**UNIVERSITE KASDI MERBAH – OUARGLA-**

Faculté des Nouvelles Technologies de l'Information et de la Communication

Département d'informatique et des technologies de l'information

**Mémoire**

**MASTER ACADEMIQUE**

Domaine : Mathématique et Informatique

Filière : Informatique

Spécialité : Informatique Industrielle

Présenté par :

**Yahiaoui Nassima**

**Thème** :

# Un système basé sur la vision par ordinateur pour la prévention des virus respiratoires (Cas du Covid-19)

Soutenu publiquement
Le : 17/06/2021

Gérer par :

- Mr AYADI Oussama Encadreur UKM Ouargla

**Année Universitaire : 2020/2021**

# Dédicace

A

*Mon père et ma mère et Mon mari*

*Mes frère et ma sœurs*

*Mes très proche amies*

*Toute ma famille*

*Tous…*

*Nassima*

# Remerciement

*Avant tous, nous remercions ALLAH, pour nous avoir données*

*Le pouvoir pour réaliser ce travail.*

*Un grand merci à mon encadreur* ***Dr.oussama ayadi*** *qui a fourni des*

*efforts énormes par ses conseils, ses orientations et ses informations.*

*Nous tenons à remercier les membres du jury et tous les enseignants*

*dudépartement MATHEMATIQUE ET INFORMATIQUE.*

*Enfin, nos plus chaleureux remerciements pour toute personne ayant*

*participé de près ou de loin à la réalisation de ce modeste travail.*

# *Sommaire :*

**INTRODUCTION GENERALE**

1. Contexte et Problématique

2. Objective

3. Organisation du mémoire

**CHAPITRE 1 : LA VISION PAR ORDINATEUR**

Introduction

1.1 La vision par ordinateur……..……..…..01

1.2 Les application de la vision par ordinateur ……...……..……..……02

1.3 La deference entre La reconnaissance d'image et la Machine Vision …..…….……...…03

Conclusion

**CHAPITRE 2 : METHODES DE CLASSIFICATION DES IMAGES**

2.1 Système d'IA de surveillance à distance sociale et de détection de masque facial pour Covid-19 …….…….……05

2.2 Explication détaillée …….…..……..…….....……..….05

2.3 Détection de personne ……..05

2.4 Détection de visage …….…….…….....06

2.6 Surveillance à distance sociale à l'aide de DBSCAN Algo…….....07

2.7 Pipeline d'architecture……..……..…….....08

2.8 Résultats…….…….…….....09

2.9Avantages et inconvénients …….....09

2.10 Portée future ……...…….....09

**CHAPITRE 3 : Les methodes de vision par ordinateur**

Introduction

3.1 BASES DE LA RECONNAISSANCE DE VISAGES ……..……12

3.2 Types de représentations du visage……..……12

3.3 Architecture de base d'un système de reconnaissance de visages ……..…….12

3.4 MODULES D'UN SYSTÈME DE FRiVS ET MÉTHODES CONSIDÉRÉES ……..13

3.4.1 Détection de visages …….…….….14

3.4.2Extraction de caractéristiques …….……14

3.4.3 Méthodes par projection …….14

3.4.4 Méthodes de vision ……..……15

3.4.5 L'analyse d'une image ……....17

**CHAPITRE 4 : Implimentation et Evaluation**

Introduction

4.2 Conception du système de detection de masque ……..18

4.3 Choix de l'environnement (langage) ……...…….19

4.4 l'interface de l'environnement PYTHON 3.9 ……..20

4.5Les phases de système de detection ……..……20

4.6 Le Code source de programme de detection de masque ……21

4.7 L'interface de l'application ……..……22

4.7.1 On a choisit des photos ……..……23

4.7.3 l'application destingue le resultat ……24

conclusion

**CONCLUSION GENERALE**

**BIBLIOGRAPHIE**

# Introduction Générale

**Introduction :**

**Contexte et Problématique :**

La maladie à coronavirus 2019 ou Covid-19 est une maladie de type zoonose virale causée par SARS-CoV-2. Les symptômes les plus fréquents sont la fièvre, la toux, la fatigue et la gêne respiratoire. Dans les formes les plus graves, notamment chez les personnes plus fragiles du fait de leur âge ou en cas de comorbidités.

De nos jours, avec la mécanisation de tous les secteurs de l'économie et surtout de la modernisation Identifier une personne à partir de son visage est une tâche aisée pour les humains.

En est-il de même pour une machine ? Ceci définit la problématique de la reconnaissance automatique de visages, qui a engendré un grand nombre de travaux de recherche au cours des dernières années.

**1.1- Objective :**

Parmi les solutions informatiques qu'on va l'adopter est la **vision par ordinateur** (aussi appelée vision artificielle ou vision numérique), c'est une branche de l'intelligence artificielle dont le principal but est de permettre à une machine d'analyser, traiter et comprendre une ou plusieurs images prises par un système d'acquisition (par exemple : caméras, etc.).

C'est une approche qui consiste à tenter d'imiter la vision humaine par le truchement de composants électroniques. Cette manière de procéder peut être perçue comme un traitement des données visuelles par le biais de modèles fondés sur la géométrie, la physique, la biologie, les statistiques et la théorie d'apprentissage (les applications de reconnaissance).

Notre objective est d'implémenter une application de reconnaissance en utilisant les principes de la vision par ordinateur et Machine Learning qu'à partir d'une base d'images.

### 1.3- Organisation du mémoire :

Notre travail s'inscrit au domaine de la detection des images, pour réaliser notre application, nous avons organisé le mémoire en quatre chapitres,incluant cette introduction générale et se termine par une conclusion générale, comme suit :

- **Le premier chapitre :**il présente quelque notion de la vision par ordinateur base sur l'image numérique et quelques descripteurs(Locaux et globaux) de l'image.
- **Dans le second chapitre :**presente le notion de reconnaissance d'image et la deference entre ce dernier notion et la vision par ordinnateur
- **Le troisième chapitre :** présente les méthodes choisies ainsi que les différentes étapes qui leurs constituent.
- **Dans le quatrieme chapitre**, nous le avons consacré à la conception et la réalisation de notre application en utilisant python et deep learning, pour faciliter l'implémentation de la detection des images.

En conclusion nous présentons les points de ce travail et quelques perspectives.

# La Vision par Ordinateur

# Introduction :

La vision par ordinateur est un domaine scientifique à part entière, mais une partie de ses applications est en lien étroit avec l'apprentissage statistique. Elle est d'ailleurs parfois nommée reconnaissance des formes.

En vision par ordinateur, l'extraction de caractéristiques visuelles (ou visual features extraction en anglais) consiste en des transformations mathématiques calculées sur les pixels d'une image numérique.

L'objectif principal de la vision par ordinateur est le développement d'un systèmeinformatique capable de doter l'ordinateur du sens de vision.

## 1.1 La vision par ordinateur :

est un domaine scientifique interdisciplinaire qui traite de la façon dont les ordinateurs peuvent acquérir une compréhension de haut niveau à partir d'images ou de vidéos numériques. Du point de vue de l'ingénierie, il cherche à comprendre et à automatiser les tâches que le système visuel humain peut effectuer.

Vue d'artiste d'un Rover automatique explorant la surface de Mars. Il est équipé sur son sommet de deux caméras vidéo lui conférant une vision stéréoscopique.

Les tâches de vision par ordinateur comprennent des procédés pour acquérir, traiter, analyser et « comprendre » des images numériques, et extraire des données afin de produire des informations numériques ou symboliques,

Dans ce contexte, la compréhension signifie la transformation d'images visuelles (l'entrée de la rétine) en descriptions du monde qui ont un sens pour les processus de pensée et peuvent susciter une action appropriée. Cette compréhension de l'image peut être vue comme l'acquisition d'informations symboliques à partir de données d'image à l'aide de modèles construits à l'aide de la géométrie, de la physique, des statistiques et de la théorie de l'apprentissage

La discipline scientifique de la vision par ordinateur s'intéresse à la théorie des systèmes artificiels qui extraient des informations à partir d'images. Les données d'image peuvent prendre de nombreuses formes, telles que des séquences vidéo, des vues de plusieurs caméras, des données multidimensionnelles à partir d'un scanner 3D ou d'un appareil de numérisation médical. La discipline technologique de la vision par ordinateur cherche à appliquer les modèles théoriques développés à la construction de systèmes de vision par ordinateur.

Les sous-domaines de la vision par ordinateur comprennent la détection d'événements, le suivi vidéo, la reconnaissance d'objets, l'apprentissage, l'indexation, l'estimation de mouvement, la modélisation de scènes 3D et la restauration d'image

La Computer Vision est une branche de l'Intelligence Artificielle qui a pour objectif de permettre aux ordinateurs d'interpréter des données visuelles (photos ou vidéo) afin d'en extraire des informations.

Les données visuelles sont issues de différents capteurs tels que des caméras de vidéosurveillance ou le smartphone de techniciens de terrain. La Computer Vision est elle-même constituée de plusieurs branches, dont les deux principales sont la reconnaissance d'image et la Machine Vision. La première branche, dans laquelle évolue Deepomatic, permet de mettre en place plusieurs pratiques comme la classification, le tagging, la détection, la segmentation ou encore l'OCR (Optical character recognition), la reconnaissance optique de caractères.

Ces tâches sont des algorithmes dits "d'apprentissage supervisé" : ils sont "entraînés" à partir d'exemples d'images annotées par les concepts qu'il faudrait pouvoir reconnaître, pour ensuite être capables de prédire les concepts d'intérêt dans des images que l'algorithme n'a encore jamais rencontrés.

