IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **Volume 11, Iss 7 , 2022**

A Review on Hand Gesture Recognition Using ML and DL Based Techniques

M Ravi Kumar[1], J Lakshmi Prasanna[2], Chella Santhosh[3], Parsipogu Glory Veronica[4]

[1,2,3,4] Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur (Dist), Andhra Pradesh, India – 522302

DOI : 10.48047/IJFANS/V11/ISS7/312

Abstract. This paper aims to present and describe several approaches to hand gesture detection via human- Interaction with a computer Computer-human interaction is a crucial part of the majority of people's day-to-day lives, especially when persons who are dumb or deaf of hearing want to communicate. In these situations, getting their message over to the rest of the world is a challenge. Therefore, these individuals can complete various jobs with ease by making an air gesture with their hands in front of a computer camera, which reads their movements into words or text that other individuals can understand. We have spoken about certain Human-Computer Interaction (HCI) strategies and methodologies that will this study's construction of the hand gesture detection system comprises two significant issues. The detection of hands comes first. Making a sign that can be utilized with only one hand at a time is another issue. This study's objective is to develop a system that can recognize, & understand hand gestures using computer vision while dealing with problematic elements including stance, orientation, position, and scale variability. This system must produce a variety of gestures, including numbers and sign language, for this project to be successful. Before any image processing is done, a webcam image is acquired and analyzed to find a hand motion. In this research, methods based on ML and DL will be used to detect hands.

Keywords: Human-Computer Interaction (HCI), Hand Gestures, ML, DL, Computer Vision, Sign Language Recognition (SLR), Feature Extraction, Classification, Pattern Recognition.

1. Introduction

The study of hand gesture recognition is a hot area right now. Experiments to validate the results of the system and application have been carried out due to the rising importance and interest in gesture detection. We have compared several strategies and methodologies used in various systems and applications in this article. These systems taught us fresh methods for identifying emotions. Although speaking is the primary mode of communication between people, there are other non-verbal forms of contact in addition [1]. The goal of developing a hand gesture recognition system is to establish a true human-computer interface via which gestures may be utilized to communicate or monitor a robot [2]. To convey information, hand gestures are used. This study illustrates many techniques for identifying and understanding hand gestures as a language for efficient communication. Recognition of hand motions using depth map technology, segmentation, and a successful marker-free approach identifying motions recorded by RGB-D cameras [3,4], EMG monitoring and depth-sensing camera recognition [5], Machine learning algorithms are being used to identify gestures based on wireless sensors, webcams, and a real-time tracking technique [6] all important and useful subjects in [7]. We can overcome the communication gap between humans and computers and make work simpler if computers could comprehend human hand movements or gestures.

Through the placement of the fingers, the center of the palm, and the shape of the hand, hand gestures constitute a form of body language. Static and dynamic hand gestures are the two categories under which they fall. The static gesture, as the name suggests, refers to the fixed shape of the hand, whereas the. Between posture and gesture, the main contrast is that, whereas gesture emphasizes hand movement, posture emphasizes the curvature of the hand. Previously,

ISSN PRINT 2319 1775 Online 2320 7876

to recognize hand motions, wearable sensors that were directly attached to gloves were put to use on the hand. Based on hand motions or finger bending, these sensors picked up a biological response. A computer that was connected to the glove was used to process the data. By using a sensor connected to a microcontroller, this glove-based sensing system might be made portable. Since the development of the data glove sensor, hand gestures for human-computer interaction (HCI) have been used. It provided easy commands for the computer interface. The precise coordinates of the positions of the palm and fingers were calculated by the gloves, which made use of a variety of sensor types to record hand motion and location. [8]. Flex sensors [9], The curvature sensor [10], optical fiber transducers [11], accelerometer sensor [12], and angular displacement sensor [13], are merely a few instances of sensors that employ the same bending angle-based methodology.

