



INSTITUT
FRANCOPHONE
INTERNATIONAL



Contrôle de présence automatique dans une salle de classe

RAPPORT FINAL

Etudiants : BERTELOT Peterson ET JANVIER Fabienne

Module : Reconnaissance des formes

Professeur : Dr. HO Tuong Vinh

Année Académique : 2019-2021 — PROMOTION 24 SIM

Table des matières

1	Introduction	2
2	Contexte	3
3	Objectifs	3
4	Contribution	4
5	Analyse du sujet	5
5.1	Le domaine d'application	5
5.2	Le problème visé	5
5.3	La situation actuelle	5
5.4	Résultats attendus	6
6	État de l'art	7
6.1	Connaissance précise et spécialisée du sujet	7
6.2	Les méthodes de reconnaissance faciale	7
6.3	Les Techniques ou méthodes existants	8
7	Solution	12
7.1	Données	13
7.1.1	La composition du jeux de données	13
7.2	Outils	14
7.2.1	Présentation de la bibliothèque OpenCV	14
7.2.2	Pré-traitement	15
7.2.3	Utilisation des Réseaux de Neurone Convolutionnels	15
7.3	Modèles conceptuels / Processus de traitement	16
7.3.1	Modèle de de conception de l'architecture de notre CNN	16
7.3.2	Processus de traitement	16
8	Implémentation et expérimentation	18
8.1	Implementation des Cnns	18
8.2	Architecture de notre réseau	19
8.3	Précision obtenue après entraînement avec les Cnns	20
8.4	Précision obtenue a partir du nombre d'époque de formations	20
9	Résultats	20
9.1	Bonnes détection	20
9.2	Mauvaise détection	21
10	Conclusion	22
11	Perspectives	22
12	Lien github du projet	22
13	Références	23

1 Introduction

Au cours des dernières années, des systèmes de reconnaissance faciale automatique ont été mis en œuvre dans une série de domaines sociaux. Bien que ces technologies soient généralement associées à la promesse d'améliorer la sécurité publique, de plus en plus d'autres applications ont vu le jour - vérifier l'identité des utilisateurs de la banque grâce à des «panneaux d'affichage intelligents». Affichez des publicités en réponse à l'humeur des passants. Il est particulièrement intéressant de savoir comment la technologie de reconnaissance faciale a été mise en œuvre dans les écoles. En effet, il existe aujourd'hui de nombreuses applications éducatives pour la reconnaissance faciale et la détection faciale, y compris les systèmes de sécurité du campus qui appellent et surveillent automatiquement les émotions et l'attention des étudiants. Jusqu'à présent, ces technologies ont rarement été contestées ou rejetées dans des pays comme les États-Unis, le Royaume-Uni et l'Australie. Après tout, les écoles de ces pays utilisent depuis longtemps des systèmes de surveillance par caméra et d'autres formes de technologie pour suivre et surveiller.

En ce sens, la reconnaissance faciale peut être considérée comme une extension logique de la tendance à la surveillance basée sur la technologie établie dans les écoles dans les années 1990.

La technologie de reconnaissance faciale la plus émergente découle également du développement parallèle de la technologie de traitement de la vision par ordinateur (dans cette technologie, la technologie d'apprentissage automatique peut être appliquée pour reconnaître et apprendre des modèles dans les flux de données d'images numériques) comme l'avancement de la technologie des appareils photo numériques. En bref, le principe de fonctionnement de la technologie de reconnaissance faciale est qu'un ordinateur extrait les traits du visage capturés sur des images vidéo numériques, puis compare les données avec des visages précédemment analysés stockés dans une base de données. Fondamentalement, ces bases de données contiennent un grand nombre de visages photographiés avec des noms associés et d'autres informations personnellement identifiables.

2 Contexte

La reconnaissance et la détection des visages est une technologie biométrique populaire ces dernières années. Elle est largement utilisée dans les applications de vérification d'identité, de contrôle d'accès et de vidéosurveillance. Il existe plusieurs méthodes de reconnaissance mondiales, locales et hybrides. Détection de visage et / ou de présence. Récemment, des chercheurs de divers domaines ont concentré leurs travaux sur les clés et les mots de passe qui ne peuvent être falsifiés, protégés et même les plus efficaces. Par conséquent, la technologie biométrique a été inventée et la technologie biométrique est devenue une mode dans des domaines qui nécessitent des niveaux élevés de sécurité et de contrôle.

Parmi toutes les technologies biométriques existantes, la reconnaissance faciale est l'une des technologies les plus utilisées et les plus appropriées, car elle nous permet d'utiliser beaucoup d'informations sur les personnes. Bien que l'informatique, en particulier le domaine du traitement d'image, ait fait de grands progrès, il reste des problèmes liés au système d'application de la technologie de reconnaissance faciale, c'est pourquoi il a produit de nombreux travaux de recherche en 2001. Le développement des 50 dernières années dans divers domaines, tels que la reconnaissance de modèles, les réseaux de neurones, la vision par ordinateur, l'infographie ...

Dans nos travaux, nous nous intéressons à ce sujet très utile et exigeant dans le domaine de la cybersécurité, de la sécurité nationale et internationale. Nous vous recommandons d'utiliser la technologie la plus avancée pour mettre en œuvre des applications de reconnaissance faciale et / ou de détection de présence.

3 Objectifs

L'idée est de placer un réseau de caméras dans la salle de classe pour détecter automatiquement les personnes présentes. Par la suite, notre application comprend un modèle qui a été formé pour reconnaître tous les élèves de la salle, et peut détecter différents visages et les placer sur la tête de chaque visage. Atteint signifie que le nom de l'individu a été détecté. Si l'individu n'est pas reconnu par l'application, il sera signalé comme intrus. Lorsque les élèves sont détectés, enfin, l'application doit dire le nombre de personnes détectées et leurs noms pour permettre aux élèves absents d'être vus. En raison du faible prix de notre système, nous utiliserons deux caméras pour les tests : une caméra externe et une caméra sur votre ordinateur.

Nous allons mener la première expérience basée sur les bibliothèques existantes dans opencv (Haardesciptor, Dnn, deepneuralnetwork, yolo), et nous essayerons également des techniques basées sur l'apprentissage profond, cette fois en utilisant des réseaux de neurones convolutifs, qui ne nous rendront pas meilleurs Le résultat, mais ce ne sera pas meilleur. Cette méthode peut faire mieux, mais le travail de recherche ne se limite pas à cela. Nous continuerons à travailler dur pour élargir la base de données d'apprentissage, car selon les articles lus, ils conviennent aux bonnes prédictions de cnns Il y a au moins 1000 échantillons d'images dans chaque catégorie, mais nous sommes près de 100 à 200 échantillons dans chaque catégorie, ce qui rendra le meilleur résultat quelque peu déformé, mais en utilisant le modèle de descripteur Haar dans opencv, nous allons créer un modèle de reconnaissance et l'exécuter en temps réel Le personnel fonctionne en temps réel et affiche les résultats de nos commentaires.

Afin d'améliorer notre travail, nous continuons à construire nos propres classes, chacune avec 1000 échantillons (nous avons ouvert 4 classes pour 4 étudiants), pour tester et utiliser Cnns pour mieux tirer des conclusions, afin de reconnaître en temps réel.

