

Anàlisi de Futbol amb Aprendentatge Profund i Visió per Computador

Marc Serra Ortega, Pol Pugibet Martínez

Abstract—La indústria de l'anàlisi i les estadístiques del futbol està en ple creixement. Avui dia, la majoria dels equips professionals de futbol fan servir eines i sistemes sofisticats per analitzar el seu entrenament, els partits i els jugadors de diverses maneres. Cada vegada es basen més les decisions en les dades i aquests sistemes. Tanmateix, aquests sistemes són costosos i necessiten dispositius portàtils per determinar les posicions dels jugadors i altres mètriques de rendiment. Aquest treball presenta una solució a aquest problema amb l'ús de tècniques d'aprenentatge profund i visió per computador per extreure automàticament estadístiques i dades de futbol a partir de gravacions de vídeo d'una sola càmera, sense necessitat d'altres dispositius o sensors de recopilació de dades.

Keywords—Sport Analysis, YOLO, Computer Vision, Deep Learning, Homography.



1 INTRODUCCIÓ

Aquest projecte introduceix una solució inovadora en l'àmbit de l'anàlisi de futbol, que combina l'ús de l'aprenentatge profund i la visió per computador per a l'anàlisi detallada del joc. Amb l'objectiu de millorar la comprensió i l'avaluació dels partits de futbol, es proposa l'implementació de mecanismes de deep learning, específicament el model YOLO (You Only Look Once), per a la detecció precisa dels jugadors, la pilota i els punts de referència al camp. A partir d'aquests resultats, s'utilitzarà una tècnica de homografia per a la modelització del partit jugat, permetent així l'extracció de diverses mètriques rellevants, com ara el comportament dels jugadors, la possessió de la pilota, entre altres aspectes clau del joc. Aquest enfocament ofereix una nova perspectiva en la manera de comprendre i analitzar el futbol, que pot contribuir significativament al desenvolupament d'estratègies més efectives i a la millora del rendiment dels equips.

Com a base de dades, farem servir el conjunt de dades *football-players-detection Image Dataset* [1], el constituida de 255 imatges de perspectiva tàctica que abasten una àmplia secció del camp i mostren tots els jugadors amb moviments suaus, amb l'objectiu de simplificar el càlcul d'homografies i la creació del model final del partit.

2 ESTAT DE L'ART

Com hem comentat, volem realitzar una analítica de vídeo a gravacions de partits de futbol. Aquesta analítica contindrà les característiques esmentades anteriorment, i potser algunes extres que veurem a continuació.

2.1 Detecció precisa dels jugadors

Aquesta és la característica més bàsica que podem trobar en projectes com aquest. Es tracta d'identificar els jugadors (diferenciant l'equip) en el vídeo, tenint així la seva posició en cada moment, podent seguir així els seus moviments. També s'identificarà l'àrbitre i la pilota. Això es farà mitjançant YOLO, com és comú en projectes de l'estil [2].



Fig. 1: Exemple de detecció dels jugadors. [3]

2.2 Detecció precisa dels punts de referència

Aquesta característica consisteix en identificar punts de referència (és a dir, les línies del camp i els marges del mateix) per tal de calcular amb precisió les posicions dels jugadors a l'àrea de joc.



Fig. 2: Exemple de detecció de punts de referència. [4]

2.3 Homografia

Finalment, es posaran en comú les dues implementacions mencionades per a realitzar la homografia, i extreure d'aquí les mètriques corresponents [5].



Fig. 3: Exemple de resultat final, amb detecció de jugadors i homografia corresponent. [6]

2.4 Extres

També s'ha trobat algunes implementacions que, tot i no ser les principals que es vol implementar, són contemplades com possibles millores un cop finalitzades les característiques principals del projecte.

