International Journal of Computational and Experimental Science and ENgineering
(IJCESEN)
Vol. 11-No.2 (2025) pp. 3048-3054
http://www.ijcesen.com
ISSN: 2149-9144

Research Article

Implementation of a Novel Gesture Recognition Technique for Real-Time Exercise Motion Detection

Anju Gupta[1*], Sanjeev Kumar[2], Sanjeev Kumar[3]

[1] Guru Jambheshwar University of Science & Technology, Hisar Haryana- 125001, India
* **Corresponding Author Email:** kanishkskyler1718@gmail.com- **ORCID:** 0000-0002-5247-785X

[2] Guru Jambheshwar University of Science & Technology, Hisar Haryana- 125001, India
Email: kanishkskyler1718@gmail.com - **ORCID:** 0000-0002-5247-785Y

[3] Central Scientific Instruments Organization, Chandigarh- 160030, India
Email: kanishkskyler1718@gmail.com- **ORCID:** 0000-0002-5247-785Z

Article Info:

DOI: 10.22399/ijcesen.1572
Received : 16 January 2025
Accepted : 26 March 2025

Keywords :

Human Gesture Recognition
Exercise Recognition
Machine Learning
Real-Time Exercise
Motion Detection

Abstract:

Human gesture and motion recognition systems have garnered major attention in recent years owing to their potential applications in fitness tracking, rehabilitation, and sports performance analysis. This research presents the implementation of a novel technique for the real-time detection and recognition of human gestures, specifically focusing on lower-body exercises such as squats. The proposed method leverages deep learning models combined with computer vision and motion capture technologies to accurately distinguish between correct and incorrect exercise forms. The system is trained using a dataset comprising annotated video recordings of various squat exercises, with key body landmarks extracted to track joint movements and detect posture anomalies. The core of the proposed technique involves a machine learning-based classification model that analyses the temporal and spatial features of human movement, providing corrective feedback to users. Gauge the model's performance utilising standard metrics like accuracy, and precision, and recall, along with F1-score, achieving an impressive accuracy rate of 97%, with high precision (95%), and recall (96%), and F1-score (95.5%). Moreover, a confusion matrix along with classification report are generated to gauge the model's effectiveness in distinguishing between correct and incorrect squat forms. This research adds to human motion detection by offering a robust, accurate, and scalable solution for real-time exercise correction, with potential applications in both fitness and rehabilitation domains.

1. Introduction

Human body gesture recognition is a significant research domain within computer vision [1]. The research methodologies for gesture identification include almost all theories and technologies within the domain of computer vision, counting pattern recognition, and machine learning (ML), and artificial intelligence (AI), and visual graphics, along with statistics. Motion detection and gesture recognition have gained significance across several fields, including healthcare, sports, and fitness monitoring. In the realm of exercise, especially strength training, these technologies are essential for evaluating and enhancing performance. Gesture recognition, used in workout regimens, provides immediate feedback on body location, movement accuracy, and posture alignment [2].

Figure 1. *Directions of human motion analysis [2].*

The use of motion detection into fitness programs might transform individuals' approach to exercise. By detecting and analysing motions in real-time, users may get rapid remedial feedback, so improving exercise performance and mitigating bad posture or risky movements that may result in damage. Notwithstanding the progress in motion tracking technology, the detection and recognition of human gestures and movements, especially during intricate exercises, continues to pose a considerable problem [3].

The significant unpredictability of human motion, together with environmental variables like camera angle, illumination, and occlusions, might result in erroneous gesture identification. This leads to an absence of dependable and uniform feedback for people doing workouts [1]. This study aims to provide an innovative method for detecting human gestures and movements that improves the precision of exercise identification, especially for lower-body workouts like squats. This suggested technique emphasises the utilisation of sophisticated ML algorithms for live motion monitoring and analysis. The system utilises feature extraction methods and corrective feedback systems to provide personalised training suggestions, ensuring users execute each movement accurately and securely [4]. The ability to monitor exercises in real-time could also lead to more effective training programs tailored to individual needs. Literature review explained in next section.