Although the techniques have produced fantastic results, they have several drawbacks that make them inappropriate for the elderly, who could feel anxious and confused because of cable connection issues. Those who have sensitive skin or burn victims may potentially have skin injury, infection, or unfavorable responses because of these sensors. A few sensors are also quite expensive. In Lamberti and Camastra's study [14], who created a computer vision technology based on colored marking gloves, some of these issues were overcome. Although no sensors were to be attached for this investigation, colorful gloves had to be worn, nonetheless. Due to these limitations, promising and affordable techniques that did not involve wearing bulky gloves were created. The term “camera vision-based sensor technologies” refers to these methods. Since open-source software libraries have developed, recognizing hand signals for application in a range of situations, such as clinical operations, is now easier than ever. [15], residence automation [16], playing games [17], virtual sceneries [18], signed language [19], tablet devices and personal computers [20], and control of the robot [21]. In essence, these methods require switching out the camera-instrumented glove. A range of cameras, including RGB cameras, TOF cameras, IR imaging devices, and night vision equipment, are used to accomplish this.

Computer vision-based algorithms have been developed for use with a variety of cameras. Algorithms try to separate & identify elements of the hands, including skin tone, look, Deep learn identification, motion, skeleton, depth, 3D model, and others. The camera kind, distance restrictions, or recognition rate are not mentioned in any review papers. To identify and classify hand gestures using a variety of technologies. This paper's goal is to provide a comparative evaluation of contemporary computer vision methods. HCI uses hand gestures in many different ways, including to speed up computer communication, create a welcoming and appealing atmosphere that will attract users, establish a non-physical connection for comfort and safety, use the computer while keeping a distance from the user, can manage complicated and virtual settings more easily [22]. Since there are several hand gesture combinations, a distinct set of gestures is utilized for each application's activities.

Additionally, from a computer's point of view, the ability to recognize hand movements is influenced by the environment (such as light, backdrop, distance, and skin color) as well as the orientation and position [23]. The stages involved in hand gesture recognition techniques range from straightforward to sophisticated. The steps are frequently organized as follows to arrive at the final gesture: first-person gesture recording, hand following, component extraction & last categorization. For applications that are related to human safety, there are many readily available methods, such as using hand signals to control and guide airplane operations. Applications which are meant to be fun, like those that use hand gestures in gaming, are also common [24].

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, **Volume 11, Iss 7 , 2022**

2. Literature Survey

Due to its adaptability and user-friendliness, one aspect of human-computer interaction is hand gesture recognition which is currently being researched. Gesture recognition is a method that reads and interprets hand gestures as commands using sensors. When speaking, people frequently make gestures with their hands, arms, or other body parts to emphasize or highlight certain ideas. To put it another way, a gesture is a physical movement that expresses meaning. Waving your hand to welcome someone is a typical example. Given the multiple uses for hand gestures in virtual reality, sign language recognition, etc., the detection of hand gestures is crucial for HCI.

2.1 Hand Gesture Recognition Analysis

Humans are capable of recognizing hand gestures through their actions. However, it is extremely difficult for a machine to recognize a human gesture. This article has addressed a few earlier studies. A simple rule classifier may be used to identify the hand motion once the fingers have been observed and identified. The rule classifier predicts the hand motion based on the number and distribution of fingers found. What fingers are identified depends on the content of the fingers.

2.2 Hand Gesture Recognition Using Webcam

This research offered a system that is meant to accept human gestures [25] as a control input for a computer program to create a new type of computer-human interaction or communication. Proposed a technique that makes use of a camera to record user motions [26], analyze them, and activate the functionality related to that particular gesture [27]. This system consists of 4 steps. The motions made possible by hand gestures in this system are accomplished using the OpenCV [6,7] library, while the Python modules utilized include PyAutoGUI [28,29] and NumPy [30,31]. Their Giono interest (ROI), which acts as a favored area while avoiding the surrounding area, will be used as the primary concept in processing the web camera image known as the backdrop [32].

