

**UNIVERSITÉ NATIONALE DU VIETNAM**  
**INSTITUT DE LA FRANCOPHONIE POUR L'INNOVATION (IFI)**

---



## **Module Travail Personnel Encadré**

---

Contrôle de présence automatique à l'aide des  
algorithmes de vision par ordinateur et  
d'apprentissage

---

### **RAPPORT FINAL**

**MASTER I: PROMOTION 23 OPTION (SIM)**

**Rédigé par** : Abdoulaye CISSE

**Encadreurs** : Dr Ho Tuong Vinh et Dr Ta Dang Chi

Année Académique 2018 - 2019

## REMERCIEMENTS

---

Je tiens tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce Modeste travail. En second lieu, je tiens à remercier mes encadreurs Dr Ho Tuong Vinh et Dr Ta Dang Chi pour leur précieux conseils et leur aide durant toute la période du travail. Mes vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail, et de l'enrichir par leurs propositions.

Je souhaite adresser mes remerciements les plus sincères aux personnes qui nous ont apporté leur aide et qui ont contribué à l'élaboration de ce mémoire ainsi qu'à la réussite de cette formidable année universitaire.

A ma famille et mes amis qui par leurs prières et leurs encouragements, m'a permis de surmonter tous les obstacles. Enfin, je tiens également à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

## Table des matières

---

Table des matières .....	1
I. INTRODUCTION .....	5
II. ANALYSE DU SUJET.....	6
III. RECHERCHE BIBLIOGRAPHIQUE .....	9
1. Connaissance précise et spécialisée du sujet .....	9
2. Les méthodes de reconnaissance faciale .....	9
3. Les Techniques ou méthodes existants .....	10
IV. SOLUTION PROPOSEE.....	15
1. Description de la solution proposée .....	15
2. Les méthodes d'apprentissage automatique .....	16
3. Les méthodes de reconnaissance faciale .....	17
V. PROCEDURE DE MISE EN PLACE.....	18
1. ANALYSE DES BESOINS .....	18
1.1. Besoins Fonctionnels .....	18
1.2. Besoins non fonctionnels .....	19
1.3. Diagramme de cas d'utilisation .....	19
1.4. Diagramme de séquence .....	20
1.5. Diagramme de classe .....	21
2. CONCEPTION.....	21
2.1. DETECTION DE VISAGE.....	22
2.1.1. Classificateur de Haar .....	22
2.1.2. Classificateur en cascade LBP .....	24
2.1.3. Comparaison Classificateur HAAR et LBP .....	26
2.2. RECONNAISSANCE DE VISAGE .....	27
2.2.1. Prétraitement .....	27
2.2.1.1. Normalisation Géométrique .....	27
2.2.1.2. Normalisation Photométrique.....	28
2.2.2. Fonctionnement des Algorithmes Utilisés .....	28
3. IMPLEMENTATION .....	33
3.1. Environnement de travail .....	33
3.2. Constitution de la base de données à partir de la Webcam.....	33
3.3. Extraction des caractéristiques.....	34

3.4. Reconnaissance faciale .....	34
VI. CONCLUSION ET PERSPECTIVE .....	37

## Liste des Tableaux

---

<u>Tableau 1</u> : Tableau Comparatif 2D des différents algorithmes .....	13
<u>Tableau 2</u> : Tableau Comparatif 3D des différents algorithmes .....	14
<u>Tableau 3</u> : Tableau Comparatif 3D expressions des différents algorithmes.....	14
<u>Tableau 4</u> : Différents algorithmes de l'apprentissage supervisé.....	16
<u>Tableau 5</u> : Comparaison classificateur HAAR et LBP .....	26

## Liste des Figures

---

<u>Figure 1</u> : Relation entre la machine Learning et la vision par ordinateur .....	6
<u>Figure 2</u> : Rapport entre le deep Learning et la vision par ordinateur .....	8
<u>Figure 3</u> : Le processus de capture jusqu'à la reconnaissance faciale.....	15
<u>Figure 4</u> : Résumé des processus référencé avec les algorithmes .....	17
<u>Figure 5</u> : Diagramme de cas d'utilisation .....	19
<u>Figure 6</u> : Diagramme de séquence .....	20
<u>Figure 7</u> : Diagramme de classe .....	21
<u>Figure 8</u> : Les étapes de la conception de notre projet .....	21
<u>Figure 9</u> : Un couple caractéristique Haar-like .....	22
<u>Figure 10</u> : Différentes visualisation .....	23
<u>Figure 11</u> : Fenêtre LBP .....	24
<u>Figure 12</u> : Conversion de LBP en binaire. ....	25
<u>Figure 13</u> : Histogramme de LBP. ....	25
<u>Figure 14</u> : Exemple de Détection de visage avec l'algorithme de Viola - Jones .....	26
<u>Figure 15</u> : Principe de fonctionnement de base d'un système de reconnaissance faciale .....	27
<u>Figure 16</u> : Illustration de l'ACP .....	28
<u>Figure 17</u> : Approximation d'un créneau par des sinusoïdales .....	29
<u>Figure 18</u> : Quelques faces propres .....	30
<u>Figure 19</u> : Illustration de la distance de Mahalanobis.....	32

## Liste des Acronymes

---

HAAR : Alfréd Haar

LBP : Local Binary Patterns (Motifs Binaires Locaux)

SVM : Support Vector Machine

LBP : Motif Binaire Local

## **I. INTRODUCTION**

Le système de présence en vigueur à l'Institut de la Francophonie pour l'Innovation fonctionne manuellement et est prise en charge par l'enseignant du cours par l'appel des identifiants. Ce processus prend plus de temps pour la participation. Généralement, 3 à 5 minutes sont gaspillées dans un système de participation conventionnel. Cette perte de temps peut également engendrer la perturbation du planning des cours, le processus d'enseignement et d'apprentissage. Mais dans ce nouveau système, il ne faut que 1 à 2 secondes de présence.

L'objectif est de mettre en place une application capable de répondre à ces critères qu'est la reconnaissance faciale.

Pour réaliser notre application, nous allons suivre une méthodologie bien structurée qui est : l'analyse du sujet, sa description et enfin la solution pour aboutir à l'implémentation de l'application.

## II. ANALYSE DU SUJET

Ce sujet porte dans le domaine de l'apprentissage automatique (machine Learning) et la vision par ordinateur qui sont des sous branches de l'intelligence artificielle. La figure ci-dessous permet de voir les relations entre la machine Learning et la vision par ordinateur

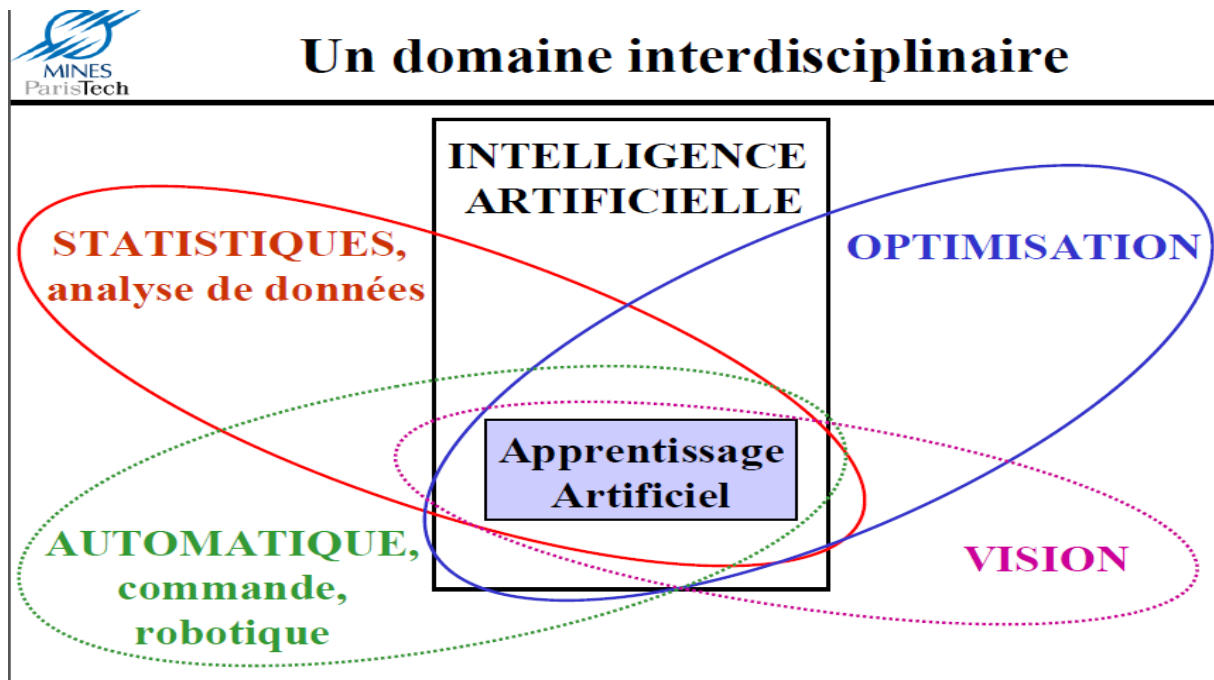


Figure 1: Relation entre la machine Learning et la vision par ordinateur

La machine Learning ainsi que la vision par ordinateur sont tous des domaines de recherche, la machine Learning quant à elle permet dans une certaine mesure à un système piloté par ordinateur (robot éventuellement) ou assisté par ordinateur, d'adopter ses analyses et ses comportements en réponse, en se fondant sur l'analyse de données empirique provenant d'une base de données ou de capteurs. Et la vision par ordinateur dont son principal but est de permettre à une machine d'analyser, traiter et comprendre une ou plusieurs images prises par un système d'acquisition (par exemple : caméras...)

Dans l'apprentissage automatique, les tâches sont généralement classées en grandes catégories. Ces catégories sont basées sur la façon dont l'apprentissage est reçu ou comment le feedback sur l'apprentissage est donné au système développé. Deux des méthodes d'apprentissage automatique les plus largement adoptées sont l'apprentissage supervisé qui forme des algorithmes basés sur des données d'entrée et de sortie étiquetées par l'homme et l'apprentissage non supervisé qui ne fournit pas à l'algorithme des données étiquetées pour lui permettre de trouver une structure et de découvrir une logique dans les données entrées.



L'apprentissage automatique est utilisé pour doter des ordinateurs ou des machines de systèmes de : perception de leur environnement : vision, reconnaissance d'objets (visages, schémas, langages naturels, écriture, formes syntaxiques...) ; moteurs de recherche ; aide aux diagnostics, médical notamment, bio-informatique, interfaces cerveau-machine ; détection de fraudes à la carte de crédit, analyse financière, dont analyse du marché boursier; classification des séquences d'ADN ; jeu ; génie logiciel ; adaptation de sites Web ; locomotion de robots ; analyse prédictive en matière juridique et judiciaire...

Les algorithmes les plus utilisés sont :

- ✓ Les machines à vecteur de support(SVM)
- ✓ Le boosting
- ✓ Les réseaux de neurones dont les méthodes d'apprentissage profond (deep Learning en anglais) pour un apprentissage supervisé ou non-supervisé
- ✓ La méthode des k plus proches voisins pour un apprentissage supervisé
- ✓ Les arbres de décision, méthode à l'origine des Random Forest, par extension également du boosting (notamment xgboost) ;
- ✓ La régression logistique
- ✓ Les algorithmes génétiques et la programmation génétique

Ces méthodes sont souvent combinées pour obtenir diverses variantes d'apprentissage. L'utilisation de tel ou tel algorithme dépend fortement de la tâche à résoudre (classification, estimation de valeurs...).

Plusieurs techniques sont utilisées dépendant du domaine. Mais nous allons nous orienter sur les techniques utilisés au niveau de la vision par ordinateur qui sont :

- ✓ Transformée de Hough
- ✓ Transformée de Radon
- ✓ Transformée de Fourier
- ✓ Apprentissage profond

Les problèmes généralement rencontrés dans le sujet sont :

- ✓ La détection des points de contraste (contours)
- ✓ La liaison des points de contours pour la formation des lignes de contours
- ✓ Le dessin des traits (droites, au sens des moindres carrés)
- ✓ La reconnaissance des objets observés

La figure ci-dessous explique le rapport entre la machine learning et la vision par ordinateur

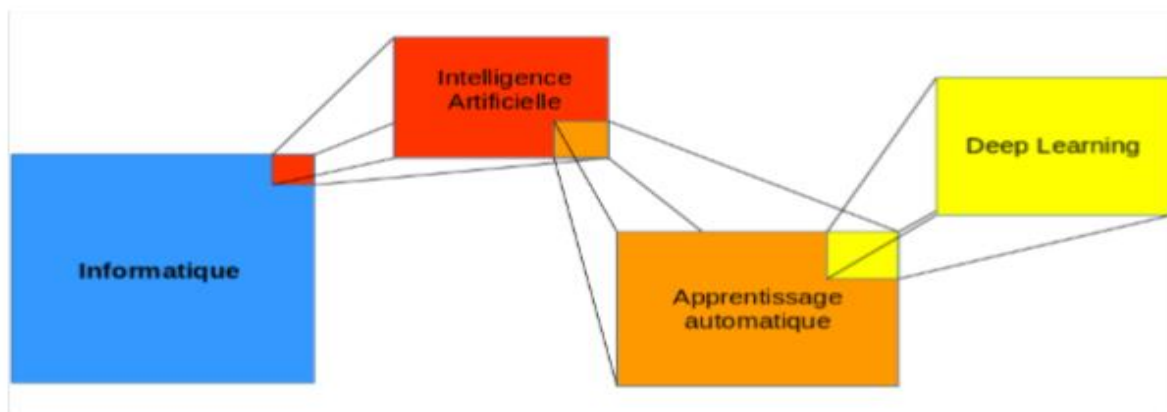


Figure 2: Rapport entre le deep Learning et la vision par ordinateur

Dans ce sujet de recherche pour aboutir à un résultat optimale, nous allons conjugués la machine Learning (technique d'apprentissage automatique) et la vision par ordinateur. Cette figure nous explique les rapports entre les différents sous domaines de l'informatique.

