Sparkling Light Publisher

Sparklinglight Transactions on Arificial Intelligence and Quantum Computing

journal homepage: https://sparklinglightpublisher.com/

Unveiling the Hidden Language of Hands: Automated Sign Language Recognition (ASLR) and Translation through the Lens of Deep Learning

Nishmitha M R [a], Varsha Puranik [b], Pramod Vishnu Naik [c], Khyathi [d], Deepthi [e]

[a] *Department of computer Science and Engineering, Shree Devi Institute of technology Mangalore, India*
[b] *Department of ICIS, Institute of Computer Science and Information Science, Srinivas University, Mangalore, India*
[c, d, e] *Department MCA, Shree Devi Institute of Technology, Mangalore, Karnataka, India*

Abstract

Aiming to close the divide in communication between hearing and deaf-people, Automated Sign Language Recognition (ASLR) is an exciting and significant field of study. This paper aims to build Deep Learning-based Automated Sign Language Recognition (ASLR) system that recognizes sign language gestures from video clips and translates them into written or spoken language. Using human pose estimation, key body joint points are extracted from each video frame to capture both spatial and temporal features. A deep neural network processes these sequences for accurate gesture recognition and translation. The system outperforms traditional methods like HMM and SVM in accuracy and generalization. Designed for real-time use, it promotes inclusive communication and offers a practical solution for bridging the gap between signers and non-signers. Deep Learning proves highly effective for ASLR and sign language translation, hitting top benchmark results in detecting and interpreting a broad spectrum of sign gestures. The system showcases promising real-time processing capabilities, making it suitable for real-world applications and interactive user interfaces

Keywords: Deep Learning, Automated Sign Language Recognition.

1. Introduction

Automated Sign Language Recognition (ASLR) has emerged as a vital research area with the potential to enhance communication and foster inclusivity for the deaf and hard of hearing community. Sign language lets millions of people around the world communicate through expressive gestures and - their everyday voice. Imagine how life-changing it would be if we had reliable tools to recognize those signs instantly. [1][2].

E-mail address of authors: nishmitha.mr@gmail.com, varshapuranik@gmail.com, pramodvishnunaik@gmail.com, kulalkhyathi@gmail.com, deepthiamin08@gmail.com

Please cite this article as: Nishmitha M R, et al., Unveiling the Hidden Language of Hands: Automated Sign Language Recognition (ASLR) and Translation through the Lens of Deep Learning, Sparklight Transactions on Artificial Intelligence and Quantum Computing (2024), 4(2), 18-27. **ISSN (Online):2583-0732.** Received Date: 2024/12/04, Reviewed Date: 2024/12/20, Published Date: 2024/12/31.

