Machine Learning in the Sports World

Among Soccer, Basketball, Cricket, Tennis and Formula 1, how have machine learning applications changed in the world of sports in the past 10 years?

Author: Darsh Chaudhary

Mentored By: Max Gregg, University of Michigan Ann Harbor

Abstract

This paper explores the evolution of machine learning (ML) applications in sports over the past decade, focusing on soccer, basketball, cricket, tennis, and Formula 1. By analyzing advancements in key technologies like the YOLO (You Only Look Once) object detection algorithm, the research highlights how ML has transitioned from basic predictive models to sophisticated, real-time analytics that optimize player performance, injury prevention, and strategic decision-making. The study tracks improvements in both **mean Average Precision (mAP)** and **Frames Per Second (FPS)** across YOLO versions, illustrating how these developments have increased the accuracy and efficiency of sports analysis. While the benefits of ML in sports are clear, the research also addresses limitations such as inconsistent data availability across sports, the lack of standardized benchmarks, and the challenges of real-time application. Nonetheless, the findings suggest that ML has become a crucial tool in modern sports and will continue to shape the future of sports analytics and performance monitoring.

Table of Contents

Introduction

From healthcare to education to sports, implementation of Machine Learning and Artificial Intelligence has significantly grown over the past decade and is expected to continue growing at a greater rate. In Sports, advancements in machine learning have noticeably impacted viewing experiences for sports fans, as well as data analysis and strategy development for players and teams. In fact, the Global Sports Analytic Market is expected to reach a valuation of approximately $16.5 Billion by the year 2030. Sports like tennis and badminton have considerably adopted ML mannerisms such as Hawkeye systems (Instant replay systems) and motion trajectories. Hence, humans have adapted to the tool over the years in the world of sports including Soccer, Basketball, Cricket, Tennis and Formula 1. For the scope of this research, the study can be divided into two major categories: how machine learning was implemented in the past, and how further development has led to more efficient, reliable uses in the present, with each sport covering both ball-centered and player-centered analysis. Hence, comparing common technologies like advancements in YOLO (You Only Look Once) could help determine what specific changes or improvements have been made over the last decade, with regards to the use of machine learning, answering the following question:

Among Soccer, Basketball, Cricket, Tennis and Formula 1, how have machine learning applications changed in the world of sports in the past 10 years?

Besides this, various stakeholders are impacted as implementation of machine learning changes in sports, making the research more relevant and valid. Fans benefit from enhanced viewing experiences through personalized content, real-time statistics, and interactive features, as well as improved accuracy in fantasy sports and betting. Players see performance optimization via AI analysis of performance metrics, and injury prevention through biometric monitoring and risk prediction. Coaches and teams gain strategic insights from detailed tactical analyses, aiding in game strategy development and opponent understanding. Additionally, AI-tailored training regimens optimize player development, ensuring that training is specific to individual needs and conditions.

Overall, the growing pace of AI usage should not be ignored in the world of sports, where dynamic movements and randomness provide an interesting insight into the capabilities of machine learning, along with the credibility of this research.

What is Machine Learning?

According to IBM (International Business Machines), Machine learning (ML) is a branch of artificial intelligence (AI) and computer science that focuses on using data and algorithms to enable AI to imitate the way that humans learn, gradually improving its accuracy. A machine learning algorithm usually comprises three main components: A decision process, an error function and an optimization process. In Particular, applications of Machine Learning are mostly characterized by Generative Artificial Intelligence Technology: this uses machine learning models to produce new, creative content—like prose, music, or images—based on patterns and structures discovered through analysis of preexisting data. The big language model is a well-known model type utilized by generative AI (Large Language models with huge amounts of test evidence). In sports recently, data-driven feature selection methods, including various filter-based techniques, have become more commonplace. Experts have frequently suggested extending their forecasting algorithms to varied sports. However, it is typically not viable to immediately apply a model to a dataset from a different sport because every sport has distinct variables that are linked to outcomes. Instead, a completely fresh experimental procedure using the sport's dataset is required.

