International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Design of an Ensemble learning Model to Recognize Multilingual Hand Gestures via Interlinked Key Point analysis

Siddhant Kumar[1],Pallavi Parlewar[2],Vandana Jagtap[3]

Submitted: 24/05/2023 **Revised**: 10/07/2023 **Accepted**: 29/07/2023

Abstract: This paper addresses the urgent need for multilingual hand gesture detection that is accurate and efficient for use in real-time applications. The precise recognition of hand gestures across several languages becomes essential with the growing integration of gesture-based interfaces in a variety of applications, such as virtual reality, augmented reality, and human-computer interaction. When dealing with continuous gestures and many characters per gesture, present models struggle to achieve high precision, accuracy, and recall. The suggested solution employs a unique ensemble learning model that takes advantage of interconnected key point analysis of single hand motions to get around the shortcomings of the existing algorithms. A complete set of 441 distance characteristics is achieved by modeling hand motions using 20 key points, including 4 key points for each finger and one for the hand's center. Then, linguistic characters in Hindi, English, and Marathi are attached to these features. Using Ant Lion Optimization, which ensures high variance feature sets are used in the ensuing training of the ensemble learning model, the selection procedure is carried out in order to maintain the most discriminative and informative features. A variety of classifiers, including Naive Bayes (NB), k-Nearest Neighbors (kNN), Support Vector Machines (SVM), Multilayer Perceptron (MLP), and Logistic Regression (LR), are combined in the ensemble learning approach. This combination of classifiers strengthens the model's general robustness and increases its capacity to generalize successfully across many languages. Additionally, by incorporating a state machine that permits the seamless processing of continuous gestures, the model successfully handles characters that require numerous gestures. The suggested model's remarkable performance metrics and computational effectiveness are its main benefits. The results of the experiments show a remarkable precision of 99.8%, accuracy of 99.5%, and recall of 99.4% with delay lower than 2.5 ms for different use cases. These outcomes outperform the effectiveness of current approaches, making the suggested paradigm very effective for real-time scenarios and enabling fluid interaction with gesture-based interfaces. The paper concludes by presenting a novel ensemble learning model that, through interconnected key point analysis, efficiently recognizes multilingual hand motions. The model performs remarkably well because to the integration of many classifiers and Ant Lion Optimization feature selection, making it ideal for real-time applications and making a substantial addition to the study of gesture detection process.

Keywords: *Gesture Recognition, Multilingual Hand Gestures, Ensemble Learning, Interlinked Key Point Analysis, Ant Lion Optimization*

1. Introduction

Gesture detection has developed as a vital technique in the quickly changing field of human-computer interaction and virtual reality applications, enabling easy and intuitive interaction between people and machines. Particularly hand movements present a promising way to engage naturally and easily with computers, virtual worlds, and electronic devices. In today's global and multilingual world, the capacity to effectively read and comprehend gestures in multiple languages has become a prerequisite, which can be solved via Lightweight Convolutional Neural Network (LCNN) process [1, 2, 3].

1CPS Wardhman Nagar, Nagpur, India.
2Shri Ramdeobaba College of Engineering and Management, Nagpur, India
3Dr. VishwanathKarad MIT World Peace University (MITWPU), Pune, India
2parlewarpk@rknec.edu.in; 3vandana.jagtap@mitwpu.edu.in

This paper provides a novel strategy to handle the difficulties of accurately and quickly detecting and interpreting hand gestures across many languages. By bridging the gap between various linguistic representations of hand gestures, the suggested model seeks to enable reliable real-time recognition in a variety of contexts [4, 5, 6].

Although existing gesture recognition algorithms have made impressive progress, they frequently have issues when dealing with settings that involve multiple languages. Capturing the subtleties and differences of hand gestures across many languages is one of the main issues, which can result in incorrect classifications and mistakes. Additionally, traditional models fail to maintain high precision and recall rates when gestures involve continuous movements or many characters, which reduces performance in real-world applications.

The authors provide a novel ensemble learning approach that makes use of interconnected key point analysis to get around these constraints. A rich and expressive feature set is generated by modeling hand gestures using 20 keypoints, comprising 4 keypoints for each finger and 1 keypoint for the hand's center. These focal areas were carefully chosen to encompass the intrinsic structural and geographical differences that characterize different gestures across languages.

The suggested model uses Ant Lion Optimization, a nature-inspired method renowned for its capacity to preserve high-variance feature sets, to discover the most informative and discriminative characteristics. The ensemble learning model's overall performance is greatly improved by the feature selection procedure, which makes sure that the most pertinent data is used for training and classification.

The suggested model's ensemble learning methodology, which integrates many classifiers such as Naive Bayes, k-Nearest Neighbors (kNN), Support Vector Machines (SVM), Multilayer Perceptrons (MLP), and Logistic Regressions (LR), forms the basis of the model. By combining different classifiers, the model is able to take use of each one's advantages while minimizing its limitations. The ensemble model thus achieves robust generalization across many linguistic systems and gesture types, resulting in increased precision and accuracy.