En tant que discipline technologique, la vision par ordinateur cherche à appliquer ses théories et ses modèles à différents systèmes. Quelques exemples de systèmes d'application de la vision par ordinateur :

- ✓ Procédés de contrôles, exemple dans la robotique industrielle.
- ✓ Navigation, exemple dans un véhicule autonome ou un robot mobile.
- ✓ Détection d'évènements, exemple pour la surveillance ou le comptage automatique de personnes.
- ✓ Organisation d'information, exemple pour indexer des bases de données d'images et de suites d'images.
- ✓ Modélisation d'objets ou d'environnements, exemple dans l'analyse d'image médicale ou la topographie.
- ✓ Interaction exemple en tant que dispositif d'entrée pour une interaction homme-machine.
- ✓ Inspection automatique exemple dans les domaines de production.

En résumé, la tâche qui nous attend est de réaliser une application qui pourra détecter et identifier le visage de chaque étudiant pour savoir si ce dernier est présent dans la salle de classe. Pour ce faire, nous utiliserons les techniques, les algorithmes et les outils liés à l'apprentissage automatique par la vision par ordinateur.

### III. RECHERCHE BIBLIOGRAPHIQUE

Cette seconde partie qu'est la Recherche Bibliographique, exigera de nous une connaissance précise et spécialisée du sujet, de comprendre les techniques existantes, de trouver les solutions possibles au sujet et enfin d'établir une comparaison des différentes approches du sujet.

#### 1. Connaissance précise et spécialisée du sujet

La mise en œuvre d'un système automatique et fiable de reconnaissance faciale est un verrou technologique qui n'est toujours pas résolu. La détection des visages joue un rôle crucial dans la reconnaissance des visages. C'est la première étape pour la reconnaissance automatique des visages. Un système de reconnaissance faciale est une application logicielle visant à reconnaître une personne grâce à son visage de manière automatique. Pour cela deux méthodes sont utilisées : la méthode 2D et la méthode 3D

#### 2. Les méthodes de reconnaissance faciale

Chacune des méthodes se base sur trois phases : la détection de visage qui repère un visage, l'analyse du visage qui donne un résultat numérique, et la reconnaissance qui compare ce résultat avec la **base de données** des visages enregistrés. Il faut donc, avant toute tentative d'identification, constituer une base de données contenant pour chaque utilisateur soit une capture initiale, soit plusieurs. Dans le cas où plusieurs captures sont effectuées, une moyenne de celle-ci est alors stockée, ce qui permet au système de mieux s'adapter à divers paramètres pouvant varier au moment de la reconnaissance (apparition de cernes/rides, luminosité ambiante, maquillage...). Une fois l'analyse effectuée, la reconnaissance effectue donc la comparaison entre le modèle obtenu avec ceux stockés en base de données. On parle alors de reconnaissance 1 à 1 si on avait une seule capture initiale stockée en base de données initialement, ou de reconnaissance 1 à N si on en avait plusieurs.

**La méthode 2D** : Considérée comme la méthode « classique » de reconnaissance faciale, cette méthode consiste à reconnaître l'utilisateur à partir d'une photo de lui. Cette photo peut être capturée par une caméra dans un but de sécurité, ou simplement être déjà enregistrée dans le cadre d'une assistance à l'utilisateur. La reconnaissance est alors effectuée par un algorithme qui peut s'appuyer sur différents éléments, comme la forme d'éléments du visage tels que les yeux et leur écartement, la bouche, le visage...

Deux catégories d'algorithmes sont alors distinguées :

- La première catégorie d'algorithmes crée une image géométrique de l'utilisateur en fonction de différents paramètres (tailles d'éléments du visage, forme et distance entre eux). Les paramètres récupérés sont alors encodés puis comparés avec ceux présents en base de données.
- La seconde catégorie d'algorithme encode numériquement l'image, en utilisant les algorithmes de Fourier, en utilisant des Eigenfaces pour créer des vecteurs de poids, ou encore via des calculs de moyennes sur certaines zones de l'image.

**La méthode 3D** : cette méthode est considérée comme une amélioration de la reconnaissance 2D. En effet, elle crée un modèle 3D depuis plusieurs photos effectuées successivement ou

depuis une vidéo, cela permettant d'avoir différents points de vue de la personne à reconnaître afin de créer le modèle 3D. De nombreux algorithmes existent en ce qui concerne l'analyse, chacun se basant sur un ou plusieurs éléments du visage (orientation du nez, du visage...) pour créer le modèle 3D correspondant au visage de l'utilisateur.

Toutefois, en raison de l'insuffisance intrinsèque de l'information faciale dans les images 2D, des algorithmes de reconnaissance de visage ont beaucoup de difficultés à la manipulation des variations d'éclairage, changements de pose, expressions faciales, temps de retard et des occlusions. La surface 3D du visage offre la possibilité de surmonter les difficultés causées par la variation de la tête de pose et d'éclairage. Avec les progrès de l'imagerie 3D techniques, les données 3D du visage sont devenues de plus en plus disponibles dans ces dernières années. Il y a maintenant plusieurs bases de données publiques de référence 3D du visage.

Dans cette section, nous présenterons un état de l'art sur les techniques de détection de visage 3D, nous donnerons une taxonomie des approches, des techniques, des algorithmes et des systèmes industriels existants. Nous exposerons aussi les méthodes de reconnaissance de visage 3D. La reconnaissance de personnes basée sur le visage 3D présente des limites. Elle ne peut faire une identification optimale en présence d'expressions faciales. Pour résoudre ce problème, une nouvelle méthode de visage 3D expressions a été introduite dont nous exposerons. Enfin, nous terminerons par une synthèse des méthodes et techniques étudiées.

### **3. Les Techniques ou méthodes existants**

Plusieurs méthodes de reconnaissance de visages ont été proposées ces dernières années, suivant deux grands axes: la reconnaissance à partir d'images fixes et la reconnaissance à partir de séquence d'images (vidéo). Les travaux portant sur ce domaine, sous différentes conditions d'éclairage, d'expressions faciales et d'orientations, peuvent être répertoriés en deux catégories distinctes suivant qu'elles se rapportent à une approche géométrique ou une approche globale. Les approches existantes peuvent être grossièrement divisées en trois groupes : les approches locales, les approches globales ainsi que les approches hybrides. Les principales méthodes de ces trois approches sont décrites dans la suite.

Nous insistons notamment sur les méthodes de réduction de dimension faisant partie des approches globales. Les méthodes de réduction de dimension entrent en effet dans le cadre plus vaste des traitements de données en général (et pas seulement des visages) et de la reconnaissance d'objets. Notons que la plupart des méthodes nécessitent une localisation précise du visage, un recadrage géométrique (pour que celui-ci apparaisse toujours dans la même position), ainsi qu'une mise à l'échelle pour que le visage ait une taille adéquate. Les techniques de reconnaissance 2D de visages robustes aux différents facteurs de variabilité (éclairage, pose, occultation) sont loin d'être développées. La reconnaissance 3D de visages constitue une alternative prometteuse pour surmonter ces problèmes, surtout depuis l'apparition de dispositifs d'acquisition 3D performant. L'avantage principal des approches basées modèle 3D réside dans le fait que le modèle 3D conserve toutes les informations sur la géométrie de visage, ce qui permet d'avoir une représentation réelle de ce dernier.

#### **Le modèle 3D**

Etant donné que ces modèles de visage 3D décrivent la géométrie du visage, ils sont plus sensibles que les images 2D de texture aux expressions faciales.

La reconnaissance du visage 3D basée sur le profil : La plupart de ces méthodes et produits sont conçus pour fonctionner avec les images de visages en frontales. Le profil de visage fournit des informations complémentaires de structure en forme de ce qui n'est pas visible dans la vue frontale, mais peut souvent être obtenu à partir d'autres vues. Selon une étude de Davidenko, le

profil silhouette des visages joue un rôle important dans la perception humaine de la tâche de reconnaissance de l'identité de genre. L'utilisation du profil de visage en biométrie est particulièrement attrayante pour les scénarios où seules les images de vision latérale sont disponibles. Le profil silhouette contient des informations plus fiables que la texture, ce qui rend les informations de profil particulièrement précieuses pour la reconnaissance.

Approches basées sur les caractéristiques géométriques faciales 3D : Dans ce cas l'étude est concentrée plus particulièrement sur les approches basées sur les courbes faciales et aux approches qui utilisent les techniques de sélection des caractéristiques, pour optimiser les performances de reconnaissance:

Approches basées sur les courbes : les méthodes d'extraction des courbes faciales 3D. Ces courbes peuvent être explicitement analysées et comparées. Plusieurs représentations de courbes ont été utilisées dans la littérature. Daoudi et al. Ils ont utilisé les courbes planes de niveaux. Les auteurs ont utilisé l'analyse de forme des courbes planes afin d'analyser les déformations du visage. En revanche, aucune étude n'a été faite pour la sélection des courbes pertinentes. Dans les mêmes auteurs ont utilisé la distance géodésique pour extraire les courbes de niveaux 3D. Selon l'étude de Bronstein et al. , les changements de la distance géodésique dus aux expressions faciales est négligeable comparée aux changements de la distance euclidienne. Donc les auteurs ont utilisé des points échantillons prélevés de l'intersection entre les courbes de niveaux et les profils radiaux, qui ont comme origine le bout du nez, ensuite ils ont calculé les distances euclidiennes entre les points correspondants des différents visages. Cette approche, permet de sélectionner des sous-ensembles de courbes faciales pour une mise en correspondance efficace du visage. Drira et al ont exploré l'utilisation de la forme de différents types de nez. Les auteurs extraient les courbes iso-géodésique, ensuite, ils ont construit les différents chemins géodésiques dans une variété Riemannienne de courbes fermées. La même approche permet de faire des calculs statistiques (moyennes) pour une recherche efficace dans des bases de visages 3D. Les mêmes auteurs ont proposé une approche similaire afin d'analyser la forme des courbes radiales 3D. Mpiperis et al. Ils proposent une para métrisation polaire, au sein de cette nouvelle para métrisation intrinsèque, les attributs du visage sont invariants aux déformations isométriques. Quand la bouche est ouverte, ce para métrisation n'est plus valable. Les auteurs proposent donc de déconnecter les lèvres sur toutes les surfaces faciales et changent ce para métrisation en se basant sur trois points de références.