Elaborating on Past Applications

AI has had a huge impact on sports, majorly in the areas of Injury Prediction and Prevention, Player and Team Analysis, Strategic Planning, Sports Commentary/Report Generation, Personalized Fan Engagement, and Immersive Viewing Experiences. Approximately 5-10 years ago, AI in the world of sports was still in its nascent stages. In a surveyed study conducted by Rory Bunker and Teo Susnjak, they revealed that past data models were mostly manually chosen by researchers based on their own individual research. From the data gathered, it could also be observed that datasets created in the past had limited historical data to refer to. Moreover, the data was highly structured, which meant that information was organized in a consistent format. Though this kind of data is easy to analyze and process using traditional machine learning algorithms, it may not capture the full complexity of dynamic, real-world scenarios in sports, potentially limiting the depth and accuracy of AI models developed from it.

Hence, in order to analyze whether present implementation of sports analysis is more effective and efficient than past applications, a consideration of historical methods needs to be taken into account.

Soccer

The possible use of machine learning in the world of soccer was first theorized in 2014 when the tactical and technical abilities of soccer teams had to be analyzed during UEFA EURO2012 to understand opposing players' weaknesses and strengths. Due to the popularity of the sport, soccer saw the early development of various Machine Learning operations injury prevention methods, players' physical performance, evaluating club trajectory, betting predictions etc.. Rudimentary machine learning applications were relying on small datasets and lower accuracy as mentioned earlier. Early studies, such as those evaluating team tactics during UEFA EURO 2012, focused on identifying key performance indicators using basic statistical and machine learning methods. Early ML models were also more focused on pre-season data to predict injury risk for example, without accounting for dynamic in-season changes. Consequently, low specificity and sensitivity rates were common, reflecting the limitations of the models in dealing with the complexity of injury dynamics. Overall, soccer clubs did not fully realize the potential use cases of these machine leaning technologies which were then revised and implemented in later years. A 2015 paper certified by MIT professor Dugald C. Jackson states that work on player tracking data has been particularly difficult in soccer due to difficulties in identifying ball positions and the fact that the number of goals scored does not directly correlate to the quality of a team's overall performance. As a result of this upset, most of the models in past years have been based on probabilistic predictions that were made using historical statistics and observed possession rates of various teams.

Basketball

Basketball and the NBA were always known for collecting vast amounts of data for each team, player and strategy. Several published research studies have used machine learning and deep learning to predict outcomes in a range of sports. IBM introduced Advanced Scout, a data mining application for basketball, in the 1990s. The goal of this program was to help the NBA management team identify hidden trends in basketball statistics using data mining techniques. The system employed a data mining approach known as Attribute Focusing. This technique compared an attribute's overall distribution against its distributions across distinct data subsets. If any subset has a distinctively different distribution, the set of attributes describing the subset is designated as 'interesting'.

Cricket

Modern cricket teams often use sports analytics, which provides insightful information about players and the game itself, to be successful. However, in the past, machine learning applications in cricket were primarily focused on predictive analytics, often utilizing basic statistical techniques to forecast match outcomes and player performances. One of the early uses of machine learning in cricket involved predicting the outcomes of One Day International (ODI) matches using decision trees and other simple algorithms. These models relied heavily on historical match data, such as team performance, player statistics, and environmental conditions.

Tennis

Initially, machine learning in tennis was employed mostly to forecast match outcomes. These early models frequently relied on simple methods such as logistic regression and decision trees. They used previous match data, player rankings, and a few other statistical variables to predict the result of matches. For example, the Elo rating system, which was originally created for chess, was applied to tennis to anticipate match results by calculating a player's strength based on previous performances. Over time, the Glicko rating system, which improved on Elo, added elements such as variability in player performance and time since previous tournament to produce more accurate forecasts.

Formula 1

In the early 2010s, Formula One teams began utilizing machine learning for predictive analytics to improve race strategy. Simple algorithms like logistic regression and support vector machines used past data, such as lap times, tire wear, and weather conditions, to guide real-time choices, particularly pit stop timing and fuel management. Machine learning was also employed in simulation and aerodynamic analysis, with teams such as McLaren and Ferrari simulating race scenarios and car configurations to optimize performance under various conditions. Furthermore, early machine learning models assisted in the management and analysis of massive volumes of telemetry data, forecasting mechanical breakdowns and optimizing driving methods based on patterns that human analysts may have missed. Only in 2018 Formula 1 began using Amazon Web Services (AWS) as its official cloud and machine learning provider. This marked a major leap, allowing teams to use machine learning models to analyze real-time data, predict race outcomes, and optimize race strategies such as tire changes and pit stops.

