An Inclusive Survey of Machine Learning based Hand Gestures Recognition Systems in Recent Applications

To cite this article: Hind Ibrahim Mohammed et al 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1076 012047

View the article online for updates and enhancements.

You may also like

- A Human-Robot Interaction for a Mecanum Wheeled Mobile Robot with Real-Time 3D Two-Hand Gesture Recognition
  Xueling Luo, Andrea Amighetti and Dan Zhang

- A deep convolutional neural network model for hand gesture recognition in 2D near-infrared images
  Celal Can, Yasin Kaya and Fatih Klç

- An Overview of Hand Gestures Recognition System Techniques
  Farah Farhana Mod Ma’asum, Suhana Sulaiman and Azilah Saparon
An Inclusive Survey of Machine Learning based Hand Gestures Recognition Systems in Recent Applications

Hind Ibrahim Mohammed¹,³, Jumana Waleed¹ and Saad Albawi²

1 College of Science, University of Diyala, Diyala, Iraq
2 College of Engineering, University of Diyala, Diyala, Iraq
3 E-mail: hindimohammed83@gmail.com

Abstract. Hand gestures represent one of the most prevalent types of body language which can be utilized for interaction and communication. Although the other types of body language represent a more general state of emotional, hand gestures capable of possessing specified linguistic content inside it. Because of the expressiveness and speed in interaction, hand gestures are commonly utilized in human-computer interaction systems (HCI), sign languages, virtual reality, and gaming. In the process of recognizing hand gestures, the complexity and diversity of gestures will extremely impact on the recognition rate and reliability. The existence of machine learning techniques can be effectively exploited in the task of improving the rate of hand gesture recognition. This paper inspected the performance of machine learning techniques in recognizing vision and sensors based hand gestures in the recently existing applications. Additionally, the widely used architecture applied in various datasets has been considered which includes the acquisition of data, preprocessing, the extraction of features, and classification.

Keywords: Vision-based hand gestures recognition, sensors-based hand gestures recognition, Machine learning techniques, Recent applications.

1. Introduction

The recognition of hand gestures represents a wide research area that is subcategorized depending on the gestures context and the technologies utilized for inputting those gestures. There are different taxonomies that impact the designed hand gestures recognition systems; environmental factors, individuals that perform the gestures, the effectiveness of the devices utilized for capturing, the gestures type (static, or dynamic), and the system application. Various techniques for analyzing the hand gestures systems have been founded: Firstly, sensors-based techniques that mostly require the person to wear sensors (input devices) such as gloves, and bracelets. These techniques have merit that identifying gestures is not distracted via the backgrounds of diverse, however, comes with a tradeoff of natural interaction lack, bulkiness, and high cost. Secondly, vision-based techniques which utilize capturing devices such as cameras or kinetic sensors for inputting information depending on the way a person perceives it’s surrounding. The effectiveness of these techniques is based on several factors such as the number of cameras and their positioning, the hand visibility and how it is isolated from the image, the efficiency of the extracted features, and the classification techniques [1].

Hand gestures remain one of the extremely natural and effective ways for people to communicate. About 65 % of human communication includes nonverbal gestures in which the hands are the most
significant part for people to communicate compared with other parts of the human body like arms, facial, and the movements of the body. Utilizing the gestures of hand for communicating with machines with no need to use any supplemental devices is very simple and appropriate in the applications of the industry such as; robot controlling, game controlling, and virtual reality for car systems and smart home [1]. This means that the recognition of hand gestures become very interesting, a highly significant and hot topic of research for developing a more efficient and robust gestures recognition systems[2][2].

The organization of this paper is structured Machine Learning based Hand Gestures Recognition Systems, Performance Analysis, Conclusion, and References

2. Machine Learning based Hand Gestures Recognition Systems

In order to present the algorithms of machine learning, it is significant to present the fundamental principles behind these algorithms. Generally, the algorithms of machine learning can be categorized into supervised, un-supervised, and semi-supervised learning. All these algorithms need data in which specific patterns or classes can be learned. These data are generally represented as training data (see Fig. 1). The supervised learning represents the state that the number of classes to be learned and the training cases assignment for these classes (labeling) are known. This algorithm of machine learning works on identifying the patterns in concepts of features constellations that can be utilized for differentiating the various classes in database. The resultant classifier can be implemented in new data with non-known class labels. The un-supervised learning represents the state that the classes and the training data assignment for these classes are unknown. The data exploration represents the main application for this type of learning in the research of clinical imaging. The semi-supervised learning represents a combination of supervised and un-supervised learning algorithms. The comprehensive assignment for the whole training data to their corresponding labels of class probably hard in many cases [2].

