

Modélisation, détection et classification d'objets urbains à partir d'images photographiques aériennes
Jérôme Pasquet

3 novembre 2016

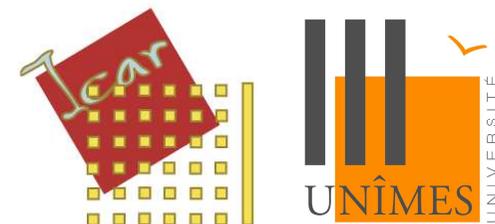
Modélisation, détection et classification d'objets urbains à partir d'images photographiques aériennes

03/11/2016

Jérôme Pasquet

Vincent Charvillat
Christophe Garcia
Matthieu Cord
Pascal Poncelet
Laurent Deruelle
Marc Chaumont
Gérard Subsol

Rapporteur
Rapporteur
Examineur
Examineur
Invité
Directeur
Co-encadrant



Motivations
●○○○○○○○○

Une méthode de référence
○○○○○○○○○○

Le réseau de SVM
○○○○○○○○○○○○

L'apprentissage profond
○○○○○○○○○○

Conclusions et perspectives
○○○○



Motivations
●○○○○○○○○○

Une méthode de référence
○○○○○○○○○○○

Le réseau de SVM
○○○○○○○○○○○○○

L'apprentissage profond
○○○○○○○○○○○

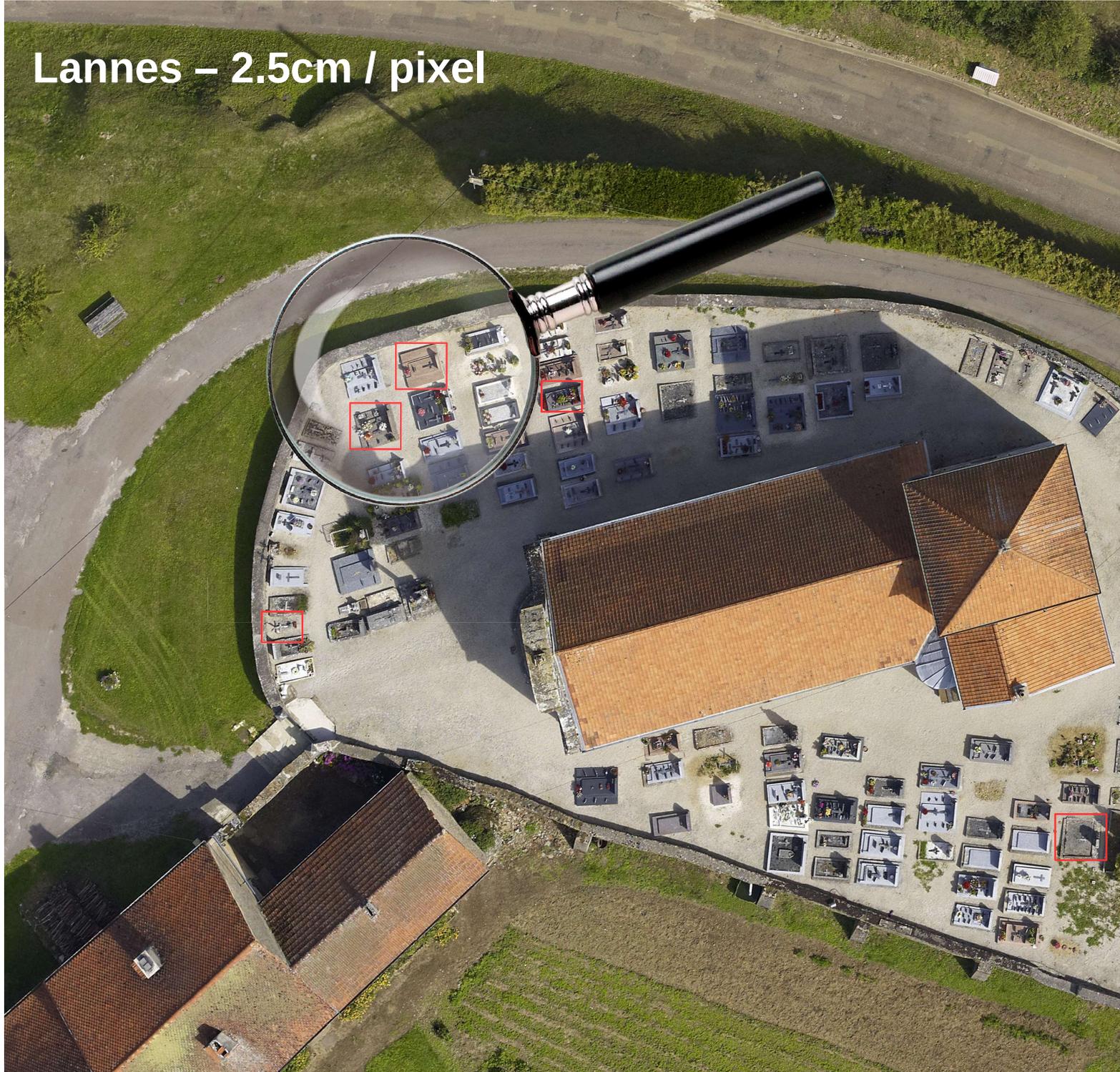
Conclusions et perspectives
○○○○



Lannes – 2.5cm / pixel



Lannes – 2.5cm / pixel



Exemples



Gandrage – 2.5cm / pixel



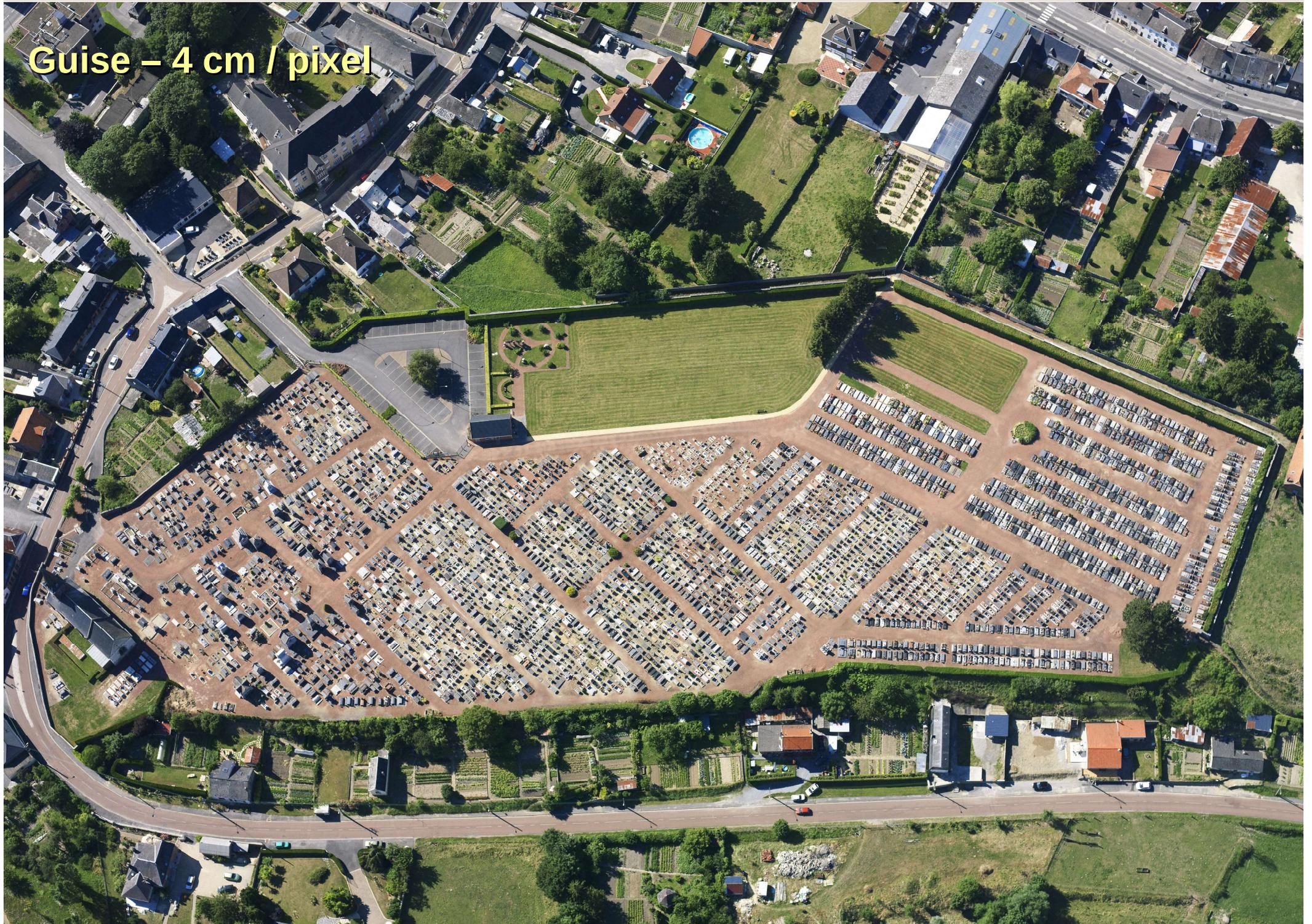
Gandrang – 2.5cm / pixel



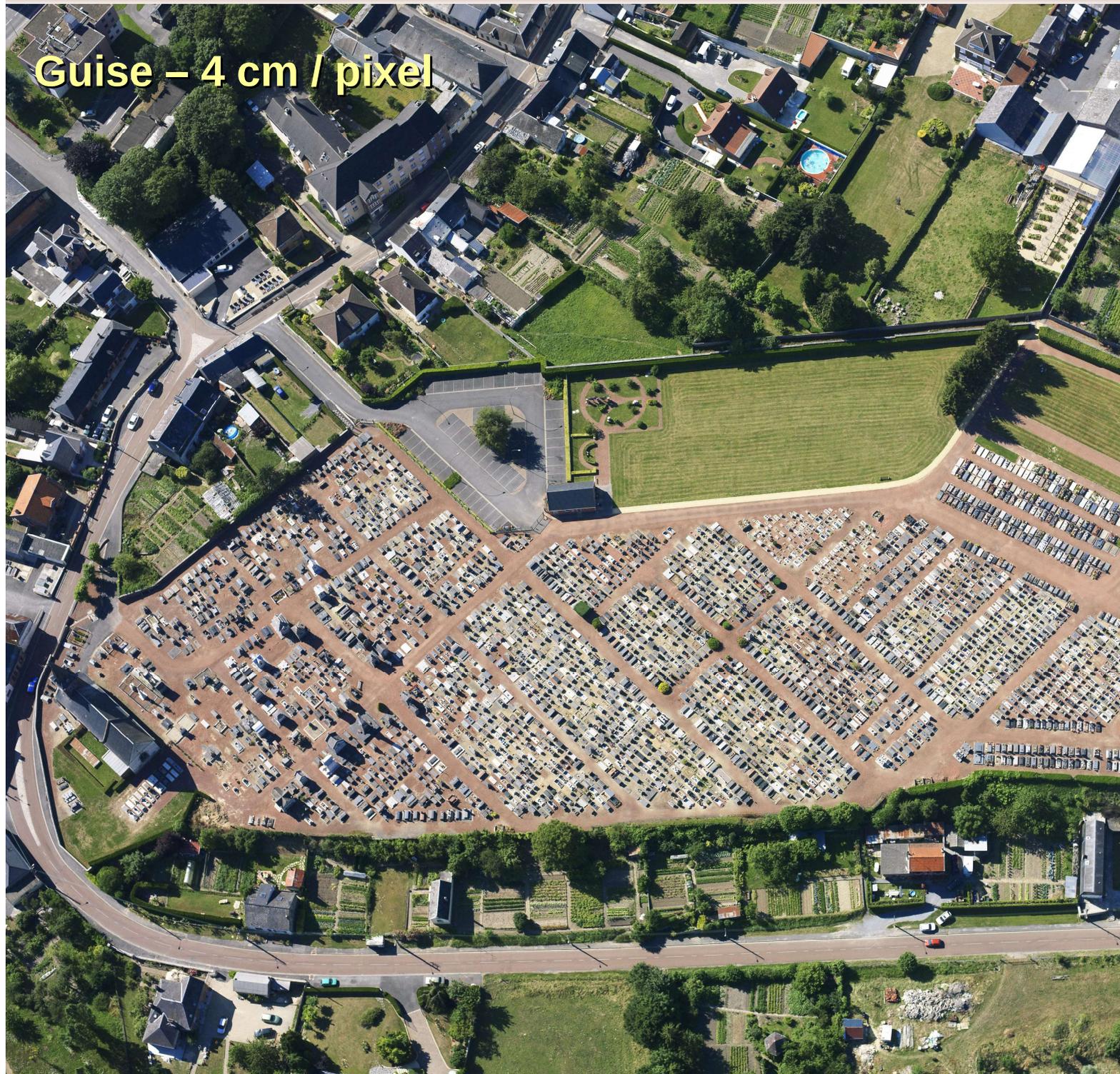
Exemples



Guise – 4 cm / pixel



Guise – 4 cm / pixel



Exemples



Objectif

- Localiser les tombes dans les images aériennes sous la forme de fenêtres englobantes.

Les données :

- 30 images couleur (24 bits)
- Très haute résolution de 2,5cm /pixel à 5 cm/pixel

Les objets urbains étudiés :

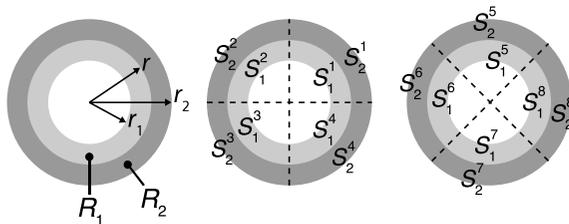
- variables en forme, texture et couleur
- ombres et occultations
- multiples dans une image
- collés les uns aux autres

Travaux existants

Méthode appliquée aux objets petits dans les images aériennes.

- Détection à partir d'un modèle paramétrique (Peeters et Etzion, 2010), (Myint et al., 2010), (Grigillo et Kanjir, 2012), (Bartoli et al., 2015)

Dans le cas de la détection de plaques d'égout :



Les tombes ont une trop grande variabilité pour ces approches !

- Peeters et Etzion - Automated recognition of urban objects for morphological attributes using GIS (ISPRS)
- Myint et al. - Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery (Remote Sensing of Environment)
- Grigillo et Kanjir - Urban object extraction from digital surface model and digital aerial images (ISPRS)
- Bartoli et al. - Manhole cover detection using a geometrical filter on very high resolution aerial and satellite images -(JURSE)

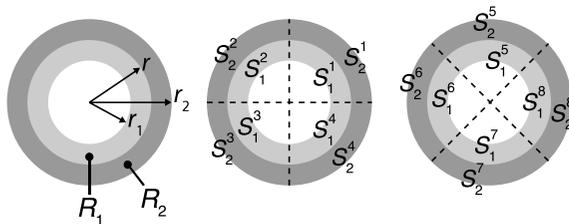
Utilisation de méthodes d'apprentissage automatique

Travaux existants

Méthode appliquée aux objets petits dans les images aériennes.

- Détection à partir d'un modèle paramétrique (Peeters et Etzion, 2010), (Myint et al., 2010), (Grigillo et Kanjir, 2012), (Bartoli et al., 2015)

Dans le cas de la détection de plaques d'égout :



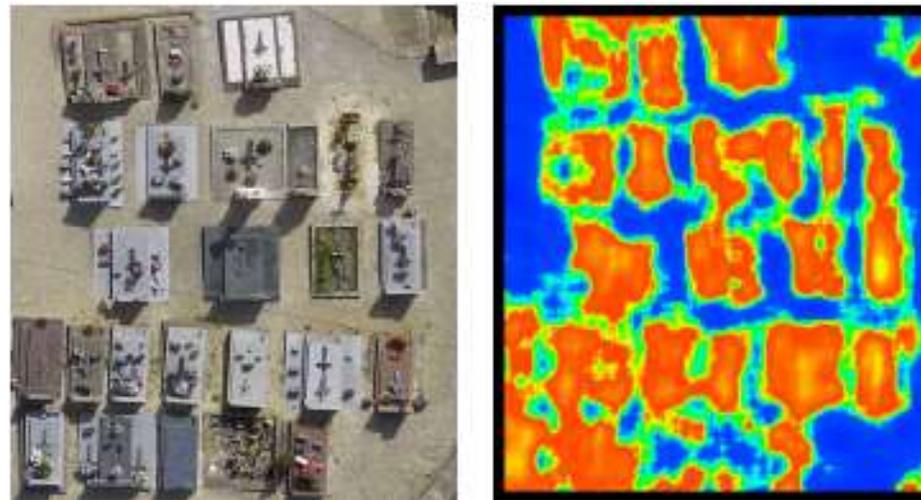
Les tombes ont une trop grande variabilité pour ces approches !