La classification consiste à reconnaître à quelle catégorie appartient une image parmi un ensemble de catégories prédéterminées. Toutes les données auront obligatoirement une seule classe. Le tagging (ou classification multi-classes) est une technique similaire à la classification dans laquelle aucune, une, ou plusieurs classes peuvent être assignées à une image.

La détection est une technique de reconnaissance d'image utilisée afin de localiser un ou plusieurs objets d'intérêt dans une image ou une vidéo. Elle permet de délimiter les objets reconnus par un rectangle sur l'image (appelé bounding box en anglais, une "boîte englobante"). La segmentation, quant à elle, est une version plus précise de la détection. Elle permet de détecter au pixel près des éléments de l'image et de cartographier très précisément chaque objet dans une image.

Enfin, l'OCR est une technique de reconnaissance d'image visant à reconnaître et traduire en fichier texte des données visuelles de textes imprimés ou dactylographiés.

## 1.2 Les application de la vision par ordinateur :

Des exemples d'applications de la vision par ordinateur comprennent des systèmes pour:

- Inspection automatique, par exemple dans les applications de fabrication;

- Aider les humains dans les tâches d'identification, par exemple, un système d'identification des espèces19;

- Contrôle des processus, par exemple, un robot industriel;

- Détection d'événements, par exemple pour la surveillance visuelle ou le comptage de personnes, par exemple dans l'industrie de la restauration; - Interaction, par exemple en tant qu'entrée d'un dispositif d'interaction ordinateur-homme;

- Modélisation d'objets ou d'environnements, par exemple, analyse d'images médicales ou modélisation topographique;

- Navigation, par exemple par un véhicule autonome ou un robot mobile

- Organisation des informations, par exemple pour l'indexation de bases de données d'images et de séquences d'images.

L'apprentissage des formes 3D a été une tâche difficile en vision par ordinateur.

Les progrès récents de l'apprentissage en profondeur ont permis aux chercheurs de créer des modèles capables de générer et de reconstruire des formes 3D à partir de cartes de profondeur ou de silhouettes à une ou plusieurs vues de manière transparente et efficace.

1.3 La deference entre La reconnaissance d'image et la Machine Vision :

La reconnaissance d'image et la Machine Vision, deux techniques de Computer Vision avec leurs avantages et inconvénients :

La *Computer Vision* et la reconnaissance d'images sont deux termes souvent associés et même confondus, alors qu'il existe bien une différence (subtile, mais existante !). Nous avons donc décidé de vous expliquer la nuance entre ces deux notions.

**La vision industrielle ou Machine Vision :** est une technique différente de la reconnaissance d'image, mais qui reste issue de la *Computer Vision.* Il s'agit d'une technique utilisée principalement dans le secteur industriel, pour la standardisation de la découpe du métal par exemple. Cette pratique utilise l'identification de points d'intérêts définis à l'avance et de la géométrie dans un algorithme complexe afin de déterminer la forme de l'objet à identifier.

La différence fondamentale entre les techniques de reconnaissance d'image et de vision industrielle est l'utilisation du Machine Learning pour la première, tandis que la seconde ne l'utilise pas Le Machine Learning permet à un ordinateur "d'apprendre" à partir de données, c'est-à-dire de résoudre des tâches sans avoir été explicitement programmé pour le faire.

## Commençons par la vision industrielle. Elle possède de nombreux avantages :

Contrairement à la reconnaissance d'image, son système n'a pas besoin d'être entraîné. La solution peut donc être mise en place sans collecte de données d'entraînement au préalable.

La solution utilise des principes géométriques. Elle peut donc s'adapter à de nombreux sujets.

Le fait d'utiliser des principes géométriques rend la solution très fiable, car ce sont des principes immuables.

Cependant, cette technique présente aussi des inconvénients. Tout d'abord, il s'agit d'un système compliqué à mettre en place. En effet, il nécessite la création d'algorithmes géométriques complexes dont l'élaboration ne peut se faire sans un expert. De plus, les principes immuables de la géométrie et la rigidité des algorithmes rend cette technique peu adaptable. Pour chaque évolution du sujet à analyser, un nouvel algorithme doit être créé.

La reconnaissance d'image, utilisant le Machine Learning donc, possède des avantages très différents :

Cette solution est très flexible et peut donc servir dans tout type de secteur. Les réseaux de neurones s'adaptent rapidement aux changements.

Il n'est pas nécessaire d'être un expert pour bénéficier de la reconnaissance d'image, il suffit de collecter des images ou des vidéos qui serviront d'exemples de situations à reconnaître. Cette solution permet donc d'être indépendant tout au long de sa mise en place.

Il s'agit d'une technique très fiable si les réseaux de neurones sont bien entrainés, et les données d'entrainement bien annotées.

Cependant, comme pour la Machine Vision, la reconnaissance d'image a aussi quelques inconvénients. En effet, afin d'être efficace, elle nécessite une moyenne de 1000 images par type d'objet à reconnaître pour être entraînée correctement. Cela veut dire que les entreprises souhaitant mettre en place une solution de reconnaissance d'image doivent déjà avoir en leur possession un gros jeu de données dès le début du projet ou à minima un plan d'action pour les collecter. De manière générale, la mise en place d'une telle solution nécessite un important travail en amont de sa mise en production, mais permet ensuite de s'adapter facilement aux évolutions des conditions de production, contrairement à la vision industrielle.

## Conclusion :

La vision par ordinateur est la construction explicite des descriptions significatives d'objets physiques à partir d'images. C'est un système pour lequel on produit des descriptions de l'environnement perçu à partir de l'analyse d'images. Ces descriptions doivent contenir les aspects des objets observés qui sont utiles pour accomplir les tâches auxquelles est dédiée la machine de vision.

# Système d'IA de surveillance à distance sociale et de détection de masque facial pour Covid-19

## 2.1 Système d'IA de surveillance à distance sociale et de détection de masque facial pour Covid-19 :

Ce projet propose un système d'IA basé sur la vision par ordinateur pour vérifier si la distanciation sociale est maintenue dans un endroit surpeuplé ou à n'importe quel endroit (par exemple: marché ou lieu de travail) couplé à un système de détection de masque pour suivre les personnes qui portent des masques. Cette solution peut être utilisée dans les caméras de vidéosurveillance et autres systèmes de vidéosurveillance. Alors que les données; tels que le nombre de personnes à proximité, le nombre de personnes violant la distanciation sociale et ne portant pas de masques faciaux ; a été utilisé pour l'analyse.

## 2.2 Explication détaillée :

Notre conformité à l'IA consiste en une surveillance de la distanciation sociale de pointe couplée à un système de détection de masque pour vérifier si les normes sont respectées ou non. Ce projet vise à surveiller les personnes qui violent la distanciation sociale sur des séquences vidéo provenant de caméras de vidéosurveillance. Utilise YOLOv3-spp pour détecter les personnes ainsi qu'un outil d'analyse de la distance sociale simulé avec un facteur de profondeur 3D basé sur la position et l'orientation de la caméra pour reconnaître les intrus potentiels. Le détecteur de visage à double injection utilisé pour détecter les visages et un modèle de classificateur de masque facial (ResNet50) sont formés et déployés pour identifier les personnes ne portant pas de masque facial.

## 2.3 Détection de personne :

La première étape consiste à capturer le flux vidéo en direct des caméras de vidéosurveillance existantes pour détecter des personnes individuelles dans chaque image à l'aide du modèle yolo v3.

YOLO (You Only Look Once) est un système de détection d'objets en temps réel à la pointe de la technologie. C'est un modèle spp basé sur la version 3 (pré-entraîné sur le jeu de données COCO), avec une résolution de 608x608 utilisé dans ce projet pour obtenir les cadres de délimitation des personnes individuelles dans une image vidéo. Pour obtenir une vitesse de traitement plus rapide, une résolution de 416x416 ou 320x320 peut être utilisée. YOLOv3-tiny peut également être utilisé pour l'optimisation de la vitesse. Cependant, cela entraînera une diminution de la précision de détection. La couche de sortie a été extraite du modèle yolo v3 qui convient à la détection de personnes. Chaque personne est détectée si le score de confiance est > 0,5.

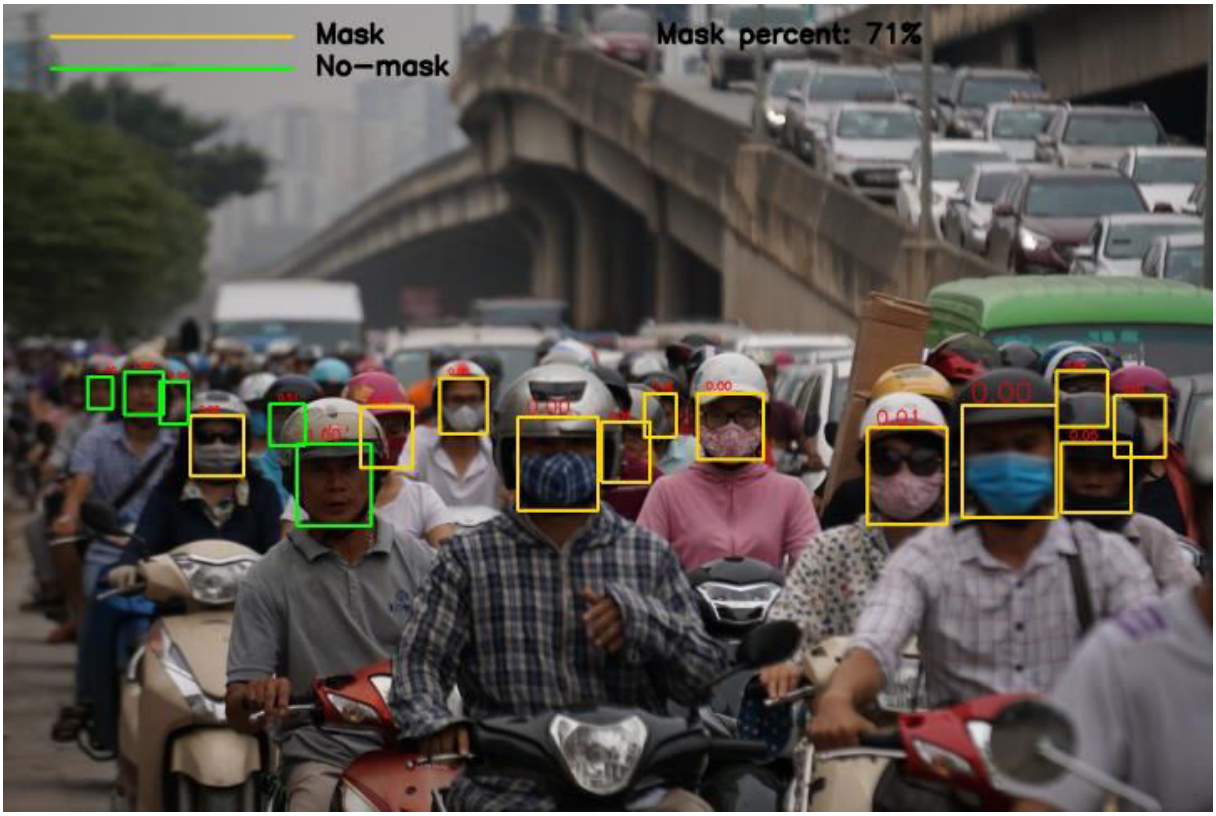
## 2.4 Détection de visage :

