SURVEY ON SIGN LANGUAGE RECOGNITION IN MEDICAL DOMAIN

Murugeshwari Adhiappan, Priyadharshini Subramaniyan, Sivasankari Kannan
AI&DS
Panimalar Engineering college Chennai

ABSTRACT

This literature survey mainly focuse on the developed advance technologies to provide more comfortable commication for deaf & dumb with normal people mainly on medical domain. Recognizing the sign language was computational tasks which involves identifying the actions so, it is very important to fill out communication gaps between hearing people with hearing impairments by using various softwares and algorithms such as 2D point based convolution network, Microsoft Kinect , 3D CNN based skeleton networks
,multiclass support vector machine,deep learning, deep convolutional long short-term memory network.This paper provides a brief survey about the various research works carried out in the medical field.

KEY WORDS:

Sign Language, Convolutional Neural Network, Artificial Neural Networks, You Only Look Once Version3 Algorithm, Deep Convolutional Long Short Term Memory Network, Hidden Markov Model, Microsoft Kinect 2, Multi-class Support Vector Machine and Support Vector Machine

INTRODUCTION

In the medical field, sign language detection entails the use of technology to identify and decipher sign language used by patients with speech or hearing problems. This technology can facilitate more effective communication and raise the standard of healthcare by bridging the communication gap between medical staff and patients who use sign language.There are several ways to detect sign language, including utilising computer vision techniques to follow the motion of hands and fingers or by using wearable sensors to record and examine sign language motions. Medical personnel can more effectively communicate with patients thanks to the processing and translation of the acquired data into text or speech.Patients who have hearing or speech problems as well as medical personnel who may find it difficult to communicate with them could benefit significantly from this technology. Patients can receive more precise diagnoses, treatments, and overall care by reducing communication barriers, improving healthoutcomes.

LITERATURE SURVEY

In paper [1] , temporal modelling and classification methods are intended to exhibit sign word detection. This model has Human pose estimation module that is considered as the major module..it is based on a light imaging system which is in Microsoft Kinect v2 sensor that is used for pose estimation. And then we have to use Principal Component Analysis as a separation for facet selection and standardization. The temporal modelling used two methods: DTW otherwise called as Dynamic Time Warping is used to model the secular factors of signs, and k-Nearest Neighbor's algorithms (k-NN) to categorize the samples. Then the regression techniques such as Random Decision Forest(RDF) have been used, Which is based on Temporal Templates(TT).

In paper [2], The algorithmic principle of machine learning, YOLO, is described by the author. The YOLO-V3 object recognition system is built on a traditional neural network. Its key benefit is that it broadcasts more frequently than other networks while still being accurate and of high quality. There are a total of 26 categories that can be assigned to each letter of the English alphabet. the visual materials required for testing and training models. The recognition kernel has the following dimensions: 1 x 1 x (B x (5 + C)). "5" stands for the five bounding box features, "B" for the range of bounding boxes that predict a cell on the feature map, and "C" for the range of categories. "5" also stands for the four bounding box features and one object confidence.The kernel size of this approach is 1 x 1 x 255 because it was trained on a coconut

tree with B = 3 and C = 80. It provides accurate predictions at three scales by dividing the initial image scale by 32, 16, and 8 respectively. In the same manner as before, the data from the previous layer is fused to an Xth convolutional layer. The third and final layer, which results in a feature map with measurements 52 by 52 by 255, is constructed in layer 106. The restoration regression is currently anticipated for the development of YOLO v3.

In [3], the author proposed hybrid model of deep convolutional long short-term memory network for Recognizing the Sign Languages which is mainly for communication to COVID-19 deaf patient in pandemic situation . Here, basically the model was developed with the two deep neural networks architecture such as LSTM for

classification network and CNN for feature extraction networks. Initially, hand gestures Videos belongs to Indian Sign Language words were given to system as input ,Videos of hand gestures were converted into sequences of images as frame then passing on to CNN model for extracting spatial features The architecture diagram proposed transfer learning technique to train the classifier for extracting image features by reusing pre trained deep learning network VGG-16 on the imagenet dataset was retrained and spatial features were extracted from video based frames, thus retraining them by HGR model .Basically VCG-16 is 16 layer CNN model composed of pooling layers which has stride with (2,2) dimension size and the convolutional layers with (3,3) dimension size as convolution filters ,small-size filters are used to provide privilege of reducing computational cost.The architecture network in VGG- 16 consisting of two contiguous blocks which has two max-pooling and convolutional layers ,then three contiguous blocks of the three max-pooling and the convolutional layers were followed .The stacked pooling and convolutional layer which are followed by softmax output layer(1000 neurons) and three fully connected layers(4096 neurons). Hidden layers were implemented with ReLU (Rectified Linear Unit) ,act as an activation functions mainly to avoid gradient problems to be vanished and enhancing the capability of learning processes

