## A Survey of Deep Learning in Sports Applications: Perception, Comprehension, and Decision (Zhao et al., 2023)

### https://arxiv.org/pdf/2307.03353v1

### Key Concepts and Taxonomy

This paper introduces a comprehensive hierarchical framework that organizes deep learning tasks in sports into three conceptual levels: Perception, Comprehension, and Decision. These levels are intended to represent an increasing level of abstraction and complexity in how AI systems understand and interact with sports data.  
  
Perception focuses on low-level tasks that process raw visual or sensor data, including player and ball localization, tracking, player re-identification, human pose estimation, and camera calibration. These tasks are crucial for building structured representations of scenes in sports videos.  
  
Comprehension tasks analyze these representations to extract semantic meaning, such as recognizing player actions (individual and group), evaluating action quality, generating summaries of matches, and captioning scenes. These tasks bridge the gap between perception and strategic understanding.  
  
Decision tasks involve predictive reasoning and content generation, such as play forecasting, match evaluation, game simulation, and generating synthetic player movements or full match videos. These models aim to enhance strategic planning and training.  
  
The paper also provides a detailed mapping of sports datasets and virtual environments, categorizing them by sport type and applicable task, forming a rich foundation for future benchmarking and research.

### Main Contributions and Findings

This work's main contribution is its holistic view of the role deep learning plays across the full pipeline of sports understanding—from visual input to strategic decision-making. It establishes a taxonomy that systematically organizes tasks into perception, comprehension, and decision, serving as a conceptual and practical guide for researchers and practitioners.  
  
It surveys over 60 datasets spanning soccer, basketball, dance, figure skating, volleyball, and more, covering a wide range of tasks including pose estimation, re-identification, captioning, and action recognition. These datasets are mapped to specific deep learning tasks, making the paper a valuable resource for dataset selection and algorithm benchmarking.  
  
The survey also tracks emerging trends such as the use of transformer-based architectures, multi-agent reinforcement learning, and multi-modal learning that incorporates visual, audio, and IoT signals. It acknowledges the growing role of generative models in tasks like video synthesis and simulation.

### Limitations and Future Directions

Despite strong advancements, the paper identifies several challenges impeding real-world deployment and performance of deep learning models in sports:  
  
- Many tasks, especially those in perception, suffer from challenges like occlusion, identity ambiguity, and low resolution in real broadcast footage.  
- A lack of standardized, large-scale benchmarks for many sports limits model comparability and generalization.  
- The full potential of IoT and wearable data remains underutilized due to poor integration with computer vision models.  
  
Future directions include the development of foundation models for sports, akin to ChatGPT in NLP or SAM in vision. These models would offer general-purpose capabilities across multiple sports-related tasks, ideally trained on large, high-quality, multi-modal datasets. The authors also emphasize the need for user-centered applications that bring elite-level analytics to amateur or casual users, supporting broader health and fitness goals.

**REFERENCE PAPERS**

**[1] Chmait & Westerbeek, 2021 – Artificial Intelligence and Machine Learning in Sport Research: An Introduction for Non-Data Scientists**

**https://link.springer.com/referenceworkentry/10.1007/978-3-030-57321-9\_24-1**

**Key Concepts and Taxonomy:**  
Introduces core AI and machine learning concepts to sports researchers with limited technical background. It outlines typical workflows and applications of ML in sports contexts.

**Main Contributions and Findings:**  
Provides a foundational, practical guide for sports science professionals to engage with AI tools. It identifies potential applications including performance prediction, injury prevention, and strategic analysis.

**Limitations and Future Directions:**  
Limited depth on specific models or advanced methodologies. Future directions include integrating real-time data and making tools accessible to a broader range of sports practitioners.

**Relevance to Primary Paper:**  
Serves as an entry point that complements the more technical taxonomy presented in the survey. Encourages interdisciplinary collaboration critical to applied AI in sports.

**[2] SMT – smt.com**

<https://www.smt.com/>

**Key Concepts and Taxonomy:**  
SMT offers broadcast and tracking technologies for sports analytics, specializing in real-time data visualization and event tagging.

**Main Contributions and Findings:**  
Provides the infrastructure behind televised sports analytics, supporting tasks such as ball/player tracking and graphical overlays.

**Limitations and Future Directions:**  
Being a commercial source, its methods are proprietary. The future lies in deeper integration with AI-based insights and fan engagement technologies.

**Relevance to Primary Paper:**  
Supports the perception and visualization layers of the deep learning taxonomy by supplying the raw data streams needed for model training and deployment.

**[3] Vizrt – vizrt.com**

<https://www.vizrt.com/>

**Key Concepts and Taxonomy:**  
A leader in broadcast graphics and real-time sports data integration for media companies and leagues.

**Main Contributions and Findings:**  
Empowers broadcasters with automated graphic rendering based on sports tracking data, enabling real-time storytelling.

**Limitations and Future Directions:**  
Limited focus on AI beyond automation. Future opportunities lie in leveraging real-time AI models for prediction and augmented commentary.

**Relevance to Primary Paper:**  
Bridges data acquisition and real-time perception modules by transforming raw input into digestible visuals.

**[4] Duarte et al., 2021 – Artificial Intelligence in Sport Performance Analysis**

<https://www.google.com/search?q=https://www.routledge.com/Artificial-Intelligence-in-Sport-Performance-Analysis/Ribeiro-Clemente-Pra%C3%A7a-Kaya/p/book/9780367420233>

**Key Concepts and Taxonomy:**  
Explores the application of AI techniques—especially computer vision and statistical learning—in sport performance evaluation.

**Main Contributions and Findings:**  
Demonstrates how AI can enhance understanding of tactical behavior, motor skill acquisition, and match dynamics through automatic annotation and video mining.

**Limitations and Future Directions:**  
Mostly limited to post-game analysis; expanding into real-time feedback and broader sports types remains a future goal.

**Relevance to Primary Paper:**  
Directly supports the comprehension and decision categories by mapping AI tools to performance evaluation pipelines.

**[5] Cust et al., 2019 – Machine and Deep Learning for Sport-Specific Movement Recognition**

<https://www.google.com/search?q=https://link.springer.com/article/10.1007/s40279-019-01183-5>

**Key Concepts and Taxonomy:**  
Systematic review of movement recognition models using both traditional machine learning and deep learning in sports.

**Main Contributions and Findings:**  
Analyzes model development across various sports, focusing on accuracy, input modalities (e.g., IMU, video), and evaluation metrics.

**Limitations and Future Directions:**  
Challenges in generalizing models across individuals and sports; need for large, standardized datasets noted.

**Relevance to Primary Paper:**  
Informs the “perception” layer by providing critical insights into movement-based recognition models across sports domains.

**[6] Bonidia et al., 2018 – Computational Intelligence in Sports: A Systematic Literature Review**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/8489370>

**Key Concepts and Taxonomy:**  
Presents a broad review of computational intelligence applications including fuzzy systems, neural networks, and evolutionary algorithms in sports.

**Main Contributions and Findings:**  
Covers diverse areas such as team formation, tactics optimization, and performance forecasting.

**Limitations and Future Directions:**  
Less emphasis on deep learning; more on classical AI. Suggests deeper integration of evolving DL techniques.

**Relevance to Primary Paper:**  
Frames historical context and bridges traditional AI with the emerging deep learning frameworks surveyed in the main paper.

**[7] Beal et al., 2019 – Artificial Intelligence for Team Sports: A Survey**

<https://arxiv.org/abs/1912.00185>

**Key Concepts and Taxonomy:**  
Focuses on AI challenges in multi-agent team sports like coordination, communication, and tactical modeling.

**Main Contributions and Findings:**  
Discusses simulation, reinforcement learning, and agent-based modeling as tools to mimic team strategies.

**Limitations and Future Directions:**  
Lacks coverage on deep learning advances and real-time applications.

**Relevance to Primary Paper:**  
Supports the “decision” layer by detailing how AI can model and influence team-level behavior in competitive sports.

**[8] Tan et al., 2016 – A Review on Badminton Motion Analysis**

<https://www.google.com/search?q=https://www.researchgate.net/publication/308969421_A_Review_on_Badminton_Motion_Analysis_using_Inertial_Sensors_and_Video_Cameras>

**Key Concepts and Taxonomy:**  
Presents methods for analyzing badminton-specific motion using computer vision and motion sensors.

**Main Contributions and Findings:**  
Discusses techniques for swing classification, shuttle tracking, and player movement analysis.

**Limitations and Future Directions:**  
Mostly classical approaches; more real-time, AI-integrated methods are needed.

**Relevance to Primary Paper:**  
Reinforces perception and comprehension pipelines in racquet sports, offering a domain-specific angle.

**[9] Wu et al., 2022 – A Survey on Video Action Recognition in Sports**

<https://arxiv.org/abs/2201.07374>

**Key Concepts and Taxonomy:**  
Comprehensive survey of datasets, architectures (e.g., CNN, RNN, Transformers), and evaluation criteria for sports action recognition.

**Main Contributions and Findings:**  
Emphasizes fine-grained recognition in dynamic scenes and compares supervised, weakly-supervised, and unsupervised techniques.

**Limitations and Future Directions:**  
Calls for better spatiotemporal modeling and multi-modal integration.

**Relevance to Primary Paper:**  
Core reference for the comprehension tier, especially action recognition and sports video understanding.

**[10] Wang et al., 2021 – A Survey of Video-Based Action Quality Assessment**

<https://arxiv.org/abs/2104.09228>

**Key Concepts and Taxonomy:**  
Surveys methods for scoring the quality of actions in sports using video input, often applied to diving, gymnastics, and figure skating.

**Main Contributions and Findings:**  
Explores regression-based, ranking-based, and hybrid models, alongside custom loss functions and attention mechanisms.

**Limitations and Future Directions:**  
Limited dataset size, difficulty modeling subjectivity in scoring.

**Relevance to Primary Paper:**  
Directly supports the action quality assessment (AQA) component under comprehension, particularly for judged sports.

**[11] Kamble et al., 2019 – Ball Tracking in Sports: A Survey**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/8718041>

**Key Concepts and Taxonomy:**  
Surveys ball tracking methods in sports using computer vision, highlighting both single-view and multi-view approaches for 2D and 3D tracking.

**Main Contributions and Findings:**  
Covers detection methods, motion modeling, and post-processing for accurate tracking. Reviews sports like soccer, basketball, and tennis. It identifies the importance of physical constraints and motion models for improving accuracy.

**Limitations and Future Directions:**  
Challenges include occlusion, motion blur, and varied backgrounds. More robust deep learning-based models with domain knowledge are needed.

**Relevance to Primary Paper:**  
Directly supports the “Perception” layer, particularly ball localization and tracking—core to building reliable analytics pipelines in sports.

**[12] Adesida et al., 2019 – Exploring the Role of Wearable Technology in Sport Kinematics and Kinetics**

<https://www.mdpi.com/1424-8220/19/7/1597>

**Key Concepts and Taxonomy:**  
Systematic review of how wearable devices measure biomechanical parameters like joint angles and ground reaction forces.

**Main Contributions and Findings:**  
Highlights the use of IMUs and force sensors in sports like running and football for injury prevention, training feedback, and performance profiling.

**Limitations and Future Directions:**  
Data sparsity and calibration remain major issues. Future work calls for standardization and integration with video data.

**Relevance to Primary Paper:**  
Extends the scope of the “Perception” layer by advocating for richer multi-modal inputs (e.g., wearables + vision).

**[13] Rana & Mittal, 2020 – Wearable Sensors for Real-Time Kinematics Analysis in Sports**

<https://www.mdpi.com/1424-8220/20/15/4144>

**Key Concepts and Taxonomy:**  
A technical overview of real-time kinematic analysis using wearable IMUs in athletic monitoring.

**Main Contributions and Findings:**  
Summarizes sensor architectures, feature extraction, real-time data processing, and challenges across sports.

**Limitations and Future Directions:**  
Calls for improved real-time inference, better wireless sync, and more datasets for benchmarking.

**Relevance to Primary Paper:**  
Enriches the data modalities feeding into deep learning pipelines, especially for pose estimation and action recognition.

**[14] Van der Kruk & Reijne, 2018 – Accuracy of Human Motion Capture Systems for Sport Applications**

<https://www.mdpi.com/1424-8220/18/10/3336>

**Key Concepts and Taxonomy:**  
Compares motion capture (MoCap) systems—optical, inertial, and markerless—for sports.

**Main Contributions and Findings:**  
Evaluates accuracy, latency, and cost-efficiency of different systems for sports like gymnastics and soccer.

**Limitations and Future Directions:**  
Markerless systems are promising but still suffer from occlusion and poor generalization.

**Relevance to Primary Paper:**  
Underpins the pose estimation task in the perception category, especially where reliable human joint data is required.

**[15] Turing, 2009 – Computing Machinery and Intelligence**

<https://academic.oup.com/mind/article/LIX/236/433/986238>

**Key Concepts and Taxonomy:**  
Seminal philosophical work introducing the concept of machine intelligence through the “Turing Test.”

**Main Contributions and Findings:**  
Raises foundational questions about AI’s capacity to simulate human cognition.

**Limitations and Future Directions:**  
Philosophical in scope—non-specific to sports or deep learning.

**Relevance to Primary Paper:**  
Provides a foundational theoretical backdrop for the AI techniques applied throughout the paper.

**[16] Wang et al., 2019 – AI Coach: Deep Human Pose Estimation and Analysis for Personalized Athletic Training**

<https://arxiv.org/abs/1908.03613>

**Key Concepts and Taxonomy:**  
Proposes a system for personalized training using deep pose estimation and movement analysis.

**Main Contributions and Findings:**  
Integrates pose estimation with activity scoring and feedback to assist amateur athletes during training.

**Limitations and Future Directions:**  
Limited generalizability across sports. Future work could improve domain transfer and real-time capabilities.

**Relevance to Primary Paper:**  
Embodies the connection between perception and decision—pose estimation used directly for action quality assessment and feedback.

**[17] Rao & Pati, 2015 – A Novel Algorithm for Detection of Soccer Ball and Player**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/7377759>

**Key Concepts and Taxonomy:**  
Classical CV approach to detect players and soccer balls using Hough transforms and color filtering.

**Main Contributions and Findings:**  
Focuses on computationally inexpensive methods for ball and player detection in constrained environments.

