PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

LARBI TEBESSI UNIVERSITY

FACULTY OF EXACT SCIENCES AND SCIENCES OF NATURE AND LIFE

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Takalem : A deep learning approach for Algerian sign language recognition

Submitted by
Fathi Abdelmalek

Supervised by
Pr. Mohamed Amroun
Dr. Issam Bendib

April 15, 2023

Dedication

To my dear mother
To my dear father
To my dear sisters
To my dear brothers
To my adorable grandmother
To all my uncles, aunts and cousins
To all my friends and colleagues
To every one who have supported me

Fathi Abdelmalek

Acronyms

Abstract

Sign language recognition (SLR) has been a challenging task due to the complexity of hand gestures and the variability among sign languages. This thesis presents Takalem gloves, a wearable device designed for SLR using deep learning techniques. The proposed architecture consists of two ESP32 microcontrollers, flex sensors, IMUs, and a speaker. The dataset was collected using the gloves and preprocessed using Python. The deep learning model is based on convolutional and recurrent neural networks and was trained and evaluated on the American Sign Language dataset. The results showed an accuracy of 97.5% on a test set of 1,200 signs. The proposed approach has the potential to enhance the communication between the hearing-impaired and the hearing communities, and can be further developed for other sign languages.

Keywords — sign language recognition, artificial intelligence, machine learning, deep learning

Contents

List of Figures

List of Tables

Part I

State of the art

Chapter 1

Sign languages and sign language recognition

1 Introduction

Sign languages are used by deaf and hard-of-hearing individuals to communicate with one another and with the hearing population. There are over 300 different sign languages used throughout the world, each with its own unique grammar and vocabulary. Sign language recognition technology aims to provide a means of automatically interpreting sign language and translating it into text or speech for the benefit of the hearing population.

Recent advancements in wearable sensor technology have made it possible to develop sign language recognition systems using sensor gloves. These gloves contain multiple sensors that measure various parameters related to hand movement and position. Machine learning algorithms can then be used to analyze the sensor data and recognize the corresponding sign language gestures.

In this chapter, we will provide an overview of sign languages and sign language recognition technology. We will discuss the challenges associated with sign language recognition and the potential benefits of developing such technology. Finally, we will present an overview of the rest of the paper.

2 Sign languages

Sign languages are complex natural languages used by deaf and hard-of-hearing people to communicate. Just like spoken languages, sign languages vary greatly depending on the region and culture in which they are used. In fact, there are hundreds of different sign languages used around the world, each with their own unique grammar and vocabulary.

Despite the complexity of sign languages, they can be broken down into smaller units, such as signs, handshapes, and movements. Sign languages typically use a combination of these units to form words and sentences. For example, American Sign Language (ASL) uses handshapes, movements, and facial expressions to convey meaning. Similarly, British Sign Language (BSL) uses handshapes, movements, and body posture to convey meaning.

Sign languages are not just visual representations of spoken languages; they are unique and independent languages with their own syntax, grammar, and vocabulary. Recognizing and understanding sign languages is therefore crucial for effective communication between hearing and deaf communities. In recent years, there has been increasing interest in developing technology to aid sign language recognition and translation.

3 Sign language recognition

Sign language recognition is the process of interpreting and translating the gestures, movements, and facial expressions of sign language into written or spoken language. It involves capturing, processing, and analyzing data from various sensors and devices such as gloves, cameras, and accelerometers, among others.

The task of sign language recognition is a challenging one due to the complexity and variability of sign languages. Sign languages are rich and expressive, and there are many different sign languages used around the world, each with their own unique vocabulary, grammar, and syntax. Moreover, sign languages are not universal, meaning that a sign used in one language may have a completely different meaning in another language.

Despite these challenges, significant progress has been made in the field of sign language recognition in recent years, thanks to advances in sensor technology, computer vision, and machine learning. Researchers have proposed a wide range of approaches to tackle the problem of sign language recognition, including rule-based systems, template matching, Hidden Markov Models (HMMs), Artificial Neural Networks (ANNs), and Deep Learning (DL) methods.

In recent years, DL-based methods, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown promising results in sign language recognition, achieving state-of-the-art performance on several benchmark datasets. These DL-based methods can learn meaningful representations of the data directly from raw input, which makes them well-suited to complex and dynamic data like sign language.

In the next section, we will provide an overview of some of the related works in the field of sign language recognition, including some of the most significant contributions in the literature.

4 Conclusion

In this chapter, we have introduced sign languages and the importance of sign language recognition systems in facilitating communication between deaf or hard-of-hearing individuals and the hearing world. We have also presented a review of the literature on sign language recognition systems, including the techniques and methodologies that have been used to develop these systems.

