Analysis of Deep learning models for Recognition and Interpretation of Indian Sign Language

ANIMESH SINGH[1],S.K.SINGH[2], AJAY MITTAL[3]
[1]Chandigarh College of Engineering and Technology, Chandigarh, India (e-mail: animeshsingh@ccet.ac.in)
[2]2Chandigarh College of Engineering and Technology, Chandigarh, India (e-mail: sksingh@ccet.ac.in)
[3]University Institute of Engineering Technology, Chandigarh, India (email: ajaymittal@pu.ac.in)

ABSTRACT Communication is an important part of our daily activities that helps us to express our feelings and emotions to others. Normal people do this by verbal and non-verbal form however for the mute and hearing impaired people sign language is the merely option. Sign language is not universal and hence the speech impaired people of India use Indian sign language for interacting with others. In this perspective, we explore the present deep learning-based Indian sign language (ISL) recognition and interpretation techniques anticipated by different researchers. The intention of this learning is to explore deep learning based techniques that influence the recognition and interpretation of India sign language that helps to develop a robust virtual interpreter. Our exploration helps other researchers to comprehend the subject matter in further facet.

KEYWORDS Machine Learning; Deep Learning; Indian Sign Language (ISL)

I. INTRODUCTION

Speech and hearing loss in a person is very common nowadays. As per World Health Organization (WHO) report [1] over 5% population of world i.e., 430 million needs attention to the problem of hearing loss and if it is not in control then this figure may increase upto 700 million by 2050. As per the record of Office of the Registrar General and Census Commissioner of India, Ministry of Home Affairs, Government of India [2] total 2.68 Crore (approx.) population are physically disable and out of this 71 Lakhs (approx.) are hearing or speech impaired. Consequently, such a big population has to be given attention and equal opportunity as compared to normal population. In India hearing and speech impaired people use Indian sign language to communicate. Indian sign language requires the use of both hands to do sign gestures. As sign language is not common to others hence it is hard for them to understand. In today's world of computing technology a virtual sign language interpreter can solve the problem of physical interpreter availability for 24*7. With the evolution of deep learning models we can easily prepare a robust state of the art model that translates sign language gestures performed by the signer to the text or speech as required by the end user. Our work break the communication hurdle between normal and hearing and speech impaired people and bridge the gaps with the help of an application build on the basis of deep learning model. For this several deep learning algorithms based papers are studied and analyzed that are earlier used by the researchers for the recognition of Indian sign language. Further, this article illustrates the literature survey, result comparison of deep learning algorithms and conclusion of the related work followed by references.

II. LITRATURE SURVEY

Due to recent development in the field of machine learning [3], [4] and artificial intelligence [5], many researchers are proposing machine learning and AI-based solutions to real-world problems [6]–[10] such as the detection of DDoS attacks [11]–[15], web application attacks [16], [17], XSS attacks [18], image recognition [19], [20] and phishing attacks [21]. Different researchers use different type of deep learning models to get a good accuracy level in Indian sign language recognition. Several deep learning models like Convolution neural network (CNN), Long Short Term Memory Networks (LSTM), Recurrent Neural Networks (RNN) etc. are used by the researchers in real time Indian sign language recognition techniques. Author [22] proposed a CNN model for static gesture and LSTM model for Dynamic gesture for 26 English alphabets, 10 numerals and 10 words and they get 99.81% accuracy for static and 99.08gestures. Author in [23] proposed a CNN model with convolutional, max pooling and ReLu layers then fully connected layers and softmax function for both single and double hand gestures and they get an accuracy of 99.72% to 99.90%. Author in [24] proposed a CNN model and uses skin colour based technique for segmentation for the dataset captured in uniform background

