

Implémentation d'un système de détection et de reconnaissance faciale

Alexandre Bonhomme

Université Paul Sabatier, Toulouse

Florent Mauftras

EFREI, Paris

Timothee Planté

EISTI, Pau

I Introduction

La reconnaissance faciale est actuellement un domaine en plein essor. Elle rentre petit à petit dans nos vies au travers de nos téléphones mobiles ou de nos ordinateurs portables par exemple. Malgré l'amélioration du taux de détection elle reste l'objet de nombreuses études actuellement.

L'objectif de notre projet sera de mettre en oeuvre un système complet permettant la détection et la reconnaissance de visage issue de plusieurs bases de données. Pour ce faire nous implémenterons différents algorithmes de reconnaissances et nous utiliserons la bibliothèque OpenCV [3] et son implémentation de l'algorithme de Viola & Jones pour la détection de visage.

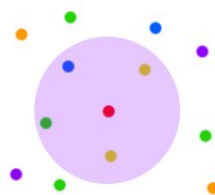


Figure 1 – K Plus Proches Voisins

Dans la Figure 1 le point de test (en rouge) va recevoir comme prédiction la classe jaune car c'est elle la plus présente dans les $K = 4$ voisins les plus proches. Le disque ici n'est que représentatif pour mettre en évidence les voisins. Les points voisins pouvant être à une distance plus ou moins importantes de notre point test.

2 Algorithmes mis en oeuvre

Au cours de notre projet nous avons confronté différents algorithmes afin de tester leurs aptitudes sur plusieurs ensembles de données différents.

2.1 K Plus Proches Voisins et Fenêtres de Parzen

L'algorithme des «K plus proches voisins» (KPPV) utilise le voisinage du point de test afin de déterminer sa classe. C'est la classe majoritaire parmi ces K points qui sera attribuée au point de test (1).

Pour cet algorithme il n'y a pas d'apprentissage des paramètres sur l'ensemble d'entraînement, seulement une optimisation des hyperparamètres K sur l'ensemble de validation.

2.1.1 Extension avec les fenêtres de Parzen

L'application que l'on a fait ici de Parzen nous permet de pondérer les points du voisinage. Plus un point sera distant du point à l'étude plus son poids sera faible. Cela est calculé à partir d'une gaussienne centrée sur le point à l'étude et d'une largeur fixée par σ , dans notre cas σ sera fixé par l'utilisateur.

2.2 Réseau de Neurones

Le réseau de neurone utilisé dans ce projet est de type Perceptron Multicouche. Ce modèle est entraîné par descente de gradient de backpropagation sur un ensemble d'entraînement afin d'optimiser les paramètres du réseau en réduisant le risque empirique régularisé. On tâchera ensuite d'optimiser les hyperparamètres de contrôle de capacité sur un ensemble de validation séparé.

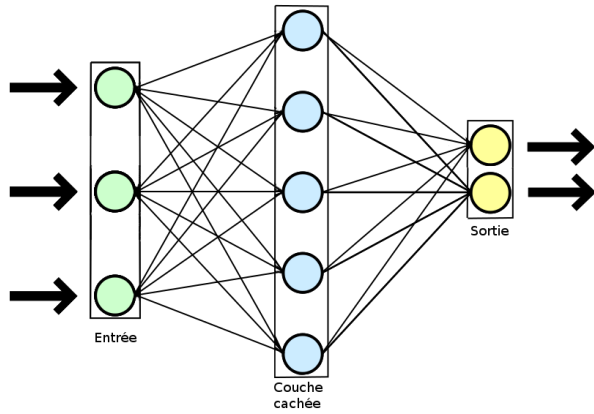


Figure 2 – Réseau de neurones

Dans notre cas nous utiliserons une architecture monocouche (cf. Figure 2) et une pénalité de type «weight decay» (1) pour la régularisation car elle défavorisera les poids les plus lourds lors de l'apprentissage des paramètres.

$$\Omega(\theta) = \|\mathbf{W}^{(1)}\|^2 + \|\mathbf{W}^{(2)}\|^2 \quad (1)$$

Cette pénalité sera pondérée par un hyperparamètre λ .

2.3 Algorithme de Viola et Jones

Afin d'effectuer efficacement la détection de visage nous avons utilisé une implémentation de l'algorithme de Viola et Jones [4] proposé par la célèbre bibliothèque de traitement d'image *OpenCV* [3].

Cet algorithme est basé sur le principe de «boosting» qui consiste à assembler plusieurs classifieurs faibles afin d'obtenir un classifieur à forte capacité. Plutôt que de travailler directement sur les pixel de l'image, Viola P. et Jones M., introduisent des caractéristiques appelées «pseudo-Haar» fortement inspirées des ondelettes de Haar. De plus il utilise une variante d'un algorithme de boosting très célèbre qui est l'«AdaBoost».

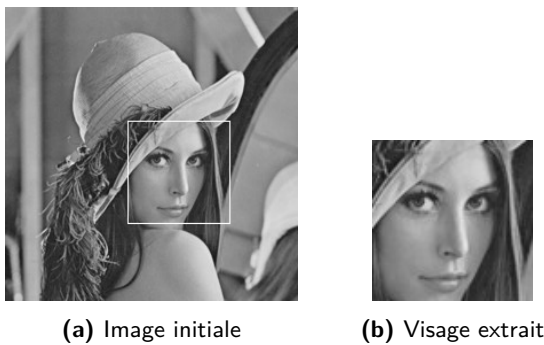


Figure 3 – Détection d'un visage par l'algorithme de Viola et Jones [4]

Cette technique est particulièrement efficace et offre d'excellents résultats pour la détection de visages (cf.