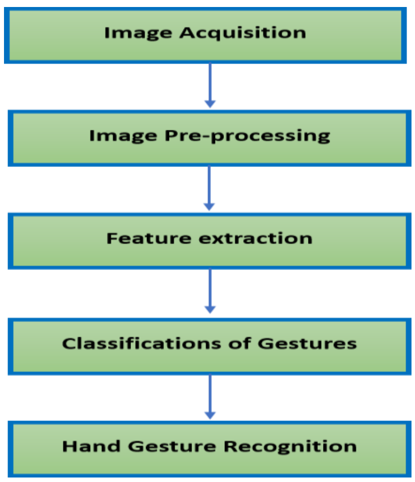

Figure 1 Hand Gesture Recognition Method

.

3. BLOCK DIAGRAM

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 11, Iss 7 , 2022

3.1 Image Acquisition

The process of removing a picture from a source, which is often a piece of hardware like a camera, sensor, etc. source for processing is referred to as image acquisition in machine vision. It is the initial stage in the workflow sequence since processing cannot take place without an image. The captured image is entirely unprocessed. The necessity to understand how images are being taken and kept in memory has developed. The most crucial step in dealing with images is to capture them before analyzing them. We refer to this as image acquisition. An appropriate camera is used to acquire images.

3.2 Image Pre-processing

Before being utilized for model training and inference, images must first undergo image pre-processing. This includes but is not limited to, adjustments to the size, orientation, and color. Although geometric transformations of images (such as rotation, scaling, and translation) are categorized here as pre-processing methods because similar techniques are applied, pre-processing is a step in the image-processing process that aims to improve the image data by suppressing unintentional distortions or enhancing certain features that are essential for subsequent processing.

3.3 Feature extraction

Deep learning and machine learning attribute extraction. Feature extraction is a process that converts unprocessed data into numerical features that may be analyzed while. It generates better outcomes compared to machine learning which is applied directly to the raw data. Pre-processing, feature extraction, and classification are all applied in three stages to the hand gesture image. The pre-processing stage prepares the image for the feature extraction stage by separating the hand motion from its context. Building user-friendly interfaces requires a variety of approaches, one of which is gesture detection. Usually, gestures can come from any body movement or position, although they frequently start with the hand or the face. Users may interact with the devices without touching them due to gesture detection.

3.4 Classifications of Gestures

3.4.1 Machine Learning Models:

The gesture is recognized using the gesture classification approach after modeling and analysis of the input hand picture. The appropriate feature parameter selection and appropriate classification method have an impact on the recognition process [33]. The machine learning models used in the tests (DT) included K-Nearest Neighbors (KNN), Logistic Regression (LR), Gaussian Naive Bayes (GNB), Multilayer Perceptron (MLP) with one hidden layer, Random Forest (RF), Decision Tree, Artificial Neural Network (ANN), Convolution Neural Network (CNN), and Support Vector Machine (SVM).

3.4.2 K-Nearest Neighbors (KNN)

The k-nearest neighbors algorithm, often known as KNN or k-NN, is a supervised learning classifier that makes predictions or classifications about how a single data point will be grouped. It can be used to address issues with classification or regression, but because it is based on the idea that adjacent similar points can be located, it is commonly employed as a classification approach. K-Nearest Neighbour is one of the most basic supervised learning-based machine learning approaches. Based on the presumption that the new case and the data are comparable to the instances that are already available, it assigns the new case to the category that resembles the existing categories the most. The K-NN algorithm categorizes new input based on similarity and records all available data. This indicates that new data can be reliably and quickly categorized using the K-NN approach. Although it may also be utilized for classification issues, most of the time it is employed for regression problems. It makes no assumptions about the underlying data because it is non-parametric. Since it keeps the dataset rather than immediately applying what it has learned to the training set, it is also known being a slow learner algorithm. However, while

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **Volume 11, Iss 7 , 2022**

classifying data, it uses a data set to carry out an operation. KNN categorizes new data into a category that is quite close to the training data by merely saving the dataset during the training phase.