and they achieved an accuracy of 99%. Author in [25] works on seven computer words gesture to propose a model to help the hearing and speech impaired people for educating computer programming. Out of these seven words three are static and four are dynamic. They use a CNN model with Adaboost optimizer to achieve an accuracy of 61.58%. Author in [26] proposes a CNN model that works on only 24 English alphabets and they get an accuracy of 95.67%. Author in [27] compares three different models VGG16, VGG16 with transfer learning and HNN models for 26 English alphabets and found that HNN achieves a better accuracy of 97%-98.52% than others. Author in [28] proposed a CNN model that works on only single handed static gestures i.e., 26 English alphabets and 9 numerals and uses skin colour detection and edge detection method for segmentation and gets an accuracy of 95%. Author in [29] propose a video calling application that compares different models like YoLo v3, YoLo v4, faster RCNN and CNN on 26 English alphabets and 9 numerals and found that CNN gives better accuracy of 93.2% than others. Author in [30] proposes a three layer CNN model approach for 24 English alphabets and 10 numerals nd achieves an accuracy of 99numerals and 97.6% for alphabets. Author in [31] tested four different combinations of CNN + LSTM model and found that LSTM + CNN works better than others and they achieve an accuracy of 97% for 11 custom collected words. Author in [32] found that CNN models work better in their custom dataset and they achieve an accuracy of 92.43%. Author in [33] use a MediaPipe approach to determine the region of interest in hand, face and pose landmarks that are useful in classification and proposed an analysis that for low complex gestures GRU works better and for high complex gestures LSTM works better for custom made 10 words in which it achieves an approximate accuracy of 99%. CNN model is used by maximum researchers however it varies in number of layers and parameters chosen.

III. RESULTS AND DISCUSSION

As pre build dataset for Indian sing language are very rare therefore many researchers do their experiments on their own customized dataset for Indian sing language. We have also created a dataset published in Mendeley [34] that are used by different researchers for real time Indian sign language recognition. Our dataset is a robust dataset that consists of the static Indian sign language gestures from all the age group people i.e., from kids to senior adults. A survey for the different methods and techniques was done by author in [35] for dataset acquisition of static gestures of Indian sign language. In this paper we try to evaluate the different deep learning models used for Indian sign language recognition with their type and number of dataset used along with the accuracy they achieved. The brief summary of this study is shown in the Table 1.

IV. CONCLUSION

This study article helps the novice researcher to do the analysis of different sign language recognition techniques and get the concept and idea to remove the gaps and evolve a new method with increased accuracy and less error. The research is increasing year by year in the field of Indian sign language recognition as it is shown in Figure 1. In ISL signing is done by one and both hands so Figure 2 shows the experiments done on single or both hands

TABLE 1: Accuracy of Different Approaches

Author	Year	Model	Dataset (Type & Numbers)	Accuracy
Bhagat et al.	2019	CNN, LSTM	Custom(26English,10Numerals and 10Words)	98.81%
Bohra et al.	2019	CNN	Custom(26English,10Numerals and 3Words)	99%
Thanasekhar	2019	CNN	Custom(7Programming words)	61.58%
Wadhawan et al.	2020	CNN	Custom(26English,10Numerals and 67Words)	99.72%
Raghavachari	2020	CNN	Custom(24English)	71.85%
Sharma	2021	HNN	Custom(26English)	97%-98.52%
Patil	2021	CNN	Custom(26English,9Numerals)	95%
Chhajed	2021	CNN, YOLOv4	Custom(26English,10Numerals and 2Words)	93.20%
Sharma	2021	CNN	Custom(24English,10Numerals)	97.6%-99%
Kothadiya	2022	LSTM + CNN	Custom(11Words)	97%
Sharma	2022	CNN	Custom ISL	92.43%
Samaan	2022	LSTM	Custom(10Words)	99%

