AI-Based Recognition and Analysis of Gatka Martial Arts Techniques Using Human Pose Estimation and Machine Learning

[1] Ms. Prabhjyot Kaur, [2] Dr. Anija

[1] Student, Department of Computer Applications, Guru Kashi University, Bathinda, India

[2] Assistant Professor, Department of Computer Applications, Guru Kashi University, Bathinda, India

Email Id: [1] prabhjyotkour9@gmail.com, [2] anijamittal@gmail.com

Accepted:20.07.2026 Published:10.08.2026 Page No. 89 – 96 DOI: 10.5281/zenodo.21802638

Abstract— *Gatka is a traditional Sikh martial art, which is defined by synchronized movement of the body, the weapons, the defensive actions and the quick transitions. Teaching and evaluation are mostly based on the expertise of teachers and the process of observation, making it hard to quantify, provide quick feedback, and systematically preserve in digital format. This paper proposes a framework called AI-GATKA which utilizes machine learning and human pose estimation to achieve the aim of recognizing and analysing basic Gatka techniques. Media Pipe Pose processes video frames to extract body landmarks such as shoulder, elbow, wrist, hip, knee and ankle positions. Landmark points are transformed to posture, joint-angle, alignment and temporal-motion features. A Support Vector Machine (SVM), Random Forest and Neural Network classifiers are then employed to determine five technique categories: basic stance, forward strike, defensive block, side movement, and circular movement. The data set used in the prototype configuration reported in the source study consists of 250 samples, with 80% of the data used for training and 20% of the data used for testing. The SVM achieved 86%, the Random Forest 91% and the Neural Network 94% accuracy in the classification. This framework shows how the use of artificial intelligence with poses can help with technique recognition, training, objective assessments and digital preservation of Gatka. The paper also outlines the main drawbacks of camera viewpoint, lighting, among practitioners, occlusion, and fast movements of weapons. Main keywords— Artificial Intelligence, Computer vision, Gatka recognition, Human action recognition, Machine Learning, MediaPipe Pose, Pose-based Analysis.*

Keywords – artificial intelligence, computer vision, Gatka recognition, human action recognition, machine learning, MediaPipe Pose, pose-based analysis.

I. INTRODUCTION

Martial arts have a deep history of transmission of Knowledge and embodied Practice. They develop their self-defence skills and their physical conditioning, discipline, concentration, coordination, timing and control of their decision making. Digital documentation may be used in addition to oral, instructor-led, and teacher-student transmission to ensure that movements can be reproduced, and measurable descriptions of the way technique is executed.

In recent years, the automatic interpretation of human motion has significantly improved thanks to the advances in artificial intelligence, computer vision and deep learning. This is a description of the fast developments in pose estimation, both two-dimensional and three-dimensional, in the fields of keypoint detection, tracking and action recognition [1, 3, 4]. MediaPipe offers a highly effective perception-pipeline architecture, which is also capable of extracting landmarks from regular RGB video in real time [2]. These have allowed pose-based applications, such as fitness, sports activity classification, healthcare, rehabilitation and human-computer interaction [5]-[6] [21]-[22] [27].

The Gatka technique is a unique recognition problem as it incorporates body position, movements with the feet, rotations and weapon lines. Landmark estimates can be degraded by self-occlusion, clothing, motion blur, background activity and overlap of the body with the weapon, and there can be variability in the technique across practitioners. Martial arts that are not traditional weapons are not well covered in existing datasets that focus on action-recognition in daily life, fitness exercises, and popular sports. Therefore, a special technique-level framework and data set is needed for the Gatka analysis.

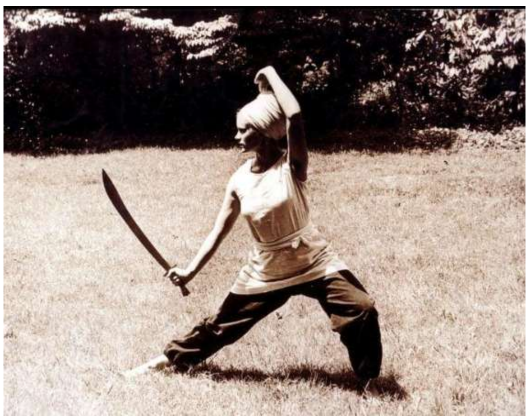


Fig. 1. Example Gatka posture illustrating coordinated stance and weapon handling.