4 Contribution

En faisant ce travail, notre contribution réside dans l'utilisation et l'extension de la reconnaissance en temps réel, car elles sont davantage utilisées pour la classification d'images dans les applications existantes qui utilisent Cnns. Mais dans notre contexte, nous avons utilisé différents algorithmes existants de reconnaissance faciale et de reconnaissance de forme pour des méthodes expérimentales, nous avons ajouté pour cela un objet temps réel étendu qui peut utiliser son algorithme sous surveillance et surveillance. Notre exemple est les étudiants en classe (dans notre exemple, nous avons utilisé 4 classes dans l'expérience, composées de 4 étudiants de l'Institut International Francophone pour la phase expérimentale) a la fin de nos travaux, nous avons obtenu des résultats satisfaisants, mais il n'y a pas eu de contestation des méthodes actuelles existantes basées sur des API qui avaient été préconçues. À l'heure actuelle, nous travaillons dur pour améliorer la composition de l'ensemble de données pour compléter le Cnn, afin que nous puissions obtenir de meilleurs résultats, et notre travail se poursuit toujours à la réforme.

5 Analyse du sujet

5.1 Le domaine d'application

La vidéosurveillance couvre plusieurs domaines d'application. La lutte contre le terrorisme est l'une des plus connues. Nous avons plusieurs enquêtes qui ont réussi non seulement dans le plan de prévention, mais aussi dans l'enquête pénale, car il permet de reconnaître des visages et de reconstituer des personnes suspectes et / ou des chemins dans des zones dangereuses (grâce aux images enregistrées dans un espace spécifique).

Dans la lutte contre la criminalité à petite et moyenne échelle, la vidéosurveillance s'est avérée la plus efficace. Grâce à la reconnaissance faciale, par exemple, il est possible de trouver et d'identifier d'éventuels auteurs de troubles autour du stade ou sur les étagères. De même, les statistiques de la criminalité montrent que les vols à l'étalage et les vols qualifiés ont considérablement diminué dans les centres commerciaux ou les rues du centre-ville équipées de caméras.

L'installation de caméras dans les transports publics (trains, bus et métros) a trois objectifs : la sécurité des passagers, la lutte contre la fraude et la prévention des dommages (non seulement dans les trains, mais aussi dans les gares, sur les quais, et les développements récents sur certains aspects)). La surveillance des lieux sensibles publics et privés sensibles aux intérêts stratégiques importants du pays et de la région (énergie, défense, transports) est également l'un des domaines d'application les plus courants. En combinant la biométrie (iris, empreinte digitale), la reconnaissance faciale et la vidéosurveillance, l'accès aux zones réservées peut être contrôlé. Plusieurs aéroports internationaux, dont l'aéroport de Londres Heathrow, sont équipés de ce type de système, conçu pour faciliter le transit et même permettre au personnel d'entrer automatiquement dans différentes zones réservées.(bagages, fret, pistes).

Dans cet extrait, nous avons vu de nombreux domaines où le système de vidéosurveillance peut être appliqué et dans le contexte de la surveillance en classe, nous l'avons ramené à notre cas.

5.2 Le problème visé

Dans la section précédente, nous avons présenté plusieurs domaines d'application du système de vidéosurveillance. Par conséquent, dans notre situation particulière, nous espérons intégrer le système de vidéosurveillance dans un environnement de contrôle efficace de la présence dans la salle de classe afin de surmonter la salle. S'il y a un problème d'intrus.

5.3 La situation actuelle

En parlant de travail actuel, il existe des applications avec des licences vendables et des API qui permettent la reconnaissance faciale dans les entreprises, les bureaux et de nombreux autres lieux publics. Ceci est un exemple de la technologie de reconnaissance faciale de l'API suivante développée par les ingénieurs de Facebook en 2011 Sauf qu'il n'y a pas de code source, tout le monde peut l'utiliser, mais il existe certains systèmes dans le domaine de la sécurité militaire. Mais en ce qui nous concerne, nous utilisons les outils et technologies open source existants pour développer et adapter un système de suivi académique de l'assiduité des élèves en classe.

5.4 Résultats attendus

À la fin du travail, nous espérons obtenir une application capable d'identifier les individus dans la classe en temps réel en utilisant un modèle pour tous les élèves de la classe. Ce modèle ne convient qu'aux personnes bien définies.

Par exemple, l'application devrait identifier spécifiquement les étudiants de la classe, mais dans notre cas, comme nous n'avons pas beaucoup de ressources matérielles pour faire de l'exercice, nous limitons la classe d'exercices à seulement 4 étudiants. À la fin de la mise en œuvre, l'application doit identifier l'intrus dans la vidéo en temps réel. Elle permet également de rechercher certaines personnes spécifiques dans la vidéo comme entrée pour inspection ou suivi. À l'avenir, nous espérons étendre notre application afin qu'ils puissent fournir des étudiants pertinents. Toutes les informations spécifiques sur son cours, telles que l'heure d'arrivée de son cours, l'heure à laquelle il a quitté la salle, la durée des études du cours, peuvent vous aider à comprendre le comportement relatif des étudiants et les aider à prendre des décisions en cas de besoin.

6 État de l'art

La deuxième partie est la recherche bibliographique, qui nous demandera d'avoir une expertise précise sur le sujet, de comprendre la technologie existante, de trouver des solutions possibles au sujet, et enfin d'établir une comparaison des différentes méthodes sur le sujet.

6.1 Connaissance précise et spécialisée du sujet

La mise en place d'un système de reconnaissance faciale automatique et fiable est un obstacle technique non résolu. La détection des visages joue un rôle essentiel dans la reconnaissance des visages. C'est la première étape de la reconnaissance faciale automatique. Le système de reconnaissance faciale est une application logicielle conçue pour reconnaître automatiquement les visages humains. Deux méthodes sont utilisées pour cela : la méthode 2D et la méthode 3D.

6.2 Les méthodes de reconnaissance faciale

Chaque méthode est basée sur trois étapes : la reconnaissance de la détection des visages qui reconnaissent les visages, l'analyse des visages qui donnent des résultats numériques et la comparaison des résultats de la reconnaissance avec une base de données de visages enregistrés. Par conséquent, il est nécessaire d'établir une base de données pour chaque utilisateur avant toute identification, qui contient une capture initiale ou plusieurs captures. Si plusieurs captures sont effectuées, la valeur moyenne est mémorisée, afin que le système puisse mieux s'adapter à divers paramètres, qui peuvent être différents en reconnaissance (cernes / rides, lumière ambiante, maquillage, etc.). Une fois l'analyse effectuée, la reconnaissance peut être comparée entre le modèle obtenu et le modèle stocké dans la base de données. Ensuite, si nous stockons initialement une capture initiale dans la base de données, alors nous parlons d'identifier 1 à 1, et si nous avons plusieurs captures initiales, c'est 1 à N.

Méthode 2D : Considérée comme une méthode de reconnaissance faciale «classique», cette méthode consiste à identifier les utilisateurs à partir de leurs photos. Pour plus de sécurité, vous pouvez prendre cette photo avec votre appareil photo ou simplement la faire enregistrer dans le cadre de l'aide de l'utilisateur. Ensuite, il est reconnu par un algorithme, qui peut être basé sur divers éléments, comme la forme des éléments du visage, comme les yeux et leur espacement, la bouche, le visage ...

Ensuite, deux types d'algorithmes sont distingués :

- Le premier type d'algorithme crée une image géométrique d'un utilisateur en fonction de différents paramètres (la taille des éléments du visage, la forme et la distance entre eux). Ensuite, encodez les paramètres récupérés et comparez-les avec les paramètres existants dans la base de données.
- Le deuxième type d'algorithme utilise l'algorithme de Fourier pour créer un vecteur de poids à l'aide de faces propres, et même encodé numériquement l'image en calculant la valeur moyenne de certaines zones de l'image.