**Limitations and Future Directions:**  
Not robust to occlusion and dynamic backgrounds. Calls for more advanced learning-based methods.

**Relevance to Primary Paper:**  
Represents early methods in the “Perception” category, providing a baseline for comparing with deep learning approaches.

**[18] Yang et al., 2018 – 3D Multiview Basketball Players Detection and Localization**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/8338302>

**Key Concepts and Taxonomy:**  
Multi-view 3D detection system for player localization in basketball using occupancy maps.

**Main Contributions and Findings:**  
Improves detection accuracy under occlusion by fusing multiple viewpoints.

**Limitations and Future Directions:**  
Expensive setup; limited scalability beyond controlled settings.

**Relevance to Primary Paper:**  
Aligns with advanced techniques in multi-camera perception for accurate player tracking and positioning.

**[19] Şah & Direkoğlu, 2019 – Evaluation of Image Representations for Player Detection in Field Sports Using CNNs**

<https://www.google.com/search?q=https://dl.acm.org/doi/10.1145/3341105.3341144>

**Key Concepts and Taxonomy:**  
Evaluates different input preprocessing and feature extraction techniques for CNN-based player detection.

**Main Contributions and Findings:**  
Finds that certain color normalization and contrast enhancements improve performance under variable lighting.

**Limitations and Future Directions:**  
Limited to detection—future work should link with re-ID and tracking.

**Relevance to Primary Paper:**  
Supports perception-related tasks by optimizing CNN inputs for robust player localization.

**[20] Gerke et al., 2017 – Soccer Player Recognition Using Spatial Constellation Features and Jersey Number Recognition**

<https://www.google.com/search?q=https://www.researchgate.net/publication/319973215_Soccer_Player_Recognition_Using_Spatial_Constellation_Features_and_Jersey_Number_Recognition>

**Key Concepts and Taxonomy:**  
Combines spatial constellation features with CNN-based digit recognition to identify players by jersey numbers.

**Main Contributions and Findings:**  
Improves recognition accuracy under partial occlusion and variable lighting.

**Limitations and Future Directions:**  
Digit segmentation remains brittle. Future integration with pose and temporal features is suggested.

**Relevance to Primary Paper:**  
Directly supports player identification (perception layer), critical for downstream analytics like re-ID and statistics

**[21] Li et al., 2018 – Jersey Number Recognition with Semi-Supervised Spatial Transformer Network**

<https://arxiv.org/abs/1802.04264>

**Key Concepts and Taxonomy:**  
Applies a semi-supervised approach with spatial transformer networks (STN) to recognize jersey numbers from sports footage.

**Main Contributions and Findings:**  
Combines a STN module with classification networks to correct perspective distortions and improve jersey digit classification with limited labels.

**Limitations and Future Directions:**  
Performance drops under heavy occlusion and small-scale digits. More advanced attention-based architectures may improve generalization.

**Relevance to Primary Paper:**  
Supports the “Perception” layer by enhancing identity recognition, a critical step in downstream player tracking and statistics systems.

**[22] Liu & Bhanu, 2019 – Pose-Guided R-CNN for Jersey Number Recognition**

<https://arxiv.org/abs/1904.06208>

**Key Concepts and Taxonomy:**  
Proposes a pose-guided region-based CNN (R-CNN) architecture to focus digit detection on jersey areas using keypoint cues.

**Main Contributions and Findings:**  
Outperforms generic object detectors in jersey number localization and recognition by exploiting pose landmarks.

**Limitations and Future Directions:**  
Still sensitive to lighting changes and fast motion blur. Future works could integrate temporal consistency or multi-view cues.

**Relevance to Primary Paper:**  
Strengthens the identity module in perception tasks, crucial for player tracking and tactical analysis.

**[23] Istasse et al., 2019 – Associative Embedding for Team Discrimination**

<https://www.google.com/search?q=https://orbi.uliege.be/handle/2268/240588>

**Key Concepts and Taxonomy:**  
Introduces associative embedding to group players into teams based on color and spatial similarity, without labeled identity data.

**Main Contributions and Findings:**  
Enables unsupervised team classification using simple visual embeddings, which can be refined by downstream clustering.

**Limitations and Future Directions:**  
Limited by jersey similarity across teams and occlusion. Calls for combining with motion or contextual data.

**Relevance to Primary Paper:**  
Serves the team recognition aspect of “Perception,” often a prerequisite for higher-order tasks like strategy modeling.

**[24] Koshkina et al., 2021 – Contrastive Learning for Sports Video: Unsupervised Player Classification**

<https://arxiv.org/abs/2104.02058>

**Key Concepts and Taxonomy:**  
Uses contrastive learning to learn embeddings for unsupervised player discrimination in amateur sports.

**Main Contributions and Findings:**  
Leverages temporal consistency and multi-view agreement to improve unsupervised classification of players and teams.

**Limitations and Future Directions:**  
Requires well-structured temporal data; sensitive to annotation drift.

**Relevance to Primary Paper:**  
Fits into the “Perception” and “Comprehension” categories as a low-label cost approach for player representation learning.

**[25] Manafifard et al., 2017 – A Survey on Player Tracking in Soccer Videos**

<https://www.google.com/search?q=https://www.researchgate.net/publication/319973059_A_Survey_on_Player_Tracking_in_Soccer_Videos>

**Key Concepts and Taxonomy:**  
Comprehensive survey covering player tracking approaches, including motion prediction, occlusion handling, and re-identification.

**Main Contributions and Findings:**  
Reviews key methods and highlights multi-object tracking and data association as core challenges in soccer analytics.

**Limitations and Future Directions:**  
Limited availability of labeled tracking datasets; multi-view integration needed.

**Relevance to Primary Paper:**  
Forms the theoretical and practical basis for “Tracking” in the perception module.

**[26] Theagarajan et al., 2018 – Soccer: Who Has the Ball?**

<https://arxiv.org/abs/1807.03556>

**Key Concepts and Taxonomy:**  
Introduces a visual analytics tool for identifying ball possession using CNNs and spatio-temporal player features.

**Main Contributions and Findings:**  
Merges object detection and scene understanding to infer which player controls the ball in crowded soccer scenes.

**Limitations and Future Directions:**  
Accuracy depends heavily on precise player localization; improvements in tracking can boost performance.

**Relevance to Primary Paper:**  
Contributes to the “Decision” layer by integrating perception (player and ball detection) with tactical inference.

**[27] Arbues-Sanguesa et al., 2020 – Using Player’s Body Orientation to Model Pass Feasibility in Soccer**

<https://arxiv.org/abs/1910.12658>

**Key Concepts and Taxonomy:**  
Exploits body pose and orientation to estimate likely passing directions and feasible tactical plays.

**Main Contributions and Findings:**  
Improves prediction models by integrating pose cues, outperforming location-only models in pass prediction.

**Limitations and Future Directions:**  
Assumes accurate pose estimation, which can be error-prone in complex scenes.

**Relevance to Primary Paper:**  
Sits at the intersection of comprehension and decision—merging body posture with strategy modeling.

**[28] Ren et al., 2015 – Faster R-CNN: Towards Real-Time Object Detection**

<https://arxiv.org/abs/1506.01497>

**Key Concepts and Taxonomy:**  
Seminal object detection model combining region proposal networks (RPNs) with CNN-based classifiers.

**Main Contributions and Findings:**  
Drastically improves object detection speed and accuracy, enabling real-time applications in sports.

**Limitations and Future Directions:**  
Struggles with small, fast objects like balls; newer models (e.g., YOLOv5, DETR) build on this foundation.

**Relevance to Primary Paper:**  
Backbone model for many sports perception systems involving player and ball detection.

**[29] Cioppa et al., 2019 – Arthus: Adaptive Real-Time Human Segmentation in Sports**

<https://www.google.com/search?q=https://orbi.uliege.be/handle/2268/242250>

**Key Concepts and Taxonomy:**  
Presents a distillation-based method for real-time player segmentation using limited annotations.

**Main Contributions and Findings:**  
Transfers knowledge from pre-trained segmentation models to low-resource sports domains.

**Limitations and Future Directions:**  
Fails under extreme occlusion; needs better generalization to unseen player appearances.

**Relevance to Primary Paper:**  
Targets “Segmentation” in perception tasks, essential for downstream pose, action, and re-ID tasks.

**[30] Vandeghen et al., 2022 – Semi-Supervised Training for Player and Ball Detection in Soccer**

<https://www.google.com/search?q=https://arxiv.org/abs/2208.11894>

**Key Concepts and Taxonomy:**  
Proposes a distillation framework for semi-supervised player and ball detection using limited labeled frames.

**Main Contributions and Findings:**  
Reduces annotation burden while maintaining accuracy, crucial for scaling to amateur sports or niche domains.

**Limitations and Future Directions:**  
Still depends on quality of teacher models; improvement possible via domain adaptation.

**Relevance to Primary Paper:**  
Provides a practical approach to tackle data scarcity in the “Perception” layer.

**[31] Sanford et al., 2020 – Group Activity Detection from Trajectory and Video Data in Soccer**

<https://arxiv.org/abs/2007.03264>

**Key Concepts and Taxonomy:**  
Combines player trajectories and video features to classify high-level group activities such as counterattacks or set plays.

**Main Contributions and Findings:**  
Integrates motion features and deep video embeddings for spatio-temporal analysis of team behavior.

**Limitations and Future Directions:**  
Sensitive to tracking quality; expansion to broader tactical contexts needed.

**Relevance to Primary Paper:**  
Fits within the “Comprehension” category—bridging individual tracking (perception) to strategic reasoning (decision).

**[32] Cioppa et al., 2021 – Camera Calibration and Player Localization in SoccerNet-V2**

<https://arxiv.org/abs/2105.04825>

**Key Concepts and Taxonomy:**  
Joint model for estimating camera parameters and localizing players in 3D space using broadcast video.

**Main Contributions and Findings:**  
Introduces a unified architecture that improves both localization and calibration, enabling accurate spatial grounding.

**Limitations and Future Directions:**  
Performance degrades under poor visibility or camera switches.

**Relevance to Primary Paper:**  
Supports both perception (localization) and downstream comprehension tasks like action spotting.

**[33] Cioppa et al., 2022 – SoccerNet-Tracking Dataset**

<https://www.soccer-net.org/tasks/tracking>

<https://arxiv.org/abs/2205.14389>

**Key Concepts and Taxonomy:**  
A large-scale benchmark dataset for multi-object tracking in soccer, providing player IDs across time.

**Main Contributions and Findings:**  
Enables robust evaluation of tracking models, supporting re-identification and trajectory prediction.

**Limitations and Future Directions:**  
Focuses on soccer; extending to other sports would broaden applicability.

**Relevance to Primary Paper:**  
Directly supports the tracking module under perception, critical for temporal modeling.

**[34] Cioppa et al., 2022 – SoccerNet-V3 Dataset**

<https://www.google.com/search?q=https://openaccess.thecvf.com/content/CVPR2022W/CVSports/html/Cioppa_Scaling_Up_SoccerNet_With_Multi-View_Spatial_Localization_and_Re-Identification_CVPRW_2022_paper.html>

**Key Concepts and Taxonomy:**  
Extends SoccerNet with annotations for multi-view correspondence, player ID, and 3D localization.

**Main Contributions and Findings:**  
A comprehensive dataset enabling multi-task learning across localization, re-ID, and action recognition.

**Limitations and Future Directions:**  
Heavy focus on soccer; lacks equivalent datasets for sports like volleyball or hockey.

**Relevance to Primary Paper:**  
Supports all three layers—perception, comprehension, and decision—via rich annotations and multi-view alignment.

**[35] Uchida et al., 2021 – Automated Offside Detection in Soccer**

<https://arxiv.org/abs/2111.01878>

**Key Concepts and Taxonomy:**  
Proposes an algorithm to detect offside situations using player and ball localization along with spatio-temporal rules.

**Main Contributions and Findings:**  
Demonstrates how structured video analysis can support rule enforcement.

**Limitations and Future Directions:**  
Dependent on precise calibration and accurate ball localization.

**Relevance to Primary Paper:**  
Connects perception-level data to decision-layer rule interpretation and event recognition.

**[36] Van Zandycke et al., 2022 – DeepSportradar-V1 Dataset for Basketball**

<https://arxiv.org/abs/2208.08190>

**Key Concepts and Taxonomy:**  
High-quality, multi-task basketball dataset including 3D localization, segmentation, and camera calibration.

**Main Contributions and Findings:**  
Facilitates research on simultaneous detection, calibration, and scene understanding.

**Limitations and Future Directions:**  
Currently focused on basketball; expansion to other domains needed.

**Relevance to Primary Paper:**  
Essential for multi-task learning across the perception layer in basketball AI applications.

**[37] Voeikov et al., 2020 – TTNet: Real-Time Table Tennis Video Analysis**

<https://arxiv.org/abs/2002.05389>

**Key Concepts and Taxonomy:**  
End-to-end deep learning model for event spotting in table tennis using ball and player trajectories.

**Main Contributions and Findings:**  
Detects strokes, rallies, and outcomes with high accuracy in real time.

**Limitations and Future Directions:**  
Limited to clean, well-structured scenes; expansion to broadcast-level footage needed.

**Relevance to Primary Paper:**  
Supports comprehension and decision tasks in racquet sports through real-time temporal localization.

**[38] Maksai et al., 2016 – Physically Constrained Interaction Modeling in Ball Sports**

<https://www.google.com/search?q=https://www.cv-foundation.org/openaccess/content_cvpr_2016_workshops/w23/papers/Maksai_Physically_Constrained_Interaction_CVPR_2016_paper.pdf>

**Key Concepts and Taxonomy:**  
Models player-ball interaction using physical constraints and probabilistic reasoning.

**Main Contributions and Findings:**  
Predicts future ball positions and player intentions based on physics-informed priors.

**Limitations and Future Directions:**  
Doesn’t scale well to multi-view or noisy real-world data.

**Relevance to Primary Paper:**  
Contributes to decision-layer tasks by integrating physics with deep prediction models.

**[39] Cheng et al., 2017 – Ball State Estimation in Volleyball**

<https://ieeexplore.ieee.org/document/8296574>

**Key Concepts and Taxonomy:**  
Estimates both physical and conceptual state of a volleyball (e.g., possession, pass type) using vision-based systems.

