

[Titre du document]

[Sous-titre du document]



*MASTER EN*

*BIG DATA AND CLOUD COMPUTING*

*Réalisé par :*

**DRYOUCH MAROUANE**

*Encadré par* :

**Mr. AMNAI MOHAMED**

***Faculté des sciences, Université Ibn Tofail, Kénitra***

Projet mémoire effectué de :

*01 Mars 2022 à 1 Septembre 2022*

# DÉDICACES

**Avec l’expression de ma reconnaissance , je dédie ce travail à ma famille , elle qui m’a doté une éducation digne , son amour fait de moi ce que je suis aujord’hui , à leurs encouragement et soutien morale et matériel durant toute ma carrière .**

**Å tous mes collègues de la filière Master en Big Data and Cloud à cause de leur aide et encouragement durant cette formation.**

# 

# REMERCIEMENT

***Avant d’entaminer, je tiens à témoigner mon profonde gratitude à toutes les personnes qui ont participé de loin ou de prés à l’élaboration de ce projet de fin d’étude.***

***Je tiens à remercier dans un premier temps, toute l’équipe pédagogique de La faculté des sciences Kénitra et tous les professeurs et les responsables de la formation Master en Big Data and Cloud.***

***Mes sincères remerciement sont adressés à :***

***Mr. Mohamed AMNAI  : Professeur en informatique à la faculté des sciences Ibn Tofail, pour m’avoir encadré durant toute la période du projet et pour ses conseils qui m’ont aidés à réaliser ce travail d’une bonne manière.***

***Ainsi, nous adressons nos extrêmes gratitudes au jury de notre soutenance de projet de fin d'étude, qui a accepté d’évaluer notre travail.***

# RÉSUMÉ

Dans le cadre de sa formation du Master en Big Data and Cloud, la faculté des sciences Ibn Tofail veille à compléter la formation de ses étudiants par un projet de fin d’études, c’est plus particulièrement un sujet de recherche qui vise à appliquer la science pour traiter des problématiques précis.

En effet, le projet de fin d’étude permet à l’étudiant d’acquérir une expérience pratique et théorique et d’élargir les champs de ses connaissances accumulées durant la période de sa formation, cette expérience est encadrée par le pilotage et le soutien de l’équipe pédagogique.

Durant ce projet de recherche nous sommes amenés à réaliser un système de reconnaissance faciale pour améliorer la performance des étudiants, dont l’objectif principale d’identifier les étudiants et leurs émotions pendant les séances des cours, afin d’analyser ces résultats pour déduire à la fin les facteurs qui influencent la réussite des étudiants et mettre des décisions efficaces pour les lutter ou les diminuer.

Ensuite nous devons concevoir un modèle de reconnaissance faciale à l’aide des algorithmes de l’apprentissage profond « **Deep Learning** », qui vont nous servir à bien extraire des incorporations faciales uniques à partir des images ou des vidéos et les comparer avec d’autres qui existent dans la base de données.

**Mots-clés :** *Apprentissage profond, reconnaissance faciale, incorporations faciales, sujet de recherche.*

# Abstract

As part of its Master's training in Big Data and Cloud, the Ibn Tofail Faculty of Sciences ensures that the training of its students should completed by a graduation project, it is more particularly a research subject for the goal to apply science to address specific issues.

Indeed, the final project of study allows the student to acquire practical and theorical experience and to broaden the fields of his knowledge accumulated during the period of his training, this experience is framed by the management and support of the educational team.

During this research project we are led to create a facial recognition system to improve student performance, the main objective of which is to identify students and their emotions during class sessions, in order to analyze these results to deduce the end the factors that influence student success and put effective decisions to eliminate or reduce them.

Then we have to design a facial recognition model using Deep Learning algorithms, which will help us to extract unique facial incorporations from images or videos and compare them with others. that exist in the database.

**Keywords :** *Deep Learning, Facial recognition, incorporations, recognition system, research subject.*

Table des matières

[DÉDICACES 1](#_Toc104995496)

[REMERCIEMENT 2](#_Toc104995497)

[RÉSUMÉ 3](#_Toc104995498)

[Abstract 4](#_Toc104995499)

[Liste des abréviations 8](#_Toc104995500)

[Introduction 9](#_Toc104995501)

[CHAPITRE 1 : CONTEXTE GÉNÉRALE DU PROJET 11](#_Toc104995502)

[**Introduction** 11](#_Toc104995503)

[**1.** **Présentation du domaine de recherche : Système de reconnaissance faciale** 12](#_Toc104995504)

[**2.** **Historique de la reconnaissance faciale** 12](#_Toc104995505)

[**3.** **Fonctionnement de la reconnaissance faciale** 13](#_Toc104995506)

[**4.** **Étude de projet** 15](#_Toc104995507)

[4.1. Présentation du projet 15](#_Toc104995508)

[4.2. Spécification des besoins 15](#_Toc104995509)

[4.3. Étude de l’existant 17](#_Toc104995510)

[4.4. Problématique 21](#_Toc104995511)

[4.5. Les solutions proposées 23](#_Toc104995512)

[4.6. Planification du projet 26](#_Toc104995513)

[**Conclusion :** 28](#_Toc104995514)

[CHAPITRE 2 : SYSTÈME DE RECONNAISSANCE FACIALE 29](#_Toc104995515)

[**Introduction** 30](#_Toc104995516)

[**1.** **Type des données utilisés** 30](#_Toc104995517)

[1.1. Image 2d : 30](#_Toc104995518)

[1.2. Image 3d 30](#_Toc104995519)

[1.3. Vidéo 31](#_Toc104995520)

[**2.** **Détection et identification des visage** 32](#_Toc104995521)

[2.1. Détection de l’objet figure 32](#_Toc104995522)

[2.2. Les caractéristiques concernées par la détection 35](#_Toc104995523)

[2.3. Les modèles de détection des traits de visages 36](#_Toc104995524)

# Liste des abréviations

# Introduction

De nos jours le monde continue de se développer sur les différents côtés, et dans ce sens la technologie présente aujourd’hui un facteur incontournable en matière de développement mondiale, cela nous oblige de se transformer à la digitalisation pour suivre ce rythme.

Enfaite l’objectif de la digitalisation est d’automatiser plusieurs traitements qui se font par l’être humain, ce qu’il va nous permet de garantir une bonne performance dans un délai très important dans des différents cas d’utilisation.

Dans notre cas, nous avons choisi une technologie biométrique qui répond à un besoin très persistant, c’est ce qu’on appelle la reconnaissance faciale qui a comme objectif principale de bien identifier les visages pour des différentes raisons (sécurité, identification des humours, gestion d’absence, gestion des accès …).

Donc vu l’importance de cette technologie, nous allons réaliser un système de reconnaissance faciale pour l’objectif d’améliorer la performance des étudiants en faisons la gestion d’absence, et l’identification de l’état psychique des étudiants durant les séances des cours, cela va nous servir de détecter quelques causes internes responsables à la défaillance éducatif des étudiants au sein de l’université.

En générale durant ce rapport nous allons citer les différentes étapes suivies pour traiter ce sujet, ils contiennent une étude théorique et sa représentation dans la pratique pour que nous réussissons à atteindre notre objectif à la fin.

# CHAPITRE 1 : CONTEXTE GÉNÉRALE DU PROJET

## **Introduction**

Dans ce chapitre nous allons essayer de faire situer la pensée et l’idée principale que l’on veut analyser et de montrer l’environnement dans lequel le sujet se situe, pour cela nous devons présenter d’une manière explicite notre sujet « Système de reconnaissance faciale », d’identifier les problématiques auxquelles nous faisons faces, ainsi les solutions proposées, les systèmes existants, le cadre générale du projet, et en fin la planification de notre projet de recherche.

## **Présentation du domaine de recherche : Système de reconnaissance faciale**

Un système de reconnaissance faciale est un logiciel qui utilise une technologie qui vise à reconnaitre et à détecter une personne à partir de son visage d’une façon automatique, il répond à un besoin très important vu les différentes erreurs et fraudes dans des différentes domaines lié à l’identification physique des personnes.

Ainsi les applications de la reconnaissance faciale sont multiples comme la robotique, la biométrie, l’identification des personnes dans les photos sur les réseaux sociaux comme le cas de Facebook, Google Image…, et aussi au niveau de la sécurité biométrique.

## **Historique de la reconnaissance faciale**

(Système de reconnaissance faciale, s.d.)

Le 7 juin 2011, Facebook lance un service de reconnaissance faciale accessible à tous. Celui-ci a été vivement critiqué par un bon nombre d'associations de protection de la vie privée et du droit à l'oubli. Sous cette pression, Facebook a dû abandonner ce service le 22 septembre 2012 en Europe9.

