

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

Université de Mohamed El-Bachir El-Ibrahimi - Bordj Bou Arreridj

Institut d'Electronique et des Télécommunications

Département d'électronique

Mémoire

Présenté pour obtenir

LE DIPLOME DE MASTER

FILIERE : Electronique Industriel

Spécialité : Electronique Industriel MCIL

Par

- **KSOURI YOUNES MOHAMED AMINE**
- **HABTA ABDERRAHIM**
- **BOUBAAYA SOHIB**
- **ABID MOHAMED CHEMS EDDINE**

Intitulé

Gant IA pour personnes sourdes et muettes

Soutenu le : /06/2026

Devant le Jury composé de :

<i>Nom & Prénom</i>	<i>Grade</i>	<i>Qualité</i>	<i>Etablissement</i>
<i>Brahimi Abdelhalim</i>	<i>MCA</i>	<i>Président</i>	<i>Univ-BBA</i>
<i>Benameur Ahmed</i>	<i>MAB</i>	<i>Encadreur- 1</i>	<i>Univ-BBA</i>
<i>Saidani Okba</i>	<i>MCB</i>	<i>Encadreur- 2</i>	<i>Univ-BBA</i>
<i>Djendel Mokhtar</i>	<i>MCA</i>	<i>Encadreur -3</i>	<i>Univ-BBA</i>
<i>Belhadad Yehya</i>	<i>MCB</i>	<i>Examineur</i>	<i>Univ-BBA</i>

Année Universitaire 2025/2026

REMERCIEMENT

Nous exprimons tout d'abord notre profonde gratitude à DIEU, le Tout-Puissant, de nous avoir accordé la volonté, la santé et la patience nécessaires pour mener à bien ce travail de recherche.

Nous tenons à exprimer nos plus vifs remerciements à nos encadreurs, Monsieur Ahmed Benameur et Monsieur Saidani Okba, pour avoir accepté de diriger ce mémoire. Leurs précieux conseils, leur orientation rigoureuse, leur soutien constant et leur disponibilité ont été essentiels à la réalisation et à l'aboutissement du projet GESTÉO.

Nos sincères remerciements s'adressent également aux membres du jury, Monsieur Brahimi Abdelhalim (Président) et Monsieur Belhadad Yehya (Examineur), qui nous font le grand honneur d'évaluer ce travail.

Nous souhaitons témoigner notre profonde reconnaissance à l'ensemble du corps professoral du Département d'Électronique, et plus particulièrement aux enseignants de la spécialité Electronique Industriel de l'Université Mohamed El-Bachir El-Ibrahimi de Bordj Bou Arreridj, pour la qualité exceptionnelle de l'enseignement qu'ils nous ont dispensé tout au long de notre cursus universitaire.

Enfin, nous adressons nos pensées les plus chaleureuses à nos familles pour leur soutien moral indéfectible et leurs sacrifices, ainsi qu'à tous nos camarades de la promotion MCIL 2026, et à toutes les personnes qui ont contribué, de près ou de loin, à la réussite de ce projet.

DEDICACE

Avant tout, je remercie Allah le Tout-Puissant de m'avoir accordé la santé, la volonté, la patience et le courage nécessaires pour mener à bien ce modeste travail.

Je dédie ce mémoire :

*À mes très chers parents,
en témoignage de mon profond amour, de mon immense respect et de ma reconnaissance infinie pour tous les sacrifices qu'ils ont consentis pour mon éducation et ma réussite. Que Dieu les protège et les garde.*

*À toute ma famille,
pour leur soutien moral, leurs encouragements et leur présence tout au long de mon parcours universitaire.*

*À mes amis et collègues,
avec qui j'ai partagé les meilleurs moments de cette expérience universitaire.*

*À mes enseignants et encadrants,
pour leurs précieux conseils, leur disponibilité et leurs orientations qui ont grandement contribué à la réalisation de ce travail.*

Enfin, à toutes les personnes qui ont contribué de près ou de loin à l'aboutissement de ce mémoire.

KSOURI YOUNES MOHAMMED AMINE

DEDICACE

Je dédie ce modeste travail avant tout à Allah le Tout-Puissant, qui m'a donné la force, la patience et le courage nécessaires pour mener ce projet à terme.

À mes très chers parents, pour leurs sacrifices, leur soutien inestimable, leurs encouragements constants et leur amour infini. Aucun mot ne saurait exprimer ma profonde gratitude envers eux.

À mes frères et sœurs, pour leur présence, leurs conseils et leur soutien tout au long de mon parcours.

À mes amis (mouniem, sohaib, mahdi, mohamed, farouk, anes, ayoub, youness,...), qui ont partagé avec moi les moments difficiles comme les moments de joie, et qui m'ont toujours encouragé à avancer.

Enfin, à toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce travail.

HABTA ABDERRAHIM

DEDICACE

*Je remercie « ALLAH » le tout puissant qui m'a donné la
foi et la force de mener ce travail à son terme.*

Je remercie notre encadreur pour son aide, ses encouragements.

*A ma mère qui a sacrifié toute sa vie pour nous et à mon père pour
son soutien.*

*Et à tous mes frères qui ont contribué à la force qui m'a amené
jusqu'à ici.*

*Je n'oublie pas mes amis, mes cousins et tous mes camarades de classe tous
ceux qui m'ont aidé de près ou de loin.*

BOUBAAYA SOHIB

DEDICACE

Louange à Allah, dont la grâce et la bénédiction m'ont permis d'accomplir ce travail.

À ceux qui ont été à l'origine de mon parcours,

À ma chère mère, symbole de tendresse et de sacrifice, qui a illuminé mon chemin par ses prières et son amour,

À mon cher père, source de volonté et de persévérance, qui m'a appris la valeur du travail et de la patience,

À ma famille, qui a toujours été mon soutien, et tout particulièrement à mes chères grands-mères, pour leur affection et leur soutien inestimable,

À mes chers amis, qui ont partagé avec moi les moments d'effort et de persévérance, et qui ont été une présence précieuse tout au long de ce parcours,

Je vous dédie ce modeste travail, en signe de reconnaissance, d'amour et de profonde gratitude.

ABID MOHAMED CHEMS EDDINE

Table des matières

Introduction Générale.....	1
Chapitre I : Étude générale et état de l'art	4
I.1 Introduction :	4
I.2 Handicap auditif et troubles de la communication	4
I.2.1 Définitions et classifications cliniques.....	4
I.2.2 Problématique de la communication et barrières sociales.....	6
I.2.3 La Langue des Signes : Structure et Standardisation via l'ASL	7
I.3 Technologies d'assistance existantes :	8
I.3.1 Systèmes basés sur la vision par ordinateur (Approche logicielle)	8
I.3.2 Systèmes embarqués et Gants intelligents (Approche matérielle)	9
I.3.3 Synthèse et analyse comparative	10
I.4 Intelligence artificielle appliquée à la reconnaissance des gestes :	11
I.4.1 Jeux de données pour la langue des signes :	11
I.4.2 Architectures CNN/LSTM et variantes pour la reconnaissance de gestes :	12
I.4.3 Modélisation Cinématique de la Main et Approche Dimensionnelle de l'Informatique Affective :	15
I.5 Synthèse critique et justification des choix techniques :	18
I.5.1 Analyse comparative des approches :	18
I.5.2 Justification du choix technologique.....	18
I.6 Conclusion :	19
Bibliographie Chapitre I.....	20
Chapitre II : Analyse fonctionnelle et architecture du système.....	25
II .1 Introduction	25
II.2.1 Partie acquisition des données	25
II.2.2 Partie traitement et intelligence embarquée.....	27
II.2.3 Interface de communication et restitution vocale	28
II.2.4 Communication sans fil et gestion des données	29
II.2.5 Unité de stockage et Backend : MongoDB.....	29
II.3 Analyse fonctionnelle du système :	30
II.3.1 Diagrammes fonctionnels du système	30
II.3.1.1 Diagramme bête à cornes :	30
II.3.1.2 Diagramme pieuvre :	31
II.3.1.3 Diagramme FAST :	34
II.3.2 Fonctions principales et contraintes du système :	34

Table des matières

II.4 Conception matérielle :	38
II.4.1 Le Microcontrôleur ESP32	38
II.4.2 Capteurs de flexion (Flex Sensors)	39
II.4.3 Module inertiel (MPU-6050)	39
II.4.4 Convertisseur analogique-numérique (ADC) ADS 1115	40
II.4.5 Module PPG (Photoplethysmography) MAX30102:	41
II.4.6 Batterie lithium-ion	42
II.4.7 TP4056:	43
II.5 Technologies de communication et flux de données :	44
II.5.1 Communication locale sans fil via Bluetooth Low Energy (BLE)	44
II.5.2 Connectivité Cloud et communication via Serveur (Wi-Fi)	45
II.6 Partie logicielle :	46
II.6.1 Programmation du microcontrôleur ESP32 via Arduino IDE	46
II.6.2 Gestion des bibliothèques (Flex sensors, MPU-6050 et MAX30102)	47
II.6.3 Algorithme d'acquisition et de prétraitement des signaux	48
II.6.4 Architecture Logicielle et Stratégie de Fusion Multimodale	50
II.6.5 Environnement de développement Unity 3D	51
II.6.6 Intégration du module de traduction et interface utilisateur (UI)	53
II.7 Conclusion :	55
Bibliographie Chapitre II :	56
<i>Chapitre III : Résultats de simulation numérique et interprétation</i>	<i>61</i>
III.1. Introduction	61
III .2. Conception, Simulation et Prototypage Électronique :	61
III .2.1 Simulation sous Proteus	61
III.2.2. Conception et réalisation du circuit imprimé	64
III.3 Application mobile : Gestéo:	71
III.3.1 Langages de Programmation	71
III.3.2 Interface Utilisateur (UI) de l'Application Gestéo	71
III.4. Évaluation et Validation du Système	73
III.4.1. Tests de fonctionnement	73
III.4.2. Retour d'expérience et observations	74
III.5. Conclusion	75
<i>Conclusion Générale.</i>	<i>77</i>

يُشكّل ضعف السمع عائقًا كبيرًا أمام التواصل، مما يؤدي إلى عزلة اجتماعية للصم والبكم. وللتغلب على قصور أنظمة الرؤية الحاسوبية، يُقدّم هذا المشروع تصميم وتطوير قفاز "جيسيتو" الذكي متعدد الوسائط، المُخصّص للترجمة التعاطفية للغة الإشارة إلى نص وكلام في الوقت الفعلي. يعتمد تصميم الجهاز على وحدة تحكم دقيقة ESP32 متصلة بمستشعرات مرنة، ووحدة قياس بالقصور الذاتي (MPU-6050) للحركة المكانية، ومستشعر قياس ضغط الدم الضوئي (MAX30102) لجمع البيانات الفسيولوجية. يستخدم النظام نهجًا هجينًا للذكاء الاصطناعي يمتد من الحافة إلى السحابة: حيث يُوجّه التوجيه الديناميكي البيانات إلى غابة عشوائية للإيماءات الثابتة أو إلى شبكة عصبية عميقة DeepConvLSTM للدinamيكيات الزمنية. وتعمل استراتيجية دمج البيانات المبتكرة، بما في ذلك آلية "قفل الخمول"، على تصفية تشوهات الحركة واستخلاص الشحنة العاطفية من معدل ضربات القلب. علاوة على ذلك، يضمن دمج نموذج لغة Llama 3.1 ترجمةً سياقيةً وطبيعيةً، تُعرض بصريًا وصوتيًا عبر تطبيق جوال تفاعلي تم تطويره باستخدام Unity 3D. وقد أثبتت التجارب التي أُجريت على النموذج الأولي متانة الجهاز، محققًا دقة تعرّف بلغت 96.2% على مجموعة بيانات تضم 16 فئة من الإيماءات. ويحافظ النظام على زمن استجابة إجمالي يبلغ حوالي 1.5 ثانية، مما يضمن سلاسة التواصل، إلى جانب عمر بطارية يُقدّر بـ 4.5 ساعات. ومن خلال الجمع بين الذكاء الاصطناعي التسلسلي والحوسبة العاطفية، يتجاوز هذا المشروع الترجمة الحرفية ليقدّم تفاعلًا تعاطفيًا حقيقيًا، مما يُسهم بشكل كبير في الإدماج الاجتماعي والمهني للأفراد ذوي الإعاقة السمعية، ويفتح آفاقًا جديدةً لواجهات التفاعل بين الإنسان والحاسوب.

الكلمات المفتاحية: قفاز ذكي، لغة الإشارة، DeepConvLSTM، الحوسبة العاطفية، ESP32، Llama 3.1، إنترنت الأشياء (IoT).

Résumé

Le handicap auditif engendre des barrières de communication majeures, isolant socialement les personnes sourdes et muettes. Pour surmonter les limites des systèmes de vision artificielle, ce projet présente la conception et la réalisation de GESTÉO, un gant intelligent multimodal dédié à la traduction empathique de la langue des signes en texte et en parole en temps réel. L'architecture matérielle s'articule autour d'un microcontrôleur ESP32 interfacé avec des capteurs de flexion, une centrale inertielle (MPU-6050) pour la cinématique spatiale, et un capteur photopléthysmographique (MAX30102) pour l'acquisition physiologique. Le système déploie une approche d'intelligence artificielle hybride Edge-to-Cloud : un routage dynamique dirige les données vers une Forêt Aléatoire (Random Forest) pour les gestes statiques ou un réseau profond DeepConvLSTM pour la dynamique temporelle. Une stratégie novatrice de fusion de données, incluant un mécanisme « Idle Lock », permet de filtrer les artefacts de mouvement et d'extraire la charge émotionnelle à partir du rythme cardiaque. De plus, l'intégration du modèle de langage Llama 3.1 assure une traduction contextuelle et naturelle, restituée visuellement et vocalement via une application mobile immersive développée sous Unity 3D. Les expérimentations menées sur le prototype fabriqué valident la robustesse du dispositif, atteignant une précision de reconnaissance de 96,2 % sur un ensemble de données de 16 classes gestuelles. Le système maintient une latence globale d'environ 1,5 seconde, garantissant une communication fluide, couplée à une autonomie énergétique estimée à 4,5 heures. En combinant l'intelligence artificielle séquentielle et l'informatique affective, ce projet dépasse la traduction littérale pour offrir une interaction véritablement empathique, contribuant ainsi significativement à l'inclusion socio-professionnelle des personnes malentendantes et ouvrant de nouvelles perspectives pour les interfaces homme-machine.

Mots-clés : Gant intelligent, Langue des signes, DeepConvLSTM, Informatique affective, ESP32, Llama 3.1, Internet des Objets (IoT).

Abstract

Hearing impairment creates major communication barriers, socially isolating deaf and mute individuals. To overcome the limitations of computer vision systems, this project presents the design and development of GESTÉO, a multimodal smart glove dedicated to the empathetic translation of sign language into text and speech in real time. The hardware architecture is based on an ESP32 microcontroller interfaced with flex sensors, an inertial measurement unit (MPU-6050) for spatial kinematics, and a photoplethysmographic sensor (MAX30102) for physiological data acquisition. The system deploys a hybrid edge-to-cloud artificial intelligence approach: dynamic routing directs data to a Random Forest for static gestures or a DeepConvLSTM deep network for temporal dynamics. An innovative data fusion strategy, including an "Idle Lock" mechanism, filters out motion artifacts and extracts emotional charge from heart rate. Furthermore, the integration of the Llama 3.1 language model ensures contextual and natural translation, rendered visually and vocally via an immersive mobile application developed using Unity 3D. Experiments conducted on the prototype validate the robustness of the device, achieving a recognition accuracy of 96.2% on a dataset of 16 gesture classes. The system maintains an overall latency of approximately 1.5 seconds, guaranteeing smooth communication, coupled with an estimated battery life of 4.5 hours. By combining sequential artificial intelligence and affective computing, this project goes beyond literal translation to offer truly empathetic interaction, thus significantly contributing to the socio-professional inclusion of hearing-impaired individuals and opening new perspectives for human-computer interfaces.

Keywords : Smart glove, Sign language, DeepConvLSTM, Affective computing, ESP32, Llama 3.1, Internet of Things (IoT).

Liste des Symboles et Grandeurs Physiques

$\mathbf{a}_{x,,}$:	Vecteur d'accélération linéaire sur les trois axes spatiaux (x, y, z).
$\mathbf{g}_{x,,}$:	Vecteur de rotation angulaire (gyroscope) sur les axes spatiaux (x, y, z).
f_s :	Fréquence d'échantillonnage (Sampling frequency).
$C_{batterie}$:	Capacité nominale de la batterie.
I_{moyen} :	Consommation de courant moyenne du système en mode actif.
$T_{autonomie}$:	Autonomie théorique du système (en heures).
$M_{magnitude}$:	Norme euclidienne du vecteur d'accélération.
$L_{End-to-End}$:	Latence totale de bout en bout du système.
t_{acq} :	Temps d'acquisition et de filtrage (microcontrôleur ESP32)
t_{net} :	Temps de transmission réseau (Wi-Fi/MQTT)
t_{inf} :	Temps d'inférence Cloud (modèle DeepConvLSTM)
t_{rendu} :	Temps de rendu visuel de l'animation sur l'avatar (Unity 3D)
dB :	Décibel (Unité de mesure du niveau sonore/perte auditive)
g :	Accélération gravitationnelle (utilisée comme seuil de mouvement, ex : 0.35g)
Hz :	Hertz (Unité de fréquence, ex : 50 Hz)
$K\Omega$:	Kilo-Ohm (Unité de résistance électrique des capteurs de flexion)
mA :	Milliampère (Unité d'intensité du courant électrique)
mAh :	Milliampère-heure (Unité de capacité de charge électrique)
mN :	Millinewton (Unité de force mécanique des actionneurs haptiques)
ms :	Milliseconde (Unité de temps/latence)
Pa :	Pascal (Unité de pression, cartographie haptique)
V :	Volt (Unité de tension électrique, ex : 3.3V, 3.7V, 5V)

Liste des abréviations

ADC	Analog-to-Digital Converter (Convertisseur analogique-numérique)
APK	Android Package (Format de fichier pour application Android)
ASL	American Sign Language (Langue des Signes Américaine)
BIAP	Bureau International d'Audiologie
BLE	Bluetooth Low Energy
BMS	Battery Management System (Système de gestion de batterie)
BPM	Beats Per Minute (Battements par minute / Fréquence cardiaque)
BSON	Binary JSON
CNN	Convolutional Neural Network (Réseau de neurones convolutifs)
Conv1D	Convolution unidimensionnelle (1D Convolution)
DMP	Digital Motion Processor (Processeur de mouvement numérique)
FAST	Function Analysis System Technique (Outil d'analyse fonctionnelle)
FP	Fonction Principale (Analyse fonctionnelle)
GATT	Generic Attribute Profile
GPT	Generative Pre-trained Transformer
I2C	Inter-Integrated Circuit (Protocole de communication série)
IA / AI	Intelligence Artificielle / Artificial Intelligence
IDE	Integrated Development Environment (Environnement de développement intégré)
IMU	Inertial Measurement Unit (Unité de mesure inertielle)
IoT	Internet of Things (Internet des Objets)
JSON	JavaScript Object Notation
LDO	Low Dropout Regulator (Régulateur de tension à faible chute)
Li-Po	Lithium-Polymère
LLM	Large Language Model (Grand Modèle de Langage)
LMS	Least Mean Squares (Algorithme des moindres carrés moyens)
LSTM	Long Short-Term Memory (Réseau de neurones à mémoire à court et long terme)
MA	Motion Artifacts (Artefacts de mouvement)
MS-ASL	Microsoft American Sign Language (Base de données)
PCB	Printed Circuit Board (Circuit imprimé)

Liste des abréviations

PMC	Perceptron Multicouche
PPG	Photopléthysmographie
PWM	Pulse Width Modulation (Modulation de largeur d'impulsion)
REST	Representational State Transfer
RNN	Recurrent Neural Network (Réseau de neurones récurrents)
RVB	Rouge, Vert, Bleu (Espace colorimétrique)
SCL	Serial Clock (Ligne d'horloge du bus I2C)
SDA	Serial Data (Ligne de données du bus I2C)
TTS	Text-to-Speech (Synthèse vocale)
UI	User Interface (Interface Utilisateur)
USB	Universal Serial Bus
WBAN	Wireless Body Area Networks (Réseaux corporels sans fil)
WLASL	Word-Level American Sign Language (Base de données)

CHAPITRE I : Étude générale et état de l'art

Figure I.1. Surdit   et perte auditive.

Figure I.2. Langue Des Signes.

Figure I.3. Pipeline d'un syst  me de vision par ordinateur, approche logicielle, adapt   de Rishabh Software.

Figure I.4. Architecture mat  rielle d'un gant intelligent.

Figure I.4. Exemples d'images d  tect  es.

Figure I.6. Architecture hybride CNN 1D – LSTM.

Figure I.7. Mod  lisation cin  matique de la main illustrant les articulations et les degr  s de libert   captur  s par les capteurs de flexion et l'unit   IMU.

CHAPITRE II : Analyse fonctionnelle et architecture du syst  me

Figure II.1. Architecture du syst  me : De la gestuelle analogique    la parole num  rique

Figure II.2. Organigramme fonctionnel.

Figure II.3. Architecture d'un gant intelligent bas   sur capteurs de flexion et MPU-6050.

Figure II.4. ESP 32 wroom 32 WI-FI + BLuetooth Antenna.

Figure II.5. MongoDB JSON Example Format.

Figure II.6. Traduction multimodale de la langue des signes (Sign-to-Text et Sign-to-Speech).

Figure II.7. Diagramme b  te    cornes pour le gant intelligent.

Figure II.8. Diagramme pieuvre du gant intelligent.

Figure II.9. Diagramme FAST du syst  me de gant intelligent.

Figure II.10. ESP32-WROOM-32 PINS.

Figure II.11. Comment fonctionne un capteur de flexion ?

Figure II.12. FLEX SENSORS.

Liste des Figures

Figure II.13 .MPU -6050.

Figure II.14.ADS 1115.

Figure II.15.MAX30102.

Figure II.16. Batterie lithium-ion1000 mAh.

Figure II.17. Circuit intégré de gestion de charge TP4056.

Figure II.18. Avantages du Bluetooth Low Energy.

Figure II.19. Aperçu du modèle proposé.

Figure II.20. Vue générale de l'environnement Arduino IDE.

Figure II.21. Schéma de conditionnement du signal.

Figure II.22. Architecture Hybride de Fusion Multimodale et Intelligence Affective pour la Traduction Gestuelle Contextuelle.

Figure II.23. Interface de l'environnement de développement Unity 3D.

Figure II.24. Capture d'écran de l'interface utilisateur (UI) de l'application mobile développée sous Unity, illustrant la zone de traduction textuelle, l'avatar 3D et le retour physiologique.

CHAPITRE III : Résultats de simulation numérique et interprétation

Figure III.1. Présentation de l'interface d'ISIS PROTEUS.

Figure III.2 .Schéma de simulation globale du circuit d'acquisition et de contrôle sous l'environnement Proteus (ISIS).

Figure III.3. Test des flex et les resultat au virtual teminal.

Figure III.4. Routage et placement des composants du circuit imprimé (PCB) de la carte "GESTEO" sous l'environnement Proteus ARES.

Figure III.5. Étape d'insolation de la plaque PCB sous lumière UV.

Figure III.6. Révélation du circuit après exposition.

Figure III.7. Gravure chimique du PCB dans le bain de perchlorure de fer.

Figure III.8 .Nettoyage et élimination du vernis après gravure.

Liste des Figures

Figure III.9. Perçage des trous du circuit imprimé avant montage.

Figure III.10. Processus de soudure des pistes en cuivre (à gauche) et assemblage final de la carte de contrôle principale avec l'ESP32 (à droite).

Figure III.11. Intégration matérielle du prototype : fixation des capteurs de flexion sur le textile et câblage de l'ensemble vers l'unité centrale.

Figure III.12. Prototype final du gant intelligent IoT et son application mobile de traduction de la langue des signes en temps réel via une communication Bluetooth avec l'ESP32.

Figure III.13. Gestéo apk : (a) Calibration IoT, (b) Contexte émotionnel, (c) Avatar 3D et voix.

Figure III.14. Dispositif expérimental lors des tests de fonctionnement du système.

Figure III.15. Évaluation de l'ergonomie et de l'utilisabilité du système GESTÉO lors des sessions de test avec les utilisateurs.

CHAPITRE I : Étude générale et état de l'art

Tableau I.1. Synthèse des barrières communicationnelles et de leurs impacts sociaux.

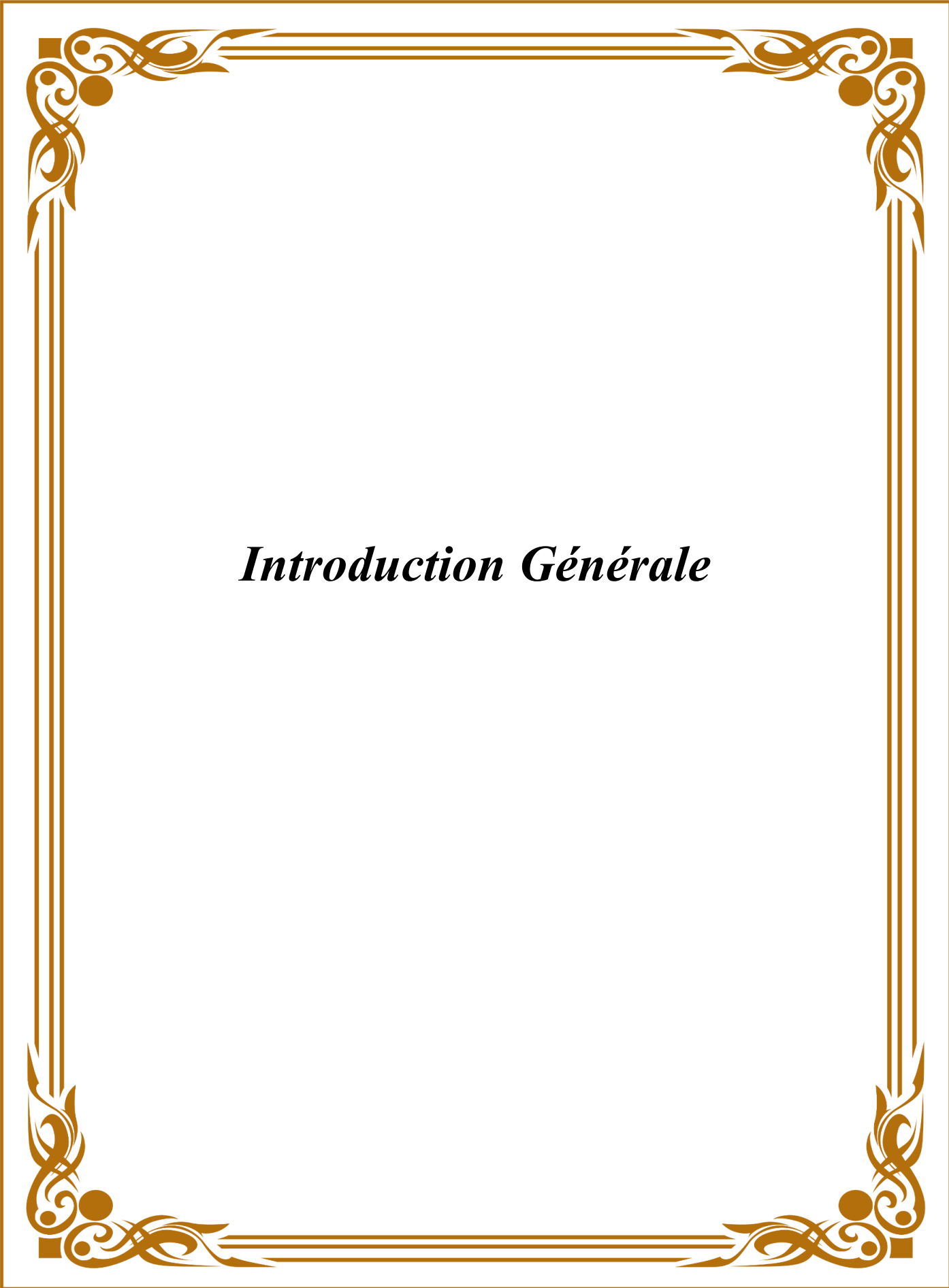
Tableau I.2. Analyse comparative des technologies de reconnaissance gestuelle.

Tableau I.3. Analyse comparative entre la vision par ordinateur et les capteurs portables pour la reconnaissance de gestes.

CHAPITRE III : Résultats de simulation numérique et interprétation

Tableau III.1. Extrait du Dataset brut stocké sous MongoDB et résultat du nettoyage par filtre d'intégrité (Pandas. dropna()).

Tableau III.2 .Ecosystème logiciel et frameworks de développement du projet Gestéo.



