

Projet de Tailleur Intelligent

NANI CHARLY & DOHA CADNEL

1 Introduction

Le monde de la mode est en constante évolution, et les consommateurs sont de plus en plus à la recherche de produits personnalisés et uniques. C'est dans ce contexte que le shopping en ligne et la couture sur mesure se sont imposés comme de nouvelles tendances, offrant une expérience d'achat inédite et répondant aux attentes d'une clientèle exigeante. La combinaison des deux tendances représentait jusqu'alors un défi technique majeur jusqu'à la publication d'articles sur l'intelligence artificielle proposant des techniques de vision par ordinateur pour prendre les mesures sans contact pour les personnes à partir des modèles de détection des objets.

Dans cette perspective, un tailleur intelligent est un concept innovant qui mêle la tradition de la couture sur mesure avec les dernières avancées en matière de technologie. Il promet une expérience de personnalisation de vêtements sans précédent, en offrant des solutions adaptées à chaque morphologie et à chaque style. Grâce à des scanners 3D ou des logiciels de reconnaissance d'images, le tailleur intelligent peut capturer les mesures du corps de manière extrêmement précise, éliminant ainsi les erreurs liées à la prise de mesures manuelle. Au-delà du client, l'industrie de la vente d'habits en ligne pourra réduire ses pertes liées au retour des vêtements dont les mesures ne respectent pas la morphologie des clients.

Si le concept est adopté d'office par les consommateurs et les entreprises, il n'en demeure pas moins qu'au niveau technologique, il représente un défi important sous plusieurs aspects. En effet, la question des mesures du corps humain suppose la reconnaissance de l'image d'une personne (la personne peut être dans un groupe, la qualité de l'image, la position de la personne...), l'identification de points clés, le calcul des mesures adaptées au domaine de la mode, la conception des patrons, le choix des modèles, le choix des tissus, etc. Dans le présent travail pratique, nous nous focalisons sur l'implémentation d'un agent intelligent permettant d'acquérir des images en 2D d'une personne à partir d'une caméra, de détecter les points clés sur le corps de la personne et de retourner des mesures respectant la taille de la personne en question. Le rapport du groupe est ainsi organisé autour des points suivants : problématique et

objectif du projet, choix techniques réalisés, les fonctionnalités disponibles, les restrictions des choix et les pistes d'améliorations.

2 Problématique et objectifs

2.1 Problématique

La question posée dans le cadre de ce projet est formulée comme suit : Comment permettre à un client de prendre ses mesures de manière précise et autonome à distance, afin de commander un vêtement sur mesure sans avoir à se déplacer ?

2.2 Objectifs

Le projet de tailleur intelligent vise à développer une solution innovante qui permettra de :

- Acquérir des images 2D d'une personne : Développer une application permettant à l'utilisateur de charger une photo existante pour le tailleur intelligent.
- Reconnaître les points clés : Utiliser des algorithmes de vision par ordinateur pour identifier automatiquement les points clés du corps (épaules, taille, hanches, etc.) sur les images.
- Prendre les mesures : Calculer les mesures corporelles pour un ensemble classique à partir des points clés identifiés, sans discrimination sur les spécificités morphologiques masculines et féminines.

Les avantages de cette solution innovante sont évidents. En effet, elle devra allier la précision des mesures à l'adaptation à la morphologie des clients. Le résultat pour l'expérience client est un meilleur confort, une personnalisation accrue, un accès à différents tailleurs sur mesure à travers le monde et un gain de temps.

3 Choix techniques

Les choix techniques sont basés sur les hypothèses suivantes :

- La disponibilité d'une image en 2D lisible pour une seule personne debout avec une seule caméra.
- L'image est présentée avec une vue frontale.
- L'éclairage de l'image est uniforme, la résolution est suffisante et l'arrière-plan est simple.

Plusieurs mesures pourront être prises pour obtenir une gamme de points de données. Par la suite, les mesures obtenues à partir de chaque séance seront comparées et la mesure la plus précise sera sélectionnée pour une utilisation ultérieure.

