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

May 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

AJSL Algerian jewish sign language

ANN Artificial Neural Network

AR Augmented Reality

ASP Algerian sign language

CNN Convolutional Neural Network

DL Deep Learning

FSL French sign language

HMM Hidden Markov Model

LSTM Long-Short Term Memory

ML Machine Learning

RNN Recurrent Neural Network

SL Sign language

SLR Sign language recognition

SVM Support Vector Machines

VR Virtual Reality

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nullam id dolor ac turpis elementum tincidunt. Suspendisse maximus, justo eu lobortis tincidunt, ex metus consectetur est, ac volutpat urna velit non dui. Aenean dapibus lobortis lorem, eu sagittis sem scelerisque a. Proin a risus vitae mauris auctor ultricies vitae sed odio. Proin ut ipsum neque. Donec pretium felis eu lacus tempor lacinia. Sed condimentum auctor luctus. Vestibulum eget posuere mauris. Pellentesque eu erat ut dolor pharetra finibus a vitae libero. Duis in massa lacus. Sed pulvinar ligula vel viverra aliquam. Sed posuere ipsum eget ligula vulputate commodo. Duis in iaculis ex, in commodo arcu. Suspendisse eget mi lorem. Nullam eu ante semper, ullamcorper nibh sit amet, fringilla purus.

***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

Deaf, mute, and hard-of-hearing people, have their language system, its Sign language (SL), there are over 300 different SL used in the world, and each one has a unique grammar and vocabulary, one of them is Algerian sign language (ASP), is used by more than 240,000 individuals around Algeria. Sign language recognition (SLR) technology aims to provide a means of automatically interpreting SL and translating it into text or speech, to facilitate the communication process between normal and deaf/mute people.

Recent advancements in wearable sensor technology have made it possible to develop SLR systems using sensor gloves. These gloves contain multiple sensors that measure various parameters related to hand movement and position. Machine Learning (ML) and Deep Learning (DL) algorithms can then be used to analyze the sensor data and recognize the corresponding SL gestures.

In this chapter, we will provide an overview of sign languages and SLR technology. We will discuss the challenges associated with SLR and the potential benefits of developing such technology.

2 Sign languages

SL is one of the methods of communication, which is defined as a set of visual symbols or gestures that are used in a very systematic way for words, concepts, or ideas of a language.[1]

Despite the complexity of sign languages, they can be broken down into smaller units, such as signs, hand shapes, and movements. Sign languages typically use a combination of these units to form words and sentences. For example, ASP uses hand shapes, movements, and facial expressions to convey meaning.

Sign languages are not just visual representations of spoken languages, they are unique and independent languages with their syntax, grammar, and vocabulary. Recognizing and understanding them 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.

2.1 Sign languages in Algeria

There is no official ASP document or reference except for a dictionary published recently by the Algerian government. It contains some signs used by the deaf and other signs borrowed from the Old French sign language (FSL).[2]

There is also Algerian jewish sign language (AJSL), an old SL which developed in several Jewish communities in the region of M'zab, Algeria, which is located in the northern part of the Sahara desert.[3]

2.2 Algerian Sign Language

ASP is an SL derived from FSL, used by the deaf/mute community of Algeria. It was officially recognized by the Algerian law as official SL in Algeria in May 2002[4]. Technically, it is a visual-gestural language that uses hand shapes, movements, and facial expressions to convey meaning. Therefore, this community is often excluded from basic communication, this has caused many deaf Algerians to go without access to education, employment opportunities, and other basic rights. For that, The government of Algeria opened many deaf schools around the country to teach them the language itself, and basic education like any normal person.