In paper [4], For deaf-mute enquiries, use the Heart-Speaker. Heart-Speaker uses MaixBit as the central controller for the entire system. The technology can give real-time sign language interpretation for deaf patients and also play music. According to the doctor's speech, the relevant sign language movie is shown to the patients. Heart- Speaker contains a 1500 mAh lithium battery that can fulfil long-term battery life needs and can be charged by USB Type-C. The device also includes a microwave radar-sensing module, which turns the screen off and enters sleep mode when there is no patient in the detecting area, extending battery life even more. This system identifies the sign language picture information recorded by the camera using the MobileNet-YOLOv3 target detection model, which improves the model's capacity to recognise tiny targets while retaining high real-time performance.

In paper [5], To recognize the action, Convolutional Neural Network is used for Image Stacking, The combination of CNN and LSTM and Long Short Term Memory with Open Pose has been performed to recognize the action by using computer based system.The proposal system is done by four categories ,they are Gathering or collecting the data, Pre-processing the data, creation and Prediction of the module. After implementation comparative has been done between the self dataset.To recognize the sign language many models are used in this paper, the models such that,CNN using Image Stacking : 12 image frames are being arranged and form an image. This single image is defined as input to the Convolutional Neural Network models. CNN+LSTM : CNN models and their characteristics are extricate from the second layer of the model and their characteristics are established to forecast the sequence in the LSTM model. LSTM with OpenPose: in this the model has been divided into four categories:Smart Frame Selection : it is initiated based on choosing a constant from recording and executed into the LSTM or Long Short Term Memory model.LSTM
: this model is created for forecast the temporary sequences of input for dependency with a long range dataset and it will find the relationship between the frames to produce a pattern for sequence prediction.

In [6], corpus is collected based on three categories: 171 signs that could be used in the bank and finance domain ,everyday usable verbs, phrases and nouns ,and 496 signs that could be used in the health domain. The samples were collected for each phrase and sign by determining the vocabulary. Then those samples were being annotated by using HamNoSys and ELAN rendered by linguists. As SLR method needs huge amounts of data,.corpus will have at least 6 repetition .The software has been developed to streamline the recording procedures for providing the modalities of Kinect v2 sensor to guide signers and during recording session online sign border annotation is allowed. Signers tended to perform 30-70 signs. In the total set of signs, without replacement these sign were sampled randomly which makes a unique session by randomization by ordering the temporal of these signs thus to reduce the significance of statistics on the effects of coarticulation. Kinetic v2 modalities such as depth map, colour video, pose information and user mask and sign borders around the annotation were saved. Once we completed the collection of corpus , then it’s will accessed mainly for educational purpose and filling out the form on license agreement available in Bosphorus Signs website

In paper [7], Instead of using SL datasets obtaining the vocabulary that comprise of quotidian and medical terms from recorded video session of realistic interaction of the patient with mental health professional by using kinect2 device and Machine version camera.The corpus includes twenty one recorded scripts in the Greek sign language. The automation looks at the semantics of uttered sentence ,motion and facial expression.For recording purpose, MC050MG-SY (machine vision camera) which gives RGB videos composed of 60fps for capturing the fast motion and decrease the blur .Microsoft Kinect 2 device allows depth data capturing and synchronous RGB video.By ELAN software recorded videos session were annotated with GSL glosses to produce XML files consists of glosses and intervals of the action .Tracking the hands , face of signer and for each body part ,set of 3D keypoints were recorded which corresponds to the predefined position in body parts.the software will produce json files contains key points.

In [8], 3D model virtual person has been practised to provide the information in the sign language by the model called Hidden Markov Models(HMM).which is used as a pattern block model for preceding the analysis in sign language recognition.To translate the text in to the sign language the Motion Capture(MoCap) has been used.this MoCap equipment provide the exact definitions of the signs that is enacted by the person.To declare the virtual human in iclone and trained it with some medical based words or sentences. For preprocessing and smoothening the data LOESS algorithm is used.To get interact between the physician and the deaf the Graphical User Interface(GUI) has been used.