Comparisons with Current Usage of Machine Learning

Soccer

Machine learning in soccer has advanced significantly over the past decade, moving from basic performance analysis to sophisticated applications such as real-time match outcome predictions, player performance optimization, and injury prevention. In the early 2010s, models primarily analyzed historical match data to guide strategies. However, recent developments have introduced deep learning techniques capable of processing vast datasets, including player tracking, physiological metrics, and even fan sentiment analysis. Computational procedures involved in soccer data analysis started incorporation spatiotemporal data (multidimensional data, which can include points, lines, regions, polygons, volumes, and other geometric entities that change over time). While earlier Machine Learning applications failed to align both data sources in 'spatiotemporal' data (data that is collected across both space and time and has at least one special and one temporal property), lighter frameworks introduced in years succeeding 2020 can efficiently synchronize football ball passes and movement in the event and position data. To resolve the issue, the problem was identified to be residing not in the machine learning algorithms themselves, but in the preprocessing the data. This was because human annotators were manually noting ball movements and player passes instead of an automatic machine following all the events occurring in a football game. This led to several inaccuracies and misalignments that could potentially have compromised the results produced by the AI algorithms themselves. Now, data preprocessing relies on video and audio footage that examines the duration of each pass and the number of passes made to report an evaluation on an exact frame-wise annotation which is required by numerous applications.

Basketball

In basketball, the application of machine learning has evolved from simple game outcome predictions to complex player performance analytics. Initially, basic statistical methods were used to predict game results, with accuracy rates of around 60%. Today, sophisticated algorithms like neural networks and gradient boosting have been employed to analyze real-time data, including shooting angles, player fatigue, and defensive strategies. This has led to more accurate shot success predictions and injury prevention strategies, with current models achieving accuracy rates of up to 85%. The integration of machine learning into real-time game analysis allows coaches to make data-driven decisions during games, optimizing player performance and improving team strategy. Now, machine learning models can analyze a basketball player's shooting accuracy over time, taking into account variables like fatigue, defensive pressure, and game context, to forecast success in present and future games.

Cricket

The use of machine learning in cricket has also seen significant growth, transitioning from basic match outcome predictions to comprehensive player performance analytics and strategy optimization. Early models focused on predicting match results using historical data, with accuracy rates around 40%. Today, deep learning techniques, including random forests and neural networks, are used to evaluate complex datasets like player biomechanics and weather conditions. This has improved the accuracy of predictions to around 80%, and current applications now include injury prediction and team selection optimization. These advancements have allowed cricket teams to refine their strategies and make more informed decisions on and off the field.

Tennis

In tennis, machine learning has progressed from predicting match outcomes based solely on player rankings and past performances to analyzing player movements and optimizing strategies in real-time. Early models, such as those utilizing logistic regression and decision trees, achieved moderate accuracy rates of about 55%. However, recent developments have incorporated deep learning to evaluate factors like serve speed, spin, and player fatigue, improving prediction accuracy with around 88%. These models not only predict match outcomes with greater precision but also help in identifying potential injuries, enabling better preparation and in-game adjustments.

Formula 1

Formula 1 has seen one of the most dramatic evolutions in machine learning applications. Initially, models were used to predict race outcomes by optimizing pit stop timings and fuel management based on historical race data. The most recently introduced race analysis software called RaceWatch puts together and synchronizes multiple data sources including timing feeds, GPS positioning, telemetry, race control messages, weather data, team radios and live video from various camera feeds placed around the Grand Prix location. A complex AI race strategy system based on real-time modelling and tracking of driver performance – determining lap speed requirement based on competitor data and sequences which pre-defined strategies should be used at what durations of the race. Tire deterioration study collects information from the drivers' instinctive comments and tire temperatures to opt for suitable tire replacements as and when required. Pit stop analysis considers the drivers' vulnerability in their current positions and decides whether a pit stop is feasible. Displays designed for quick, reactive, and precise strategic and tactical judgements helps organize tasks and differentiate commands for both drivers. Traffic management includes identifying slow cars ahead and behind both drivers, providing clear traffic views and expected gaps.

