AI-Based Bowling Action Analysis and Correction

Priyanshu Ranjan
Project KITTY

May 24, 2025

Abstract

Cricket is a sport where biomechanics play a critical role, especially in bowling actions. Correct bowling technique is vital not only for performance but also for injury prevention. Among all phases of bowling, the run-up, jump, and release mechanics are crucial to ensure efficiency and legal delivery. The motivation for choosing bowling action analysis stems from the need to provide athletes with real-time feedback and correction without requiring expensive motion capture systems or coaching staff.

Medium pace bowlers often face a high risk of overuse injuries, due to repetitive strain and suboptimal biomechanics during the bowling action. Minor flaws in technique like poor alignment, insufficient jump control, or unstable head posture can go unnoticed but accumulate over time, leading to chronic issues.

An AI-based coaching system can act as a preventive tool by analyzing each delivery frame by frame, identifying risky deviations from ideal form, and providing personalized corrective feedback. This enables medium pacers to correct technique early, reduce injury risk, and consistently improve performance without always needing a human coach on hand.

This project focuses on analyzing the run-up and delivery phase of cricket bowlers using only video input, extracting pose-based features, and providing actionable feedback using a deep learning model and baseline biomechanical curves.

1 Architecture and Model Design

The proposed pipeline includes the following components to analyze and provide corrective feedback for Right-Arm medium pace bowlers:

1.1 Data Scraping and Preprocessing

Match footage was sourced from the official IPL website and curated cricket clips on Reddit. A custom script was used to download segmented .ts files from master M3U8 streams, which were then concatenated and converted to .mp4 format. These videos were standardized in resolution and playback speed slowed to 0.5× to improve pose estimation accuracy using MediaPipe.

To extract bowling actions, the videos were segmented based on the detection of shoulder and ankle movement, as these joints play a critical role in a bowler's delivery stride and follow-through. Using MediaPipe's pose tracking, joint velocities were analyzed to detect motion phases indicative of a bowling action. Relevant segments were then manually reviewed and selected to ensure precision.

An alternative approach was also explored, involving the automatic detection of frames where the pitch is visible from the bowler's end camera. This method was investigated with the intent of training a model capable of segmenting delivery phases based on pitch visibility, potentially enabling fully automated action extraction. The extracted segments were divided into good and bad bowling clips.

1.2 Pose Extraction

To improve pose estimation accuracy and focus only on the bowler, each frame is first passed through a YOLOv8 object detection model trained on a custom Roboflow dataset to detect cricketers.

After the bowler is identified, the bounding box is cropped from the original frame and passed to MediaPipe Pose to extract 3D body landmarks. This targeted processing increases accuracy and reduces noise from background players or crowd.

These landmarks, along with timestamps and frame indices, were stored in structured JSON files. The dataset was organized into separate directories for good and bad bowling samples.

To address privacy and ethical considerations, all data was fully anonymized focusing exclusively on biomechanical motion patterns without capturing or storing any personally identifiable information (PII). Only the extracted pose keypoints were retained, omitting any visual or identifying cues of the players. These JSON files were securely uploaded to a Firebase Realtime Database, and the model was trained exclusively on data retrieved from this database to ensure a privacy-preserving training pipeline.

1.3 Feature Engineering

From the landmark sequences, the following time-series features are extracted:

- Left and Right Elbow Angle
- Shoulder Alignment (horizontal deviation)
- Jump Height (ankle vertical position)
- Head Stability (variance in head Y-position)
- Hip-Shoulder Separation (torque indicator)

1.4 Model: LSTM with Attention

An LSTM with attention mechanism is used to capture temporal dependencies in the feature sequence:

- Input: (Timestamps $\times$ Features)
- LSTM Layer(s)
- Attention Layer
- Dense output with sigmoid activation (binary: good/bad)

The attention mechanism can also be used to highlight the most critical moments during the bowling action that influenced the decision, making the model more interpretable in future work.

1.5 Feedback Mechanism

For bad predictions:

- The user feature curves are compared against the baseline 'good' curves (average temporal curves for each key feature) stored as JSON.
- Signed deviations are computed and interpreted to generate feedback.
- Time-localized feedback is also visualized on video.

