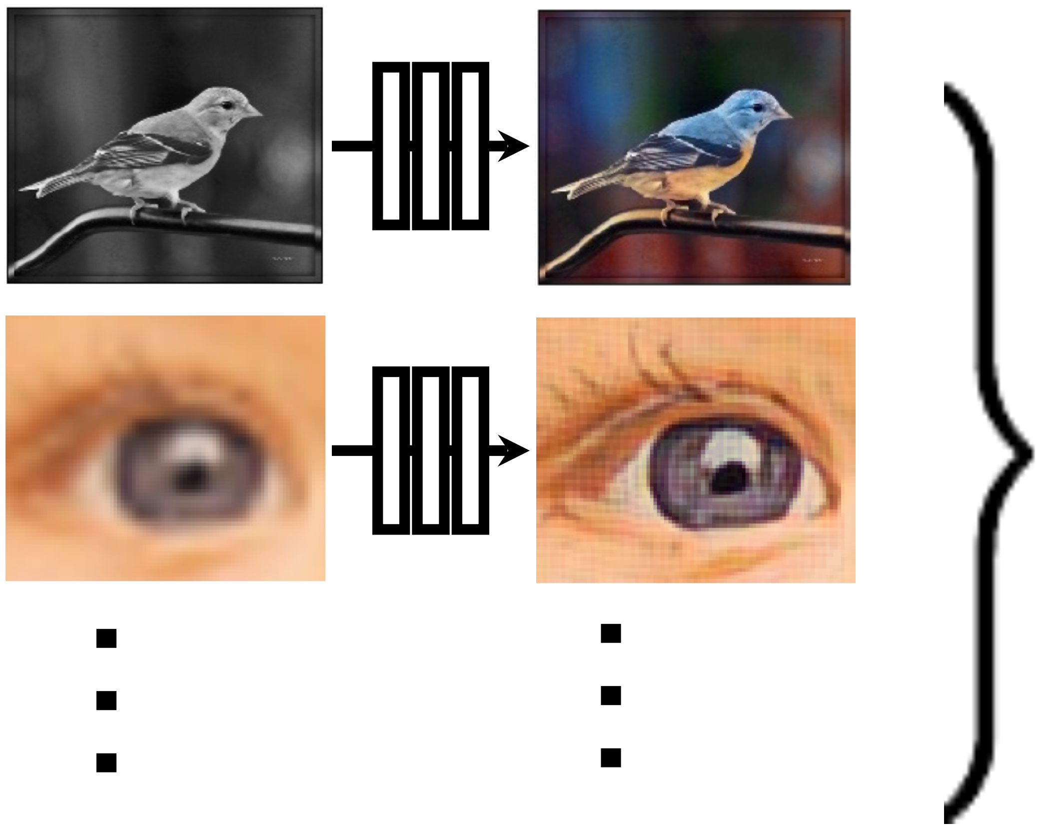
A. Owens

Revenons à la colorisation…



Quelle fonction de perte?

Image générées

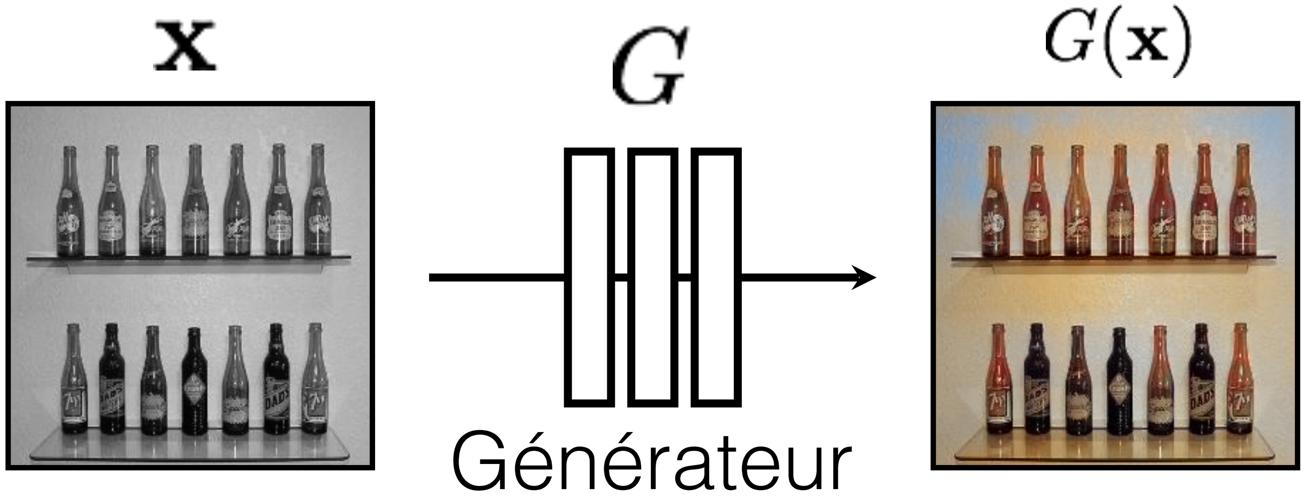


Réseaux génératifs adversariaux
Generative Adversarial Network
(GANs)