Les utilisateurs doivent porter des vêtements appropriés ou ajustés pendant la capture pour obtenir des informations plus précises, sinon le jeu des vêtements pourrait être observé comme faisant partie du corps humain. La méthode utilisée est basée sur la détection des points clés. Elle a été proposée par Prasad Dhore, Aparna Pande, Shital Mehta et Saili Sable dans un article publié en novembre 2022 dans la revue Neuquantology volume 20. L'article a fait une revue critique des modèles sur les poses humaines à partir de la vision par ordinateur. La méthode est classée dans la catégorie des méthodes bottom-up. Elle permet de détecter d'abord les points clés dans un premier temps puis de les affecter à une personne dans un second temps.

3.1 Acquisition d'images

Pour l'acquisition d'images, nous avons choisi les bibliothèques Python OpenCV et PIL pour le traitement préliminaire des images.

3.2 Modèle pour le traitement de l'image

Le modèle de détection des poses que nous avons utilisé est basé sur des architectures de réseaux de neurones convolutifs (CNN). Il est conçu pour identifier et localiser les points clés du corps humain. Nous avons opté pour le modèle pré-entraîné de la famille de OpenPose avec Caffe pour l'inférence et l'entraînement des modèles. L'architecture comporte les couches suivantes :

- Couches de convolution pour extraire les caractéristiques visuelles des images.
- Couches de pooling pour réduire la taille des représentations afin de diminuer la complexité du modèle et améliorer la généralisation.
- Couches entièrement connectées pour produire les coordonnées des articulations en sortie.
- Couches de perte pour calculer la différence entre les coordonnées prédites et les coordonnées cibles.

Il faut noter que le modèle CNN multi-stage comporte deux branches. La première branche prédit un ensemble de vecteurs de confiance 2D des emplacements des parties du corps (par exemple, coude, genou, etc.), tandis que la deuxième branche

prédit un ensemble de champs vectoriels 2D d'affinités de pièces, qui codent le degré d'association entre les pièces. Cette deuxième branche permet notamment d'associer l'épaule gauche au genou gauche et à l'œil gauche. La prédiction qui n'est pas associée à l'élément d'une paire est rejetée. Il faut signaler aussi que la base de données d'entraînement est dénommée Coco (Common Object in Context).

Elle comporte un ensemble de données de 200 000 images étiquetées pour reconnaître 17 points clés.

3.3 Prétraitement des Images

Les images sont redimensionnées et normalisées pour être compatibles avec l'entrée du réseau. Cela permet de standardiser les données d'entrée, facilitant ainsi l'apprentissage du modèle.

3.4 Extraction de Caractéristiques

Les réseaux de neurones convolutifs (CNN) sont utilisés pour extraire des caractéristiques visuelles de haut niveau. Ces réseaux sont composés de plusieurs couches :

- Convolution : Applique des filtres pour détecter les motifs visuels tels que les bords, les textures, etc.
- Activation : Utilise des fonctions non linéaires comme ReLU (Rectified Linear Unit) pour ajouter de la non-linéarité et permettre au modèle d'apprendre des représentations complexes.
- Pooling : Réduit la dimensionnalité des cartes de caractéristiques tout en préservant les informations les plus importantes, ce qui aide à contrôler le surapprentissage et à réduire le temps de calcul.

3.5 Détection des Points Clés

La sortie des couches de convolution est utilisée pour prédire les positions des points clés. Cette étape peut impliquer :

- Heatmaps : Génération de cartes de chaleur pour chaque point clé, indiquant la probabilité que le point clé soit présent à chaque emplacement de la carte.
- Argmax : Localisation des positions avec la probabilité maximale sur les cartes de chaleur pour déterminer les coordonnées des points clés.

3.6 Post-traitement

Les points clés détectés peuvent être ajustés et affinés pour une meilleure précision. Des techniques de post-traitement peuvent être utilisées pour s'assurer que les points clés suivent des configurations anatomiquement plausibles.