Méthode 3D : Cette méthode est considérée comme une amélioration de la reconnaissance 2D. En fait, il peut créer un modèle 3D à partir de plusieurs photos ou vidéos prises en continu, afin de pouvoir reconnaître différentes perspectives de personnes et créer un modèle 3D. Il existe de nombreux algorithmes d'analyse, et chaque algorithme crée un modèle 3D correspondant au visage de l'utilisateur basé sur un ou plusieurs éléments du

visage (nez, direction du visage, etc.).

Cependant, en raison des déficiences inhérentes aux informations faciales dans les images 2D, les algorithmes de reconnaissance faciale ont de grandes difficultés à traiter les changements d'éclairage, les changements de posture, les expressions faciales, le temps de retard et l'occlusion. La surface faciale 3D offre la possibilité de surmonter les difficultés causées par les changements de posture et l'éclairage de la tête. Avec l'avancement de la technologie d'imagerie 3D, les données faciales 3D sont devenues de plus en plus disponibles ces dernières années. Il existe plusieurs bases de données publiques de référence de visage 3D.

Dans cette section, nous présenterons le dernier état de la technologie de détection de visage 3D et classeront les méthodes, technologies, algorithmes et systèmes industriels existants. Nous discuterons également des méthodes de reconnaissance faciale 3D. La reconnaissance de personnes à partir de visages 3D présente certaines limites. Dans le cas des expressions faciales, elle ne peut pas obtenir une meilleure reconnaissance. Pour résoudre ce problème, nous allons introduire une nouvelle méthode d'expression faciale 3D. Enfin, nous résumons les méthodes et techniques étudiées.

6.3 Les Techniques ou méthodes existants

Ces dernières années, plusieurs méthodes de reconnaissance faciale ont été proposées, qui ont principalement deux aspects : la reconnaissance à partir d'images fixes et la reconnaissance à partir de séquences d'images (vidéo). Dans ce domaine, le travail sous différents éclairages, expressions faciales et conditions d'orientation peut être divisé en deux catégories différentes selon qu'il s'agit de méthodes géométriques ou de méthodes holistiques.

Les méthodes existantes peuvent être grossièrement divisées en trois catégories : les méthodes locales, les méthodes globales et les méthodes hybrides. Les principales méthodes de ces trois méthodes sont décrites ci-dessous.

Nous insistons particulièrement sur la méthode de réduction de dimensionnalité, qui fait partie de l'approche globale. En fait, les méthodes de réduction de dimensionnalité appartiennent en fait au cadre général du traitement des données (pas seulement des visages) et de la reconnaissance d'objets. Veuillez noter que la plupart des méthodes nécessitent un positionnement précis du visage, un recadrage géométrique (pour qu'il apparaisse toujours dans la même position) et une mise à l'échelle pour que le visage ait une taille suffisante. La technologie de reconnaissance faciale 2D robuste à différents facteurs variables (éclairage, posture, dissimulation) est loin d'être développée. La reconnaissance faciale 3D est une méthode alternative prometteuse pour surmonter ces problèmes, d'autant plus depuis l'émergence de puissants équipements d'acquisition 3D. Le principal avantage de la méthode basée sur le modèle 3D réside dans le fait que le modèle 3D conserve toutes les informations sur la géométrie du visage, afin que celle-ci puisse être véritablement représentée.

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

Reconnaissance faciale 3D basée sur le profil : la plupart de ces méthodes et produits sont conçus pour les images frontales. Les contours du visage fournissent d'autres informations structurelles sous une forme qui n'est pas visible dans la vue frontale, mais

qui peut souvent être obtenue à partir d'autres vues. Selon une étude de Davidenko, les contours du visage jouent un rôle important dans la perception humaine de la tâche d'identification de l'identité de genre. Lorsque seules des vues de côté sont disponibles, l'utilisation des contours du visage en biométrie est particulièrement intéressante. Le contour de contour contient des informations plus fiables que la texture, ce qui rend les informations de contour particulièrement précieuses pour la reconnaissance.

Méthodes basées sur les caractéristiques géométriques faciales 3D : Dans ce cas, la recherche se concentre davantage sur les méthodes basées sur les courbes faciales et les méthodes qui utilisent des techniques de sélection des caractéristiques pour optimiser les performances de reconnaissance

Méthode basée sur les courbes : méthode d'extraction des courbes faciales 3D. Ces courbes peuvent être clairement analysées et comparées. Plusieurs représentations de courbes ont été utilisées dans la littérature. Daoudi et coll. Ils ont utilisé des lignes de contour plates. L'auteur utilise l'analyse de la forme de la courbe plane pour analyser les déformations faciales. En revanche, il n'y a pas de recherche pour sélectionner la courbe de corrélation. Dans le même auteur, les distances géodésiques ont été utilisées pour extraire les contours 3D. Selon les recherches de Bronstein et al. Comparé au changement de distance euclidienne, le changement de distance géodésique dû aux expressions faciales est négligeable. Par conséquent, les auteurs ont utilisé des points d'échantillonnage obtenus à partir de l'intersection entre la ligne de contour et le contour radial, qui partait de la pointe du nez, puis ils ont calculé la distance euclidienne entre les points correspondants sur différentes faces. Cette méthode permet de sélectionner un sous-ensemble de courbes faciales pour une correspondance faciale efficace. Drira et al. ont exploré l'utilisation de différents types de formes de nez. L'auteur a extrait des isogèdes, puis a établi différents chemins géodésiques dans une série de courbes fermées de Riemann. La même méthode peut être utilisée pour le calcul statistique (valeur moyenne) pour une recherche efficace dans la base de données de visages 3D. Le même auteur a proposé une méthode similaire pour analyser la forme de la courbe radiale 3D. Mpiperis et coll. Ils ont proposé une méthode de paramétrage polaire Dans cette nouvelle méthode de paramétrage inhérente, les propriétés de la surface sont invariantes à la déformation isométrique. Lorsque la bouche est ouverte, ce paramétrage n'est plus valide. Par conséquent, l'auteur recommande de déconnecter toutes les lèvres du visage et de modifier ce paramétrage en fonction de trois points de référence.

Méthode basée sur la sélection des caractéristiques : Plusieurs méthodes ont été proposées pour analyser la capacité de reconnaissance de différentes parties du visage. Darniyal et al. ont proposé un algorithme dans lequel un visage est représenté comme un vecteur de distances entre des points d'intérêt (points de repère) sur le visage. Ils ont sélectionné des points à partir de combinaisons possibles de points utilisés / non utilisés par la force brute et ont comparé les taux de reconnaissance. Ils ont conclu que la meilleure option était les points autour des yeux et du nez. Kakadiaris et Al. Ils utilisent un modèle de visage annoté, qui se déforme élastiquement pour s'adapter à chaque visage, lui permettant automatiquement d'annoter ses différentes zones (comme le nez, les yeux, la bouche, etc.). Afin de résoudre le problème des expressions faciales, l'auteur classe les visages à l'aide de coefficients d'ondelettes, qui représentent des zones non affectées par les expressions (par exemple, les yeux et le nez). Cependant, le taux de reconnaissance de l'ensemble du visage est le plus élevé, ce qui signifie que le rejet de certaines zones affectées par des déformations faciales entraînera la perte des informations de reconnais-