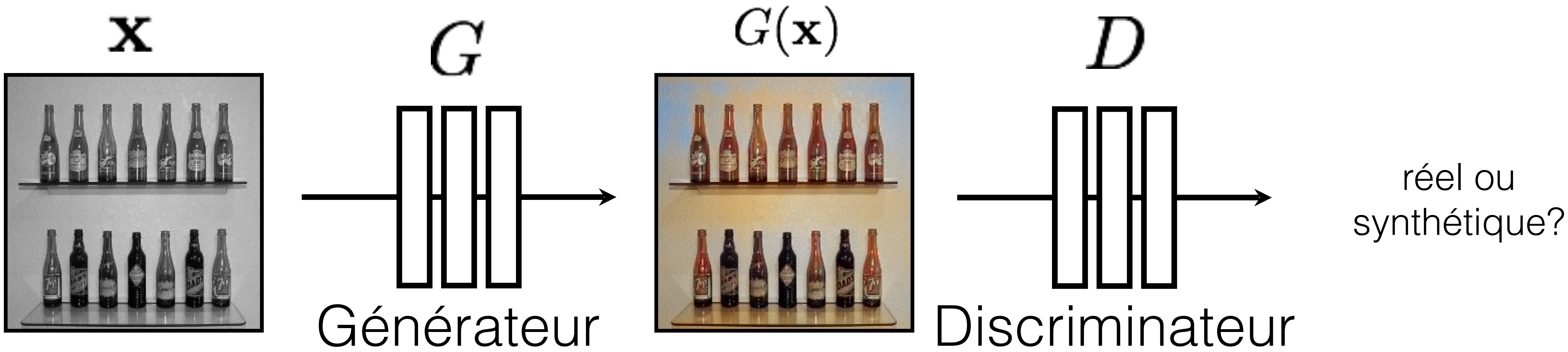
Classificateur:
généré vs réel

Image réelles

[Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville, Bengio 2014]



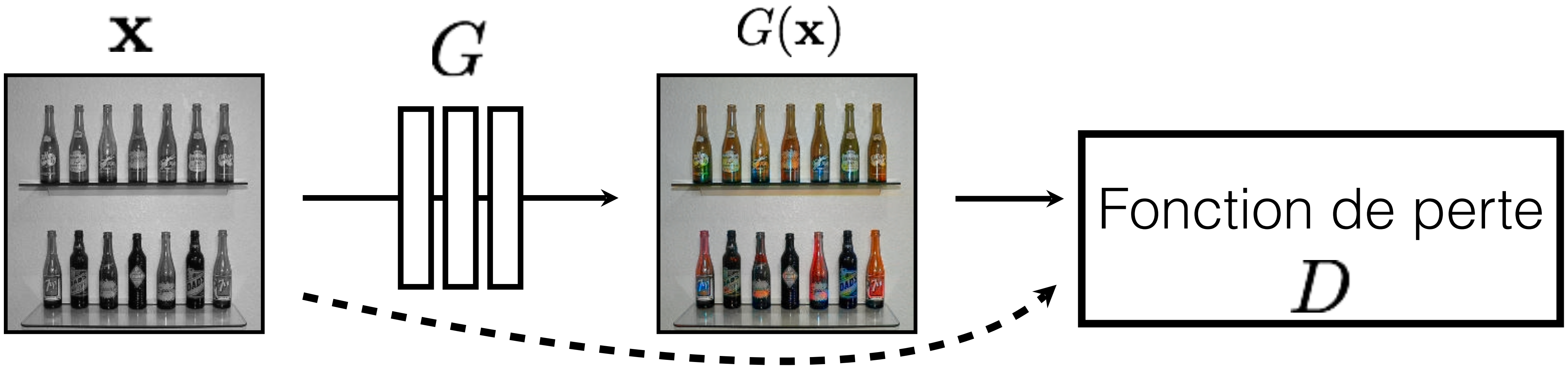
[Goodfellow et al., 2014]



G essaie de générer des images pour tromper **D**

D essaie d’identifier les images générées par **G**

[Goodfellow et al., 2014]



$+ L1$

Du point de vue de **G**, **D** est une fonction de perte.

Au lieu d'être déterminée à la main (L2, etc.), elle est *apprise*.

[Goodfellow et al., 2014]
[Isola et al., 2017]

Gris vers couleur

Images de [Russakovsky et al. 2015]

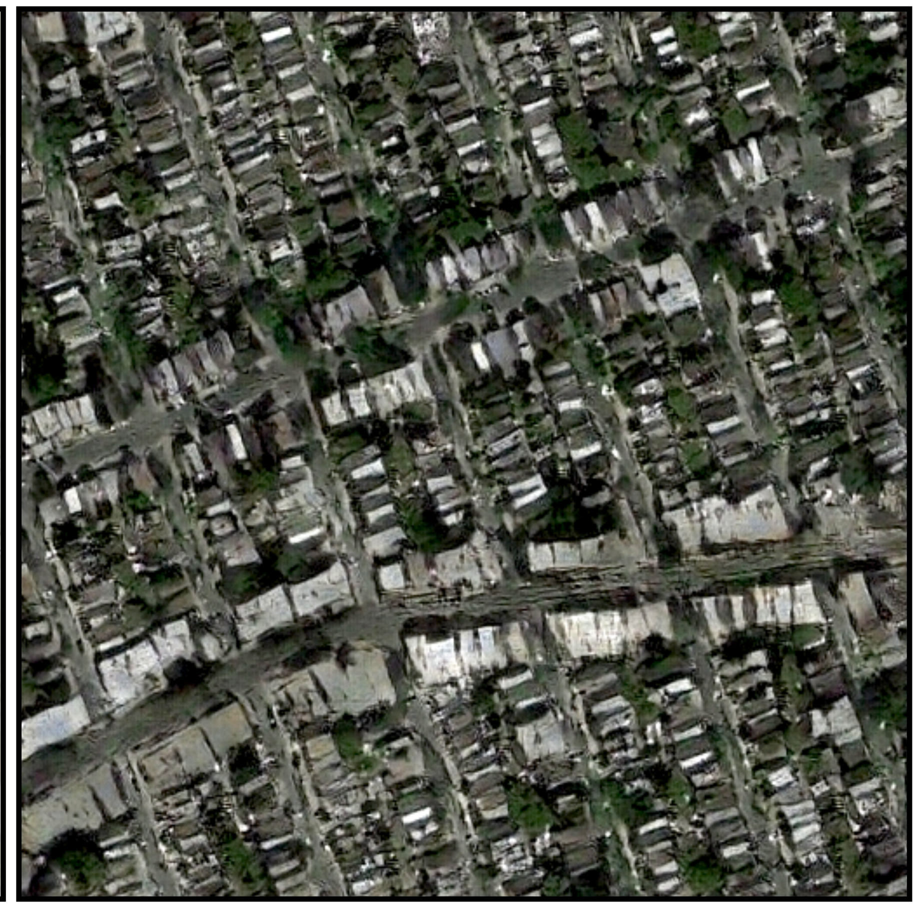
Gris vers couleur

Images de [Russakovsky et al. 2015]

