

Data Dunk: Beeldanalyse en gebruik van Machine Learning algoritmes voor het identificeren van spelers en bijhouden van scores en statistieken in basketbalwedstrijden.

Bachelorproef, 2023-2024

Kyana Marckx

E-mail: kyana.marckx@student.hogent.be

Co-promotor: Nog te bepalen

Samenvatting

Deze studie onderzoekt de integratie van beeldanalyse en machine learning in basketbalwedstrijden voor het automatisch identificeren en vastleggen van spelers, objecten, scores en statistieken. Het heeft tot doel het werk van tafellofficials te stroomlijnen en geautomatiseerde analyses te bieden aan spelers en coaches in de snel evoluerende basketbalomgeving. De methodologie omvat het verzamelen van gegevens, data-annotatie, splitting en parsing van de geannoteerde data, het opstellen van modellen, training, model finetuning en front-end-ontwikkeling. Het gewenste resultaat, "Data Dunk", is een systeem dat automatisch scores, spelers, ballen, etc. in realtime kan identificeren. Het visualiseren van resultaten op een website of in een spreadsheet levert waardevolle inzichten op, waaronder gedetailleerde statistieken zoals individuele punten, rebounds, speeltijd en vrije worpen. Dit onderzoek vertegenwoordigt een technologische vooruitgang en een strategische stap naar een slimmere en efficiëntere basketbalervaring voor tafellofficials, spelers en coaches.

Keuzerichting: AI & Data Engineering

Slutelwoorden: Beeldanalyse, Basketbal, Machine Learning, Deep Learning

Inhoudsopgave

1	Introductie	1
2	State-of-the-art	2
3	Methodologie	2
3.1	Data Verzameling	2
3.2	Data Annotatie	3
3.3	Data Splitting en Parsing	3
3.4	Opstellen en Trainen van Model	3
3.5	Model Finetuning	3
3.6	Front-end Ontwikkeling	3
4	Verloop	3
4.1	Flowchart	3
4.2	Gantt-chart	3
5	Verwacht resultaat, conclusie	4
6	Baanbrekende innovaties	4
	Referenties	4

1. Introductie

In de snel evoluerende basketbalomgeving bieden technologieën zoals video-analyse, artificial intelligence, deep learning en machine learning nieuwe inzichten op een dieper niveau van het spel. Dit onderzoek richt zich op het optimaliseren van de taken van tafellofficials en het leveren van waardevolle, geautomatiseerde analyses voor spelers en coaches.

Handmatige processen voor het registreren van scores, spelersbewegingen en statistieken zijn tijdrovend en foutgevoelig. De kernvraag is hoe geavanceerde technologieën deze processen kunnen verbeteren.

Het onderzoek streeft naar een realtime werkend model dat automatisch scores, spelers, de bal en andere relevante elementen herkent en volgt tijdens basketbalwedstrijden. Met behulp van camera's en geavanceerde algoritmes wordt getracht een tool te ontwikkelen die tafellofficials ondersteunt, en spelers & coaches voorziet van gedetailleerde statistieken. Het resultaat is een systeem dat deze gegevens visualiseert op een website of exporteert naar een overzichtelijke spreadsheet, waardoor de waarde van verzamelde informatie wordt gemaximaliseerd.

De titel "Data Dunk" omarmt de belofte van diepgaande inzichten via beeldanalyse en het gebruik van machine learning en deep learning in basketbal. Dit onderzoek markeert niet alleen technologische vooruitgang, maar is ook een strategische stap naar een slimmer en efficiënter basketbalgebeuren voor alle betrokkenen.

In de loop van het onderzoek zullen er oplossingen geformuleerd worden op meer concrete en specifieke situaties, resultaten en doelen.

De volgende vragen zijn een goede start om dit onderzoek op te delen of verder uit te breiden:

- Welke specifieke uitdagingen en obstakels zijn er bij het identificeren van spelers in dynamische en snel veranderende situaties zoals een basketbalwedstrijd?
- In hoeverre is het mogelijk om onderscheid te maken tussen verschillende soorten scores, zoals vrije worpen, tweekunten en driepunten?
- Hoe kan de functionaliteit voor handmatige aanpassingen/bijstellingen van scores tijdens de live wedstrijd geïntegreerd worden om zo mogelijkheden voor verbetering te bieden?
- Welke bronnen en datasets zijn het meest geschikt voor het trainen van het model en het verbeteren van de nauwkeurigheid van de beeldanalyse?
- Hoe kan het model omgaan met variaties in verlichting, camerahoeken en andere omgevingsfactoren die typisch zijn aan live basketbalwedstrijden?

