International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Comprehensive Survey on Indian Sign Language Generation Systems

Param Chordiya[a], Kedar Dalvi[b], Aneesh Berde[c]

[a]Pune Institute of Computer Technology, Dhankwadi, Pune – 411043, India
[b]Pune Institute of Computer Technology, Dhankwadi, Pune – 411043, India
[c]Pune Institute of Computer Technology, Dhankwadi, Pune – 411043, India

A B S T R A C T

The research focuses on the generation of Indian Sign Language (ISL) using Generative Adversarial Networks (GANs). Sign language is an important means of communication for those who are deaf or have speech problems, and ISL is utilized for this in India. The paper presents a description of the numerous GANs used in ISL to generate hand gestures, as well as the tools and methods used in the process. Furthermore, the difficulties involved with ISL generation are explored as well as suggested study topics for future investigation.

Keywords: Indian Sign Language (ISL), Generative Adversarial Networks (GANs), Machine Learning (ML), Deep Learning (DL)

1. Introduction

People are always inspired to create new ways to interact with machines by needs and new technologies. This interaction can be for a specific purpose or as a framework that can be used in a variety of applications. Sign language recognition is a critical area in which ease of interaction with humans or machines will benefit many people. Indian Sign Language (ISL) is a distinct visual language used by India's deaf and hard of hearing communities [2]. As technology advances, there is a rising interest in developing systems that can generate ISL images from written text.

The study presents an overview of the many types of GANs utilized for ISL picture production, such as Vanilla GANs, Conditional GANs, and Cycle-consistent GANs. The study also goes into the many methods and algorithms that are used in this process, such as natural language processing techniques and image processing approaches. The paper also discusses the obstacles of ISL generation, such as the requirement for big datasets and the intricacy of sign language motions. Additionally, the study investigates potential future research areas in this subject, such as the creation of multi-modal systems capable of producing ISL from both textual and visual inputs [3].

Ultimately, the purpose of this paper is to provide a thorough overview of the present state of the art in ISL image production utilizing GANs [2]. The insights offered in this research have the potential to be extended to additional sign languages and used to improve communication and accessibility for the deaf and hard of hearing people in India.

2. Background

The hard of hearing community uses sign language, visual language, to convey themselves through gestures, facial expressions, and body language. The most used sign language in India is called Indian Sign Language (ISL). The poor knowledge of sign language among hearing people, however, causes a serious communication gap between the deaf and hearing groups. The deaf community may experience social exclusion, restricted access to chances for education and employment, and a decline in quality of life because of this communication gap.[5]

In India, the deaf community uses Indian Sign Language (ISL), a specialized visual language. However, there is a significant communication gap between the hearing and deaf groups due to the low level of sign language literacy among hearing individuals [7]. The generation of sign language generating systems (SLGS), which can automatically create sign language animations or movies from spoken language inputs, is a suggested technological remedy to reduce this gap. Despite the recent growth in interest in ISL SLGS, there are now just a handful of comprehensive surveys that provide an overview of the research on this topic.[8] This survey article provides a comprehensive analysis of the most recent ISL in an effort to close this gap.

The survey paper identifies the main obstacles that researchers and developers in the field of ISL SLGS must overcome, including the dearth of standardized sign language data and the difficulty of sign language grammar [9]. It also highlights the routes this field of study will take in the future, such as the development of interactive systems and the inclusion of multimodal inputs.

3. Methodology

First, a thorough literature review will be carried out. This will entail looking for and reviewing relevant research papers and books on ISL image generation using GANs. Following that, data on the various types of GANs utilized for ISL image generation, such as Vanilla GANs, conditional GANs, and cycle consistent GANs, will be collected [1]. The information will also disclose the tools and procedures used in the process.

The data gathered will be evaluated to detect patterns and trends in the use of GANs for ISL image generation. This will entail comparing and contrasting the various types of GANs, as well as their individual strengths and limitations. The data analysis results will be presented in a clear and plain manner [3]. The findings will be presented in tabular form, with relevant charts and graphs utilized to highlight key findings.