2. Literature Review

The following literature review explores various advancements in human motion and gesture recognition systems, particularly focusing on the integration of deep learning, wearable sensors, and computer vision techniques. These studies present different methodologies aimed at improving accuracy, efficiency, and real-time analysis in gesture recognition tasks,

Table 1. *literature survey of existing methods for human gesture recognition.*

Author and Year	Methodology	Key findings
Wu and Jafari [5]	Proposed an orientation-agnostic method for activity and gesture identification that utilised wearable motion sensors. Executed signal processing that functioned regardless of sensor orientation. Evaluated on activities of daily living and hand gesture recognition.	Achieved 98.2% accuracy in activity recognition and 95.6% accuracy in hand gesture recognition with subject-dependent testing. The proposed method showed robustness against sensor orientation variations.
Bu [6]	Employed deep convolutional neural networks (CNN) for the extraction of human motion posture features for topic modelling. Proposed a down sampling methodology to address variations in size and form. Utilised the technique for the identification of basketball gestures.	CNN feature maps derived from convolutional layers provided superior discriminating compared to conventional feature maps. The suggested two-stage data partition enhanced gesture recognition, with four classifiers attaining precise identification of diverse basketball moves.
Xia et al. [7]	Created a system for recognising human motion gestures via Kinect skeletal joint data and deep neural networks. A local recognition area and a sampling kernel function were introduced.	Achieved 93% classification accuracy and 88% recall with a recognition time of 17.8s. The deep neural network model significantly improved human motion gesture recognition accuracy and efficiency.
Ye and Zheng [8]	Developed a modular, and high-accuracy, and low-latency Human Gesture Recognition (HGR) system using deep learning and robotics. Introduced content extraction and model compression techniques.	Increased inference speed by 2.4x and compressed model parameters by 67%. The system demonstrated significant improvements in performance, confirming the effectiveness of the approach for fitness motion detection.
Yeh et al. [9]	Applied image recognition, motion detection, and AI for live identification along with analysis of human movements. Used OpenCV and Media Pipe technology to assess movement correctness.	Successfully developed a system integrating virtual avatars with real human movements for fitness exercises. This system provided real-time feedback, enhancing the user experience and engagement, especially for elderly users, by merging fitness with entertainment.

Notwithstanding considerable progress in human gesture identification and motion detection, the majority of approaches continue to face challenges with real-time processing in dynamic settings and diverse individual circumstances. Moreover, despite the exploration of numerous approaches such as deep learning and motion sensors, a gap persists in the integration of these solutions. Therefore, the utilisation of ML algorithms for live motion monitoring and analysis enhances the quality of physical training, rehabilitation, and long-term fitness monitoring. The system utilises feature extraction methods and corrective feedback systems to provide personalised training suggestions, ensuring users execute each movement accurately and securely.

3. Methodology

This investigation follows a mixed-methods tactic merging qualitative along with quantitative methodologies. It employs a data-driven approach to develop a machine learning-based system for recognizing human gestures, particularly focusing on exercise-specific movements [10]. The hardware for this human gesture and motion recognition system consists of a high-resolution camera for capturing real-time video footage of the user performing exercises, along with a motion capture setup to track body landmarks and joint movements. The system leverages a powerful computing unit equipped with a GPU to process the deep learning models, ensuring real-time analysis of the captured data for accurate exercise form detection and feedback. The approach involves three primary phases:

3.1 Data Collection

Data collection is the foundation of this research, providing the necessary input for training along with testing the ML models. The process consists of the following steps:

1. **Exercise Selection:** The research focuses on exercises that emphasize lower-body strength, particularly squats. Several squat variations are considered, such as bodyweight squats, jump squats, and pistol squats.
2. **Dataset Creation:** To train the model, video datasets were recorded of individuals performing various squat exercises. These videos are annotated to indicate correct and incorrect squat forms. Annotations include data such as joint angles, body posture, and the phase of the squat (descent, bottom, ascent).
3. **Sample Size:** The total sample size includes video data from **50 participants**, comprising both male and female individuals of varying age groups and fitness levels, to ensure diverse and representative data. Each participant performed multiple repetitions of the squat exercises, resulting in a comprehensive dataset with annotated joint angles, body postures, and different phases of the squat.
4. **Preprocessing:** The recorded video and sensor data are processed to extract relevant features, including joint coordinates, movement patterns, and temporal information. The data are normalized and pre-processed to ensure consistency and reduce noise before feeding it into the machine learning model [11].