However, due to its unique and dynamic nature, the accurate recognition and interpretation of sign language gestures have posed significant challenges. Traditional approaches to ASLR have relied on handcrafted features and machine learning techniques, but they often struggle to capture the intricacies and temporal dependencies inherent in sign language. Deep learning breakthroughs in recent years have totally transformed computer vision, opening exciting doors to tackle the tricky challenges of sign language recognition. This research presents an innovative approach to ASLR using Deep Learning techniques. The primary objective is to design a robust and real-time ASLR system capable of accurately recognizing a wide range of sign language gestures, enabling seamless and efficient communication between the deaf and hearing communities. In daily life, hearing-impaired individuals rely on hand signals, a structured collection of hand gestures with distinct meanings, as their primary means of communication. For these individuals, automated Sign Language Recognition (ASLR) is essential, as it can greatly improve their quality of life. However, despite its significance, ASLR remains a challenging and largely unexplored research area. One of the primary obstacles in ASLR is the complexity and diversity of sign language. There are hundreds of different signs, many of which may differ only slightly in hand motion, shape, or location. Moreover, some signs involve intricate finger movements and overlapping hand positions, further complicating the recognition process. The combination of these factors, along with the individual variances in signing styles, makes ASLR extremely difficult for existing computer vision algorithms. Furthermore, the lack of large-scale gesture recognition datasets and the variability of sign language symbols among different signers based on their ethnic backgrounds add to the challenges. Without access to comprehensive and diverse data, developing an accurate and reliable Sign-Language Recognition (SLR) system becomes even more arduous. In light of these difficulties, this research tackles ASLR's toughest challenges head-on using cutting-edge deep learning methods. We tap into the strengths of models like Convolutional Neural Networks (CNNs) to detect key visual features and Recurrent Neural Networks (RNNs) to track dynamic gestures over time, we seek to enhance the accuracy and robustness of ASLR for various sign language gestures. In this research, we present a novel methodology that effectively combines the strengths of vision- based and direct measurement approaches while addressing the limitations of each. Specifically, we propose an advanced SLR system that processes video sequences to extract precise body and hand skeletal information, subsequently utilized for reliable SLR. The video sequences undergo initial skeletal data extraction, and the subsequent analysis and classification are supported by a robust deep-learning architecture, as detailed in our investigations [9][10]. Our approach encompasses an all- encompassing ASLR system that is unobtrusive, relying solely on video sequences without the necessity for sensors that could restrict signers' movements. Furthermore, the accuracy and dependability of the ASLR results are significantly enhanced by harnessing highly discriminative skeletal data within our proposed method. Our research also focuses on collecting a comprehensive dataset that encompasses a large range of signers from different ethnic backgrounds. This dataset will enable us to train the ASLR system on a diverse set of signing styles and enhance its adaptability so it performs reliably across diverse users and signing variations. Through this research, we aspire to contribute to the advancement of ASLR technology, providing a more accurate and trustworthy system that empowers hearing-impaired individuals to communicate effortlessly and efficiently with the world around them. By addressing the existing challenges and incorporating innovative solutions, we envision a future where ASLR becomes a valuable tool in promoting inclusivity and helping deaf and hard-of-hearing folks connect more easily with the world. Past efforts in automated sign language recognition (ASLR) have made real strides, exploring all sorts of clever ways to crack the code on those fluid, expressive gestures. Early research in ASLR focused on traditional machine learning techniques, such as Hidden Markov Models (HMMs) [3] and SVM's, leveraging handcrafted features and temporal modeling. However, these methods often struggled to capture the intricate nuances and variations in sign language. With the emergence of deep learning, the landscape of ASLR shifted towards leveraging CNN's model [4] was one of the first to gain major concern [5][6][7], RNN's [8], and Transformer-based architectures. DL models have demonstrated remarkable success in capturing spatial and temporal features, significantly making better accuracy of ASLR systems [1][2]. On top of that, earlier studies have experimented with blending different data types, incorporating depth data and facial expressions to enhance recognition. The creation of large-scale sign language datasets, along with benchmark datasets, has been crucial in advancing ASLR research and evaluating model performance. Moreover, several studies have tackled real-time ASLR to support interactive sign language exchanges across diverse applications. As ASLR continues to progress, researchers remain dedicated to addressing challenges related to individual signing style variations, limited data availability, and ethical considerations to create more inclusive and accurate ASLR systems that empower the deaf and hard of hearing community. The research journey commences with the establishment of a comprehensive dataset comprising diverse sign language gestures, performed by various signers in diverse environments. To effectively represent sign language gestures, this work employs a cutting-edge deep learning model for human pose estimation. It detects and extracts key-points that highlight crucial body joints and movements in each sign language video. The temporal sequences of key-points are then extracted, and relevant feature engineering is applied to capture the spatial and dynamic characteristics of sign gestures. A specialized deep learning architecture, optimized for sequential data, models the temporal dynamics in key-point sequences. Harnessing deep learning's potential in ASLR, this research advances assistive technologies to boost inclusivity and seamless communication for those with hearing impairments. This study's findings promise to reshape communication pathways, with more precise, efficient, and accessible sign language recognition that truly bridges the divide between deaf and hearing communities.

