

Badminton AI Analysis: LangGraph Orchestration Pipeline

ShreeRaj Mummidivarapu

June 13, 2025

What is Agentic AI?

- **Traditional ML/AI Systems (e.g., CNNs, Vision Transformers, LLMs):**
 - **CNNs (Convolutional Neural Networks):** Excellent for image classification, object detection. Primarily pattern recognition.
 - **Vision Transformers (ViTs):** Leverage self-attention for image tasks, capturing global dependencies. Still largely reactive.
 - **LLMs (Large Language Models):** Powerful for text generation, understanding. Can exhibit emergent reasoning but lack inherent agency or persistent state.
 - **Common Characteristics:**
 - *Monolithic & Reactive:* Designed for specific tasks, respond to input without internal goals or long-term planning.
 - *Limited Self-Correction:* Require retraining for significant behavioral changes.
 - *No Persistent State:* Each interaction is often independent, lacking memory across sessions.

Agentic AI Systems

- **Composed of autonomous agents with specific roles.**
- Possess capabilities such as:
 - Planning
 - Memory
 - Tool use
 - Self-reflection
- Can break down complex tasks, orchestrate actions, and adapt to new information.
- Aim for more human-like problem-solving, decision-making, and continuous learning.
- **Key Capabilities:** Planning, memory, tool use, self-reflection, and dynamic adaptation.
- **Examples:** Autonomous research agents, complex task automation systems, adaptive control systems.

LangGraph Pipeline for Badminton Analysis (1/2)

- **Multimodal Data Integration:**
 - Effective integration of video analysis (pose metrics) and audio transcription in a unified pipeline.
 - Sequential processing with comprehensive error handling for reliable results.
- **LangGraph for Sports Analysis:**
 - Pioneering the use of LangGraph for structured pipeline orchestration in sports analytics.
 - Foundation for future expansion to more complex agent interactions and workflows.
- **Actionable, Granular Feedback:**
 - Focus on generating highly specific, actionable feedback for players and coaches.
 - Moves beyond descriptive statistics to prescriptive recommendations.
- **Scalable & Extensible Architecture:**
 - Modular pipeline design allows for easy addition of new analysis capabilities in future iterations.

LangGraph Pipeline for Badminton Analysis (2/2)

- **Holistic System View:**

- Our system seamlessly combines video (player movement) and audio (speech transcription) for comprehensive understanding.

- **Linear Pipeline Orchestration:**

- Utilizes a four-node LangGraph pipeline for video processing, audio processing, and report generation.
- Sequential processing with state management for efficient data flow between components.

- **Automated, Granular Reporting:**

- Generates detailed, objective reports with actionable insights.
- Identifies strengths and areas for improvement to support targeted coaching decisions.

How It Works: High-Level Overview

Input:

- Badminton match video.

Processing:

- Video frames are processed for pose estimation and audio is transcribed for speech content.
- LangGraph orchestrates a linear pipeline with four processing nodes for efficient data flow.

Output:

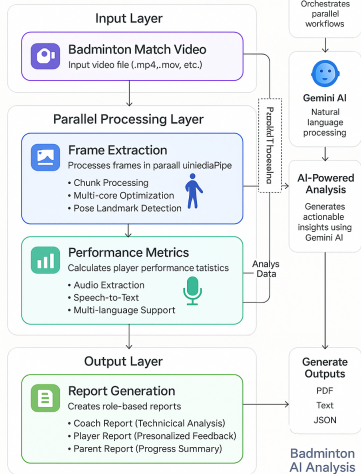
- Comprehensive text and PDF reports with actionable insights and strategic recommendations.

System Pipeline Overview

System Architecture Diagram

Badminton AI Analysis – System Architecture

End-to-end architecture diagram for the parallel badminton video analysis system



System Pipeline: High-Level Walkthrough (1/2)

- **1. Video Input:**

- Raw match footage (e.g., MP4 files).
- Supported formats: MP4, AVI, MOV with H.264/H.265 encoding.
- Optimal resolution: 1080p (1920×1080) at 60fps for detailed motion capture.

- **2. Frame & Audio Extraction:**

- Video is processed to extract individual frames for visual analysis.
- Audio track is separated and processed for speech transcription.
- Preprocessing includes frame resizing and color conversion.