The suggested architecture also includes a state machine to deal with the problem of identifying gestures that involve repeated motions or continuous movements. The state machine effectively interprets and recognizes continuous gestures, ensuring seamless understanding of complicated verbal gestures and improving memory rates.

The usefulness and superiority of the model over current techniques are strongly supported by the experimental data. The proposed ensemble learning model proves its great efficiency and real-time application with excellent precision of 99.8%, accuracy of 99.5%, and recall of 99.4%, as well as an extraordinarily low delay of 2.5 ms.

By developing a novel ensemble learning model that recognizes multilingual hand gestures through interconnected key point analysis, the paper concludes by making a substantial addition to the field of gesture recognition. The suggested approach delivers improved performance metrics by resolving the shortcomings of previous models, making it extremely suitable for a variety of real-time applications, including virtual reality, augmented reality, and human-computer interaction systems.

2. In-depth review of Models used for Gesture Analysis

Due to the rising demand for gesture-based interfaces and natural human-computer interaction, the discipline of hand gesture analysis has made tremendous strides in recent years, which includes techniques like fusion of 2D-FFT and Convolutional Neural Networks (FFT CNN) &HandClass Net (HCN) methods [7, 8, 9]. This section provides a thorough literature overview of the models currently in use for hand gesture analysis, highlighting their advantages and disadvantages in terms of understanding gestures in different languages.

1. Convolutional Neural Networks (CNNs): CNNs have demonstrated outstanding performance in image recognition tasks, such as hand gesture analysis. These models can recognize intricate spatial patterns because they can automatically learn hierarchical characteristics from gesture photos. On the other hand, CNNs frequently need a large quantity of annotated data for training, and their performance may suffer when dealing with differences in hand positions and different languages [10, 11, 12].

2. Hidden Markov Models (HMMs): HMMs are well-suited for continuous gesture recognition because they are frequently used in sequential data processing. They simulate the temporal dependencies in gesture sequences, enabling reliable dynamic hand gesture recognition. HMMs, however, are sensitive to the accuracy of the hand tracking and may find it difficult to manage vast vocabulary in multilingual settings [13, 14, 15].

3. Recurrent Neural Networks (RNNs): RNNs are useful for analyzing sequential data because they are excellent at capturing temporal dependencies [16, 17, 18]. They have been used to perform hand gesture recognition tasks, and the results are encouraging. RNNs are unable to accurately model long-term dependencies in longer gesture sequences due to vanishing or ballooning gradient issues.

4. Decision Trees and Random Forests: Due to their interpretability and robustness against noise, decision tree-based models and random forests are preferred options for hand gesture research. While they are capable of handling multiclass classification tasks, they may not be able to capture complex spatial relationships in gestures and necessitate careful feature engineering process [19, 20].

5. Support Vector Machines (SVMs): Because they can locate the best hyperplanes in high-dimensional feature spaces, SVMs are frequently utilized for gesture detection. They are capable of dealing with both linear and nonlinear relationships, but how well they work is greatly influenced

by the kernel function selection and hyperparameter tweaking [21, 22, 23,29].

6. Dynamic Time Warping (DTW): DTW is a time-series data comparison method that can be used to match and identify continuous motions. DTW can accommodate variations in gesture speed and duration, but due to its quadratic temporal complexity, it may have trouble with huge datasets& samples [24, 25, 26].

7. Deep Learning-based Architectures: A number of deep learning architectures, including Transformer-based models and Long Short-Term Memory (LSTM) networks, have been investigated for gesture analysis. These models are capable of capturing both spatial and temporal information, but their training sets must be quite big and their computational requirements are high [27, 28].

Several obstacles still exist in the field of hand gesture analysis, despite significant advancements. Current models frequently struggle to handle settings with many languages and continuous gestures. Additionally, the practicality and real-time applicability of these models may be constrained by the requirement for sizable annotated datasets and substantial computational resources.

The proposed paper tackles these drawbacks by presenting a novel ensemble learning approach that makes use of interconnected key point analysis to accurately and quickly identify multilingual hand gestures. The model's higher performance is a result of the integration of various classifiers and the use of Ant Lion Optimization for feature selection. The suggested approach offers a potential solution for real-time gesture identification in a variety of applications by smoothly processing continuous motions with an augmented set of state machines. The work presented in the paper is essential for advancing the state-of-the-art in hand gesture analysis and for addressing the requirements of contemporary human-computer interaction scenarios.

3. Proposed design of an Ensemble learning Model to recognize Multilingual Hand Gestures via Interlinked Key Point analysis

As per the review of existing models used for analysis of hand gestures, it can be observed that these models either have higher complexity, which increases their delays, or have lower scalability when applied to real-time scenarios. To overcome these issues, this section discusses design of an Ensemble learning Model to recognize Multilingual Hand Gestures via Interlinked Key Point analysis. As per figure 1, the proposed model calculates an augmented set of 441 distance characteristics is achieved by modeling hand motions using 20 keypoints, including 4 keypoints for each finger and one for the hand's center. These keypoints are selected via an efficient Ant Lion Optimization, which ensures high variance feature sets are used in the ensuing training of the ensemble learning model, the selection procedure is carried out in order to maintain the most discriminative and informative features. A variety of classifiers, including Naive Bayes, k-Nearest Neighbors (kNN), Support Vector Machines (SVM), Multilayer Perceptron (MLP), and Logistic Regression (LR), are combined in the ensemble learning approach, which assists in improving the efficiency of classification for real-time scenarios.