![Figure 1. Overview of the techniques of machine learning.](image)

Many researchers have suggested and developed many techniques that use the term Learning Machine or a Kernel-based Extreme Learning Machine without specifying an algorithm for hand gesture recognition as D. Lu et al. (2016) used a YoBu data glove to gather data for gestures. In this system, an extreme learning machine (ELM) is utilized for recognizing gestures that have not yet established in the related application [3]. V. Gajjar et al. (2017) Presented Machine Learning based hand gesture recognition, with the use of Computer Vision, different types of gesture applications has been created [4], therefore the best value in accuracy value V. Gajjar et al. (2017) and table 1 shows the researches details.
Table 1. Comparison between the systems based on ML in [3] and [4].

|                         | D. Lu et al. (2016)                           | V. Gajjar et al. (2017)               |
|-------------------------|-----------------------------------------------|--------------------------------------|
| **The data set**        | Established a dataset that involves ten types of static gestures | Sabastian Marcel – Hand Gesture dataset. |
| **Preprocessing**       | Utilized a novel YoBu data glove for collecting data | Disjointing the elementary colors precisely |
| **Feature extraction**  | 54-dimensional features                        | Haar – Like features                  |
| **Classification**      | Kernel-based SVM and ELM                       | Machine Learning.                     |
| **Accuracy**            | ELM-Kernel achieved accuracy higher than SVM and ELM | The accuracy is approximately 96%.    |

There are several Machine Learning techniques, As follows, the widely used techniques are:

1- Support Vector Machine (SVM): that’s one of the simplest linear classification algorithm. It analyzes the data based on separating different classes by finding a hyper plane which can best maximize the margin among different classes, hyper planes as a fundamental concept for decision boundaries and for separating the different classes. SVM is one of classifiers most frequently utilized and has highly acclaimed with tasks performed on data that can linearly separable. When it comes to non-linearly separable data, the efficiency of SVM falls short, the solution to this problem is to use kernel functions to transform data into wider dimension space with this step data that can be separated linearly. In addition to adjusting kernel parameters, the main idea of this classifier is to define a convenient kernel function. [5].

2- Naive Bayesian (NB): One of the most common classification approaches is Naive Bayesian algorithms derived from Bayesian theory which is interested with computing the probability of a specific pattern (X), given some set of observations. However, to test pattern X, The posterior probability of a class or hypothesis (that is, the probability of a particular pattern that belongs to class i and gives its observed feature values) of a class or hypothesis is computed [6].

3- K Nearest Neighbors (KNN): This classifier is a case-based machine learning algorithm that is depending on a function of similarity or distance for different pairs of observations like the function of Euclidean distance. This classifier is applied to considerable applications due to its efficiency, easy to implement, and non-parametric characteristics. Moreover, with this algorithm, the time of classification is highly long and it is hard to obtain the optimal K value. In general, the preferable alternate of k to be selected bases on the data. Additionally, the noise effect on the classification is minimized via the bigger values of k, however, produce boundaries between classes less distinct. Based on utilizing different heuristic algorithms, a preferable ‘k’ can be chosen. for overcoming the above-said disadvantages, the conventional KNN with various K values for various classes must be modified instead of using a fixed value for the whole classes [7].

4- Decision Trees (DT): The class of supervised learning falls under decision tree algorithm. One of the benefits of a decision tree could utilize to solve both problems of classification and degradation. The DT utilizes an impersonation tree for solving a problem at every terminal node matches with a group naming and appears the characteristics at an internal node of the tree. The key challenge in Decision Tree is to define the characteristic of a root node in all steps. The procedure is to know a selection of characteristics [6].

5- Random Forest: as its name suggests, made up of a big number of singular decree trees that act accordingly a group. Every tree in a random jungle broadcasts a category foretelling and the group that gets the maximum proposal to be predictive of the model. The basic ideas beyond random forests are simple but robust is the maxim of concentrates. In data science, the purpose why a
random forest model doing so well is that a great number from comparatively unrelated models that act such a committee will excel every from the singular component types [6].

6- Neural Networks: It is capable of performing a number of regressing and/or classifying processes at once. In spite of the fact that usually each network only performs one operation. Thus, in the network would have a single output variable in most situations, but in the case of multi-state classification techniques, this can lead to a number of output units (the post processing step needs to take care of mapping output units to output variables). Three key principles are based on the Artificial Neural Network (ANN): the input and activation functions of the device, the network architecture, and the weight of each input relationship. ANN is an information processing system specific to biological neural networks that has unique performance characteristics. The mathematical models of human cognition or neural biology have been developed as generalizations. Network architecture is called the arrangement of neurons into layers and the connection patterns within and between layers. Neural networks are also known as either single or multi-layer layers. The ANN class comprises various architectures, including Coevolutionary Neural Networks (CNN) [7].