Un autre projet de reconnaissance de visage, appelé DeepFace, développé par le groupe de recherche de Facebook a également commencé à voir le jour début 2015. Cet outil a pour but de reconnaître un visage, quelle que soit son orientation, ainsi que d'y associer une personne. DeepFace permet donc de faire correspondre un très grand nombre de photos différentes d'une même personne, même si celle-ci n'est pas identifiée explicitement. Pour réaliser cela, le système est composé d'un réseau de neurones artificiels entraîné via l'apprentissage profond sur des millions d'images10.

Toutefois, cette technologie évolue rapidement et cherche à se démocratiser via le développement de nombreux cas d'usage. Depuis fin 2018, les annonces se multiplient faisant entrer progressivement la reconnaissance faciale dans le quotidien, tous les systèmes montrant une avancée spectaculaire, tant au niveau de l'efficacité des algorithmes employés qu'au niveau de la masse de données pouvant être traitée par les programmes. En 2019, Cydral Technology diffuse un système utilisable depuis un simple téléphone et propose de ce fait le tout premier moteur de recherche grand public11 fonctionnant à l'image de la recherche inversée proposée par Google mais pour l'identification de profils exclusivement.

En janvier 2020, le New York Times révèle l'existence de la société Clearview AI qui commercialise auprès d'agences de sécurité une application de reconnaissance faciale qui permet d'identifier quelqu'un à partir d'une photographie12. L'auteur de l'enquête souligne que cette application remet en cause notre conception de la vie privée13,14.

En juin 2020, la société espagnole Herta Security remporte un label d'excellence pour une technologie capable d'analyser des visages et de surveiller la foule. Cette technologie pourrait être utilisée en cas de nouvelle vague de Covid-19. Ce logiciel, nommé Aware, est capable de vérifier la distanciation sociale et d'identifier les personnes à différents points de contrôles15

## **Fonctionnement de la reconnaissance faciale**

La reconnaissance faciale est très connue au niveau de l’identification des personnes et par exemple : le déverrouillage des comptes …, mais il existe d’autres cas spéciaux d’utilisation qui apparaissent selon l’évolution de la technologie à chaque fois, alors nous citons les étapes nécessaires dans le fonctionnement de la reconnaissance faciale :

1. **Stockage des données sous forme des images :**

Il faut tout d’abord avoir une base de données contiennent des images pour les comparer avec les images entrantes ou bien pour déduire les états psychiques des personnes selon les signes de leurs visages.

1. **Détecter les visages :**

Cette étape se fait dans le coté client, il permet de récupérer une image pour faire une comparaison avec les images qui existent au sein de la base de données, et généralement des systèmes de reconnaissance faciale cette étape doit être faite en temps réel pour garantir la fiabilité des images entrantes, et c’est grâce aux algorithmes de l’intelligence artificielle et l’apprentissage automatique ces traitements peuvent avoir un taux de fiabilité très élevé en temps réel.

1. **Analyser le visage :**

Généralement la majorité des technologies de reconnaissance faciale font une analyse des images en 2d car c’est plus pratique qu’une analyse en 3d, elles se basent sur la distance entre les yeux, les sourcils, la forme des pommettes, la profondeur des orbites, la distance entre le front et le mention…, ainsi d’autres détails pour bien identifier les spécifications du visage.

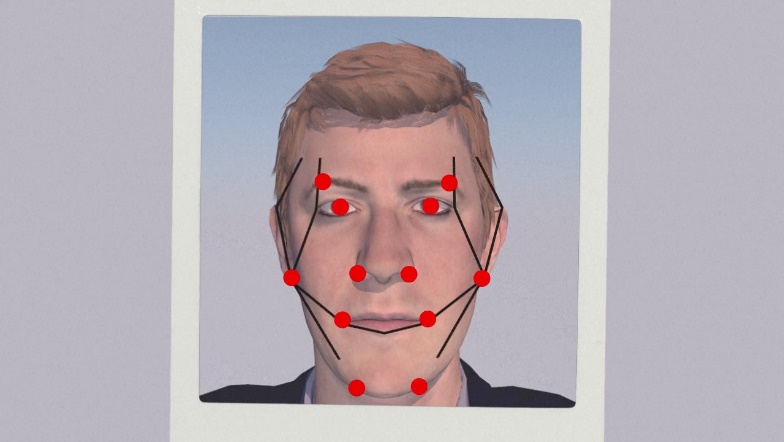


Figure : Exemple des points concernés par l'analyse du visage

1. **Conversion de l’image en des données :**

Le processus de capture de visage fait la transformation d’une image en ensemble des informations numériques selon les caractéristiques de chaque visage, cette étape est appelée l’empreinte faciale qui se base sur des formules mathématiques et qui se représente sous forme d’un code numérique qui est unique pour chaque personne.

1. **Trouver une correspondance :**

Dans cette partie le système cherche une correspondance entre l’image traité et les images qui existent dans la base de données, par exemple le FBI à un accès à une base de données qui contient plus de 650 millions images sur de nombreuses base de données de l’état, et sur Facebook…, pour qu’il puisse identifier la personne concernée.

## **Étude de projet**

### Présentation du projet

Dans le cadre de notre recherche, notre projet sera destiné à l’université **Ibn Tofail**, C’est une université marocaine crée en 1985, elle contient presque 21375 étudiants et se situe à la ville Kénitra.

Alors nous devons établir un système de reconnaissance faciale pour l’université dans le but principale est et d’identifier étudiants, puis il détecte ses émotions et réactions, ce système peut avoir des différentes utilisations dans l’espace universitaire ou bien les services pédagogiques, et voilà un ensemble de ces utilisations possibles :

* **Évaluation de la performance des étudiants**
* **Détection de la fraude durant les examens**
* **Contrôle des accès aux espaces privés**
* **Gestion des absences**
* **Gestion de la bibliothèque**

### Spécification des besoins

Un système de reconnaissance faciale au sein de l’université Ibn Tofail représente des besoins intensifs pour plusieurs raisons, pour cette raison nous distinguons deux types des besoins :

* **Besoins fonctionnels**

Dans notre système nous avons trois acteurs principaux :

Tableau : Les besoins fonctionnels

|  |  |
| --- | --- |
| Acteur | Rôle |
| Administrateur | * Responsable de la mise à jour de la base des données qui contient les informations et les images des étudiants. * Il fait la gestion des étudiants et filière * Il fait la gestion des emplois du temps. |
| Professeur | * La consultation des rapports d’absence des étudiants, ou des statistiques sur les émotions des étudiants durant les séances des cours (*dans le cas d’une intégration dans un système*). |
| Étudiant | * L’interaction directe avec le système (identification). * La consultation des rapports personnelles (*dans le cas d’une intégration dans un système*) |

* **Besoins non fonctionnels**

Tableau : Les besoins non fonctionnels

|  |  |
| --- | --- |
| Besoin | Utilité |
| La rapidité de traitement | En effet, vu le nombre important des transactions quotidiennes, il est impérativement nécessaire que la durée d'exécution des traitements s'approche le plus possible du temps réel. |
| La portabilité | Dans notre cas le système que nous devons réaliser doit impérativement facile à l’adaptation et l’installation pour que nous pouvons l’intégrer avec d’autres systèmes d’une manière simple. |
| Le rendement et l’efficacité | Notre système doit être avant tout performant c'est à-dire à travers ses fonctionnalités, répond à toutes les exigences des utilisateurs d'une manière optimale. |
| La convivialité | Notre système doit être facile à utiliser. En effet, les interfaces utilisateurs doivent être conviviales c'est-à-dire simples, ergonomiques et adaptées à l'utilisateur. |
| La maintenabilité | Ce besoin est très important car notre système doit tourner toujours, et il a un impact sur plusieurs processus au sein de l’université, alors il doit être stable et facile au modification et maintenance. |

### Étude de l’existant

#### Le système existant à l’université

Alors au sein de l’université Ibn Tofail cette identification au sein des séances des cours ou des examens, bibliothèque …, se fait d’une manière classique, dans laquelle l’étudiant doit se présenter avec sa carte d’étudiant, puis les administrateurs ou les professeurs font une vérification manuelle pour savoir si la figure de la personne physique est correspondante à celle trouvé dans la carte, ainsi ils récupèrent leurs informations après la recherche par son numéro d’apogée dans un système informatique.

(Identification par carte, s.d.)