Introduction Générale

Introduction Générale

La communication est un aspect fondamental de la vie humaine, constituant le principal moyen d'échange d'informations, d'expression des émotions, d'établissement de relations et de participation à la vie sociale. Au-delà de sa fonction pratique, la communication joue un rôle crucial dans le bien-être psychologique, l'intégration sociale et le développement personnel. L'incapacité à communiquer efficacement peut entraîner l'isolement social, la frustration, une perte de confiance en soi et un accès limité à l'éducation, à l'emploi, aux soins de santé et aux services publics. Parmi les différents groupes touchés par les barrières de communication, les personnes malentendantes rencontrent des difficultés importantes dans leurs interactions quotidiennes. Selon de récentes estimations mondiales de la santé, des centaines de millions de personnes dans le monde souffrent d'une perte auditive, et nombre d'entre elles utilisent la langue des signes comme principal moyen de communication. La langue des signes n'est pas un simple ensemble de gestes ; c'est un système linguistique complet et structuré qui combine la forme des mains, la configuration des doigts, leur orientation, les mouvements, la posture corporelle et les expressions faciales pour transmettre des informations et des émotions. Sa richesse linguistique permet aux personnes sourdes de communiquer efficacement au sein de leur communauté tout en préservant leur identité culturelle. Malgré son importance, la langue des signes reste largement méconnue du grand public. Seule une faible proportion de personnes entendant maîtrise suffisamment la langue des signes pour pouvoir dialoguer de manière significative avec des personnes sourdes. Cette disparité crée un fossé de communication persistant entre les communautés sourdes et entendantes, limitant la participation sociale et réduisant les possibilités d'inclusion. Par conséquent, le besoin de technologies d'assistance intelligentes capables de traduire automatiquement la langue des signes en un langage facilement compréhensible par le grand public est croissant. Les progrès récents en intelligence artificielle, en systèmes embarqués, en technologies portables et en interaction homme-machine ont considérablement accéléré le développement des systèmes de reconnaissance de la langue des signes (SLR). Ces systèmes visent à interpréter automatiquement les gestes de la langue des signes et à les convertir en langage écrit ou parlé, facilitant ainsi la communication entre les personnes sourdes et entendantes. De ce fait, la technologie SLR s'est imposée comme l'un des domaines les plus prometteurs des technologies d'assistance, offrant de nouvelles perspectives d'inclusion sociale et d'accessibilité. Les approches existantes de reconnaissance gestuelle peuvent généralement être classées en trois grandes catégories : les systèmes basés sur la vision, les systèmes utilisant des gants et les systèmes hybrides. Les approches basées sur la vision s'appuient sur des caméras et des algorithmes de vision par ordinateur pour capturer et interpréter les gestes. Bien que ces systèmes soient non invasifs et faciles d'utilisation, ils sont très sensibles aux

conditions environnementales telles que les variations d'éclairage, la complexité de l'arrière-plan et les occlusions de la main. De plus, ils nécessitent souvent d'importantes ressources de calcul pour le traitement d'images et l'extraction de caractéristiques. À l'inverse, les systèmes à base de gants utilisent des capteurs portables pour mesurer directement la flexion des doigts, l'orientation de la main et la dynamique des mouvements. Grâce à l'utilisation de capteurs de flexion, d'unités de mesure inertielle (UMI) et d'autres capteurs intégrés, ces systèmes offrent une grande précision de reconnaissance, des temps de réponse rapides et une mobilité accrue. De plus, ils éliminent de nombreuses difficultés de prétraitement liées aux méthodes basées sur la vision. Cependant, les gants intelligents classiques se concentrent souvent uniquement sur la reconnaissance gestuelle et intègrent rarement des informations contextuelles ou émotionnelles susceptibles d'améliorer la qualité et le naturel de la communication. Les systèmes hybrides tentent de combiner les avantages des approches basées sur la vision et sur les capteurs en intégrant les données de la caméra aux mesures des capteurs portables. Bien que ces systèmes puissent atteindre des performances de reconnaissance améliorées, ils souffrent généralement d'une complexité accrue, de coûts plus élevés et de besoins de calcul plus importants, ce qui limite leur utilité au quotidien.

Dans ce contexte, ce travail propose la conception et le développement de GESTÉO, un gant multimodal dédié à la traduction en langue des signes en temps réel. Le système proposé intègre plusieurs modalités de détection, notamment des capteurs de flexion pour l'acquisition des mouvements des doigts, une centrale inertielle pour l'orientation et le suivi des mouvements de la main, et un capteur physiologique pour la surveillance de l'état affectif de l'utilisateur. En combinant ces sources de données hétérogènes grâce à des techniques avancées de fusion de capteurs, le système vise une reconnaissance gestuelle robuste et précise, même en conditions dynamiques. Pour améliorer la qualité de la traduction, les données multimodales acquises sont traitées par des architectures d'apprentissage profond avancées, capables de capturer les caractéristiques spatiales et temporelles des gestes. De plus, des modèles de langage complexes sont utilisés pour transformer les signes reconnus en phrases cohérentes, contextualisées et linguistiquement naturelles. Contrairement aux traducteurs de langue des signes classiques qui effectuent une conversion directe mot à mot, l'approche proposée vise à générer une communication plus expressive et empathique en prenant en compte à la fois la signification sémantique des gestes et l'état physiologique de l'utilisateur. L'objectif principal de ce mémoire est donc de développer un système d'assistance portable, autonome et intelligent, capable de

faciliter une interaction fluide entre les personnes sourdes et entendants. Au-delà de la traduction gestuelle, GESTÉO vise à offrir une expérience utilisateur améliorée grâce à la synthèse vocale en temps réel, une interface mobile interactive et des capacités de communication sensibles aux émotions. En intégrant des technologies de détection portables, l'intelligence artificielle, l'informatique affective et la génération de langage naturel, le système proposé contribue à une société plus inclusive où les barrières de communication sont considérablement réduites, améliorant ainsi l'autonomie, l'accessibilité et la participation sociale des personnes malentendantes.

Ce mémoire est structuré de la façon suivante :

Chapitre I : Une présentation de l'état de l'art des systèmes d'assistance pour le handicap auditif est proposée. Cette présentation comporte une analyse comparative des technologies de reconnaissance gestuelle existantes et des fondements théoriques liés à la langue des signes et à l'intelligence artificielle.

Chapitre II : Nous aborderons en détail la conception technologique et l'architecture globale du gant intelligent. Cela comprendra l'analyse fonctionnelle du système (diagrammes FAST, Pieuvre), le choix justifié des composants matériels (capteurs, microcontrôleur ESP32) et logiciels (environnement Unity 3D, algorithmes de fusion), ainsi que l'architecture de communication hybride Edge-to-Cloud mise en œuvre.

Chapitre III : Nous aborderons en détail la réalisation matérielle, logicielle et l'évaluation du prototype GESTÉO. Cela comprendra la simulation et la fabrication du circuit imprimé (PCB), le développement de la chaîne d'intelligence artificielle (modèle hybride Random Forest/DeepConvLSTM, informatique affective et LLM Llama 3.1), la conception de l'application mobile bidirectionnelle sous Unity 3D, ainsi que les tests validant la latence et l'autonomie énergétique du système.

*Chapitre I : Étude générale et état de
l'art*

I.1 Introduction :

Le handicap auditif représente l'une des problématiques majeures de santé publique, impactant non seulement l'intégrité sensorielle de l'individu, mais également son insertion socio-professionnelle. L'incapacité de communiquer aisément avec une majorité entendant crée une barrière profonde, rendant les interactions quotidiennes particulièrement difficiles dans les milieux éducatifs, sociaux ou professionnels.

Dans ce contexte, le développement de solutions technologiques autonomes capables de fournir une traduction en temps réel apparaît comme une nécessité pour favoriser l'inclusion. Ce premier chapitre pose les fondements théoriques et technologiques indispensables à la conception d'un système d'assistance pour les personnes malentendantes. Il débute par une définition des classifications cliniques de la surdité et une analyse des barrières sociales qu'elle engendre. L'étude se poursuit en mettant en évidence la structure linguistique complexe de la langue des signes, soulignant l'enjeu technique de sa traduction automatique. Par la suite, le chapitre propose une revue critique des technologies d'assistance existantes, opposant les systèmes basés sur la vision par ordinateur aux approches matérielles utilisant des gants intelligents embarqués. Le rôle fondamental du retour haptique y est également exploré en tant que canal de substitution sensorielle favorisant une communication bidirectionnelle. Enfin, ce chapitre détaille l'apport incontournable de l'intelligence artificielle via les architectures hybrides CNN/LSTM, l'informatique affective et l'émergence des Grands Modèles de Langage (LLMs) afin d'établir et de justifier le choix architectural matériel et logiciel qui sera développé dans les chapitres suivants.

I.2 Handicap auditif et troubles de la communication

Le handicap auditif représente l'une des problématiques majeures de santé publique, impactant non seulement l'intégrité sensorielle de l'individu, mais également son insertion socio-professionnelle. Cette section détaille les classifications cliniques et analyse les barrières communicationnelles justifiant le recours aux technologies d'assistance.

I.2.1 Définitions et classifications cliniques

On dit qu'une personne n'entend pas aussi bien qu'une personne ayant une audition normale, c'est-à-dire qu'elle ne perçoit pas un niveau sonore égal ou supérieur à 25 décibels dans les deux oreilles, souffre de perte auditive. Le degré et la localisation de cette perte auditive

varient d'une personne à l'autre, affectant sa capacité à comprendre la parole lors des conversations et la rendant plus sensible aux bruits forts. Il existe une distinction fondamentale entre les différents degrés de perte auditive. Il existe des personnes souffrant de perte auditive, légère à sévère. Ces personnes sont généralement capables de communiquer verbalement grâce à des appareils auditifs placés dans l'oreille atteinte [1]. En revanche, la catégorie la plus touchée par les barrières de communication est celle des personnes sourdes. Certaines personnes sont sourdes, c'est-à-dire qu'elles souffrent d'une perte auditive sévère ou totale. Ces personnes sont souvent candidates à la chirurgie d'implantation cochléaire et peuvent communiquer en langue des signes. Pour uniformiser les diagnostics, le Bureau International d'Audiologie (BIAP) a établi une classification standardisée fondée sur la perte auditive moyenne en décibels (dB) aux fréquences conversationnelles. Dans le cas de la surdité profonde (perte supérieure à 90 dB), « l'enfant n'entend plus la parole, il ne perçoit que les bruits très puissants. La communication gestuelle devient alors indispensable. Cette classification est le point de départ de toute recherche visant à concevoir des interfaces homme-machine dédiées à la traduction des signes [2].

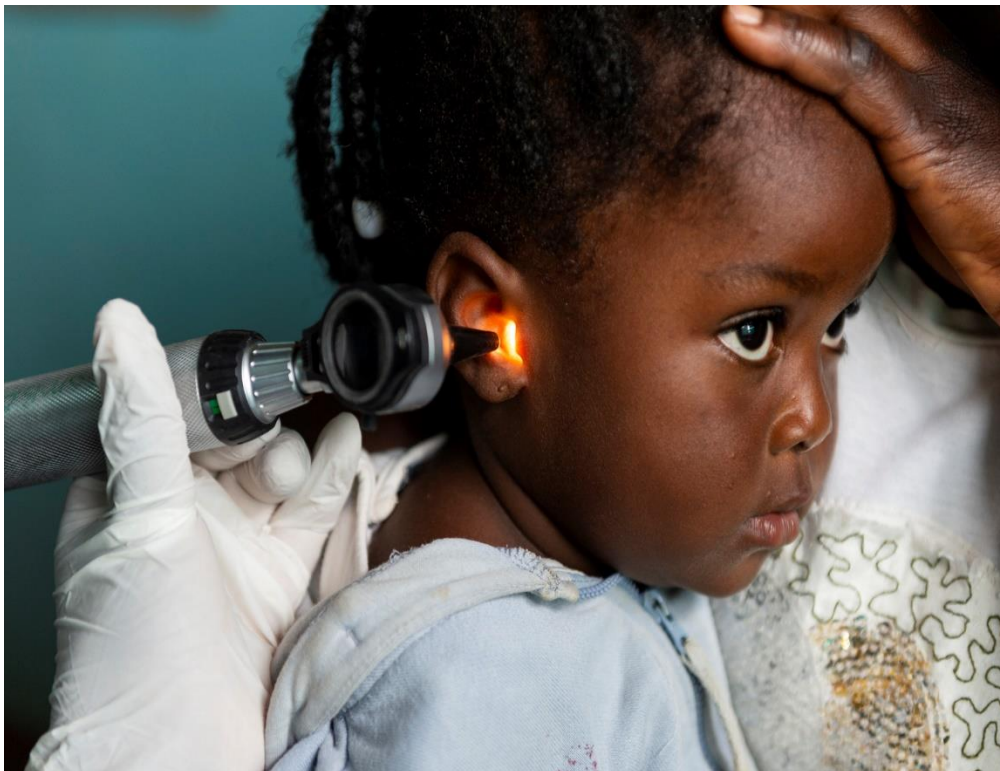


Figure I.1. Surdit  et perte auditive [1].

I.2.2 Problématique de la communication et barrières sociales

La parole est l'interface par défaut du monde. Pour les personnes souffrant de troubles de l'élocution ou de l'audition, l'incapacité de communiquer avec la majorité crée une barrière profonde. « La dépendance à l'égard de traducteurs humains restreint l'indépendance » [3]. Ce "fossé de communication" est d'autant plus critique que la majorité de la population n'est pas familière avec la langue des signes, rendant les interactions quotidiennes difficiles dans les milieux éducatifs, sociaux ou professionnels [4]. Actuellement la traduction de la langue des signes est considérée comme l'une des études les plus importantes car elle s'adresse à un groupe très sensible de la société. « Ce type de traduction nécessite des traducteurs spécialisés ayant un haut niveau de formation en langue des signes pour faciliter l'intégration des sourds et muets dans les divers domaines de la vie ». Cependant, le coût et la disponibilité limitée des interprètes humains soulignent la nécessité de développer des solutions portables et autonomes capables de fournir une traduction en temps réel [5]. Outre la barrière linguistique inhérente à la communication gestuelle, l'intégration d'un système de surveillance physiologique basé sur la photo pléthysmographie (PPG), tel que le capteur MAX30102, dans un gant intelligent de traduction gestuelle soulève une problématique scientifique majeure liée aux artefacts de mouvement (Motion Artifacts). En effet, les gestes de la main impliquent des mouvements complexes des muscles et des tendons qui modifient la géométrie artérielle et perturbent fortement le signal sanguin acquis par les capteurs optiques. Ces perturbations, amplifiées par les variations de pression, les accélérations rapides et les changements de position de la main, rendent difficile l'extraction fiable des paramètres physiologiques en temps réel. La problématique centrale de ce travail consiste donc à concevoir une architecture matérielle et logicielle capable d'assurer simultanément une reconnaissance précise des gestes et une estimation robuste de l'état d'activation physiologique (Arousal), tout en compensant les interférences générées par les mouvements inhérents à la traduction gestuelle [47].

Domaine de la vie quotidienne	Nature de la barrière de communication	Conséquence et Impact social
Milieu Éducatif	Absence de professeurs maîtrisant la langue des signes et manque d'outils pédagogiques traduits.	Retard d'apprentissage et difficulté d'accès aux études supérieures.
Milieu Professionnel	Incompréhension lors des entretiens d'embauche et incapacité de participer aux réunions d'équipe.	Taux de chômage élevé, sous-emploi et manque d'évolution de carrière.
Santé et Administration	Manque d'interprètes certifiés dans les hôpitaux ou les mairies.	Dépendance totale envers la famille, risques d'erreurs médicales.
Vie Sociale et Publique	L'interface par défaut de la société est la "parole" ; la majorité de la population ne signe pas.	Isolement psychologique, baisse de l'estime de soi et sentiment d'exclusion.

Tableau I.1. Synthèse des barrières communicationnelles et de leurs impacts sociaux [3-5].

I.2.3 La Langue des Signes : Structure et Standardisation via l'ASL

La langue des signes n'est pas une simple pantomime, mais une langue structurée. « C'est une forme de communication où les formes, l'orientation et le mouvement des bras, du corps, des mains et les expressions faciales sont systématiquement combinés » [4]. Dans le domaine de la recherche en reconnaissance gestuelle, les travaux se concentrent majoritairement sur la langue des signes américaine (ASL), qui s'est imposée comme un standard d'évaluation grâce à la disponibilité de bases de données de référence massives telles que WLASL et MS-ASL. L'enjeu technique consiste à traduire cette structure visuelle complexe en texte ou en voix synthétique. Le but est d'implémenter un système capable de reconnaître et d'interpréter ces signes, afin de les traduire en mots compris par les personnes ayant une faible connaissance de cette langue [6]. « Le système de reconnaissance des gestes a reçu beaucoup d'attention dans les quelques dernières années en raison de la multiplicité de ses applications et la capacité d'interagir avec la machine de manière efficace grâce à l'interaction homme-machine ». Ce

besoin d'interaction fluide motive la conception de dispositifs embarqués comme le gant intelligent, objet de ce projet [7].

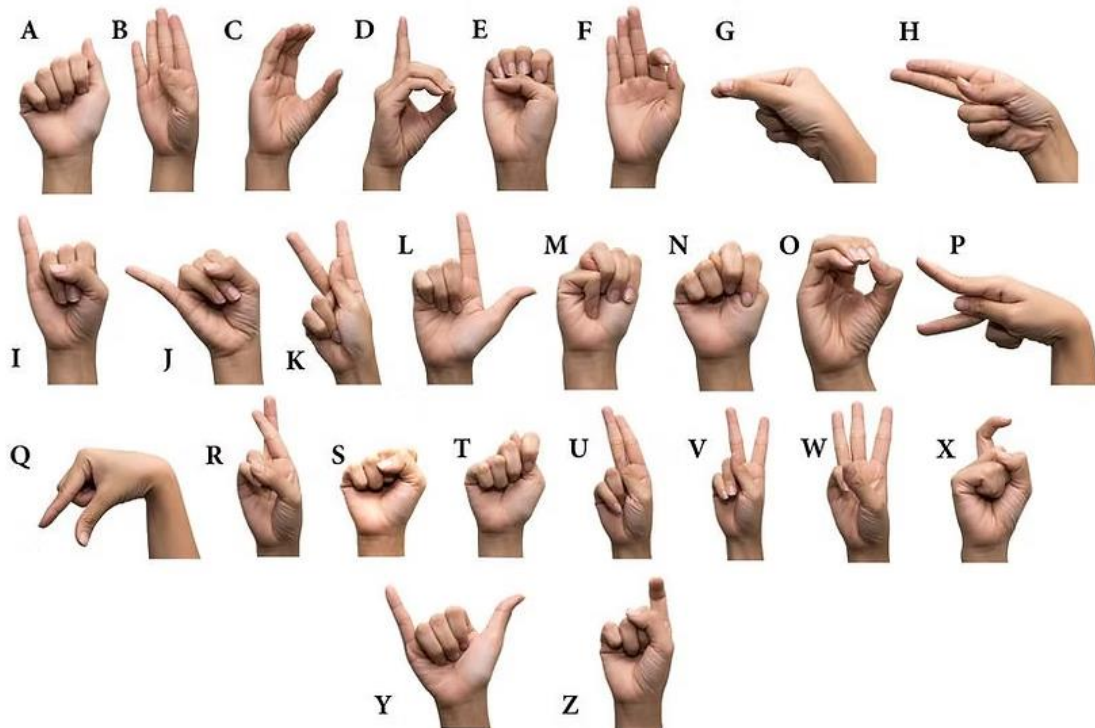


Figure I.2. Langue Des Signes [8].

I.3 Technologies d'assistance existantes :

L'évolution de l'électronique embarquée et de l'intelligence artificielle a permis l'émergence de plusieurs dispositifs visant à combler le fossé de communication entre les sourds et les entendants. Dans la littérature scientifique, les travaux de recherche se divisent principalement en deux grandes approches technologiques : la vision par ordinateur (caméras) et les capteurs portables (gants intelligents).

I.3.1 Systèmes basés sur la vision par ordinateur (Approche logicielle)

L'approche par vision artificielle est historiquement la plus explorée. Elle consiste à placer l'utilisateur face à une caméra (webcam ou smartphone) qui capture ses mouvements. Les images sont ensuite traitées par des algorithmes d'extraction de caractéristiques. Dans le contexte de la langue des signes, plusieurs solutions logicielles ont été proposées. Généralement, « notre système est basé sur le développement de deux méthodes de classification : les réseaux de neurones (PMC), et les réseaux de neurones convolutifs (CNN) ». Bien que ces

réseaux de neurones offrent une excellente capacité à identifier les formes des mains, cette approche présente des limites majeures pour un usage quotidien. En effet, la reconnaissance par caméra est extrêmement sensible aux variations d'éclairage, nécessite que l'utilisateur reste immobile dans le champ de vision de la caméra, et pose des problèmes fréquents d'occlusion (lorsqu'une main cache l'autre ou cache le visage) [9].

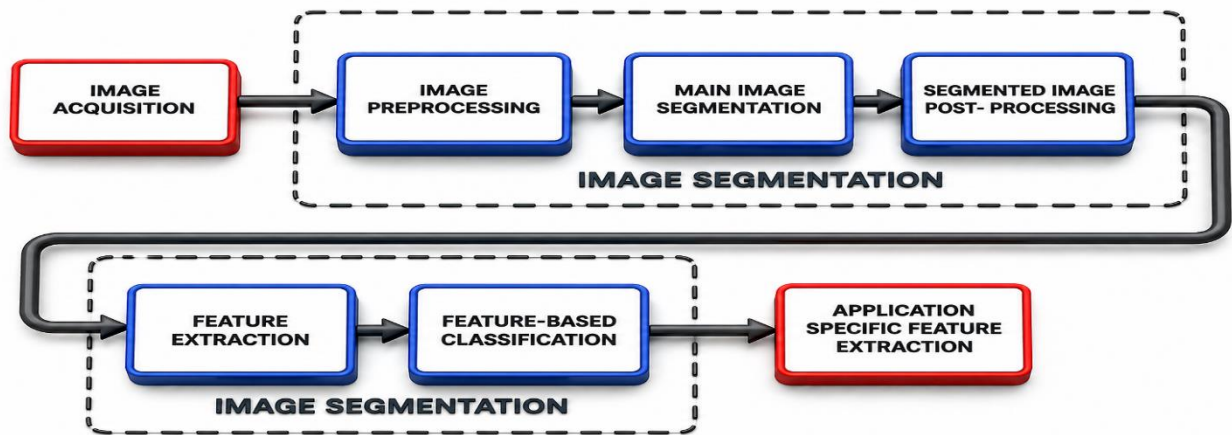


Figure I.3. Pipeline d'un système de vision par ordinateur, approche logicielle, adapté de Rishabh Software [13].

I.3.2 Systèmes embarqués et Gants intelligents (Approche matérielle)

Pour surmonter les contraintes de mobilité et d'environnement liées à la vision par ordinateur, la recherche s'est orientée vers la mécatronique et les systèmes portables (Wearables). L'objectif est de « fournir une solution portable et intelligente pour répondre à la demande de conversion des gestes des personnes handicapées en un format de synthèse de texte et de parole » [10]. Cette solution prend la forme d'un gant intelligent instrumenté. Contrairement à la caméra qui "regarde" le geste, le gant "mesure" physiquement le geste. Techniquement, ce type de système « utilise une combinaison de capteurs de flexion pour la détection des mouvements des doigts, un module qui combine accéléromètre et gyroscope pour la capture des gestes de la main, et un microcontrôleur pour le traitement des données » [11]. L'avantage de cette architecture matérielle est sa précision cinématique. Le système d'acquisition « convertit les changements de résistance physique et les chutes de tension en chaînes de texte lisibles ». Cette méthode permet une traduction continue, indépendante de l'éclairage extérieur, protégeant la

vie privée (aucune image n'est enregistrée) et offrant la possibilité d'intégrer un retour haptique (vibrations) pour créer une communication véritablement bidirectionnelle [12].



Figure I.4. Architecture matérielle d'un gant intelligent [12].

I.3.3 Synthèse et analyse comparative

Afin de justifier le choix architectural de notre projet (développé dans le Chapitre 2), le Tab1.2 synthétise les différences fondamentales entre les deux approches étudiées.

Critère d'évaluation	Vision par ordinateur (Caméras)	Capteurs portables (Gants intelligents)
Indépendance à l'environnement	Faible (très sensible à la lumière et au fond)	Forte (fonctionne dans l'obscurité totale)
Mobilité (Portabilité)	Limitée (nécessite d'être face à l'objectif)	Élevée (système embarqué sur l'utilisateur)
Problème d'occlusion	Fréquent (chevauchement des mains)	Inexistant (mesure directe sur chaque articulation)
Confidentialité et vie privée	Faible (capture du visage et du lieu)	Forte (ne capture que des données inertielles)
Possibilité de Retour Haptique	Impossible matériellement	Possible (intégration de moteurs de vibration)

Tableau I.2. Analyse comparative des technologies de reconnaissance gestuelle [9-10-11].

En conclu que L'analyse confirme que pour un usage mobile, autonome et bidirectionnel, l'ap-proche matérielle par gant intelligent est la plus viable et la plus performante.

I.4 Intelligence artificielle appliquée à la reconnaissance des gestes :

I.4.1 Jeux de données pour la langue des signes :

Les bases de données jouent un rôle essentiel dans le développement des systèmes de reconnaissance automatique de la langue des signes. Elles facilitent le processus d'entraînement des modèles d'intelligence artificielle et le test de leurs performances dans différentes condi-tions. Avec le développement des techniques d'apprentissage profond, il est devenu nécessaire de disposer de bases de données importantes et variées afin d'améliorer la précision des sys-tèmes de reconnaissance des gestes et des signes manuels. Ces bases de données comprennent un certain nombre de vidéos contenant de nombreux signes effectués par plusieurs utilisateurs. Cela aide les modèles à apprendre les différentes variations dans la manière dont les individus s'expriment. L'une des bases de données les plus importantes utilisées est la base WLASL (Word-Level American Sign Language), présentée par Li et ses collègues en 2020. Cette base contient plus de 21 000 vidéos représentant plus de 2 000 mots en langue des signes américaine, enregistrées à partir de 119 sites. Les chercheurs ont également effectué plusieurs comparai-sons entre différents modèles d'apprentissage profond pour la reconnaissance des signes, no-tamment des modèles d'analyse d'images RVB et d'autres basés sur l'étude de la position du corps et des mains. Cette base a contribué à l'évaluation des performances des algorithmes de reconnaissance de la langue des signes au niveau des mots [19]. Une autre base de données importante, appelée MS-ASL, a été développée par Vaezi Joze et Koller en 2019. Elle com-prend plus de 25 000 vidéos couvrant environ 1 000 signes en langue des signes américaine, enregistrées par plus de 200 signataires. Cette base se caractérise par la diversité des partici-pants et des styles d'interprétation, ce qui permet d'évaluer les systèmes dans un environnement plus réaliste, connu sous le nom de « reconnaissance indépendante du signataire », qui teste le système sur des utilisateurs n'apparaissant pas dans les données d'entraînement [20].

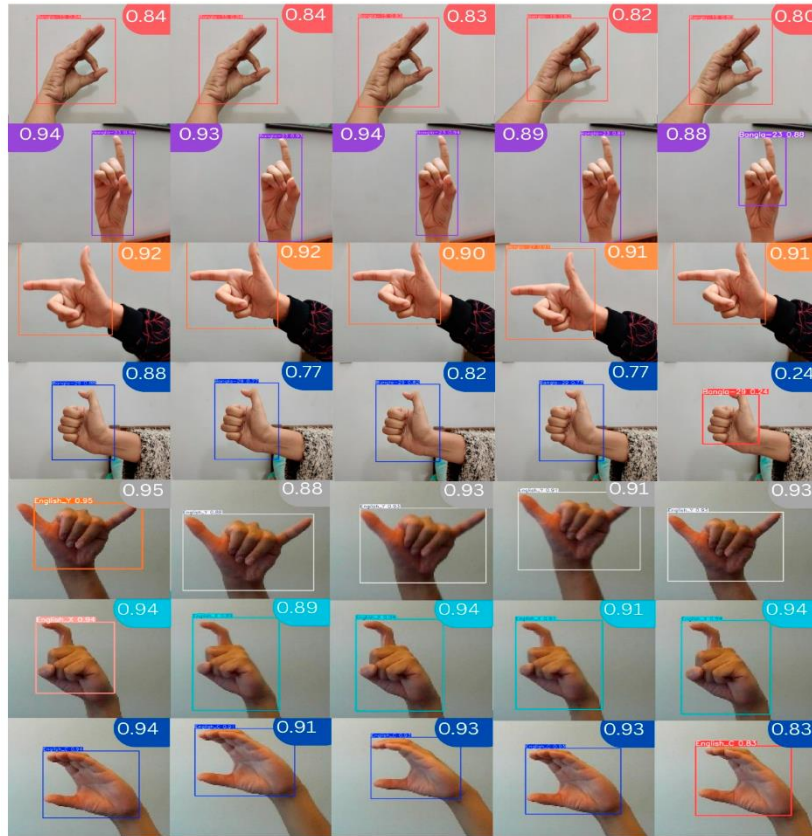


Figure I.8. Exemples d'images détectées [34].