The KNN method is a supervised ML technique for gesture detection. KNN is utilized for classification, and each data point's categorization is influenced by the classification of its neighbors. In KNN, the closest neighbors are determined using the Euclidean distance [34]. This calculation is done using several tiny distances to achieve the minimal Euclidean distance [35]. Accuracy rises concurrently with the k-value increase. The Euclidian distance formula is often applied. When utilizing KNN, utilizing a threshold value created by averaging the nearest data points. The performance is entirely determined by a threshold value, a similarity measurement, and the distance to the nearest neighbor. The quantity of neurons in the covert layer determines the hidden layer's size, which results in an accuracy assessment.

3.4.3 Logistic Regression (LR)

It is among the most well-known Machine Learning algorithms used with the Supervised Learning approach. Using a specific collection of independent factors is utilized to forecast the categorical dependent variable. The outcomes of a dependent variable with a categorical component are predicted by logistic regression. The outcome must therefore be a discrete or categorical value. It offers the probabilistic values that lie between 0 and 1 rather than the exact values between 0 and 1. It can be either True or False, 0 or 1, or Yes or No. The method of application is the primary distinction between logistic regression and linear regression. In contrast to logistic regression, which is used to address regression-related issues, linear regression addresses classification-related issues. Logistic regression, which focuses on demonstrating how informative factors relate to a discrete response variable, combines regression methods. The difference between conventional straight regression and logistic regression, the response variable Y is shown to be continuous. The response variable in logistic regression is discrete. [36]. The probability-based logic of logistic regression, which predicts values between 0 and 1, is referred to be a predictive analytical technique. To convert the outcome into categorical numbers, it makes use of the sigmoid activation function. This sigmoid function is also known as a logistic function.

3.4.4 Gaussian Naive Bayes (GNB)

Based on the Bayes theorem, the probabilistic machine learning technique known as Naive Bayes is employed for numerous classification applications. The generalization of naive Bayes is called Gaussian naive Bayes. The Gaussian or normal distribution is the most straightforward to implement among the several functions used to estimate data distribution, as you just need to calculate the training data's mean and standard deviation. A variation of Naive Bayes that handles continuous data and adheres to the Gaussian normal distribution is called Gaussian Naive Bayes. As a result, continuous-valued features are supported, and each is modeled as following a Gaussian (normal) distribution. Assuming that the data is characterized by a Gaussian distribution with no covariance (independent dimensions) between dimensions is one method for building a straightforward model. This model fitted by determining the average and standard deviation points inside each label, which is all that is necessary to create such a distribution.

The Naive Bayes variation known as Gaussian Naive Bayes supports continuous data and adheres to the Gaussian normal distribution. This allows features with continuous values and models them all as fitting into a Gaussian (normal) distribution. Assuming that the data is distributed as a Gaussian distribution with independent dimensions—covariance—allows for the creation of a straightforward model. To define such a distribution, obtaining each point's mean and standard deviation within each label is all that is necessary, which can be used to fit the model. It only utilizes continuous data and supports the Gaussian normal distribution. Furthermore, it uses supervised machine learning and is grounded in the Bayes theorem [37]. Since only the mean and standard deviation from the training dataset need to be estimated, this

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, **Volume 11, Iss 7 , 2022**

one is easy to utilize. The probability of each class frequency will be used to calculate the input values. Each class's mean and standard deviation must also be noted and saved.

3.4.5 Multilayer Perceptron (MLP)

This restriction was addressed with the development of the Multilayer Perceptron. It is a NN with a non-linear input-to-output mapping. A multilayer perceptron has one or more hidden layers, an input, and an output, each of which has multiple neurons stacked on top of one another. Multilayer perceptron neurons can have any arbitrary activation function, unlike perceptron neurons, which must have an activation function that enforces a threshold, such as ReLU or sigmoid. Given that inputs and starting weights are mixed in a weighted sum and are both subject to the activation function, Multilayer Perceptron falls within the category of feedforward algorithms. However, the distinction is that each linear combination is carried over to the following layer. Each layer's internal representation of the data and output is communicated to the layer below it. This continues until the output layer after passing through all hidden levels.