**Main Contributions and Findings:**  
Combines object detection with rule-based logic to estimate semantic state.

**Limitations and Future Directions:**  
Struggles with ambiguous frames and partial occlusion.

**Relevance to Primary Paper:**  
Sits at the edge of perception and comprehension, enriching raw detections with semantic context.

**[40] Parisot & De Vleeschouwer, 2019 – Consensus-Based Trajectory Estimation for Ball Tracking**

<https://ieeexplore.ieee.org/document/8693444>

**Key Concepts and Taxonomy:**  
Uses consensus models over calibrated multi-camera systems to reconstruct accurate 3D ball trajectories.

**Main Contributions and Findings:**  
Improves robustness to occlusion and noisy detections in multi-view sports environments.

**Limitations and Future Directions:**  
High reliance on precise camera calibration.

**Relevance to Primary Paper:**  
Supports robust 3D ball tracking, key for understanding gameplay events and player interactions.

**[41] Sköld, 2015 – Estimating 3D Trajectories from Monocular Video Sequences**

<https://www.google.com/search?q=https://www.diva-portal.org/smash/get/diva2:827011/FULLTEXT01.pdf>

**Key Concepts and Taxonomy:**  
Proposes methods for reconstructing 3D trajectories of objects, such as balls, from single-view video input.

**Main Contributions and Findings:**  
Utilizes motion constraints and camera parameters to infer the 3D motion path of sports objects like tennis balls and basketballs.

**Limitations and Future Directions:**  
Struggles with occlusion and requires calibrated camera assumptions.

**Relevance to Primary Paper:**  
Supports perception and decision layers by reconstructing ball dynamics for further event recognition or rule enforcement.

**[42] Van Zandycke & De Vleeschouwer, 2022 – 3D Ball Localization from a Single Calibrated Image**

<https://arxiv.org/abs/2203.17243>

**Key Concepts and Taxonomy:**  
Enhances 3D ball localization using calibrated single-view images, focusing on size estimation and visual cues.

**Main Contributions and Findings:**  
Bridges the gap between traditional ballistic models and learned visual inference for estimating ball location and depth.

**Limitations and Future Directions:**  
Still vulnerable under heavy occlusion; more robust integration with tracking models is recommended.

**Relevance to Primary Paper:**  
Vital to perception tasks involving 3D ball localization, useful in event detection like offside calls or scoring validation.

**[43] Liu & Wang, 2022 – MonoTrack: Shuttle Trajectory Reconstruction from Monocular Badminton Video**

<https://arxiv.org/abs/2203.10268>

**Key Concepts and Taxonomy:**  
Presents a deep learning method for reconstructing 3D shuttle trajectories from single-view badminton videos.

**Main Contributions and Findings:**  
Introduces spatiotemporal modeling for fast-moving and tiny objects, leveraging recurrent inference and motion constraints.

**Limitations and Future Directions:**  
Requires well-labeled datasets; limited generalization to other sports without domain-specific training.

**Relevance to Primary Paper:**  
Strengthens the “Ball Tracking” component in perception, especially in racquet sports with unique flight dynamics.

**[44] Naik et al., 2022 – DeepPlayerTrack: Tracking Players and Referees Using Jersey Color Recognition**

<https://www.google.com/search?q=https://link.springer.com/chapter/10.1007/978-981-19-1525-5_53>

**Key Concepts and Taxonomy:**  
Proposes jersey color-based ID management in a deep tracking system to maintain consistent identities for players and referees.

**Main Contributions and Findings:**  
Combines color-masking, appearance-based cues, and tracking-by-detection for robust temporal linking.

**Limitations and Future Directions:**  
Struggles when jerseys are similar in color or when lighting conditions change.

**Relevance to Primary Paper:**  
Improves player tracking in the perception module, helping maintain accurate identity tracking over long sequences.

**[45] Naik & Hashmi, 2023 – YOLOv3-SORT for Detection and Tracking in Soccer**

<https://ieeexplore.ieee.org/document/10126487>

**Key Concepts and Taxonomy:**  
Applies YOLOv3 with the SORT algorithm to jointly detect and track players and balls in soccer broadcasts.

**Main Contributions and Findings:**  
Lightweight pipeline that balances accuracy and real-time performance by combining fast detection with Kalman-based tracking.

**Limitations and Future Directions:**  
Fails in cluttered scenes or during occlusion-heavy events like corners.

**Relevance to Primary Paper:**  
Forms a baseline perception model for tracking in soccer analytics systems.

**[46] Bewley et al., 2016 – SORT: Simple Online and Real-Time Tracking**

<https://arxiv.org/abs/1602.00763>

<https://github.com/abewley/sort>

**Key Concepts and Taxonomy:**  
Introduces the SORT framework—a simple, real-time object tracker using Kalman filtering and Hungarian matching.

**Main Contributions and Findings:**  
Offers high-speed tracking for multi-object scenarios without learning-based modules.

**Limitations and Future Directions:**  
Identity switch and occlusion remain challenging; commonly enhanced with deep re-ID modules (e.g., DeepSORT).

**Relevance to Primary Paper:**  
Foundational for many tracking modules under perception, especially in real-time applications.

**[47] Hurault et al., 2020 – Self-Supervised Small Soccer Player Detection and Tracking**

<https://arxiv.org/abs/2012.00559>

**Key Concepts and Taxonomy:**  
Leverages self-supervised pretraining to improve detection and tracking of small players in low-resolution soccer videos.

**Main Contributions and Findings:**  
Fine-tunes generic object detectors on soccer-specific datasets without full annotations.

**Limitations and Future Directions:**  
Depends on the quality of pseudo-labels; more sophisticated augmentation and label refinement could enhance performance.

**Relevance to Primary Paper:**  
Supports perception-layer tracking in low-data or low-resolution settings, extending usability of AI in amateur sports.

**[48] Zhang et al., 2020 – Deep Player Identification in Multi-Camera Sports Videos**

<https://www.google.com/search?q=https://www.researchgate.net/publication/348270830_Deep_Player_Identification_in_Multi-Camera_Sports_Videos>

**Key Concepts and Taxonomy:**  
Uses deep embeddings and temporal consistency across cameras for robust player re-identification.

**Main Contributions and Findings:**  
Combines facial cues, pose estimation, and team information across views to identify players over time.

**Limitations and Future Directions:**  
Limited scalability to large teams or unknown identities.

**Relevance to Primary Paper:**  
Supports re-identification in the perception layer, critical for tactical modeling and statistics.

**[49] Sun et al., 2021 – DanceTrack: Multi-Object Tracking in Uniform Appearance Sports**

<https://arxiv.org/abs/2111.13532>

<https://github.com/DanceTrack/DanceTrack>

**Key Concepts and Taxonomy:**  
Introduces DanceTrack dataset and benchmark for tracking individuals with similar appearance and dynamic motion (e.g., dance, synchronized sports).

**Main Contributions and Findings:**  
Stresses reliance on motion cues and skeleton-based features rather than appearance.

**Limitations and Future Directions:**  
Pose estimation noise can affect tracking; real-world sports scenes may introduce new challenges.

**Relevance to Primary Paper:**  
Demonstrates the value of motion and pose in tracking tasks where visual identity is weak—relevant for volleyball, dance, and gymnastics.

**[50] Burić et al., 2019 – Player Tracking in Sports Videos**

<https://www.google.com/search?q=https://www.researchgate.net/publication/334547926_Player_tracking_in_sports_videos_A_survey_and_a_new_benchmark>

**Key Concepts and Taxonomy:**  
Surveys and benchmarks traditional and deep learning methods for player tracking in field sports.

**Main Contributions and Findings:**  
Highlights key challenges like occlusion, re-ID, camera switches, and the importance of spatial consistency.

**Limitations and Future Directions:**  
Calls for more unified datasets and better identity management strategies.

**Relevance to Primary Paper:**  
Serves as an early foundational review for perception-layer tracking and its integration with higher-order tasks.

**[51] Wang et al., 2022 – Multi-Level Temporal Transformer for Action Quality Assessment**

<https://arxiv.org/abs/2202.04690>

**Key Concepts and Taxonomy:**  
Proposes a hierarchical Transformer-based model to capture short- and long-term dependencies in action sequences.

**Main Contributions and Findings:**  
Introduces multi-scale temporal attention for better temporal resolution in AQA tasks like diving and gymnastics.

**Limitations and Future Directions:**  
Sensitive to noisy video segments; lacks support for multi-modal inputs (e.g., pose, commentary).

**Relevance to Primary Paper:**  
Advances comprehension tasks, particularly Action Quality Assessment (AQA), enhancing score prediction precision.

**[52] Tang et al., 2020 – Uncertainty-Aware Score Distribution Learning for AQA**

<https://arxiv.org/abs/2003.10098>

**Key Concepts and Taxonomy:**  
Introduces score distribution modeling with uncertainty estimation in regression-based AQA.

**Main Contributions and Findings:**  
Models action scores as distributions rather than point estimates, improving robustness to subjective labels.

**Limitations and Future Directions:**  
Limited to diving tasks; lacks generalization across broader sports.

**Relevance to Primary Paper:**  
Improves the reliability of deep regression techniques in evaluating athletic performance.

**[53] Zheng et al., 2021 – PoseFormer: Transformer for 3D Human Pose Estimation**

<https://github.com/zczcwh/PoseFormer>

<https://arxiv.org/abs/2103.15392>

**Key Concepts and Taxonomy:**  
A pure transformer architecture that estimates 3D human poses from 2D sequences.

**Main Contributions and Findings:**  
Outperforms CNN-based methods by modeling long-range joint dependencies without convolutions.

**Limitations and Future Directions:**  
Training data intensive; performance may degrade in crowded sports scenes.

**Relevance to Primary Paper:**  
Supports the perception layer for sports involving detailed motion tracking (e.g., gymnastics, diving).

**[54] Fang et al., 2022 – Learning from Synthetic Data for Pose Estimation in Sports**

<https://arxiv.org/abs/2203.16781>

**Key Concepts and Taxonomy:**  
Applies domain adaptation from synthetic datasets to improve pose estimation in sports videos.

**Main Contributions and Findings:**  
Uses transfer learning to reduce the need for annotated sports data.

**Limitations and Future Directions:**  
Bridging the reality gap between synthetic and real data remains a challenge.

**Relevance to Primary Paper:**  
Provides a practical solution to data scarcity in training sports-specific pose estimation models.

**[55] Wang et al., 2021 – Multi-View Action Quality Estimation for Synchronized Sports**

<https://arxiv.org/abs/2108.06205>

**Key Concepts and Taxonomy:**  
Leverages synchronized multi-camera inputs for fine-grained AQA tasks like synchronized swimming and dance.

**Main Contributions and Findings:**  
Fuses multi-view pose and visual features using attention-based architectures for improved scoring.

**Limitations and Future Directions:**  
High infrastructure cost limits general applicability.

**Relevance to Primary Paper:**  
Extends comprehension in team-based aesthetics-focused sports via advanced AQA frameworks.

**[56] Xu et al., 2022 – FineDiving Dataset for Fine-Grained AQA in Diving**

<https://www.google.com/search?q=https://github.com/xujing-ps/FineDiving>

<https://arxiv.org/abs/2203.09744>

**Key Concepts and Taxonomy:**  
Provides 3,000 samples of diving videos annotated with sub-action labels, difficulty, and performance scores.

**Main Contributions and Findings:**  
First dataset offering rich annotation of fine-grained motion phases in diving (e.g., takeoff, flight, entry).

**Limitations and Future Directions:**  
Focuses solely on diving; similar datasets needed for other judged sports.

**Relevance to Primary Paper:**  
Foundation for fine-grained AQA and supervised learning in Olympic-style sports.

**[57] Qi et al., 2022 – FSD-10: A Dataset for Fine-Grained Sports Action Recognition in Figure Skating**

<https://arxiv.org/abs/2207.13596>

**Key Concepts and Taxonomy:**  
Includes 10 fine-grained figure skating actions with detailed scores.

**Main Contributions and Findings:**  
Supports development of models for fine-grained recognition and scoring.

**Limitations and Future Directions:**  
Limited number of actions and samples; needs expansion.

**Relevance to Primary Paper:**  
Enables deeper action classification and assessment in figure skating.

**[58] Li et al., 2022 – SportsVQA: Video Question Answering Dataset for Sports**

<https://www.google.com/search?q=https://github.com/SportsVQA/SportsVQA-official>

<https://arxiv.org/abs/2205.10263>

**Key Concepts and Taxonomy:**  
Builds a video question-answering benchmark with sports-specific knowledge grounding.

**Main Contributions and Findings:**  
Includes questions on tactics, players, and results with multi-modal video inputs.

**Limitations and Future Directions:**  
Difficult to generalize to new sports domains; heavily reliant on annotations.

**Relevance to Primary Paper:**  
Supports high-level comprehension and reasoning, extending sports AI to educational and analytical tools.

**[59] Zhang et al., 2021 – Ball Action Spotting via Spatio-Temporal Graph Networks**

<https://arxiv.org/abs/2112.08332>

**Key Concepts and Taxonomy:**  
Applies graph convolutional networks (GCNs) to detect ball-centric events like passes, serves, or spikes.

**Main Contributions and Findings:**  
Models relationships between player actions and ball dynamics in volleyball and basketball.

**Limitations and Future Directions:**  
Accuracy affected by tracking quality; better temporal reasoning is needed.

**Relevance to Primary Paper:**  
Sits at the junction of perception and comprehension—critical for real-time analytics and replay segmentation.

**[60] Giancola et al., 2018 – Action Spotting in Soccer with Deep Learning**

<https://arxiv.org/abs/1807.05193>

**Key Concepts and Taxonomy:**  
Introduces a task and benchmark for spotting discrete events (e.g., goals, fouls) in untrimmed soccer broadcasts.

**Main Contributions and Findings:**  
Demonstrates early attempts at learning temporal anchors for precise action spotting.

**Limitations and Future Directions:**  
Anchor-based methods may miss subtle or overlapping events; newer Transformer-based spotting methods now outperform.

**Relevance to Primary Paper:**  
Foundational work in sports video temporal segmentation, supporting downstream comprehension tasks.

**[61] K. Habel et al., 2022 – Clip-ReIdent: Contrastive Training for Player Re-identification**

<https://arxiv.org/abs/2208.07472>

**Key Concepts and Taxonomy:**  
Proposes a player ReID system using contrastive loss and CLIP-based representations to distinguish players under challenging conditions.