Le détecteur de visage à double injection (DSFD) est utilisé tout au long du projet pour détecter les visages. Les détecteurs de visage courants tels que les Haar-Cascades ou le MTCNN ne sont pas efficaces dans ce cas d'utilisation particulier car ils ne sont pas capables de détecter les visages couverts ou à faible résolution. DSFD est également bon pour détecter les visages dans une large gamme d'orientations. Il est un peu lourd sur le pipeline, mais produit des résultats précis. Les coordonnées du visage détecté sont ensuite transmises au modèle de classification de masque.

## 2.5 Classificateur de masque facial :

Un modèle ResNet50 légèrement modifié (avec des couches de base pré-entraînées sur imagenet) est utilisé pour classer si un visage est correctement masqué ou non. La combinaison de certaines couches AveragePooling2D et Dense (avec abandon) se terminant par un classificateur Sigmoïde ou Softmax est ajoutée au-dessus des couches de base. Différentes architectures peuvent être utilisées à cette fin, mais les architectures complexes doivent être évitées pour réduire le surapprentissage. Le modèle doit être entraîné sur des tonnes de données pertinentes avant de pouvoir l'appliquer en temps réel et de s'attendre à ce qu'il fonctionne. Il a besoin de beaucoup de puissance de calcul et je veux dire beaucoup! Nous pouvons essayer nos modèles entraînés sur un petit ensemble de données dans nos machines locales, mais cela ne produirait pas les résultats souhaités. J'ai donc utilisé des modèles open source pré-entraînés pour le moment. Nous utilisons donc le modèle formé par l'équipe de Thang Phamdans ce but. Il s'agit essentiellement d'un modèle ResNet50 avec un sommet modifié.

Un modèle ResNet50 légèrement modifié (avec des couches de base pré-entraînées sur imagenet) est utilisé pour classer si un visage détecté est masqué correctement ou non. Une combinaison de quelques couches de mise en commun moyenne 2D et dense (avec abandon) se terminant par un classificateur Sigmoïde ou Softmax est ajoutée au-dessus des couches de base. Si le score de confiance est inférieur à 0,5, cela signifie que le masque facial est présent alors que si le score de confiance est supérieur à 0,5, cela signifie que le masque facial n'est pas présent. Nous avons observé que le déploiement du réseau pré- entraîné ResNet50 nous a apporté un modèle de haute précision ainsi qu'un temps de formation rapide.




## 2.6 Surveillance à distance sociale à l'aide de DBSCAN Algo :

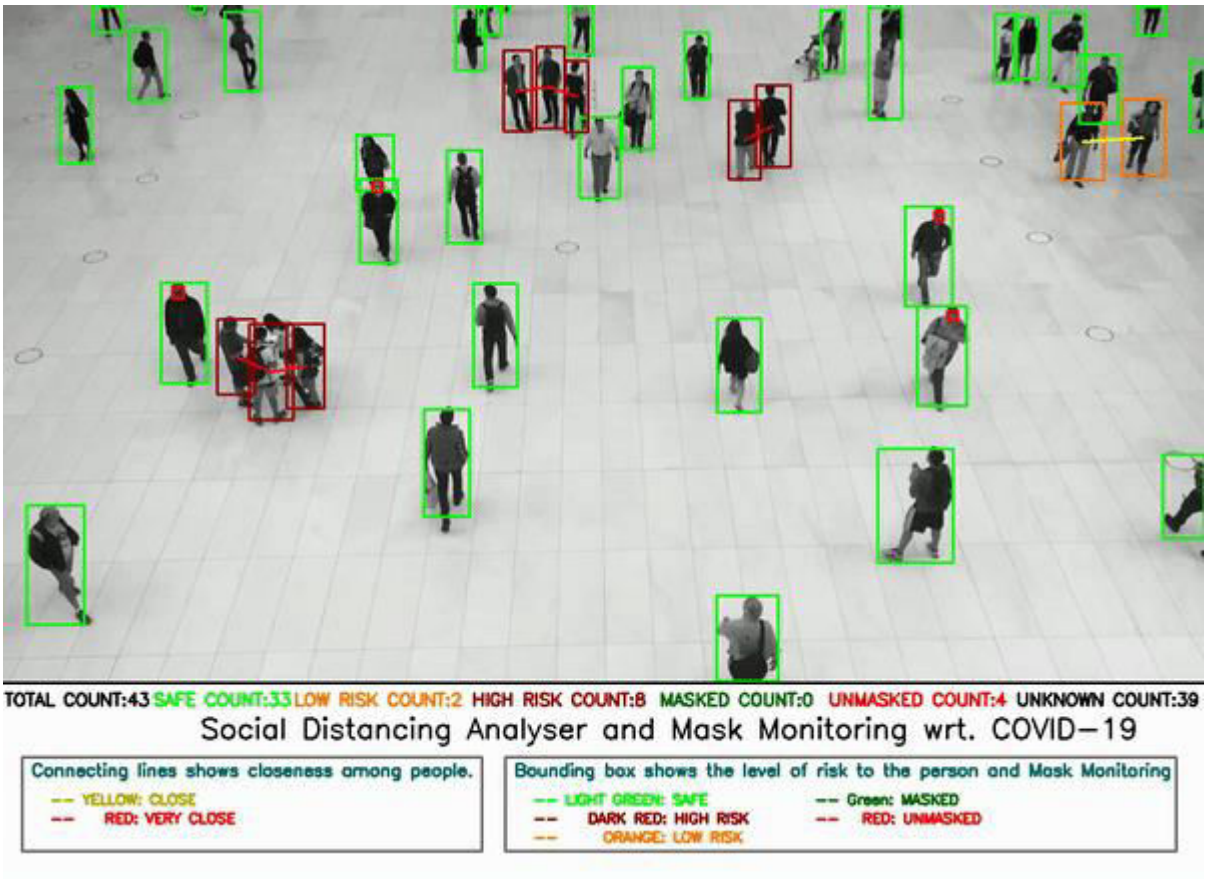
L'algorithme de clustering DBSCAN est utilisé pour reconnaître les intrus potentiels en calculant la distance entre chaque cadre de délimitation et en émettant une alerte si deux personnes ou plus se trouvent sous une certaine distance de seuil, également connue sous le nom de distance de sécurité minimale devant être maintenue entre les personnes. Ce paramètre affecterait les résultats, il est donc ajusté en fonction de la séquence capturée par la caméra de vidéosurveillance. Le tracker de distanciation sociale fonctionne sur la règle des 6 pieds et suit également les personnes qui enfreignent cette règle. Cela complète l'aspect de distanciation sociale de la solution.

Une fois la personne détectée, nous suivons le centroïde afin d'identifier de manière unique la personne et de vérifier la distance avec les personnes adjacentes à l'aide de l'algorithme de clustering de données DBSCAN (Density based spatial clustering of applications with noise)

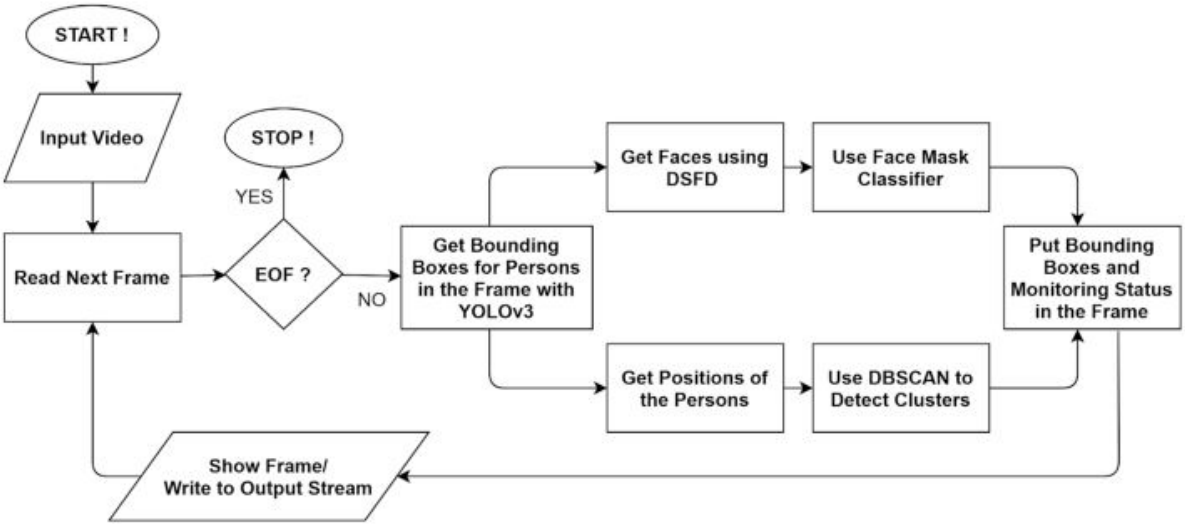
importé du module scikit-learn. En passant la distance de sécurité minimale en unités de pixels comme paramètre de l'algorithme ci-dessus, nous définissons les critères pour déterminer si la distance sociale est maintenue ou non. Si la distance calculée entre les centroïdes des cadres de délimitation est inférieure aux critères, les deux ID sont ajoutés à la

Système d'IA de surveillance à distance sociale et de détection de masque facial pour Covid-19

pour le même.