For gesture categorization, the multi-layer perceptron (MLP) network was employed [38, 39, 40, 41]. To determine finger positioning, a direct gesture categorization method employing an MLP network was used. The wrist position was used to standardize the standardize fingers' position. The normalization by the middle of the wrist &hand distance from the camera further enables the neural network's capacity to be independent of the hand's position and orientation. The logical sigmoid function was employed in the network's hidden and output layers. Network entrances were recorded as in the database normalized fingertip's locations in the range of (0,1), and using this information, MLP achieved categorization. Every network output contained a number indicating the group membership rate in the range (0, 1) The improved back propagation mistakes approach with changeable learn rate and momentum parameter was employed for the MLP network training.

3.4.6 Random Forest (RF)

Favored algorithm for machine learning A component of the supervised learning approach is Random Forest. It can be applied to address classification & regression-related ML problems. It is based on the concept of ensemble learning, a technique for combining numerous classifiers to enhance models and solve difficult situations performance. The RF classifier, as its name suggests, averages several decision trees applied to various subsets of the input dataset to improve the predictability of the dataset. The random forest uses predictions from each tree and predicts the outcome based on the votes of most projections rather than relying solely on one decision tree. The greater number of trees in the forest prevents higher accuracy and overfitting. Given the various trees are used by the random forest to estimate the dataset's class, some decision trees may predict the right output while others may not. But when all the trees are considered, as a result, the following are two hypotheses for an improved Random Forest classifier: The feature variable in the dataset should have some real values so that the classifier can forecast correct outcomes rather than an assumed result. There must be extremely little connection between each tree's predictions. A classifier called a random forest [36,42,43] decision trees are used in succession to label the sample data. To evaluate the algorithm's performance, there are many decision trees available. Additionally, the algorithm created each of these trees to generate a forecast. Random forest methods are chosen because of their benefits, such as their effectiveness in big datasets, lack of overfitting, use of numeric and categorical variables, ease of application in multi-class environments, and need for fewer parameters than other reducing techniques.

3.4.7 Decision Tree (DT)

Although it is frequently used for dealing with classification challenges, a supervised learning technique called decision trees can be used to deal with regression & classification issues. It is a tree-structured classifier, where each leaf node represents the classification outcome and inside

ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, Volume 11, Iss 7 , 2022

nodes represent the features of a dataset. The Decision Node and Leaf Node are the two nodes of a decision tree. Making decisions is done via decision nodes, which have numerous branches, Leaf nodes, on the other hand, are the results of choices and do not have any further branches. The features of the submitted dataset are utilized to run the test or reach conclusions. It is a graphical representation for assembling every viable solution to a decision or issue based on established circumstances. A decision tree is what it is named because, like a tree, the root node is the starting point, and it grows on succeeding branches to take the form of a tree-like structure. The Classification and Regression Tree Algorithm, or CART algorithm, is used to build a tree. A decision tree just asks a question and then categorizes the answers (Yes/No) into subtrees. Choosing the best approach for the dataset and task at hand is the most crucial consideration when creating a machine learning model. The following are the two justifications for using the decision tree: Decision trees are frequently created to closely mirror how individuals think while making decisions, making them easy to understand. Due to its tree-like structure, the decision tree's reasoning is understandable. The learned function is represented by a decision tree in the discrete-valued target function approximation technique known as decision tree learning [44]. The foundational principle of building a tree is to identify a characteristic that maximizes the information gained for each node.