- Peeters et Etzion - Automated recognition of urban objects for morphological attributes using GIS (ISPRS)
- Myint et al. - Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery (Remote Sensing of Environment)
- Grigillo et Kanjir - Urban object extraction from digital surface model and digital aerial images (ISPRS)
- Bartoli et al. - Manhole cover detection using a geometrical filter on very high resolution aerial and satellite images -(JURSE)

Utilisation de méthodes d'apprentissage automatique

Travaux existants

- Détection à partir d'un modèle paramétrique
- Apprentissage et classification pixel (Chaumont et al., 2013), (Salah, 2014), (Aldavert et al., 2010), (Yang et al., 2009)

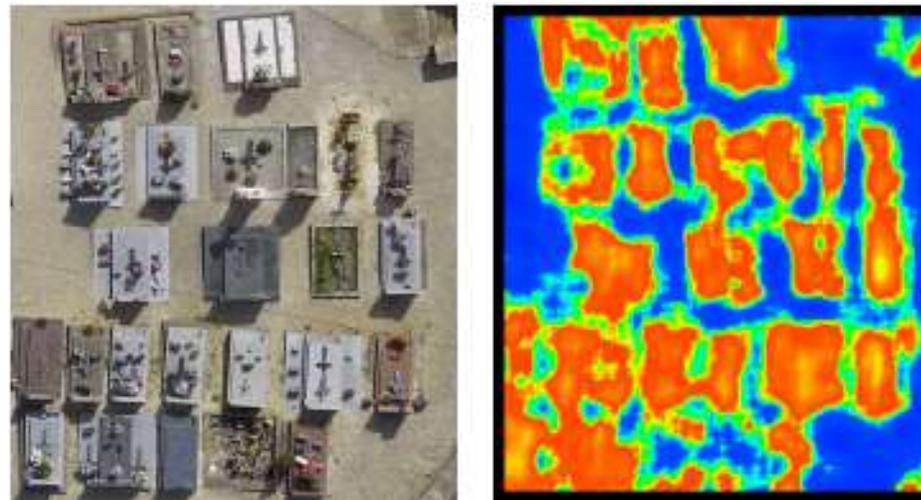


- Chaumont et al. - Automatic localization of tombs in aerial imagery : application to the digital archiving of cemetery heritage (Digital Heritage - International Congress)
- Salah - Combining pixel-based and object-oriented support vector machines using bayesian probability theory - (ISPRS)
- Aldavert et al. - Fast and Robust Object Segmentation with the Integral Linear Classifier - (CVPR)
- Yang et al. - Tree detection from aerial imagery - (International Conference on Advances in Geographic Information Systems)

Les tombes collées ne peuvent être différenciées

Travaux existants

- Détection à partir d'un modèle paramétrique
- Apprentissage et classification pixel (Chaumont et al., 2013), (Salah, 2014), (Aldavert et al., 2010), (Yang et al., 2009)



- Chaumont et al. - Automatic localization of tombs in aerial imagery : application to the digital archiving of cemetery heritage (Digital Heritage - International Congress)
- Salah - Combining pixel-based and object-oriented support vector machines using bayesian probability theory - (ISPRS)
- Aldavert et al. - Fast and Robust Object Segmentation with the Integral Linear Classifier - (CVPR)
- Yang et al. - Tree detection from aerial imagery - (International Conference on Advances in Geographic Information Systems)

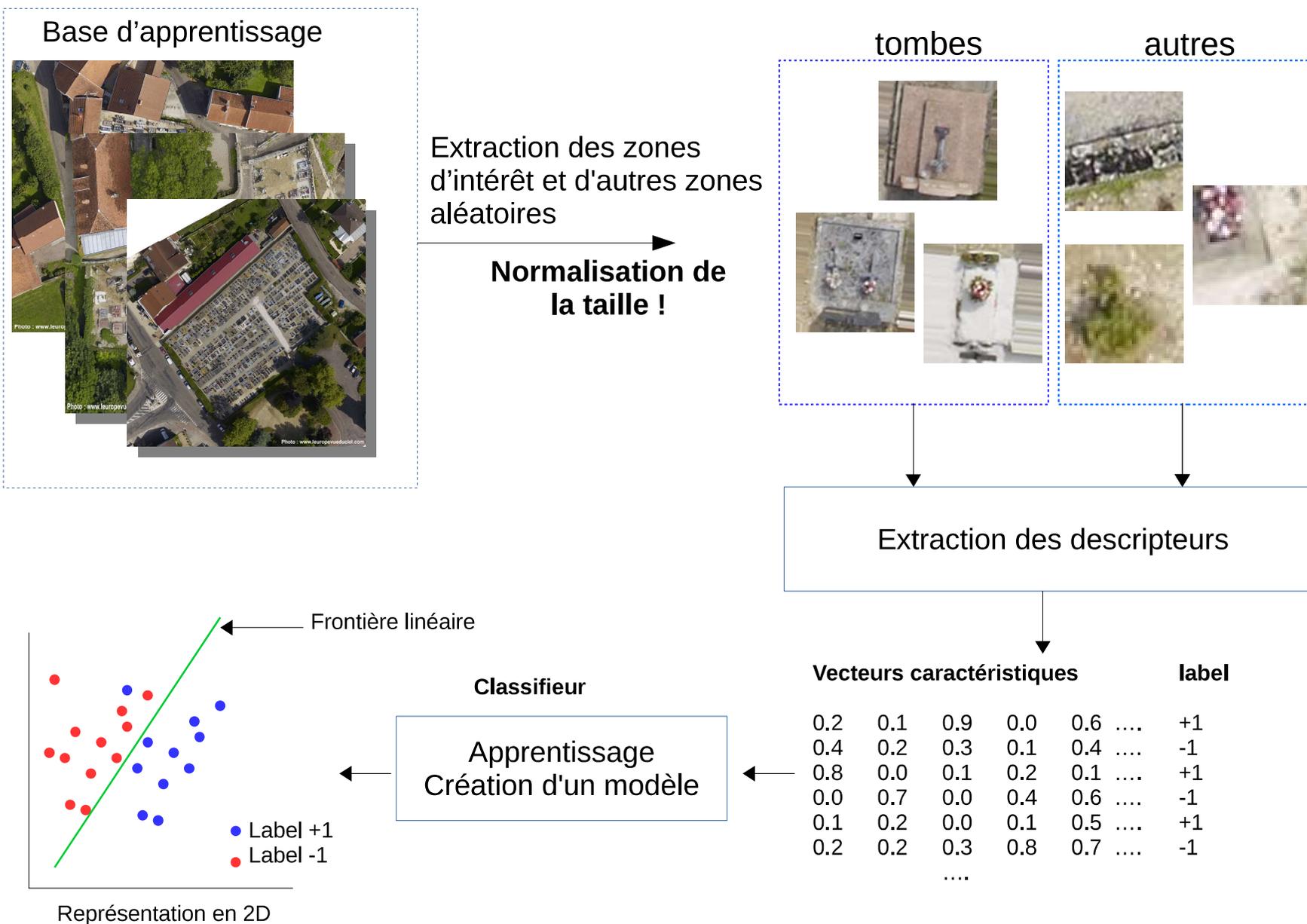
Les tombes collées ne peuvent être différenciées

Sommaire

- ① Motivations
- ② Méthode de référence
 - Méthodologie générale d'apprentissage
 - Méthodologie générale d'évaluation
 - Descripteur et classifieur utilisés
 - Évaluation
- ③ Proposition du réseau de SVM
 - Problème de fusion des résolutions
 - Proposition du réseau de SVM
 - La gestion de la couleur
 - Accélération du réseau
- ④ Combinaison d'un CNN et d'un réseau de SVM
 - Rappels sur les CNN
 - Le CNN utilisé et sa base d'apprentissage
 - Évaluation de l'approche
- ⑤ Conclusions et perspectives

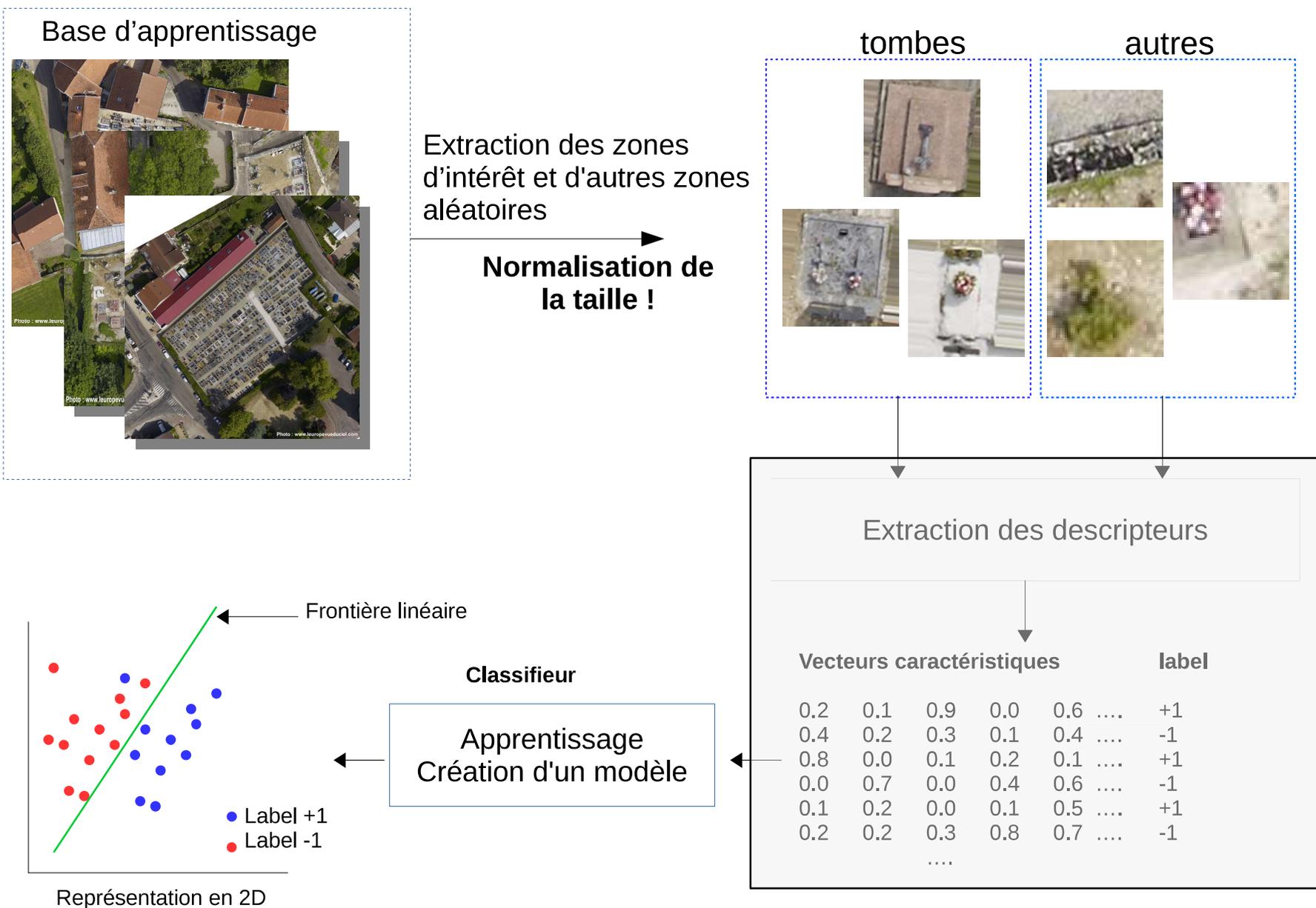
Méthodologie générale de la segmentation objet

Phase d'apprentissage



Méthodologie générale de la segmentation objet

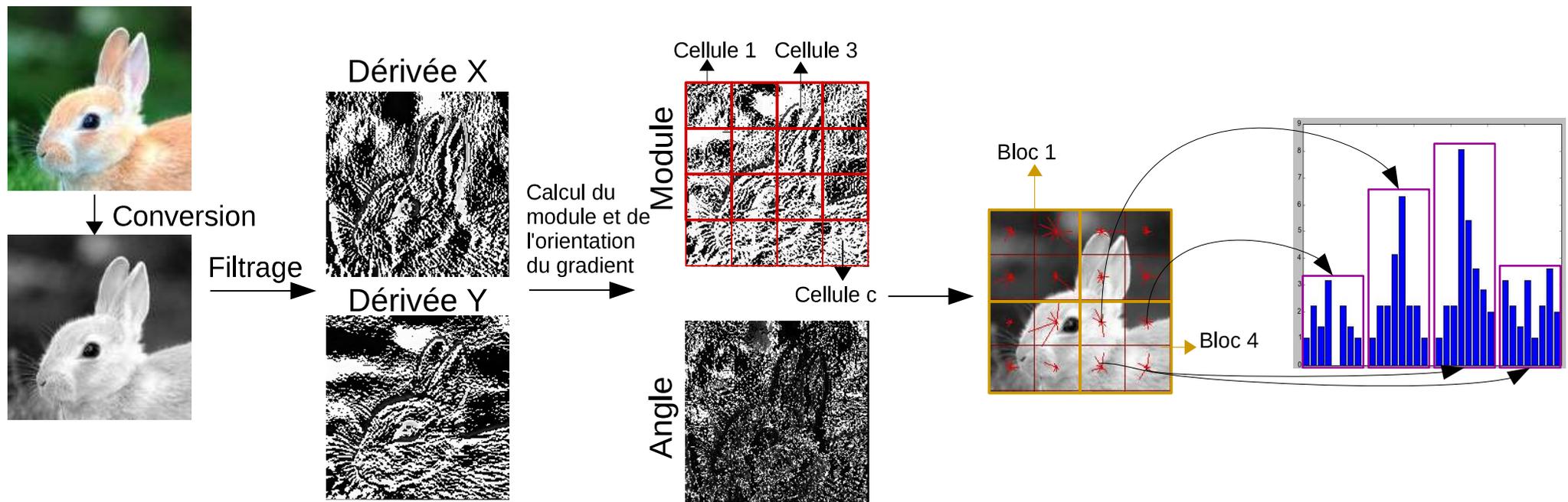
Phase d'apprentissage



Descripteur utilisé

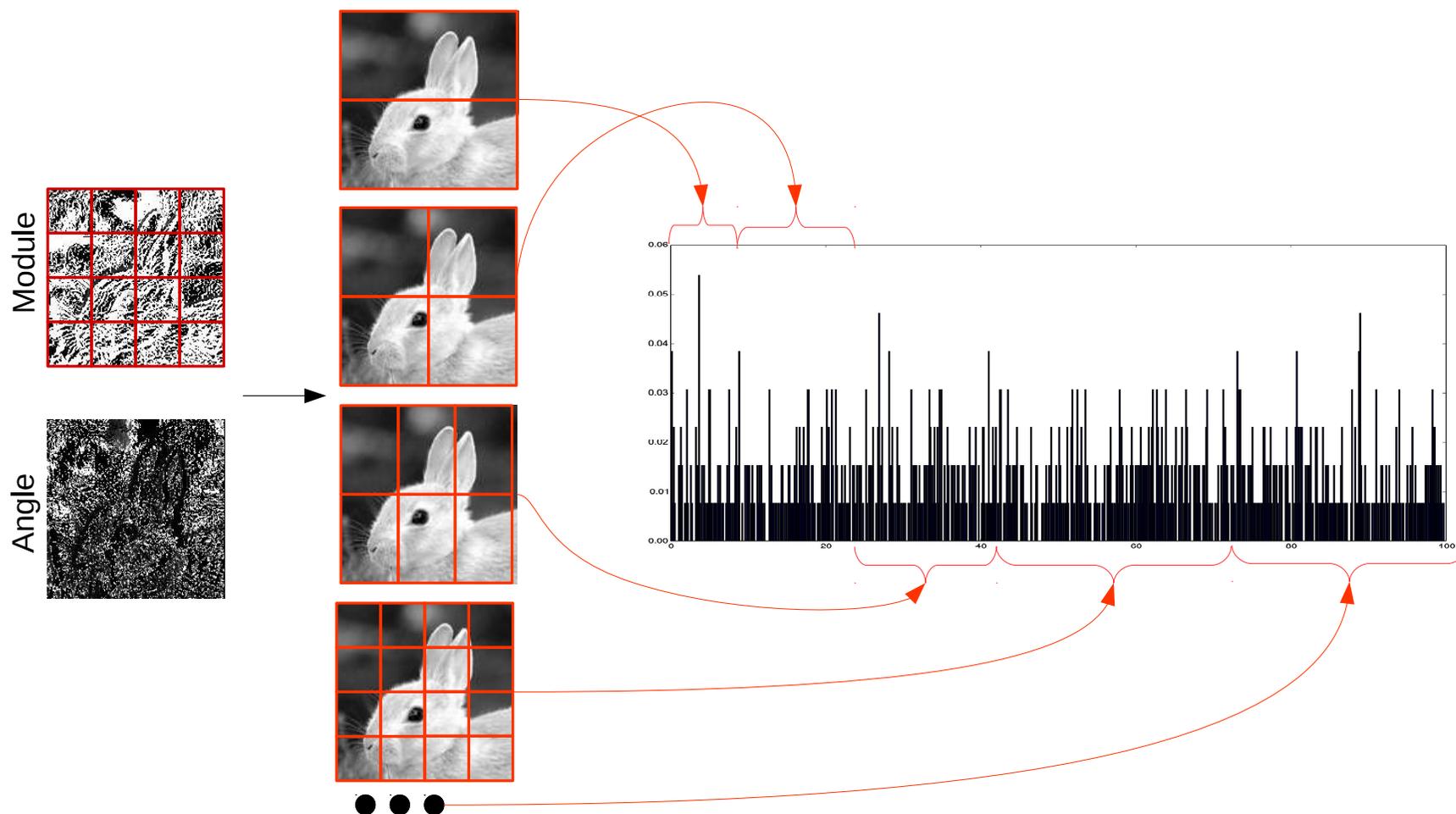
- Un descripteur rapide à extraire
- Permettant la description de forme

⇒ Utilisation des **histogrammes de gradient orienté (HOG)** (Dalal et Triggs, 2005) : calcul de la distribution des orientations dans une zone en niveau de gris.