Aquests extres consisteixen en: fer ús de deepSORT per assignar un identificador a cada jugador, guardant aquest atribut fins i tot quan el jugador no aparegui a pantalla per quan torni a aparèixer [7]; identificació del dorsal dels jugadors [8]

3 PROPOSTA

La nostra proposta per al projecte es centra en l'arquitectura general de la implementació, enfocada en les característiques principals per a dur a terme una anàlisi exhaustiva de futbol. Aquesta anàlisi inclou la detecció dels participants en el partit (jugadors, àrbitres i pilota) i la identificació dels punts característics del camp de futbol, a partir dels quals es calcularà l'homografia per modelitzar el partit.

Tant per a la detecció dels participants del partit, com els punts característics del camp hem optat per utilitzar el model YOLO. Aquest model és molt versatil en la detecció de diferents objectes i té una gran capacitat de processar informació en temps real, fet que el converteix en una eina ideal per a l'aplicació d'aquest projecte en entorns de producció. Entrenarem dos models específics:

- Model de Detecció de Participants:** Centrat en la detecció de jugadors, àrbitres i pilota. Aquest model ens permetrà identificar i seguir aquests elements al llarg del partit.
- Model de Detecció de Punts Característics del Camp:** Dissenyat per detectar els punts clau del camp de futbol (per exemple, les línies de banda, els gols, el cercle central, etc.) que són essencials per al càlcul de la matriu d'homografia.

Un cop obtinguts els punts característics calculem l'homografia que ens permetrà transformar les coordenades de la imatge capturada per la càmera a coordenades del camp de futbol. Això és crucial per obtenir una representació precisa de les posicions relatives dels jugadors i la pilota en el camp.

Per a la classificació automàtica dels jugadors s'utilitzara el model d'aprenentatge no supervisat K-means que determinara a partir del color de la camiseta de cada jugador el equip al que pertany.

Amb tota la informació recollida, modelitzarem el partit de futbol. Aquesta modelització ens permetrà crear una representació detallada del joc, facilitant així l'anàlisi tàctica i estratègica. El projecte té com a objectiu proporcionar una plataforma des de la qual es puguin extreure estadístiques i informació clau que ajudin a entendre i millorar el rendiment dels jugadors i de l'equip. Entre les estadístiques que es poden generar es troben:

- Percentatge de possessió de la pilota per part de cada equip.
- Distàncies recorregudes per cada jugador.
- Zones del camp més utilitzades per cada equip.
- Velocitats màximes i mitjanes dels jugadors.

4 EXPERIMENTS, RESULTATS I ANÀLISI

El projecte s'ha desenvolupat en dues fases principals. Inicialment, vam implementar les eines necessàries per al càlcul de l'homografia, experimentant amb diverses tècniques tradicionals de visió per computador. En particular, vam provar una combinació de preprocessament d'imatges i l'algoritme SIFT per correlacionar els punts característics del camp en cada fotograma amb els punts corresponents en el pla del camp final, on es modelitzaria el partit de futbol. No obstant això, aquesta implementació no va donar els resultats esperats. Això ens va portar a optar per l'ús d'un model específic de YOLO per detectar els punts característics del camp, facilitant així el càlcul posterior de la matriu d'homografia.

Per a la detecció dels jugadors, àrbitres, pilota i punts característics, vam implementar el model YOLOv1 des de zero, basant-nos en l'article original, *You Only Look Once: Unified, Real-Time Object Detection* [10], i utilitzant PyTorch. Es va implementar l'arquitectura convolucional descrita en l'article, juntament amb les funcions de cost corresponents. Addicionalment, es van desenvolupar funcions de mètriques i visualització específiques per a aquest model. Vam provar la capacitat del model per detectar diferents elements en imatges a través de l'overfitting, comprovant que funcionava eficaçment. No obstant això, la limitada capacitat computacional i la manca de temps per entrenar el model des de zero, així com la precisió reduïda de YOLOv1 en la detecció d'elements petits, van fer que desistíssim d'utilitzar-lo. Malgrat aquests contratemps, aquesta implementació ens va permetre comprendre detalladament el funcionament del model YOLO i aprofundir en l'aprenentatge de les xarxes convolucionals.

