EasyChair Preprint

№ 3012

An Improved Hand Gesture Recognition System Based on Optimized MSVM and SIFT Feature Extraction Algorithm

Sonia Heer and Darpan Anand

March 22, 2020

An Improved Hand Gesture Recognition System Based on Optimized MSVM and SIFT Feature Extraction Algorithm

Sonia[1] and Er. Darpan Anand [2]

[1] Chandigarh University, Gharuan, Mohali, India.
[2] Chandigarh University, Gharuan, Mohali, India.
Soniaheer13@gmail.com,darpan.e8545@cumail.in

Abstract. In today's robotics and machine translation tasks are performing the main role in hand gestures. Gesture or detection implementation will help humans in several means. Gesture recognition systems are used in various fields such as DNN, ML and NNs etc. The applications of this research are sign language translation, music creation and Robot remote controlling, etc. In this research work proposed HGR methods are using feature extraction and optimized MSVM classification. This process gives the applications of edge detection, interference, filters, SIFT-ALO algorithm and binary during image pre-processing, in which these approaches add to better extraction and selection. In proposed work, implemented an optimized MSVM method are two classes such as training and testing. Optimized M-SVM is performed on the ASL gesture dataset along with existing SURF and SIFT techniques. This research work is used for the simulation tool (MATLAB) and calculated performance metrics like processing time, error rate and accuracy rate with the MSVM value at 99.1 percent as compared to existing feature extraction methods.

Keywords: Hand Gesture Recognition System, Multi-SVM, SIFT-ALOA algorithm.

1 Introduction

It is easy and simple data communication between human and human. Once, it's near on humans and machine it is a very difficult job since machines identified all types of languages human can talk and understand [1]. Thus, for enhancing those transferring characteristics which machines can implement a few interactive methods such as GR. This is a famous research-topic in the area of CS and its uses regarded to DL technology. HGR is needed for modeling, control boundaries, no-touch and in- vehicles. Such DL: Deep Learning technologies give drivers to drive through at a similar time connecting with various panels such as cooling and sound [2]. HCI describes the path how the human communicates to the robot and because the mechanism is not helpful till an HTs the mechanism for a certain job. There are normally two features that would be verified when implementing a man-machine interaction model as described in Mechanism usage and evaluation [3]. The postures may be in any format such as pixel and hand image or human-defined pose that need minimum difficulty. Various methods are being implemented by the corporations for gaining essential data for detection hand-made gesture recognition-models. Few representative works with main equipment's like data glove and color caps to propose difficult data about HG offered by the human [4]. GR: Gesture Recognition is normally divided into 2 classes such as (i) Dynamic and (ii) Static [5]. The SGs are those that only need the handling of a particular hand gesture image as the input of the classifier, the main benefits of this method is the minimum cost. In DGs need the handling of hand gesture image series more tough GR methods. In the literature survey, it can explore various recognition scenarios based on UL: Unsupervised and SL: Supervised learn- ing methods. They may describe the SVM, CNN and KNN classifier algorithms. The proposed work, its valued gesture image bases of 16-gestures, certain edge detection and improved feature extraction methods and the use of MSVM of the classification method. With the research flowchart, it determined that with the basic flow of MSVM, it is possible to attain better outcomes for the ASL database classification technique and to associate the research methods with other prior methods in the survey and other HGR procedures. In research work has implemented an enhanced feature extraction method using the SIFT-ALO

algorithm. SIFT algorithm is used to extract the features and the ALO algorithm is used to feature selection in the Hand gesture recognition system. After that in research work has implemented the MSVM algorithm to classify the hand gestures categories. In this proposed work, compute and evaluate the performance metric such as time, accuracy and error rates (FAR, FRR).

