

**REPUBLIQUE DU CAMEROUN**

**Paix -Travail – Patrie**

**UNIVERSITE DE YAOUNDE I**

**FACULTE DES SCIENCES**

**DEPARTEMENT  
D'INFORMATIQUE**

**BP/P.O.Box 812 Yaoundé-**



**REPUBLIC OF CAMEROON**

**Peace - Work – Fatherland**

**UNIVERSITY OF YAOUNDE I**

**FACULTY OF SCIENCE**

**COMPUTER SCIENCES  
DEPARTMENT**

**BP/P.O.Box 812 Yaoundé-**

**SECURITE DES PERSONNES DANS UN ESPACE  
(UNIVERSITE DE YAOUNDE I) BASE SUR LA DETECTION DE  
MOUVEMENTS**

**GROUPE 6**

<b>NOMS &amp; RENOMS</b>	<b>Matricule</b>
DONGMO NGUIMKENG BIBICHE LAURE	21S2812
MAFRE MEIDIE SANOU SONIA	18T2775
NYA NJIKE ARMEL	21S2802
KAPATA NGANKAM DARLIN CANIS	21S2810

**Sous l'encadrement de :**

**Dr AMINOU**

**Année académique: 2021-2022**

## SOMMAIRE

### SOMMAIRE

#### PREMIÈRE PARTIE : CAHIER DE CHARGE

I.	PRÉSENTATION DU PROJET.....	3
	Contexte.....	3
	Objectifs.....	3
	Cadre.....	3
	Enjeux.....	4
II.	DESCRIPTION FONCTIONNELLE.....	4
	Description de l'application.....	4
	Les fonctionnalités de l'application.....	4
	Identification.....	4
	Démarrer le système.....	4
	suivi.....	5
III.	SYSTÈME D'ALERTE DES MOUVEMENTS À RISQUE .....	5
	Explication.....	5
	Le type de mouvements a risque.....	6
IV.	LES ÉTAPES DE LA SÉCURISATION .....	7
	Détecter.....	7
	Signaler.....	7
V.	BESOINS NON FONCTIONNELS.....	7

#### DEUXIÈME PARTIE : MANUEL D'IMPLÉMENTATION

VI.	CONCEPTION .....	8
-----	------------------	---

A- les diagrammes

- 1.Diagramme de cas d'utilisation.....
- 2.diagramme de classe.....
- 3.diagramme de séquences.....

B- les interfaces.....

- 1.interface principale.....
- 2. ....

C-outils et technologie

VII. SOLUTION PROPOSEE.....9

A-MODÈLE DE SCIENCE DE DONNEES.....9

Fonctionnement du modèle.....9

Méthodologie et données ... .....9

Impact du modèle sur l'application .....9

B-QUELQUES CAPTURES DU CODE SOURCE.....9

TROISIÈME PARTIE : MANUEL DE DÉPLOIEMENT

- 1.diagramme de déploiement.....

QUATRIÈME PARTIE : MANUEL D'UTILISATION

- 1.guide d'installation.....

## CONCLUSION

### Annexes

#### références :

- [<https://www.researchgate.net/>] state of the art of the human activity recognition
- [<https://www.tensorflow.org/?hl=fr>] : building tensorflow with CUDA
-

# **PREMIÈRE PARTIE : CAHIER DE CHARGE**

## **I-PRESENTATION DU PROJET**

### **1. Contexte**

L'université de Yaoundé I, est une institution d'enseignement supérieur d'étude et de recherche, constituée par la réunion de divers établissements. Compte tenu de sa superficie et du nombre important d'étudiants, l'université de Yaoundé I ne dispose d'aucun dispositif pour garantir la sécurité de tous car l'accès n'est pas restreint uniquement aux étudiants et aux personnels. L'université aimerait avoir un système garantissant la sécurité des étudiants et du personnel administratif au sein du campus universitaire en temps réel.

### **2. Objectifs**

Ce cahier de charges a pour objectif de fournir des informations nécessaires pour la mise sur pied d'un outil optimal et dynamique qui facilitera la gestion de la sécurité des personnes basé sur la détection de mouvement au sein de l'université de Yaoundé I.

### 3. Cadre

La gestion de cette sécurité consiste à mettre à la disposition du personnel sécurité de l'université une interface permettant à partir des sources vidéo d'alerter sur des situations à risques à partir des analyses des séquences vidéo recueillies.

