UCLouvain

LOUVAIN
School of Management

Louvain School of Management

Etude de Faisabilité : Modèle de détection automatique dans le football grâce à l'intelligence artificielle

Auteur : Louis Van Twembeke
Promoteur : Manuel Kolp
Année académique 2023-2024
Travail de fin d'études en vue d'obtenir le diplôme en sciences de gestion avec pour finalité finance internationale

Remerciements

Premièrement, j’aimerais remercier mon promoteur, le professeur Kolp qui m’a permis de choisir ce sujet de mémoire et qui a fait preuve de flexibilité afin de me permettre d’écrire une thèse qui avait de l’intérêt pour moi.

Je tiens également à remercier le Royal Union Saint-Gilloise et en particulier Olivier Taquin, mon maître de stage, qui m’a donné la chance de vivre un stage unique ainsi que pour les contacts précieux qu'il m'a fournis pour ce mémoire.

Je suis profondément reconnaissant envers mes parents, qui m'ont toujours soutenu tout au long de mes études et ont tout mis en œuvre pour me mettre dans les meilleurs conditions.

Je remercie aussi chaleureusement toutes les personnes, amis et camarades, avec qui j'ai étudié et trouvé la motivation nécessaire pour avancer.

Enfin, je souhaite exprimer ma reconnaissance envers tous les médecins qui m'ont aidé durant mes études sans qui je pense, la réussite ces dernières n'aurait pas été possible.

Table des matières

During the preparation of this master's thesis, the author(s) utilized [ChatGPT] for the following purpose:

1.[REASON 1] : J'ai utilisé ChatGPT pour m'aider à développer mon code et reformuler mon texte. J'ai utilisé cette IA comme un outil d'apprentissage afin de mieux comprendre certains concepts parfois complexes et d'obtenir des explications plus claires avec un vocabulaire varié.

After using [ChatGPT], the author(s) diligently reviewed and edited the content produced by the tool. We take full responsibility for the final content presented in this thesis.

By signing this declaration, we affirm that the content of this master's thesis reflects our original work, augmented by the responsible use of AI.

Introduction

Contexte

Les statistiques sportives occupent une place croissante dans le football moderne. Cependant, seules les grandes équipes professionnelles disposant de budgets conséquents peuvent se permettre d'investir dans des outils statistiques avancés. Il est bien connu que les statistiques constituent un outil de management puissant pour une équipe, améliorant ainsi les performances, la gestion du recrutement, la gestion des blessures, les stratégies tactiques et la préparation physique. Il est donc crucial, pour faire progresser le sport, que le plus grand nombre puisse accéder à ces outils. Néanmoins, les clubs de divisions inférieurs n'ont pas accès aux statistiques. Ces plus petits clubs sont financés par des plus petits investisseurs qui recherchent souvent une rentabilité immédiate et hésitent à investir dans des outils statistiques relativement coûteux. Cela crée une inégalité significative entre les clubs des différentes divisions. Le manque d'accès à des outils statistiques avancés n'impacte pas seulement les performances des joueurs et des équipes, mais a également des répercussions sur le développement global du football à une échelle plus large. Si les clubs de toutes les divisions semi-professionnelle pouvaient bénéficier de la démocratisation des outils statistiques, cela améliorerait la qualité du football belge. Par exemple, en Angleterre, les cinq premières divisions ont accès aux statistiques, ce qui élève le niveau général du pays et contribue à en faire le meilleur championnat du monde. Les statistiques permettent également de justifier certaines décisions. En effet, elles rationalisent les choix faits par les entraîneurs, offrant ainsi une base chiffrée qui rend les décisions plus facilement communicables et compréhensibles.

Question de recherche

Nous nous sommes donc demandé comment nous pourrions apporter ces outils aux équipes que nous qualifierions de 'second rang', c'est-à-dire celles évoluant dans les divisions semi-professionnelles. Actuellement, les statistiques sont encore annotées manuellement par des humains. Ces derniers sont indispensables pour identifier les joueurs, les arbitres, le gardien de but, le ballon, etc., et pour interpréter correctement les différentes données statistiques. Cependant, avec les récents progrès prometteurs de l'intelligence artificielle, nous nous sommes posé la question suivante : **« est-il possible de créer une solution capable de détecter automatiquement les différents acteurs d'un match grâce à l'IA ? »**

Dans cette thèse, nous commencerons par explorer la possibilité de détecter automatiquement les différents acteurs d'un match. Nous examinerons ensuite les différentes méthodes pour y parvenir et si cela s'avère réalisable, nous les mettrons en œuvre de manière approfondie.

