A Technical Review on Hand Gesture Recognition

[1]Nandani Sharma, [2]Anurag Prajapati, [3]Deepak Kumar, [4]Saurabh Kumar, [5]Shivani Dubey

Department of Information Technology
Greater Noida Institute of Technology, Greater Noida, India

Abstract: Hand gesture recognition and its numerous have an ability to communicate with machines effectively via hci interface, the system has gotten a lot of attention in recent years. This study provides an overview of state-of-the-art hand gesture recognition systems. A major goal of gesture recognition research is to develop systems that can identify and use specific human gestures to communicate data. In this paper, we take a technical review and analyze the three algorithms and compare those algorithms and find out which algorithm suits best for Hand Gesture Recognize protocol, we have taken the datasets of all of the three algorithms and compare those datasets and the algorithm which have a high accuracy amongst them will be proven as the best, we found that SVM works very well amongst all of the algorithms.

Index Terms – KNN, SVM, Naïve Bayes, Machine Learning

I. INTRODUCTION

Handshake detection for social contact is a hot topic in machine learning and computer vision research. A major aim of hand gesture is to create systems that can recognized and use particular human gestures to communicate data. We provide a comparative research in this publication. of three algorithms for static hand gesture recognition .SVM works very well and has high accuracy rate of 95.7%[1]while KNN (K-Nearest Neighbor) has accuracy rate of 91.5% [1]and Naïve Bayes has very low accuracy rate of 82.6%[1].

A. K-Nearest Neighbour

It is a simple machine learning technique that uses a supervised learning approach. Assume new cases/data are equivalent to existing cases, and assign new cases to the type that most closely resembles existing groups. Store all accessible data and classify new data points based on similarity. This means that when new data arrives, the K-NN method can be used to quickly classify it into the appropriate group. It can be used for both regression and classification, but is typically used for classifiers. This is a nonparametric method. That is, no assumptions are made about the underlying data. It is also called a delayed learning algorithm because it does not learn from experience. Instead of saving the learning algorithm immediately, save the dataset and perform actions during classification. During the training phase, the ANN algorithm simply stores the data set and classifies it into categories that are very similar to the additional information. To determine the correct K for the data, we run the KNN algorithm multiple times with different K values to reduce the number errors while maintaining the ability to make accurate predictions when the algorithm receives unseen data. Select K to decrease. As we drop the When we reduce the value of K to one, our projections become less stable. Our forecasts become more stable as we increase the value of K. recognition. Human contact, amazing gaming tech, gestures, and other applications fall under this category [2]. The 91.77 solution might be real-time or offline. created a real-time HGR systems that rely on differential image content produced using a convex lens where hand tracking has been firstly accomplished in a confined setting. In particular, a real-time biometric scanner for receiving hand photographs is built. In the detecting section of their system, they create a depth map using the SAD approach based on right-left pictures collected with a stereo camera. This system detects the foreground item and performs hand movements. The HGR algorithm then utilizes the difference in image entropy between the source images and the median picture. Studies were carried out utilizing a hand gesture data (size 240) to estimate the capability of the proposed system [3].

B. Support Vector Machine (SVM)

A classifier is a well-known pattern matching approach that uses a learning algorithm. SVM separates subspaces by class, so it can handle currently uncertain information efficiently, but it is not suitable for sample clustering. Support Vector Machines (SVM) is a class of linear machines with very good properties. In the field of pattern recognition, the main concept of SVM is to construct separating hyper planes that act as texture descriptors. A hyper plane is currently constructed to optimize the range of differences between positive and negative examples. SVM uses a systematic technique based on statistical learning theory known as structural risk aversion to reduce the maximum constraint on the error function (i.e. the object over there is the risk or the model is the training data , minimize the loss of this object). We now have an ordered set of related groups used to organize data by similarity. It consists of symbols in add to adjectives in add to permits results of the survey to get put into appropriate categories in order to create important information. There is now a handy tool for developing compile data. There is presently an abstract meta class that describes (classifies) a group of instances with similar characteristics. Support vector machine (SVM), which was introduced by Vap kin in the 1990s, is today one of the top machine learning methods. SVM is a