**Main Contributions and Findings:**  
Focuses on class-agnostic learning, enabling transfer across datasets and sports by leveraging pretrained vision-language features.

**Limitations and Future Directions:**  
Still limited by clothing similarity and occlusion; deeper temporal consistency needed.

**Relevance to Primary Paper:**  
Supports the perception module by enabling more robust player tracking and identity matching across cameras and frames.

**[62] A. Maglo et al., 2022 – Semi-Interactive Transformer-Based ReID System**

<https://arxiv.org/abs/2209.00030>

**Key Concepts and Taxonomy:**  
Introduces a semi-supervised player re-identification framework using transformers with minimal manual annotation.

**Main Contributions and Findings:**  
Combines visual cues and user-provided interactions to build a lightweight, scalable ReID system.

**Limitations and Future Directions:**  
Depends on the quality and frequency of user feedback.

**Relevance to Primary Paper:**  
Directly ties into scalable perception systems for team sports analytics where full labels are impractical.

**[63] Vats et al., 2021 – Weakly Supervised Jersey Number Recognition in Ice Hockey**

<https://arxiv.org/abs/2111.13036>

**Key Concepts and Taxonomy:**  
Utilizes transformers with cross-entropy loss for jersey number classification under weak supervision.

**Main Contributions and Findings:**  
Demonstrates successful jersey digit classification with sparse annotations, even under occlusion and fast motion.

**Limitations and Future Directions:**  
Focused on hockey; generalization to more dynamic and occlusion-heavy sports remains limited.

**Relevance to Primary Paper:**  
Improves low-label cost perception tasks—vital for automated athlete identity systems in high-speed sports.

**[64] Yan et al., 2022 – Dual Data Augmentation for Instance Segmentation**

<https://arxiv.org/abs/2208.09709>

**Key Concepts and Taxonomy:**  
Addresses occlusion in instance segmentation through a dual copy-paste data augmentation pipeline.

**Main Contributions and Findings:**  
Enhances segmentation performance in team sports by injecting synthetic occlusions into training samples.

**Limitations and Future Directions:**  
May reduce realism if augmentations are not well aligned; lacks temporal coherence.

**Relevance to Primary Paper:**  
Strengthens perception tasks (segmentation) for crowded and high-occlusion environments like basketball or hockey.

**[65] Yan et al., 2022 – Strong Segmentation Pipeline for MMsports Challenge**

<https://arxiv.org/abs/2208.09709>

**Key Concepts and Taxonomy:**  
Presents a highly optimized pipeline combining instance segmentation and object detection for team sports.

**Main Contributions and Findings:**  
Achieves SOTA performance in crowded scenes by enhancing model precision and using fine-tuned backbone architectures.

**Limitations and Future Directions:**  
Resource-intensive training and inference; less suited for real-time settings.

**Relevance to Primary Paper:**  
Refines segmentation within perception—essential for body parsing, re-identification, and pose estimation.

**[66] Ghiasi et al., 2021 – Copy-Paste Data Augmentation for Instance Segmentation**

<https://arxiv.org/abs/2012.07177>

**Key Concepts and Taxonomy:**  
Introduces a simple and effective augmentation technique for instance segmentation via copy-pasting objects between images.

**Main Contributions and Findings:**  
Boosts performance without extra data collection or model changes; effective across many datasets.

**Limitations and Future Directions:**  
Works best when spatial and contextual consistency is preserved.

**Relevance to Primary Paper:**  
Foundational technique for data augmentation in perception tasks like player segmentation and action recognition.

**[67] Zhang et al., 2023 – Recognition of Throwing Actions in Basketball Using Segmentation**

<https://www.mdpi.com/2076-3417/13/4/2253>

**Key Concepts and Taxonomy:**  
Proposes a segmentation-driven method to classify throwing styles in basketball using pose and mask-based features.

**Main Contributions and Findings:**  
Demonstrates how segmentation masks help disambiguate action phases in fast movements like shots or passes.

**Limitations and Future Directions:**  
Dataset-specific; may not transfer to other sports without adaptation.

**Relevance to Primary Paper:**  
Sits between perception and comprehension, linking visual segmentation to semantic action understanding.

**[68] He et al., 2017 – Mask R-CNN**

<https://arxiv.org/abs/1703.06870>

**Key Concepts and Taxonomy:**  
A seminal framework for instance segmentation combining object detection and pixel-wise segmentation.

**Main Contributions and Findings:**  
Accurate and extensible model that serves as the base for most modern segmentation methods in sports analytics.

**Limitations and Future Directions:**  
Computationally expensive for real-time applications.

**Relevance to Primary Paper:**  
Backbone model for many perception modules involving segmentation and action parsing.

**[69] Chai et al., 2023 – Unsupervised Domain Adaptation for 3D Pose Estimation**

<https://arxiv.org/abs/2303.11182>

**Key Concepts and Taxonomy:**  
Develops a domain adaptation method that improves 3D human pose estimation in sports without requiring labeled target data.

**Main Contributions and Findings:**  
Transfers from lab-based to sports environments by mixing domain-specific features using adversarial training.

**Limitations and Future Directions:**  
More generalizable domain alignment strategies needed for broader application.

**Relevance to Primary Paper:**  
Improves pose estimation in real sports scenes—key to perception and AQA tasks.

**[70] Cao et al., 2021 – OpenPose: Real-Time Multi-Person 2D Pose Estimation**

<https://github.com/CMU-Perceptual-Computing-Lab/openpose>

<https://arxiv.org/abs/1812.08008>

**Key Concepts and Taxonomy:**  
Landmark method for real-time human pose estimation from 2D video using part affinity fields.

**Main Contributions and Findings:**  
Achieves accurate joint detection in multi-person scenarios; used across many sports analytics pipelines.

**Limitations and Future Directions:**  
Fails under severe occlusion; lacks depth awareness.

**Relevance to Primary Paper:**  
Core tool in perception-layer tasks including pose tracking, action classification, and quality evaluation.

**[71] Promrit & Waijanya, 2019 – Badminton Skill Practice Model Using Posture Detection and One-Shot Learning**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/8933221>

**Key Concepts and Taxonomy:**  
Develops a motion posture detection method to assess badminton basic skills using video embedding and one-shot learning.

**Main Contributions and Findings:**  
Demonstrates real-time posture recognition to provide immediate feedback during skill practice.

**Limitations and Future Directions:**  
Limited to specific predefined poses; lacks adaptation to more complex game scenarios.

**Relevance to Primary Paper:**  
Contributes to the decision-making layer for personalized training systems using pose feedback in racquet sports.

**[72] Suda et al., 2019 – Volleyball Trajectory Prediction Using Setter Skeletal Motion**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/8975877>

**Key Concepts and Taxonomy:**  
Utilizes skeletal motion of volleyball setters to predict ball trajectory before a pass or spike.

**Main Contributions and Findings:**  
Shows strong correlation between body pose and future ball paths using kinematic modeling.

**Limitations and Future Directions:**  
Focused on controlled environments; expansion to full-match scenes remains future work.

**Relevance to Primary Paper:**  
Strengthens comprehension tasks in volleyball through anticipatory modeling of play events.

**[73] Shimizu et al., 2019 – Tennis Shot Direction Prediction via Player Pose**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/8975876>

**Key Concepts and Taxonomy:**  
Predicts tennis shot direction based on the pose and positioning of players before contact.

**Main Contributions and Findings:**  
Improves forecast accuracy by modeling biomechanics, enabling smarter defensive player positioning.

**Limitations and Future Directions:**  
Performance drops with occlusion or inaccurate pose estimation.

**Relevance to Primary Paper:**  
Directly supports play forecasting in the decision layer for tennis applications.

**[74] Wu & Koike, 2020 – FuturePong: Real-Time Table Tennis Trajectory Forecasting Using Pose Prediction**

<https://dl.acm.org/doi/10.1145/3334480.3382855>

**Key Concepts and Taxonomy:**  
Real-time trajectory forecasting system based on pose-based neural networks.

**Main Contributions and Findings:**  
Predicts ball and paddle dynamics with low latency, enabling responsive practice environments.

**Limitations and Future Directions:**  
Limited data diversity; performance may suffer with different skill levels or play styles.

**Relevance to Primary Paper:**  
Enables game simulation and decision-making through motion prediction in fast-paced sports.

**[75] Sak et al., 2014 – LSTM Architectures for Acoustic Modeling**

<https://www.google.com/search?q=https://www.isca-speech.org/archive/archive_papers/interspeech_2014/i14_1960.pdf>

**Key Concepts and Taxonomy:**  
Introduces several LSTM architectures for large-scale sequence modeling, including bidirectional and projection-based models.

**Main Contributions and Findings:**  
Efficiently handles long-term dependencies in time-series data, including sequential sports actions.

**Limitations and Future Directions:**  
Generic model, needs domain adaptation for sports-specific features.

**Relevance to Primary Paper:**  
Backbone temporal modeling tool for pose forecasting, event spotting, and AQA tasks in comprehension and decision layers.

**[76] Einfalt et al., 2019 – Frame-Level Event Detection in Athletics Using Pose-Based CNNs**

<https://www.google.com/search?q=https://www.researchgate.net/publication/336340245_Frame-Level_Event_Detection_in_Athletics_Using_Pose-Based_Features_and_Convolutional_Neural_Networks>

**Key Concepts and Taxonomy:**  
Combines convolutional networks and pose features to detect event phases (e.g., takeoff, landing) in athletics.

**Main Contributions and Findings:**  
Captures event boundaries in track & field with high frame precision using pose sequences.

**Limitations and Future Directions:**  
Domain-specific; not easily transferable to other sports.

**Relevance to Primary Paper:**  
Improves comprehension in sports with rhythmic or cyclical motions through fine-grained phase detection.

**[77] Thilakarathne et al., 2022 – Group Activity Recognition in Volleyball Using Pose Tracking**

<https://www.mdpi.com/2076-3417/12/21/10899>

**Key Concepts and Taxonomy:**  
Utilizes tracked poses of multiple volleyball players to identify collective actions such as rallies or blocks.

**Main Contributions and Findings:**  
Models spatial dependencies and temporal transitions of player groups via pose features.

**Limitations and Future Directions:**  
Pose tracking quality impacts group classification.

**Relevance to Primary Paper:**  
Extends comprehension from individual to group action recognition in team sports.

**[78] Sun et al., 2017 – Taichi Dataset for Fine-Grained Pose Assessment**

<https://arxiv.org/abs/1704.08226>

**Key Concepts and Taxonomy:**  
Introduces a dataset and framework for recognizing and assessing complex Taichi movements using pose analysis.

**Main Contributions and Findings:**  
Enables assessment of form accuracy and flow in slow-movement sports via pose alignment.

**Limitations and Future Directions:**  
Focused on structured motions; may not generalize to dynamic sports.

**Relevance to Primary Paper:**  
Supports pose-based action quality assessment in sports emphasizing posture precision.

**[79] Hong et al., 2021 – Video Pose Distillation for Fine-Grained Sports Action Recognition**

<https://arxiv.org/abs/2108.06733>

**Key Concepts and Taxonomy:**  
Employs pose distillation to reduce the sample requirement for fine-grained recognition in niche sports.

**Main Contributions and Findings:**  
Achieves high accuracy with few-shot learning using pose-aware video summarization.

**Limitations and Future Directions:**  
Dependent on high-quality pose estimation during pretraining.

**Relevance to Primary Paper:**  
Enhances model generalization in comprehension tasks across sports with limited data.

**[80] An et al., 2022 – Multi-Granularity Network for Player Re-identification**

<https://arxiv.org/abs/2209.00031>

**Key Concepts and Taxonomy:**  
Proposes a hierarchical feature fusion system that considers granular visual cues and attention for player ReID.

**Main Contributions and Findings:**  
Improves discrimination between players with similar appearance using part-based and temporal cues.

**Limitations and Future Directions:**  
Limited validation in extreme occlusion scenarios.

**Relevance to Primary Paper:**  
Extends perception capabilities for long-term tracking and identity resolution.

**[81] Wang et al., 2021 – Game Impact Metric via Deep Reinforcement Learning**

<https://www.google.com/search?q=https://www.researchgate.net/publication/348255554_Learning_a_Game_Impact_Metric_for_Role-Based_Team_Sports_from_Spatiotemporal_Data>

**Key Concepts and Taxonomy:**  
Introduces a new player evaluation metric for basketball using Q-learning in a deep reinforcement learning framework.

**Main Contributions and Findings:**  
Assesses the contextual impact of player actions in dynamic game situations, offering more meaningful performance evaluation.

**Limitations and Future Directions:**  
Model performance depends heavily on high-fidelity tracking and labeled reward signals.

**Relevance to Primary Paper:**  
Directly supports the decision module through actionable, data-driven match evaluation.

**[82] Luo et al., 2020 – Inverse Reinforcement Learning in Ice Hockey**

<https://arxiv.org/abs/2007.03058>

**Key Concepts and Taxonomy:**  
Combines Q-learning with inverse reinforcement learning to estimate player utilities in a multi-agent ice hockey simulation.

**Main Contributions and Findings:**  
Proposes a unique ranking system based on learned strategies and team interactions.

**Limitations and Future Directions:**  
Limited generalizability to other sports beyond ice hockey due to environmental complexity.

**Relevance to Primary Paper:**  
Informs match evaluation and decision-making by providing interpretable player valuations.

**[83] Liu & Schulte, 2018 – Context-Aware Player Valuation in Ice Hockey**

<https://arxiv.org/abs/1802.04938>

**Key Concepts and Taxonomy:**  
Uses Q-function learning to assess individual contributions of players in specific in-game contexts.

**Main Contributions and Findings:**  
Introduces Game Impact Metric (GIM), capturing subtle yet critical player influence during gameplay.

**Limitations and Future Directions:**  
Only applicable in data-rich sports; lacks robustness to low-resolution tracking.

**Relevance to Primary Paper:**  
Pivotal for the decision layer by enriching performance analytics and coaching metrics.

**[84] Yanai et al., 2022 – QBall: Basketball Game Modeling via DDPG**

<https://www.google.com/search?q=https://www.jstage.jst.go.jp/article/pjsai/JSAI2022/0/JSAI2022_2Pin1-39/_article/-char/ja/>

**Key Concepts and Taxonomy:**  
Extends Deep Deterministic Policy Gradient (DDPG) for modeling player performance and team interactions in basketball.