3.4.8 Artificial Neural Network (ANN)

The biological neural networks that determine the structure of the human brain are the source of the term "artificial neural network". The same as how neurons are interconnected in the real brain, ANN'slikewise includes different layers of the networks, neurons that are connected. Nodes are the name given to these neurons. ANNs, which can replicate the neuronal network found in the human brain, are employed in artificial intelligence to give computers the capacity to interpret information and make judgments in a manner that is comparable to that of a person. An ANN is created by computer programming to operate exactly like a network of connective tissue in the brain. The human brain contains about1000 billion neurons. Between 1,000 to 100,000 association points are present in each neuron. Data is distributed and stored in the human brain, allowing us to simultaneously access many pieces of information from memory as needed. The human brain is said to contain a huge number of incredible parallel processors. Take a look at an illustration of a digital logic gate with inputs and outputs. Two inputs are required for the "OR" gate. If either one or both of the inputs are "On," the output will also be "On." If both inputs are "Off," the output will also be "Off." In this case, the output is dependent on the input. Our brains do not carry out the same function, because our brain's neurons are always "learning the connection between outputs and inputs is constantly changing.

A network of FFNNs is a linked node that can store and provide access to material. The backpropagation method is a popular choice for neural network training. An error value is computed once the output produced is compared to the desired one. The backpropagation process, which uses this computed error number to start, propagates this error back to the network and modifies the weights to lower the error at each try [45]. ANN is a straightforward electrical model that resembles the neural network in the human brain. Learning occurs in the brain as a result of human experiences [46], [47]. ANN is a system that functions similarly to the biological nervous system in how it processes information. The system is made up of a sizable number of unified processing components that collaborate to find solutions [46], [47]. It is particularly constructed for learning processes that classify data or recognize patterns [46], [47]. Some of the numerous advantages of employing ANNs include their capacity for self-learning and their ability to manage massive amounts of data. A specialist in the study of neural networks that have been taught to analyze the data that has been presented to it.

3.4.9 Convolution Neural Network (CNN)

Convolutional neural networks, also referred to as CNNs or convnets, are a component of machine learning. It is one of several artificial neural network models used for various tasks and data sets. For deep learning algorithms, a CNN is a specific kind of network design that is used for tasks like image recognition and pixel data processing. Although there are different kinds of

ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, Volume 11, Iss 7 , 2022

neural networks in deep learning, CNNs are the preferred network design for identifying and recognizing objects. They are therefore ideally suited for computer vision (CV) activities and for applications where accurate object identification is crucial, such as face and self-driving automobile systems. A CNN is a particular kind of deep neural network created for analyzing organized arrays of data, such as photographs. In the input image, lines, gradients, circles, or even eyes and faces are very well recognized by CNN. It is because of this feature that convolutional neural networks are so reliable for computer vision. It does not require any preparation and may run straight on an underdone image. A convolutional NN is an FFNN, which hardly ever has up to 20 layers. Convolutional neural networks' strength comes from a particular type of layer called the convolutional layer. Each of the convolutional layers in this CNN architecture can recognize increasingly complex forms. Three or four convolutional layers are sufficient to recognize handwritten numerals, but 25 layers are required to recognize human faces. The goal of this field is to train machines to perceive the environment similarly way humans do, and to utilize that understanding for a variety of tasks, a few examples are image and video recognition, image analysis and classification, media reconstruction, recommendation engines, and natural language processing. A multi-layered feed-forward neural network, which is created by stacking numerous undetectable layers on top of one another in a particular way, is used to create a convolutional neural network. Given its sequential design, CNN might be able to learn hierarchical properties.

Convolutional layers are frequently used in CNN, followed by activation layers, grouping layers, and hidden layers. The pre-processing is necessary for a ConvNet, which is identical to that of the corresponding pattern of neurons in the human brain, inspired by the organization of the visual cortex. Due to its capacity for gathering the fundamental element values based on the informational image familiarizing the reader with the comparison of various instances by using a variety of examples in its preparation, CNN is renowned for recognizing and has preferred results over other procedures [48,35]. In any event, the speed of equipment computing has traditionally limited its growth. Recently, as semiconductor manufacturing has advanced, illustration preparation units' computation speeds have increased, and the CNN Network has been able to expand faster due to a constraint in equipment handling speed [49]. Applying this CNN involves the following steps: first, an image is inputted; Additionally, processing and filtering must be carried out; & third, following the classification, the results are acquired. Before being built as a layered architecture with many convolutional layers using kernels (or filters) and a pooling procedure, each model must first be trained, then tested.

