

● État de l'Art et Solution Proposée ●

Reconnaissance de Visage

Auteurs du Rapport : Fotsing Sikadie Gervais S., Ginel Dorleon, Eliodor Ednalson Guy M. (P21)

Superviseur : HO Tuong Vinh

Module : Reconnaissance de Forme - M2 - IFI

I. INTRODUCTION

La reconnaissance de visage est l'une des applications les plus pertinentes de l'analyse d'image. C'est un véritable défi de construire un système automatisé qui équivaut à la capacité humaine pour reconnaître les visages. Bien que les humains identifient assez bien les visages connus, nous ne sommes pas très qualifiés quand nous devons faire face à une grande quantité de visages inconnus. Les ordinateurs, avec une mémoire et une vitesse de calcul presque illimitée, devrait surmonter les limitations humaines. La reconnaissance du visage est un domaine évolutif, changeant et qui s'améliore constamment. Au cours de ces dernières années, la reconnaissance du visage a attiré beaucoup d'attention et ses recherches ont rapidement augmenté non seulement pour les ingénieurs mais aussi pour les neuro-scientifiques, puisqu'elle comporte de nombreuses applications potentielles dans la communication par vision par ordinateur et le système automatique de contrôle d'accès. En particulier, la détection de visage est une partie importante, première étape, de la reconnaissance du visage. C'est dans cette même ordre d'idée que nous nous sommes proposés d'approfondir dans ce domaine dans le but de comprendre les différents algorithmes, les méthodes anciennes, les travaux réalisés afin d'implémenter au final une solution pouvant démontrer réellement l'importance de la reconnaissance de visage comme sous-domaine de la reconnaissance de formes. Dans ce rapport, nous présentons notre démarche pour la construction d'un système de reconnaissance de visage.

II. DÉFINITION - CONCEPTS ET MOTS CLÉS

Dans ce domaine, divers termes sont utilisés et portent parfois à équivoque, alors on tient ci-dessous à faire une clarification de ces différents termes et concepts.

A. Reconnaissance de visage

C'est le fait de déterminer automatiquement si deux faces sont susceptibles de correspondre à la même personne. La reconnaissance faciale est ce domaine de la vision par ordinateur consistant à reconnaître automatiquement une personne à partir d'une image de son visage. Il s'agit de l'identifier, qui est-ce ? ou de l'authentifier est-ce la bonne personne ? après analyse automatique de son

visage. Par exemple, à l'heure actuelle, Google Face fournit des fonctionnalités pour la détection de visage et non la reconnaissance de visage.

B. Détection de visages

C'est le processus consistant à localiser automatiquement les visages humains dans les médias visuels (images numériques ou vidéo). Un visage détecté est signalé dans une position avec une taille et une orientation associées. Une fois qu'un visage est détecté, on peut chercher des repères tels que les yeux et le nez.

C. Face Tracking

C'est une technique qui permet d'augmenter la détection des visages aux séquences vidéo. Tout visage apparaissant dans une vidéo pendant un certain temps peut être suivi. C'est-à-dire que les visages qui sont détectés dans des images vidéo consécutives peuvent être identifiés comme étant la même personne. Notez que ce n'est pas une forme de reconnaissance de visage ; ce mécanisme ne fait que des inférences basées sur la position et le mouvement du (des) visage (s) dans une séquence vidéo.

III. PROBLÈMES RENCONTRÉS PAR CE DOMAINE

La première étape et la plus importante de la reconnaissance du visage est la détection de visage. En fait, détecter les visages humains et extraire du visage les caractéristiques d'une image sans contrainte sont un processus difficile. Il est très difficile de localiser les positions des visages dans une image avec précision. Le visage humain est un objet tridimensionnel (3-D), et pourrait être sous une perspective déformée et d'une illumination irrégulière. Il existe plusieurs variables qui nuisent les performances de la détection, y compris le port de lunettes, coloration cutanée, genre, poils faciaux et expressions faciales, la variation de la pose (avant, non frontale), l'occlusion, l'orientation de l'image.