2. Literature Survey

Marouane Benmoussa et al., 2019 [6] discussed the HCI expected a nice deal of care during this final period. Prior authors had curved to behavioral communication systems such as gesture human-machine edges. Current works were trying to resolve the issue of HGR: Hand Gestures Recognition utilizes ML techniques. Few of them were play- ing to get accurate performance. But, some of them were compelled into account mainly necessaries to relate the workflow of LM: Learning Model, normally data unbalance, model selection and generalize feature selection. In this research section, they implemented an ML technique for real-time classification and recognition of sixteen movements of user-hands utilizing KS that compliments like necessaries. The hand recognition was generated only when there was a moveable HG. This technique was depending on the training of the SVM method on hand depth information from which BoWs (Bag-of-Words) of Scale Invariant Feature Transformation and SURF features were extracted. The information was reserved managed and the structure kernel and metrics were chosen utilizing CVP: Cross-Validation Process. This technique attained 98 percent overall parameters utilizing the field under the ROC: Receiver operating characteristic curve calculates.

2.1 Gesture using SIFT and SURF Algorithm

HGs train images could define by sets of key-points created by SURF and SIFT, then the no. of key-points from the images are various and reduce reasonable ordering. To identify this issue, they utilize the BoWs technique. A BoW, which is a single of the most important techniques CV. Normally, the vital points in the pictures are the VWs: Vision-Words. These sections are characteristics and they are discriminators and even to change and scale-variations.

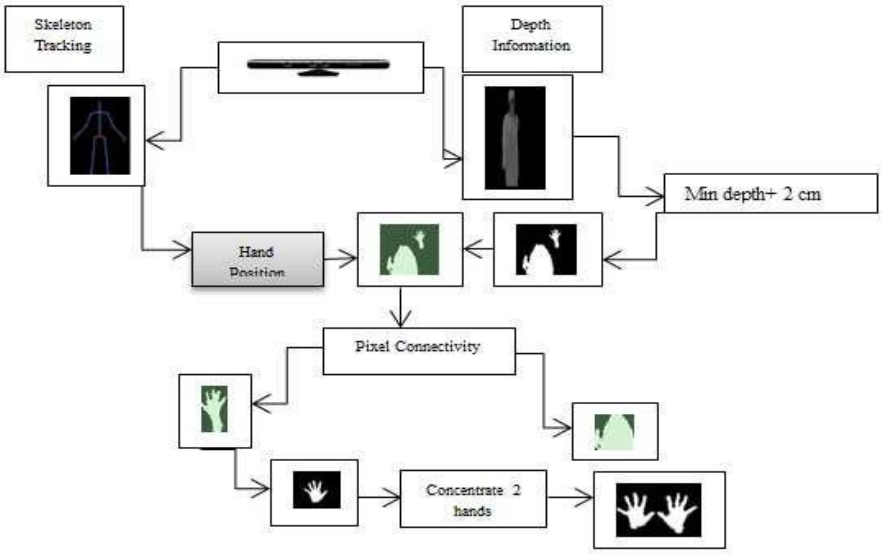

Fig 1. Three Phases of Segmentation

These characteristics are fetched utilizing SIFT and SURF technique forming a dataset of VWs then every novel picture is going to be defined as a geometric of VWs that looks in the picture. It is completed utilizing the VQC method (Vector Quantization Clustering), where an individual, group defines a VW that correspond to a valuable LP common by the Key-points in that group. Fig 1 defines a workflow for the model of the technique that utilizes the Bag-of-Words technique to classified hand signs. The block 1st, the fetched key sections are grouped and create a code-words dictionary, then calculate a histogram utilizing VWs for all the train samples. These histograms are then served to SVM classification. The block 2nd, test pictures are interchanged top-histogram demonstration by including various key-

sections seemed in each group and classified the consequences utilizing the train Support Vector Machine method [7].

Jay Prakash et al., 2019 [8] described as a social and technology was developed day-by-day, the use of devices such as wireless mobiles, electrical computers was also increased. In daily life, they utilize no. of the path of communication which contains language, lettering by somebody moveable, but while in case of robot they quiet held with writing and language so they essential few improvements so that they could interact with robots in somebody moveable also. Hand Gesture was a non-vocal route of communication which utilizes hand-motion, several body postures, facial emotions. So, to create a smart robot they were permitting the machine to take knowledge by recognizing various HGs. HGs were utilized as input in the DIP system.