In paper [9], sign language was recognized by CNN framework is otherwise known as Convolutional Neural Network framework is used for pattern perception, image identification etc. Based on this framework, two models such as SGD or Stochastic Gradient Descent and Adam optimizer have been analyzed for sign recognition. Keras library has created the subsequent network and compute the complexity layer. Another library is Sci-kit learn which is for plot the confusion matrix in the model.

In [10],The system carries out many stages of image processing, such as real-time picture capture, skin segmentation, hand tracking, feature extraction, and artificial neural networks. This method is based on taking a picture using a webcam when taking an image in real time. The picture has undergone pre-processing in order to be processed as an image**.** The image processing includes grayscale conversion and thresholding. the camera should be installed fully to visible full hands including edges. the input of the camera should be normal in the resolution of 640 X 480. based on different devices the camera id should be differ so, the capturing the video should be before Skin segmentation is performed for to identify the skin region of a hand gesture in a video. The image is converted into YCbCr domain because RGB colour is not affection for skin segmentation. skin colour pixels are detected based on the hand region. In hand tracking the ROI is fast and error free. The eye can track the hand's position in the area and also notice changes in facial gesture. The gesture is considered an input to the system if the captured image is not changing at that exact moment. And that particular gesture has now been acknowledged. In Feature Extraction it classifies into edges template matching, moment invariant based features. In this edge feature is very broad and plain. The mathematical methods known as "edge detection" are used to locate areas in digital photographs where the brightness of the image suddenly changes. In ANN, the major objective is to recognise gesture patterns by building an ANN with learning skills that can generalise and create results from many types of glove input data, even if those input patterns change significantly from the trained input patterns. As a learning/training procedure, the back propagation algorithm is used to establish the network's weights. This application uses a feed forward ANN architecture. The image is then captured as a 6x7 frame, creating 42 inputs for the input layer of an artificial neural network. The output will then range from 0 to 1 when the ANN compares the inputs with the desired outputs.

In the paper[11], It explains Thresholding, Picture Preprocessing, Noise Reduction, and Camera Interfacing. The first and most crucial step in the entire sign recognition process is in camera interfacing. To take pictures with a webcam, camera interfacing is needed. By recording gestures, the inbuilt camera can recognise hand positions and actions. Pre-processing an image involves removing unwanted noise, adjusting brightness and contrast, and cropping the image as necessary. This method combines colour filtering, segmentation, and image enhancement. Erosion and dilatation are the two types of noise reduction.It may be used to eliminate noise, divide up a picture into smaller fragments, join sections that are out of alignment, and even identify intensity holes or bumps. Erosion is a technique for removing noise from a photograph. Moreover, you may add pixels to the borders and join broken-up pixels together using dilation. The thresholding process may be used to convert grayscale photos into binary images. As a result, while thresholding photographs, bright and dark zones are established.

In [12], In this study, the hidden Markov model (HMM) and SVM for ASL recognition are utilised. It is challenging to recognise dynamic hand movements, and researchers have been working on it for the past ten years. The appearance of the same sign varies when used by two persons. For ASL classification, the authors proposed a deep learning-based approach. A number of methods, including as motion gloves, image processing, and leap motion controllers, have been tested for ASL detection. To recognise 3D motion, the authors proposed an ANN-based model based on 50 ASL phrases. It is a time-consuming and expensive computing approach. In spite of several researchers' efforts, inter-class differences, sign complexity, and substantial inter-class similarities make ASL detection a challenging problem.They used the provided technique to identify different finger, thumb, and hand positions. They used these motions to recognise sign language as well. Using the SVM model as a classification method was used. This classifier gave the group the highest confidence rating. Hand motions are another assignment for this class. There are 28 static hand gestures used by Arce. The recommended SVM method is used to recognise these static hand motions. This approach successfully identified 28 hand gestures and digits 0 through 9. The leap motion controller used 10 digits and 26 sign language alphabets. Six sets of 23 traits each make up the features. The findings indicate the significance of the distance between two nearby fingertips. They used a DNN-based technique to recognise sign language. For sign