IV. Domaine d'Application

La reconnaissance de visage trouve son application dans divers domaines de la vie quotidienne, comme par exemple :

- Contrôle d'accès (ATM, Aéroport, etc.)
- Surveillance
- Suggestion automatique d'identification de visage de Facebook
- Game, Exemple : CrowsComing
- la justice criminelle (recherche de malfaiteurs dans des bases de données)
- L'identification et la vérification (banque, commerce électronique)
- La biométrie
- L'indexation de vidéo

V. OUTILS ET TRAVAUX EXISTANTS

Parlant des applications de détection du visage, il existe essentiellement deux approches. L'approche classique consiste en l'utilisation locale. Et la deuxième approche émergente consiste à utiliser le pouvoir d'Internet et à appeler des services Web. Ci dessous nous présentons quelques travaux existants. [1]

A. Deep Face

DeepFace est un système de reconnaissance faciale d'apprentissage approfondi créé par un groupe de recherche de Facebook. Il identifie les visages humains dans les images numériques. Il emploie un réseau neuronal de neuf couches avec plus de 120 millions de poids de connexion et a été formé sur quatre millions d'images téléchargées par les utilisateurs de Facebook. Le système est 97 pourcents précis, comparativement à 85 pourcents pour le système d'identification de la prochaine génération du FBI.

B. FaceNet

Technologie de reconnaissance faciale développée par Google. Le système frôle les 100 pourcents de précision au test de référence Labeled Faces in The Wild (LFW), et dépasse 95% sur la base YouTube Faces DB. En utilisant un réseau neuronal artificiel et un nouvel algorithme.

C. Microsoft Project Oxford

Les algorithmes de visage basés sur le cloud de Microsoft pour détecter et reconnaître les visages humains dans les images. L'enregistrement des besoins est actuellement gratuit jusqu'à 5000 usages/mois.

D. OpenFace

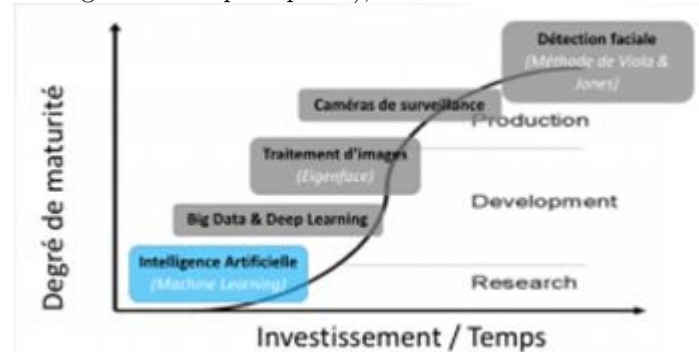
Fabriquée par Brandon Amos, Bartosz Ludwiczuk et Mahadev Satyanarayanan, OpenFace fournit une reconnaissance de visage gratuite et open source avec l'utilisation des réseaux neuronaux profonds.

E. GaussianFace

GaussianFace, est un algorithme de reconnaissance de visage développé par deux chinois : Lu Chaochao et Tang Xiaoou. Cet algorithme a pu réaliser une précision relativement supérieure à celle des humains, avec un pourcentage de 98,52 % contre 97,53 % pour les humains. Ces inconvénients sont la capacité de mémoire requise pour son fonctionnement et le temps nécessaire pour l'exécution du processus.

VI. MÉTHODES, TECHNIQUES ET APPROCHES

La Reconnaissance Faciale repose aujourd'hui sur quatre technologies. La courbe de Foster ci-dessous donne à voir l'état actuel de ces sous-domaines en terme de stade de développement. Le traitement automatique de visage (calcul de distances), l'Eigenface ("représentation des éléments caractéristiques d'une image de visage à partir d'images modèles en niveau de gris"), l'analyse de points particuliers (similaire à l'Eigenface en plus précis), les réseaux de neurones.