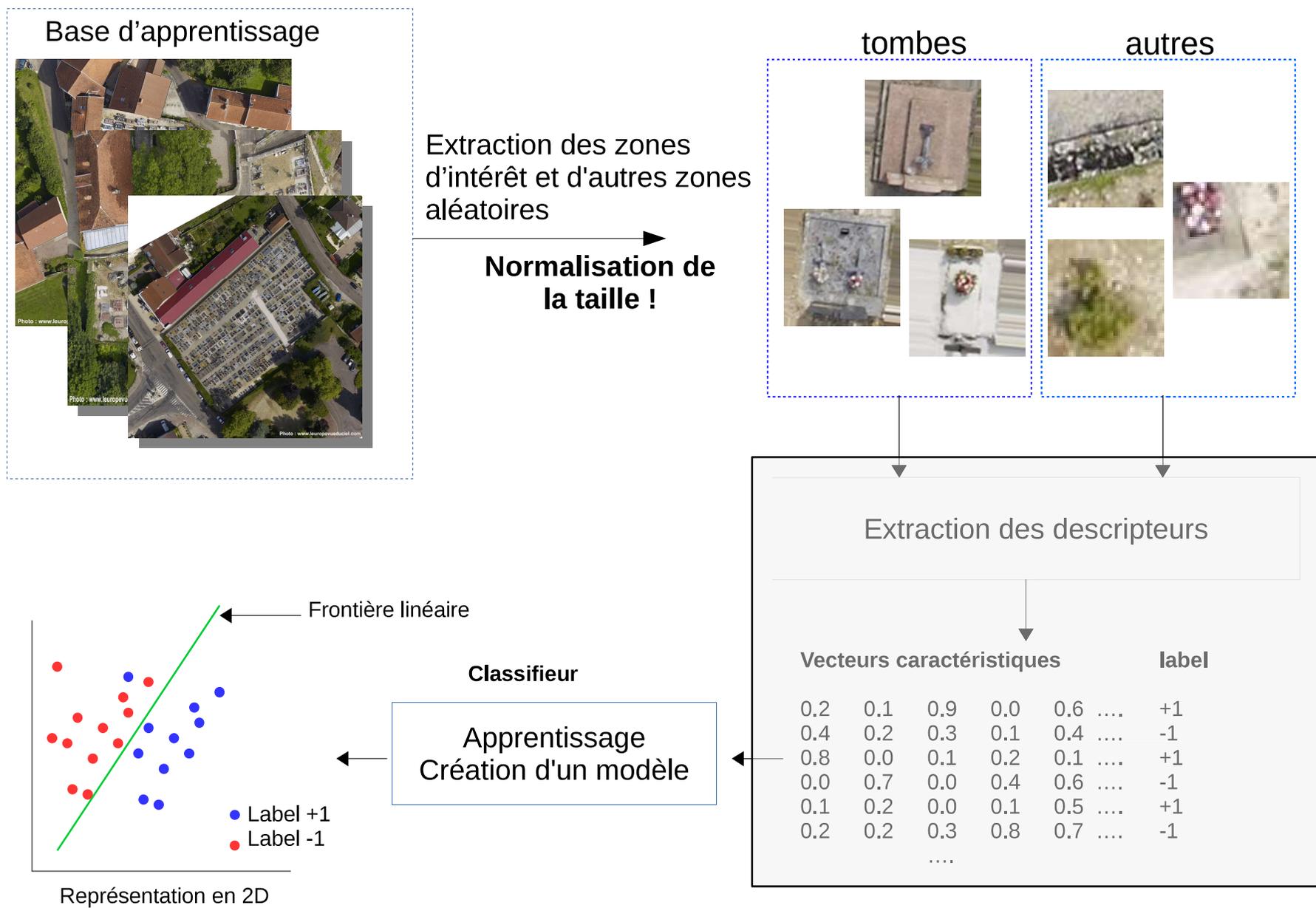


Descripteur utilisé

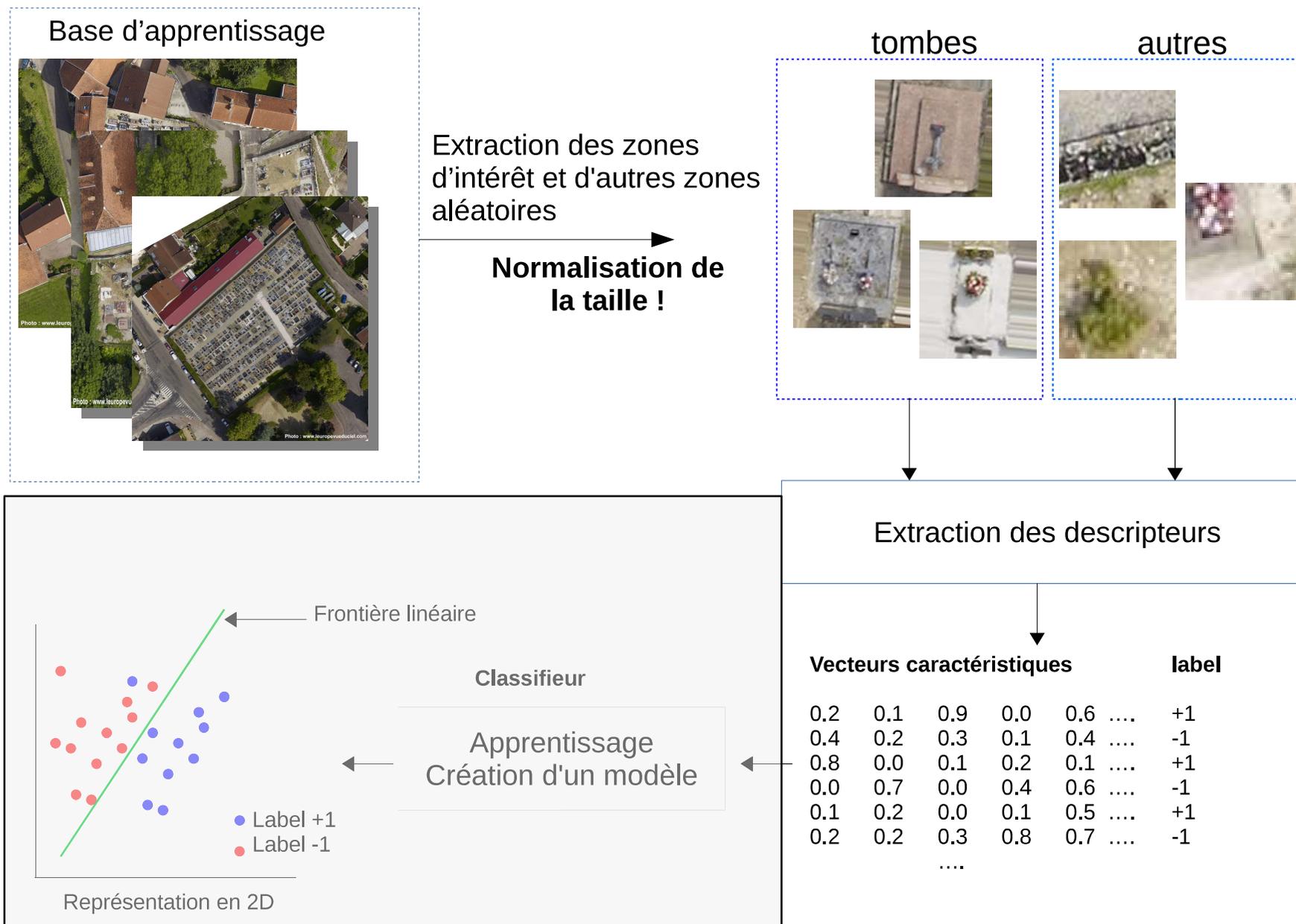
⇒ Extension aux histogrammes de gradient orienté multi-résolutions (Zhu et al., 2006)



Méthodologie générale de la segmentation objet



Méthodologie générale de la segmentation objet

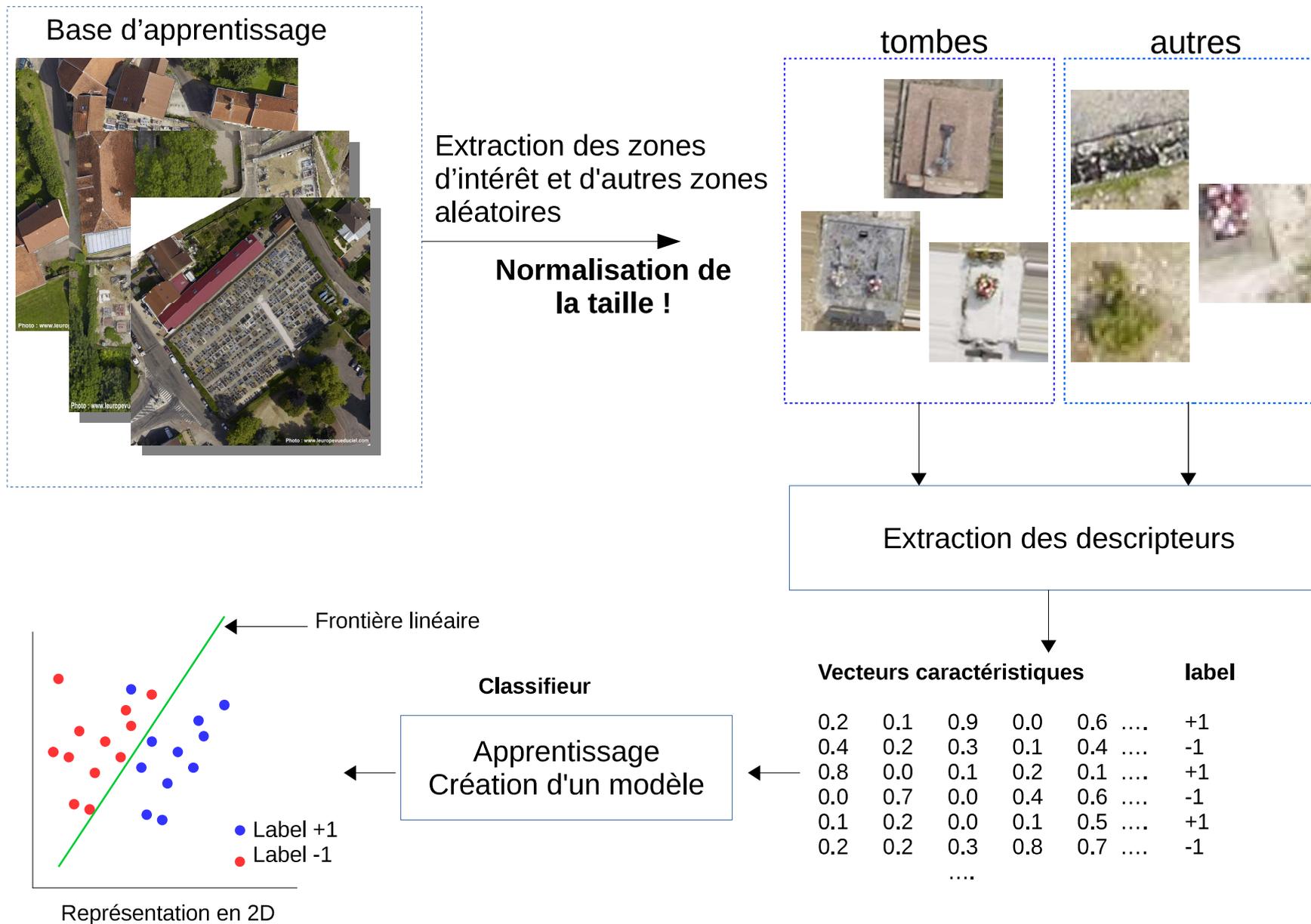


Classifieur utilisé

⇒ **Utilisation du SVM** (Chang et al. 2011)

- Traitement de descripteurs de grandes dimensions
- Aucune supposition a priori (non paramétrique)
- Robustesse aux données aberrantes (et sur-apprentissage)
- Fonctionnel avec peu d'échantillons (voir moins d'échantillons que la dimension du descripteur)
- Utilisable dans des cas complexes (non linéaires)

Méthodologie générale de la segmentation objet



Méthodologie générale de la segmentation objet

Phase de tests

Fenêtre glissante



Extraction du vecteur caractéristiques

0.2 0.4 0.2 0.0 ...

classification

Label -1

Modèle linéaire

Représentation du vecteur caractéristiques

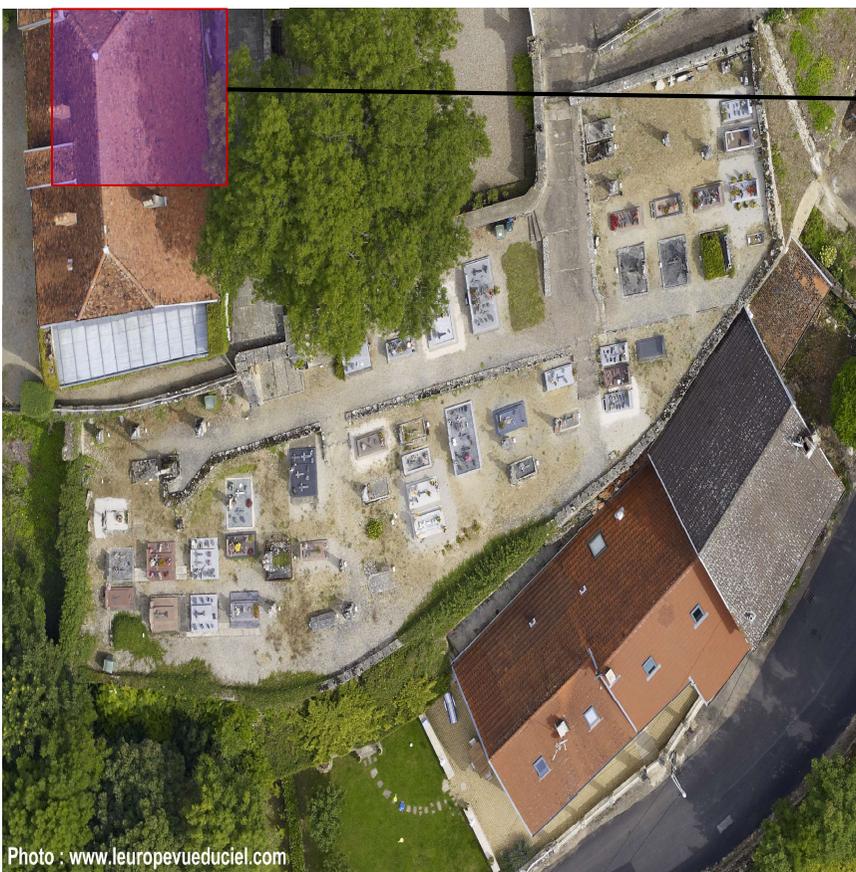
Label +1

Représentation en 2D

Label -1 ⇒ Pas un objet d'intérêt

Méthodologie générale de la segmentation objet

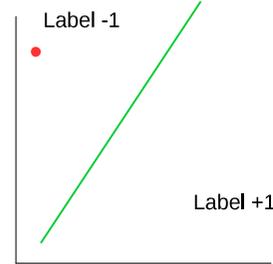
Phase de tests



Extraction du vecteur caractéristiques

0.2 0.1 0.3 0.2 ...