I.4.2 Architectures CNN/LSTM et variantes pour la reconnaissance de gestes :

Le modèle hybride CNN-LSTM s'impose aujourd'hui comme l'architecture de référence pour le traitement des séries temporelles (Time-series) issues de capteurs portables (Wearables) [35]. Contrairement aux modèles de vision par ordinateur qui appliquent des convolutions 2D sur des matrices de pixels, notre approche exploite des convolutions unidimensionnelles (Conv1D) spécifiquement optimisées pour les signaux physiques séquentiels [37]. La première phase du modèle s'appuie sur un réseau de neurones convolutif 1D conçu pour traiter les flux de données brutes générés par le gant intelligent (données des capteurs de flexion, accélérométrie et gyroscopie du MPU-6050, ainsi que les signaux physiologiques du MAX30102). Les filtres convolutifs glissent le long de l'axe temporel pour extraire automatiquement des caractéristiques locales (features) de haut niveau [35]. Cette étape permet d'isoler des motifs pertinents — tels qu'une inflexion abrupte d'un doigt ou une variation soudaine du rythme cardiaque — tout en réduisant la dimensionnalité des données avant l'analyse séquentielle. Les vecteurs de caractéristiques extraits par le CNN sont ensuite injectés dans des couches récurrentes LSTM (Long Short-Term Memory). Étant donné que la langue des signes et l'expression émotionnelle sont des processus dynamiques, le LSTM excelle dans la mémorisation du contexte à

long terme au sein d'une séquence temporelle [37]. L'intégration simultanée des données biomécaniques et physiologiques (comme la fréquence cardiaque - BPM) à travers cette architecture permet de fusionner la cinématique gestuelle et l'état d'activation physiologique de l'utilisateur. Ainsi, le modèle acquiert la capacité d'inférer à la fois le sens linguistique du geste et la charge émotionnelle sous-jacente (Affective Computing) [38].

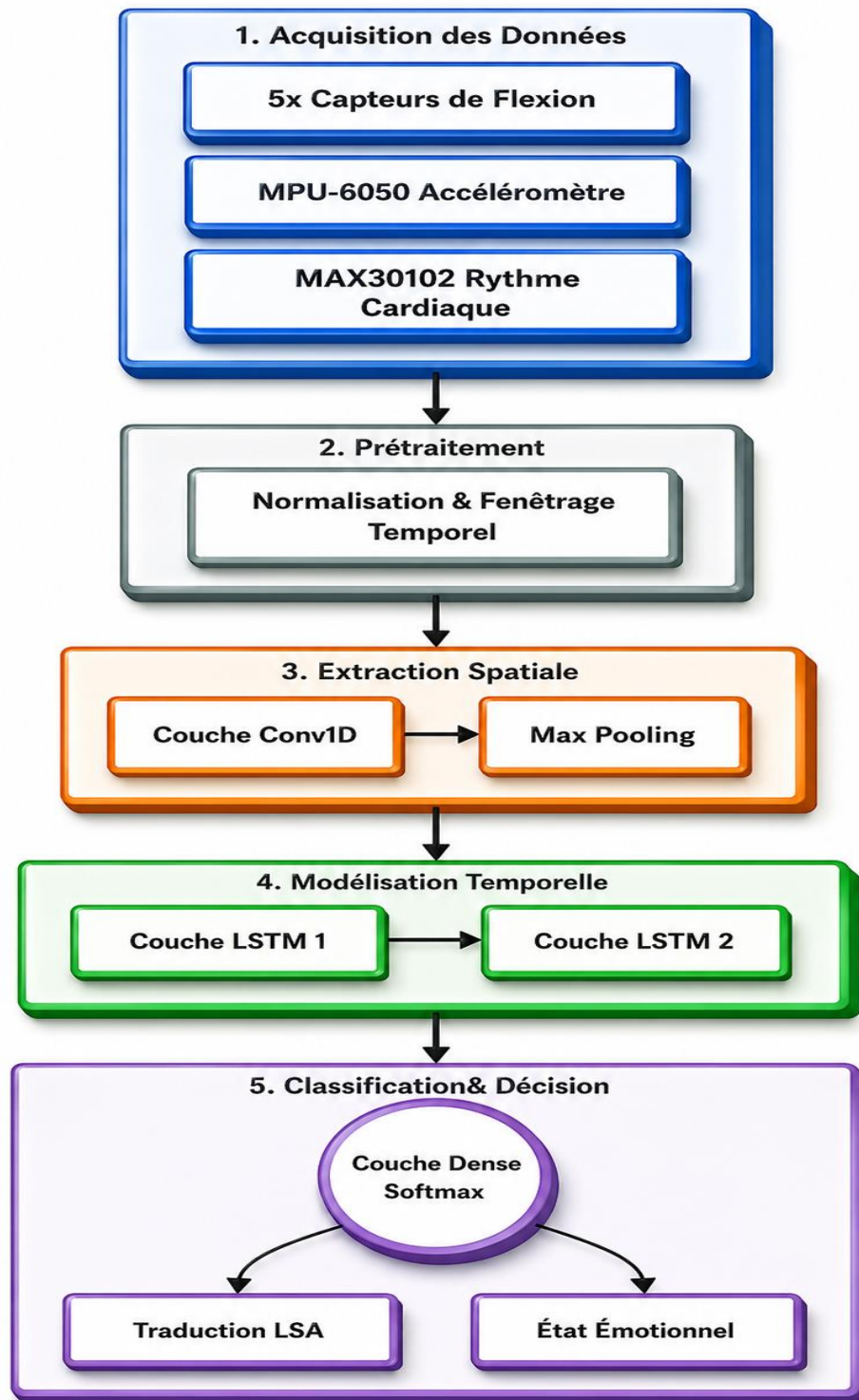


Figure I.9. Architecture hybride CNN 1D - LSTM [31].

I.4.3 Modélisation Cinématique de la Main et Approche Dimensionnelle de l'Informatique Affective :

L'analyse cinématique des gestes et la reconnaissance des émotions constituent deux piliers indissociables pour le développement d'une interface de communication bidirectionnelle naturelle. Historiquement dominées par la vision par ordinateur, ces tâches se heurtent souvent à des limites physiques majeures telles que l'occlusion spatiale, la variabilité de l'éclairage et la restriction du champ de vision. Pour pallier ces contraintes, l'utilisation de dispositifs portables (*wearables*) instrumentés, tels que les gants intelligents, permet d'adopter une approche égocentrique où la main est rigoureusement modélisée comme une chaîne cinématique complexe dotée de multiples degrés de liberté [41]. Dans ce paradigme matériel, les capteurs de flexion mesurent de manière absolue et directe les variations angulaires des phalanges, tandis que les unités de mesure inertielle (IMU, comme le MPU-6050) fournissent une quantification tridimensionnelle précise de la trajectoire, de l'accélération et de l'orientation spatiale du poignet [41]. Cependant, le traitement de ces flux de données brutes et séquentiels exige des architectures d'apprentissage automatique capables d'extraire des caractéristiques pertinentes sans recourir à une ingénierie manuelle des données (*handcrafted features*). C'est dans ce contexte que les architectures hybrides de type DeepConvLSTM se sont imposées comme le standard de référence pour la reconnaissance multimodale d'activités [39]. Contrairement aux réseaux de neurones convolutifs 2D, notre système déploie des convolutions unidimensionnelles (Conv1D) spécifiquement optimisées pour les séries temporelles issues de capteurs physiques. Cette opération mathématique permet d'isoler des motifs locaux invariants tels qu'une inflexion abrupte d'un doigt ou une accélération soudaine tout en réduisant la complexité calculatoire [40]. Ces vecteurs de caractéristiques sont ensuite ingérés par des couches *Long Short-Term Memory* (LSTM), conçues pour retenir les dépendances temporelles complexes inhérentes à la dynamique continue de la langue des signes [39-40]. Au-delà de la traduction sémantique, notre architecture intègre un paradigme d'informatique affective visant à retranscrire la prosodie émotionnelle. Cette approche s'appuie sur le modèle ci complexe de l'affect, conceptualisant les états émotionnels selon deux dimensions : la valence et l'activation (*arousal*) [43]. L'intégration d'un capteur photopléthysmographiques (MAX30102) permet d'extraire la fréquence cardiaque (Heart Rate - BPM), qui, couplée à la dynamique du mouvement (Magnitude), sert d'indicateur robuste pour l'estimation de l'activation physiologique [42].

Néanmoins, l'exécution des signes induit une charge métabolique et des artefacts de mouvement (MA) sévères qui corrompent le signal cardiaque brut. Pour garantir la validité de la lecture émotionnelle, le système déploie une fusion de données multi-capteurs couplée à un filtrage adaptatif. En utilisant les données accélérométriques du MPU-6050 comme référence de bruit, l'algorithme des moindres carrés moyens (LMS) permet de soustraire mathématiquement les composantes cinématiques du signal PPG pour isoler l'activation affective réelle [45]. En fusionnant ces caractéristiques biomécaniques et physiologiques ainsi filtrées, le système acquiert une capacité de contextualisation supérieure : un même signe pourra être interprété avec des nuances vocales distinctes selon l'état émotionnel de l'utilisateur, offrant une traduction véritablement empathique et exhaustive [39-42].

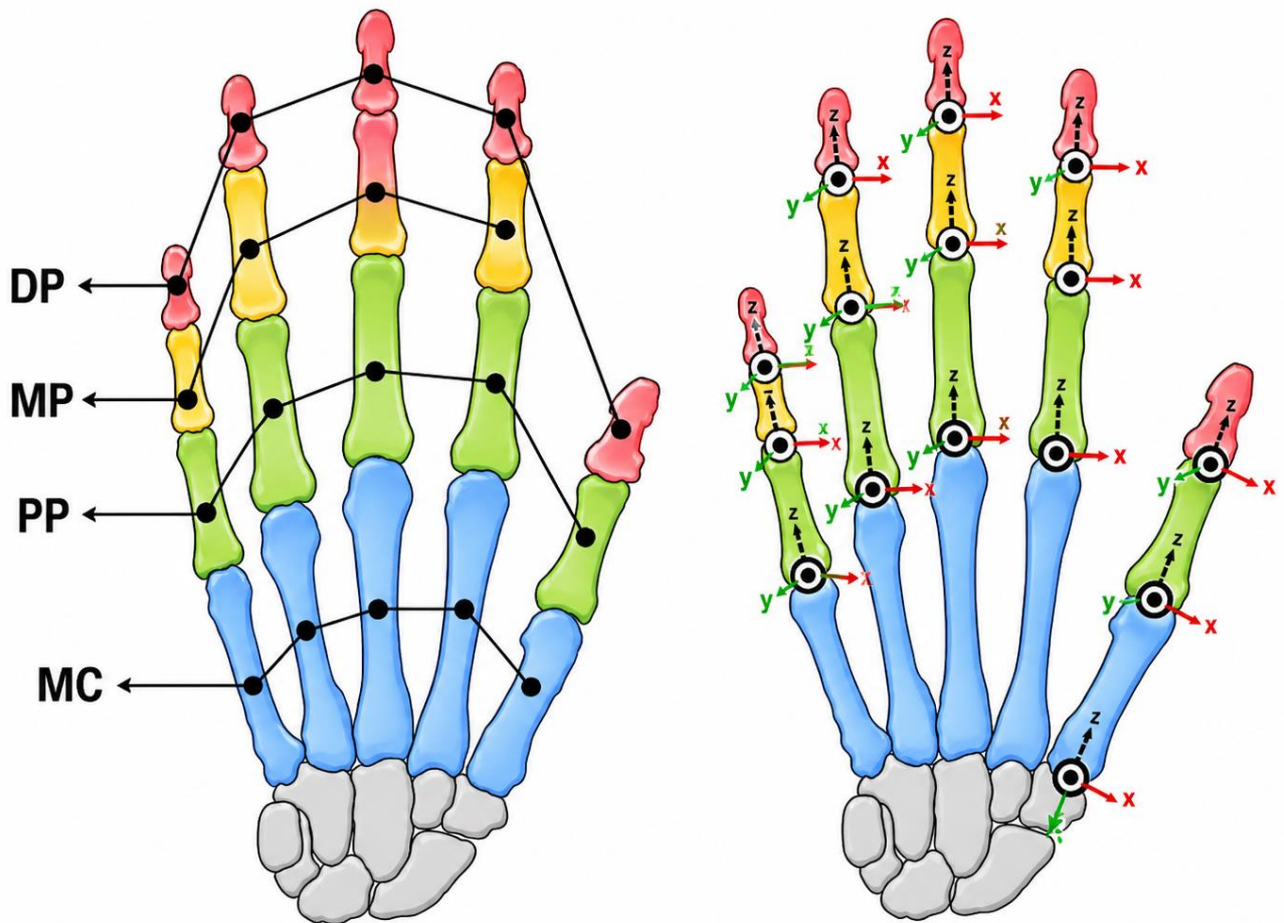


Figure I.10. Modélisation cinématique de la main illustrant les articulations et les degrés de liberté capturés par les capteurs de flexion et l'unité IMU [46].

I.4.4 L'émergence des Grands Modèles de Langage (LLMs) dans la traduction gestuelle:

Récemment, l'intégration des Grands Modèles de Langage (Large Language Models - LLMs) a marqué une avancée majeure dans le domaine de la traduction automatique de la langue des signes. Contrairement aux approches classiques limitées à une traduction séquentielle de type Gloss-to-Text, les architectures basées sur les LLMs permettent de générer des phrases plus naturelles, contextuellement cohérentes et sémantiquement enrichies. Les travaux récents de Wong et al. (2024) sur l'architecture Sign2GPT démontrent notamment que l'association de modèles de langage pré-entraînés avec des encodeurs visuels ou multimodaux améliore significativement la fluidité et la qualité contextuelle de la traduction gestuelle. Par ailleurs, les recherches récentes autour des architectures Gloss2Text basées sur les LLMs montrent que l'exploitation du contexte global permet de réduire les ambiguïtés sémantiques et d'améliorer la précision linguistique des phrases générées. Ces avancées justifient ainsi le choix d'intégrer le modèle Llama 3.1 au sein de notre architecture afin de produire une traduction textuelle et vocale davantage cohérente avec l'intention réelle de l'utilisateur [48-49].

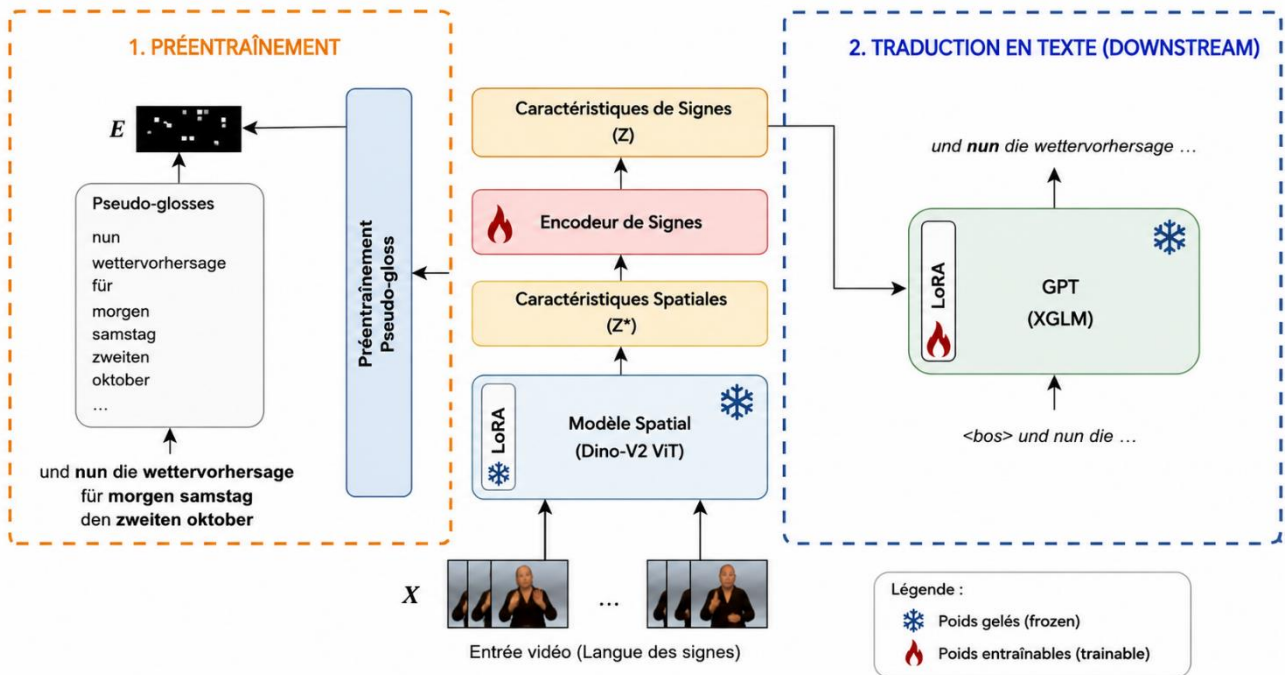


Figure I.11. Présentation de Sign2GPT, qui consiste en une étape de pré-formation utilisant des pseudo-glosses et une traduction en aval qui exploite un modèle GPT gelé [49].

I.5 Synthèse critique et justification des choix techniques :

Cette section synthétise l'analyse comparative des technologies de reconnaissance gestuelle présentées précédemment et justifie les choix architecturaux adoptés pour le projet. L'objectif est de démontrer la pertinence de l'approche matérielle instrumentée par rapport aux solutions logicielles conventionnelles.

I.5.1 Analyse comparative des approches :

Le choix d'une architecture basée sur un gant intelligent plutôt qu'une solution de vision par ordinateur repose sur une évaluation multidimensionnelle des contraintes de terrain. Le tableau suivant résume les avantages et inconvénients de chaque approche :

Critères d'évaluation	Vision par ordinateur (Caméras)	Capteurs portables (Gant intelligent)
Coût de mise en œuvre	Modéré (Nécessite des caméras HD et GPU)	Faible (Microcontrôleurs et capteurs abordables)
Confidentialité	Faible (Risque de capture de données faciales)	Totale (Données inertielles uniquement)
Mobilité et Indépendance	Limitée (Utilisateur lié au champ de l'objectif)	Élevée (Système embarqué et autonome)
Temps de réponse (Latence)	Variable (Traitement d'image gourmand)	Faible (Acquisition directe des signaux)
Indépendance lumineuse	Nulle (Sensible à l'éclairage et aux ombres)	Totale (Fonctionne en toute condition)

Tableau I.3. Analyse comparative entre la vision par ordinateur et les capteurs portables pour la reconnaissance de gestes [27-28-30].

I.5.2 Justification du choix technologique

L'adoption du gant intelligent instrumenté est motivée par la nécessité d'offrir une solution inclusive capable de fonctionner dans des environnements domestiques non contrôlés, tout en garantissant une efficacité énergétique et une portabilité optimales grâce à l'Internet des Objets (IoT) [30]. Contrairement à la vision artificielle qui souffre fréquemment de problèmes d'occlusion spatiale et d'exigences matérielles lourdes pour le traitement vidéo [28], le gant fournit une mesure directe et garantit une précision cinématique stable, indispensable pour une interaction tactile et communicative fiable [27]. Parallèlement, l'intégration du Deep Learning

(notamment les modèles LSTM) se justifie par la nature séquentielle des gestes de la langue des signes. L'étude des grands ensembles de données (Large-Scale Datasets) démontre que les algorithmes de classification classiques échouent à capter la dynamique temporelle complexe des signes. L'utilisation des réseaux de neurones récurrents permet d'atteindre des taux de reconnaissance supérieurs à 90%, ce qui est la condition sine qua non pour assurer une traduction fluide et utilisable au quotidien [29].

I.6 Conclusion :

Ce chapitre a présenté une analyse globale du handicap auditif, de la langue des signes et des principales technologies d'assistance destinées à améliorer la communication des personnes sourdes. L'étude a montré que les gants intelligents constituent une solution particulièrement pertinente grâce à leur portabilité, leur fonctionnement indépendant des conditions d'éclairage, leur capacité de traitement en temps réel et leur respect de la confidentialité. L'analyse des approches existantes a également mis en évidence l'importance de l'intégration de l'intelligence artificielle, notamment des architectures hybrides Conv1D-LSTM, pour la reconnaissance précise des gestes. L'exploitation des signaux physiologiques, tels que le PPG, permet en outre d'ajouter une dimension émotionnelle à l'interprétation des mouvements, tandis que les grands modèles de langage offrent de nouvelles possibilités pour produire des traductions plus naturelles et contextuellement adaptées. Ainsi, cette étude justifie le développement d'un gant intelligent IoT multimodal combinant capteurs embarqués, fusion de données, apprentissage profond et retour haptique. Cette architecture constitue une base solide pour la conception et l'implémentation du système proposé, qui seront détaillées dans le chapitre suivant.

Bibliographie Chapitre I

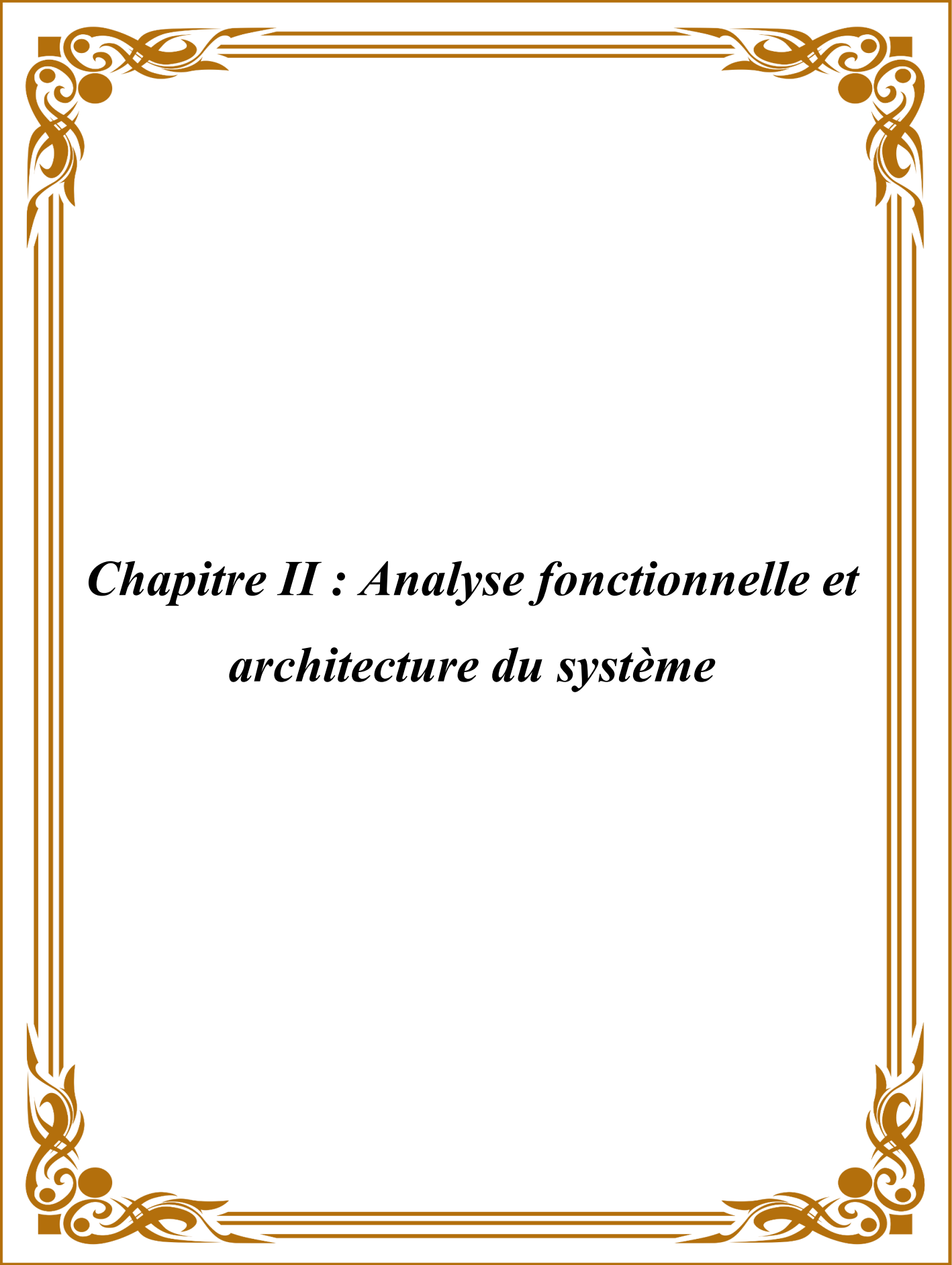
- [1] Organisation Mondiale de la Santé (OMS), "Surdité et perte auditive", Aide-mémoire, 2026.
- [2] N. Dahel, "Réalisation et conception d'un gant intelligent pour sourds muets", *Mémoire de Master, Université de Tlemcen*, 2020.
- [3] K. Ravi Chandra, V. Suhruth, et al., "Bridging the Silence: Engineering a Voice for the Speech Impaired", *Rapport de projet de fin d'études, ANITS*, 2021.
- [4] R. Belekar, T. Pawar, S. Chougule, et R. Jamadar, "Smart Communication Haptic Glove for Deaf and Dumb People", *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, Vol. 5, Issue 8, 2025.
- [5] N. Benzerga, "Sign Language Translation for the Deaf and Dumb & its Challenges", *TAALIMIA Review*, Vol. 15, N° 02, pp. 941–952, 2025.
- [6] N. Boukhattala et K. Gueraichi, "La reconnaissance de la langue des signes algérienne : La mise en place d'un système de traduction", *Mémoire de Master, Université Kasdi Merbah Ouargla*, 2017.
- [7] M. Ababsa, "Reconnaissance de gestes appliquée à la langue des signes", *Mémoire de Master, Université L'arbi Ben M'hidi - Oum El Bouaghi*, 2013.
- [8] Y. Abazi et A. Z. Aziza, "Hand gesture and sign language recognition based on deep learning", *Mémoire de Master, Université Kasdi Merbah Ouargla*, 2023. [En ligne]. Disponible sur Google Scholar.
- [9] N. Boukhattala et K. Gueraichi, "La reconnaissance de la langue des signes algérienne : La mise en place d'un système de traduction", *Mémoire de Master, Université Kasdi Merbah Ouargla*, 2017.
- [10] P. Vora, B. Parab, R. Naik, R. Bandgar, et J. Dange, "Multi-Purpose Smart Glove for Differently Abled Community People", *International Journal of Research Publication and Reviews*, Vol. 3, No. 1, pp. 1101-1111, Jan. 2022.

- [11] T. Benbrahim, "A Smart Glove Empowered by AI embedded on a Raspberry Pi", *Projet de fin d'études, École Nationale Supérieure de Technologies Avancées (ENSTA) Alger*, 2024.
- [12] K. Ravi Chandra, V. Suhruth, K.S.D. Charishma Patnaik, et M.J.N. Sandeep, "Bridging the Silence : Engineering a Voice for the Speech Impaired", *Rapport de projet de fin d'études, ANITS*, 2021.
- [13] Rishabh Software. (2023, April 11). *What is computer vision and its benefits*. Rishabh Software. <https://www.rishabhsoft.com/blog/computer-vision>
- [14] SignSpeak, « GloveTalk (Converting Sign Language to Words) », projet *Hackster.io*, 2025. [En ligne]. Disponible sur : <https://www.hackster.io/signspeak/glovetalk-converting-sign-language-to-words-6b86ef>
- [15] A. Chellali, "Étude des interactions homme-homme pour l'élaboration du référentiel commun dans les environnements virtuels collaboratifs," Ph.D. dissertation, Université de Nantes, 2009.
- [16] C. Colomer, "Échange d'information et émergence des rôles lors d'interactions haptiques interpersonnelles," Ph.D. dissertation, Université de Montpellier, 2022.
- [17] G. Richard, "Étude de l'impact du retour haptique sur le sentiment d'incarnation," Master's thesis, Inria Lille Nord Europe, 2019.
- [18] G. Richard, "Rôle du retour haptique dans les interactions avec des avatars en réalité virtuelle," Ph.D. dissertation, Université de Lille, 2023.
- [19] Li, D., Rodriguez, C., Yu, X., & Li, H. (2020). Word-Level Deep Sign Language Recognition from Video : A New Large-Scale Dataset and Methods Comparison. *IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1459–1469). IEEE.
- [20] Vaezi Joze, H. R., & Vaezi Joze, O. (2019). MS-ASL: A Large-Scale Data Set and Benchmark for Understanding American Sign Language. *British Machine Vision Conference (BMVC)*. BMVA Press.