## 2.7 Pipeline d'architecture :



## 2.8 Résultats

Comme nous pouvons le voir, le système identifie les personnes dans le cadre et place des cadres de délimitation vert clair, rouge foncé ou orange si elles sont en sécurité (Aucun risque), à risque élevé ou à risque faible respectivement. Les lignes de connexion entre les personnes indiquent le niveau de proximité entre les personnes (le rouge indique très proche et le jaune indique proche). Après avoir détecté la personne, le système détecte également le visage et identifie si la personne est masquée ou non en mettant un cadre de délimitation vert ou rouge. Le statut est affiché dans une barre en bas, montrant tous les détails. Ce sont des résultats assez décents compte tenu de la complexité du problème que nous avons entre nos mains.

## 2.9Avantages et inconvénients :

### 2.9.1 Avantages :

Comme nous pouvons le voir, le système d'IA identifie les personnes dans le cadre et leurs visages (s'il est visible) et met des cadres de délimitation verts ou rouges s'ils sont respectivement en sécurité ou non et détecte correctement les visages masqués. Le statut est affiché dans une barre en haut, montrant tous les détails. Ce sont des résultats assez décents compte tenu de la complexité du problème que nous avons entre nos mains et peuvent être déployés dans des solutions d'IA de surveillance CCTV.

### 2.9.2 Désavantages :

Le modèle doit être entraîné sur des tonnes de données pertinentes avant de pouvoir l'appliquer en temps réel et de s'attendre à ce qu'il fonctionne. Il a besoin de beaucoup de puissance de calcul et je veux dire beaucoup! Nous pouvons essayer nos modèles entraînés sur un grand ensemble de données dans des machines virtuelles cloud, ce qui pourrait produire des résultats souhaitables. Pour l'instant, j'ai utilisé des modèles open source pré-entraînés.

## 2.10 Portée future :

Certaines optimisations peuvent être faites sous forme de vectorisation. Pour obtenir la position d'une personne, il existe différentes approches. L'un d'eux utilisant simplement les centres des boîtes englobantes, celui utilisé dans ce projet. L'autre utilise la transformation de perspective d'OpenCV pour obtenir une vue d'ensemble des positions, mais ce genre de besoin nécessite un cadre de points de référence assez précis. Son utilisation augmente également un peu la complexité du système. Cependant, s'il est mis en œuvre correctement, il produira sans aucun doute de meilleurs résultats. Pour l'instant, nous nous en tenons à la première approche. N'oubliez pas qu'il y a toujours place à amélioration !

# BASES DE LA RECONNAISSANCE DE VISAGES

# Introduction

L'intelligence artificielle, et notamment la reconnaissance d'image, aura bientôt une place prépondérante dans nos vies quotidiennes. Les possibilités induites par cette technologie sont considérables, car elle va améliorer les services existants et en créer de nouveaux, impulser de nouvelles opportunités économiques et ainsi redéfinir les standards de nos industries.

La reconnaissance d'image connaît déjà de nombreuses utilisations au quotidien, et dans des domaines très variés. Par exemple, elle permet de rendre votre appareil photo numérique plus intelligent ou d'améliorer le contrôle de qualité des produits de grande consommation sur les lignes de production. Les possibilités induites par cette technologie sont immenses.

# 3.1 BASES DE LA RECONNAISSANCE DE VISAGES :

Les systèmes de reconnaissance de visages (FR) sont caractérisés par plusieurs composantes

tout dépendamment du mode d'application dans lequel ils sont employés, des données utilisées pour l'entraînement des modèles de visages, des données employées durant le mode d'opération, ainsi que des préalables d'utilisation imposés par leur contexte d'application.

À ce chapitre, certaines des différentes modalités applicables aux systèmes de FR seront introduites afin de mieux situer le reste du projet, les motivations de certaines approches, et les développements réalisés.

## 3.2 Types de représentations du visage :

Lorsque observés à très haut niveau, les systèmes FR sont habituellement catégorisés en trois

grandes familles afin de parvenir à définir une méthode de représentation des visages en mémoire [8, 26, 122, 150], de sorte à pouvoir exploiter ces représentations pour entraîner un système de classification qui pourra à son tour identifier ou vérifier l'identité d'un individu.

Ces familles regroupent (1) les approches holistiques (template matching), c'est-à-dire dont le

visage est globalement représenté par un ensemble de caractéristiques discriminantes d'un visage à l'autre, (2) les approches locales (feature-based), qui extraient des caractéristiques particulières à certains points d'intérêt du visage comme les yeux, le nez, la bouche et le pourtour de la mâchoire, et finalement (3) les approches hybrides qui exploitent en symbiose les deux techniques précédentes.

## 3.3 Architecture de base d'un système de reconnaissance de visages :

L'ensemble des sections précédentes démontrent clairement à quel point il y possible d'obtenir des systèmes de FR extrêmement variés, malgré que ceux-ci accomplissent tous un

objectif de nature relativement similaire. Il en est de même pour les architectures permettant de représenter le séquencement des étapes principales employées en reconnaissance de visages.

Dit autrement, malgré que les algorithmes utilisés aux plus bas niveaux de chaque application de FR puissent grandement varier d'un cas à l'autre, en considérant non seulement tous les types de catégories présentées jusqu'à présent, mais également les modes spécifiques d'opérations envisageables, l'architecture à plus haut niveau est en tout point similaire d'un cas à l'autre.

La mise en place un système de FR complet requiert trois étapes majeures, soit (1) la détection de visage (face detection, FD) à partir des images ou séquences vidéo, (2) l'extraction de caractéristiques (feature extraction, FE) permettant de décrire et encoder l'apparence d'un visage sous forme numérique ou binaire, et finalement, (3) la classification de visages (face classification, FC) à proprement parlé pour accomplir la FR qui permet d'identifier ou vérifier la présence d'un individu dans scène en entrée [8, 122, 150]. La Figure 3.1 représente cette architecture sous sa forme la plus traditionnelle et simplifiée.

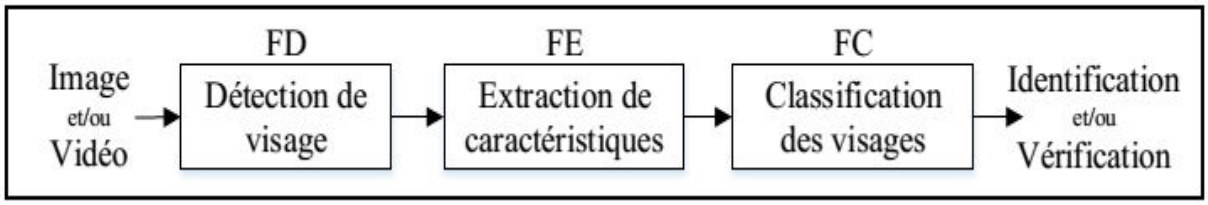



Figure 3.1 Architecture fondamentale des systèmes de reconnaissance de visage

## 3.4 MODULES D'UN SYSTÈME DE FRiVS ET MÉTHODES CONSIDÉRÉES :

Ce chapitre présentera une vue d'ensemble sur les multiples modules qui composent un système de FRiVS spécifiquement pour le cas de S2V-FR entraîné par SSPP. Certains algorithmes qui ont pu avoir été considérés lors de la réalisation du projet, ainsi que quelques justifications des choix, seront discutés lorsqu'opportuns. Il faut noter que toutes les méthodes ne sont pas employées dans le résultat final, mais seulement les techniques jugées comme plus adéquates selon la problématique visée et les défis rencontrés. Ainsi, cette section donnera un aperçu de méthodes existantes, mais ne sera en aucun point une liste exhaustive de toutes les possibilités associées à chacune des catégories explorées. Les lecteurs seront donc invités à se référer aux références bibliographiques ou autres publications en littérature afin d'obtenir plus de détails sur les technologies et algorithmes existants qui pourraient les intéresser davantage.

### 3.4.1 Détection de visages :

Il existe un très grand nombre de méthodes permettant d'accomplir la première étape d'un système de FR, c'est-à-dire la détection de la position de régions d'intérêt des visages sur les images ou vidéos. La technique la plus répandue est sans doute l'algorithme proposé par les auteurs Viola-Jones [129, 130, 131], portant le même nom. La détection VJ consiste à appliquer plusieurs faibles classificateurs en cascade (d'où le nom alternatif cascade classifier de l'algorithme) qui ont été entraînés à progressivement valider avec de plus en plus de rigueur la présence de caractéristiques Haar bien précises et qui représentent la réponse typique d'un visage.