2. Background And Related Work

Background: ASLR is a vital technology designed to bridge the communication gap between deaf and hearing communities. As a visual-gestural language, it conveys meaning via intricate hand gestures, nuanced facial expressions, and fluid body movements. ASLR holds immense potential in enabling real-time and accurate interpretation of sign language, empowering deaf individuals to interact seamlessly in various domains, including education, employment, and social interactions. In recent years, significant progress has been made in developing automated methods for various linguistic tasks through the utilization of advanced algorithms capable of learning from past experiences [11]. One area that stands to benefit greatly from automation is Sign Language Recognition (SLR), as it has the potential to significantly enhance the quality of life for numerous individuals who rely on sign language as their primary means of communication on a daily basis [12]. The successful implementation of automated SLR capabilities could open up a wide range of specialized services tailored to the needs of the deaf and hard of hearing community. However, Ensuring the accuracy and reliability of these ASLR tools is paramount to avoid creating confusing or dysfunctional responses. In this section, we delve into the background of several important approaches that have been employed in automated SLR research, highlighting their significance and potential impact in advancing the field. By understanding and improving these approaches, we aim to create more meaningful and effective ASLR systems that can truly empower sign language users and promote inclusiveness in communication.Traditional ASLR approaches often relied on handcrafted-features and traditional machine learning algorithms like HMM's and Support Vector Machines(SVM's) Nevertheless, the intricacy and diversity of sign language gestures were difficult for these techniques to capture. With the advent of Deep Learning, particularly CNN's, Recurrent Neural Networks (RNNs), and Transformer-based architectures, there has been a significant shift in ASLR research.Deep learning models excel at automatically learning hierarchical features from raw data, effectively capturing the spatial hierarchies and temporal dynamics vital for precise sign language recognition.

Related Work: Numerous studies have explored ASLR using Deep Learning techniques, aiming to improve accuracy, real-time performance, and generalization across different signers and sign languages. Some key contributions in the related work include:

A. Deep Learning Architectures:

Researchers have investigated the effectiveness of various DL architectures in ASLR. CNNs have shown promise in capturing spatial patterns from video sequences, while RNNs and Transformers have proven useful for modeling temporal dependencies in sequential sign gestures.

B. Large-Scale Datasets:

Training deep learning models and evaluating ASLR systems have been made possible by the availability of large-scale sign language datasets, such as RWTH-BOSTON-50, Phoenix-14, and ASLLVD. These datasets encompass diverse signers and sign languages, facilitating model generalization.

C. Transfer Learning and Data Augmentation:

Transfer-learning techniques, where models pretrained on one sign language dataset are fine-tuned on another, have been explored to address data scarcity challenges. Data-augmentation techniques have also been employed to augment training data synthetically.

D. Multimodal Fusion:

Some studies have investigated the fusion of multiple modalities, such as combining depth information or facial expressions with video data, to enhance ASLR performance and improve adaptability to different signing styles.

E. Real-Time ASLR:

Research efforts have been dedicated to developing real-time ASLR systems, enabling interactive sign language communication in applications like virtual assistants and sign language interpretation tools.

F. Ethical Considerations:

As ASLR involves processing video data of signers, ethical considerations regarding privacy, data protection, and user consent have been discussed to ensure responsible deployment of ASLR systems. By analyzing the background and related work in ASLR using Deep-Learning, this work aims to propose an innovative methodology that combines the advantages of vision-based and direct measurement approaches while addressing the limitations of each method. The proposed system seeks to advance the accuracy, unobtrusiveness, and reliability of ASLR, contributing to more inclusive and efficient communication channels for the deaf and hard of hearing community.

3. Methodology

In this work, we propose a novel ASLR system that leverages Deep Learning techniques to recognize sign language gestures accurately and in real- time. The proposed methodology combines vision-based methods with precise skeletal data, making it a robust and efficient approach for sign-language recognition.

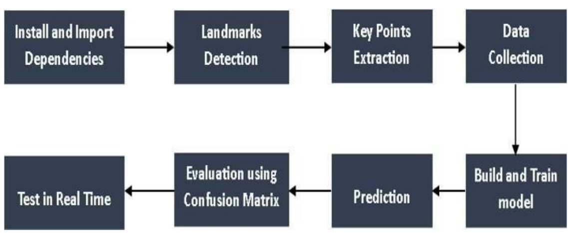

Fig. 1. SRL System-Illustarte the structure of the Sign Language Recognition(SRL) System

