International Journal of Research Publication and Reviews, Vol 4, no 5, pp 2250-2259, May 2023

International Journal of Research Publication and Reviews

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

Deep Learning Methods for Sign Language Prediction & Recognition using N-gram Model

Shrikant N. Sarda

Department of Computer Science and Engineering, Research Scholar, Shri. JJTU, Rajasthan.

A B S T R A C T

Sign language is a three-dimensional visual-spatial language used by persons who are deaf or hard of hearing to communicate. Sign language, unlike spoken languages, can only be represented by gestures. Sign language enables communication between those who are deaf and the rest of society. Sign Languages are widely seen as the most important gestures to be identified in the Gesture Recognition (GR) hierarchy. Research has resulted in the development of several Sign Language Recognition (SLR) systems, although these systems are confined to recognizing solitary sign motions. This study presented a deep learning-based self-adaptive technique for SLR and prediction using the N-gram model. Recognition is performed separately by Hidden Markov Model (HMM), which is a numerical Markov model in which the represented scheme is thought to be a Markov procedure with unobservable ("hidden") conditions. The best recognition accuracy has been achieved as high as 97.85% with HMM.

Keywords: Deep Learning, sign language recognition, N-gram model, Hidden Markov Chain, Gesture Recognition, Prediction.

Introduction

Sign Language is the most widely used, but not predominant, mode of communication among significant groups of individuals in society. Over 466 million people are deaf according to a World Health Organization (WHO) report that would be released in 2020 [1]. Various groups use multiple kinds of sign languages which including the USA, Argentina, Poland, Germany, Greek, Spain [2], China, Korea, Iran [3], and so on. Building a strong system that could translate spoken languages into sign languages and then vice versa is essential to facilitating reciprocal communication among the hearing and deaf populations. As a result, a cooperative system must include recognition of the creation of sign language [4].

1. Automatic Sign Language Recognition (ASLR) is a significant issue for facilitating interaction with those who are sensory or gustatory handicapped. Sign Languages are widely seen as the most important gestures to be identified in the Gesture Recognition (GR) hierarchy. Such significance motivates automated identification of signs as a multidisciplinary development subject to tackle various disciplines such as Human-Computer Interaction, Computer Vision, Pattern Recognition, Neural Networks, Machine Learning, Fuzzy Systems, and others. Since hand motions could be seen in several ways, it's difficult to accurately recognize sign language [5].

2. Sign language translation objectives are to either create a video of sign construction or to extract a corresponding verbal language phrase from a video of anyone continuously signing translation [6]. However, much of the current work in computer vision has intensive on recognizing the arrangement of sign glosses [7]. Continuous Sign Language Recognition (CSLR) is relatively more than the occupied conversion to a spoken language equivalent to Sign Language Translation (SLT). Such differences are crucial spoken and sign languages have quite similar grammar. Generally said, the mapping around language with signs is complicated, and there is no straightforward word-to-sign plotting. Figure 1 depicts an evaluation of end-to-end sign language [8].

There is a wide range of sign language technology from the capture of signs to their genuine portrayal to enhance communication among hearing-impaired and hearing individuals, including among speaking persons and hearing-impaired. Sign language capture entails the correct extraction of mouth expressions, hand, and body utilizing suitable sensing devices in marker-based and marker-less settings [9]. The precision of sign language recording technology is presently restricted by sensor resolution and discriminating capability, as well as the fact that occlusions and quick hand motions offer substantial hurdles to reliable sign capture [10].

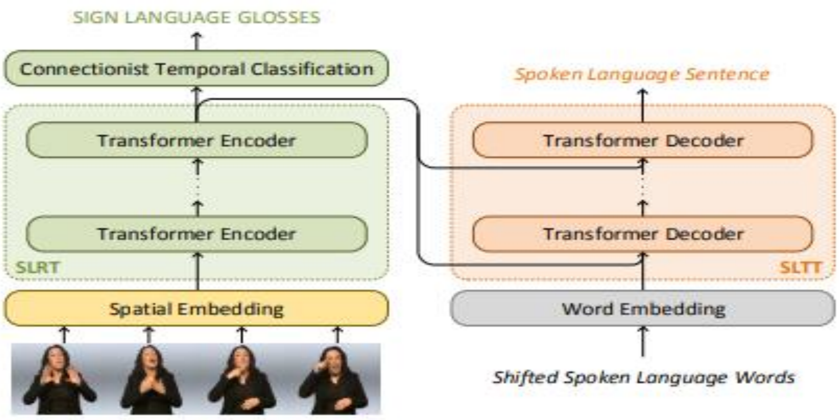

Figure 1. An evaluation of end-to-end Sign Language [8]

Sign Language Recognition (SLR) includes the creation of sophisticated machine learning procedures that reliably categorize social speeches as single signals and whole sentences [11]. Sign language is understood by just a small percentage of individuals. After that, communication among the Deaf population then the hearing majority has become even more difficult. Due to the difficulty of writing spoken languages, written communication is laborious for Deaf people. Furthermore, the form of communication is impersonal and sluggish in face-to-face talks [12].

