JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A REVIEW OF VARIOUS APPROCHES FOR HUMAN HAND GESTURE RECOGNITION AND CLASSIFICATION OF SIGN LANGUAGE DATA

[1]R. Siji, [2]Dr. S. Nithya
[1]Research Scholar, [2]Assistant professor
[1]Department of Computer Science, [2]Department of Computer Science
[1]Park's College Tirupur, Tamil Nadu, India, [2]A.V. P College of Arts and ScienceTirupur, Tamil Nadu, India.

Abstract:For the hearing-impaired and deaf society sign language (SL) is a crucial interaction tool. SL, is viewed as a highly organized and ordered form among the different gesture types in the interactive hand gesture taxonomies. People with hearing impairments communicate through signs in visual space rather than through speech or sounds [1]. SL examines non-manual signs, such as facial expressions and other body poses, that transmit semantic information in addition to hand movements. Understanding of SL is one particular area of interest [2,3]. Current methods utilized in research on hand gesture and SL recognition are examined in this study. Hand gestures are utilized as signs, and films and pictures can be used to recognize these motions. Based on the characteristics of the gesture image, hand movements are recognized and categorized. Different approaches for extracting features are used to find them, and machine learning (ML) and deep learning (DL)methods are employed for categorizing them [4]. Reviewing the primary findings of a comparison of feature extraction techniques of comparable systems employed in image-based hand gesture identification according to accuracy rate is the paper's primary contribution. In summary, the work aims to offer readers a thorough overview of automated gesture and SL identification, while also advancing study endeavors in this domain.

***Index Terms*: Classification, feature extraction, hand gesture recognition, sign language recognition, recognition accuracy**.

I. INTRODUCTION

The majority of hearing-impaired individuals, particularly those with moderate to extensive impairments, utilize SL or gestured communication to express meaning instead of employing sounds that can be heard. Communicating with people who share identical SL is possible for a disabled person by employing hand gestures and facial reactions. Individuals utilize hand gestures to convey their ideas and emotions as well as to reaffirm data that is said in casual conversation. A framework of organized hand gestures that combines visual motions and signals for communication is called SL [5]. SL provides the speech-impaired community with practical skills for everyday engagement. Several bodily parts, including the hand, fingers, head, and facial expression, are utilized in SL to convey data. Nonetheless, the hearing community rarely employs SL much, and even fewer can comprehend it. This presents a real obstacle to interaction that has not yet been entirely resolved among the deaf population and the general public [6].

Several factors render hand gesture identification a highly difficult task. First, the system's capacity to process inputs that differ significantly from those employed throughout its creation phase. When developing hand gesture detection infrastructure, factors like noise in the surrounding environment, the linguistic and environmental heterogeneity of signers, and so forth, might not have been takenintoaccount. This is so that the segmentation and tracking processes will go more smoothly. Typically, impose constraints on the signers' surroundings. Managing the transition motions among two signs presents an additional obstacle in the creation of hand gesture identification since it can be challenging to determine when each sign's hand gesture begins and ends [7,8]. A system may recognize something incorrectly or poorly if it is unable to distinguish the limit between two signs. This complexity makes it seem as though many academics have not given continuous SL as much attention as they need for vision-based hand gesture identification, that has restricted applicability in the actual world.

Natural language processing, computer vision, linguistics, and pattern matching are all involved in the joint study field of SL identification. Its goal is to develop different techniques for recognizing signs that have previously been made and interpreting their meaning [9]. Human-computer interface (HCI) oriented SL recognition technologies are made to facilitate productive and interesting interactions. These systems use an integrated approach that includes SL testing, SL linguistics, SL technology, and data collecting. A system like this can be implemented in public spaces like banks, resorts, trains, and offices to help persons with hearing loss understand new ideas and information and manage their emotions.