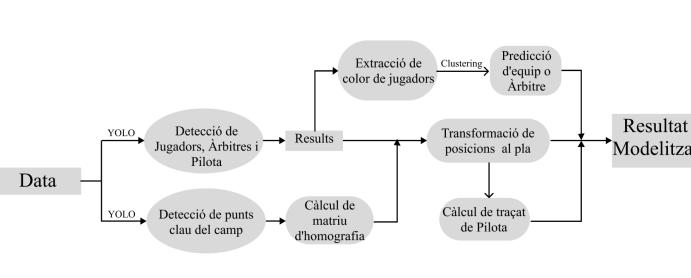


Fig. 4: Proposta de resolució del projecte

Adicionalment es vol implementar unes funcionalitats extres:

• Seguiment (Tracking):

Implementarem tècniques de seguiment per mantenir la identificació dels jugadors i de la pilota a mesura que es desplacen pel camp. Això ens permetrà obtenir informació específica de cada jugador, com la seva posició i moviment en el temps i analitzar la possessió de la pilota per part de cada equip.

• Càlcul del Flux Òptic (Optical Flow):

Incorporarem el càlcul del flux òptic per determinar la velocitat de cada jugador durant el partit. Aquesta tècnica ens permetrà mesurar els desplaçaments i acceleracions dels jugadors, proporcionant dades addicionals per a l'anàlisi de rendiment.

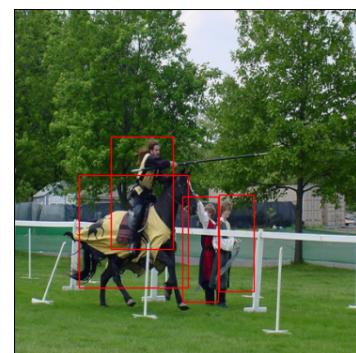


Fig. 5: Predicció del model YOLOv1 implementat des de zero.

Així doncs, vam optar per fer fine-tuning amb els models pre-entrenats YOLOv5[11] d'*Ultralytics* per a la detecció de jugadors, pilota i àrbitres, i el model YOLOv8[12], també de la mateixa plataforma, per a la detecció dels punts característics del camp.

Aquesta elecció es va basar principalment en la millor precisió del model *YOLOv5*[11] en la detecció de la pilota en comparació amb el model *YOLOv8*[12]. D'altra banda, per a la detecció de punts característics en el camp, el model *YOLOv8*[12] presentava una millor eficiència i precisió, a més de la capacitat de detectar postures o esquelets de punts, característica que resultava molt prometedora per a la detecció de punts característics del camp i el seu etiquetatge.

Vam utilitzar el conjunt de dades *football-players-detection Image Dataset* [1] per al ajust del model *YOLOv5*[11] per a la detecció de jugadors pilota i àrbitre i tot i no contenir una gran quantitat d'exemples es va mostrar suficient per als vídeos de prova que vàrem triar. Per realitzar el fine tuning del model *YOLOv8*[12] vam haver de crear un dataset des de zero on vam anotar els punts característics de cada frame per tal d'entrenar el model.

Un cop identificats els jugadors i àrbitres, fem un tracking amb *ByteTrack* [13], de la llibreria supervision. *ByteTrack* [13] és un algorisme de seguiment d'objectes multiobjecte que utilitza la informació de detecció del YOLO per a associar les caixes delimitadores dels objectes en fotogrames consecutius, creant identificadors únics per cada objecte i mantenint la correspondència al llarg del vídeo.

Un cop identificats els jugadors i àrbitres, calia classificar cada jugador en un equip. Per a això, vam implementar un sistema que utilitza un extractor de color per a cada jugador detectat i un model K-means per determinar l'equip. Vam extreure el color de cada jugador utilitzant una àrea petita centrada en la bounding box proporcionada pel model de detecció, i seguidament vam predir l'equip de cada jugador utilitzant el model K-means.

El següent pas va ser utilitzar les deteccions del model de punts característics per calcular l'homografia del camp de futbol. Aquest procés, tot i semblar relativament senzill, va suposar un dels reptes més complicats a causa dels canvis constants de perspectiva de les càmeres i la inconsistència en el nombre de punts visibles en les gravacions.

