

A Review on Automatic Sign Language Detection in the Field of Image Processing

Shrikant Singh¹, Devendra Rewadikar², Ankur Taneja³

PG Student, Department of Computer Science & Engineering, SAM College of Engineering & Technology, Bhopal, Madhya Pradesh, India¹

Assistant Professor, Department of Computer Science & Engineering, SAM College of Engineering & Technology, Bhopal, Madhya Pradesh, India^{2,3}

ABSTRACT: Sign language is a gesture based expressive language that is used for communicating people those who cannot communicate verbally. Sign languages have their own linguistic structure, grammar and features, and are independent of the rules that govern spoken languages. They are visual languages; those who rely on hand gestures as well as physical and facial expressions. Sign language may differ in different countries, so it is important to know the visual traits. Indian sign language is one of them where deaf people used it for removing the communication barrier. There are lots of researches have been done in this field that may be effective or may not. Most of the researcher uses Convolutional Neural Network, Support Vector Machine, Hidden Markov Model and many more for recognizing gestures effectively. The applications of automatic sign language recognition are diverse: sign-to-text translation, assistive technology for hearing-impaired individuals, teaching system for sign language students, among others.

KEYWORDS: Automatic Sign Language Recognition, Support Vector Machine, Convolutional Neural Network, Hidden Markov Model, Machine Learning.

I. INTRODUCTION

There are two general categories for sorting hand gesture detection methods, depending on whether they are vision-based or include data gloves. The latter record the presence sensors that transfer physical signals into electrical signals that describe hand posture, or simply wearable property that helps track the subject's hands through color segmentation. Hand segmentation is primarily concerned with skin detection and normalization, commonly tackled with parametric (Gaussian modelling) and non-parametric (histograms) techniques.

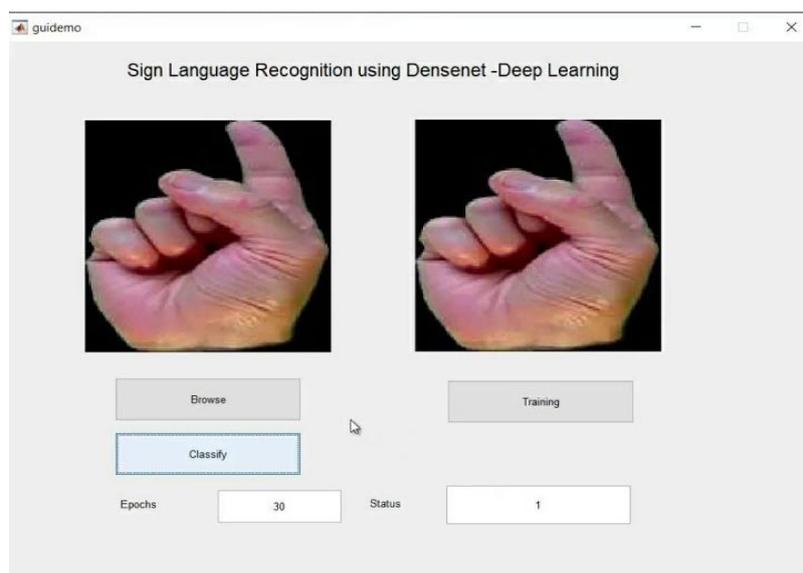


Fig. 1. Automatic Sign Language Recognition [1]

Automatic sign language recognition can be implemented using various methods that intend to classify skin color, hand gesture either using skeletons or foreground extraction methods. Sign language is employed by deaf and hard hearing people to exchange information between their own community and with people. Computer recognition of signing deals from sign gesture acquisition and continues till text/speech generation. Sign gestures are often classified as static and dynamic. However static gesture recognition is easier than dynamic gesture recognition but both recognition systems are important to the human community. The signing recognition steps are described during this survey. The info acquisition, data preprocessing and transformation, feature extraction, classification and results obtained are examined.