classification



Représentation en 2D

Label -1 ⇒ Pas un objet d'intérêt

Méthodologie générale de la segmentation objet

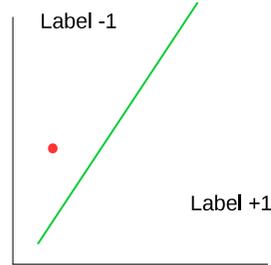
Phase de tests



Extraction du vecteur caractéristiques

0.1 0.1 0.4 0.4 ...

classification



Représentation en 2D

Label -1 ⇒ Pas un objet d'intérêt

Méthodologie générale de la segmentation objet

Phase de tests



Changement de résolution



Méthodologie générale de la segmentation objet

Phase de tests



Le processus reprend sur une autre résolution

Méthodologie générale de la segmentation objet

Phase de tests



On effectue les tests à chaque position

Méthodologie générale de la segmentation objet

Phase de tests



Classification label +1
→ **Une tombe est trouvée**

Méthodologie générale de la segmentation objet

Phase de tests



Changement
de résolution

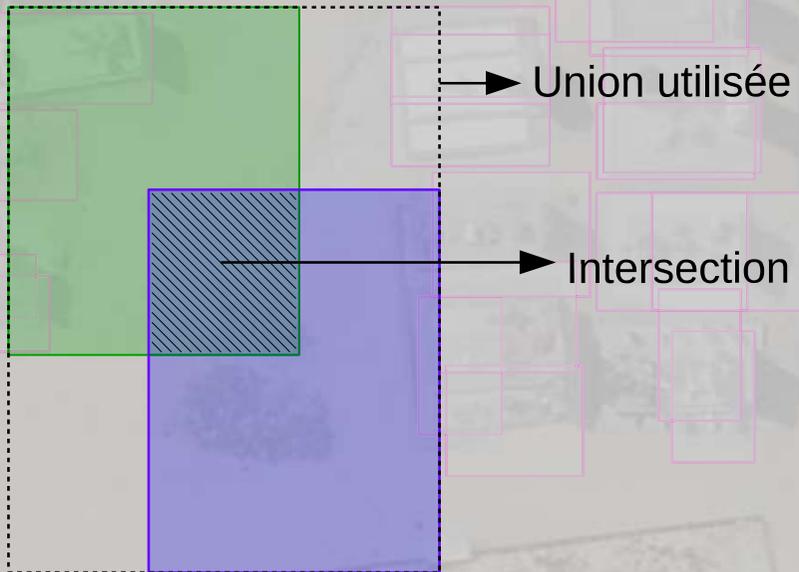


...

On ré-itére le processus sur toutes les résolutions



Règle de fusion des boxes



$$a = \frac{\text{Aire}(\text{ } \cap \text{ })}{\text{Aire}(\text{ } \cup \text{ })}$$

$$\text{Aire}(\text{ } \cup \text{ })$$

→ La fenêtre associée au plus haut taux de détectabilité est conservée



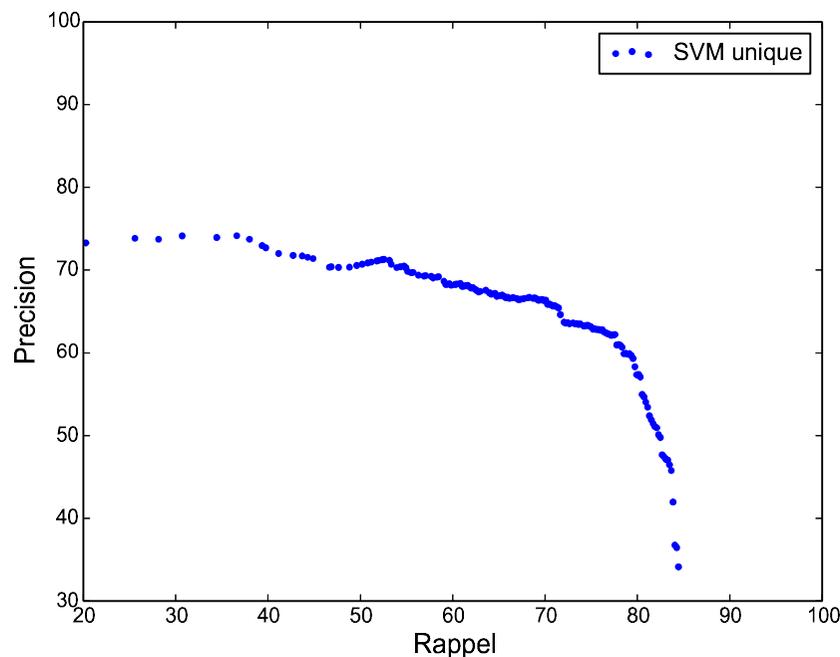
Résultats de référence

Descripteur

- Un descripteur HOG
- 81 résolutions
- grande dimension > 34000

Classifieur

- SVM (RBF)
- Paramètres déterminés par validation

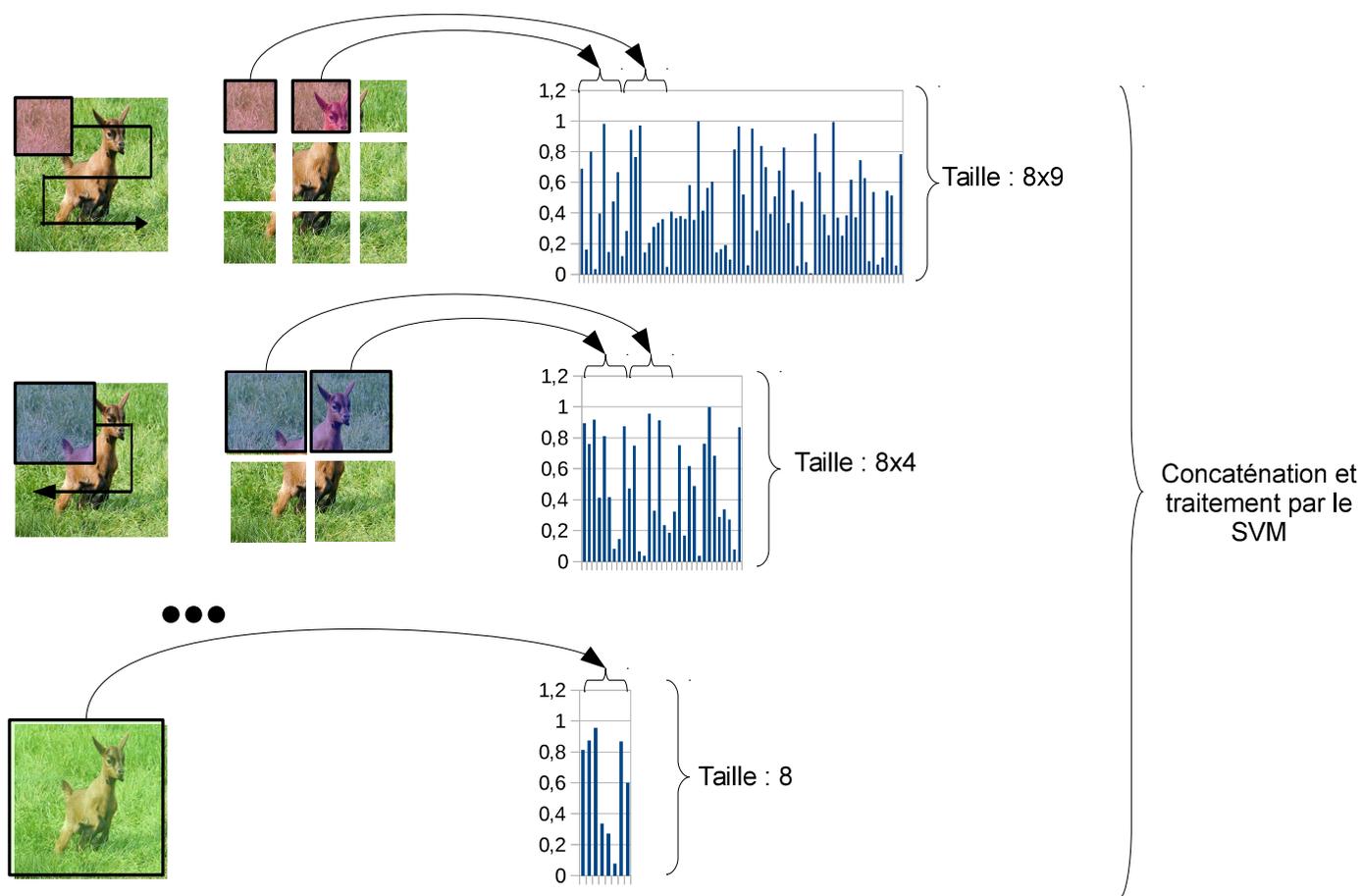


Le paramètre variant est le seuil de détection.

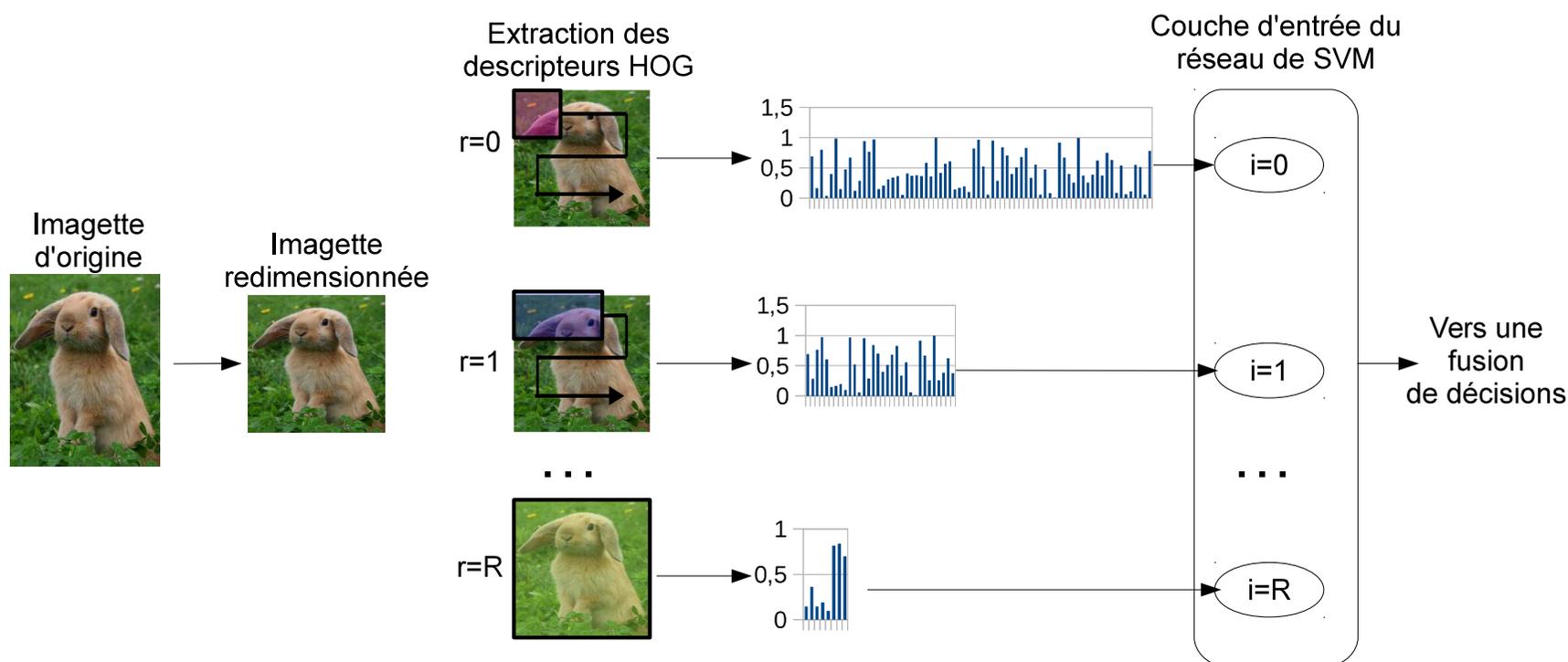
Sommaire

- ① Motivations
- ② Méthode de référence
 - Méthodologie générale d'apprentissage
 - Méthodologie générale d'évaluation
 - Descripteur et classifieur utilisés
 - Évaluation
- ③ **Proposition du réseau de SVM**
 - Problème de fusion des résolutions
 - Proposition du réseau de SVM
 - La gestion de la couleur
 - Accélération du réseau
- ④ **Combinaison d'un CNN et d'un réseau de SVM**
 - Rappels sur les CNN
 - Le CNN utilisé et sa base d'apprentissage
 - Évaluation de l'approche
- ⑤ **Conclusions et perspectives**

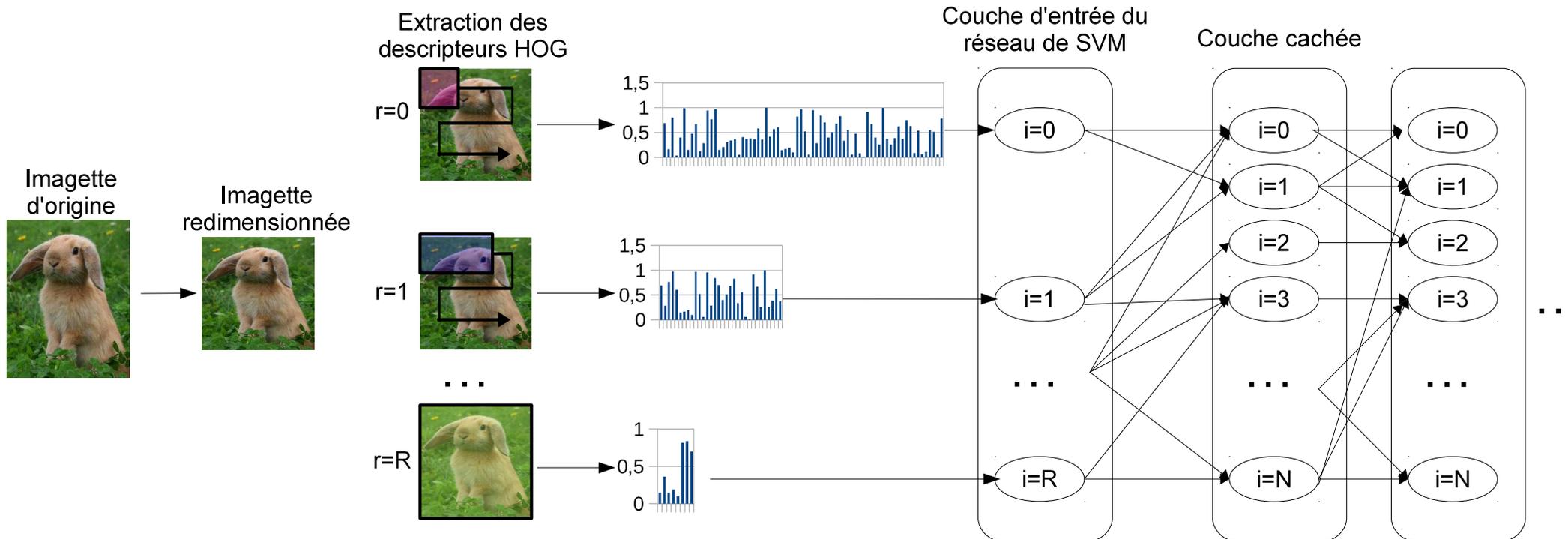
Le problème de la transmission des résolutions HOG au SVM