Carte vers satellite

Input

Output

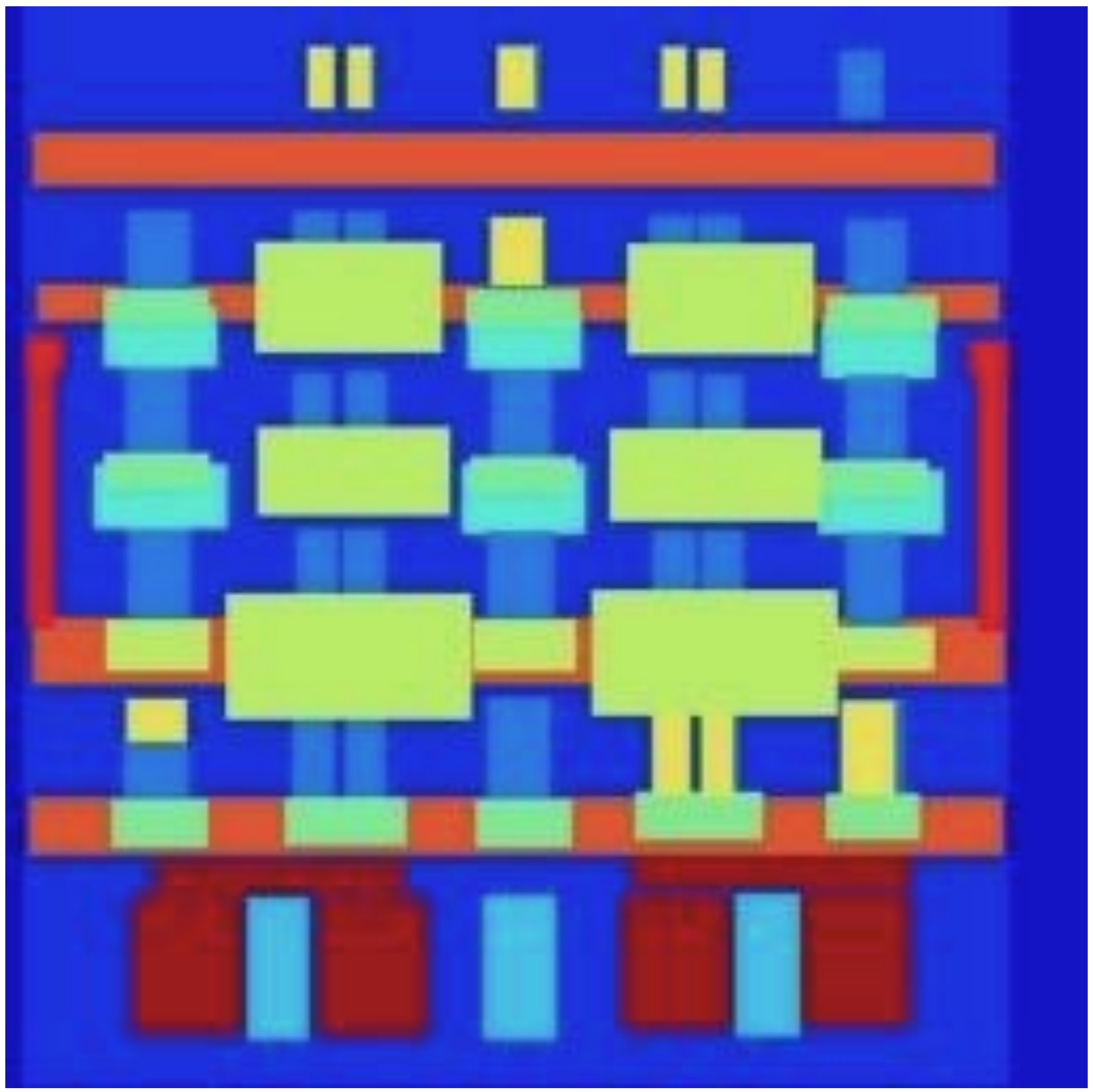


Groundtruth

Images de [maps.google.com]

