

**People's Democratic Republic of Algeria**  
**Ministry of Higher Education and Scientific Research**  
**20 August 1955 University – Skikda**  
**Faculty of Science**  
**Department of Computer Science**  
**Specialty: MASTER2 IA**

**Project Title**

---

# **Sign Language Translation Glasses**

---

**A project to obtain a Master's Degree**

**+ Startup Certificate under Ministerial Decree 1275**

**Presented by :**

- **Medjbouri ismahan**

**Supervised by**

- **Dr.Cheikh mohamed**

**2024-2025**

## Thanks

*In the name of Allah, the Most Gracious, the Most Merciful.*

*Praise be to Allah, who illuminated my path with knowledge and granted me patience and perseverance to navigate these many years of study. With a heart overflowing with gratitude, I stand today to thank everyone who has been a pillar of support throughout this journey.*

*First and foremost, I extend my deepest thanks to my immediate family. Your countless sacrifices, sincere prayers, and unwavering love have been the fundamental support propelling me towards success. Every day, I see my dream materialize through your eyes, and this is the greatest recognition I could ask for.*

*I will never forget my second family, the members of **Skikda I Tech Club**. You were not merely colleagues; you were brothers and sisters who significantly contributed to the success of my project at every step, offering assistance with great care and affection. Every piece of advice, every word of encouragement, and every moment we shared will remain etched in my memory.*

*To my dear friends, your continuous support and encouragement were a source of strength for me. Thank you all.*

*With profound appreciation and respect, I express my sincere gratitude and acknowledgment to our research director, **Dr.Cheikh Mohamed**. Your invaluable advice, constant availability, and continuous encouragement have been a source of inspiration throughout this project.*

*Finally, we extend our sincere thanks and appreciation to the honorable members of the examination committee for graciously reviewing this work.*

## **Abstract**

Deaf and mute people are significantly affected by communication barriers, restricting social integration and everyday interactions. The "Hope Smart Glasses" project offers a novel assistive technology that converts sign language into spoken words in real time.

The AI glasses contain an inbuilt camera to record hand movements, which are processed through deep learning algorithms and translated into audible speech through built in speakers. At the same time, a companion mobile application allows hearing individuals to read sign language by pointing the camera of their smartphone at the user.

Such a system enables deaf and mute people to communicate, independent of an intermediary, in social settings, in an educational environment, and within professional spheres. Our intention is to promote independence, promote autonomy, This is more than a technical solution it's a human mission. mission that says: "Every person has the right to be heard, to express, and to be understood."

We're using artificial intelligence not just to solve a problem but to give voice to those who were denied one.

## **Resume**

Les personnes sourdes et muettes sont considérablement affectées par les barrières de communication, ce qui restreint leur intégration sociale et leurs interactions quotidiennes. Le projet "Hope Smart Glasses" propose une nouvelle technologie d'assistance qui convertit la langue des signes en mots parlés en temps réel.

Les lunettes AI contiennent une caméra intégrée pour enregistrer les mouvements des mains, qui sont traités par des algorithmes d'apprentissage profond et traduits en discours audible grâce à des haut parleurs intégrés. Parallèlement, une application mobile compagnon permet aux personnes entendantes de lire la langue des signes en dirigeant la caméra de leur smartphone vers l'utilisateur.

Un tel système permet aux personnes sourdes et muettes de communiquer, sans intermédiaire, dans des contextes sociaux, éducatifs et professionnels. C'est plus qu'une solution technique, c'est une mission humaine. Une mission qui dit : "Chaque personne a le droit d'être entendue, de s'exprimer et d'être comprise."

Nous utilisons l'intelligence artificielle non seulement pour résoudre un problème, mais pour donner une voix à ceux qui en étaient privés. Nous utilisons l'intelligence artificielle non seulement pour résoudre un problème, mais aussi pour donner une voix à ceux qui en étaient privés.

## ملخص

الأشخاص الصم والبكم يتأثرون بشكل كبير بحواجز التواصل، مما يعيق الاندماج الاجتماعي والتفاعلات اليومية. مشروع "نظارات الأمل الذكية" يقدم تقنية مساعدة جديدة تحول لغة الإشارة إلى كلمات منطوقة في الوقت الفعلي. تحتوي نظارات الذكاء الاصطناعي على كاميرا مدمجة لتسجيل حركات اليد، والتي تتم معالجتها من خلال خوارزميات التعلم العميق وتُترجم إلى كلام مسموع عبر مكبرات الصوت المدمجة.

في الوقت نفسه، تتيح تطبيقات الهاتف المحمول المرافقة للأشخاص السامعين قراءة لغة الإشارة من خلال توجيه كاميرا هواتفهم الذكية نحو المستخدم.

مثل هذا النظام يمكن الصم والبكم من التواصل، دون الحاجة إلى وسيط، في الأوساط الاجتماعية، وفي البيئة التعليمية، وداخل المجالات المهنية. هذا أكثر من مجرد حل تقني إنه مهمة إنسانية. مهمة تقول: "لكل شخص الحق في أن يُسمع، أن يُعبر، وأن يُفهم".

نحن نستخدم الذكاء الاصطناعي ليس فقط لحل مشكلة، بل لمنح صوت لأولئك الذين حُرّموا منه.

# Content

<b>General introduction.....</b>	<b>14</b>
<b>Chapter1: Medical and social context</b>	
<b>1 Introduction .....</b>	<b>16</b>
<b>2 The ear .....</b>	<b>16</b>
<b>2.1 Ear anatomy .....</b>	<b>16</b>
<b>2.2 Outer Ear .....</b>	<b>17</b>
<b>2.3 Middle Ear .....</b>	<b>18</b>
<b>2.4 Inner Ear .....</b>	<b>18</b>
<b>2.5 Mechanism of Hearing .....</b>	<b>19</b>
<b>2.6 Mechanism of Balance.....</b>	<b>20</b>
<b>3 Hearing loss and deafness .....</b>	<b>20</b>
<b>4 History of Sign Languages .....</b>	<b>22</b>
<b>5 Assistive technologies .....</b>	<b>25</b>
<b>5.1 Smart Gloves for Sign Language Translation .....</b>	<b>25</b>
<b>5.2 Portable Translators for Sign Language .....</b>	<b>26</b>
<b>5.3 Translator Robots.....</b>	<b>28</b>
<b>6 Conclusion .....</b>	<b>30</b>
<b>Chapter 2: Artificial Intelligence and sign language recognition technologies</b>	
<b>1 Introduction.....</b>	<b>32</b>
<b>2 Definition of Artificial Intelligent.....</b>	<b>32</b>
<b>3 Machine Learning.....</b>	<b>33</b>
<b>3.1 Definition .....</b>	<b>33</b>
<b>3.2 Types of machine learning .....</b>	<b>33</b>
<b>3.2.1 Supervised learning.....</b>	<b>34</b>

3.2.2	Unsupervised learning.....	36
3.2.3	Semi supervised learning .....	38
3.2.4	Reinforcement learning.....	38
<b>4</b>	<b>Deep learning.....</b>	<b>38</b>
<b>4.1</b>	<b>Definition .....</b>	<b>38</b>
<b>4.2</b>	<b>Deep Learning Operating Mechanism .....</b>	<b>39</b>
<b>4.3</b>	<b>Deep Learning in Machine Learning Paradigms.....</b>	<b>39</b>
4.3.1	Supervised Learning .....	39
4.3.2	Unsupervised Learning .....	40
4.3.3	Reinforcement Learning .....	40
<b>4.4</b>	<b>Deep learning Algorithms .....</b>	<b>40</b>
4.4.1	Convolutional Neural Networks (cnns) .....	40
4.4.2	Long Short Term Memory Networks (lstm) .....	41
4.4.3	Recurrent Neural Networks (rnns) .....	42
<b>5</b>	<b>Fundamentals of Pattern Recognition.....</b>	<b>43</b>
<b>5.1</b>	<b>Definition of Pattern Recognition .....</b>	<b>43</b>
<b>5.2</b>	<b>Process of Pattern Recognition .....</b>	<b>43</b>
<b>5.3</b>	<b>Role in Artificial Intelligence .....</b>	<b>44</b>
<b>6</b>	<b>Sign Language Recognition .....</b>	<b>44</b>
<b>6.1</b>	<b>Importance .....</b>	<b>44</b>
<b>6.2</b>	<b>Challenges .....</b>	<b>45</b>
<b>6.3</b>	<b>Deep Learning in Sign Language Recognition .....</b>	<b>45</b>
6.3.1	Traditional Techniques .....	45
6.3.2	Deep Learning Approaches .....	46
<b>6.4</b>	<b>Practical Development Process of a Sign Language Recognition System .....</b>	<b>47</b>

<b>6.5</b>	<b>Evaluation, Practical Challenges, and Future Prospects .....</b>	<b>48</b>
6.5.1	Evaluation Metrics .....	49
6.5.2	Practical Challenges .....	49
6.5.3	Future Prospects .....	49
<b>7</b>	<b>Conclusion .....</b>	<b>50</b>
<b>Chapter 3 : Conception and implementation</b>		
<b>1</b>	<b>Introduction.....</b>	<b>52</b>
<b>2</b>	<b>Conception.....</b>	<b>52</b>
2.1	Project Overview.....	52
2.2	Working Principle.....	54
2.3	Challenges and Solutions.....	55
2.4	Organigramme .....	56
2.5	Future Work and Enhancements .....	57
<b>3</b>	<b>Implementation .....</b>	<b>57</b>
<b>3.1</b>	<b>Hardwar layer .....</b>	<b>57</b>
3.1.1	ESP32 CAM Camera .....	57
3.1.2	speakers .....	58
3.1.3	Wi Fi .....	58
3.1.4	Bluetooth.....	59
<b>3.2</b>	<b>Mobile application layer.....</b>	<b>59</b>
3.2.1	Flutter.....	59
3.2.2	Dart .....	60
3.2.3	Vs code .....	60
3.2.4	Flutter TTS.....	61
<b>3.3</b>	<b>firebase and AI processing layer.....</b>	<b>61</b>

3.3.1	fire base .....	61
3.3.2	AI processing .....	62
<b>4</b>	<b>Project implementation .....</b>	<b>65</b>
<b>5</b>	<b>Conclusion .....</b>	<b>70</b>
<b>Chapter 4: Commercial Overview</b>		
<b>1</b>	<b>Section 1: Project Presentation.....</b>	<b>72</b>
1.1	Field of Activity .....	72
1.2	Genesis and Evolution of the Idea .....	72
1.3	Project Objectives .....	73
1.4	Bidirectional Use Cases .....	74
1.5	Project Team .....	76
1.6	Implementation Environment .....	76
1.7	Value Proposition .....	76
1.8	Project Development Plan.....	78
<b>2</b>	<b>Section 2: Innovation Aspects .....</b>	<b>78</b>
2.1	Nature of Innovation .....	78
2.2	Areas of Innovation .....	79
<b>3</b>	<b>Section 3: Invention Description.....</b>	<b>79</b>
3.1	Patent Title .....	79
3.2	Abstract .....	79
3.3	Technical Fields .....	79
3.4	Objective of the Invention .....	80
3.5	Core of the Invention .....	80
<b>4</b>	<b>Section 4: Claims.....</b>	<b>82</b>
<b>5</b>	<b>Section 5: Illustrations (to be included in final thesis).....</b>	<b>82</b>

**General conclusion..... 92**  
**References..... 94**

## Table of figures

Figure 1 : Ear anatomy.....	17
Figure 2 : Expected increase in the prevalence of hearing loss, 2019 2050 .....	20
Figure 3: Arabic sign language alphabets.....	23
Figure 4: English sign language alphabets.....	24
Figure 5:signaloud Gloves .....	25
Figure 6:Uni Handheld Translator.....	27
Figure 7:ASLAN Robot (Belgaum).....	29
Figure 8 : Venn diagram of the components of Artificial Intelligence.....	32
Figure 9 : types of Machine Learning.....	33
Figure 10 :Neural network architecture .....	39
Figure 11: Convolutional Neural Networks architecture.....	41
Figure 12:LSTM architecture.....	42
Figure 13 : Rnns architecture.....	43
Figure 14 : The organigram of the project.....	56
Figure 15 : ESP32 CAM Camera.....	58
Figure 16 :Speakers.....	58
Figure 17 : WIFI icon.....	58
Figure 18: Bluetooth icon .....	59
Figure 19 : Flutter logo.....	59
Figure 20 : Dart logo.....	60
Figure 21: Vs code logo.....	60
Figure 22 : firebase logo .....	61
Figure 23 :python logo.....	62
Figure 24 : tensorflow logo.....	63
Figure 25: jupyter logo.....	63
Figure 26 : collab logo .....	64
Figure 27 : kaggle logo .....	64
Figure 28 : Splash screen .....	65

Figure 29 : Choose method screen.....	66
Figure 30 : Glasses user method screen.....	67
Figure 31 : App user method screen .....	68
Figure 32 : The glasses design.....	69
Figure 33 : Splash screen .....	83
Figure 34 : Choose method screen.....	84
Figure 35 : Glasses user method screen.....	85
Figure 36 : App user method screen .....	86
Figure 37 : The glasses design.....	87

## List of tables

Table 1: Hearing levels .....	21
Table 2:Project Development Plan .....	78
Table 3 : Business Model Canvas.....	89
Table 4:Annual Cost Table – Hope Project .....	90
Table 5 : Revenue Projection Table – Hope Project.....	91

## **General introduction**

Human communication is one of the fundamental pillars of social life. It is the essential means through which people express ideas, emotions, and exchange knowledge. Despite the tremendous technological advancements achieved globally, millions of deaf and mute individuals still face significant challenges in daily interaction due to the absence of fast and effective tools to interpret their natural language sign language.

In this context, the idea of the Hope Smart Glasses project emerged. This graduation project falls within the field of assistive technologies and aims to develop a wearable device that allows deaf and mute individuals to translate sign language into spoken speech in real time, empowering them to communicate independently with their surroundings without the need for a human interpreter.

The project leverages Artificial Intelligence (AI) specifically Deep Learning alongside mobile and cloud technologies. It consists of smart glasses equipped with a miniature camera that captures hand gestures and sends them to a mobile application, which processes and vocalizes the interpreted words. Additionally, a hearing person can use the mobile app alone to point the camera at a deaf person and instantly receive a translation of their gestures.

This thesis presents all stages of the project's development, starting from theoretical and technical foundations, through design and implementation, and finally exploring the commercial and innovative aspects all with the ultimate goal of providing a practical and inclusive solution that bridges the communication gap between the deaf community and the rest of society.

**Chapter 1 :**  
**Medical and Social**  
**Context**

## **1 Introduction :**

People with hearing impairments, who confront particular everyday problems, are finding fresh hope as a result of the enormous technology revolution taking place in our world. Among the challenges this population faces include decreased employment possibilities, social isolation, and restricted access to information.

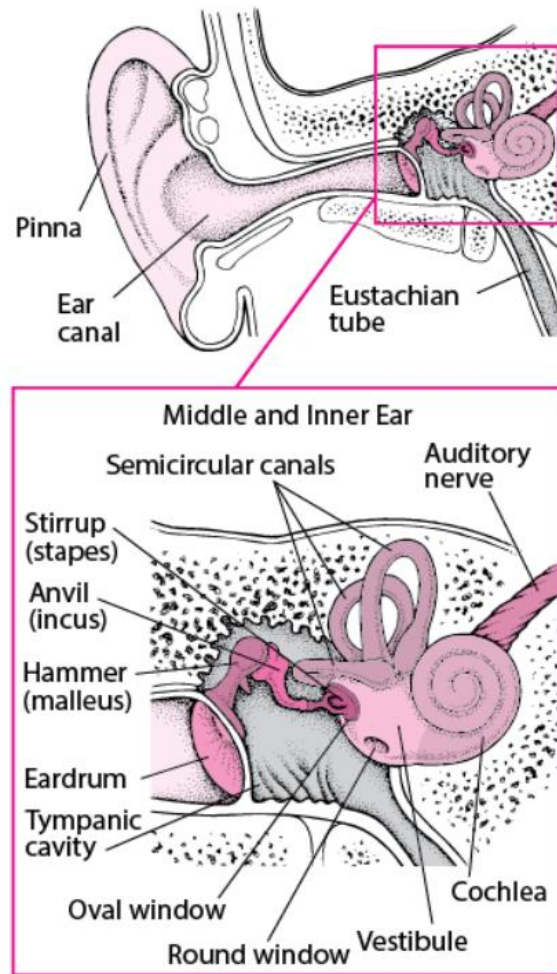
Rapid technical development, however, gives optimism. Significant advancements in cochlear implants, hearing aids, and smart phone apps tailored to the requirements of the deaf and hard of hearing have been made in recent years. In addition to enhancing hearing, these contemporary technologies also promote improved communication, greater independence, and new avenues for social engagement.