Fig. 2. Common Sign Language Gestures [2]

Sign language may be a visual-gestural language employed by deaf and hard-hearing people for communication purposes. Three dimensional spaces and therefore the hand movements are used (and other parts of the body) to convey meanings. it's its own vocabulary and syntax which is only different from spoken languages/written language. Spoken languages use the oratory faculties to supply sounds mapped against specific words and grammatical combinations to convey meaningful information. Then the oratory elements are received by the auditory faculties and processed accordingly. signing uses the visual faculties which is different from speech . speech makes use of rules to supply comprehensive messages; similarly signing is additionally governed by a posh grammar. a symbol language recognition system consists of a simple , efficient and accurate mechanism to rework signing into text or speech. The computerized digital image processing and a good sort of classification methods wont to recognize the alphabet flow and interpret signing words and phrases. signing information are often conveyed using gestures of hands, position of head and body parts. Four essential components during a gesture recognition system are: gesture modeling, gesture analysis, gesture recognition and gesture-based application systems [3].

II. RELATED WORK

There are several researches have been done in the field of automatic sign language recognition. This paper is intended to review those researches and find out the common challenges with them. The accuracy of the system is often important because correct recognition can better emulate the system with people and incorrect recognition may distort the communication. Bhumika Gupta et al. [4] proposed a system which is based on K-nearest correlation for extracting features. System tried to recognize all alphabets by their gestures using double hand. But Classification method is generally used for classifying regular or irregular objects from an image to better predict the object model. Hand gestures or sign languages are training based data models where supervised learning is important for precise recognition. Segmentation is not a method that can built sign language recognition system because it requires deep analysis what gesture it belongs.



Fig. 3. K-Nearest Correlated Classification [4]

The proposed technique seeks to properly classify the Indian signing gestures from static images by first categorizing them into single-handed or double-handed then employing a fusion of HOG and SIFT descriptors for a classification algorithm. The correlation of those features is employed with a KNN classifier and both sorts of features are passed thereto together. This system, though gives substantial accuracy for recognition of this gestures, has its limitations. The technique is currently restricted to detect static gestures. An extension would be the event of a model that recognizes dynamic gestures also [4]. G.Anantha Rao et al. [5] proposed a system which is based on CNN. The system depends upon the training samples and machine iterations. The weakness of CNNs lays within the amount of knowledge you provide to them. If you provide them with less, expect the CNNs to perform poorly. CNNs have many parameters and with small dataset, would run into an over-fitting problem because they needs massive amount of knowledge to quench the thirst. So, you give much data, CNNs is more strong and willing to offer you better performance, but when you give less data then CNNs becomes extremely weak. Large dataset may increases the time and space complexity that may requires much time to execute and respond poorly. Snehal Madhukar Daware et al. [6] proposed a system which is based on morphological operations. Morphological operation is an image processing tool that can work with grayscale image or binary image. It filters the image either by eliminating the unnecessary pixels that is called erosion or filling pixels where it loses that is called dilation. But by using this operation or algorithm a complete gesture cannot be portrayed because it can only have active pixels and erode the rest one, it means that it transform or segment an image in black and white pixels (either 0 or 1) that does not dignify the hand gesture.



Fig. 4. K-convex hull algorithm [6]

Distinctive calculations are performed for the extraction of the attributes of hand motion acknowledgment framework. This incorporates the K convex hull for finger recognition, unconventionality, extending of the objects. Experimental results portray that convex K-shell method gives more exact fingertip recognition. Moreover,

different calculations are performed to urge more data precision within the acknowledgment framework. Feature extrication and arrangement is processed by extricating frames from stored video [6].

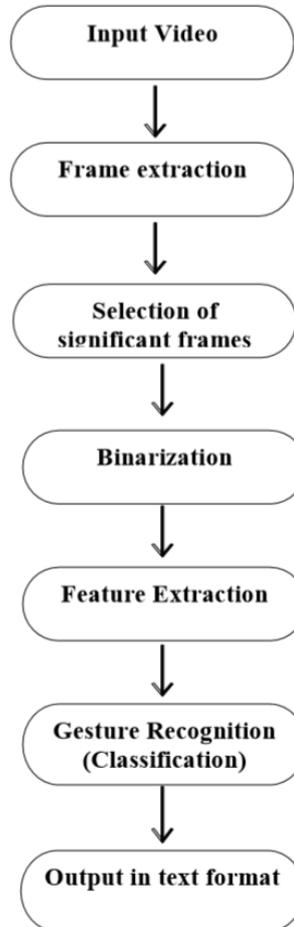


Fig. 5. Flow Chart [6]