2.2 Template Matching

This technique for HGR and postures used a simulation technique to define the essential no. of template of a particular gesture to be full that shall be protected on the dataset for the similarity procedure of the method. If this system would not be able to recognize the posture offered with the templates an extra template must be stored and trained in the dataset until the HGR system accurately recognizes the postures. The promoter will sum-up all the time in second under a particular no. the similar no. of template gesture and process is shown in fig 2 [9].

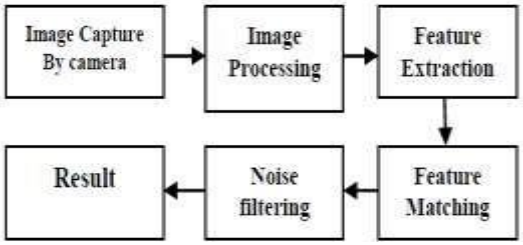

Fig 2. Flow Chart of HGR Process

Hanwen Huang et al., 2019 [10] defined GR, while had been searching for various years, was still a limitation issue such as complex background, digital camera angles and brightness situations kind the issue more complex. In this work, defines a robust and speedy technique for HGR depend on RGB video. Initially, they recognize the skin-based on their color. They fetch the region and edge the hand contour. Lastly, they detect the hand gesture. The outcomes of simulation, determine that the research technique was effective to detect postures with increase accuracy rate than the state-of-the-art.

2.3. Hand Region Segmentation and Contour Extraction Method

They eliminate the interference in the hand images. They required extracting regions. They measure each section group as an edge. Between these regions, there is only a single region, which defines HR. Facial and Hand regions are the huge twice contours. The issue of searching the hand from the regions develops the difficulty of diving hands from facials. Then, they gather a hundred image samples of facial and hand regions. The VggNet is used for the classification process and training model shown in figure 3. In this network, the model is based on DNN established by the Visual- geometry cluster and scholars at google-deep-Mind. This model finds the association between the complexity and evaluation of CNNs. By loading 3*3 matrix is used tiny convolution advantages and 2*2 Max-Pooling layer. This model has positively built sixteen or 19 layers DCNNs. In this research work has used 16 –layers in the vggNet model [11].

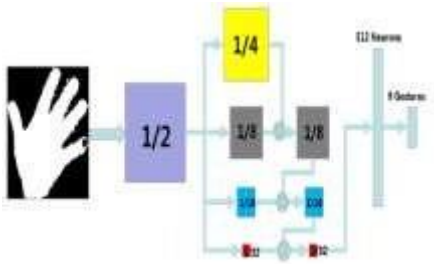

Fig 3. Training structure in VggNet Model

Table 1. Analysis the various methods and performance metrics

Author Name	Method	Parameters	Citation
Marouane et al., 2019	SIFT SURF SVM	Average time in sec = 0.3 and 0.1 sec Speed = 0.5 and 0.25s AUC = 98% and 91%	[6]
Jay Parkash et al., 2019	Template Matching Naïve Bayes PCA	Price Efficient	[8]
Hanwen et al., 2019	VggNet Model	Accuracy = 98.1%	[10]

3. Research Challenges and Methodology in HGR System

During the study of the literature survey we have found out some ways in real-time human communication by posture verification needs future enhancement to recognize hand gesture data accurate and effective design of handshape or size. It looks at difficulties in sign recognition for variable performance by different performers. From the survey, we have found the problem or proposed work in which we are going to continue our work of Hand Gesture Recognition for feature extraction and classify through Multi Support Vector Machine.

3.1 Challenges: The HGR has various challenges described in detailed such as (i) Rotation (ii) Size (iii) Illumination changes and (iv) Location Issues [12].

(i) Rotation Issue:

The degree of freedom is the main problem. Then, the degree of freedom is modified and gesture initial image might be different and by output, the image can change.

(ii) Size Issue:

The human being hand features have different shapes and sizes such as child have small hands and younger have big hands, it could create issues in the system.