Étiquette vers façade

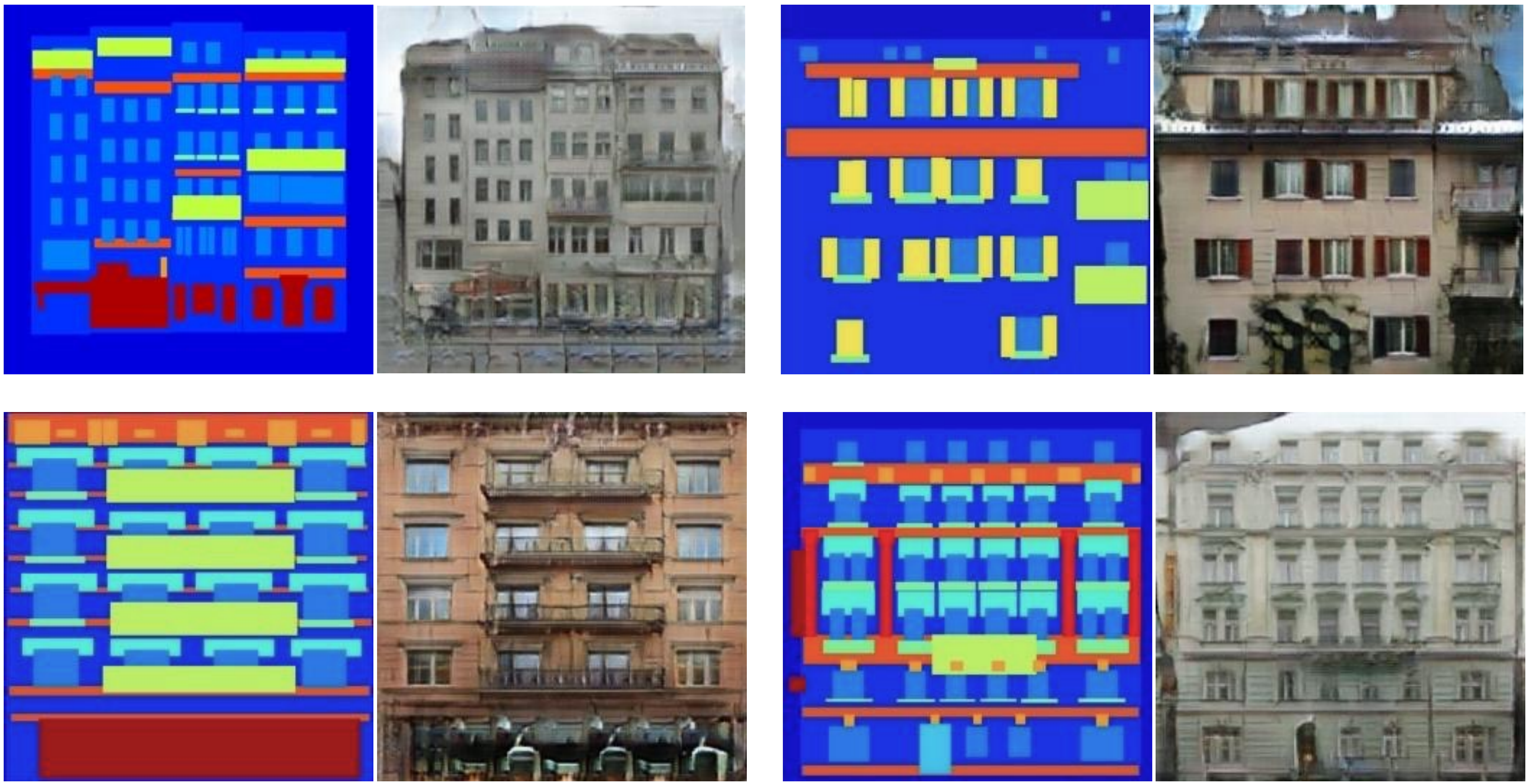
Entrée



Sortie

Données de [Tylecek, 2013]

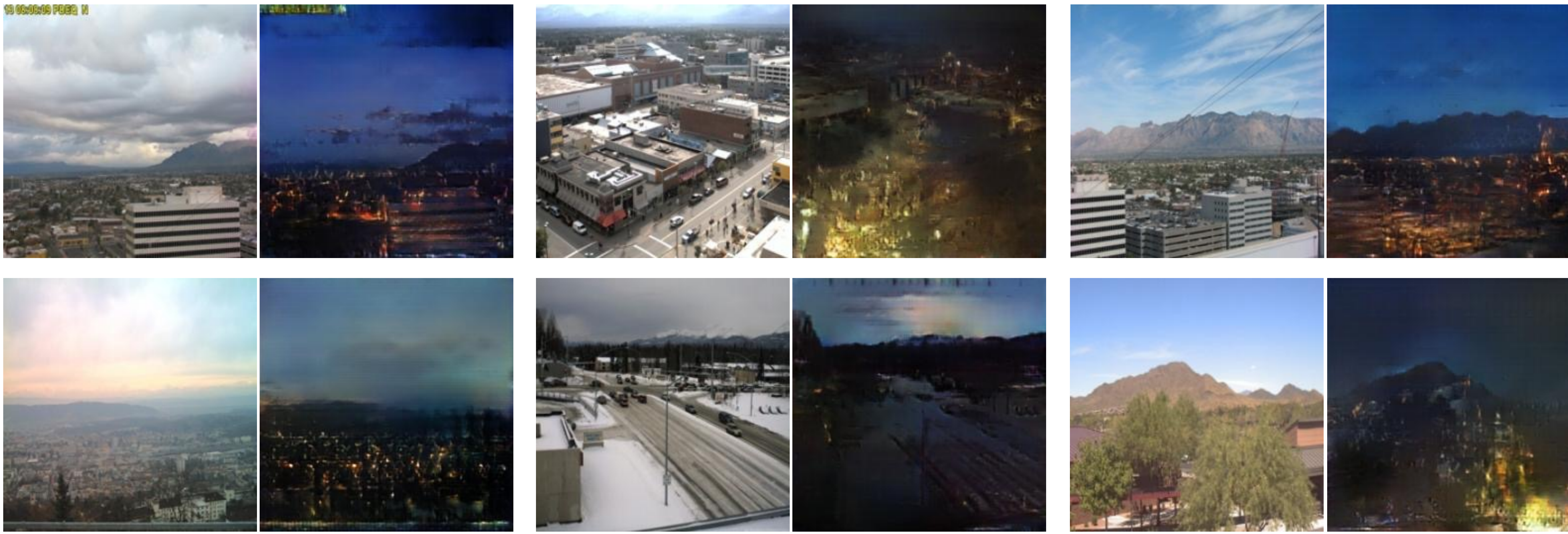
Étiquette vers façade



Données de [Tylecek, 2013]

C’est le jour et la nuit!