1.1 Sign Notations

There are a variety of systems that could be used to signify Sign Languages, but no standard notation has ever been established. An illustration is required for the alternate phase among translation and recognition from Sign to Text, as previously indicated. Some of the current systems are covered slightly below [13].

- **Stokoe Notation**

Stoke notation is the first phonemic script used in American Sign Language. Linguist William Stokoe created the Stokoe notation in 1965 [14]. Stokoe notation concentrates primarily on manual components (hand gestures) and then excludes non-manual elements like head movements, shoulder movements, facial expressions, and eye gaze [15]. Stokoe notation is linearly encoded in American Standard Code For Information Interchange (ASCII). In his sign language dictionary, William Stokoe grouped 55 constraints into three categories such as movement, location, and hand shape. Initially, Stokoe designed just the tabula, designator, and signification for his notation system. The present structure of the Stokoe notation system comprises four crucial criteria that are measured when creating a valid sign, which is described below:

- Dez and designator and handshape.
- Tab and tabula and hand location.
- Alignment of hand.
- Sig and Significance and action or motion [16].

- **Hamburg Notation System**

Signed language writing was pioneered in 1984 by academics at the University of Hamburg. The Hamburg Notation System (HamNoSys) was created as a phonetically based representation scheme. The system was originally handwritten, but an appliance-readable Unicode is accessible from the Campus of Hamburg. Signing Gesture Markup Language (SiGML) is an XML encrypting of HamNoSys that is also accessible. It was created as part of the ViSiCast project. The writing notation for any sign language is Ham No Sys. It is possible to use this notation system to create both non-manual and manual signs. Figure 2 shows the structure of the Hamburg Notation [17].

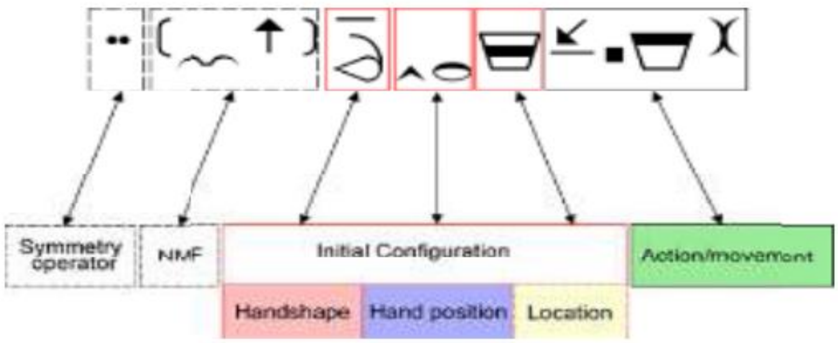

Figure 2: Structure of Hamburg Notation [18]

- **ASL-Phabet**

In the 1990s, Samuel Supalla, a deaf linguist started employed on writing American Sign Language. At the Salk Institute in California, a streamlined signwriting system based on the Sign Font writing system. It collaborated with the Canadian Cultural Society dictionary for children using his writing notations; ASL-phabet does not contain non-manual elements and is not meant for full sentence writing. His writing notation makes use of a total of 10 different hand forms, 5 different places, and 10 different movement patterns. The series of symbols represents hand form, position, and movement. Figure 3 shows the signs used in ASL-phabet notations:

Figure 3: Sign used in ASL-phabet Notation system [18]

- **SignWriting**

SignWriting is fast establishing the standard for sign languages. Most sign languages throughout the globe employ the International SignWriting Alphabet (ISWA), which was established in 2010 and comprises all the symbols required to write the hand forms, motions, facial expressions, and bodily gestures [19]. Notation for writing signs' visual language is called SignWriting. Its components enable the graphic display of any sign language, the depiction of the linguistic construction of sign languages in a graphical system. Valérie Sutton created SW as a component of a wider scheme of drive notation known as Sutton Movement Writing and Shorthand (SMWS), which could record any sign language and movement in dance as well as in sports and physiotherapy. In Brazil, the usage of SignWriting is still quite restricted, possibly as most data and learning heritage is documented in other written languages, such as Portuguese. Figure 4 depicts the examples of signwriting [20].

Figure 4:Examples of SignWriting [20]

1.2 N-gram Model

Language modeling uses n-gram models are probabilistic representations of the manuscripts that make usage of a small number of words, or history dependencies, where n denotes the number of words that are involved in the dependency relationship. In automatic speech recognition, n-grams are significant to classical some of the physical procedures of normal language. For example, the model uses word dependences to allocate a greater possibility to "how are you today" than to "are you how today you," even though both phrases comprise accurate identical arguments. In several basic unigram language representations, n-gram models with n are equal to 1, and models without term dependencies resulted in high retrieval. The introduction of bigram models n-gram models with n is equal to 2 would enable the scheme to represent through term dependencies and treat New York inversely from independent existences of New and York, perhaps boosting recovery presentation. The algorithm could employ trigram representations to detect direct instances of the New York metro. Figure 5 depicts the basic structure of the N-gram Model [21].