2 Baseline Curve Comparison

Each biomechanical feature is time-aligned and compared to its corresponding baseline — a smoothed average of several good bowling samples.

Steps Involved:

- Normalize and resample player's feature vector to match the baseline temporal resolution.
- Compute signed difference at each timestep.
- Highlight time intervals where deviation exceeds a pre-defined threshold.

Feedback Generation:

- For each significant deviation, generate a textual correction (e.g., "Keep head steadier after jump").

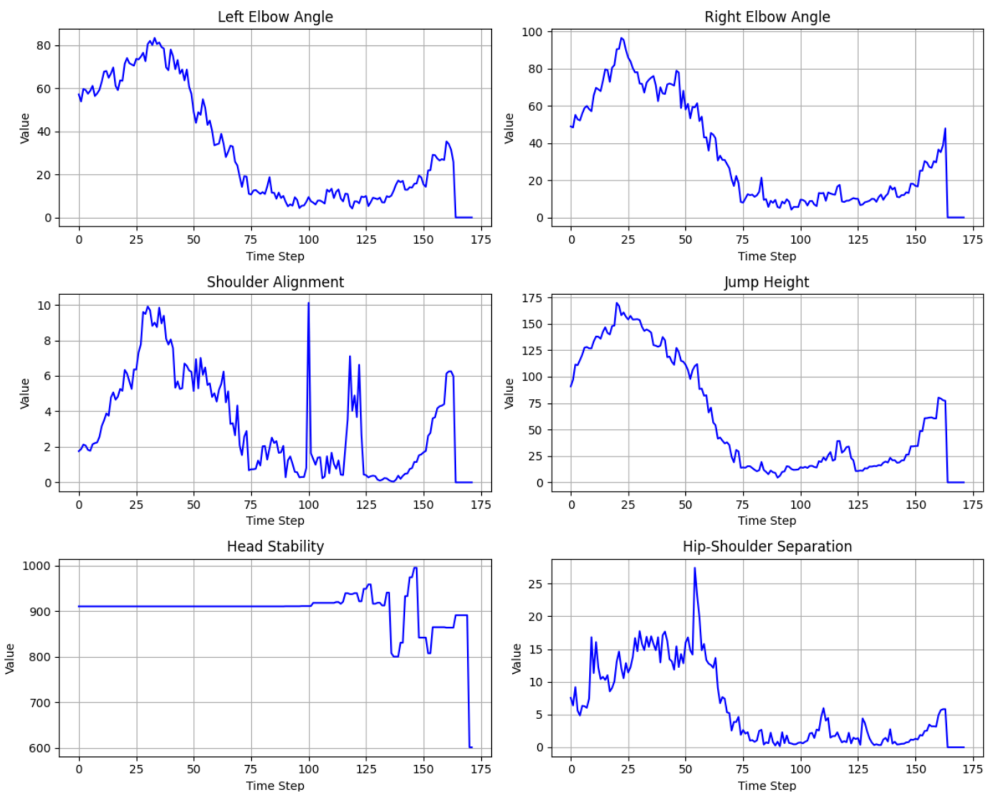

Figure 1: Baseline curves for different features

3 Qualitative and Quantitative Results

3.1 Quantitative Performance

The model was trained on labeled video data consisting of good and bad bowling samples.

- **Training Accuracy:** 87%
- **Validation Accuracy:** 92%
- **Inference Time:** ∼0.2s per sample