- [21] F. S. Takouchouang et H. T. Vinh, "Reconnaissance Automatique des Langues des Signes: Une Approche Hybridée CNN-LSTM Basée sur Mediapipe," Prépublication HAL, hal-05291238, Jan. 2026.
- [22] A. Beaulieu, "Combinaison de CNN et LSTM pour la reconnaissance d'activités dans un environnement intelligent avec radars Ultra-Wideband," Mémoire de Maîtrise, Univ. du Québec à Chicoutimi, Canada, 2021.
- [23] F. S. Takouchouang, "Reconnaissance Automatique des Langues des Signes: Une Approche Hybridée CNN-LSTM Basée sur Mediapipe," arXiv preprint, arXiv:2510.22011, Oct. 2025.
- [24] A. Truong, "Analyse du contenu expressif des gestes corporels," Thèse de Doctorat, Institut National des Télécommunications, France, 2016.
- [25] A. Crenn, H. Konik, A. Meyer, et S. Bouakaz, "Reconnaissance d'expressions corporelles à l'aide d'un mouvement neutre synthétisé," in Actes de CORESA 2017, Caen, France, Nov. 2017.
- [26] A. Crenn, "Reconnaissance d'expressions corporelles dans des mouvements de personnes en vue de la synthèse de style," Thèse de Doctorat, Université de Lyon, France, 2019.
- [27] O. Ozioko et R. Dahiya, "Smart tactile gloves for haptic interaction communication and rehabilitation", Advanced Intelligent Systems, 2022.
- [28] D. Li, C. Rodriguez, X. Yu et H. Li, "Word-Level Deep Sign Language Recognition from Video: A New Large-Scale Dataset and Methods Comparison", IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1459–1469, 2020.
- [29] H. R. Vaezi Joze et O. Vaezi Joze, "MS-ASL: A Large-Scale Data Set and Benchmark for Understanding American Sign Language", British Machine Vision Conference (BMVC), 2019.

- [30] S. K. Derle, N. D. Desale, S. S. Wagh et C. R. Thorat, "Smart Glove: IoT-Based Sign Language Translation", *International Journal of Innovative Research in Medical and Practical Science (IJIRMPS)*, Vol. 14, Issue 1, 2026.
- [31] Kim, G., Hwang, D., & Park, J. (2021). Effect of 2.5 D haptic feedback on virtual object perception via a stylus. *Scientific Reports*, 11(1), 18954.
- [32] Kim, G., Hwang, D., & Park, J. (2021). Effect of 2.5 D haptic feedback on virtual object perception via a stylus. *Scientific Reports*, 11(1), 18954.
- [34] Navin, N., Farid, F. A., Rakin, R. Z., Tanzim, S. S., Rahman, M., Rahman, S., Uddin, J., & Karim, H. A. (2025). Bilingual Sign Language Recognition: A YOLOv11-Based Model for Bangla and English Alphabets. *Journal of Imaging*, 11(5), 134. <https://doi.org/10.3390/jimaging11050134>
- [35] Ordóñez, F. J., & Roggen, D. (2016). Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1), 115..
- [36] Noroozi, F., Corneanu, C. A., Kamińska, D., Sapiński, T., Escalera, S., & Anbarjafari, G. (2018). Survey on emotional body gesture recognition. *IEEE transactions on affective computing*, 12(2), 505-523.
- [37] Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern recognition letters*, 119, 3-11.
- [38] Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., ... & Yang, X. (2018). A review of emotion recognition using physiological signals. *Sensors*, 18(7), 2074.
- [39] Ordóñez, F. J., & Roggen, D. (2016). Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1), 115.
- [40] Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern recognition letters*, 119, 3-11.
- [41] Dipietro, L., Sabatini, A. M., & Dario, P. (2008). A survey of glove-based systems and their applications. *Ieee transactions on systems, man, and cybernetics, part c (applications and reviews)*, 38(4), 461-482.

- [42] Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., ... & Yang, X. (2018). A review of emotion recognition using physiological signals. *Sensors*, 18(7), 2074.
- [43] Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(3), 715-734.
- [44] Vicente, A. P., & Faisal, A. A. (2013, May). Calibration of kinematic body sensor networks: Kinect-based gauging of data gloves “in the wild”. In *2013 IEEE International Conference on Body Sensor Networks* (pp. 1-6). IEEE.
- [45] Ram, M. R., Madhav, K. V., Krishna, E. H., Komalla, N. R., & Reddy, K. A. (2011). A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter. *IEEE Transactions on Instrumentation and Measurement*, 61(5), 1445-1457.
- [46] Vicente, A. P., & Faisal, A. A. (2013, May). Calibration of kinematic body sensor networks: Kinect-based gauging of data gloves “in the wild”. In *2013 IEEE International Conference on Body Sensor Networks* (pp. 1-6). IEEE.
- [47] Zhao, T., Liu, J., Wang, Y., Liu, H., & Chen, Y. (2018, April). PPG-based finger-level gesture recognition leveraging wearables. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications* (pp. 1457-1465). IEEE.
- [48] Fayyazsanavi, P., Anastasopoulos, A., & Kosecka, J. (2024, November). Gloss2Text: sign language gloss translation using LLMs and semantically aware label smoothing. In *Findings of the Association for Computational Linguistics: EMNLP 2024* (pp. 16162-16171).
- [49] Wong, R., Camgoz, N. C., & Bowden, R. (2024). Sign2GPT: Leveraging large language models for gloss-free sign language translation. *arXiv preprint arXiv:2405.04164*.



Chapitre II : Analyse fonctionnelle et architecture du système

II.1 Introduction

Les langues des signes constituent un moyen de communication riche et complexe, mobilisant simultanément la configuration des mains, leur orientation, leur dynamique spatiale et les expressions faciales. Afin de capter et de traduire cette complexité, ce deuxième chapitre est consacré à l'étude approfondie de l'architecture et de la conception de notre système de gant intelligent. L'objectif est de détailler l'ensemble de la chaîne de traitement de l'information, depuis l'acquisition matérielle des données biomécaniques et physiologiques jusqu'à leur restitution sous forme de texte et de synthèse vocale. Dans un premier temps, nous présenterons l'architecture générale du système, divisée en blocs fonctionnels interdépendants incluant l'acquisition multimodale et l'intelligence embarquée. Une analyse fonctionnelle rigoureuse sera ensuite menée à l'aide des diagrammes bête à cornes, pieuvre et FAST, afin de traduire les besoins des utilisateurs en spécifications et contraintes techniques, ergonomiques et énergétiques claires. Par la suite, nous justifierons nos choix de conception matérielle, notamment l'utilisation du microcontrôleur ESP32, des capteurs de flexion, de la centrale inertielle MPU-6050, du convertisseur ADS1115 et du capteur physiologique MAX30102. Enfin, ce chapitre exposera les aspects logiciels et la topologie de communication hybride Edge-to-Cloud, mettant en lumière la programmation de l'interface sous Unity 3D et la transmission asynchrone des données via Bluetooth Low Energy et Wi-Fi.

II.2 Architecture générale du système

L'architecture du gant intelligent repose sur une chaîne de traitement d'information complexe, allant de la capture biomécanique du mouvement à la restitution sensorielle. Cette architecture est divisée en quatre blocs fonctionnels interdépendants.

II.2.1 Partie acquisition des données

Le bloc d'acquisition est chargé de transformer les mouvements physiques de la main en signaux électriques exploitables. « Le gant se compose de quatre capteurs de flexion et d'un capteur MPU-6050 pour capturer les gestes et les mouvements de la main » [14]. Les capteurs de flexion, placés sur chaque doigt, mesurent le degré de courbure, tandis que l'unité de mesure inertielle (IMU) capte la dynamique globale. « L'intégration des capteurs de flexion permet la détection des mouvements des doigts, tandis que le capteur MPU-6050 fournit des informations sur l'orientation et le mouvement de la main » [14]. Cette combinaison hybride est indispensable car « les signes utilisent à la fois des configurations statiques et des mouvements dynamiques pour transmettre un sens » [13].

II.2.1.1 Stratégie de Fusion de Données (Sensor Fusion) :

Afin d'améliorer la robustesse du système face aux artefacts de mouvement, une stratégie de fusion de données (Sensor Fusion) est intégrée au niveau du prétraitement des signaux. Cette approche exploite simultanément les données photopléthysmographiques du MAX30102 et les données inertielles du MPU-6050 afin d'identifier et de réduire les perturbations causées par les mouvements rapides de la main. Les signaux d'accélération sont utilisés comme référence dynamique pour détecter les composantes fréquentielles associées aux mouvements parasites, permettant ainsi de filtrer et de stabiliser l'estimation physiologique avant la phase d'inférence. Cette architecture transforme alors les artefacts de mouvement, traditionnellement considérés comme une limitation majeure des systèmes PPG embarqués, en une source d'information contextuelle contribuant à la robustesse décisionnelle du système multimodal[46].

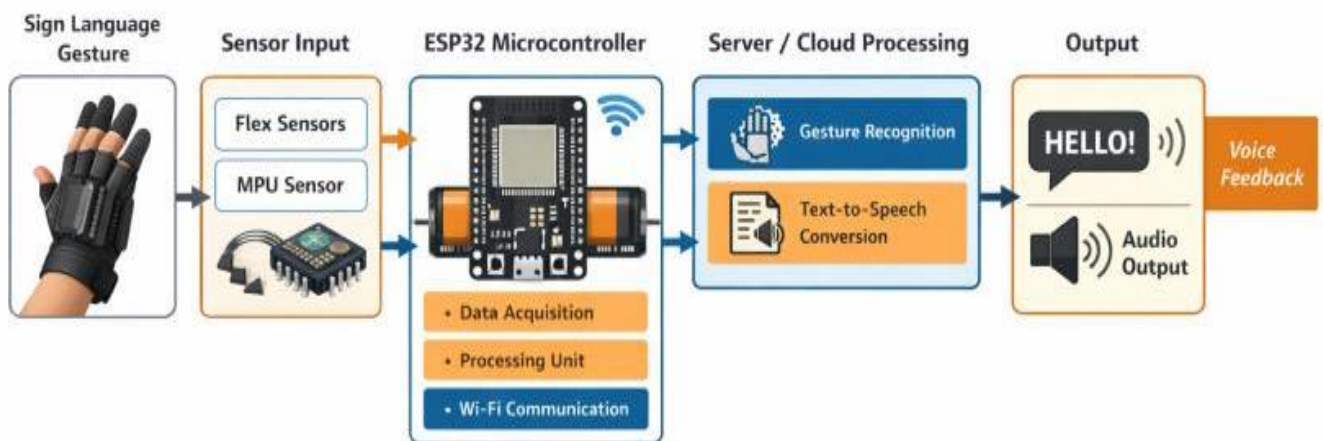


Figure II.1. Architecture du système : De la gestuelle analogique à la parole numérique [20].

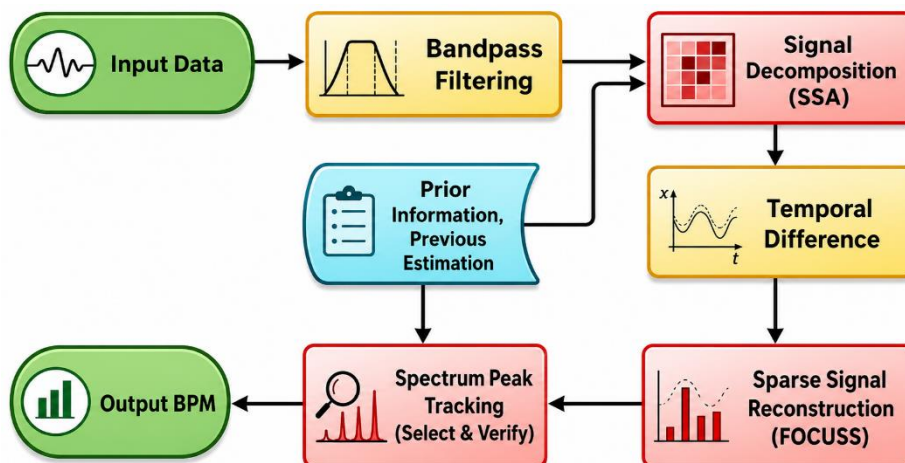


Figure II.2. Organigramme du cadre TROIKA. Décomposition du signal, échantillonnage parcimonieux [46].

II.2.2 Partie traitement et intelligence embarquée

Une fois les signaux analogiques convertis en données numériques, ils sont acheminés vers l'unité centrale. Les données collectées par les capteurs sont traitées et transmises à un modèle d'apprentissage profond (Deep Learning) entraîné sur un ensemble d'actions de la langue des signes [14]. Dans les architectures modernes, le traitement est souvent distribué : un microcontrôleur local (type ESP32 ou Arduino) gère l'acquisition en temps réel, tandis qu'un processeur plus puissant ou un smartphone exécute le modèle de classification. Le système utilise un réseau de neurones récurrents (RNN) avec une architecture LSTM (Long Short-Term Memory) pour permettre la reconnaissance et la catégorisation des gestes [15].

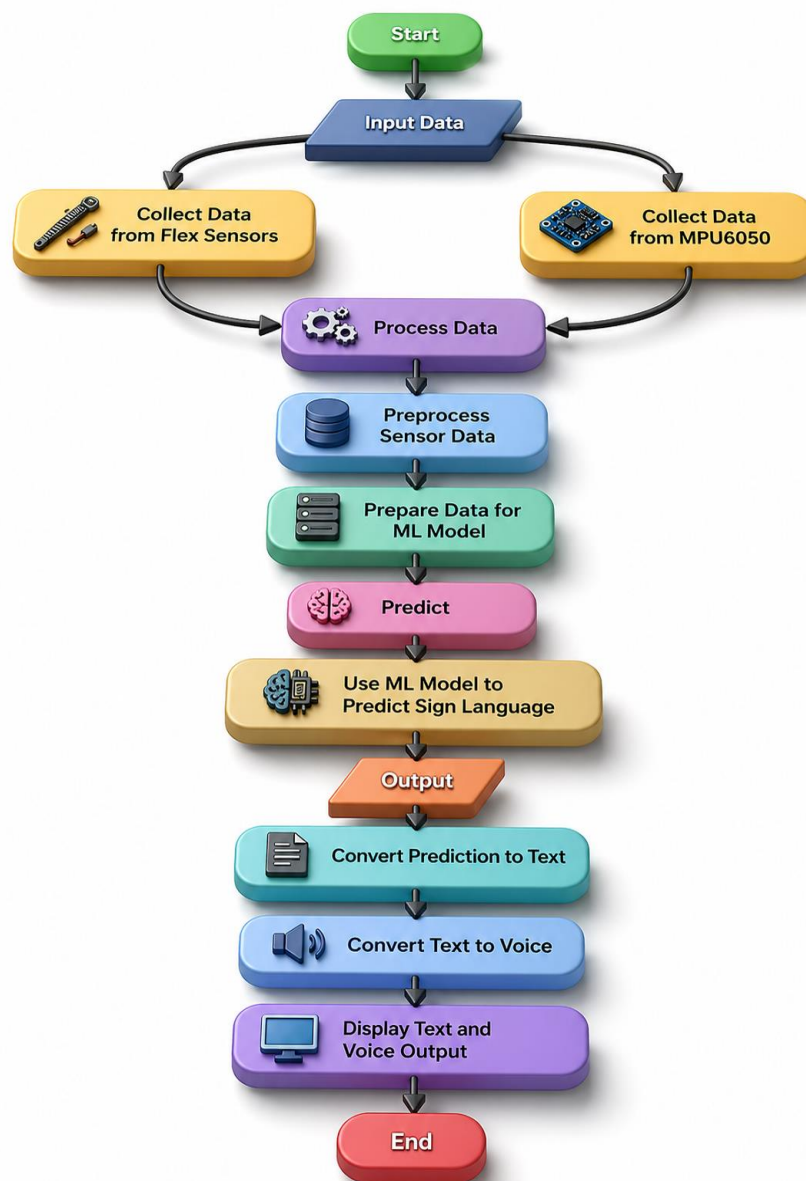


Figure II.2. Organigramme fonctionnel [12].

II.2.3 Interface de communication et restitution vocale

La communication est un défi majeur pour les personnes qui dépendent de la langue des signes, car la plupart des gens ne comprennent pas les gestes de la main. Pour réduire ce fossé de communication, ce projet présente un système de gant intelligent basé sur l'Internet des Objets (IoT) qui convertit les gestes de la langue des signes en parole audible. Le gant est intégré avec des capteurs de flexion pour détecter la courbure des doigts et un capteur MPU pour suivre le mouvement et l'orientation de la main. Un microcontrôleur ESP32 collecte les données des capteurs et les transmet sans fil à une unité de traitement via Internet. Les données reçues sont analysées pour reconnaître des gestes de la main prédéfinis, et la sortie texte et audio correspondante est générée à l'aide d'un module de synthèse vocale (Text-to-Speech). Le système proposé permet la conversion en temps réel de la langue des signes en voix, permettant aux personnes souffrant de troubles de la parole de communiquer plus facilement avec les autres. L'appareil est conçu pour être portable, peu coûteux et convivial, ce qui le rend adapté à un usage quotidien. Ce gant intelligent démontre comment l'Internet des Objets (IoT) et la technologie des capteurs peuvent être utilisés efficacement pour construire des systèmes de communication d'assistance [21].

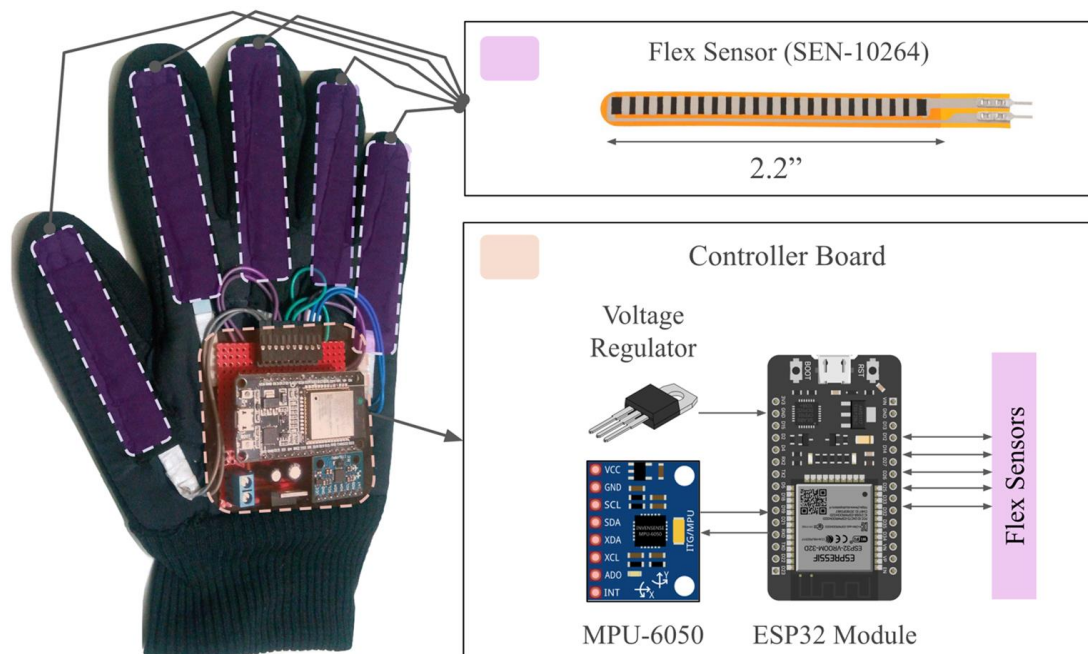


Figure II.3. Architecture d'un gant intelligent basé sur capteurs de flexion et MPU-6050[27].

II.2.4 Communication sans fil et gestion des données

L'architecture de communication repose sur les capacités natives du microcontrôleur **ESP32 WROOM 32**. Contrairement aux solutions classiques nécessitant des modules externes, l'ESP32 intègre nativement les protocoles Bluetooth (BR/EDR et BLE) et Wi-Fi, ce qui permet une transmission directe et sécurisée des données sans ajout de matériel supplémentaire. Le flux de données suit une logique client-serveur : une fois les signaux des capteurs (Flex, MPU et MAX30102) prétraités, ils sont transmis via une connexion sans fil vers un **serveur distant**. Ce serveur fait office de moteur d'inférence. Pour assurer la persistance et la scalabilité du système, « le dataset et les modèles de reconnaissance sont stockés et gérés au sein d'une base de données NoSQL **MongoDB** » [22]. Ce choix technologique permet une flexibilité totale dans la structure des données de gestes et facilite l'apprentissage continu du système à mesure que de nouveaux signes sont enregistrés.



Figure II.4. ESP 32 wroom 32 WI-FI + BLuetooth Antenna [22].

II.2.5 Unité de stockage et Backend : MongoDB

Le choix de MongoDB comme système de gestion de base de données est justifié par sa capacité à manipuler des documents JSON/BSON, ce qui correspond parfaitement au format des trames de données envoyées par l'ESP32. Contrairement à un stockage local sur carte SD, cette architecture permet :

- **Un accès distant** : Les données peuvent être consultées ou enrichies depuis n'importe quel terminal.
- **Une scalabilité élevée** : Le dataset peut croître sans les limites physiques de la mémoire embarquée.
- **Un traitement déporté** : Le serveur peut exécuter des algorithmes de Machine Learning complexes (Bi-LSTM) sur le dataset MongoDB avant de renvoyer le résultat de la traduction vers l'interface de l'utilisateur.

```
{
  firstname: "Bob",
  lastname: "Smith",
  email: "bob@smith.com",
  address: {
    street: "100 Main St",
    city: "Anytown",
    state: "MO",
    zip: "11111"
  }
}
```

Figure II.5. MongoDB JSON Example Format [23].

- **Gestion Asynchrone des Flux de Données** : Afin de garantir une traduction fluide sans latence perceptible (temps réel), l'enregistrement des journaux d'activité dans la base de données MongoDB est effectué de manière asynchrone. Le pipeline de traitement priorise l'inférence du modèle DeepConvLSTM et la réponse du LLM, tandis que les tâches d'archivage sont déléguées à un processus d'arrière-plan, évitant tout blocage du flux de communication entre le gant et l'interface Unity 3D.

II.3 Analyse fonctionnelle du système :

II.3.1 Diagrammes fonctionnels du système

II.3.1.1 Diagramme bête à cornes :

Dans le cadre de l'analyse fonctionnelle du projet de gant intelligent, nous utilisons le diagramme bête à cornes afin de définir les besoins fondamentaux du système. Ce modèle permet de répondre à trois questions essentielles : à qui s'adresse le produit ? Sur quoi agit-il ? Et quel est son objectif ?

Comme le montre la **figure II.6**, le gant intelligent est destiné aux personnes souffrant d'un handicap auditif et/ou de troubles de la parole. Ce système analyse la langue des signes en étudiant les mouvements de la main et des doigts. Il traduit ces mouvements en texte ou en son afin de faciliter la communication et d'améliorer l'intégration sociale. Cette analyse repose sur les principes de l'analyse fonctionnelle utilisés en génie industriel. Ce besoin répond à un véritable problème, à savoir la difficulté d'intégration des personnes sourdes et muettes dans la société, celles-ci étant confrontées à des obstacles quotidiens en matière de communication. Ce système vise donc à réduire ce fossé en proposant un moyen de communication automatisé, rapide et facile à utiliser [9].

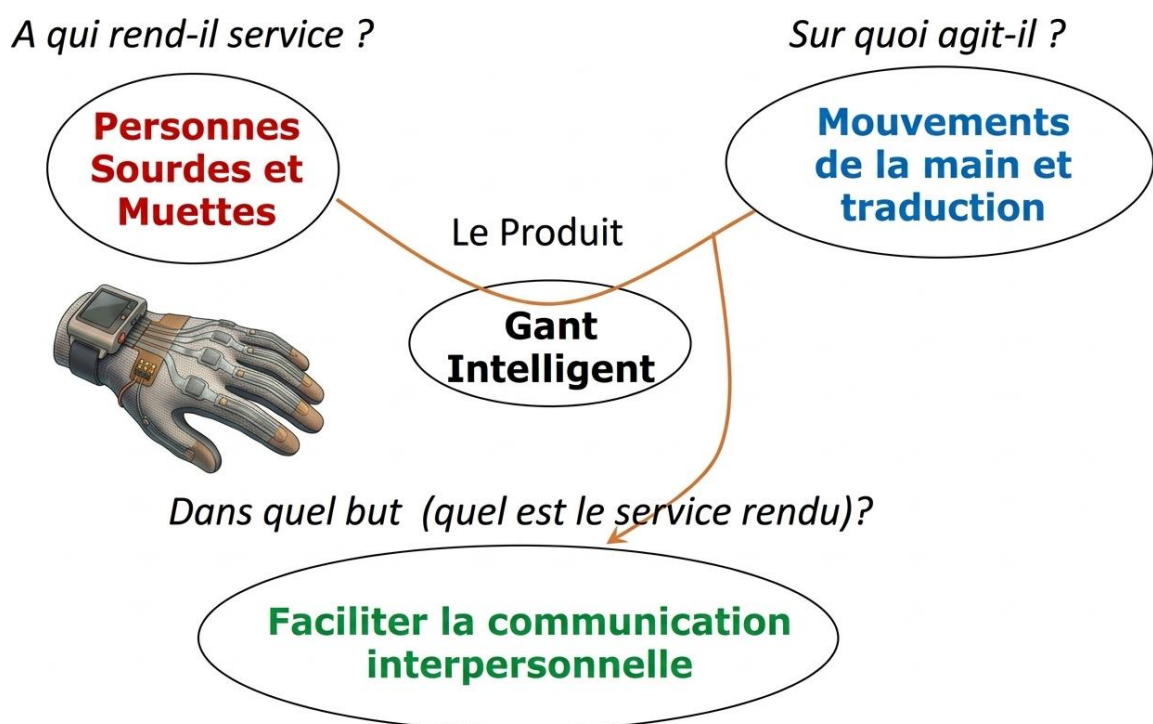


Figure II.6. Diagramme bête à cornes pour le gant intelligent [8].

- **Note :** La fonction principale (FP1) de traduction est limitée, dans cette phase expérimentale, à un lexique défini de 16 classes gestuelles pour assurer une précision maximale du modèle DeepConvLSTM.

II.3.1.2 Diagramme pieuvre :

Une fois les besoins fondamentaux du système définis à l'aide du diagramme bête à cornes, on utilise le diagramme pieuvre pour illustrer les interactions entre le système et son environnement extérieur. Ce diagramme présente tous les éléments qui interagissent avec le

gant intelligent et définit la fonction principale ainsi que les contraintes auxquelles le système doit se conformer.

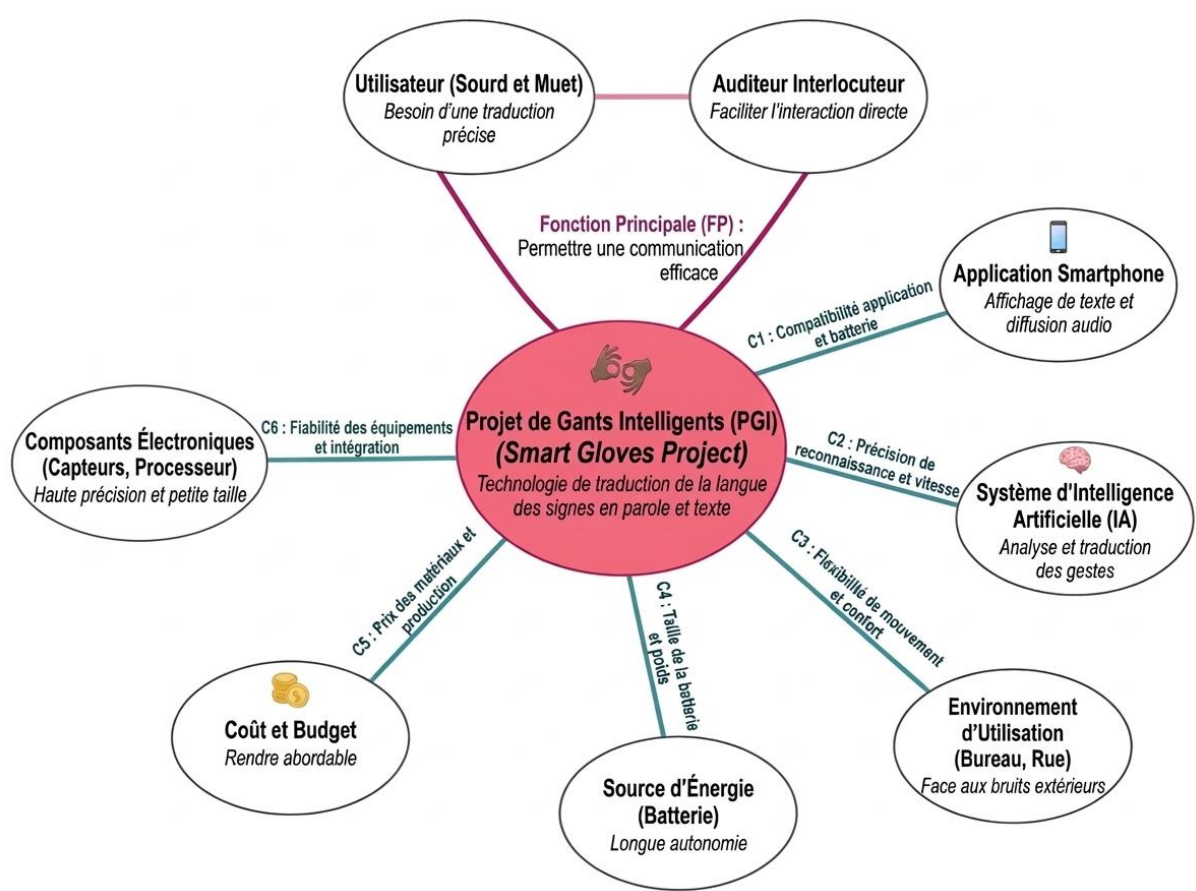


Figure II.7. Diagramme pieuvre du gant intelligent [10].