**Main Contributions and Findings:**  
Merges deep RL and game simulation for evaluating both offensive and defensive decisions.

**Limitations and Future Directions:**  
Training requires extensive expert data and domain constraints.

**Relevance to Primary Paper:**  
Enhances simulation and match modeling capabilities under the decision framework.

**[85] Zhao et al., 2019 – Multi-Agent Learning in Team Sports**

<https://www.google.com/search?q=https://arxiv.org/abs/1908.10982>

**Key Concepts and Taxonomy:**  
Explores deep multi-agent reinforcement learning (MARL) in simulated sports environments.

**Main Contributions and Findings:**  
Demonstrates emergent team strategies through competition-based MARL.

**Limitations and Future Directions:**  
Coordination and reward sparsity pose challenges; needs domain-specific constraints.

**Relevance to Primary Paper:**  
Lays groundwork for future multi-agent simulation systems in tactical coaching and decision support.

**[86] Jia et al., 2020 – FEVER Basketball: An Asynchronous Multi-Agent Environment**

<https://arxiv.org/abs/2012.06201>

**Key Concepts and Taxonomy:**  
Presents a complex basketball simulation environment for single and multi-agent reinforcement learning.

**Main Contributions and Findings:**  
Allows realistic modeling of asynchronous player interactions in virtual training.

**Limitations and Future Directions:**  
Complexity hinders reproducibility and real-time deployment.

**Relevance to Primary Paper:**  
Supports decision-layer applications in simulated training and virtual coaching.

**[87] Huang et al., 2021 – TiKick: Multi-Agent Learning in GFootball**

<https://arxiv.org/abs/2104.03362>

**Key Concepts and Taxonomy:**  
Extends GFootball simulation by learning full game control from single-agent demonstrations using offline data.

**Main Contributions and Findings:**  
Demonstrates transfer learning across agents and environments.

**Limitations and Future Directions:**  
May overfit to training play styles; lacks generalization across strategies.

**Relevance to Primary Paper:**  
Supports scalable decision modeling in simulated team sports.

**[88] Lin et al., 2023 – TiZero: Curriculum Learning for Multi-Agent Soccer**

<https://arxiv.org/abs/2304.04565>

**Key Concepts and Taxonomy:**  
Employs curriculum learning and self-play for mastering full games in multi-agent football.

**Main Contributions and Findings:**  
Improves long-term planning and coordination across agents through adaptive learning.

**Limitations and Future Directions:**  
Still faces challenges in real-world transfer and non-transitivity.

**Relevance to Primary Paper:**  
Paves the way for decision-level learning in autonomous sports simulations.

**[89] Yu et al., 2021 – MA-PPO for Cooperative Multi-Agent Games**

<https://arxiv.org/abs/2103.01955>

**Key Concepts and Taxonomy:**  
Applies Multi-Agent Proximal Policy Optimization (MA-PPO) to cooperative environments, demonstrating strong emergent coordination.

**Main Contributions and Findings:**  
Achieves SOTA in complex team-based sports simulations using centralized training and decentralized execution.

**Limitations and Future Directions:**  
Scalability to large teams and complex environments is still under evaluation.

**Relevance to Primary Paper:**  
Enables robust policy learning for team-based sports decision systems.

**[90] Wen et al., 2022 – Sequence Modeling for Multi-Agent Reinforcement Learning**

<https://arxiv.org/abs/2205.14953>

**Key Concepts and Taxonomy:**  
Reframes MARL as a sequence modeling problem using attention mechanisms and autoregressive decoding.

**Main Contributions and Findings:**  
Simplifies policy training and improves coordination in dynamic sports scenarios.

**Limitations and Future Directions:**  
Still faces limitations with noisy inputs and adversarial interactions.

**Relevance to Primary Paper:**  
Connects the decision and comprehension layers via predictive modeling of agent behaviors.

**[101] Chen et al., 2021 – CTR-GCN: Channel-wise Topology Refinement Graph Convolution for Skeleton-Based Action Recognition**

<https://arxiv.org/abs/2107.08318>

**Key Concepts and Taxonomy:**  
Introduces an adaptive graph convolutional network that dynamically learns topological dependencies among human joints.

**Main Contributions and Findings:**  
Improves skeleton-based action recognition by refining graph structure at the channel level rather than globally.

**Limitations and Future Directions:**  
Requires high-quality pose inputs; may be affected by noisy keypoints in broadcast sports.

**Relevance to Primary Paper:**  
Enhances individual action recognition in perception/comprehension pipelines of sports systems.

**[102] Yuan et al., 2021 – GAR DIN: Spatio-Temporal Dynamic Inference Network for Group Activity Recognition**

<https://arxiv.org/abs/2104.02251>

**Key Concepts and Taxonomy:**  
Proposes a model combining dynamic spatial-temporal message passing for group activity recognition.

**Main Contributions and Findings:**  
Outperforms static GCN-based methods in team sports by accounting for changing spatial configurations of players.

**Limitations and Future Directions:**  
Computationally expensive in long sequences; may need simplification for real-time analysis.

**Relevance to Primary Paper:**  
Supports comprehension-layer tasks such as strategy recognition in sports like volleyball or soccer.

**[103] Duan et al., 2021 – PoseC3D: Revisiting Skeleton-Based Action Recognition with 3D CNNs**

<https://arxiv.org/abs/2104.13586>

**Key Concepts and Taxonomy:**  
Leverages 3D CNNs directly on pose sequences for action recognition, departing from graph-based models.

**Main Contributions and Findings:**  
Achieves strong performance on various fine-grained sports datasets by capturing temporal motion cues via convolutions.

**Limitations and Future Directions:**  
Heavy reliance on accurate skeleton estimation; limited interpretability.

**Relevance to Primary Paper:**  
Provides an alternative to GCNs for action recognition in the comprehension pipeline.

**[104] Xiang et al., 2018 – S3D: Stacking Segmental P3D for Action Quality Assessment**

<https://www.google.com/search?q=https://bmvc2018.org/contents/papers/0503.pdf>

**Key Concepts and Taxonomy:**  
Combines Pseudo-3D convolutions and stacked temporal blocks to extract motion-aware features for AQA.

**Main Contributions and Findings:**  
Captures spatiotemporal features for accurate scoring of sports performances like diving and gymnastics.

**Limitations and Future Directions:**  
Architecture complexity and limited pose interpretability.

**Relevance to Primary Paper:**  
Strengthens AQA modeling in comprehension through advanced temporal segmentation.

**[105] Parmar & Morris, 2019 – C3D-LSTM for Action Quality Assessment**

<https://arxiv.org/abs/1705.04494>

**Key Concepts and Taxonomy:**  
Proposes a deep learning model that combines 3D CNN (C3D) for spatial feature extraction with LSTM for temporal modeling.

**Main Contributions and Findings:**  
Effective in estimating AQA scores for multiple sports (e.g., diving, skating) from full video sequences.

**Limitations and Future Directions:**  
Sequence length sensitivity; generalization to unseen action classes is weak.

**Relevance to Primary Paper:**  
Important AQA baseline referenced in multiple newer comprehension methods.

**[106] Parmar & Morris, 2019 – Multi-Task Learning for AQA**

<https://arxiv.org/abs/1904.07849>

**Key Concepts and Taxonomy:**  
Extends AQA models to include three tasks: action classification, quality scoring, and comment generation.

**Main Contributions and Findings:**  
Improves performance and generalization through shared representations across related tasks.

**Limitations and Future Directions:**  
Limited to sports with predefined action sequences.

**Relevance to Primary Paper:**  
Adds multi-task learning capability to comprehension models for broader application in judged sports.

**[107] Xu et al., 2019 – C3D-MSLSTM for Scoring Figure Skating Videos**

<https://www.google.com/search?q=https://www.researchgate.net/publication/334865187_C3D-MSLSTM_A_Novel_Framework_for_Scoring_Figure_Skating_Videos>

**Key Concepts and Taxonomy:**  
Introduces a multi-scale LSTM on top of C3D features for modeling hierarchical motion patterns.

**Main Contributions and Findings:**  
Allows evaluation of fine-to-coarse action segments, aiding nuanced performance assessments.

**Limitations and Future Directions:**  
Limited to figure skating; needs domain transfer evaluation.

**Relevance to Primary Paper:**  
Refines AQA in sports requiring high-precision motion assessments.

**[108] Tang et al., 2020 – I3D-USDL: Uncertainty-Aware Score Distribution Learning for AQA**

<https://arxiv.org/abs/2003.06830>

**Key Concepts and Taxonomy:**  
Uses I3D features and predicts score distributions rather than point estimates.

**Main Contributions and Findings:**  
Incorporates uncertainty modeling, improving robustness to subjective judging biases.

**Limitations and Future Directions:**  
Model complexity; fine-tuning needed for different sports.

**Relevance to Primary Paper:**  
Promotes fairness and generalizability in AQA systems through distributional learning.

**[109] Wang et al., 2021 – TSA-Net: Tube Self-Attention for AQA**

<https://arxiv.org/abs/2104.09228>

**Key Concepts and Taxonomy:**  
Applies attention mechanisms along temporal video tubes to model subtle motion patterns.

**Main Contributions and Findings:**  
Captures long-range dependencies and motion saliency for improved score prediction.

**Limitations and Future Directions:**  
Less effective in real-time scenarios due to computational demands.

**Relevance to Primary Paper:**  
Advances comprehension in AQA via self-attention, influencing many modern Transformer-based models.

**[110] Qi et al., 2022 – Weakly Supervised Video Fight Detection**

<https://arxiv.org/abs/2206.13084>

**Key Concepts and Taxonomy:**  
Proposes a two-stage weakly supervised model for detecting violent events in sports videos.

**Main Contributions and Findings:**  
Improves detection using temporal attention and contrastive learning without dense labels.

**Limitations and Future Directions:**  
Focused on binary fight/no-fight detection; lacks granularity.

**Relevance to Primary Paper:**  
Links perception and decision by detecting anomalies or fouls for automated refereeing or highlight generation.

**[111] Carreira & Zisserman, 2017 – Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset**

<https://arxiv.org/abs/1705.07750>

**Key Concepts and Taxonomy:**  
Introduces I3D (Inflated 3D ConvNet), which inflates 2D ConvNets into 3D to process video spatiotemporal data.

**Main Contributions and Findings:**  
Achieves strong results on action recognition tasks and introduces the large-scale Kinetics dataset for video understanding.

**Limitations and Future Directions:**  
Model is computationally expensive and not optimized for real-time applications.

**Relevance to Primary Paper:**  
I3D is a foundational model widely used across perception and comprehension layers, including AQA and summarization.

**[112] Monfort et al., 2019 – Moments in Time Dataset: One Million Videos for Event Understanding**

<http://moments.csail.mit.edu/>

<https://arxiv.org/abs/1801.03150>

**Key Concepts and Taxonomy:**  
Proposes a large-scale dataset covering one million videos annotated for dynamic event understanding.

**Main Contributions and Findings:**  
Supports temporal reasoning in event classification and understanding of high-level concepts in video.

**Limitations and Future Directions:**  
Does not focus on sports specifically; broad categories may overlook fine-grained actions.

**Relevance to Primary Paper:**  
Serves as a strong pretraining source for models later fine-tuned on sports comprehension tasks.

**[113] Wang et al., 2022 – Multi-task Representation Learning for Action Recognition Pre-training**

<https://arxiv.org/abs/2203.09034>

**Key Concepts and Taxonomy:**  
Explores human-centered prior-guided multi-task learning to improve feature reuse and task efficiency in video pre-training.

**Main Contributions and Findings:**  
Improves transferability of representations across sports comprehension tasks.

**Limitations and Future Directions:**  
Limited empirical testing on domain-specific (e.g., Olympic) datasets.

**Relevance to Primary Paper:**  
Supports efficient comprehension model training across action recognition and quality assessment domains.

**[114] Shao et al., 2020 – FineGym: A Hierarchical Video Dataset for Fine-Grained Action Understanding**

<https://sdolivia.github.io/FineGym/>

<https://arxiv.org/abs/2004.06704>

**Key Concepts and Taxonomy:**  
Presents a hierarchically annotated gymnastics dataset for fine-grained action recognition and temporal localization.

**Main Contributions and Findings:**  
Captures both coarse and fine action granularity, facilitating structured recognition models.

**Limitations and Future Directions:**  
Restricted to gymnastics; expansion to other complex sports is needed.

**Relevance to Primary Paper:**  
Key dataset for action quality assessment and sequence modeling in judged sports.

**[115] Sun et al., 2017 – Taichi Dataset for Fine-Grained Action Recognition**

<https://arxiv.org/abs/1704.08226>

**Key Concepts and Taxonomy:**  
Introduces a Taichi dataset with fine-grained action classes for evaluating form-based actions.

**Main Contributions and Findings:**  
Captures subtle transitions and symmetrical poses critical for pose analysis.

**Limitations and Future Directions:**  
Small dataset; generalizability to dynamic sports is limited.

**Relevance to Primary Paper:**  
Supports development of pose-sensitive models in comprehension tasks such as form assessment.

**[116] Cheng et al., 2020 – SlowOnly for Skeleton-Based Action Recognition**

<https://arxiv.org/abs/2001.00222>

**Key Concepts and Taxonomy:**  
Simplifies the SlowFast architecture by using a single slow pathway to process 3D skeletal data.

**Main Contributions and Findings:**  
Provides lightweight, accurate skeleton-based recognition without requiring multimodal input.

**Limitations and Future Directions:**  
Not optimized for real-time inference; limited to single-person actions.

**Relevance to Primary Paper:**  
Useful in skeleton-based recognition pipelines common in perception modules.

**[117] Cheng et al., 2021 – MSG3D: Multi-Scale Graph Convolution Network for 3D Skeleton-Based Action Recognition**

<https://arxiv.org/abs/2012.11580>

**Key Concepts and Taxonomy:**  
Implements a multi-scale spatial-temporal graph model to capture detailed dynamics of skeletal movements.