language recognition, we employed 24 ASL alphabets and a fine-tuned deep CNN algorithm In paper[13] ,A Deep learning method is used for recognizing dynamic and static Arabic signs from recorded video sessions. Recording the procedure , verification and selection of tasks methods are used for building own ASL dataset with public KSU Arabic Speech Dataset .Initially,80 signs were proposed by King Saud University students from Department of HoH.The procedure of recording is done by signers in three times phases .In phase 1
,supervision of sign language translator is utilized for recorded the signs by HoH department deaf students but not satisfied so proceed to the second phase, where a new application was launched to record the signs at KSU..Next Sony handheld camera was used to record continuous as standby at diverse angle resulting in dataset that contains 450GB handy-cam Sony recordings ,kinect1 {RGB
+ Depth} and KinectV2 {RGB + Depth + Skeleton}. The videos were edited manually frame by frame by use of the superposed key points.

In [14], proposed system contains three phases such as training phase where multiclass support vector machine (MSVM) was used for training the class, testing phase includes extracting the features from input gesture and in recognition phase , outcome of the most probable groups were identified for recognizing the gestures.Then Preprocessing is done in two steps , converting grayscale image into the binary image as segmentation of input image was being performed by Global threshold algorithm which is used to trackle classification problem with segmentation problem by dividing into two classes such as hand with pixel as 1 and background with pixel as 0.finally, filtering is done to eliminate the noise on segmented image.Next extraction of features were carry out by Hu Invariant Moment, mainly used for disjoint shapes and it composed of 7 values calculated by normalizing the central moments
,then Structural Shapes Descriptor were used for mainly extracting the features by approaches namely Compactness, Spreadness, Solidity ,Aspect ratio, Elongation and Orientation.MSVM is used to classify hand gesture by converting binary classifier to multiclass classifier with class i is considered as positive and j as negative .Computing the minimum distance of vector to represent binary pattern.The extracted features from input image was recognized and easily retrieved.

In paper [15], machine learning algorithms were used to recognize Portuguese Sign by selecting the hand features.Here the prototype is proposed consists of the two modules namely extracting necessary features and classification of sign language gestures ,in which user is allowed to be positioned on font of kinect cameras ,performing hand gestures and get interpreted by system and classifications were displayed on the interface and the solution is implemented upon the set of some assumption.Feature vectors were obtained ,then it being normalized and SVM is trained for classifying it.

In [16], hand gestures and the coresponding signs were implemented by a person with an Android smartphone . Indian Sign Lignuages were being captured , then obtained frames were transmitted and moved to further process .It was preprocessed by pre-processing techniques namely detecting the face and eliminating unwanted features, Skin colour segmenting for identifying skin colours in image , Morphology operations and Object stabilisation using Facial reference. Features were extracted from tracked hand g Grid-based fragmentation technique and then fedding onto classifier.After recognizing the hand pose classes in frames were send back to an android devices to classify them , intermediate motion and their pattern was stored in tuples. Next encoding it for Hidden Markov Model (HMM) to fed them into it , HMM chain with highest score in gesture is determined by forward-backward algorithm and produce the recognized gesture.

In paper [17], the traditional models or algorithms had been used.To classify and analysis the regression in the problem the K-Nearest Neighbour(KNN) algorithm is used,to enhance the certainty of the classifier the Simple Multi Attribute Rating Technique(SMART) framework is defined. To get the high accuracy level in the sign recognition,the association of Hidden Markov Model and Convolutional Neural Network is used. the deep learning algorithms has been defined that is Long Short Term Memory which is the elongation of Recurrent Neural Network otherwise called as RNN to expand memory of the model.here the algorithms has been compared based on the accurate level of the test and Support Vector Machine (SVM) has high accuracy when compared to the other deeplearning algorithms.

In [18], the hand gesture has been recognised by preprocessing the input by locating the signer with the help of viola jones algorithm then the hand gesture has been predicted. deep learning model such as an openpose framework is used to discover the 2D images to
enrich the machine realization of human activities to normalize and crop the image.For feature learning the Convolutional 3D(C3D) architecture with Fully connected(FC) layer is practised. In advance MLP that is MultiLayer Perceptron is to be integrate the characteristics of the vector and theclassification.