Even with the existence of deaf schools, the teachers themselves are not qualified neither master ASP, all of them are hearing individuals who hold different degrees which are not related to deaf education or SL, and they have never been trained to use the language before they get hired, some teachers attend training courses to learn alphabet only. And this makes it difficult for those pupils to get basic education and go even to middle or high school and college, most of them are marginalized and can only be manual workers, they are denied access to a high-quality education that meets their special needs to improve their lives and live as equal to their peers.[2]

3 Sign language recognition

3.1 Definition

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

The task of SLR is a challenging one due to the complexity and variability of sign languages. They 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 SLR in recent years, thanks to advances in sensor technology, computer vision, ML and DL. Researchers have proposed a wide range of approaches to tackle the problem of SLR, including rule-based systems, template matching, Hidden Markov Model (HMM), Artificial Neural Network (ANN), and DL methods.

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

3.2 SLR applications and use

SLR technology has the potential to empower deaf and hard-of-hearing individuals by enabling them to communicate more effectively with the wider community. Moreover, it can be used in other areas alongside basic use, here are the main areas of use of SLR technology.

3.2.1 SL translation

SL translation is one of the primary applications of SLR technology. By capturing and analyzing SL gestures, SLR systems can convert them into written or spoken language, enabling effective communication between deaf and hearing individuals. This technology can be utilized in various contexts, such as educational settings, customer service centers, healthcare facilities, and public institutions. For example, in education, SLR translation can support deaf students by providing real-time SL interpretation during lectures, ensuring they have access to the same educational content as their hearing peers. In customer service, SLR translation can facilitate communication between deaf customers and service representatives, improving accessibility and customer satisfaction. Additionally, SLR translation can be integrated into translation apps or devices, allowing deaf individuals to communicate with individuals who do not understand SL, bridging the communication gap.

3.2.2 Virtual and augmented reality

SLR technology can enhance the immersive experience of virtual and augmented reality environments by incorporating SLR capabilities. In VR/AR applications, SLR enables users to interact with the virtual world using SL gestures, making the experience more intuitive and inclusive for deaf users. For instance, in a Virtual Reality (VR) game, SLR can recognize and interpret

SL commands as input, allowing players to control their characters or perform actions using SL gestures. In Augmented Reality (AR), SLR can be used to overlay real-time SL translations onto the user's field of view, enabling seamless communication between deaf and hearing individuals in AR scenarios. This integration of SLR in VR/AR not only enhances entertainment experiences but also opens up new possibilities in training simulations, remote collaboration, and interactive storytelling.

3.2.3 Research and linguistics

SLR technology plays a crucial role in linguistic research and the study of sign languages. By capturing and analyzing SL data, SLR systems provide valuable insights into the structure, grammar, and syntax of sign languages. Researchers can use SLR to examine the linguistic patterns and variations within SL communities, contributing to the documentation and preservation of sign languages. SLR can also assist in studying the cognitive aspects of SL processing and acquisition. Furthermore, SLR can be used to create SL corpora and databases, which serve as valuable resources for linguistic analysis and comparison across different SL systems. Overall, SLR technology empowers researchers and linguists to delve deeper into the intricate nature of sign languages, leading to a better understanding and appreciation of deaf culture and communication.

4 Deep Learning

DL, a subfield of ML, has revolutionized various domains by enabling the automatic learning of intricate patterns and representations from raw data. In SLR, DL proves to be a promising approach due to its ability to process and analyze complex visual and temporal information.

ANN serve as the foundation of DL models. These computational structures are inspired by the biological neurons in the human brain and consist of interconnected layers of artificial neurons. CNN are widely employed in image recognition tasks, including the analysis of hand gestures in sign language. CNNs leverage convolutional layers to automatically learn spatial features from input images, enabling the model to discern intricate patterns and variations.

In addition to CNNs, RNN play a vital role in sign language recognition. RNNs are particularly effective in handling sequential data, which is inherent in SL gestures. RNNs can capture the temporal dependencies in a sequence of gestures and retain information over time. Long-Short Term Memory (LSTM) networks, a type of RNN, excel in modeling long-range dependencies and have demonstrated promising results in SLR tasks.