The hand gesture recognition systems are capable of using a Principal Component Analysis (PCA) algorithm alone or with ANN techniques. T.-N. Nguyen et al. (2015), proposed a new approach to recognize sign language that used to highlight the skin in the image. The pre-processing improves image quality and gives your hands armless. The characteristics of each hand are then extracted using the PCA method, with the use of an artificial neural network. [7]. And M. K. Ahuja and A. Singh (2015), focused on the vision-based hand gesture recognition system by proposing a system based on skin color model and thresholding methods with a database-driven hand gesture recognition along with an efficient template matching using PCA [8]. These algorithms have different resolutions, in different conditions such as lighting, distance, motion, camera quality, background, and contrast, therefore seen that accuracy of M. K. Ahuja and A. Singh (2015) is the highest. The table 2 shows the details of these systems,

| Table 2. Comparison between the systems based on PCA and ANN in [8] and [9], |
|---------------------------------------------------------------|
| **The data set** | The letters of American sign language |
| **Preprocessing** | Enhancing the quality of image and getting the hands with no arms |
| **Feature extraction** | Eigen space, eigenvalue, eigenvector, skin color and Locate the wrist and separate |
| **Classification** | PCA and ANN |
| **Accuracy** | 94.30% |
| **M. K. Ahuja and A. Singh (2015)** | |
| **The data set** | Four gestures from four subjects with 5 various poses per gesture |
| **Preprocessing** | Applying skin color model in YCbCr color space. |
| **Feature extraction** | skin color and thresholding |
| **Classification** | PCA |
| **Accuracy** | average accuracy =91.25% |

There are lots of researchers who worked for designing and developing the recognition hand gestures systems by utilizing only the SVM algorithm. A. I. Maqueda et al. (2015) developed a robust vision-based hand-gesture recognition system and created a new database for testing by delivered a bank of SVM classifiers to perform the identification of gestures. Excellent results have been achieved using a dynamic filter to increase its rate of detection by outperforming other existing systems. [9]. G. Pombozapuinez and J. H. Terriza (2016) The gesture recognition was achieved using SVM to identify the various kernel forms (radial, polynomial and sigmoid) to achieve optimal conditions for gesture learning and
recognition as well as precise determination.[10]. S. Sapienza et al. (2018) presented a minimal complexity algorithm based on the Average Threshold Crossing (ATC) technology. Using ElectroMyo Graphic surface signals of three forearm muscles, the number of threshold crossing events produced by a fully adapted acquisition board, are used to detect four different motions of the wrist: flexion, extension, abduction and grasp.[11]. M. Tavakoli et al. (2018) presented a system of 2-channels EMG with a high dimensional space of features and utilizing SVM as a classification algorithm. Additionally, the system tolerance for rejecting unsolicited gestures through the body movements has been evaluated. This system was performed for ensuring two methods; One is based on an SVM threshold, and another is based on a locking gesture being applied. Up to 5 gestures (hand closing, hand opening, wrist flexion, wrist extension, and double wrist flexion) can be recognized by the resulting device.[12]. G. Poon et al. (2019) A learning-based method was suggested for bi-manual gesture recognition by SVM. Several cameras from different viewing angles used to gather hand data will compensate for the distortion of one image by another to solve the issue of self-occlusion. [13]. And P. Parvathy et al. (2020) proposed a system to train and test using Sebastian Marcel static hand posture database which is available online using SVM [1]

In the above papers, we see that the accuracy value using the SVM algorithm began to gradually increase for the period from (2015-2020) to reach its highest in 2020 with a value of =96.50%, Table 3 shows the details of these systems.

Table 3. Comparison between the systems based on SVM in [1] and [10-14].

| A. I. Maqueda et al. (2015) | (ASL) Dataset, NTU Dataset | Searching for the possible hand in a frame of a video sequence poses | global spatial information | SVM | accuracy of the global system is very high |
|----------------------------|-----------------------------|------------------------------------------------------------------|---------------------------|-----|------------------------------------------|
| G. Pomboza-junez and J. H. Terriza (2016) | The EMG data Supplied by the MYOR armband | Recognizing gestures related to the classes in the gesture’s library. | Hudgins’ time domain (T) features | SVM | 79.36% |
| G. Pomboza-junez and J. H. Terriza (2016) | EMG data delivered by the MYO armband | Collecting the data using EMG sensor and converting it into discrete signals | Gaussian Radial Basis ( RBF), Polynomial and Sigmoid | SVM | The best one accuracy is RBF |
| M. Benmoussa and A. Mahmoudi (2018) | Vocabulary dataset | keep the data balanced | ‘Scale Invariant Features Transform (SIFT), and Speeded Up Robust Features (SURF)’ | SVM | accuracy of the system is approximately high |
| S. Sapienza et al. (2018) | Acquisition of the signals for 10 subjects who repeated every gesture ten times | Recording the signals, computing the features and predicting the gestures. | Average Threshold Crossing (ATC) based hand movement recognition algorithm | SVM | |
### Accuracy: 92.87%