Figure : Exemple d'une vérification manuelle de l'identité

#### Les solutions existantes sur le marché

Puisque cette technologie répond à un besoin intensif, alors nous trouverons plusieurs sociétés qui offre des solutions liées à ce sujet, pour cela nous citons quelques-unes :

1. **Amazon Rekognition :**

*(Système de reconnaissance faciale, s.d.)*

Amazon Rekognition est l'un des noms les plus fiables du jeu logiciel de reconnaissance faciale :

* Il peut identifier des objets et des scènes en leur donnant des étiquettes.
* Amazon a l'avantage de disposer d'un ensemble de données massif. Cela garantit la précision.
* Rekognition permet aux utilisateurs d'ajouter des étiquettes personnalisées aux objets et aux scènes pour répondre aux besoins de l'entreprise. Cela garantit des taux de réussite de correspondance plus élevés. Il suffit à l'entreprise de fournir des images des objets et des lieux à identifier.
* Amazon propose une modération de contenu. Il dispose d'un ensemble de données testé qu'il utilise pour signaler le contenu inapproprié. Il fournit également une détection de texte pour reconnaître les noms.
* Sa dernière offre est la détection d'EPI qui identifie les travailleurs portant des EPI pour des raisons de sécurité.
* Il permet également la recherche et la vérification des visages dans un référentiel privé de photos. Veuillez noter que des frais de stockage sont encourus dans ce cas.
* Rekognition fournit également la détection et l'analyse des visages dans les vidéos - en direct ou stockées.

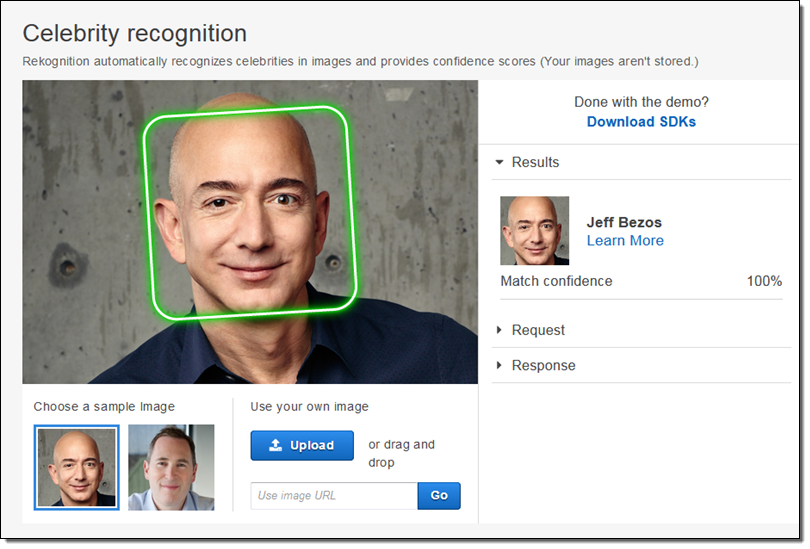


Figure : Amazon Rekognition

**Clientèle :** la clientèle actuelle d'Amazon Rekognition comprend des médias, des sociétés d'analyse de marché, des sites de commerce électronique et des solutions de crédit.

**Support client** : la solution fournit des documents, des exemples et des didacticiels.

**Modèle de tarification** : Le modèle de tarification est basé sur les services utilisés et le nombre d'images analysées. Dans le cadre de l'offre gratuite d'AWS, le traitement d'image peut être utilisé gratuitement. Le niveau gratuit dure 12 mois et vous permet d'analyser 5 000 images par mois et de stocker 1 000 métadonnées de visage par mois. Après cela, le coût mensuel est calculé en utilisant le nombre d'images traitées (plus les frais de stockage, le cas échéant). Des packages similaires basés sur des niveaux sont disponibles pour le traitement vidéo et les étiquettes personnalisées.

**Commentaires éditoriaux :** Amazon Rekognition peut être utilisé par des organisations qui n'ont aucune expertise en apprentissage automatique. C'est une solution coûteuse, donc c'est pour les entreprises avec un budget à revendre. Une chose que les utilisateurs avertissent de garder à l'œil est l'addition des coûts de stockage.

1. **Beta face :**

*(Les logiciels de reconnaissance faciale, s.d.)*

Betaface se concentre principalement sur l'analyse des images et des vidéos et la reconnaissance des visages et des objections.

* Il propose trois types de services : les SDK de reconnaissance faciale, les services de développement de logiciels client et les services Web hébergés.
* Les services peuvent être une simple détection de visage ou une reconnaissance, une identification et une vérification faciales complexes.
* Il utilise des mesures biométriques pour le suivi des caractéristiques faciales sur les images et les vidéos.
* Il peut reconnaître l'émotion et l'ethnicité.
* Il peut également suivre la peau, les cheveux, les traits du visage et la forme de la coiffure.
* Elle fournit des solutions logicielles de vidéosurveillance et de sécurité.
* **Clientèle :** la clientèle de Betaface comprend des agences de publicité et de divertissement Web, des producteurs de contenu multimédia et des développeurs de logiciels B2B.
* **Support client :** Betaface fournit une assistance à l'intégration. Il aide également à la prise en charge complète du service Web et de l'infrastructure de base de données, le cas échéant.
* **Modèle de tarification :** Betaface vous permet d'utiliser son API gratuitement pour 500 images par jour et 0,035 € par image par la suite. Il propose également trois plans d'abonnement mensuels - le plan de base à 245 $ pour le traitement de 40 000 images, le plan premium à 490 $ pour 100 000 images et le plan ultra à 1 595 $ pour 300 000 images. Des frais supplémentaires sont facturés pour le traitement de chaque image supplémentaire en fonction du plan d'abonnement.
* **Commentaires éditoriaux** **:** Betaface peut être utilisé par des organisations qui sont un peu férues de technologie et qui ont peut-être leurs propres développeurs qui cherchent à intégrer leur application avec un FRS. Cette solution peut convenir aux petites entreprises.

1. **BioID**

BioID est une solution conforme au RGPD qui fournit la biométrie en tant que service. Il fournit des services FRS basés sur le cloud auxquels votre produit peut accéder à l'aide d'API. Le logiciel propose trois produits :

*Web service BioID :* Il s'agit d'une offre SaaS qui peut être déployée sur site ou sur le cloud.

*Détection de la vivacité :* il s'agit d'un service de reconnaissance permettant de détecter la présence de l'utilisateur à l'aide de la reconnaissance faciale, oculaire et vocale. Il est utilisé pour prévenir la fraude en ligne et les menaces d'identité.

*PhotoVerify :* Cette solution combine la technologie de détection de visage avec le service Liveness Detection de BioID pour vérifier les photos utilisées comme preuves d'identité.

* **Clientèle :** cette solution s'adresse aux entreprises qui ont besoin d'identification et de signature électroniques. Il fournit des services aux services financiers, KYC avec biométrie, location de voitures en libre-service, soins de santé et
* organisateurs d'examens en ligne.
* **Support client :** BioID dispose d'une solide documentation en ligne de ses API.
* **Modèle de tarification :** les services de BioID peuvent être achetés à un niveau de paiement par utilisateur. Vous pouvez contacter BioID pour les prix.
* **Commentaires éditoriaux :** Les organisations disposant de leurs propres applications ou produits peuvent envisager BioID. Il est idéal pour les entreprises où la vérification d'identité est essentielle, comme les locations de voitures. Les entreprises présentes sur tous les continents peuvent envisager cette solution car elle est conforme au RGPD.

Et enfin il existe plusieurs logiciels sur le marché qui ont actif dans le domaine de la reconnaissance faciale.

### Problématique

Cette partie va nous aider à bien savoir notre objectif de recherche, dans laquelle nous devons détecter les problématiques liées aux systèmes existants pour que notre recherche sert à répondre à ces problématiques.

Et pour cette raison nous citons en dessus les problématiques concernant les systèmes existants au sein de l’université Ibn Tofail selon chaque cas d’utilisation.

|  |  |
| --- | --- |
| Cas d’utilisation | Problématique |
| Gestion des absences | La gestion des absences d’une façon manuelle consomme beaucoup du temp et d’énergie par les professeurs à chaque séance, ce qui est n’est pas souhaitable si on veut profiter le maximum possible de la séance. |
| Détection de la fraude durant les examens | L’organisation et le déroulement des examens au sein des établissements présente un grand défi et surtout en ce qui concerne la fraude et le fait de tricher durant les examens, donc cette opération se fait généralement par des professeurs et parfois il est difficile de contrôler toute la classe pendant la période d’examen. Ainsi la vérification de l’identification des étudiants manuellement durant les examens produit plusieurs erreurs, et si le nombre des étudiants est élevés, il est probablement qu’il sera des conflits au niveau du pointage durant l’examen. |
| Contrôle d’accès aux espaces privés | Au sein d’un établissement universitaire existe plusieurs espaces qui ont un accès restreint, par exemple :   * Les laboratoires * Les Archives qui contiennent des documents critiques.   Alors la gestion de ces espaces présente une problématique réelle. |
| Évaluation de la performance des étudiants | Au sein d’un établissement universitaire existe plusieurs problèmes au niveau de l’apprentissage des étudiants et aussi de la manière par laquelle les cours sont présentés et les méthodes utilisées. |
| Gestion de la bibliothèque | Au sein de chaque université il existe une bibliothèque ou plusieurs, cela présente un grand défi au niveau de la gestion vu le volume des livres et documents existants, ainsi parfois des étudiants non-inscrits s’entrent avec une carte d’un autre étudiant, ainsi au niveau de l’association de l’étudiant avec le livre loué se fait d’une manière manuelle ce qu’il risque d’avoir beaucoup d’erreur (perte des livres, le non-respect du délai). |

Figure : Les problématiques trouvés

Et concernant les problématiques liés aux existants du marchés :

* La plupart des solutions disponibles sur le marché sont très coûteux, surtout que le nombre des étudiants à l’université est très grand, ce qu’il permettra un grand taux de consommation des données, cela va produire des tarifs d’abonnement très élevés.
* Le traitement des données personnelles des étudiants comme les photos présente un grand risque de confidentialité s’il se fait en dehors du système de l’université.