(iii) Location Issue:

Although the define input image if hands location different such as hand place in corner of the screen, all points which recognizing hand location does not lie on hand image, then it could generate an issue to consider input from people.

(iv) Illuminate variation Issue:

The brightness variations can affect HG input as it could vary extracted skin contour.

3.2 Methodology

In this section, elaborates the HGRS is separated into 4 sections like;

- Sign Image Dataset Collection
- Sign Image pre-processing
- Segmentation in Gesture images
- Feature Extraction using SIFT method and
- Optimized MSVM classification algorithm.
- Performance Metrics.

The collection of ASL: American Sign Language gesture dataset from the online repository site. Firstly, upload the sign image from the trained dataset folder. Conversion applied in the resized image that is rgb2grayscale format. It reduces the format of the converted image in 2D. It identifies noises in the grayscale image. It applies the filtration method to remove the interference in the noisy image. After that image converts into a binary pattern and detects the edge in the binary image. SIFT algorithm is used to extract

the features in the key-point format. In a hybrid optimized MSVM algorithm to classify the gestures and recognition rate. At the last, evaluate and compared the performance metric such as accuracy, processing time and error rate.

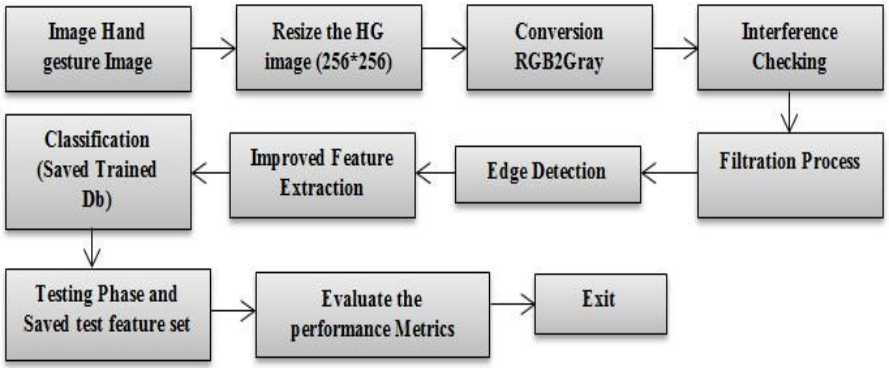

Fig 4. Diagram of research methodology in Work Flow

4. Experiment Analysis

In this research work of the method was completed in MATLAB 2016a. The ASL Database [13] is used for recognizing the gestures and posture of the hand's images shown in fig 7. The database contained 16 English Alphabets categories, 80 percent of train the hand gesture samples and 20 percent is used for testing. The research method evaluates a 99% accuracy rate in verifying the test image in the database. The research algorithm workflow is shown in fig 4. Such a huge no. of HG samples offered different sizes and shapes as can be defined in fig 5, which defines various alphabetic samples of gesture 'a'. Some other samples of HGs could define in fig 6.

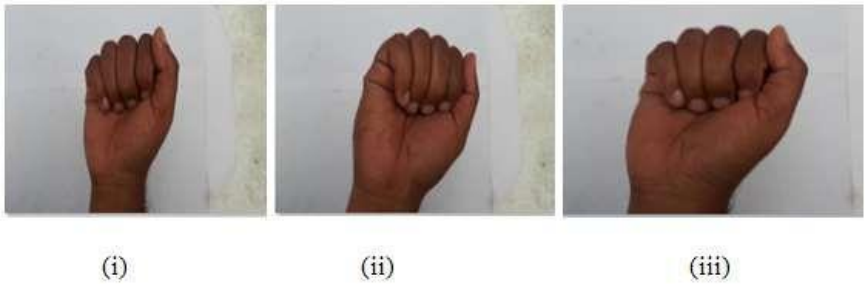

Fig 5. Hand Gesture of 'a' sample

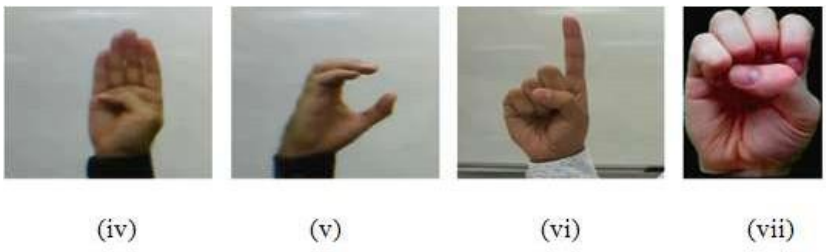

Fig 6. Other HG samples L→ R: Gesture 'b', 'c', 'd' and 'e'.