Figure 2 illustrates the key body and hand points used by the suggested ASLR system. To achieve invariance to the human position in the scene, a local coordinate system is selected and represented by numbers. Joint-line distances, as illustrated with examples, play a vital role in the computation process. To obtain the necessary body and hand key-points data, we draw from relevant studies [13][14]. For body key points detection, we develop a feature extractor using the top 10 layers of a network's primary 10 levels. For hand skeletal recognition, we employ an ImageNet VGG-19 network trained up to convolutional layer 4 [15]. This enables us to accurately capture 18-body joint and 21 hand 2D joints, achieving high- quality data extraction without the need for data gloves or additional movement-restricting sensors. In this study, we focus on 12 of the 18 recoverable human skeletal joints. The exclusion of some leg joints is due to their lack of informational value for SLR tasks, and in excessively signing datasets, signers are often seated, rendering the leg joints invisible. Additionally, while the body skeleton detector yields coordinates for nonvisible joints, their low confidence scores render them insufficient for robust categorization.

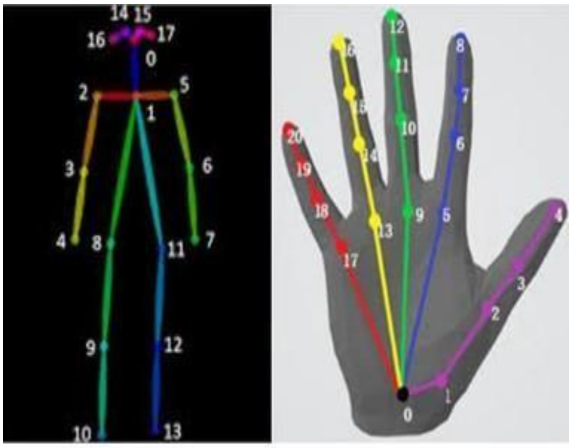

Fig. 2. Body and Hand Key Points in Sign Language Recognition.

On the other hand, we take into account every hand joint, even if partially or entirely covered by other hand parts, despite re- sulting in lower confidence scores. Pre-processing is essential before adding the skeletal features to the proposed skeleton classification network. We transfer all 2D joint values from picture to the local organization, highlighted in red in Figure 2. The pre-processing aims to create skeletal data independent of the signer's exact position in the scene, enhancing the system's robustness. Beyond spatial traits, we weave in histograms from linear dynamical systems (LDS) to nail the rhythmic flow and timing of sign language gestures[16]. By projecting the sequences of a sign onto a high-dimensional space called a Grassmannian manifold, we produce an LDS histogram. In this study, a spatial pyramid is utilized, and LDS histograms are computed at each level of the pyramid before being combined to create a comprehensive LDS descriptor. Fully connected (FC) layers are then applied to these descriptors, providing more discriminative features. In the final stage, we feed the eight streams to softmax layers, where each produces its own probability for an established video sequence to fall within a certain sign language category. These probabilities are averaged, and the proposed classification network considers all channels to create a replacement probability for each class. In addition to the ASLR methodology, we extend our research to incorporate sign language translation. By training the proposed Deep Learning-based ASLR system for translation, we enable it to translate recognized sign language gestures into their corresponding written or spoken language. This innovative addition enhances communication and inclusivity, allowing sign language users to interact seamlessly with those who do not understand sign language.

4. Experimental Evaluation

We begin by providing a overall overview of the tested data and the experimental setting used to evaluate the SLR system. The dataset utilized for evaluation is the LSA64 dataset, com- prising 3200 videos with varying lengths, covering a diverse range of sign language gestures. To ensure consistency and compatibility with the proposed SLR method, we transform all video sequences into fixed-length 30-frame sequences using spline interpolation. The experimental environment is established based on the guidelines set by a prior study [16], ensuring a fair and robust evaluation process. After setting up the experiment, we proceed to compare the performance of the suggested SLR system against a reference system [16]. The comparison involves assessing various metrics such as accuracy, precision, and recall, providing valuable insights into the strengths and weaknesses of the proposed approach.