Approches basées sur la sélection des caractéristiques : Plusieurs méthodes ont été proposées pour analyser le pouvoir discriminant des différentes parties du visage. Daniyal et Al. Ils ont proposé un algorithme dans lequel le visage est représenté comme un vecteur de distance entre les points d'intérêts (landmarks) des visages. Ils ont sélectionné les points par brute-forcing des combinaisons possibles des points utilisés/non-utilisés et ils ont comparé les taux de reconnaissance. Ils ont conclu que la meilleure sélection correspond aux points situés autour des yeux et du nez. Kakadiaris et Al. Ils utilisent un modèle de visage annoté qui est déformé élastiquement pour s'adapter à chaque visage, ainsi lui permettre automatiquement l'annotation de ses différentes zones comme le nez, yeux, bouche, etc. Pour résoudre le problème des expressions faciales, les auteurs classent les visages en utilisant les coefficients d'ondelettes qui représentent les zones non affectées par les expressions, telles que les yeux et le nez. Cependant, le meilleur taux de reconnaissance est atteint pour le visage entier, ce qui implique que le rejet de certaines zones, affectées par les déformations du visage, conduit à la perte de l'information discriminante. Un appariement de plusieurs régions du nez a été proposé. Les premiers résultats montrent une amélioration par rapport à une seule grande région frontale. Faltnier et al. Ils proposent de recalculer, en utilisant l'algorithme ICP, différentes parties du visage puis de les fusionner pour améliorer les performances biométriques. Récemment, Wang et al. Ils calculent la carte de différence de forme signée (SSDM), calculée entre deux visages alignés, comme une représentation intermédiaire pour la comparaison de formes. Basé sur le SSDM, trois types

d'opérateurs ont été utilisés pour coder à la fois la similarité locale et le changement des caractéristiques de différentes formes du visage. Les caractéristiques locales les plus discriminantes ont été sélectionnées, par AdaBoost, en les utilisant comme des classifieurs faibles. Les caractéristiques individuelles sont du type : Haar-like, Gabor et local binary pattern (LBP). En utilisant les caractéristiques similaires, Li et Al dans proposent de concevoir un dispositif commun et un système de classement afin de collecter différentes caractéristiques géométriques bas-niveau, telles que la courbure, et les classer selon leurs sensibilités aux expressions faciales. Pour les approches basées sur les courbes, l'analyse de la forme du visage, devient l'analyse des courbes faciales 3D. Toutes les courbes du visage donnent de meilleures performances. Approche visages 3D par combinaison de courbes et techniques de sélection des caractéristiques, On propose un nouvel algorithme de reconnaissance du visage 3D qui combine la représentation des visages 3D par une collection de courbes et de techniques de sélection des caractéristiques. Après l'acquisition des visages 3D on effectue quelques prétraitements afin d'extraire la partie informative du visage et corriger certaines imperfections telles que les points aberrants, les trous ou encore la présence de parties indésirables, comme les cheveux et les habits. On extrait les courbes faciales radiales et de niveaux. Puis, selon la nature des courbes (fermées ou ouvertes), on utilise l'un des cadres géométriques pour analyser les formes des courbes faciales. Cette approche géométrique est nécessaire pour comparer et analyser la forme des courbes faciales (ouvertes et fermées).

### **Reconnaissance de visages 3D expressions**

L'interface homme machine du futur utilisera de nouvelles sources d'informations telles que les expressions faciales de six cas (colère, dégoût, joie, peur, surprise, tristesse), la posture humaine ou parole.

#### **Système de reconnaissance d'expression faciale existant**

Ce système consiste à évaluer la possibilité de reconnaissance des six émotions universelles, en analysant les déformations des caractéristiques faciales telles que les yeux, les sourcils, et la bouche. On considère que ces caractéristiques contiennent suffisamment les informations pour la reconnaissance.

**Méthode basée sur la segmentation** : Dans l'étape de segmentation, le système va localiser les régions des yeux, des sourcils, et de la bouche. L'algorithme est présenté dans la thèse de Hammal. Dans l'étape d'extraction de données, les squelettes sont extraits. En fait, ce sont les contours des yeux, de la bouche, des sourcils. A partir des squelettes, on va choisir cinq distances pour caractériser les déformations. Les caractéristiques considérées contiennent les contours des yeux, des sourcils et des lèvres. Finalement, on obtient les squelettes d'expressions. A partir du squelette d'expression, on peut déterminer les déformations de caractéristiques faciales. Cinq distances sont définies pour chaque squelette. À partir de ces distances caractéristiques, on va utiliser le modèle de croyance transférable pour la reconnaissance d'expressions faciales. Après cela, chaque expression faciale est caractérisée par une combinaison d'états symboliques. Une combinaison d'états est définie pour chaque expression

**Méthode par masque pour le visage** : L'idée c'est d'utiliser un masque pour l'enregistrement dans une reconnaissance de visage. Ce système n'est pas nouveau. Par exemple, Blanz et Al utilisent un masque 3D pour enregistrer la vue frontale avec la vue de profil. Certains masques existants, sont bien conçus pour le montage surface déformable selon l'une des actions musculaires. Toutefois, lorsqu'on travaille avec des images face planes, il est difficile d'atteindre d'un bon enregistrement pour tous les sommets densément placés. En outre, bien que ces masques soient à morphologie souple. Les masques triangulaires sont avantageux dans un mappage de texture

Méthode par le modèle d'apparence (AAM) : Franck et Al ont travaillé sur la reconnaissance et synthèse d'expressions faciales par modèle d'apparence. Ils traitent une application nouvelle des modèles actifs d'apparence pour l'analyse et la synthèse de visages expressifs, ainsi que pour la reconnaissance d'expressions faciales. Ils considèrent les six expressions universelles que sont les expressions de joie, colère, peur, dégoût, tristesse et surprise. Après une description de ce type de modèle (calculé à l'aide de trois ACP ou d'une seule), ils proposent une méthode originale d'analyse et de synthèse permettant, à partir d'une seule photo, d'annuler l'expression d'un visage quelconque, puis de synthétiser une expression faciale artificielle sur ce même visage. Ils proposent pour cela deux approches permettant la modélisation des expressions faciales. Enfin, l'intérêt du modèle actif d'apparence pour la reconnaissance automatique d'expressions faciales est montré.

En somme pour obtenir un meilleur résultat, nous allons conjugués quelques algorithmes de des deux méthodes (2D, 3D et 3D expressions) et également la méthode par masque pour le visage.

Méthodes	BDD	Taille Images	Nb. Images	Time lapse	Taux (%)	Expression	Ill.	Position
Direct LDA	ORL	112 -92	200-200	N	90:8	O	O	O
DF-LDA	ORL	112 -92	200-200	O	96		O	N
	UMIST	112 -92	160-415	N	98		N	N
DCV	Yale	126-152	150-150	N	97:33		O	N
	AR	229-299	350-350	O	99:35			
ICA	FERET	60-50	425-421	O	89	O	N	N
PDBNN	SCR	80-20	320-1280	N	100	O	O	O
	FERET	200-200		N	99	O	O	N
	ORL			N	96		O	O
RBF	ORL	160-120	300-300		98:1	O		O
HMM	FERET	128-128	500-500	N	97	O	N	N
Gabor EFM	FERET	128-128	200-100	N	99	O	N	N
	ORL	128-128	200-200	N	100	O	N	O
EBGM	FERET	256-384	250-250	N	80	O		O
WPA	MIT	480-640	155-155		80:5	O	O	
	FERET	256-384	200-400		89			
IFS	ORL	112-92	200-	N	95			
IFS	MIT	480-640	90-90		90			O

Tableau 1: Tableau Comparatif 2D des différents algorithmes



Measure	Application	Method	Dimension	Reliability	Reference
Euclidean distance	Face recognition	PCA	3D	88,8%	Gupta et al 2007
		LDA	3D	97,60%	Gupta et al
	Face detection + landmark localization	Active Shape models(ASM)	2D	Fix error mean 6,73 pixels	Prabhu et al 19955
	Face expression recognition	PNN	3D	88%	Soyel et al 2008
		LDA +Gabor waveless	2D	74,1%	
		LDA+topographic		79,2%	
Geodesic distance	Face recognition	PCA	3D	89,9%	Gupta et al 2007
		New algorithm	3D	3,1%(EER for midly occluded face) : 5,5%	Bronstein et al 2006
Ratios	Face recognition	PCA	2D	75%-98% depending on the similitary  measure	Shi et al 2006
	Face expression recognition	LDA	3D	83,6%	Tang et al 2008
		Support Vector machine (SVM)		87,1%	

Tableau 2: Tableau Comparatif 3D des différents algorithmes

Source	Année	Base	Méthode	Taux de reconnaissance
Zilu Ying Beihang university, Beijing	2009	JAFFE	LBP et LE	70.48%
Ligang Zhang. Queensland University of Technology.Australia	2011	Le naturel d'expression visible et infrarouge du visage NVIE	FAP SIFT+mRMS	76.1% <b>83.4 %</b>
Anisha Halder. Jadavpur University	2011		GT2FS	96.67%
Amit Konar1. Jadavpur University	2012		IT2FS	88.66%
Rizwan Ahmed Khan. Université de Lyon	2012	CK (Cohn-Kanade)	SVM C4.5 Decision Tree RF	97.3% 97.6% <b>99.5%</b>

Tableau 3: Tableau Comparatif 3D expressions des différents algorithmes



## IV. SOLUTION PROPOSEE

Précédemment dans les deux premières parties du TPE, nous avons analysé le sujet pour comprendre ses enjeux, avons mené une recherche bibliographique sur les travaux existants et enfin dans cette troisième section nous allons proposer une solution du plan de travail qui nous emmènera à la partie pratique. Dans cette section, nous devons décrire la solution proposée, la justifier et enfin la méthodologie pour sa mise en pratique. Elle doit définir plus en détail le choix des outils, le choix des algorithmes et effectuer une comparaison des solutions existantes. Enfin, nous devons spécifier dans cette section comment notre travail sera élaboré clairement.

### 1. Description de la solution proposée

La première étape de notre étude constituera à former une base de donnée contenant pour chaque utilisateur (étudiant) plusieurs captures du visage. Une fois cela fait, l'enseignant active l'application installée sur son téléphone et il prendra un cliché de la classe entière. Cette image sera stockée pour un traitement ultérieur (reconnaissance de chaque visage sur l'image).

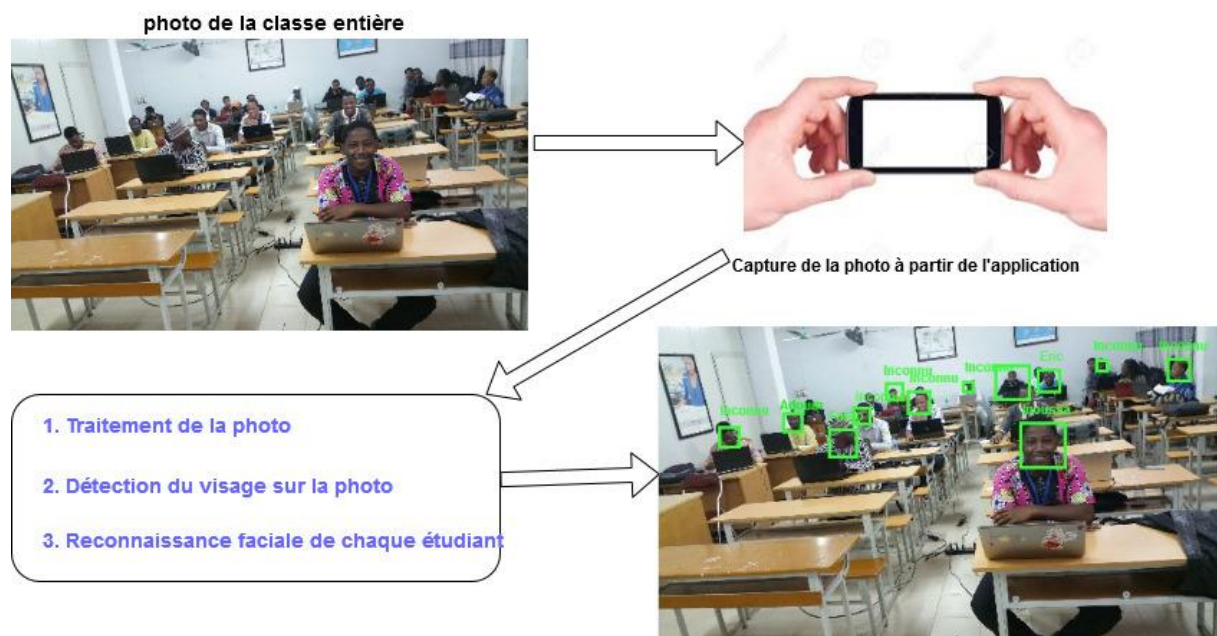


Figure 3: Le processus de capture jusqu'à la reconnaissance faciale

## 2. Les méthodes d'apprentissage automatique

Nous avons essentiellement trois types d'apprentissages automatiques : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement. Dans notre cas nous utiliserons la méthode supervisée car on va fournir à l'ordinateur des données en entrée pour entraîner le modèle pour que l'algorithme devienne capable de prédire sur de nouvelles données non annotées une fois entraîné. Le type de données (image) est une valeur discrète appelée classification ce qui nous pousse à utiliser la méthode supervisée.