Structure de la thèse

Cette thèse se compose de deux chapitres principaux. Dans le premier chapitre, nous passerons en revue la littérature afin de prendre connaissance des travaux déjà existants. Le second chapitre exposera notre contribution, en détaillant la motivation de la recherche, les théories et modèles, le traitement des données, ainsi que les résultats des tests. Nous terminerons par une section sur les limites et une conclusion.

---

1. Revue de la littérature

Dans ce premier chapitre, nous passerons en revue différents travaux existants sur l'intelligence artificielle et ses applications dans le sport, en mettant un accent particulier sur le football. Cette revue de la littérature nous permettra également de démontrer dans quels cas le football à recourt à l'intelligence artificielle. Nous explorerons aussi les modèles de détections utilisés dans la reconnaissance d'images, non seulement dans le sport mais aussi dans d'autres domaines. Nous conclurons ce premier chapitre en mettant en avant l'importance des statistiques dans le monde du football.

1.1 Apport de l'IA dans le football

Dans notre premier article [1], les auteurs soulignent l'importance de l'intelligence artificielle dans le sport. En effet, cette révolution aide au développement de stratégies et de tactiques de plus en plus sophistiquées. À l'origine, les tactiques étaient construites sur l'avis subjectif de l'entraîneur. Cela a fortement changé avec l'apport de l'IA, les entraîneurs se basant désormais en grande partie sur les données. L'IA joue un rôle crucial dans ces données car, d'une part, elle aide à la collecte et d'autre part elle est capable de traiter et d'analyser une énorme quantité de données beaucoup plus rapidement qu'un humain. En effet,

Le football, en particulier, est un sport rempli de différentes tactiques. Chaque équipe a son propre dispositif, son propre pressing, etc. L'IA permet d'analyser tous les paramètres de l'équipe adverse. Certaines choses sont évidentes pour un analyste vidéo, comme le dispositif mis en place par l'équipe adverse, le pressing utilisé, etc. Cependant, d'autres aspects ne sont pas directement visibles, comme les appels des joueurs hors ballon, que l'analyste vidéo pourrait manquer en se concentrant sur la balle. Dans l'article, il est expliqué que l'IA a plusieurs rôles significatifs. Tout d'abord, elle joue un rôle dans l'analyse et l'adaptation tactiques. En effet, l'IA est devenue un outil qui aide les équipes à s'adapter à l'adversaire. L'IA peut retracer l'historique et analyser en temps réel les données de l'équipe adverse, fournissant ainsi une série d'informations permettant à l'entraîneur, par exemple à la mi-temps, de changer sa tactique en fonction de l'équipe adverse et de réagir immédiatement.

L'IA a également un rôle dans l'optimisation de la performance des différents joueurs. Elle analyse les entraînements et les matchs de chaque joueur et, sur la base de cette analyse, les différents entraîneurs peuvent créer des programmes d'entraînement spécifiques pour chacun. L'IA permet donc aux équipes d'avoir des capacités prédictives vis-à-vis de leurs adversaires et d'anticiper

leurs différentes stratégies. Cela a un impact significatif sur la prise de décision tactique, que ce soit avant le match ou même en temps réel.

Dans un autre article [2], les auteurs sont bien conscient que le big data peut fortement aider les équipes à la prise de décisions mais ils soulignent également le fait que l'énorme quantité de données peut aussi être un obstacle en soi. L'IA est donc là justement pour aider à travailler efficacement sur ces données tellement précieuses. En effet, selon [3], l'intelligence artificielle relève ce défi en automatisant l'analyse des données ,semblable à l'exploration de données, en identifiant les modèles et les tendances qui pourraient autrement passer inaperçus.

1.2 Qu'existe - t - il en matière d'IA dans le sport et en particulier le football ? Dans quels cas est-elle utilisée ?

Dans cette section, nous allons vous présenter les différentes manières dont l'intelligence artificielle est utilisée. Nous parlerons dans un premier temps de la prédiction de l'issue d'un match, ensuite du rôle que peut avoir l'intelligence artificielle dans l'arbitrage et pour finir vous retrouverez un passage parlant de l'intelligence artificielle dans l'analyse sportive.

1.2.1 IA dans la Prédiction de l'issue d'un match

L'IA est fortement utilisée dans la prédiction du résultat. En effet, l'humain à certaines limites car il doit étudier un large tas d'information alors que l'IA peut surpasser ces problèmes. Les réseaux de neurones artificielles transforment un set d'inputs en un output en essayant d'imiter les réseaux de neurones biologiques [4]. La puissance des ANNs provient de la non-linéarité des neurones cachés dans l'ajustement des poids qui renvoie à la décision finale [5]. C'est-à-dire que le système de neurones va envoyer des fonctions d'activation non-linéaire . Ensuite en fonction de son résultat, si la prédiction est correcte ou incorrecte alors les poids sont ajustés. Ainsi, de cette manière, l'intelligence artificielle pourra s'entrainer et c'est ça qui va la rendre si puissante.