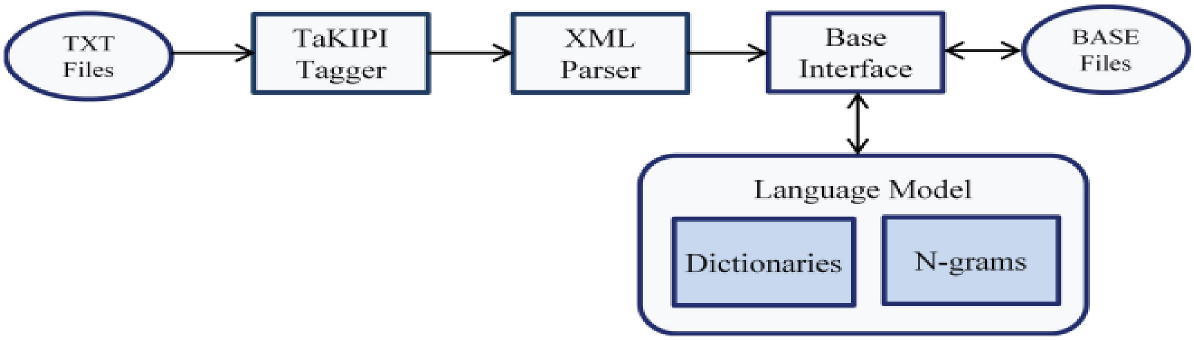

Figure 5: Basic structure of the N-gram Model [22]

Literature Review

The following section discusses the various studies based on a Deep learning-based self-adaptive technique for sign language recognition and prediction using N- the gram model. Several researchers explained their findings as seen below:

Huang Y. et al., (2020) [23] stated that a deep learning-based technique for detecting fake news researchers establish that the suggested model outperforms state-of-the-art methods, with a maximum precision of 99.4 percent. The researchers additionally analyze the problem of cross-domain intractability and get the best efficiency of 72.3 percent. According to the author, it is still better to improve the ensemble learning model in dealing with cross-domain intractableness.

Ibrahim N. et al., (2020) [24] explained the conclusion integration of a Convolutional Neural Network (CNN) into an HMM by using a Bayesian framework to understand the CNN's outputs. A hybrid strategy was investigated in terms of the impact of both CNN- and HMM structures. Finally, researchers demonstrated that using ensembles of hybrid CNN-HMMs could improve efficiency even more.

Borg M. et al., (2020) [25] examined a two-stage deep learning (DL)-based ASLR scheme. A typical computer vision pipeline or Open Pose hand key points are used as inputs to the system. Hand motion information inherent in a low-rank trajectory space could be extracted using a factorization technique as part of a pipelined deep learning strategy. Many recent efforts using DL in ASLR have been outperformed by approach by a margin of 10.2% or more Only one recent study (1.3 percent) outperforms the suggested system.

Zhang Y. et al., (2019) [26] computed the enhanced three-way CNN for Sentence-level Sentiment Classification using Deep Learning and then the Traditional feature-based. The researcher developed a three improved CNN model, which was identified as 3W-CNN. Using the improved model, 3W-CNN could be realized as a combination approach for improving CNN. According to the research findings, three-way choices have the potential to boost sentiment categorization accuracy even more.

Wei C. et al., (2019) [27] found deep construction with numerous classifiers for CSLR. To recognize the words then n-grams in a phrase, the author suggested an n-gram classifier (NGC) component and a word-independent classifiers (WIC) module. The method was put to the test on a Chinese benchmark for CSLR, then the outcomes suggest that it was effective and superior.

Mittal A. et al., (2019) [28] computed that an SLR model for seamless gesture recognition could recognize a chain of connected gestures. A new CSLR framework based on the leap motion sensor. The suggested approach was evaluated using 942 signed phrases in Indian Sign Language (ISL). These symbol phrases were identified with the help of 35 distinct sign terms. On signed phrases and isolated symbol words, the precision level was 72.3 percent and 89.5 percent, respectively.

Kındıroglu A. et al., (2019) [29] suggested the development of Temporal Accumulative Features (TAF), a visual representation based on poses that aggregates joint heatmaps across sign language movies, and the application of these features to the categorization of single sign signals. Using the Deeplab v3+ algorithm, the researcher constructed a handshape-based segmentation approach and added the segmented handshapes to the TAF method. The Bosphorus Sign dataset has demonstrated the effectiveness of TAF models for the recognition of sign language.