This research aims to present an AI-GATKA framework to extract the pose landmarks, movement features and machine readable labels of the technique of Gatka from the videos. The contribution is threefold: (1) it provides a pose-centred pipeline of the five basic Gatka techniques; (2) it puts forward a comparison between conventional and neural classification models, based on the reported prototype data; and (3) it enumerates applications for training support, objective assessment, and cultural preservation.

II. BACKGROUND & RESEARCH MOTIVATION

A. Gatka Martial Art

Gatka is a martial art tradition of the Sikhs that combines the use of weapons, Defensive and Response, Footwork, Body Alignment, Timing and Tactical Positioning. The upper and lower arm/legs and the weapon must work continuously together to perform well. Many of the details, such as the orientation of the joints, stance width, balance, hand path or timing of movement can vary slightly and affect the quality and safety of a technique.

- ***Weapon handling:*** Movement of traditional weapons in a controlled manner, in the correct direction, speed, accuracy and safety.
- ***Defensive movement:*** Defensively moving patterns in response to an incoming action such as blocking and avoidance patterns, and counter-movement patterns.
- ***Footwork:*** Effective and stable repositioning, maintaining balance and attacking and defending.
- ***Posture and coordination:*** Coordination and alignment of the trunk, arms, legs and weapon throughout a technique.
- ***Combat transition:*** Transition from stance, attack, defense and recovery at various distances and times.

Fig. 2. Gatka practitioners performing weapon-based movement sequences.

B. Research Motivation

Use of traditional evaluation is helpful because it is based on the knowledge of context, intention, and safety, as well as stylistic variation, that is gained by the instructor's years of experience. However, manual observation is not easily standardized, scaled and archived. A computer assisted system can provide consistent body geometry and timing of movement measurements, and can still leave the instructor at the conclusion as the final decision maker. AI-GATKA is thus not a replacement, but rather an assistant that improves teaching, based on measurable evidence, a longitudinal record and consistent preliminary feedback.

The research is motivated by four needs: digital preservation of technique examples; automated extraction of movement descriptors; objective comparison of learner and reference performances; and accessible training support when expert supervision is not continuously available. Several pose-based systems have been explored to correct exercise [7] – [14], to screen the biomechanics of an exercise [17] – [20] and to analyze injury risk [23] – [27] and sports performance [28]. The above developments suggest that a system designed accordingly can be modified to meet the demands of structural and cultural needs of Gatka.

III. RELATED WORK

Zheng et al. reviewed deep-learning methods for human pose estimation and provided an overview of the current development of keypoint detection, multi-person tracking and benchmark test [1]. MediaPipe is a framework of reusable components to build real-time perception pipelines on diverse platforms introduced by Lugaresi et al. [2]. Song et al. surveyed the relationship of pose estimation and action recognition, pose representations and temporal modelling strategies [3]. Zhou et al. also investigated deep models utilizing convolutional networks, graph neural networks, and Transformers in the areas of pose estimation, tracking, and action recognition [4].

In the domain of fitness and sports, pose-estimation (for exercise) and action-recognition techniques for exercise movements were proposed by Fu et al. [5] and Sport PoseNet used MediaPipe landmarks and a deep image model, achieving 92.67% validation accuracy on a custom multi-sport dataset, respectively [6]. A series of related studies have focused on the assessment of squat, which can be interpreted, temporal models for prediction of injuries, tracking of the skeleton, lower-limb movements, and correction of the movement [7]-[14]. There are various sport-specific applications such as pose-based tackle detection, skiing analysis, running joint-angle assessment, landing-error scoring and athlete tracking [15]-[20]. Real-time systems for pose tracking and skeletal analysis, as well as graph-based analysis of the skeletons and biomechanical study of different seasons, show that the trend towards data-driven sports coaching is ongoing [21]-[27].