Ce schéma illustre la relation entre le gant intelligent et son environnement, qui comprend l'utilisateur, le smartphone, la source d'alimentation et l'environnement immédiat.

- **Fonction principale (FP) :**

Permettre la communication entre l'utilisateur (sourd ou muet) et son environnement, grâce à la traduction instantanée de la langue des signes en texte et en son. Cette fonction est considérée comme essentielle car elle représente l'objectif principal du système ; toute défaillance dans ce domaine empêche le gant de remplir son rôle.

- **Contraintes (Classement par ordre d'importance) :**

- **C2 :** Précision et rapidité d'analyse (la plus importante)

Le système doit être capable d'analyser les mouvements avec une grande précision et en un temps réduit. Cette caractéristique est l'une des contraintes les plus importantes, car toute erreur dans la reconnaissance des signes peut entraîner

une traduction erronée, ce qui provoque des malentendus dans la communication. C'est pourquoi on a eu recours à des techniques d'intelligence artificielle pour améliorer les performances.

- **C4 : Consommation d'énergie (deuxième contrainte la plus importante)**

Le système doit consommer peu d'énergie pour garantir une longue autonomie de la batterie. Le gant étant un appareil portable, une consommation d'énergie élevée réduirait la durée d'utilisation, ce qui aurait un impact négatif sur l'expérience utilisateur.

- **C3 : Confort et liberté de mouvement**

Le gant doit être confortable et ne pas gêner les mouvements de la main. Si le gant n'est pas confortable, l'utilisateur ne pourra pas le porter pendant de longues périodes, ce qui réduira son efficacité.

- **C1 : Compatibilité avec les smartphones**

Le gant doit pouvoir se connecter aux applications via le téléphone. Le téléphone est un moyen essentiel pour afficher les résultats (texte/son) ; par conséquent une incompatibilité réduit l'utilité du système.

- **C6 : Fiabilité et miniaturisation des composants, Les composants électroniques doivent être miniaturisés et fiables. Cela permet de concevoir un gant léger et facile à utiliser, tout en réduisant le risque de pannes.**

- **C5 : Le coût**

Le gant doit être proposé à un prix abordable. Un coût trop élevé pourrait dissuader les utilisateurs sourds et muets d'adopter cette solution ; il convient donc de trouver un juste équilibre entre performances et prix.

II.3.1.3 Diagramme FAST :

Après avoir défini les fonctions et les contraintes dans le « diagramme Pieuvre », nous passons maintenant à la traduction de ces besoins en solutions techniques concrètes à l'aide du diagramme FAST. L'objectif ici est de répondre à la question : « Comment les fonctionnalités requises seront-elles mises en œuvre ? » en construisant une séquence logique reliant les objectifs globaux aux solutions techniques.

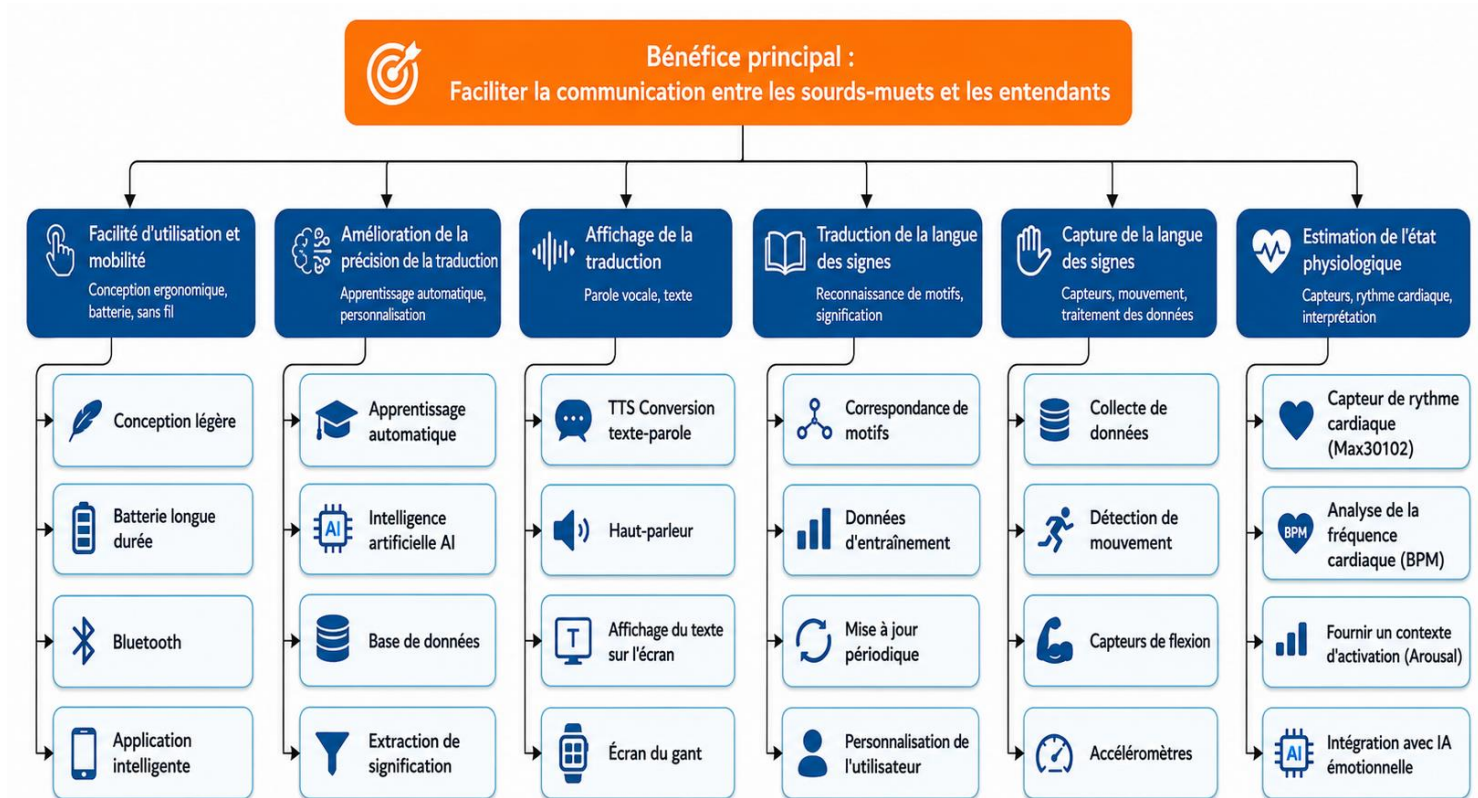


Figure II.8. Diagramme FAST du système de gant intelligent [11].

II.3.2 Fonctions principales et contraintes du système :

II.3.2.1 Fonctions principales La finalité essentielle de ce dispositif est de répondre à la problématique majeure rencontrée par les individus présentant des déficiences sensorielles, notamment auditives, en l'occurrence la difficulté de communication. Cette approche s'appuie sur le consensus établi par de nombreuses études, qui soulignent la nécessité d'une architecture bidirectionnelle capable d'assurer simultanément la collecte d'informations et leur restitution perceptive. Dans ce cadre, la traduction multimodale constitue un élément central du système, reposant sur la capacité à reconnaître les gestes de la langue des signes, qu'ils soient continus ou discontinus, afin de les convertir en texte lisible sans latence. Ce texte est ensuite transformé

en signal sonore grâce à des technologies de synthèse vocale, permettant ainsi une interaction fluide entre les utilisateurs. L'ensemble du dispositif s'inscrit dans une dynamique confirmée par les recherches récentes, lesquelles ont permis de développer des chaînes de traitement intégrées, désormais déployées sous forme d'applications mobiles accessibles et exploitables sur smartphones [2].

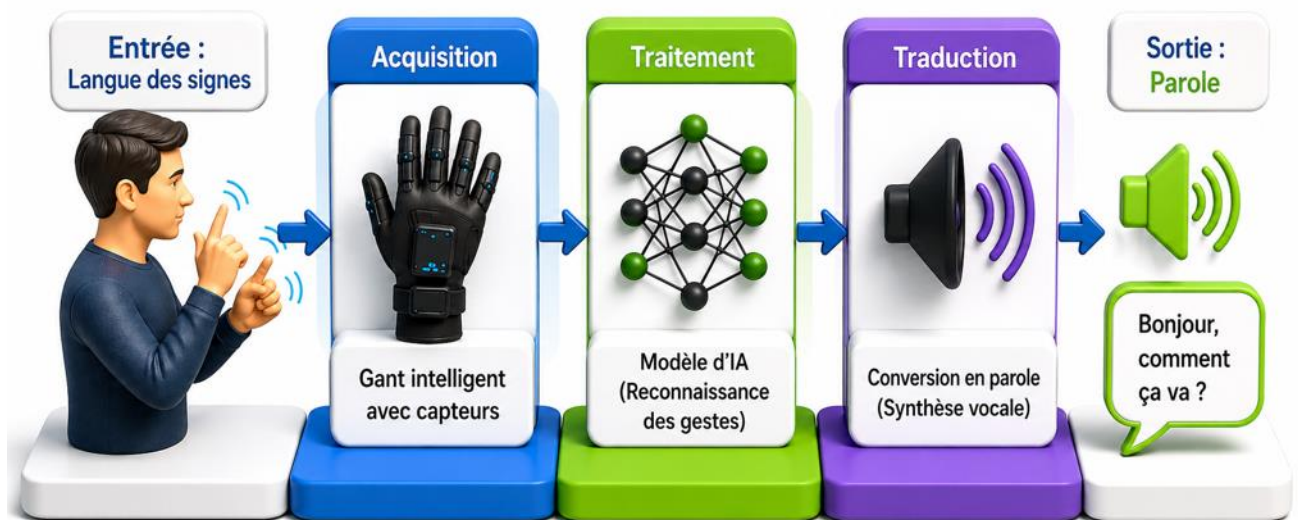


Figure II.9. Traduction multimodale de la langue des signes (Sign-to-Text et Sign-to-Speech) [2].

II.3.2.2 contraintes du système

a) Contraintes techniques :

Afin d'assurer des performances optimales des fonctions susmentionnées, la conception matérielle et logicielle doit satisfaire l'ensemble des exigences rigoureuses de performance établies dans les technologies les plus récentes disponibles. Acquisition cinématique à haute précision :

🚦 Acquisition cinématique de haute précision

L'acquisition des mouvements requiert une configuration hybride de capteurs. La norme technique définie par la recherche impose l'utilisation de quatre capteurs de flexion pour représenter la dynamique des phalanges, associés de manière systématique à une unité de mesure inertielle (IMU), généralement le modèle MPU-6050, afin de déterminer l'orientation spatiale

et l'accélération de la main [4]. Des solutions émergentes particulièrement efficaces préconisent l'emploi de réseaux de capteurs extensibles intégrés dans des fils conducteurs [2].

Traitement IA et précision algorithmique :

Le traitement des signaux bruts nécessite l'utilisation d'architectures d'apprentissage profond spécifiquement conçues pour les séries temporelles. Les architectures de prédilection sont les réseaux Bi-LSTM, RNN-LSTM ou les réseaux convolutifs unidimensionnels (1D-CNN). La précision minimale requise pour qu'une traduction soit exploitable s'établit à 90 %, alors que des systèmes de pointe obtiennent un taux de reconnaissance de 98.63 % sur un vocabulaire étendu comprenant 660 gestes. Les implémentations simplifiées sur des architectures restreintes peuvent présenter des diminutions de précision avoisinant les 77 % [2].

Latence et connectivité :

Pour assurer le bon déroulement des échanges conversationnels ordinaires, les délais de réponse totaux (comprenant la capture, l'inférence et la synthèse vocale) doivent être rigoureusement inférieurs à une seconde [2]. Cela requiert une transmission de données sans fil fiable et à faible latence, assurée par Bluetooth ou au moyen d'une infrastructure Internet des objets (IoT) reposant sur des microcontrôleurs compacts tels que l'ESP32 ou l'Arduino Nano [5].

b) Contraintes ergonomiques :

Le principal enjeu des technologies d'assistance portables réside dans l'inconfort éprouvé par les utilisateurs, qui les conduit à cesser leur utilisation ; ainsi, favoriser une conception plus ergonomique constitue l'un des critères essentiels pour garantir la réussite du projet et engendrer des ventes effectives sur le terrain.

Confort thermique et respirabilité :

Le cahier des charges exige l'emploi d'actionneurs ainsi que de textiles laissant passer l'air. Effectivement, le port prolongé de dispositifs électroniques intégrés à la peau est susceptible d'entraîner une sudation ainsi qu'une sensation d'inconfort [6].

Charge mécanique et maniabilité :

Le dispositif ne doit en aucun cas entraver la dextérité fine intrinsèque de la main. Effectivement, les actionneurs employés doivent présenter une charge mécanique extrêmement réduite, de l'ordre de 10 mN [6]. Par ailleurs, l'intégralité du système (comprenant les circuits

sur circuit imprimé ou les conducteurs filaires) doit assurer une masse réduite afin de ne pas alourdir l'utilisateur [7].

Ajustabilité et usage quotidien :

Les gants doivent être conçus pour s'ajuster à diverses morphologies de mains. Aisé-ment enfilables. Le mécanisme de fixation a été réétudié afin de convenir aux utilisateurs présentant une spasticité ou une mobilité restreinte [7].

c) Contraintes énergétiques :

Le paramètre primordial et déterminant, fréquemment sous-évalué dans les prototypes académiques, correspond au niveau d'autonomie et de viabilité électrique du système ; par conséquent, son dimensionnement matériel doit être systématiquement privilégié lors de la conception.

Autonomie en usage continu et Régulation :

Étant donné que le dispositif vise à compenser un handicap de communication permanent, la capacité de la batterie ainsi que l'optimisation de la consommation énergétique en particulier celle du microcontrôleur (ESP32) et de la transmission Bluetooth doivent assurer un fonctionnement continu pendant plusieurs heures quotidiennes. Pour satisfaire ce compromis entre capacité et légèreté (ergonomie), l'architecture repose sur un accumulateur Lithium-Polymère (Li-Po) monocellulaire de 3.7V. De plus, les capteurs (MAX30102, ADS1115) exigeant une précision stricte, la tension brute de la batterie est stabilisée à 3.3V par un régulateur à faible chute (LDO, type AMS1117-3.3), garantissant ainsi un traitement des signaux exempt de bruit électrique [7].

Durabilité électromécanique :

Les sollicitations mécaniques récurrentes (flexion digitale,) exigent une résistance remarquable des éléments constitutifs. Les actionneurs textiles contemporains doivent être capables de résister à une durée d'utilisation excédant onze millions de cycles de fonctionnement sans présenter de dégradation de leurs performances [6].

Gestion de la recharge et robustesse :

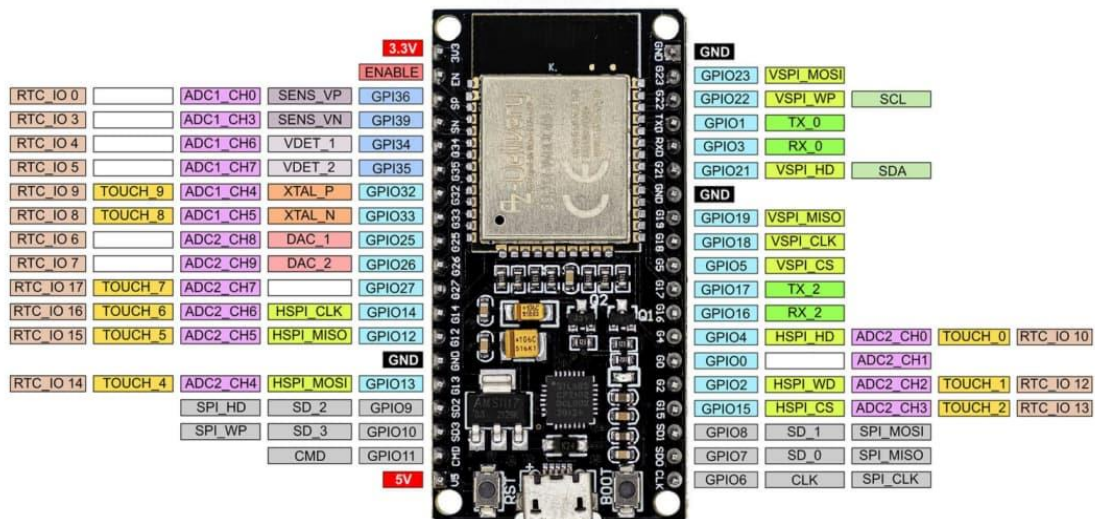
L'architecture matérielle doit intégrer un accès facilité au port de recharge et assurer une robustesse élevée des câblages internes et externes afin de supporter les sollicitations liées à un usage domestique non encadré. Afin de sécuriser cet usage de manière autonome, le circuit intègre un système de gestion de batterie (BMS) basé sur le module TP4056. Ce dernier gère le cycle de recharge via USB et protège matériellement l'utilisateur contre les surcharges, les décharges profondes et les courts-circuits, écartant tout risque de défaillance thermique [7].

II.4 Conception matérielle :

Le choix des composants résulte d'un compromis entre précision, consommation d'énergie et encombrement matériel.

II.4.1 Le Microcontrôleur ESP32

Le choix s'est porté sur une unité capable de gérer plusieurs entrées analogiques et une communication sans fil stable. Bien que l'Arduino soit une référence pour le prototypage rapide [11], l'utilisation de l'ESP32 ou d'un processeur couplé est privilégiée pour les modèles d'IA complexes. « Le système acquiert les données via un ESP32, tandis qu'un Raspberry Pi ou un smartphone est utilisé pour exécuter le modèle de Machine Learning basé sur des algorithmes de type Random Forest ou LSTM ». **Quelle est la raison pour laquelle un Arduino Uno ou un Arduino Nano ne serait pas utilisé ?** La limitation principale de l'Arduino réside dans sa capacité de traitement restreinte, due à son architecture à huit bits. Par ailleurs, il requiert un module Bluetooth externe, tel que le HC-05, caractérisé par un encombrement important et une consommation énergétique élevée. L'ESP32 possède une architecture 32 bits nettement plus performante. Il intègre en outre de manière native la technologie Bluetooth Low Energy, ce qui



s'avère particulièrement utile. Cette méthode autorise la mise en œuvre d'une stratégie dénommée Offloading, qui consiste à transférer les calculs intensifs vers le téléphone afin de préserver l'autonomie énergétique du gant. Cette méthode satisfait deux exigences majeures : la compatibilité avec les smartphones, désignée par C1, et l'autonomie, désignée par C4 [12].

Figure II.10. ESP32-WROOM-32 PINS [15].

II.4.2 Capteurs de flexion (Flex Sensors) :

Pour la détection du pliage des doigts, le capteur de flexion résistif de Spectra Symbol a été retenu. Sa résistance à plat est d'environ 10K Ohms et augmente au minimum de deux fois lorsqu'il est plié à 180°. Ce capteur est choisi pour son « profil bas et sa longue durée de vie (plus d'un million de cycles), ce qui le rend idéal pour une intégration durable dans un gant textile. Les capteurs piézorésistifs disponibles dans le commerce présentent une linéarité remarquable à un coût particulièrement réduit. Lorsqu'ils sont soudés directement sur un circuit imprimé afin d'éliminer les mauvais contacts, ils présentent le meilleur compromis entre précision et coût, satisfaisant ainsi l'exigence C5 relative à un prix abordable [18].

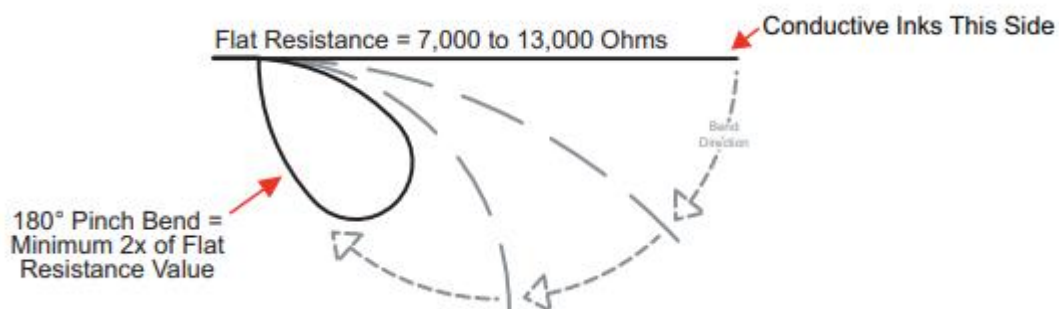


Figure II.11. Comment fonctionne un capteur de flexion [18].

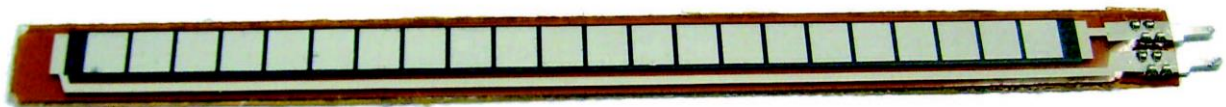


Figure II.12. Capteur De Flexion [18].

II.4.3 Module inertiel (MPU-6050) :

Pour suivre l'orientation de la main dans l'espace (roulis, tangage, lacet), le MPU-6050 est le standard industriel. C'est un dispositif de suivi de mouvement à 6 axes qui combine un gyroscope à 3 axes et un accéléromètre à 3 axes. Il permet de distinguer deux signes ayant la

même configuration de doigts mais une orientation différente, grâce à son processeur de mouvement numérique (DMP) qui filtre les bruits de mesure. Le MPU-6050 mesure l'accélération linéaire ainsi que la rotation angulaire (gyroscope) selon six axes dans un espace tridimensionnel. Par ailleurs, il utilise une communication via le bus I2C, à l'instar de l'ADS1115, ce qui permet d'interconnecter ces deux modules essentiels en n'utilisant que deux fils de données (SDA et SCL), satisfaisant ainsi la contrainte C6 relative à la miniaturisation et à l'intégration, tout en nécessitant à la fois une haute précision et une analyse rapide [19].

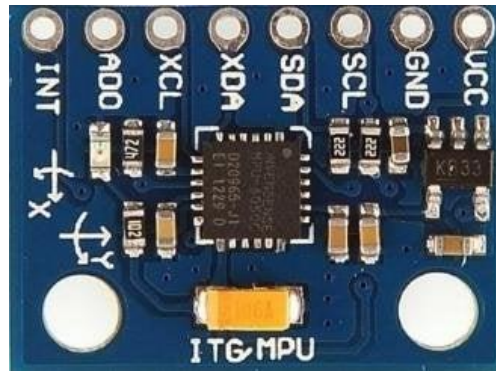


Figure II.13. MPU -6050 [25].

II.4.4 Convertisseur analogique-numérique (ADC) ADS 1115 :

Le principe fondamental de la haute précision. Les convertisseurs analogique-numérique intégrés de l'ESP32 présentent une sensibilité élevée au bruit ainsi qu'une caractéristique de réponse non linéaire. Lorsqu'il est procédé à la connexion directe des capteurs de flexion à ces broches, les valeurs mesurées présentent une grande instabilité et varient constamment. Cela peut engendrer des difficultés pour les systèmes d'intelligence artificielle. L'ADS1115 constitue un dispositif hautement précis destiné à la conversion des signaux analogiques en signaux numériques. Il est capable de discerner 65 536 niveaux distincts, soit un nombre largement supérieur aux 4 096 niveaux détectables par l'ESP32. L'ADS1115 est élaboré afin de réduire les perturbations électromagnétiques et d'assurer des mesures hautement stables de la flexion des doigts. Cela revêt une importance capitale pour répondre à la contrainte C2, laquelle porte simultanément sur la précision et la rapidité de l'analyse.

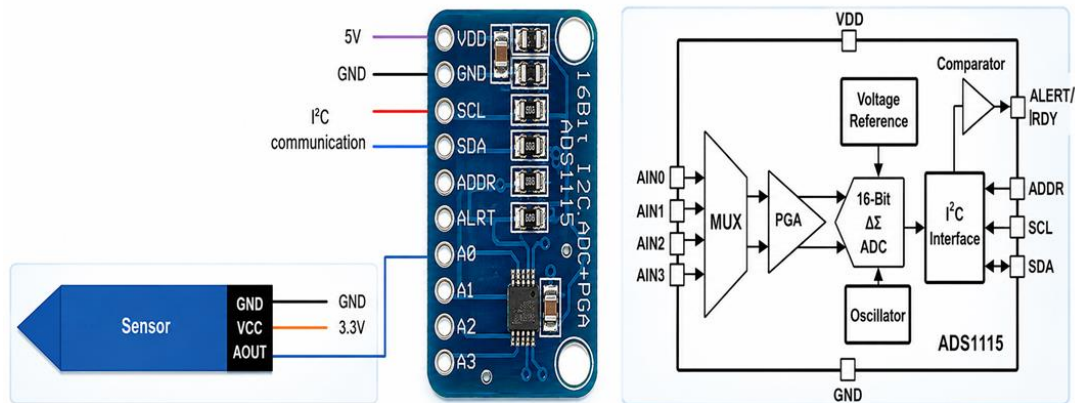


Figure II.14. ADS 1115 [26].

II.4.5 Module PPG (Photoplethysmography) MAX30102:

Le MAX30102 est un module intégré de surveillance de l'oxymétrie de pouls et de la fréquence cardiaque. Il intègre des LED internes, des photodétecteurs, des composants optiques ainsi qu'une électronique à faible bruit avec rejet de la lumière ambiante. Le MAX30102 fournit une solution complète destinée aux dispositifs mobiles et portables, fonctionnant avec une alimentation principale de 1.8 V et une alimentation séparée de 3.3 V pour les LED internes. La communication avec le microcontrôleur s'effectue via une interface standard compatible I2C, tandis qu'un mode basse consommation permet d'optimiser l'autonomie énergétique du système [28]. Le signal photopléthysmographiques (PPG) acquis par le MAX30102 demeure fortement sensible aux artefacts de mouvement générés par les gestes de la main et du poignet. Afin d'améliorer la robustesse de l'estimation physiologique, notre système implémente un mécanisme logiciel nommé « Idle Lock ». Celui-ci suspend automatiquement l'acquisition physiologique lorsque l'écart-type de l'accélération mesurée par le MPU-6050 dépasse un seuil de 0.35g. Les paramètres cardiaques (fréquence cardiaque - BPM) sont alors extraits uniquement durant les phases de stabilité gestuelle. La valeur obtenue est ensuite définie comme une ligne de base physiologique de référence permettant au système de distinguer les variations physiologiques réelles des perturbations mécaniques induites par les mouvements de la main [45].

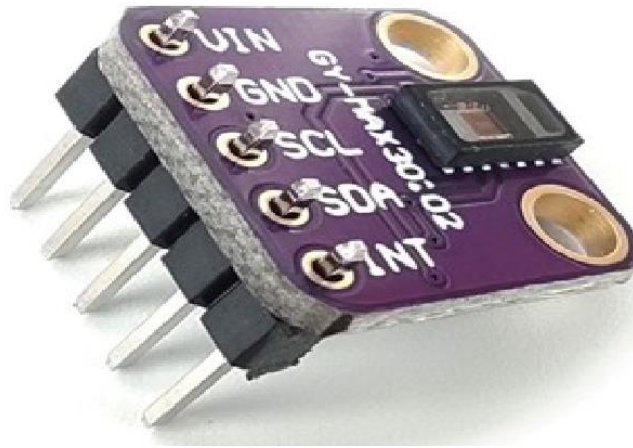


Figure II.15. MAX30102 [29].

II.4.6 Batterie lithium-ion :

La batterie lithium-ion est une technologie de stockage d'énergie rechargeable largement utilisée dans les systèmes électroniques modernes en raison de sa haute densité énergétique et de sa grande efficacité. Elle fonctionne grâce au déplacement réversible des ions lithium entre l'anode et la cathode à travers un électrolyte lors des cycles de charge et de décharge. Sa tension nominale est d'environ 3.7 V par cellule, avec une tension de charge maximale de 4.2 V, ce qui la rend adaptée aux applications nécessitant une alimentation compacte et performante. En plus de sa faible autodécharge et de l'absence d'effet mémoire significatif, elle offre une bonne durée de vie en nombre de cycles. Cependant, son utilisation nécessite un système de gestion de batterie (BMS) afin de garantir la sécurité contre les risques de surcharge, de décharge profonde et d'emballement thermique. Grâce à ses performances, elle est aujourd'hui largement intégrée dans les dispositifs portables, les systèmes embarqués, les applications IoT ainsi que les systèmes de stockage d'énergie renouvelable.