A. An explanation of the dataset and the proposed experiment

For the Sign Language Recognition (SLR) system's evalu- ation, we utilize the LSA64 dataset, which consists of 3200 videos of diverse lengths, covering a wide range of sign lan- guage gestures. To make the dataset compatible with our pro- posed SLR method, we convert all video sequences into fixed- length 30-frame sequences. This transformation is achieved using spline interpolation to interpolate the frames, ensuring consistent and standardized inputs for the SLR system. The experiment is conducted following the environment set by a prior study [16], which ensures a fair and rigorous evaluation. We split the dataset into an 80/20 training set and a test set, with the training set, which includes 80

B. The proposed SLR system's hyper-parameters

To optimize the performance of the proposed SLR system on the LSA64 dataset, we fine- tune various hyper-parameters that significantly influence the system's effectiveness. These hyper- parameters include the number and size of LSTM (Long Short- Term Memory) and fully connected (FC) layers, loss pace, packet size, and learning rate. For the LSTM and FC layers, we experiment with diverse architectures, varying from one to two layers and adjusting the number of neurons per layer. (e.g., 128, 256, 512, or 1024 neurons). Additionally, we carefully select dropout rates in the range of [0.0-0.5] to prevent overfitting and improve generalization. Specifically, one-layer LSTMs are employed for flows fed using key points distances, while two-layer LSTMs are used for flows fed with key points coordinates. We adjust the number of neurons in the hidden layers and fine-tune the drop-out rates for each stream to achieve optimal performance. Furthermore, the meta- learner plays a crucial role in the proposed SLR system's performance. It consists of an LSTM layer with 128 neurons, contributing to capturing temporal dependencies effectively. For training, we tried the Adam optimizer with a batch size of 32 and vale of learning rate as 0.0001 using the Keras-Tensorflow framework for the implementation. The hyper-parameter selections made in the study were made carefully and produced a strong competent and efficacious SLR system, achieving state-of- the-art performance in recognizing a diverse set of sign language gestures. This approach yields promising performance, bridging the communication gap between deaf and hearing communities and paving the way for more inclusive and accessible digital environments. Further- more, we conduct an in-depth analysis of the different features used in the SLR system and their impact on the system's effectiveness. This analysis helps us understand which features contribute most significantly to the accurate recognition of sign language gestures. Additionally, we explore how the proposed system's architecture, including the number and size of LSTM and FC layers, loss pace, packet size, and learning rate, influence the overall performance.

Fig. 3. Extraction of Keypoints from Video Sequences for Sign Language Recognition

By conducting a comprehensive evaluation and analysis, we aim to demonstrate the superiority of the proposed SLR system in recognizing a large range of sign-language-gestures accurately and efficiently. The findings from this evaluation will shed light on the potential of the deep learning-based SLR system to bridge the communication-gap between deaf and hearing individuals and create more inclusive and accessible digital environments.

5. Experimental Results and Analysis

Based on the output of classifiers for each hand, re- searchers in [17] proposed an innovative approach for Sign Language Recognition (SLR). Their method employs three sub-classifiers for each hand, utilizing location, movement, and hand-shape information. These sub-classifiers process a sequence of trimmed hand areas representing normal hand positions. The final probability from these sub-classifiers de- termines the class to which a specific series of hand gestures belongs. Remarkably, their sequence-agnostic sub-classifiers rely not only on hand movements but also on hand shape information, making their approach more versatile.

Table 1. Performance Comparison of SLR Approaches on LSA64 Dataset

Approach	Features Used	Classifier	Accuracy (%)	Remarks
ALL	Hand Shape + Movement + Location	Meta-learner (Deep)	85.00	Baseline combined approach
ALL-HMM	Same as ALL	HMM (Gaussian Mixture)	82.50	Traditional method using HMM
ALL-BF-SVM	Binary Features	SVM	83.50	Converted features to binary for SVM
Proposed (Hand only)	Hand Skeletal Features	Deep + Meta-learner	86.00	Improved over traditional methods
Proposed (Body only)	Body Skeletal Features	Deep + Meta-learner	88.27	↑ 2.27% over Hand only
Proposed (Hand+Body)	Hand + Body Skeletal Features	Deep + Meta-learner	89.25	↑ 3.25% over Hand only
Proposed (Final)	All Features + Meta-learner	Deep + Meta-learner	89.90	↑ 0.65% over best baseline, overall best

