Post Big-Four Team Members: Xujia Qin, Hao (Daniel) Luo

Problem Description

This project focuses on predicting the **future movement trajectory of tennis players** using body pose and ball position data from match videos. It has real-world value in automating sports broadcasting, improving player performance analysis, and enabling AI-driven camera systems.

The motivation comes from gaps in current tracking algorithms and the increasing availability of structured tennis video data, making it a timely chance to apply deep learning and computer vision for smarter sports analytics. Our motivation also comes from our personal interest in tennis and a desire to apply computer vision and deep learning to better understand the game.

Background and Related Work (Review of 5 Research Papers)

1. **AlShami et al. [2024]** propose a solution to this gap by leveraging Transformer-based encoder-decoder models that incorporate player body joints, past centroid positions, and ball location to predict the future trajectory of players. Their approach introduces a more holistic understanding of motion and timing, offering greater accuracy and utility for applications such as automated sports broadcasting and player behavior analysis.
2. **Fernando et al. [2019]** introduce a Memory-augmented Semi-Supervised GAN (MSS-GAN) that leverages neuroscience-inspired episodic and semantic memory modules to model tennis player behavior, achieving superior shot prediction accuracy on 2012 Australian Open data and enabling opponent strategy analysis through hierarchical feature learning.
3. **Wei et al. [2013]** present a framework using Hawk-Eye spatiotemporal data to predict tennis shot locations. Their Dynamic Bayesian Network models player behavior by analyzing movement, shot speed, and positioning, predicting court regions and impact points with a Gaussian Mixture Model. Adaptive pre-game and online methods enhance accuracy across opponents. Tested on top players, it achieves about 1.7-meter error and strong AUC scores, aiding coaching, broadcasting, and strategy.
4. **In the work by Kienzle et al. [2024]**, the authors developed a method to predict ball spin and 3D trajectory from standard table tennis broadcast videos. They trained a neural network exclusively on synthetic data with physically accurate simulations, using targeted augmentations to enable generalization to real footage. The approach successfully transferred from synthetic to real data without requiring real training samples, though accuracy was limited by motion blur and annotation errors.
5. **In the work by Xiao et al. [2024],** the authors created predictive models for ball trajectories in sports like ping pong where position, velocity, and spin data are often noisy. They created an end-to-end framework that jointly trains a dynamics model and factor graph estimator using Gram-Schmidt process for roto-translational invariant representations. Their approach outperformed data augmentation methods alone but required careful factor graph initialization for optimal real-world performance.