Un réseau de SVM



Un réseau de SVM



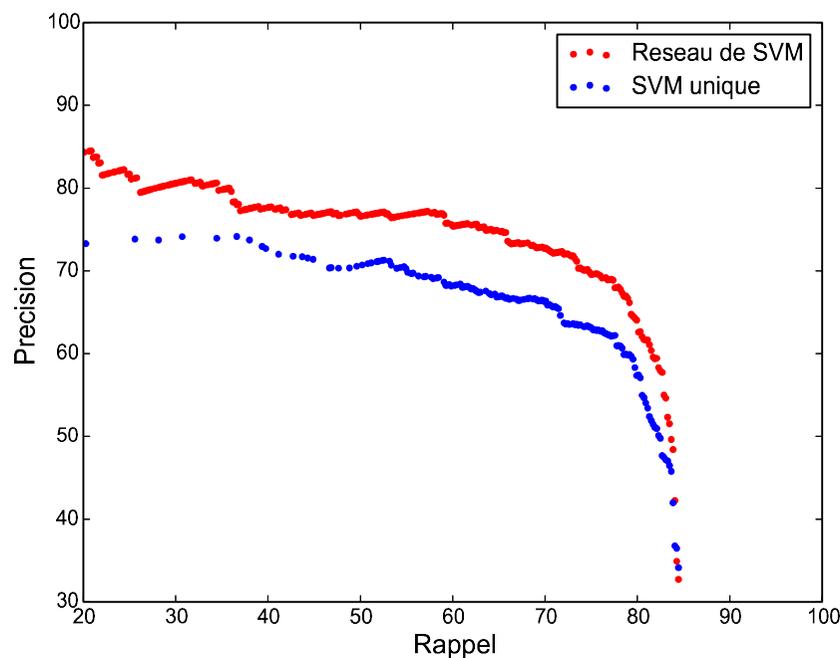
Résultats du réseau de SVM

Descripteur

- Un descripteur HOG
- 81 résolutions

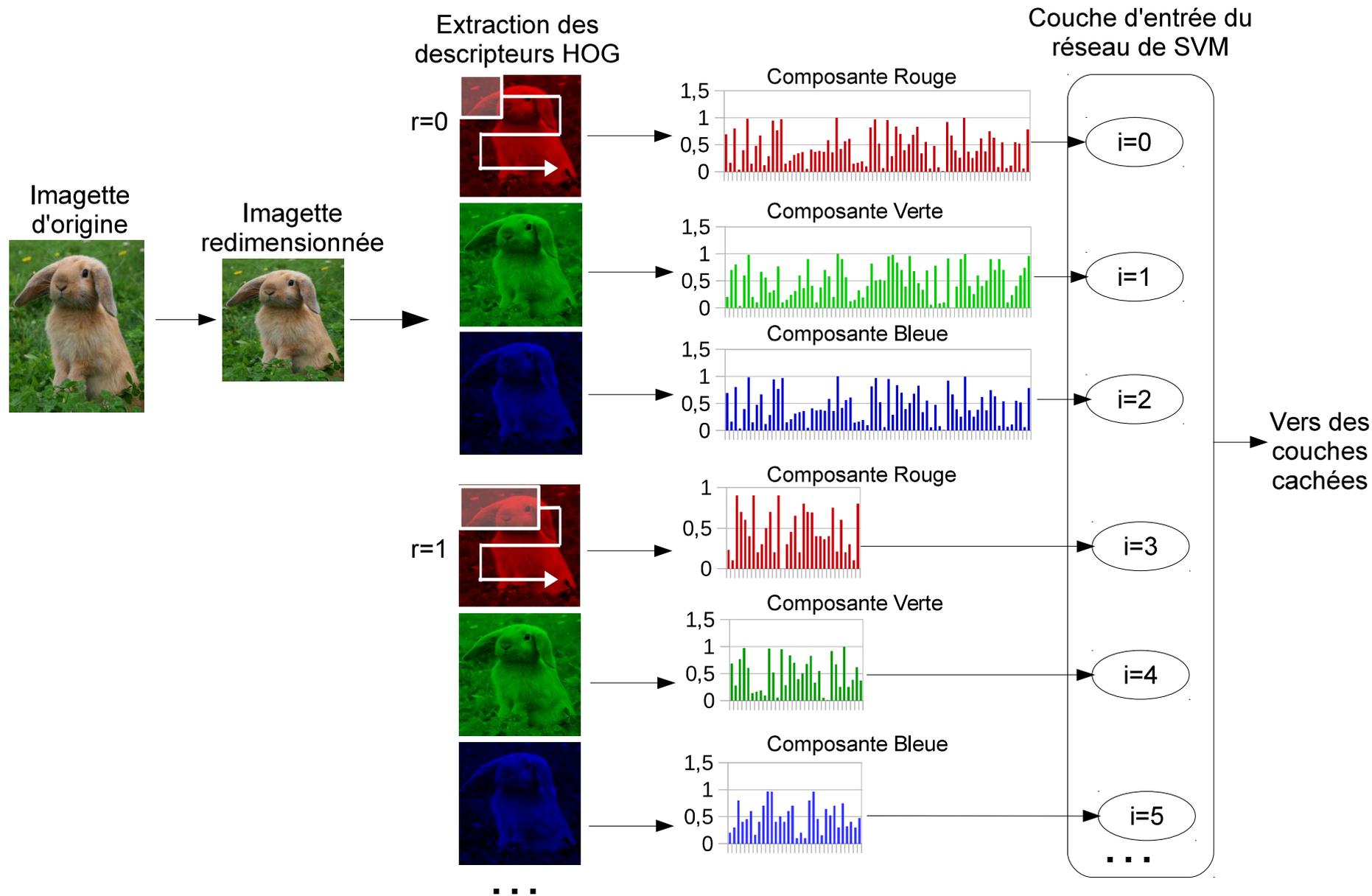
Classifieur

- SVM (linéaire)
- Deux couches de 300 SVM



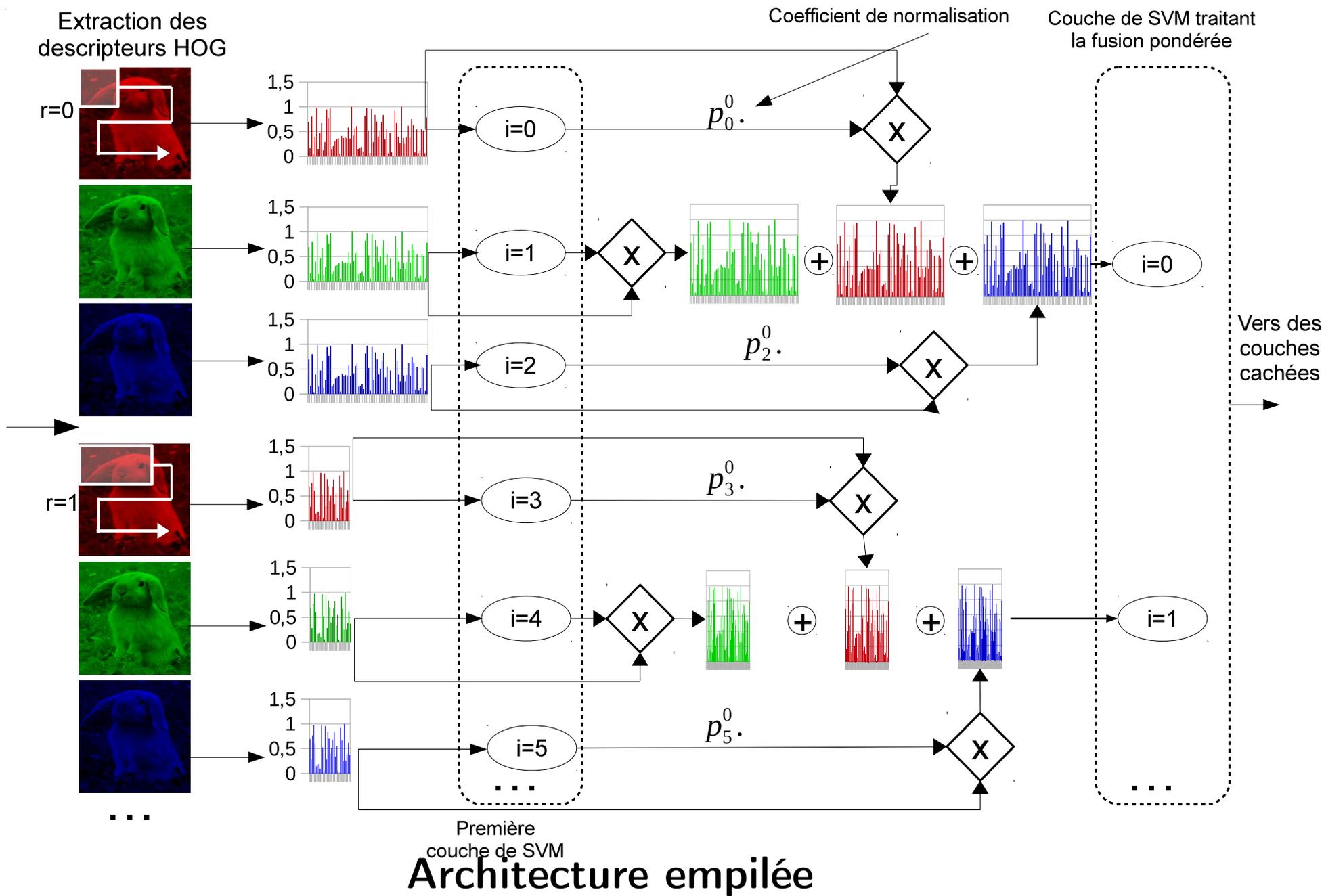
Le réseau de SVM a un gain d'environ 15%.