<b>Apprentissage Supervisé</b>	
Méthodes d'apprentissages	Caractéristiques
Boosting	Regroupe de nombreux algorithmes, s'appuie sur des ensembles de classifieurs binaires, optimise les performances
Machine à vecteur de support	Destiné à résoudre des problèmes de discrimination et de régression
Réseau de neurones	Inspiration naturelle, processeur élémentaire calcul une sortie unique
Méthode des K plus proches voisins	Base de données d'apprentissage, Entrée $Y$ / $S(Y) = S$
Arbre de décision	s'appuie sur les observations
Classification naïve bayésienne	Requiert peu de données d'entraînement, suppose que l'existence d'une caractéristique est indépendante de l'existence d'autres caractéristiques
Espace de version	La technique des espaces de versions sert à déterminer, dans un espace d'hypothèses, lesquelles peuvent correspondre à un ensemble d'exemples pris dans l'espace des données

Tableau 4 : Différents algorithmes de l'apprentissage supervisé

### 3. Les méthodes de reconnaissance faciale

Pour notre approche, nous allons utiliser une stratégie de reconnaissance basée sur la combinaison 2D+3D. La première étape consiste à la détection de visage, pour cela nous utiliserons l'algorithme de Viola-Jones dû à sa rapidité de détection, il obtient un taux de détection de 88,8 % comparativement aux autres algorithmes de détection de visage. Après cette étape, nous utiliserons comme extraction des caractéristiques des descripteurs tels que LBP/HOG/CNN ces descripteurs permettent entre autre d'analyser les textures présentes dans l'image, de décrire le visage qu'elle contient, d'améliorer la performance de reconnaissance de visage, de classifier le visage d'individu... Ensuite nous entamerons la troisième étape qu'est la classification des visages en utilisant comme algorithme le 1-NN pour la règle du plus proche voisin et l'utilisation de machine à vecteur de support (Support Vector Machine, SVM). Enfin selon un seuil de reconnaissance, une décision sera prise si l'image a été acceptée ou refusée. L'environnement de développement serait Spyder, les librairies qui seront utilisés sont OpenCV, dlib, face\_recognition et comme langage de programmation python. La figure ci-dessous résume tout ce que nous avons détaillé ci-dessus.

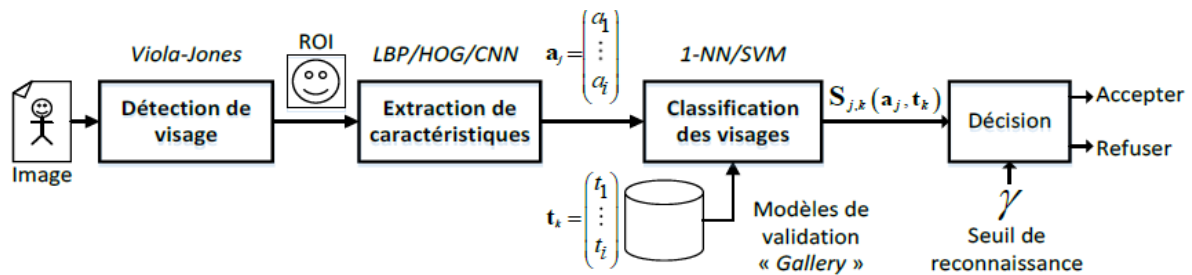


Figure 4: Résumé des processus référencé avec les algorithmes

## V. PROCEDURE DE MISE EN PLACE

Ce module comprend trois étapes principales: le prétraitement d'image, la détection de visage et la reconnaissance de visage. Lors du prétraitement de l'image, l'image capturée par la caméra sera débranchée, redimensionnée si nécessaire et convertie en niveaux de gris pour faciliter le traitement ultérieur. L'algorithme de détection des visages en cascade de Haar est ensuite appliqué pour détecter les visages dans la salle de classe, tandis que les approches Eigenface et Fisherface sont utilisées pour former et reconnaître les visages des étudiants.

Le principal problème dans la détection et la reconnaissance d'un objet est relatif aux différentes représentations possibles de celui-ci. Ainsi la détection et la reconnaissance du visage dépendent de plusieurs facteurs, dont les plus étudiés sont :

- ✓ La position : sur une image, un visage peut être vu de face, de profil, ou d'un angle quelconque.
- ✓ L'expression faciale : l'apparence d'un visage dépend aussi de son expression.
- ✓ La présence d'attributs : une personne peut avoir un chapeau, des lunettes, une moustache, une barbe, une cicatrice....
- ✓ Les conditions extérieures : la couleur, l'intensité de l'éclairage, la taille, la texture sont différentes sur chaque image.
- ✓ L'occultation : une partie du visage peut être cachée par un autre objet ou par une autre personne.
- ✓ La couleur : Les êtres humains ont des couleurs de peau différentes, d'où la différence de la valeur du pixel représentant la peau de chaque personne.

### 1. ANALYSE DES BESOINS

Nous allons définir l'analyse des besoins fonctionnels et non fonctionnels du module à développer. Cette analyse sera appuyée par les diagrammes de cas d'utilisation et de séquence système

#### 1.1. Besoins Fonctionnels

Notre application se répartit en deux modules fondamentaux :

- ✓ Module de détection de visage (obtenir des visages à partir la caméra)
- ✓ Module de reconnaissance de visage

Ces modules représentent les fonctionnalités les plus importantes de notre application dont les besoins fonctionnels sont

#### ❖ La détection de visage

Le système doit détecter le visage dans une image. La détection consiste à délimiter une zone rectangulaire contenant la partie de l'image reconnue comme un visage.

#### ❖ La construction d'une base de descripteurs associés aux visages

Dans cette étape nous extrairons de l'image les informations qui seront sauvegardées en mémoire pour être utilisées plus tard dans la phase de reconnaissance. Les informations doivent être choisies de manière à être discriminantes et non redondantes.

### ❖ La reconnaissance de visage

La reconnaissance de visage permet de savoir si un nouvel individu est déjà enregistré dans la base ou non. Cette phase est l'aboutissement du processus, elle peut être valorisée par un taux de fiabilité qui est déterminé par le taux de justesse de la décision.

## 1.2. Besoins non fonctionnels

### ❖ Interface utilisateur

L'application devra être cohérente du point de vue de l'ergonomie. La qualité de l'ergonomie sera un facteur essentiel, étant donné l'utilisation intensive qui sera faite de l'application.

### ❖ Sécurité

L'authentification, l'identification de l'utilisateur pour accéder à l'interface de l'utilisateur.

### ❖ Documentation

Une documentation écrite devra accompagner le système de reconnaissance faciale, qui décrit de manière claire et précise son fonctionnement.

### ❖ Rapidité

Le temps d'exécution de l'application ne devra pas dépasser 5 secondes.

### ❖ Configuration requise par l'application

La configuration minimale adoptée est celle du Pentium III 800 Mhz avec 1 Go de RAM. La résolution d'affichage minimale sera de 800x600 en 256 couleurs.

Après avoir identifié les besoins fonctionnels et non fonctionnels du module détection et reconnaissance de visages, nous allons identifier les principaux cas d'utilisation du système, afin de bien concevoir notre solution.

## 1.3. Diagramme de cas d'utilisation

Les cas d'utilisation sont un moyen qui permet de capturer les exigences fonctionnelles d'un système. Ils décrivent les interactions entre les acteurs et le système et constituent une méthode pour recueillir et décrire les besoins des acteurs

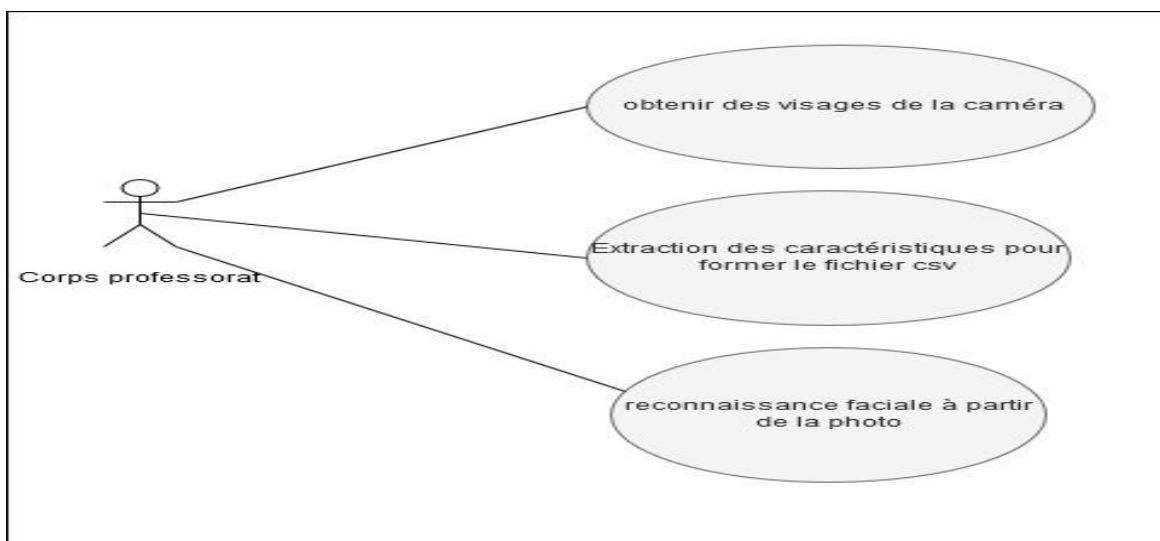


Figure 5: Diagramme de cas d'utilisation

### Identification d'un étudiant:

- ✓ Acteur : Corps professorat
- ✓ Description : Le professeur doit faire une capture de plusieurs photos de chaque étudiant pour former la base de données. Ensuite il y a l'extraction pour former le fichier csv dans lequel on aura les caractéristiques de chaque étudiant et enfin la reconnaissance faciale à partir d'une photo prise dans la salle de classe par le téléphone de l'enseignant dont l'application est initialement installée, cette dernière permettra de reconnaître les visages sur la photo et d'établir une liste de noms des faces présentes sur la photo, c'est cette liste qui constituera la liste de présence des étudiants.

### 1.4.Diagramme de séquence

Ce diagramme décrit la formation de la base de données par un utilisateur :

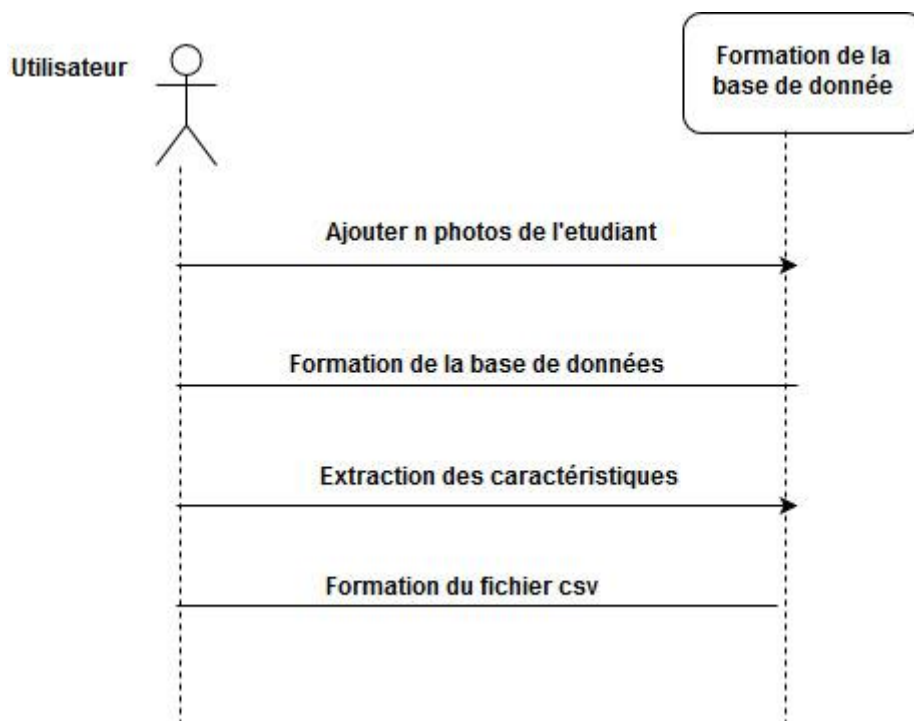


Figure 6 : Diagramme de séquence

## 1.5. Diagramme de classe

Ce diagramme présente les classes et les interfaces des systèmes ainsi que les différentes relations entre elles.