collection of linked supervised learning algorithms that are used for classification as well as regression. There is now guided learning. Inferring from training data data is a deep learning task. A algorithm for supervised learning examines the training data while also producing an approximated useful and essential as a class. SVM provides great efficiency, nice security provided about over fitting, and an adequate core that can now function well even when the data is not discrete in the base subspace. A novel learning machine for two group multiclass classification is the support vector network. The following concept is conceptually implemented by the machine: Furthermore, input vectors are not uniformly transferred to an extremely high - dimensional space. A piecewise linear surface is presently being built in this object's feature space. Special qualities of the decision surface guarantee that the learning machine has a strong representational power. The concept of support vector was previously developed for the limited case at whatever location the training data is now split without faults produced while being more efficient than the prior art. We may extend our work in a variety of ways because we are utilizing a simple version of this classifier. We think that our work paves the path for the development of more advanced generated models-based early identification approaches [4].

C. Naïve Bayes Algorithm

Currently exists a which is a training algorithm currently exists Built on the Bayesian network and used to handle class challenges. That item over there is now mostly utilized in text categorization and contains a tall dataset. The naïve bayes classification encoder is now one of the simplest and most successful classified methods for developing rapid models of machine learning that object there at. I am now able to make rapid forecasts. There are currently classification techniques that make predictions based on item probabilities. Spam filtering, sentiment analysis, and article classification are among the current applications of nav gradient descent. The theorem, often known as Bayes' Rule or Bayes' Law, is a mathematical formula used to calculate the probability of an assertion given prior knowledge, determined by conditional probabilities. The following formula applies to Bayes' theorem: The information now exists filtered before to the sensors signal being processed in nave classifier in add to NN method. A daily average filter is currently used to minimize noise. Participant underwent each position thirty times for data retrieval during the training phase. In the naïve bayes classification technique, the mean and standard deviation of the rms signal from the training phase are now saved in the individual's data [5].

$$P(A|B) = P(B|A) \, P(A)/ \, P(B)$$

In the above formula, hypothesis A is B. P(B|A) refers to probability. Probability: The probability of proving a hypothesis to be correct. P(A) stands for Probability Value, which is the possibility of an assertion before testimony. (B) Partial likelihood is also known as probability of evidence. Naive Bayes models fall into three types:

    1. **Gaussian**:

Gaussian models are based on the assumption that the features have a normality test. This means that the theory assumes that this morality is drawn from the Gaussian core, if the variables accept attribute variables rather than real numbers

    2. **Bernoulli**

If the data are multinomial, the Nave Bayes multinomial classifier is used. Typically used to indicate which category a particular document belongs to when it is difficult to categorize the document B. Athletics, Business or Arts. The Bernoulli classifier works similarly to the classifier predictor, except that the predictor variables are autonomous Boolean variables. For example, whether a particular term appears in the text. This approach is also known for performing text classification tasks.

II. MACHINE LEARNING

Deep learning is currently a topic of study that provides computers the ability to learn without being pattern recognition. Ml is now one of the most intriguing solutions that item that has ever come upon. As the name implies, it provides the machine that item over there the chance to learn, making it more comparable to people. Deep learning is presently in active usage, maybe in many more areas than one would imagine. Computer Software is divided into three types [6]:

A. Supervised MachineLearning:

A learning algorithm is a form of computer vision, where a computer learns from very good training data and a machine predicts an output based on that object. The tagged data shows that some input data is already tagged with the correct output. With the enrichment method, the training data provided to the machine works in the same way that the supervisor of that object tells the computer to correctly predict the output. Use the same principles that your child learns under the guidance of a tutor. There are now methods of feeding machine learning models with input data and appropriate output data, known as reinforcement methods [7].

B. Unsupervised machine learning:

Unsupervised learning, as the name suggests, is a machine learning technique in which the algorithm is not currently guided by a training data set. Instead, computers find hidden patterns and insights in available data. It became possible to exist compared to what happens in the human brain when learning new things. The training process is a form of pattern recognition that uses an unlabeled data set to create a system that can manipulate that piece of data without supervision. Learning algorithms are not readily applicable to regression or classification problems. Because, like deep classification, we have all the input data but no matching output data. The goal of unsupervised learning is to have input data but no matching output data. The goal of unsupervised learning is to discover the basic organization of data, group object data according to their properties, and present them in a compressed way over the raw data.

