

Implémentation d'un système de détection et de reconnaissance faciale

Alexandre Bonhomme

Université Paul Sabatier, Toulouse

Florent Maufras

EFREI, Paris

Timothée Planté

EISTI, Pau

I Introduction

Le but de notre projet est de prendre en entrée une image comportant un visage d'en faire la détection sur l'image puis d'en tester la reconnaissance sur une base de donnée.

Au final toute image en entrée devrait mener vers un résultat indiquant si le visage qu'elle comporte est reconnu parmi les visages appris par le modèle que l'on a entraîné sur un nombre x d'individu.

2 Algorithmes mis en oeuvre

2.1 K plus proches voisins et Fenêtres de Parzen

L'algorithme KNN correspond à l'algorithme recherchant les K plus proches voisins (nearest neighbor) et qui permet de classifier en fonction de la classe majoritaire dans ces K voisins sélectionnés.

L'entraînement d'un tel algorithme correspond à l'apprentissage simple des données de l'ensemble d'entraînement. Pour tout nouvel élément test entrant il nous suffit alors de chercher les K plus proches voisins tiré de l'ensemble d'entraînement et d'attribuer comme classe la classe la plus représentée dans son voisinage.

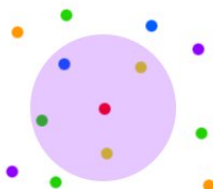


Figure 1 – K Plus Proches Voisins

Dans cet exemple le point de test (point rouge) va recevoir comme prédiction de classe la classe jaune car

c'est elle la plus présente dans les $K = 4$ voisins les plus proches (zone de voisinage de ce point mis en évidence par le disque). 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.

2.1.1 Extension avec les fenêtres de Parzen

L'application que l'on a fait ici de Parzen nous permet ici de rajouter un poids à nos point. 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 , σ dans notre cas est fixé par l'utilisateur.

2.2 Réseau de Neurones

L'algorithme du réseau de neurones (Neural Networks) est un modèle qui comporte plusieurs couches et l'entraînement d'un tel système est la tendance à réduire la perte/l'erreur à l'aide de plusieurs itérations. On part d'un état avec des poids définis aléatoirement et on tend alors par une réduction des erreurs à chaque itération vers un état d'apprentissage de notre modèle.

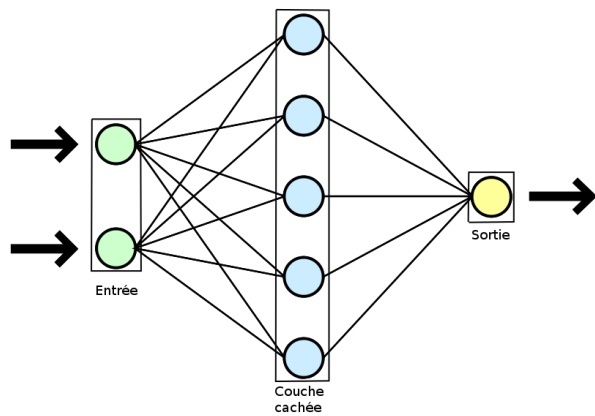


Figure 2 – Réseau de neurones

La réduction de l'erreur revient à la minimisation du risque empirique et la réduction de ce dernier peut-être réalisé de diverses manières :

back propagation consiste en une propagation de la correction liée à l'erreur depuis la sortie du système vers l'entrée

feed forward consiste en une propagation de l'erreur depuis l'entrée vers la sortie

3 Ensembles de données utilisés

3.1 Description

Nos données sont découpées en deux ensemble : l'ensemble ORL et l'ensemble LFW. Ces ensembles d'images ont été tiré de base de données d'institut.

Dans le cas du jeu de donnée ORL, l'institut a prit diverses photographies en noir et blanc d'individus avec différentes postures et ce sur plusieurs années pour chacun. Ces clichés sont enregistrés au format pgm.

Pour le second jeu de données, LFW, il s'agit de photographies tirées d'internet et sur lesquels un traitement de détection de visage (Viola et Jones) a été appliqué. Il récupérait alors les visages ainsi repéré en élargissant le cadre autour du visage. Toutes ces images sont en couleurs et au format jpg.

3.2 Prétraitements

Toutes les images sont alors traité et ce de façon différente en fonction de l'ensemble de données depuis lequel on la sélectionne.

Une photographie en couleur, de l'ensemble LFW, sera tout d'abord transformée en une image en niveau de gris puis ensuite on recadra toutes les images selon une taille unique sur toutes les photographies et ce autour des visages. La taille unique est définie par la sélection des maximum observés selon la largeur et la hauteur des cadres obtenus lors de la détection des

visages. En effet lors de la détection du visage nous obtenons un encadrement de ce dernier en fonction d'une position, d'une hauteur et d'une largeur.



(a) Image initiale

(b) Visage extrait

Figure 3 – Détection d'un visage par l'algorithme de Viola et Jones [1]

La transformation alors commune à toutes les images est la vectorisation (4) de ces dernières. Cela revient à passer d'une image matricielle à une image sous forme d'un vecteur selon un des deux axes de la matrice. L'axe choisit est celui des lignes.

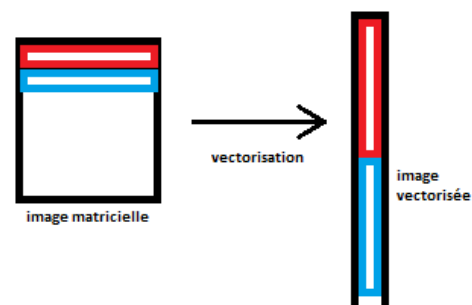


Figure 4 – Vectorisation d'une image

4 Résultats

5

Un des premiers points spécifiques à notre projet et le fait que ce dernier a été réalisé par une équipe de trois personnes ayant suivi des formations différentes. Nous n'avions pas les mêmes disponibilités non plus pour avancer le projet du fait que nous étions inscrits à des cours différents aussi. Cette raison a mené à la mise en place d'une gestionnaire de version avec une centralisation via un site web qui offre ce service : <https://bitbucket.org/> Ceci à donc rajouter une difficulté au projet qui était la gestion de l'équipe et la répartition des tâches au sein de cette dernière. L'usage d'un logiciel tel que Git pour gérer les versions en local et via aussi la passerelle distante.

L'usage du python était aussi une des difficultés auquel nous avons voulu nous confronté étant donné que nous avons tous trois découvert ce langage au cours des derniers mois nous voulions essayé de le pratiquer dans en longueur au cours de ce projet.

Le sujet lui-même représente une difficulté quant à la recherche d'un bon résultat étant donné qu'encore à ce jour des entreprises et chercheurs travaillent encore au développement et à l'amélioration de système de reconnaissance faciale car aucun système n'est encore réellement fiable dans ce domaine.

6 Conclusion

Références

- [1] VIOLA P. AND JONES M., *Rapid object detection using a boosted cascade of simple features*. Computer Vision and Pattern Recognition, 2001.
- [2] TURK M. AND PENTLAND A., *Eigenfaces for recognition*. J. Cognitive Neuroscience, 1991.