



HAL
open science

Exploration des supports vidéos en analyse des données à la recherche d'efficacité sportive en compétition footballistique

Eya Kalboussi, Nadia Ndhaief, Nidhal Rezg

► To cite this version:

Eya Kalboussi, Nadia Ndhaief, Nidhal Rezg. Exploration des supports vidéos en analyse des données à la recherche d'efficacité sportive en compétition footballistique. SAGIP 2023 (1er Congrès Annuel de la Société Française d'Automatique de Génie Industriel et de Productique), Jun 2023, Marseille, France. ⟨hal-04254259⟩

HAL Id: hal-04254259

<https://hal.science/hal-04254259v1>

Submitted on 23 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

SAGIP 2023

Data analysis of soccer players effectiveness based on video media

Exploration des supports vidéos en analyse des données à la recherche d'efficacité sportive en compétition footballistique

Eya Kalboussi, Nadia Ndhaief, Nidhal Rezg

Laboratoire de Génie Informatique de Production et de Maintenance, Université de Lorraine (LGIPM), France

eya.kalboussi@univ-lorraine.fr, nadia.ndhaief@univ-lorraine.fr,
nidhal.rezg@univ-lorraine.fr

Mots-clés : *Requêtes XML, Tracking, Analyse de performance des joueurs.*

1 Introduction

Aujourd'hui, les données sont devenues de plus en plus accessibles, indépendamment de leur provenance ou de leur domaine d'application. Notamment, les données et les statistiques du domaine de football. Pour cela, le traitement des données à l'aide de techniques de la science des données tels que l'apprentissage automatique et l'apprentissage profond dans ce domaine est devenue primordiale afin d'aider les parties prenantes à prendre la décision stratégique pour le recrutement ou opérationnelle pour la gestion de l'équipe. En se référant à la littérature, les chercheurs ont acquis au fil des années une solide expertise dans le domaine de football. Les chercheurs ont utilisé une approche de science des données pour minimiser le temps nécessaire à la sélection des joueurs d'une équipe en tenant compte du coût et des compétences des joueurs [1]. Andriyanov [2] a effectué une comparaison entre deux algorithmes de regroupement de données ; le premier est basé sur le modèle de mélange gaussien. Quant au deuxième, il s'agit d'un algorithme de réseaux de neurones. Toutefois, le modèle de mélange gaussien semble plus précis. Les chercheurs ont utilisé une variété d'algorithmes allant de Rodrigues et Pinto [3], qui ont testé diverses méthodes d'apprentissage automatique qui utilisent de multiples statistiques des matchs précédents et les attributs des joueurs des deux équipes comme entrée pour prédire le résultat d'un match de football. Le meilleur modèle est le forêt aléatoire, qui a une précision de 25,26%. Anfilets et al., [4] ont utilisé un réseau neuronal multicouches profond pour prédire le résultat d'un match avec une précision de 61,14% dans l'objectif de prédire le résultat du match s'il s'agit d'une victoire, perte ou tirage. Majumdar et al., [5] ont étudié la prédiction des blessures dans le football pour aider à comprendre l'adaptabilité aux programmes d'entraînement, évaluer l'épuisement et la récupération, et réduire le risque de blessure et de maladie. Rathi et al., [6] ont illustré le potentiel de l'Intelligence Artificielle (IA) et ses limites dans le football. Les avantages de l'IA comprennent l'amélioration de la prise de décision, du processus de sélection des joueurs et de la prédiction du prix des billets. Toutefois, elle présente certaines limites, comme son coût élevé, l'absence de sentiments et d'émotions, la dépendance croissante à l'égard des machines, le manque de créativité originale et de pensée originale.

D'ailleurs, la détermination de la valeur marchande des joueurs, constitue aujourd'hui une préoccupation majeure. Mustafa et Sakir [7] ont proposé à cet effet une méthode quantitative

objective basée sur l'apprentissage automatique. Zanganeh et al., [8] ont utilisé une méthode d'identification et de reconnaissance des observations relatives aux objets, événements, ou scènes en utilisant l'apprentissage par transfert. Dans les travaux antérieurs, les techniques d'IA sont les plus couramment utilisées dans le contexte de l'amélioration des performances de domaine de football. À cet égard, notre travail vise à résoudre les plus grands problèmes rencontrés par les entraîneurs pour la sélection de joueurs grâce à un système intelligent de manipulation et d'analyse de données basé sur le requêtage et le tracking. Notre approche prouve que nous pouvons optimiser efficacement les matchs de football, améliorer les performances du football en compétition et donc améliorer de manière significative la prise de décision.

2 Méthode proposée

Dans ce travail, un outil d'exploitation et d'analyse des données de football a été développé dont l'objectif est de faciliter la sélection des futurs recrutements à l'entraîneur, mais aussi améliorer les performances de l'équipe et de prendre les meilleures décisions. Nous proposons de simplifier l'exploration de données dont les fonctionnalités sont :

- Visualisation de données : Vue globale et détaillée et agrégation temporelle avec la possibilité de modification de types de graphiques.
- Dashboarding et reporting : Génération des tableaux de bord interactifs.