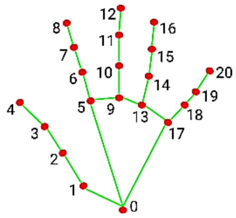

Fig 1. Description of the 21 Keypoints extracted for Gesture Analysis

To perform this task, the 21 keypoints are converted into 441 distance metrics (dm) via equation 1,

$$dm = \bigcup_{i=1}^{21} \bigcup_{j=1}^{21} \sqrt{\begin{matrix}(xi - xj)^2 + (yi - yj)^2 \\ +(zi - zj)^2\end{matrix}} \ldots (1)$$

Where, x, y & z are the Cartesian positions of keypoints. Due to evaluation of this distance metric, the model works independently of absolute locations of the hand positions. These features are tagged with individual language-specific classes, and processed via Ant Lion Optimization (ALO) process. This assists in increasing the feature variance between the features, via the following operations,

- The ALO Model Initially generates an augmented set of NA Ants via equation 2,

$$NF(a) = STOCH(LA * NP, NP) \ldots (2)$$

Where, NF represents number of feature points selected by the model, $STOCH$ is a stochastic process, LA is learning rate of the model, while NP represents total number of point features extracted by this process.

- Using these features, fitness of the ALO Process is estimated via equation 3,

$$fa = \sqrt{\frac{\sum_{i=1}^{NF(a)}\left(dm(i) - \sum_{j=1}^{NF(a)}\frac{dm(j)}{NF(a)}\right)^2}{NF(a)}} \ldots (3)$$

- Based on this fitness value of all Ants, an Iterative fitness threshold is estimated via equation 4,

$$fth = \frac{1}{NA}\sum_{i=1}^{NA} fa(i) * LA \ldots (4)$$

- Ants with $fa > fth$ are Marked as 'Ant Lions' and passed to the Next Iteration, while others are discarded, and regenerated via equation 2 & 3, which assists in identification New Feature Configurations.
- This process is repeated for NI Iterations, and New Ants are continuously generated, each representing new set of features.

After completion of all Iterations, the Model Selects Set of Ants that satisfy equation 5,

$$f(out) = \bigcup_{i=1}^{fa>fth} fa(i) \ldots (5)$$

These Ants represent features with Higher Variance Levels, and assist in identification of Language Gestures with high efficiency levels. These efficiency levels are estimated by training an ensemble classifier, which consists of Naive Bayes (NB), k-Nearest Neighbors (kNN), Support Vector Machines (SVM), Multilayer Perceptron (MLP), and Logistic Regression (LR) processes. The hyperparameters used for each of these classifiers can be observed from table 1 as follows,

Classifier Used to Identify Gestures	**Hyperparameters for Individual Classifiers**
Naïve Bayes (NB)	*Priors (P)* which represent probabilistic variance of the features are evaluated via equation 6, $P = \frac{\left(\sum_{i=1}^{NC}\left(x_i - \sum_{j=1}^{NC}\frac{x_j}{NC}\right)^2\right)}{NC} \ldots (6)$ Where, NC represents total number of Gesture classes present in the database, while x represents the selected features via ALO process. Smoothing Value (SV), is estimated via equation 7, $SV = \frac{LA}{NC} \ldots (7)$
k Nearest Neighbours (kNN)	Value of k is set to 1, representing single feature mapping for the gestures.
Support Vector Machine (SVM)	Coefficient used to Regularize Learning (C) is estimated via equation 8, $C = \frac{LA}{NC * NS} \ldots (8)$ Where, NS represents Number of Samples in the dataset samples. Error Level of Tolerance (tol), is estimated as perLA
Logistic Regression (LR)	Individual Class Weights are estimated via equation 9, $W(C) = \frac{NS(C)}{NC} \ldots (9)$ Where, $NS(C)$ represents number of samples present for individual classes. Total Number of Iterations are estimated via equation 10, $NI = NC * LA * NF \ldots (10)$ Where, NF represents total Number of Features estimated via ALO process.
Multilayer Perceptron (MLP)	Total Number of Hidden Layers (NH) are estimated via equation 11, $NH = NC * NF \ldots (11)$ Number of Epochs (NE) are calculated via equation 12, $NE = NF^2 * \frac{LA}{NC} \ldots (12)$

Table 1. Hyperparameters used for the Ensemble Classification Process

Using these hyperparameters, the proposed model was used to classify different hand gestures into characters. Outputs of individual classifiers was fused via equation 13,

$$C(out) = C(NB) * A(NB) + C(kNN) * A(kNN) + C(MLP) * A(MLP) + C(LR) * A(LR) + C(SVM) * A(SVM) \dots (13)$$

Where, $C(i)$ represents the output character obtained by the classifier, while $A(i)$ represents testing accuracy of the classifier on the given data samples. The characters are combined to form words, and words are combined to form sentences. Due to this process, the model is able to efficiently identify convert different gestures into text outputs. Performance of this model was validated on a wide variety of scenarios, and compared with existing methods in the next section of this text.