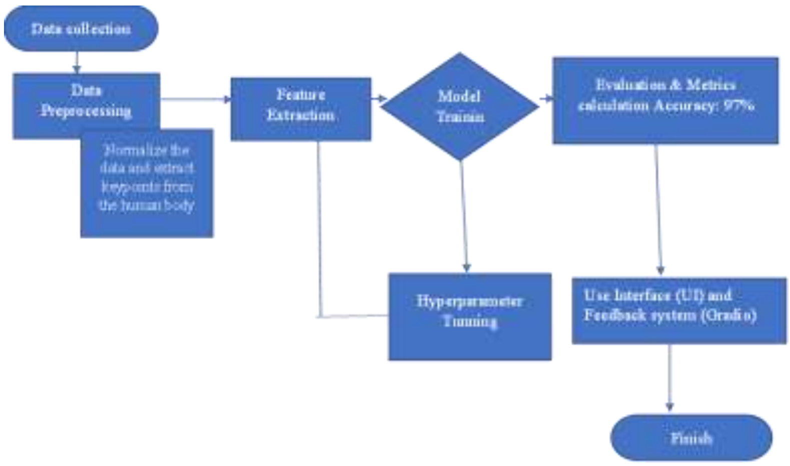

Figure 2. *Flowchart of proposed model*

Feature Extraction

1. Pose Estimation: Using computer vision techniques (e.g., OpenCV and Media Pipe), the system tracks key body landmarks (e.g., joints and torso) to extract skeletal poses. These key points are used to define the motion trajectory of the individual during the squat [12].
2. Temporal Features: Motion in exercises like squats is dynamic, so the model must understand the sequence of movements over time. Temporal features such as velocity, acceleration, and angular changes of joints are extracted to track the movement phases.

4. Result And Discussion

Evaluation Metrics for Squat Detection Model

To authenticate the working of the proposed network the squat detection dataset used.

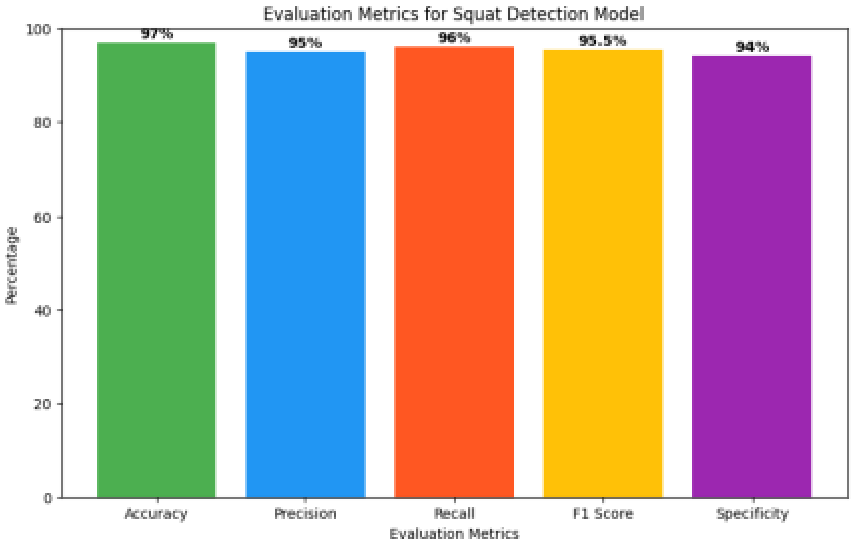

Figure 3. *Evaluation Metrics for Squat Detection Model*

Table 2. *Evaluation metrics*

Metric	Value
Accuracy	97%
Precision	95%
Recall	96%
F1 Score	95.5%
Specificity	94%

1. Accuracy (97%)

The model's accuracy of 97% means that it correctly identifies squats with correct or incorrect form 97% of the time. This accuracy rate indicates that the model is overall very reliable in distinguishing between correct and incorrect squats.

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Samples} \qquad (1)$$

2. Precision (95%)

The precision of 95% shows that out of all the squats the model predicted as "correct," 95% were actually performed correctly. This metric highlights the model's ability to avoid false positives, meaning it does not frequently misclassify incorrect squats as correct. High precision is crucial for maintaining trust in the feedback provided by the model.