Il existe de nombreux algorithmes et de méthodes pour la reconnaissance de visage pour localiser un visage humain dans une scène - plus facile et plus difficile. Voici une liste des techniques les plus récentes publiées pour détection de visage :

D'après l'article intitulé « Naive-Deep Face Recognition : Touching the limit of LFW Benchmark or Not? » paru en 2015, on a appliqué une technique basée sur le deep learning et l'apprentissage sur de grandes masses de données. Dans cet article, les auteurs ont montré comment le big data impacte la performance de la reconnaissance d'image. Ils ont construit un système nommé Megvii Face Recognition System qui accomplit une performance de 95.50 % sur la base d'image LFW benchmark (qui est une base comportant des photos des visages de personnes en plusieurs exemplaires éventuellement destinée aux test pour la détection de visage), ce qui a dépassé la performance des systèmes parues avant 2015. [2]

Dans un deuxième article étudié qui s'intitule « Face recognition in real-world images » paru en 2017, les systèmes de reconnaissance sont conçus pour fonctionner dans des situations bien contrôlées. Toutefois,

les images du monde réel ont des orientations très diverses, des expressions faciales variées et des conditions d'illuminations allant du plus sombre au plus clair. Les algorithmes de reconnaissance faciale traditionnels fonctionnent mal sur de telles images. Dans leurs papiers, ils ont présenté une méthode adaptée aux conditions réelles qui peuvent être entraînée en utilisant très peu d'exemples d'entraînement et qui est peu gourmand en calcul. Leur méthode consiste en l'utilisation d'un algorithme RSC modifié.[3]

D'autres méthodes et algorithmes classiques comme :

- Analyse en Composante Principale (ACP) Inventée en 1901 par Karl Pearson, ACP implique le calcul de la décomposition de valeur propre d'une matrice de covariance de données ou une décomposition de valeur singulière d'une matrice de données, généralement après avoir centré les données pour chaque attribut. Les images de base trouvées par ACP dépendent uniquement des relations par paires entre pixels dans la base de données d'images.
- Reconnaissance de visages par SVM Les machines à vecteurs support (SVM en anglais) sont entraînées par une base d'apprentissage puis servent par la suite à détecter les visages dans les images. La particularité de la méthode réside dans la façon dont la base d'apprentissage est définie. En effet les SVM permettent de séparer de façon optimale les données en deux classes, il importe donc de faire le bon choix de représentation des images pour profiter de la robustesse et de l'efficacité de la méthode.

VII. SOLUTION PROPOSÉE : 2 APPROCHES

Dans notre étude nous nous intéressons à la reconnaissance basée sur les images fixes prises en entrée. Nous présentons deux approches pour réaliser ce système :

- 1) Reconnaissance avec la méthode d'Analyse Discriminante Linéaire de Fisher
Cette approche comporte deux phases :
 - a) La phase d'apprentissage : C'est la phase où le système apprend le visage à partir d'une ou plusieurs images du même visage, elle s'effectue en utilisant l'algorithme de Fisher. A la fin de cette phase, on aura pour chaque visage un modèle unique qui le caractérise.
 - 2) La phase de test : Elle consiste à identifier un visage de la base de test à partir des images qui se trouvent dans la base d'apprentissage. L'Analyse Discriminante Linéaire de Fisher permet d'effectuer une réduction de la dimension spécifique. Inventée par le grand statisticien Sir R. A. Fisher. Il l'a utilisée avec succès pour classer les fleurs dans son article de 1936. Afin de trouver la combinaison de fonctionnalités qui sépare le mieux entre les classes, l'Analyse Discriminante Linéaire maximise le rapport entre les classes et la dispersion des classes,

au lieu de maximiser la dispersion globale. L'idée est simple : les mêmes classes devraient se regrouper ensemble, tandis que les classes différentes sont aussi loin que possible les unes des autres dans la représentation de la dimension inférieure. Contrairement à l'ACP, il permet d'effectuer une véritable séparation de classes. Cette approche cherche les directions qui sont efficaces pour la discrimination entre les données.

Les combinaisons résultantes peuvent être employées comme classificateur linéaire, ou généralement dans la réduction de caractéristiques avant la classification postérieure. Avec l'Analyse Discriminante de Fisher, chaque visage, qui se compose d'un grand nombre de pixel, est réduit à un plus petit ensemble de combinaisons linéaires avant la classification. Chacune des nouvelles dimensions est une combinaison linéaire des valeurs de pixel, qui forment un Template. Les combinaisons obtenues en utilisant LDA s'appellent les Fisherfaces.