**C. M. B and M. E. Benalc (2018)**

| The data set | the data of 10 healthy volunteers |
|--------------|----------------------------------|
| Preprocessing| Rectifying, filtering, and detecting the activity of muscle. |
| Feature extraction | Utilizing a sub-window for observing a signal segment that is seen through the main window |
| Classification | ANN with 3-layers for classifying each sub-window |
| Accuracy | 90.7% |

**G. Poon et al. (2019)**

| The data set | The multichannel-based EMG data |
|--------------|---------------------------------|
| Preprocessing| designing of three a support structure electrodes. |
| Feature extraction | Two-channel EMG module with high-dimensional space for features |
| Classification | EMG system with SVM |
| Accuracy | Between 95% and 100% for a trained user |

**P. Parvathy et al. (2020)**

| The data set | Sebastian Marcel static hand posture database |
|--------------|-----------------------------------------------|
| Preprocessing| Removing noise by utilizing filters such as; Kalman filter, morphological filter, and other filters like Gaussian and median filters |
| Feature extraction | Discrete wavelet transforms (DWT), with modified Speed Up Robust |
| Classification | SVM |
| Accuracy | 96.50% |

Other researchers used various classification algorithms with the SVM algorithm such as KNN, NB, (DT), and Neural Networks for detecting the hand gesture and simple output gesture. In order just two gestures can be observed, such as open and closed hand gestures, or in sign language, two groups can include two. Alejandro Betancourt et al. (2015) present the dataset for hand-detection and incorporate a dynamic perspective using SVM, DT, and Random Forest (RF) but Studies on which SVM-HOG is the best hand-detection combination [14]. X. Yingxin et al. (2016), proposed a robust method for hand gesture recognition based on CNN, which is utilized to automatically extract the spatial and semantic feature of hand gesture [15]. P. Pławiak et al. (2016) experimented on ten people performed (22) hand body language, gestures by using machine learning algorithms (Probabilistic neural network, SVM, and KNN) trained and tested by a 10-fold cross-validation technique [16]. J. L. F. C et al. (2017) proposed a system proposed to create two CNNs, each with a different number of layers, depth, and number of parameters per layer, to compare the accuracy and error results obtained. In contrast to the normal systems that are based on machine learning, the CNNs showed better performance. Good use of digital image processing techniques can help to better define the area comprising the hand gesture. [17]. M. E. Benalcazar et al. (2017), proposed a new real-time system for hand gesture recognition based on the KNN and dynamic time warping algorithms. This system can learn to recognize any gesture of the hand [18]. D. A. Maharani et al.(2018) discussed the hand gesture recognition as input command for Bioloid Premium Robot using two methods, K-Means clustering and SVM with DAG decision, these study resulted in Multiclass SVM with DAG decision performs better than the K-Means clustering method [19]. M. Benmoussa and A. Mahmoudi (2018), Proposed a machine learning system for the
real-time recognition of 16 gestures of the user’s hands using the Kinect sensor, recognition is only activated when there is a moving hand gesture that extracts a strategy based on the training of a Support Vector Machine model on hand depth data from which SIFT and SURF descriptor word bags are taken.[20]. C. M. B and M. E. Benalc (2018), proposed a hand gesture recognition model based on superficial electromyography signals used a feed forward neural network to label very sub-window observation, found that the recognition accuracy of the proposed model improves when the values were combined them of the preprocessed signal with the results of applying a bag of functions [21]. F. Wahid et al. (2018), Used many algorithms for machine learning, KNN, Discriminant Analysis (DA), NB, RF, and SVM for recognizing three different hand gestures. This proposed system works on improving the gesture recognition accuracy [22]. Cheng et al. (2019), proposed a new search algorithm for spike firing time that can minimize the search interval and a pre-sparing neural network is built to efficiently boost efficiency by using EMG signals in pattern recognition tasks. [23]. B. Abhishek et al. (2020), Described how the system implementation is performed based on the images captured. Hand detection is done using the Open CV and TensorFlow object detector and optimized for interpretation of gestures by the computer to perform actions like switching the pages, scrolling up or down the page . [24]. B. Yu et al. (2020), presented a model in order to capture the low-level features of the original video frame and its subsequent optical flow, shallow two-stream CNNs are used. In addition, an attentive function fusion module is designed to selectively combine useful information based on the attention mechanism from the previous two streams. [25]. T. Zhang et al. (2020), proposed a two-stage hand gesture recognition method. At the first stage, hand pose estimation is developed to locate the hand key points using the convolution pose machine and the second stage. The Fuzzy Gaussian mixture models (FGMMs) are tailored to reject the non-gesture patterns and classify the gestures based on the estimated hand key points [26]. L. Chen et al. (2020), designed a compact CNN model, which not only improves the classification accuracy but also reduces the number of parameters in the model [27]. And A. Sharma et al. (2020), proposed a novel system in which the HOG, PCA, and Local Binary Patterns are used. For preprocessing, this system used canny edge detection, ORB and bag of word technique. Several classifiers (RF, SVM, NB, Logistic Regression (LR), KNN, and Multilayer Perceptron) are passed through the preprocessed data to achieve successful performance. [27]. Table 4 shows the details of the above systems.

