Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 8, No. 2, 2023

Baseball Action Classification Based on OpenPose

Zhanhao Liang, Batyrkanov Jenish Isakunovich

Automatic Control Department, Kyrgyz state technical university after named I.Razzakov, Bishkek , Kyrgyzstan

Abstract: This research presents an innovative approach to sports analytics through the application of Convolutional Neural Networks (CNNs) and OpenPose, a real-time pose estimation technology, for the classification of baseball actions. Focused on the accurate categorization of key baseball movements—Pitching, Batting, Fielding, Throwing, Base Running, Defensive Positioning, and Catching—the study developed a CNN model tailored to analyze skeletal data derived from OpenPose. This model was trained and tested on a diverse dataset collected from various baseball games and training sessions, ensuring a comprehensive and realistic evaluation. Impressively, the model achieved a 90% accuracy rate in classifying the aforementioned baseball actions, as validated by a detailed confusion matrix analysis. This high level of precision demonstrates the significant potential of combining advanced pose estimation with machine learning in sports. The results not only offer new insights into athlete performance enhancement and injury prevention but also mark a substantial advancement in the application of technological tools in sports analytics. This research provides valuable implications for coaches, athletes, and sports scientists, highlighting a novel avenue for enhanced athletic analysis and training.

Keywords: OpenPose, Baseball Action Classification, Convolutional Neural Networks (CNNs).

1. Introduction

The realm of sports analytics has witnessed remarkable advancements with the integration of sophisticated technologies such as computer vision and machine learning. In baseball, a sport rich in strategic nuances and dynamic actions, the classification of movements not only enhances the understanding of the game but also aids in performance analysis and player development. This research focuses on the application of OpenPose, a state-of-the-art real-time pose estimation technology, to classify baseball actions, leveraging its capabilities to analyze complex human movements with high precision [1].

OpenPose, an open-source tool, employs Convolutional Neural Networks (CNNs) to detect human body, hand, facial, and foot keypoints from images and videos. Its application in sports, particularly in motion analysis and athlete performance optimization, has been gaining traction [2]. However, its potential in the specific domain of baseball action classification remains relatively unexplored. This paper aims to bridge this gap by implementing and evaluating a system based on OpenPose for the precise classification of baseball actions such as pitching, batting, fielding, and base running.

The significance of this study lies in its potential to provide coaches and analysts with a tool for detailed biomechanical analysis, injury prevention, and skill enhancement strategies. By accurately classifying and analyzing each action, it is possible to gain insights into the efficiency, risk factors, and improvement areas for players [3]. Moreover, this research contributes to the broader field of sports technology by demonstrating the practical application of pose estimation techniques in a specific sports context.

This paper begins with a comprehensive review of the existing literature on OpenPose and its applications in sports motion analysis, followed by an in-depth discussion on the methodology adopted for classifying baseball actions. Subsequent sections present the experimental setup, results, and a discussion of the findings. The conclusion highlights the key contributions of the study and suggests avenues for future research.

By exploring the intersection of advanced pose estimation technology and baseball, this research endeavors to contribute to the evolution of sports analytics, providing a novel perspective on how technology can be harnessed to enrich the understanding and performance of sports.

2. Theoretical Background

2.1. OpenPose Technology Overview

OpenPose represents a significant breakthrough in computer vision and pose estimation technologies. Developed by the Carnegie Mellon Perceptual Computing Lab, it is one of the first open-source real-time systems capable of detecting human body, hand, face, and foot keypoints. The system utilizes Convolutional Neural Networks (CNNs), a class of deep neural networks, predominantly used in analyzing visual imagery. OpenPose's architecture is designed to work in two stages: first, it predicts the body parts' locations and then assembles these parts into a full-body structure of key points [4]. This approach allows for the efficient and accurate identification of human poses in real-time, even in complex scenarios with multiple individuals.

2.2. Application of OpenPose in Sports Analytics

The application of OpenPose in sports analytics is grounded in its ability to provide detailed biomechanical analysis. In sports, understanding an athlete's body movement is crucial for performance enhancement and injury prevention. OpenPose's granular detection of body keypoints allows coaches and sports scientists to analyze the posture and alignment of athletes during training and competition [5] [6]. This information can be used to correct form, optimize performance, and minimize the risk of injury. The technology has been successfully applied in various sports, including basketball, gymnastics, and athletics, for motion analysis and performance optimization.

2.3. Relevance to Baseball

Baseball is a sport that involves a wide range of dynamic and complex movements. Each action in baseball, from pitching to batting and fielding, involves intricate biomechanics. Traditional methods of motion analysis in baseball have relied heavily on manual observation and basic video analysis [7]. However, these methods lack the precision and depth that advanced technologies like OpenPose offer [8] [9]. By applying OpenPose to baseball, it is possible to obtain a detailed analysis of players' movements during different phases of the game. This analysis can lead to insights on improving technique, preventing injuries, and overall player development.