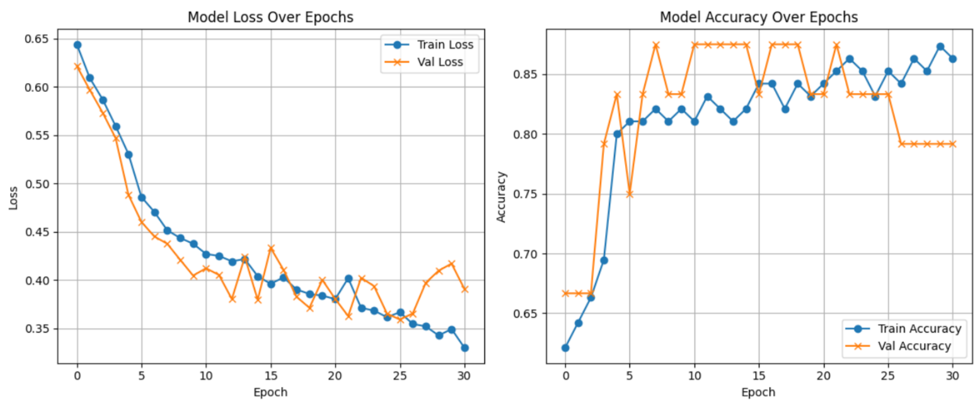

Figure 2: Model training

3.2 Qualitative Feedback Examples

Input Video: Frame-wise pose sequence from a bowler. **Model Output:** Predicted class = **Bad** (Prediction: 0.22)

Generated Feedback:

- **Left Elbow Angle:** Above ideal — try to relax your elbow more during the run-up.
- **Right Elbow Angle:** Above ideal — try to relax your elbow more during the run-up.
- **Jump Height:** Below ideal — try to increase your jump height for more momentum.
- **Head Stability:** Below ideal — you might be too rigid; allow slight natural head movement.
- **Hip-Shoulder Separation:** Above ideal — your separation is excessive; try to sync your hips and shoulders better.

Visual Overlay: Video frames showing pose skeleton

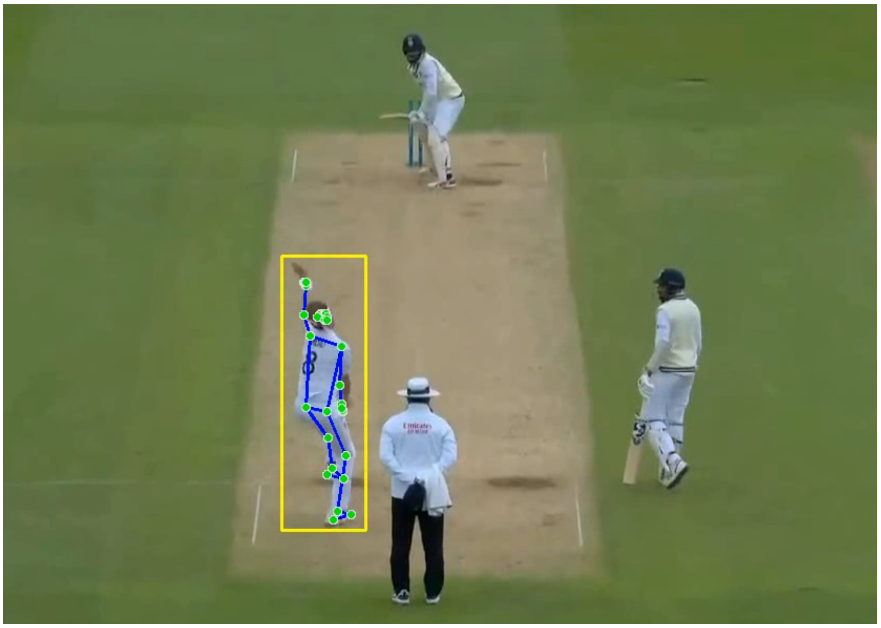

Figure 3: Visual showing captured pose skeleton with bounding box

4 Conclusion

This project successfully demonstrates a lightweight, interpretable system for analyzing cricket bowling actions using deep learning and pose estimation. The integration of attention-based LSTM and biomechanical feedback provides both classification and actionable suggestions, making it highly applicable in coaching and performance improvement contexts.

5 Future Work

- Expand feedback and analysis for more bowling styles
- Expand to multi-camera or 3D enhanced data
- Extend feedback to foot placement and wrist movement
- Incorporate sensors to better analyze bowling style
- Deploy as a mobile app for field use
- Incorporate reinforcement learning for personalized training plans