More than one machine learning algorithm was used in these papers, and most of them had SVM as a common factor. As for the accuracy values, two or more algorithms per paper are very close, but the highest accuracy is achieved by using SVM.

### Table 4. Comparison between the systems based on different Machine Learning techniques in [15-28].

| **Alejandro Betancourt et al. (2015)** |
|---|
| **The data set** | UNIGE-HANDS dataset |
| **Preprocessing** | With and without hands, a balanced number of frames |
| **Feature extraction** | Changes in illumination, camera motion and hands occlusions. |
| **Classification** | SVM, DT, and RF |
| **Accuracy** | SVM-HOG is of high accuracy. |

| **X. Yingxin et al. (2016)** |
|---|
| **The data set** | Cambridge Hand Gesture Dataset and self-constructed dataset |
| **Preprocessing** | Removing variable illuminations inherent in the original hand gestures data by using Canny edge detection |
| **Feature extraction** | Dynamic hand gesture that includes three primitive hand shapes and three primitive motions |
| **Classification** | CNN, SVM |
| **Accuracy** | CNN achieved a higher recognition accuracy than SVM |

| **P. Plawiak et al. (2016)** |
|---|
| Data Set | Preprocessing | Feature Extraction | Classification | Accuracy |
|----------|---------------|--------------------|----------------|----------|
| J. L. F. C et al. (2017) | Data grouped using the DG5 VH and glove device | Removing the redundant information | SVM, GA, PNN, and CNN | SVM is the best 98.24% |
| The data set | LSP | Smoothing and contrast enhancement, Segmentation and Delimitation, and extracting the object of interest | CNN | The First CNN = 95.37% and the second CNN = 96.20% |
| M. E. Benalcázar et al. (2017) | Thirty gestures were recorded for Two seconds | Reducing the noise, and smoothing every channel | The k-nearest neighbor and dynamic time warping algorithms | The model performs better (86% accurate) than the Myo system (83%) |
| D. A. Maharani et al. (2018) | 6-people for 3-distances (2, 3, and 4) m and 3-slopes position (-45, 0, 45) | Using the statistical data (sum, mean, median, and variance) from the start and end times to produce the preprocessing data. | The camera and proximity sensor contained in Kinect v2 are used to obtain a xy-z coordinate axis | SVM = 95.15%, K-Means needed 4.45ms recognition time and accuracy = 77.42% |
| M. Benmoussa and A. Mahmoudi (2018) | Vocabulary dataset | keep the data balanced | K-means and SVM classifiers | 98% |
| F. Wahid et al. (2018) | Image for 10 healthy males using a MYO armband. | Encapsulating the electrical potential of all Spatio-temporal patterns. | Normalization and DA, KNN, NB, SVM, and RF | 96.38% |
| Cheng et al. (2019) | The EMG data set | Reducing the dimension of data, counteracting the interference of noise, and possessing a considerable compression space for improving the computational capacity | spiking neural network (SNN) | 97.40% |
| B. Abhishek et al. (2020) | | In addition to one time-frequency-domain feature (standard deviation of DWT), the acquisition of four time-domain features (integral EMG, zero crossings, root mean square, and channel energy percentage) | | |
The data set
EGO dataset and Jester dataset.

Preprocessing
To compare the captured images with the dataset for the right-hand gestures movements.