### Les solutions proposées

Pour résoudre les problèmes cités, nous devons proposer plusieurs solutions qui seront l’objectif de ce mémoire de recherche, Alors nous proposons comme solution principale :

* La réalisation d’un système de reconnaissance faciale qui doit être capable à se communiquer avec les autres systèmes existants au sein de l’université, cela va répondre au problème de la portabilité, et aussi va éviter l’utilisation des solutions existantes sur le marché qui demande un budget très important.

Ainsi nous citons plusieurs solutions précises qui répondent aux problèmes rencontrés dans des différents services au sein de l’université, ces solutions nécessitent la première solution pour qu’ils seront réalisés :

Tableau : Les solutions proposés

|  |  |
| --- | --- |
| La solution | Les exigences de la solution |
| Évaluation de la performance des étudiants | Dans ce cas nous avons des exigences selon deux types des méthodes des études :   * *Cours présentiel :* * *Il faut avoir une caméra de surveillance au sein de chaque classe ou Amphi.* * *Il faut avoir un système qui fait la gestion des emplois du temps et les séances selon chaque classe ou Amphi et aussi selon chaque module ou élément pour savoir le groupe des étudiants, les matières et les professeurs ciblés.* * *Cours à distance :* * *Il faut avoir une plateforme pour la gestion des cours en ligne.*   Un problème de gestion des emplois du temps ne doit pas être apparu car il peut influencer toute la solution surtout dans les cas où un professeur veut remporter une séance ou l’heure d’entrer et sortie n’a pas bien respecté. |
| Détection de la fraude durant les examens | Pour réaliser cette solution il faut avoir :   * *Des caméras installées à chaque classe / Amphi ou bien une caméra portable destinés aux examens pour minimiser les charges.*   *Un système informatique qui fait la gestion des examens et il met seulement la liste des étudiants convoqués à l’examen concernés* |
| Contrôle d’accès aux espaces privés | Dans ce cas il faut avoir :   * *Une caméra de surveillance installé dans les locaux concernés*   *Un système informatique qui gère les responsables et les employés autorisés pour chaque local.*  Un problème de gestion des emplois du temps ne doit pas être apparu car il peut influencer toute la solution surtout dans les cas où un professeur veut remporter une séance ou l’heure d’entrer et sortie n’a pas bien respecté. |
| Gestion des absences | Pour réaliser cette solution nous sommes besoin de :   * *Une tablette ou ordinateur programmable à l’entrée de chaque classe ou Amphi.*   *Un système de gestion des emplois du temps par séance, matière, filière…* |
| Gestion de la bibliothèque | Pour réaliser cette solution nous sommes besoin de :   * *Avoir une tablette programmable à l’entrée et une autre à la sortie.* * *Avoir une machine pour la distribution de livre, et il identifier l’étudiant à chaque opération (prendre-déposer) avec la précision de la marge du délai, cette dernière va aussi aider un étudiant d’avoir l’information à propos la date de disponibilité du livre.*   **Exemple :**  Book vending machines! | Vending machine, Books, Library  Figure : Bibliothèque électronique |

### Planification du projet

La planification du projet sert à identifier et faire l’ordonnancement des tâches que nous devons réaliser dans un cadre temporel, cela va nous permet de bien estimer le cout de réalisation de chaque tâche.

Ensuite le tableau suivant représente l’ensemble des tâches principales pour la réalisation du projet :

***DF (début à fin) :*** *signifie que la t*â*che ne peut pas terminer tant que la précédente n’a pas démarré.*

***FD (fin à début) :*** *signifie que la t*â*che ne peut débuter que lorsque la précédente sera terminée.*

***DD (début à début) :*** *signifie qu’une t*â*che ne peut débuter que lorsque la précédente sera débuter.*

***FF (fin à fin) :*** *signifie que les 2 tâches doivent se terminer en même temps.*

Tableau : Planification du projet

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Tâche | Durée (en jours) | | Dépendances |
| Étape de cadrage et de planification | | | 1. **Étude de projet** | | 7 |  |
| 1. **Réalisation du contexte générale du projet** | | 14 | DF (A) |
| Étape de conception | | | 1. **Étude théorique du projet** | | 14 | FD (B) |
| 1. **Étude et conception de la solution** | | 7 | FD (C) |
| Étape de réalisation du projet | | | 1. **Réalisation de la solution** | | 24 | FD (D) |
| 1. **Mise en œuvre des tests** | | 4 | FD (E) |
| 1. **Intégration et déploiement du solution** | | 4 | FD (F) |
| Étape de clôture | | | 1. **Réalisation du rapport** | | 74 | DD (A, B, C, D, E, F, G) |
| 1. **Réalisation de la présentation** | | 7 | FD (G, H) |
| 1. **Préparation de la soutenance** | | 7 | DD (I) |

Et pour bien visualiser ces tâches et son ordonnancement, nous utiliserons l’outil Gantt destiné pour la gestion du projet. Il contient des affectations des dates pour préciser une tâche doit être commencer et terminer.

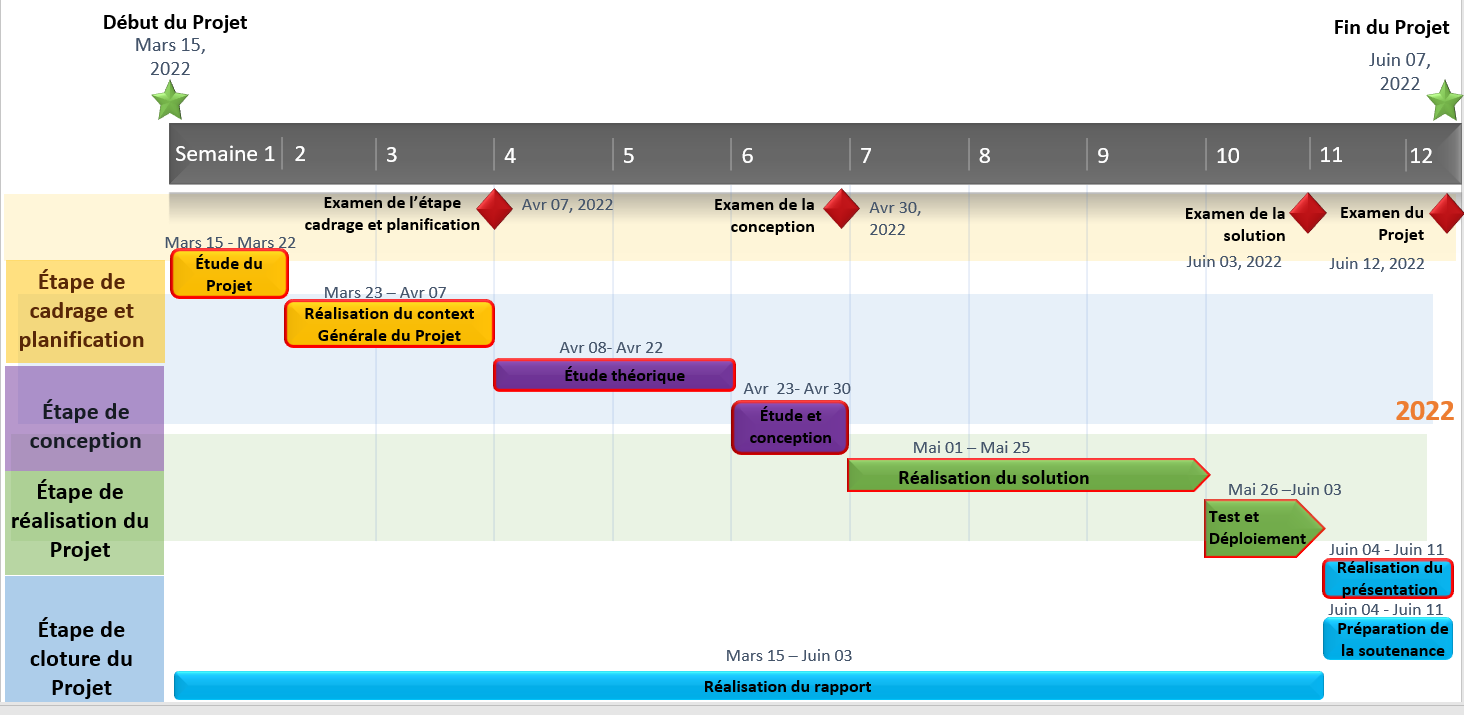


Figure : Diagramme de Gantt

Dans ce diagramme existe des tâches qui sont encadrées par une bordure en rouge, cela signifie que ce sont des tâches critiques que nous ne pouvons pas les retarder durant le cycle de vie du projet, car ils vont affecter tout le déroulement du projet.