Fig 7. ASL hand gesture Dataset Images

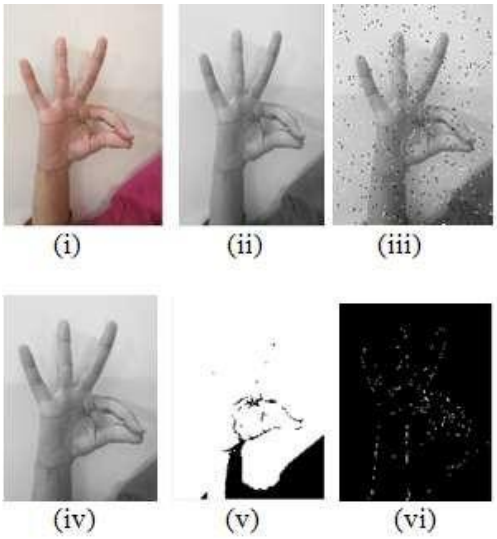

Fig 8. (i) Input Hand Gesture Image (ii) Conversion HG image (iii) Noisy HG Image (iv) Filtered HG Image (v) Binary HG Images and (vi) Edge HG image.

Fig 8 shows (i) uploaded the hand gesture input train and test image database from the knowledge domain. The properties of uploaded images are pixel size 864 * 1152, Bit Depth 24, Horizontal and Vertical resolution 96 dpi. (ii) The uploaded color image covert into a grayscale image. In conversion, grayscale format image is reducing the dimensionality of the color image. (iii) In this image is identifying the noise-level. (iv) If noise presented in the uploaded image and then applying the filtration method to remove the interference in the image. (iv) After that, it converted a smooth image into binary format. (vi) Lastly, it detected the edge in the binary image.