Feature extraction
Specify the centroid at its geometric center that divides the picture into 2-halves.

Classification
CNN

Accuracy
Back propagation provides better accuracy.

B. Yu et al. (2020)
The data set
The Jester dataset.

Preprocessing
Calculation of the vertical and horizontal optical flow for data representation

Feature extraction
Capturing information from temporal and spatial dimensions

Classification
Two-stream CNN

Accuracy
95.77%.

T. Zhang et al. (2020)
The data set
Rendered Hand Pose (RHD) dataset

Preprocessing
Applying the machine of convolutional pose for estimating the hand pose that are capable of efficiently localizing the key points of the hand even in complex background.

Feature extraction
CNN architecture

Classification
SVM, MLP, FGMM

Accuracy
Accuracy of SVM is only 65%, and MLP is only 47.5% while FGMM can achieve 95%

L. Chen et al. (2020)
The data set
The Ninapro DB5 Dataset and the Myo Dataset

Preprocessing
Verifying the performance of the LCNN is better than CNN-LSTM

Feature extraction
Utilizing the transform of continuous wavelet for processing data as input of EMGNet

Classification
LDA and SVM

Accuracy
The accuracy was higher than traditional machine learning algorithms

A. Sharma et al. (2020)
The data set
(ASL) images

Preprocessing
Histogram of Gradients, PCA and Local Binary Patterns.

Feature extraction
Oriented Fast and Rotated Brief (ORB)

Classification
RF, SVM, NB, LR, KNN, Multilayer Perceptron

Accuracy
higher than the existing models

3. Performance Analysis
When the recognition systems of hand gestures are depending on the algorithms of machine learning that can be implemented in various domains, numerous applications are oriented. In general, neural networks and SVMs perform extremely better when working with continuous and multi-dimensions' features. However, logic-based systems achieve better when working with categorical/discrete features. For the models of neural networks and SVMs, large size of sample is needed for achieving its highest accuracy of prediction, while, NB may require somewhat small datasets.

Generally, there is an approval that KNN is more sensitive to non-relevant features. This attribute can be demonstrated via the manner the algorithm works. Furthermore, the existence of non-relevant features capable of making neural networks training non-efficient, and impractical as well. Most the
algorithms of DT are not capable of performing well with issues that need diagonal separation. The instance space division is orthogonal for the axis of one variable and parallel to the whole other axes. So, the obtaining regions after separation are all hyper rectangles. SVMs and ANNs provide well achievements when multiple collinearities is providing and a non-linear relation provides between the input and output features.

The algorithms of NB need a small space of storage through the training and classification phases, the strict minimum which represents the memory required for storing the conditional and prior probabilities. The basic algorithm of KNN utilizes considerable storage of space for the training stage, and its implementation space is at least as large as its training space. Generally, the implementation space is considerably smaller than the training space, because the obtaining classifier is mostly a highly condensed summary of the data. Moreover, KNN and NB can be utilized with ease as incremental learners while rule algorithms cannot. NB is very robust to missing values because these are easily neglected in the probabilities of computing without influence on the last decision. In contrast, neural networks and KNN need full records for completing their work.

Generally, DT and NB possess various operational profiles, if one is not accurate, then the other is high accurate and conversely. On the other hand, rule classifiers and DTs possess the same operational profile. Also, ANN and SVM possess the same operational profile. Various datasets with various variables type and instances specify the kind of algorithm that will execute well. There is no single algorithm of learning are capable of uniformly outperforming other learning algorithms through all datasets. Despite the high performance of some of the recent methods discussed in this research, hand gesture recognition is still an evolving topic that needs more experiments. The hand gesture recognition method also needs to be extended to cover all of the areas of information technology and artificial intelligence, such as tablets, smartphones, gaming consoles, smart televisions, laptops, and desktops. Applications such as home control systems, healthcare systems, gaming technologies, automobiles, televisions, home automation, and robotics are expected to be able to use hand gesture recognition to represent the communication between the user and the devices. Furthermore, some of the applications are very sensitive and in need of having a high recognition accuracy almost close to 100% to be able to use them without causing any damage or danger to human lives; such as applications of the health field, the transportation field, and the flight operation field.[28]

4. Conclusion

In the last two decades, the valuable efforts in the researches of hand gestures recognition paved the way to natural systems of human-computer interaction. There are still many unsolved issues like reliable identifying of the gesturing phase, sensitivity to speed, shape, and size variations, and problems owing to occlusion keep hand gestures recognition researches still much active. In this paper, a survey of vision-and sensors-based hand gestures recognition systems that are based on machine learning techniques reported in recent years have been provided.