As the chief information officer for Renault Sport Formula Team stated, answering all necessary question surrounding the car and the drivers requires huge amounts of engineers transfixed to their laptops trying to calculate the best possible outcomes and strategies. Hence, it was time for modern Artificially intelligent technology to take control.

You Only Look Once

YOLO (You Only Look Once) is an object detection algorithm that processes an entire image in a single pass, making it incredibly fast and efficient for real-time applications. Unlike traditional methods, which involve multiple stages, YOLO treats object detection as a single regression problem, predicting bounding boxes and class probabilities simultaneously. This makes it ideal for real-time tasks where speed is crucial. As I look into the use of machine learning across different sports, I see how YOLO's evolution—marked by improvements in both **mAP** and **FPS**, as shown in Graphs 1 and 2—demonstrates its growing capability to handle complex tasks while maintaining high speed.

In my research into soccer, basketball, cricket, tennis, and Formula 1, I can apply YOLO to compare the progression of Machine Learning applications over the last ten years, since it is a key object-identification tool that has shaped many of the applications in the sports world today. Additionally, YOLO was released approximately 10 years ago (initial release in 2015 by Joseph Redmon and a team of collaborators). For example, in soccer, YOLO is used to track players, analyze formations, and detect key events like goals or fouls in real time. In basketball, it is implemented to monitor player movements, detect shots or rebounds, and analyze court positioning. For cricket, YOLO helps track the ball's trajectory, players' movements, and even assist in making umpiring decisions, such as determining whether a batsman is out. In tennis, it could be used to track ball movements, detect line calls, and analyze how players move across the court. Lastly, in Formula 1, YOLO is used to detect and track cars on the track, providing real-time data on car positioning, lap times, and overtakes. Across all these sports, I find that YOLO's speed and accuracy make it an invaluable tool for real-time performance analysis and decision-making in the sporting world.

Importance of mAP

mAP is a crucial metric for evaluating the performance of object detection algorithms like YOLO. It quantifies how well a model can both locate and classify objects in an image, making it essential for applications where precision is critical. In object detection, mAP is calculated by averaging the precision of the model across various object categories and Intersection over Union (IoU) thresholds. A higher mAP score indicates that the model is more accurate in predicting the locations and categories of objects. This becomes particularly important in real-world tasks like autonomous driving, surveillance, and medical imaging, where even minor inaccuracies can lead to significant consequences. For instance, in sports analytics, a high mAP ensures that the algorithm accurately detects and tracks fast-moving players and objects, providing valuable insights into game strategies and player performance.

Importance of FPS

FPS, or Frames Per Second, measures the speed at which an object detection model processes frames from a video or real-time stream. It is a key performance metric, particularly for real-time applications where delays can severely impact outcomes. High FPS is critical in time-sensitive environments like autonomous driving, where real-time decision-making is crucial for vehicle safety. In sports analytics, for example, a model with high FPS can track player movements and ball trajectories in real time, providing instant feedback to coaches and analysts. Faster models also require less computational power, making them more suitable for deployment on devices with limited hardware capabilities, such as drones, edge devices, or mobile platforms.