System Pipeline: High-Level Walkthrough (2/2)

- **3. Pipeline Orchestration with LangGraph:**

- Extracted data flows through a four-node linear pipeline orchestrated by LangGraph.
- Pipeline includes video processing, audio transcription, data integration, and report generation.
- State management ensures efficient data flow between processing steps.

- **4. Report Generation:**

- Consolidated data is processed by Google Gemini API to generate comprehensive text and PDF reports.
- Reports include performance metrics, observations based on pose data, and actionable feedback.
- Customizable templates based on user role (player, coach, analyst).

Detailed Technical Implementation

LangGraph Pipeline: Architecture and Integration

- **Pipeline Architecture:**

- Linear processing pipeline with four main nodes:
 - Video processing
 - Audio processing
 - Data integration
 - Report generation
- Simple Directed Acyclic Graph (DAG) with sequential flow and error-handling edges.
- Supports both synchronous and asynchronous execution using `asyncio`.

- **Integration with Python Ecosystem:**

- Integrates with `asyncio` for non-blocking execution.
- Compatible with:
 - MediaPipe (vision)
 - Google Web Speech API (audio)
 - Gemini API (report generation)

Key Components: Vision and Audio Nodes

- **Video Processing Node:**
 - **Primary Function:** Analyzes video frames for player pose
 - **Technical Implementation:** MediaPipe Pose model for human pose detection and tracking.
 - **Key Capabilities:** Pose estimation, elbow angle calculation, wrist distance measurement.
 - **Output:** Structured JSON with timestamped keypoints and performance metrics.
- **Audio Processing Node:**
 - **Primary Function:** Extracts and transcribes speech from the video's audio track.
 - **Technical Implementation:** PyDub for audio extraction; Google Web Speech API for transcription.
 - **Key Capabilities:** Multi-language support, silence-based segmentation.
 - **Output:** Transcribed text of spoken content.

Report Generation: Technical Implementation

- **Data Aggregation:**
 - **Input Sources:** Combines pose metrics from video analysis and speech transcription from audio.
 - **Data Sampling:** Processes first 100 pose metrics to manage context size for LLM processing.
 - **JSON Formatting:** Structures data in standardized format for AI model consumption.
- **Natural Language Generation:**
 - **AI Model:** Google's Gemini 1.5 Flash model generates contextual, role-specific reports.
 - **Role-Based Prompting:** Custom system prompts tailored to coach, student, or parent perspectives.
 - **Personalization:** Adapts language, technical depth, and focus areas based on target audience.
- **Multilingual Support:**
 - **Language Options:** Reports available in multiple languages (English, Hindi, Tamil, Telugu, Kannada).

Novelty and Innovation

Multimodal Data Integration:

- Effective integration of video analysis (pose metrics) and audio transcription in a unified pipeline.
- Sequential processing with comprehensive error handling for reliable results.

LangGraph for Sports Analysis:

- Pioneering the use of LangGraph for structured pipeline orchestration in sports analytics.
- Foundation for future expansion to more complex agent interactions and workflows.

Actionable, Granular Feedback:

- Focus on generating highly specific, actionable feedback for players and coaches.
- Moves beyond descriptive statistics to prescriptive recommendations.

Scalable & Extensible Architecture:

- Modular pipeline design allows for easy addition of new analysis capabilities in future iterations.

System Implementation: Technical Stack

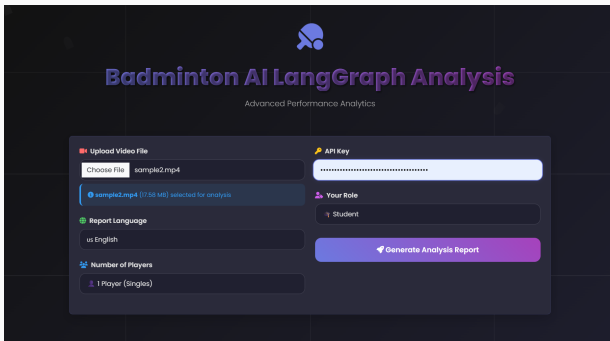
Frontend Technologies:

- **Web Interface:** HTML5, CSS3, JavaScript with responsive design.
- **Video Upload:** Custom file uploader with format validation and progress tracking.
- **Report Viewer:** Interactive PDF viewer with annotation capabilities.