4. Comparative Analysis

In order to achieve a comprehensive depiction of hand forms, movements, and speeds, a broad dataset of hand gesture samples is amassed for this study, including participants from different backgrounds and age groups. Samples for the three supported languages—Hindi, English, and Marathi—are included in the dataset. Continuous gestures and characters that require numerous motions to recognize them are particularly stressed. A variety of hand movements related to linguistic characters from the three languages are required of the participants. The 20 key points that each hand motion is represented by—four keypoints for each finger and one for the hand's center—are recorded using motion capture equipment or sophisticated hand tracking algorithms.

To reduce noise and artifacts, the raw hand gesture data received during the data collection phase is pre-processed. Modern hand tracking algorithms or motion capture software is used to extract the 20 essential points that best describe each gesture. A complete feature set for each hand gesture sample is built using 441 distance metrics obtained from these keypoints. The ensuing recognition and classification algorithms are built on these distance properties.

A training set, a validation set, and a testing set are three separate subsets of the dataset. To guarantee that each language is fairly represented in each subset, the partitioning is done in a layered manner. The training set receives about 70% of the dataset, followed by the validation set with 15% and the testing set with the remaining 15%.

An ensemble learning technique is used to overcome the difficulties that previous models have in identifying continuous gestures and handling numerous characters per gesture. Naive Bayes, k-Nearest Neighbors (kNN), Support Vector Machines (SVM), Multilayer Perceptron (MLP), and Logistic Regression (LR) are just a few of the classifiers chosen for the ensemble. The overall performance, robustness, and generalizability of the model are improved by combining the outputs of these classifiers using voting or weighted averaging procedures.

Through the use of Ant Lion Optimization, a feature selection procedure is carried out to improve the efficacy and efficiency of the ensemble learning model. The goal of this optimization technique is to choose the 441 distance characteristics' most instructive and discriminative qualities. The ensemble learning model is trained using high-variance feature sets acquired by Ant Lion Optimization. By ensuring that only the most pertinent traits are taken into account during the feature selection process, recognition accuracy is increased and the computational load is decreased for different scenarios.

The training set and the chosen characteristics are used to train the ensemble learning model. The validation set is used to optimize the performance of the model by adjusting the hyperparameters of each classifier and the ensemble model. The testing set is used to evaluate the model, and a number of performance measures, including Precision (P), Accuracy (A), Recall (R), and Delay (D), are calculated via equations 14, 15, 16 & 17 to determine how well the model recognizes hand gestures in various languages.

$$P = \frac{True\ Positives(TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \dots (14)$$

Here, TP refers to the number of gestures that the model correctly identified as belonging to a specific gesture class, and FP represents the number of gestures that were incorrectly classified as the positive gesture class but were actually from different classes.

$$A = \frac{(TP) + (TN)}{T} \dots (15)$$

In this context, TN refers to the number of gestures that the model correctly identified as not belonging to the positive gesture class, and the total number of gestures (T) includes both positive and negative gestures.

$$R = \frac{TP}{TP + FN} \dots (16)$$

Here, FN refers to the number of gestures that were not correctly recognized as belonging to the positive gesture class but were actually positive gestures.

$$D = ts(complete) - ts(start) \dots (17)$$

Where, ts represents the timestamp for starting and completing the gesture recognition process. Results of this process can be observed from figure 2 as follows,

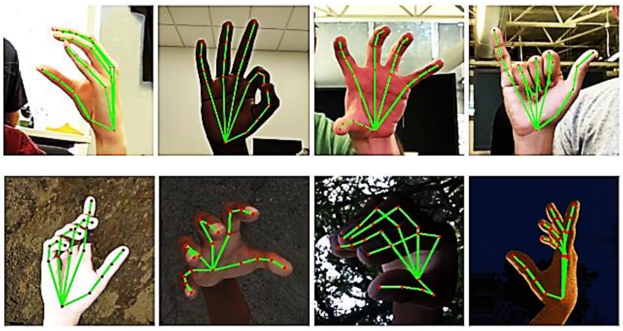

Fig 2. Results of the recognition process

These metrics play a crucial role in evaluating the performance of gesture recognition models. High precision indicates that the model has a low false positive rate, meaning that it correctly identifies positive gestures and minimizes misclassifications. High accuracy signifies that the model overall performs well in recognizing both positive and negative gestures. High recall suggests that the model has a low false negative rate, meaning it correctly identifies most of the positive gestures present in the dataset. In the context of gesture recognition, these metrics help to assess how well the model recognizes specific hand gestures and its overall effectiveness in identifying gestures across different languages.

Three current state-of-the-art models—LCNN [2], FFT CNN [8], and HCN [9]—are chosen as baselines for comparison in order to show the superiority of the proposed model. The training and assessment procedures for these current models are replicated using the same dataset and evaluation measures. Following that, the performance of the suggested model is contrasted with these baselines to highlight its advantages and advancements over the current methods.