## **2 The ear :**

### **2.1 Ear anatomy :**

The ear is a vital organ that plays a crucial role in our ability to hear and maintain our sense of balance. It consists of three primary components: the outer ear, the middle ear, and the inner ear. The outer ear serves the important function of collecting sound waves from the environment and funneling them into the ear canal. Once the sound reaches the middle ear, this part amplifies the sound waves, making them louder and clearer.

Subsequently, the inner ear takes on the task of converting these amplified sound waves into electrical signals that the brain can interpret. Together, these three parts work harmoniously to enable us to perceive a wide range of sounds and to maintain equilibrium in our daily activities.[1]



**Figure 1 : Ear anatomy**

### 2.2 Outer Ear :

The outer ear collects and directs sound waves toward the eardrum. Its components are [1]

- **Pinna(Auricle):**

The external, visible part of the ear. It is made of cartilage and skin, shaped to capture sound waves effectively and direct them into the ear canal. The pinna also helps with sound localization, allowing us to identify the direction of sounds.

---

## Chapter1 : Medical and social context

---

- **Ear Canal (External Auditory Canal):**

A narrow passageway that leads to the eardrum. It amplifies certain frequencies of sound and contains small hairs and glands that produce earwax, which helps protect the ear by trapping dirt and deterring insects.

- **Eardrum (Tympanic Membrane):**

A thin, flexible membrane that separates the outer ear from the middle ear. It vibrates in response to sound waves, converting them into mechanical energy for the middle ear.

### 2.3 Middle Ear :

The middle ear amplifies sound and transmits it to the inner ear. It contains:

- **Ossicles:** These are the three smallest bones in the human body:
  - **Malleus (Hammer):** Attached to the eardrum and vibrates with it.
  - **Incus (Anvil):** Connects the malleus to the stapes and transmits vibrations between them.
  - **Stapes (Stirrup):** The smallest bone, which presses on the oval window of the inner ear to transmit vibrations.

The ossicles amplify sound vibrations by about 20 times, ensuring the sound is strong enough to move the fluids in the inner ear.

- **Eustachian Tube:**

A narrow passage connecting the middle ear to the back of the throat. It equalizes pressure on both sides of the eardrum, which is necessary for clear hearing and to prevent damage to the eardrum.[1]

### 2.4 Inner Ear :

The inner ear is where sound waves are converted into electrical signals and balance is maintained. Key components include:[1]

# Chapter1 : Medical and social context

---

- **Cochlea:**

A spiral shaped, fluid filled structure that plays a critical role in hearing. Inside the cochlea are:

- **Basilar Membrane:** Divides the cochlea and vibrates in response to sound.
- **Hair Cells:** Specialized sensory cells located on the basilar membrane. They have tiny hair like structures (stereocilia) that bend with vibrations, converting mechanical energy into electrical signals. High frequency sounds are detected near the base of the cochlea, and low frequency sounds are detected near the apex.

- **Vestibular System:**

This system is responsible for balance and spatial orientation. It includes:

- **Semicircular Canals:** Three loop shaped structures filled with fluid that detect rotational movements of the head.
- **Otolith Organs (Utricle and Sacculle):** Detect linear movements and the effects of gravity.

- **Auditory (Cochlear) Nerve:**

This nerve transmits electrical signals from the cochlea to the brain, where they are interpreted as sounds. It also has connections to the vestibular system for balance.

## 2.5 Mechanism of Hearing :

The hearing process involves several steps that convert sound waves into neural signals Sound waves are captured by the pinna and travel through the ear canal to the eardrum.

- The eardrum vibrates, and these vibrations are transmitted to the ossicles in the middle ear.
- The ossicles amplify the vibrations and transfer them to the cochlea via the oval window.
- The fluid in the cochlea moves, creating waves along the basilar membrane.
- The movement of the membrane bends the hair cells, triggering the release of electrical signals.

These signals travel via the auditory nerve to the brain's auditory cortex, where they are interpreted as sound. [2]

## 2.6 Mechanism of Balance:

Balance relies on inner ear structures detecting movement and orientation.

- Head movements cause fluid in the semicircular canals to shift, bending sensory hair cells within them.
- The hair cells send signals to the brain about the direction and speed of movement.
- The otolith organs detect gravity and linear acceleration, helping maintain posture and equilibrium.

## 3 Hearing loss and deafness :

According to World Health Organization (WHO) projections, at least 700 million people would need some kind of rehabilitation treatment to deal with different health concerns by 2050. Furthermore, it is predicted that 2.5 billion individuals globally would suffer from hearing loss to some degree. In addition to causing large financial losses because of the exclusions people will experience in terms of communication, education, and employment opportunities, the widespread prevalence of hearing impairment and the need for rehabilitation will also have a significant impact on the general health and well being of those who are affected.

As a result of these disabilities, people and society will probably incur more expenses as they attempt to function in a world that demands excellent communication and participation in a variety of activities.[2][3]

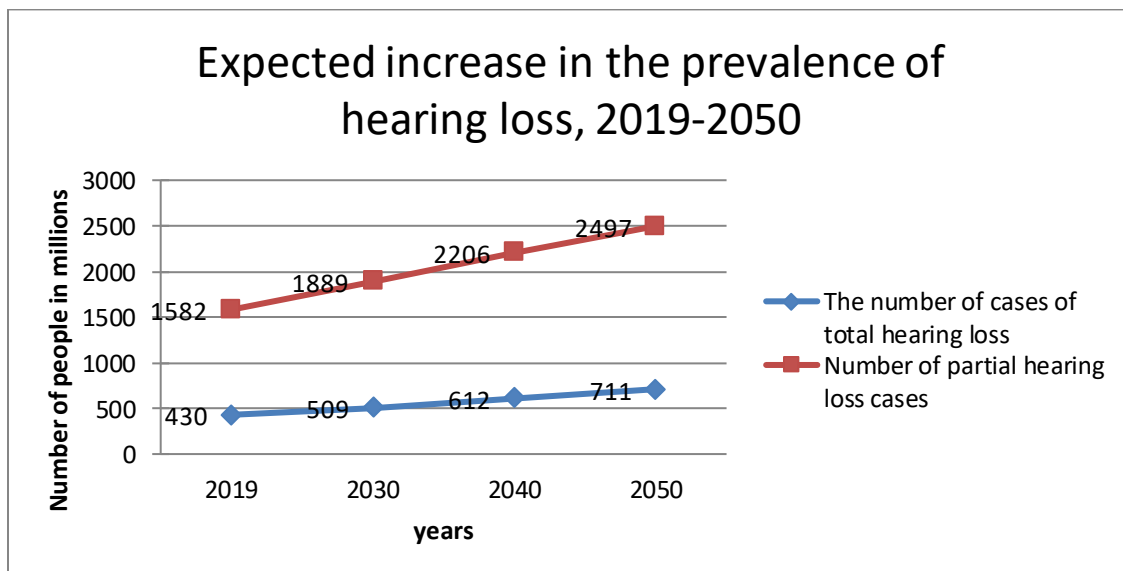


Figure 2 : Expected increase in the prevalence of hearing loss, 2019 2050

## Chapter1 : Medical and social context

---

Hearing loss is defined as the inability to hear as well as someone with normal hearing, which is defined as hearing thresholds of 20 db or better in both ears. Mild, moderate, severe, and profound hearing loss are all possible. One or both ears may be affected, making it difficult to hear loud noises or conversational speech. [2]

Decibel Range	Hearing Loss Severity
26 40 db	Mild hearing loss
40 70 db	Moderate hearing loss
70-90 db	Severe hearing loss
>90 db	Profound hearing loss (Deafness)

Table 1: Hearing levels

Due to the differences in hearing levels, each kind communicates differently; some utilize sign languages, while others employ lip reading, also referred to as oralism.

- **Oralism :**

Oralism is the practice of educating someone to read the speaker's lip; those who are hard of hearing typically employ this technique, which necessitates that the learner understand the mouth motions required to enunciate each letter or word. However, this approach may be difficult to use when identifying words and noises that are silent and lack recognizable mouth movements.[4]

# Chapter1 : Medical and social context

---

- **Sign Languages :**

The structured set of hand movements with distinct meanings that deaf people use to communicate in daily life is known as sign language. Being a visual language, It communicates through the body, facial, and hand gestures. Due to its greater diversity and range of motion, sign languages are the most often used form of communication among those with hearing impairments. Because sign languages are so widespread, a nation or region may have more than one sign language. [4]

#### **4 History of Sign Languages :**

Since the beginning of time, and even before language and civilization evolved Humans have expressed and communicated their views through both facial and body movements and feelings. However, the use of signs and gestures as a distinct language did not emerge until recently because it was always considered a component of cultural gestures and language. [5]

Since fluency in speech and language was considered a sign of high intelligence, deaf and hearing impaired people have endured years of discrimination, shame, and being treated as inferior to others because they are unable to speak. Sign language use was also stigmatized and denigrated. This even reached the point where some laws attempted to outlaw the use of sign language in deaf schools.

Despite all the negative efforts, some progress was made later. For example, in the 16th century, education for the deaf was first introduced in Paris, France, in 1760. Charles Michel De L'Eppe founded the first public school for the deaf, the Institut National de Jeunes Sourds de Paris, and he also created a dictionary of sign language words and an alphabet in French. This sparked initiatives to establish sign languages as internationally recognized languages (which runs counter to the myth that there is a single universal sign language) and to establish additional deaf schools that use and teach sign language as their primary communication method. These initiatives are still ongoing today, as certain sign languages [6]

Sign languages were able to thrive at these public schools for the deaf, which were first founded in France and later spread over most of Europe and North America. However, as we mentioned earlier, many people saw previous sign languages as primitive systems of gestures, and signing was outlawed in the majority of schools after a noticeable trend toward oralism in

## Chapter1 : Medical and social context

the late 19th century. However, outside of schools and in deaf organizations and clubs, sign languages flourished, and by the middle of the 20th century, oralism started to decline.

However, despite all of this history and academic interest in sign language, linguistics study didn't start until 1960. Since then, research on more than 100 sign languages has been carried out worldwide, demonstrating the significant growth of the field of sign language linguistics. We know that historical ties have frequently arisen from migration, global politics, and the export of educational systems, even though the genetic relationships among the world's sign language have not been fully studied . [7]

- Arabic sign language alphabets [8] :



Figure 3: Arabic sign language alphabets

# Chapter1 : Medical and social context

- English sign language alphabets [9]

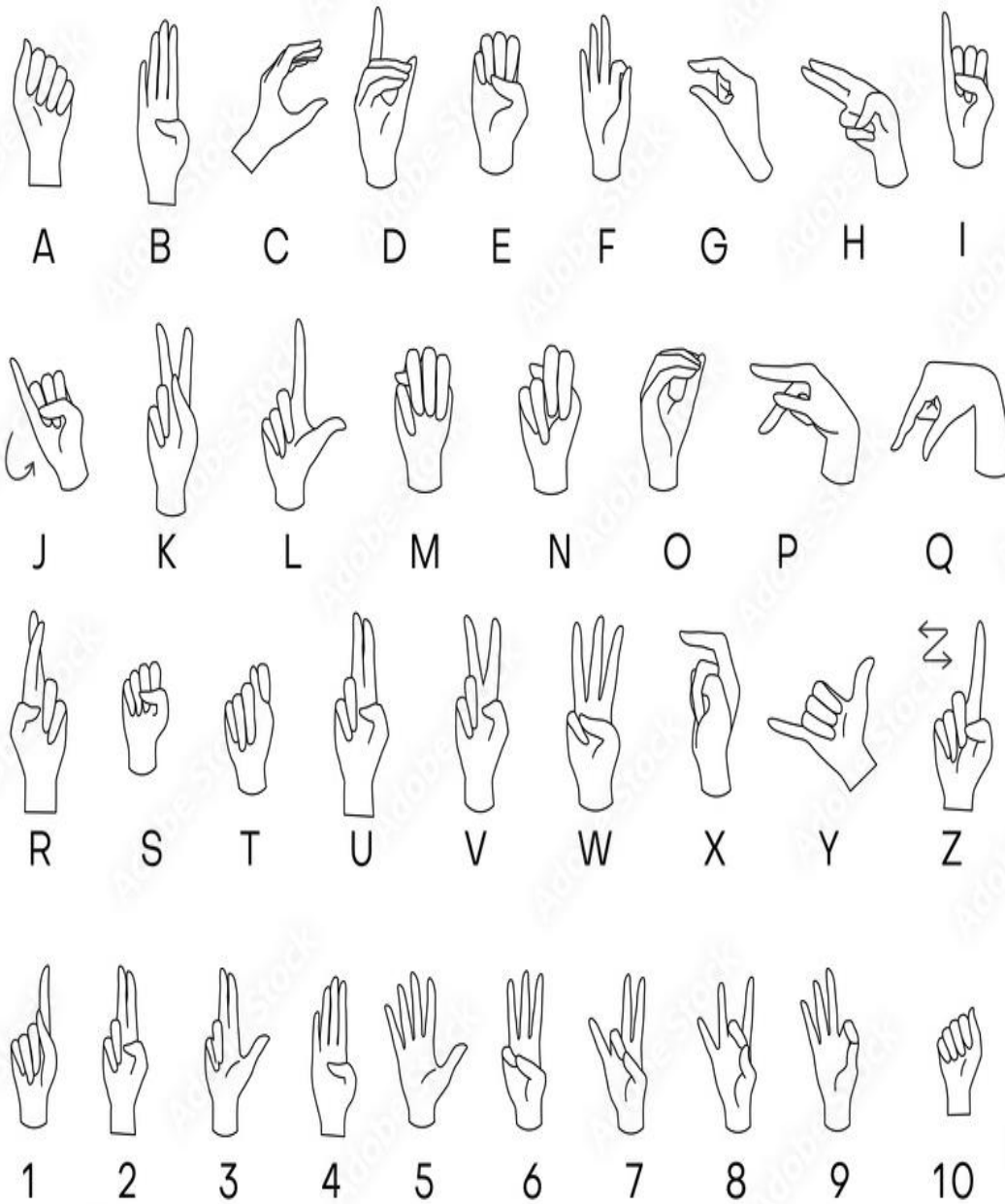


Figure 4: English sign language alphabets

## 5 Assistive technologies :

### 5.1 Smart Gloves for Sign Language Translation :

Smart gloves are an innovative solution designed to bridge the communication gap for the deaf and mute communities by translating sign language gestures into text or speech. These gloves are equipped with sensors that detect the movements of the hands and fingers, and the data is processed by a connected device (like a smartphone) to generate readable text or spoken words.[10]

- **Description and Examples:**

- **Signaloud Gloves :** Developed by researchers, these gloves use sensors to track finger and hand movements and convert them into speech. They rely on motion sensors and machine learning algorithms to analyze sign language gestures and generate real time speech translation.
- **Brightsign Gloves (UK):** Similar to signaloud, these gloves are designed to translate sign language into text, which can then be displayed on a screen or converted into voice. The gloves are used in combination with an app or wearable device that processes the data.[11]

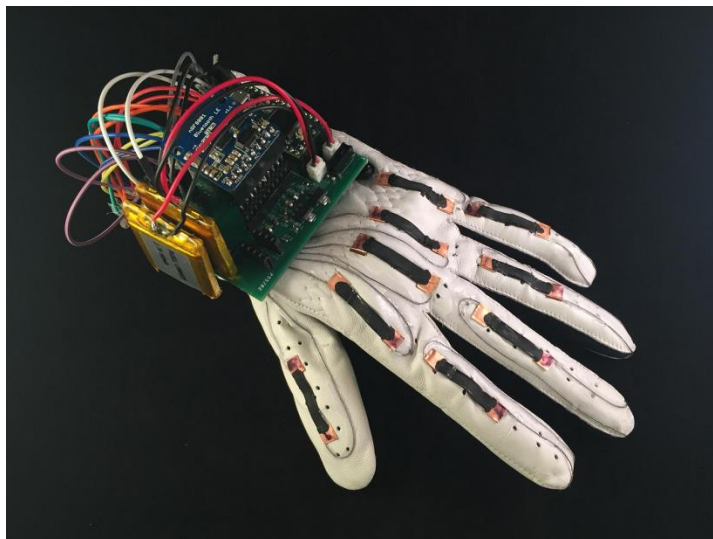


Figure 5:signaloud Gloves

- **Advantages:**
  - **Real Time Translation:** Provides immediate translation of sign language into speech or text, facilitating communication between deaf and hearing individuals.
  - **Mobile Integration:** Smart gloves often sync with mobile applications, enhancing portability and ease of use in various situations.
- **Challenges:**
  - **Constant Wearing Requirement:** The user must always wear the gloves to communicate, which may not be practical in many real life situations. This can be especially inconvenient for people who are not used to wearing such devices regularly or in hot or cold environments.
  - **Limited Coverage of Regional Sign Languages:** Many smart gloves focus on specific sign languages, like American Sign Language (ASL). This limits their usefulness in regions where other forms of sign language are used, such as British Sign Language (BSL), Indian Sign Language (ISL), or Arabic Sign Language. The technology may need to be adapted or expanded to recognize a wide variety of local sign languages, which is a significant challenge for global applicability.
  - **Accuracy and Precision:** Sign language involves complex movements and expressions that may vary slightly depending on the user. The gloves must accurately track a wide range of nuanced movements, which may be difficult for some devices to process, especially in cases of non standard or more complex signs.