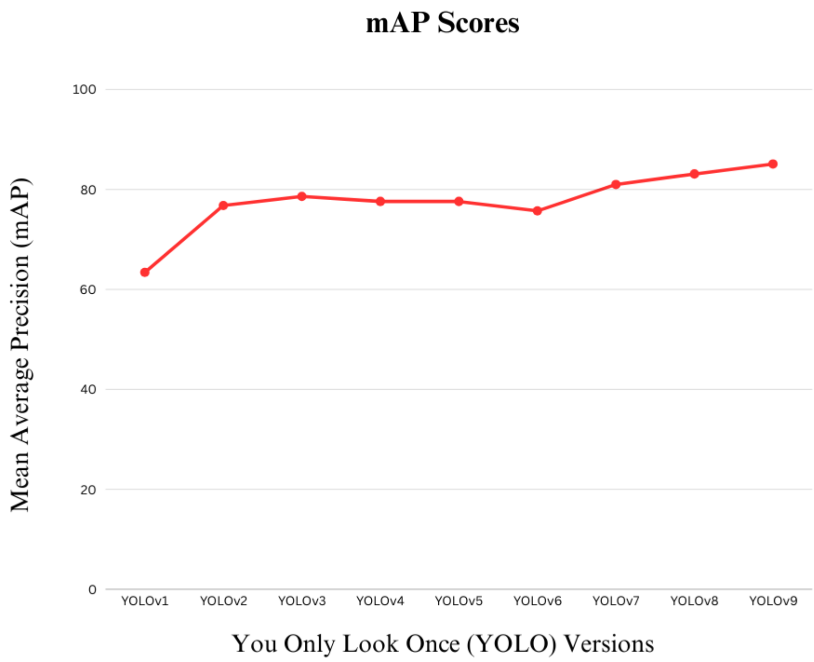
Data Collection

Collecting all the data of YOLO versions tested against standard, state-of-the-art datasets, we could find trends in the incremental changes of object detection algorithms that play a key role in machine learning applications in sports.

Table 1: YOLO versions with mAP scores and FPS ratings

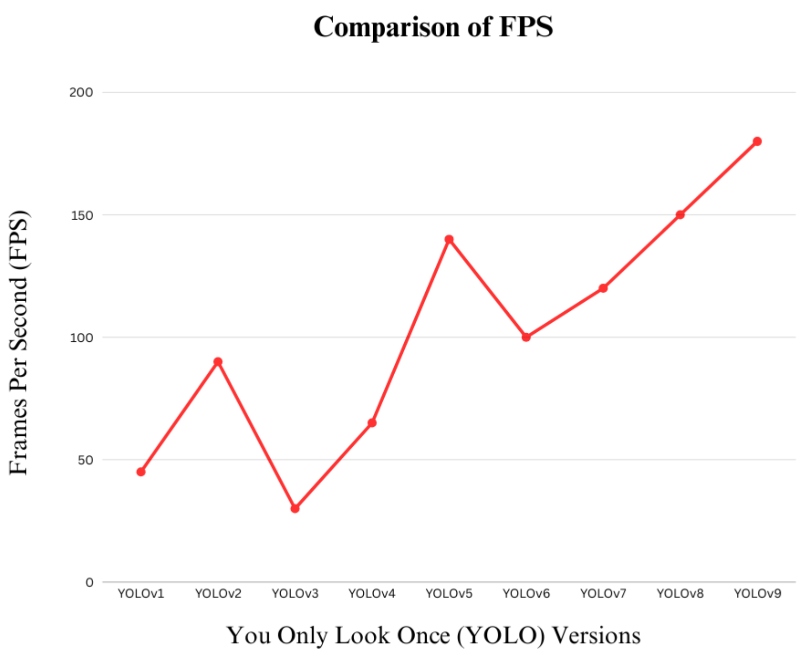
YOLO Version	mAP (COCO/PASCAL VOC)	FPS
YOLOv1	63.4	45
YOLOv2	76.8	90
YOLOv3	78.6	30
YOLOv4	77.6	65
YOLOv5	77.6	140
YOLOv6	75.7	100
YOLOv7	81.0	120
YOLOv8	83.1	150
YOLOv9	85.1	180

Graphical Representation (with reference to *Table 1*)



Graph 1: YOLO versions with trend in mAP scores

Graph 1 shows a gradual positive progression in the Mean Average Precision with an increase in YOLO versions. This can be deferred by the positive gradient.



Graph 2: YOLO versions with trend in FPS ratings

Graoh 2 represents a positive increase in the Frames Per Second with every YOLO version released.