3.4.10 Support Vector Machine (SVM)

Classification and regression problems are resolved using SVM, among the most popular supervised learning methods. However, Machine Learning Classification issues account for the majority of its usage. To swiftly categorize new data points in the future, The SVM algorithm seeks to specify the optimal decision boundary or line that may categorize an n-dimensional space. This optimal decision boundary is known as a hyperplane. The extreme points and vectors that are useful to create the hyperplane are chosen via SVM. The SVM approach is based on support vectors, which are used to illustrate these extreme circumstances. A supervised machine learning methodology is the Support Vector Clustering (SVC) approach. The kernel-based learning paradigm is the foundation of SVC. SVC analyses a pattern using both regression and classification methods. To find a hyperplane, SVC divides a data set into two different data classes [34]. For us, this hyperplane acts as a binary classifier. We must first look for the closest hyperplane data points that can employ this information as a support vector in this. These vectors are extremely important for the growth of hyperplanes [49]. The position and appearance of the hyperplane are also impacted by support vectors.

3.5 Method of Hand Gesture Recognition for Human-Computer Interaction (HCI)

In this study, a strong and transparent plan is given. The skin color and labeling method is used for hand recognition; it separates the hand region from the backdrop [50]. The patterns of hand gestures will determine how a signature is generated [50, 51]. The final stage will involve

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **Volume 11, Iss 7 , 2022**

classifying the class labels because this is predictable. The initial set of data [52] from hand photographs captured with a regular camera under the same conditions [53] was applied in this experiment. The background of the image should be very clear because it is the major emphasis [54]. However, the effectiveness of the strategy suggested in this paper depends on how well hand gestures are detected [54].

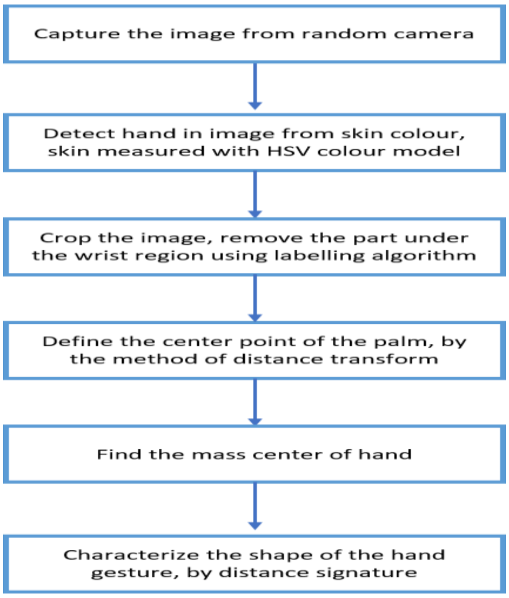

Figure 2. Algorithm for Proper Detection of Hand Gesture.

5. Conclusions

In this paper, there are several approaches and strategies for gesture recognition, including KNN, LR, GNB, MLP, RF, DT, ANN, CNN, and SVM for feature representation. Recognition of hand gestures rectifies a problem with interface systems. Since there is no need for hardware devices, controlling things manually is more natural, simple, adaptable, and affordable. It's not necessary to resolve hardware device problems either. The preceding sections made it evident that using a camera sensor to build reliable and resilient algorithms required certain criteria for how to manage frequent problems and provide a trustworthy result. Each of the aforementioned approaches has benefits and drawbacks, and they may succeed brilliantly in some circumstances while completely failing in others. Even to capture the structure of the hand, several approaches and algorithms are needed for feature extraction. Depending on the application required, a particular recognition algorithm is chosen. The gesture system's application areas are presented in this work.