sance. L'appariement de plusieurs zones du nez a été proposé. Les premiers résultats ont montré qu'il y avait une amélioration dans une seule grande zone du lobe frontal. Faltmier et coll. Ils recommandent d'utiliser l'algorithme ICP pour enregistrer différentes parties du visage, puis de les fusionner pour améliorer les performances biométriques. Récemment, Wang et al. Ils ont calculé une carte de différence de forme signée (SSDM) calculée entre deux faces alignées comme représentation intermédiaire de la comparaison de forme. Sur la base de SSDM, trois types d'opérateurs ont été utilisés pour coder la similitude locale et les changements de caractéristiques de différentes formes de visage. AdaBoost sélectionne les caractéristiques locales les plus distinctives et l'utilise comme un classificateur faible. Les types de chaque fonctionnalité sont : le type Haar, le type Gabor et le modèle binaire local (LBP). En utilisant des caractéristiques similaires, Li et Al ont suggéré de concevoir un dispositif commun et un système de classification pour collecter différentes caractéristiques géométriques de bas niveau (comme la courbure) et les classer en fonction de leur sensibilité aux expressions faciales. Pour les méthodes basées sur les courbes, l'analyse de la forme du visage devient l'analyse des courbes faciales 3D. Toutes les courbes du visage ont de meilleures performances. Grâce à la méthode des faces 3D qui combine des courbes et des techniques de sélection de caractéristiques, nous proposons un nouvel algorithme de reconnaissance de visage 3D qui combine la représentation des faces 3D à travers une série de courbes et des techniques de sélection de caractéristiques. Après l'acquisition du visage 3D, un prétraitement est effectué pour extraire la partie informationnelle du visage et corriger certains défauts, tels que des valeurs aberrantes, des trous ou la présence de parties indésirables (comme les cheveux et les vêtements). Extraire les courbes faciales radiales et horizontales. Ensuite, selon la nature de la courbe (fermée ou ouverte), utilisez l'un des cadres géométriques pour analyser la forme de la courbe faciale. Cette méthode géométrique est nécessaire pour comparer et analyser la forme des courbes faciales (ouvertes et fermées).

Reconnaître les expressions faciales 3D : La future interface homme-machine utilisera de nouvelles sources d'informations, telles que les expressions faciales dans six situations (colère, dégoût, joie, peur, surprise, tristesse), la posture humaine ou la parole.

Système de reconnaissance des expressions faciales existant : le système comprend l'analyse de la déformation des traits du visage (tels que les yeux, les sourcils et la bouche) pour évaluer la possibilité de reconnaître six émotions courantes. On pense que ces caractéristiques contiennent suffisamment d'informations pour être identifiées.

Méthode basée sur la subdivision : Dans l'étape de subdivision, le système localise les zones des yeux, des sourcils et de la bouche. L'algorithme est proposé dans l'article de Hammal. Dans l'étape d'extraction des données, les os sont extraits. En fait, ce sont les contours des yeux, de la bouche et des sourcils. A partir des os, nous choisirons cinq distances pour caractériser la déformation. Les traits considérés incluent les contours des yeux, des sourcils et des lèvres. Enfin, nous avons le squelette d'expression. À partir du squelette de l'expression faciale, nous pouvons déterminer la déformation des traits du visage. Chaque squelette définit cinq distances. À partir de ces distances caractéristiques, nous utiliserons un modèle de croyance transférable pour reconnaître les expressions faciales. Après cela, chaque expression faciale est caractérisée par une combinaison d'états de symboles.

Méthode du masque : L'idée est d'utiliser un masque pour l'enregistrement dans la reconnaissance faciale. Ce système n'est pas nouveau. Par exemple, Blanz et Al ont utilisé

des masques 3D pour enregistrer des vues frontales et longitudinales. Certains masques existants sont soigneusement conçus pour s'adapter à une surface déformable en fonction d'une action musculaire. Cependant, lors du traitement d'images plates, il est difficile d'obtenir un bon enregistrement de tous les sommets densément placés. De plus, bien que ces masques aient des formes flexibles. Les masques triangulaires présentent des avantages dans le mappage de texture.

Méthode du modèle d'apparence (AAM) : Franck et Al se sont engagés à identifier et à synthétiser les expressions faciales à travers des modèles d'apparence. Ils ont traité de la nouvelle application des modèles d'apparence active dans l'analyse et la synthèse des expressions faciales et la reconnaissance des expressions faciales. Ils ont considéré six expressions universelles de joie, de colère, de peur, de dégoût, de tristesse et de surprise. Après avoir décrit ce type de modèle (en utilisant trois PCA ou en utilisant un seul calcul PCA), ils ont proposé une méthode originale d'analyse et de synthèse qui permet de supprimer toute expression faciale d'une photo puis de la mettre en haut Synthétiser des expressions faciales artificielles. Ce même visage. Ils ont fait des recommandations pour ces deux méthodes qui permettent la modélisation des expressions faciales. Enfin, il montre l'intérêt du modèle d'apparence active pour la reconnaissance automatique des expressions faciales.

En bref, afin d'obtenir de meilleurs résultats, nous commenterons quelques algorithmes des deux méthodes (expressions 2D, 3D et 3D), ainsi que la méthode de masquage de 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

FIGURE 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 1995
	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%	

FIGURE 2 – Tableau Comparatif 2D des différents algorithmes

Source	Année	Base	Méthode	Taux de reconnaissance
Zilu Ying Beihang university, Beijing	2009	JAFPE	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% 83.4 %
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% 99.5%

FIGURE 3 – Tableau Comparatif 2D des différents algorithmes

7 Solution

Ce chapitre sera divisé en deux parties principales, donc la première partie présentera le processus d'acquisition de données, le traitement et la composition de l'ensemble de données, et enfin il sera utilisé pour la reconnaissance de la forme faciale et la reconnaissance faciale dans la deuxième partie.

7.1 Données

7.1.1 La composition du jeu de données

Dans cette partie, nous présentons le processus d'acquisition de données, qui nous permettra de construire des modèles de reconnaissance faciale et de reconnaissance de personne. Nous pouvons généralement décrire le processus d'acquisition de données comme suit :