The survey article will also provide a critical analysis of the findings. The talk will focus on the issues of ISL detection and creation, as well as prospective future research areas in this subject. Lastly, the survey report will summarize the key findings and make recommendations for future study in ISL picture generation using GANs. The methods described above will provide a thorough examination of the many types of GANs utilized for ISL picture generation, their distinct strengths and limitations, and prospective future research areas in this subject.

4. System Architecture

GANs consist of two main components: A generator and a discriminator. The generator takes a random noise vector as input and generates synthetic data, such as images, that should resemble the real data. The discriminator is a binary classifier that takes as input either real or synthetic data and predicts whether the input is real or synthetic.

During training, the generator and discriminator are trained in an adversarial manner. The generator attempts to generate synthetic data that the discriminator cannot differentiate from the real data. At the same time, the discriminator tries to correctly classify the input data as real or synthetic.

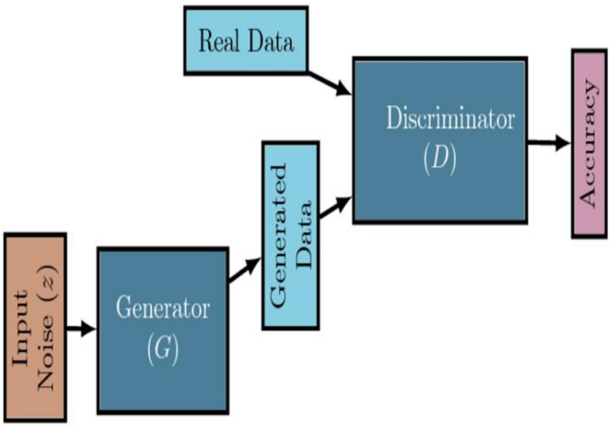

Figure 1 GANs Network

4.1 Generator

The generator takes two inputs, a random noise vector and the conditioning information. The random noise vector is used to introduce some degree of randomness into the generated data, while the conditioning information helps to guide the generator towards generating data that is more specific and targeted. The generator then produces synthetic data that should resemble real data. The generator is typically a deep neural network that maps the input noise vector and conditioning information to the output data.

4.2 Discriminator

The discriminator takes two inputs: the conditioning information and either real or synthetic data. The discriminator is a binary classifier that predicts whether the input data is real or synthetic. The conditioning information is used to help the discriminator make better predictions by providing additional context and information about the data being generated. The discriminator is also typically a deep neural network that maps the input data and conditioning information to a binary classification.

While processing, the generator tries to generate synthetic data that the discriminator cannot differentiate from the real data, taking into account the conditioning information. The discriminator tries to correctly classify the input data as real or synthetic, again taking into account the conditioning information. The process is repeated iteratively until the generator produces synthetic data that is indistinguishable from the real data, while also taking into account the conditioning information. The C-GANs system architecture is a powerful approach to generate synthetic data that is more specific and targeted, by taking into account additional conditioning information. The use of conditioning information has enabled C-GANs[9] to be used for various applications, such as image-to-image translation, text-to-image synthesis, and video generation.

5. Limitations and Future Scope

5.1 Limitations

Data scarcity: One of the most significant issues for ISL SLGS is the scarcity of high-quality data. Generating labelled datasets for training machine learning models is time-consuming and expensive. Furthermore, it is difficult to obtain data that adequately portrays the diversity of sign languages and signing communities.

Lack of standardization: Sign languages fluctuate greatly between regions and communities. This makes developing standardized models that may be utilized for diverse sign languages or signing communities challenging. Further research on cross-lingual and cross-cultural sign language generation is required.

Naturalness and expressiveness: Creating realistic and expressive sign language animations remains a major difficulty. The complex nuances of sign language, such as facial emotions and body language, may be missed by current approaches and methodologies. Enhancing the naturalness and expressiveness of sign language animations is a critical topic of study [6].