Des exemples de ce type de caractéristiques Haar sont présentés à la Figure 3.2

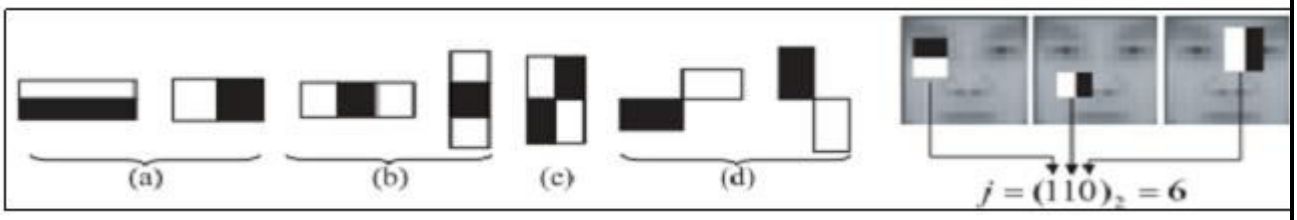



Figure 3.2 :Représentations de caractéristiques Haar employées par Viola-Jones [84]

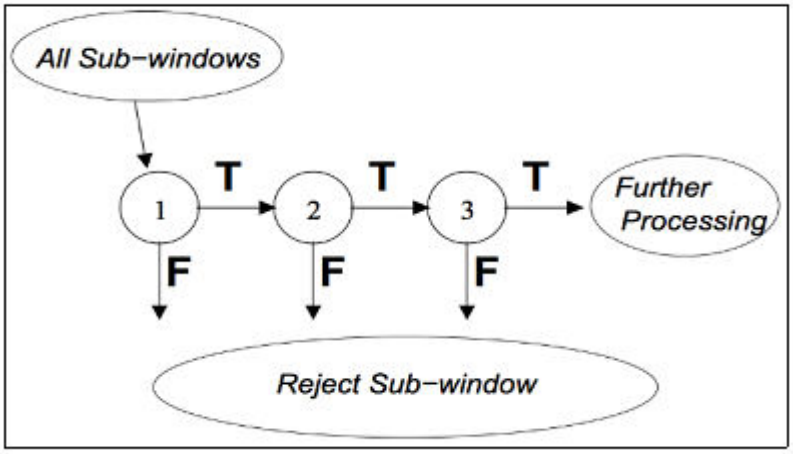



Figure3.3 Classificateurs faibles en cascade de l'algorithme Viola-Jones [130]

## 3.4.2Extraction de caractéristiques :

Afin d'accomplir la reconnaissance de visages, il est nécessaire de représenter ceux-ci dans un hyper-espace augmenté de caractéristiques, dont des classificateurs entraînés à partir

d'exemples appropriés pourraient classifier adéquatement. Ainsi, les techniques d'extraction

de caractéristiques (FE) discriminantes pour représenter les visages sont requises pour un système de FRiVS. De manière générale, trois sous-classes de descripteurs de la catégorie template-matching sont définies en littérature [67, 69], soit (1) les méthodes par projection, (2) les méthodes de vision et, (3) les méthodes par apprentissage de caractéristiques. Celles -ci seront présentées aux sous-sections qui suivent.

## 3.4.3 Méthodes par projection :

Les approches par projection en FR regroupent autant les techniques génériques permettant la

réduction de dimensionnalité des vecteurs descripteurs caractérisant les visages, tels que PCA

ou LDA, que les méthodes très spécifiques pour la reconnaissance de visages comme les représentations par eigenfaces ou fisherfaces. Ces techniques ont maintes fois été appliquées au domaine de la reconnaissance de visages [14, 69, 72, 88, 118, 124, 125, 135]. Le principe général des techniques se base sur le fait que plusieurs des caractéristiques existantes dans l'hyper-espace servant à caractériser les visages sont extrêmement redondantes. Ainsi, étant donné la corrélation élevée existant à travers certaines combinaisons de features, il devient possible de les projeter selon un système d'axes différent, à plus petite dimensionnalité, et où chaque dimension (ou nouvelle feature générée) est bien plus discriminante et non corrélée aux autres données qu'initialement.

Ainsi, les pixels formant l'image de la personne à reconnaître peuvent être similairement projetés afin de trouver des combinaisons pondérées des valeurs de pixels pour former des sous-représentations de celui-ci. En ne sélectionnant que les quelques premières représentations les plus décorrélées, nous obtenons des visages discriminants pouvant être classifiés. Ces méthodes sont particulièrement efficaces pour les cas où la dimensionnalité des vecteurs générés est trop grande et que l'on veut la réduire. En effet, les grands vecteurs utilisés pour la classification causent habituellement le problème de curse of dimensionality [16, 40, 51, 67], soit lorsque les caractéristiques ne deviennent plus suffisamment discriminantes pour identifier des individus particuliers. Cela est causé par le fait que l'information redondante qui est distribuée à travers trop de sous éléments provoque de très faibles variations à travers l'ensemble des valeurs. Les faibles variations de caractéristiques deviennent alors à peine percevables à travers le bruit naturellement généré par les variations de conditions d'acquisition sur un même visage. Cette catégorie de descripteurs est donc principalement basée sur l'analyse des données et les manières de les représenter autrement, ou les compresser, sans toutefois perdre la majorité des informations pertinente à la classification de visages.

## 3.4.4 Méthodes de vision :

Les approches de FE basées sur les méthodes de vision sont celles qui sont fondées sur les

variations de gradients des pixels formant l'image du visage, de techniques morphologiques

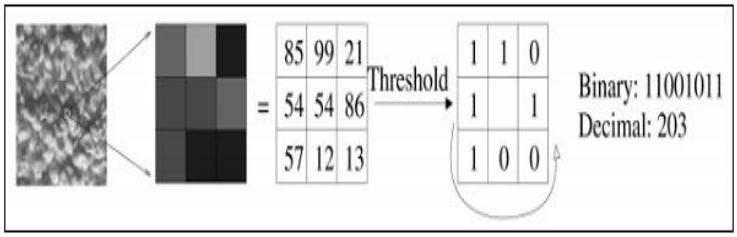
afin d'appliquer des masques sur les pixels, ou des analyses en fonction des couleurs de l'image. Cette catégorie regroupe ainsi les approches traditionnelles de feature engineering, qui sont obtenues en se basant sur des propriétés sélectionnées selon des objectifs précis. Étant donné la variété de la catégorie, il existe une abondance de méthodes, mais les quelques exemples suffiront à démontrer pourquoi elles sont considérées handcrafted.

Une première approche de ce type qui est souvent ressortie en littérature pour la FR est le

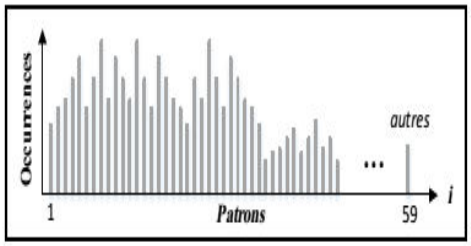
descripteur LBP [2, 53, 148, 149, 151, 152]. Celui-ci est conceptualisé dans l'objectif de caractériser des patterns de textures très spécifiques, à partir des tons de gris dans l'image. La technique se base sur un seuillage des pixels autour d'une valeur centrale dans une fenêtre

localisée, afin d'encoder binairement si les pixels voisins sont supérieurs ou inférieurs à celleci, d'où le nom Local Binary Patterns. Ce procédé d'encodage est représenté à la Figure3.4.

Les occurrences de combinaisons binaires observées sur l'image sont ensuite accumulées dans un histogramme, tel que montré à la Figure 2.13, après filtrage à un maximum de 59 features pour supprimer les patrons irréalistes selon les observations accomplies dans le mode réel.




Encodage des pixels par extraction de caractéristiques LBP [2]




Histogramme de caractéristiques LBP

Figure 3.4 : encodage des pixeles et l'histogramme de caracteristique

## 3.4.5 L'analyse d'une image :

Comme son nom l'indique, une image numérique est une représentation numérique qui prend souvent la forme de matrices à dimensions M x N, où M et N

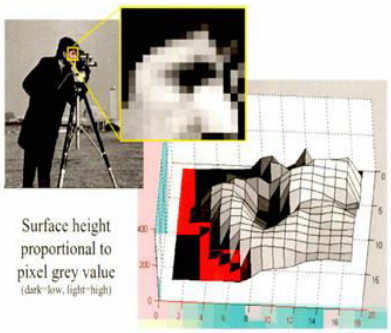
sont la hauteur et largeur d'une image. La matrice d'une image contient, pour chaque emplacement interne, une structure permettant de représenter l'intensité moyenne de la lumière à cet endroit que l'on appelle un pixel. Un pixel n'est toutefois pas une représentation parfaite de l'information visuelle de ce qu'elle représente, puisque la limitation des capteurs cause une perte d'information.

En effet différent facteurs p uv nt influencer la perception de la couleur t de la forme d'un objet, tels que la quantité de lumièr atteignant l'objet, la r'flexion à sa surface ou mêm la qualité du y tème de capture d image pour l'acqui ition

des donné s.

Cette perte possible d information e t xacerbée lor de l'utili ·ation d capteurs numériques. Le ystème de vi ion du cerveau humain 'appuie sur la moitié du cortex du cerveau (Wordenweber et al. 2007) oit environ 30 milliard de neurones et plu ieurs trilliards de synap e et utili e différents systèmes encore peu connus par 1 chercheurs, d'finis par divers modèles opposés (Gib on, 1966; Gregory, 1970) afin d enrichir leur données par les informations déjà connues telles que sa repré ntation 3D, s textur s, ses motif t bi n plus encore.

Une caméra ne peut offrir eule ce processu et transmet donc une perception ubjec ive incomplète de 1 environnement perçu. Chaqu pixel de l'image peut alors seulement repré enter 1 inten it, moyenn de luminance pour son emplacem nt nomm' l x.y t toute information supplémentair nécessite 1 utilisation de plusieurs système additionnels pour le enrichir.