The conclusion is that there are several models reported in the scientific literature that attempts to utilize ML to address the issue of HGR. The research results from each of these models are shown, along with the importance of classification accuracy and processing speed in some works. The dataset creation varies depending on the number of samples collected, the form, origin, and type of sensor used, and the types and numbers of gestures that the models analyzed. This work cannot be compared as a result. The models also show several techniques that are applied in the modules for data gathering, pre-processing, feature extraction, classifiers, and post-processing. Thus, the parameters in the modules for data collecting, post-processing, and post-processing were changed to tackle the issue from the standpoint of supervised learning.

References

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **Volume 11, Iss 7 , 2022**

1. Listed in alphabetical order (A-Z) of author's last name, year. Title of the Article. Journal's 1. Z. Zafrulla, H. Brashear, T. Starner, H. Hamilton, and P. Presti, "American sign language recognition with the Kinect," in Proceedings of the 13th international conference on multimodal interfaces. ACM, 2011, pp. 279–286.
2. Chen WL, Wu CH, Lin CH. "Depth-based hand gesture recognition using hand movements and defects". In: Next-Generation Electronics (ISNE); 2015 International Symposium on; 2015 May 4; Taiwan, Taipei; New York: IEEE;2015; p. 1-4.
3. Ohn-Bar, E. & Manubhai Trivedi, M. (2014). "Hand gesture recognition in real time for automotive interfaces: A multi-modal vision-based approach and evaluations". IEEE Transactions on Intelligent Transportation Systems, Vol. 15, pp. 2368–2377.
4. Z. Ren, J. Yuan, J. Meng, and Z. Zhang, — "Robust part-based hand gesture recognition using Kinect sensor," IEEE Trans. Multimedia, vol. 15, no. 5, pp. 1110–1120,2013.
5. Itseez, "Open-Source Computer Vision Library" - itseez2015opencv, 2015, accessed at https://github.com/itseez/opencv on 01/2/2017.
6. Ahuja, M.K.; Singh, A. "Static vision-based Hand Gesture recognition using principal component analysis". In Proceedings of the 2015 IEEE 3rd International Conference on MOOCs, Innovation, and Technology in Education (MITE), Amritsar, India,1–2 October 2015; pp. 402–406.
7. Shrote, S.B.; Deshpande, M.; Deshmukh, P.; Mathapati, S. "Assistive Translator for Deaf & Dumb People". Int. J. Electron. Commun. Comput. Eng. 2014,5, 86–89.
8. Kramer, R.K.; Majidi, C.; Sahai, R.; Wood, R.J. "Soft curvature sensors for joint angle proprioception". In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; pp. 1919–1926.
9. Fujiwara, E.; dos Santos, M.F.M.; Suzuki, C.K. "Flexible optical fiber bending transducer for application in glove-based sensors". IEEE Sens. J.2014, 14, 3631–3636. [CrossRef]
10. Gupta, H.P.; Chudgar, H.S.; Mukherjee, S.; Dutta, T.; Sharma, K. "A continuous hand gestures recognition technique for human-machine interaction using accelerometer and gyroscope sensors". IEEE Sens. J. 2016, 16, 6425–6432. [CrossRef]
11. Jesperson, E.; Neuman, M.R. "A thin film strain gauge angular displacement sensor for measuring finger joint angles". In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, New Orleans, LA, USA, 4–7 November 1988; pp.807–vol.
12. Lamberti, L.; Camastra, F. "Real-time hand gesture recognition using a color glove". In Proceedings of the International Conference on Image Analysis and Processing, Ravenna, Italy, 14–16 September 2011; pp. 365–373.
13. Wachs, J.P.; Kölsch, M.; Stern, H.; Edan, Y. "Vision-based hand-gesture applications". Commun. ACM 2011, 54,60–71. [CrossRef]
14. Desai, S.; Desai, A. "Human-Computer Interaction through hand gestures for home automation using Microsoft Kinect". In Proceedings of the International Conference on Communication and Networks, Xi'an, China,10–12 October 2017; pp. 19–29.
15. Kaur, H.; Rani, J. A review: "Study of various techniques of Hand gesture recognition". In Proceedings of the 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 4–6 July 2016; pp. 1–5.