C. Reinforcement Learning:

Learning algorithm is presently a subfield within machine learning. It now exists in terms of taking appropriate behaviors to maximize benefit in a certain scenario. It is now utilized by many software and robots to determine the best feasible action or course to take in a given scenario. Learning algorithm differs from deep classification in that the training data in reinforcement methods contains the solution key, thus the model is already trained with right answer, whereas in relevance feedback [8].

III. LITERATURE

Several researchers suggested a method based on an introduced this concept for data-driven hand sign identification. The studies show a static hand gesture detection system utilizing digital image analysis. The author presented a real-time eyesight solution for hand sign identification in several applications that involve human contact. They demonstrated an application that aids deaf and dumb people in communicating with the world at large using hand signals. It is now utilized in virtual world control systems. Additionally, it is employed in sign translating, robot tv remote, and musical work. With the advancement of modern technology, in addition to the fact that people commonly use hand gestures to convey the motivation when communicating. Hand gesture recognition has come to be recognized as an important part of human-computer interaction (Human Computer Interaction). It allows computer capture, interpret and execute commands. The aim of this study is to conduct a comprehensive literature review to identify the main methods, applications and barriers in hand gesture identification [9]. For further than eight years, researchers have been working on recognition of hand gestures research (prasahan 2014). Jordan suggested a device that item there atdetects the number of fingers bending using hand gloves in 1977(Praveen & Shreya ,2015). Moreover, in 1983 Gary Grimes created a technique for identifying if the thumb is now touching other portion of the hands or fingertips (Laura Angelo, 2008.In 1990, given the limited computer power capacities the systems established then offered promising outcomes. The subject of recognition of hand gestures is now quite broad and a substantial portion of hand gestures is now in the previous two or three years. In this item over their study, we surveyed the most recent studies on hand gesture recognition where objects over there really were done. We will also contrast the various methods, applications, and problems given by the examined work. The rationale for whatever illogical cause the most recent research papers from the lee. databases used to existence chosen to exist researched currently exists that item over there all of us wished to create a valid basis of the current scenario in addition to hand gesture technologies recognition. According to all the several research, the results as well as an analysis of the task under consideration The next generation goes over the future of recognition of hand gestures [10].

IV. PROPOSED SYSTEM

The action recognition system gets input finger movements via the built-in webcam, which has a sensitivity of 320 x 240 megapixels. The photographs are being acquired in a scale based designed to light the reflect the presence, which is currently occupied against a black background in order to eliminate shadows. Images are now recorded at a set distance (usually 1.5 - 2 ft) between the camera and the signer. Currently, motions are made using the palm side of the right hand. The acquired video is presently being analyzed for hand motion sensors, in addition to being done using sad. Then the separation of hand now existing carried done. The segmentation hand picture is being used to find characteristics. These functionalities are presently available.
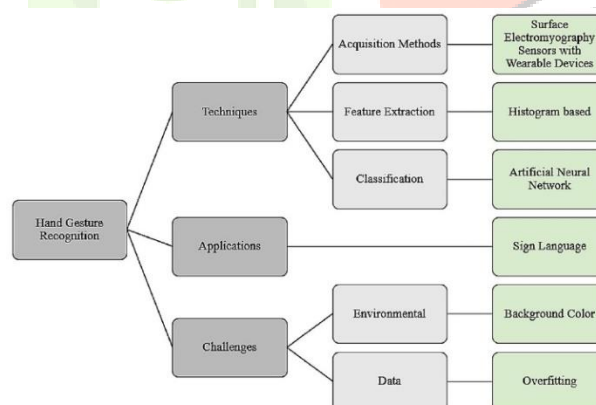

Fig.1 Structure of Hand Gesture Recognition

V. METHODOLOGY

The three stages: the first is from before the, followed by extracting features. Grading is the final stage. The first phase is hand separation, which seeks to extract hand gestures from the backdrop while also reducing sounds using specific filters. This object recognition step also involves edge information to determine the final form of the hand. The following stage, which consists of Next stage, which forms the majority of this item over here study, is currently allocated to the image retrieval problem, where two extracting features approaches, namely, hand contouring and compound moments, are now applied. Because they employed distinct ways to extract the features, namely a demarcation for hands contour and a region-based in complexity moment, these two extraction techniques were utilized in this item here to study [11]. The feature extraction techniques address issues related to hand gestures such as resizing, movement, and spin. During the classification step, wherever human brains are now utilized to recognize the motion picture based on its selected features, we all study various challenges associated to identification in contrast to the neural network computation convergence. Because of its capacity to function in parallel, wavelet transform has been widely