Currently available SLR techniques generally categorised into two classes: deep neural network-based and traditional machine learning (ML) techniques that rely on image attributes. The earlier method first segments the hand shape from SL images or video utilizing conventional image segmentation methods. Then, it applies ML techniques (like SVM, HMM, and k-NN algorithm) to execute categorization according to the image elements (like HOG and SIFT) [10]. The following are its drawbacks: The representational power of these traits is restricted. The extraction of representative semantic analysis from complicated content is challenging, and the real-time performance of step-by-step gesture detection is not good. The latter technique locates the hand and simultaneously classifies motions by training a target identification neural network on the attributes of the video frame. Deep neural network-based target identification networks frequently achieve maximum accuracy and recognition speed as well as superior immediate efficiency regarding to traditional image analysis and ML techniques, making them the go-to technique for dynamic target identification. Static and dynamic SL identification are other subcategories of SL identification [11]. The previous type does not carry dynamic data; instead, it recognizes gestures by analysing hand posture. The latter consists of a classification issue; it recognizes gestures according to the video series and incorporates hand movements. Although dynamic SL recognition is significant and valuable than static SL recognition, it is far more difficult to accomplish.

II. OVERVIEW OF SIGN LANGUAGE RECOGNITION SYSTEM

There are two stages to the identification method: testing and training. The training dataset applied for learning the classifier during the training phase. The scholar has the option of creating the database from scratch or using one that already exists. To take the training photos, you may utilize an external webcam or the built-in webcam in laptops. The majority of SL detection algorithms categorize signs on hand motions; that is, they do not take into account face expressions. The development of the database, feature extraction, and classifier training are crucial processes in the training phase. Preprocessing, feature extraction, classification, and image/video acquisition are all included in the testing step.

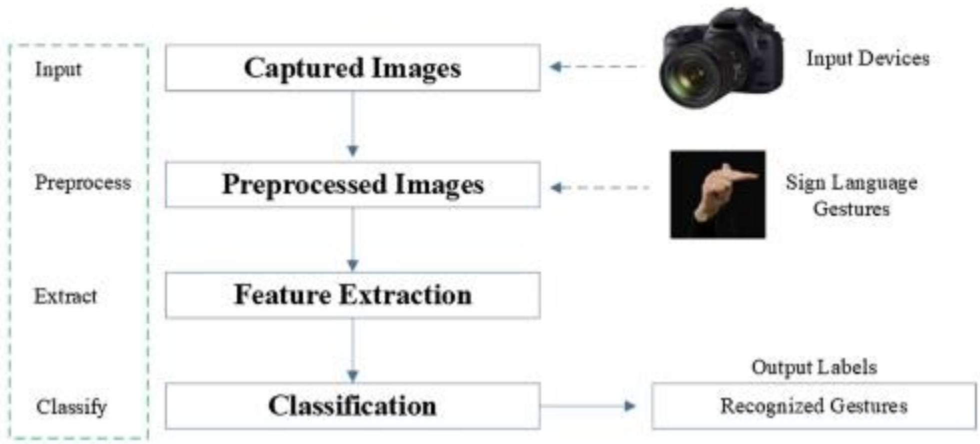

FIGURE 1. SIGN LANGUAGE RECOGNITION SYSTEM

2.1 Preprocessing

To extract the region of interest (ROI), the training images go through a preprocessing step. If only hand motions are factored into account, the ROI can be hands; if facial gestures are taken intoaccount, it can be both hands and face. Preprocessing typically entails segmentation, morphological filtering, picture enhancement, scaling, and filtering. Any of the widely used techniques can be applied to image enhancement and filtering. The segmentation process requires choosing the method that best fits the input photos or video. Frequently employed methods for segmentation include background subtraction, Otsu's thresholding, skin color-based segmentation, and motion-based segmentation [12]. In order to extract the ROI, the test photos or videos are preprocessed within the testing phase.

2.2 Feature Extraction

Although the feature vectors attained from this step serve as the classifier's inputs, feature extraction is the important phases in the detection of SL [13]. Feature extraction algorithms ought to identify shapes with a high degree of reliability and robustness, regardless of variations in the brightness, position, orientation, and size of the item inside a video or image. Pixel collections are utilized for displaying objects in images. The characteristics of these pixel groups must be described in order to recognize objects. There are various methods to acquire the features: Fourier descriptors, wavelet decomposition, Haar wavelets, texture features, orientation histogram, and scale invariant feature function. Principal Component Analysis (PCA) is sometimes used to minimise the dimensionality by employing the ROI pixels as the feature vector. The classifier is trained using the feature vector that was derived in this way utilizing any feature extraction technique.