In paper [19], It has been a popular and challenging area of research to identify sign language motions in real-time video and successfully classify them into one of a number of categories. This field has been the focus of extensive research, so we also considered how we may advance it by working on it for our senior year major project. This idea has also been the subject of research by Liang et al., which has helped us throughout the implementation. The
single-sentence characterization

of the work carried out by this suggested system is the process of identifying and categorizing a sign language gesture. Together withthis, a text to ASL fingerspelling capability is also offered that makes the two-way communication from sign to text and vice versa

In [20], It was developed to process sign language using the KullbackLeibler divergence HMM (KL- HMM). The decoding procedure is the same as the standard HMM-based technique, where the log likelihood of state is replaced by the KL- divergence between the feature data and the condition categorical distribution. The feature observation in this method is a stack of posterior characteristics known as zt. The stack of posterior features can theoretically be expanded to mimic new input channels, such as mouthing and facial expression, as and when necessary or available. This is vital to keep in mind. A group of discrete units or subunits that correspond to the information channel are present in each posterior feature vector. If only target language resources are used for top- down modelling and only auxiliary or non-target language resources are used for bottom-up modelling, resource constraints in voice recognition can be addressed. The literature reports the use of two distinct methods to investigate the automatic derivation of hand movement components: first, unsupervised segmentation and clustering; and second, the use of sign annotations from HamNoSys. An HMM-based method that extracts signers' light supervision-dependent hand movement components has been developed in recent years. This study expands on that strategy to model hand movement data for multilingual sign language recognition in a language-independent manner.

CONCLUSION

In this sign language recognition survey paper , various techniques and tools had been studied and analyzed. Sign language recognition system had been developed to classify the static signs and apprehend dynamic actions by recording the sequences of continuous frames .Many researchers are building their SLR by self-made small datasets and most of the published paper with vision-based approaches show better outcomes than appearance- based approaches .The Classification techniques were used for identifying sign languages that gets varied among many researchers, their ideas and the limitation was compared with another method
.CNN –BILSTM model and HMM provide good recognition accuracy. A comparison between various methods proposed are shown in table1

TABLE 1: COMPARISONBETWEEN VARIOUS METHODS

Methods	Accuracy	Purpose
Sign language detection for communicating deaf in hospitals.	97.88 % overall	To use a spatio-Temporal Modelling techniques for sign language recognition
Hand Sign Recognition for Medical Field Genesis	95% accuracy overall	To apply YOLO Machine learning algorithmic rules & CNN to determine the sign language recognition in medical genesis
Deep Neural Network for detection of Sign action for COVID-19 Patients	97% and 99.34±0.66%	CNN - BILSTM Model is Proposed for sign Word detection in hospitalization
Sign Language detection for	90.77%	To use MobileNet –

Medical Consultation	overall accuracy	YOLO in CNN model to recognize sign language.
Video Recognition method for Indian Sign Language in medical field	97 ± 2% accuracy	To extract of OpenPose + LSTM model for sign action detection
A Turkish Sign word detection in healthcare	92.5% overall accuracy	To use Microsoft Kinetic v2 for recording sign languages
Visual Sign Language dataset for homecare and healthcare service sector.	96.5% level of accuracy	Deep learning algorithm such as NLP and deep neural language models like ELMO and BERT has been used for sign language detection
To recognize the Mexican sign language for assisting physician in healthcare field	Accuracy scores average rates of 99% and 88%	By using HMM model , MoCap and LOESS for predicting the sign language and GUI used as interact between physicians.