En un principi vam entrenar un model per tal que simplement detectés els punts característics, però no tenia la consistència necessària per a poder realitzar una homografia decent.



Fig. 6: Exemple de deteccions del model de punts característics.

Una altra opció que vàrem valorar al principi com a millora, i finalment vam intentar implementar com a alternativa, va ser la d'utilitzar un model *YOLOv8*[12] especialitzat en Pose Estimation. Aquest enfocament permetria aprofitar les distàncies i distribucions consistentes entre els diferents punts característics del camp de futbol, a més d'obtenir l'etiquetatge corresponent per a cada punt i també facilitaria una correlació més precisa i directa entre els punts característics, permetent un càlcul més precís de l'homografia i l'extracció de mètriques de manera més eficient. Per entrenar aquest model vam utilitzar el dataset *football-field-detection de Roboflow* [9]. Lamentablement, el model final va resultar igualment inconsistent, tenint un problema d'*underfitting* a causa de la falta de potència computacional i temps.

Finalment, per culpa de les mencionades inconsistències del



Fig. 7: Exemple de deteccions del model de Pose Estimation.

model a l'hora de detectar i classificar els punts, hem hagut de seleccionar els punts manualment en una perspectiva tàctica. Tot i que creiem que no estem molt lluny d'implementar el model d'homografia automàtic.



Fig. 8: Exemple de resultat final.

Pel que fa a les estadístiques, hem implementat una funcionalitat de càlcul i guardat tant de les distàncies totals en metres recorregudes pels jugadors com de les velocitats mínima i màxima de cada un. Aquestes distàncies i velocitats es calculen mitjançant les diferències entre les coordenades del jugador en l'homografia, adaptant-les després a les distàncies reals. Es desen en un arxiu CSV on cada columna són les estadístiques d'un jugador, l'identificador del qual serà el valor de la primera fila.

4.1 Anàlisi de Resultats

Els resultats obtinguts dels models de detecció mostren una detecció eficaç dels jugadors i els àrbitres, indicant un alt grau de precisió i robustesa en aquests casos. No obstant això, es va observar una manca de precisió en la detecció de la pilota, atribuïble probablement a la insuficient representació d'aquest objecte durant l'entrenament del model. Aquesta limitació es veu agreujada per problemes de rendiment en el seguiment de la pilota (tracking), ja que en molts fotogrames la pilota no apareix o bé la pilota apareix borrosa a causa del temps d'exposició de la càmera que capture el vídeo.

Per abordar aquestes limitacions, es recomana un entrenament més exhaustiu i específic del model *YOLOv8*[12], utilitzant un conjunt de dades amb més exemples i major diversitat d'imatges que incloguin la pilota en diferents contextos i condicions de visibilitat. A més, es podrien implementar tècniques avançades de detecció i discriminació de la pilota, com ara l'ús de models especialitzats per identificar objectes petits i ràpidament moviments. Això permetria diferenciar millor la pilota dels altres elements del camp, com els punts de penal, i augmentaria significativament la precisió de la detecció de la pilota.

Respecte a la detecció de punts característics de camp, la implementació es podria millorar entrenant millor i amb molt més de temps el model *YOLOv8*[12] especialitzat amb Pose Estimation, ja que no el vam poder entrenar de manera eficaç per la falta de temps i de capacitat computacional. Aquest model permetria extreure de manera molt més precisa i etiquetada els keypoints, la qual cosa ens permetria fer l'homografia de manera robusta en qualsevol classe de perspectiva.



Fig. 9: Exemple de possible resultat utilitzant Pose Estimation.

Aquestes millores podrien incrementar la qualitat general de l'anàlisi del partit, proporcionant dades més precises i fiables per a l'anàlisi tècnica i estratègica. L'augment en la precisió de la detecció de la pilota facilitaria una millor comprensió del joc, ajudant a extreure estadístiques més detallades i útils per a l'avaluació del rendiment dels jugadors i dels equips.