However, this development does not seem to have been reflected in the existing literature, which is still lacking of a special Gatka dataset or a technique-recognition pipeline that specifically considers weapon-based movement. The majority of systems rate general activities or isolate exercises. This leaves a gap in the research to be filled with culturally specific data collection, handling of the rapid movements of the weapons, technique-level labelling and feedback appropriate to a traditional martial-art learning context.

TABLE I

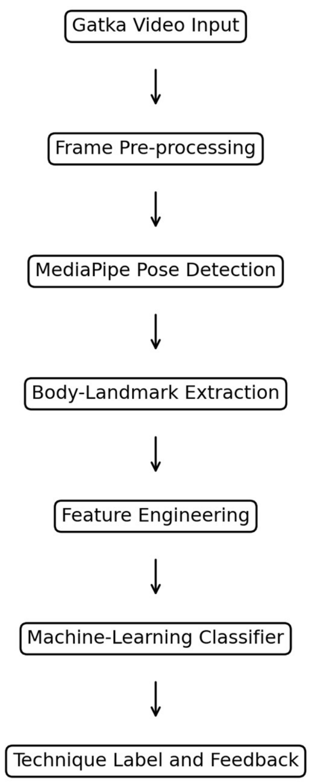
SUMMARY OF REPRESENTATIVE POSE-BASED STUDIE

Ref.	Data/Domain	Core Method	Reported Outcome or Gap
[1]	Pose benchmarks	Deep pose-estimation survey	Broad taxonomy; no Gatka-specific dataset
[2]	RGB perception pipelines	MediaPipe framework	Real-time modular processing; general-purpose
[3], [4]	Pose/action datasets	CNN, GNN, Transformer methods	Strong general action coverage; limited traditional martial arts
[5]	Fitness movements	Pose estimation and action recognition	Sports-oriented recognition; no weapon-based analysis
[6]	1,010 sports images	MediaPipe + ResNet-200	92.67% validation accuracy; common sports only
[7]-[14]	Exercise and biomechanics	Pose, temporal and biomechanical models	Demonstrates correction and risk-assessment potential

Ref.	Data/Domain	Core Method	Reported Outcome or Gap
[15]-[27]	Sport-specific motion	ML, deep learning and skeletal tracking	Supports coaching applications but not Gatka technique labels

IV. PROPOSED AI-GATKA FRAMEWORK

The pipeline proposed in this work consists of a series of stages including frame preparation, pose detection, feature extraction and classification of the technique. Design is based on the concept of pose coordinates, rather than on raw image pixels. This lowers the dimensionality of the input, simplifies the features to be more interpretable and enables the system to focus on the practitioner's geometry, and not on irrelevant visual background information.



A. Data Collection

Video is taken of Gatka practitioners doing the basic movements with the help of video cameras in a controlled and varied environment. The proposed categories are: basic stance, forward strike, defense block, side move and circular move. Videos will feature different heights of practitioners, clothing, tempo of movement, distance of camera and different styles of executing the movement. Technical validity of the labels should be assigned or confirmed by qualified Gatka instructors.

B. Human Pose Detection

First, each video is divided into frames and then each frame is processed using Media Pipe Pose [2]. There are 33 landmarks identified by the model, with some landmarks being of particular interest for stance, arm trajectory, trunk orientation, and foot placement: shoulders, elbows, wrists, hips, knees, and ankles. For each landmark a horizontal coordinate and a vertical coordinate is stored, as well as a depth coordinate and visibility and/or confidence information.

C. Feature Extraction

The significant stream becomes numerical attributes of the stream. The position features include normalized elbow and shoulder locations, hip and knee alignments, and trunk inclination and bilateral symmetry; the angle features include elbow, shoulder, hip and knee positions; the alignment features include trunk inclination and bilateral symmetry; and the temporal features include velocity, acceleration, direction change and frame to frame displacement. Normalization with respect to hip centre or shoulder width, or body height, decreases the sensitivity to the distance between the camera and the participant or the size of the participant.