**Main Contributions and Findings:**  
Outperforms standard GCNs by learning both local and global temporal cues.

**Limitations and Future Directions:**  
Training is complex and data-intensive.

**Relevance to Primary Paper:**  
Advances skeleton-based modeling crucial for body-centric sports analytics.

**[118] Shi et al., 2019 – Two-Stream Adaptive Graph Convolutional Networks (2s-AGCN)**

<https://arxiv.org/abs/1904.02807>

**Key Concepts and Taxonomy:**  
Proposes a two-stream model leveraging both joint positions and motion dynamics for skeleton action recognition.

**Main Contributions and Findings:**  
Improves robustness to occlusion and spatial variation in pose sequences.

**Limitations and Future Directions:**  
Performance can drop in low-resolution sports video.

**Relevance to Primary Paper:**  
Supports perception systems requiring accurate player pose recognition.

**[119] Shan et al., 2020 – Fineskating: Figure Skating Dataset for Action and Quality Assessment**

<https://arxiv.org/abs/2007.03150>

**Key Concepts and Taxonomy:**  
Releases a dataset with hierarchical action annotations and scoring for figure skating routines.

**Main Contributions and Findings:**  
Supports AQA research with competition-level videos and multi-level granularity.

**Limitations and Future Directions:**  
Small number of unique skaters; limited diversity.

**Relevance to Primary Paper:**  
Directly enhances comprehension-layer models in judged sports.

**[120] Liang et al., 2022 – Context-Aware Score Prediction for AQA with Transformers**

<https://arxiv.org/abs/2207.13501>

**Key Concepts and Taxonomy:**  
Leverages Transformer encoders to model temporal dependencies for action scoring.

**Main Contributions and Findings:**  
Integrates visual and contextual information for nuanced quality predictions.

**Limitations and Future Directions:**  
Not easily adaptable to sports with highly variable action sequences.

**Relevance to Primary Paper:**  
Represents state-of-the-art in fine-grained comprehension tasks like AQA.

**[121] Bertasius et al., 2019 – Am I a Baller? Basketball Performance Assessment from First-Person Videos**

<https://arxiv.org/abs/1710.03357>

**Key Concepts and Taxonomy:**  
Utilizes first-person video input to assess basketball performance through event detection and skill evaluation.

**Main Contributions and Findings:**  
Applies convolutional-LSTM networks to detect and evaluate sequences of actions in basketball gameplay.

**Limitations and Future Directions:**  
Limited to ego-centric views; does not generalize well to third-person broadcast scenarios.

**Relevance to Primary Paper:**  
Provides a novel perspective on player performance analysis relevant for comprehension and decision layers.

**[122] Yu et al., 2021 – Group-Aware Contrastive Regression for Action Quality Assessment**

<https://arxiv.org/abs/2103.11116>

**Key Concepts and Taxonomy:**  
Introduces a contrastive learning framework that incorporates group context to assess action quality.

**Main Contributions and Findings:**  
Demonstrates that inter-video contrast improves the accuracy of score predictions in judged sports.

**Limitations and Future Directions:**  
Requires group-level annotations and fine-grained temporal segmentation.

**Relevance to Primary Paper:**  
Advances AQA modeling by integrating social context and contrastive cues.

**[123] Li et al., 2018 – End-to-End Learning for Action Quality Assessment**

<https://arxiv.org/abs/1805.11404>

**Key Concepts and Taxonomy:**  
Proposes a C3D-based end-to-end AQA model combining regression and ranking losses.

**Main Contributions and Findings:**  
Improves both score prediction and relative ranking accuracy in evaluated actions.

**Limitations and Future Directions:**  
Overfits to small datasets; needs broader validation.

**Relevance to Primary Paper:**  
Key early work in comprehension-layer score prediction from spatiotemporal features.

**[124] Parisi et al., 2016 – Human Motion Assessment with Recurrent Self-Organizing Networks**

<https://www.frontiersin.org/articles/10.3389/fnbot.2016.00019/full>

**Key Concepts and Taxonomy:**  
Uses a recurrent self-organizing neural network to learn and compare body motion sequences.

**Main Contributions and Findings:**  
Enables online and unsupervised evaluation of movement quality in continuous time.

**Limitations and Future Directions:**  
Performance varies with motion complexity; limited fine-grained score regression.

**Relevance to Primary Paper:**  
Provides a base for real-time comprehension tasks like form evaluation in sports.

**[125] Kim & Yong, 2017 – EvaluationNet: Can Human Skill Be Evaluated by Deep Networks?**

<https://arxiv.org/abs/1711.05316>

**Key Concepts and Taxonomy:**  
Models action as a sequence of structured units using LSTMs for skill evaluation.

**Main Contributions and Findings:**  
Learns hierarchical relationships among movements to assess performance.

**Limitations and Future Directions:**  
Requires pre-segmented action sequences; lacks interpretability.

**Relevance to Primary Paper:**  
Contributes to early AQA research via structured action modeling.

**[126] Wang et al., 2021 – TSA-Net: Tube Self-Attention for Action Quality Assessment**

<https://arxiv.org/abs/2104.09228>

**Key Concepts and Taxonomy:**  
Uses tube self-attention to model inter-frame dependencies for AQA.

**Main Contributions and Findings:**  
Improves both score regression and robustness to background noise using sparse feature interactions.

**Limitations and Future Directions:**  
Computationally heavy; not ideal for mobile deployment.

**Relevance to Primary Paper:**  
SOTA comprehension model integrating attention-based aggregation.

**[127] Tang et al., 2020 – I3D-USDL: Uncertainty-Aware Score Distribution Learning for AQA**

<https://arxiv.org/abs/2003.06830>

**Key Concepts and Taxonomy:**  
Models score distribution with uncertainty in deep learning predictions using I3D features.

**Main Contributions and Findings:**  
Enables probabilistic scoring aligned with real-world subjectivity in judged sports.

**Limitations and Future Directions:**  
Needs large-scale score distribution annotations.

**Relevance to Primary Paper:**  
Pioneers uncertainty modeling in comprehension-layer AQA models.

**[128] Xu et al., 2019 – C3D-MSLSTM: Multi-Scale LSTM for Scoring Sports Videos**

<https://www.google.com/search?q=https://www.researchgate.net/publication/334865187_C3D-MSLSTM_A_Novel_Framework_for_Scoring_Figure_Skating_Videos>

**Key Concepts and Taxonomy:**  
Combines C3D visual features with hierarchical LSTM for detailed temporal modeling.

**Main Contributions and Findings:**  
Captures both micro (frame-level) and macro (action-level) motion patterns.

**Limitations and Future Directions:**  
Specific to structured sports like skating or gymnastics.

**Relevance to Primary Paper:**  
Enables fine-grained modeling in long-sequence AQA comprehension.

**[129] Parmar & Morris, 2019 – C3D-AVG-MTL: Multitask Learning for Action Quality Assessment**

<https://arxiv.org/abs/1904.07849>

**Key Concepts and Taxonomy:**  
Introduces a multitask framework predicting action class, score, and commentary.

**Main Contributions and Findings:**  
Improves score generalization and facilitates interpretability.

**Limitations and Future Directions:**  
Limited performance on unseen sports due to domain-specific features.

**Relevance to Primary Paper:**  
Fuses AQA and semantic prediction in comprehension-layer pipelines.

**[130] Xiang et al., 2018 – S3D: Stacking Segmental P3D for Action Quality Assessment**

<https://www.google.com/search?q=https://bmvc2018.org/contents/papers/0503.pdf>

**Key Concepts and Taxonomy:**  
Divides complex actions into distinct temporal phases and models each using P3D modules.

**Main Contributions and Findings:**  
Supports detailed motion analysis with phase-level interpretability.

**Limitations and Future Directions:**  
Rigid segmentation assumptions may limit flexibility.

**Relevance to Primary Paper:**  
Foundation for many temporal segmentation-based AQA models in comprehension tasks.

**[131] Han et al., 2022 – Dual-AI: Dual-Path Actor Interaction Learning for Group Activity Recognition**

<https://arxiv.org/abs/2203.04521>

**Key Concepts and Taxonomy:**  
Introduces a dual-path transformer model to capture temporal and spatial actor interactions in team sports.

**Main Contributions and Findings:**  
Demonstrates superior accuracy in group activity recognition using role-aware attention and actor interaction modeling.

**Limitations and Future Directions:**  
Sensitive to actor misalignment and occlusion; requires high-quality detections.

**Relevance to Primary Paper:**  
Supports comprehension-layer tasks like team behavior modeling in volleyball and soccer.

**[132] Xu & Yin, 2023 – MLP-AIR: MLP-Based Actor Interaction Relation Learning**

<https://arxiv.org/abs/2301.10022>

**Key Concepts and Taxonomy:**  
Proposes a multi-layer perceptron approach for learning interaction relations between actors in group activities.

**Main Contributions and Findings:**  
Reduces complexity while achieving competitive accuracy in sports action classification tasks.

**Limitations and Future Directions:**  
Lacks temporal interpretability compared to transformer-based counterparts.

**Relevance to Primary Paper:**  
Promotes lightweight alternatives for comprehension tasks involving multi-agent sports.

**[133] Bettadapura et al., 2016 – Contextual Cues for Generating Basketball Highlights**

<https://www.google.com/search?q=https://dl.acm.org/doi/10.1145/2983524.2983533>

**Key Concepts and Taxonomy:**  
Uses semantic and contextual analysis to select highlight-worthy segments in basketball videos.

**Main Contributions and Findings:**  
Employs cues like player reactions and crowd noise to detect key moments beyond explicit actions.

**Limitations and Future Directions:**  
Limited generalization across sports due to domain-specific features.

**Relevance to Primary Paper:**  
Influences summarization models for sports video highlight generation.

**[134] Heilbron et al., 2017 – Semantic Context Cascade (SCC) for Efficient Action Detection**

<https://arxiv.org/abs/1709.09198>

**Key Concepts and Taxonomy:**  
Applies a multi-stage model that uses hierarchical context reasoning for event detection in sports videos.

**Main Contributions and Findings:**  
Efficiently detects temporal actions through progressive context propagation.

**Limitations and Future Directions:**  
Best suited for well-segmented clips; suffers with noisy or occluded data.

**Relevance to Primary Paper:**  
Advances comprehension-layer spotting and highlight selection techniques.

**[135] Felsen et al., 2017 – What Will Happen Next? Forecasting Player Moves in Sports Videos**

<https://arxiv.org/abs/1711.04258>

**Key Concepts and Taxonomy:**  
Introduces a model to predict short-term future movements of players using deep generative techniques.

**Main Contributions and Findings:**  
Enables early action recognition and anticipation in sports scenarios with high accuracy.

**Limitations and Future Directions:**  
Forecast horizon is limited; works best with structured plays.

**Relevance to Primary Paper:**  
Core to the decision-layer task of play forecasting.

**[136] Cioppa et al., 2018 – Semantic Interpretation of Main Camera Streams in Soccer**

<https://arxiv.org/abs/1709.09198>

**Key Concepts and Taxonomy:**  
Builds a bottom-up semantic parser to extract meaningful events from continuous soccer video streams.

**Main Contributions and Findings:**  
Performs robust interpretation of camera shots for segmenting soccer events like goals and fouls.

**Limitations and Future Directions:**  
Designed for soccer; limited transferability.

**Relevance to Primary Paper:**  
Pioneering work in summarization and event detection within long-form soccer matches.

**[137] Tsunoda et al., 2017 – Hierarchical LSTM for Football Action Recognition**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/8296317>

**Key Concepts and Taxonomy:**  
Applies hierarchical LSTM models to learn high-level game phases and detailed actions from broadcast video.

**Main Contributions and Findings:**  
Demonstrates robust hierarchical encoding of game structure in football.

**Limitations and Future Directions:**  
Performance drops on sports with less consistent structure.

**Relevance to Primary Paper:**  
Strengthens comprehension-layer tasks for structured team sports.

**[138] Cai et al., 2019 – Pose and Flow-Based Temporal Hockey Action Recognition**

<https://www.google.com/search?q=https://openaccess.thecvf.com/content_WACV_2020/html/Cai_Pose-and-Flow-based_Temporal_Hockey_Action_Recognition_WACV_2020_paper.html>

**Key Concepts and Taxonomy:**  
Fuses pose and optical flow information to model temporal action progression in hockey.

**Main Contributions and Findings:**  
Improves recognition accuracy in fast-paced sports with player occlusion and abrupt motion.

**Limitations and Future Directions:**  
Pose estimation challenges reduce performance in cluttered frames.

**Relevance to Primary Paper:**  
Hybrid perception-comprehension model; foundational for complex action recognition.

**[139] Sanabria et al., 2019 – Multimodal Summarization Architecture for Soccer Games**

<https://www.google.com/search?q=https://www.researchgate.net/publication/337191196_Multimodal_Summarization_Architecture_for_Soccer_Games_with_Audio-Visual_Attention>

**Key Concepts and Taxonomy:**  
Proposes a deep network combining audio, text, and video features to summarize soccer matches.

**Main Contributions and Findings:**  
Achieves coherent highlight generation using cross-modal alignment.

**Limitations and Future Directions:**  
Struggles with sparse audio events and ambiguous crowd noise.

**Relevance to Primary Paper:**  
Advances summarization by integrating multimodal cues, aligning with comprehension goals.

**[140] Li et al., 2018 – Tennis Shot Classification Using Wearable Sensors**

<https://www.mdpi.com/1424-8220/18/10/3223>

**Key Concepts and Taxonomy:**  
Classifies tennis shot types using IMU data from wrist-worn wearables.

**Main Contributions and Findings:**  
Achieves high accuracy in real-time shot recognition using lightweight sensor fusion.

**Limitations and Future Directions:**  
Limited to laboratory-controlled environments; sensor placement sensitive.

**Relevance to Primary Paper:**  
Demonstrates wearable AI use in the perception layer for skill analysis.

**[141] Giancola et al., 2018 – SoccerNet: A Scalable Dataset for Action Spotting in Soccer Videos**

<https://www.soccer-net.org/>

<https://arxiv.org/abs/1807.05193>

**Key Concepts and Taxonomy:**  
Introduces the large-scale SoccerNet dataset, annotated for action spotting tasks in full-match soccer videos.

**Main Contributions and Findings:**  
Provides benchmarks and metrics to detect events like goals, substitutions, and cards with second-level precision.