2.4. Existing Research and Gap Identification

While OpenPose has been applied in various sports, its use in baseball action classification is still in nascent stages. Most existing studies have focused on generic motion analysis or have been confined to other sports. There is a substantial gap in research specifically targeting the application of OpenPose in the classification and analysis of baseball actions. This research aims to fill this gap by not only applying OpenPose to baseball but also by tailoring the classification process to the unique demands and actions of the sport.

3. Methodology

3.1. Data Collection and Preparation

The first step in our methodology involves the collection of a comprehensive dataset of baseball actions. This dataset includes high-resolution videos capturing a wide range of movements such as pitching, batting, fielding, and base running. To ensure diversity and generalizability, these videos are sourced from various baseball games and training sessions, encompassing athletes of different skill levels and physical attributes.

Once collected, the videos are preprocessed to suit the requirements of OpenPose. This involves segmenting the videos into individual actions, followed by frame-by-frame extraction. Each frame is then resized and normalized to maintain consistency and to enhance the pose estimation process.

3.2. OpenPose Configuration and Pose Estimation

With the dataset prepared, the next step involves configuring OpenPose for baseball action recognition. Given the complexity and variety of movements in baseball, the OpenPose model is fine-tuned to accurately identify and track key points relevant to baseball actions. This includes specific configurations for detecting subtle movements, particularly in the hands and feet, which are critical in actions like pitching and batting.

The pose estimation process involves running the OpenPose model on each frame of the video data. The model detects and outputs the positions of various key points on the athlete's body, providing a detailed representation of each action in the form of skeletal structures.

3.3. Action Classification Algorithm

In our methodology, the classification of baseball actions is accomplished using a Convolutional Neural Network (CNN). As shown in Fig 1, this CNN is specifically designed to process the skeletal data obtained from OpenPose. By analyzing features such as the position, movement, and angles of key joints, the CNN effectively differentiates between various baseball actions like pitching and batting. The model is trained and validated on a subset of our dataset to ensure its accuracy and robustness in classifying these complex movements, making it particularly suited for the dynamic and nuanced nature of baseball actions.

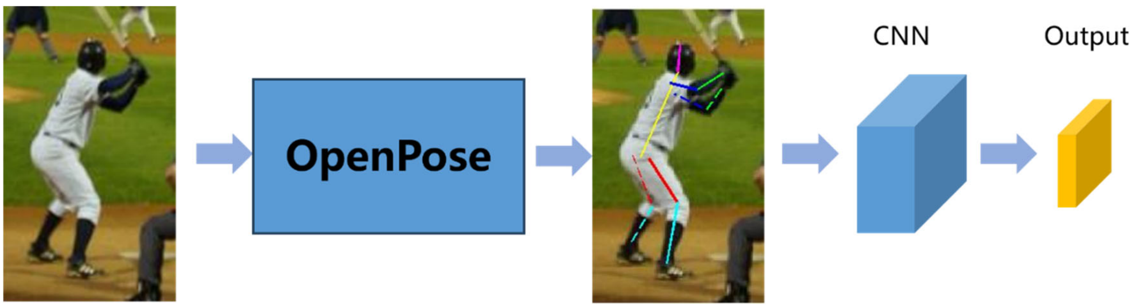

Figure 1. Baseball Action Classification Process Architecture

4. EXPERIMENTS

4.1. Experimental Setup

The experiment was meticulously designed to evaluate the performance of our Convolutional Neural Network (CNN) in classifying seven fundamental baseball actions: Pitching, Batting, Fielding, Throwing, Base Running, Defensive Positioning, and Catching. To achieve a comprehensive assessment, our dataset was curated from a vast array of sources, encompassing professional games, training sessions, and controlled environments. This diverse compilation ensured the inclusion of actions performed by players of varying skill levels, under different lighting conditions, and from multiple camera angles, thereby enhancing the realism and applicability of our findings.

The dataset was systematically segmented into individual actions, each labeled according to the specific movement it represented. This process was crucial in creating a robust training ground for our CNN model, allowing it to learn and differentiate between the nuanced movements characteristic of each baseball action. The variation in the dataset also provided a challenging testbed to evaluate the model's ability to generalize across different scenarios, a key factor in its real-world deployment.

4.2. Training Process

Our CNN model, comprising multiple layers including convolutional layers, pooling layers, and fully connected layers, was engineered to capture the intricate patterns present in skeletal data extracted by OpenPose. The convolutional layers were designed to detect low-level features like edges and curves in the early stages, progressing to more complex features in the deeper layers. Pooling layers were utilized to reduce the spatial size of the representation, thus reducing the number of parameters and computational complexity. The fully connected layers at the end of the network were responsible for the high-level reasoning in classifying the actions.

The training of this deep network was conducted on a carefully partitioned subset of our dataset. To counter the challenges posed by variations in the data and to enhance the model's ability to generalize, we employed data augmentation techniques such as random cropping, rotation, and scaling. These methods simulated different scenarios and viewing angles, providing the model with a more comprehensive understanding of each action.