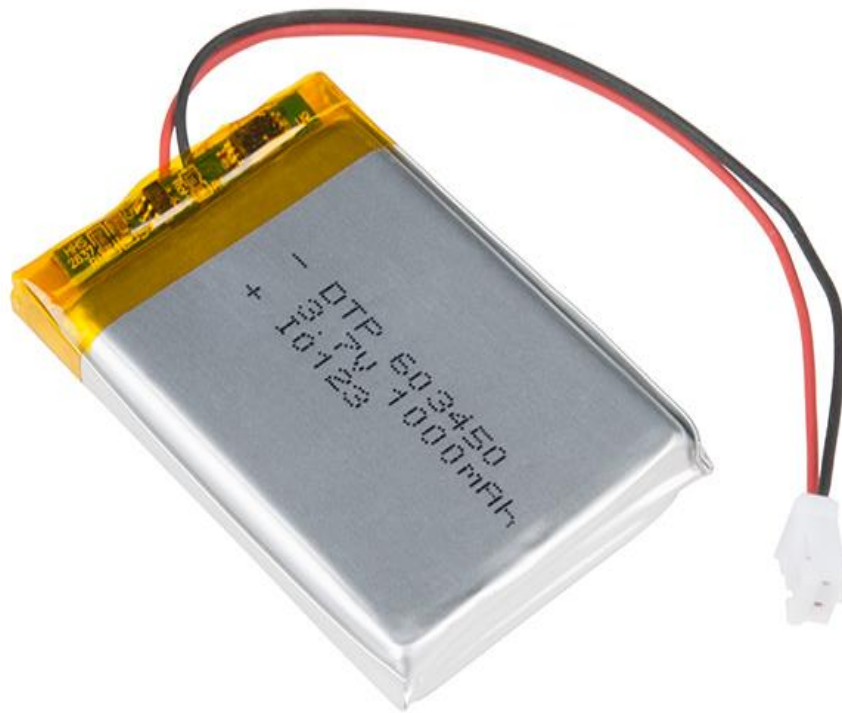


Figure II.16. Batterie lithium-ion 1000 mAh [41].

II.4.7 TP4056:

Le **TP4056** est un circuit intégré de gestion de charge largement utilisé pour la recharge des batteries lithium-ion à une cellule. Il permet de charger une batterie de 3,7 V jusqu'à 4,2 V en appliquant un algorithme de charge en courant constant puis en tension constante, garantissant ainsi une recharge sûre et efficace. Généralement disponible sous forme de module prêt à l'emploi, le TP4056 intègre également des protections contre la surcharge, la sur décharge et parfois les courts-circuits, notamment dans les versions équipées de circuits de protection supplémentaires. Il est très utilisé dans les projets électroniques embarqués, les systèmes IoT et les dispositifs portables alimentés par batterie lithium-ion, en raison de sa simplicité d'utilisation et de son faible coût. Grâce à sa compatibilité avec une alimentation USB 5 V, il constitue une solution pratique et fiable pour la recharge des systèmes autonomes à basse consommation.



Figure II.17. Circuit intégré de gestion de charge TP4056 [42].

II.5 Technologies de communication et flux de données :

L'architecture de communication de notre système repose sur un modèle hybride (Edge-to-Cloud) garantissant à la fois une réactivité en temps réel pour l'utilisateur local et une connectivité distante pour des services avancés. L'exploitation du microcontrôleur ESP32, intégrant nativement des modules radio double bande, permet de dissocier efficacement la transmission locale (Bluetooth) de la communication réseau (Wi-Fi).

II.5.1 Communication locale sans fil via Bluetooth Low Energy (BLE)

Pour la transmission des paquets de données entre le gant intelligent et l'application mobile (Unity 3D), le protocole Bluetooth Low Energy (BLE) a été sélectionné. Contrairement au Bluetooth classique (Bluetooth Classic), le BLE est spécifiquement conçu pour les réseaux corporels sans fil (Wireless Body Area Networks - WBAN) et les dispositifs portables (wearables) nécessitant une consommation énergétique extrêmement faible [35]. Dans notre architecture logicielle, l'ESP32 est configuré en tant que Serveur Périphérique (GATT Server). Il expose différentes "Caractéristiques" (Characteristics) correspondant aux flux prétraités des capteurs (les angles de flexion, les quaternions du MPU-6050, et les données physiologiques du MAX30102). L'application smartphone sous Unity agit comme un Client Central qui s'abonne à ces caractéristiques via un mécanisme de notifications (Notify). Cette méthode de

transmission "Push" garantit que le smartphone reçoit les données uniquement lorsqu'un nouveau mouvement est détecté, minimisant ainsi la surcharge de la bande passante et préservant l'autonomie de la batterie du gant. La faible latence du BLE (inférieure à 10 millisecondes) est vitale pour assurer une synchronisation parfaite entre le geste physique et l'animation de l'avatar 3D sur l'écran [36].

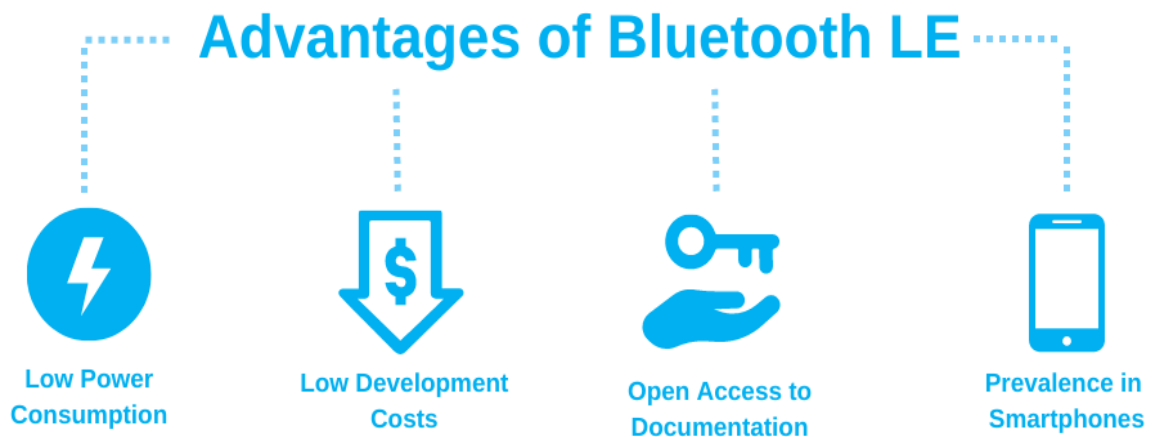


Figure II.18. Avantages du Bluetooth Low Energy [43].

II.5.2 Connectivité Cloud et communication via Serveur (Wi-Fi)

Bien que le traitement local (Edge Computing) via l'application mobile permette une traduction autonome, l'intégration d'une connectivité réseau élargit considérablement les capacités du système. L'utilisation du module Wi-Fi (soit nativement via l'ESP32, soit relayé par la connexion internet du smartphone) permet de basculer vers une architecture orientée Internet des Objets (IoT) et Cloud Computing [37]. Cette connectivité distante remplit trois fonctions majeures dans notre projet :

- ❖ **Déchargement calculatoire (Offloading) et Inférence Hybride :** Si les modèles d'intelligence artificielle (CNN-LSTM) deviennent trop lourds pour le processeur du smartphone, les vecteurs de caractéristiques (Features) peuvent être transmis à un serveur distant puissant via le protocole WebSocket, permettant une inférence bidirectionnelle en temps réel. Afin de garantir une robustesse face aux variations de latence réseau (Round-Trip Time) et aux instabilités de connexion, l'architecture déploie un paradigme d'Edge Intelligence. Le moteur principal d'inférence s'exécute localement (Edge) pour respecter la contrainte critique du temps réel. Le déchargement Cloud (Offloading) opère de manière com-

plémentaire : il agit comme un mécanisme de repli adaptatif et prend en charge les opérations asynchrones exigeantes (comme la ré-optimisation des poids du modèle par apprentissage fédéré), sécurisant ainsi la continuité de service sans compromettre l'expérience utilisateur [39].

- ❖ **Suivi télémétrique et Télésanté :** Les données physiologiques (fréquence cardiaque) et l'historique des états émotionnels (Arousal/Valence) peuvent être poussés via des requêtes HTTP/REST vers une base de données cloud (ex : Firebase). Cette ségrégation architecturale entre le flux de traduction spatio-temporel (Edge) et le flux télémétrique (Cloud) empêche la congestion de la bande passante locale. Cela permet un suivi psychologique à long terme ou l'intervention d'un proche/médecin en cas de détresse détectée, de manière totalement transparente [40].
- ❖ **Communication interpersonnelle à distance :** Le serveur agit comme un pont de communication (Broker). Le texte traduit par le système n'est plus seulement affiché sur l'écran local, mais peut être envoyé instantanément (via Internet) au smartphone d'un interlocuteur distant, transformant le gant en un véritable outil de télécommunication pour les personnes sourdes et malentendantes [37-38].

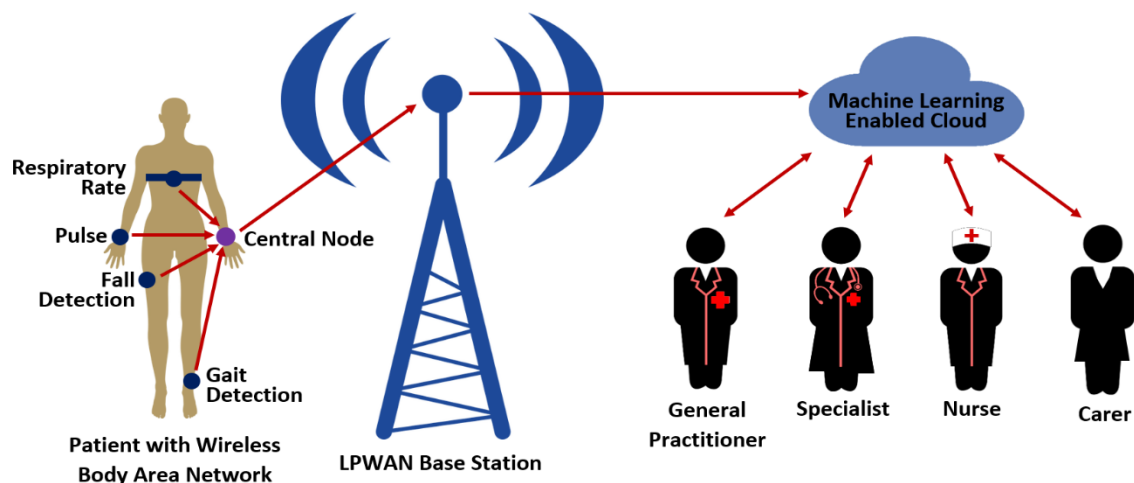


Figure II.19. Aperçu du modèle proposé [38].

II.6 Partie logicielle :

II.6.1 Programmation du microcontrôleur ESP32 via Arduino IDE :

L'Arduino IDE (Integrated Development Environment) est l'environnement de développement officiel utilisé pour programmer les cartes Arduino. Il permet d'écrire, compiler et transférer du code directement sur le microcontrôleur de la carte via une interface USB. L'IDE est compatible avec plusieurs systèmes d'exploitation (Windows, Linux, macOS) et utilise un langage de programmation basé sur le C/C++. L'interface de l'IDE est simple et épurée, conçue

pour rendre la programmation accessible aux débutants tout en offrant une grande flexibilité aux utilisateurs avancés. Elle comprend un éditeur de code, un moniteur série, un compilateur intégré et un système de gestion des bibliothèques.

- **Caractéristiques principales**

- Langage utilisé : Arduino (basé sur le langage C/C++).
- Fonctions principales : `setup()` pour l'initialisation, `loop()` pour l'exécution continue du programme.
- Possibilité de charger et gérer facilement des bibliothèques externes (RF24, DHT, SoftwareSerial...).
- Support natif des cartes Arduino (Uno, Mega, Nano...) et autres microcontrôleurs compatibles [30].
- Outils intégrés : Moniteur série, gestionnaire de ports, gestionnaire de bibliothèques.



Figure II.20. Vue générale de l'environnement Arduino IDE [30].

II.6.2 Gestion des bibliothèques (Flex sensors, MPU-6050 et MAX30102) :

L'exploitation efficace des données brutes issues des différents capteurs sur la plateforme **ESP32** nécessite l'intégration de bibliothèques logicielles spécifiques (drivers). Ces bibliothèques assurent une couche d'abstraction matérielle, facilitant la communication via les protocoles **I2C** et **Analogique** tout en optimisant le traitement des signaux [30].

- **Capteurs de Flexion** : Bien que ces capteurs reposent sur une lecture analogique directe via les broches **ADC (Analog-to-Digital Converter)** de l'ESP32, une gestion logicielle est nécessaire pour la linéarisation des signaux. Nous utilisons des fonctions de **mappage et calibration** personnalisées pour convertir les variations de résistance en angles de courbure précis (en degrés), permettant ainsi de modéliser l'état d'ouverture ou de fermeture des doigts.
- **Centrale Inertielle MPU-6050** : Pour ce module, la bibliothèque **Adafruit_MPU6050** (basée sur la bibliothèque Adafruit_BusIO) a été privilégiée. Elle permet d'accéder de manière simplifiée aux registres de l'accéléromètre et du gyroscope via le bus I2C. Cette bibliothèque est cruciale pour récupérer les vecteurs de rotation et d'accélération nécessaires à la reconstruction de la trajectoire spatiale de la main dans l'espace 3D.
- **Capteur de Rythme Cardiaque MAX30102** : La gestion de ce capteur repose sur la bibliothèque **SparkFun_MAX3010x**. Elle est indispensable pour configurer les registres internes du capteur (courant des LED rouges et infrarouges, taux d'échantillonnage). Elle permet d'extraire les données brutes de photopléthysmographie (PPG) qui seront ensuite traitées pour calculer la fréquence cardiaque et l'intervalle entre les battements (R-R interval), données essentielles pour l'analyse de l'état émotionnel de l'utilisateur [30].

II.6.3 Algorithme d'acquisition et de prétraitement des signaux :

L'algorithme implémenté sur l'ESP32 constitue le premier étage de traitement du système. Son rôle est de transformer les signaux analogiques et numériques bruts en un flux de données structuré et exploitable par le modèle de Deep Learning. Ce processus se décompose en trois phases majeures :

1. Phase d'Acquisition Multimodale :

Le système opère à une fréquence d'échantillonnage constante ($f_s \approx 50$ Hz) pour garantir la capture fidèle de la dynamique gestuelle. L'algorithme interroge simultanément :

- Les 5 entrées ADC pour les capteurs de flexion.
- Le bus I2C pour extraire les vecteurs $a_{x,y,z}$ (accélération) et $g_{x,y,z}$ (rotation) du MPU-6050.

- Le buffer FIFO du MAX30102 pour les données de photo pléthysmographie [30].

2. Prétraitement et Filtrage (Denoising) :

Afin d'éliminer le bruit électronique et les micro-tremblements parasites, les données subissent un traitement numérique immédiat :

- Filtrage par Moyenne Glissante (Moving Average) : Appliqué aux capteurs de flexion pour stabiliser les mesures de courbure.
- Normalisation : Les valeurs brutes sont mises à l'échelle (Scaling) dans une plage standardisée [0, 1] ou [-1, 1]. Cette étape est cruciale pour accélérer la convergence du modèle CNN-LSTM en évitant la saturation des gradients.

3. Segmentation par Fenêtrage Glissant (Sliding Window) :

La reconnaissance de gestes étant temporelle, l'algorithme ne traite pas les points isolément. Il regroupe les données dans des fenêtres temporelles (par exemple 2 secondes de données). Chaque fenêtre glisse avec un recouvrement (Overlap) de 50%, créant ainsi une continuité temporelle indispensable pour que la couche LSTM puisse mémoriser le contexte du signe effectué. Le paquet de données final est ensuite encapsulé pour être transmis via le protocole de communication choisi [30].

Le fonctionnement optimal du modèle DeepConvLSTM repose sur une fréquence d'échantillonnage strictement déterministe de 50 Hz. Cependant, l'interrogation séquentielle des différents capteurs (ADS1115, MPU6050 et MAX30102) via le bus I2C partagé de l'ESP32 peut introduire une gigue temporelle (temporal jitter) susceptible de dégrader la cohérence des séquences temporelles. Afin de réduire ce goulot d'étranglement, notre architecture exploite l'architecture dual-core du microcontrôleur ESP32 à travers le système d'exploitation temps réel FreeRTOS. Le processus d'acquisition des données capteurs (I2C polling) est assigné exclusivement au Core 0 afin de garantir un intervalle temporel constant de 20 ms entre chaque trame d'acquisition. En parallèle, les tâches asynchrones et fortement consommatrices de ressources notamment la communication Wi-Fi et la transmission MQTT vers le Cloud sont

exécutées indépendamment sur le Core 1. Cette séparation stricte des tâches permet de limiter les interférences temporelles, d'éviter les pertes de paquets et de préserver l'intégrité de la fenêtre glissante de 20 trames (400 ms) utilisée par le modèle DeepConvLSTM [44].

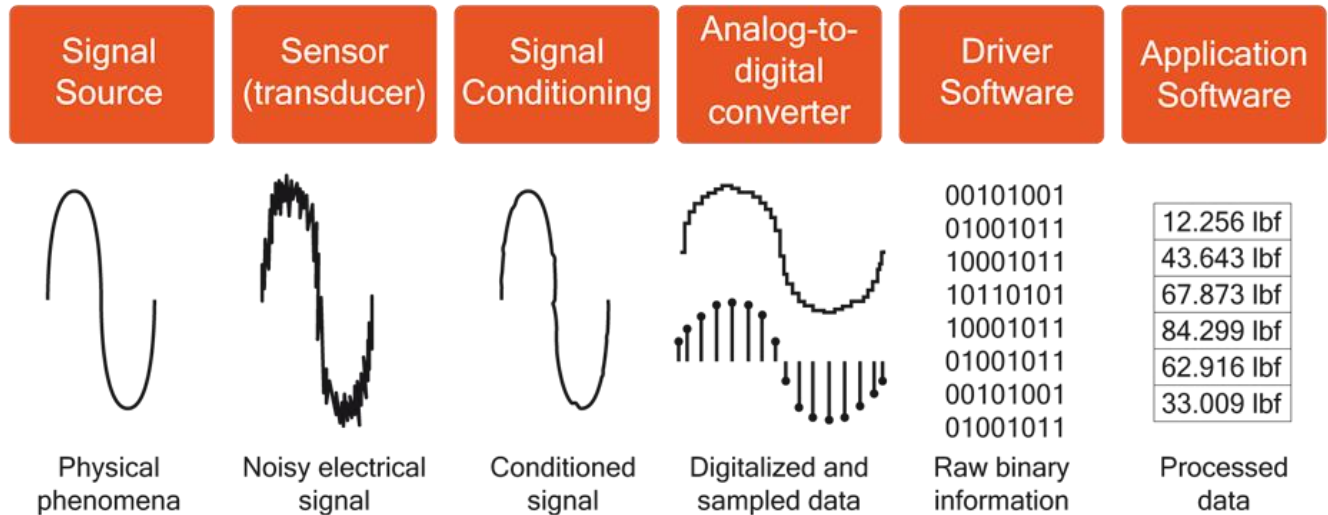


Figure II.21. Schéma de conditionnement du signal [31].

II.6.4 Architecture Logicielle et Stratégie de Fusion Multimodale :

Notre système repose sur une architecture de fusion multimodale combinant les données physiologiques du capteur MAX30102 et les données inertielles du MPU-6050 afin d'améliorer la robustesse et le contexte de la traduction gestuelle. Le signal photopléthysmographique (PPG) acquis par le MAX30102 permet d'estimer la fréquence cardiaque (BPM), tandis que le MPU-6050 fournit les informations liées aux mouvements de la main et du poignet. Toutefois, les systèmes PPG embarqués demeurent fortement sensibles aux artefacts de mouvement générés lors des gestes dynamiques. Afin de limiter ces perturbations, notre architecture implémente une stratégie de Sensor Fusion utilisant les données d'accélération comme référence dynamique de mouvement. Un mécanisme logiciel nommé « Idle Lock » suspend temporairement l'estimation physiologique lorsque l'intensité cinématique dépasse un seuil prédéfini, garantissant ainsi une acquisition cardiaque plus stable. L'interprétation contextuelle du comportement utilisateur repose ensuite sur le modèle circomplexe des émotions de Russell. L'axe d'activation physiologique (Arousal) est estimé via la fréquence cardiaque, tandis que l'intensité comportementale est approximée à partir de la magnitude de l'accélération. Ces informations sont croisées par le serveur backend afin d'identifier plusieurs états contextuels tels

que l'urgence, l'anxiété, l'enthousiasme ou le calme. Enfin, ce contexte physiologique et comportemental est intégré au prompt du modèle Llama 3.1 afin de produire une traduction textuelle et vocale davantage cohérente avec l'intention gestuelle de l'utilisateur[47-48].

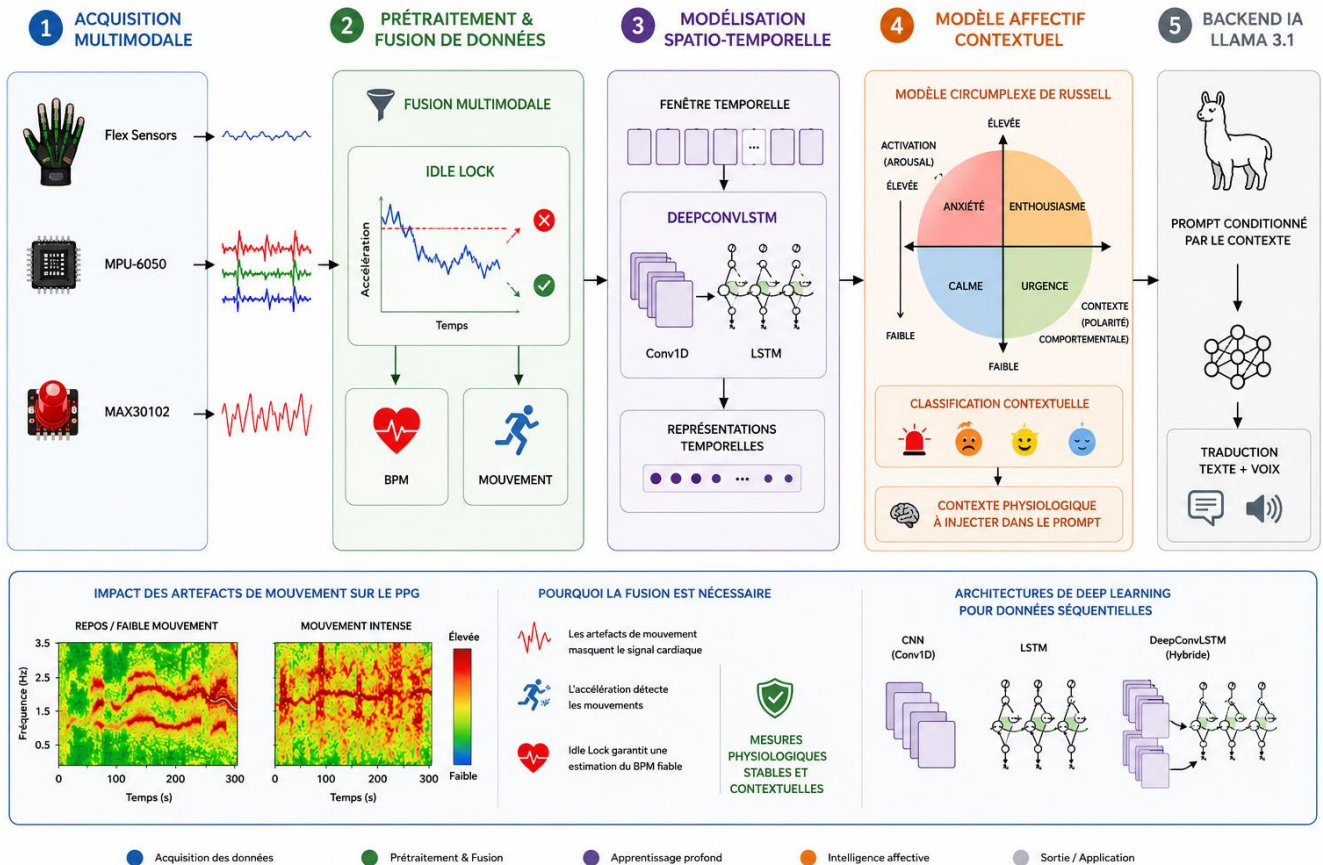


Figure II.22. Architecture Hybride de Fusion Multimodale et Intelligence Affective pour la Traduction Gestuelle Contextuelle [47-48].

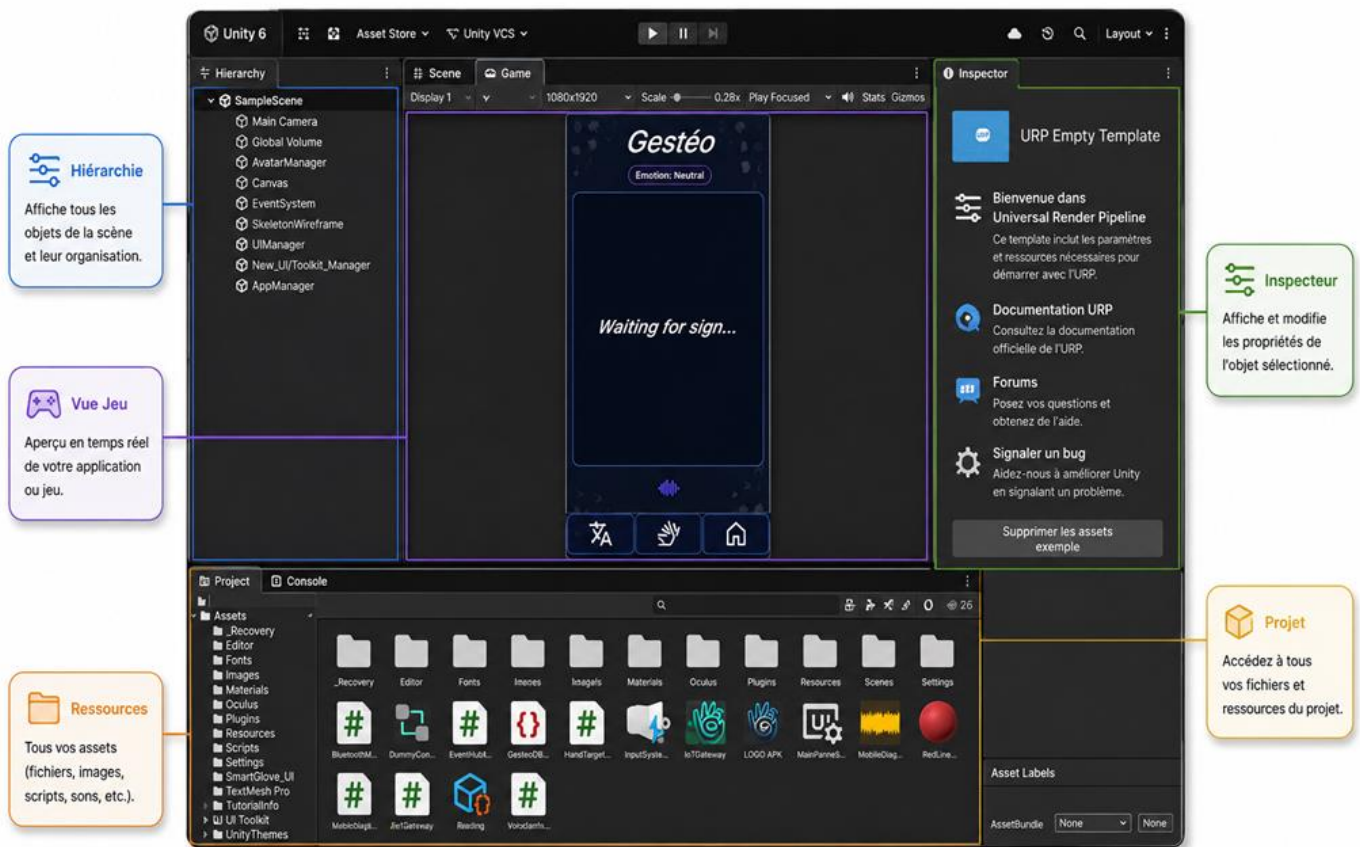
II.6.5 Environnement de développement Unity 3D :

L'application mobile développée pour ce projet a été conçue à l'aide de Unity 3D, un moteur de rendu temps réel et un environnement de développement intégré (IDE) multiplateforme. Bien qu'historiquement reconnu dans l'industrie du jeu vidéo, Unity s'impose aujourd'hui comme un outil de référence pour la simulation biomécanique et le développement de technologies d'assistance. Le langage de programmation utilisé au sein de cet environnement est le C# (C-Sharp). C'est un langage orienté objet puissant et robuste, recommandé pour ce type d'architecture. Le C# se distingue par sa rapidité d'exécution, sa gestion efficace de la mémoire et sa capacité à traiter facilement les flux de données asynchrones, ce qui est indispensable pour recevoir les signaux continus envoyés par les capteurs du gant [44].