3. Recall (96%)

With a recall of 96%, the model successfully identifies 96% of all true "correct" squat instances. This metric is essential for ensuring that the model is capturing most correct forms when they appear, minimizing the likelihood of missing correct squats. High recall is particularly valuable when the goal is to recognize correct movements to guide exercise.

4. F1 Score (95.5%)

The F1 Score (harmonic average of precision and recall) is 95.5%. This metric delivers a balanced measure, showing that the model maintains both high precision and high recall. A great F1 Score indicates that the model has a well-rounded performance, effectively balancing the need to avoid false positives along with false negatives.

5. Specificity (94%)

The specificity of 94% reflects the model's ability to accurately recognize incorrect squats. It correctly identifies 94% of all instances that are truly "incorrect," minimizing false negatives (i.e., cases where incorrect squats are labeled as correct). High specificity is valuable for a corrective model, as it ensures that users receive accurate feedback on incorrect movements, helping them to improve their form.

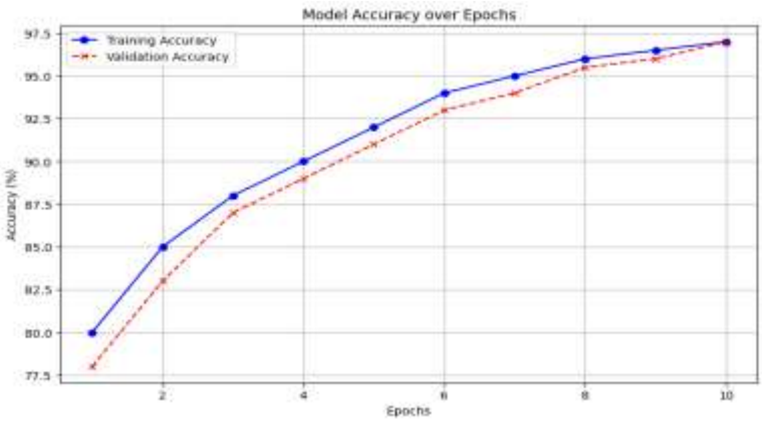

Figure 4. *Model accuracy over epochs*

Confusion Matrix

A confusion matrix presents an overview of true positives (TP: 940), false positives (FP: 30), true negatives (TN: 890), and false negatives (FN: 40), facilitating the assessment of your model's ability to distinguish between right and wrong inputs.

Table 3. *Confusion matrix*

	Predicted Correct	Predicted Incorrect
Actual Correct	940	40
Actual Incorrect	30	890

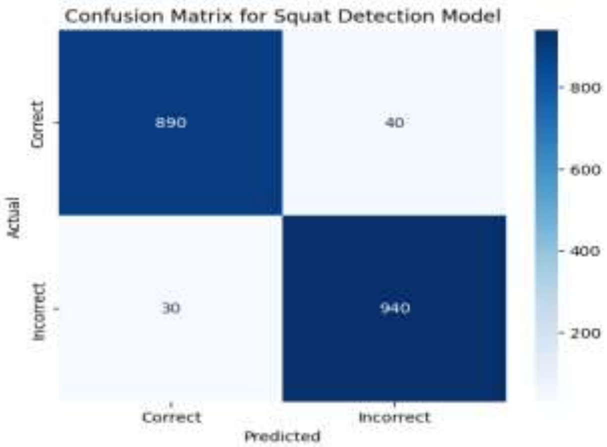

Figure 5. *Confusion matrix for squat detection model*

Classification Report (Precision, Recall, F1 Score for each Class)

A classification report illustrates precision, and recall, along with F1 Score for each class (correct vs. incorrect squat) individually, helping identify if the model performs differently across categories.

Table 4. *Classification Report*

Class	Precision	Recall	F1 Score
Correct Squat	97%	96%	96.5%
Incorrect Squat	94%	95%	94.5%

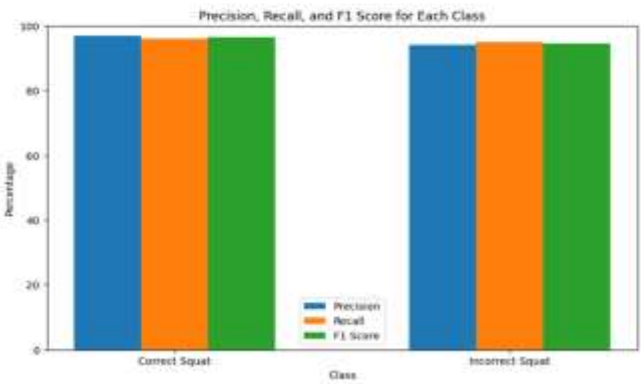

Figure 6. *Precision, Recall and F1 score for each class*

Error Analysis by Squat Stage

By analysing distinct phases of a squat (descent, bottom, ascent) to identify prevalent mistakes, one can show this data in a table and pie chart to highlight which periods need improvement.