### 5.2 Portable Translators for Sign Language :

Portable translators are small, handheld devices designed to translate sign language into spoken language, providing an immediate and accessible way for people who are deaf or mute to communicate with hearing individuals. These devices offer a more flexible alternative to other translation methods, such as smart gloves or apps, since they are compact, easy to carry, and allow for real time communication. [12]

## Chapter1 : Medical and social context

---

- **Description and Examples:**

- **Uni Handheld Translator:** This device is a portable solution that captures the user's sign language and converts it into audio or text. It typically uses cameras, sensors, or motion detectors to understand the hand gestures and map them to corresponding words. The translations are displayed on a screen or converted to speech. \$399.99
- **Kasa Translate:** Another example of a portable translator designed for sign language interpretation. It is a compact device that can translate sign language into both text and audio. These devices are often built with high quality cameras and sensors that allow for the detection of precise hand movements.



Figure 6:Uni Handheld Translator

- **Advantages:**

**Portability:** These devices are compact and easy to carry around, making them suitable for use in everyday situations. Unlike smart gloves, which require the user to wear them at all times, portable translators can be pulled out and used whenever necessary.

## Chapter1 : Medical and social context

---

- **Ease of Use:** Many portable translators are designed to be intuitive and simple, allowing individuals with limited technical knowledge to use them effectively. They may involve pressing a button or simply making a gesture to activate the translation.
- **Versatility:** These devices can be used in a variety of environments, such as at work, in educational settings, or while traveling, making them a versatile solution for people who need to communicate in different contexts.
- **Challenges:**
  - **High Cost:** Portable translators, especially high quality ones, can be expensive. This may limit their accessibility for people with lower incomes or for widespread adoption in communities with higher rates of hearing impairment.
  - **Limited Language Coverage:** Many portable translators are designed for specific sign languages (such as American Sign Language or British Sign Language). They may not support all regional variations or less commonly used languages, which can make them less useful in diverse linguistic communities.
  - **Accuracy and Speed:** The accuracy of the translation can vary depending on the device's technology. Some portable translators may struggle with recognizing more complex or nuanced signs, and the translation may be slower or less precise in these cases. Additionally, background noise or fast movements can interfere with the device's ability to capture the signs accurately.

### 5.3 Translator Robots:

Translator robots are designed to bridge the communication gap between spoken language and sign language by converting verbal language into visual gestures or sign language. These robots use advanced technology to recognize spoken language and then translate it into appropriate sign language movements that can be understood by people who are deaf or hard of hearing. [13]

## Chapter1 : Medical and social context

---

- **Description and Example:**

**ASLAN Robot (Belgium):** team of engineering students from the University of Antwerp are building a humanoid robot that will have the ability to translate speech into sign language.

Sponsored by the European Institute for Otorhinolaryngology, the robot titled Project Aslan aims to support the short supply of sign language interpreters across the world. The project uses 3D printing combined with readily available components to make the robot affordable and easily manufacturable. Using 3D Hubs, a network of 3D printing services the Aslan robot will be able to be produced in over 140 countries.

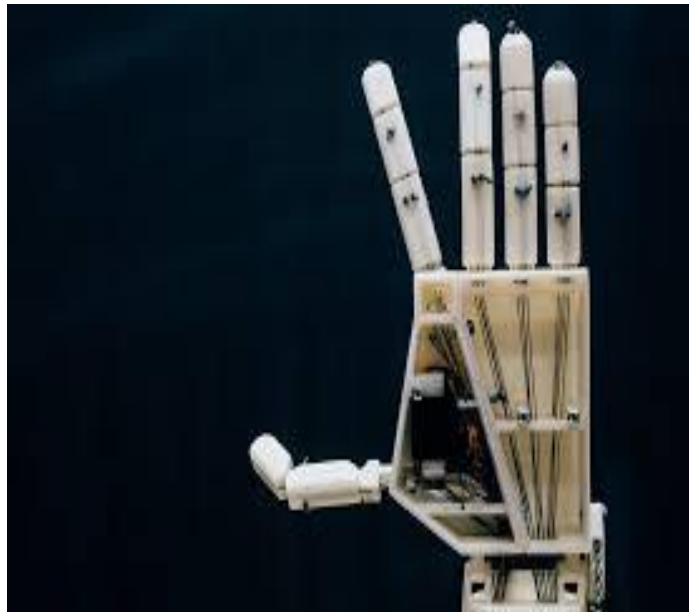


Figure 7:ASLAN Robot (Belgaum)

# Chapter1 : Medical and social context

---

- **Challenges:**

- **Bulky Design:** One of the significant challenges for translator robots is their large size and complexity. Most current robots are too bulky, making them impractical for everyday use in social or public settings. They are often not designed to be portable or convenient to use, limiting their effectiveness outside of controlled environments like labs or demonstration settings.
- **Cost and Accessibility:** These robots are typically expensive to produce and maintain. The high cost may prevent widespread adoption, especially in regions where funding or resources for assistive technologies are limited.
- **Limited Interactivity:** While these robots can translate spoken language into sign language, they are typically not able to engage in dynamic two way communication as effectively as human interpreters or more interactive assistive devices. Their ability to interpret non verbal cues or emotional context in a conversation is limited.
- **Complexity in Sign Language Translation:** Translating sign language is not a straightforward task, as different regions have distinct sign languages with their own rules, vocabulary, and dialects. Adapting robots to handle such variations adds complexity to the development process.

## 6 Conclusion :

This chapter provided a foundational understanding of the human auditory system, the causes and consequences of hearing loss, and the role of sign languages in bridging communication gaps for the deaf and mute community.

We explored the anatomy of the ear, the mechanism of hearing and balance, and the historical and social significance of sign language. In addition, we reviewed various assistive technologies that inspired the development of innovative solutions such as Hope smart glasses.

This chapter sets the stage for the technical exploration in the following chapters, where we delve into the artificial intelligence systems and hardware implementation that drive our project.

# **Chapter 2 :**

## **Artificial Intelligence and Sign Language Recognition Technologies**

### 1 Introduction:

Hand gesture recognition is a key technology in developing devices that can interpret sign language in real time. With the rapid progress of Artificial Intelligence (AI), especially deep learning, researchers are now able to create wearable devices such as smart glasses that can translate sign language into text or speech. This chapter explains how machine learning and deep learning contribute to these advancements.

### 2 Definition of Artificial Intelligent:

Artificial intelligence or AI for short is the field that englobes both Machine Learning (ML) and Deep Learning (DL), it concerns the science of recreating human intelligence in machines and making them able to do tasks that humans can. AI is the root and main domain of all other subsequent fields and domains that aim to train or teach machines how to reason rationally.

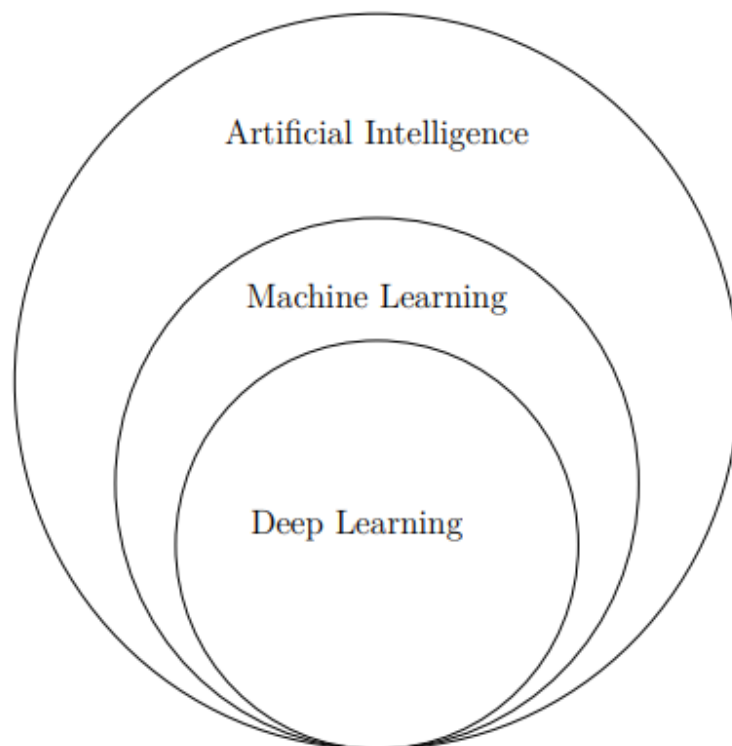


Figure 8 : Venn diagram of the components of Artificial Intelligence

### 3 Machine Learning:

#### 3.1 Definition

Machine learning, or ML, is the second layer in relation to artificial intelligence (AI), if we consider AI to be the first layer. As the name implies, this discipline of study focuses on creating programs and algorithms that let the computer to recognize patterns and learn without being specifically programmed to do so. This is accomplished by providing the software with a sizable amount of data, known as training data, and letting the machine learn from it.

Machine learning (ML) is frequently used to tackle complex issues with many inputs that are challenging (and perhaps impossible) for humans to handle; in these situations, a computer intervention would save a significant amount of time and money.[15]

#### 3.2 Types of machine learning :

Generally, ML is separated into 4

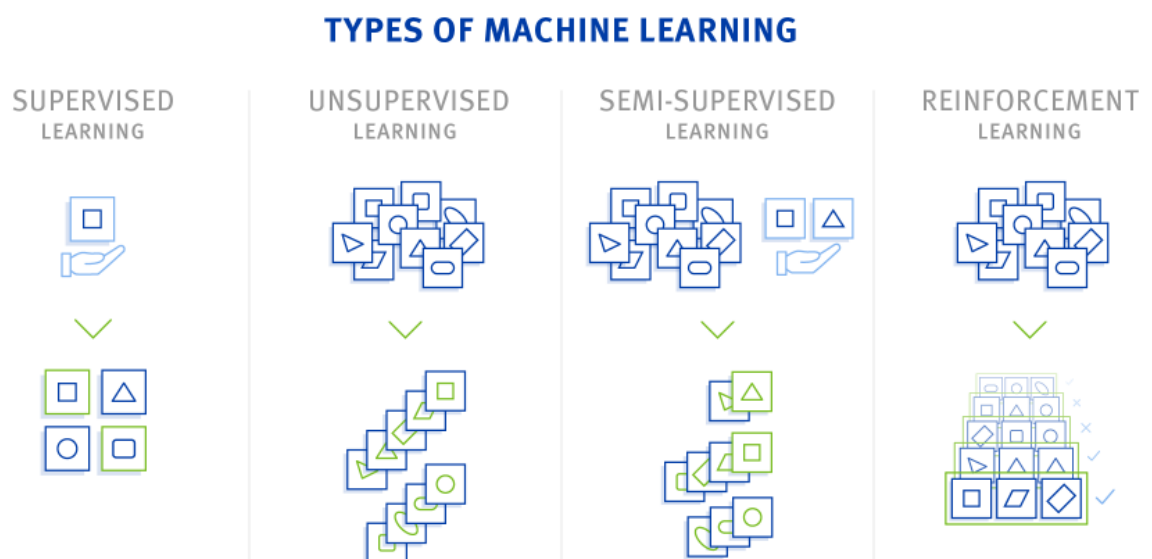


Figure 9 : types of Machine Learning

### 3.2.1 Supervised learning:

Algorithms are trained on datasets where each example is paired with a target or response variable, known as the label. The goal is to learn a mapping function from input data to the corresponding output labels, enabling the model to make accurate predictions on unseen data. Supervised learning problems are generally categorized into two main types Classification and Regression. Most widely used supervised learning algorithms are: [15][16]

- **Linear Regression :** [17]

Is used to predict a continuous value by finding the best fit straight line between input (independent variable) and output (dependent variable)

- Minimizes the difference between actual values and predicted values using a method called “least\_squares” to best fit the data.
- Predicting a person’s weight based on their height or predicting house prices based on size.

- **Logistic Regression :**

Logistic regression predicts probabilities and assigns data points to binary classes (e.g., spam or not spam).

- It uses a logistic function (S shaped curve) to model the relationship between input features and class probabilities.
- Used for classification tasks (binary or multi class).
- Outputs probabilities to classify data into categories.

- **Decision Trees :**

A decision tree splits data into branches based on feature values, creating a tree like structure.

- Each decision node represents a feature; leaf nodes provide the final prediction.
- The process continues until a final prediction is made at the leaf nodes
- Works for both classification and regression tasks.

- **Support Vector Machines (SVM) :**

Svms find the best boundary (called a hyperplane) that separates data points into different classes.

- Uses support vectors (critical data points) to define the hyperplane.
- Can handle linear and non linear problems using kernel functions.
- Focuses on maximizing the margin between classes, making it robust for high dimensional data or complex patterns.

- **K Nearest Neighbors (k NN) :**

KNN is a simple algorithm that predicts the output for a new data point based on the similarity (distance) to its nearest neighbors in the training dataset, used for both classification and regression tasks.

- Calculates distance between point with existing data points in training dataset using a distance metric (e.g., Euclidean, Manhattan, Minkowski)
- Identifies k nearest neighbors to new data point based on the calculated distances.
  - For **classification**, algorithm assigns class label that is most common among its k nearest neighbors.
  - For **regression**, the algorithm predicts the value as the average of the values of its k nearest neighbors.

- **Naive Bayes :**

Based on Bayes' theorem and assumes all features are independent of each other (hence "naive")

- Calculates probabilities for each class and assigns the most likely class to a data point.
- Assumption of feature independence might not hold in all cases ( rarely true in real world data )
- Works well for high dimensional data.

- **Random Forest :**

Random forest is an ensemble method that combines multiple decision trees.

- Uses random sampling and feature selection for diversity among trees.
- Final prediction is based on majority voting (classification) or averaging (regression).

- **Neural Networks ( Including Multilayer Perceptron) :**

Neural Networks, including Multilayer Perceptrons (mlps), are considered part of supervised machine learning algorithms as they require labeled data to train and learn the relationship between input and desired output; network learns to minimize the error using backpropagation algorithm to adjust weights during training.

- **Multilayer Perceptron (MLP):** Neural network with multiple layers of nodes.

Used for both classification and regression ( **Examples:** image classification, spam detection, and predicting numerical values like stock prices or house prices)

### 3.2.2 Unsupervised learning:

Performs better on transactional data that is unlabeled. When an unlabeled dataset is given to the algorithm, it looks for similarities between data groups to create clusters or data groups; in other words, the system is unsure of the "right answer" in this situation; therefore, the algorithm must determine what is being displayed for the purpose of exploring the data and thus finding some structure in it.[15]

- **Clustering :**

Clustering algorithms group data points into clusters based on their similarities or differences. The goal is to identify natural groupings in the data. Clustering algorithms are divided into multiple types based on the methods they use to group data. These types include Centroid based methods, Distribution based methods, Connectivity based methods, and Density based methods.

- **Centroid based Methods:** Represent clusters using central points, such as centroids or medoids.
  - K Means clustering: Divides data into k clusters by iteratively assigning points to nearest centers, assuming spherical clusters.
  - K Means++ clustering

- K Mode clustering
- Fuzzy C Means (FCM) Clustering
  
- **Distribution based Methods :**
  - Gaussian mixture models (gmms) : Models clusters as overlapping Gaussian distributions, assigning probabilities for data points' cluster membership.
  - Expectation Maximization Algorithms
  - Dirichlet process mixture models (dpmms)
- **Connectivity based methods :**
  - Hierarchical clustering : Builds a tree like structure (dendrogram) by merging or splitting clusters, no predefined number.
    - Agglomerative Clustering
    - Divisive clustering
  - Affinity propagation
- **Density Based methods :**
  - DBSCAN (Density Based Spatial Clustering of Applications with Noise) : Forms clusters based on density, allowing arbitrary shapes and detecting outliers, with distance and point parameters.
  - OPTICS (Ordering Points To Identify the Clustering Structure)

- **Dimensionality Reduction :**

Dimensionality reduction is used to simplify datasets by reducing the number of features while retaining the most important information.