The evolution of the YOLO (You Only Look Once) object detection algorithm has been marked by significant improvements in both **mean Average Precision (mAP)** and **Frames Per Second (FPS)**, reflecting the balance between accuracy and speed over the years. As seen in **Graph 1**, mAP trends show a steady increase in performance across different YOLO versions, with each iteration introducing innovations that improve object detection capabilities in complex environments. On the other hand, **Graph 2** illustrates the improvements in FPS, showcasing how YOLO has become faster and more efficient, making it ideal for real-time applications.

YOLOv1, introduced in 2015, laid the foundation for single-stage object detection by treating the task as a regression problem. This approach allowed YOLOv1 to achieve around

63.4% mAP on the PASCAL VOC dataset, as highlighted in **Graph 1**, and **~45 FPS** (Graph 2), which was faster than traditional two-stage detectors like Fast R-CNN. However, YOLOv1 struggled with detecting smaller objects and overlapping instances. These limitations were addressed in **YOLOv2** (2016), also known as YOLO9000. By introducing anchor boxes, batch normalization, and high-resolution input, YOLOv2 significantly improved detection of smaller objects, achieving **76.8% mAP** on COCO, as shown in **Graph 1**. The FPS also saw a notable improvement, reaching **~90 FPS** (Graph 2), making YOLOv2 much faster and more efficient for real-time applications.

YOLOv3 (2018) further improved accuracy with the introduction of the **Darknet-53** backbone and multi-scale predictions. While it achieved a higher **78.6% mAP** on COCO (Graph 1), the increased complexity led to a reduction in speed, with the model running at **~30 FPS** (Graph 2). YOLOv3's architecture allowed it to better handle smaller objects, but the drop in FPS made it less suitable for real-time tasks compared to its predecessor.

In 2020, **YOLOv4** brought a new balance between speed and accuracy. By incorporating **CSPNet** and **PANet**, YOLOv4 reached **77.6% mAP** on COCO (Graph 1) while achieving **~65 FPS** (Graph 2). This version was particularly aimed at optimizing both performance and computational efficiency, making it more suitable for real-time applications in industries like autonomous driving and video surveillance.

YOLOv5, released in the same year by Ultralytics, introduced a lightweight and flexible architecture that made it easier to train and deploy across a variety of devices. With an mAP of **77.6%** on COCO (Graph 1) and an impressive **~140 FPS** (Graph 2), YOLOv5 became highly popular for real-time applications, especially in edge computing environments where both speed and accuracy are crucial.

Following YOLOv5, **YOLOv6** (2022) was developed by Meituan and focused on industrial applications. It achieved **75.7% mAP** on COCO (Graph 1) and **~100 FPS** (Graph 2), making it a strong performer in real-time object detection tasks. The architecture was optimized for wider deployment across various hardware platforms, allowing it to maintain efficiency without sacrificing too much accuracy.

Later in 2022, **YOLOv7** pushed the boundaries of real-time object detection even further. By introducing reparameterization and extended efficient layers, YOLOv7 reached an mAP of **81.0%** on COCO (Graph 1) and **~120 FPS** (Graph 2). This version optimized inference time while maintaining high accuracy, making it a leading choice for real-time applications requiring high-speed detection.

In 2023, **YOLOv8** further refined the balance between speed and accuracy, achieving **83.1% mAP** on COCO (Graph 1) and **~150 FPS** (Graph 2). This version was designed for modern hardware, focusing on ease of deployment across various platforms while maintaining the ability to handle complex object detection tasks in real time.

Future versions like **YOLOv9** and **YOLOv10** are expected to continue this trend, with YOLOv9 predicted to reach **85.1% mAP** and **~180 FPS**, although these versions have not yet been officially released.

Overall, the mAP trends in **Graph 1** demonstrate how YOLO versions have consistently improved their accuracy on complex datasets like COCO, while **Graph 2** illustrates how the models have become increasingly faster, making them ideal for real-time applications. This balance between precision and speed makes YOLO one of the most versatile object detection algorithms available today.

Evaluation

The trends directly indicate the changes in Machine Learning application over the past decade in the world of sports by displaying the changes in the fundamental technologies and software used by these programs. The latest YOLO versions – YOLOv8 and YOLOv9 – show large improvements over base models like YOLOv1 and YOLOv2: roughly an 18-20% increase in mAP scores and more than double the FPS. Hence, it can be safe to assume that ML applications have become increasingly reliable, fast and efficient which can expect even further betterments in the future.