2.3 Classification:

To categorize the input signs into distinct modules in SL recognition, a classifier is required. Throughout the training phase, the classifier is trained utilizing the feature vector that was retrieved from the training database. The trained classifier identifies the class associated with the sign when a test input is provided, at which point it plays the sound or displays the text. Videos or pictures utilized as test inputs. The prevalent classifiers are K Nearest Neighbor (KNN), Convolution Neural Network (CNN) [14], Deep Neural Network, Fuzzy Systems, Artificial Neural Networks (ANN), Multiclass Support Vector Machines (SVM), Hidden Markov Models (HMM). The recognition rate serves as a gauge for the classifier's effectiveness.

III. Literature Review

Mohandes et al [15] suggested a novel method for ArSLR that makes usage of the recently issued LMC. The hand and fingers are detected by this device, which then tracks them to provide location and motion data. suggest utilizing the LMC as the ArSLR system's mainframe. The system comprises three stages: preprocessing, feature extraction, and classification, in addition to data collecting. Additionally, contrast the MLP neural networks' efficiency with that of the Nave Bayes classifier. When the suggested approach is applied to the Arabic sign alphabets, classification accuracy with NB is 98%, and with the MLP, it is over 99%.

Bodda et al [16] explained the conversion of Word and Alphabet signs. Existing glove-based designs have hardware-dependent problems that are addressed by new methods and algorithms that are presented and put into practice. With distributed processing units throughout, the entire device is made to be modular, promoting modular enhancement, lowering complexity, and enhancing subsystem interdependence. Error correction and gesture recognition both employ decision trees. Additionally, expect that the preceding architecture and design will serve as the foundation for future developments in sensor-based sign language translation study as well as research into smart gloves other cybernetic devices.

Raghuveera et al [17] introduced A successful method that converts intelligible English text and speech from the input hand gesture in IS. The Microsoft Kinect is employed by the system to record hand movements. The dataset that was utilized includes 4600 depth and RGB photos from the Kinect Xbox 360, together with 140 distinct ISL movements from 21 people. These gestures encompass fingerspelling, double-handed signs, and single-handed signs. The hand region is precisely segmented and retrieved utilizing Speeded Up Robust attributes, Histogram of Oriented Gradients, and Local Binary Patterns in order to identify the hand posture. To increase the average detection accuracy to 71.85%, the framework combines three feature classifiers that were trained utilizing SVM. The algorithm converts the series of identified hand signals into the most accurate approximations of valid English phrases.

Islam et al [18] represented a more accurate HGR algorithm built on ASL identification. This method utilizes a mobile video camera to capture ASL gesture images against a black background to extract attributes. Five features are extracted by the system throughout the processing phase: rotation, elongated Ness, eccentricity, fingertip finder, and pixel segmentation. A suggested method for feature extraction essentially integrates the convex hull and K curvature methods. It is known as the "K convex hull" approach and is quite accurate at detecting fingertip. An ANN is trained using thirty feature vectors that correctly identify 37 signs of the American alphabet and numerals. This is done with the assistance of a feed forward and back propagation technique and is useful for HCI systems. This system's overall gesture detection rate in a real-time setting is 94.32%.

Mishra et al [19] introducedapplication of convolutional neural networks in conjunction with ML and computer vision. This work attempts to promote communication between individuals who are deaf or mute and others. A system that combines motion capture and gesture recognition is developed to translate hand motions into voice in order to accomplish this goal. People who are deaf or mute will benefit from this method since it will improve their ability to communicate with others.