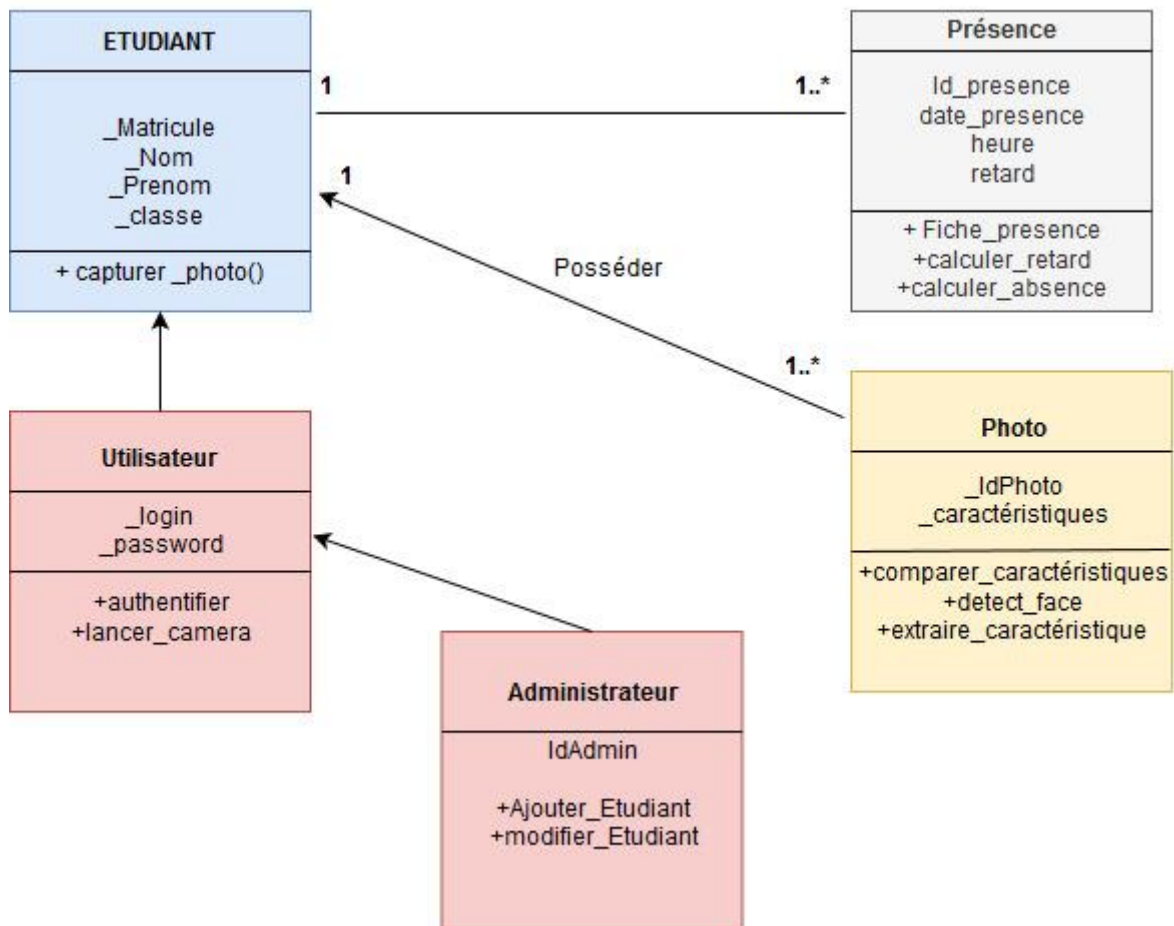


Figure 7: Diagramme de classe

## 2. CONCEPTION

La figure ci-dessous explique les grandes lignes de la conception de notre projet

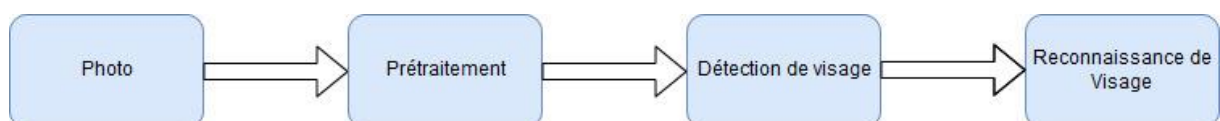


Figure 8: Les étapes de la conception de notre projet

## 2.1. DETECTION DE VISAGE

Un programme informatique qui détermine si une image est une image positive (image de visage) ou négative (image de non visage) est appelé un **classificateur**. Un classificateur est formé sur des centaines de milliers d'images de visage et non visage pour apprendre à classer correctement une nouvelle image. OpenCV nous fournit deux classificateurs de détection de visage préformés et prêts à être utilisés:

- ✓ Classificateur de Haar (algorithme de Viola-Jones)
- ✓ Classificateur LBP

Ces deux classificateurs traitent les images en niveaux de gris, essentiellement parce que nous n'avons pas besoin d'informations sur les couleurs pour déterminer si une image a un visage ou non. Comme ils sont préformés à OpenCV, leurs fichiers de connaissances sont également fournis avec OpenCV `opencv / data /`.

Pour exécuter un classificateur, nous devons d'abord charger les fichiers de connaissances, comme s'il n'y avait aucune connaissance, comme un nouveau-né.

Chaque fichier commence par le nom du classificateur auquel il appartient. Par exemple, un **classificateur en cascade Haar** commence comme `haarcascade_frontalface_alt.xml`.

### 2.1.1. Classificateur de Haar

#### 2.1.1.1. Fonctionnement

Le classificateur de Haar est une approche basée sur l'apprentissage automatique, un algorithme créé par Paul Viola et Michael Jones, qui (comme mentionné précédemment) sont formés à partir de nombreuses images positives (avec des visages) et négatives (sans visages).

Cela commence par extraire les caractéristiques de Haar de chaque image comme indiqué par les fenêtres ci-dessous:

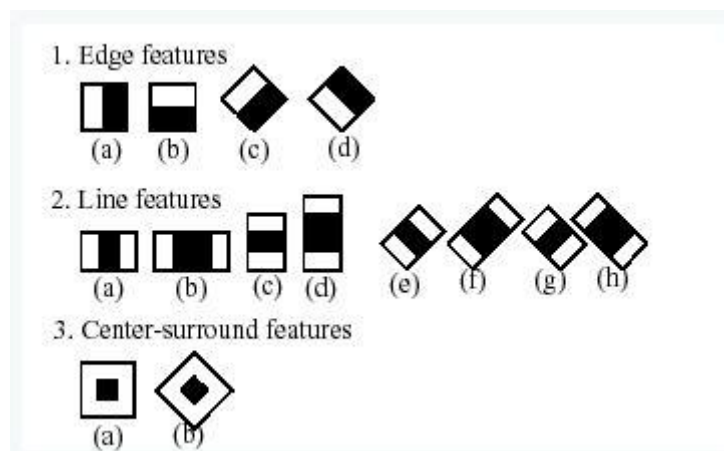


Figure 9: Un couple caractéristique Haar-like



Chaque fenêtre est placée sur l'image pour calculer une seule caractéristique. Cette caractéristique est une valeur unique obtenue en soustrayant la somme des pixels sous la partie blanche de la fenêtre de la somme des pixels sous la partie noire de la fenêtre.

Désormais, toutes les tailles possibles de chaque fenêtre sont placées à tous les emplacements possibles de chaque image pour calculer de nombreuses fonctionnalités.

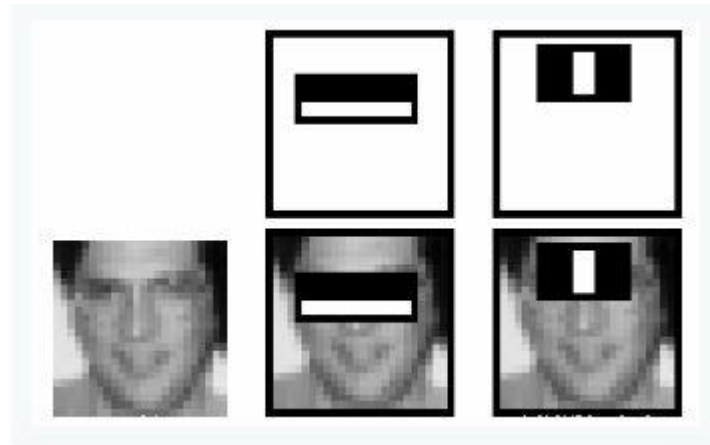


Figure 10: Différentes visualisation

Par exemple, dans l'image ci-dessus, nous extrayons deux caractéristiques. La première se concentre sur le fait que la région des yeux est souvent plus sombre que la région du nez et des joues. La deuxième caractéristique repose sur la propriété que les yeux sont plus sombres que le pont du nez.

Mais parmi toutes ces caractéristiques calculées, la plupart d'entre elles ne sont pas pertinentes. Par exemple, lorsqu'elles sont utilisées sur la joue, les fenêtres deviennent inutiles car aucune de ces zones n'est plus foncée ou plus claire que les autres zones situées sur la joue, tous les secteurs sont identiques.

Par conséquent, nous éliminons rapidement les fonctionnalités non pertinentes et gardons uniquement celles pertinentes avec une technique sophistiquée appelée **AdaBoost**. AdaBoost est un processus de formation à la détection de visage, qui sélectionne uniquement les fonctionnalités connues pour améliorer la précision de classification (visage / non-visage) de notre classificateur.

En fin de compte, l'algorithme considère le fait que généralement la majeure partie de la région d'une image est une région sans visage. Compte tenu de cela, il est préférable de disposer d'une méthode simple pour vérifier si une fenêtre est une région sans visage, et si ce n'est pas le cas, supprimez-la immédiatement et ne la traitez plus. Nous pouvons donc nous concentrer principalement sur la zone où se trouve un visage.

### 2.1.2. Classificateur en cascade LBP

Comme tout autre classificateur, les modèles binaires locaux, ou LBP en bref, doivent également être formés sur des centaines d'images. LBP est un descripteur visuel / texture et, heureusement, nos visages sont également composés de micro-motifs visuels.

Ainsi, les caractéristiques LBP sont extraits pour former un vecteur de caractéristiques qui classifie un visage à partir d'un non-visage.

"Mais comment trouve-t-on les fonctionnalités de LBP?"

Chaque image d'entraînement est divisée en plusieurs blocs, comme indiqué dans l'illustration ci-dessous.

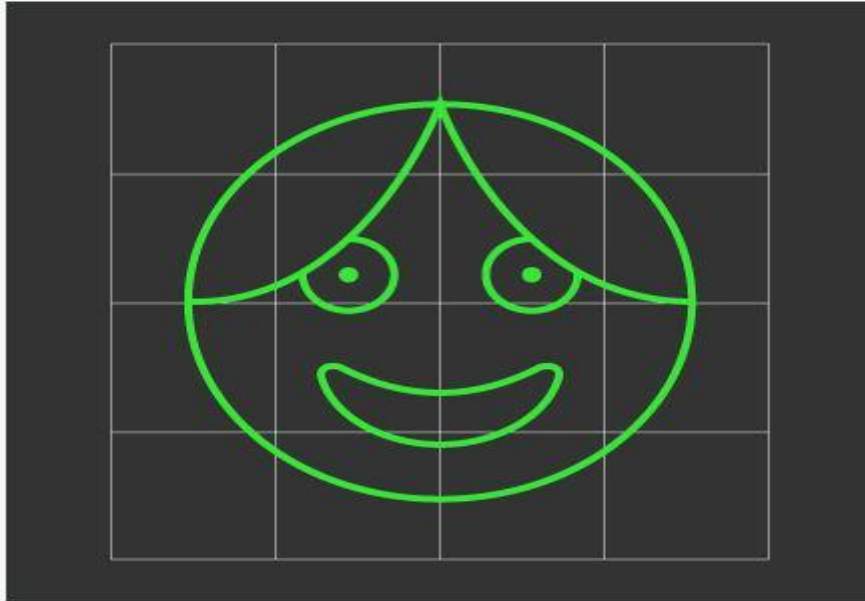


Figure 11: Fenêtre LBP

Pour chaque bloc, LBP examine 9 pixels (fenêtre  $3 \times 3$ ) à la fois et s'intéresse particulièrement au pixel situé au centre de la fenêtre.

Primo, il compare la valeur de pixel centrale à la valeur de pixel de chaque voisin sous la fenêtre  $3 \times 3$ . Pour chaque pixel voisin supérieur ou égal au pixel central, il définit sa valeur sur 1 et les autres sur 0.

Secundo, il lit les valeurs de pixel mises à jour (qui peuvent être 0 ou 1) dans le sens des aiguilles d'une montre et forme un nombre binaire. Tertio, il convertit le nombre binaire en un nombre décimal, et ce nombre décimal est la nouvelle valeur du pixel central. Nous faisons cela pour chaque pixel d'un bloc.

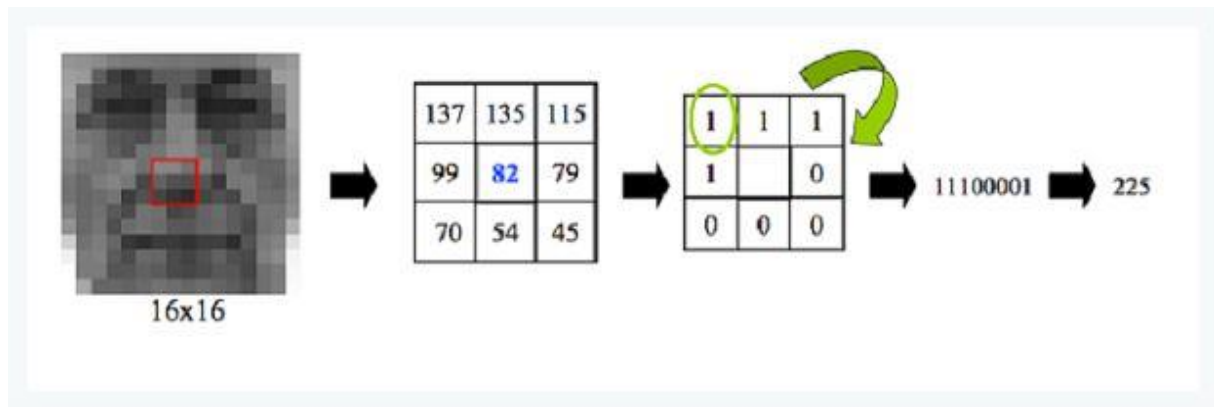


Figure 12 : Conversion de LBP en binaire.