## **Conclusion :**

En général dans ce chapitre, nous avons cité les différentes informations et le contexte générale de notre projet de recherche, puis nous avons détecté les problématiques liées à ce sujet en donnant ensuite un ensemble des solutions possibles, et à la fin nous avons mis une planification pour savoir les taches que nous devons les faire pour réussir ce projet et son ordonnancement.

Alors maintenant nous avons une idée claire sur ce que nous devons faire, il nous reste maintenant d’attaquer le sujet de recherche au niveau scientifique et faire une étude théorique concrète, c’est ce que nous découvrirons dans le prochain chapitre.

# CHAPITRE 2 : SYSTÈME DE RECONNAISSANCE FACIALE

## **Introduction**

Dans ce chapitre nous devons introduire les composants qui caractérisent un système de reconnaissance faciale, tout dépendamment aux exigences du contexte ou le système sera appliquer, ainsi les données utilisées au niveau de l’entraînement des modèles de visages, puis les données appliquées durant l’opération système.

## **Type des données utilisés**

Dans un système de reconnaissance faciale nous utiliserons des données précises pour faire les traitements nécessaires, et parmi ces données nous trouverons :

### Image 2d :

C’est une image qui se représente sur un axe de deux dimensions (X, Y), ce qui nous permet d’identifier des positions précis (X, Y) lié aux formes des composants du visage comme les yeux, nez, la mâchoire, la bouche …au sein de l’image, se sont des caractéristiques qui ont un aspect unique pour chaque personne.

(image visage, s.d.)



Figure : Image visage 2d

### Image 3d

C’est une image représentée sur un axe géométrique tridimensionnelles (X, Y, Z). Les traitements de la reconnaissance faciale appliqués aux images 3d peut atteindre une précision plus efficace que celle traité à la base de deux dimensions

(image 3d, s.d.)

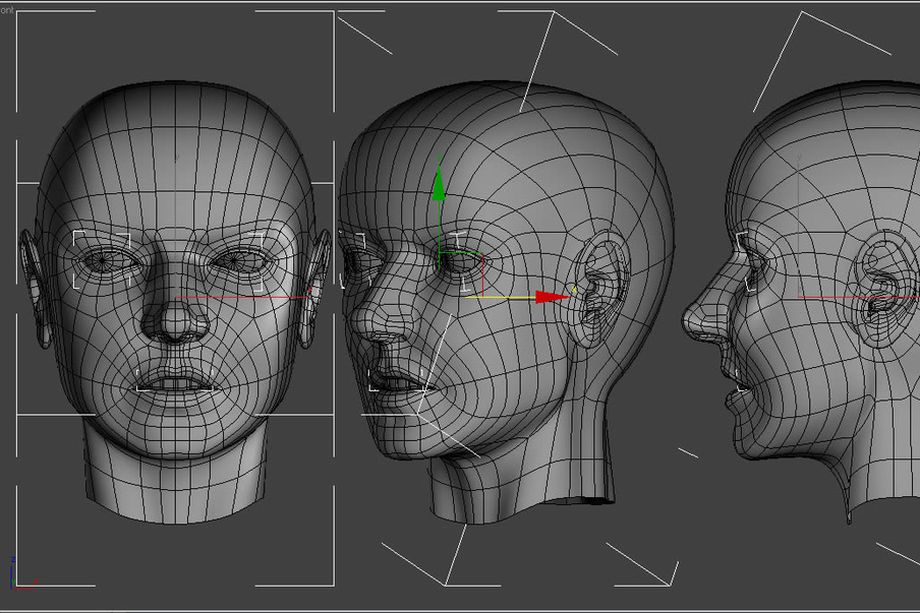


Figure : Image 3d

La principale limitation technologique de la méthode de reconnaissance faciale 3d c’est l’acquisition d’une image 3D qui nécessite généralement une caméra télémétrique, mais par exemple nous pouvons construire un objet 3d à partir plusieurs images capturées de différents angles par un caméra.

### Vidéo

La technologie de la reconnaissance faciale peut être appliqué sur une vidéo pour identifier une ou plusieurs personnes identifiés au sein d’un vidéo et retrouvé à travers des bases des données contenants des photos ou des vidéos dans ce cas nous comparons des vidéos par d’autres vidéos, on se basant sur les caractéristiques de ces figures, et il y a aussi une possibilité de faire l’inverse on identifiant des individus au sein d’une photo à la base des vidéos existants dans notre base de données.

Le défi principal dans ce cas est la qualité et la résolution des images, pour avoir de la souplesse lors de l’application des algorithmes.

(reconnaissance faciale par video, s.d.)

Figure : Identification par vidéo



## **Détection et identification des visage**

La détection du visage est la première étape dans un système de reconnaissance faciale pour l’identification des objets de type visage au sein des images ou des vidéos, alors il doit être réaliser d’une bonne façon, pour cela nous avons réalisé un modèle de détection qui se base sur les algorithmiques de l’apprentissage approfondi (Deep Learning).

### La collection et l’annotation des images

Avoir des images est très important pour entrainer notre modèle, pour cela nous devons commencer par collecter plusieurs images qui présentent des situations auxquelles notre système sera confronté, pour cela il faut avoir plusieurs images et faire la classification, comme l’exemple suivant :



Figure : Exemple des images collectées

Cette figure représente un exemple des images qu’on veut utiliser pendant l’entrainement de notre modèle, dans notre cas nous avons utilisés 90 images qui contient (60 images pour la partie d’entrainement et 16 images pour la partie du test, puis 14 images pour la partie de validation).

Et comme il est remarqué sur la figure il y a plusieurs situations différentes, par exemple : des images qui contient le visage sur des différentes parties de la photo et d’autres images sans visage.

Après le fait de collecter les images, il faut faire des annotations en prenant les images et annoter les visages sur chaque image comme il est mentionné sur la figure suivante :

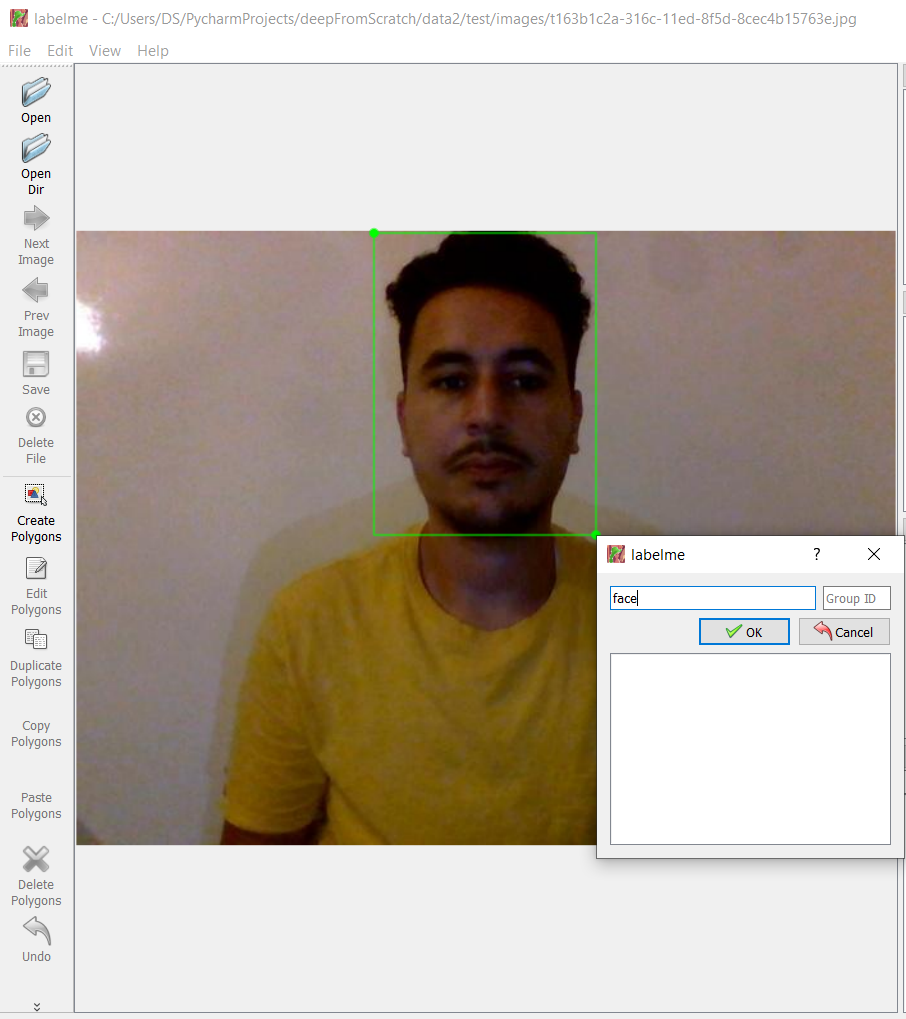


Figure : Exemple d'annotation du visage au sein d'une image collectée.