D. Technique Classification

The extracted feature vectors are then given to the supervised classifiers. On moderate dimensional features a Support Vector Machine (SVM) would be a good baseline, Random Forest would provide non-linear decision boundaries and feature-importance estimates, and a Neural Network (NN) could learn more complex relationships between posture and motion descriptors. Model selection should be done based on the cross-validated accuracy, precision, recall, F1-score, latency and robustness, and not accuracy alone.

V. METHODOLOGY

A. Dataset and Pre-processing

The prototype data set reported is 50 samples per 5 techniques of Gatka which makes a total of 250 samples. The data set is split into 80% training data and 20% testing data. The videos are cleaned of corrupted or irrelevant frames and frames are standardized prior to landmark extraction. Ideally, the split between train and test sets should be "practitioner independent", meaning that a particular practitioner should not be in both sets: otherwise, the accuracy measured may not reflect technique recognition for general practitioners, but rather for that practitioner.

TABLE II

PROPOSED TECHNIQUE CLASSES

Technique	Operational Description
Basic stance	Ready posture before an offensive or defensive action
Forward strike	Weapon movement directed toward the opponent
Defensive block	Protective action intended to intercept or redirect an attack
Side movement	Lateral repositioning while preserving balance and readiness
Circular movement	Rotational movement coordinated with weapon handling

B. Landmark and Angle Representation

The landmark vector for landmark i at frame t is P_t = (x_i, y_i, z_i, v_i) and v_i is the landmark's visibility. The cosine of the angle at B, between the vectors BA and BC, is calculated, and this angle is used to determine the angle at B, between A-B-C. Time derivatives of landmarks or angles are used to get velocity and acceleration descriptors. Low confidence landmark points can be interpolated, smoothed or omitted based on a quality criterion.

C. Model Training and Evaluation

Feature scaling process is used if needed, primarily for training SVM and Neural Networks. The hyperparameters should be determined by cross validating on the training set. Final model is tested on the test set that was not used in the training process. Apart from overall accuracy, a confusion matrix and class-wise precision, recall and F1-score are recommended as similar techniques might be misinterpreted even if the overall accuracy is high.

VI. EXPERIMENTAL SETUP

The proposed implementation is based on Python, OpenCV, MediaPipe Pose, NumPy and Pandas, and scikit-learn which are employed for frame processing, extracting landmarks, preparing the features, and implementing the conventional machine-learning models. Prototype development can be done on a standard desktop or laptop computer, with a minimum of an Intel Core i5 processor, 8 GB of RAM, and an HD camera. The performance of the deployment in real time is dependent on resolution, frame rate, complexity of the model and hardware acceleration.

TABLE III

SOFTWARE & HARDWARE CONFIGURATION

Component	Specification
Programming language	Python
Computer vision	OpenCV
Pose estimation	MediaPipe Pose
Machine learning	Scikit-learn
Data processing	NumPy and Pandas
Processor	Intel Core i5 or higher
Memory	Minimum 8 GB RAM
Camera	HD webcam or equivalent
Operating system	Windows 10/11 or Linux

VII. RESULTS & DISCUSSION

According to the source manuscript the Support Vector Machine, Random Forest and Neural

Network had prototype accuracy of 86%, 91% and 94%. The Neural Network is thus the most accurate of the three, with the reported accuracy of 78.3% compared to 75.3% for Random Forest and 70.3% for SVM. The pattern is consistent as a neural model can accommodate the non-linear relationships between a number of joint and motion features. The accuracy values reported however should be considered in conjunction with the sampling protocol, class balance, independence of the participants and the lack of class-wise metrics.

TABLE IV

REPORTED MODEL ACCURACY

Model	Accuracy
Support Vector Machine	86%
Random Forest	91%
Neural Network	94%

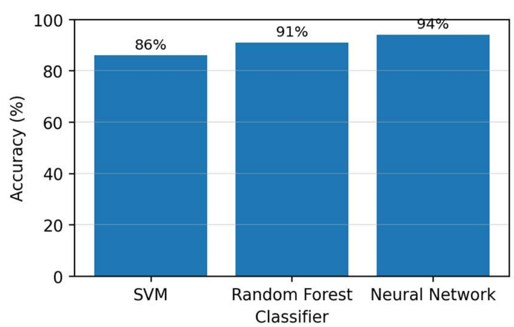


Fig. 4. Reported classification accuracy of the three models.