Représentation des intensités de couleurs pour une image. Source : Robert Collins, Université de Penn State [1]

Figure3.5 : representation des l'intensite

# Implementation

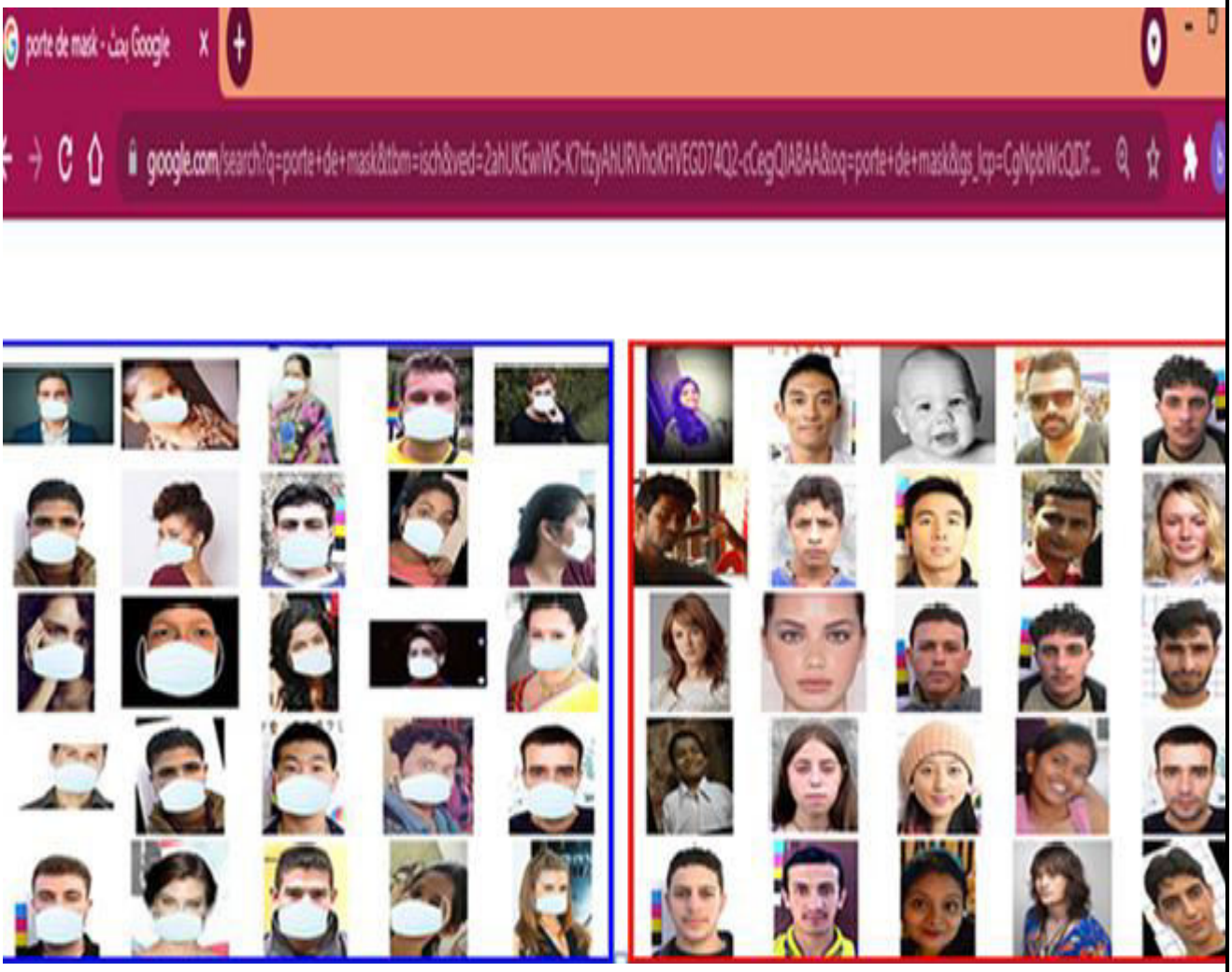
## 4.1Introduction :

La **maladie à coronavirus 2019** ou **Covid-19** est une maladie de type zoonose virale causée par SARS-CoV-2. Les symptômes les plus fréquents sont la fièvre, la toux, la fatigue et la gêne respiratoire. Dans les formes les plus graves, notamment chez les personnes plus fragiles du fait de leur âge ou en cas de comorbidités.

Après avoir présenté dans le chapitre précédent les méthodes, l'objectif de ce chapitre est d'étudier, définir et de mettre en place un système de detection de masque. Ce chapitre donne une vue plus détaillée sur le travail réalisé pour la detection de masque.

## 4.2 Conception du système de detection de masque :

La première étape consiste à choisir une collection d'images à partir d'un moteur de recherche internet (ex. : Google) comme illustré dans la figure suivante :



## 4.3 Choix de l'environnement (langage) :

Pour assurant une implémentation robuste du notre application,je avonsopté l'environnement **PYTHON 3.9 et deep learning.**

Python et deep learning sent des langages dynamique et interactif orienté sur les objets. Il vous permet d'utiliser modules, exceptions, syntaxe dynamique, classes et des types de données de haut niveau. Il est souvent comparé à d'autres langages comme Perl, Ruby ou Java.

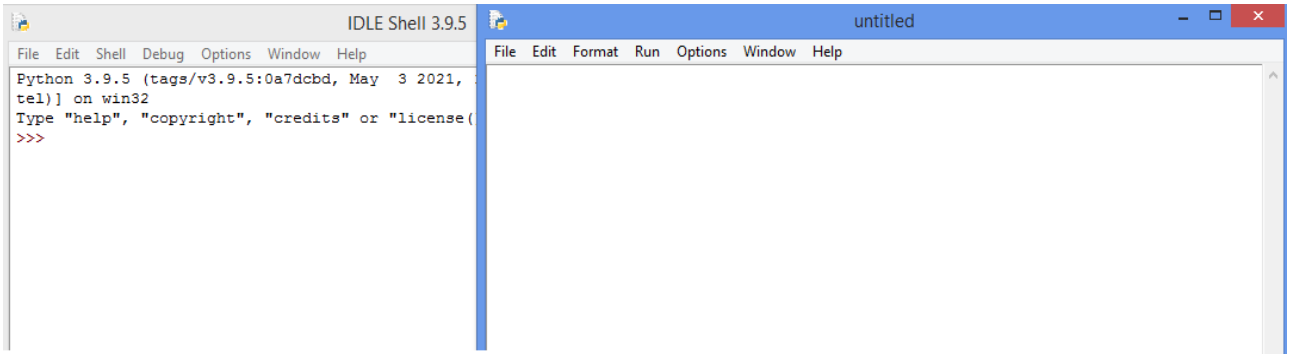
La bibliothèque standard couvre des domaines comme les processeurs de chaîne de caractères (expressions régulières, Unicode, comparaison de tailles de fichiers), protocoles Internet (HTTP, FTP, SMTP, POP, XML), interface des systèmes (appels système, systèmes de fichiers, sockets TCP/IP, connexion de serveurs web).

sent définissent comme un langage facile à utiliser. Il utilise une syntaxe élégante qui permet au code source d'être facilement lu. Il est multiplateforme, donc il fonctionne sous Windows, MacOS, Linux, etc.

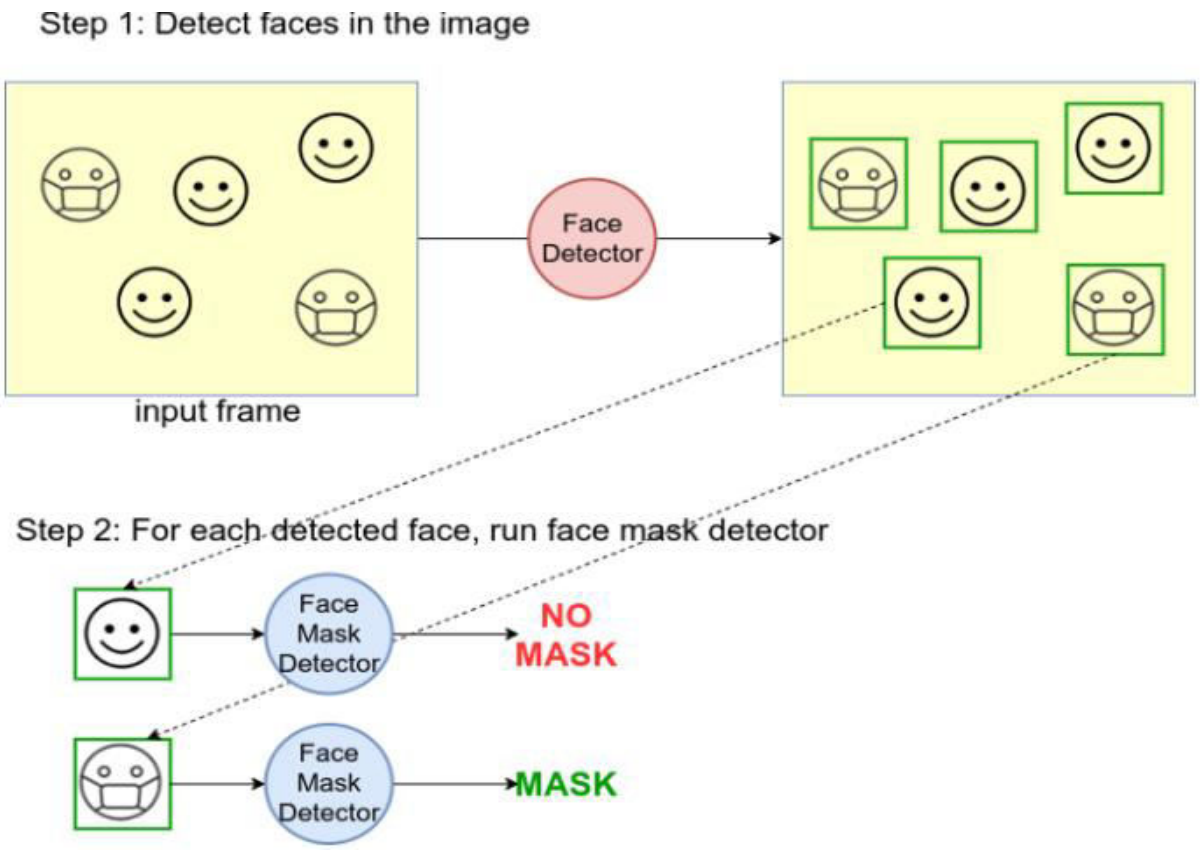
et cela pour de nombreuses raisons :