REFERENCES

1. Necati Cihan Camgoz,Ahmet Alp Kındıroglu, Lale Akarun,"Sign Language Recognition for Assisting the Deaf in Hospitals", International Workshop on Human Behavior Understanding, Oct 2016
2. Dr. Uttara Gogate, Vahaz Kazi, Aakash Jadhav, Tirthesh Bari,"Hand Sign Recognition for Medical Field Genesis", Ijraset Journal For Research in Applied Science and Engineering Technology, April 2022
3. Adithya Venugopalan, Rajesh Reghunadh, "Applying Hybrid Deep Neural Network for the Recognition of Sign Language Words Used by the Deaf COVID-19 Patients", Arabian Journal for Science and Engineering, April 2022
4. Kun Xia,Weiwei Lu,Hongliang Fan and Qiang Zhao, "A Sign Language Recognition System Applied to Deaf- Mute Medical Consultation",National Library of Medicine-National Center for Biotechnology Information(NCBI), Nov 2022
5. Jash Gandhi,Parth Gandhi,Aayush Gosar,Sheetal Chaudhari, "Video Recognition Techniques for Indian Sign Language in Healthcare Domain", 2nd International Conference for Emerging Technology (INCET)Belgaum, India, May21-23, 2021
6. Necati Cihan Camgoz,Alp Kindiroglu,Serpil Karabüklu,Meltem Kelepir, "A Turkish Sign Language Recognition Corpus in Health and Finance Domains"10th Edition of the Language Resources and Evaluation Conference (LREC),May 2016
7. D. Kosmopoulos,I. Oikonomidis,C. Constantinopoulos N. Arvanitis,K. Antzakas,
8. Bifis,G.Lydakis, A.Roussos,A. Argyros,"Towards a visual Sign Languagedataset for home care services",Foundation for Research and Technology – Hellas (FORTH), Heraklion, Greece.15th IEEE International Conference on Automatic Face andGesture Recognition.,Nov 2020
9. Candy Obdulia Sosa-Jiménez, Homero Vladimir Ríos-Figueroa, Ana Luisa Solís- González-Cosío, "A Prototype for Mexican Sign Language Recognition and Synthesis in Support of a Primary Care Physician",Journals and magazines Access
10. (Volume: 10),Dec 2022
11. Surejya Suresh,Mithun Haridas T.P,Supriya M.H, "Sign Language Recognition System Using Deep Neural Network ", IEEE 5th International Conference on Advanced Computing & Communication Systems (ICACCS),2019
12. PriyankaC Pankajakshan,Thilagavathi,"Sign Language RecognitionSystem",IEEE Sponsored 2nd International conference on Innovations in Information Embedded andCommunication Systems(ICIIECS'15)Mar2015
13. shish S. Nikam,Aarti G. Ambekar, "Sign Language Recognition Using Image Based Hand Gesture Recognition Techniques", IEEE 2016 Online International Conference on Green Engineering and Technologies (IC- GET),2016
14. Abdul Mannan, Ahmed Abbasi, Abdul Rehman Javed, Anam Ahsan,Thippa Reddy Gadekallu,Qin Xin, "Hypertuned Deep Convolutional Neural Network for Sign Language Recognition",National Library of Medicine-National Center for Biotechnology Information(NCBI),Apr 2022
15. Mohammed A.bencherif, Mohammed Algabri , Mohamed A.Mekhtiche , Mohammed faisal, Mansour alsulaiman, Hassan Mathkour , Muneer Al- hammadi,(Member, IEE), Hamid ghaleb, "Arabic Sign Language Recognition System Using

2D Hands and Body Skeleton Data " ,IEEE Access journals,Mar 2021

16. Karishma Dixit,Anand Singh Jalal, "Automatic Indian Sign Language Recognition System", 3rd IEEE International Advance Computing Conference (IACC),Feb 2013
17. Paulo Trigueiros, Fernando Ribeiro, Luís Paulo Reis, "Vision-based Portuguese Sign Language Recognition System ",published on WorldCIST,Apr 2014
18. Kartik Shenoy, Tejas Dastane, Varun Rao, Devendra Vyavaharkar,"Real-time Indian Sign Language (ISL) Recognition", Shenoy, K., Dastane, T., Rao, V., & Vyavaharkar, D. (2018). Real-time Indian Sign Language (ISL) Recognition. 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT).
19. Ahmed Sultan , Walied Makram , Mohammed Kayed, and Abdelmaged Amin Ali, "Sign language identification and recognition: A comparative study", Published by De Gruyter Open Access May 25, 2022
20. Muneer Al-Hammadi, Ghulam Muhammad, Wadood Abdul,Mansour Alsulaiman, Mohammed A. Bencherif, Tareq S. Alrayes, Hassan Mathkour, And Mohamed Amine Mekhtiche,"Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation",IEEE Access journals.Nov 2020
21. Amrita Thakur, Pujan Budhathoki, Sarmila Upreti, Shirish Shrestha, Subarna Shakya, "Real Time Sign Language Recognition and Speech Generation", Journal of Innovative Image Processing (JIIP) 2020
22. Sandrine Tornay; Marzieh Razavi; Mathew Magimai.-Doss, "Towards Multilingual Sign Language Recognition", ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)