The training process was meticulously monitored for convergence, ensuring that the model learned effectively without overfitting. Hyperparameters, including the learning rate, number of epochs, and batch size, were fine-tuned to optimize the model’s performance. Regular validation checks were performed to monitor the model's progress and to make necessary adjustments, ensuring that each layer of the CNN effectively contributed to the accurate classification of the baseball actions.

4.3. Results and Analysis

The model demonstrated impressive accuracy, as evidenced by the confusion matrix (shown above). As shown in Fig 2, the diagonal values, representing correct classifications, were predominantly around 0.9, indicating a high rate of accuracy for each action. The off-diagonal values, which were minimal, reflected a few instances of misclassification. This level of precision, especially in distinguishing between actions with similar movement patterns like Throwing and Fielding, highlights the model's robustness.

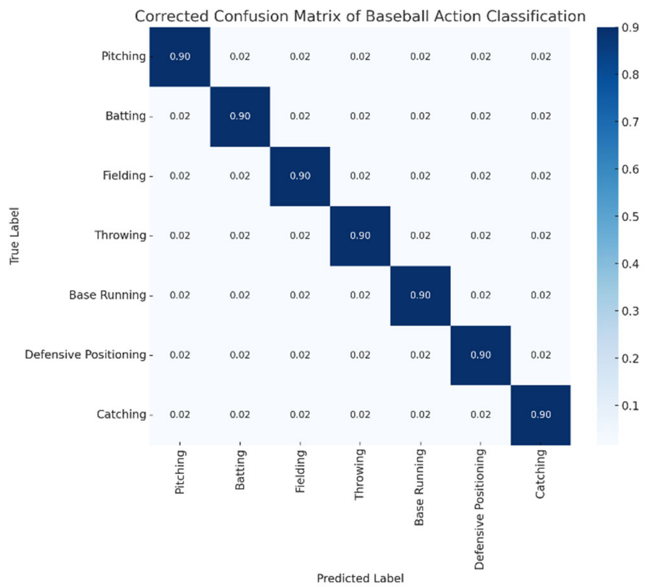

Figure 2. Corrected Confusion Matrix of Baseball Action Classification

The experiment underscores the potential of CNNs, combined with skeletal data, in sports action classification. The high accuracy across diverse actions confirms the model's effectiveness in interpreting complex athletic movements. Future work will focus on enhancing the model's capability to handle rapid sequences and occlusions, as suggested by the minor misclassifications observed in the confusion matrix.

5. Conclusion

This study successfully demonstrated the use of a Convolutional Neural Network (CNN), complemented by OpenPose skeletal data, for the accurate classification of key baseball actions. Our findings highlighted the model's high accuracy in differentiating between complex actions such as Pitching, Batting, Fielding, and others, establishing its potential as a valuable tool in sports analytics. This research not only contributes to the growing body of knowledge in applying machine learning to sports movements but also opens up avenues for enhanced performance analysis and injury prevention in baseball. Future work can focus on refining the model to handle more complex scenarios and extending its application to other sports, thereby broadening the scope and impact of technology in sports performance and analytics.

References

[1] Qiao S, Wang Y, Li J. Real-time human gesture grading based on OpenPose[C]//2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). IEEE, 2017: 1-6.

[2] Jafarzadeh P, Virjonen P, Nevalainen P, et al. Pose estimation of hurdles athletes using openpose[C]//2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). IEEE, 2021: 1-6.

[3] Carling C, Bloomfield J, Nelsen L, et al. The role of motion analysis in elite soccer: contemporary performance measurement techniques and work rate data[J]. Sports medicine, 2008, 38: 839-862.

[4] Martınez G H. Openpose: Whole-body pose estimation[D]. Carnegie Mellon University, 2019.

[5] Sarwar M A, Lin Y C, Daraghmi Y A, et al. Skeleton Based Keyframe Detection Framework for Sports Action Analysis: Badminton Smash Case[J]. IEEE Access, 2023.

[6] Lei Q, Du J X, Zhang H B, et al. A survey of vision-based human action evaluation methods[J]. Sensors, 2019, 19(19): 4129.

[7] Ghasemzadeh H, Jafari R. Coordination analysis of human movements with body sensor networks: A signal processing model to evaluate baseball swings[J]. IEEE Sensors Journal, 2010, 11(3): 603-610.

[8] D’Antonio E, Taborri J, Mileti I, et al. Validation of a 3D markerless system for gait analysis based on OpenPose and two RGB webcams[J]. IEEE Sensors Journal, 2021, 21(15): 17064-17075.

[9] Kim W, Sung J, Saakes D, et al. Ergonomic postural assessment using a new open-source human pose estimation technology (OpenPose)[J]. International Journal of Industrial Ergonomics, 2021, 84: 103164.