5 CONCLUSIONS

En aquest projecte, hem desenvolupat una metodologia integral per a l'anàlisi exhaustiva de partits de futbol, centrant-nos en la detecció de participants i punts característics del camp per al càlcul de l'homografia. L'ús del model YOLO ens ha permès identificar jugadors, àrbitres, pilota i punts clau del camp amb gran precisió, gràcies a la seva versatilitat i capacitat per processar informació en temps real.

El projecte es va desenvolupar en dues fases principals. Inicialment, vam experimentar amb tècniques tradicionals de visió per computador per al càlcul de l'homografia, però aquestes no van oferir els resultats esperats. Per aquest motiu, vam decidir utilitzar un model YOLO per a la detecció dels punts característics del camp. La implementació del model YOLOv1 des de zero ens va proporcionar una comprensió profunda del seu funcionament, tot i les limitacions computacionals i de temps.

Posteriorment, vam optar per fer fine-tuning amb els models preentrenats YOLOv5[11] i YOLOv8[12], obtenint millors resultats en la detecció de jugadors, pilota i punts característics del camp. La combinació d'aquests models ens va permetre millorar significativament la precisió i eficiència de les deteccions, especialment amb l'ús de YOLOv8[12] per a la detecció de postures i esquelets de punts, facilitant així el càlcul de l'homografia.

La classificació automàtica dels jugadors es va realitzar mitjançant un model K-means, basat en l'extracció del color de les samarretes. A més, vam implementar tècniques de seguiment per mantenir la identificació dels jugadors i la pilota al llarg del partit, i vam incorporar el càlcul del flux òptic per analitzar la velocitat i moviment dels jugadors.

En conclusió, aquest projecte proporciona una plataforma robusta per a l'anàlisi tècnica i estratègica del futbol, permetent extreure estadístiques i informació clau que poden ajudar a millorar el rendiment dels jugadors i de l'equip. Les estadístiques generades inclouen el percentatge de possessió de la pilota, les distàncies recorregudes pels jugadors, les zones del camp més utilitzades i les velocitats màximes i mitjanes dels jugadors. Amb futures millores, com la implementació de models especialitzats en Pose Estimation, es podrà obtenir una precisió encara més gran en el càlcul de l'homografia i l'extracció de mètriques.

REFERÈNCIES

- [1] Roboflow. 2024. Football Players Detection [Dataset]. Roboflow Universe.
- [2] Newman, J., Sumsion, A., Torrie, S., Lee, D. "Automated Pre-Play Analysis of American football formations using deep learning," *Electronics*, 12(3), 726, 2023.
- [3] Roboflow, "YOLOv9 Tutorial: Train Model on Custom Dataset — How to Deploy YOLOv9," YouTube, 2024.
- [4] Öberg, F. "Football analysis using machine learning and computer vision. Master's Thesis," Luleå University of Technology. DiVA, 2021, p. 34
- [5] Descoins, Alan, and Marvid, Diego. "Automatically measuring soccer ball possession with AI and video analytics," Tryolabs Blog, 2022.
- [6] Machine Learning With Hamza, "Computer Vision for Football Analysis in Python with Yolov8 & OpenCV," YouTube, 2023.
- [7] Kalinowski, Nick. "Exploring Basketball Spacing Through Computer Vision & Broadcast Data," Medium, 2023.
- [8] Azevedo, Daniel. "Player Detection using Deep Learning," Analytics Vidhya, Medium, 2021.
- [9] Roboflow. 2024. Football Field Detection [Dataset]. Roboflow Universe.
- [10] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2015). You Only Look Once: Unified, Real-Time Object Detection. Retrieved from arXiv:1506.02640 [cs.CV]
- [11] Ultralytics. (2023). YOLOv5 - Ultralytics YOLOv5 Documentos.
- [12] Ultralytics. (2023). YOLOv8 - Ultralytics YOLOv8 Documentos.
- [13] Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W., Wang, X. (2022). ByteTrack: Multi-Object Tracking by Associating Every Detection Box