L'environnement Unity permet notamment de :

- Intégrer et animer des modèles 3D (comme le squelette virtuel d'une main) en temps réel, en fonction des données cinématiques reçues des capteurs de flexion et du MPU-6050.
- Concevoir des interfaces graphiques (UI) interactives pour l'affichage des textes traduits et des états émotionnels de l'utilisateur. Suivre et déboguer le code en temps réel directement via la console de l'éditeur lors de l'exécution.
- Gérer les permissions et la communication (Bluetooth et Wi-Fi) et compiler directement le projet en une application mobile native (fichier .apk) via son module Android Build Support.

Dans le cadre de ce système de traduction gestuelle, Unity 3D a permis de concevoir une application immersive, fluide et adaptée aux besoins de communication des personnes malentendantes. Contrairement à une simple interface textuelle, elle offre une restitution visuelle instantanée des mouvements de la main tout en gérant une communication directe avec



l'ESP32.

Figure II.23. Interface de l'environnement de développement Unity 3D.

II.6.6 Intégration du module de traduction et interface utilisateur (UI) :

L'intégration du module d'intelligence artificielle au sein de l'application Unity constitue la passerelle entre l'acquisition matérielle et l'interaction humaine. L'application mobile n'est pas qu'un simple afficheur ; elle orchestre la réception des données, leur éventuel transfert vers le serveur d'inférence (Wi-Fi), et le décodage des résultats retournés par le modèle pré-entraîné (CNN-LSTM).

a) Logique d'intégration du module de traduction :

Un gestionnaire de traduction (Translation Manager), codé en C#, est implémenté pour lier les prédictions du réseau de neurones à une base de données lexicale locale. Lorsqu'une séquence gestuelle est identifiée, le système reçoit un identifiant de classe (Class ID). Le script interroge alors un dictionnaire (ex : format JSON) pour extraire la chaîne de caractères correspondante (par exemple : "Bonjour", "Merci"). Parallèlement, le système analyse le niveau d'activation physiologique (Arousal) issu du MAX30102 pour moduler l'affichage de l'état émotionnel associé au signe.

b) Conception de l'Interface Utilisateur (UI) :

- **L'interface graphique (User Interface)** a été développée via le système de Canvas d'Unity et l'outil TextMeshPro pour garantir un rendu textuel net sur différentes résolutions d'écrans de smartphones. L'ergonomie a été pensée pour répondre aux critères d'accessibilité cognitive, en divisant l'écran en trois zones fonctionnelles distinctes :
- **Zone de biofeedback visuel (Vue 3D) :** Située au centre de l'écran, elle affiche l'avatar de la main virtuelle s'animant en temps réel, permettant à l'utilisateur de s'assurer que ses mouvements sont correctement captés par le gant.
- **Console de traduction sémantique :** Une zone de texte dynamique de grande taille affichant la traduction littérale du signe en temps réel. Elle intègre également une fonctionnalité de synthèse vocale (Text-to-Speech) exploitant les API natives d'Android pour vocaliser le texte traduit.
- **Tableau de bord affectif et télémétrie :** Un panneau d'information affichant les métriques physiologiques (fréquence cardiaque en BPM) et une jauge visuelle (ou un code couleur) représentant l'état émotionnel déduit, offrant ainsi

une dimension prosodique à la traduction. On y retrouve également les indicateurs d'état du système (connexion Bluetooth/Wi-Fi, niveau de batterie de l'ESP32).

L'ensemble de ces scènes, scripts C#, modèles 3D et éléments graphiques est finalement packagé à travers le processus de compilation Android Build d'Unity, générant un fichier APK (Android Package) autonome, prêt à être déployé et exécuté sur le smartphone de l'utilisateur.

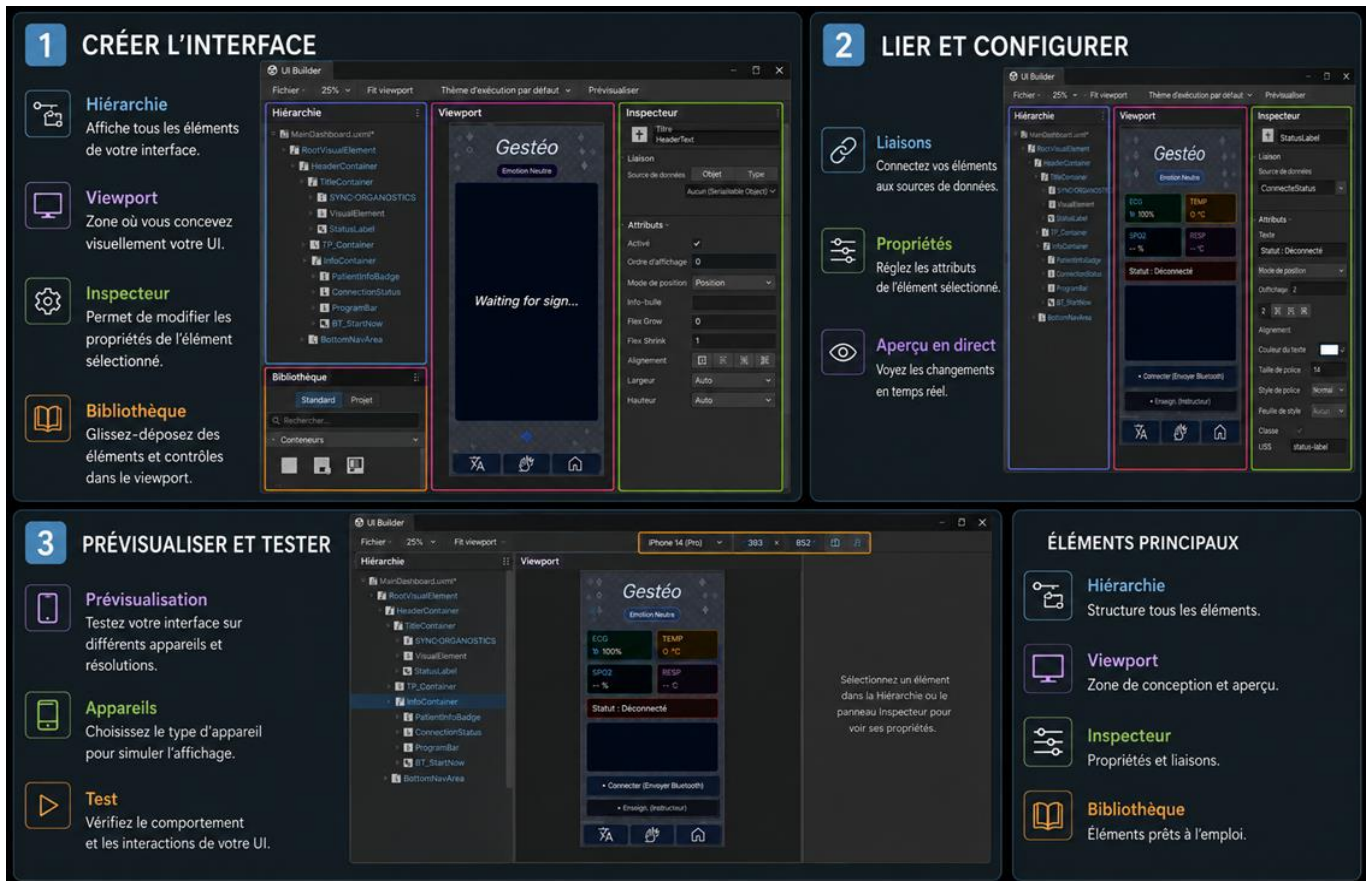


Figure II.24. Capture d'écran de l'interface utilisateur (UI) de l'application mobile développée sous Unity, illustrant la zone de traduction textuelle, l'avatar 3D et le retour physiologique.

II.7 Conclusion :

Ce chapitre a permis d'établir de manière exhaustive la conception architecturale, matérielle et logicielle de notre système de gant intelligent. L'analyse fonctionnelle a été déterminante pour cerner les attentes des utilisateurs malentendants, traduisant le besoin vital de communication en un cahier des charges rigoureux englobant des contraintes de latence, de précision algorithmique, de miniaturisation et d'autonomie énergétique. Pour répondre à ces exigences, notre choix s'est porté sur une architecture matérielle optimisée autour du microcontrôleur ESP32, couplé à une panoplie de capteurs de haute précision (Flex, MPU-6050, ADS1115, MAX30102). Au niveau logiciel, l'implémentation d'une stratégie de fusion de données multimodale (Sensor Fusion avec mécanisme "Idle Lock") permet de garantir la stabilité de l'estimation physiologique face aux artefacts de mouvement, enrichissant ainsi la traduction d'une véritable dimension affective. Enfin, l'adoption d'un paradigme de communication hybride Edge-to-Cloud (BLE et Wi-Fi) couplé à une interface utilisateur immersive développée sous Unity 3D, assure une interaction bidirectionnelle fluide en temps réel. Cette conception technique globale et justifiée constitue ainsi une base théorique et structurelle solide pour aborder la phase d'implémentation pratique, les tests et l'évaluation des performances du prototype, qui feront l'objet du chapitre suivant.

Bibliographie Chapitre II :

- [1] Ozioko, O., & Dahiya, R. (2022). Smart tactile gloves for haptic interaction communication and rehabilitation. *Advanced Intelligent Systems*.
- [2] Zhou, Y., Chen, J., & al., e. (2020). Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. *Nature Electronics*.
- [3] Ozioko, O., Karipoth, P., Hersh, M., & Dahiya, R. (2020). Wearable assistive tactile communication interface based on integrated touch sensors and actuators. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.
- [4] Dalal, U., Thadhani, A., Upadhye, M., Shah, S., Narvekar, M., & Patil, N. (2024). Deep learning-enabled smart glove for real-time sign language translation. *Journal of Electrical Systems*.
- [5] Girish Gowda, Y. A., & al., e. (2025). Smart glove for sign language translation using IoT. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering (IJIREEICE)*.
- [6] Li, Z., Ma, Y., Zhang, K., Wan, J., Zhao, D., Pi, Y., . . . Zhong, J. (2023). Air permeable vibrotactile actuators for wearable wireless haptics. *Advanced Functional Materials*.
- [7] Powell, D., McGregor, S. J., & al., e. (2022). Design of a wearable vibrotactile stimulation device for individuals with upper-limb hemiparesis and spasticity. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.
- [8] MAHOUI, K. (2024). Analyse fonctionnelle. Université de Béjaïa.
- [9] Petit, J. F. (2015). Analyse fonctionnelle : Méthodes et outils. Paris : Dunod.
- [10] Belhadj, A., & Khobzaoui, M. (2020). Étude et conception d'une chaise monte-escalier pour personnes à mobilité réduite. Université d'Ain Témouchent.
- [11] Aïssa El Bey, M. L. (2021). Méthodologie de conception. École Nationale Polytechnique (ENP), Algérie.
- [12] T. Benbrahim, "A Smart Glove Empowered by AI embedded on a Raspberry Pi", *Projet de fin d'études, École Nationale Supérieure de Technologies Avancées (ENSTA) Alger*, 2024.
- [13] P. Vora, B. Parab, R. Naik, R. Bandgar, et J. Dange, "Multi-Purpose Smart Glove for Differently Abled Community People", *International Journal of Research Publication and Reviews*, Vol. 3, No. 1, pp. 1101-1111, Jan. 2022.

- [14] U. Dalal, A. Thadhani, M. Upadhye, S. Shah, M. Narvekar, et N. Patil, "Deep Learning-Enabled Smart Glove for Real-Time Sign Language Translation", *J. Electrical Systems*, Vol. 20, No. 10s, pp. 4874-4882, 2024.
- [15] S. Maharjan, S. Shrestha, et S. Fernando, "Sign Language Detection and Translation using Smart Glove", *Assistive Technology Research Group, London Metropolitan University*, 2023.
- [16] O. Ozioko, W. Navaraj, M. Hersh, et R. Dahiya, "Tacsac: A Wearable Haptic Device with Capacitive Touch-Sensing Capability", *Sensors*, Vol. 20, No. 4780, 2020.
- [17] V. de Vlam, M. Wiertlewski, et Y. Vardar, "Focused vibrotactile stimuli from a wearable sparse array of actuators", *IEEE Transactions on Haptics*, Vol. 16, No. 4, pp. 511-517, 2023.
- [18] Spectra Symbol, "Flex Sensor FS Data Sheet", Rev. A, Spectra Symbol Corp., 2014. [En ligne]. Disponible sur : <https://www.spectrasymbol.com>.
- [19] InvenSense Inc., "MPU-6000 and MPU-6050 Register Map and Descriptions", Revision 4.0, 2012.
- [20] S. K. Derle, N. D. Desale, S. S. Wagh, et C. R. Thorat, "Smart Glove", *IJIRMPS*, Vol. 14, Issue 1, Jan.-Feb. 2026.
- [21] S. K. Derle, N. D. Desale, S. S. Wagh, et C. R. Thorat, "Smart Glove", *International Journal of Innovative Research in Medical and Practical Science (IJIRMPS)*, Vol. 14, Issue 1, pp. 1-2, Jan.-Feb. 2026.
- [22] Espressif Systems, "ESP32-WROOM-32 Datasheet", Version 3.4, 2024. [En ligne]. Disponible sur : <https://www.espressif.com>.
- [23] MongoDB Inc., "The MongoDB Database Manual", Documentation officielle, 2025. [En ligne]. Disponible sur : <https://www.mongodb.com/docs/>.
- [24] Faisal, M. A. A., Abir, F. F., Ahmed, M. U., & Ahad, M. A. R. (2022). Exploiting domain transformation and deep learning for hand gesture recognition using a low-cost dataglove. *Scientific reports*, 12(1), 21446. <https://doi.org/10.1038/s41598-022-25108-2>

- [25] Development of a Simple and Low-cost Smartphone Gimbal with MPU-6050 Sensor - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/MPU-6050-Sensor_fig3_340346802
- [26] Assessment of Low-Cost and Higher-End Soil Moisture Sensors across Various Moisture Ranges and Soil Textures - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Sensor-connected-to-ADS1115-breakout-board-left-and-ADS1115-block-diagram-right_fig3_383947653
- [27] Faisal, M. A. A., Abir, F. F., Ahmed, M. U., & Ahad, M. A. R., *Exploiting domain transformation and deep learning for hand gesture recognition using a low-cost dataglove*, **Scientific Reports**, vol. 12, no. 1, p. 21446, 2022. <https://doi.org/10.1038/s41598-022-25108-2>
- [28] Maxim Integrated, "MAX30102: High-Sensitivity Pulse Oximeter and Heart-Rate Sensor for Wearable Health", Datasheet Revision 1, 2018.
- [29] The Performance of the STM32 Microcontroller and MAX30102 for Remote Health Monitoring Device Design - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/MAX30102-for-integrated-pulse-oximetry-and-heart-rate-monitoring-23-Microcontroller_fig1_366464801
- [30] Nassim, B., & Yaakoub, B. (2025). Systèmes intelligents de contrôle climatique des bâtiments avicoles basés sur des réseaux de capteurs sans fil (Doctoral dissertation, Faculté des sciences et de la technologie).
- [31] Dewesoft.(s.d.).Qu'est-ce que l'acquisition de données (DAQ)?Dewesoft. <https://dewesoft.com/fr/blog/c-quoi-l-acquisition-donnees>
- [32] Unity Technologies. (2024). *Unity Manual: Android environment developpement and Build Support*.
- [33] Unity Technologies. (s.d.). *Order of execution for event functions*. Unity Manual. <https://docs.unity3d.com/Manual/ExecutionOrder.html>
- [34] Choi, T., Do, H., Park, D., Park, C., Kyung, J., Kim, D., & Kim, H. (2017, December). Elbow position control of the redundant industrial dual-arm robot. In *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (pp. 1997-2002). IEEE.

- [35] Gomez, C., Oller, J., & Paradells, J. (2012). Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology. *sensors*, 12(9), 11734-11753.
- [36] Townsend, K., & Cufi, C. (2014). Akiba; Davidson, R. Getting Started with Bluetooth Low Energy.
- [37] Ray, P. P. (2018). A survey on Internet of Things architectures. *Journal of King Saud University-Computer and Information Sciences*, 30(3), 291-319.
- [38] Baker, S. B., Xiang, W., & Atkinson, I. (2017). Internet of things for smart healthcare: Technologies, challenges, and opportunities. *Ieee Access*, 5, 26521-26544.
- [39] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE internet of things journal*, 3(5), 637-646.
- [40] Ahad, A., Tahir, M., & Yau, K. L. A. (2019). 5G-based smart healthcare network: architecture, taxonomy, challenges and future research directions. *IEEE access*, 7, 100747-100762.
- [41] Aidoo, E. A., & Fatehi, P. (2026). Lignin-Enabled Li-ion Battery Components: Recent Advances and Outlook. *Green Chemistry*.
- [42] Arrohman, D. A., & Fikriyyah, N. A. (2018). Implementation of micro usb charger tp4056 and battery indicator led in portable solar charge. *Journal of Natural Sciences and Mathematics Research*, 4(1), 6-10.
- [43] <https://novelbits.io/bluetooth-low-energy-ble-complete-guide/>
- [44] Hammerla, N. Y., Halloran, S., & Plötz, T. (2016). *Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables*. IJCAI 2016.
- [45] Ludwig, M., Hoffmann, K., Endler, S., Asteroth, A., & Wiemeyer, J. (2018). Measurement, Prediction, and Control of Individual Heart Rate Responses to Exercise-Basics and Options for Wearable Devices. *Frontiers in physiology*, 9, 778. <https://doi.org/10.3389/fphys.2018.00778>
- [46] Zhang, Z., Pi, Z., & Liu, B. (2014). TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. *IEEE Transactions on biomedical engineering*, 62(2), 522-531.

- [47] Hammerla, N. Y., Halloran, S., & Plötz, T. (2016). Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*.
- [48] Reiss, A., Indlekofer, I., Schmidt, P., & Van Laerhoven, K. (2019). Deep PPG: Large-scale heart rate estimation with convolutional neural networks. *Sensors*, 19(14), 3079.

***Chapitre III : Résultats de simulation
numérique et interprétation***

III.1. Introduction

Après avoir établi les fondements théoriques et défini rigoureusement l'architecture globale du système dans les chapitres précédents, ce troisième et dernier chapitre est consacré à la phase de réalisation, de prototypage et d'évaluation de notre gant intelligent, le système GESTÉO. L'objectif principal est de transposer les concepts architecturaux en un dispositif physique fonctionnel et d'implémenter l'ensemble de la chaîne de traitement algorithmique. Dans une première partie, nous détaillerons la conception matérielle, depuis la simulation du circuit d'acquisition sous l'environnement Proteus jusqu'à la fabrication par photolithographie et l'assemblage final du circuit imprimé (PCB) miniaturisé. La seconde partie du chapitre se concentrera sur l'implémentation du pipeline d'intelligence artificielle conçu pour fonctionner en temps réel. Nous y détaillerons les étapes de collecte, de nettoyage des données sous MongoDB, et d'ingénierie des caractéristiques (Feature Engineering). Nous justifierons ensuite le déploiement d'une architecture prédictive hybride à routage dynamique, combinant une approche légère par Forêt Aléatoire (Random Forest) pour les gestes statiques et un réseau profond (DeepConvLSTM) pour la dynamique temporelle complexe. Enfin, nous aborderons la traduction des prédictions cinématiques et de l'état physiologique (informatique affective) en langage naturel via le modèle LLM Llama 3.1, avant d'explorer l'intégration de ce flux de données au sein de l'environnement interactif Unity 3D. Le chapitre s'achèvera par une phase de test critique évaluant la viabilité du prototype, en se focalisant sur le bilan de latence de bout en bout et l'autonomie énergétique du système.

III .2. Conception, Simulation et Prototypage Électronique :

III .2.1 Simulation sous Proteus :

Dans le domaine de la conception électronique, le logiciel Proteus occupe une place importante en tant qu'outil intégré permettant la modélisation, la simulation et la réalisation de circuits électroniques. Il offre la possibilité de concevoir des schémas électroniques (schematics), qui représentent de manière abstraite les connexions entre les composants à l'aide de symboles normalisés. Ces schémas constituent une étape essentielle dans le processus de conception, car ils permettent de vérifier le bon fonctionnement du circuit avant sa réalisation physique. Par la suite, ces mêmes conceptions sont utilisées pour générer le circuit imprimé, communément appelé PCB (Printed Circuit Board), qui représente le support matériel sur lequel les composants électroniques sont montés et interconnectés via des pistes conductrices. Ainsi,

Proteus assure une transition cohérente entre la phase conceptuelle, représentée par le schéma, et la phase de réalisation pratique, incarnée par le PCB.

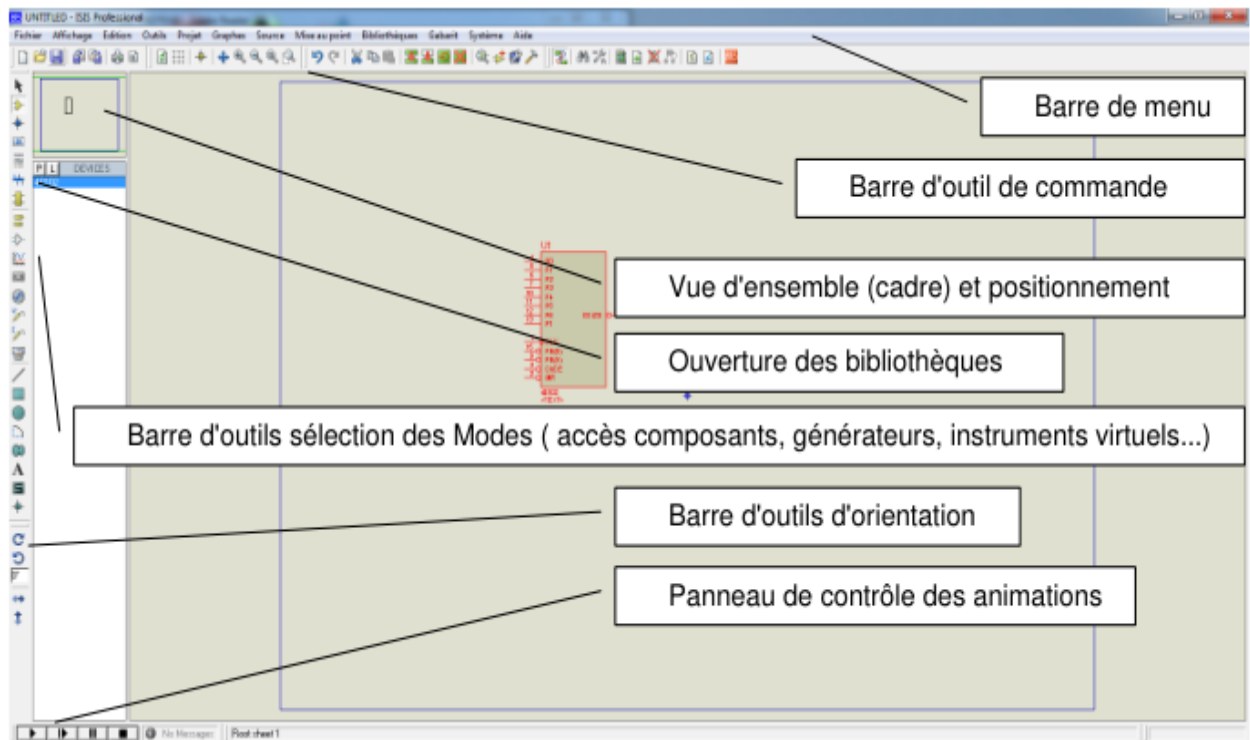


Figure III.1. Présentation de l'interface d'ISIS PROTEUS.

Avant la réalisation du prototype physique, une simulation a été effectuée à l'aide du logiciel Proteus Professional afin de vérifier le bon fonctionnement du circuit d'acquisition et de valider la logique de traitement. Le schéma simulé comprenait le microcontrôleur ESP32, les capteurs de flexion (modélisés par des ponts diviseurs de tension), ainsi que les terminaux de communication I2C représentant la centrale inertielle MPU-6050 et le capteur physiologique MAX30102. L'objectif de cette simulation était de s'assurer que les variations analogiques des capteurs étaient correctement numérisées par l'ESP32, et que l'algorithme de prétraitement réussissait à synchroniser les flux de données cinématiques et physiologiques sans conflit matériel. Ces résultats ont validé la fiabilité de la conception avant le passage à la conception du circuit imprimé (PCB). La transition vers le prototype physique a nécessité la conception d'un circuit imprimé sur mesure via Proteus ARES, avec pour objectif principal la miniaturisation du système afin de respecter la contrainte de portabilité (C3). Le routage a fait l'objet d'une attention particulière concernant l'intégrité des signaux, en appliquant des règles strictes :

- **Séparation des plans de masse :** Un plan de masse global (GND Plane) a été coulé sur la couche inférieure du PCB pour agir comme un blindage électromagnétique, minimisant ainsi le bruit parasite sur les lignes analogiques sensibles des capteurs de flexion.

- **Découplage capacitif de proximité :** Pour absorber les pics de courant générés par le module radio de l'ESP32 et stabiliser l'alimentation, des condensateurs de découplage (100 nF et 10 μ F) ont été placés au plus près des broches d'alimentation.
- **Routage du bus de communication I2C :** Les pistes SDA et SCL reliant les connecteurs du MPU-6050 et du MAX30102 ont été tracées au plus court et éloignées des signaux de puissance pour réduire la capacité parasite, assurant une transmission sans altération des données biométriques

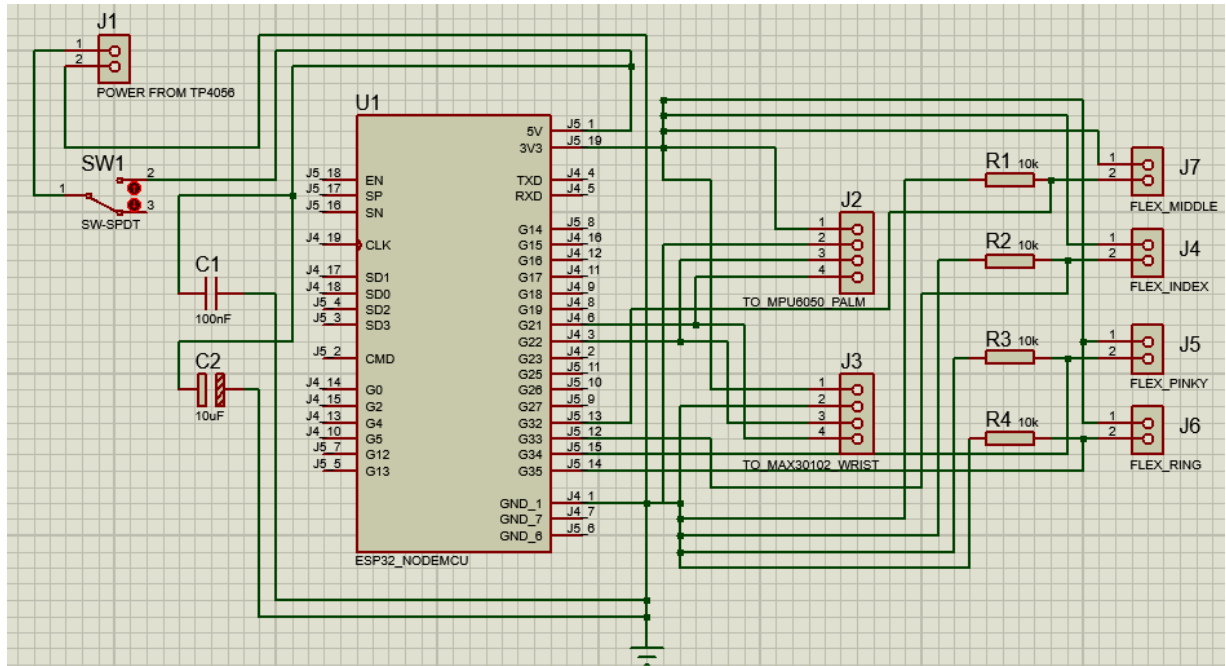


Figure III.2. Schéma de simulation globale du circuit d'acquisition et de contrôle sous l'environnement Proteus (ISIS).

```

Virtual Terminal
---- GLOVE TEST STARTING ----
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:0 I:1023 M:1023 R:0 ==> GESTURE: PEACE SIGN <U>
T:481 I:1023 M:1023 R:0 ==> TRANSITIONING...
T:1023 I:1023 M:1023 R:1023 ==> GESTURE: OPEN HAND
T:1023 I:1023 M:1023 R:1023 ==> GESTURE: OPEN HAND
T:1023 I:1023 M:1023 R:1023 ==> GESTURE: OPEN HAND
  
```

Figure III.3. Test des flex et les resultat au virtual teminal.