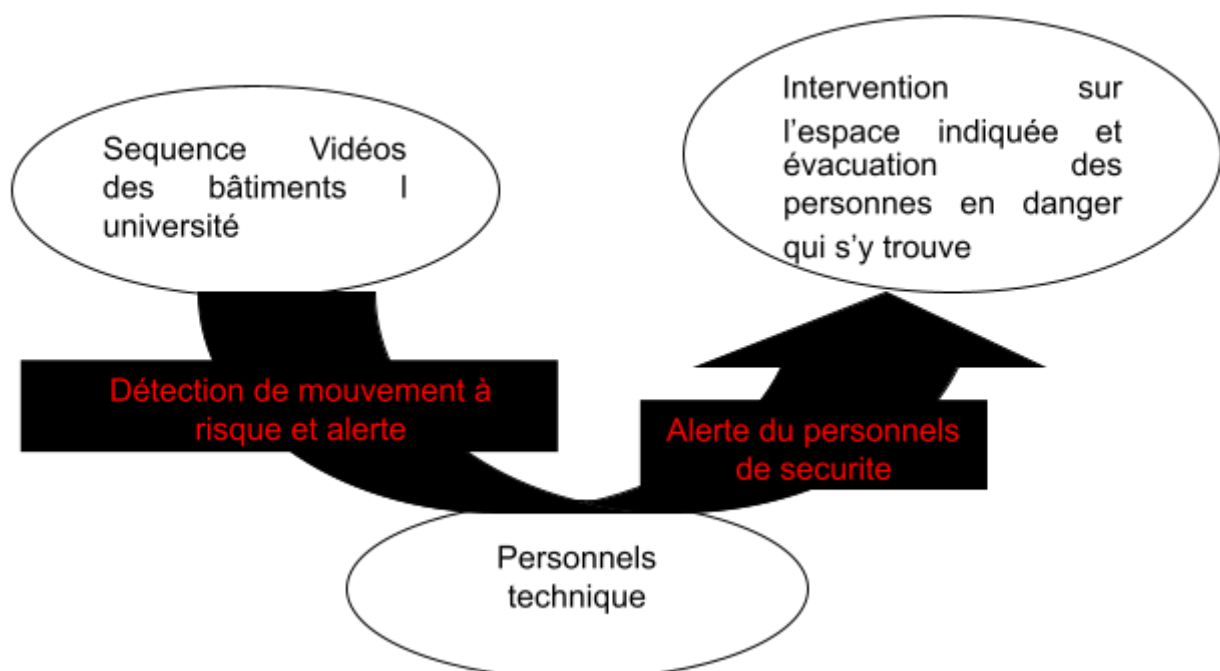
### 4. Enjeux

L'implémentation de cet outil de sécurité des personnes au sein de l'université de Yaoundé I à travers la détection de mouvement aura pour enjeux de fournir un service de sécurité de bonne qualité en visualisant en temps réel les bâtiments autorisés aux étudiants et aux personnels du campus.

## II. DESCRIPTION FONCTIONNELLE

### 1. Description de l'application

L'application qui sera développée doit fournir les fonctionnalités suivantes présentées dans le diagramme de cas d'utilisation ci-après :



## 2. Les fonctionnalités de l'application

### 2.1. Identification

L'identification du personnel technique est indispensable pour accéder à l'application. Elle permet de tracer toutes les actions effectuées par le personnel dans le but d'assurer au mieux la sécurité du système. Sur cette page, le personnel doit remplir un formulaire de login renseignant les champs : nom du personnel et mot de passe.

### 2.2. Démarrer le système

Le démarrage du système va lancer automatiquement les tâches de tracking de mouvement et d'objets, de classification, d'évaluation et de gestion des alertes en cas de détection de mouvements suspects.

#### i. Surveillance et protection contre les mouvements à risque

Détection de mouvements risques dans les endroits ou bâtiments ou s'y trouvent des étudiants ou personnels de l'université

Dans le cas d'une détection de mouvements à risque engendré par des personnes les opérations suivantes doivent être effectuées :

- Détection de l'alerte
- Emission par le personnel technique d'un signal d'alarme vocale par voie téléphonique pour avertir le personnel de sécurité sur l'espace indiquée
- Intervention du personnel de sécurité en charge de la sécurité

Ici notre modèle sera en mesure d'apprendre à détecter les mouvements à risque dans des espaces précis de l'université de Yaoundé I.