Mujahid et al [20] suggesteda small framework for gesture detection without extra image enhancement, image filtering, or preprocessing that depends on YOLO v3 and DarkNet-53 CNN. The suggested algorithm effectively recognized gestures in low-resolution picture mode and attained excellent accuracy even in a complicated context. A labeled dataset of hand motions in both Pascal VOC and YOLO format was utilized to assess the suggested approach. and obtained improved outcomes by employing suggested YOLOv3 centered system to extract hand characteristics and identify hand gestures, with F-1 scores of 96.70%, 94.88, 98.66, and 97.68, for accuracy, precision, and recall.

Kasukurthi et al [21] proposed an illustration to identify letters in American Sign Language from RGB pictures. Before the Deep Neural Network was trained, the training images were scaled and preprocessed. Thus, to enable the model to function with an accuracy of 83.29% on mobile devices, it was trained utilizing a squeeze net framework.

Alzohairi et al [22] developedan automatically identify ArSL alphabets by employing an image-based approach. In particular, several visual cues are examined in order to develop a precise ArSL alphabet recognizer. One-Versus-All SVM receives the retrieved visual descriptors. The findings of the evaluation reveals that the HOG descriptor works better than the other descriptors under consideration. Therefore, the suggested approach uses the ArSL gesture algorithms that One-Versus-All SVM acquires by utilizing HOG descriptors.

Jiang[23] suggesteda technique utilizing parameter-optimized MGSVM and GLCM. The hand outlines were separated from the background after digital camera or keyframes from a video were employed to capture the SL visuals. Every picture was resized to N×N proportions and transformed into a grayscale image. A grayscale image with 256 intensity values was decreased to 8 in order to produce a gray-level co-occurrence matrix. MGSVM received the reduced and extracted features, and a 10-fold cross validation was utilized to perform the categorization. With 450 isolated Chinese SL images across 30 categories, outcomes showed that the GLCM–MGSVM outperformed GLCM-DT in classification accuracy, with an accuracy of 85.3%. Chinese SL was found to be effectively classified by the GLCMMGSVM.

Shin et al [24] identified American sign language by utilizing hand photos captured by a webcam. Two sorts of features were produced from the calculated coordinates of the joints gathered for categorization in this project: the separation among the joint points and the angles among vectors and 3D axes. The media pipe hands method was employed to calculate the hand joints from RGB images of hands captured by a web camera. Light GBM and SVM were the classifiers employed to categorize the characters. The Massey dataset, ASL Alphabet dataset, and finger spelling A dataset were the three character datasets that were utilized for identification. The Massey dataset yielded 99.39% of its findings, the ASL Alphabet dataset yielded 87.60%, and the Finger Spelling A dataset produced 98.45%. The suggested method for automatically recognizing American SL is more efficient than earlier research, doesn't require any unique sensors or equipment, and less expensive.

Wadhawan and Kumar[25] provides reliable simulation of static signs employing CNN inspired by DL for the purpose of recognizing SL. 35,000 sign pictures of 100 static signs in total were gathered from several users. The suggested system's effectiveness is assessed using about 50 CNN simulations. The suggested method has been found to attain the greatest training accuracy of 99.72% and 99.90% on colored and grayscale images, accordingly. The findings are analyzed based on various optimizers. F-score, recall, and precision have all been used to assess the suggested method's efficiency. Additionally, the system shows how effective it is compared to previous efforts that simply took a few hand gestures into account for identification.

Dudhal et al[26] establisheda basic CNN-based technique for recognizing Indian SL. The suggested solution utilized anywhere because it only requires a laptop and webcam. For image classification, CNN is employed. For feature extraction, SIFT is hybridized with Gaussian blur image smoothing and adaptive thresholding. A dataset of 5000 photos, 100 images for each of 50 motions, has been generated in response to the unavailability of the ISL dataset. The Python package known as Keras is used for both system implementation and testing. While CNN with adaptive thresholding obtained 91.84% accuracy, the optional CNN with hybrid SIFT implementation yields 92.78% accuracy.