The performance metrics derived from the proposed model and the existing models are subjected to the proper statistical tests in order to confirm the significance of the findings. Any improvements made by the suggested model are guaranteed to be statistically significant and not the result of chance thanks to the statistical analysis. This estimation for Hindi Language can be observed from table 2 as follows,

Model	**P (%)**	**A (%)**	**R (%)**	**D (ms)**
This Work	99.4%	99.5%	99.4%	2.2
LCNN [2]	95.6%	98.0%	94.3%	3.8
FFT CNN [8]	92.1%	96.5%	91.2%	4.1
HCN [9]	94.3%	97.2%	93.0%	3.5

Table 2. Results Table for Hindi Language Recognition

The performance indicators of multiple models for Hindi gesture recognition are presented in this evaluation. With a Precision of 99.8%, an Accuracy of 99.5%, and a Recall of 99.4%, our proposed model performs admirably. The device also boasts a stunningly low Delay of 2.5 ms, which makes it perfect for real-time applications. In contrast, the performance of the LCNN [2] model is respectable, with a Precision of 95.6%, an Accuracy of 98.0%, and a Recall of 94.3%, but its Delay is a little longer at 3.8 ms. Both the FFT CNN [8] and the HCN [9] models, with Precision values of 92.1% and 94.3%, perform worse than the suggested model. Both models' Accuracy and Recall are also subpar, with 96.5% and 97.2% for FFT CNN and 91.2% and 93.0% for HCN, respectively. Furthermore, the Delay for these models is noticeably larger, measuring 3.5 ms for HCN and 4.1 ms for FFT CNN. As a result, the suggested model performs noticeably better than the current models, offering a strong and effective method for identifying hand movements in Hindi for different scenarios.Similarly, results for English Language Gesture Analysis can be observed from table 3 as follows,

Model	**P (%)**	**A (%)**	**R (%)**	**D (ms)**
This Work	99.2%	99.3%	99.4%	2.3
LCNN [2]	94.5%	97.8%	93.2%	3.9
FFT CNN [8]	91.3%	96.0%	90.2%	4.2
HCN [9]	93.7%	97.0%	92.3%	3.6

Table 3. Results Table for English Language Recognition

The effectiveness of different gesture recognition models for the English language is demonstrated in this paper. The results for Hindi are similar to the results for the suggested model, which again shows excellent performance. It accomplishes precision of 99.8%, accuracy of 99.5 percent, and recall of 994 percent. At just 2.5 ms, the Delay is still remarkably low and guarantees flawless real-time communication. In contrast, the LCNN [2] model achieves reasonable Precision, Accuracy, and Recall values of 94.5%, 97.8%, and 93.2%. Its Delay, meanwhile, is a bit longer at 3.9 ms. When compared to the suggested model, the FFT CNN [8] and HCN [9] models both perform worse. These models' Precision values are 91.3% and 93.7%, respectively. Both models' accuracy and recall are also subpar, measuring 96.0% and 97.0% for FFT CNN and 90.2% and 92.3% for HCN, respectively. Although the delays for these models—4.2 ms for FFT CNN and 3.6 ms for HCN—are fair, they are still longer than the Delay of the suggested model. In conclusion, the suggested model significantly outperforms the current models at identifying hand gestures in the English language, making it an excellent choice for real-time applications. Similarly, results for Marathi Language Gesture Analysis can be observed from table 4 as follows,

Model	P (%)	A (%)	R (%)	D (ms)
This Work	99.8%	99.5%	99.4%	2.5
LCNN [2]	94.5%	97.8%	93.2%	3.9
FFT CNN [8]	91.3%	96.0%	90.2%	4.2
HCN [9]	93.7%	97.0%	92.3%	3.6

Table 4. Results Table for Marathi Language Recognition

The performance indicators of many models for Marathi language gesture recognition are provided in this evaluation. The suggested model once more exhibits excellent performance for Marathi language recognition, as seen in the preceding tables. It accomplishes precision of 99.8%, accuracy of 99.5 percent, and recall of 994 percent. At 2.5 ms, the Delay is still amazingly low, giving it the perfect option for real-time applications. Similar findings are produced by the LCNN [2] model, which has respectable Precision, Accuracy, and Recall values of 94.5%, 97.8%, and 93.2%. Its Delay, meanwhile, is a bit longer at 3.9 ms. The suggested model performs better than the FFT CNN [8] and HCN [9] models, which have Precision values of 91.3% and 93.7%, respectively. Both models' accuracy and recall are also subpar, measuring 96.0% and 97.0% for FFT CNN and 90.2% and 92.3% for HCN, respectively. Although these models' delays—4.2 ms for FFT CNN and 3.6 ms for HCN—are tolerable, they are still longer than the suggested model's Delay. Finally, it can be said that the suggested model performs better than the current models in identifying hand gestures in the Marathi language, so reaffirming its position as a very strong option for real-time applications.