Farag I. et al., (2019) [30] stated that segmentation of continuous motion data using a new approach for use with Japanese Sign Language continuous sentence phrases. The algorithm distinguishes between static and dynamic signal stages and achieves division precision comparable to prior state-of-the-art methods. A two-stage classifier for continuous sign motion data based on their technique should then be used to improve identification accuracy in this demanding environment.

Pu J. et al., (2018) [31] computed that real-world CSLR using a new deep neural network with an iterative optimization technique. For visual feature extraction, the authors employ a 3-Dimensional Residual Network (3D- ResNet). On a massive continuous sign language benchmark, readers test the hypotheses. Results show a decreased World Error Rate (WER) and the authors' suggestion had outperformed the current best practice.

Zheng L. et al., (2017) [32] stated that a comprehensive introduction to deep learning is based on SLR techniques. The researchers address many sorts of such algorithms for recognition from the perspectives of accessible modalities offered by depth sensors, as well as feature classification and extraction. In the field of sign language research, current developments in deep learning have complete a significant contribution.

Camgoz N. et al., (2017) [33] explained that sequence-to-sequence learning architecture learns middle depictions to attendant the knowledge process. The author had accomplished a hand figure recognition network using the Sub-Networks (SubNet) construction since hands are one of the most information networks of a sign. SubNets that learn intermediate depictions help the system simplify and improved, according to CSLR.

Cui R. et al., (2017) [34] found that Deep convolutional neural networks for CSLR are suggested in this study. The deep neural network architecture was being trained using a staged optimization process. The improved feature representations allow for optimizing the sequence learning model and enterprise a less managed detection network for regulation.

Pigou L. et al., (2017) [35] examined that Deep residual networks could recognize patterns in movies of continuous hand gestures and sign language using conventional Red, Blue, and Green (RGB) cameras and little no preprocessing. The Corpus Nederlandse Gebarentaal (NGT) had a highest ten framewise precision of 73.3 percent while the Corpus Vlaamse Gebarentaal (VGT) had a highest ten framewise precision of 55.7 percent. Through the ChaLearn LAP Continuous Gesture Dataset (ConGD) Challenge, users earned a mean Jaccard index of 0.3164.

Kumar P. et al., (2017)(1) [36] explained that by using a Coupled Hidden Markov Model (CHMM), the author offers a unique multi-sensory combination framework for SLR. Existing data fusion algorithms were used to evaluate the dataset. With CHMM, the highest recognition accuracy has been reached at 90.80%. CHMM-based strategy outperforms previous data fusion algorithms in terms of recognition performance.

Kumar P. et al., (2017)(2) [37] found that a new multimodal outline for isolated SLR has been developed using sensor devices. Hidden Markov Model (HMM) then Bi-directional Long Short-Term Memory Neural Network (BLSTM-NN) built sequential classifiers conduct recognition independently. With HMM, researchers found enhancements of 2.26 percent ‘single hand’ and 0.91 percent ‘both hands’, while with BLSTM-NN classifiers, researchers found improvements of 2.88 percent ‘single hand’ then 1.67 percent both hands.

Li K. et al., (2016) [38] showed the architecture of SLR systems and a novel technique for representing the transition of data between signs in CSLR. The researchers developed a continuous SLR modeling system based on Automated Speech Recognition (ASR) hidden Markov modeling approaches. The good findings show that the intended SLR framework could be used to construct real-world SLR applications.

Gattupalli S. et al., (2016) [39] examined that using deep learning algorithms and CNN, researchers evaluated deep learning-based stance approximation for SLR. By doing user-independent trials on the dataset, the author assesses the presence of two deep learning-based stance approximation approaches. Researchers also use transfer learning to increase posture estimate accuracy, and the results show that transfer learning could help.

Koller O. et al., (2016) [40] stated that the end-to-end embedding of a CNN into an HMM, including Bayesian interpretation of the CNN's outputs. An end-to-end embedding enables to outperform the state-of-the-art on 3 standard CSLR responsibilities by among 38% and 15% qualified then up to 13.3% complete.

Koller O. et al., (2015) [41] evaluated that a deep convolutional neural network was used to represent the robustness of mouth forms in the situation of SLR. The authors suggest a technique for building a CNN in an unsupervised, unlabeled approach. The researcher suggestively recovers the existing state-of-the-art mouth shape categorization presentation.

Koller O. et al., (2015) [42] stated that the statistical recognition method recognizes a wide vocabulary of continuous sign language across several signers using the Constrained Maximum Likelihood Linear Regression (CMLLR) technique. Experimental results demonstrate the significance of tracking for the detection of hand and face landmarks in sign language. This project attempts to provide novices to the field a place to start.

There is a wide range of authors who used the technique and presented their discoveries as shown in Table 1.

Table 1: Summarize table of compared reviewed literature.