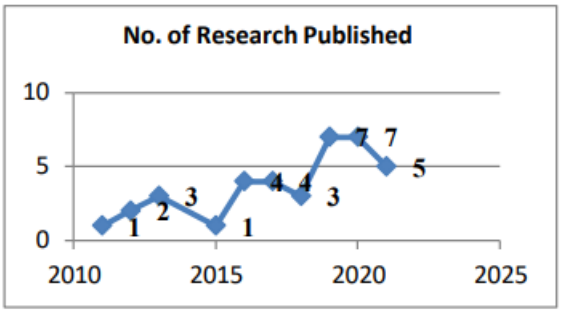

FIGURE 1: Number of Published work

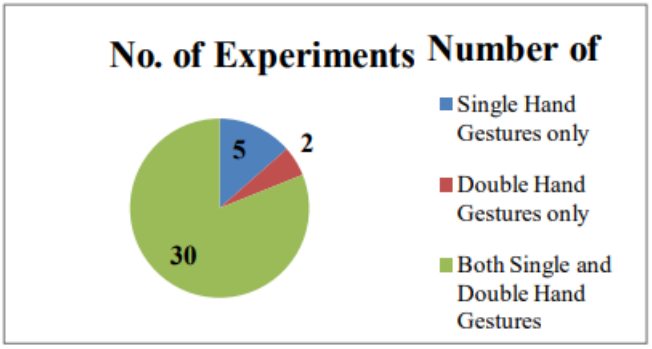

FIGURE 2: Number of Experiments

REFERENCES

[1] W. H. Organiazation, "deafness-and-hearing-loss," https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss, 2022, [Online; accessed 19-July-2022].

[2] C. TABLES, "Office of the Registrar General Census Commissioner, India," https://censusindia.gov.in/census.website/data/census-tables, 2022, [Online; accessed 19-July-2022].

[3] C. L. Stergiou and et al., "Secure machine learning scenario from big data in cloud computing via internet of things network," in Handbook of computer networks and cyber security. Springer, 2020, pp. 525–554.

[4] S. R. Sahoo and et al., "Hybrid approach for detection of malicious profiles in twitter," Computers & Electrical Engineering, vol. 76, pp. 65–81, 2019.

[5] M. Hammad, M. H. Alkinani, B. Gupta, A. El-Latif, and A. Ahmed, "Myocardial infarction detection based on deep neural network on imbalanced data," Multimedia Systems, vol. 28, no. 4, pp. 1373–1385, 2022.

[6] N. Kumar and et al., "A novel framework for risk assessment and resilience of critical infrastructure towards climate change," Technological Forecasting and Social Change, vol. 165, p. 120532, 2021.

[7] A. Tewari and et al., "A lightweight mutual authentication protocol based on elliptic curve cryptography for iot devices," International Journal of Advanced Intelligence Paradigms, vol. 9, no. 2-3, pp. 111–121, 2017.

[8] B. B. Gupta and et al., "Blockchain-assisted secure fine-grained searchable encryption for a cloud-based healthcare cyber-physical system," IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 12, pp. 1877–1890, 2021.

[9] M. Kaur and et al., "Secure and energy efficient-based e-health care framework for green internet of things," IEEE Transactions on Green Communications and Networking, vol. 5, no. 3, pp. 1223–1231, 2021.

[10] A. Tewari and et al., "A mutual authentication protocol for iot devices using elliptic curve cryptography," in 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2018, pp. 716–720.

[11] M. Chhabra and et al., "A novel solution to handle ddos attack in manet," 2013.

[12] A. Tewari and et al, "Secure timestamp-based mutual authentication protocol for iot devices using rfid tags," International Journal on Semantic Web and Information Systems (IJSWIS), vol. 16, no. 3, pp. 20–34, 2020.

[13] K. Alieyan and et al., "Dns rule-based schema to botnet detection," Enterprise Information Systems, vol. 15, no. 4, pp. 545–564, 2021.

[14] A. Dahiya and et al., "A reputation score policy and bayesian game theory based incentivized mechanism for ddos attacks mitigation and cyber defense," Future Generation Computer Systems, vol. 117, pp. 193–204, 2021.

[15] A. Mishra and et al., "Classification based machine learning for detection of ddos attack in cloud computing," in 2021 IEEE International Conference on Consumer Electronics (ICCE). IEEE, 2021, pp. 1–4.

[16] S. Gupta and et al., "Detection, avoidance, and attack pattern mechanisms in modern web application vulnerabilities: present and future challenges," International Journal of Cloud Applications and Computing (IJCAC), vol. 7, no. 3, pp. 1–43, 2017.

[17] A. Tewari and et al., "An analysis of provable security frameworks for rfid security," in Handbook of computer networks and cyber security. Springer, 2020, pp. 635–651.