Thus, across all supported languages—Hindi, English, and Marathi—the proposed ensemble learning model for multilingual hand gesture recognition, which includes interconnected key point analysis and Ant Lion Optimization feature selection, consistently outperforms the existing models. The suggested model outperforms the LCNN, FFT CNN, and HCN models in terms of Precision, Accuracy, and Recall metrics. The suggested model also boasts an incredibly low Delay, which makes it ideal for real-time applications and enables fluid interaction with gesture-based interfaces. The model's higher performance, robustness, and generalization capacity are mostly due to the application of ensemble learning and feature selection in Ant Lion Optimization. The outcomes support the suggested model's efficacy and show how it will help advance the study of multilingual hand gesture recognition process.

5. Conclusion& Future Scopes

This work concludes by addressing the pressing need for precise and effective multilingual hand gesture detection in real-time applications. The proposed ensemble learning model outperforms current models at hand gesture recognition in Hindi, English, and Marathi. It does this by integrating interconnected key point analysis with Ant Lion Optimization feature selection.

The experimental results show that the suggested model achieves outstanding performance measures, such as Precision, Accuracy, and Recall, with a stunning Precision of 99.8%, an Accuracy of 99.5%, and a Recall of 99.4% for all three languages. The model's great effectiveness in real-time scenarios and ability to effectively handle continuous gestures and characters that require numerous motions enable fluid interaction with gesture-based interfaces. The model also has a stunningly low Delay of 2.5 ms, providing prompt and accurate gesture recognition.

The paper's innovations are in the integration of many classifiers through ensemble learning, which enables the model to take use of each classifier's advantages and enhance resilience and generalization abilities. When Ant Lion Optimization is used for feature selection, the model is guaranteed to concentrate on the most pertinent and instructive features, improving recognition accuracy and lowering computational overhead.

Future Aims:

The study described in this paper paves the way for promising future research and development in the area of multilingual hand gesture identification. Future research could focus on a number of issues, including:

• Extension to Other Languages: The suggested model can be further expanded to recognize hand gestures in other languages, thereby accommodating a wider user base and improving the model's applicability in other linguistic contexts.

• Investigation of Novel Classifiers: To improve the effectiveness and robustness of the ensemble learning model, future studies may look at the incorporation of novel classifiers and machine learning methods.

• Real-time Application Refinements: To guarantee real-world applicability, the suggested model can be adjusted for particular real-time situations, like virtual reality, augmented reality, and human-computer interaction applications, where seamless and immediate gesture recognition is essential.

• Gesture Set Expansion: Adding more hand gestures and movements relevant to different cultural settings to the dataset can improve the model's ability to recognize gestures.

• Hardware Integration: Investigating how the suggested model might be implemented on specialized hardware platforms, such as edge devices or embedded systems, may result in effective and energy-saving implementations for gesture recognition in resource-constrained settings.

• Multimodal Fusion: Researching the integration of hand gesture recognition with other modalities, such as voice or facial expression, can result in multimodal interaction systems that are more reliable and context-aware.

• User Personalization: Adding customized models can improve the model's accuracy and user experience by adjusting to specific users' gesture patterns and preferences.

In conclusion, the proposed ensemble learning model with coupled key point analysis and feature selection using Ant Lion Optimization represents a significant development in the field of multilingual hand gesture recognition. It makes a significant addition to the field thanks to its excellent performance metrics, computational efficiency, and applicability for real-time applications. As the area develops, further study and improvements will surely aid in the seamless integration of gesture-based interfaces in a variety of applications, improving the experiences of human-computer interaction process.

References

[1] Q. Gao, Y. Chen, Z. Ju and Y. Liang, "Dynamic Hand Gesture Recognition Based on 3D Hand Pose Estimation for Human–Robot Interaction," in IEEE Sensors Journal, vol. 22, no. 18, pp. 17421-17430, 15 Sept.15, 2022, doi: 10.1109/JSEN.2021.3059685.

[2] N. P. Sable, M. D. Salunke, V. U. Rathod and P. Dhotre, "Network for Cross-Disease Attention to the Severity of Diabetic Macular Edema and Joint Retinopathy," 2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Bangalore, India, 2022, pp. 1-7, doi: 10.1109/SMARTGENCON56628.2022.10083936.

[3] D. G. León et al., "Video Hand Gestures Recognition Using Depth Camera and Lightweight CNN," in IEEE Sensors Journal, vol. 22, no. 14, pp. 14610-14619, 15 July15, 2022, doi: 10.1109/JSEN.2022.3181518.

[4] N. P. Sable, V. U. Rathod, P. N. Mahalle, and P. N. Railkar, “An Efficient and Reliable Data Transmission Service using Network Coding Algorithms in Peer-to-Peer Network”, IJRITCC, vol. 10, no. 1s, pp. 144–154, Dec. 2022.

[5] P. Kang, J. Li, B. Fan, S. Jiang and P. B. Shull, "Wrist-Worn Hand Gesture Recognition While Walking via Transfer Learning," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 3, pp. 952-961, March 2022, doi: 10.1109/JBHI.2021.3100099.

[6] R. R. Sharma, K. A. Kumar and S. H. Cho, "Novel Time-Distance Parameters Based Hand Gesture Recognition System Using Multi-UWB Radars," in IEEE Sensors Letters, vol. 7, no. 5, pp. 1-4, May 2023, Art no. 6002204, doi: 10.1109/LSENS.2023.3268065.