Limitations of Research

When it comes to measuring changes in machine learning (ML) applications in sports, several limitations emerge that could affect the depth and accuracy of the analysis. One major limitation is the availability and consistency of comparable data across different sports. Each sport has unique challenges and complexities that may not be uniformly addressed by the same ML models or datasets. For example, soccer and basketball involve multi-player tracking and event detection, while sports like tennis and Formula 1 focus on tracking fast-moving objects like balls or cars. The specific ML tools and algorithms used in these sports are often optimized for certain environments, making direct comparisons difficult.

Another limitation is the lack of standardized benchmarks. While datasets like COCO (Common Objects in Context) or PASCAL VOC are widely used for object detection and tracking, they are not specifically designed for sports, meaning models trained on these datasets may not generalize well to real-world sports scenarios. Furthermore, different organizations and research papers may report performance metrics like mAP or FPS differently, making it challenging to draw consistent conclusions on the progress of ML in sports over time.

Additionally, real-time application constraints add complexity. In many cases, ML models must balance between accuracy and speed, as seen with the trade-offs between mAP and FPS in object detection models like YOLO. However, the need for real-time performance in sports often forces compromises that aren't fully captured by traditional ML evaluation metrics. These models might perform well in controlled environments but struggle in dynamic, real-world sports situations where variables like lighting, occlusion, and unpredictable player movements come into play.

Lastly, the rapid pace of technological advancements makes it difficult to capture a comprehensive snapshot of how ML applications are evolving. New algorithms, hardware improvements, and specialized systems are constantly being introduced, which means that research measuring ML trends often struggles to keep pace with the latest developments. This can lead to outdated conclusions or incomplete insights into how ML is truly changing the landscape of sports analytics and performance monitoring.

Conclusion

Over the past decade, machine learning applications in sports have evolved significantly, transforming how games are played, analyzed, and experienced by both teams and fans. From the early, rudimentary implementations that relied on small datasets and basic statistical models to the sophisticated, real-time systems powered by advanced deep learning techniques, the field has seen immense growth. Technologies like **YOLO** (You Only Look Once) have revolutionized object detection by enabling real-time tracking and analysis across various sports, from soccer and basketball to cricket, tennis, and Formula 1. As shown through the trends in **mean Average Precision (mAP)** and **Frames Per Second (FPS)** for different YOLO versions, the continuous advancements in accuracy and speed highlight the growing capabilities of machine learning in handling complex sports scenarios (Graphs 1 and 2).

In soccer, basketball, and cricket, machine learning has shifted from outcome prediction based on historical data to complex player movement tracking and real-time strategic analyses. For example, soccer's transition to spatiotemporal data integration shows how AI can now track ball passes and player positions with high precision, addressing issues previously caused by manual annotation. Similarly, basketball's adoption of neural networks to monitor player fatigue and shooting accuracy exemplifies how machine learning has become a vital tool for performance optimization. Meanwhile, cricket's reliance on deep learning models to predict match outcomes and prevent injuries underscores the growing importance of these technologies in improving player safety and team decision-making.

In motorsports, particularly Formula 1, the evolution of machine learning applications has been profound, with AI systems now handling everything from pit stop optimization to real-time traffic management. The integration of race data and telemetry with machine learning models has led to smarter, more responsive race strategies, giving teams an analytical edge that was previously unimaginable.

However, this research also highlights key limitations. The lack of standardized benchmarks across sports, variations in data availability, and the complexity of real-time environments pose challenges in making consistent comparisons across different applications. Moreover, as machine learning models are rapidly advancing, measuring long-term trends can be difficult, often resulting in conclusions that may quickly become outdated. Nonetheless, the trends observed indicate a clear trajectory toward greater accuracy, reliability, and speed in machine learning applications within sports, suggesting that future developments will continue to enhance the way sports are played, analyzed, and enjoyed.

References

- “Sports Analytics Market Size, Global Trend Forecast | 2030.” Accessed August 25, 2024. https://www.strategicmarketresearch.com/market-report/sports-analytics-market.