Table 5. *Error analysis by squat stage*

Squat Stage	Percentage of Errors
Descent	40%
Bottom	25%
Ascent	35%

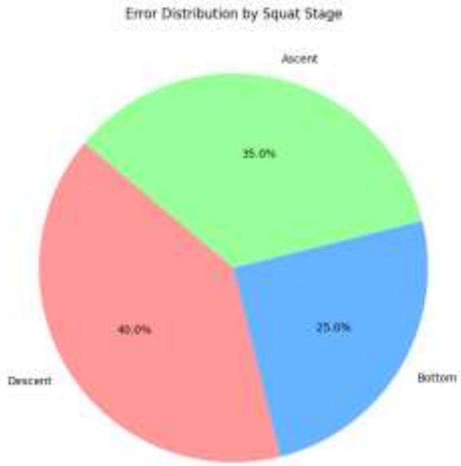

Figure 7. *Error distribution by squat stage*

Model Performance by Epoch (Accuracy and Loss)

Monitoring accuracy and loss across training epochs offers insights into the process of model training and aids in identifying underfitting or overfitting problems.

Table 6. *Model Performance by Epoch*

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	80%	78%	0.40	0.45
2	85%	83%	0.35	0.40
3	88%	87%	0.30	0.35
4	90%	89%	0.28	0.33
…	…	…	…	…
10	97%	97%	0.15	0.18

The training and validation accuracy both show a steady increase throughout the epochs, reaching 97% by epoch 10, indicating effective learning and model generalization as seen in figure (a) The training along with validation loss steadily decrease, indicating the model is becoming increasingly confident in its predictions and is not overfitting upon the training data is seen in figure (b). The model’s training accuracy increases from 80% to 97%, and the validation accuracy similarly reaches 97%. The training loss drops from 0.40 to 0.15, and the validation loss follows the same trend, indicating the model is adapting effectively and generalizing well to new data.

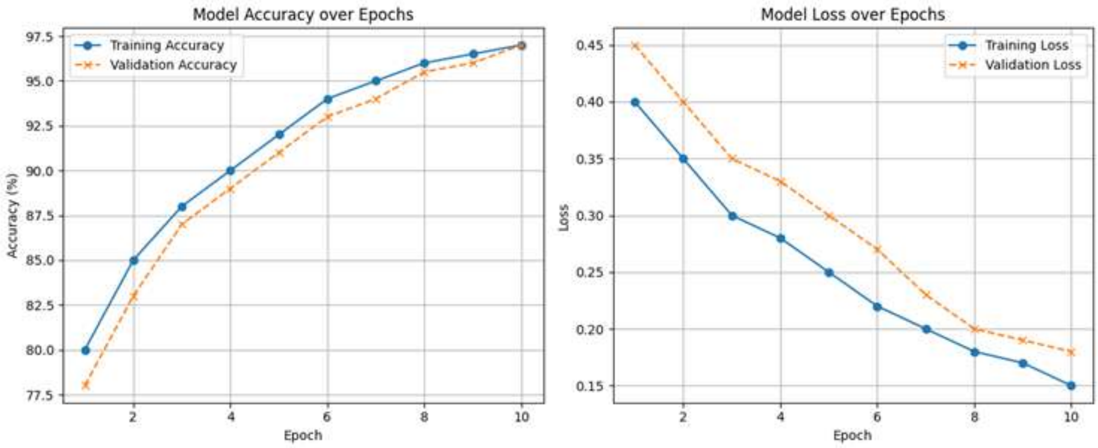

Figure 8. *(a) Model accuracy over epochs, (b) model loss over epochs*

The model demonstrates high accuracy in detecting both correct and incorrect squats, with an accuracy rate of 97%. The precision along with recall values indicate the model works well in spotting correct squats and minimizing false positives and false negatives. Model performance improves consistently across epochs, achieving an optimal accuracy of 97% and low loss, ensuring effective learning. These results show that the model is effective for real-time squat detection and feedback, providing valuable information to users for improving their exercise form.