used as a classifier, particularly for practical uses. Application scenarios are strategies for identifying, clarifying, and organizing technical specifications in systems engineering. A use case is a collection of possible sequences of events between computers and humans in a particular environment, tied to a particular purpose of the actions performed by a user to fulfill a use case activity.
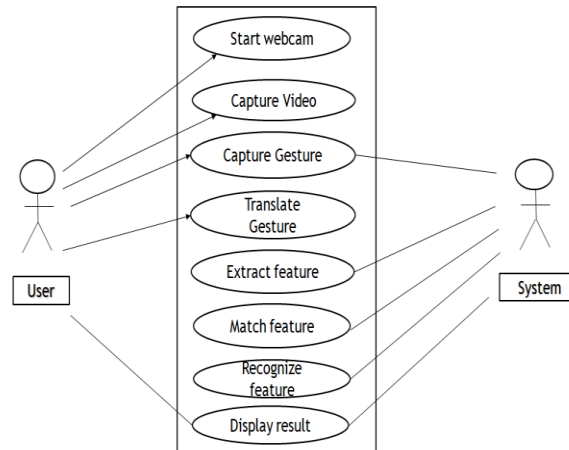

Fig.2 Use case Diagram for Feature Extraction

A. Pre-Processing and Feature Extraction

Hand segment, along with extracting features, is now a critical stage in computer vision algorithms for hand gesture recognition which was before stage prepares the input picture and extracts information that will be utilized later by the classification algorithms.

- **Hand Motion Detection**

  To determine if movement is there in the live video stream, the live video frames given by the camera must be compared using a motion detection algorithm. For motion detection, the Sum all absolute errors (SAD) approach is used in this paper. SAD is a popular simple method for determining the similarity of picture fragments. The SAD algorithm is broken down into the following steps: 1.A threshold greater than the camera noise level is established. Calculate an absolute error between every pixel in this image and the pixel in the previous image. These disparities are added together to get the Sum of absolute difference, which is then standardized. After gathering two SAD data, the variation value is determined, if the variation is less than the limit, no movement is observed. When no motion is observed, the current picture is sent to the usually manifests.

- **Segmentation**

  In this object over here step, hand currently exists segmented from the background. The image in rgb colour currently exists converted into grey scale image which currently, there is now a black and white picture that has been transformed. Filling is now applied to the picture, which results in the filling of the gaps within the hand area. Glob research is now used to extract the biggest white region in a binary picture, similar to the hand. At this existing moment in time, lateral direction presently there is conducted, during which the high system of the gestures (the lengthiest axis) presently there is recognised in add - ton to presently exists then repositioned by the correct angle to rectify it to the upright posture, so that deny there the hand sign image will appears to exist vertical. Architectural filters dilation and erosion are now utilised to soften the edges [12].

- **Feature Extraction**

  In The digital picture now existing handled using thinning operation in this object over here step, which is a morpho procedure that item over there now exists used to eliminate chosen number of pixels from binary image. This item over here generates another binary picture, the skeletal image. This picture over here is presently utilised for obtaining the end points. The endpoint now exists a point that object over there includes only one converts neighbour's in as well as representing the termination pixel of the thin segment, in accordance with the connectivity neighbours. displays the finish points.
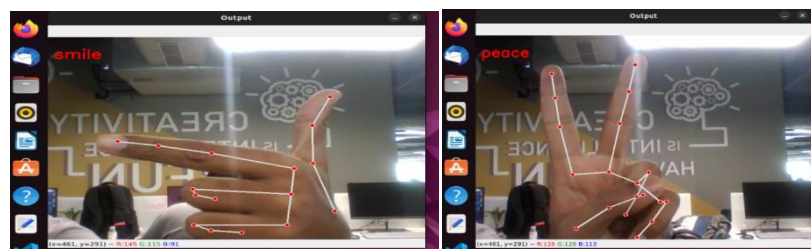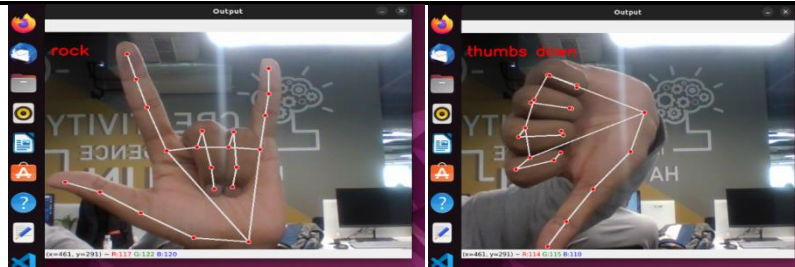
Fig.3 Feature Extraction Process