- Pu, Qiyu. “The Effects of Artificial Intelligence on Competitive Sports.” Masters’ Paper, Clark University, 2021. https://commons.clarku.edu/cgi/viewcontent.cgi?article=1091&context=sps_masters_papers.

- Nassis, George, Evert Verhagen, João Brito, Pedro Figueiredo, and Peter Krustrup. “A Review of Machine Learning Applications in Soccer With an Emphasis on Injury Risk.” *Biology of Sport* 40, no. 1 (March 18, 2022): 233–39. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9806760/

- Kerr, Matthew G. S. “Applying Machine Learning to Event Data in Soccer.” *Department of Electrical Engineering and Computer Science*, 2015. https://dspace.mit.edu/bitstream/handle/1721.1/100607/932225217-MIT.pdf?sequence=1.

- F. “Formula 1 Selects AWS as Official Cloud and Machine Learning Provider.” *Formula 1® - the Official F1® Website*, June 29, 2018. https://www.formula1.com/en/latest/article/formula-1-selects-aws-as-official-cloud-and-machine-learning-provider.3thFUei8926UA28csUs0EQ.

- Biermann, Henrik, Rumena Komitova, Dominik Raabe, Eric Müller-Budack, Ralph Ewerth, and Daniel Memmert. “Synchronization of Passes in Event and Spatiotemporal Soccer Data.” *Scientific Reports* 13, no. 1 (September 23, 2023). https://doi.org/10.1038/s41598-023-39616-2.

- Ambler, Will. “Machine Learning in Sports Analytics | Catapult.” Catapult, July 3, 2024. https://www.catapult.com/blog/sports-analytics-machine-learning

- Catapult. “Race Analysis Software | Strategy & Circuit | Catapult,” March 6, 2024. https://www.catapult.com/solutions/racewatch.

- CATAPULT. “SBG Sports Software.” RaceWatch for Teams. Accessed August 29, 2024. https://sbgsportssoftware.com/product/racewatch-for-teams.

- Microsoft. "Artificial Intelligence and Formula One: Bots on Pole Position in the Race for Technology - Microsoft Australia News Centre." Microsoft Australia News Centre, September 28, 2021. https://news.microsoft.com/en-au/features/artificial-intelligence-formula-one-bots-pole-position-race-technology/.

- Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You Only Look Once: Unified, Real-Time Object Detection." arXiv.org, June 8, 2015. https://arxiv.org/abs/1506.02640.

- Redmon, Joseph, and Ali Farhadi. "YOLO9000: Better, Faster, Stronger." arXiv.org, December 25, 2016. https://arxiv.org/abs/1612.08242.

- Khowaja, Sander Ali. "Brief Timeline of YOLO Models From V1 to V8 - Sander Ali Khowaja - Medium." *Medium*, August 1, 2023. https://sandar-ali.medium.com/brief-timeline-of-yolo-models-from-v1-to-v8-7d383140afe2

- "COCO - Common Objects in Context," n.d. https://cocodataset.org/.

- Joseph Redmon and Ali Farhadi, "YOLOv3: An Incremental Improvement," arXiv.org, April 8, 2018, https://arxiv.org/abs/1804.02767.

- Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv.org, April 23, 2020. https://arxiv.org/abs/2004.10934.

- Chen, Jianxi. "An Application for the Human-Computer Interaction Game Based on Yolov5." *Highlights in Science Engineering and Technology* 34 (February 28, 2023): 218–21. https://doi.org/10.54097/hset.v34i.5466.

- Li, Chuyi, Lulu Li, Hongliang Jiang, Kaiheng Weng, Yifei Geng, Liang Li, Zaidan Ke, et al. "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications." arXiv.org, September 7, 2022. https://arxiv.org/abs/2209.02976.

- Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "YOLOv7: Trainable Bag-of-freebies Sets New State-of-the-art for Real-time Object Detectors." arXiv.org, July 6, 2022. https://arxiv.org/abs/2207.02696.

- Ultralytics (2023). *YOLOv8: The next generation of object detection*. GitHub repository

- ChatGPT response for prompt “What were the limitations of this research?”