The hand gestures recognition systems that are depending on the algorithms of machine learning are suffering over fit with minimal sample size and requiring suitable signal pre-conditioning for preferable recognizing. Additionally, with unseen or untrained data, the task of classification for the learning algorithm yields minimal performance.

Lots of researches on gestures recognition concentrate on pre-processing, extracting features, and selecting the suitable machine learning algorithms based on the volume of data acquired owing to its influence on the accuracy of recognition.

This paper had successfully presented the most prominent techniques, applications, and challenges in hand gesture recognition. These include the gesture acquisition methods, the feature extraction process, the classification of hand gestures, the applications that were recently proposed in the field of sign language, robotics, and other fields, the environmental surroundings challenges, and the datasets.
challenges that face researchers in the hand gesture recognition process, and the future of hand gesture recognition in building efficient human-machine interaction and Hand detection is done using OpenCV and TensorFlow object detector.

References
[1] P. Parvathy, K. Subramaniam, G. K. D. P. Venkatesan, P. Karthikaikumar, J. Varghese, and T. Jayasankar, “Development of hand gesture recognition system using machine learning,” J. Ambient Intell. Humaniz. Comput., 2020, doi: 10.1007/s12652-020-02314-2.
[2] N. Mirehi, M. Tahmasbi, and A. Tavakoli, “Hand gesture recognition using topological features,” no. January, 2019.
[3] D. Lu, Y. Yu, and H. Liu, “Gesture Recognition Using Data Glove: An Extreme Learning Machine Method,” pp. 1349–1354, 2016.
[4] V. Gajjar, “Hand Gesture Real Time Paint Tool – Box : Machine Learning Approach,” 2017 IEEE Int. Conf. Power, Control. Signals Instrum. Eng., pp. 856–860, 2017.
[5] M. J. Cheok, Z. Omar, and M. H. Jaward, “A review of hand gesture and sign language recognition techniques,” Int. J. Mach. Learn. Cybern., vol. 10, no. 1, pp. 131–153, 2019, doi: 10.1007/s13042-017-0705-5.
[6] F. Y. Osisanwo, J. E. T. Akinsola, O. Awodele, J. O. Hinmikaiye, O. Olakanmi, and J. Akinjobi, “Supervised Machine Learning Algorithms : Classification and Comparison,” no. July, 2017, doi: 10.14445/22312803/IJCTT-V48P126.
[7] T.-N. Nguyen, H.-H. Huyhn, and J. Meunier, “Static Hand Gesture Recognition Using Principal Component Analysis Combined with Artificial Neural Network,” J. Autom. Control Eng., vol. 3, no. 1, pp. 40–45, 2015, doi: 10.12720/joace.3.1.40-45.
[8] M. K. Ahuja and A. Singh, “Static vision based Hand Gesture recognition using principal component analysis,” Proc. 2015 IEEE 3rd Int. Conf. MOOCs, Innov. Technol. Educ. MITE 2015, pp. 402–406, 2016, doi: 10.1109/MIT E.2015.7375353.
[9] A. I. Maqueda, C. R. Del-Blanco, F. Jaureguizar, and N. Garcia, “Human-computer interaction based on visual hand-gesture recognition using volumetric spatiograms of local binary patterns,” Comput. Vis. Image Underst., vol. 141, pp. 126–137, 2015, doi: 10.1016/j.cviu.2015.07.009.
[10] G. P. and J. H. Terriza, “Hand Gesture Recognition based on sEMG signals using Support Vector Machines,” pp. 230–234, 2016, doi: 10.1109/ITAIC.2019.8785542.
[11] S. Sapienza et al., “On-Line Event-Driven Hand Gesture Recognition Based on Surface Electromyographic Signals,” Proc. - IEEE Int. Symp. Circuits Syst., vol. 2018-May, 2018, doi: 10.1109/ISCAS.2018.8351065.
[12] M. Tavakoli, C. Benussi, P. Alhais Lopes, L. B. Osorio, and A. T. de Almeida, “Robust hand gesture recognition with a double channel surface EMG wearable armband and SVM classifier,” Biomed. Signal Process. Control, vol. 46, pp. 121–130, 2018, doi: 10.1016/j.bspc.2018.07.010.
[13] G. Poon, K. C. Kwan, and W. M. Pang, “Oclusion-robust bimanual gesture recognition by fusing multi-views,” Multimed. Tools Appl., vol. 78, no. 16, pp. 23469–23488, 2019, doi: 10.1007/s11042-019-7660-6.
[14] G. Azzopardi and N. Petkov, “A Dynamic Approach and a New Dataset for Hand-detection in First Person Vision,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9256, 2015, pp. 274–287.
[15] X. Yingxin, L. Jinghua, W. Lichun, and K. Dehui, “A Robust Hand Gesture Recognition Method via Convolutional Neural Network,” Proc. - 2016 Int. Conf. Digit. Home, ICDH 2016, pp. 64–67, 2017, doi: 10.1109/ICDH.2016.023.
[16] P. Pławiak, T. Sośnicki, M. Niedźwiecki, Z. Tabor, and K. Rzecki, “Hand Body Language Gesture Recognition Based on Signals From Specialized Glove and Machine Learning Algorithms,” IEEE Trans. Ind. Informatics, vol. 12, no. 3, pp. 1104–1113, 2016, doi: 10.1109/TII.2016.2550528.
[17] C. J. L. Flores, A. E. G. Cutipa, and R. L. Enciso, “Application of convolutional neural networks
for static hand gestures recognition under different invariant features,” *Proc. 2017 IEEE 24th Int. Congr. Electron. Electr. Eng. Comput. INTERCON 2017*, pp. 5–8, 2017, doi: 10.1109/INTERCON.2017.8079727.