### 2.3. Suivi

#### i. Consultation des Alertes

Sur cette page, le personnel peut visualiser toutes les différentes alertes contenant la date, heure, le type d'alertes et de la localisation. Ainsi, on peut filtrer les alertes par type de localisation, par date, par heure.

#### ii. Archivage des alertes

Si une alerte n'a pas encore été archivée, la date de retour ne sera pas encore renseignée dans l'historique des alertes.

Chaque alerte traitée devra figurer dans l'historique des alertes qui ont été traitées ce qui permettra au système de faire une étude statistique pour ressortir les espaces ou les bâtiments où on détecte le plus souvent de risque.

### **III. SYSTÈME D'ALERTE DES MOUVEMENTS A RISQUE**

#### **1. Explication**

Une alerte est un signal qui prévient d'un danger imminent, en particulier d'une attaque ou d'un danger grave.

Dans notre cas l'alerte sera déclenchée lorsqu'un mouvement dangereux, suspect, inquiet sera observé par le système ceci pour permettre la protection des étudiants et du personnel administratif de campus en temps réel. Lorsque le système détectera une alerte, le personnel technique alertera le personnel de sécurité sur l'espace indiqué pour un contrôle de la situation.

#### **2. Le type de mouvements à risque**

Le type d'alerte signalé sera fonction de la nature du mouvement détecté et considéré comme mouvement dangereux. Comme type de mouvements à risque, nous pouvons énumérer :

- Mouvement de Bagarres : affrontement ou agression physique entre les personnes dans un espace du campus universitaire (coups de poings, coups de pieds). présence d'animaux libres (chien), combat, foule en mouvement ...

**NB : Le potentiel risque d'un mouvement dépendant fortement de l'environnement où il est détecté.**

Ces différentes situations ou caractéristiques nécessitent une reconnaissance par le système automatique.

- Mouvement de panique : par un élément déclencheur ou par la peur, une personne ou un petit groupe de personnes vont entraîner les autres derrière elles, et les gens qui chutent risquent d'être piétinés qui est caractérisé par les symptômes physiques à savoir :



- Palpitations cardiaques ou accélération du rythme cardiaque
  - Tremblements ou secousses musculaires parfois généralisées à tout le corps
  - Symptômes respiratoire (étouffement ou essoufflement)
  - Vomissement
- Déplacement vers un espace interdit : dans le cas où une personne se déplace vers un espace déjà connu comme un espace à accès interdit (bâtiments administratifs, bureaux administratifs) pour les étudiants de l'université et uniquement autorisé aux personnels administratifs de l'université. Ce type de mouvement est caractérisé par :
    - Lorsqu' une personne reste trop longtemps devant un espace interdit (en général devant une porte d'accès interdit)
    - Lorsqu' on observe un étudiant se rapproche d'un de ces bâtiments ou bureaux à une heure tardive (à plus 22h)
  - Mouvement d'agressions : coups portés au moyen d'objets divers et variés (armes à feu, armes blanches, ustensiles de cuisine, chaises), morsures. Ce type de mouvement est caractérisé par :
    - Possession d'objets dangereux (armes à feu, armes blanches, ustensiles de cuisine)

## **IV. LES ÉTAPES DE LA SÉCURISATION**

### **1. Détecter**

Détecter les mouvements à risque pour agir convenablement selon des procédures bien déterminées afin d'assurer la sécurité des personnes et de limiter les dégâts dans les espaces contrôlés.

### **2. Signaler**

En ce qui concerne le signalement, ici le système fournira les informations sur l'espace ou le mouvement a été détecté. Les informations porteront sur le jour, l'heure, le type d'alerte détecté, le nom du bâtiment ou du secteur concerné et son plan de localisation pour une intervention rapide et efficace qui sera utile pour la sauvegarde des différentes alertes détectées.