Alawwad et al [27] introduced a novel ArSL recognition algorithm that uses a Faster R-CNN to localize and recognize the Arabic SL alphabet. In particular, quicker R-CNN is intended to identify the hand's location in a given image by extracting and mapping its features. The segmentation effort meant to identify the hand region and the selection of pertinent characteristics for representing the sign visual descriptors are both made easier by the suggested method. The ResNet-18 and VGG-16 systems were utilized for the development and evaluation of the suggested Faster R-CNN based sign identification framework, and a genuine ArSL picture dataset was gathered. The suggested method produced 93% accuracy and demonstrated the suggested algorithm's resilience to significant background fluctuations in the recorded scenes.

Sharma and Kumar[28] addressed the problems with dynamic ASL identification, 3-D CNNs—a more sophisticated CNN replacement is used to identify patterns in volumetric data, such as videos. Utilizing the Boston ASL LVD dataset, which consist over 3300 English words stated by six distinct signers, CNN is trained to categorise 100 words. The framework is tested utilizing the remaining 30% of the dataset, with the remaining 70% being utilized for training. The suggested effort works better than the current modern algorithms of f-measure (3.9%), recall (4.3%), and precision (3.7%). The suggested work can be employed in real-time applications, as seen by its computing time of 0.19 seconds per frame.

Aparna and Geetha [29] proposed a CNN and LSTM DL system for SL identification. The design passes the pretrained CNN framework for feature extraction to the LSTM for spatiotemporal data capture. To improve accuracy, a further LSTM is stacked. Less temporal information is captured by the DL approach. Fewer studies deal with the detection of SL utilizing DL systems like CNN and LSTM. A dataset of ISL was utilized to test the method and after testing with the ISL dataset, they reported their success evaluation. Studies on DL algorithms capturing temporal information indicates that this is still an unsolved issue.

A thorough synopsis of all the previously examined work is presented in a table, outlining the approaches, benefits, and drawbacks.

Table 1: Comparison Table for Recognition and Classification of Sign Language Data with Existing Methods

Author	Methods	Merits	Demerits
Mohandes et al [15]	Multilayer Perceptron (MLP) neural networks	Using the Arabic sign alphabets, the suggested system provides 99% categorization accuracy.	The model is not fit for all type of sign language.
Bodda et al [16]	Decision Tree	Perfect model for gesture recognition and error correction.	Still having hardware-dependent issues.
Raghuveera et al [17]	Histogram of Oriented Gradients and Local Binary Patterns, SVM	Clustering method play vital role in the system.	The framework accuracy (71.85%) has to be increased.
Islam et al [18]	ANN	Perfect model for detect fingertip withhigh accuracy.	Need to focus on real time dataset.
Mujahid et al [20]	YOLO v3 and DarkNet-53 convolutional neural networks.	Noticed gestures even in low-resolution picture mode.	It will focus on validating the performance model.
Kasukurthi et al [21]	Deep Neural Network	It can operate on mobile.	Achieved less accuracy result.
Alzohairi et al [22]	Histograms of Oriented Gradients (HOG),SVM	The model is fit for Arabic Sign Language (ArSL) alphabets.	Getting less accuracy for other types of sign language dataset.
Jiang [23]	Medium Gaussian support vector machine	Extracted features are utilized to improve the	Consume high execution time to extract the

	(MGSVM).	classification part.	features.
Shin et al [24]	SVMand light gradient boosting machine (GBM)	The Massey dataset yielded 99.39% of its findings, the ASL Alphabet dataset yielded 87.60%, and the Finger Spelling A dataset produced 98.45%.	The model is computationally expensive.
Wadhawan and Kumar [25]	DL-based convolutional nneural networks (CNN).	The suggested technique's effectiveness is assessed employing about 50 CNN simulations.	which only a selectnumber hand signals are taken as factors for identification.
Dudhal et al [26]	Scale invariant feature transformation (SIFT),CNN.	High volume of dataset images are possible to use this model.	Achieve improved algorithm accuracy and efficiency.
Alawwad et al [27]	Faster Region-based Convolutional Neural Network (R-CNN), VGG-16 and ResNet-18 models	the suggested algorithm's resilience to significant background fluctuations in the scenes that were collected.	Need to add additional parameters to prove the model.
Sharma and Kumar [28]	3-D CNNs.	The computing time (0.19 sec/frame) is low compare to other model.	Need to additional algorithm to support the classification.
Aparna and Geetha [29]	Convolutional neural network (CNN) and long short-term memory (LSTM)	LSTM is capturing spatio-temporal information for better classification.	Limited sign language dataset is possible to use this model.