Training and optimization are essential steps in the deep learning workflow. During the training phase, a neural network is exposed to labeled training data, and its internal parameters are adjusted to minimize the discrepancy between predicted and actual outputs. Backpropagation, a gradient-based optimization algorithm, is commonly used to update the model's parameters. The training process is iterative, and the model undergoes multiple epochs to improve its performance gradually.

Transfer learning, a technique widely employed in DL, has shown promise in SLR. It involves leveraging pre-trained models trained on large-scale datasets and fine-tuning them on smaller SL datasets. Transfer learning allows the model to benefit from the learned representations in the pre-trained models, leading to improved performance and reducing the need for extensive training data.

Despite the remarkable potential of DL in SLR, it comes with its own set of challenges and limitations. One major challenge is the scarcity of annotated SL datasets, which hinders the training process and necessitates domain-specific data collection efforts. Overfitting, a phenomenon where the model performs well on training data but fails to generalize to unseen data, is another challenge that requires careful regularization techniques.

Moreover, DL models often demand significant computational resources, including high-performance computing units and memory capacity. The interpretability of DL models is another concern, as they are often seen as "black boxes" that make it challenging to understand the internal decision-making process.

In conclusion, DL techniques offer immense potential for SLR, enabling the automatic extraction of meaningful representations from visual and temporal data. This section has provided a comprehensive overview of the key concepts, methodologies, and challenges associated with DL in the context of SLR. By leveraging DL, we can unlock new possibilities for accurate and real-time interpretation of SL gestures.

5 Conclusion

In this chapter, we have introduced sign languages and the importance of SLR 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 SLR systems, including the techniques and methodologies that have been used to develop these systems.

From our review, it is clear that SLR is a challenging task that requires the use of sophisticated techniques such as ML, 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 SLR 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 SLR and the challenges that lie ahead.

Chapter 2

Related works

1 Introduction

In this section, we provide an introduction to the topic of SLR and its importance in facilitating communication for the deaf and hard of hearing community. We present an overview of the goals and objectives of this study, highlighting the significance of developing accurate and efficient SLR systems. Additionally, we outline the specific objectives of this research and provide an overview of the subsequent sections in this chapter.

1.1 Background

In this subsection, we provide background information on sign languages, their unique characteristics, and their importance as a means of communication for individuals with hearing impairments. We discuss the complexity of sign languages and the challenges associated with their recognition, emphasizing the need for advanced technologies to aid in real-time sign language interpretation.

1.2 Motivation

In this subsection, we discuss the motivation behind the development of sign language recognition systems. We highlight the limitations of traditional communication methods for individuals with hearing impairments and the potential impact that accurate and efficient sign language recognition can have on their daily lives. We also discuss the growing interest in using machine learning and artificial intelligence techniques to enhance sign language recognition capabilities.

1.3 Objectives

In this subsection, we outline the specific objectives of this research project. We identify the main goals and targets that we aim to achieve, such as designing a wearable sign language recognition system, developing robust machine learning models, and evaluating the system's performance through experiments and user studies. We also highlight the potential applications and benefits of the proposed system in real-world scenarios.

1.4 Overview of the Chapter

In this subsection, we provide an overview of the subsequent sections in this chapter. We briefly describe the content and organization of each section, outlining the main topics and discussions that will be covered. This serves as a roadmap for the reader, giving them a clear understanding of what to expect in the rest of the chapter.

By following this structure, you can introduce the context, motivation, objectives, and the overall organization of Chapter 2 in your thesis. Feel free to customize the content and add more details based on the specific focus and requirements of your research.

2 Methods of sign language recognition

In this section, we explore various methods and techniques that have been employed in the field of SLR. We provide an overview of different approaches, highlighting their strengths, weaknesses, and suitability for real-time recognition systems. The objective is to review the existing literature and gain insights into the state-of-the-art methods used for SLR.

2.1 Computer vision based approaches

2.2 Sensors based approaches

2.3 Summary

In summary, we highlighted the strengths and weaknesses of different SLR methods, setting the stage for the subsequent section where we synthesize the literature and identify research gaps.