- Code facile à comprendre et très lisible et une aide très bien faite.
- Propose un ensemble complet d'algorithmes standard de référence.
- Programmation infiniment plus rapide pour le calcul et pour l'affichage.
- Une librairie très riche (descripteurs d'images intégrés).
- Fonctions et d'applications destinées au traitement d'images, à l'analyse, à la visualisation et au développement d'algorithmes.
- Permet de réaliser de nombreuses opérations, parmi lesquelles l'analysed 'images, la segmentation d'images, l'amélioration d'images, la suppression du bruit, le recalaged'images.

## 4.4 l'interface de l'environnement PYTHON 3.9 :




## 4.5Les phases de système de detection




**Phase1** : la selection des visage dans les images

**Phase2** : la detection des visages porte des mask et les visages ne porte pas de mask

**Phase3** :destinguer les resultat

# 4.6 Le Code source de programme de detection de masque

```
try/except
import cv2
import os
import numpy as np
import tensorflow as tf
# Detection de visages
prototxt_path = os.path.join('model_data/deploy.prototxt')
caffemodel_path = os.path.join('model_data/weights.caffemodel')
model = cv2.dnn.readNetFromCaffe(prototxt_path, caffemodel_path)
#Chargement du modèle permettant de détecter le port du masque
modelMasque = tf.keras.models.load_model("QSTOMIT-MASQUE.model")
#Capture de la caméra (idCamera)
cap = cv2.VideoCapture(0)
while True:
 image = cap.read()
   blob = cv2.dnn.blobFromImage(cv2.resize(image, (300, 300)), 1.0, (300, 300), (104.0, 177.0, 123.0))
   model.setInput(blob)
   detections = model.forward()
   h = image.shape[0]
   w = image.shape[1]
   for i in range(0, detections.shape[2]):
      box = detections[0, 0, i, 3:7] * np.array([w, h, w, h])
      (startX, startY, endX, endY) = box.astype("int")

      confidence = detections[0, 0, i, 2]
      # If confidence > 0.5, save it as a separate file
      if (confidence > 0.5):
         frame = image[startY:endY, startX:endX]
         #Appel du modèle appris pour la detection de masque
         capture = cv2.resize(frame, (224, 224))
         capture = capture.reshape((1, capture.shape[0], capture.shape[1], capture.shape[2]))
         predict = modelMasque.predict(capture)
         pasDeMasque = predict[0][0]
         avecMasque = predict[0][1]
         if (pasDeMasque > avecMasque):
            cv2.rectangle(image, (startX, startY), (endX, endY),(0, 0, 255), 2)
            cv2.putText(image, "PAS DE MASQUE", (startX, startY-10),cv2.FONT_HERSHEY_SIMPLEX, 0.45, (0, 0, 255), 2)
         else:
            cv2.rectangle(image, (startX, startY), (endX, endY),(0, 255, 0), 2)
            cv2.putText(image, "OK", (startX, startY), cv2.FONT_HERSHEY_SIMPLEX, 0.45, (0, 255, 255), 2)
   # Affichage de l'image
   cv2.imshow('img', image)
   k = cv2.waitKey(30) & 0xff
    if k==27:
      break
cap.release()
```

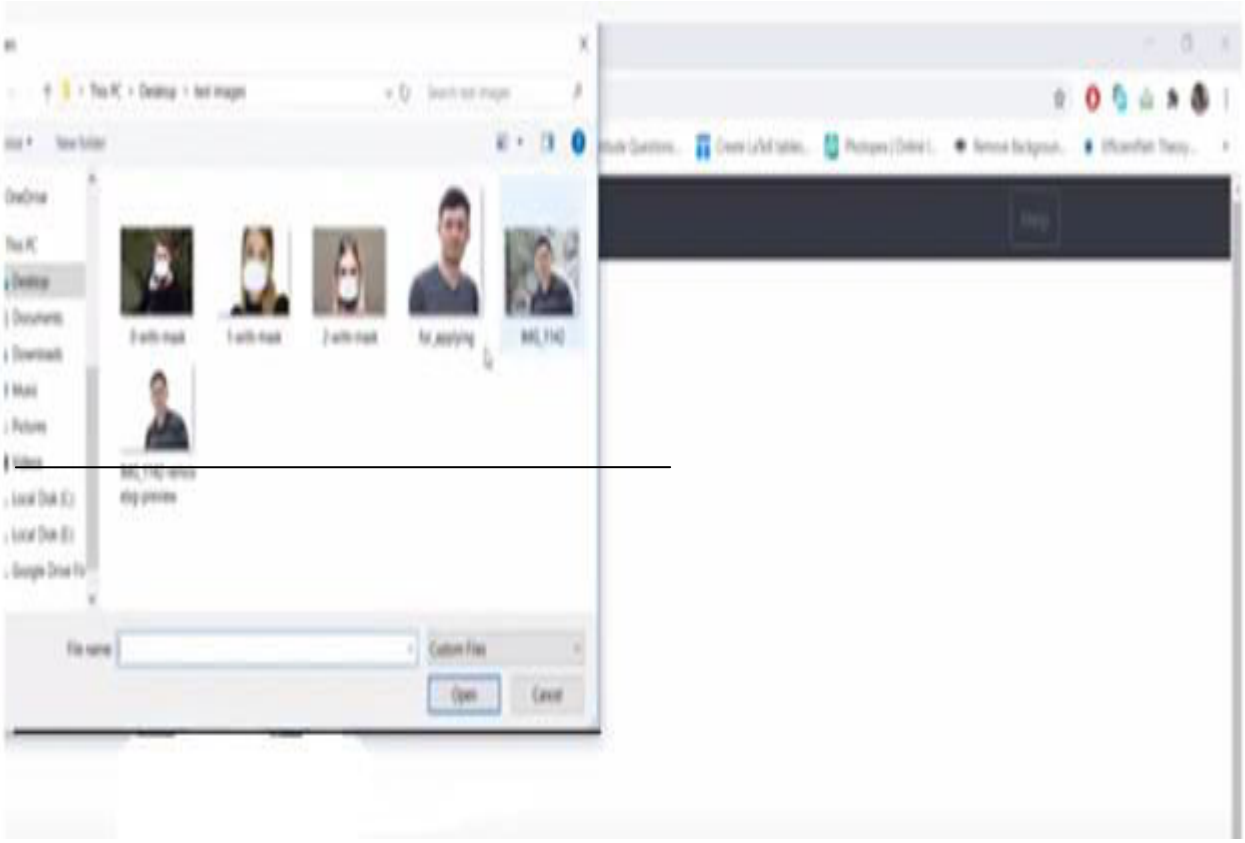
## 4.7 L'interface de l'application :

1 :ouvre le dossier pour choisir les image.

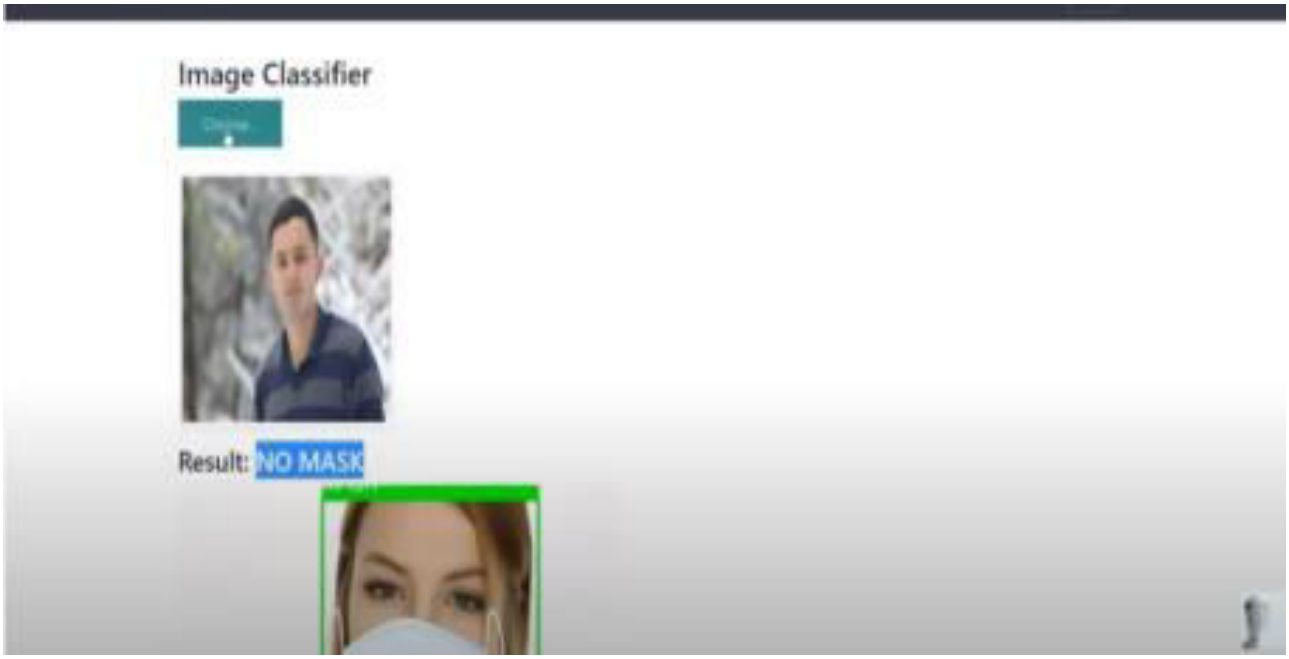
2 :la detection des images porte ou nom le mask