Author	Objective	Technique Used	Outcomes	Research Gap
Huang Y. et al., (2020) [23]	**Fake News Detection Using a Self-Adaptive Harmony Search Algorithm-Based Ensemble Learning Model.**	**Deep Learning Model**	**Ensemble learning could yet be improved to solve the problem of cross-domain intractability.**	**A more thorough examination of grammar would provide more important information for preprocessing, resulting in improved false news detection accuracy.**
Ibrahim N. et al., (2020) [24]	**The inserting of a CNN into an HMM from start to finish, including interpreting the CNN outputs in a Bayesian context.**	**Hybrid CNN-HMM**	**Research demonstrated that using ensembles of hybrid CNN-HMMs could improve performance even more.**	**It intends to broaden its approach to include all essential modalities.**
Borg M. et al., (2020) [25]	**Two-stage ASLR based on DL A typical computer vision pipeline or Open Pose hand key points could be used as inputs to the system.**	**Deep Learning, Factorization Algorithm, and Recurrent Neural Networks.**	**Several recent efforts using DL in ASLR have been outperformed by System by a margin of 10.2 percent or more.**	**It would be expanding SU RNNs to grip non-manual indications like facial language and then mouthing.**
Zhang Y. et al., (2019) [26]	**Sentiment Classification at the Sentence Level Using Enhanced Three-Way CNN**	**Traditional feature-based and Deep Learning**	**On 4 benchmark datasets, 3W-CNN performed well.**	**To identify additional actual assurance purposes and to test the system on various models to determine its efficiency and applicability.**
Wei C. et al., (2019) [27]	**Multi-class classifier deep architecture for continuous recognition of sign language.**	**NGC and WIC**	**Extensive testing on a large-scale benchmark confirmed the efficiency of the recommended techniques.**	**The effectiveness of grammatical rule-based classifiers was not significant.**
Mittal A. et al., (2019) [28]	**The continuous SLR or modified LSTM model for continuous sequences of signs identifies a series of related movements.**	**LSTM**	**Signed sentences have an accuracy rate of 72.3 percent while isolated sign words have an accuracy rate of 89.5 percent.**	**For better model learning, additional training data could be included to increase recognition performance.**
Kındıroglu A. et al., (2019) [29]	**TAF was used to represent and recognize solitary sign language actions.**	**Deeplab v3+**	**TAF representations were an effective strategy for SLR in the Bosphorus Sign dataset.**	**To enhance the accuracy by 3% to 81.37 percent.**
Farag I. et al., (2019) [30]	**A new technique for segmenting continuous motion data for use with Japanese Sign Language continuous sentence phrases.**	**CNN**	**Static and dynamic motion phases could be differentiated by the algorithm, which achieves segmentation accuracy comparable to that of earlier state-of-the-art methods.**	**It should be tested on various data sets to see how it performs on other sign languages and data types.**
Pu J. et al., (2018) [31]	**Deep neural architecture with iterative optimization for real-world continuous recognition of signed languages.**	**3D-ResNet**	**It was tested on a large continuous sign language standard dataset from RWTH-PHOENIX.**	**The framework would be used to collect data on numerous hands, facial expressions, and lip motions.**
Zheng L. et al., (2017) [32]	**A whole assessment of deep learning-based SLR methods.**	**Deep Learning**	**The latest breakthroughs in deep learning for SLR have made a significant influence on the sign**	**To increase SLR accuracy and efficiency.**

			language investigation community.	
Camgoz N. et al., (2017) [33]	A revolutionary deep learning strategy to simultaneously align and recognize difficulties.	Sequence-to-Sequence Learning	The network generalizes effectively using SubUNets that learn intermediate representations.	Researchers should look towards hierarchical SubUNets, where each professional scheme contains lower-level professional schemes.
Cui R. et al., (2017) [34]	A deep framework for CSLR.	RNN and CNNWIC	Authors update the order learning model's feature depictions and then develop a regularization weakly managed discovery network.	To make the algorithm multi-modal to include more signals that would help it progress even further.
Pigou L. et al., (2017) [35]	In CSL and gesture language movies, deep residual networks could learn patterns with essentially little preprocessing and the usage of conventional RGB cameras.	CNN	With the Corpus NGT, framewise precision was 73.3 percent, but with the Corpus VGT, framewise accuracy was 55.7 percent.	The last suggestion would be to use language classical in the SLR scenario since projected glosses tend to be connected.
Kumar P. et al., (2017)(1) [36]	A unique multi-sensor fusion outline for SLR	CHMM	With CHMM, the greatest recognition precision has been as extreme as 90.80 percent.	The architecture could be used to collect data on numerous hands, facial expressions, and lip motions.
Kumar P. et al., (2017)(2) [37]	Sensor-based SLR in a new multimodal framework.	HMM and BLSTM-NN	Using a combination of HMM and BLSTM-NN for single-handed and double-handed recognition, overall accuracy rates of 97.85 percent and 94.55 percent have been achieved.	To develop the framework, capture facial emotions, and lip movements.
Li K. et al., (2016) [38]	In CSLR, a novel method to represent transition information between signals has been developed SLR.	HMM	It is possible to create real-world SLR systems using the suggested SLR architecture.	It is now possible to build large-language CSLR systems using huge sign language corpora and meaningful n-gram modeling.
Gattupalli S. et al., (2016) [39]	Deep Learning-based Pose Estimation for SLR is being evaluated.	Deep Learning and CNN	By doing user-independent tests on the dataset, researchers assess the presentation of two deep learning-based stance approximation approaches.	The dataset could be used to evaluate current and future techniques.
Koller O. et al., (2016) [40]	End-to-end inserting of a CNN into an HMM with Bayesian output interpretation.	CNN and HMM	For CSLR, the state of the art on 3 benchmark tasks varies from 15% to 38% relative and up to 13.3% absolute.	All relevant modalities should be included in the scope of the approach.