Application à la gestion couleur

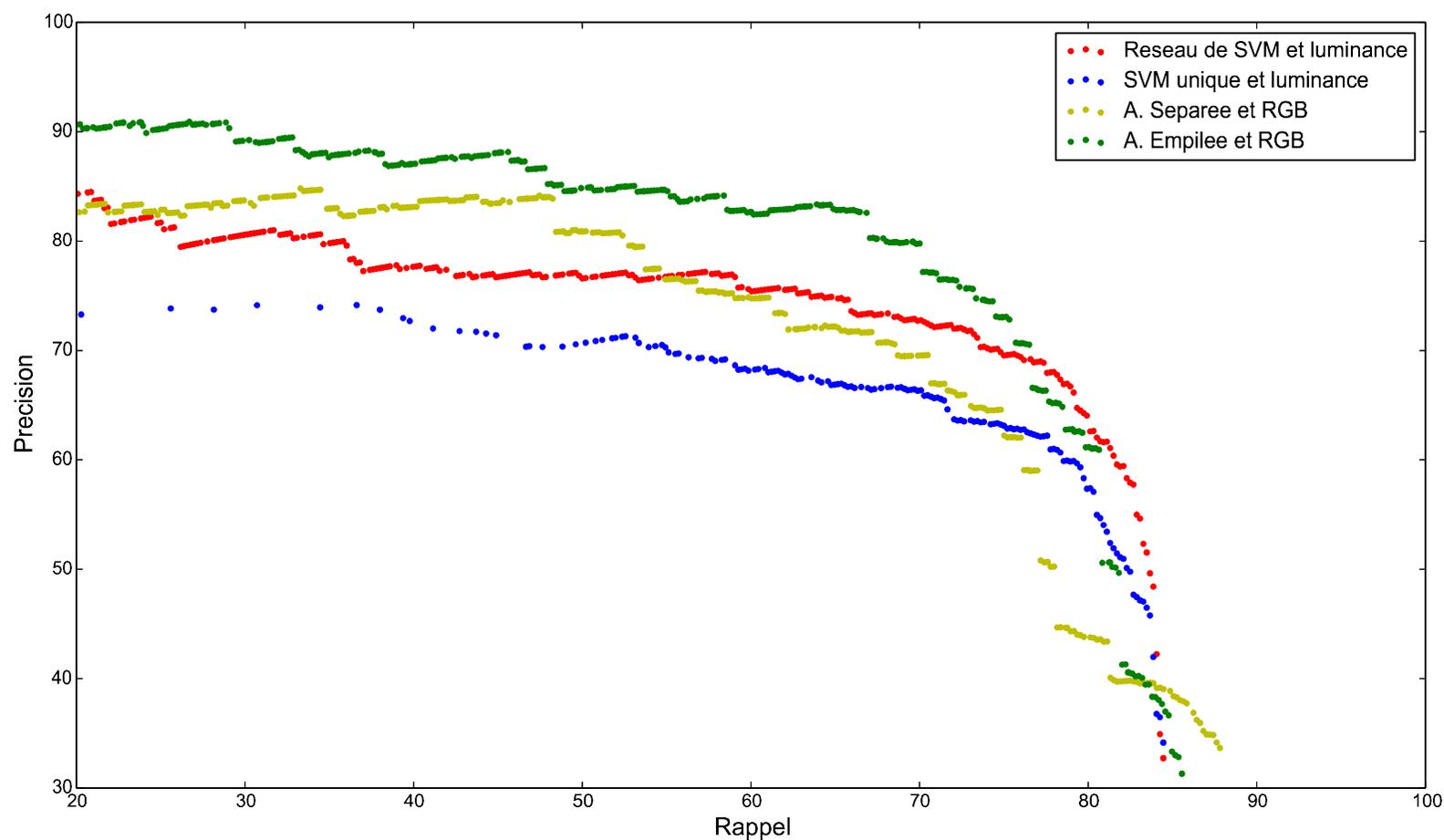


Architecture séparée

Structure de gestion de la couleur



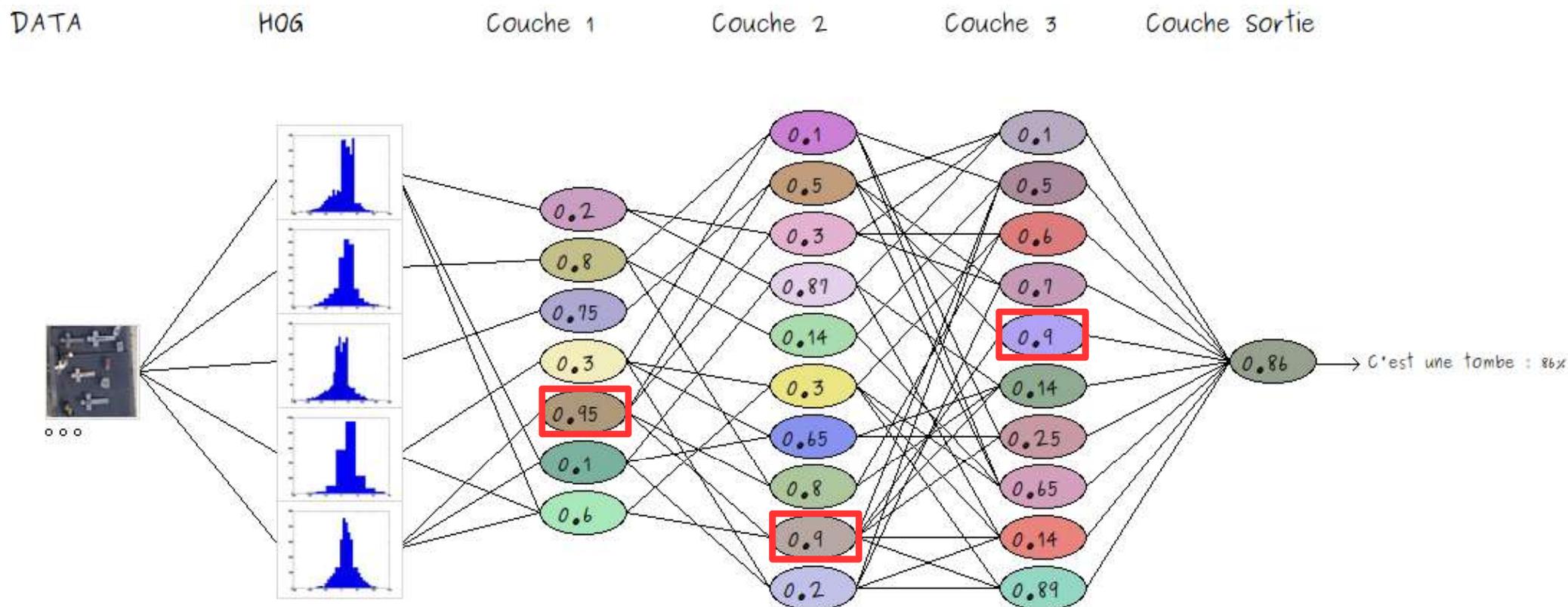
Résultats



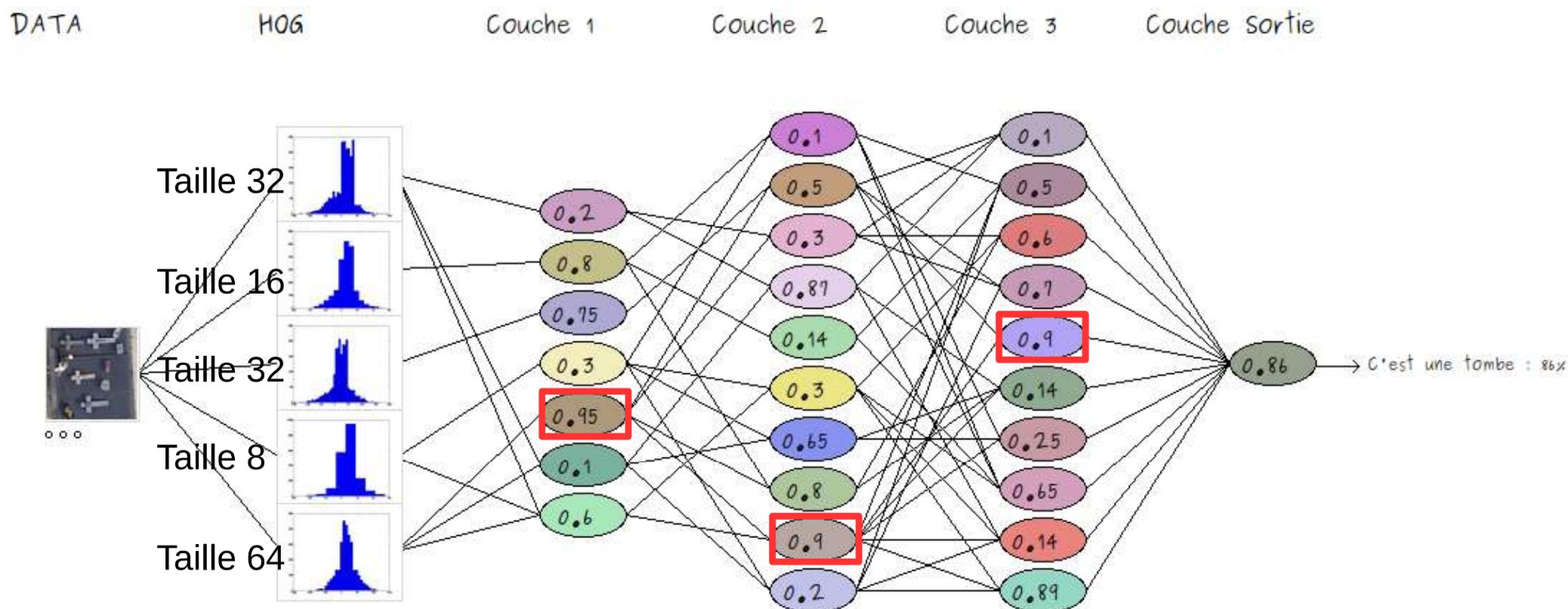
- La structure du réseau permet un meilleur traitement de la couleur
- le coût calculatoire augmente de plus de 35%

Exemple d'activation d'un réseau SVM

Analyse des SVM avec des scores élevés

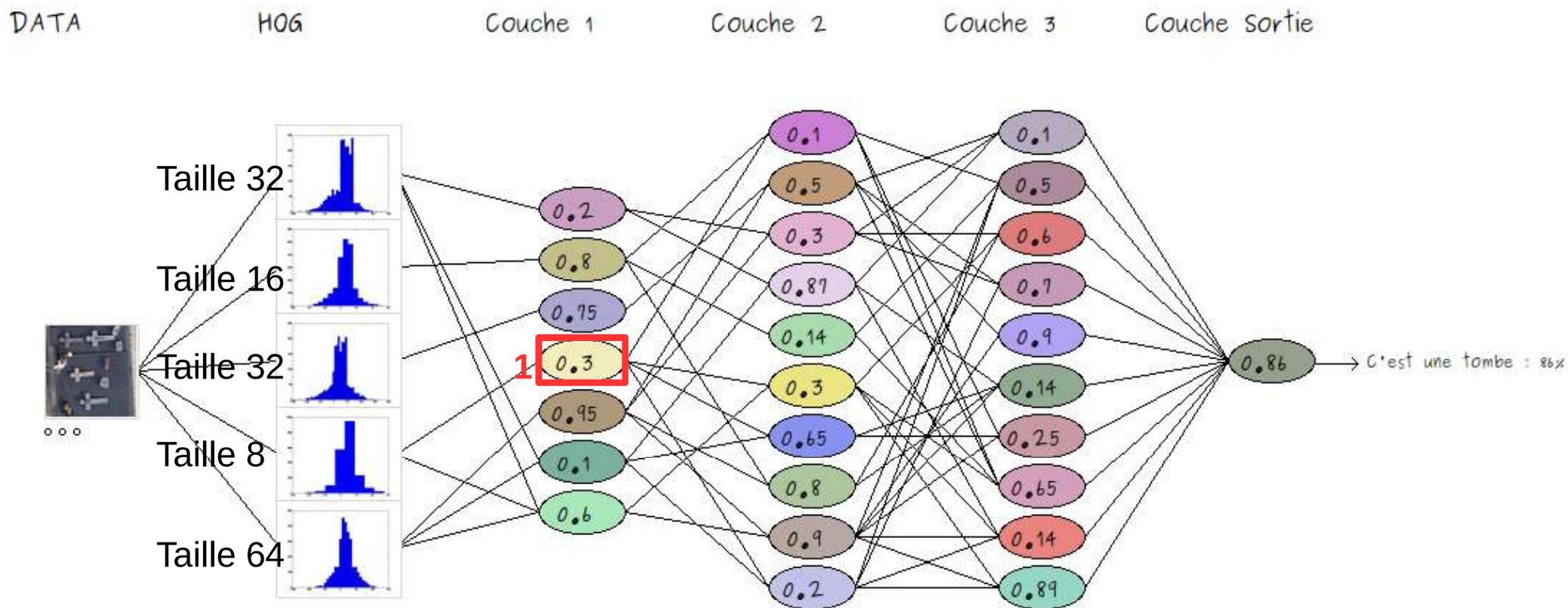


Coût d'activation d'un SVM

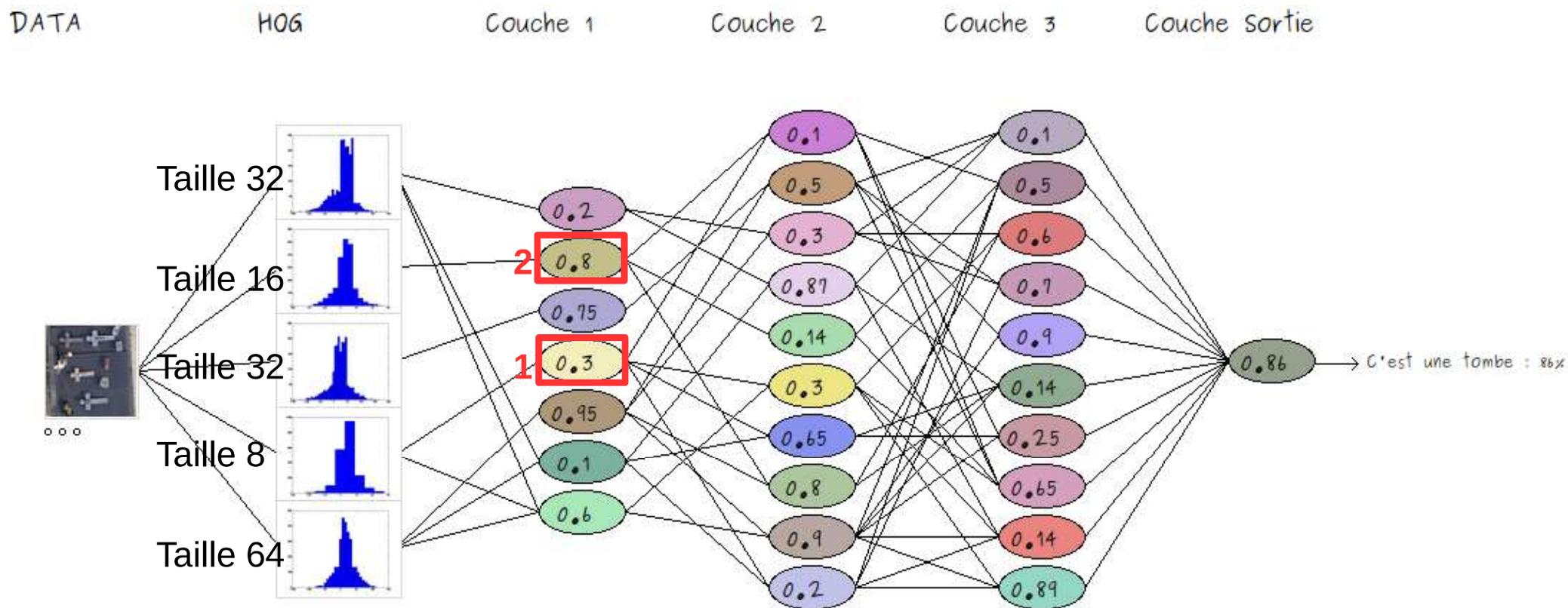


Le coût d'activation et le nombre d'opérations nécessaire au SVM pour être activé.