## V. BESOINS NON FONCTIONNELS

Du point de vue technique, le système doit être :

- Autonome, indépendant d'aucun autre système-hôte
- Facilement configurable à travers un écran offrant à la fois simplicité de la configuration et la visualisation des informations que le système doit fournir au responsable technique
- Fiable, avec une probabilité de fausse alarme infiniment petite
- Alimenté par le courant de secteur avec une alimentation de secours non interruptible
- D'un coût relativement abordable par rapport aux systèmes disponibles dans ce domaine
- Les informations sur les séquences vidéos enregistrées doivent être confidentielles et Les données de l'application seront protégées contre tout accès non autorisé
- L'application doit présenter les fonctionnalités attendues pour la gestion de la sécurité des étudiants et du personnel de l'université
- Interface de L'application doit être simple et facile à manipuler
- L'application doit être facilement maintenable et évolutive
- L'application doit être compatible avec les différents systèmes
- L'application doit pouvoir gérer la surcharge lorsqu'il y'a plusieurs alertes sur le réseau
- Le système devra avoir un temps de réponse satisfaisants en temps réel 🎬  
L'application doit pouvoir sauvegarder et mémoriser les différents mouvements effectués
- L'application doit être très facile à utiliser
- Le chargement d'une page doit être très rapide
- L'application doit être tolérante face aux pannes

# DEUXIÈME PARTIE : MANUEL D'IMPLÉMENTATION

## V. CONCEPTION

## VI. SOLUTION PROPOSÉES

### A-MODÈLE DE SCIENCE DE DONNÉES

#### 1.FONCTIONNEMENT DU MODÈLE

Le système développé est un modèle basé précisément sur le Deep Learning. En particulier, les réseaux de neurones convolutifs (CNN), qui sont utilisés pour détecter les caractéristiques des séquences de la vidéo, tandis que les réseaux de neurones récurrents (RNN) sont utilisés pour analyser ces caractéristiques et classifier les activités possibles dans la vidéo. Les techniques d'apprentissage en profondeur ont été d'une très grande utilité dans le contexte de la classification au cours des dernières années en raison de la croissance de la puissance de calcul. Pour ce projet, nous utiliserons :

- Les réseaux de neurones convolutifs (CNN) pour la compréhension des images. Les réseaux de neurones convolutifs sont un type spécial de réseaux de neurones artificiels conçus pour traiter une grande quantité de données d'entrée (images, audio ou vidéo). En raison de la grande quantité de données d'entrée, CNN réduit les informations en examinant des régions individuelles des données dans le but de récupérer les caractéristiques pertinentes.

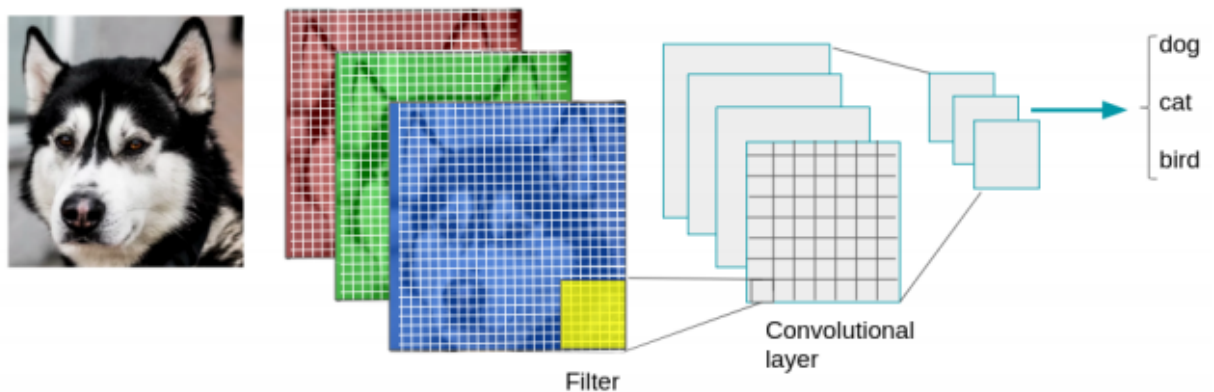


fig 1 : processus de convolution dans un CNN

Pour notre projet, nous allons utiliser la partie extraction de caractéristiques d'un modèle pré-entraîné : processus appelé apprentissage par transfert axé sur le stockage des connaissances et leur application à un problème différent mais connexe. Le modèle que nous allons utiliser est Inception v3 (pour ce qui est des CNN) car il a une excellente précision de classification et un faible coût de calcul. (Il existe d'autres modèles qui atteignent de meilleures performances comme Inception.)