- **Association Rule :**

Find patterns (called association rules) between items in large datasets, typically in market basket analysis (e.g., finding that people who buy bread often buy butter). It identifies patterns based solely on the frequency of item occurrences and co occurrences in the dataset.

- **Apriori algorithm** : Finds frequent itemsets by iterating through data and pruning non frequent item combinations.
- **FP Growth (Frequent Pattern Growth)** : Efficiently mines frequent itemsets using a compressed **FP tree** structure without candidate generation.

### 3.2.3 Semi supervised learning :

Is typically employed for applications that are comparable to those of supervised learning. In reality, this means that there is a small amount of labeled data and a high volume of unlabeled data. Tools like regression, classification, and prediction can be used with this type of learning. Since basic instances in this kind of learning, such as recognizing a face on a webcam, allow for a fully labeled training procedure, it is highly helpful to equate the labeling cost as being very high.[15]

### 3.2.4 Reinforcement learning:

Operates by integrating a dynamic feedback loop in which the computer receives a score based on its performance and seeks to maximize that score. Developing a model that can precisely handle issues and spot trends is the aim of any strategy.[15]

## 4 Deep learning:

### 4.1 Definition:

Deep learning is the third layer in relation to artificial intelligence (AI), that focuses on algorithms called artificial neural networks that are modeled after the structure and operations of the brain. [18]

Deep learning models automatically build hierarchical representations of data through several processing layers, in contrast to typical machine learning techniques that frequently demand for manual feature extraction. They can process and extract complex patterns from high dimensional data, including audio, video, and pictures, because to this capabilities. [18]

## 4.2 Deep Learning Operating Mechanism :

Layers of linked nodes, or neurons, make up a neural network, which works together to process incoming data. Data moves through several layers in a fully connected deep neural network, where each neuron carries out nonlinear modifications to enable the model to learn complex data representations.

Data enters a deep neural network through the input layer and then travels through hidden layers that use nonlinear functions to alter the data. The model's forecast is produced by the last output layer. [18]

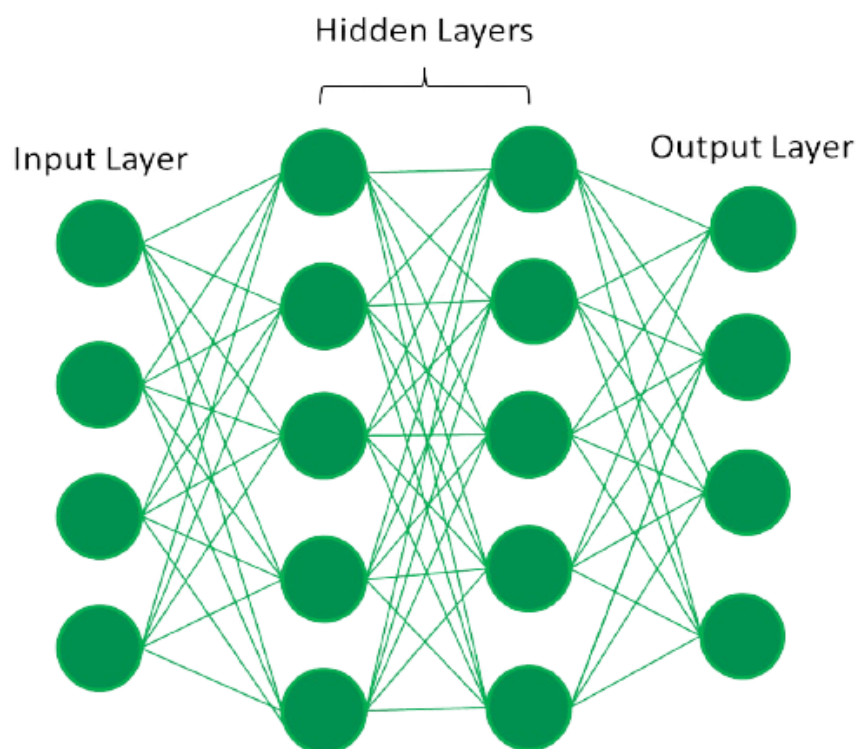


Figure 10 :Neural network architecture

## 4.3 Deep Learning in Machine Learning Paradigms:

### 4.3.1 Supervised Learning:

Cnns and rnns are algorithms that use labeled data to perform tasks like language translation and picture recognition.[18]

### 4.3.2 Unsupervised Learning:

Auto encoders and generative models are two methods that neural networks use to find patterns in unlabeled data for tasks like anomaly detection and grouping.[18]

### 4.3.3 Reinforcement Learning:

Algorithms like DQN and DDPG are used in robotics and gaming to teach agents to make decisions by maximizing rewards [18]

## 4.4 Deep learning Algorithms:

### 4.4.1 Convolutional Neural Networks (cnns) :

CNN's popularly known as convnets majorly consists of several layers and are specifically used for image processing and detection of objects.

It was developed in 1998 by Yann leCun and was first called lenet. Back then, it was developed to recognize digits and zip code characters. Cnns have wide usage in identifying the image of the satellites, medical image processing, series forecasting, and anomaly detection.

Cnns process the data by passing it through multiple layers and extracting features to exhibit convolutional operations. The Convolutional Layer consists of Rectified Linear Unit (relu) that outlasts to rectify the feature map.

The Pooling layer is used to rectify these feature maps into the next feed. Pooling is generally a sampling algorithm that is down sampled and it reduces the dimensions of the feature map. Later, the result generated consists of 2 D arrays consisting of single, long, continuous, and linear vector flattened in the map. The next layer i.e., called Fully Connected Layer which forms the flattened matrix or 2 D array fetched from the Pooling Layer as input and identifies the image by classifying it.[19] [20]

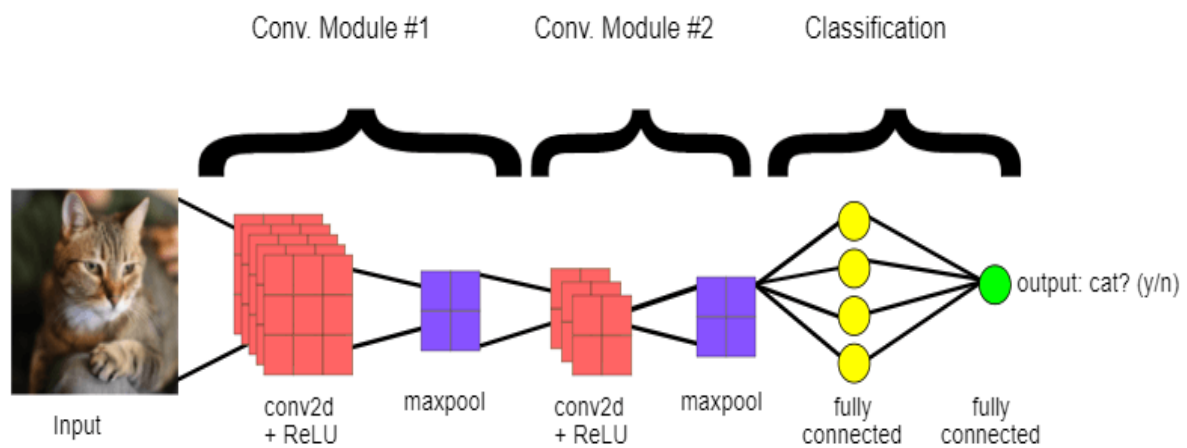


Figure 11: Convolutional Neural Networks architecture

### 4.4.2 Long Short Term Memory Networks (Lstms) :

Lstms can be defined as Recurrent Neural Networks (RNN) that are programmed to learn and adapt for dependencies for the long term. It can memorize and recall past data for a greater period and by default, it is its sole behavior.

Lstms are designed to retain over time and henceforth they are majorly used in time series predictions because they can restrain memory or previous inputs. This analogy comes from their chain like structure consisting of **four** interacting layers that communicate with each other differently. Besides applications of time series prediction, they can be used to construct speech recognizers, development in pharmaceuticals, and composition of music loops as well.

LSTM work in a sequence of events. First, they don't tend to remember irrelevant details attained in the previous state. Next, they update certain cell state values selectively and finally generate certain parts of the cell state as output. Below is the diagram of their operation.[19]

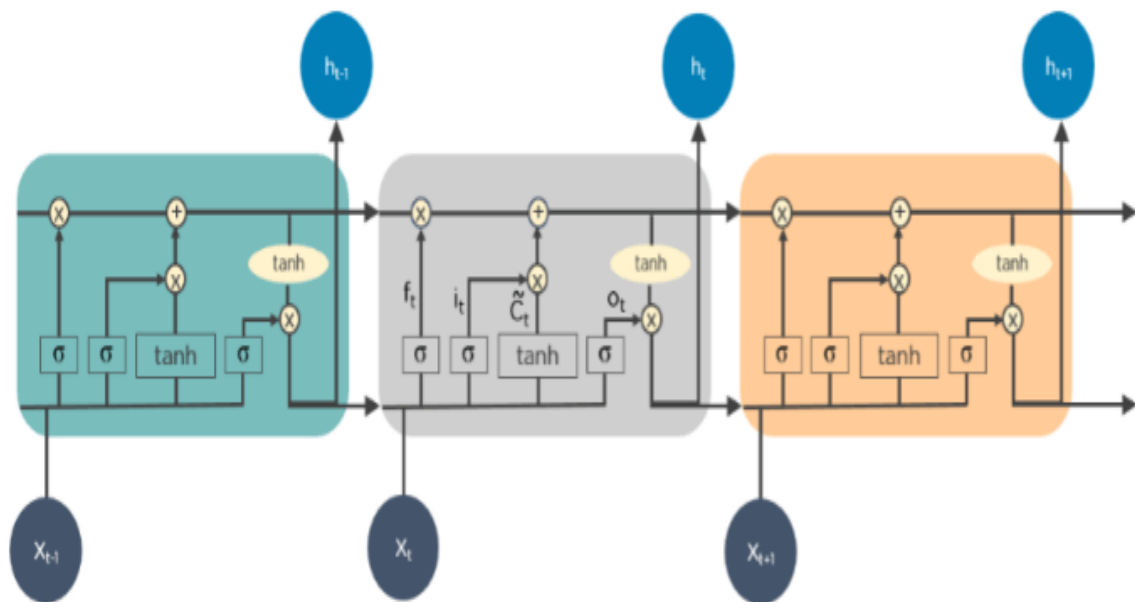


Figure 12:LSTM architecture

#### 4.4.3 Recurrent Neural Networks (rnns) :

Recurrent Neural Networks or rnns consist of some directed connections that form a cycle that allow the input provided from the lstms to be used as input in the current phase of rnns. These inputs are deeply embedded as inputs and enforce the memorization ability of lstms lets these inputs get absorbed for a period in the internal memory.

Rnns are therefore dependent on the inputs that are preserved by lstms and work under the synchronization phenomenon of lstms. Rnns are mostly used in captioning the image, time series analysis, recognizing handwritten data, and translating data to machines.

Rnns follow the work approach by putting output feeds ( $t$  1) time if the time is defined as  $t$ . Next, the output determined by  $t$  is feed at input time  $t+1$ . Similarly, these processes are repeated for all the input consisting of any length.

There's also a fact about rnns is that they store historical information and there's no increase in the input size even if the model size is increased. Rnns look something like this when unfolded. [19]

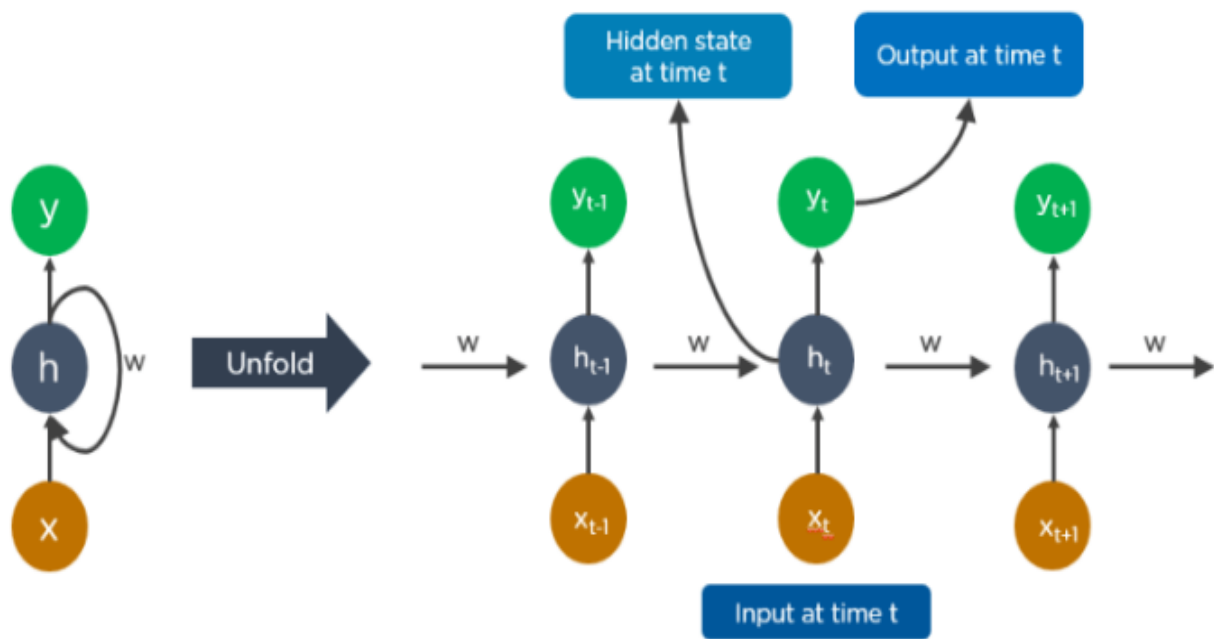


Figure 13 : Rnns architecture

## 5 Fundamentals of Pattern Recognition:

Pattern recognition is a cornerstone of artificial intelligence that involves identifying and classifying regularities within data. This section outlines the key principles and techniques underlying this field. [21]

### 5.1 Definition of Pattern Recognition :

Pattern recognition transforms raw inputs (images, audio signals, or temporal sequences) into structured representations using mathematical models and statistical methods.

### 5.2 Process of Pattern Recognition :

The process begins with data acquisition, followed by preprocessing to enhance the signal to noise ratio. Features are then extracted and used by classification algorithms to identify patterns in high dimensional spaces.

### 5.3 Role in Artificial Intelligence :

- **Techniques:** Early methods included clustering, principal component analysis, and simple supervised models. With advancements in computing, modern systems use complex nonlinear models to capture intricate data relationships.
- **Feature Extraction:** In image processing, features such as edges, textures, and shapes are identified and used to reduce data complexity while enhancing classification accuracy.
- **Deep Learning Impact:** Deep neural networks automatically learn hierarchical features, improving performance across tasks from natural language processing to image and video analysis.
- **Literature Review:** Numerous studies in journals like *Pattern Recognition* and *IEEE Transactions on Neural Networks* document the evolution from traditional methods to current deep learning approaches.

### 6 Sign Language Recognition :

This section specifically addresses sign language recognition as a unique application of pattern recognition.

#### 6.1 Importance :

- **Social Impact:** Sign language is a complete language system crucial for communication among the deaf and hard of hearing community. Reliable recognition systems can bridge communication gaps and promote inclusivity.
- **Accessibility:** Transforming sign language into text or speech enables better integration into educational, professional, and public service settings.
- **Empowerment:** Providing real time translation fosters independence and improves interaction with society at large.
- **Cultural Sensitivity:** Recognizing the nuances of regional and cultural variations in sign language is essential for developing systems that respect and accommodate diversity.
- **Research Significance:** Studies have shown that high performance recognition systems can significantly enhance quality of life, making the development of such systems a priority in assistive technology research.

### 6.2 Challenges :

- **Variability of Gestures:** Sign language involves dynamic, continuous hand movements and facial expressions that vary widely among individuals and contexts.
- **Environmental Interference:** Factors such as lighting conditions, background clutter, and occlusions can degrade the performance of segmentation and feature extraction processes.
- **Temporal Dynamics:** The sequential nature of sign language requires models that can capture contextual dependencies between gestures, making time series analysis essential.
- **Data Requirements:** Building robust models demands large, annotated datasets that cover a wide range of signing styles and environmental conditions.
- **Real Time Processing:** Ensuring low latency is critical for interactive applications, which challenges the computational efficiency of complex deep learning models.

### 6.3 Deep Learning in Sign Language Recognition :

This section discusses how the field has evolved from classical computer vision methods to modern deep learning techniques.