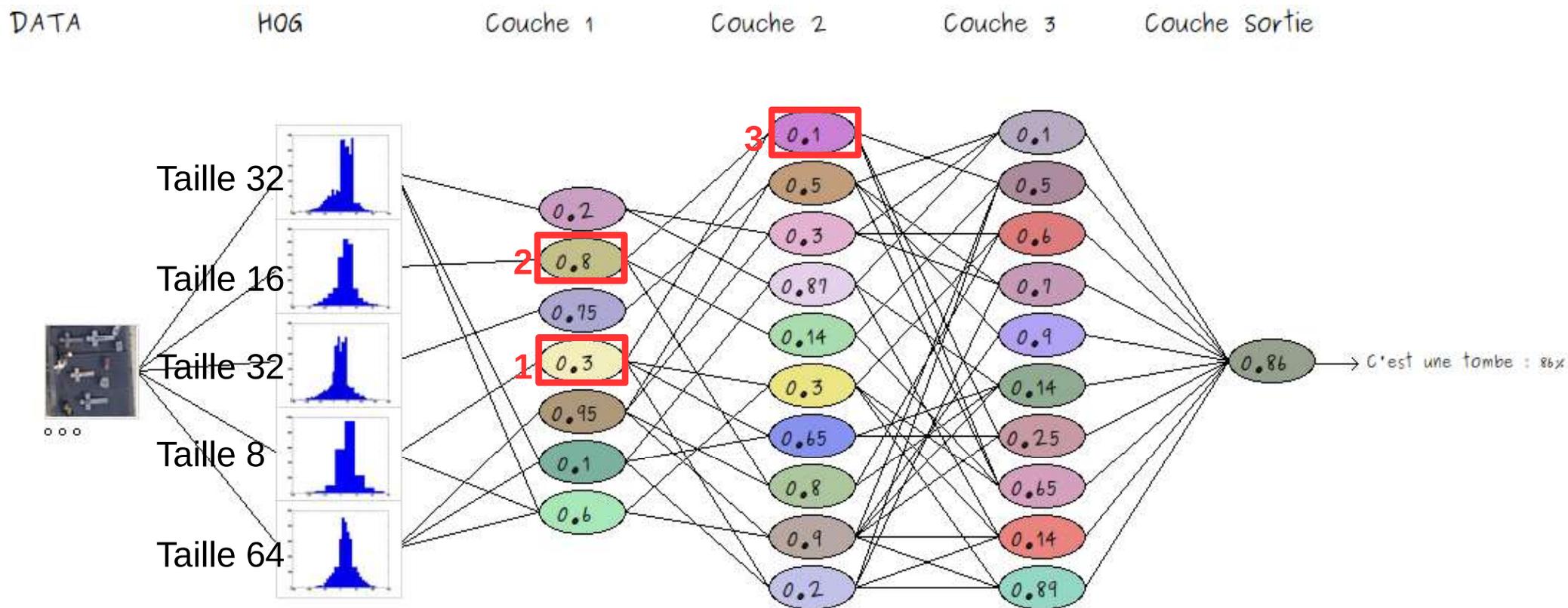
Chemin d'activation



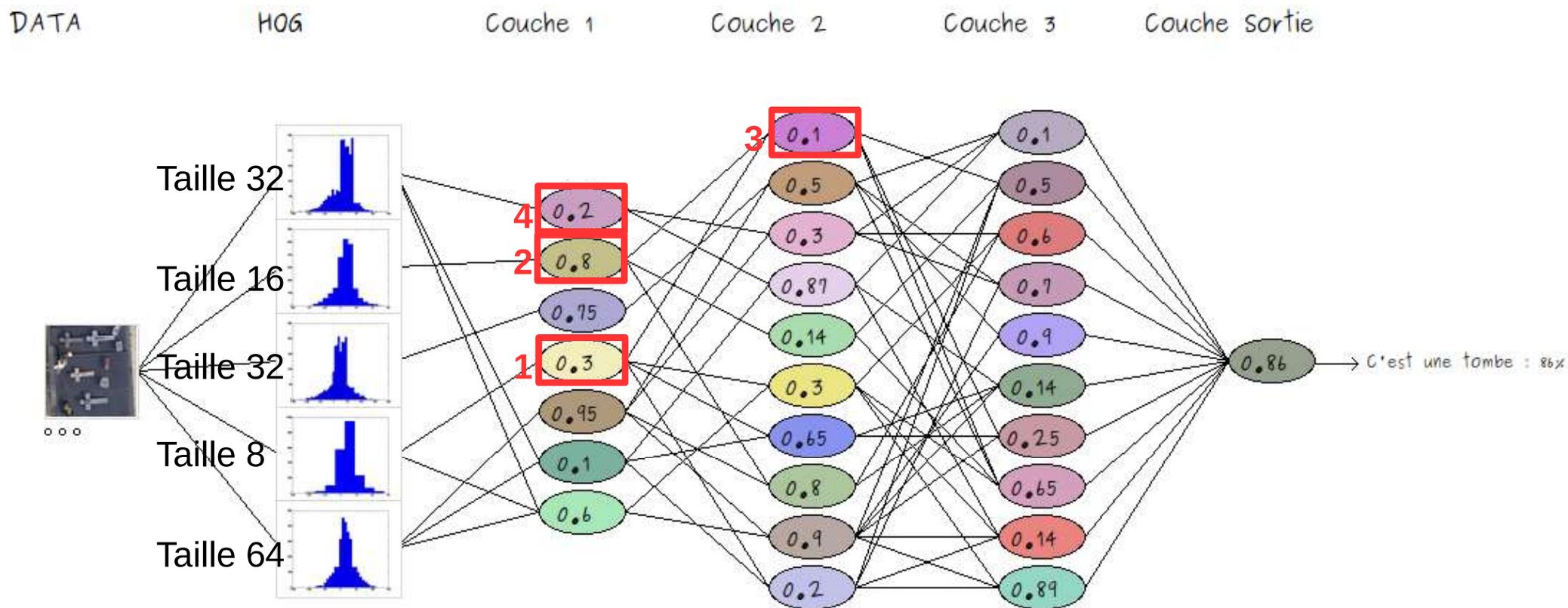
Chemin d'activation



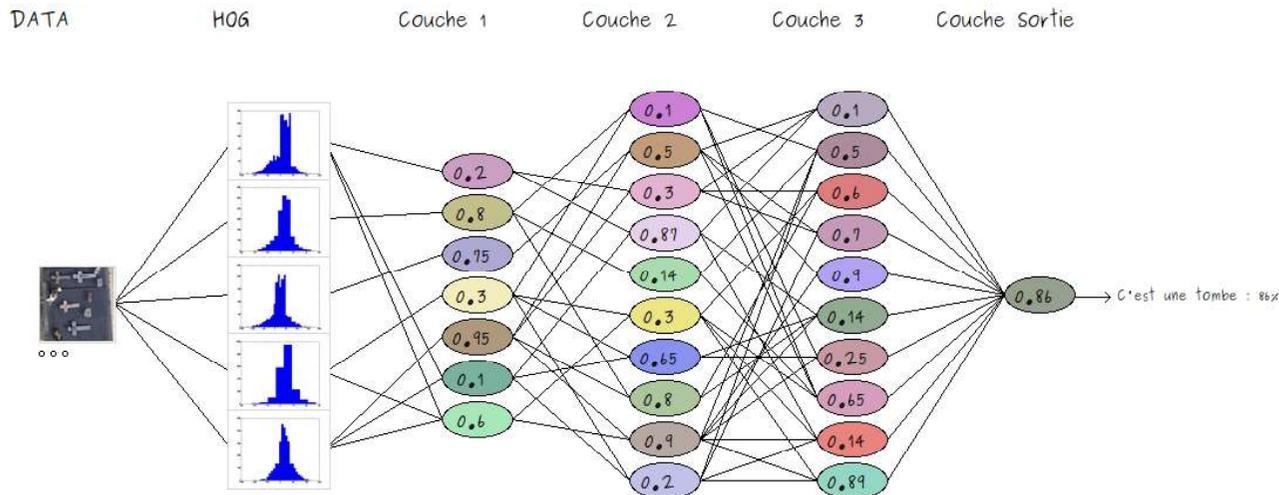
Chemin d'activation



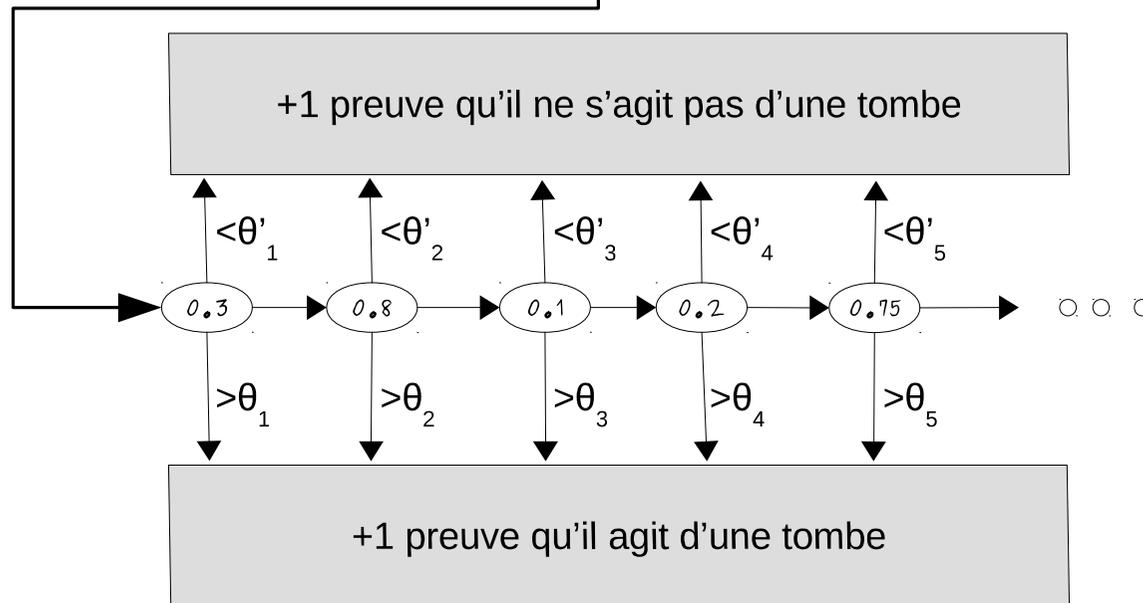
Chemin d'activation



Chemin d'activation



Tri par complexité d'activation



Les SVM avec une précision faible sont rejetés du chemin d'activation.

Résultats

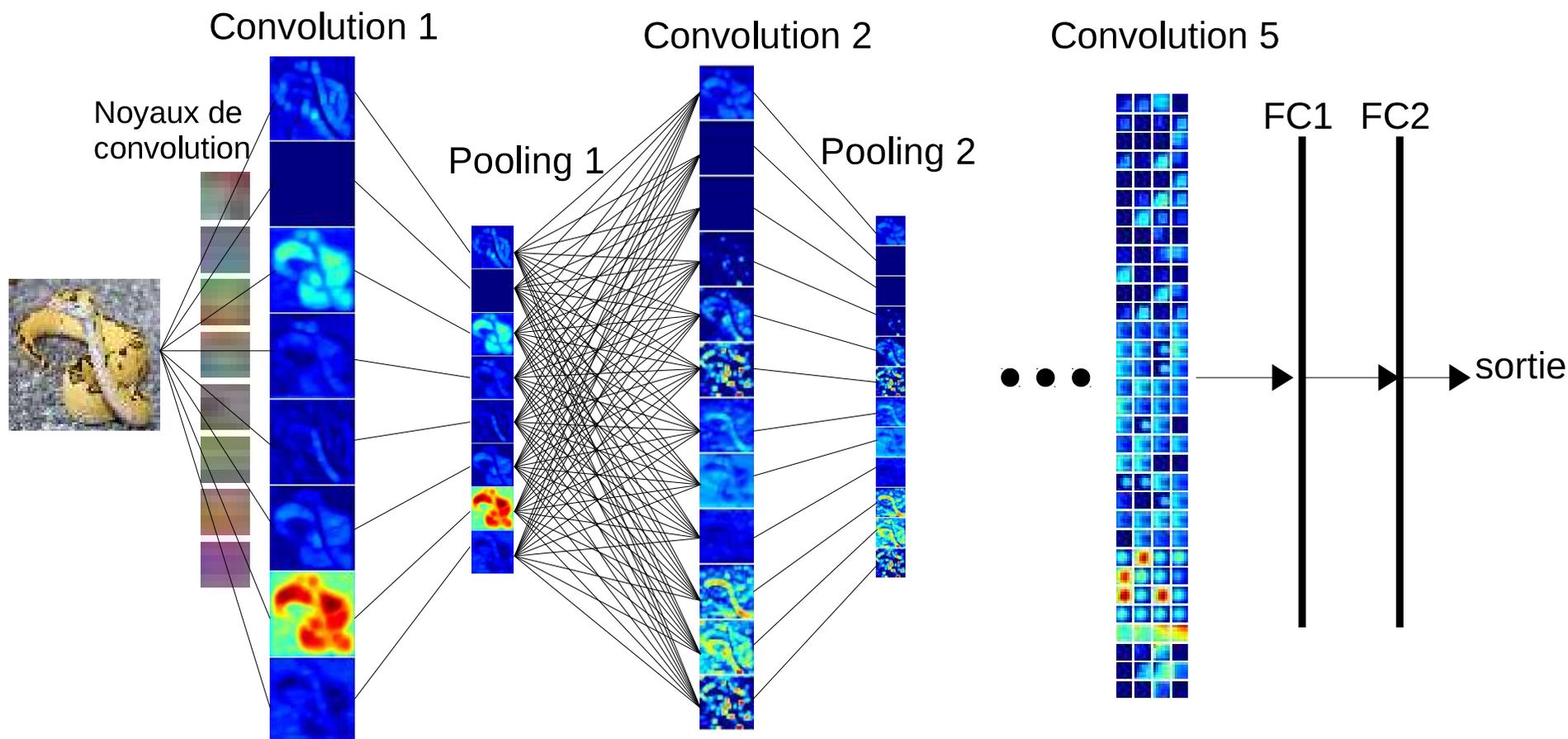
Approche	Temps hh:mm	Précision moyenne
SVM seul	15:55	60.0%
Réseau de SVM	24:35	75.0%
Réseau de SVM (avec chemin d'activation)	4:25	76.0%

Réduction significative du temps nécessaire à la phase d'évaluation.

Sommaire

- ① Motivations
- ② Méthode de référence
 - Méthodologie générale d'apprentissage
 - Méthodologie générale d'évaluation
 - Descripteur et classifieur utilisés
 - Évaluation
- ③ Proposition du réseau de SVM
 - Problème de fusion des résolutions
 - Proposition du réseau de SVM
 - La gestion de la couleur
 - Accélération du réseau
- ④ **Combinaison d'un CNN et d'un réseau de SVM**
 - Rappels sur les CNN
 - Le CNN utilisé et sa base d'apprentissage
 - Évaluation de l'approche
- ⑤ **Conclusions et perspectives**

Principe de classification

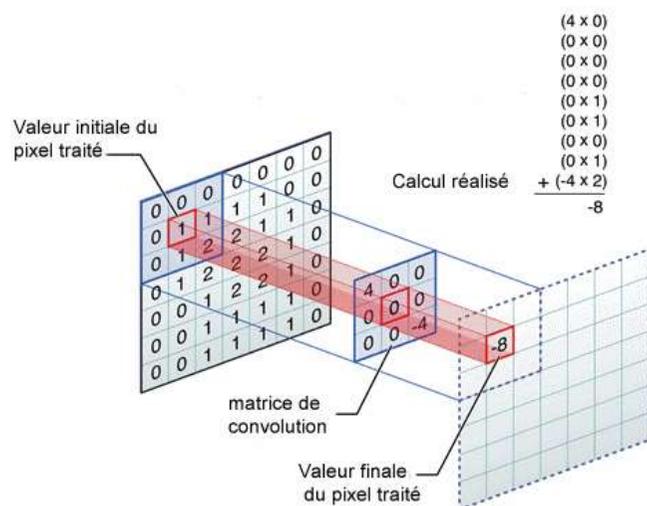


⇒ Des millions de paramètres libres.

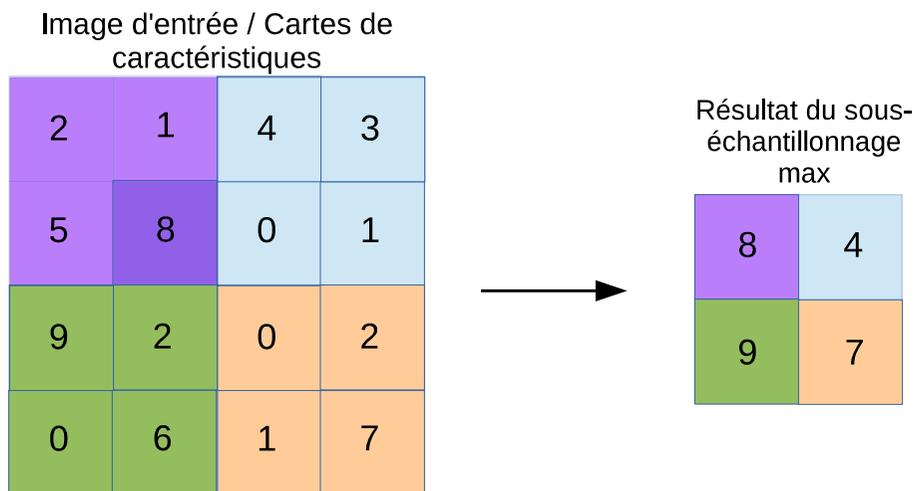
⇒ Détermination des poids par rétro-propagation d'erreur.

Les convolutions et sous-échantillonnages

Couche de convolution



Couche de sous-échantillonnage



Le CNN utilisé

Adaptation du réseau AlexNet pour nos image de dimension 64×64 pixels.
Krizhevsky et al. - ImageNet Classification with Deep Convolutional Neural Networks - NIPS 2012

Augmentation de la taille de la base d'apprentissage

Base d'apprentissage $\times 10^6$	I	II	III	IV
# tombes	1.34			
# autres	1.6			
#semi-tombes		1.34		
#toits			0.4	
#murs				0.7
Précision (rappel=70%)	12%	23%	34%	45%

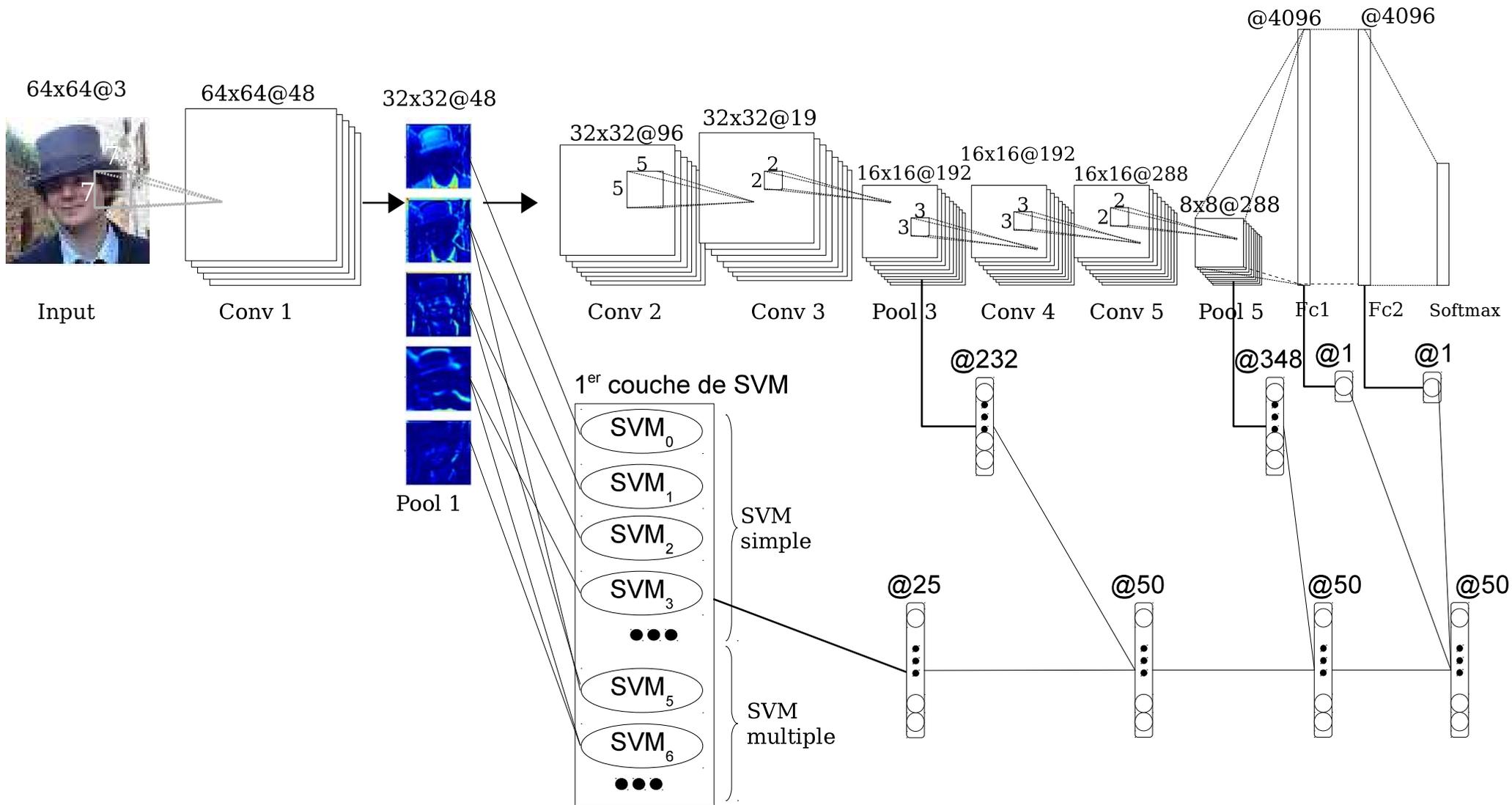
Temps de calculs très important !

Augmentation de la taille de la base d'apprentissage

Base d'apprentissage $\times 10^6$	I	II	III	IV
# tombes	1.34			
# autres	1.6			
#semi-tombes		1.34		
#toits			0.4	
#murs				0.7
Précision (rappel=70%)	12%	23%	34%	45%

Temps de calculs très important !

Adapter un réseau de SVM

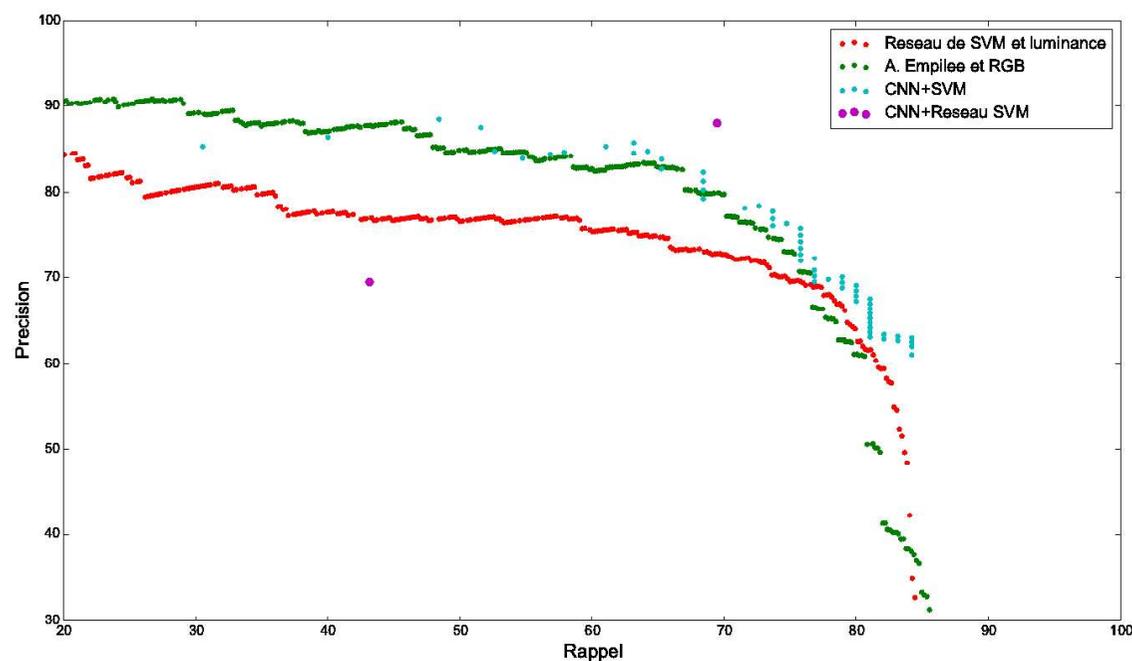


Accélération via un chemin d'activation

Résultats et gain calculatoire

θ_{lim}	Rappel	Précision	Gain calculatoire (%)
0	43.1	69.4	98.6
1	47.3	86.5	98.4
2	52.6	90.9	98,2
5	68.4	90.2	97.8
10	70.5	84.8	97.2
11	70.5	82.7	97.1
27	70.52	77.9	95.0

Détermination du meilleur paramètre θ_{lim} sur une base de validation.