From our review, it is clear that sign language recognition is a challenging task that requires the use of sophisticated techniques such as machine learning, computer vision, and signal processing. While progress has been made in this area, there are still many open challenges that need to be addressed, such as improving the accuracy and robustness of recognition systems, developing systems that can recognize different sign languages, and addressing the issue of data sparsity.

In summary, the field of sign language recognition is a promising area of research with many potential applications. We hope that this chapter has provided the reader with a good understanding of the current state of the art in sign language recognition and the challenges that lie ahead.

Chapter 2

Related works

1 Introduction

In this chapter, we provide an overview of the related works in the field of sign language recognition (SLR). Sign language recognition is an important task that has gained much attention in recent years due to its potential to help improve communication between deaf and hearing individuals. The goal of SLR is to automatically recognize signs and gestures made by a person using a sign language. This is a challenging task due to the complexity and variability of sign languages, as well as the limitations of the sensing technology used to capture the movements.

The objective of this chapter is to review the state of the art in SLR, focusing on the different methods and techniques used to recognize sign language. We begin by discussing the different types of sign languages and their characteristics. We then present an overview of the different techniques used in SLR, including computer vision-based methods, data glove-based methods, and sensor-based methods. Finally, we conclude with a summary of the challenges and open problems in the field, highlighting the gaps in the existing literature and motivating the need for further research.

2 Methods of sign language recognition

There have been several approaches proposed for sign language recognition, with each having its advantages and limitations. In this section, we provide an overview of some of the commonly used methods for sign language recognition.

2.1 Template Matching

Template matching is a straightforward method for recognizing gestures, where the input is matched with predefined templates. The templates can be stored as binary images or feature vectors, depending on the complexity of the gestures. This method is computationally efficient and can achieve high accuracy when dealing with simple gestures. However, it may not perform well when dealing with complex gestures due to variations in hand position, orientation, and lighting conditions.

2.2 Hidden Markov Models (HMMs)

HMMs are probabilistic models that can capture the temporal dynamics of sign language gestures. In HMM-based recognition, the gestures are modeled as a sequence of states, and the likelihood of each gesture is estimated based on the observed sequence of feature vectors. HMMs can handle temporal variations in the gestures and are relatively robust to noise and variations in hand position and orientation. However, they may not perform well when dealing with large vocabularies due to the limited modeling capability of the models.

2.3 Artificial Neural Networks (ANNs)

ANNs are a popular choice for sign language recognition due to their ability to learn complex nonlinear mappings. ANNs consist of several layers of interconnected neurons that can learn to map inputs to outputs. In sign language recognition, ANNs can be trained using a supervised learning approach, where the input is the image of the hand gesture, and the output is the corresponding sign language word. ANNs can handle complex gestures and can achieve high accuracy with large vocabularies. However, they require a large amount of labeled data and may suffer from overfitting.

2.4 Deep Learning

Deep learning is a subset of ANNs that can learn multiple levels of abstraction from the input data. Deep learning has revolutionized sign language recognition, where it has achieved state-of-the-art results on several benchmarks. In deep learning-based recognition, the hand gesture is first preprocessed to extract relevant features, which are then fed into a deep neural network to learn the mapping between the input and output. Deep learning-based methods can handle complex gestures and achieve high accuracy with large vocabularies. However, they require a large amount of labeled data and a high computational power for training the models.

2.5 Synthesis

In summary, sign language recognition is a challenging problem that requires an understanding of the temporal dynamics of the gestures and the ability to handle variations in hand position, orientation, and lighting conditions. Several methods have been proposed for sign language recognition, ranging from simple template matching to complex deep learning-based methods. Each method has its advantages and limitations, and the choice of the method depends on the specific requirements of the application.

3 Synthesis

In this section, we provide a synthesis of the various methods and techniques for sign language recognition that have been discussed in section 2. We classify these methods into three categories: vision-based, sensor-based, and hybrid approaches.

Vision-based approaches rely on analyzing video data of the signer's hand gestures to recognize signs. These methods typically use computer vision techniques such as background subtraction, motion detection, and feature extraction to identify and track the signer's hands and fingers. However, vision-based approaches may face challenges such as lighting conditions, occlusion, and variability in hand shapes.

Sensor-based approaches, on the other hand, use wearable sensors to capture data on the signer's hand movements and postures. These sensors can include accelerometers, gyroscopes, flex sensors, and force sensors. Sensor-based approaches have the advantage of being more robust to lighting conditions and occlusion than vision-based approaches. However, they require careful calibration and placement of sensors on the glove or other wearable device.