Input Output Input Output Input Output



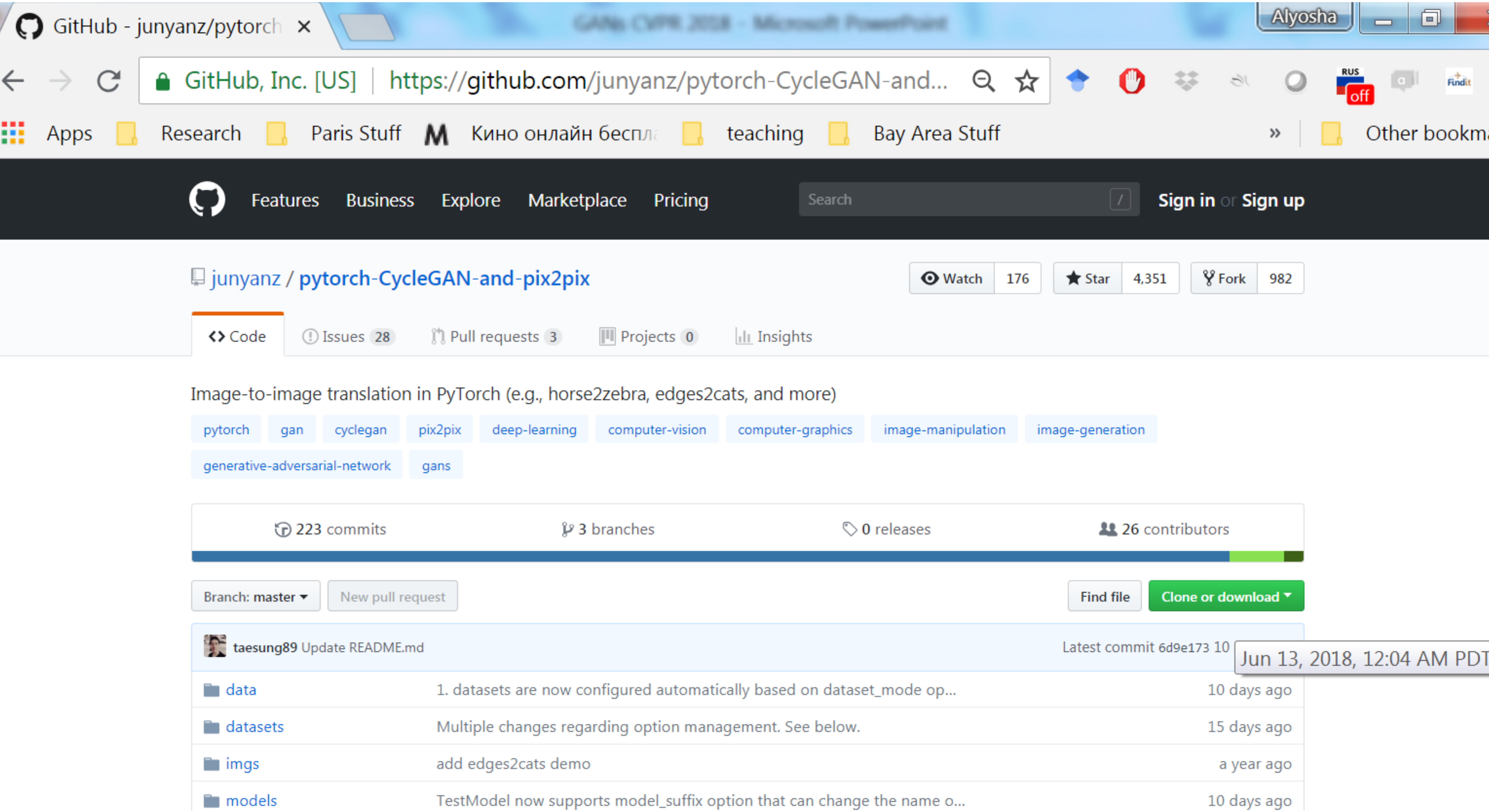
Données de [Laffont et al., 2014]

Arêtes vers images

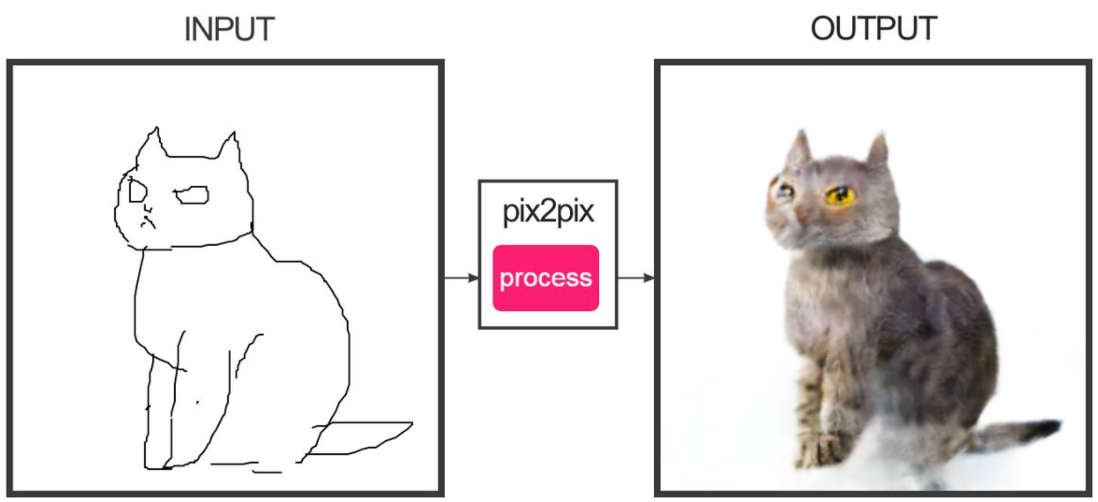
Arêtes de [Xie and Tu, 2015]

Dessins vers images

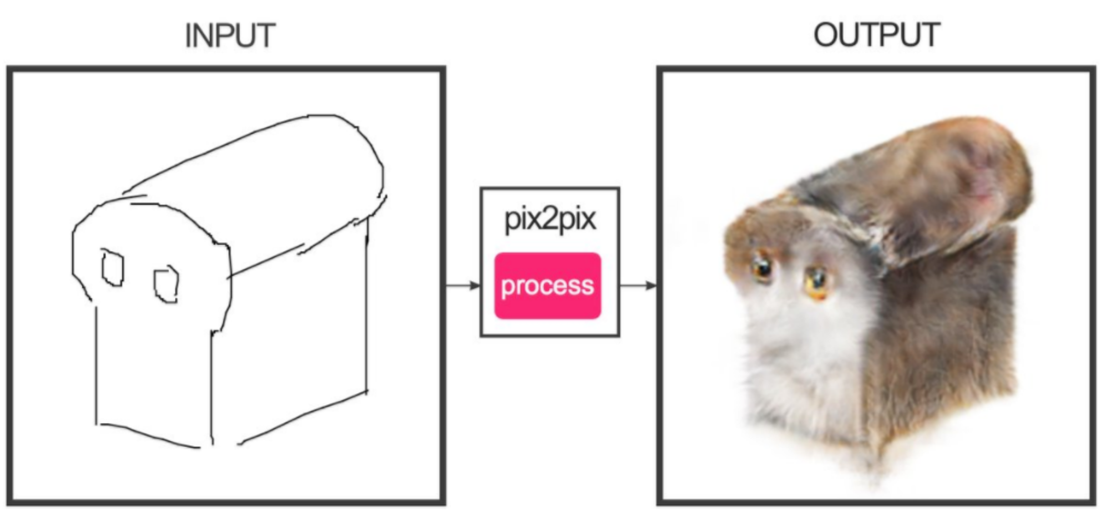
Données de [Eitz, Hays, Alexa, 2012]