Muthu Mariappan H et al. [7] proposed a system which is based on preprocessing and classification techniques such as morphological operations and fuzzy classification resp. Clustering is known as the process of grouping of similar data items together, while the items in the other clusters are as dissimilar as possible. In fuzzy clustering, the data items may belong to more than one cluster. Among several fuzzy clustering algorithms, fuzzy c-means clustering (FCM) algorithm is used most widely, and this can be used for both supervised learning and unsupervised learning, depending upon the needs. Clustering after erosion may erode the useful information from an image that may distort the sign language gesture which results inaccuracy.



a)

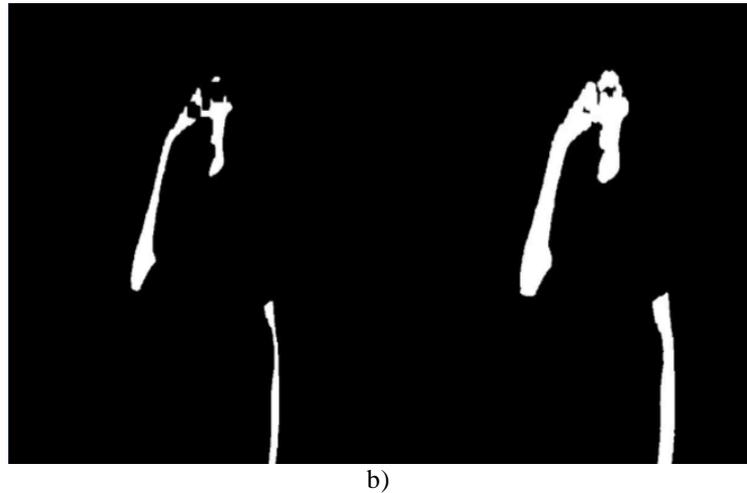


Fig. 6. a) Hand Posture RGB Image b) Erosion (left) and Dilation (right) [7]

K. Revanth et al. [8] proposed a system which is based on Support Vector Machine for classifying sign language gestures from an image. System also uses skin masking technique for segmenting region of interest and eliminating the background by focusing the skin color. For skin segmentation; system uses OpenCV. On perceptual tasks (vision, speech and so on), they are massively outclassed by deep neural networks. On structured data, they are outperformed by gradient boosted trees. Other disadvantages include not giving class probabilities and being rather cumbersome for multiclass problems (you need one model per class).

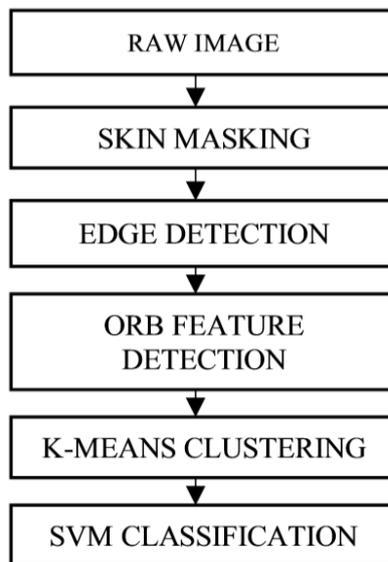


Fig. 7. Work Flow [8]

Sruthi C. et al. [9] proposed a system which is based on deep learning methodology i.e. CNN. This work addresses Indian signing static alphabet recognition problem with a vision based approach. A Convolutional Neural Network (CNN) which may be a deep learning technique is employed to make a model named signet, which may recognize signs, supported supervised learning on data the entire process are often divided into CNN training and model testing. A concentric layer consists of a set of matrices, multiplied by the output of the previous layer in a process called convention, in which these features may be core features to detect certain features (such as edges, Color grade or pattern) or complex one (e.g. shape, nose, normal) or mouth). So, those matrices are called filters or kernels. Sign language recognition is a complex issue because it may have similar gestures for different sign languages. A strong learning module is required to achieve the target with high level of accuracy.