**Limitations and Future Directions:**  
Focused on professional soccer; lacks diversity across leagues and formats.

**Relevance to Primary Paper:**  
One of the cornerstone datasets for video comprehension and event detection in soccer.

**[142] Cioppa et al., 2020 – A Context-Aware Loss Function for Action Spotting in Soccer Videos**

<https://arxiv.org/abs/2006.14337>

**Key Concepts and Taxonomy:**  
Proposes a custom loss function to penalize temporal errors in event detection based on proximity to ground truth.

**Main Contributions and Findings:**  
Improves the temporal precision of action spotting networks in sports like soccer.

**Limitations and Future Directions:**  
Task-specific loss may not generalize to other sports without modification.

**Relevance to Primary Paper:**  
Supports comprehension and decision layers by refining temporal localization of key events.

**[143] Hong et al., 2022 – Spotting Temporally Precise, Fine-Grained Events in Video**

<https://arxiv.org/abs/2203.08332>

**Key Concepts and Taxonomy:**  
Combines visual similarity and temporal consistency to detect high-precision events in continuous video.

**Main Contributions and Findings:**  
Outperforms baselines in spotting subtle sports actions like fouls and passes.

**Limitations and Future Directions:**  
Requires dense temporal annotations; limited scalability.

**Relevance to Primary Paper:**  
Advances action spotting for comprehension and officiating support in sports AI.

**[144] Darwish & El-Shabrway, 2022 – STE: Spatio-Temporal Encoder for Action Spotting in Soccer**

<https://arxiv.org/abs/2208.14081>

**Key Concepts and Taxonomy:**  
Presents a lightweight encoder that fuses temporal and spatial features for spotting soccer events.

**Main Contributions and Findings:**  
Achieves real-time event spotting with competitive accuracy.

**Limitations and Future Directions:**  
Focused on soccer; needs extension to other team sports.

**Relevance to Primary Paper:**  
Enables efficient comprehension-layer event detection for real-time decision support.

**[145] Cartas et al., 2022 – Graph-Based Soccer Action Spotting with Unsupervised Player Classification**

<https://arxiv.org/abs/2208.08778>

**Key Concepts and Taxonomy:**  
Builds a graph of player trajectories and classifies players unsupervised to detect soccer actions.

**Main Contributions and Findings:**  
Increases explainability and performance in identifying complex team behaviors.

**Limitations and Future Directions:**  
Depends on reliable player detection and tracking.

**Relevance to Primary Paper:**  
Integrates graph neural networks into comprehension-layer models for tactical insight.

**[146] Zhu et al., 2022 – Transformer-Based System for Action Spotting in Soccer Videos**

<https://arxiv.org/abs/2208.08985>

**Key Concepts and Taxonomy:**  
Applies Transformer architectures to model long-range dependencies in soccer gameplay.

**Main Contributions and Findings:**  
Outperforms RNN-based spotting methods in detecting sparse yet impactful events.

**Limitations and Future Directions:**  
Memory-intensive; requires tuning for varying match formats.

**Relevance to Primary Paper:**  
Brings state-of-the-art NLP techniques into sports comprehension pipelines.

**[147] Soares & Shah, 2022 – Action Spotting Using Dense Detection Anchors**

<https://arxiv.org/abs/2208.11438>

**Key Concepts and Taxonomy:**  
Revisits dense anchor-based detection for temporally locating events in soccer videos.

**Main Contributions and Findings:**  
Achieves high precision on the SoccerNet challenge benchmarks.

**Limitations and Future Directions:**  
Dense anchoring increases computational load.

**Relevance to Primary Paper:**  
Enhances precision for comprehension-layer temporal spotting tasks.

**[148] Soares et al., 2022 – Temporally Precise Action Spotting Using Dense Anchors**

<https://arxiv.org/abs/2208.11438>

**Key Concepts and Taxonomy:**  
Improves anchor-based action spotting by tuning temporal granularity.

**Main Contributions and Findings:**  
Reduces false detections in continuous match footage while maintaining recall.

**Limitations and Future Directions:**  
Still lacks interpretability in predictions.

**Relevance to Primary Paper:**  
Contributes to video summarization and refereeing automation in decision systems.

**[149] Pan et al., 2019 – Action Assessment by Joint Relation Graphs**

<https://arxiv.org/abs/1908.08339>

**Key Concepts and Taxonomy:**  
Introduces a graph-based method using relative joint motion for scoring athletic actions.

**Main Contributions and Findings:**  
Models local dependencies between joints, boosting action quality assessment accuracy.

**Limitations and Future Directions:**  
Works best on pose-rich datasets; performance degrades with pose noise.

**Relevance to Primary Paper:**  
Strong candidate for comprehension-layer pose-based skill evaluation.

**[150] Parisi et al., 2016 – Human Motion Assessment in Real Time Using Recurrent Self-Organization**

<https://www.frontiersin.org/articles/10.3389/fnbot.2016.00019/full>

**Key Concepts and Taxonomy:**  
Applies recurrent self-organizing networks to analyze human motion in an online setting.

**Main Contributions and Findings:**  
Provides real-time motion profiling without needing explicit labels.

**Limitations and Future Directions:**  
Requires high consistency in motion sequences; may not generalize to chaotic plays.

**Relevance to Primary Paper:**  
Early attempt at unsupervised comprehension for sports skill analysis.

**[151] Kim & Yong, 2017 – EvaluationNet: Can Human Skill Be Evaluated by Deep Networks?**

<https://arxiv.org/abs/1711.05316>

**Key Concepts and Taxonomy:**  
Models human action as structured sequences and uses LSTM networks for encoding motion.

**Main Contributions and Findings:**  
Captures temporal dynamics of structured action units in tasks like diving and gymnastics, enabling automatic scoring.

**Limitations and Future Directions:**  
Limited to structured and staged actions; lacks robustness to noisy or incomplete sequences.

**Relevance to Primary Paper:**  
Early and influential model in the action quality assessment (AQA) category of comprehension tasks.

**[152] Yu et al., 2021 – Group-Aware Contrastive Regression for AQA**

<https://arxiv.org/abs/2103.11116>

**Key Concepts and Taxonomy:**  
Combines contrastive learning with regression to enable action ranking and scoring based on contextual information.

**Main Contributions and Findings:**  
Improves score discrimination in group settings, crucial for team-based judged sports.

**Limitations and Future Directions:**  
Requires accurate video-level feature extraction; may be complex to adapt for low-resolution or occluded frames.

**Relevance to Primary Paper:**  
Refines comprehension-layer models for comparative performance scoring.

**[153] Li et al., 2018 – End-to-End Learning for Action Quality Assessment**

<https://arxiv.org/abs/1805.11404>

**Key Concepts and Taxonomy:**  
Utilizes 3D CNNs and end-to-end supervised learning for AQA, incorporating ranking loss and MSE.

**Main Contributions and Findings:**  
Delivers precise regression of skill scores for actions like gymnastics or skating.

**Limitations and Future Directions:**  
Struggles with generalization to diverse sports due to fixed temporal length input.

**Relevance to Primary Paper:**  
Forms a baseline for AQA model benchmarking in comprehension systems.

**[154] Bertasius et al., 2019 – Am I a Baller? Basketball Performance Assessment from First-Person Videos**

<https://arxiv.org/abs/1710.03357>

**Key Concepts and Taxonomy:**  
Assesses basketball actions through egocentric videos using a convolutional LSTM pipeline.

**Main Contributions and Findings:**  
Demonstrates ability to detect and evaluate shooting and dribbling performance in first-person view.

**Limitations and Future Directions:**  
First-person data can introduce jitter and occlusions; model needs calibration for third-person tasks.

**Relevance to Primary Paper:**  
Applies comprehension techniques to player-centric skill evaluation scenarios.

**[155] Agyeman et al., 2019 – Soccer Video Summarization Using Deep Learning**

<https://www.google.com/search?q=https://dl.acm.org/doi/10.1145/3347450.3357161>

**Key Concepts and Taxonomy:**  
Applies a 3D ResNet-LSTM pipeline to extract highlights from lengthy soccer matches.

**Main Contributions and Findings:**  
Detects five action classes to enable concise highlight generation.

**Limitations and Future Directions:**  
Summarization quality depends on frame-level annotations and model calibration.

**Relevance to Primary Paper:**  
Supports summarization layer in sports comprehension by combining vision and temporal modeling.

**[156] Rafiq et al., 2020 – Scene Classification for Sports Video Summarization Using Transfer Learning**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/9064789>

**Key Concepts and Taxonomy:**  
Uses transfer learning with pre-trained AlexNet for scene classification in cricket videos.

**Main Contributions and Findings:**  
Provides lightweight summarization suitable for sports with less structured video feeds.

**Limitations and Future Directions:**  
Limited to static scene understanding; lacks motion or context modeling.

**Relevance to Primary Paper:**  
Provides scalable scene summarization in low-resource sports settings.

**[157] Khan et al., 2020 – Content-Aware Summarization of Broadcast Sports Videos**

<https://www.mdpi.com/2076-3417/10/7/2550>

**Key Concepts and Taxonomy:**  
Combines audio and visual cues for automatic generation of sports video summaries.

**Main Contributions and Findings:**  
Captures audience excitement and game dynamics to improve summarization relevance.

**Limitations and Future Directions:**  
Audio signals can be noisy or ambiguous in non-stadium sports settings.

**Relevance to Primary Paper:**  
Enriches comprehension with multimodal information for summarization tasks.

**[158] Shingrakhia & Patel, 2022 – Hybrid ML Models for Cricket Video Summarization**

<https://www.google.com/search?q=https://www.scitepress.org/Link.aspx%3Fdoi%3D10.5220/0010839800003120>

**Key Concepts and Taxonomy:**  
Combines stacked RNNs and deep belief networks (DBNs) for multimodal sports video classification.

**Main Contributions and Findings:**  
Accurately detects key cricket match events using hybrid models.

**Limitations and Future Directions:**  
Tailored to cricket; adaptation to other sports requires re-architecture.

**Relevance to Primary Paper:**  
Enhances summarization systems by fusing audio-visual and sequential modeling.

**[159] Li et al., 2022 – Hierarchical Structure-Aware Summarization for Soccer Videos**

<https://arxiv.org/abs/2208.10651>

**Key Concepts and Taxonomy:**  
Proposes structure-adaptive summarization using coarse-to-fine action proposals.

**Main Contributions and Findings:**  
Aligns predicted actions with hierarchical match structure (e.g., halves, goals, substitutions).

**Limitations and Future Directions:**  
Highly engineered for soccer; generalization needs further validation.

**Relevance to Primary Paper:**  
Innovative summarization approach combining decision- and comprehension-layer features.

**[160] Chai & Wang, 2022 – Deep Vision Multimodal Learning: Methodology, Benchmark, and Trend**

<https://arxiv.org/abs/2205.13258>

**Key Concepts and Taxonomy:**  
Reviews multimodal learning strategies combining video, audio, and text, with applications in sports understanding.

**Main Contributions and Findings:**  
Highlights challenges in multimodal alignment, representation, and fusion.

**Limitations and Future Directions:**  
Lacks sports-specific implementations; theoretical overview with few benchmarks.

**Relevance to Primary Paper:**  
Frames future research opportunities in multimodal comprehension and summarization tasks.

**[161] Wang et al., 2018 – Weakly Supervised Learning for Sports Video Event Detection**

<https://dl.acm.org/doi/10.1145/3240508.3240546>

**Key Concepts and Taxonomy:**  
Applies multi-instance learning and weak supervision to detect key sports actions without frame-level annotations.

**Main Contributions and Findings:**  
Reduces reliance on costly annotations and enables scalable event detection across broadcast sports.

**Limitations and Future Directions:**  
Detection precision is lower than fully supervised methods; needs improved temporal resolution.

**Relevance to Primary Paper:**  
Supports comprehension tasks involving spotting actions using limited supervision.

**[162] Yu et al., 2020 – Self-Supervised Representation Learning for Tactical Pattern Discovery in Soccer**

<https://arxiv.org/abs/2005.07412>

**Key Concepts and Taxonomy:**  
Leverages self-supervised contrastive learning to discover tactical formations and transitions in team sports.

**Main Contributions and Findings:**  
Enables unsupervised learning of team formations, revealing tactical transitions in soccer.

**Limitations and Future Directions:**  
Focused on top-down data; player pose dynamics not deeply integrated.

**Relevance to Primary Paper:**  
Contributes to decision-layer modeling of team strategies via learned latent features.

**[163] Pan et al., 2020 – Sports Video Captioning via Attentive Motion and Group Relationship Modeling**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/9046049>

**Key Concepts and Taxonomy:**  
Introduces a hierarchical model that captures motion and group activity for generating descriptive captions of sports clips.

**Main Contributions and Findings:**  
Improves coherence and relevance in video-to-text generation using attention-based motion modeling.

**Limitations and Future Directions:**  
Requires dense annotations; challenges with long-form sports content remain.

**Relevance to Primary Paper:**  
Enhances comprehension layer with video captioning tied to team dynamics.

**[164] Yu et al., 2018 – Fine-Grained Video Captioning for Sports Narrative**

<https://arxiv.org/abs/1805.09935>

**Key Concepts and Taxonomy:**  
Focuses on generating temporally and semantically aligned captions for sports scenes using multi-level modeling.

**Main Contributions and Findings:**  
Captures subtle transitions in sports routines; facilitates richer narrative generation.

**Limitations and Future Directions:**  
Limited to individual sports (e.g., diving); harder to generalize to team dynamics.

**Relevance to Primary Paper:**  
Contributes to sports-specific captioning under the comprehension task taxonomy.

**[165] Wang et al., 2018 – Double-Team Strategy Learning in NBA with Deep Reinforcement Learning**

<https://www.google.com/search?q=https://ieeexplore.ieee.org/document/8489373>

**Key Concepts and Taxonomy:**  
Applies deep RL to identify optimal moments for double-teaming a player in basketball.

**Main Contributions and Findings:**  
Shows that learned strategies can outperform handcrafted rules in simulated NBA games.

**Limitations and Future Directions:**  
Depends on synthetic reward functions and structured positional data.

**Relevance to Primary Paper:**  
Exemplifies decision-layer modeling with tactical reinforcement learning.

**[166] Luo, 2020 – Inverse Reinforcement Learning for Valuing Actions and Players**

<https://arxiv.org/abs/2007.03058>

**Key Concepts and Taxonomy:**  
Uses inverse reinforcement learning to assign value to player actions based on observed team outcomes.