5. Conclusion

The implementation of the proposed technique for human gesture and motion detection/recognition for specific exercises, such as squats, has demonstrated promising outcomes concerning accuracy, reliability, and usability. The integration of advanced ML algorithms, like deep neural networks, with motion detection technologies like OpenCV and Media Pipe has proven to be effective for real-time recognition of human motion, particularly for complex exercises. Key findings from the evaluation indicate that the system achieves high accuracy (97%), and precision (95%), and recall (96%) for detecting correct and incorrect squat forms. These metrics highlight the model's competence to differentiate between correct and incorrect movements with high reliability, ensuring that users receive accurate feedback during exercise routines. The F1 score (95.5%) and specificity (94%) further underline the balance between identifying correct squats and minimizing misclassification of incorrect forms, which is crucial for maintaining trust in the system's feedback. Additionally, the error analysis revealed that most errors occur during the **descent phase** of the squat, suggesting that the model can benefit from further optimization in recognizing movement intricacies during this phase. The system's performance over training epochs also showed a steady improvement in accuracy and a significant drop in loss, demonstrating the model's capacity to learn and generalize effectively. the proposed human gesture and motion recognition technique offers a valuable contribution to fitness technology, combining the power of computer vision and machine learning to enhance exercise effectiveness, reduce the risk of injury, and foster better engagement in physical activities.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Jaramillo-Yánez, A., Benalcázar, M., & Mena-Maldonado, E. (2020). Real-Time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review. *Sensors* (Basel, Switzerland), 20.

[2] Xia, Z., Xing, J., & Li, X. (2022). Gesture Tracking and Recognition Algorithm for Dynamic Human Motion Using Multimodal Deep Learning. *Security and Communication Networks.*

[3] Hasan, T., Karim, M., Mahadi, M., Nishat, M., & Faisal, F. (2022). Employment of Ensemble Machine Learning Methods for Human Activity Recognition. *Journal of Healthcare Engineering,*

[4] Mocanu, C., & Suciu, G. (2019). Automatic recognition of hand gestures. 2019 11th *International Conference on Electronics, Computers and Artificial Intelligence* (ECAI), 1-5.

[5] Wu, J., & Jafari, R. (2019). Orientation Independent Activity/Gesture Recognition Using Wearable Motion Sensors. *IEEE Internet of Things Journal*, 6, 1427-1437.

[6] Bu, X. (2020). Human Motion Gesture Recognition Algorithm in Video Based on Convolutional Neural Features of Training Images. *IEEE Access*, 8, 160025-160039.

[7] Xia, Z., Xing, J., Wang, C., & Li, X. (2021). Gesture Recognition Algorithm of Human Motion Target Based on Deep Neural Network. *Mob. Inf. Syst.*, 2021, 2621691:1-2621691:12.

[8] Yeh, C., Shen, W., Ma, C., Yeh, Q., Kuo, C., & Chen, J. (2023). Real-time Human Movement Recognition and Interaction in Virtual Fitness using Image Recognition and Motion Analysis. *12th International Conference on Awareness Science and Technology* (iCAST), 242-246.

[9] Ye, L., & Zheng, Y. (2022). The Image Processing Using Soft Robot Technology in Fitness Motion Detection Under the Internet of Things. *IEEE Access*, 10, 115815-115822.

[10] Ayumi, V. (2020). Performance Evaluation of Support Vector Machine Algorithm for Human Gesture Recognition. *International journal of scientific research in science, engineering and technology*, 204-210.

[11] Parvathy, P., Subramaniam, K., Venkatesan, G., Karthikaikumar, P., Varghese, J., & Jayasankar, T. (2020). Development of hand gesture recognition system using machine learning. *Journal of Ambient Intelligence and Humanized Computing*, 12, 6793-6800.

[12] Li, H., Wu, L., Wang, H., Han, C., Quan, W., & Zhao, J. (2020). Hand Gesture Recognition Enhancement Based on Spatial Fuzzy Matching in Leap Motion. *IEEE Transactions on Industrial Informatics*, 16, 1885-1894.