There are several things which can decrease the quality of the recognition. Body-only pose models may not be a direct representation of fast-moving weapons. Wrists, elbows or lower-limb landmarks could be obscured by extreme camera angles and/or self-occlusion. Landmark confidence may be decreased with lighting and background variation. Lastly, the practitioner can perform the same technique by a valid style of variation. Practitioner-independent testing, landmark-confidence filtering, temporal sequence modelling and integration of weapon detection/tracking are motivators for these issues.

VIII. APPLICATIONS

The proposed framework is able to facilitate interactive training by comparing the movement of learners with examples validated by instructors and give a general idea of the deviation in the stance, alignment and timing of the movement of the learner. It can help teachers to create uniform initial measurements and session histories. Video and landmark repository could also aid in Digital Preservation of Gatka techniques. Further validation could be expanded to mobile coaching, remote practice review, competition analysis and Augmented/Virtual reality learning environments.

IX. LIMITATIONS & FUTURE WORK

Some of the main drawbacks of the prototype are the small size and balanced nature of the dataset, the lack of evidence of practitioner diversity, possible data leakage if performers are present in both the training set and the test set, and the lack of detailed precision, recall, F1-score and confusion-matrix results. The body landmarks provided by MediaPipe are not a direct model of the weapons, but they are used to estimate them. The estimation of the body landmarks is crucial for Gatka. The need for real-time feedback, culturally relevant technique annotations and quality scoring by the instructor also needs further development.

Future research needs to create a large multi-view, expert-annotated dataset for Gatka, assess practitioner-independent and cross-environment generalization, integrate weapon detection and tracking of the trajectory, compare recurrent, temporal-convolutional, graph-based and Transformer models, quantify the latency on mobile hardware, and undertake usability studies with learners and teachers. Feedback should be able to be explained, and should not be presented as algorithmic output (or even "automatic feedback") instead of expert feedback.

X. CONCLUSION

We have developed a framework using AI to recognize and analyse basic Gatka technique using Media Pipe Pose and machine-learning classification. The framework converts a video into body landmarks, posture and motion feature and technique prediction. The results of the prototype experiments in the report indicate that the Neural Network outperforms Random Forest and SVM, and indicate that these techniques should be tested with more robust validation, using more diverse data, and with weapon-aware modelling. The fusion of traditional Gatka knowledge with pose-based AI provides a stepping stone towards technology-assisted training, precise movement analysis, and responsible digital preservation.

REFERENCES

[1] C. Zheng, W. Wu, C. Chen, T. Yang, S. Zhu, J. Shen, et al., "Deep learning-based human pose estimation: A survey," ACM Computing Surveys, vol. 56, no. 1, pp. 1–37, 2023.

[2] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, et al., "MediaPipe: A framework for building perception pipelines," arXiv:1906.0817, 2019.

[3] L. Song, G. Yu, J. Yuan, and Z. Liu, "Human pose estimation and its application to action recognition: A survey," Journal of Visual Communication and Image Representation, vol. 76, Art. no. 103055, 2021.

[4] L. Zhou, X. Meng, Z. Liu, M. Wu, Z. Gao, and P. Wang, "Human pose-based estimation, tracking and action recognition with deep learning: A survey," arXiv:2310.13039, 2023.

[5] H. Fu, J. Gao, and H. Liu, "Human pose estimation and action recognition for fitness movements," Computers & Graphics, vol. 116, pp. 418–426, 2023.

[6] M. V. Saakshi, D. Rao, M. Nidhi, A. Hurkadli, and S. Varur, "SportPoseNet: Leveraging pose estimation and deep learning for sports activity classification," 2024.

[7] P. S. et al., "SquatNet: An interpretable deep learning framework for real-time biomechanical analysis and injury-risk assessment in squat exercises," publication details not provided in the source manuscript.