Figure 3) et par extension la détection de motifs (ou d'objets) au sein d'une image.

3 Ensembles de données utilisés

Afin d'entraîner nos algorithmes et de mesurer leurs performances nous avons essentiellement utilisé deux jeux de données.

3.1 Our Database of Faces (ORL [2])

Le premier ensemble de données utilisé est issu du laboratoire de l'entreprise *AT&T*. C'est un ensemble de taille moyenne qui regroupe au total 40 sujets différents avec 10 images par sujet. Les photos ont été prises sous différentes postures et durant plusieurs années et ont une taille de 92×112 pixels.

Les images sont en niveaux de gris (NdG) et au format PGM.

3.2 Labeled Faces in the Wild Home (LFW [5])

Pour le second jeu de données, il est mis à disposition par l'université du Massachusetts et il s'agit de photographies tirées d'Internet et sur lesquelles l'algorithme de Viola et Jones [4] a été appliqué afin d'obtenir une image de taille 250×250 pixels centrée sur le visage.

Cet ensemble est relativement grand puisqu'il contient 13233 images de 5749 sujets différents. Les images sont en couleur et au format JPG.

3.3 Prétraitements des données

3.3.1 Cas particulier : l'ensemble LFW

Contrairement à l'ensemble ORL, LFW nécessite des prétraitements supplémentaires. En effet, nous avons utilisé l'ensemble LFW pour simuler une entrée réelle dans notre système. Pour ce faire nous avons appliqué l'algorithme de Viola et Jones afin de détecter les visages au sein des images présentes dans l'ensemble. Nous avons ainsi pu extraire (après divers traitements de redimensionnement) une image de taille 92×115 pixels en plusieurs NdG.

3.3.2 Cas général

Appartir des images en différents NdG nous permet de construire une représentation vectorielle de celle-ci (cf. Figure 4).

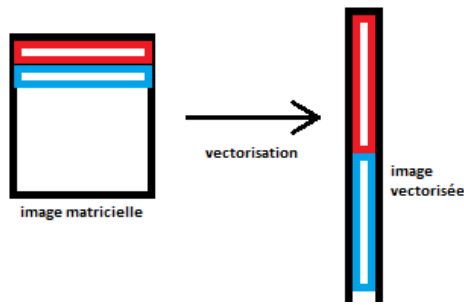


Figure 4 – Vectorisation d'une image

Cependant cette représentation est très lourde car elle contient de nombreuses caractéristiques. Afin de réduire la dimensionnalité de nos données d'entraînement nous avons appliqué l'algorithme du *PCA* («Principal Components Analysis»), plus connu sous le nom de «Eigenfaces» [1] lorsqu'il est appliqué à la reconnaissance faciale.

3.3.3 Réduction de la dimensionnalité : PCA

Nous avons implémenté le PCA dans sa forme classique.

On normalise l'ensemble de données en soustrayant le «visage» moyen :

$$X_N = X - \bar{x} \quad (2)$$

À la suite de quoi nous utilisons une «astuce mathématique» qui consiste à calculer la Décomposition en Valeurs Singulières (ou SVD pour «Singular Value Decomposition») afin de déterminer les vecteurs propres de la matrice de covariance. On obtient alors :

$$X_N = USU^T \quad (3)$$

Où les colonnes de U contiennent les vecteurs propres et S^2 les valeurs propres associées. On ne conserve alors qu'un nombre M de vecteurs propres (ie. M colonnes de la matrice U). Après réduction du nombre de colonnes on obtient une matrice U' (aussi appelée, dans notre cas, «Eigenspace») dans laquelle on projette l'ensemble de données X afin de déterminer la matrice de poids W associée :

$$W = \langle (U')^T, X \rangle \quad (4)$$

Ainsi, pour chaque nouvel exemple x on calculera son vecteur de poids w par projection dans l'Eigenspace U' .

4 Résultats

5 Particularités du projet

Une des caractéristiques de notre projet est qu'il a été réalisé par une équipe de trois personnes ayant suivi des formations différentes. De plus nous avons dû

mettre en place un gestionnaire de version (Git) afin de pouvoir travailler efficacement malgré nos disponibilités très disparates.

L'usage du langage Python a aussi ajouté son lot de complications car nous avons tous les trois découvert ce langage au début du semestre.

La dernière difficulté, et non des moindres, concerne le sujet en lui-même. En effet, la reconnaissance faciale est encore un sujet d'étude actuellement et aucun système n'est encore réellement fiable dans ce domaine.

6 Conclusion

Références

- [1] TURK M. AND PENTLAND A., *Eigenfaces for recognition*. J. Cognitive Neuroscience, 1991.
- [2] SAMARIA F. AND HARTER A., *Parameterisation of a stochastic model for human face identification*. 2nd IEEE Workshop on Applications of Computer Vision, 1994.
- [3] BRADSKI G., *The OpenCV Library*. Dr. Dobb's Journal of Software Tools, 2000.
- [4] VIOLA P. AND JONES M., *Rapid object detection using a boosted cascade of simple features*. Computer Vision and Pattern Recognition, 2001.
- [5] HUANG G. B., RAMESH M., BERG T. AND LEARNED-MILLER E., *Labeled Faces in the Wild : A Database for Studying Face Recognition in Unconstrained Environments*. University of Massachusetts, Amherst, 2007.