2. State-of-the-art

Artificial Intelligence (AI), en meer bepaald Computer Vision (CV), is tegenwoordig een hot topic in de Computer Science Industry. Computer Vision is een gebied in de computerwetenschap dat erop is gericht om met computers objecten en personen op foto's en video's te identificeren, aldus Azure (2024). Om meer in detail te gaan, zal er op videobeelden Object Tracking (OT), en op afbeeldingen Object Classification (OC) en Object Identification (OI) gehanteerd worden (Géron, 2023). De opkomst van AI heeft de manier waarop we basketbal analyseren ingrijpend veranderd. Het biedt vele voordelen in de sportwereld, maar er moet natuurlijk ook rekening gehouden worden met beperkingen en valkuilen. Artificial General Intelligence (AGI) maakt in 2023 zijn opmars. In tegenstelling tot de huidige AI, die gericht is op het verbeteren van Human Intelligence (HI), vertoont AGI menselijke cognitieve vaardigheden en creativiteit. Hoewel AI zoals ChatGPT¹ en Dall-E 3² snel evolueert, overtreft het nog steeds de ontwikkelingen in de sportsector. In plaats van te focussen op futuristische AI-toepassingen, is het essentieel om eerst bestaande AI-voordelen helemaal te benutten. Vragen over verantwoordelijkheid, beheer & ethiek zijn cruciaal voor de toekomst van AGI in sport. In dit stadium is het noodzakelijk om bestaande AI-voordelen grondig te begrijpen en te optimaliseren (Robertson, 2023).

¹<https://openai.com/blog/chatgpt>

²<https://openai.com/dall-e-3>

LinkedIn-artikelen zoals "The AI Revolution: Transforming the Future of Basketball" benadrukken dat AI niet alleen spelers en coaches, maar ook scheidsrechters voorziet van nauwkeurigere statistieken, waardoor het spel objectiever wordt beoordeeld. In zijn artikel bespreekt Finlay (2023) de impact van AI op basketbal.

Een ander LinkedIn-artikel, "Courting Success: AI-Driven Analysis for Game-Changing Basketball Insights" (Mahamoud, 2023), richt zich op geautomatiseerde analyse, waarbij AI wordt gebruikt voor betere training en slimmere spelstrategieën. Hier wordt de kracht van AI gecombineerd met menselijke intelligentie.

Bedrijven zoals Greenfly³, HomeCourt⁴ en SportsVisio⁵ bieden momenteel diverse praktische toepassingen aan van AI in basketbal(analyse).

Academisch onderzoek, zoals het Stanford-onderzoek over "Player Tracking and Analysis of Basketball Plays", geschreven door Cheshire e.a. (2015), laat zien hoe AI wordt ingezet om spelers te volgen en hun bewegingen op het veld te analyseren, met technieken zoals Histogram of Oriented Gradients (HOG).

Het IEEE-artikel van Yan e.a. (2023) over "Basketball Shooting Analysis" legt specifiek de nadruk op het gebruik van Convolutional Neural Networks (CNN) en Recurrent Neural Networks (RNN) voor gedetailleerde schietanalyse en voorspellingen.

Het kan daarnaast breder bekeken worden dan alleen de toepassingen in basketbal. In de bredere sportindustrie zijn er soortgelijke situaties met dergelijke producten en oplossingen. Het perfecte voorbeeld is het bedrijf MyPitch⁶; zij gebruiken AI in voetbalwedstrijden om game highlights te genereren voor de spelende club via hun platform, houden statistieken van de spelers bij zoals het aantal (succesvolle) passes, goals, ...

Dit alles samen biedt een kort overzicht van de huidige stand van zaken en toekomstige mogelijkheden van AI in basketbalanalyse, waarbij menselijke expertise wordt versterkt door geavanceerde (machine learning) technologieën.

3. Methodologie

3.1. Data Verzameling

Om een effectief model op te stellen en te trainen, is de eerste stap het verzamelen van voldoende data. Dit wordt bereikt door gebruik te maken van bestaande beeldfragmenten uit basketbalwedstrijden.

³https://www.greenfly.com/landing/ai-vision-demo-landing-page/?utm_medium=cpc&utm_source=adwords&utm_campaign=search_sports_vertical&gad_source=1

⁴<https://www.homecourt.ai>

⁵<https://sportsvisio.com>

⁶<https://www.mypitch.app>

Deze zullen afkomstig zijn van diverse bronnen zoals YouTube, officiële basketbalteam-sites (zowel nationaal als internationaal), zelf-gemaakte beelden, ... op verschillende posities, met diverse kwaliteiten en belichting. De beelden omvatten video-opnames waarvan korte frames worden genomen, evenals foto's.