User-centered design: It is critical to create ISL SLGS that match the needs and preferences of signing communities. More user-centered design techniques that include input from signing communities and take into account their specific language and cultural aspects are required.

5.2 Future Scope

Using multimodal inputs: Sign languages are communicated not just through hand movements but also through facial expressions, body language, and other non-manual elements. Further study on combining these multimodal inputs into ISL SLGS is required.

Creating customized models: Creating personalized models that adapt to individual signing techniques and preferences may aid in improving the naturalness and expressiveness of sign language animations. Further study on generating individualized models for ISL SLGS is required [4].

GAN-based approaches are being advanced: GAN-based approaches have showed potential in generating realistic sign language animations. Further research on advancing GAN-based techniques is required, including the development of more effective architectures and training strategies.

6. Conclusion

We reviewed the present level of research on Indian Sign Language Generation systems in this survey article (ISL SLGS). With the development of numerous techniques and procedures based on machine learning and computer vision, we have witnessed remarkable improvement in recent years. However, other problems remain, including the restricted availability of high-quality data, the lack of standards, and the need for more realistic and expressive sign language animations.

In the future, there will be various intriguing research paths, including combining multimodal inputs, constructing personalized models, and expanding GAN-based techniques. Additionally, it is critical to assess and improve ISL SLGS accessibility for signing communities. We can continue to advance the state of the art in ISL SLGS and help to establish more inclusive and accessible communication tools for signing communities by tackling these difficulties and pursuing these approaches.

Acknowledgements

We would like to express our gratitude to all those who have contributed to this review paper, including the authors of the cited references and the reviewers who provided valuable feedback.

REFERENCES

1. Rokade, Yogeshwar & Jadav, Prashant. (2017). Indian Sign Language Recognition System. International Journal of Engineering and Technology. 9. 189-196. 10.21817/ijet/2017/v9i3/170903S030.
2. K. Shenoy, T. Dastane, V. Rao and D. Vyavaharkar, "Real-time Indian Sign Language (ISL) Recognition," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 2018, pp. 1-9, doi: 10.1109/ICCCNT.2018.8493808.
3. Patil, R.C., Patil, V., Bahuguna, A., & Datkhile, G. (2021). INDIAN SIGN LANGUAGE RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK. Strad Research.
4. Asifullah Khan, Anabia Sohail, Umme Zahoora, & Aqsa Saeed Qureshi (2020). A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review, 53(8), 5455–5516.
5. Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan and Y. Zheng, "Recent Progress on Generative Adversarial Networks (GANs): A Survey," in IEEE Access, vol. 7, pp. 36322-36333, 2019, doi: 10.1109/ACCESS.2019.2905015.

6. K. Zhu, X. Liu and H. Yang, "A Survey of Generative Adversarial Networks," 2018 Chinese Automation Congress (CAC), Xi'an, China, 2018, pp. 2768-2773, doi: 10.1109/CAC.2018.8623645.
7. Liu, L., Ouyang, W., Wang, X. et al. Deep Learning for Generic Object Detection: A Survey. Int J Comput Vis 128, 261–318 (2020). https://doi.org/10.1007/s11263-019-01247-4.
8. Madakannu, Arun & Selvaraj, Arivazhagan. (2020). DIGI-Net: a deep convolutional neural network for multi-format digit recognition. Neural Computing and Applications. 32. 10.1007/s00521-019-04632-9.
9. Q. Xuan, Z. Chen, Y. Liu, H. Huang, G. Bao and D. Zhang, "Multiview Generative Adversarial Network and Its Application in Pearl Classification," in IEEE Transactions on Industrial Electronics, vol. 66, no. 10, pp. 8244-8252, Oct. 2019, doi: 10.1109/TIE.2018.2885684.
10. C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 105-114, doi: 10.1109/CVPR.2017.19.