IV. Inferences From The Existing Work

The main problem of existing methods of MLmethods and DL methods are high computational cost. Recognizing and comprehending SL is a crucial component of the SL translation work. Research on gesture detection is presently ongoing due of its broad range of applications, including virtual reality communication between humans and computers, remote control robotics, and SL recognition. However, there are still obstacles in the way of developing a reliable and accurate system, including hand occlusion, the existence of affine transformation, database scalability, varying background illumination, and high computing cost.

V. Conclusion And Future Work

Large segments of society utilize SL and other sign-based communication methods often. The DL techniques have led to a notable advance in accuracy in the SL identification field in recent years. The study examined the ML and DL frameworks for SL recognition that have been suggested recently, according to a suggested taxonomy. Recently, a lot of concepts were suggested by scholars. Subsequent research in this field will concentrate on the concept of creating an improved DL framework for the SL classification job using various methods for gesture recognition. It is recommended that future research employing benchmarked databases enable direct comparison of the techniques employed.

Reference

[1]Cui, R. Liu, H. and Zhang, C. 2019. A deep neural framework for continuous sign language recognition by iterative training. IEEE Transactions on Multimedia, 21(7):1880-1891.

[2]Goldin-Meadow, S. and Brentari, D. 2017. Gesture, sign, and language: The coming of age of sign language and gesture studies. Behavioral and brain sciences, 40.

[3]Padmalatha, E. Sailekya, S. Reddy, R.R. Krishna, C.A. and Divyarsha, K. 2019. Sign language recognition. International Journal of Recent Technology and Engineering (IJRTE), 8(3):2128-2137.

[4]Wang, Z. Zhao, T. Ma, J. Chen, H. Liu, K. Shao, H. Wang, Q. and Ren, J. 2020. Hear sign language: A real-time end-to-end sign language recognition system. IEEE Transactions on Mobile Computing, 21(7): 2398-2410.

[5]Koller, O. Zargaran, S. Ney, H. and Bowden, R. 2018. Deep sign: Enabling robust statistical continuous sign language recognition via hybrid CNN-HMMs. International Journal of Computer Vision, 126:1311-1325.

[6]Kumar, P., Gauba, H., Roy, P.P. and Dogra, D.P., 2017. Coupled HMM-based multi-sensor data fusion for sign language recognition. Pattern Recognition Letters, 86 :1-8.

[7]Dignan, C. Perez, E. Ahmad, I. Huber, M. and Clark, A. 2022. An AI-based approach for improved sign language recognition using multiple videos. Multimedia Tools and Applications, 81(24) :34525-34546.

[8]Mohammad, N.A. and Sek, T.K. 2021. Development of a Malaysian Sign Language Interpreter by using Image Recognition Techniques for the Community to Understand the Deaf. Evolution in Electrical and Electronic Engineering, 2(2): 629-636.

[9]Pigou, L. Van Den Oord, A. Dieleman, S.Van Herreweghe, M. and Dambre, J. 2018. Beyond temporal pooling: Recurrence and temporal convolutions for gesture recognition in video. International Journal of Computer Vision, 126 :430-439.

[10]Wong, W.K. Juwono, F.H. and Khoo, B.T.T. 2021. Multi-features capacitive hand gesture recognition sensor: A machine learning approach. IEEE Sensors Journal, 21(6) :8441-8450.

[11]Hu, Y. Wong, Y. Wei, W. Du, Y. Kankanhalli, M. and Geng, W. 2018. A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition. PloS one, 13(10).