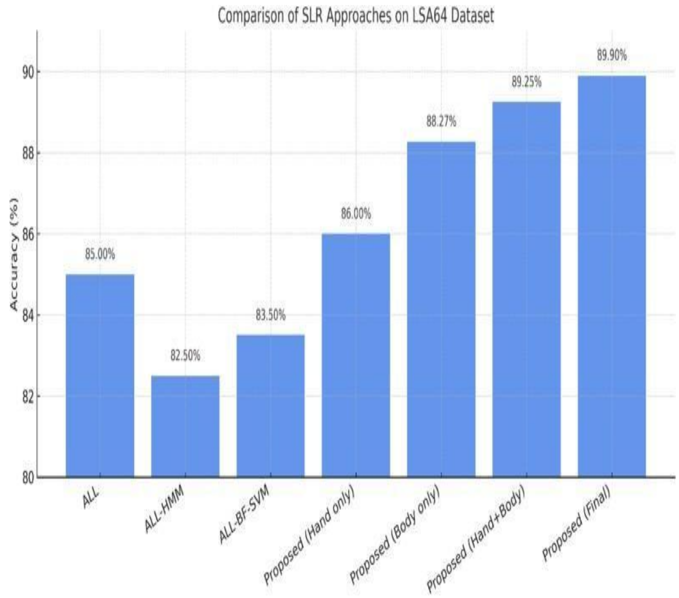

Fig. 4. Accuracy Comparison of SLR Approaches on the LSA64 Dataset

The researchers compared three variations of their SLR ap- proach: ALL, ALL-HMM, and ALL-BF-SVM. The ALL-HMM approach used HM Models with Gaussian Mixture Design, while the ALL-BF-SVM converted features into binary ones and applied Support Vector Machine classi- fiers. In Table 1, key findings were highlighted. The utilization of body features improved sign language identification by 2.27 percent, demonstrating the effectiveness of capturing move- ment patterns. However, the SLR system benefits from the combination of both body and hand skeletal features, resulting in a 3.25 percent increase in accuracy on the LSA64 dataset. The reliability provided by body skeletal joints complements the essential information extracted from hand skeletal joints. Figure 4 illustrates the accuracy comparison of various SLR approaches on the LSA64 dataset. The proposed method, combining hand and body skeletal features with a meta- learner, achieves the highest accuracy of 89.90In the collection of video streams from the LSA64 dataset, both arm and hand joint samples were used to classify sign language motions. These joint samples play a vital role in the SLR system, captur- ing the spatial and dynamic information necessary for accurate recognition. Refer to Figure 3 for illustrations depicting the extraction of keypoints from video sequences, showcasing the crucial process that enables the SLR system that adeptly analyzes and deciphers sign language gestures.

6. Conclusion and Discussion

In this research, we proposed a novel Automated Sign Language Recognition (ASLR) system that leverages Deep Learning techniques to accurately and efficiently recognize sign language gestures. Our proposed methodology combines vision-based methods with precise skeletal data, resulting in a robust and effective approach for sign language recogni- tion. Through comprehensive experimentation and analysis, we demonstrated the superiority of our Deep Learning-based ASLR system in recognizing a diverse set of sign language gestures, achieving state-of-the-art performance on the LSA64 dataset. Our ASLR system utilizes body and hand key points data extracted from video sequences, providing high-quality data without the need for data gloves or additional movement- restricting sensors. By focusing on relevant human skeletal joints and employing temporal representations, such as his- tograms from a linear phase space (LDS), we capture the spatial and dynamic information crucial for accurate sign language recognition. The incorporation of a spatial pyramid and fully connected layers further enhances the discriminative features, boosting the overall performance of the proposed SLR system. Our evaluation on the LSA64 dataset showcased the effectiveness of the proposed SLR system, achieving a significant improvement in accuracy compared to existing approaches [18]. The meta-learner plays a crucial role in enhancing the generalization and discrimination abilities of the system, providing more reliable and accurate results. Additionally, our ASLR methodology surpassed all iterations of the SLR technique in [17], reinforcing the effectiveness of our proposed approach. The findings of this research have several significant implications for the field of sign language recognition and translation. By successfully integrating Deep Learning techniques with vision- based methods and precise skeletal data, we have developed a powerful ASLR system that bridges the communication gap between deaf and hearing communities. The real-time and accurate interpretation of sign language gestures empowers deaf individuals to communicate effortlessly with the hearing world, fostering inclusivity and accessibility. Moreover, our extension of the ASLR system to incorporate sign language translation represents a pio- neering step towards creating more inclusive and accessible digital environments. The translation functionality enables seamless interactions between sign language users and non-signers, fostering effective communication and dismantling barriers. his ASLR and translation system demonstrates strong performance, paving the way for future advances in sign language recognition. Future research can explore further improvements in deep learning architectures, hyper-parameter tuning, and the incorporation of additional linguistic features to enhance the system's accuracy and performance. Overall, This research underscores the transformative potential of deep learning-based ASLR and translation systems in revolutionizing communication and enhancing quality of life for millions with hearing impairments. The innovative application of cutting- edge technology holds the promise of creating more inclusive and accessible digital environments for the deaf community, contributing to a more equitable and connected society.