[18] B. B. Gupta, S. Gupta, and P. Chaudhary, "Enhancing the browser-side context-aware sanitization of suspicious html5 code for halting the dom-based xss vulnerabilities in cloud," International Journal of Cloud Applications and Computing (IJCAC), vol. 7, no. 1, pp. 1–31, 2017.

[19] L. Zou and et al., "A novel coverless information hiding method based on the average pixel value of the sub-images," Multimedia tools and applications, vol. 78, no. 7, pp. 7965–7980, 2019.

[20] B. B. Gupta, S. Yamaguchi, and D. P. Agrawal, "Advances in security and privacy of multimedia big data in mobile and cloud computing," Multimedia Tools and Applications, vol. 77, no. 7, pp. 9203–9208, 2018.

[21] A. K. Jain and et al., "Phish-safe: Url features-based phishing detection system using machine learning," in Cyber Security. Springer, 2018, pp. 467–474.

[22] N. K. Bhagat, Y. Vishnusai, and G. Rathna, "Indian sign language gesture recognition using image processing and deep learning," in 2019 Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2019, pp. 1–8.

[23] A. Wadhawan and P. Kumar, "Deep learning-based sign language recognition system for static signs," Neural computing and applications, vol. 32, no. 12, pp. 7957–7968, 2020.

[24] T. Bohra, S. Sompura, K. Parekh, and P. Raut, "Real-time two way communication system for speech and hearing impaired using computer vision and deep learning," in 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, 2019, pp. 734–739.

[25] B. Thanasekhar, G. D. Kumar, V. Akshay, and A. M. Ashfaaq, "Real time conversion of sign language using deep learning for programming basics," in 2019 11th International Conference on Advanced Computing (ICoAC). IEEE, 2019, pp. 1–6.

[26] C. Raghavachari and G. S. Sundaram, "Deep learning framework for fingerspelling system using cnn," in 2020 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2020, pp. 469–473.

[27] A. Sharma, N. Sharma, Y. Saxena, A. Singh, and D. Sadhya, "Benchmarking deep neural network approaches for indian sign language recognition," Neural Computing and Applications, vol. 33, no. 12, pp. 6685–6696, 2021.

[28] R. Patil, V. Patil, A. Bahuguna, and G. Datkhile, "Indian sign language recognition using convolutional neural network," in ITM Web of Conferences, vol. 40. EDP Sciences, 2021, p. 03004.

[29] R. R. Chhajed, K. P. Parmar, M. D. Pandya, and N. G. Jaju, "Messaging and video calling application for specially abled people using hand gesture recognition," in 2021 6th International Conference for Convergence in Technology (I2CT). IEEE, 2021, pp. 1–4.

[30] P. Sharma and R. S. Anand, "A comprehensive evaluation of deep models and optimizers for indian sign language recognition," Graphics and Visual Computing, vol. 5, p. 200032, 2021.

[31] D. Kothadiya, C. Bhatt, K. Sapariya, K. Patel, A.-B. Gil-González, and J. M. Corchado, "Deepsign: Sign language detection and recognition using deep learning," Electronics, vol. 11, no. 11, p. 1780, 2022.

[32] S. Sharma and S. Singh, "Recognition of indian sign language (isl) using deep learning model," Wireless Personal Communications, vol. 123, no. 1, pp. 671–692, 2022.

[33] G. H. Samaan, A. R. Wadie, A. K. Attia, A. M. Asaad, A. E. Kamel, S. O. Slim, M. S. Abdallah, and Y.-I. Cho, "Mediapipe's landmarks with rnn for dynamic sign language recognition," Electronics, vol. 11, no. 19, p. 3228, 2022.

[34] A. e. a. Singh, "Static gestures of indian sign language (isl) for english alphabet, hindi vowels and numerals, mendeley data."

[35] A. Singh, S. K. Singh, and A. Mittal, "A review on dataset acquisition techniques in gesture recognition from indian sign language," Advances in Data Computing, Communication and Security, pp. 305–313, 2022.