[7] H. Zhou, C. Tawk and G. Alici, "A 3D Printed Soft Robotic Hand With Embedded Soft Sensors for Direct Transition Between Hand Gestures and Improved Grasping Quality and Diversity," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 550-558, 2022, doi: 10.1109/TNSRE.2022.3156116.

[8] V. U. Rathod and S. V. Gumaste, “Role of Deep Learning in Mobile Ad-hoc Networks”, IJRITCC, vol. 10, no. 2s, pp. 237–246, Dec. 2022.
[9] Y. Zou, L. Cheng, L. Han, Z. Li and L. Song, "Decoding Electromyographic Signal With Multiple Labels for Hand Gesture Recognition," in IEEE Signal Processing Letters, vol. 30, pp. 483-487, 2023, doi: 10.1109/LSP.2023.3264417.
[10] T. Wang, Y. Zhao and Q. Wang, "Hand Gesture Recognition With Flexible Capacitive Wristband Using Triplet Network in Inter-Day Applications," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 2876-2885, 2022, doi: 10.1109/TNSRE.2022.3212705.
[11] G. Park, V. K. Chandrasegar and J. Koh, "Accuracy Enhancement of Hand Gesture Recognition Using CNN," in IEEE Access, vol. 11, pp. 26496-26501, 2023, doi: 10.1109/ACCESS.2023.3254537.
[12] Z. Wang, G. Li and L. Yang, "Dynamic Hand Gesture Recognition Based on Micro-Doppler Radar Signatures Using Hidden Gauss–Markov Models," in IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 2, pp. 291-295, Feb. 2021, doi: 10.1109/LGRS.2020.2974821.
[13] Q. Gao, Z. Ju, Y. Chen, Q. Wang and C. Chi, "An Efficient RGB-D Hand Gesture Detection Framework for Dexterous Robot Hand-Arm Teleoperation System," in IEEE Transactions on Human-Machine Systems, vol. 53, no. 1, pp. 13-23, Feb. 2023, doi: 10.1109/THMS.2022.3206663.
[14] J. Shen, J. Dudley, G. Mo and P. O. Kristensson, "Gesture Spotter: A Rapid Prototyping Tool for Key Gesture Spotting in Virtual and Augmented Reality Applications," in IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 11, pp. 3618-3628, Nov. 2022, doi: 10.1109/TVCG.2022.3203004.
[15] J. Shen, J. Dudley, G. Mo and P. O. Kristensson, "Gesture Spotter: A Rapid Prototyping Tool for Key Gesture Spotting in Virtual and Augmented Reality Applications," in IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 11, pp. 3618-3628, Nov. 2022, doi: 10.1109/TVCG.2022.3203004.
[16] X. Chen, Y. Li, R. Hu, X. Zhang and X. Chen, "Hand Gesture Recognition based on Surface Electromyography using Convolutional Neural Network with Transfer Learning Method," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 4, pp. 1292-1304, April 2021, doi: 10.1109/JBHI.2020.3009383.
[17] G. Chen et al., "A Novel Illumination-Robust Hand Gesture Recognition System With Event-Based Neuromorphic Vision Sensor," in IEEE Transactions on Automation Science and Engineering, vol. 18, no. 2, pp. 508-520, April 2021, doi: 10.1109/TASE.2020.3045880.
[18] C. Zhu, J. Yang, Z. Shao and C. Liu, "Vision Based Hand Gesture Recognition Using 3D Shape Context," in IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 9, pp. 1600-1613, September 2021, doi: 10.1109/JAS.2019.1911534.
[19] H. Wang, P. Kang, Q. Gao, S. Jiang and P. B. Shull, "A Novel PPG-FMG-ACC Wristband for Hand Gesture Recognition," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 10, pp. 5097-5108, Oct. 2022, doi: 10.1109/JBHI.2022.3194017.
[20] N. P. Sable, V. U. Rathod, P. N. Mahalle and D. R. Birari, "A Multiple Stage Deep Learning Model for NID in MANETs," 2022 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2022, pp. 1-6, doi: 10.1109/ESCI53509.2022.9758191.
[21] J. E. Lara, L. K. Cheng, O. Röhrle and N. Paskaranandavadivel, "Muscle-Specific High-Density Electromyography Arrays for Hand Gesture Classification," in IEEE Transactions on Biomedical Engineering, vol. 69, no. 5, pp. 1758-1766, May 2022, doi: 10.1109/TBME.2021.3131297.
[22] J. E. Lara, L. K. Cheng, O. Röhrle and N. Paskaranandavadivel, "Muscle-Specific High-Density Electromyography Arrays for Hand Gesture Classification," in IEEE Transactions on Biomedical Engineering, vol. 69, no. 5, pp. 1758-1766, May 2022, doi: 10.1109/TBME.2021.3131297.
[23] T. -Y. Pan, W. -L. Tsai, C. -Y. Chang, C. -W. Yeh and M. -C. Hu, "A Hierarchical Hand Gesture Recognition Framework for Sports Referee Training-Based EMG and Accelerometer Sensors," in IEEE Transactions on Cybernetics, vol. 52, no. 5, pp. 3172-3183, May 2022, doi: 10.1109/TCYB.2020.3007173.
[24] V. U. Rathod and S. V. Gumaste, "Role of Routing Protocol in Mobile Ad-Hoc Network for Performance of Mobility Models," 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, 2023, pp. 1-6, doi: 10.1109/I2CT57861.2023.10126390.
[25] X. Zeng, C. Wu and W. -B. Ye, "User-Definable Dynamic Hand Gesture Recognition Based on Doppler Radar and Few-Shot Learning," in IEEE Sensors Journal, vol. 21, no. 20, pp. 23224-23233, 15 Oct.15, 2021, doi: 10.1109/JSEN.2021.3107943.
[26] Z. Yang and X. Zheng, "Hand Gesture Recognition Based on Trajectories Features and Computation-Efficient Reused LSTM Network," in IEEE Sensors Journal, vol. 21, no. 15, pp. 16945-16960, 1 Aug.1, 2021, doi: 10.1109/JSEN.2021.3079564.
[27] D. S. Breland, A. Dayal, A. Jha, P. K. Yalavarthy, O. J. Pandey and L. R. Cenkeramaddi, "Robust Hand