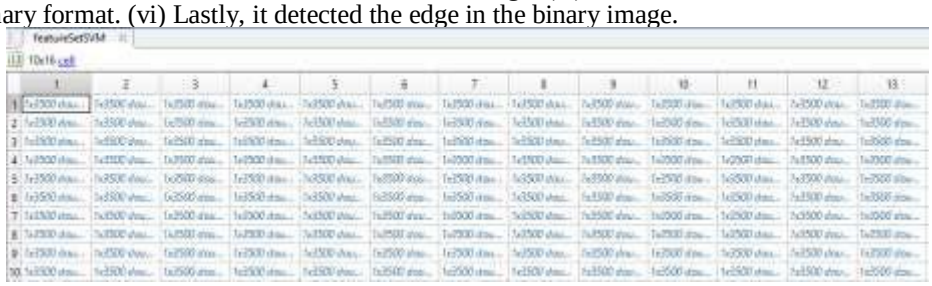

Fig 9. Extracted Features with SIFT algorithm

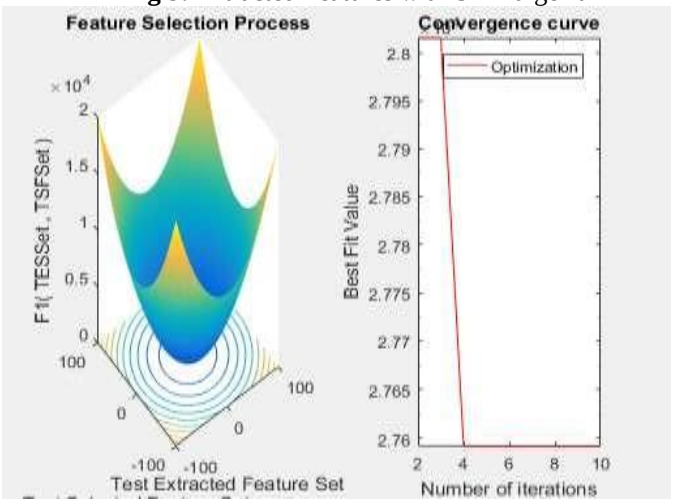

Fig 10. Ant Lion Optimization using feature selection

Fig 9 defines the feature set with all types of local features in numeric format. In this section, has implemented a SIFT algorithm to extract the features in the form of key- points. It extracted the intersect key points from the gray scale image at which local nearest features. This method is used for a large number of feature extractions in the uploaded image. Fig 10 shows the graph based on the ALO method is used for feature selection. This

method consists of a search by a random selection of the extracted feature set. It has compressed the extracted feature set. After that Optimized MSVM classification method has implemented to recognize the hand gesture in alphabetic format (A-W).

Table 2. Performance Parameters in Optimized MSVM algorithm in HGR.

Sr no.	Parameters Name	Proposed Values
1	Accuracy rate (%)	99.1
2	FAR	0.022
3	FRR	0.003216
4	MSE	0.0322
5	Processing Time	0.036

Table 3. Comparative performance analysis with O-MSVM, SURF and SIFT methods

Parameter Values	SIFT	SURF	OMSVM
Accuracy rate (%)	91	98	99
Processing time (second)	0.118	0.29	0.036

Table 2 defines the performance parameters with the Optimized MSVM method. In the proposed method parameter performance, such as accuracy rate 99%, processing time 0.0369 and MSE value is 0.0322. Table 3 shows the comparative performance analysis with proposed (OMSVM) and existing (SIFT and SURF) methods. In the research algorithm, the accuracy rate is maximum as compared to other feature extraction methods (SIFT and SURF) and the processing time is lesser than other methods.

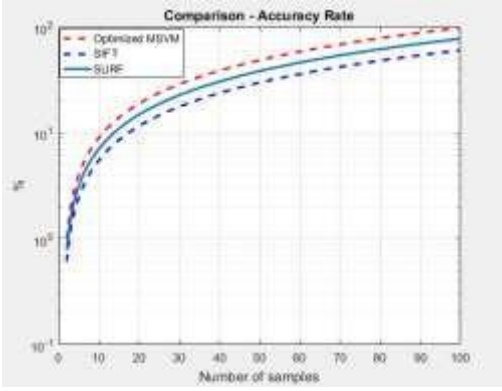

Fig 11. Comparison based on SIFT,SURF and Optimized MSVM algorithm with Accuracy Rate (%)

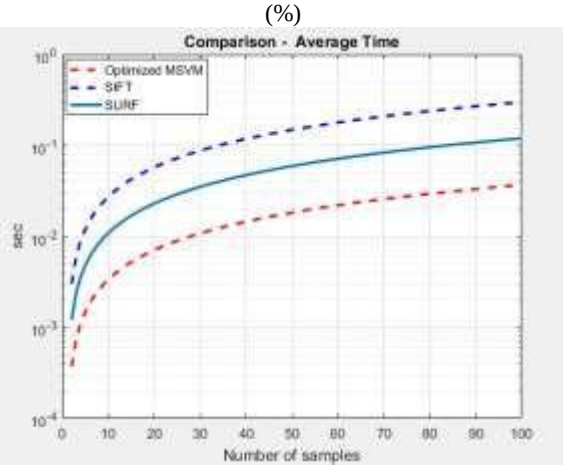

Fig 12. Comparison based on SIFT, SURF and Optimized MSVM algorithm with Processing Time in seconds.

The above figure 11 and 12 shows the comparison between proposed an optimized MSVM method and various feature extraction methods such as SIFT and SURF algorithm with accuracy rate are calculated. In optimized MSVM accuracy rate is higher than other feature extraction methods. In hand gesture recognition, the rate is maximized with the optimized MSVM algorithm. Second

comparison shows the processing time with proposed and existing feature extraction methods. In research algorithm, the processing time is less as compared to the existing feature extraction methods.