De deliverable in deze fase is een reeks beelden die zal worden gebruikt tijdens de annotatie.

3.2. Data Annotatie

De verzamelde data wordt vervolgens geannoteerd door labels en bounding boxes toe te voegen aan de beelden, die verschillende elementen in het beeld aanduiden. Een annotatie-tool wordt gebruikt om de beelden te annoteren en de labels & bounding boxes op te slaan in een apart corresponderend bestand.

Het resultaat van de fase is een verzameling geannoteerde beelden die dienen als training data voor het model.

3.3. Data Splitting en Parsing

Na annotatie is het cruciaal om de data op te splitsen in een training set, een test set en een validatie set, waarbij de beelden met hun bijbehorende labels en bounding boxes in dezelfde set worden geplaatst. Om de data te gebruiken voor het trainen van het model wordt deze eerst nog geparsed naar een formaat dat bruikbaar is voor het model.

De uitkomst van deze fase is een training set, een test set en een validatie set in de correcte bruikbare formaten.

3.4. Opstellen en Trainen van Model

Met alle beschikbare data kunnen er verschillende modellen opgesteld en getraind worden om zo het meest effectieve resultaat te bekomen. Het model zal geëvalueerd worden aan de hand van verschillende metrics zoals de accuracy, Intersection over Union (IoU), precision & recall, Average Precision (AP), Mean Average Precision (mAP) en de F1 score⁷. Ook de snelheid van de modellen zal in acht genomen worden aangezien dit cruciaal is voor realtime statistieken. Uit de combinatie hiervan zal dan blijken welke van de geteste modellen de beste resultaten levert.

Dit leidt uiteindelijk tot 1 getraind finaal model.

3.5. Model Finetuning

Na de trainingsfase en evaluatieresultaten wordt de aandacht verlegd naar hyperparameter tuning. Dit is cruciaal om het geselecteerde model nog een stapje hoger te brengen.

In deze fase is het model gefinetuned en klaar voor gebruik en implementatie.

⁷<https://labeledyourdata.com/articles/object-detection-metrics>

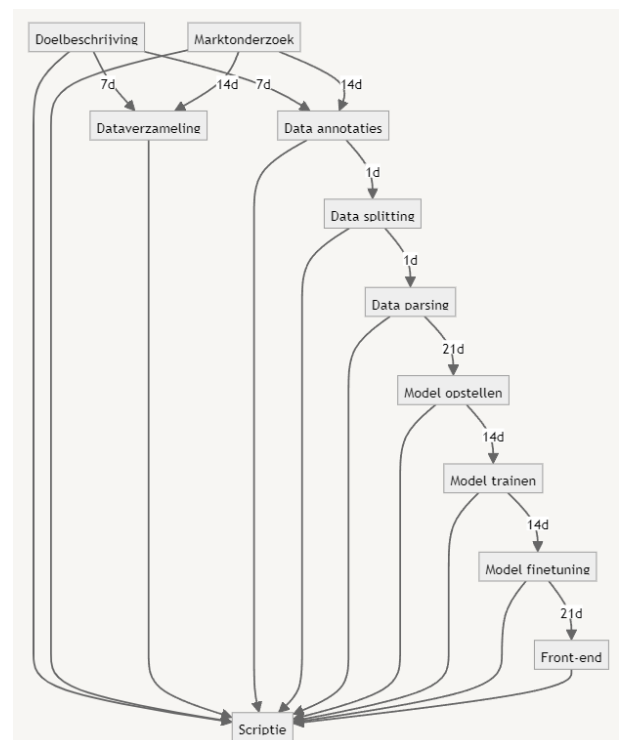
3.6. Front-end Ontwikkeling

Om makkelijk en gebruiksvriendelijk het model te kunnen benutten, zal er een front-end worden ontwikkeld. Deze front-end maakt het mogelijk om het model effectief in gebruik te nemen en resultaten te visualiseren. Daardoor kunnen tafelofficials, spelers en coaches tijdens en na basketbalwedstrijden informatie en statistieken verkrijgen. Opties voor de gebruikersinterface (GUI) zijn een zelf gemaakte website of een online spreadsheet. Indien er tijdens het onderzoek nog andere opties aan het licht komen, zullen deze ook overwogen worden.

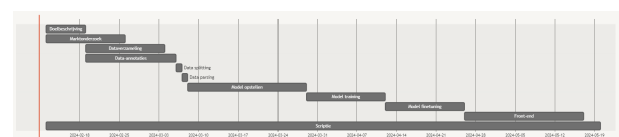
De uitkomst van deze fase is een front-end waarmee het model kan worden gebruikt en resultaten visueel kunnen worden weergegeven.