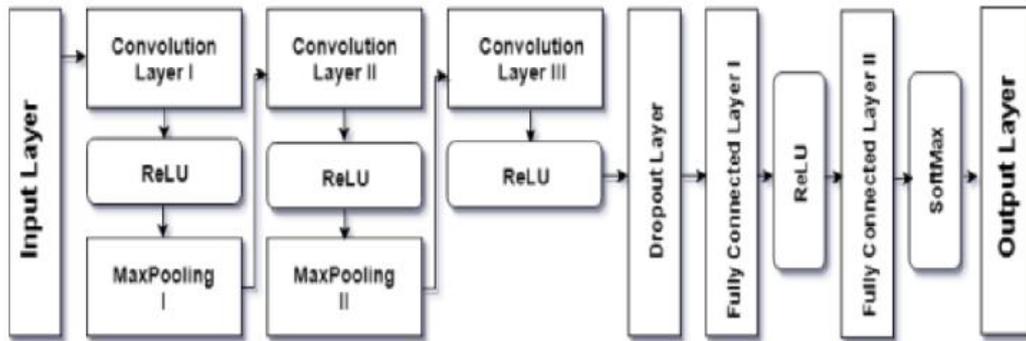


Fig. 8. Diagrammatic Representation of Proposed Signet Architecture [9]

Salma Hayani et al. [10] proposed a system that uses Convolutional Neural Network (CNN) for recognizing sign language gestures. System recognizes Arabic sign languages of different categories of letters and numbers. Most of the signs may have similar gestures that seriously confuse the system to obtain the correct result. Template matching may return false result when machine has not been trained with sufficient samples and system is not effective for complex gestures and back propagation is not effectively work with CNN that is why the correct gesture recognition rate may degrade due to ineffectiveness. CNN do not encode the position and orientation of the object into their predictions. They completely lose all their internal data about the pose and the orientation of the object and they route all the information to the same neurons that may not be able to deal with this kind of information. Accuracy is often important for proper communication.

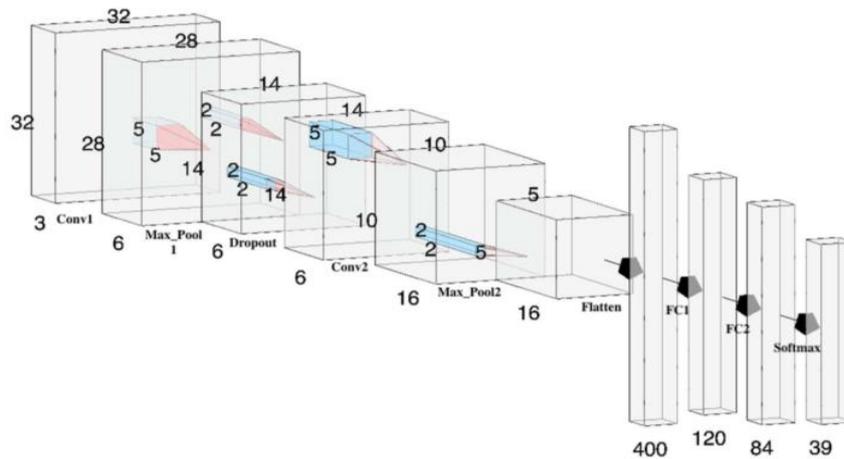


Fig. 9. Proposed System Architecture [10]

III. CONCLUSION & FUTURE SCOPE

This paper is successfully reviewed some systems which have been implemented till now where there are some flaws over there. Most of the system uses machine learning methods to train the system with various samples for a single hand gesture. But a large dataset can consume the large amount of memory that increases the execution time where it is very important to communicate as earlier as possible with high level of accuracy. Some of the system uses morphological operation which is a weak technique for recognizing any sign language gesture because a gesture may have various distinguish internal edges that reflect those posture uniquely but eroding the image may eliminate or mask the gesture completely that increases the incorrect recognition that indirectly affect the precision rate. In future a system can be developed that may have good accuracy rate with less false alarm or recognition that acquire less execution time.

REFERENCES

- [1] https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3D6ScSDs5MUIM&psig=AOvVaw22_IPDLStdBEATyp4g0pIj&ust=1590135603647000&source=images&cd=vfe&ved=0CAIQjRxqFwoTCIjGrffCxOkCFQAAAAAdAAAAABAD
- [2] <https://towardsdatascience.com/sign-language-recognition-in-pytorch-5d72688f98b7>
- [3] https://www.researchgate.net/publication/262187093_Sign_language_recognition_State_of_the_art