Source: López & Ruiz; Modèles binaires locaux appliqués à la détection et à la reconnaissance de visages.

Ensuite, il convertit chaque valeur de bloc en un histogramme. Nous avons donc maintenant un histogramme pour chaque bloc d'une image, comme ceci:

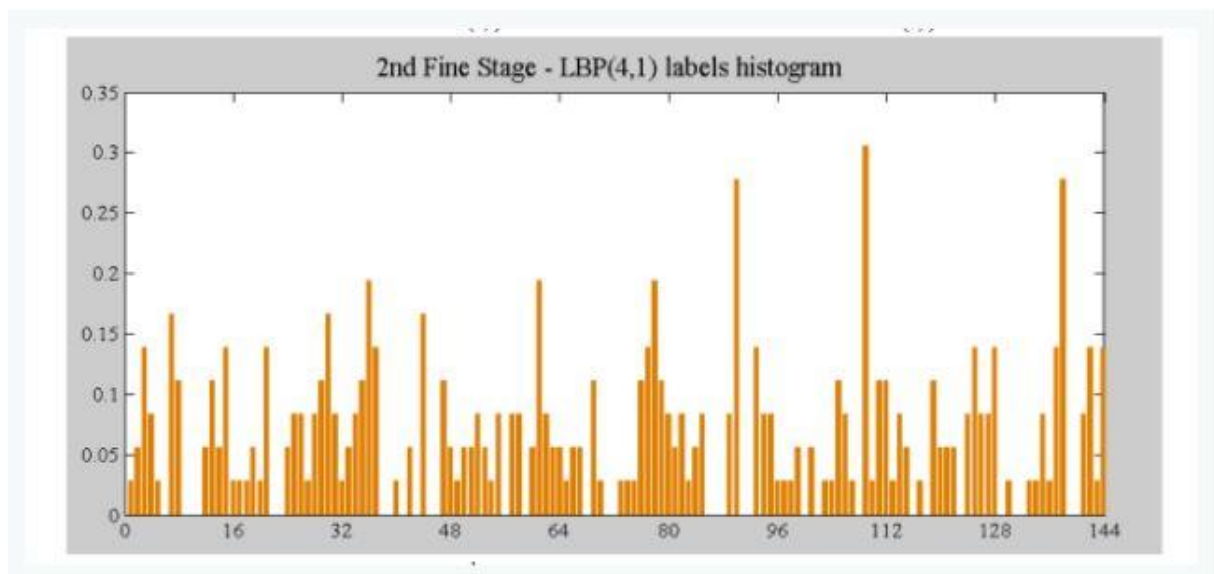


Figure 13 : Histogramme de LBP.

Source: López & Ruiz; Modèles binaires locaux appliqués à la détection et à la reconnaissance de visages.

Enfin, il concatène ces histogrammes de blocs pour former un vecteur d'une caractéristique pour une image, qui contient toutes les caractéristiques qui nous intéressent. Voici comment nous extrayons les caractéristiques LBP d'une image.

### 2.1.3. Comparaison Classificateur HAAR et LBP

Le Tableau ci-dessous décrit les avantages et les désavantages de chaque classificateur

Algorithmes	Avantages	Désavantages
Haar	<ul style="list-style-type: none"><li>✓ Haute précision de détection</li><li>✓ taux de faux positifs faible</li></ul>	<ul style="list-style-type: none"><li>✓ Calcul complexe et lent</li><li>✓ Temps d'entraînement long</li><li>✓ moins précis sur les visages noirs</li></ul>
LBP	<ul style="list-style-type: none"><li>✓ Calcul simple et rapide</li><li>✓ Temps d'entraînement court</li><li>✓ Robuste en cas de changement brusque d'illumination</li><li>✓ Robuste pour l'occlusion</li></ul>	<ul style="list-style-type: none"><li>✓ Moins précis</li><li>✓ Taux de faux positif élevé</li></ul>

Tableau 5 : Comparaison classificateur HAAR et LBP

Chaque classificateur de détection de visage OpenCV a ses avantages et ses inconvénients, mais les différences majeures concernent la précision et la vitesse. Donc, au cas où des détections plus précises sont nécessaires, le classificateur de Haar est la voie à suivre. Ce mauvais garçon est plus approprié dans les technologies telles que les systèmes de sécurité ou le harcèlement criminel haut de gamme. Mais le classificateur LBP est plus rapide et doit donc être utilisé dans les applications mobiles ou les systèmes intégrés.

Dans notre cas, nous utiliserons l'algorithme Haar de Viola – Jones

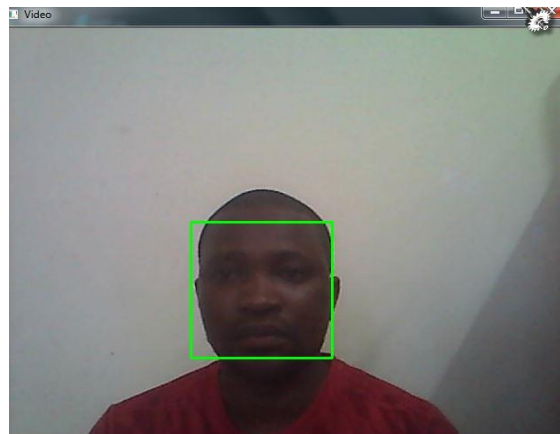


Figure 14 : Exemple de Détection de visage avec l'algorithme de Viola - Jones

## 2.2. RECONNAISSANCE DE VISAGE

Comme nous avons déjà fait la comparaison entre les algorithmes au niveau de la recherche bibliographique, et également avons défini les méthodes et techniques à appliquer pour la reconnaissance faciale. Nous allons décrire dans cette partie, la méthodologie pour obtenir un seuil de reconnaissance acceptable avec les algorithmes utilisés.

C'est dans cette phase que nous allons attribuer un nom à chaque visage. Pour ce faire une phase de prétraitement est nécessaire

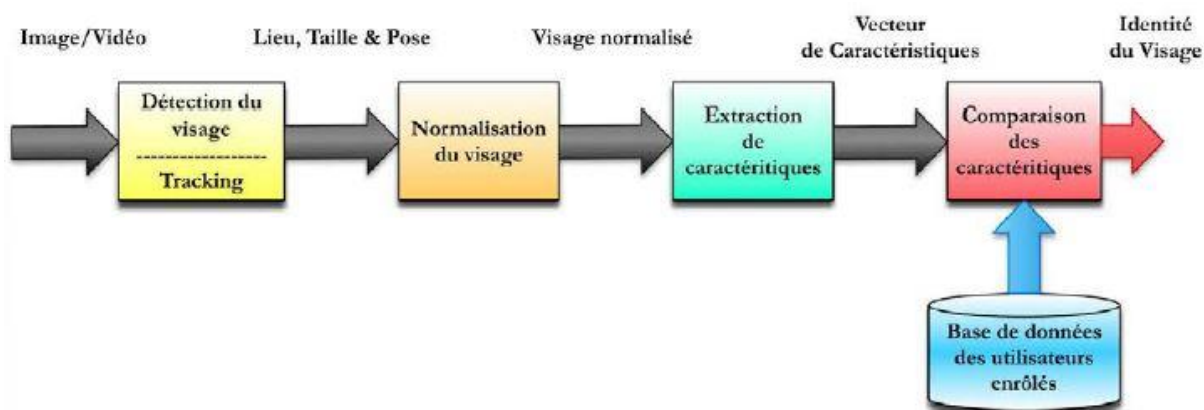


Figure 15: Principe de fonctionnement de base d'un système de reconnaissance faciale

### 2.2.1. Prétraitement

La phase de prétraitement vient après la phase de détection. Elle permet de préparer l'image du visage de telle sorte qu'elle soit exploitable. On l'appelle aussi phase de normalisation puisqu'elle ramène à un format prédéfini toutes les images extraites de l'image brute. Elle consiste généralement en un centrage du visage dans l'image et une élimination des zones non informatives. Pour garantir la bonne performance du système de reconnaissance de visages, il est important que toutes les images soient de taille identique, à la même échelle et au même format concernant les couleurs (par exemple, les images couleur sont parfois converties en niveaux de gris). Ceci améliore incontestablement le fonctionnement de l'étape d'extraction de signatures et par conséquent la qualité de cette dernière. La normalisation est constituée de deux processus : géométrique et photométrique. La normalisation géométrique est nécessaire parce que la taille du visage à l'intérieur de l'image acquise peut varier en fonction de la distance entre le module d'acquisition et la personne. L'étape de normalisation photométrique tente d'éliminer ou de réduire les effets de l'illumination de l'image.

#### 2.2.1.1. Normalisation Géométrique

En utilisant un algorithme de reconnaissance se basant sur la réduction de l'espace, nous ne pouvons pas négliger un point très important qui est la normalisation géométrique des images

de visage. Cette normalisation géométrique consiste à extraire la zone du visage de l'image originale, ensuite une rotation du visage est effectuée afin d'aligner l'axe des yeux avec l'axe horizontal. Enfin, une réduction proportionnelle à la distance entre les centres des deux yeux est appliquée. On obtient alors une image de visage dont la distance entre les centres des yeux est fixe. Les dimensions de l'image du visage sont calculées à partir de la distance à obtenir entre les centres des deux yeux.

#### **2.2.1.2. Normalisation Photométrique**

Dans le domaine de la reconnaissance par le visage, un certain nombre de méthodes de normalisation de l'illumination ont été présentées. Nous avons opté pour la méthode de l'égalisation d'histogramme qui a amélioré le taux d'identification de notre système de reconnaissance.

#### **2.2.2. Fonctionnement des Algorithmes Utilisés**

Nous cherchons à effectuer de la reconnaissance faciale à partir d'une photo préalablement prise par l'utilisateur, nous partons donc de la remarque, qui nous simplifie grandement notre démarche, selon laquelle nous pratiquons de la reconnaissance à deux dimensions.

La méthode de reconnaissance faciale Eigenfaces emploie la technique de l'analyse en composante principale, qui marque une différence notable avec les méthodes plus classiques, appelées méthodes géométriques ou locales, qui se basent sur les particularités du visage analysé, et dont les défauts résident dans son manque de précision, ainsi que sa sensibilité aux informations qui ne sont pas pertinentes.

La méthode que nous avons utilisé est qualifiée de globale, puisque l'ensemble du visage est alors analysé.

Notre technique de reconnaissance va donc utiliser la méthode d'analyse en composantes principales (également dite ACP). De manière simple, nous visons la diminution de la dimension de l'espace dans lequel nous allons travailler, et nous pourrions alors simplifier les données à notre disposition et leur interprétation.

Ainsi nous pourrions prendre en compte les informations importantes qui nous permettront de reconnaître un visage parmi d'autres avec un bon taux de réussite.

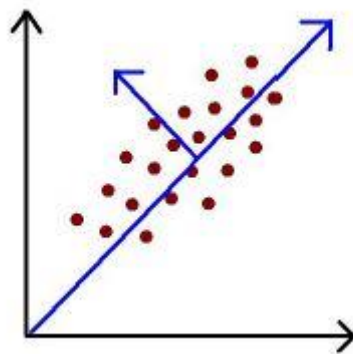


Figure 16 : Illustration de l'ACP

Faisons un simple parallèle entre ma méthode des faces propres et les séries de Fourier. Les séries de Fourier nous permettent généralement de représenter un signal périodique à l'aide de sommes de cosinus et de sinus.

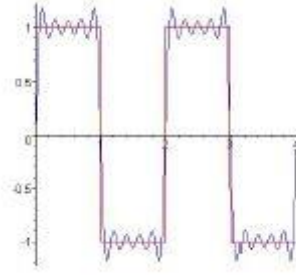


Figure 17 : Approximation d'un créneau par des sinusoïdales

Dans le cas de notre projet, notre approche consiste à représenter un visage comme étant la combinaison linéaire d'un ensemble d'images, ces dernières formant une base de référence. Mathématiquement, cela revient à parvenir à l'équation :

$$\Phi_i = \sum_{i=1}^n p_i d_i$$

Où  $d_i$  représente le visage propre, et  $p_i$  le coefficient associé.

Nous allons chercher à trouver les visages propres ; tout d'abord, nous devons prendre un nombre  $M$  de visages d'apprentissage. Chacune de ces images, qui sont en pratique des matrices  $N \times N$  sont alors transformées en une unique vecteur colonne de longueur  $N^2$ .