3.7 Mesures du tailleur

Nous avons choisi certaines mesures clés pour coudre un haut et un pantalon pour un homme ou un haut à manches longues et une jupe longue pour une femme. La distance euclidienne est adoptée comme norme de distance. Ainsi, les mesures suivantes sont retenues à partir des points clés retenus pour la couture : (à compléter)

4 Description des fonctionnalités de l'application

L'application de détection de pose humaine codée offre plusieurs fonctionnalités puissantes pour analyser les poses dans les images. Les fonctionnalités suivantes sont proposées :

4.1 Capture Vidéo et Chargement d'Images

Chargement d'Images : L'application permet de charger une image de test depuis un chemin spécifié. Cette image est utilisée pour effectuer l'inférence et la détection des points clés.

4.2 Prétraitement de l'Image

Conversion en Tenseur : L'image est convertie en tenseur à l'aide de `tf.convert_to_tensor` pour être compatible avec les opérations de TensorFlow.

Ajout d'un Axe de Lot : L'image est redimensionnée pour correspondre aux attentes du modèle, en ajoutant une dimension supplémentaire (souvent appelée batch dimension) via une fonction comme `tf.expand_dims`.

4.3 Inférence du Modèle

Chargement du Modèle TensorFlow : Le modèle TensorFlow pré-entraîné est chargé à partir du chemin spécifié. Exécution de l'Inférence : Le modèle est utilisé pour effectuer l'inférence sur l'image d'entrée. Les sorties du modèle sont converties en tableaux numpy pour une manipulation facile.

4.4 Détection des Points Clés

Extraction des Points Clés et Scores de Confiance : Les points clés détectés et leurs scores de confiance sont extraits des sorties du modèle. Les points clés représentent des parties spécifiques du corps humain (comme les épaules, coudes, genoux, etc.). **Affichage des Points Clés :** Les points clés détectés sont affichés sur l'image avec des cercles verts et des annotations indiquant leur indice.

4.5 Calcul des Mesures Corporelles

Calcul Automatique des Mesures : L'application calcule automatiquement des mesures corporelles telles que la largeur des épaules, la longueur de manche, le tour de poitrine, etc., en utilisant les points clés détectés.

Gestion des Points Absents : Si certains points clés sont absents (par exemple en raison d'une faible confiance), ils sont identifiés et marqués comme absents.

4.6 Affichage des Mesures et des Points Absents

Encadré pour les Mesures : Les mesures calculées sont affichées dans un encadré noir. **Affichage des Points Absents :** Les points absents sont indiqués

4.7 Affichage des Résultats dans la Console

Mesures Calculées : Les mesures calculées sont également affichées dans la sortie de la console pour une lecture facile.

Scores de Confiance : Les scores de confiance pour chaque point clé sont définis par le seuil de confiance ici 0.1.

5 Avantages et limites du modèle

Le modèle proposé comporte les avantages suivants :

- **Précision Elevée :** Les modèles de détection de pose basés sur CNN, comme ceux utilisés dans OpenPose, atteignent des niveaux élevés de précision en détectant les points clés avec une grande fiabilité.
- **Flexibilité :** Ces modèles sont capables de fonctionner dans divers environnements, que ce soit en intérieur ou en extérieur, et sont robustes face à différentes variations d'éclairage, de pose, et d'arrière-plan.
- **Applications Multiples :** Les modèles de détection de pose peuvent être utilisés dans des domaines variés, y compris le sport (analyse de la performance), la médecine (rééducation, suivi des patients), la sécurité (surveillance), et les jeux vidéo (interfaces utilisateur naturelles).

- Scalabilité : Les architectures CNN sont bien adaptées pour être entraînées sur des jeux de données volumineux et peuvent être optimisées pour fonctionner efficacement sur du matériel GPU, ce qui les rend scalables pour des déploiements à grande échelle.
- Interopérabilité : Grâce à des bibliothèques populaires comme TensorFlow et OpenCV, ces modèles peuvent être intégrés facilement dans diverses applications et systèmes, facilitant ainsi leur adoption.