OPTA STATS PERFORM est l'un des principaux fournisseurs de données que nous avons utilisé sur les compétitions de football dans le monde. Nous avons eu accès à deux types de données qui sont les données d'événements relatives à chacune des actions d'un match, telles que les passes, les frappes, les remplacements, les décisions des arbitres, etc ainsi que les données de tracking incluant des caractéristiques liées à la position des joueurs. Ces données sont souvent stockées dans un fichier XML et txt respectivement. Nous avons choisi d'abord de travailler avec les données d'événements. Ces dernières ont été mappées vers une base de données SQL de manière structurée afin de faciliter les processus d'exploration et de traitement des données. Le mappage du fichier XML s'est fait à l'aide de la fonction OPENXML permettant d'itérer facilement sur les éléments du fichier. Ensuite, nous avons construit des divers requêtes XML visant à avoir des statistiques sur les matchs, sur les équipes, sur les joueurs et d'autres comme le nombre de ballons interceptés ainsi que récupérés. Ces statistiques ont pour objectif d'analyser les performances des joueurs dans une vidéo de football. Pour les données de tracking, nous avons généré un fichier CSV à l'aide du module csv de la bibliothèque standard de Python afin d'examiner les caractéristiques de performance de chaque joueur et de les classer en fonction de l'indicateur choisi par l'entraîneur. Notre méthode proposée est présentée dans la figure 1.

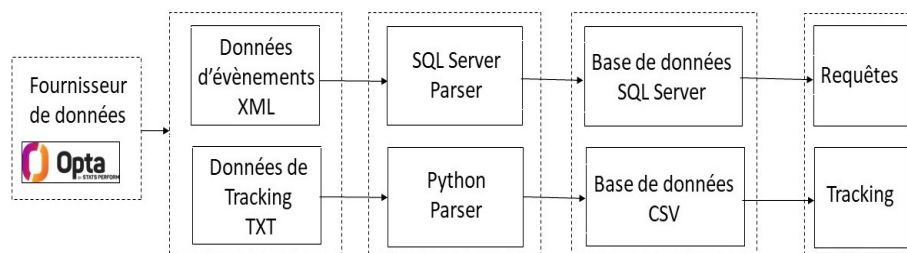


FIG. 1 – Système d'analyse de performances des joueurs de Football

La figure 2 ci-dessous présente un aperçu de visualisation des informations extraites à partir des données d'événements et de tracking d'un seul joueur. Cette visualisation montre la performance de ce joueur comme le nombre de courses, la distance parcourue dans ces plages de vitesse, le nombre de mètres gagnés par la passe ainsi que le nombre de ballons interceptés et récupérés par zone.



Player ID = 587783

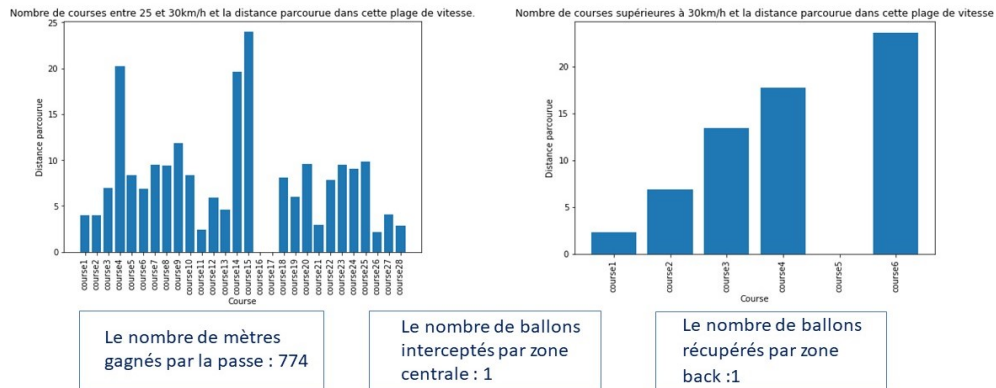


FIG. 2 – Un aperçu du tableau de bord

3 Conclusions et perspectives

Dans le cadre de ce travail, nous avons développé un outil d'exploitation de la vidéo des matchs de Football pour l'aide à la sélection des joueurs. Nous avons présenté la méthode proposée pour l'analyse vidéo de football. En effet, nous avons établi qu'il est en fait possible d'extraire des événements spécifiques et de réaliser de filtres selon les préférences. Nous avons également la possibilité d'avoir des informations utiles sur les joueurs de football. En outre, la perspective de ce travail réside dans le développement des différents modèles de réseaux de neurones permettant de prédire l'information utile ce qui permet aux entraîneurs d'analyser leurs vidéos plus rapidement depuis des différents types de base de données.

Références

- [1] Rajesh, P., Alam, M., & Tahernezehadi, M. (2020, July). A data science approach to football team player selection. In *2020 IEEE International Conference on Electro Information Technology (EIT)* (pp. 175-183). IEEE.
- [2] Andriyanov, N. (2020). Comparative analysis of football statistics data clustering algorithms based on deep learning and Gaussian mixture model. In *CEUR Workshop Proceedings* (Vol. 2667, pp. 71-74).
- [3] Rodrigues, F., & Pinto, Â. (2022). Prediction of football match results with Machine Learning. *Procedia Computer Science*, 204, 463-470.
- [4] Anfilets, S., Bezobrazov, S., Golovko, V., Sachenko, A., Komar, M., Dolny, R., ... & Oso-linskyi, O. (2020). Deep multilayer neural network for predicting the winner of football matches. *International Journal of Computing*, 19(1), 70-77.
- [5] Majumdar, A., Bakirov, R., Hodges, D., Scott, S., & Rees, T. (2022). Machine Learning for Understanding and Predicting Injuries in Football. *Sports Medicine-Open*, 8(1), 1-10.
- [6] Rathi, K., Somani, P., Koul, A. V., Manu, K. S. (2020). Applications of artificial intelligence in the game of football : The global perspective. *Researchers World*, 11(2), 18-29.
- [7] Al-Asadi, M. A., & Tasdemir, S. (2022). Predict the value of football players using FIFA video game data and machine learning techniques. *IEEE Access*, 10, 22631-22645.
- [8] Zanganeh, A., Jampour, M., & Layeghi, K. (2022). IAUFDF : A 100k images dataset for automatic football image/video analysis. *IET Image Processing*, 16(12), 3133-3142.