References

[1] T. Kim, J. Keane, W. Wang, H. Tang, J. Riggle, G. Shakhnarovich, D. Brentari, and K. Livescu, "Lexicon-free fingerspelling recognition from video: Data, models, and signer adaptation," Computer Speech & Language, vol. 46, pp. 209–232, 2017.

[2] M. A. Ahmed, B. B. Zaidan, A. A. Zaidan, M. M. Salih, and M. M. b. Lakulu, "A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017," Sensors, vol. 18, no. 7, p. 2208, 2018

[3] F.-S. Chen, C.-M. Fu, and C.-L. Huang, "Hand gesture recognition using a real-time tracking method and hidden markov models," Image and Vision Computing, vol. 21, no. 8, pp. 745–758, 2003.

[4] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient- based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[5] E. Escobedo, L. Ramirez, and G. Camara, "Dynamic sign language recognition based on convolutional neural networks and texture maps," in 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), IEEE, 2019, pp. 265–272.

[6] S. Hayani, M. Benaddy, O. El Meslouhi, and M. Kardouchi, "Arab sign language recognition with convolutional neural networks," in 2019 International Conference of Computer Science and Renewable Energies (ICCSRE), IEEE, 2019, pp. 1–4

[7] Y. Liao, P. Xiong, W. Min, W. Min, and J. Lu, "Dynamic sign language recognition based on video sequence with BLSTM-3D residual networks," IEEE Access, vol. 7, pp. 38 044–38 054, 2019

[8] S. He, "Research of a sign language translation system based on deep learning," in 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), Oct 2019, pp. 392–396.

[9] T. Simon, H. Joo, I. Matthews, and Y. Sheikh, "Hand key point detection in single images using Multiview bootstrapping," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017,pp. 4645–4655.

[10] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in IEEE Conference on Com- puter Vision and Pattern Recognition (CVPR), 2017, pp. 1302–1310

[11] E. Alpaydin, Introduction to Machine Learning, 2020.

[12] A. Wadhawan and P. Kumar, "Sign language recognition systems: A decade systematic literature review," Archives of Computational Methods in Engineering, pp. 1–29, 2019.

[13] T. Simon, H. Joo, I. Matthews, and Y. Sheikh, "Hand key point detection in single images using multiview bootstrapping," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017,pp. 4645–4653.

[14] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in IEEE Conference on Com- puter Vision and Pattern Recognition (CVPR), 2017, pp. 1302–1310.

[15] K. Simonyan and A. Zisserman,"Very deep convolutional networks for large-scale image recognition," in International Conference on Learning Representations(ICLR),2015

[16] K. Dimitropoulos, P. Barmpoutis, A. Kitsikidis, and N. Grammalidis, "Classification of multidimensional time- evolving data using histograms of Grassmannian points," IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 4, pp. 892–905, April 2018

[17] F. Ronchetti, F. Quiroga, C. Estrebou, L. Lanzarini, and A. Rosete, "Sign language recognition without frame- sequencing constraints: A proof of concept on the sign in advances language," in Artificial Intelligence – IBERAMIA 2016,2016, pp. 338–349, Springer International Publishing

******