#### 6.3.1 Traditional Techniques :

Before the rise of deep learning, traditional methods were widely used for gesture recognition. These techniques relied heavily on manual feature extraction and performed best in controlled environments:

- **Handcrafted Features:** Early systems relied on manual extraction of features like histograms of oriented gradients (HOG) or scale invariant feature transform (SIFT) descriptors to capture hand shapes and movements.
- **Rule Based Classification:** Techniques such as Support Vector Machines (**svms**) and Hidden Markov Models (**hmms**) were used to classify gestures based on these features, though their performance was limited by the quality of manual feature selection.
- **Controlled Environment Performance:** While these methods worked under controlled settings, they struggled to adapt to the variability and noise found in real world environments.

### 6.3.2 Deep Learning Approaches :

Deep learning has significantly advanced the field of sign language recognition by automating feature extraction and improving accuracy. The following models have proven especially effective:

- **Convolutional Neural Networks (cnns):** cnns automatically extract spatial features from raw images, capturing detailed visual information about hand shapes without the need for manual intervention.
- **Recurrent Neural Networks (rnns) and lstms:** These models are adept at processing sequential data, making them ideal for capturing the temporal dynamics of sign language.
- **Hybrid Models:** Integrating **cnns** for spatial feature extraction with **lstms** for temporal sequence modeling has shown impressive improvements in accuracy, often exceeding 90% in controlled experiments.
- **Challenges and Solutions:**

Although deep learning models have advanced the field, they require large datasets and are sensitive to environmental factors. Researchers continue to explore data augmentation, transfer learning, and sensor fusion to address these issues.

### 6.4 Practical Development Process of a Sign Language Recognition System :

This section outlines the step by step process for developing a practical system.

- **Data Collection :**
  - **Dataset Creation:** High quality, annotated video data of sign language gestures is crucial. Recordings should capture multiple signers, varied lighting conditions, and diverse backgrounds.
  - **Existing Databases:** Utilizing established datasets such as RWTH PHOENIX Weather can provide a valuable foundation for model training.
  - **Diversity:** Ensuring the dataset reflects the full variability of sign language is key to building a robust model.
  
- **Preprocessing :**
  - **Segmentation:** Techniques such as thresholding and edge detection are applied to isolate the hand from the background.
  - **Image Enhancement:** Preprocessing steps include noise reduction and normalization to standardize input images, thereby improving model performance.
  - **Multi Sensor Integration:** In some cases, additional sensors (e.g., depth cameras) are used to complement visual data, further enhancing segmentation accuracy.

- **Feature Extraction :**
  - **Automatic Extraction via cnns:** Modern systems leverage cnns to automatically learn spatial features, reducing the reliance on manual feature engineering.
  - **Hierarchical Representations:** Lower layers capture basic visual cues, while higher layers abstract more complex patterns such as finger positions and orientations.
  - **Robustness:** The features extracted must be invariant to changes in lighting, scale, and orientation to be effective across varied conditions.
  
- **Temporal Analysis :**
  - **Sequential Modeling:** LSTM networks process the sequence of feature vectors derived from CNN outputs, capturing the temporal dynamics of sign language gestures.
  - **Contextual Dependencies:** This stage ensures that the recognition system understands not just isolated gestures but also their context within a continuous signing sequence.
  - **Integration:** The output of the temporal analysis is a cohesive representation that is ready for final classification.
  
- **Integration and Deployment :**
  - **Classification and Output:** The final classification layer maps the processed features to specific sign language words or phrases, which can then be converted into text or speech.
  - **User Interface:** A real time interface is essential for interactive applications, providing immediate feedback to the user.
  - **System Optimization:** Rigorous testing and continuous feedback from end users help refine the system, ensuring it performs reliably in real world settings.

### 6.5 Evaluation, Practical Challenges, and Future Prospects :

This section covers how the system is evaluated and discusses both current challenges and future directions.

### 6.5.1 Evaluation Metrics :

To objectively assess the performance of a sign language recognition system, a combination of quantitative and qualitative evaluation metrics is employed. These evaluations ensure the system meets both technical accuracy and user expectations:

- **Quantitative Measures:** Performance is measured using metrics such as accuracy, precision, recall, and F1 score on test datasets.
- **Qualitative Feedback:** User trials assess the system's responsiveness, reliability, and overall user satisfaction in real world scenarios.
- **Robustness Testing:** Evaluations are performed under various environmental conditions to ensure the system's adaptability and consistency.

### 6.5.2 Practical Challenges :

Despite advancements in deep learning, several practical obstacles continue to affect real-time deployment and usability of sign language recognition systems. Understanding these challenges is critical for improving performance and scalability:

- **Real Time Processing:** Deep learning models can be computationally intensive, potentially introducing latency in processing video streams.
- **Environmental Variability:** Issues such as poor lighting, motion blur, and background noise continue to challenge the segmentation and recognition processes.
- **Data Scalability:** Expanding the system to accommodate diverse sign languages, dialects, and user specific variations requires extensive data collection and model retraining.
- **Integration Complexity:** Merging data from multiple sensors (visual, inertial, depth) to enhance accuracy adds layers of complexity that must be managed carefully.

### 6.5.3 Future Prospects :

The future of sign language recognition lies in more intelligent, flexible, and integrated systems. Researchers and developers are focusing on emerging strategies to enhance performance, accessibility, and global deployment:

- **Multi Modal Sensor Fusion:** Future systems may integrate additional sensor modalities to further improve accuracy and robustness.

- **Model Efficiency:** Advances in model optimization, including hardware acceleration (gpus, AI chips) and edge computing, will likely reduce processing latency.
- **Adaptive Learning:** Techniques like transfer learning and few shot learning can help systems adapt to new sign languages or variations with minimal additional data.
- **Wider Deployment:** As the technology matures, sign language recognition systems are expected to become integral components of communication aids, educational tools, and public service interfaces, greatly enhancing inclusivity.

### 7 Conclusion :

In conclusion, this chapter demonstrated the importance of leveraging advanced pattern recognition and deep learning techniques for sign language recognition. It outlined the critical challenges posed by gesture variability and environmental factors, while highlighting the potential of cnns and lstms in bridging these gaps. By integrating spatial and temporal analysis, such systems can significantly enhance communication and foster inclusivity for individuals with special needs. Ongoing research and development will further refine these approaches for effective real world applications.

# **Chapter 3:**

## **Conception and implementation**

### **1 Introduction:**

The continuous evolution of assistive technologies is one of the most important factors in enhancing the quality of life for individuals with disabilities. The most overlooked community is that of deaf and mute individuals, who are often impeded by communication barriers in everyday life, such as at school, in the workplace, during doctor patient consultations, or even in casual social interactions.

While there are other communication tools and mobile apps to help them out, these kinds of solutions rely on manual effort, are not in real time, and insist on constant attention to a screen making them inconvenient and unsuitable for improvised interactions.

This project, " Hope Sign Language Translation Glasses," aims to fill these voids with a wearable technology that interprets sign language gestures in real time and converts them to spoken words in different languages. This system aims to provide voice for deaf and mute individuals, facilitate their free conversation with others, and eventually gain independence and acceptance in society.

### **2 Conception:**

#### **2.1 Project Overview:**

Deaf and mute people are placed at a significant disadvantage in day to day life because they are not being spoken to; this can mean they are unable to get education, health care, employment opportunities and to be fully involved in social situations. Traditional sign language interpreters are not always available and as many people around the globe cannot read sign language, inclusive communication is missing to an unacceptable level.

The second reason why most individuals (most disadvantaged ones) cannot use existing text based helper apps is that they are not reading. Some current smart glasses already in existence on the market are too expensive, too big or complicated to use on a regular basis.

But more importantly the design of the “hope” glasses is focused on these things:

- Real time sign language recognition using a deep learning model trained on a large dataset of sign language gestures.

## Chapter3 : Conception and implementation

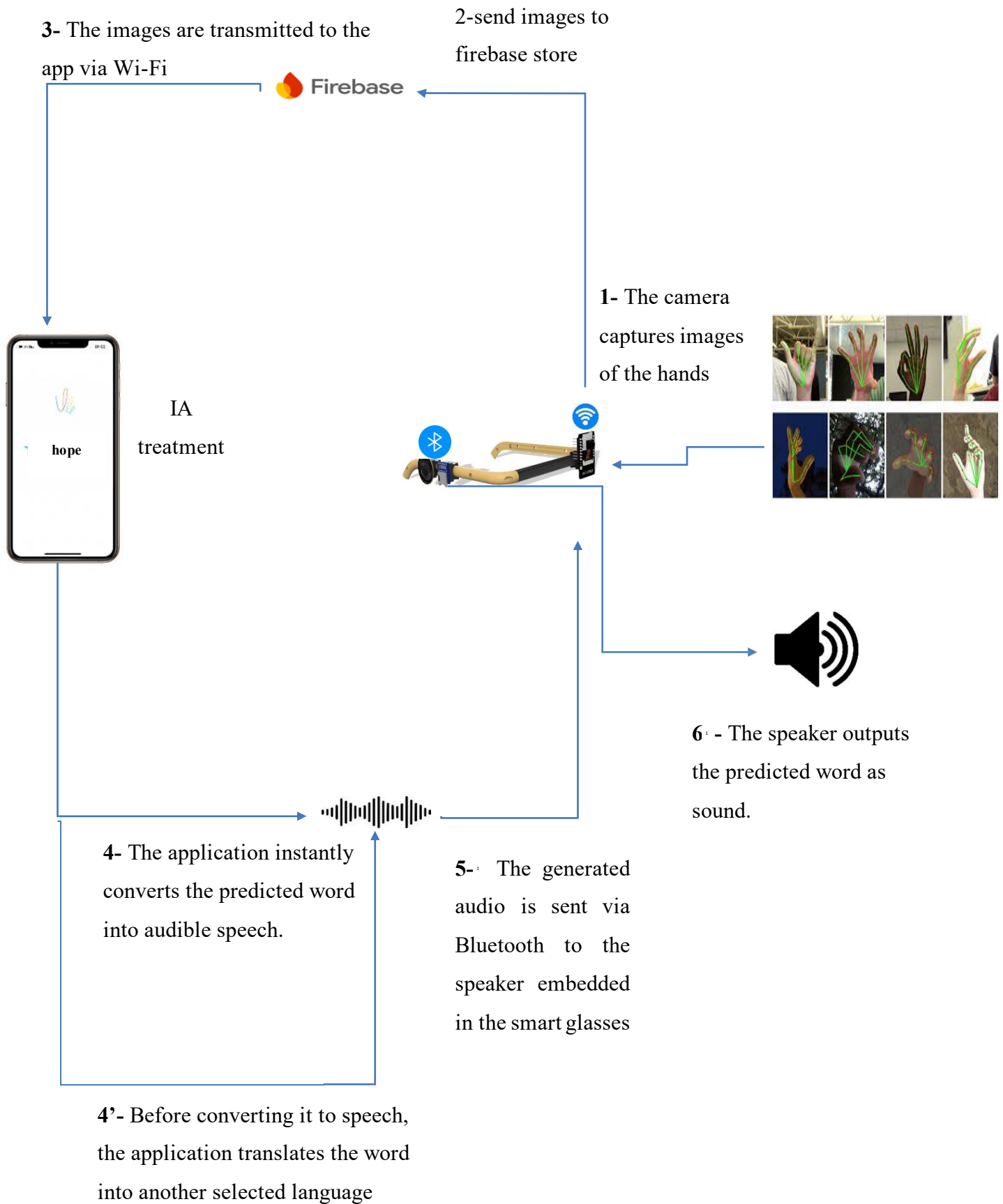
---

- Audio output of recognized gestures. That would allow deaf/mute people to “speak“ through the glasses.
- Multi language support (when the spoken output can be translated into different languages depending on the user's choice).
- On device processing using tflite, ensuring real time performance and eliminating dependency on internet connection or cloud services.
- Hands free operation with built in camera and fully automatic interaction without the user touching the device.
- Mobile application integration
- Ergonomic design, as the glasses will look like normal eyewear so social stigma wont fall on them.
- Low cost and accessibility, focusing on affordability for users in developing countries or with limited resources.

With these features hope glasses aim to bridge the communication gap with deaf and mute people and offer them a practical, intuitive and economical solution.

# Chapter3 : Conception and implementation

## 2.2 Working Principle:



### 2.3 Challenges and Solutions:

- Real Time Performance:
  - **Challenge:** Delay in cloud processing could affect user experience.
  - **Solution:** Optimized image resolution and used efficient model architecture.
- Limited Dataset:
  - **Challenge:** Difficulty finding a large, high quality dataset for sign language gestures.
  - **Solution:** Created a custom dataset using recorded gestures from volunteers and applied data augmentation techniques to increase variability.
- Model Integration with Flutter:
  - **Challenge:** Initially, Firebase ML was used to host the trained .tflite model and connect it with the Flutter app. However, this approach led to several issues, including high latency in prediction, reliance on internet connectivity, and difficulty in achieving real time performance all of which negatively impacted the user experience.
  - **Solution:** To overcome these limitations, the model was instead embedded directly into the Flutter application using the tflite\_flutter plugin. This shift enabled fully offline, on device inference, significantly reducing response time and ensuring consistent, real time translation of sign language without relying on external services.
- Multi language Support:
  - **Challenge:** Difficulty in training the model to recognize multiple spoken languages.
  - **Solution:** The model is trained in English, and external translation libraries from Google and Flutter are used to translate the output text into various languages before sending it to the speech engine.

## Chapter3 : Conception and implementation

### 2.4 Organigramme :

The diagram below illustrates the various functions integrated into our smart glasses and how it works.

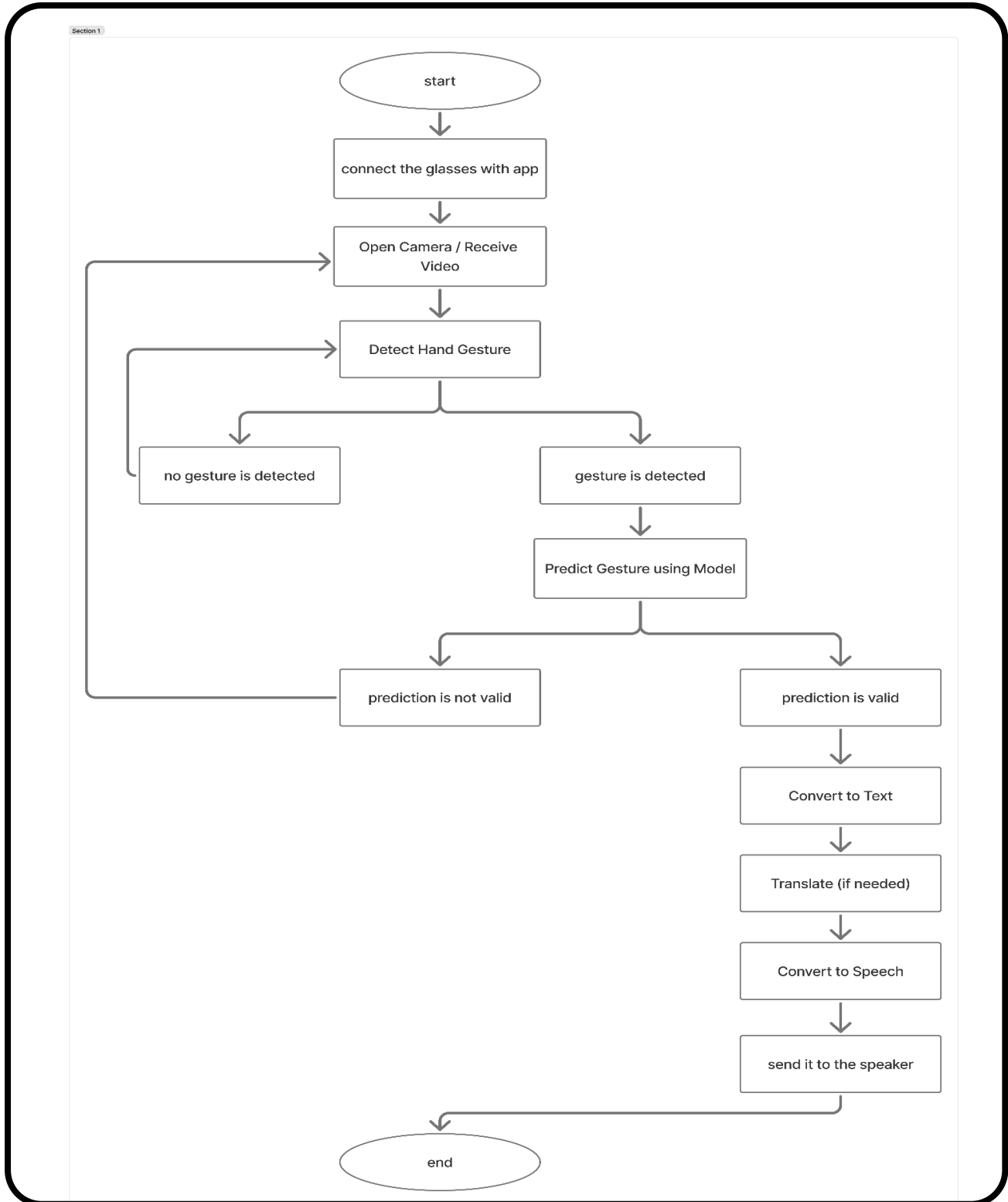


Figure 14 : The organigram of the project

### 2.5 Future Work and Enhancements:

This project is designed with scalability in mind. Future upgrades include:

- Bi directional Communication: Enable the app to recognize voice input and convert it to sign language .
- Expanded Vocabulary: Continuously train and update the model with new signs and expressions.
- Facial Expression Recognition: To improve the context and emotional tone of sign language.
- Emergency Features: Include SOS voice alerts or automatic location sharing in urgent situations.