Matrice  $N \times N$  initiale :

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{1i} & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & \alpha_{2i} & \alpha_{2N} \\ \dots & \dots & \dots & \dots \\ \alpha_{N1} & \alpha_{N2} & \alpha_{Ni} & \alpha_{NN} \end{pmatrix}$$

transformée:

$$\begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{N1} \\ \vdots \\ \alpha_{1N} \\ \vdots \\ \alpha_{NN} \end{pmatrix}$$

Nous devons par la suite déterminer le visage moyen, déduit des  $M$  visages d'apprentissages.

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i$$

Ce visage moyen va servir dans l'analyse d'images, on soustrait en effet ce visage moyen aux visages d'apprentissages, ce qui nous laisse alors les informations propres à ce visage, nous récupérons alors dans  $\phi_i$  uniquement les informations qui sont particulières à ce visage d'apprentissage.

$$\Phi_i = \Gamma_i - \psi$$

Où  $\Phi_i$  représente le  $i^{\text{ème}}$  visage auquel on a soustrait le visage moyen.

A présent nous devons calculer la matrice de covariance D. Elle correspond à

$$D = QQ^T$$

$$\text{Avec } Q = [\Phi_1, \Phi_2, \dots, \Phi_M]$$

Nous devrions calculer les vecteurs propres  $d_i$  de la matrice D. Mais cela représente pour nous  $N^2$  vecteurs propres de dimension  $N^2$  chacun. C'est à présent que nous allons réduire l'information en limitant les composantes avec lesquelles nous travaillerons, en accord avec le principe de l'analyse en composantes principales.

Nous allons donc considérer la matrice  $E = Q^T Q$ , dont nous trouverons les vecteurs propres  $e_i$ . Cette matrice est de taille  $M \times M$  ce qui nous simplifiera donc les choses étant donné que nous aurons M vecteurs propres de taille M chacun. Le passage de la matrice D à la matrice E n'est pas anodin, nous utilisons en effet le fait que les vecteurs propres de ces deux matrices sont liés de manière assez proche.

En effet, nous avons comme relation,

$$Ee_i = Q^T Q e_i = \lambda_i e_i$$

Avec  $\lambda_i$  la valeur propre associée au vecteur propre  $e_i$

En multipliant cette équation par la matrice Q, il vient  $Q E e_i = Q Q^T Q e_i$

Nous voyons alors apparaître la matrice D donne :  $Q E e_i = D Q e_i = \lambda_i Q e_i$

Nous en déduisons donc qu'avec  $e_i$  vecteur propre de la matrice E associé à la valeur propre  $\lambda_i$ , nous avons par conséquent  $Q e_i$  est un vecteur propre de la matrice D associé à la même valeur propre  $\lambda_i$ .

Ainsi, nous avons  $d_i$  vecteur propre de D, avec

$$d_i = Q e_i$$

Ce sont les valeurs propres qui leur sont associées qui nous permet ensuite de classer les vecteurs en fonction de leur capacité à caractériser les variations entre les images.

Lorsque l'on les visualise (ces vecteurs sont à l'origine des matrices de taille  $N \times N$ ), les faces propres sont ce que l'on pourrait appeler des images aux airs fantomatiques. Mais gardons à l'esprit que sont les vecteurs propres de la matrice de covariance des images d'apprentissage des visages



Figure 18 : Quelques faces propres

Les M vecteurs propres que nous avons alors obtenus nous permettrons donc d'approximer au mieux les visages d'apprentissage en utilisant les visages propres de plus grande importance.



L'avantage de réduire le nombre de visages propres est d'une part de nécessiter de moins d'espace mémoire, mais aussi de réduire les calculs, leur temps d'exécution ; cependant nous perdons sans aucun doute de l'information et donc l'information moins précise, mais les résultats ne s'en verront pas vraiment modifiés, étant donné que nous ne nous donnons qu'une mission d'identification.

Nous ne cherchons pas à reconstruire le visage du sujet à partir de nos visages propres, mais seulement à le reconnaître.

Parmi les  $M$  vecteurs propres trouvés, nous allons seulement conserver un nombre  $L$ , qui seront les plus significatifs.

Nous allons trouver maintenant le poids associé à chacun des visages propres. Les images servant à l'apprentissage, auquel on a enlevé l'image moyenne, sont en fait combinaison linéaire des visages propres.

$$\Phi_i = \sum_{i=1}^L p_i d_i$$

Pour trouver le poids associé, nous faisons pour chacune des coordonnées correspondant à un visage d'apprentissage

$$p_i = d_i^T \Phi_i$$

Ce qui nous permet d'obtenir pour chacun des  $M$  visages d'apprentissages un vecteur  $\Pi_i$ , où  $i$  représente le  $i^{\text{ème}}$  visage, et qui nous informe sur le coefficient appliqué à chacun des visages propres.

$$\Pi_i = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_L \end{pmatrix}$$

Passons à présent au travail à effectuer pour la reconnaissance d'un visage d'un sujet.

Une fois l'image prise, l'image (vecteur colonne  $\Gamma$ ) obtenue est soustraite à l'image moyenne  $\psi$  :

$$\Phi = \Gamma - \psi$$

Puis nous trouvons les coordonnées de cette image dans l'espace réduit des faces propres

$$p_i = d_i^T \Phi_i$$

Ce qui nous donne au final :

$$\Pi = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_L \end{pmatrix}$$

Il nous faut maintenant interpréter la projection de l'image à analyser pour identifier le sujet. Pour cela nous allons utiliser une mesure de distance particulière, la distance de Mahalanobis.

L'intérêt de cette distance réside dans le fait qu'elle va accorder un poids moins important aux composantes bruitées, et qu'elle permet séparer efficacement les axes pour lesquels l'information peut être mieux classifiée.

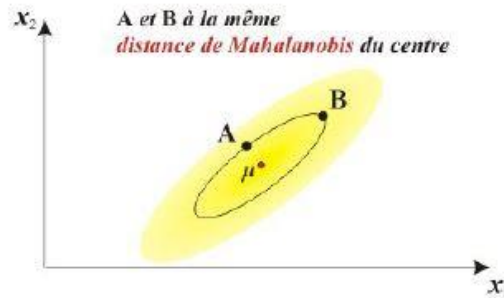


Figure 19: Illustration de la distance de Mahalanobis

Elle est définie par :

$$d(a, b) = \sqrt{(a - b)^T Q^{-1} (a - b)}$$

Avec Q covariance des variables.

Nous cherchons donc :  $m = \min ||\mathbf{II} - \mathbf{II}_i||$

Puis, nous comparons la valeur de m trouvée à une valeur seuil Delta, qui aurait dû être déterminée à partir d'essais sur des images choisies aléatoirement, qui peuvent aussi bien représenter des visages qu'autres choses, puis en comparant ces valeurs aux valeurs obtenues avec des visages d'apprentissage, et décider du seuil que nous avons choisi. Mais le choix de ce seuil dépend de trop nombreuses conditions (prise de vue des images, niveau de précision souhaité pour la reconnaissance, etc..) c'est pour cela qu'en pratique nous avons décidé de ne pas prendre de seuil.

L'utilisation d'un seuil nous aurait permis de déterminer si oui ou non l'image analysée correspond à un visage présent dans la base de données. Si alors la valeur de m était inférieure à celle du seuil, l'image correspond au visage qui a donné cette valeur la plus basse.

Dans notre cas, le fait de ne pas choisir de seuil a eu pour conséquence que lorsqu'un visage non présent dans la base de données était testé, il était tout de même reconnu par le programme.

### 3. IMPLEMENTATION

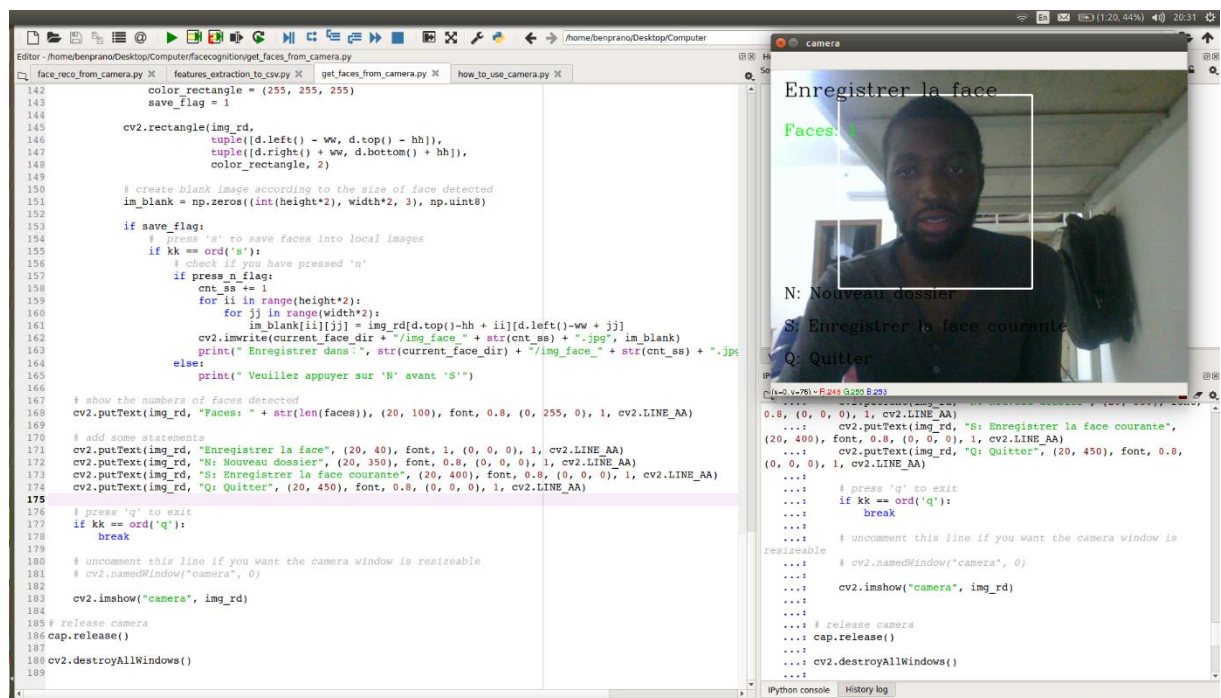
#### 3.1. Environnement de travail

Afin de mener à bien ce projet, nous avons utilisé un ensemble de matériels dont les principales caractéristiques sont les suivantes :

- ✓ Processeur : Intel® Core™ i3-2330M CPU @2.20GHz
- ✓ RAM : 6.00 Go
- ✓ OS : Linux Ubuntu 16.04 LTS
- ✓ Environnement de développement : Spyder
- ✓ Langage : Python
- ✓ Libraries: OpenCV, dlib, face\_recognition

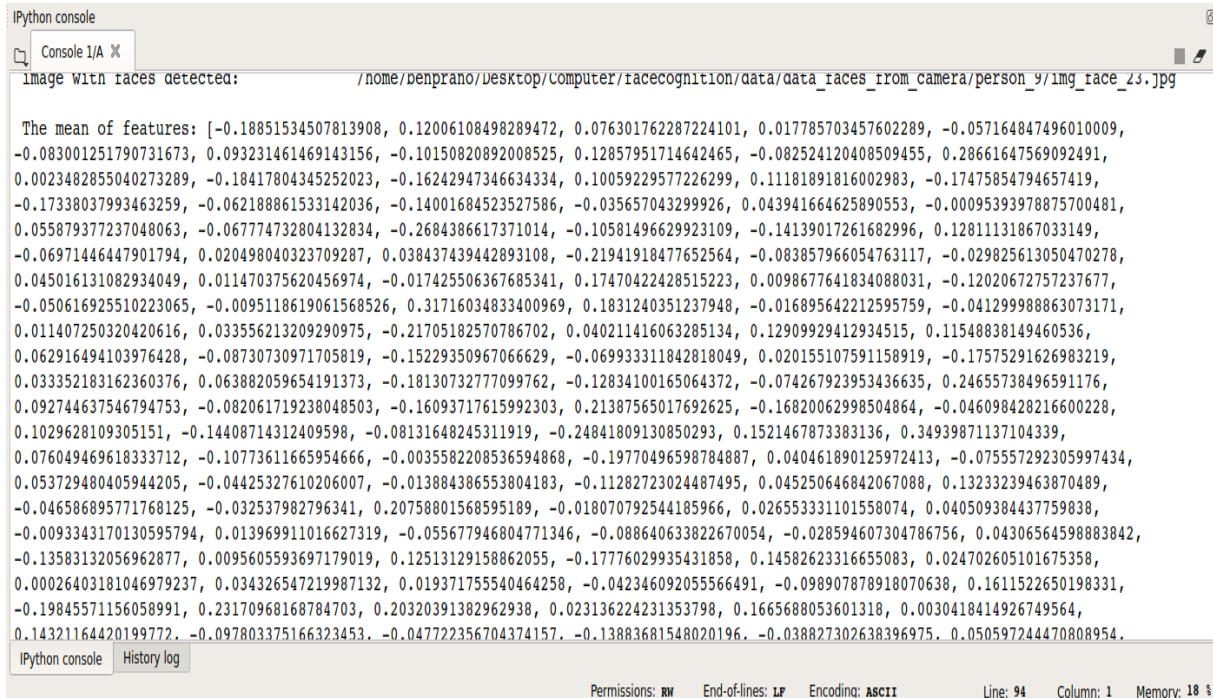
#### 3.2. Constitution de la base de données à partir de la Webcam