Cette étape nécessite d’encadrer les visages qui se trouvent dans la partie d’entrainement, puis il faut générer un fichier spécifique pour chaque image qui contient les cordonnées des visages annotés.

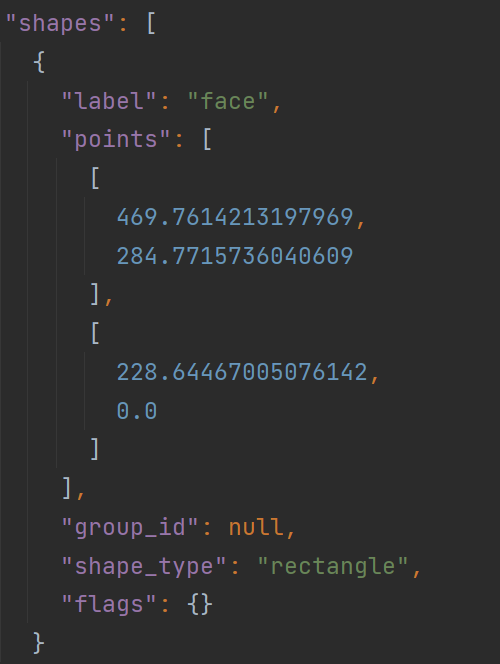


Figure : Exemple des points annotées d'un visage

### L’augmentation des images

L’augmentation des images est une technique très utilisée dans le domaine d’apprentissage approfondi, cette technique permet de modifier l’aspect de l’image sans changer l’aspect de l’image, par exemple : (diminuer la luminosité, faire une rotation, ajuster la taille de l’image …)

#### Pourquoi on utilise l’augmentation des images ?

Il existe plusieurs algorithmes puissants et complexes qui nécessitent d’avoir des données pour les entrainer, et dans le cas des données sous forme des images les techniques d’augmentation peuvent créer des variations d’images qui vont améliorer le processus d’apprentissage du modèle puisqu’il en résulte en plus d’échantillons d’entrainement.

En plus l’application de l’augmentation d’une bonne façon permet d’avoir un modèle plus performant et robuste, et voilà quelques techniques les plus utilisés dans l’augmentation des images :

* **Retournement (Flip)**

Il signifie l’inversion des lignes ou des colonnes de pixels dans le cas d’un flip vertical ou horizontal.

Exemple :



Figure : Exemple du retournement d'image (Flip)

* **Rotation**

Cette technique ne permet pas de préservait la taille de l’image sauf la rotation de 90° dans le cas d’un rectangle.

* **Scale :**

L’image peut être mise à l’échelle vers l’extérieur ou l’intérieur, dans le cas de la mise à l’échelle vers l’extérieur la taille finale de l’image sera plus grande.

* **Crop :**

Cette technique permet de redimensionner une partie de l’image pour qu’il sera de la même taille de toute l’image et par exemple :



Figure : Exemple du Crop d'image

* **Le bruit Gaussien (over fitting ) :**

Cette technique est très utilisée lorsque le modèle tente d’apprendre des caractéristiques de haute fréquence et des patterns que se produisent beaucoup et qui ne sont pas très utile pour le traitement, donc l’ajout d’une quantité intéressante du bruit gaussien permet au modèle de bien améliorer sa capacité d’apprentissage.

Exemple :

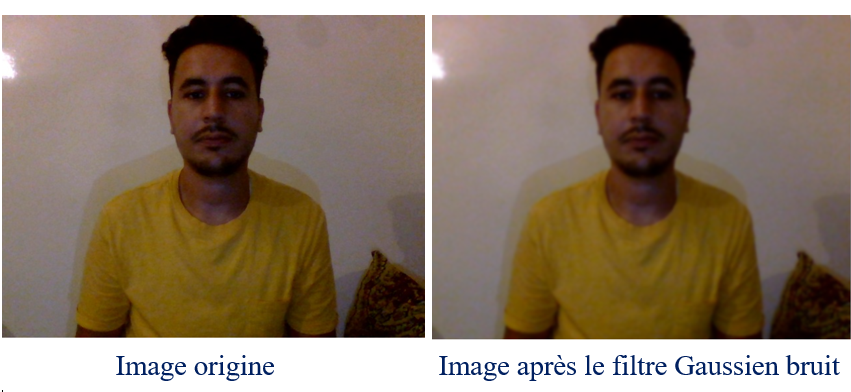


Figure : Le bruit de Gaussien

* **Luminosité :**

Il existe plusieurs techniques pour modifier la luminosité de l’image, soit par l’assombrissement ou bien l’éclaircissement des images, et même avec les deux pour que l’apprentissage du modèle soit sur des images formés sur différents niveaux d’éclairage.

### La réalisation du modèle de détection du visage

Après la collection et la préparation des données maintenant nous devons commencer à entrainer notre modèle de détection en utilisant l’apprentissage approfondi (Deep Learning), alors dans cette partie on va citer les différentes étapes nécessaires pour la construction du modèle et faire une étude fondamentale et pratique en relation avec les concepts et techniques utilisés.

#### Les réseaux de neurones convolutif (CNN)

Un réseau neuronal convolutif se compose d'une couche d'entrée, de couches cachées et d'une couche de sortie. Dans tout réseau neuronal à anticipation, toutes les couches intermédiaires sont dites cachées car leurs entrées et sorties sont masquées par la fonction d'activation et la convolution finale. Dans un réseau neuronal convolutif, les couches cachées comprennent des couches qui effectuent des convolutions. Cela inclut généralement une couche qui effectue un produit scalaire du noyau de convolution avec la matrice d'entrée de la couche. Ce produit est généralement le produit interne de Frobenius et sa fonction d'activation est généralement ReLU. Lorsque le noyau de convolution glisse le long de la matrice d'entrée pour la couche, l'opération de convolution génère une carte de caractéristiques, qui à son tour contribue à l'entrée de la couche suivante. Viennent ensuite d'autres couches telles que les couches de regroupement, les couches entièrement connectées et les couches de normalisation.

*(Wikepédia, s.d.)*

Donc généralement un réseau de neurones est un système composé de neurone qui sont répartis dans plusieurs couches, et il est très souvent utilisé dans les problèmes de statistique et classification.

##### Fonctionnement

Pour extraire les caractéristiques d’après une image, il utilise des opérations de filtrage par convolution, dans lequel la première couche filtre l’image en se basant sur plusieurs noyaux de convolution et renvoie des mapping de caractéristique (Feature maps) qui sont ensuite normalisés avec une fonction d’activation.

Alors ce dernier processus peut se répéter pour plusieurs fois, donc on filtre les (features maps) obtenues avec nouveaux noyeau puis nous faisons autrefois la normalisation et le redimensionnement en se basant sur une fonction d’activation. Pour obtenir à la fin des valeurs du dernier Feature Maps qui seront concaténés dans un vecteur, ce vecteur sera la sortie du premier bloc et l’entrée du deuxième.

Et finalement le deuxième bloc se retrouve à la fin de tous les réseaux de neurones, les valeurs du vecteur en entrée sont transformées à l’aide de plusieurs fonctions d’activation et combinaison linéaire pour renvoyer à la fin un vecteur qui contient autant d’éléments par exemple :

L’élément x appartient à la classe A et chaque élément à une valeur entre 0 et 1 et la somme globale vaut 1, ces probabilités sont calculés par la dernière couche de se bloc et le réseau en générale, et il utilise comme fonction d’activation : une fonction logistique (la classification binaire) et la fonction softmax ( la classification multi-classe).