4. Verloop

4.1. Flowchart



4.2. Gantt-chart



5. Verwacht resultaat, conclusie

Als laatste stap om het succes van het uitgevoerde onderzoek te beoordelen, zal een front-end worden ontwikkeld waarmee het model kan worden toegepast en de resultaten kunnen worden gevisualiseerd. In deze visuele weergave is het mogelijk om zowel live als achteraf de basketbalwedstrijd te analyseren.

Live-analyse omvat het volgen van spelers, scheidsrechters, de bal, enz., terwijl achteraf gegenereerde statistieken de individuele punten van de spelers, rebounds, speeltijd per speler, aantal genomen vrije worpen en meer omvatten. Er zal ook gekeken worden in hoeverre het mogelijk is om (als tafelofficial) manueel de scores aan te passen indien er te veel of te weinig punten aan de ploegen zijn toegekend door het systeem.

Voor de analyse achteraf wordt een functionaliteit toegevoegd waarmee coaches en spelers een opname van de wedstrijd kunnen uploaden. Deze opname wordt verwerkt door het model en de resulterende analyse wordt gepresenteerd in de front-end. Dit stelt coaches en spelers in staat om te begrijpen waar eventuele struikelpunten in de wedstrijd lagen en te reflecteren op gemaakte tactische beslissingen.

De gecollecteerde statistieken en inzichten uit de analyses kunnen bijgevolg dienen als waardevolle input voor het aanpassen van de doorde-weekse trainingen en het verbeteren van prestaties in toekomstige wedstrijden.

6. Baanbrekende innovaties

Bij succesvolle afronding van dit project openen er nieuwe deuren voor innovaties. Dit zijn een aantal van de mogelijkheden:

- **Game Highlights Creatie:** het model kan worden uitgebreid om automatisch hoogtepunten van de wedstrijd te creëren waardoor boeiende momenten snel toegankelijk zijn voor de clubs en hun fans
- **Livestream Opties:** de implementatie van functies voor live streaming naar populaire platforms zoals YouTube, Twitch, Facebook, enz., waardoor fans de wedstrijd realtime en van thuis uit kunnen volgen
- **Referee Handgebaren Analyse:** door handgebaren van scheidsrechters in 'close sight' te monitoren, kan het model lichaamstaal interpreteren om verschillende soorten fouten te herkennen en deze aan specifieke spelers te koppelen

- **Uitgebreide Spelersactie Analyse:** het model kan worden getraind om gedetailleerde acties van spelers bij te houden zoals lopen, stilstaan, dribbelen, verdedigen, enz., om meer diepgaande en persoonlijke statistieken te genereren
- **Intelligent Area Prediction:** de toepassing van intelligente voorspellingen op de belangrijkste locaties op het speelveld vast te leggen via camera's waardoor specifieke zones en hotspots kunnen worden geïdentificeerd

Deze voorgestelde innovaties breiden de functionaliteiten van het huidige model uit en verbeteren de algehele basketbalervaring voor zowel fans als professionals.

Referenties

- Azure, M. (2024). *Wat is Computer Vision?* Verkregen februari 8, 2024, van <https://azure.microsoft.com/nl-nl/resources/cloud-computing-dictionary/what-is-computer-vision#objectclassificatie>
- Cheshire, E., Halasz, C., & Perin, J. K. (2015). *Player Tracking and Analysis of Basketball Plays* [onderzoeksrap.]. Stanford. Verkregen november 18, 2023, van https://web.stanford.edu/class/ee368/Project_Spring_1415/Reports/Cheshire_Halasz_Perin.pdf
- Finlay, D. (2023). *The AI Revolution: Transforming the Future of Basketball*. Verkregen november 18, 2023, van <https://www.linkedin.com/pulse/ai-revolution-transforming-future-basketball-david-finlay/>
- Géron, A. (2023, januari 20). *Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow*.
- Mahamoud, G. (2023). *Courting Success: AI-Driven Analysis for Game-Changing Basketball Insights*. Verkregen november 18, 2023, van <https://www.linkedin.com/pulse/courting-success-ai-driven-analysis-game-changing-guled-mahamoud/>
- Robertson, S. (2023). *Does AI have a role in the future of sport?* Verkregen februari 8, 2024, van <https://www.vu.edu.au/about-vu/news-events/news/does-ai-have-a-role-in-the-future-of-sport>
- Yan, W., Jiang, X., & Liu, P. (2023). *A Review of Basketball Shooting Analysis Based on Artificial Intelligence*. *IEEE Access*, 11, 87344–87365. <https://doi.org/10.1109/ACCESS.2023.3304631>