Koller O. et al., (2015) [41]	**Deep convolutional neural networks are used to construct robust mouth forms in the situation of SLR.**	**Deep Convolutional Neural Network**	**Over the present state of the art, the researchers obtain a considerable increase in mouth shape categorization performance.**	**Estimation accuracy varies during CNN training with poorly supervised frame labels.**
Koller O. et al., (2015) [42]	**A statistical technique for continuous sign language identification across several signers using a big vocabulary.**	**CMLLR**	**For sign language identification, the relevance of following hands and facial landmarks is shown experimentally.**	**To increase recognition via the use of class language models.**

Comparison and Discussion

In this study, researchers have reviewed some techniques and come out with the accuracy of these techniques. Mittal A. et al., (2019) [28] used the LSTM technique, the LSTM is a more sophisticated RNN (sequential network) that enables data to persist. The author used this technique and proposed that the accuracy was 72.3%. Pigou L. et al., (2017) [35] used CNN – Corpus NGT, the achievable target of the Corpus NGT is to collect and store video data in the Dutch sign language ensuring internet access to all interested parties and long-term availability for all participants. A modified technique suggested the accuracy was 73.3%. Pigou L. et al., (2017) [35] used CNN – Corpus VGT, which gives an accuracy was 55.7%. Kumar P. et al., (2017)(1) [36] used CHMM, it enhances the fundamental HMM paradigm by assuming different but correlated state sequences underneath the observed variables, thereby "coupling" the state processes. Using this technique accuracy has come to 90.80%. Kumar P. et al., (2017)(2) [37] used HMM, which is a numerical Markov model in which the represented scheme is thought to be a Markov procedure with unobservable ("hidden") conditions and the accuracy was 97.85%. Kumar P. et al., (2017)(2) [37] used BLSTM-NN, the neural network (NN) that enabled neural network-based analysis of sequences with long-distance dependencies and the accuracy was 94.55%. The significant model discovered during studies was based on HMM having an accuracy of 97.85% as presented in Table 2.

Table 2: Evaluation of existing studies

Author	Year	Technique	Accuracy
Mittal A. et al., [28]	**2019**	**LSTM**	**72.3%**
Pigou L. et al., [35]	**2017**	**CNN – Corpus NGT**	**73.3%**
Pigou L. et al., [35]	**2017**	**CNN – Corpus VGT**	**55.7%**
Kumar P. et al., (1) [36]	**2017**	**CHMM**	**90.80%**
Kumar P. et al., (2) [37]	**2017**	**HMM**	**97.85%**
Kumar P. et al., [37]	**2017**	**BLSTM-NN**	**94.55%**

The analysis of results obtained and studied through Table 2 is shown in Figure 6. Graphical analysis shows that HMM technique used by Kumar P. et al., (2017) [37] presents an accurate and feasible option that can be used for sign language prediction using the N-gram model in real life.

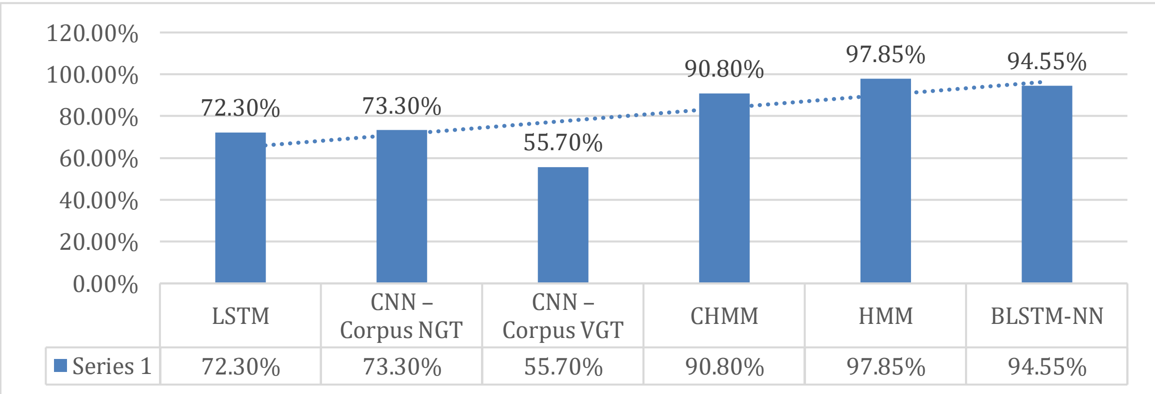

Figure 6: Graphical analysis of previous studies.