VI. COMPARISON OF DATASET (KNN, SVM & NAÏVE BAYES ALGORITHM)

K-Nearest Neighbors is currently one of the simplest machine learning algorithms based on supervised learning. In addition to classifying new cases, it assumes connections between new specific instances and existing cases, and classifies categories in which there are already objects that are most closely compared to possible classes. It collects all available data and classifies new information primarily according to similarity. The dot here means that the KNN method can be used to quickly classify new data into the appropriate category. It can now be used for both regression and classification, but is typically used for classification problems. This is a nonparametric method. That is, no assumptions are made about the underlying data. It is available today and is called a slow learning algorithm because it does not learn immediately from the training data set. It not only stores information, but also performs activities while sorting. In the learning phase, KNN simply stores the data set in order when new data is received, classifying this element there category that objects over there now existing in, which is quite comparable to the new data.
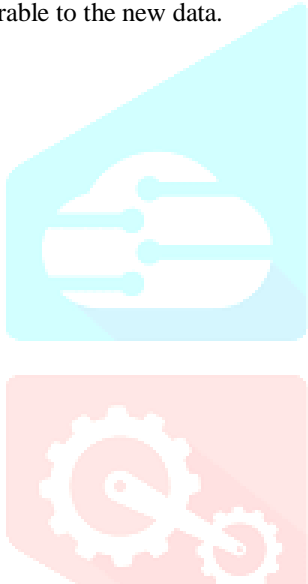
KNN DATASET

	TR UE1	TR UE2	TR UE3	TR UE4	TR UE5	TR UE6	CLASS PRECISIO
PRED1	51	0	5	0	0	0	91.07
PRED2	0	60	0	6	0	0	90.90
PRED3	3	0	34	0	0	0	91.8
PRED4	0	6	0	39	0	0	86.66
PRED5	3	0	0	0	63	0	95.45
PRED6	0	0	2	0	0	61	96.82
CLASS RECALL	94.41	95.52	86.36	83.63	92.94	96.90	

Fig.4 Feature Extraction Process

The currently existing SVM study series is an excellent approach to pattern matching. Because it splits the feature space by categories, SVM is currently good for handling unknown data, but not currently good for grouping sample data. Support vector machines (SVM) currently have linear machines with some very good properties. In the field of pattern recognition, the main concept of classifiers is to create hyper planes that act as decision boundaries. Hyper planes are now constructed such that the separation margin between positive and negative instances is maximized. SVM uses a fundamental technique known from statistical learning theory known as structural risk aversion to minimize the maintenance of misclassifications (i.e., reduce the 'risk' or 'cost' of this factor parameter estimation heuristics that look for model parameters that the model only arises from training examples). There is now an ordered collection of related groups used to organize data based on similarity. It consists of both code and descriptor. Additionally, survey responses can be broken down into logical units to generate actionable data. We now have useful tools for designing statistical surveys. We currently have an aggregate meta class that describes (categorizes) a set of instances with similar properties [13].

SVM DATASET

	TR UE1	TR UE2	TR UE3	TR UE4	TR UE5	TR UE6	CLASS PRECISIO
PRED1	59	0	5	0	0	0	92.18
PRED2	0	65	0	6	0	0	91.54
PRED3	3	0	37	0	0	0	92.5
PRED4	0	0	0	33	0	0	86.82
PRED5	3	0	0	0	67	0	95.70
PRED6	0	0	0	0	0	64	92.75
CLASS RECALL	92.85	91.85	89.36	88.63	93.03	95.70	

Fig.5 SVM Data Set

The naive bayes method is a supervised learning technique that is based on the Bayes rule and is used to solve classification issues. That product over is indeed currently largely used for text classification and comes with a large training set. That item over there is now mostly used in text categorization and contains a steeply training dataset. The naive bayes classifier is presently one of the simplest and most successful classification algorithms for developing rapid models of machine learning that can generate quick predictions for objects. There is now a classification algorithm, which predicts based on the likelihood of an item. Some notable instances of naive bayes algorithm now include spam filtration, emotive evaluation, in additional to classification articles [13].