### 3 Implementation:

To develop the Hope glasses, a rigorous selection of materials and software tools was made to meet functional requirements and performance standards. This section outlines the key hardware components and software development environments chosen for this innovative project.

#### 3.1 Hardwar layer:

##### 3.1.1 ESP32 CAM Camera:

The ESP32 CAM is a Wi Fi module equipped with an ESP32 microcontroller and an OV2640 camera. It is designed for use in Internet of Things (iot) projects that require live video capture and wireless data transmission. Below are the key features of the ESP32 CAM:

- Dual core ESP32 microcontroller clocked at 160 mhz,
- 802.11b/g/n Wi Fi connectivity,
- Integrated antenna,
- 2 MP OV2640 camera,
- Supports video capture up to 1600×1200 pixels,
- Supports JPEG, BMP, GIF, PNG, and YUV422 formats,
- Includes a restful API based programming interface,
- Supports MQTT, HTTP, and websocket communication protocols.

[22]



Figure 15 : ESP32 CAM Camera

### 3.1.2 speakers :

The speaker outputs the predicted word as sound.



Figure 16 :Speakers

### 3.1.3 Wi Fi :

Wi Fi is a wireless networking technology that allows devices such as computers (laptops and desktops), mobile devices (smart phones and wearables), and other equipment (printers and video cameras) to interface with the Internet.



Figure 17 : WIFI icon

### 3.1.4 Bluetooth:

Bluetooth technology allows devices to communicate with each other without cables or wires. Bluetooth relies on short range radio frequency, and any device that incorporates the technology can communicate as long as it is within the required distance



Figure 18: Bluetooth icon

### 3.2 Mobile application layer:

#### 3.2.1 Flutter:

Flutter is Google's portable UI toolkit for crafting beautiful, natively compiled applications for mobile, web, and desktop from a single codebase. Flutter works with existing code, is used by developers and organizations around the world, and is free and open source.



Figure 19 : Flutter logo

### 3.2.2 Dart:

Dart is a client optimized language designed for building fast applications on any platform. Its goal is to provide the most productive programming language for cross platform development, combined with a flexible execution platform for application frameworks.

Languages are defined by their technical envelope the design choices made during development that shape a language's capabilities and strengths. Dart is built with a technical envelope specifically suited for client side development, prioritizing both developer experience (hot reload with sub second stateful updates) and high quality production performance across a wide range of compilation targets (web, mobile, and desktop).

Dart also serves as the foundation for Flutter. It powers Flutter applications with its language and runtime environments, while also supporting essential development tasks such as code formatting, static analysis, and unit testing.



Figure 20 : Dart logo

### 3.2.3 Vs code :

Visual Studio Code (vscode) is a source code editor and integrated development environment ( IDE ) from Microsoft. It is open source and cross platform, meaning it runs on Windows, Linux, and Mac. It was designed for web developers, but it supports many other programming languages such as C++, C#, Python, Java, etc. It offers many features like syntax highlighting, auto completion, error highlighting, code navigation, debugging, version control, integration with Git , and many more. It is also extensible using a wide variety of community developed extensions, allowing developers to customize the editor according to their needs. [23]



Figure 21: Vs code logo

### 3.2.4 Flutter TTS:

Flutter Text to Speech (TTS) is a highly useful dependency for converting text into speech in Flutter app development. This package enables your Flutter application to read text aloud in multiple languages and with customizable voices. It is simple to use and offers extensive configuration options. [24]

### 3.3 firebase and AI processing layer:

#### 3.3.1 fire base :

Provides detailed documentation and cross platform app development sdks, to help build and ship apps for ios, Android, the Web, Flutter, Unity, and C++. [25]



- **Storage :**

Firestore storage is a service that developers can use for various purposes including uploading, storing, and downloading files generated by app users. The best thing about this type of storage system is that there is no need for server side coding.

- **Function :**

Firestore Cloud Functions is a serverless computing service provided by Firebase, which allows you to run backend code in response to events triggered by Firestore services such as changes to data in the Firestore Realtime Database, new user sign ups via Firestore Authentication, and updates to files in Firestore Storage.

- **Realtime:**

Firestore's first product was the Firestore Realtime Database, an API that synchronizes application data across ios, Android, and Web devices, and stores it on Firestore's cloud. The product assists software developers in building real time, collaborative applications.

- **Machine learning:**

A set of ready to use cloud based apis for common mobile use cases: recognizing text, labeling images, and recognizing landmarks. Unlike on device apis, these apis leverage the power of Google Cloud's machine learning technology to give a high level of accuracy.

### 3.3.2 AI processing

- **Python:** Python is a popular programming language. It was created by Guido van Rossum, and released in 1991. [26]

It is used for:

- Web development (server side),
- Software development,
- Mathematics,
- System scripting.



Figure 23 :python logo

- **Tensorflow :**

The tensorflow platform helps developers implement best practices for data automation, model tracking, performance monitoring, and model retraining. Using production level tools to automate and track model training over the lifetime of a product, service, or business process is critical to success. [27]



Figure 24 : tensorflow logo

- **Jupyter:**

Jupyter Notebook is used to create interactive notebook documents that can contain live code, equations, visualizations, media and other computational outputs. Jupyter Notebook is often used by programmers, data scientists and students to document and demonstrate coding workflows or simply experiment with code.



Figure 25: jupyter logo

- **Collab:**

Colab is a hosted Jupyter Notebook service that requires no setup to use and provides free of charge access to computing resources, including gpus and tpus. Colab is especially well suited to machine learning, data science, and education



Figure 26 : collab logo

- **Kaggle:**

Kaggle is a platform for data science competitions, where data scientists and machine learning engineers can compete to create the best models for solving specific problems or analysing certain data sets.

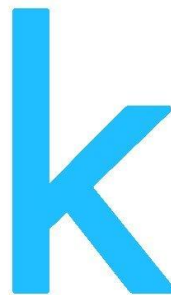


Figure 27 : kaggle logo

### 4 Project implementation:

- **Software layer :**
  - **Splash screen :**

This is the app's welcome screen that briefly shows the Hope logo while the app loads in the background.



Figure 28 : Splash screen

- **Choose method screen :**

This screen lets users pick how they'll use the app: as a glasses user (for deaf users) or as an app user (for hearing users), ensuring a tailored experience.

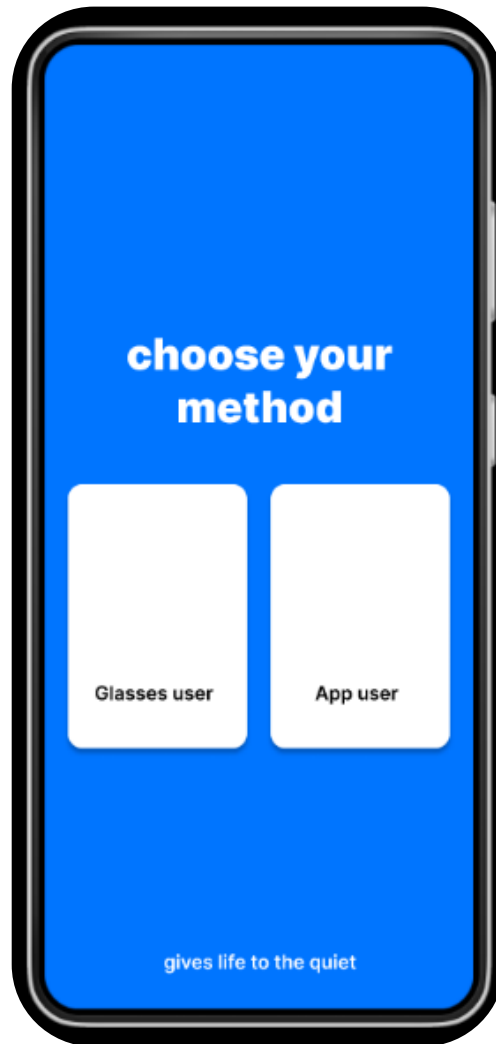


Figure 29 : Choose method screen

- **Glasses user method :**

In this mode, the camera sends hand gestures to the app, which predicts the word and plays it aloud allowing the user to “speak” through signs.

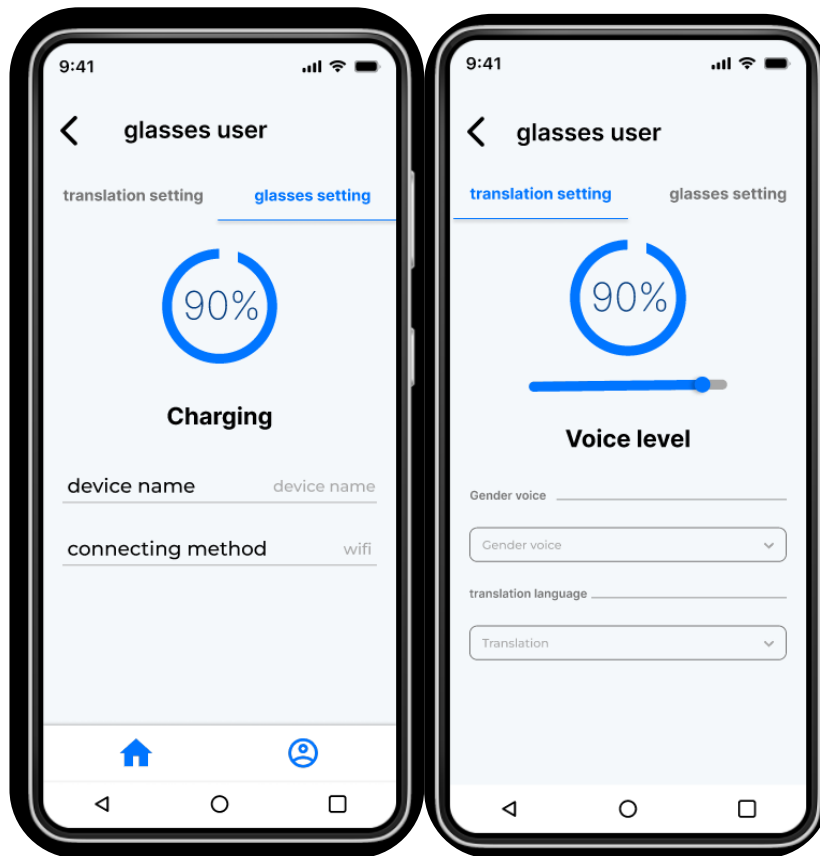


Figure 30 : Glasses user method screen

- **App user method :**

This mode is for hearing users. The phone's camera captures sign language from a deaf person and displays:

- Live video feed
- Real-time word prediction
- Optional voice playback

It turns the app into a visual interpreter, helping bridge communication gaps.

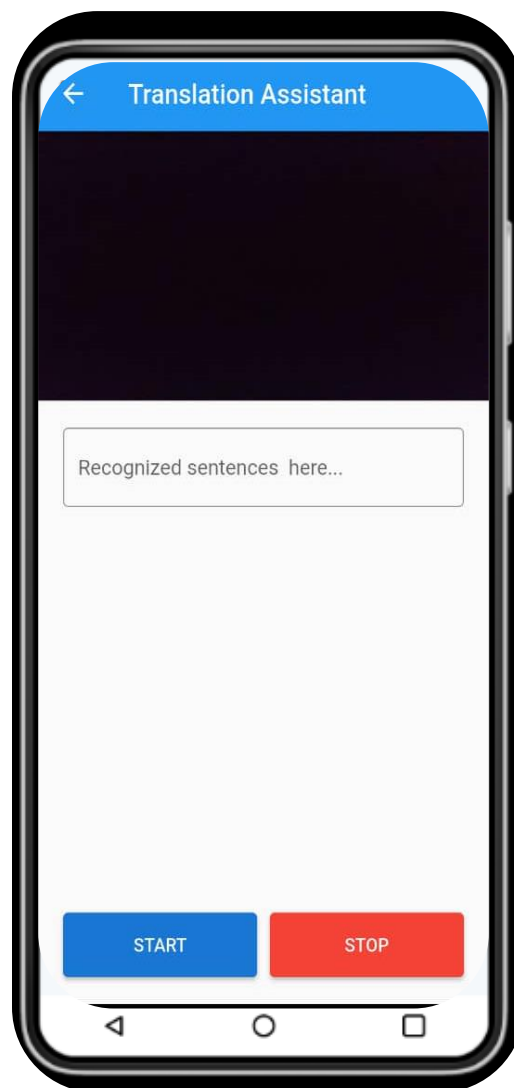


Figure 31 : App user method screen

## Chapter3 : Conception and implementation

---

- **Hardware layer :**

Hope smart glasses include a camera, ESP32-CAM processor, and speaker, all built into a lightweight frame. These components work together to capture gestures, process them, and provide audio output in real time.

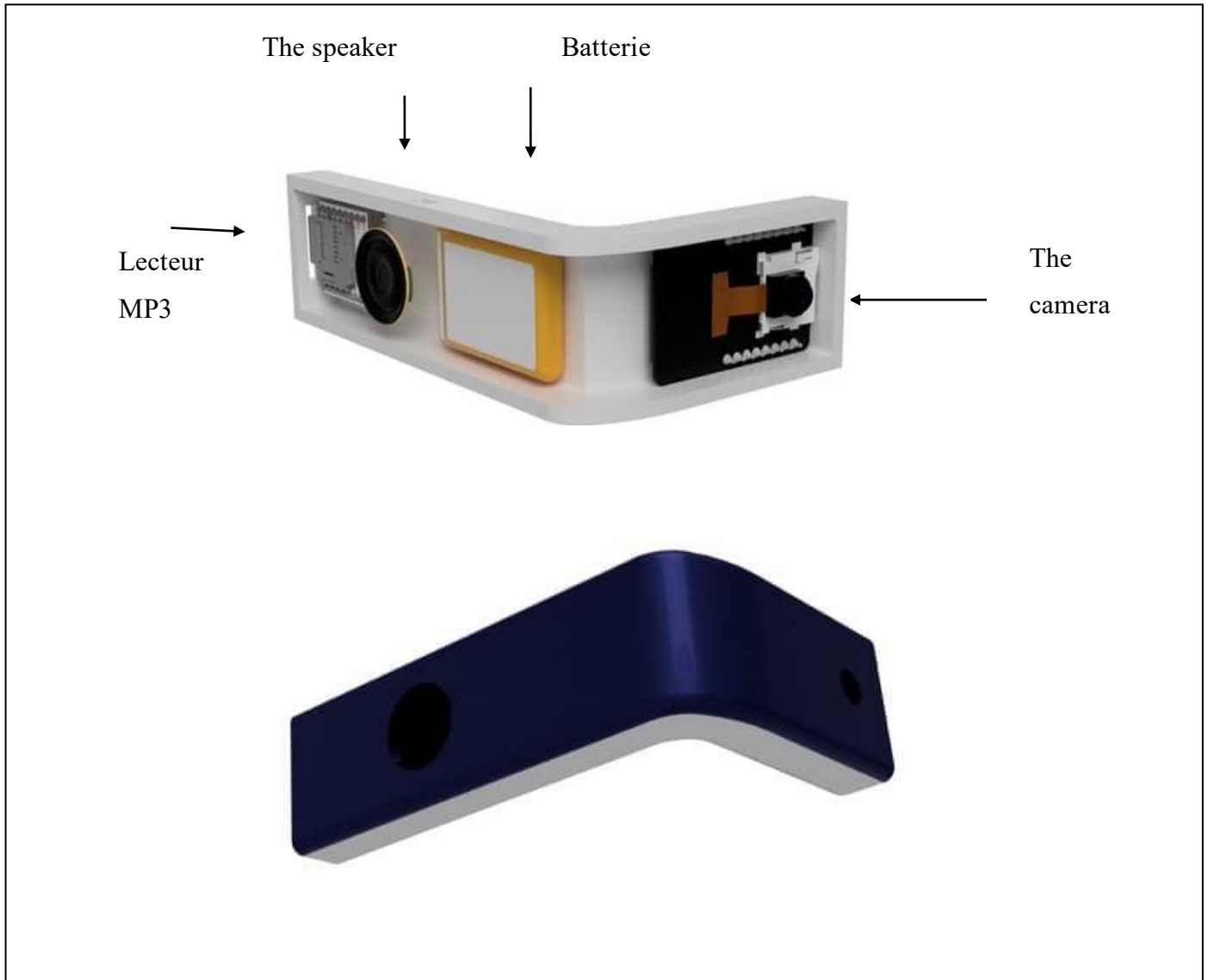


Figure 32 : The glasses design

### **5 Conclusion :**

This chapter detailed the technical foundation of the Hope Smart Glasses project, highlighting the integration between hardware components, mobile application layers, and artificial intelligence.