4. Conclusion

Researchers have worked hard to create SLR methods that would enable the spoken expression of this CSLR language. Speech notation is critical to the creation of any sign language translation system. Non-manual assistance and universality are critical components of any sign language speech notation. In every sign language, facial expressions are critical, and hence the non-manual component of the sign language is essential. This study presented a deep learning-based self-adaptive technique for SLR and prediction using the N-gram model. Considering different techniques namely, LSTM, CNN-Corpus

NGT, CNN-Corpus VGT, CHMM, HMM, and BLSTM-NN have been reviewed. Even though non-manual assistance is present in many speech notations the recognition has been carried out by P. et al., (2017) [37] using HMM, which seems to be the most accurate technique having an accuracy of 97.85% using HMM approach. Finally, by combining data acquired using HMM, we have achieved better accuracy. In the future, the framework might be utilized to capture multiple-hand data combined with facial expressions and lip motions for emotion-based gesture recognition.

REFERENCES

1. World Health Organization Deafness and hearing loss. http://www.who.int/mediacentre/factsheets/fs300/en/, 2021.
2. Start ASL. Spanish sign language. https://www.startasl.com/spanish-sign-language-ssl/, 2021.
3. Ethnologue. Persian sign language. https://www.ethnologue.com/language/psc, 2021.
4. Rastgoo, Razieh, Kourosh Kiani, Sergio Escalera, and Mohammad Sabokrou. "Sign language production: A review." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3451-3461. 2021.
5. Elakkiya, R. "3 SU subunit sign modeling for continuous sign language recognition" Anna University (2018): 1-34.
6. Necati Cihan Camgoz, Simon Hadfield, Oscar Koller, Hermann Ney, and Richard Bowden. Neural Sign Language Translation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
7. Triantafyllos Afouras, Joon Son Chung, Andrew Senior, Oriol Vinyals, and Andrew Zisserman. Deep Audio-visual Speech Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2018.
8. Camgoz, Necati Cihan, Oscar Koller, Simon Hadfield, and Richard Bowden. "Sign language transformers: Joint end-to-end sign language recognition and translation." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 10023-10033. 2020.
9. Kudrinko, K.; Flavin, E.; Zhu, X.; Li, Q. Wearable Sensor-Based Sign Language Recognition: A Comprehensive Review. IEEE Rev. Biomed. Eng. 2020, 14, 82–97.
10. Nimisha, K.; Jacob, A. A brief review of the recent trends in sign language recognition. In Proceedings of the IEEE 2020 International Conference on Communication and Signal Processing (ICCSP), Virtual, 16–18 March 2021; pp. 186–190.
11. Papastratis, Ilias, Christos Chatzikonstantinou, Dimitrios Konstantinidis, Kosmas Dimitropoulos, and Petros Daras. "Artificial intelligence technologies for sign language." Sensors 21, no. 17 (2021): 5843.
12. Pigou, Lionel, Sander Dieleman, Pieter-Jan Kindermans, and Benjamin Schrauwen. "Sign language recognition using convolutional neural networks." In European Conference on Computer Vision, pp. 572-578. Springer, Cham, 2014.
13. Hutchinson, Jessica. "Literature review: analysis of sign language notations for parsing in machine translation of SASL." Rhodes University, Grahamstown, South Africa (2012).
14. L. Goyal and V. Goyal, "Development of Indian Sign Language Dictionary using Synthetic Animations," Indian J. Sci. Technol., vol. 9, no. 32, pp. 1–5, 2016.
15. Y. Bouzid and M. Jemni, "TuniSigner: A virtual interpreter to learn signwriting," in 14th International Conference on Advanced Learning Technologies, 2014, pp. 601–605.
16. Dhanjal, Amandeep Singh, and Williamjeet Singh. "Comparative analysis of sign language notation systems for Indian sign language." In 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), pp. 1-6. IEEE, 2019.
17. Goyal, Lalit, and Vishal Goyal. "Development of Indian sign language dictionary using synthetic animations." Indian Journal of Science and Technology 9, no. 32 (2016): 1-5.
18. Goyal, Lalit. "Review and Comparison of." An International Journal of Engineering Sciences issue June (2015): 70-80.
19. Stiehl, D., L. Addams, Luiz S. Oliveira, Cayley Guimarães, and A. S. Britto. "Towards a SignWriting recognition system." In 2015 13th International Conference on Document Analysis and Recognition (ICDAR), pp. 26-30. IEEE, 2015.
20. Iatskiu, Carlos EA, Laura Sánchez García, Rafael Dos Passos Canteri, and Diego Roberto Antunes. "The low use of signwriting computational tools from hci perspective." In International Conference on Universal Access in Human-Computer Interaction, pp. 373-382. Springer, Cham, 2015.
21. Hiemstra, Djoerd. "N-Gram models." In Encyclopedia of Database Systems. Springer, 2017
22. Banasiak, Dariusz, Jarosław Mierzwa, and Antoni Sterna. "Extended N-gram Model for Analysis of Polish Texts." In International Conference on Man–Machine Interactions, pp. 355-364. Springer, Cham, 2017.
23. Huang, Yin-Fu, and Po-Hong Chen. "Fake news detection using an ensemble learning model based on self-adaptive harmony search algorithms." Expert Systems with Applications 159 (2020): 113584.
24. Ibrahim, Nada B., Hala H. Zayed, and Mazen M. Selim. "Advances, challenges and opportunities in continuous sign language recognition." Journal of Engineering and Applied Sciences 15, no. 5 (2020): 1205-1227.
25. Borg, Mark, and Kenneth P. Camilleri. "Phonologically-meaningful subunits for deep learning-based sign language recognition." In European Conference on Computer Vision, pp. 199-217. Springer, Cham, 2020.
26. Zhang, Yuebing, Zhifei Zhang, Duoqian Miao, and Jiaqi Wang. "Three-way enhanced convolutional neural networks for sentence-level sentiment classification." Information Sciences 477 (2019): 55-64.
27. Wei, Chengcheng, Wengang Zhou, Junfu Pu, and Houqiang Li. "Deep grammatical multi-classifier for continuous sign language recognition." In 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), pp. 435-442. IEEE, 2019.
28. Mittal, Anshul, Pradeep Kumar, Partha Pratim Roy, Raman Balasubramanian, and Bidyut B. Chaudhuri. "A modified LSTM model for continuous sign language recognition using leap motion." IEEE Sensors Journal 19, no. 16 (2019): 7056-7063.
29. Kindiroglu, Ahmet Alp, Ogulcan Ozdemir, and Lale Akarun. "Temporal accumulative features for sign language recognition." In 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), pp. 1288-1297. IEEE Computer Society, 2019.