[18] M. E. Benalcazar *et al.*, “Real-time hand gesture recognition using the Myo armband and muscle activity detection,” *2017 IEEE 2nd Ecuador Tech. Chapters Meet. ETCM 2017*, vol. 2017-Janua, pp. 1–6, 2018, doi: 10.1109/ETCM.2017.8247458.

[19] D. A. Maharani, H. Fakhrurroja, Riyanto, and C. Machbub, “Hand gesture recognition using K-means clustering and Support Vector Machine,” *ISCAIE 2018 - 2018 IEEE Symp. Comput. Appl. Ind. Electron.*, pp. 1–6, 2018, doi: 10.1109/ISCAIE.2018.8405435.

[20] M. Benmoussa and A. Mahmoudi, “Machine learning for hand gesture recognition using bag-of-words,” *2018 Int. Conf. Intell. Syst. Comput. Vision, ISCV 2018*, vol. 2018-Janua, pp. 1–7, 2018, doi: 10.1109/ISACV.2018.8354082.

[21] I. Pozzi, R. Nusselder, and D. Zambrano, “Real-Time Hand Gesture Recognition Based on Electromyographic Signals and Artificial Neural Networks,” vol. 1, no. 1, pp. 284–293, 2018, doi: 10.1007/978-3-030-01418-6.

[22] M. F. Wahid, R. Tafreshi, M. Al-Sowaidi, and R. Langari, “Subject-independent hand gesture recognition using normalization and machine learning algorithms,” *J. Comput. Sci.*, vol. 27, pp. 69–76, 2018, doi: 10.1016/j.jocs.2018.04.019.

[23] L. Cheng, Y. Liu, Z. G. Hou, M. Tan, D. Du, and M. Fei, “A Rapid Spiking Neural Network Approach with an Application on Hand Gesture Recognition,” *IEEE Trans. Cogn. Dev. Syst.*, vol. PP, no. 1, p. 1, 2019, doi: 10.1109/TCDS.2019.2918228.

[24] B. Abhishek, K. Krishi, M. Meghana, M. Daaniyaal, and H. S. Anupama, “Hand gesture recognition using machine learning algorithms,” *Comput. Sci. Inf. Technol.*, vol. 1, no. 3, pp. 1734–1737, 2020, doi: 10.11591/csit.v1i3.p116-120.

[25] B. Yu, Z. Luo, H. Wu, and S. Li, “Hand gesture recognition based on attentive feature fusion,” *Concurr. Comput.*., no. May, pp. 1–9, 2020, doi: 10.1002/cpe.5910.

[26] T. Zhang, H. Lin, Z. Ju, and C. Yang, “Hand Gesture Recognition in Complex Background Based on Convolutional Pose Machine and Fuzzy Gaussian Mixture Models,” *Int. J. Fuzzy Syst.*, vol. 22, no. 4, pp. 1330–1341, 2020, doi: 10.1007/s40815-020-00825-w.

[27] L. Chen, J. Fu, Y. Wu, H. Li, and B. Zheng, “Hand gesture recognition using compact CNN via surface electromyography signals,” *Sensors (Switzerland)*, vol. 20, no. 3, pp. 1–14, 2020, doi: 10.3390/s20030672.

[28] M. Yasen and S. Jusoh, “A systematic review on hand gesture recognition techniques, challenges and applications,” *PeerJ Comput. Sci.*, vol. 2019, no. 9, pp. 1–30, 2019, doi: 10.7717/peerj-cs.218.