GitHub - junyanz/pytorch
https://github.com/junyanz/pytorch-CycleGAN-and...
GitHub, Inc. [US]
Apps
Research
Paris Stuff
Кино онлайн беспл
teaching
Bay Area Stuff
Other bookm
Features
Business
Explore
Marketplace
Pricing
Search
Sign in or Sign up
junyanz / pytorch-CycleGAN-and-pix2pix
Watch 176
Star 4,351
Fork 982
Code
Issues 28
Pull requests 3
Projects 0
Insights
Image-to-image translation in PyTorch (e.g., horse2zebra, edges2cats, and more)
pytorch
gan
cyclegan
pix2pix
deep-learning
computer-vision
computer-graphics
image-manipulation
image-generation
generative-adversarial-network
gans
223 commits
3 branches
0 releases
26 contributors
Branch: master
New pull request
Find file
Clone or download
taesung89 Update README.md
Latest commit 6d9e173 10
Jun 13, 2018, 12:04 AM PDT
data
1. datasets are now configured automatically based on dataset_mode op...
10 days ago
datasets
Multiple changes regarding option management. See below.
15 days ago
imgs
add edges2cats demo
a year ago
models
TestModel now supports model_suffix option that can change the name o...
10 days ago

#edges2cats [Christopher Hesse]