NAÏVE BAYES DATASET

	TR UE1	TR UE2	TR UE3	TR UE4	TR UE5	TR UE6	CLASS PRECISIO
PRED1	56	0	5	0	0	0	93.38
PRED2	0	62	0	6	0	0	91.17
PRED3	3	0	34	0	0	0	91.80
PRED4	0	0	0	33	0	0	86.82
PRED5	3	0	0	0	65	0	95.58
PRED6	0	2	0	0	0	61	96.80
CLASS RECALL	90.32	96.87	87.17	84.61	93.03	95.70	

Fig.6 Naïve Bayes Data Set

VII. RESULT ANALYSIS BASED ON ALGORITHM'S DATA SET

Accuracy currently exists There is just one measure for assessing prediction model. Loosely, quality is the percentage of predications made by our model. Technically, precision is defined as: For a classifier, accuracy is the proportion of correct predictions divided by the total number of decisions. Accuracy can be derived in the form of strengths and weaknesses as follows:

Accuracy = tp + tntp + tn + fp + fn, where tp = true positives, tn = true rejects, fp = false positives, and fn = false negatives.

Table.1 Comparison of Accuracy Rate of algorithms

Classifying	KNN	Naïve bayes	SVM
Accurate (%)	91.5	82.9	96.7

CONCLUSION & FUTURE SCOPE

In this object over here currently are discussed for hand gesture recognition, these methods currently are svm, naïve bayes in addition to ken (k-nearest neighbors) the experimentation currently exists tested on various dataset which justifies that object over there the batch type surpasses methodology by becoming resistant to scale variation and without requiring any predetermined template for acknowledgements. works really well in hand gesture identification when combined by the other two systems ken and nave classifier. the evaluation object over here contract currently exists to Create a real-time motion categorization system such that we can recognize motions in natural illumination conditions. In order to reach this aim, a real-time display of affection system for identifying moves is now being devised. Make disable person capable to communicate with able person. Make a system to establish a way of sharing thought in addition to ideas of disabled person. Use to detect, recognizes in addition to interpret the hand gesture through computer vision so that object over there they also currently are able to communicate with others.

REFERENCES

[1] Georgians, in addition to a.m. Petrie. On time action one to report or move accordingly.

[2] Mitra,S. andT.Acharya,Gesturerecognition: ASurvey according to the machine learningManandCybernetics2007.

[3] In addition to rds. Jadon, a review of Vision based hand gesture for real time reporting 2009. 2(2): p. 405-410.

[4] Faria, b.M., n. Lau, in addition to l.P... 2009. Pivo de varia, Portugal.

[5] gillion, net, Gesture Recognition for Musician Computer Interaction, in Music Department2011, faculty of arts, Humanities in addition to Social Sciences: Belfast. p. 206.

[6] Faria, b.M., et al., Machine learning Approaches and facial expression (cis)2010:

[7] mannini, a. in addition to a.M. Sabatini, classify human activity according to the presence of mind, 2010.

[8] Maldonado-Báscon, S., et al., Based Character Recognition neural networks 2007.

[9] E. Stergiopoulou, N. Papamarkos. (2009). "Hand Gesture Recognition Using Neural Network Shape Adaptation Techniques," Elsevier Engineering Applications for Artificial Intelligence.

[10] Mokhar M. Hasan、Pramod K. Mishra、(2012)

[11] N. A. Ibrahim, R. Z. Khan (2012). Vision-Based Gesture Recognition Using Neural NetworksApproaches:ReviewInternational Journal of Human Computer Interaction

[12] P. Molchanov, S. Gupta, K. Kimand J."Hannung mit 3D-Convolutional Neural Networks", Proceedings of IEEE Conference on Computer Vision 2015

[13] L. Perez und J. Wang, Effectiveness of data Augmentation in image Classification using deep learning ", arXiv-Preprint arXiv:1712.04621.2017.