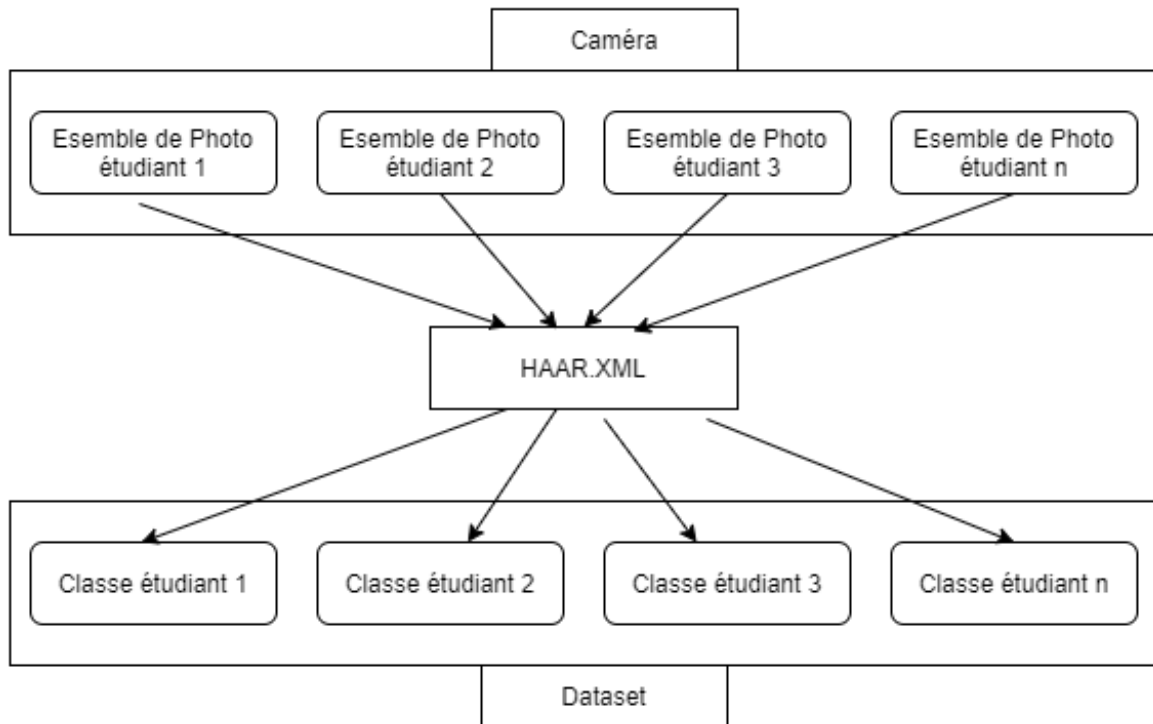


FIGURE 4 – Collecte de données d'entraînement

Dans la figure ci-dessus, nous essayons de récupérer des informations pour chaque étudiant, car notre application est destinée à la surveillance dans une salle de classe, nous avons donc choisi 4 étudiants car nos ressources matérielles ne nous permettent pas de prendre en charge de nombreuses classes pour entraîner notre modèle de reconnaissance faciale.

Les étapes que nous avons suivies sont les suivantes :

1. Capture des visage de chaque étudiant
2. Enregistrez la capture de l'image correspondant au visage de chaque personne.

Dans notre première étape d'enregistrement de visage pour chaque étudiant, l'objectif est de collecter différentes formes correspondant aux visages de chaque étudiant et d'utiliser tous leurs visages pour construire un modèle de reconnaissance faciale. Nous avons réalisé des enregistrements pour 4 étudiants afin de faciliter le traitement.

Après avoir enregistré les différentes formes de visage pour chaque étudiant, nous utilisons la fonction `CascadeClassifier()` disponible dans `opencv`, qui accepte le paramètre `haarcascade_frontalface_alt2.xml`, qui est un fichier xml contenant le descripteur de visage de la personne. Effectuer la même opération séquentiellement comprend des descripteurs utilisés pour décrire les contours du visage humain.

Le descripteur de Haar permet d'extraire la forme d'un objet d'une image. Il existe dans de nombreux types, tels que les personnes, les corps, les mains, les yeux, et bien d'autres types.

Nous le rendrons légèrement plus haut dans la suite. Les détails les plus avancés.

Vous pouvez trouver de nombreux autres paramètres mis en évidence par cette fonctionnalité dans notre code source.

Voici l'arborescence de notre dossier base d'image correspondant aux différents visages détectés (.png) dans la capture d'écran ci-dessous.

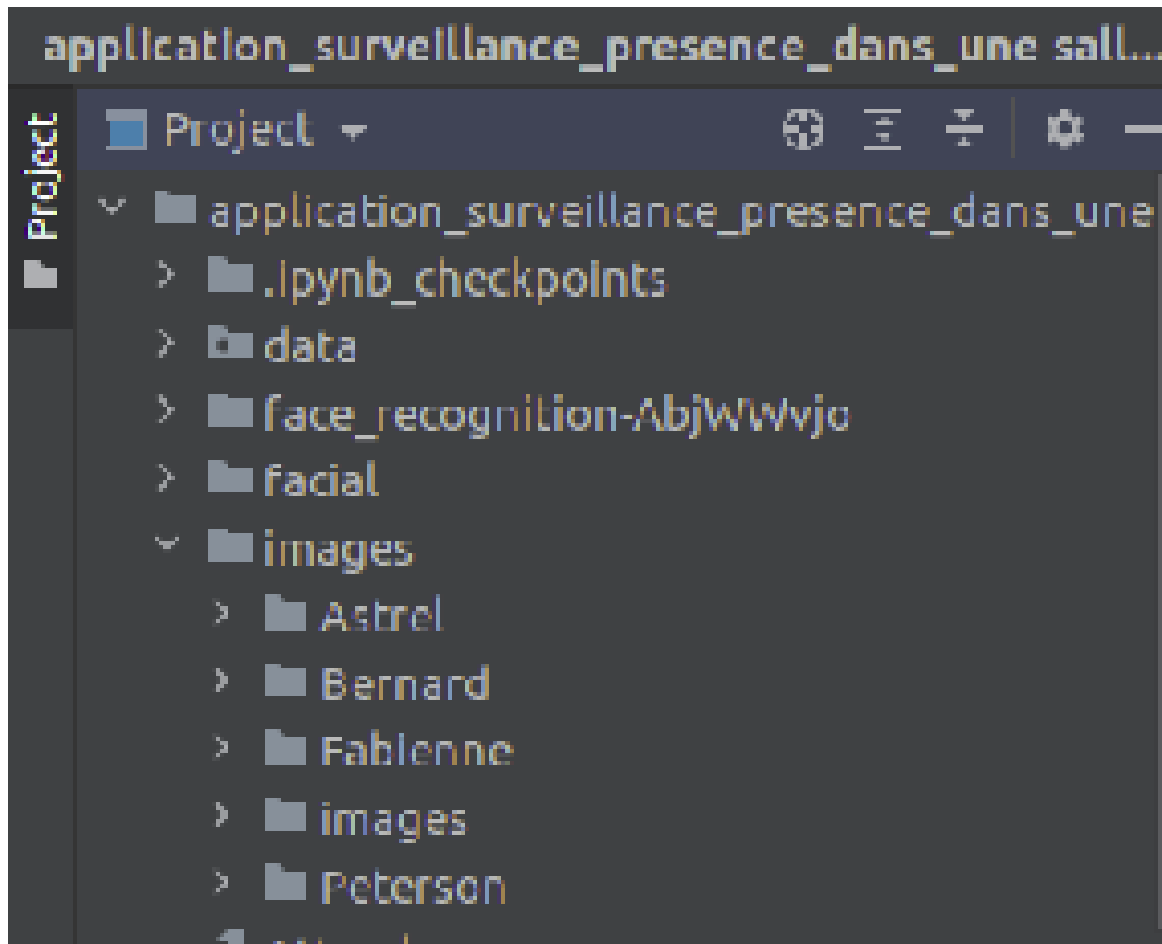


FIGURE 5 – Arborescence du dataset

7.2 Outils

Pour la détection de visage, nous avons utilisé les outils existants dans la bibliothèque OpenCV, car elle a implémenté une variété d'outils de traitement d'image, en particulier la méthode "Viola-Jones" que nous avons choisie pour la détection de visage.

7.2.1 Présentation de la bibliothèque OpenCV

OpenCV [OPENCV, 2007] (Open Source Computer Vision Library), est une bibliothèque de traitement d'images et de vision par ordinateur C / C ++, récemment optimisée par python et fournie par Intel pour Windows et Linux. Il est «open source» et comprend de multiples solutions de traitement d'image et d'analyse de mouvement. Par conséquent, nous utiliserons l'environnement suivant :

1. Framework : Anaconda
2. IDE : Visual studio
3. Langage : Python
4. Librairies : Keras, Scikit learn, numpy, pickle etc.

7.2.2 Pré-traitement

Pour les opération de prétraitement sur nos image, nous utilisons les fonctions :

cvtColor() : Il prend l'image ou la capture comme paramètre, puis la convertit en niveau de gris pour manipuler correctement les paramètres de couleur, car avec les images en couleur, elles contiendront plus de fonctionnalités, qui peuvent être déformées et entraîner un temps de traitement de capture plus long.