- [4] B. Gupta, P. Shukla and A. Mittal, "K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion," 2016 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, 2016, pp. 1-5, doi: 10.1109/ICCCI.2016.7479951.
- [5] G. A. Rao, K. Syamala, P. V. V. Kishore and A. S. C. S. Sastry, "Deep convolutional neural networks for sign language recognition," 2018 Conference on Signal Processing And Communication Engineering Systems (SPACES), Vijayawada, 2018, pp. 194-197, doi: 10.1109/SPACES.2018.8316344.
- [6] Daware, Snehal & Kowdiki, Manisha. (2018). Morphological Based Dynamic Hand Gesture Recognition for Indian Sign Language. 343-346. 10.1109/ICIRCA.2018.8597417.
- [7] H. Muthu Mariappan and V. Gomathi, "Real-Time Recognition of Indian Sign Language," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862125.
- [8] K. Revanth and N. S. M. Raja, "Comprehensive SVM based Indian Sign Language Recognition," 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2019, pp. 1-4, doi: 10.1109/ICSCAN.2019.8878787.
- [9] S. C.J. and L. A., "Signet: A Deep Learning based Indian Sign Language Recognition System," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0596-0600, doi: 10.1109/ICCSP.2019.8698006.
- [10] S. Hayani, M. Benaddy, O. El Meslouhi and M. Kardouchi, "Arab Sign language Recognition with Convolutional Neural Networks," 2019 International Conference of Computer Science and Renewable Energies (ICCSRE), Agadir, Morocco, 2019, pp. 1-4, doi: 10.1109/ICCSRE.2019.8807586.
- [11] A. Subhash Chand, A. S. Jalal and R. Kumar Tripathi, "A survey on manual and non-manual sign language recognition for isolated and continuous sign," Applied Pattern Recognition, 2016.
- [12] H. Cooper, B. Holt and R. Bowden, "Sign language recognition," Springer, 2011.
- [13] S. Ong and S. Ranganath, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005.
- [14] M. Al-Rousan and M. Hussain, "Automatic recognition of Arabic sign language finger spelling," International Journal of Computers and Their Applications, 2001.
- [15] K. Assaleh and M. Al-Rousan, "Recognition of Arabic sign language alphabet using polynomial classifiers," EURASIP Journal on Applied Signal Processing, 2005.
- [16] O. Al-Jarrah and A. Halawani, "Recognition of gestures in Arabic sign language using neuro-fuzzy systems," Artificial Intelligence, 2001.
- [17] T. Shanableh, K. Assaleh and M. Al-Rousan, "Spatio-temporal feature-extraction techniques for isolated gesture recognition in Arabic sign language," IEEE Transactions on Systems, Man, and Cybernetics Part B (Cybernetics), 2007.
- [18] M. AL-Rousan, K. Assaleh and A. Tala'a, "Video-based signerindependent Arabic sign language recognition using hidden Markov models," ELSEVIER, 2009.
- [19] M. Maraqa, F. Al-Zboun, M. Dhyabat and R. Abu Zitar, "Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks," Intelligent Learning Systems and Applications, 2012
- [20] R. Alzohairi, R. Alghonaim, W. Alshehri, S. Aloqeely, M. Alzaidan and O. Bchir, "Image based Arabic Sign Language Recognition System," International Journal of Advanced Computer Science and Applications, 2018.
- [21] Y. LeCun, L. Bottoux, Y. Bengio and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," IEEE, November 1998.
- [22] A. Krizhevsky, I. Sutskever and G. Hinton, "Imagenet classification with deep convolutional neural networks," Proceedings of the 25th international conference on neural information processing systems, 2012.
- [23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint, 2014.
- [24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinov, C. Hill and A. Arbor, "Going Deeper with Convolutions," IEEE Xplore, 2015.
- [25] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2015.
- [26] D. Coomans and D. Massart, "Alternative k-nearest neighbour rules in supervised pattern recognition: Part 1. k-Nearestneighbour classification by using alternative voting rules," AnalyticaChimicaActa 136, 15-27, 1982.
- [27] C. Williams and M. Seeger, "Using the Nyström method to speed up kernel machines," 2001.
- [28] Y-W. Chang, C.-J. Hsieh, K-W. Chang, M. Ringgaard and C-J. Lin, "Training and testing low-degree polynomial data mappings via linear SVM," Journal of Machine Learning Research, april 2010.
- [29] J-P. Vert, K. Tsuda and B. Schölkopf, "A primer on kernel methods," Kernelmethods in computationalbiology, 47, 35-70, 2004.