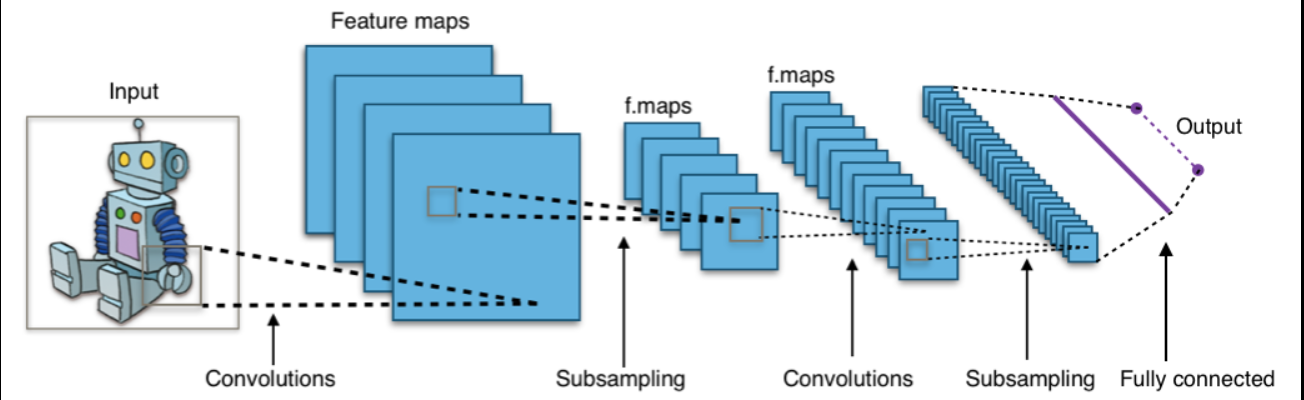


Figure : CNN architecture

*(Wikepedia-CNN-Architecture, s.d.)*

Maintenant on va essayer de présenter le rôle de chaque couche mentionnée dans la figure 16.

* **La couche de Convolution :**

Le rôle principal de cette première couche c’est de repérer la présence d’un ensemble de caractéristiques (features) dans les images reçues en entrées, cela permet de calculer le produit de convolution entre la caractéristique (feature) et chaque portion de l’image, donc un feature c’est un filtre dans ce contexte.

Alors pour chaque paire (image, feature) on obtient (Feature Maps) qui nous indique ou se situe les caractéristiques dans l’image.

* **La couche de pooling et subsampling :**

Cette couche revoit plusieurs features Map et il applique pour chacun l’opération de Pooling.

L’objectif principale de l’opération de Pooling c’est de réduire la taille des images en gardant leurs caractéristiques imporants. Cela permet aussi de réduire le nombre de paramètre et de calculs au sein du réseau.

* **La couche de correction Relu ( Rectified Linear Unit ):**

Relu c’est la fonction réelle non-linéaire représenté par l’expression suivante :

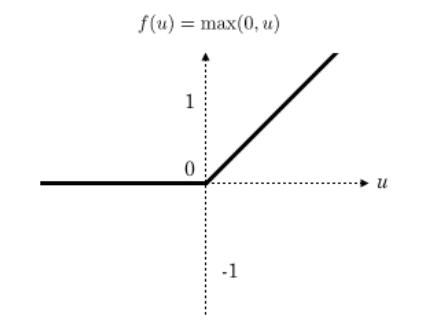


Figure : Allure de la fonction ReLU

Donc cette couche de correction ReLU permet de remplacer chaque valeur négative reçues en entrées par des zéros. Elle joue le rôle d’une fonction d’activation.

* **La couche Fully Connected :**

Ce type de couche reçoit comme entrée un vecteur et il retourne en sortie un vecteur. Pour cela, elle applique une combinaison linéaire puis éventuellement une fonction d'activation aux valeurs reçues en entrée.

Cette dernière couche *fully-connected* permet de classifier l'image en entrée du réseau : elle renvoie un vecteur de taille NN, où NN est le nombre de classes dans notre problème de classification d'images. Chaque élément du vecteur indique la probabilité pour l'image en entrée d'appartenir à une classe.

**Par exemple :**

Dans le cas de notre problème on veut distinguer entre une image contient un visage ou non, le vecteur final sera de taille 2 : le premier élément (respectivement, le deuxième) donne la probabilité d'appartenir à la classe "visage" (respectivement "vide"). Ainsi, le vecteur [0.9 0.1] signifie que l'image a 90% de chances de représenter un visage.

Chaque valeur du tableau en entrée "vote" en faveur d'une classe. Les votes n'ont pas tous la même importance : la couche leur accorde des poids qui dépendent de l'élément du tableau et de la classe.

Pour calculer les probabilités, cette couche multiplie chaque élément en entrée par un poids, fait la somme, puis applique une fonction d’activation (logistique N=2, softmax si N>2)

Ce traitement revient à multiplier le vecteur en entrée par la matrice contenant les poids. Le fait que chaque valeur en entrée soit connectée avec toutes les valeurs en sortie explique le terme fully-connected.

**Explication :**

Dans le cas de N=2, ça signifie que nous avons deux éléments, alors pour ce cas on va avoir deux valeurs 0 et 1, pour cela on utilise la fonction logistique qui a comme expression mathématique :

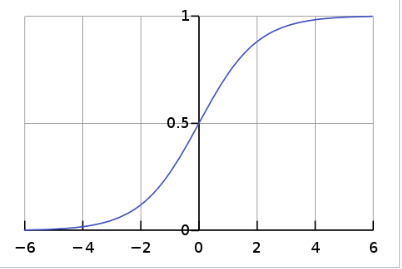


Figure : Courbe de fonction logistique

Dans la figure 18 nous avons la courbe d’une fonction logistique dans le cas où :

Et pour le cas de N>2 nous devons utiliser une fonction d’activation Softmax qui prend un vecteur Z avec K nombres réels et le normaliser en une distribution de probabilité constituée de K probabilités proportionnelles aux exponentielles des nombres d'entrée. Autrement dit, avant d'appliquer Softmax, certaines composantes vectorielles pourraient être négatives ou supérieures à 1 et leurs sommes ne nous peut pas être équivalent à 1, mais après l’application de softmax, chaque composant sera dans l'intervalle (0,1), et la somme de tous les composants sera 1, de sorte qu'ils puissent être interprétés comme des probabilités. De plus, les composants d'entrée plus grands correspondront à des probabilités plus grandes, et voilà la formule mathématique de Softmax :

#### L’architecture des réseaux de neurones VGG16

VGG16 est un type de CNN qui est considéré comme l'un des meilleurs modèles de vision par ordinateur à ce jour. Les créateurs de ce modèle ont évalué les réseaux et augmenté la profondeur en utilisant une architecture avec de très petits filtres de convolution (3 × 3), ce qui a montré une amélioration significative par rapport aux configurations de l'art antérieur. Ils ont poussés la profondeur à 16-19 couches de poids, ce qui en fait environ - 138 paramètres entraînables.

Il est capable de faire la classification de 1000 images en 1000 catégories avec une précision de 92.7%.