7.2.3 Utilisation des Réseaux de Neurone Convolutionnels

L'architecture générale de fonctionnement d'un réseau de neurones convolutifs est la suivante

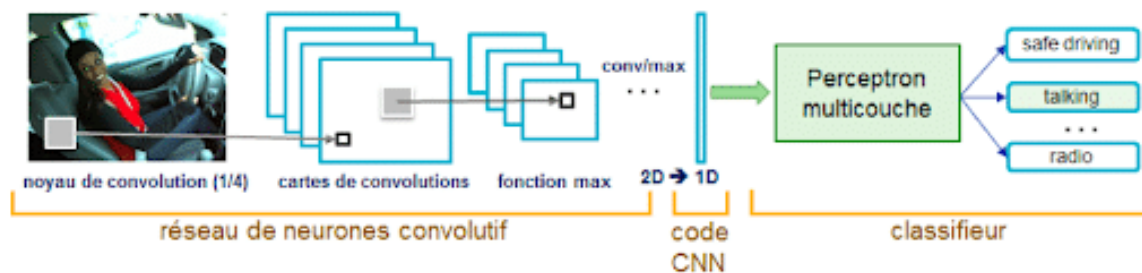


FIGURE 6 – architecture générale de fonctionnement d'un réseau de neurones convolutifs

Dans ce cas, nous utiliserons les fonctions suivantes :

convolution2d() : pour ajouter des couches de convolution

maxpooling2d() : pour réduire la taille des différentes images reçu des couches de convolutions

flatten() : permet d'aplatir les images sous forme de tableau array qui pourrait être enfin pris en paramètre pendant l'entraînement du modèle.

dropout() : fonction permettant de diminuer le taux d'apprentissage du modèle afin d'éviter un surapprentissage

sequential() : permet d'initialiser notre réseau de neurones.

7.3 Modèles conceptuels / Processus de traitement

Dans cette section, nous présenterons le modèle conceptuel de données, les algorithmes utilisés et notre processus de traitement.

7.3.1 Modèle de de conception de l'architecture de notre CNN

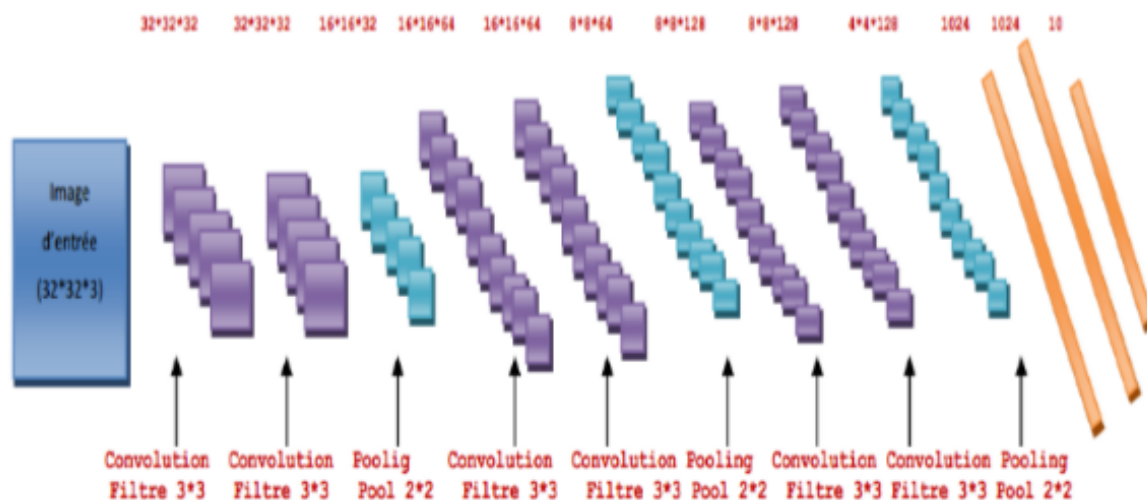


FIGURE 7 – Notre modèle utilisant les CNNs.

7.3.2 Processus de traitement

Ici, nous présentons le processus du traitement à la reconnaissance des photos dans la vidéo.



FIGURE 8 – Processus de traitement

Ci-dessus, nous avons présenté le modèle conceptuel des applications de détection de visage basé sur des descripteurs HAAR et des réseaux de neurones. Nous pouvons voir le pipeline de fonctionnement général de notre application dans ce diagramme plus détaillé.