[8] Weng et al., "Real-time squat pose assessment and injury risk prediction based on enhanced temporal convolutional neural networks," International Journal of Medical and All Body Health Research, 2024, doi: 10.54660/ijmbhr.2024.5.1.53-62.

[9] Bauer et al., "Unlocking the future of injury prevention: AI-powered pose estimation in real time linked with biomechanical simulation," 2025, doi: 10.1109/AMLDS63918.2025.11159438.

[10] Portafaix et al., "Improving accuracy and runtime of skeletal tracking of lower limbs for athletic jump mechanics assessment," 2021, doi: 10.1109/EMBC46164.2021.9629726.

[11] Jubair et al., "Machine learning for real-time exercise correction and injury prevention: A systematic review," 2025, doi: 10.22541/au.174466167.73334339/v1.

[12] Ren et al., "IoT-based 3D pose estimation and motion optimization for athletes: Application of C3D and OpenPose," 2024, doi: 10.48550/arXiv.2411.12676.

[13] Armstrong et al., "Validation of human pose estimation and human mesh recovery for extracting clinically relevant motion data from videos," arXiv:2503.14760, 2025.

[14] "Prediction of lower-limb kinematics from a vision-based system using deep-learning approaches," in Proc. IEEE EMBC, 2022, doi: 10.1109/EMBC48229.2022.9871577.

[15] Nishio et al., "Objective detection of high-risk tackle in rugby by combination of pose estimation and machine learning," Lecture Notes in Computer Science, 2023, doi: 10.1007/978-3-031-29168-5_15.

[16] Zwölfer et al., "Deep learning-based 2D keypoint detection in alpine ski racing—A performance analysis of state-of-the-art algorithms applied to regular skiing and injury," publication details not provided in the source manuscript.

[17] Ino et al., "Validity and reliability of OpenPose-based motion analysis in measuring knee valgus during the drop vertical jump test," Journal of Sports Science and Medicine, 2024, doi: 10.52082/jssm.2024.515.

[18] Krishnan et al., "Can BlazePose accurately assess joint angles in outdoor running environments?" 2025, doi: 10.2478/bhk-2025-0002.

[19] Hébert-Losier et al., "The 'DEEP' landing error scoring system," Applied Sciences, vol. 10, no. 3, Art. no. 892, 2020, doi: 10.3390/app10030892.

[20] Tang, "Real-time body movement tracking of athletes via a CNN-based approach," Applied and Computational Engineering, 2024, doi: 10.54254/2755-2721/67/20240609.

[21] Baca et al., "Ubiquitous computing in sports and physical activity—Recent trends and developments," Sensors, vol. 22, Art. no. 8370, 2022, doi: 10.3390/s22218370.

[22] Zhu et al., "Transformative skeletal motion analysis: Optimization of exercise training and injury prevention through graph neural networks," Frontiers in Neuroscience, 2024, doi: 10.3389/fnins.2024.1353257.

[23] Katre et al., "Pose tracking, analysis and impact estimation in real time," publication details not provided in the source manuscript.

[24] Joseph et al., "Temporal and sex-related differences in knee biomechanics over the course of the varsity athletic season: Pre- and postseason knee kinematics in collegiate varsity athletes using Kinect," Orthopaedic Journal of Sports Medicine, 2025, doi: 10.1177/23259671251386445.

[25] Kumar et al., "Prediction of lower-limb kinematics from a vision-based system using deep-learning approaches," in Proc. IEEE Engineering in Medicine and Biology Society, 2022, doi: 10.1109/EMBC48229.2022.9871577.

[26] Kanase et al., "Pose estimation and correcting exercise posture," ITM Web of Conferences, vol. 40, Art. no. 03031, 2021, doi: 10.1051/itmconf/20214003031.

[27] Xi et al., "Enhancing human pose estimation in sports training: Integrating a spatiotemporal Transformer for improved accuracy and real-time performance," Alexandria Engineering Journal, 2024, doi: 10.1016/j.aej.2024.08.072.