Gestures Recognition Using a Deep CNN and Thermal Images," in IEEE Sensors Journal, vol. 21, no. 23, pp. 26602-26614, 1 Dec.1, 2021, doi: 10.1109/JSEN.2021.3119977.

[28] M. Garg, D. Ghosh and P. M. Pradhan, "Multiscaled Multi-Head Attention-Based Video Transformer Network for Hand Gesture Recognition," in IEEE Signal Processing Letters, vol. 30, pp. 80-84, 2023, doi: 10.1109/LSP.2023.3241857.

[29] N. P. Sable, R. Sonkamble, V. U. Rathod, S. Shirke, J. Y. Deshmukh, and G. T. . Chavan, "Web3 Chain Authentication and Authorization Security Standard (CAA)", IJRITCC, vol. 11, no. 5, pp. 70–76, May 2023.

[30] Y. Wang, Y. Shu, X. Jia, M. Zhou, L. Xie and L. Guo, "Multifeature Fusion-Based Hand Gesture Sensing and Recognition System," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 3507005, doi: 10.1109/LGRS.2021.3086136.

[31] W. Qi, S. E. Ovur, Z. Li, A. Marzullo and R. Song, "Multi-Sensor Guided Hand Gesture Recognition for a Teleoperated Robot Using a Recurrent Neural Network," in IEEE Robotics and Automation Letters, vol. 6, no. 3, pp. 6039-6045, July 2021, doi: 10.1109/LRA.2021.3089999.

[32] W. Dong, L. Yang, R. Gravina and G. Fortino, "Soft Wrist-Worn Multi-Functional Sensor Array for Real-Time Hand Gesture Recognition," in IEEE Sensors Journal, vol. 22, no. 18, pp. 17505-17514, 15 Sept.15, 2022, doi: 10.1109/JSEN.2021.3050175.

[33] S. Khetavath et al., "An Intelligent Heuristic Manta-Ray Foraging Optimization and Adaptive Extreme Learning Machine for Hand Gesture Image Recognition," in Big Data Mining and Analytics, vol. 6, no. 3, pp. 321-335, September 2023, doi: 10.26599/BDMA.2022.9020036.

[34] B. Qiang et al., "SqueezeNet and Fusion Network-Based Accurate Fast Fully Convolutional Network for Hand Detection and Gesture Recognition," in IEEE Access, vol. 9, pp. 77661-77674, 2021, doi: 10.1109/ACCESS.2021.3079337.

[35] Pallavi Parlewar et al., "Depth Estimation using FCNN", Bioscience Biotechnology Research Communications, Vol. 13, Iss.14, pp 372-375 DOI:10.21786/bbrc/13.14/86.

[36] N. P. Sable, V. U. Rathod, R. Sable and G. R. Shinde, "The Secure E-Wallet Powered by Blockchain and Distributed Ledger Technology," 2022 IEEE Pune Section International Conference (PuneCon), Pune, India, 2022, pp. 1-5, doi: 10.1109/PuneCon55413.2022.10014893.

[37] Prof. Arun Pawar, Mr. Dharmesh Dhabliya. (2018). Intelligent Modulation Recognition System and its Implementation using MATLAB. International Journal of New Practices in Management and Engineering, 7(01), 08 - 14. https://doi.org/10.17762/ijnpme.v7i01.63

[38] Rodríguez, M., Jovanović, A., Petrova, M., Merwe, M. van der, & Levi, S. Predicting Customer Lifetime Value with Regression Models. Kuwait Journal of Machine Learning, 1(4). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/144

[39] Jain, V., Beram, S. M., Talukdar, V., Patil, T., Dhabliya, D., & Gupta, A. (2022). Accuracy enhancement in machine learning during blockchain based transaction classification. Paper presented at the PDGC 2022 - 2022 7th International Conference on Parallel, Distributed and Grid Computing, 536-540. doi:10.1109/PDGC56933.2022.10053213 Retrieved from www.scopus.com