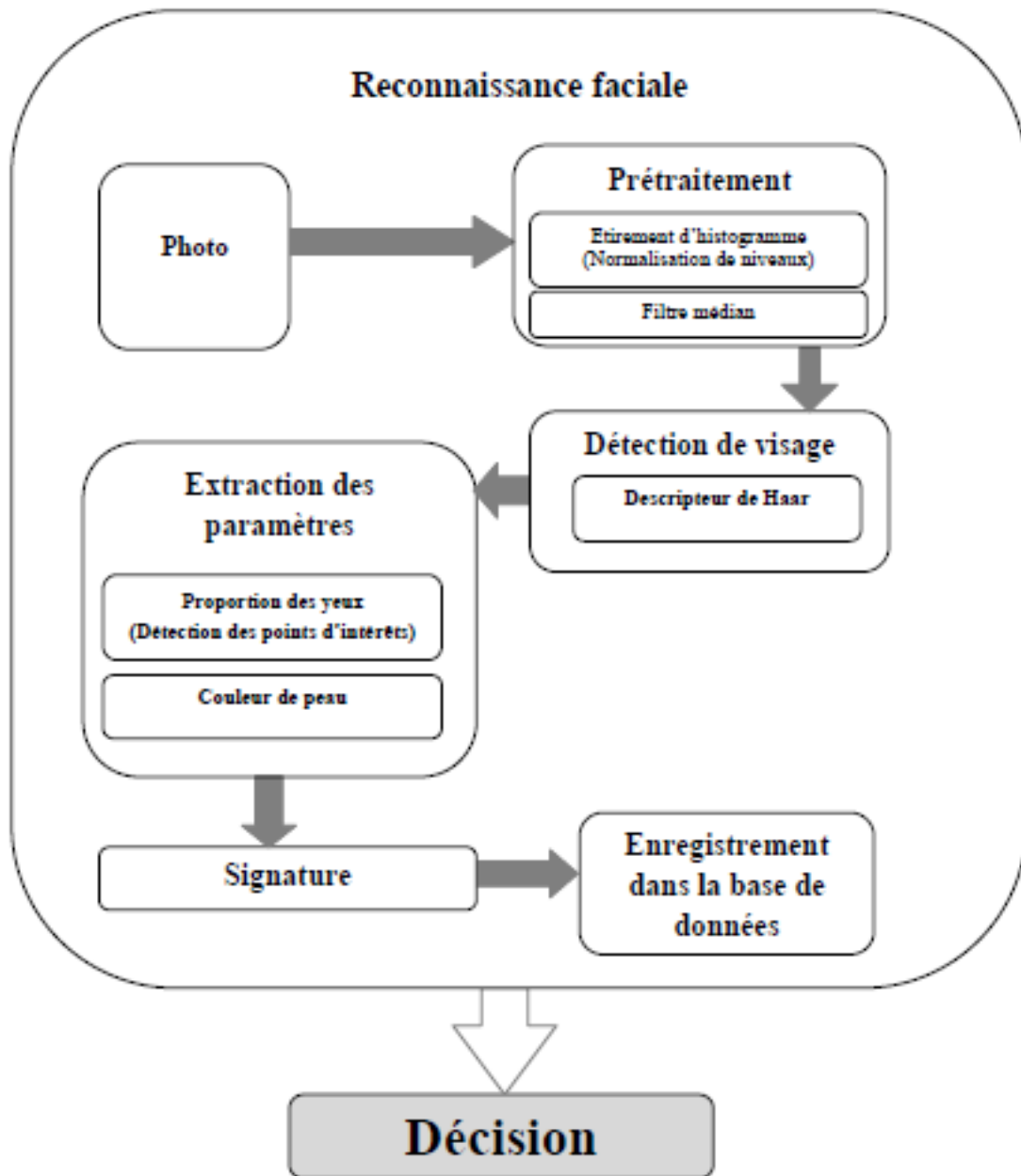


FIGURE 9 – Architecture générale du système

8 Implémentation et expérimentation

Dans ce chapitre, nous définirons l'architecture des deux modèles que nous utilisons en nous basant sur les images capturées ci-dessus. À cette fin, nous utiliserons les bibliothèques et `cnn`s `Tensor ow` et `Keras` pour l'apprentissage et la classification, ainsi que certaines fonctions de la bibliothèque `opencv` pour améliorer les performances du modèle. Nous utiliserons des techniques simples et efficaces, telles que l'augmentation des données et `dropout` pour les `Cnns`.

8.1 Implementation des Cnns

TensorFlow : est un cadre de programmation pour les calculs numériques, qui a été open source par Google en novembre 2015. Son nom est notamment inspiré du fait que le fonctionnement actuel des réseaux de neurones se fait principalement au travers de tables de données multidimensionnelles appelées Tensors. Un tenseur bidimensionnel est égal à une matrice. Aujourd'hui, les principaux produits de Google sont basés sur TensorFlow : Gmail, Google Photos et la reconnaissance vocale.

Keras : est une API de réseau neuronal de haut niveau écrite en Python qui peut s'exécuter sur TensorFlow ou Theano. Son développement se concentre sur l'expérimentation rapide. Être capable de tirer des résultats de l'idée dans les plus brefs délais est la clé d'une bonne recherche. Il fait partie des travaux de recherche du projet ONEIROS (Open Neuroelectronic Intelligent Robot Operating System) dont le principal auteur et mainteneur est l'ingénieur Google François Chollet. En 2017, l'équipe TensorFlow de Google a décidé de prendre en charge Keras dans la bibliothèque principale TensorFlow. Chollet a expliqué que Keras a été conçu comme une interface plutôt que comme un environnement d'apprentissage de bout en bout. Il fournit un ensemble d'abstraction de niveau supérieur et plus intuitif et peut facilement configurer un réseau neuronal indépendant de la bibliothèque informatique principale. Microsoft travaille également à ajouter le backend CNTK à Keras.

Python : est une interprétation de haut niveau (pas d'étape de compilation) avec une sémantique dynamique, un langage de programmation orienté objet. Un grand nombre de développeurs et de programmeurs en ont une forte demande. Python est un langage simple, facile à apprendre et peut réduire considérablement les coûts de maintenance du code. Les bibliothèques (packages) Python encouragent la modularité et la réutilisabilité du code. La plupart des plates-formes fournissent gratuitement Python et ses bibliothèques (ainsi que son code source ou ses binaires) et peuvent être redistribuées gratuitement.

Scikit-learn : est une bibliothèque Python gratuite dédiée à l'apprentissage automatique. Il est développé par de nombreux contributeurs, notamment académiques, par des établissements d'enseignement supérieur et de recherche français, comme Inria et Télécom ParisTech. Il comprend la fonction d'estimation des forêts aléatoires, la régression logistique, les algorithmes de classification et les machines vectorielles de support. Il vise à s'harmoniser avec d'autres bibliothèques Python gratuites (y compris NumPy et SciPy).

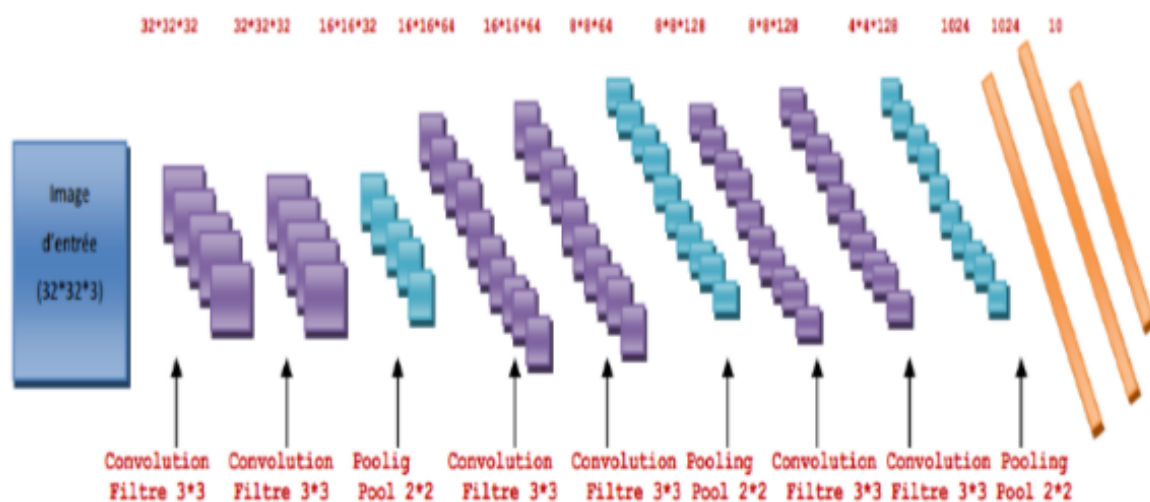
8.2 Architecture de notre réseau

Au cours de notre expérience, nous avons créé un modèle avec une architecture différente (améliorable) basée sur l'image capturée ci-dessus. Ensuite, nous introduisons l'architecture du réseau : Notre modèle se compose de cinq couches de convolution, deux couches de maxpool et trois couches entièrement connectées. La taille de l'image d'entrée est de $150 * 150$, et l'image entre d'abord dans la première couche convolutionnelle. Cette couche est composée de 32 filtres d'une taille de $3 * 3$. Après chacune de nos couches convolutives, il y a une fonction d'activation ReLU, qui oblige le neurone à renvoyer une valeur positive. Après cela, 32 filtres d'une taille de $32 * 32$ sera créé. Carte des caractéristiques de convolution.

Les 32 cartes d'entités précédemment obtenues sont données en entrée de la deuxième couche convolutive. La deuxième couche convolutive est également composée de 32 filtres. La fonction d'activation RELU est appliquée sur la couche convolutive, puis nous appliquons Maxpool pour réduire le nombre et le calcul des paramètres de taille pour les petites images. À la sortie de cette couche, nous aurons 32 cartes d'entités d'une taille de $16 * 16$.

Les mêmes étapes sont répétées pour les couches 3, 4 et 5. Ces couches sont composées de 64 filtres et la fonction d'activation ReLU est toujours appliquée à chaque convolution. La couche Maxpool est appliquée après la cinquième convolution. À la sortie de cette couche, nous aurons 64 cartes d'entités d'une taille de $8 * 8$. La taille du vecteur de caractéristiques généré par convolution est de 4096.