3 :la zone des resultat

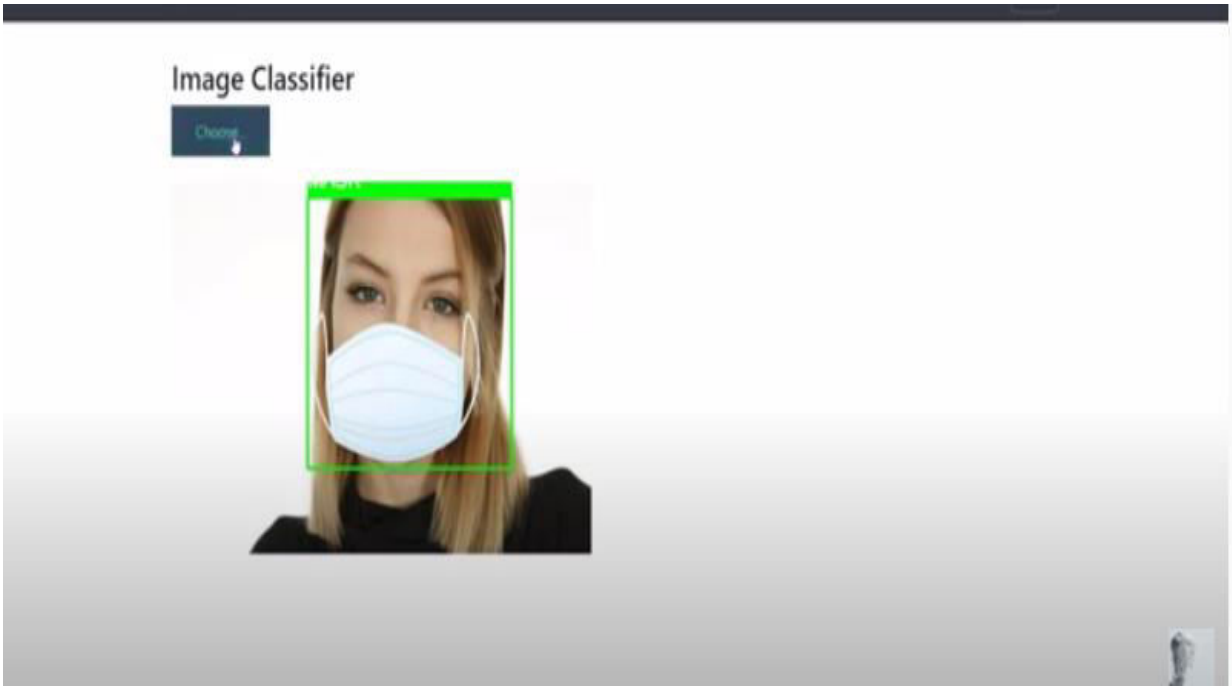
### 4.7.1 On a choisit des photos :



### 4.7.2 La detection de porte de mask :



### 4.7.3 l'application destingue le resultat




## conclusion :

Nous avons présenté dans ce chapitre notre contribution en essayant de répondre aux besoins des utilisateurs liés au domaine de detection de porte de mask, nous y avons développé les principales étapes d'un processus de detection à savoir .

# Conclusion Generale

## Conclusion générale :

Malgré les efforts et les travaux intensifs réalisés dans le domaine de classification des images, aucune méthode n'est jugé fiable à 100%, Mais au fur et à mesure les chercheurs essayent d'améliorer les scores pour de meilleurs résultats.

Dans ce mémoire, nous avons abordé le problème de classification des images. Plus précisément, nous nous sommes focalisés sur la classification basée sur le contenu visuel.

Le but de ce travail été d'ajouter une valeur minimale aux travaux de recherche du domaine de classification des images distinctement, et au domaine de l'imagerie et la reconnaissance des formes d'une façon générale. Pour atteindre cet objectif ,nous avons commencé par un état de l'art sur la vision par ordinateur et les caractéristiques des images, Ensuite, nous avons étudié les différentes méthodes d'extractions des caractéristiques visuelles. Enfin nous nous sommes intéressés à l'intégration de nos méthodes proposées à un système de classification des images.

On a choisi aussi des méthodes comme méthode de classification, ce choix est justifié par leur principal avantage pour tous les méthodes supervisées, qui est l'aspect d'"*apprentissage*" qui permet de développer des systèmes dynamiques et évolutifs.

Nous avons développé une interface qui permet àl'utilisateur de proposer facilement un modèle et d'afficher ses résultats de classification, cette application peut être intégré ou utilisé avec des machines qui possèdent un environnement d'acquisition des images. De plus ce travail peut être utilise dans des autre domaines (analyse des images, traitement, classification…etc.).

Enfin, ce travail reste ouvert pour des travaux de comparaison avec d'autres méthodes de classification (SVM, méthodes neuronales…..) et d'autres types de descripteurs (descripteurs de Fourier, ondelette …).

# BIBLIOGRAPHIE

**[1]. [BEDOUHENE]** BEDOUHENE. Recherche d'images par le contenu. Université Mouloud Mammeri, Tizi-Ouzou. Mémoire De Magister.

**[2]. [TAIBAOUI 13]** Mohamed Taibaoui et Djabbar Djaafer, La découverte des concepts sémantiques cachés avec plusieurs niveaux d'abstraction pour la recherche d'images,
Mémoire de Master, Université d'Ouargla, 2013.

**[3]. [ABED, H., & ZAOUI, L]**. (2011). Système D'Indexation et de Recherche d'Images par le contenu. Système D'Indexation et de Recherche d'Images par le contenu, 3.

**[4]. [BEN CHIKH, N., & BEN ELZZINE, R.]** (2011, 06). LA RECHERCHE D'IMAGES PAR LA SEMANTIQUE. LA RECHERCHE D'IMAGES PAR LA SEMANTIQUE. Mémoire de Master informatique académique, département mathématique at informatique, Université kasdiMarbah Ouargla.

**[5]. [LAOUBI, A.]** (2014, 06 19). Recherche d'images sémantique basée sur la sélection automatique des concepts. Mémoire MASTER ACADEMIQUE. Ouargla, Département d'informatique et des technologies de l'information, Algérie : UNIVERSITE KASDI MERBAH OUARGLA.

**[6]. [LESLOUS, O., & BOUCETTA, M.]** (2011, 07 27). Recherche d'images par le contenu visuel : recherche d'images par le contenu visuel :Mémoire de master. Ouargla, départementmathématique et informatique,UniversitéKasdi Merbah.Ouargla

**[7]. [LÊ THILAN]**(2005)**.** indexation et recherche d'images par le contenu. Mémoire de master. Institut polytechnique de Hanoï.

**[8]. [Laalmi Houria, Bassou Oumessàd]**, Classification Des Feuilles de Plantes à Base de Moment de Hu, Mémoire de master, Université KASDI MARBAH OUARGLA, 06/2014.

**[9]. [Mme Benyahia Souad Née Belaidi]**, Application de la classe des méthodes d'apprentissage statistique SVM (support vector machine) pour la reconnaissance des formes dans les images, Mémoire de magister, Université Abou Bakr Belkaid– Tlemcen, 2012.

**[10]. [Web, 01]** http://fr.wikipedia.org/wiki/Histogramme(Consulté le 14/04/2016)

**[11]. [Web, 02]**http://fr.wikipedia.org/wiki/Les formats d'images (22/03/2016).

**[12]. [Web, 03]** http://www.oms.fr (Consulté le 09/03/2016).

**[13]. [Web, 04]**http://www.pages-perso-julie-carreau.univ-montp2.fr (Consulté le 11/03/2016).
**[14]. [Web, 05]**http://www.grappa.univ-lille3.fr/~ppreux/fouille.(Consulté le 11/03/2016).

**[15]. [Web, 06]**http://www.iro.umontreal.ca/~vincentp/ift3390.(Consulté le 15/03/2016).

**[16]. [Web, 07]**https://tel.archives-ouvertes.fr/tel-00541059.(Consulté le 19/05/2016).

**[17]. [Web, 08]**http://bnt.insa-rouen.fr/.(Consulté le 17/04/2016).

**[18]. [Web, 09]**http://www.creatis.insa-lyon.fr.(Consulté le 06/05/2016).

**[19]. [Web, 10]**http://www.inria.fr.(Consulté le 17/05/2016).

**[20]. [Web, 11]**http://perso.telecom-paristech.fr/~gramfort/liesse_python/3-Scipy.html.(Consulté le 19/04/2016).

**[21]. [Web, 12]**http://scikit-learn.org/stable/index.html.(Consulté le 22/02/2016).

**[22]. [Web, 13]**http://www.loria.fr/~rougier/teaching/matplotlib/matplotlib.html.(Consulté le 05/05/2016).

**[23]. [Web, 14]**http://jrjohansson.github.io/.

**[24]. [Web, 15]**http://sysdiag.cnrs.fr.(Consulté le 27/01/2016).

**[25]. [Web, 16]**http://www.mathworks.com.(plusieurs fois).