Il y a deux manières de constituer la base de données soit en faisant plusieurs prises à partir de la webcam comme le montre la figure ci-dessous soit en prenant les photos de l'étudiant dans plusieurs conditions et en l'insérant dans la base de donnée



### 3.3.Extraction des caractéristiques

Comme la montre la figure ci-dessous, à partir de photos enregistrées dans la base de données, l'algorithme calcule la matrice moyenne et extrait les caractéristiques de chaque photo



```
Python console
Console 1/A X
image with faces detected: /home/benprano/desktop/computer/facecognition/data/data_faces_from_camera/person_9/img_face_23.jpg

The mean of features: [-0.18851534507813908, 0.12006108498289472, 0.076301762287224101, 0.017785703457602289, -0.057164847496010009,
-0.083001251790731673, 0.093231461469143156, -0.10150820892008525, 0.12857951714642465, -0.082524120408509455, 0.28661647569092491,
0.002348285040273289, -0.18417804345252023, -0.16242947346634334, 0.10059229577226299, 0.11181891816002983, -0.17475854794657419,
-0.17338037993463259, -0.062188861533142036, -0.14001684523527586, -0.035657043299926, 0.043941664625890553, -0.00095393978875700481,
0.055879377237048063, -0.067774732804132834, -0.2684386617371014, -0.10581496629923109, -0.14139017261682996, 0.12811131867033149,
-0.06971464447901794, 0.020498040323709287, 0.038437439442893108, -0.21941918477652564, -0.083857966054763117, -0.029825613050470278,
0.045016131082934049, 0.011470375620456974, -0.017425506367685341, 0.17470422428515223, 0.0098677641834088031, -0.1202067275237677,
-0.050616925510223065, -0.0095118619061568526, 0.31716034833400969, 0.1831240351237948, -0.016895642212595759, -0.041299988863073171,
0.011407250320420616, 0.033556213209290975, -0.21705182570786702, 0.040211416063285134, 0.12909929412934515, 0.11548838149460536,
0.062916494103976428, -0.08730730971705819, -0.15229350967066629, -0.069933311842818049, 0.020155107591158919, -0.17575291626983219,
0.033352183162360376, 0.063882059654191373, -0.18130732777099762, -0.12834100165064372, -0.074267923953436635, 0.24655738496591176,
0.092744637546794753, -0.082061719238048503, -0.16093717615992303, 0.21387565017692625, -0.16820062998504864, -0.046098428216600228,
0.1029628109305151, -0.14408714312409598, -0.08131648245311919, -0.24841809130850293, 0.1521467873383136, 0.34939871137104339,
0.07604946961833712, -0.10773611665954666, -0.0035582208536594868, -0.19770496598784887, 0.040461890125972413, -0.075557292305997434,
0.053729480405944205, -0.04425327610206007, -0.013884386553804183, -0.11282723024487495, 0.045250646842067088, 0.13233239463870489,
-0.046586895771768125, -0.032537982796341, 0.20758801568595189, -0.018070792544185966, 0.026553331101558074, 0.040509384437759838,
-0.0093343170130595794, 0.013969911016627319, -0.055677946804771346, -0.088640633822670054, -0.028594607304786756, 0.04306564598883842,
-0.13583132056962877, 0.0095605593697179019, 0.12513129158862055, -0.17776029935431858, 0.14582623316655083, 0.024702605101675358,
0.00026403181046979237, 0.034326547219987132, 0.019371755540464258, -0.042346092055566491, -0.098907878918070638, 0.1611522650198331,
-0.19845571156058991, 0.23170968168784703, 0.20320391382962938, 0.023136224231353798, 0.1665688053601318, 0.0030418414926749564,
0.14321164420199772, -0.097803375166323453, -0.047722356704374157, -0.13883681548020196, -0.038827302638396975, 0.050597244470808954.
```

### 3.4.Reconnaissance faciale

L'algorithme d'Eigenfaces et Fisherface compare la photo prise par le téléphone de l'utilisateur via l'application avec les caractéristiques extraites contenues dans le fichier csv si les calculs matricielles sont approximativement égaux la ou les personnes sont reconnues et étiquetés.

Nous pouvons constater ces résultats sur les images ci-dessous :

Photo prise avant la reconnaissance faciale





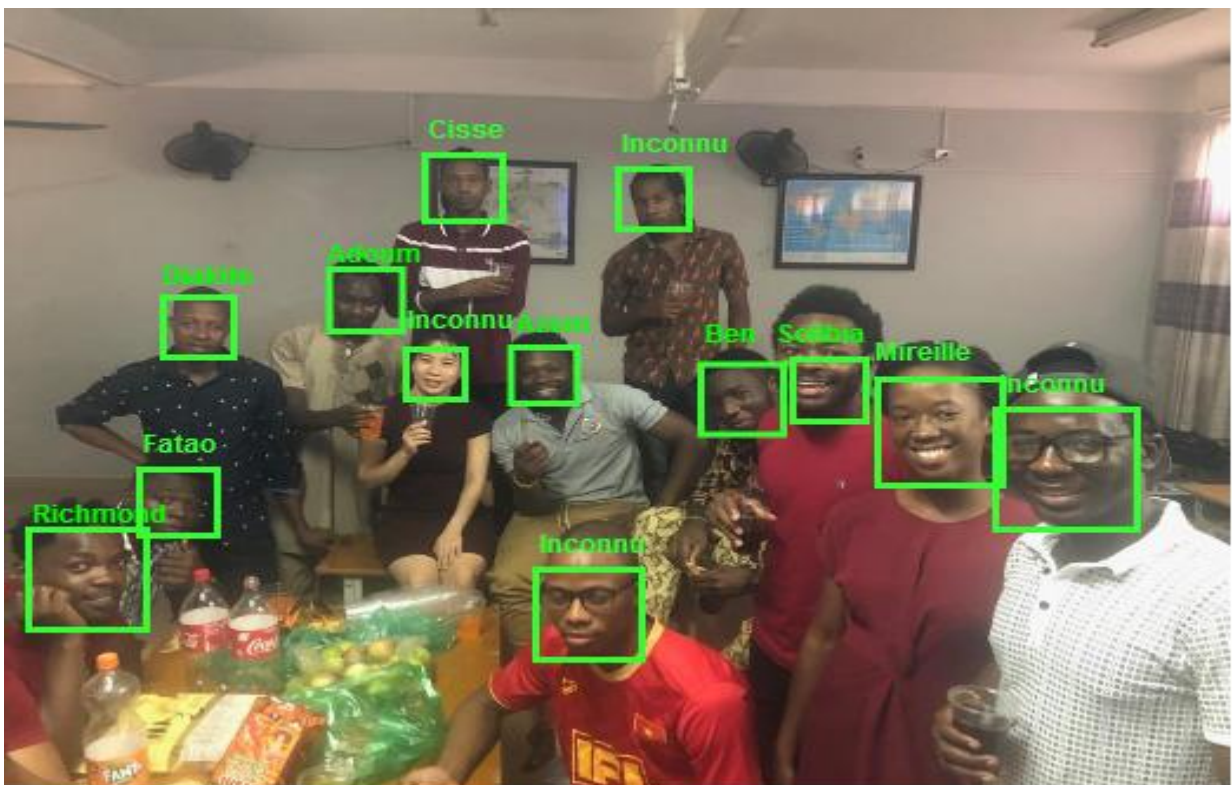
Après avoir appliqué la reconnaissance faciale



2<sup>ieme</sup> exemple de photo :  
Avant la reconnaissance



Après la reconnaissance faciale





## **VI. CONCLUSION ET PERSPECTIVE**

Dans ce projet, nous nous sommes intéressés sur le problème de gestion manuelle des présences des étudiants au sein de l'Institut Francophone International (IFI). Cela étant, l'objectif de notre travail était celui de concevoir et développer une application de gestion des présences des étudiants basées sur la technologie biométrique de reconnaissance faciale.

Au cours de ce travail, une analyse des différentes techniques de reconnaissance développées au cours de ces dernières années présentée et cela pour mettre en évidence les particularités ainsi que les avantages et les inconvénients de chacune d'entre elles. Cependant, nous avons découvert des approches holistiques, locales et hybrides pour la reconnaissance faciale. Compte tenu de la nature de notre système de gestion de présence, il nous a fallu utiliser une technique de reconnaissance qui serait très rapide et robuste du fait que le système fonctionnera à temps réel.

En ce sens, la première partie de l'application qui consiste à localiser les visages emploie l'algorithme de Viola et Jones, largement reconnu comme méthode fonctionnant en temps réel et fournissant des résultats robustes et fiables.

La seconde partie de l'application s'occupe de la reconnaissance des visages localisés.

Notre travail consiste à mettre au point un algorithme efficace destiné à reconnaître un individu par son visage en utilisant la méthode Eigenface qui se base sur une analyse en composante principale (ACP). L'ACP est une méthode mathématique qui peut être utilisée pour simplifier un ensemble de données, en réduisant sa dimension. Elle est utilisée pour représenter efficacement les images de visages, qui peuvent être approximativement reconstruites à partir d'un petit ensemble de poids et d'une image de visage standard. Il faut savoir qu'Eigenfaces est considérée comme la méthode la plus simple et la plus précise de reconnaissance de visage, mais beaucoup d'autres méthodes ou combinaisons de méthodes multiples (beaucoup plus compliqué) peuvent être utilisées.

Ainsi, la réalisation de ce projet nous a permis de comprendre plus en profondeur le développement d'un logiciel, de la conception à la réalisation en passant par la réflexion logique et algorithmique. Si nous avions plus de temps, nous aurions aimé développer un peu plus l'interface graphique afin de compléter son rôle. Il aurait été également intéressant d'affiner le moteur de reconnaissance, en trouvant un moyen de prendre des images correctement cadrées et en déterminant les seuils de façon rigoureuse.

### **Perspectives**

Réaliser une application complète avec l'interface graphique, mener des recherches sur les variations de pose et d'éclairage pour aboutir à une reconnaissance faciale optimale.

## Références

- ✓ [https://fr.wikipedia.org/wiki/Syst%C3%A8me\\_de\\_reconnaissance\\_faciale](https://fr.wikipedia.org/wiki/Syst%C3%A8me_de_reconnaissance_faciale)
- ✓ Évaluation de méthodes de reconnaissance de visages pour l'identification d'individus à partir d'une image de référence : Francis C. Migneault, professeur-superviseur : Éric Granger, Département de génie de la production automatisée, Laboratoire d'imagerie de vision et d'intelligence artificielle
- ✓ CONCEPTION DE SYSTÈME DE RECONNAISSANCE DE VISAGES SPATIO-TEMPORELLE SUR VIDÉOS À PARTIR D'UNE SEULE IMAGE DE RÉFÉRENCE par Francis CHARETTE MIGNEAULT
- ✓ Automated Classroom Monitoring With Connected Visioning System : Jian Han Lim, Eng Yeow Teh, Ming Han Geh and Chern Hong Lim, Faculty of Computing and Information Technology Tunku Abdul Rahman University College, Setapak, Malaysia
- ✓ [www.google.com](http://www.google.com)
- ✓ <http://www.natural-solutions.eu/blog/le-deep-learning-pour-les-nuls>
- ✓ [https://fr.wikipedia.org/wiki/Vision\\_par\\_ordinateur](https://fr.wikipedia.org/wiki/Vision_par_ordinateur)
- ✓ <https://www.supinfo.com/articles/single/6041-machine-learning-introduction-apprentissage-automatique>
- ✓ [https://fr.wikipedia.org/wiki/Apprentissage\\_profond](https://fr.wikipedia.org/wiki/Apprentissage_profond)
- ✓ [https://fr.wikipedia.org/wiki/Apprentissage\\_automatique](https://fr.wikipedia.org/wiki/Apprentissage_automatique)
- ✓ [1] H. Kopetz, "Internet of things," in Real-time systems. Springer, 2011, pp. 307–323.
- ✓ [2] R. H. Weber and R. Weber, Internet of things. Springer, 2010, vol. 12.
- ✓ [3] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," IEEE transactions on pattern analysis and machine intelligence, vol. 15, no. 10, pp. 1042–1052, 1993.
- ✓ [4] K.-C. Lee, J. Ho, and D. J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," IEEE Transactions on pattern analysis and machine intelligence, vol. 27, no. 5, pp. 684–698, 2005.
- ✓ [5] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on. IEEE, 1991, pp. 586–591.
- ✓ [6] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs.fisherfaces: Recognition using class specific linear projection," IEEE