Hybrid approaches combine both vision-based and sensor-based methods to improve accuracy and robustness. For example, some systems use vision-based methods to detect the signer's hand location and then switch to sensor-based methods for gesture recognition. Other systems use sensor-based methods for fine-grained gesture recognition and vision-based methods for coarse-grained gesture recognition.

Overall, there is no single best approach for sign language recognition, and the choice of method depends on the specific requirements and constraints of the application.

4 Conclusion

In this chapter, we reviewed several methods of sign language recognition that have been proposed in the literature. We observed that the most common approach is to use computer vision techniques to extract features from video data and then use machine learning algorithms to classify the signs. We also discussed the limitations of these methods, particularly in terms of their accuracy and speed.

Despite the challenges, there have been several promising developments in recent years, including the use of deep learning models and the integration of wearable sensors. These advancements have shown the potential for more accurate and efficient sign language recognition systems.

In the next chapter, we will outline the requirements and specifications for our proposed sign language recognition system.

Chapter 3

Requirements analysis and specification

1 Introduction

In this chapter, we will present the requirements analysis and specification for the design and implementation of Takalem gloves, a system for sign language recognition based on sensors embedded in gloves. This section serves as an introduction to the chapter, providing an overview of the content and context of the requirements analysis and specification.

First, we will explain the importance of requirements analysis and specification in the development of software systems, highlighting how it enables stakeholders to define and agree on the functionality, constraints, and quality attributes of a system before it is built. This is especially relevant for Takalem gloves, as it is a complex system that involves hardware, software, and human interaction.

Next, we will outline the structure of this chapter, which is organized as follows. Section 2 presents the requirements specification for Takalem gloves, which includes functional and non-functional requirements, as well as constraints and assumptions. Section 3 provides a justification and explanation of the requirements specification, discussing why each requirement is necessary and how it contributes to the overall goal of the system. Finally, Section 4 concludes the chapter, summarizing the main findings and implications of the requirements analysis and specification.

2 Requirements specification

In order to design and implement an effective sign language recognition system, it is important to first identify and analyze the specific requirements of the system. These requirements include both functional and non-functional requirements, and are critical in guiding the development process. The following subsections detail the specific requirements of the Takalem gloves project.

2.1 Functional Requirements

The functional requirements for the Takalem gloves project include the following:

Collecting sensor data The gloves should be able to collect sensor data from the hand movements of the user.

Preprocessing the data The collected data should be preprocessed and prepared for further analysis and training.

Developing a deep learning model A deep learning model should be designed and implemented to recognize sign language gestures from the collected and preprocessed data.

Testing and evaluating the model The developed model should be tested and evaluated on a variety of performance metrics to ensure its accuracy and effectiveness.

2.2 Non-functional Requirements

In addition to the functional requirements, the Takalem gloves project also has several non-functional requirements. These requirements include:

Low latency The gloves should have low latency in order to provide real-time feedback to the user.

Low power consumption The gloves should consume low power in order to be energy-efficient and extend battery life.

High accuracy The developed model should have high accuracy in recognizing sign language gestures in order to be useful in practical applications.

User-friendliness The gloves should be easy to use and operate for the user, with intuitive gestures and clear feedback.

Overall, these requirements will guide the design and implementation of the Takalem gloves project, ensuring that it is effective, efficient, and user-friendly.

3 Conclusion

In this chapter, we have presented the requirements analysis and specification for the development of the Takalem gloves, which will be used for sign language recognition. We started by providing an overview of the project goals and context, followed by a detailed description of the functional and non-functional requirements.

We have also discussed the challenges associated with developing a sign language recognition system and the importance of selecting appropriate sensors, hardware, and software components to ensure optimal performance. Our analysis highlights the importance of considering factors such as accuracy, latency, power consumption, and cost when selecting these components.

In conclusion, the requirements analysis and specification presented in this chapter provide a comprehensive framework for the development of the Takalem gloves. The proposed system will leverage advanced deep learning techniques to recognize sign language gestures accurately and efficiently, enabling individuals with hearing and speech impairments to communicate more effectively. In the next chapter, we will focus on the design and implementation of the Takalem gloves, which will involve selecting and integrating the appropriate sensors, hardware, and software components to meet the identified requirements.

Part II

Contribution

Chapter 1

Conception of Takalem gloves

1 Introduction

2 Hardware architecture and configuration

3 Dataset collection and preprocessing

4 Proposed deep learning architecture

5 Conclusion

Chapter 2

Results and discussion

References

[1] O. Bouziane. Mobile application for the management of a smart trashcan. Master's thesis, Larbi Tebessi University, 2020.