* **Architecture du VGG16 :**

*(Image architecture vgg16, s.d.)*

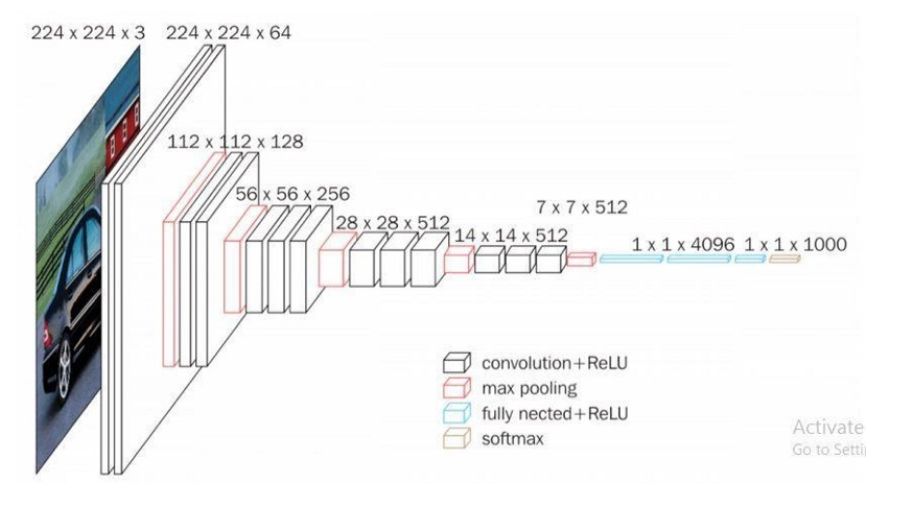


Figure : Architecture de VGG16

*(Image les couches de vgg16, s.d.)*

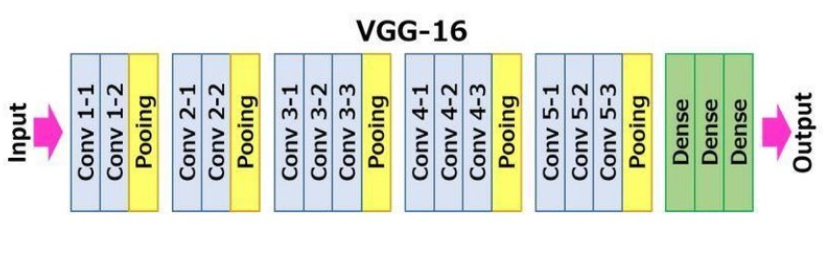


Figure : Les couches de VGG16

* Le 16 dans VGG16 fait référence à 16 couches qui ont des poids. Dans VGG16, il y a 13 couches convolutives, contient 5 couches Max Pooling et 3 couches Dense qui totalisent jusqu'à 21 couches, mais il n'a que 16 couches de poids, c'est-à-dire une couche de paramètres apprenables.
* VGG16 prend la taille du tenseur d'entrée comme 224, 244 avec 3 canaux RVB
* La chose la plus importante et unique à propos de VGG16 est qu'au lieu d'avoir un grand nombre d'hyper-paramètres, ils se sont concentrés sur des couches de convolution de filtre 3x3 avec la foulée 1 et ont toujours utilisé le même rembourrage et la même couche Maxpool du filtre 2x2 de la foulée 2.
* Les couches de convolution et de MaxPool sont disposées de manière cohérente dans toute l'architecture
* Conv-1 Layer contient 64 filtres, Conv-2 contient 128 filtres, Conv-3 contient 256 filtres, Conv 4 et Conv 5 contiennent 512 filtres.
* Trois couches entièrement connectées (FC) suivent une pile de couches de convolution : les deux premières ont 4096 canaux chacune, la troisième effectue une classification ILSVRC à 1000 voies et contient donc 1000 canaux (un pour chaque classe). La couche finale est la couche soft-max.

**Une classification ILSVRC :**

*ImageNet Large Scale Visual Recognition**Challenge* (ILSVRC) consiste à classer les images obtenues à partir de Flickr et d'autres moteurs de recherche sur 1000 catégories d'objets possibles. Cette opération est une référence standard pour l'apprentissage approfondi.

#### Définir la fonction de perte et les optimiseurs

Une fonction de perte est une fonction qui compare les valeurs de sortie cible et prédite ; mesure à quel point le réseau de neurones modélise les données de formation. Lors de la formation, nous visons à minimiser cette perte entre les sorties prévues et cibles.

Pour cela nous allons utiliser un optimiseur pour modifier les poids de chaque époque minimiser la fonction de perte et réduire la perte globale et améliorer la précision du modèle.

* **Époque (epoch) ?**

Une époque signifie entraîner le réseau de neurones avec toutes les données d'entraînement pour un cycle. Dans une époque, nous utilisons toutes les données exactement une fois. Une passe avant et une passe arrière ensemble comptent pour une seule passe :

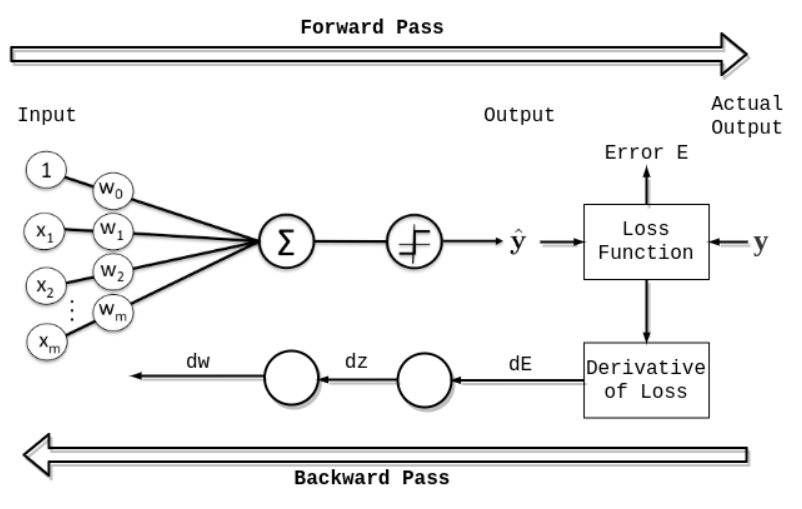
(Image epoch, s.d.; Image mécanisme d'une époque, s.d.)**

Figure : Mécanisme d'une époque

Une époque est composée d'un ou plusieurs lots, où nous utilisons une partie de l'ensemble de données pour entraîner le réseau de neurones. Nous appelons le passage à travers les exemples de formation dans un lot une itération.

Une époque est parfois mélangée à une itération. Pour clarifier les concepts, considérons un exemple simple où nous avons 1000 points de données comme présenté dans la figure ci-dessous :

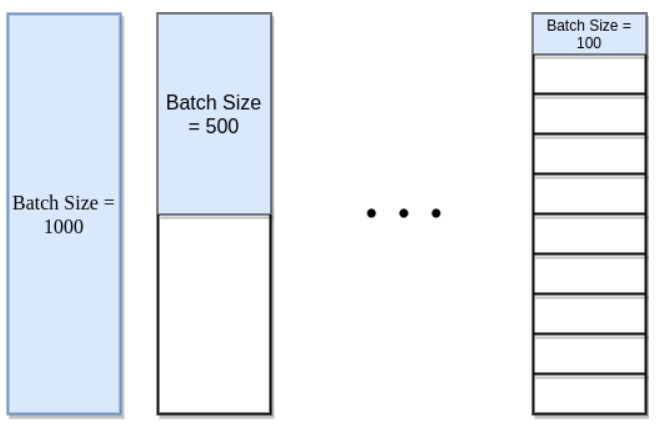
(Image epoch, s.d.; Image itération par époque, s.d.)**

Figure : Exemple nombre d'itération par époque selon la taille du lot (batch)

Si la taille du lot est de 1000, nous pouvons compléter une époque avec une seule itération. De même, si la taille du lot est de 500, une époque prend deux itérations. Ainsi, si la taille du lot est de 100, une époque prend 10 itérations pour se terminer. Simplement, pour chaque époque, le nombre d'itérations requis multiplié par la taille du lot donne le nombre de points de données.

* **Adam Optimiseur :**

Dans notre cas nous avons utilisé Adam optimiseur comme algorithme d’optimisation, puisqu’il permet de remplacer de la procédure classique de descente de gradient stochastique ou bien l’algorithme d’optimisation différentiable pour mettre à jour itérativement les pondérations du réseau en fonction des données d'apprentissage, il est très populaire et le plus utilisé à cause de ces bonnes résultats, et voilà un exemple dans la figure de comparaison entre Adam algorithme et les autres algorithmes d’optimisation

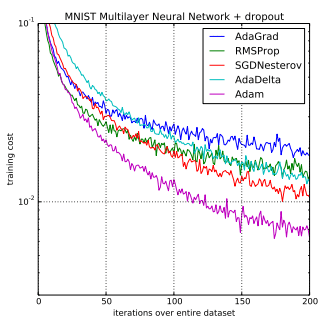
*(Comparaison de l'algorithme Adam et les autres algorithmes d'optimisation, s.d.)*

Figure : Comparaison d'Adam avec d'autres algorithmes d'optimisation

Alors après que nous défini de quoi s’agit une époque et algorithme d’optimisation Adam, maintenant la fonction de perte se compose en deux parties : (localization loss) pour la prédiction du décalage de la boîte englobante et la perte de classification (classification loss) pour les probabilités de classe conditionnelles.

### Réalisation du modèle et résultat

Durant la partie précédente, nous avons cité les algorithmes d’apprentissage approfondi sur lesquels nous avons basé pendant la réalisation de notre Modèle de détection des visages au sein d’une image.

Donc après la collection des images comme on a vu déjà et la classification et l’augmentation des données, on a passé vers l’entrainement de notre modèle et ensuite nous avons défini notre fonction de perte (localization loss) et (classification loss) en ajoutant l’optimiseur Adam, et lors de la construction de notre modèle nous avons utilisé 6 époques, et nous avons obtenues le résultat suivant dans la dernière époque au niveau de la fonction de perte :

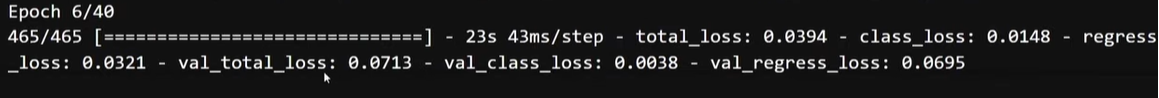


Figure : Résultat de (Localization Loss) et (Classification Loss) dans la dernière époque

Et comme il est mentionné dans la figure 24 nous avons obtenu un totale de perte pour les données d’entrainements équivalent à 0.0394 qui est un très bon résultat pour les données qui existe déjà