30. Farag, Iva, and Heike Brock. "Learning motion disfluencies for automatic sign language segmentation." In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 7360-7364. IEEE, 2019.
31. Pu, Junfu, Wengang Zhou, and Houqiang Li. "Dilated convolutional network with iterative optimization for continuous sign language recognition." In IJCAI, vol. 3, p. 7. 2018.
32. Zheng, Lihong, Bin Liang, and Ailian Jiang. "Recent advances of deep learning for sign language recognition." In 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pp. 1-7. IEEE, 2017.
33. Cihan Camgoz, Necati, Simon Hadfield, Oscar Koller, and Richard Bowden. "Subunets: End-to-end hand shape and continuous sign language recognition." In Proceedings of the IEEE international conference on computer vision, pp. 3056-3065. 2017.
34. Cui, Runpeng, Hu Liu, and Changshui Zhang. "Recurrent convolutional neural networks for continuous sign language recognition by staged optimization." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7361-7369. 2017.
35. Pigou, Lionel, Mieke Van Herreweghe, and Joni Dambre. "Gesture and sign language recognition with temporal residual networks." In Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 3086-3093. 2017.
36. Kumar, Pradeep, Himaanshu Gauba, Partha Pratim Roy, and Debi Prosad Dogra. "Coupled HMM-based multi-sensor data fusion for sign language recognition." Pattern Recognition Letters 86 (2017): 1-8 (1).
37. Kumar, Pradeep, Himaanshu Gauba, Partha Pratim Roy, and Debi Prosad Dogra. "A multimodal framework for sensor-based sign language recognition." Neurocomputing 259 (2017): 21-38 (2).
38. Li, Kehuang, Zhengyu Zhou, and Chin-Hui Lee. "Sign transition modeling and a scalable solution to continuous sign language recognition for real-world applications." ACM Transactions on Accessible Computing (TACCESS) 8, no. 2 (2016): 1-23.
39. Gattupalli, Srujana, Amir Ghaderi, and Vassilis Athitsos. "Evaluation of deep learning-based pose estimation for sign language recognition." In Proceedings of the 9th ACM international conference on PErvasive technologies related to assistive environments, pp. 1-7. 2016.
40. Koller, Oscar, O. Zargaran, Hermann Ney, and Richard Bowden. "Deep sign: Hybrid CNN-HMM for continuous sign language recognition." In Proceedings of the British Machine Vision Conference 2016. University of Surrey, 2016.
41. Koller, Oscar, Hermann Ney, and Richard Bowden. "Deep learning of mouth shapes for sign language." In Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 85-91. 2015.
42. Koller, Oscar, Jens Forster, and Hermann Ney. "Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers." Computer Vision and Image Understanding 141 (2015): 108-125