Après ces cinq couches de convolution, nous utilisons un réseau de neurones composé de trois couches entièrement connectées. Les deux premières couches comportent chacune 1024 neurones, la fonction d'activation utilisée est ReLU, et la troisième couche est softmax, ce qui permet de calculer la distribution de probabilité de 6 classes (le nombre de classes dans la bibliothèque d'images ci-dessus). Dans la section précédente, nous avons présenté l'architecture du réseau, et dans la section suivante, nous avons présenté la configuration du réseau. Veuillez noter que notre architecture utilisant CNN offre la possibilité d'utiliser les trois attributs de couleur de l'image.



8.3 Précision obtenue après entraînement avec les Cnns

```
57/57 [=====] - 82s 1s/step - loss: 0.1213 - acc: 0.9593 - val_loss: 0.9203 - val_acc: 0.7645
Epoch 98/100
57/57 [=====] - 84s 1s/step - loss: 0.1183 - acc: 0.9622 - val_loss: 1.3506 - val_acc: 0.6911
Epoch 99/100
57/57 [=====] - 84s 1s/step - loss: 0.0751 - acc: 0.9709 - val_loss: 1.3502 - val_acc: 0.6881
Epoch 100/100
57/57 [=====] - 82s 1s/step - loss: 0.0735 - acc: 0.9726 - val_loss: 1.1567 - val_acc: 0.7339
Out[1]: <keras.callbacks.History at 0x7fd8e0ed5fd0>
In [2]:
```

FIGURE 10 – Précision obtenue après entraînement avec les Cnns

8.4 Précision obtenue a partir du nombre d'époque de formations

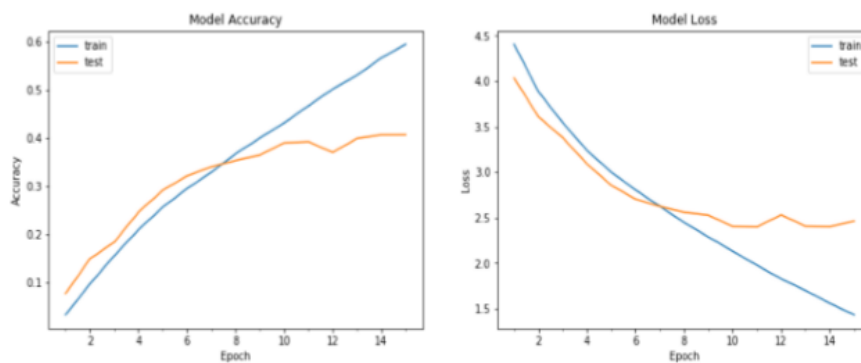
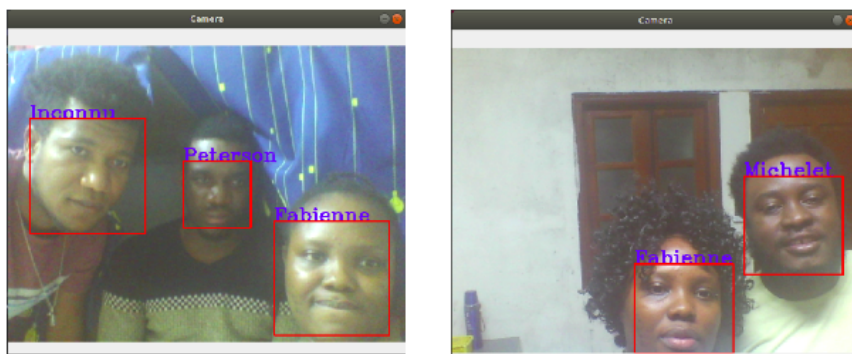


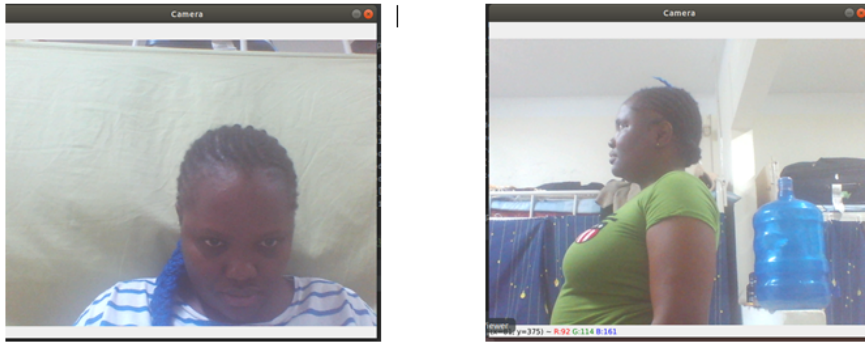
FIGURE 11 – Précision obtenue après entraînement avec les Cnns

9 Résultats

9.1 Bonnes détection



9.2 Mauvaise détection



Dans cette image, nous voyons notre application, qui a réussi à détecter et à créer une boîte englobante lorsqu'un visage était détecté, mais elle était invalide car comme nous l'avons précisé ci-dessus, le descripteur HAAR ne prenait pas en compte la personne abaissée Face ou l'angle d'inclinaison, il est donc impossible de détecter tous les visages. C'est aussi le point le plus important. Il est également important de noter que des amis inconnus postulent sur la tête de tous les élèves. C'est parce que toutes ces personnes ne sont pas reconnues par notre modèle.

10 Conclusion

Dans ce projet, nous avons étudié le problème du contrôle de l'assiduité des étudiants dans les établissements d'enseignement. Cependant, notre objectif est de concevoir et de développer des applications de gestion de l'assiduité des étudiants basées sur la technologie de reconnaissance faciale.

Au cours de ces travaux, différentes technologies de reconnaissance développées ces dernières années ont été analysées pour mettre en évidence leurs caractéristiques, avantages et inconvénients respectifs. Cependant, nous avons découvert des méthodes de reconnaissance faciale holistiques, partielles et hybrides. Compte tenu de la nature du système de gestion des présences, puisque le système fonctionnera en temps réel, nous devons utiliser une technologie de reconnaissance très rapide et puissante.

En ce sens, la première partie de l'application de reconnaissance de visage utilise l'algorithme Viola and Jones, qui est largement considéré comme une méthode qui fonctionne en temps réel et peut fournir des résultats fiables et fiables. La deuxième partie de l'application traite de la reconnaissance faciale partielle.

Notre travail consiste à développer un algorithme efficace qui vise à reconnaître les visages humains en utilisant la méthode des faces propres basée sur l'analyse en composantes principales (ACP). La PCR est une méthode mathématique qui peut simplifier l'ensemble de données en réduisant la quantité de données. Il est utilisé pour représenter efficacement l'image du visage, qui peut être grossièrement reconstruite à partir d'un petit ensemble de poids et d'images de visage standard. Veuillez noter que les faces propres sont considérées comme la méthode de reconnaissance faciale la plus simple et la plus précise, mais de nombreuses autres méthodes ou une combinaison de plusieurs méthodes (beaucoup plus complexes) peuvent être utilisées.

Par conséquent, la réalisation de ce projet nous donne une compréhension plus approfondie du développement logiciel, de la conception à la mise en œuvre, en passant par la logique et la pensée algorithmique.

11 Perspectives

Notre perspective est de mettre en place une application complète via une interface graphique, et d'étudier les changements de posture et d'éclairage pour obtenir la meilleure reconnaissance faciale.

12 Lien github du projet

https://github.com/Fabienne703/Reconnaissances_des_Formes

13 Références

https://fr.wikipedia.org/wiki/Syst%C3%A8me_de_reconnaissance_faciale

<http://www.natural-solutions.eu/blog/le-deep-learning-pour-les-nuls>

<https://www.larousse.fr/encyclopedie/divers/vid%C3%A9osurveillance/101594>

https://link.springer.com/content/pdf/10.1007%2F978-3-540-24670-1_36.pdf