**Main Contributions and Findings:**  
Quantifies action utility in complex, interdependent game settings (e.g., basketball).

**Limitations and Future Directions:**  
IRL models can be unstable without strong priors; interpretability remains an issue.

**Relevance to Primary Paper:**  
Bridges perception and decision tasks by assigning interpretable value to actions.

**[167] Liu & Schulte, 2018 – Deep RL in Ice Hockey for Context-Aware Player Evaluation**

<https://arxiv.org/abs/1802.04938>

**Key Concepts and Taxonomy:**  
Evaluates player contributions in hockey using deep RL models that consider temporal and spatial context.

**Main Contributions and Findings:**  
Performs nuanced player valuation by tracking puck location, possession, and game state.

**Limitations and Future Directions:**  
Limited by the quality of tracking data and action sparsity.

**Relevance to Primary Paper:**  
Expands the decision layer into player evaluation under uncertainty.

**[168] Yanai et al., 2022 – Q-Ball: Modeling Basketball Games with Deep Reinforcement Learning**

<https://www.google.com/search?q=https://www.jstage.jst.go.jp/article/pjsai/JSAI2022/0/JSAI2022_2Pin1-39/_article/-char/ja/>

**Key Concepts and Taxonomy:**  
Trains agents using deep RL to simulate basketball gameplay and assess strategy efficiency.

**Main Contributions and Findings:**  
Models offensive and defensive dynamics jointly for high-fidelity simulations.

**Limitations and Future Directions:**  
Hard to generalize to open-world or low-structure games.

**Relevance to Primary Paper:**  
Advances strategic decision-making simulation using learned policy networks.

**[169] Lillicrap et al., 2015 – Continuous Control with Deep Reinforcement Learning**

<https://arxiv.org/abs/1509.02971>

**Key Concepts and Taxonomy:**  
Presents the DDPG algorithm for continuous action control in reinforcement learning.

**Main Contributions and Findings:**  
Enables fine-grained, real-time control in simulated sports environments.

**Limitations and Future Directions:**  
Struggles with exploration and sparse reward settings in sports.

**Relevance to Primary Paper:**  
Foundational for RL-based models used in tactical and motor control in sports.

**[170] StatsPerform – Optical Tracking for Football Performance Analysis**

**Key Concepts and Taxonomy:**  
Commercial platform providing high-resolution optical tracking data for football analytics.

**Main Contributions and Findings:**  
Offers player-level, ball-level trajectory data used for coaching, scouting, and prediction.

**Limitations and Future Directions:**  
Access and cost barriers limit widespread academic adoption.

**Relevance to Primary Paper:**  
Essential for data-driven perception and decision modeling across sports AI.

**[171] Second Spectrum – Player Tracking Technology**

**Key Concepts and Taxonomy:**  
Commercial-grade optical tracking system widely used in professional sports for capturing fine-grained positional data.

**Main Contributions and Findings:**  
Enables high-resolution player and ball movement analysis through computer vision and sensor fusion.

**Limitations and Future Directions:**  
Proprietary and expensive; limited public datasets.

**Relevance to Primary Paper:**  
Essential for perception and forecasting models based on spatiotemporal data.

**[172] Wei et al., 2016 – Forecasting the Next Shot Location in Tennis**

<https://www.google.com/search?q=https://www.researchgate.net/publication/313410582_Forecasting_the_Next_Shot_Location_in_Tennis>

**Key Concepts and Taxonomy:**  
Uses fine-grained spatiotemporal tracking data to predict shot location in tennis using SVMs and handcrafted features.

**Main Contributions and Findings:**  
Establishes baseline accuracy for shot prediction using historical position and motion patterns.

**Limitations and Future Directions:**  
Limited by model capacity and non-adaptive to player behavior shifts.

**Relevance to Primary Paper:**  
Supports early decision-layer research in play forecasting using structured data.

**[173] Fernando et al., 2019 – Memory-Augmented Deep Generative Models for Shot Forecasting**

<https://www.google.com/search?q=https://openreview.net/forum%3Fid%3DHJxB2sRcFQ>

**Key Concepts and Taxonomy:**  
Introduces a memory-augmented architecture to generate probabilistic forecasts of shot outcomes in tennis.

**Main Contributions and Findings:**  
Outperforms deterministic baselines; captures multimodal uncertainty in player decisions.

**Limitations and Future Directions:**  
Focused on tennis; generalization to team sports is complex.

**Relevance to Primary Paper:**  
Enhances play forecasting with advanced generative modeling.

**[174] Wei et al., 2016 – Predicting Shot Outcomes in Tennis Using Style and Context Priors**

<https://ieeexplore.ieee.org/document/7552504>

**Key Concepts and Taxonomy:**  
Combines player-specific tendencies and game context to predict shot outcomes in tennis.

**Main Contributions and Findings:**  
Accurately forecasts winning probabilities and risk-reward dynamics based on historical data.

**Limitations and Future Directions:**  
Requires large-scale annotated data per player.

**Relevance to Primary Paper:**  
Blends perception and decision tasks in a unified forecasting framework.

**[175] Le et al., 2017 – Data-Driven Ghosting Using Deep Imitation Learning**

<https://arxiv.org/abs/1709.07185>

**Key Concepts and Taxonomy:**  
Proposes deep imitation learning to model defensive player behavior in NBA games.

**Main Contributions and Findings:**  
Generates realistic “ghost” players that simulate defense against unseen offensive strategies.

**Limitations and Future Directions:**  
Heavily dependent on structured tracking data.

**Relevance to Primary Paper:**  
Pioneering work in simulating team tactics through deep learning in decision tasks.

**[176] Power et al., 2017 – Risk-Reward Analysis of Soccer Passes**

<https://dl.acm.org/doi/10.1145/3097983.3098051>

**Key Concepts and Taxonomy:**  
Quantifies the effectiveness of soccer passes using spatiotemporal metrics and risk-reward models.

**Main Contributions and Findings:**  
Introduces data-driven methods to evaluate strategic decisions in soccer passing.

**Limitations and Future Directions:**  
Context modeling (e.g., defensive pressure) remains limited.

**Relevance to Primary Paper:**  
Advances tactical decision modeling by quantifying passing decisions.

**[177] Wang et al., 2022 – ShuttleNet: Stroke Forecasting in Badminton**

<https://arxiv.org/abs/2202.04692>

**Key Concepts and Taxonomy:**  
Combines rally progress and player style using position-aware neural networks to predict return strokes in badminton.

**Main Contributions and Findings:**  
Improves stroke forecasting by jointly modeling spatial and temporal player dynamics.

**Limitations and Future Directions:**  
Focused on two-player scenarios; generalization to team sports is limited.

**Relevance to Primary Paper:**  
Integrates perception and decision-making through predictive modeling.

**[178] Martins et al., 2021 – RSoccer: RL Framework for Small-Size Robot Soccer**

<https://arxiv.org/abs/2112.00412>

**Key Concepts and Taxonomy:**  
Develops a modular RL simulation environment for small-size robot soccer teams.

**Main Contributions and Findings:**  
Supports flexible agent training and benchmarking in controlled robotic soccer tasks.

**Limitations and Future Directions:**  
Limited visual realism; focused on low-dimensional inputs.

**Relevance to Primary Paper:**  
Framework for decision-layer training and benchmarking of RL sports agents.

**[179] Liu et al., 2019 – Emergent Coordination Through Competition**

<https://www.google.com/search?q=https://openreview.net/forum%3Fid%3DBkxYv2AcY7>

**Key Concepts and Taxonomy:**  
Studies emergent team behavior in competitive multi-agent settings using reinforcement learning.

**Main Contributions and Findings:**  
Shows how complex coordination can arise without explicit instructions via self-play.

**Limitations and Future Directions:**  
Still lacks realism for full-scale sports simulations.

**Relevance to Primary Paper:**  
Provides theoretical underpinnings for learning coordinated strategies in team sports.

**[180] Liu et al., 2021 – From Motor Control to Team Play in Simulated Football**

<https://arxiv.org/abs/2105.12196>

**Key Concepts and Taxonomy:**  
Integrates low-level motor control with high-level strategy learning in humanoid football environments.

**Main Contributions and Findings:**  
Demonstrates that DRL can bridge fine motor control with coordinated team play.

**Limitations and Future Directions:**  
High computational cost; transfer to real-world settings is limited.

**Relevance to Primary Paper:**  
Spearheads end-to-end decision systems in sports simulations with strategic depth.

**[181] Kurach et al., 2020 – Google Research Football: A Novel Reinforcement Learning Environment**

<https://github.com/google-research/football>

<https://arxiv.org/abs/1907.11180>

**Key Concepts and Taxonomy:**  
Introduces GFootball, a simulated soccer environment tailored for reinforcement learning (RL) agents.

**Main Contributions and Findings:**  
Provides complex, high-level soccer scenarios with diverse action spaces for multi-agent training and evaluation.

**Limitations and Future Directions:**  
Early versions lacked support for multi-agent cooperation and opponent diversity.

**Relevance to Primary Paper:**  
Foundational platform for decision-layer simulation and benchmarking in soccer AI.

**[182] Zhao et al., 2019 – Multi-Agent Learning in Team Sports Games**

<https://www.google.com/search?q=https://arxiv.org/abs/1908.10982>

**Key Concepts and Taxonomy:**  
Presents hierarchical multi-agent RL for modeling strategic behavior in team sports.

**Main Contributions and Findings:**  
Supports imitation and policy-based learning in simulated multi-agent settings with hierarchical planning.

**Limitations and Future Directions:**  
Coordination scalability and communication limits remain major hurdles.

**Relevance to Primary Paper:**  
Key contribution to the multi-agent simulation and policy-learning stream in sports decision-making.

**[183] Jia et al., 2020 – Fever Basketball: A Complex, Flexible Multi-Agent Environment**

<https://arxiv.org/abs/2012.06201>

**Key Concepts and Taxonomy:**  
Introduces an asynchronous RL environment for basketball with real-time multi-agent interaction.

**Main Contributions and Findings:**  
Supports diverse training paradigms: single/multi-agent, synchronous/asynchronous decision loops.

**Limitations and Future Directions:**  
Primarily simulation-based; challenges remain in transferring learned policies to real-world data.

**Relevance to Primary Paper:**  
Valuable for simulation, training, and evaluation of complex basketball strategies.

**[184] Li & Zhu, 2020 – WeKick: Google Football Competition Winning Model**

<https://arxiv.org/abs/2007.04296>

**Key Concepts and Taxonomy:**  
Describes the WeKick agent built for the Google Research Football competition using imitation and distributed training.

**Main Contributions and Findings:**  
Demonstrates state-of-the-art control and tactics for single-agent play in a soccer environment.

**Limitations and Future Directions:**  
Not generalizable to multi-agent tasks; lacks opponent modeling.

**Relevance to Primary Paper:**  
Benchmark agent influencing follow-up works like TiKick and TiZero in decision-layer environments.

**[185] Huang et al., 2021 – TiKick: Multi-Agent Football from Single-Agent Demonstrations**

<https://arxiv.org/abs/2104.03362>

**Key Concepts and Taxonomy:**  
Extends WeKick into a multi-agent football system using offline learning on replay data.

**Main Contributions and Findings:**  
Successfully replays full games using coordinated multi-agent policies derived from single-agent demos.

**Limitations and Future Directions:**  
Requires curated replay data; behavior diversity is constrained.

**Relevance to Primary Paper:**  
Pushes toward multi-agent generalization in sports game simulation.

**[186] Lin et al., 2023 – TiZero: Curriculum-Based Self-Improving Multi-Agent RL**

<https://arxiv.org/abs/2304.04565>

**Key Concepts and Taxonomy:**  
Proposes a curriculum-learning and self-play-based approach for learning soccer from scratch.

**Main Contributions and Findings:**  
Enables complex behaviors without imitation, outperforming prior systems like TiKick in robustness.

**Limitations and Future Directions:**  
High training cost and policy convergence challenges remain.

**Relevance to Primary Paper:**  
Advanced framework for autonomous multi-agent training in sports RL.

**[187] Yu et al., 2021 – The Surprising Effectiveness of MAPPO in Cooperative Games**

<https://arxiv.org/abs/2103.01955>

**Key Concepts and Taxonomy:**  
Evaluates MAPPO (Multi-Agent Proximal Policy Optimization) for team cooperation scenarios.

**Main Contributions and Findings:**  
Outperforms specialized algorithms in cooperative, multi-agent tasks including sports simulations.

**Limitations and Future Directions:**  
Struggles in highly non-stationary settings with dynamic team formations.

**Relevance to Primary Paper:**  
Key optimizer for decision-layer policy learning in team sports contexts.

**[188] Wen et al., 2022 – MARL is a Sequence Modeling Problem**

<https://arxiv.org/abs/2205.14953>

**Key Concepts and Taxonomy:**  
Reframes multi-agent RL as sequence modeling using transformers and autoregressive decoders.

**Main Contributions and Findings:**  
Achieves superior sample efficiency and generalization on cooperative tasks like soccer.

**Limitations and Future Directions:**  
Limited real-world deployment due to temporal resolution requirements.

**Relevance to Primary Paper:**  
Innovative cross-pollination of NLP and sports decision learning.

**[189] Ecoffet et al., 2021 – First Return, Then Explore**

<https://www.google.com/search?q=https://www.nature.com/articles/s41586-021-03262-x>

**Key Concepts and Taxonomy:**  
Proposes a novel RL exploration strategy that balances exploration with returnability.

**Main Contributions and Findings:**  
Achieves better coverage in sparse reward settings, such as exploration in simulated soccer.

**Limitations and Future Directions:**  
Sensitive to trajectory design and memory depth.

**Relevance to Primary Paper:**  
Enhances policy robustness and completeness in simulated sports exploration.

**[190] Tendulkar et al., 2020 – Feel the Music: Generating Dance for Input Songs**

<https://arxiv.org/abs/2010.02701>

**Key Concepts and Taxonomy:**  
Synthesizes human motion aligned with music using adversarial and sequence modeling.

**Main Contributions and Findings:**  
Generates high-fidelity, musically synchronized dance sequences with natural transitions.

**Limitations and Future Directions:**  
Focuses on aesthetics; lacks athletic intent modeling.

**Relevance to Primary Paper:**  
Illustrates motion generation potential for creative or performance sports (e.g., gymnastics, figure skating).

Bottom of Form