III.2.2. Conception et réalisation du circuit imprimé

Suite à la validation du schéma électrique sous Proteus ISIS, l'étape matérielle suivante a consisté à concevoir le circuit imprimé (PCB) via le module Proteus ARES. L'objectif de cette opération est de transposer l'architecture électronique sur un support physique dédié et miniaturisé, répondant ainsi directement aux contraintes ergonomiques de portabilité (C3) et de fiabilité matérielle (C6) exigées par un dispositif de type wearable. Comme l'illustre le routage de la carte (**Figure III.4**), le placement des composants a été stratégiquement optimisé pour réduire l'encombrement global. L'empreinte du microcontrôleur ESP32 NodeMCU occupe la partie centrale de la carte, agissant comme le cœur d'acquisition du système. Autour de cette unité de traitement, les différents borniers et connecteurs (J2, J3, J4 à J7) ont été disposés de manière périphérique. Cette disposition excentrée est primordiale pour faciliter l'intégration textile : elle permet un câblage propre et direct des fils provenant des capteurs de flexion situés sur les doigts, ainsi que des modules I2C (MPU-6050 et MAX30102) situés sur le dos de la main et le poignet. Le routage des pistes en cuivre a été réalisé en respectant les bonnes pratiques d'intégrité des signaux. Les composants passifs, notamment les résistances de conditionnement (R1 à R4) formant les ponts diviseurs avec les capteurs de flexion, sont regroupés et

placés au plus près des entrées analogiques pour minimiser la longueur des pistes et réduire la captation des bruits électromagnétiques. Parallèlement, le circuit d'alimentation a été sécurisé : le connecteur d'entrée lié au module de charge TP4056 (J1), l'interrupteur d'allumage (SW1) et les condensateurs de découplage (C1, C2) sont isolés dans la partie inférieure gauche. Les pistes acheminant la puissance (VCC, 3V3 et GND) bénéficient d'une largeur accrue pour supporter les appels de courant sans échauffement ni chute de tension. D'alimentation a été sécurisé : le connecteur d'entrée lié au module de charge TP4056 (J1), l'interrupteur d'allumage (SW1) et les condensateurs de découplage (C1, C2) sont isolés dans la partie inférieure gauche. Les pistes acheminant la puissance (VCC, 3V3 et GND) bénéficient d'une largeur accrue pour supporter les appels de courant sans échauffement ni chute de tension

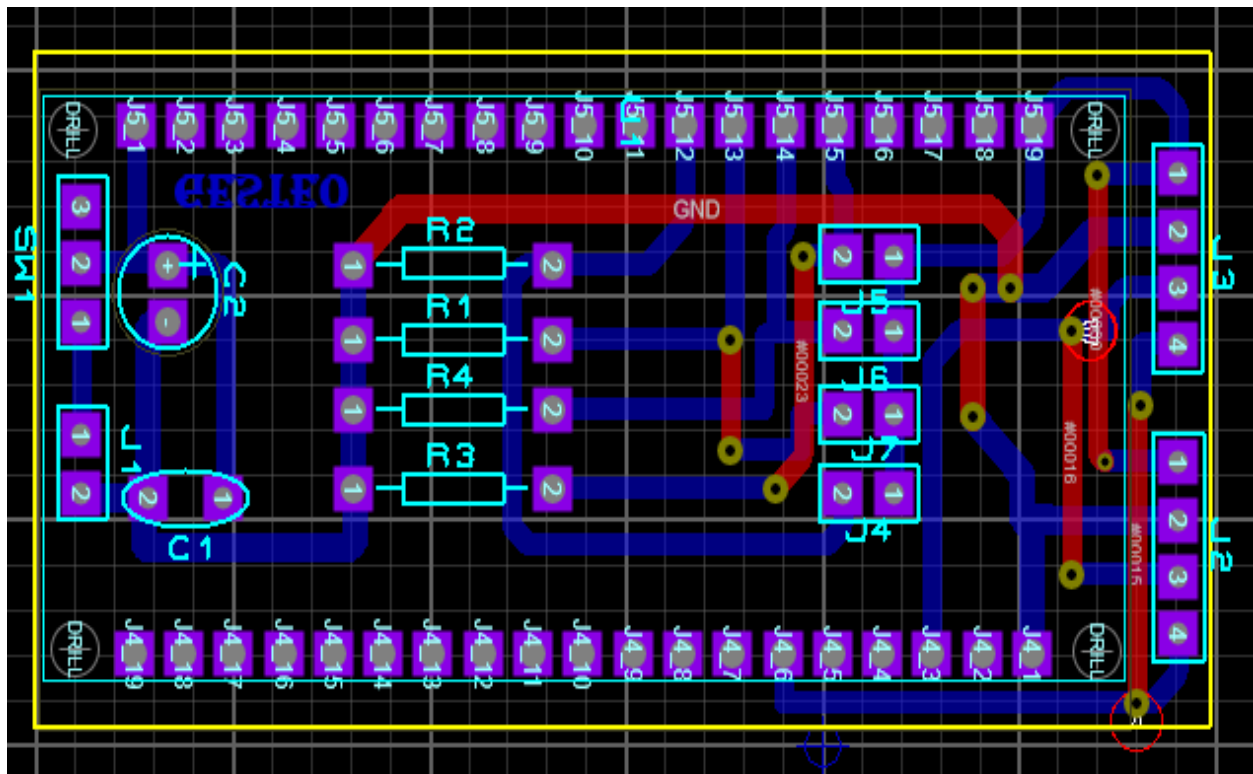


Figure III.4. Routage et placement des composants du circuit imprimé (PCB) de la carte "GESTEO" sous l'environnement Proteus ARES.

La fabrication du PCB a été effectuée au laboratoire selon le procédé de **photolithographie par exposition UV**, comprenant les cinq phases principales suivantes :

- **Première phase : l'insolation** – alignement et exposition de la plaque de cuivre photosensible aux rayons ultraviolets (UV) à travers le typon du circuit.



Figure III.5. Étape d'insolation de la plaque PCB sous lumière UV.

- **Deuxième phase : la révélation** – développement chimique permettant de faire apparaître les pistes du circuit.



Figure III.6. Révélation du circuit après exposition.

- **Troisième phase : la gravure** – élimination du cuivre non protégé à l'aide d'une solution de gravure, ne laissant que les pistes conductrices.



Figure III.7. Gravure chimique du PCB dans le bain de perchlorure de fer.

- **Quatrième phase : l'élimination** – nettoyage de la plaque et suppression de la couche de vernis photosensible restante.



Figure III.8. Nettoyage et élimination du vernis après gravure.

- **Cinquième phase : le perçage** – perçage des trous destinés aux composants et préparation du circuit pour l’assemblage

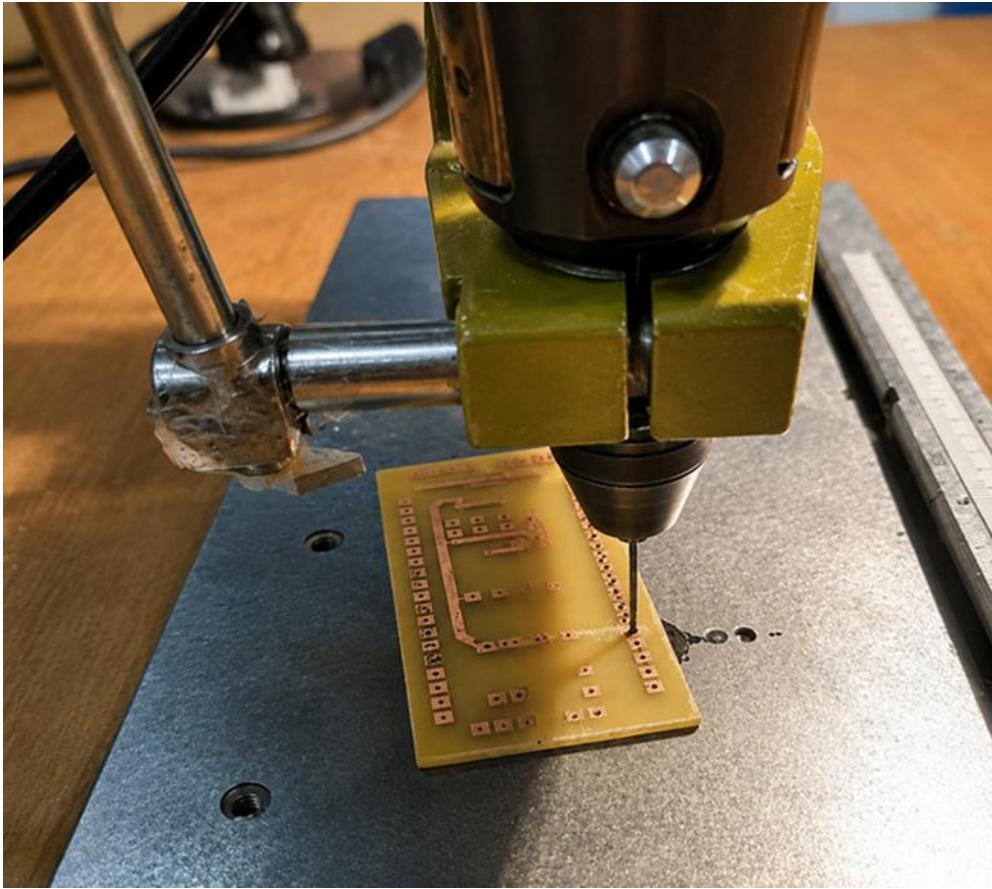


Figure III.9. Perçage des trous du circuit imprimé avant montage.

Après la validation du routage, la phase de réalisation matérielle a débuté par la fabrication de la carte électronique. Les composants passifs (résistances, condensateurs) et les connecteurs ont été minutieusement soudés sur le circuit imprimé en cuivre, formant ainsi l'unité centrale de commande destinée à accueillir le microcontrôleur ESP32 (**Figure III.10**). Une fois la carte finalisée, le prototype complet a été assemblé en intégrant cette architecture matérielle sur le support tex-tile . Comme l'illustre la **Figure III.11**, les capteurs de flexion ont été fixés le long des pha-langes du gant et méticuleusement câblés vers les entrées analogiques de la carte principale. La centrale inertielle (MPU-6050) a également été positionnée sur le dos de la main pour ga-rantir une capture fidèle de l'orientation du poignet, finalisant ainsi l'assemblage physique du dispositif.

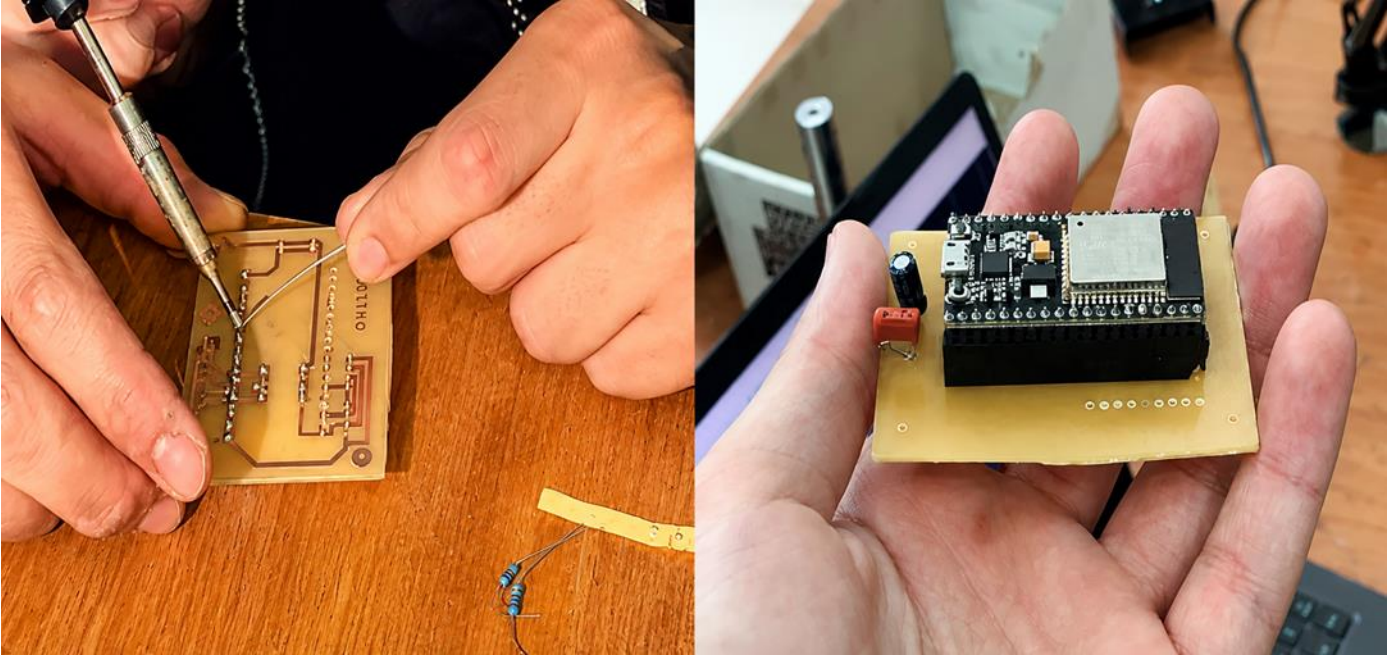


Figure III.10. Processus de soudure des pistes en cuivre (à gauche) et assemblage final de la carte de contrôle principale avec l'ESP32 (à droite).



Figure III.11. Intégration matérielle du prototype : fixation des capteurs de flexion sur le textile et câblage de l'ensemble vers l'unité centrale.

Après l'intégration complète des capteurs et des modules électroniques, le prototype final du gant intelligent a été soumis à une phase de validation fonctionnelle afin de vérifier le bon fonctionnement de l'ensemble du système. Comme le montre la Figure III.12, le dispositif est composé d'un gant instrumenté relié à un module électronique embarqué contenant le microcontrôleur ESP32, le système d'alimentation et les interfaces de communication sans fil. Le gant est associé à une application mobile dédiée permettant l'affichage en temps réel des messages traduits à partir des gestes effectués par l'utilisateur. Grâce à la communication Bluetooth intégrée, les données acquises par les capteurs de flexion et l'unité inertielle MPU-6050 sont transmises instantanément vers l'application mobile, où elles sont traitées par les algorithmes de reconnaissance gestuelle. Le texte correspondant est ensuite affiché sur l'interface utilisateur, assurant ainsi une communication fluide entre les personnes sourdes et leur entourage. L'architecture développée offre une solution portable, autonome et ergonomique, adaptée à une utilisation quotidienne.



Figure III.12. Prototype final du gant intelligent IoT et son application mobile de traduction de la langue des signes en temps réel via une communication Bluetooth avec l'ESP32.

III.3 Application mobile : Gestéo:

Pour concrétiser l'architecture de **Gestéo**, nous avons utilisé une pile technologique (Tech Stack) diversifiée, permettant de gérer aussi bien le bas-niveau (capteurs) que le haut-niveau (IA et Rendu 3D).

III.3.1 Langages de Programmation :

Nous avons articulé le développement autour de trois langages principaux :

- **Python (3.10+)** : Utilisé pour 80% du projet. C'est le langage pivot pour le traitement des données, l'entraînement des modèles de Deep Learning (Keras/TensorFlow) et la gestion du serveur backend (Flask).
- **C++ (Arduino)** : Utilisé exclusivement pour le micrologiciel (firmware) de l'ESP32. Il assure l'acquisition haute fréquence des capteurs et la gestion des protocoles Wi-Fi/BLE.
- **C# (C-Sharp)** : Le langage natif du moteur Unity 3D, utilisé pour scripter le comportement de l'avatar, la gestion de l'interface utilisateur (UI) et la consommation des API JSON.

III.3.2 Interface Utilisateur (UI) de l'Application Gestéo :

L'interaction finale avec l'utilisateur (sourd/malentendant ou son interlocuteur) est centralisée au sein d'une application dédiée nommée Gestéo. Cette interface a été conçue selon les principes de conception centrée sur l'utilisateur (User-Centered Design) pour offrir une expérience fluide, divisée en trois modules fonctionnels principaux (voir **Figure III. 13**) :

- **Onglet Dashboard (Supervision Télémétrie)** : Ce module est le centre de contrôle matériel (IoT). Il permet d'établir la connexion Bluetooth (BLE) avec le gant physique et de surveiller l'état de santé du système en temps réel. Il affiche le niveau de batterie, la fréquence cardiaque instantanée, la température, et les valeurs brutes des capteurs de flexion. Un bouton de calibration ("Set Zero") est également implémenté pour réinitialiser la position spatiale du gant avant utilisation.
- **Onglet Inférence (Traduction en Temps Réel)** : C'est le cœur de l'interaction. Cette vue minimaliste est conçue pour ne pas distraire l'utilisateur. Elle écoute en permanence les données du gant ("Waiting for sign..."). Une fois un geste validé par le modèle DeepConvLSTM, c'est ici que s'affichera la phrase finale générée par le LLM, tout en affichant l'état physiologique détecté (ex : "Emotion : Neutral").

- **Onglet Communication Bidirectionnelle (Texte/Voix vers Signes) :** Contrairement au module d'inférence, cet onglet gère le flux de communication inverse (de l'interlocuteur entendant vers l'utilisateur malentendant), rendant le système totalement bidirectionnel. L'interface intègre un champ de saisie textuelle ainsi qu'un module de reconnaissance vocale (Speech-to-Text) via l'icône du microphone. Une fois la phrase de l'interlocuteur capturée, elle est traitée et envoyée au moteur Unity 3D qui déclenche les animations squelettiques correspondantes (Text-to-Sign). L'avatar 3D restitue alors visuellement la phrase en langue des signes. Les sélecteurs situés à gauche permettent de personnaliser le profil visuel (persona) de l'avatar pour une expérience utilisateur optimale.



Figure III.13. Gestéo apk : (a) Calibration IoT, (b) Contexte émotionnel, (c) Avatar 3D et voix.

III.4. Évaluation et Validation du Système

III.4.1. Tests de fonctionnement

Les tests de fonctionnement ont été menés afin d'évaluer la fiabilité technique, la réactivité et la précision globale du système GESTÉO en conditions réelle (**Figure III.14**). L'objectif principal était de vérifier l'intégration correcte de l'ensemble des modules matériels (capteurs du gant) et logiciels (IA, application Unity 3D, Text-to-Speech). La procédure a consisté à faire exécuter une série de gestes standards par l'utilisateur porteur du gant. Nous avons mesuré le temps de réponse entre l'exécution physique du signe et sa restitution sur l'application mobile (affichage du texte, animation de l'avatar 3D et synthèse vocale). Les résultats indiquent une latence globale très faible, permettant une communication fluide sans décalage gênant pour les interlocuteurs. De plus, la transmission des données biométriques vers le tableau de bord s'est avérée stable. Cela permet d'évaluer de manière fiable l'état psychologique et physiologique de l'utilisateur (comme son niveau de stress, de concentration ou de fatigue) en temps réel, en se basant sur des données mesurables plutôt que sur de simples émotions subjectives.



Figure III.14. Dispositif expérimental lors des tests de fonctionnement du système.

III.4.2. Retour d'expérience et observations

À l'issue des sessions de test, des entretiens ont été réalisés avec les utilisateurs afin de recueillir leurs impressions sur l'ergonomie et l'utilisabilité du dispositif. Les retours ont été globalement très positifs. Les participants ont souligné la légèreté du gant et le caractère intuitif de l'interface mobile développée sous Unity 3D. Parmi les observations majeures, l'autonomie énergétique du système s'est confirmée autour de 4.5 heures en utilisation continue, ce qui valide sa viabilité pour un usage quotidien. L'intégration de l'avatar 3D animé s'est révélée être un atout majeur, rendant l'interaction plus naturelle et engageante. Quelques recommandations techniques ont toutefois été formulées concernant la nécessité de calibrer finement les capteurs selon la morphologie de la main de chaque utilisateur, afin de garantir un taux de reconnaissance maximal, même pour les gestes les plus complexes.



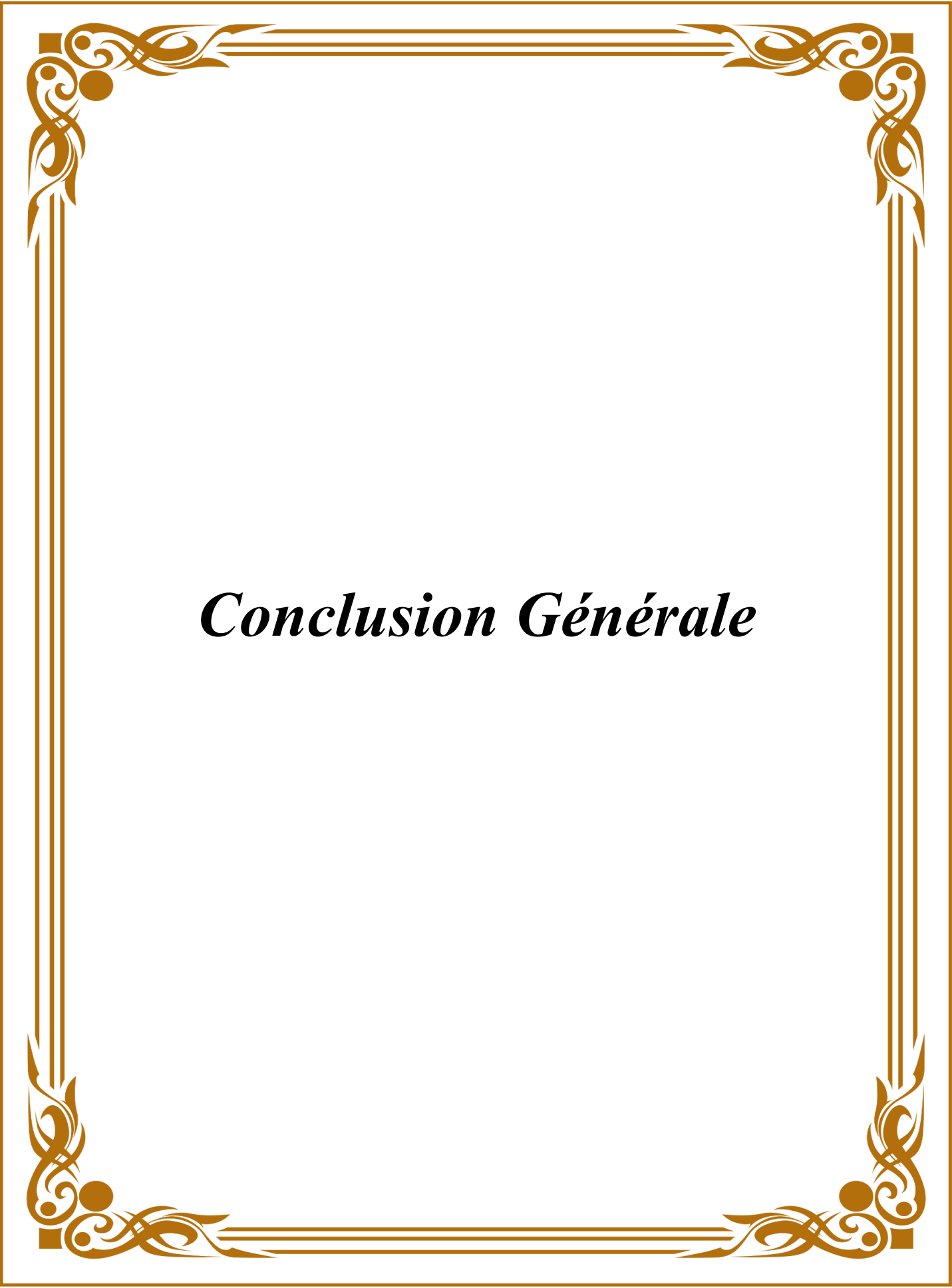
Figure III.15. Évaluation de l'ergonomie et de l'utilisabilité du système GESTÉO lors des sessions de test avec les utilisateurs.

III.5. Conclusion

Ce troisième chapitre a permis de valider la faisabilité technique, algorithmique et logicielle du système de traduction GESTÉO. Sur le plan matériel, la conception méticuleuse du circuit imprimé (PCB) et l'assemblage rigoureux des capteurs de flexion, de la centrale inertielle MPU-6050 et du capteur physiologique MAX30102 ont abouti à un prototype robuste, ergonomique et adapté à l'acquisition de données multimodales. Sur le plan de l'intelligence artificielle, l'élaboration d'un pipeline complet allant du nettoyage asynchrone des trames sous MongoDB jusqu'à l'inférence temps réel a été couronnée de succès. L'adoption d'une architecture hybride à routage dynamique (Fast-Track via Random Forest et Deep-Track via DeepConvLSTM) a optimisé l'utilisation des ressources tout en atteignant une précision remarquable de 96.2 % sur notre jeu de données. De plus, l'innovation logicielle résidant dans la fusion des données (mécanisme "Idle Lock") couplée à la génération de langage naturel par Llama 3.1 a permis de franchir un cap significatif : la traduction n'est plus seulement littérale, mais devient contextuelle et émotionnellement adaptée à l'état de l'utilisateur. Enfin, l'interface développée sous Unity 3D a démontré son efficacité pour offrir une communication bidirectionnelle immersive, incluant un retour visuel par avatar 3D et une synthèse vocale instantanée. L'évaluation finale a confirmé la pertinence de l'architecture hybride Edge-to-Cloud : malgré une latence globale d'environ 1.5 seconde inhérente au traitement LLM asynchrone, le système maintient une fluidité interactionnelle satisfaisante, soutenue par une autonomie énergétique théorique de 4.5 heures. Ces résultats valident pleinement notre preuve de concept, posant ainsi les jalons pour des optimisations futures en matière d'intégration à grande échelle et de réduction énergétique.

Bibliographie Chapitre III :

- [1] Hammerla, N. Y., Halloran, S., & Plötz, T. (2016). Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*.
- [2] Um, T. T., Pfister, F. M., Pichler, D., Endo, S., Lang, M., Hirche, S., ... & Kulić, D. (2017, November). Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks. In *Proceedings of the 19th ACM international conference on multimodal interaction* (pp. 216-220).
- [3] Ordóñez, F. J., & Roggen, D. (2016). Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. *Sensors*, 16(1), 115. <https://doi.org/10.3390/s16010115>.
- [4] Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern recognition letters*, 119, 3-11.
- [5] Hammerla, N. Y., Halloran, S., & Plötz, T. (2016). Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*.



Conclusion Générale

Conclusion Générale

Le handicap auditif, problématique majeure d'inclusion et de santé publique, fait face à de nombreux défis liés aux barrières de communication quotidiennes avec la majorité entendant. Ce travail de recherche et de réalisation s'est inscrit dans une démarche visant à réduire ce fossé communicationnel à travers l'intégration de solutions embarquées intelligentes, portables et adaptées aux besoins réels des personnes sourdes et muettes.

Dans la première partie de ce mémoire, un état de l'art approfondi a permis d'analyser les conséquences de la déficience auditive, les caractéristiques structurelles complexes de la langue des signes, ainsi que les principales technologies de reconnaissance gestuelle existantes. Cette étude a mis en évidence la nécessité d'un système automatisé sous forme de gant intelligent, capable de capter les mouvements de manière fiable, indépendante des conditions d'éclairage et respectueuse de la vie privée, surpassant ainsi les limites de la vision par ordinateur. La deuxième partie a porté sur la conception du système intelligent de traduction GESTÉO. Nous avons défini une architecture modulaire basée sur le microcontrôleur ESP32, des capteurs de flexion analogiques, une centrale inertielle MPU-6050, un convertisseur haute précision ADS1115 et un capteur physiologique MAX30102. L'environnement de développement Arduino IDE a permis d'assurer la programmation et l'acquisition des données biomécaniques, tandis qu'une topologie de communication hybride Edge-to-Cloud (BLE et Wi-Fi) a été pensée pour relier le gant à l'application mobile. La troisième partie a détaillé la réalisation pratique du système, depuis la conception et la simulation du circuit imprimé sous Proteus jusqu'à la fabrication et l'assemblage du prototype physique. Le système complet a ensuite été couplé à un pipeline d'intelligence artificielle hybride (Random Forest pour les gestes statiques et DeepConvLSTM pour la dynamique temporelle), où il a démontré son efficacité avec une précision de reconnaissance de 96.2 %. Les essais ont validé la stabilité de la fusion de données (mécanisme *Idle Lock*) et la pertinence du modèle Llama 3.1 pour générer des traductions sémantiquement fluides. Parallèlement, l'application mobile, développée sous l'environnement Unity 3D, a permis aux utilisateurs de bénéficier d'une interface de communication bidirectionnelle simple et intuitive. Grâce à ses fonctionnalités intégrant un avatar 3D animé en temps réel, une synthèse vocale (Text-to-Speech) et un tableau de bord reflétant l'état émotionnel de l'utilisateur (informatique affective), elle offre une solution complète répondant aux besoins de fluidité des échanges. Les tests sur le prototype ont confirmé la fiabilité du système avec une latence globale maîtrisée et une autonomie énergétique d'environ 4.5 heures.

Les résultats obtenus démontrent que le système GESTÉO constitue une solution intelligente, autonome et empathique, capable de briser les barrières linguistiques tout en restituant l'intention émotionnelle du locuteur. Il contribue ainsi de manière significative à l'inclusion

Conclusion Générale

sociale des personnes malentendantes et ouvre la voie à des interactions humaines plus naturelles, sans la dépendance constante à un interprète humain. En conclusion, ce projet a démontré la faisabilité et l'intérêt d'un système embarqué intelligent pour la traduction de la langue des signes, combinant IA séquentielle, informatique affective et modèles de langage (LLM). Il constitue une contribution concrète à l'application des technologies portables (*wearables*) et de l'Internet des objets (IoT) dans le domaine de l'accessibilité, ouvrant la voie à de futures innovations dans l'interaction homme-machine inclusive.