@gods_tail

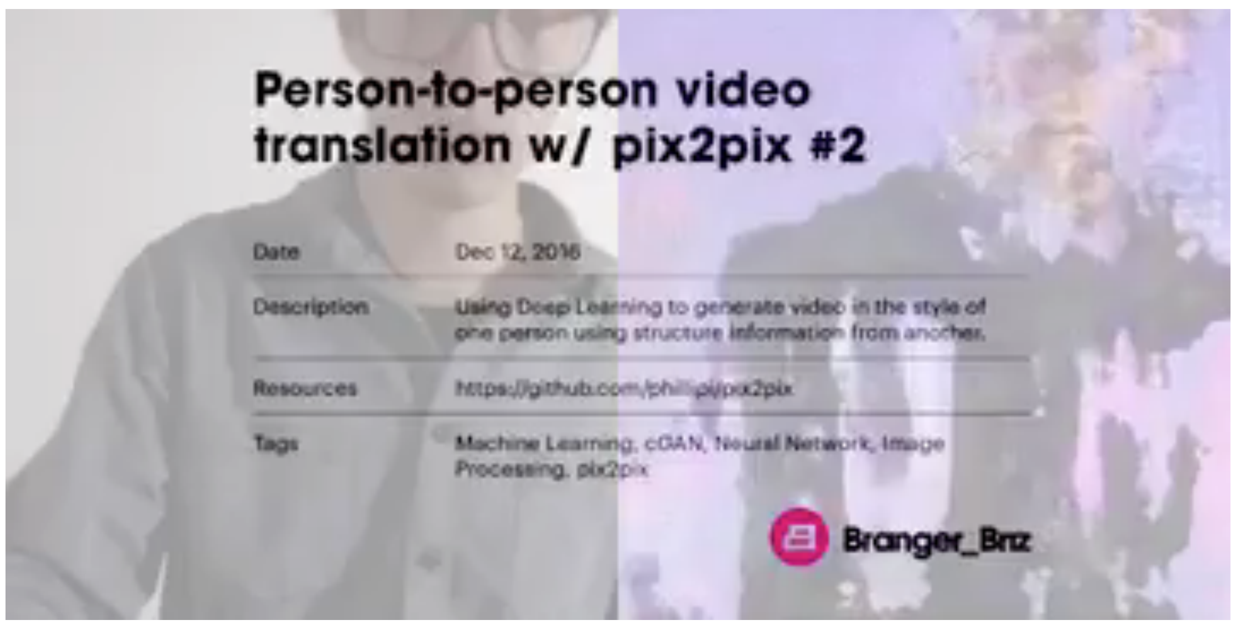


Ivy Tasi @ivymyt

Vitaly Vidmirov
@vvid

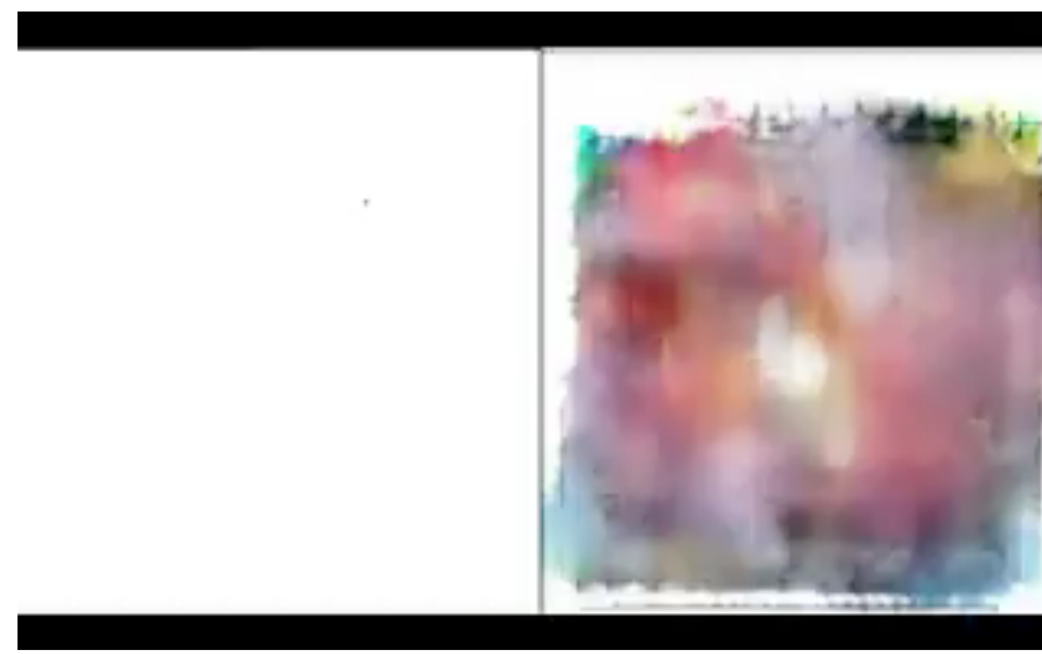
@ka92

Twitter-driven research: #pix2pix

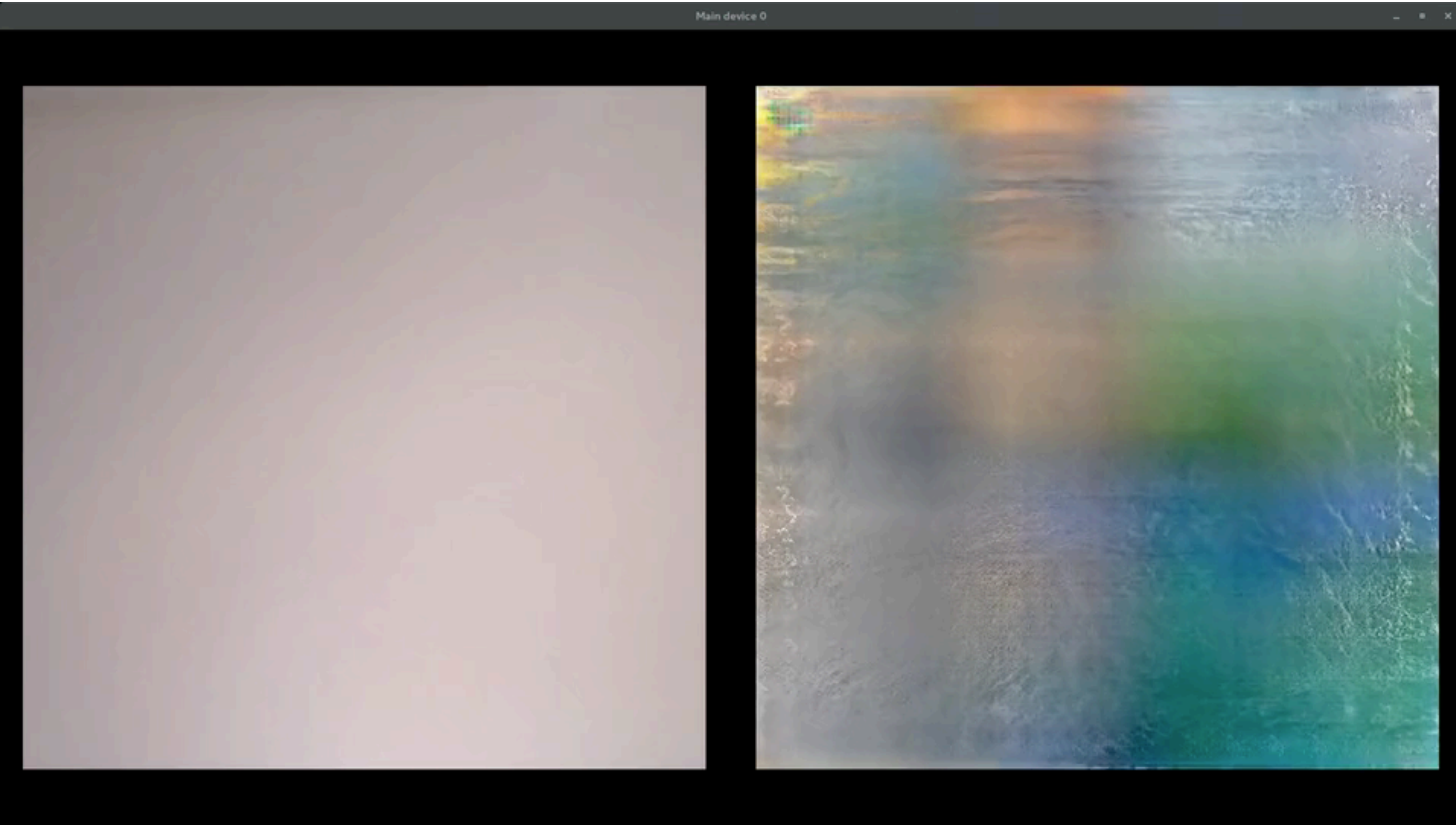


Brannon Dorsey @brannondorsey

Mario Klingemann @quasimondo

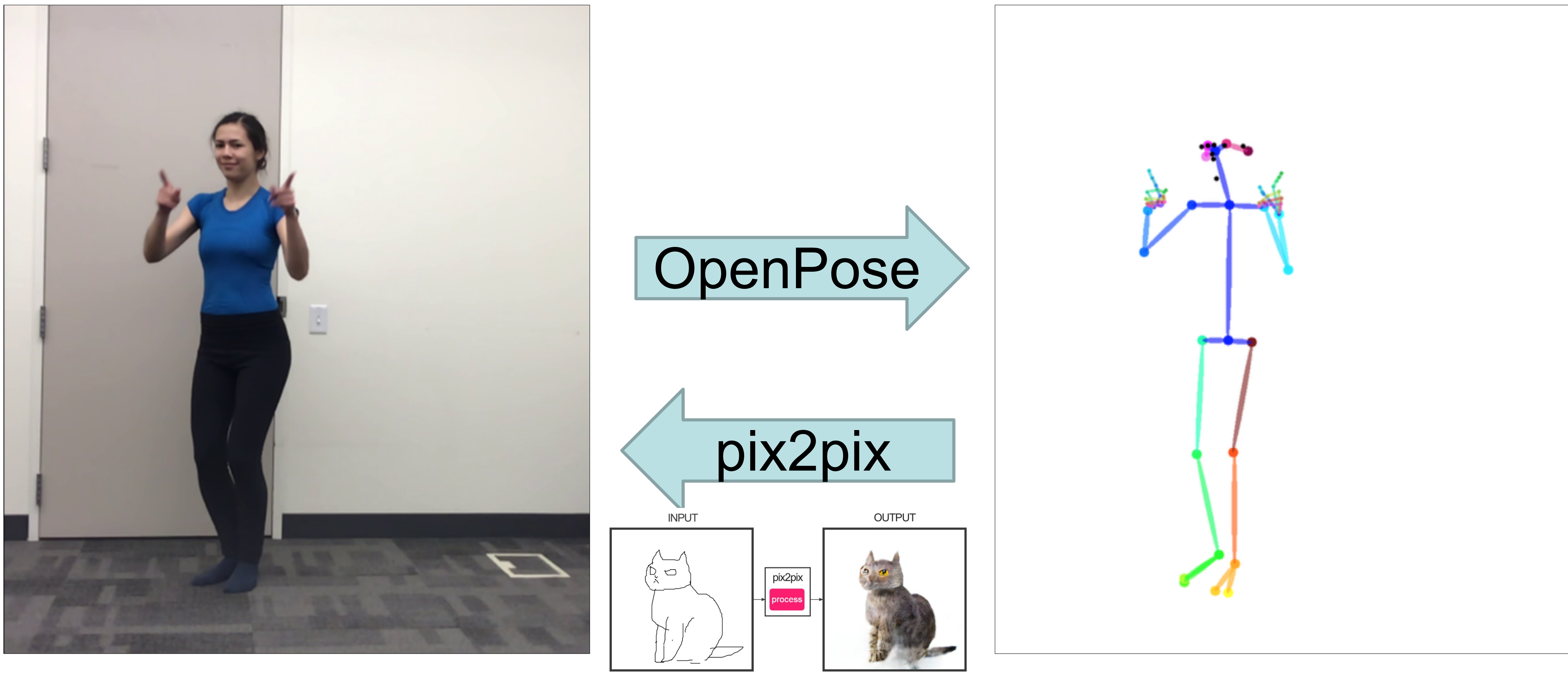