3 Synthesis

SLR is a complex task that requires a combination of computer vision techniques, ML algorithms, and domain-specific knowledge. Researchers have explored different methodologies and strategies to tackle this challenge. The synthesis of related works allows us to identify common trends, key techniques, and advancements in the field.

Several studies have focused on the utilization of computer vision techniques for hand gesture detection and tracking. Techniques such as background subtraction, hand region segmentation, and hand shape analysis have been employed to extract relevant visual features from input videos or images. These features serve as the basis for subsequent classification or recognition processes.

Furthermore, ML algorithms play a pivotal role in SLR. Various approaches, including traditional ML methods and DL techniques, have been explored. Traditional ML algorithms such as Support Vector Machines (SVM), decision trees, and random forests have shown promising results in certain scenarios. These algorithms require handcrafted features and extensive feature engineering, which can be time-consuming and resource-intensive.

DL models, on the other hand, have demonstrated remarkable performance in SLR tasks. CNNs and RNNs have been widely adopted. CNNs excel in capturing spatial features and have proven effective for hand gesture recognition. RNNs, particularly LSTM networks, have been successful in capturing the temporal dynamics of SL gestures.

From the analysis of related works, it is evident that DL models, particularly CNNs and RNNs, have emerged as powerful tools for SLR. CNNs excel in capturing spatial features from visual inputs, while RNNs effectively model the temporal dynamics of SL gestures. The integration of these models has shown promising results, leading to improved accuracy and real-time performance.

To address the scarcity of labeled SL datasets, researchers have explored various strategies, such as data augmentation, transfer learning, and semi-supervised learning. Data augmentation techniques generate additional training samples by applying transformations, deformations, or adding noise to existing data. Transfer learning allows leveraging pre-trained models on large-scale datasets and fine-tuning them on SL data. Semi-supervised learning combines a small amount of labeled data with a larger pool of unlabeled data to enhance model performance.

Furthermore, the availability of labeled SL datasets has been a crucial factor in the advancement of the field. Researchers have developed dedicated datasets for training and evaluating SLR models. However, the limited size and diversity of these datasets pose challenges in training robust models that generalize well across different sign languages and variations.

The evaluation and benchmarking of SLR methods are crucial to assess their effectiveness and compare different approaches. Researchers have utilized various evaluation metrics such as accuracy, precision, recall, and F1 score to measure the performance of their models. Datasets specifically designed for SLR, such as RWTH-BOSTON-50 and Jochen Triesch's dataset, have been widely used for evaluation purposes.

Despite the progress made, there are still several challenges and opportunities for future research. One key challenge is the robustness of SLR models in handling variations in hand shape, orientation, and articulation. Developing models that can adapt to individual user variations and different SL dialects is an ongoing area of research.

Another avenue for exploration is the integration of multimodal inputs, such as combining visual information from gloves or cameras with other sensory inputs like speech or haptic feedback. This multimodal approach has the potential to enhance the accuracy and usability of SLR systems, providing more comprehensive and natural communication interfaces.

4 Conclusion

In conclusion, the analysis and synthesis of related works provide a comprehensive overview of the state of the art in SLR. The advancements in computer vision, ML, and DL techniques have paved the way for more accurate and efficient recognition of SL gestures. However, there are still challenges to overcome, and further research is needed to address the limitations and improve the usability of SLR systems.

Part II

Contribution

Chapter 1

Conception of Takalem gloves

Chapter 2

Results and discussion

1 Introduction

2 Evaluation criteria

3 Experiments

4 Discussion

5 Conclusion

References

[1] H. Abdelouafi and M. Omari, "Teaching sign language to the deaf children in adrar, algeria."

[2] H. Abdelouafi, "Challenges of deaf education in algeria."

[3] S. Lanesman, *Algerian Jewish Sign Language: its emergence and survival.*

[4] "Algerian sign language - wikipedia."