5. Conclusion and Future Scope

The Hand Gesture Recognition System is performing a major role in today's robotics and machine translation tasks. This research makes a message between human and mechanism is very easy and understandable. The applications of this research are sign language translation, music creation and Robot remote controlling, etc. The main focus of the proposed algorithm is on size, positions of gesture and recognition rate. After studying various existing techniques such as feature extraction, and classification techniques, we developed pre-processing methods such as image resizing, grayscale conversion, filters, edge detection, and binary process. The proposed hybrid optimized feature extraction and classification technique is developed and simulated. The overall research is based on various performance parameters such as accuracy rate, processing time and error rate. Optimized M-SVM is performed on the ASL gesture dataset along with existing SURF and SIFT techniques. The performance is calculated after the simulation the proposed algorithm shows better performance than the previously developed algorithms. This research MATLAB is used for simulation and achieved an accuracy rate at 99.1 percent, an error rate of 0.024 in the 0.01 processing time. As per the comparison in chapter 6, the proposed optimized MSVM shows better performance in all the cases.

In further work, it can develop a color segmentation method and deep learning method to improve the tracking and capturing of the gesture in any shape or size.

References

1. Donsong Ya, Yongjun Xie, Design on Computer System based on hand gesture recogni- tion, 2018 IEEE15th international conference on networking, Sensing and control, pp 1-4 (2018).
2. Prof Quan Min Zhu, Hand Gesture Recognition System , International Journal of Comput- er applications in Technology, 57(4), pp. 1-6(2018).
3. Hong Dai, Research on SVM improves algorithm for large data classification. IEEE 3rd international conference on big data analytics, pp 181-185(2018).
4. Varun, Kollipara Sai, I. Puneeth, and T. Prem Jacob. "Hand gesture recognition and im- plementation for disables using CNN'S." In *2019 International Conference on Communi- cation and Signal Processing (ICCSP)*, pp. 0592-0595. IEEE(2019).
5. K. Pisharady and M. Saerbeck, Recent methods and databases in vision-based hand ges- ture recognition: a review, Computer Vision and Image Understanding, 141(4356), pp. 152–165(2015).
6. Benmoussa, M., and Mahmoudi, A., Machine learning for hand gesture recognition using bag-of-words. In 2018 International Conference on Intelligent Systems and Computer Vi- sion (ISCV) (pp. 1-7). IEEE (2018).
7. Runqing Zhang, Yue Ming, and Juanjuan Sun. Hand gesture recognition with surf-bof based on gray threshold segmentation. In Signal Processing (ICSP), 2016 IEEE 13th Inter- national Conference on, pages 118–122. IEEE (2016).
8. Jay Prakash, Uma Kant Gautam Hand Gesture Recognition, International Journal of Re- cent Technology and Engineering (IJRTE), vol 7(6), doi : 90300476C19\19©BEIESP ((2019).
9. Vladimir I. Pavlovic, Rajeev Sharma, Thomas S. Huang. Visual Interpretation of Hand Gestures for Human-Computer Interaction;, . s.l. : IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997. https://pdfs.semanticscholar.org/8dd3/d5ef2e1c40433417cd90e21ce9a0468a7626.pdf
10. Huang, H., Chong, Y., Nie, C., & Pan, S. Hand Gesture Recognition with Skin Detection and Deep Learning Method. In Journal of Physics: Conference Series (Vol. 1213, No. 2, p. 022001). IOP Publishing(2019).
11. K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, Proc. Int'l Conf. Learning Represent.,(2015).
12. Shigeo, Abe. Support Vector Machines for Pattern Classification. Kobe University.
13. A.L.C. Barczak, N.H. Reyes, M. Abastillas, A. Piccio and T. Susnjak, A New 2D Static Hand Gesture Colour Image Dataset for ASL Gestures, Vol. 15, Res. Lett. Inf. Math. Sci., pp. 12–20(2011).