Bertrand Gondouin @bgondouin



© Memo Akten, “Learning to See: Gloomy Sunday”

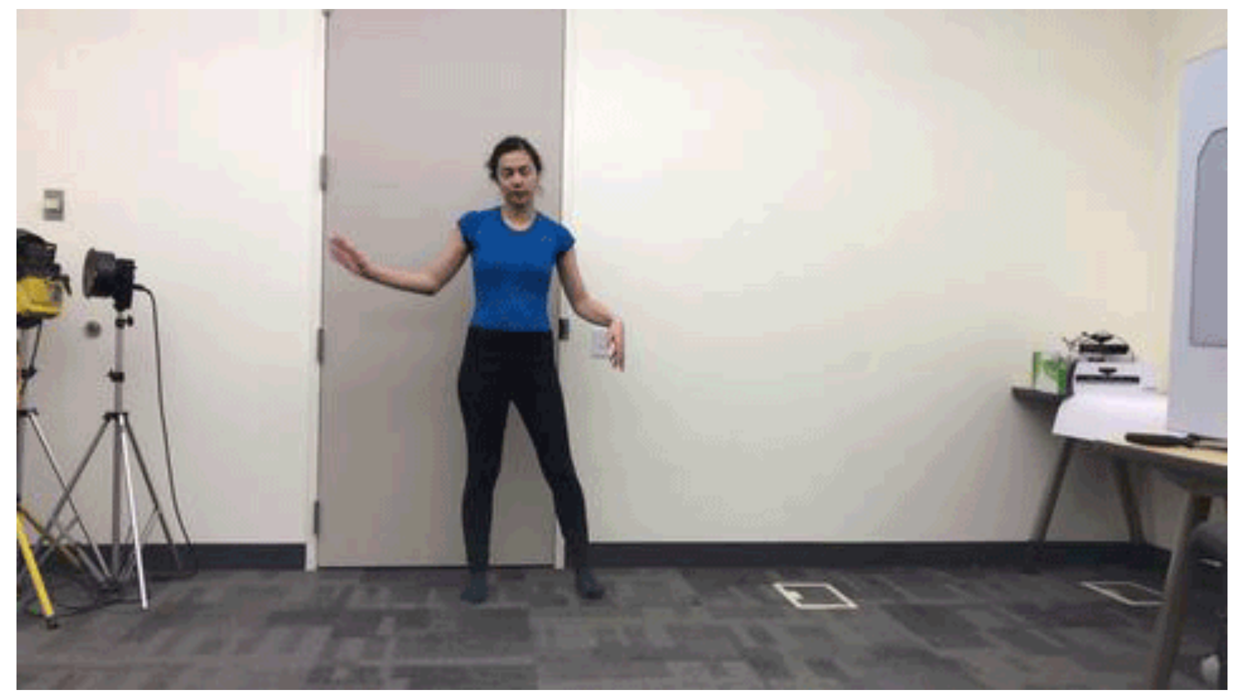
«Fais comme moi»



Everybody Dance Now

Caroline Chan, Shiry Ginosar, Tinghui Zhou, Alexei A. Efros
UC Berkeley

Source



Destination

Résultats

https://www.youtube.com/watch?v=PCBTZh41Ris&feature=youtu.be