- Les réseaux de neurones récurrents (RNN) pour une compréhension temporelle des données telles que vidéos ou textes. Les méthodes décrites ci-dessus sont conçues pour classer des données d'entrée indépendantes. Mais que se passe-t-il lorsque nous traitons avec des données de séries chronologiques ? Pour répondre à ces exigences, un nouveau type de réseau de neurones a été conçu pour modéliser les données de séquence temporelle. Ces réseaux sont appelés réseaux de neurones récurrents (RNN) et permettent aux informations de persister, en y ayant des boucles. Dans ce modèle, un réseau LSTM (Long Short-Term Memory) est proposé pour suivre la partie caractéristique du Réseau Inception v3. La taille du vecteur de caractéristiques renvoyé par le réseau Inception v3 est de 2048, donc une couche LSTM de la même taille est proposée pour se souvenir de chaque caractéristique de la séquence de vecteurs.

#### ❖ Extraction de caractéristiques.

Le modèle en général étant composé de deux sous modèles différents, un CNN et un réseau LSTM. Le CNN, dont le rôle est d'obtenir les caractéristiques des cadres, est déjà entraîné par Google, donc seul le RNN doit être entraîné. Afin d'entraîner le réseau LSTM, les données vidéo converties en vecteurs de caractéristiques(extraction de caractéristiques), par le modèle CNN (Inception V3) vont être les entrées de ce réseau. Ce processus peut prendre plusieurs heures selon la machine (donc des jours pour des PC classiques).

## ❖ préparation de données

Une fois toutes les caractéristiques obtenues, la prochaine étape consiste à alimenter le réseau LSTM avec celles-ci, ce qui n'est pas simple. La première approche logique est d'alimenter toute la séquence d'activités (dans un format vectoriel de caractéristiques) au réseau LSTM. Mais c'est une approche naïve car chaque activité a une durée différente et, par conséquent, une quantité différente de vecteurs de caractéristiques, pourtant Tensor Flow indique qu'il n'est pas possible d'avoir des séquences de longueur variable, chaque séquence doit être de même longueur. Ainsi, la forme des données doit ressembler à :

Video x images x caractéristiques

La première approche consistait à utiliser la fonction `pad_sequences()` où, selon Keras cette fonction "Transforme une liste de séquences en séquences de longueur fixe (maxlen). Les séquences plus courtes que maxlen sont complétées par des zéros à la fin. Les séquences plus longues que maxlen sont tronquées afin qu'elles correspondent à la longueur souhaitée." Cette approche était correcte, mais pouvait encore être améliorée. Au lieu de remplir avec des zéros (qui n'ajoutent pas d'informations) des séquences qui sont plus courtes, il pourrait répéter la séquence sous forme de boucle jusqu'à ce qu'elle soit adaptée à la longueur désirée. Et pour les séquences plus longues, au lieu de récupérer juste la première partie de la séquence, il peut s'agir de récupérer une représentation de la séquence entière en stockant certaines images d'une manière organisée. Cette approche améliore la précision du système de manière significative.

## ❖ Apprentissage du modèle

Pendant l'entraînement, le jeu de données est séparé en jeu d'entraînement et en jeu de validation. Toutes les données d'entraînement sont transmises au réseau de neurones pour l'entraîner, et les données de validation sont transmises pour mesurer la précision. Pour chaque itération ou époque d'entraînement, Keras affiche différents paramètres qui montre l'amélioration de l'apprentissage :

ACC : la précision de la prédiction de la classe cible comme étant la classe correcte.

Top k categorical accuracy : La précision de la prédiction de la classe cible entre les 5 premières classes prédites.

loss : La fonction de perte choisie. Ce sera l'entropie croisée catégorique.

Val acc : précision pour l'ensemble de validation.

Val top k categorical accuracy : Précision catégorique Top k pour l'ensemble de validation.

Val loss : précision de perte pour l'ensemble de validation.

## II-MÉTHODOLOGIE ET DONNÉES

Le système sera implémenté à l'aide de tensorflow, un outil d'apprentissage en profondeur très populaire qui simplifie le processus de construction du réseau de neurones en offrant un APT qui communique avec des frameworks sophistiqués. Notre modèle aura besoin de séquence vidéo pour s'entraîner. Une analyse des jeux de données disponibles a été réalisée afin de choisir le plus adapté pour la mise en œuvre.