- 2) Reconnaissance avec les Réseaux de Neurones Convolutionnels (RNC ou CNN)

La deuxième approche proposée est celle de reconnaissance de visage avec les réseaux de neurones convolutifs. L'approche avec CNN a été testée sur la base de données MINST et fournit un taux d'erreur de 0,23% en Février 2012 et 97, 6% de performance. Les réseaux de neurones convolutifs sont basés sur le perceptron multicouche. C'est-à-dire que le CNN est basé sur un classifieur linéaire de type réseau neuronal formel organisé en plusieurs couches au sein desquelles une information circule de la couche d'entrée vers la couche de sortie uniquement ; il s'agit donc d'un réseau de type feedforward. Chaque couche est du réseau est alors constituée d'un nombre variable de neurones, les neurones de la couche de sortie correspondant toujours aux sorties du système.[4]

Le réseau de neurones convolutionnels utilisé ici fonctionne comme un réseau de reconstruction ; il réalise une projection non linéaire du visage présenté en entrée sur un sous - espace puis reconstruit un visage de référence choisi au préalable. Les réseaux de neurones convolutionnels offrent l'avantage d'apprendre automatiquement l'extraction des caractéristiques dans les premières couches, puis de les classer dans les dernières. Ainsi, aucun choix d'algorithmes d'extraction n'est effectué. Ils sont de plus conçus pour être plus robustes aux changements de pose ainsi qu'aux changements thermiques des zones d'un visage liés aux conditions extérieures (comme la température), étant données leurs fonctions d'activation non linéaires.

- Démarche avec le CNN

Notre démarche avec le CNN prend en entrée une image de taille fixe et la passe à travers une succession de couches de convolution C_i , de sous-échantillonnage S_i et de neurones complètement connectés F_i . Chaque couche calcule son opération spécifique (convolution et sous-échantillonnage) avec son vecteur de poids, ajoute ensuite un biais puis passe le résultat dans une fonction d'activation. La sortie du réseau est une image, de même taille que l'entrée, qui est reconstruite par la dernière couche du réseau. Cette démarche a été inspirée par les travaux de Lecun et al. et adaptée à notre problème.[5]

VIII. CONCLUSION

La reconnaissance de visage est un domaine de recherche en pleine expansion et qui propose de multiples applications pratiques. Les méthodes utilisées pour sa mise en œuvre sont diverses et se basent pour certaines directement sur les caractéristiques de l'image et pour d'autre sur les algorithmes d'apprentissage ou encore des approches statistiques. Le présent rapport fait un état de l'art des techniques, les méthodes, les algorithmes et outils utilisés dans la reconnaissance de visage ainsi que l'approche de la solution proposée. Il n'existe pas encore de méthodes ou approche parfaite. Cependant, nous optons tirer le meilleur de ces approches afin de fournir un système de détection avec des performances remarquables. Le domaine de la reconnaissance de visage est à la recherche de meilleures applications et/ou de méthodes afin d'améliorer ses performances et précisions. Dans une prochaine étape, nous présenterons les résultats obtenus après implémentation de la solutions proposée en se basant sur ces différentes approches évaluant leurs avantages et inconvénients.

Références

- [1] <https://cmusatyalab.github.io/openface/overview>, OpenFace-Overview, consulté le 16 Mai 2017
- [2] Fontaine, X., Achanta, R., Süssstrunk, S. (2017). Face Recognition in Real-world Images. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (No. EPFL-CONF-224338)
- [3] Lawrence, A., Ashwin, N. M., Manikantan, K. (2017). Face Recognition Using Background Removal Based on Eccentricity and Area Using YCbCr and HSV Color Models. In Proceedings of the International Conference on Signal, Networks, Computing, and Systems (pp. 33-43). Springer India.
- [4] Sun, X., Wu, P., Hoi, S. C. (2017). Face detection using deep learning : An improved faster rcnn approach. arXiv preprint arXiv :1701.08289.
- [5] Y. LeCun, L. Bottou, Y. Bengio, and P. Hader. Gradient-based learning applied to document recognition. In Intel ligent Signal Proessing, pages 306351. IEEE Press, 2001.