[12]Shivashankara, S. and Srinath, S. 2018. American sign language recognition system: an optimal approach. International Journal of Image, Graphics and Signal Processing, 11(8).

[13]Bantupalli, K. and Xie, Y. 2018, December. American sign language recognition using deep learning and computer vision. In 2018 IEEE International Conference on Big Data (Big Data) 4896-4899.

[14]Cui, Q. Zhou, Z. Yuan, C. Sun, X. and Wu, Q.J. 2018. Fast American Sign Language Image Recognition Using CNNs with Fine-tuning. Journal of Internet Technology, 19(7) :2207-2214.

[15]Mohandes, M. Aliyu, S. and Deriche, M. 2014, June. Arabic sign language recognition using the leap motion controller. In 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE) pp. 960-965.

[16]Bodda, S.C. Gupta, P. Joshi, G. and Chaturvedi, A. 2020. A new architecture for hand-worn Sign language to Speech translator. arXiv preprint arXiv:2009.03988.

[17]Raghuveera, T. Deepthi, R. Mangalashri, R. and Akshaya, R. 2020. A depth-based Indian sign language recognition using microsoftkinect. Sādhanā, 45 :1-13.

[18]Islam, M.M. Siddiqua, S. and Afnan, J. 2017, February. Real time hand gesture recognition using different algorithms based on American sign language. In 2017 IEEE international conference on imaging, vision & pattern recognition (icIVPR) pp. 1-6.

[19]Mishra, S.K. Sinha, S. Sinha, S. and Bilgaiyan, S. 2019. Recognition of hand gestures and conversion of voice for betterment of deaf and mute people. In Advances in Computing and Data Sciences: Third International Conference, ICACDS 2019, Ghaziabad, India, April 12–13, 2019, Revised Selected Papers, Part II 3: 46-57.

[20]Mujahid, A. Awan, M.J. Yasin, A., Mohammed, M.A. Damaševičius, R. Maskeliūnas, R. and Abdulkareem, K.H. 2021. Real-time hand gesture recognition based on deep learning YOLOv3 model. Applied Sciences, 11(9).

[21]Kasukurthi, N. Rokad, B. Bidani, S. and Dennisan, D.A. 2019. American sign language alphabet recognition using deep learning. arXiv preprint arXiv:1905.05487.

[22]Alzohairi, R. Alghonaim, R. Alshehri, W. and Aloqeely, S. 2018. Image based Arabic sign language recognition system. International Journal of Advanced Computer Science and Applications, 9(3).

[23]Jiang, X., 2020. Isolated Chinese sign language recognition using gray-level Co-occurrence Matrix and parameter-optimized Medium Gaussian support vector machine. In Frontiers in Intelligent Computing: Theory and Applications: Proceedings of the 7th International Conference on FICTA (2018), 2: 182-193.

[24]Shin, J. Matsuoka, A. Hasan, M.A.M. and Srizon, A.Y. 2021. American sign language alphabet recognition by extracting feature from hand pose estimation. Sensors, 21(17).

[25]Wadhawan, A. and Kumar, P. 2020. Deep learning-based sign language recognition system for static signs. Neural computing and applications, 32 :7957-7968.

[26]Dudhal, A. Mathkar, H. Jain, A. Kadam, O. and Shirole, M. 2019. Hybrid SIFT feature extraction approach for Indian sign language recognition system based on CNN. In Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB) 727-738.

[27]Alawwad, R.A. Bchir, O. and Ismail, M.M.B. 2021. Arabic sign language recognition using faster R-CNN. International Journal of Advanced Computer Science and Applications, 12(3).

[28]Sharma, S. and Kumar, K., 2021. ASL-3DCNN: American sign language recognition technique using 3-D convolutional neural networks. Multimedia Tools and Applications, 80(17) :26319-26331.

[29]Aparna, C. and Geetha, M., 2020. CNN and stacked LSTM model for Indian sign language recognition. In Machine Learning and Metaheuristics Algorithms, and Applications: First Symposium, SoMMA 2019, Trivandrum, India, December 18–21, 2019, Revised Selected Papers 1: 126-134.