### Les DataSet

Pour ce projet, nous avons répertorié 3 ensembles de données qui sont :

- Les deux premiers ensembles de données explicitement conçus pour évaluer la détection de combat. ○ Le premier jeu de données ("Films") consiste en 200 clips vidéo dans lesquels des combats ont été extraits de films d'action. Les vidéos non-combat ont été extraites des ensembles de données de reconnaissance de l'action publique. ○ Le deuxième jeu de données ("Hockey") est constitué de 1000 clips à une résolution de 720x576 pixels, divisés en deux groupes, 500 combats et 500 non-combats, extraits des matchs de hockey de la Ligue nationale de hockey (LNH). Chaque clip était limité à 50 images et la résolution réduite à 320x240. Contrairement à l'ensemble de données Hockey, dont le format et le contenu étaient relativement uniformes, l'ensemble de données sur les films comportait une plus grande variété de scènes capturées à différentes résolutions. Cet ensemble de données a été redimensionné à une taille uniforme.
- L'UCF101 est un ensemble de données de vidéos d'action réalistes collectées sur YouTube, comportant 101 catégories d'action. UCF1 donne la plus grande diversité en termes d'actions et avec la présence de grandes variations dans le mouvement de la caméra, l'apparence et la pose de l'objet, l'échelle de l'objet, le point de vue, l'arrière-plan encombré et les conditions d'éclairage, c'est l'ensemble de données le plus difficile à ce jour. Pour ce cas, c'est encore plus difficile puisqu'il comprend également 50 actions issues du sport. À notre connaissance, il s'agit de l'ensemble de données le plus vaste et le plus difficile dans lequel un algorithme de détection de combat ait été testé.
- Activity Net est une vaste collection de vidéos Youtube, étiquetées avec les activités qui apparaissent sur eux. Pour ce travail, la version 1.3 du jeu de données

sera utilisée. Le nombre total d'échantillons de vidéos de l'ensemble de données est de 19994 (849 heures de vidéo) étiqueté avec 203 activités différentes. Chaque vidéo n'est pas limitée à une activité, le jeu de données compte en moyenne 1,41 catégories d'activité par vidéo. Intégration Nous projetons faire une classification 1 vs all, ce qui nous permettra de restreindre la quantité de donnée à utiliser

### III- IMPACT DU MODÈLE SUR L'APPLICATION

L'interface principale de notre application est celle qui présente la sortie vidéos des caméras de surveillance, est l'élément qui va interagir avec notre modèle en ce sens que c'est sur cette interface qu'on pourra visionner les images de vidéosurveillance qui seront analysées en arrière plan par notre modèle et dans le cas précis d'une détection de situations à risques ou dangereuses, une alerte (sonore ou visuel) signale au personnel technique de la situation et automatiquement l'alerte donc on pourra visualiser les caractéristiques(types, heures, ...) en effectuant un clic dessus s'affiche sur le panneau vertical à gauche des écrans. Toutes les alertes relayées par le système(modèle) sont sauvegardées et pourront être visualisées dans la partie historique des alertes de notre application.

## TROISIÈME PARTIE : MANUEL DE DÉPLOIEMENT

- Dépendances logiciel

Toutes les dépendances logicielles sont installées automatiquement à l'installation du logiciel.

- Dépendances matérielles

mémoire : min-> 4Go

Processeur : -> core i3

## Conclusion

Dans ce travail, nous avons implémenté un système de reconnaissance de l'activité humaine en utilisant un modèle de Deep Learning. Le système développé peut reconnaître plusieurs activités dans les vidéos RVB en extrayant des caractéristiques avec un CNN, et gère la séquentialité avec un RNN pour prédire l'activité cible. Les principaux points forts de cet article sont : la manipulation de données **d'ensembles de données distincts** pour s'entraîner le modèle, l'utilisation des réseaux de neurones récurrents LSTM et d'autre savoir faire en conception logiciel.

**En tant que travaux futurs, nous visons:**

- **augmenter la précision du système en tirant parti de la diversité qui des ensembles de données existant tels que : “ActivityNet” pour obtenir un système de reconnaissance d'activité encore plus robuste.**
- **Un meilleur modèle pourrait être mis en œuvre en améliorant le processus d'extraction de caractéristiques, en faisant varier le nombre de couches, de neurones, du taux d'apprentissage.**
- **Améliorer l'expérience utilisateur de la solution logiciel utilisé**