Backend Technologies:

- **Server:** Flask for web application serving and file handling.
- **Video Processing:** OpenCV, MediaPipe for pose estimation and tracking.
- **Audio Processing:** Google Web Speech API for audio transcription.
- **Pipeline Orchestration:** LangGraph for linear pipeline definition and state management.
- **LLM Integration:** Google Gemini API for report generation with custom prompt templates.
- **Report Generation:** Custom templates for role-based and multilingual reports.

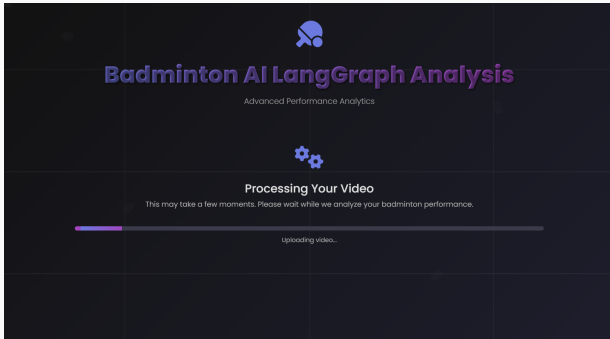
Demo Step 1: Upload Video



The screenshot displays the 'Badminton AI LangGraph Analysis' web interface. At the top, there is a logo of a badminton racket and the title 'Badminton AI LangGraph Analysis' in a purple font, with the subtitle 'Advanced Performance Analytics' below it. The main content area is a dark grey panel with several input fields and a button. On the left, under the heading 'Upload Video File', there is a 'Choose File' button next to the text 'sample2.mp4'. Below this, a blue bar indicates 'sample2.mp4 (17.58 MB) selected for analysis'. Further down, the 'Report Language' is set to 'us English' and the 'Number of Players' is set to '1 Player (Singles)'. On the right side, there is an 'API Key' field with a yellow key icon and a masked input field. Below that, the 'Your Role' is set to 'Student'. At the bottom right, there is a large purple button labeled 'Generate Analysis Report'.

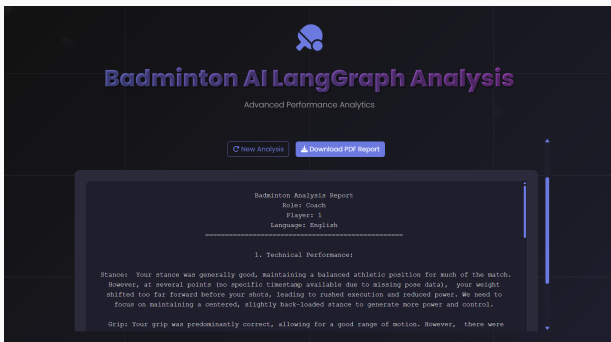
- User uploads a badminton match video and all necessary through the web interface.
- Supported formats include MP4, AVI, MOV.
- The system initiates pre-processing after receiving all information

Demo Step 2: Processing Status



- System displays processing stages with real-time progress indicators.
- Separate modules handle video frame extraction, pose detection, and audio transcription.
- Users are informed of each module's completion status.

Demo Step 3: Interactive Report



- An interactive, browser-based report is generated after processing.
- Key metrics like elbow angles, wrist distances, and speech insights are displayed.
- Visual overlays and summaries make interpretation intuitive.

Demo Step 4: Download PDF Report



- Users can explore specific sections in detail through the interface.
- A downloadable PDF report summarizes key findings for offline review.
- Report layout adapts to role — coach, player, or parent.

Demo Walkthrough: User Experience Summary

- **Complete Journey Overview:**
 - **Step 1: Upload** — Upload badminton match video via a simple interface.
 - **Step 2: Process** — Monitor real-time progress with clear indicators.
 - **Step 3: Report View** — See insights in a rich, interactive web view.
 - **Step 4: Export** — Download professional reports for long-term use.
- **Focus on Usability:** Designed for non-technical users (coaches, athletes).
- **Future Enhancement:** Mobile-friendly UI and real-time streaming support.

Thank You! Questions?