From the camera based image capture to the real time prediction and audio feedback, each element was carefully designed to create a seamless user experience. The challenges faced during development such as connectivity, prediction speed, and hardware limitations were addressed through thoughtful solutions and efficient system architecture.

The result is a functional and innovative prototype that demonstrates the practical potential of technology to break communication barriers and support inclusivity.

# **Chapter 4 :**

# **Commercial Overview**

### 1 Section 1: Project Presentation

#### 1.1 Field of Activity :

The "Hope" project works in a place where helpful tech, smart gear you wear, and smart computer systems meet. It looks to solve a big problem in our societies: making it easy for people who cannot hear or speak to talk with those who can hear. In many groups, people who talk with their hands face real talk roadblocks day by day in schools, at public places, in hospitals, or just when hanging out with others. Even though people who translate and old ways of learning hand talk help a bit, these ways don't always work well, aren't easy to use everywhere, or can't be used by everyone.

The "Hope" system brings a new fix by using live computer view and word handling in two ways: smart glasses for deaf people and a phone app for those who can hear. These smart glasses have a small camera that sees hand signs, works on them with an AI model, and then turns them into spoken words. At the same time, people who can hear use the phone app to get the sign language just by aiming their phone's camera at the person signing. This lets both sides talk without needing to know sign language or wear any extra gear.

This big leap closes a long gap by cutting out the need for others in talking, and it helps real social mix and self reliance. The work does more than fix a health or word issue it changes the way we talk by turning hand moves into sound and letting the quiet signs of many be known right away.

#### 1.2 Genesis and Evolution of the Idea :

The start of the "Hope" plan comes from deep care, observation, and a strong desire to promote inclusion. It all began with a key but simple truth: in a world so connected, communication should be a fundamental right not a privilege limited by physical ability or language barriers. While observing behavior in schools, healthcare facilities, and public institutions, the team noticed that deaf and mute individuals often felt uncertain, excluded, or overlooked.

Even in everyday settings like stores, buses, or social gatherings, they encountered invisible barriers that prevented them from expressing needs, thoughts, or emotions.

---

## Chapter4: Commercial Overview

---

As the idea matured, it became clear that many existing solutions placed the full burden of communication on the deaf person. They were expected to adapt, carry devices, or use unfamiliar interfaces, while hearing individuals remained passive. This approach felt both unfair and ineffective.

Another major concern was the slowness of current technologies. Many tools required high end devices, constant internet access, or manual switching between functions. Conversations that should have been fluid and emotionally rich became slow and rigid.

At this critical stage, with guidance and insightful feedback from our supervisor, the idea was not only selected but also thoughtfully refined and shaped into a meaningful solution. Their input helped clarify the vision, address real world challenges, and ensure that the project was both technically sound and socially impactful.

"Hope" evolved to become more than a translator it became a powerful two-way bridge. It is a system where deaf individuals can naturally communicate using real sign language via smart glasses, while hearing people can understand those signs without prior training, simply by using their smartphones. This vision of equal empowerment where both parties contribute and benefit formed the foundation of the "Hope" ecosystem: a leap forward in respect, accessibility, and shared responsibility in communication.

### **1.3 Project Objectives :**

The core objectives of the "Hope" project are:

- To eliminate communication barriers between deaf individuals and the hearing population.
- To empower deaf mute users with the ability to express themselves clearly and independently using wearable technology.
- To enable hearing individuals to understand sign language users instantly without prior knowledge.
- To support bilingual and multilingual communication by allowing voice output in multiple languages.

## Chapter4: Commercial Overview

---

- To develop a scalable and practical assistive technology that can be integrated into daily life, education, and public services.

### **1.4 Bidirectional Use Cases :**

The “Hope” system has been meticulously designed to be both inclusive and flexible, operating through two complementary modes that address the needs of both deaf mute and hearing users. These modes are independent yet interoperable, creating a seamless communication bridge between users who otherwise speak different languages.

## Chapter4: Commercial Overview

---

- **Glasses Mode (for deaf mute users):** This mode allows deaf or mute individuals to naturally communicate using their native sign language without changing their behavior or learning new technologies:
  - The user wears smart glasses equipped with a compact yet powerful embedded camera positioned above the eyes to closely track hand gestures and movements.
  - The camera continuously captures hand signs in real time as the user communicates, even in outdoor or variable lighting environments, thanks to adaptive image preprocessing techniques.
  - The video stream is fed into a deep learning model trained specifically to recognize a wide range of sign language gestures. This model can run either locally on a lightweight edge processor (e.g., ESP32 with tensorflow Lite) or via a high performance Firebase ML backend depending on the configuration.
  - Upon recognition, the gesture is converted into a corresponding word or phrase, which is then processed by a text to speech (TTS) engine.
  - The message is instantly spoken out loud using a miniaturized speaker integrated within the glasses frame, allowing hearing individuals nearby to understand the user in real time.
  - The system is hands free, fully autonomous, and optimized for low latency, ensuring uninterrupted and natural conversations in dynamic environments such as classrooms, hospitals, or public venues.
- **App Mode (for hearing users):** This complementary mode is tailored for hearing individuals who wish to understand sign language users, without requiring them to wear or operate any special device:
  - The user opens the “Hope” mobile application and activates the camera.
  - By pointing the smartphone camera at the deaf individual using sign language, the app starts capturing the hand gestures in real time.
  - An AI powered computer vision system processes each frame on device or in the cloud, depending on performance requirements and connectivity.
  - Detected signs are interpreted into text and displayed clearly on the screen. Optionally, the translated text is also spoken aloud using the smartphone's speaker.

---

## Chapter4: Commercial Overview

---

- This interaction mode enables hearing users be it teachers, medical staff, family members, or strangers to understand and respond to deaf individuals without needing to learn sign language themselves.

The coexistence of these two modes is what makes “Hope” a truly bidirectional solution. It doesn't rely on one party making all the effort. Instead, both deaf and hearing individuals are given intuitive tools adapted to their context, fostering equality in dialogue and restoring the natural flow of human connection.

### **1.5 Project Team :**

The project is led by Medjbouri Ismahan as part of a university graduation thesis and Dr.CHEIKH Mohamed as supervisor. The core team includes:

- AI Developer: Responsible for training and optimizing the gesture recognition model.
- Mobile App Developer: In charge of building the cross platform Flutter application.
- UX/UI Designer: Focused on creating an intuitive user interface.
- Academic Advisor: Providing technical guidance and oversight.

Together, the team represents a blend of technical, design, and scientific expertise.

### **1.6 Implementation Environment :**

The prototype was developed in a university laboratory using open source tools and platforms such as Python, tensorflow, Firebase, and Flutter. Hardware components like the ESP32 CAM and portable audio modules were integrated to ensure mobility and low cost manufacturing. Future iterations will aim to optimize performance, reduce latency, and explore partnerships with startups or ngos for broader deployment.

### **1.7 Value Proposition :**

“Hope” offers a transformative leap in assistive communication by delivering a dual mode communication platform that is both inclusive and user centric. Its value lies not just in the technological components, but in how it reshapes the relationship between deaf and hearing individuals.

## Chapter4: Commercial Overview

---

- Real time, bidirectional communication: Most existing tools only enable one way interpretation usually requiring the deaf individual to adapt. “Hope” breaks this mold by providing two independent, interactive communication paths. Whether the deaf user wears smart glasses or the hearing user employs the app, both parties can initiate conversations equally and spontaneously.
- Inclusive access for all users: Unlike systems that require specific training or manual input, “Hope” is designed with simplicity and intuitiveness at its core. The deaf user simply signs as usual, and the hearing user needs only to open an app and point the camera. This natural interaction lowers the barrier to entry, allowing broader adoption across age groups, literacy levels, and technical skills.
- **Ergonomic and user friendly design:** The glasses are lightweight, discreet, and designed to blend with common eyewear aesthetics. The app interface is minimalist, responsive, and tailored for non expert users, which makes it viable for real world use in busy environments like classrooms, hospitals, and public offices.
- **Multilingual support and personalization:** Voice output can be generated in various languages depending on user preference, facilitating communication across cultural boundaries. This makes “Hope” a globally scalable solution, capable of being used in international organizations, tourist hubs, and cross cultural exchanges.
- Cost effectiveness and scalability: Built using open source frameworks and affordable components, “Hope” can be mass produced and deployed in public institutions with limited resources. It also opens doors for **ngos**, schools, and healthcare providers to enhance accessibility without incurring prohibitive costs.

More than a product, “Hope” represents a philosophy of inclusion an embodiment of equity, dignity, and empowerment. It shifts the paradigm from assistance to autonomy, enabling people who use sign language to engage with the world on their own terms while allowing society at large to meet them halfway. This innovation doesn’t just improve communication it redefines what it means to connect.

### 1.8 Project Development Plan

Phase	Activities	Duration
Needs Assessment	Researching user requirements	1 week
AI Model Training	Training and testing the gesture model	4 weeks
App Development	Building and integrating Flutter app	6 weeks
Hardware Integration	Developing the wearable smart glasses	2 week
Testing and Feedback	Testing with real users and gathering data	1 week
Final Adjustments	Refinement based on user feedback	2 week

Table 2:Project Development Plan

## 2 Section 2: Innovation Aspects :

### 2.1 Nature of Innovation :

“Hope” introduces multiple layers of innovation:

- A dual system architecture: wearable and mobile.
- Real time gesture recognition with low latency feedback.
- Integration with cloud services for model updates and language translation.
- Use of lightweight, open source components to keep costs low.
- Support for voice customization and multilingual output.

### 2.2 Areas of Innovation :

The project impacts several domains:

- **Health and Rehabilitation:** Enhancing independence for the hearing impaired.
- **Education:** Promoting inclusive classrooms.
- **Public Service Accessibility:** Enabling better communication in hospitals, transport, and administration.
- **AI and Wearable Tech:** Pioneering real time human AI interaction in a social context.

### 3 Section 3: Invention Description

#### 3.1 Patent Title :

Sign Language Translation Glasses (hope)

#### 3.2 Abstract :

This invention proposes a hybrid system comprising smart wearable glasses and a mobile application for real time translation of sign language to spoken words. The glasses are equipped with a vision sensor that captures hand gestures and transmits data to an AI module, which interprets the gestures and outputs them as audible speech. The mobile app provides an alternative interface for hearing users to understand sign language using their phone camera. The system supports multiple languages and offers bidirectional use for both deaf and hearing individuals. It serves as a breakthrough in assistive communication.

#### 3.3 Technical Fields :

- Artificial Intelligence
- Computer Vision
- Assistive Technology
- Wearable Devices
- Mobile Computing

### 3.4 Objective of the Invention :

The primary objective of the "Hope" invention is to break down the long standing communication barrier that separates the deaf and hearing communities. This barrier often limits access to education, healthcare, employment, and social integration for individuals who use sign language as their primary means of communication. The invention aims to provide an intelligent, responsive, and mobile solution that allows both parties to interact in real time without the need for human interpreters, printed text, or pre existing language knowledge.

Unlike conventional systems that offer one way translation or static interpretation, the goal of "Hope" is to foster dynamic, natural dialogue. It allows deaf users to express themselves using native gestures while ensuring hearing users receive fluent voice based translations. In reverse, it empowers hearing users to understand sign language via a simple app interface. This seamless interaction ensures that neither party is disadvantaged, thus reinforcing inclusion, autonomy, and equality in communication.

### 3.5 Core of the Invention :

At the heart of the invention are five foundational technologies and design principles that make the “Hope” system functional, scalable, and inclusive:

- Real time image analysis for gesture recognition: Using computer vision and deep learning, the system captures and processes sign language gestures instantaneously, allowing uninterrupted conversations even in fast paced environments.
- Voice synthesis for output: Once a gesture is recognized, the corresponding text is passed to a Text to Speech (TTS) engine that produces audible speech in a clear, human like voice. The output can be personalized in terms of language.
- Dual interaction modes: The solution consists of both a wearable device (smart glasses) and a mobile application. The glasses are tailored for deaf users to transmit their signs as voice, while the app enables hearing users to receive visual signs as translated text or speech.
- On device AI model with future scalability: The learning model is embedded directly within the mobile application using tensorflow Lite, enabling fast, offline predictions without relying on cloud services. While the current version runs entirely on device for

## Chapter4: Commercial Overview

---

optimal speed and privacy, the system is designed with future scalability in mind allowing the model to be retrained and updated periodically with new gestures, languages, and regional variations as needed.

- User focused modular design: Components such as the TTS engine, recognition model, language settings, and interface are modular. This allows developers and institutions to adapt the system to specific use cases whether in education, healthcare, or public service without redesigning the entire platform.
- Real time image analysis for gesture recognition.
- Voice synthesis for output.
- Dual interaction modes: glasses and mobile app.

### 4 Section 4: Claims

- A dual mode intelligent communication system capable of translating sign language gestures into real time audible speech using wearable and mobile technologies.
- Smart glasses integrated with an embedded camera, real time gesture recognition module, and voice output component, allowing deaf or mute users to engage in fluid conversations with hearing individuals.
- A mobile application that leverages the smartphone's camera to detect, interpret, and vocalize sign language gestures, requiring no wearable hardware.
- Integration with cloud based machine learning infrastructure for continuous updates, gesture set expansion, and multilingual adaptability.
- Personalization options enabling users to select output language, voice tone, speed, and interaction preferences, ensuring accessibility across diverse linguistic and cultural settings.

### 5 Section 5: Illustrations (to be included in final thesis)

- **Software layer :**
  - **Splash screen :**

This is the app's welcome screen that appears for a few seconds when the application launches. It features the "Hope Glasses" logo centered on the screen, often with a smooth animation or loading indicator. Its purpose is to introduce the user to the brand and ensure a clean transition into the app. It sets a professional tone and gives a brief moment to initialize background processes before interaction begins.



Figure 33 : Splash screen

## Chapter4: Commercial Overview

---

- **Choose method screen :**

This screen allows the user to select how they intend to use the application. It presents two main buttons:

- **"Glasses User"**: For deaf or mute individuals wearing the smart glasses.
- **"App User"**: For hearing users who want to understand sign language via the app.

Each button includes an icon or short explanation to make the choice clear. The goal is to offer a personalized experience depending on the user's role in the communication process.

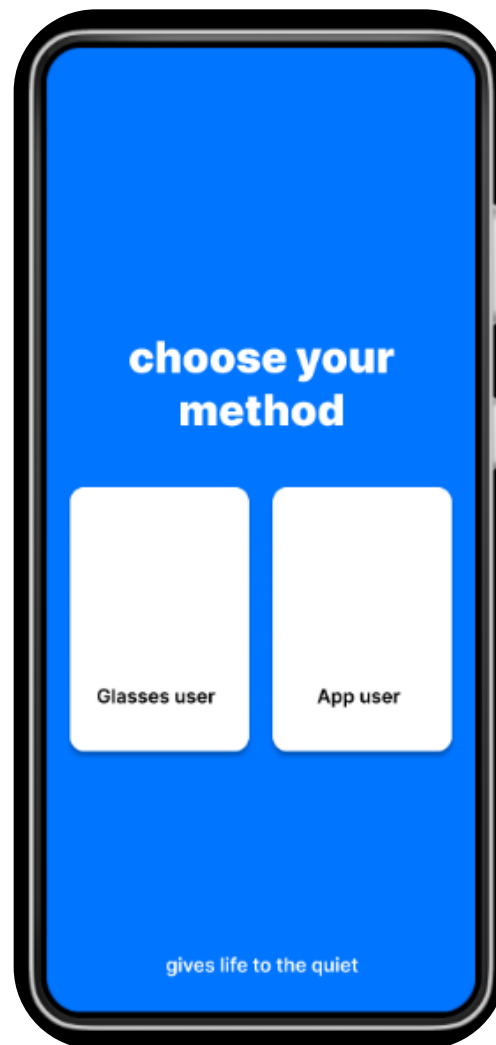


Figure 34 : Choose method screen

## Chapter4: Commercial Overview

---

- **Glasses user method :**

This screen is designed for users who are wearing the smart Hope Glasses. The camera on the glasses captures hand signs in real time and sends them to the app. The screen shows:

- A live feed
- The predicted word or letter from the sign language
- Automatic voice output that says the translated word aloud in the selected language

This empowers deaf users to communicate verbally with others in real time, using only their hand gestures and glasses.