Récapitulatif des propositions de la thèse

Comparaison approche pixel et approche objet

⇓ 1 et 2

Gestion de la multi-résolution par réseau de SVM

⇓ 3

Gestion de la couleur par un réseau de SVM

⇓ 4

Accélération de la phase de tests

⇓ 3

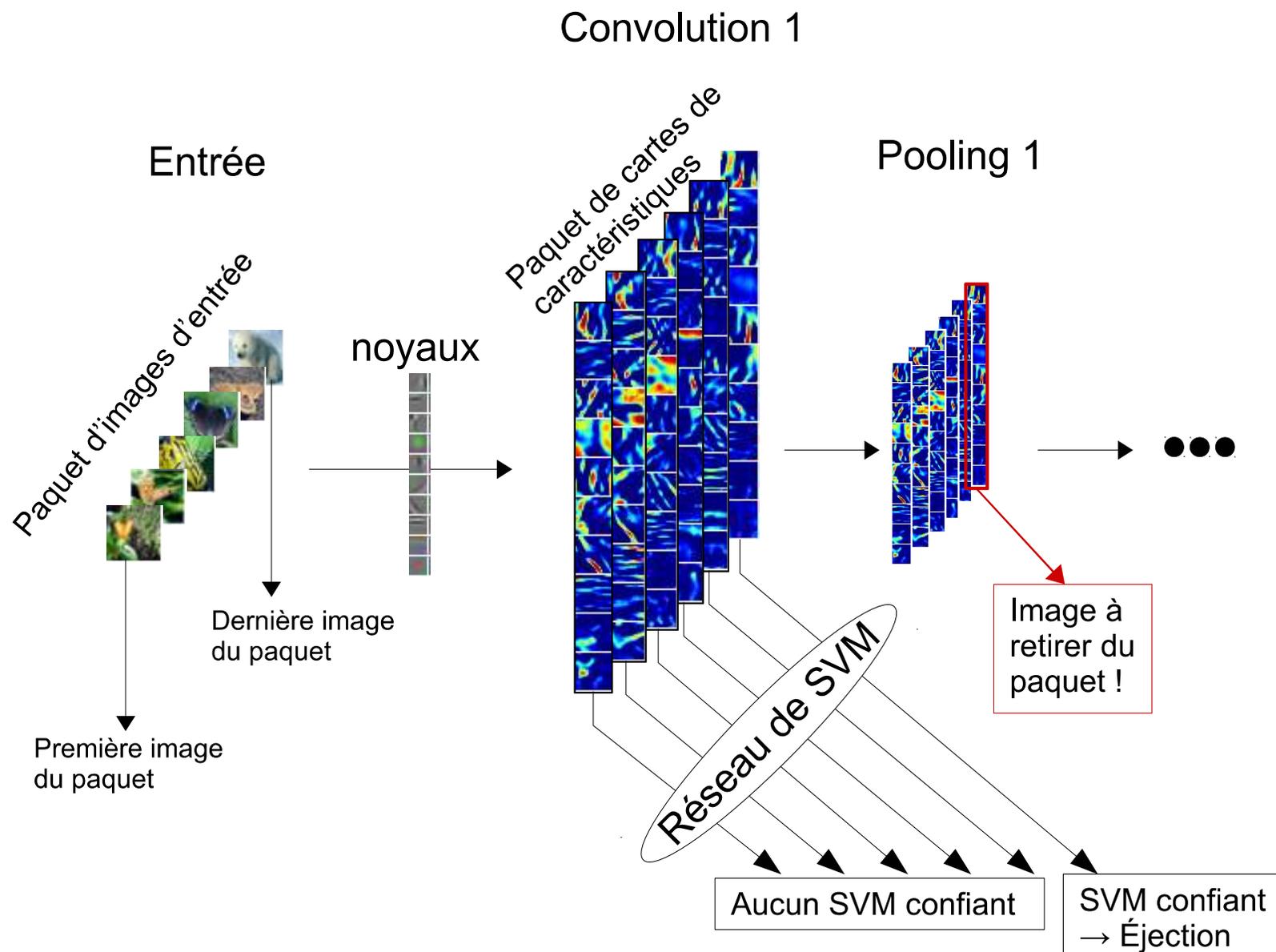
Amélioration et réduction des coûts d'activation de CNN

⇓ 5

- 1- Comparaison de la segmentation pixel et segmentation objet pour la détection d'objets multiples et variables dans des images - Pasquet et al. - Coresa 2014
- 2- Detection of manhole covers in high-resolution aerial images of urban areas by combining two methods - Pasquet et al. - JSTARS 2016
- 3- An efficient multi-resolution deep learning approach for object detection in aerial images - Pasquet et al. - MLSP 2015
- 4- Optimizing color information processing inside a deep convolutional svm network - Pasquet et al. - Electronic Imaging 2015
- 5- Speeding-up a convolutional neural network by connecting an svm network - Pasquet et al. - ICIP2016

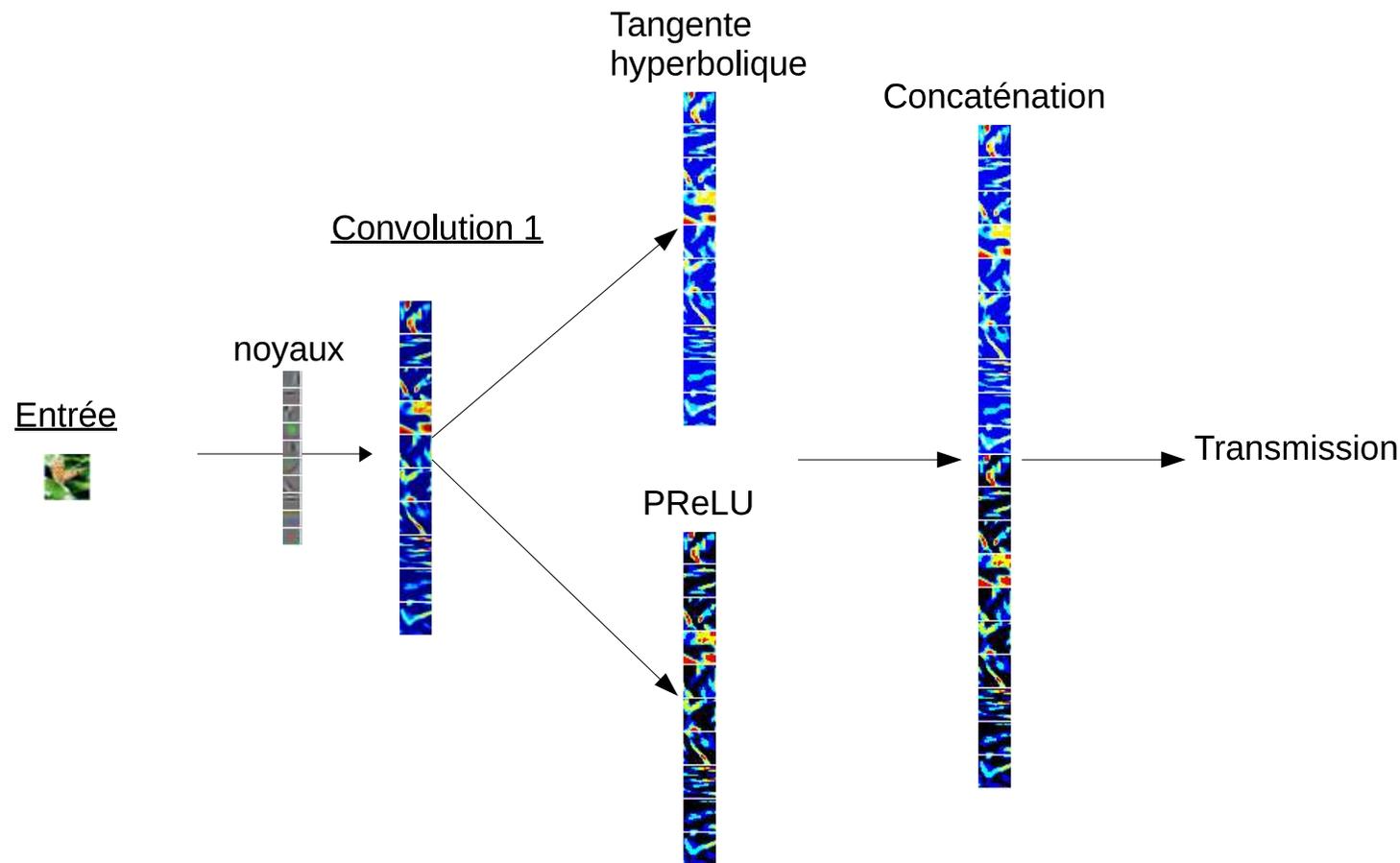
Implémentation du chemin d'activation

Optimisation du CNN par paquet d'images afin de mieux paralléliser les calculs GPU.



Amélioration de la fonction d'activation

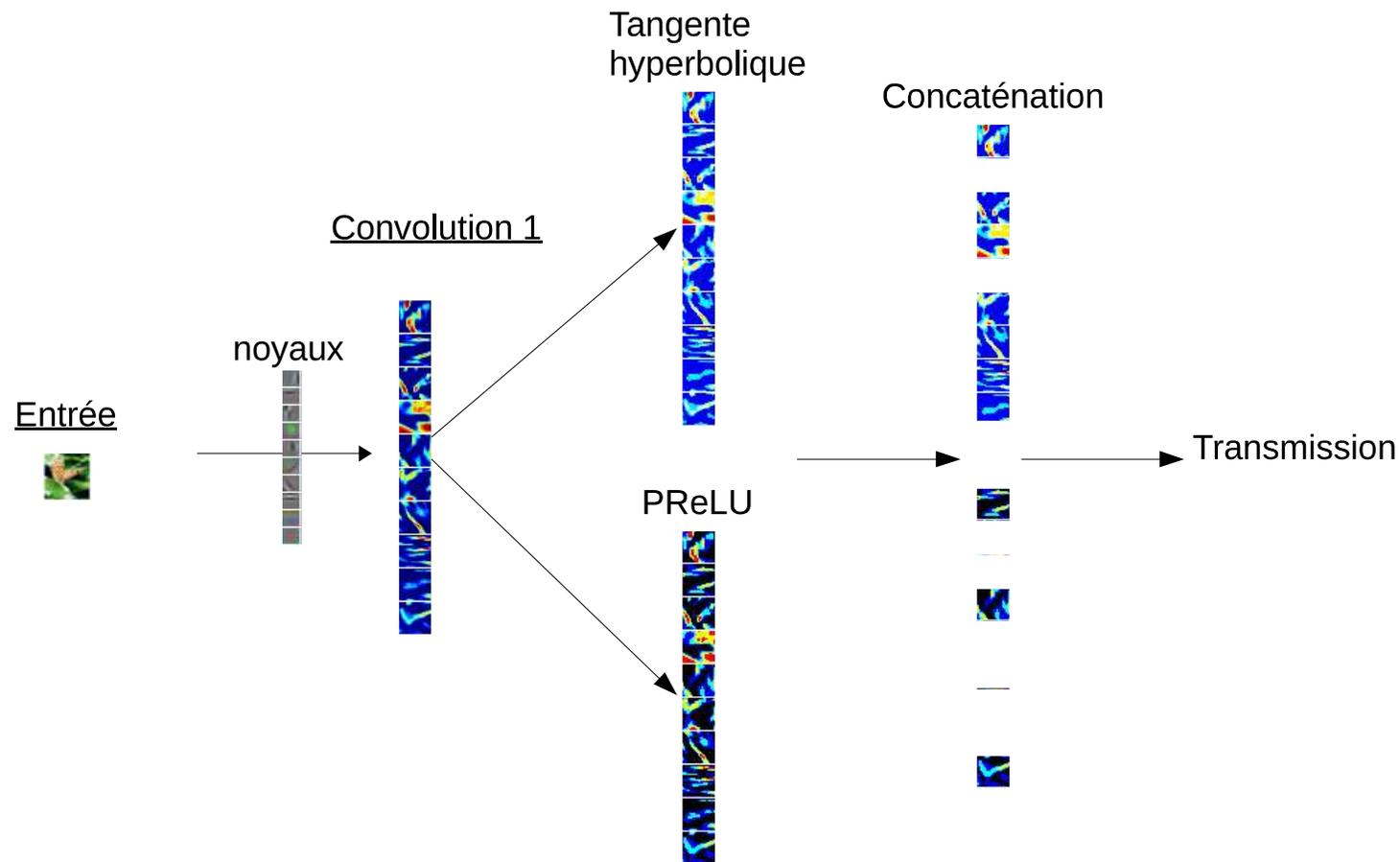
Objectif : permettre au réseau de s'adapter en utilisant la meilleure fonction d'activation.



Papier en cours d'écriture avec Lionel Pibre et al.

Amélioration de la fonction d'activation

Objectif : permettre au réseau de s'adapter en utilisant la meilleure fonction d'activation.



Papier en cours d'écriture avec Lionel Pibre et al.

Amélioration des performances du réseau de SVM

- L'architecture empilée pour fusionner des descripteurs différents
- Créer un réseau SVM prenant en entrée un CNN et des caractéristiques HOG
- Accumulation de preuves continues (non binaires)
- Utiliser le réseau de SVM pour la gestion d'autres types de données
- Étendre la détection à d'autres objets urbains

Merci de votre attention

Revue internationale

- Detection of manhole covers in high- resolution aerial images of urban areas by combining two methods. Journal of Selected Topics in Applied Earth Observations and Remote Sensing - JSTARS 2016
- New classification method to detect quasars based on deep learning - Journal A&A - en cours de soumission

Conférence Internationale

- Speeding-up a convolutional neural network by connecting an SVM network. ICIP 2016
- Optimizing color information processing inside an SVM network. Electronic Imaging 2016.
- Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover source-mismatch. Electronic Imaging 2016 - L. Pibre, J. Pasquet et al.
- An efficient multi- resolution SVM network approach for object detection - MLSP 2015
- Detection of manhole covers in high-resolution aerial images of urban areas by combining two methods. JURSE 2015
- Steganalysis with cover-source mismatch and a small learning database - EUSIPCO 2014
- Automatic localization of tombs in aerial imagery : application to the digital archiving of cemetery heritage, Digital Heritage 2013 - M. Chaumont, J. Pasquet et al.

Conférence nationale

- Comparaison de la segmentation pixel et segmentation objet pour la détection d'objets multiples et variables dans des images - Coresa 2014
- Des millions d'images pour la stéganalyse : inutiles - Coresa 2013
- Étude des réseaux de neurones sur la stéganalyse - Coresa 2016 - L. Pibre, J. Pasquet et al.

Merci de votre attention

Revue internationale

- Detection of manhole covers in high- resolution aerial images of urban areas by combining two methods. Journal of Selected Topics in Applied Earth Observations and Remote Sensing - JSTARS 2016
- New classification method to detect quasars based on deep learning - Journal A&A - en cours de soumission

Conférence Internationale

- Speeding-up a convolutional neural network by connecting an SVM network. ICIIP 2016
- Optimizing color information processing inside an SVM network. Electronic Imaging 2016.
- Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover source-mismatch. Electronic Imaging 2016 - L. Pibre, J. Pasquet et al.
- An efficient multi- resolution SVM network approach for object detection - MLSP 2015
- Detection of manhole covers in high-resolution aerial images of urban areas by combining two methods. JURSE 2015
- Steganalysis with cover-source mismatch and a small learning database - EUSIPCO 2014
- Automatic localization of tombs in aerial imagery : application to the digital archiving of cemetery heritage, Digital Heritage 2013 - M. Chaumont, J. Pasquet et al.

Merci à