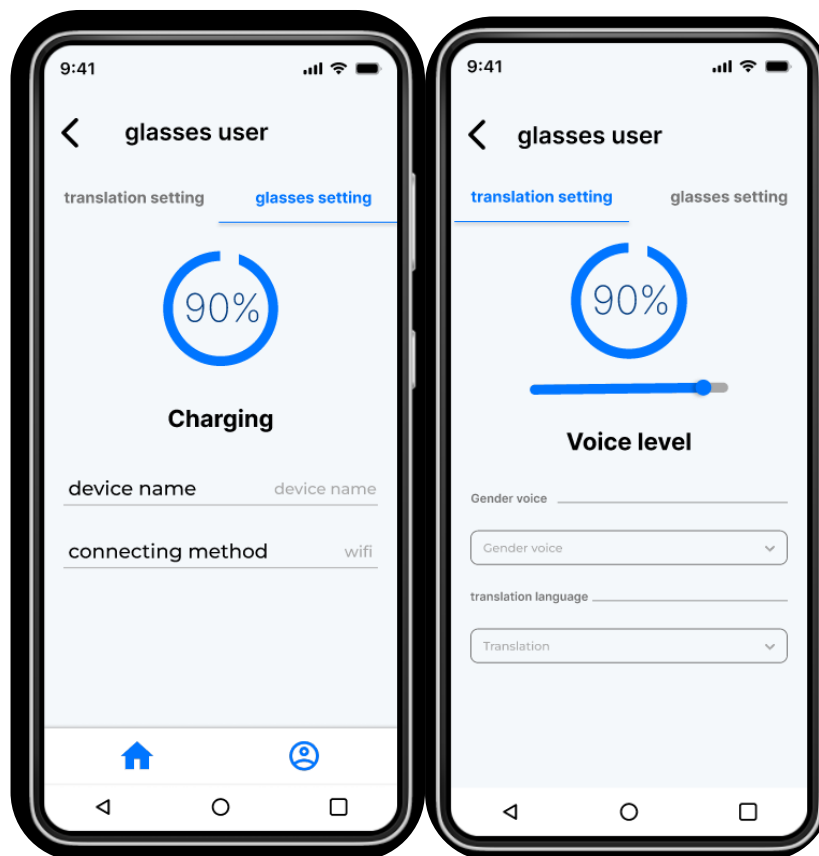


Figure 35 : Glasses user method screen

## Chapter4: Commercial Overview

---

- **App user method :**

This screen is intended for hearing individuals who do not wear the glasses but want to understand what a deaf person is signing. The phone's rear camera is used to capture the hand signs of the other person. The screen displays:

- The live video feed
- The real time prediction of the signed word
- An optional voice playback of the word or sentence

It helps break communication barriers by turning the app into a smart visual interpreter for sign language, enhancing inclusion and understanding.

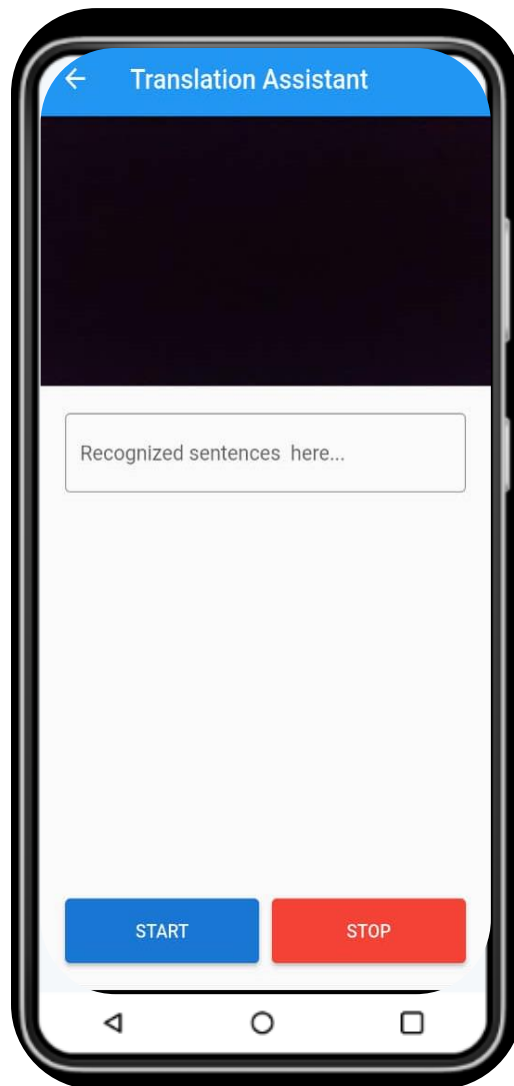


Figure 36 : App user method screen

## Chapter4: Commercial Overview

- **Hardware layer :**

The hardware layer of the Hope smart glasses consists of compact, lightweight components integrated into the frame to ensure comfort, mobility, and real-time functionality. It includes a front-facing camera for capturing hand gestures, a microcontroller (ESP32-CAM) for processing, and a speaker for audio feedback all designed to operate seamlessly and efficiently in everyday environments.

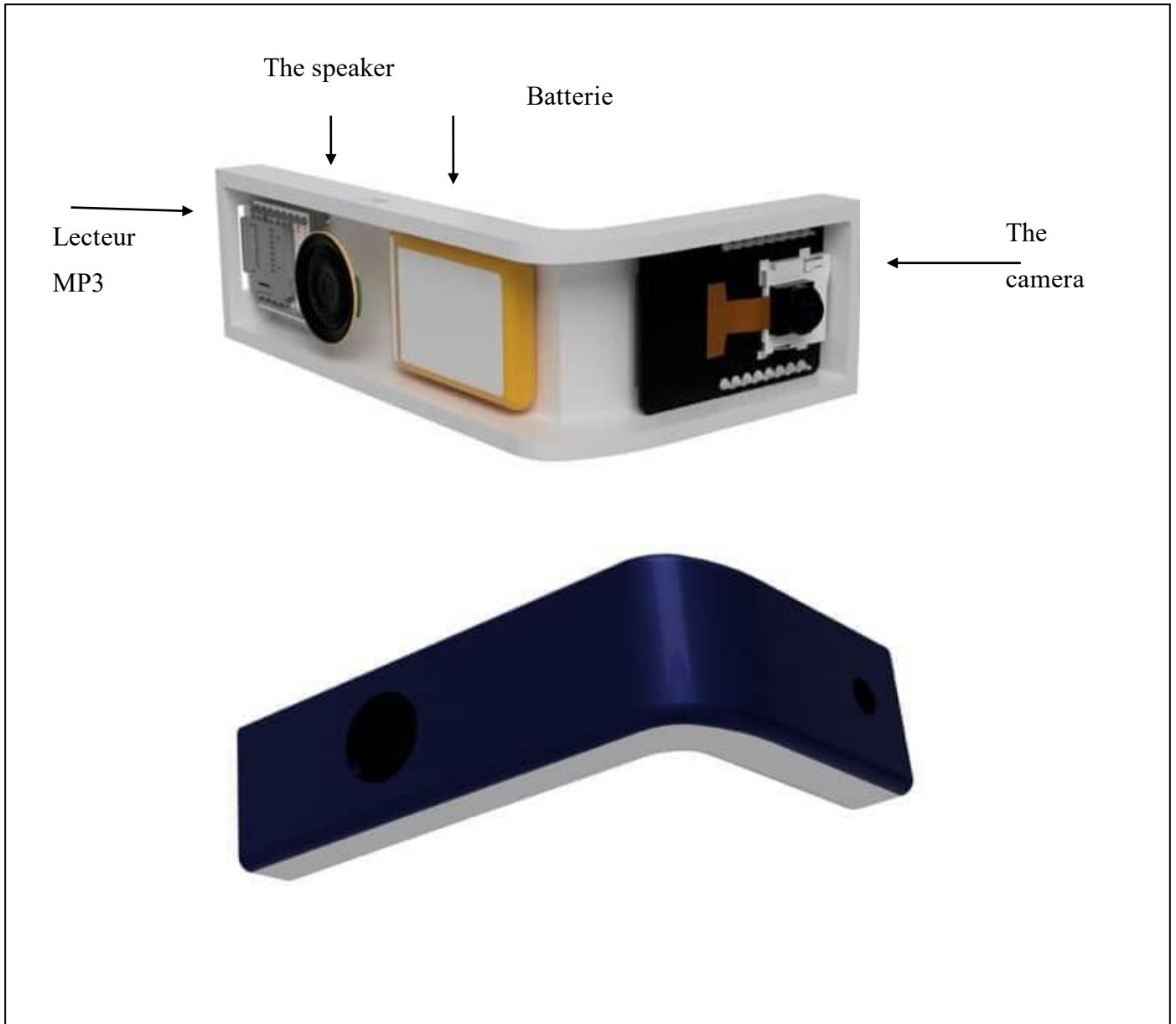


Figure 37 : The glasses design



# Chapter4: Commercial Overview

<b>Business Model Canvas</b>				
<p><b>Key Partners</b></p> <ul style="list-style-type: none"> <li>Deaf and mute individuals</li> <li>Schools and educational institutions for special needs</li> <li>Healthcare centers and hospitals</li> <li>Government entities and non profit organizations</li> </ul>	<p><b>Key Activities</b></p> <ul style="list-style-type: none"> <li>Developing AI models for sign language recognition</li> <li>Integrating the model into smart glasses and mobile apps</li> <li>Hardware design and assembly</li> <li>Data management and training using Firebase ML</li> </ul>	<p><b>Value Propositions</b></p> <ul style="list-style-type: none"> <li>Enables hearing impaired individuals to communicate instantly with others</li> <li>Converts sign language into audible speech in multiple languages</li> <li>User friendly design: portable and lightweight</li> <li>Provides real time translation support</li> </ul>	<p><b>Customer Relationships</b></p> <ul style="list-style-type: none"> <li>Continuous technical support</li> <li>Regular system updates via the internet</li> </ul>	<p><b>Customer Segments</b></p> <p><b>B2C:</b></p> <ul style="list-style-type: none"> <li>Deaf and mute individuals</li> </ul> <p><b>B2B:</b></p> <ul style="list-style-type: none"> <li>Specialized educational institutions</li> <li>Hospitals and rehabilitation centers</li> <li>Government agencies and nonprofit organizations</li> <li>Sign language learning and training centers</li> </ul>
<p><b>Key Resources</b></p> <ul style="list-style-type: none"> <li>Software development team (AI, Flutter, Firebase)</li> <li>Hardware engineering team (smart glasses)</li> <li>Training datasets for AI models</li> <li>Manufacturing and assembly equipment</li> </ul>		<p><b>Channels</b></p> <ul style="list-style-type: none"> <li>Official e commerce website</li> <li>Online sales platforms (Amazon, Jumia, etc.)</li> <li>Medical and optical device distributors</li> <li>Product demos at exhibitions and tech fairs</li> <li>Awareness campaigns and NGO partnerships</li> </ul>		
<p><b>Cost Structure</b></p> <ul style="list-style-type: none"> <li>Smartglass manufacturing (lenses, cameras, batteries)</li> <li>Software/application development</li> <li>Cloud storage and data processing</li> <li>Marketing and distribution</li> </ul>		<p><b>Revenue Streams</b></p> <ul style="list-style-type: none"> <li>Sales of smart glasses</li> <li>Subscriptions for premium features in the application and cloud updates</li> <li>Supply contracts with institutions and governments</li> </ul>		

Table 3 : Business Model Canvas

## Chapter4: Commercial Overview

---

**Annual Cost Table – Hope Project**

Year	Activities	Unit Cost (DZD)	Number of Glasses	App Development Cost (DZD)	Total Cost (DZD)
Year 1	Production + App Development + Maintenance	30,000	100	10,000	3,010,000
Year 2	Production + App Development + Maintenance	30,000	300	10,000	9,010,000
Year 3	Production + App Development + Maintenance	30,000	500	10,000	15,010,000
Year 4	Production + App Development + Maintenance	30,000	700	10,000	21,010,000

Table 4: Annual Cost Table – Hope Project

## Chapter4: Commercial Overview

---

### Revenue Projection Table – Hope Project :

Year	Unit Selling Price (DZD)	Number of Units Sold	Total Revenue (DZD)
Year 1	30,000	100	3,000,000
Year 2	30,000	300	9,000,000
Year 3	30,000	500	15,000,000
Year 4	30,000	700	21,000,000

Table 5 : Revenue Projection Table – Hope Project

# **General conclusion**

## **General Conclusion:**

This project sought to merge technological innovation with social awareness by addressing a real and persistent problem faced by millions around the world: the lack of immediate and accessible communication between the deaf and hearing individuals.

Throughout this thesis, we demonstrated how artificial intelligence and deep learning techniques can be harnessed to create a smart device capable of translating sign language into speech quickly and efficiently, without the need for complex hardware or prior training by the communication partner. The Hope system was designed to be simple, portable, and practical focusing primarily on empowering the sign language user, rather than placing the burden on the hearing person.

The final chapter also explored the innovative and commercial potential of the project, highlighting its scalability and applicability in various sectors such as education, healthcare, and public services.

We hope this project can serve as a stepping stone toward the development of more inclusive smart technologies tools that break isolation, empower individuals, and promote communication as a universal human right. Our greatest aspiration is to see Hope Glasses one day used in real world settings, making a tangible difference in people's lives.

## References

- [1]. Formeister, Eric J. *Ears*. S.l. : duke University School of Medicine, 2022.
- [2]. <https://www.mayoclinic.org/diseases-conditions/hearing-loss/symptoms-causes>
- [3]. World Health Organization. *Deafness and Hearing Loss - Fact Sheet*. Available at: <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>.
- [4]. S. Srivastava, A. Gangwar, R. Mishra, and S. Singh. *Sign language recognition*. 2022.
- [5]. Nendauni, L. R. The development of sign language: A synopsis overview,. 2021.
- [6]. World Federation of the Deaf. “The legal recognition of national sign languages.”.
- [7]. R. Pfau, M. Steinbach, and B. Woll. ,Sign Language: An International Handbook.
- [8]. <https://images.app.goo.gl/v2uclxyz76jxjgy18>.
- [9]. <https://images.app.goo.gl/djghwcuxw5yg2nkq7>.
- [10]. The Index Project. *Award Nominee: Sign Language Device*. Available at: <https://theindexproject.org/award/nominees/2051>
- [11]. <https://www.brightsignglove.com/>.
- [12]. <https://vasco-translator.com/translators/>.
- [13]. The Index Project. *Award Nominee: Smart Translation Tool*. Available at: <https://theindexproject.org/award/nominees/2692>Petapixel.
- [14]. Starner, T., Weaver, J., & Pentland, A. (1998). Real-time American Sign Language Recognition using Desk and Wearable Computer Based Video.
- [15]. Koll, M. Et al. (2016). Deep Learning for Sign Language Recognition: A Comparative Study.
- [16]. <https://www.geeksforgeeks.org/machine-learning-algorithms/>

- [17]. Lee, C., & Tan, Y. (2018). Sequential Modeling of Hand Gestures using LSTM Networks for Sign Language Translation.
- [18]. <https://www.tpointtech.com/deep-learning-algorithms>
- [19]. Journal of Artificial Intelligence Research, 2021.
- [20]. Journal of Pattern Recognition, 2020.
- [21]. M. A. H. S. -. P. Info, «ESP32-CAM Serveur Web de diffusion vidéo avec ESP32-CAM,» Robotique, [En ligne]. Available:  
<https://www.robotique.tech/tutoriel/serveur-web-dediffusion-video-avec-esp32-cam/>
- [22]. <https://bility.fr/definition-visual-studio-code/>
- [23]. Toolify, «Le Flutter Tutorial pour Text To Speech (TTS),» Toolify.ai, 2024. [En ligne]. Available: <https://www.toolify.ai/fr/ai-new-fr/le-flutter-tutorial-pour-text-to-speechhttps>
- [24]. ChanakaDev. *Firestore Cloud Functions Guide*. Available at:  
<https://medium.com/@ChanakaDev/firebase-cloud-functions-b90008ef088b>
- [25]. W3Schools. *Introduction to Python*. Available at:  
[https://www.w3schools.com/python/python\\_intro.asp](https://www.w3schools.com/python/python_intro.asp)