Dans cet article [6], les auteurs ont démontrés comment les réseaux bayésiens sont utilisés pour prédire le résultat d'un match. Ces réseaux bayésiens ont ensuite largement été utilisés dans le domaine de l'IA. Le score d'un match est souvent déterminé par une somme d'indicateurs. Les chercheurs [7] grâce à l'utilisation de techniques d'apprentissage automatique explicables, ont pu déterminer comment chaque indicateur influence le score. Les résultats ont révélé quels indicateurs ont le plus d'impacts sur l'issue d'un match. Cela peut donc clairement aider les entraineurs dans l'amélioration de la performance de leur équipe.

1.2.2 L'intelligence artificielle dans l'arbitrage

Dans l'article [8] ,on peut lire que l'intelligence artificielle peut aider les arbitres dans leurs décisions mais pas dans 100% des cas. Prenons le cas des phases à interprétations comme une faute de main ou une poussée. L'intelligence artificielle n'a actuellement aucun moyen de savoir si la main ou la poussée est volontaire. D'autant plus que toute poussée dans le football n'est pas fautive, certaines sont autorisée jusqu'à une certaine limite. La poussée était-elle trop forte à tel point qu'elle a pu déstabiliser le joueur ? La faute de main était - elle intentionnelle ou pas ? Ces questions, l'IA ne peut pas encore y répondre. L'humain est donc primordial pour prendre ce genre de décisions. L'IA pourrait par contre aider l'arbitre dans d'autres situations. Dans un match de football, d'autres facteurs extérieurs que le jeu influence l'arbitre comme l'ambiance, l'importance financière de la rencontre, la pression médiatique, etc. Ces derniers peuvent parfois influencer l'arbitre dans ses décisions. De tels facteurs pourraient être neutralisés en utilisant l'IA mais du travail est encore à réaliser pour être réellement utilisé. L'IA est actuellement disponible pour la détection des positions de hors-jeu. Dès qu'une situation des hors-jeu potentiel est détectée, les arbitres dans le VAR reçoivent les images. Une confirmation humaine est quand même nécessaire mais de tels jugements sont donc possible avec la technologie.

Il est également important de préciser que l'IA est bénéfique pour l'arbitrage mais selon [9] , elle présente certains dangers comme des décisions non motivées, une absence de prise en compte de l'aspect humain, etc.

1.2.3 l'Intelligence artificielle dans l'analyse sportive

L'intelligence artificielle est devenue un outil incontournable dans l'analyse sportive. Selon une étude [10], la révolution des données sportives permet à l'analyse de données de rationaliser les charges en se positionnant comme une aide précieuse à la prise de décision. Grâce à ses capacités d'analyse de grandes quantités de données et à la précision de ses modèles prédictifs, l'IA permet aux équipes et aux entraîneurs de mieux comprendre les performances des joueurs, de prévoir les résultats et d'optimiser les stratégies de jeu. Dans cette section, nous examinerons les applications de l'IA dans l'analyse et le scouting, l'annotation pour les analystes vidéo, et la détermination des positions des joueurs en fonction de leurs statistiques.

Analyse et scouting

Des auteurs [10] ,ont mis au point un solution dans l'analyse sportive avec l'intégration de l'IA. Dans leur étude, les auteurs proposent une solution novatrice permettant aux staffs sportifs de poser des questions dans une interfaces conversationnelle. Les entraineurs peuvent donc poser des questions complexes et l'interfaces va leur répondre sous forme de visualisations et analyses détaillées. Par exemple, un entraineur peut demander des informations spécifiques sur les performances d'un joueur, du style « montre-moi une liste de vidéos contenant tous les tirs de ce joueur en dehors de la surface » et ensuite l'interface conversationnelle va lui renvoyer les résultats demandés. Cette interface fonctionne sous forme d'un langage naturel étant donné qu'un staff sportif n'a pas de compétence informatique. La base de leur modèle est construite sur le flux XML F24 d'une société qui s'appelle Opta [11]. En fait, à l'heure d'aujourd'hui, les statistiques sont enregistrées sous la forme d'évènements. Le match est retransmis en direct à des personnes (souvent sous-traité en inde) et le rôle de ces personnes est de noter chaque évènement. Ensuite, une fois tous les événements enregistrés, ces derniers sont transmis sous forme de données brutes et le rôle justement de la solution proposée par les auteurs est de structurer ces données afin qu'elles soient disponibles dans un langage naturel pour les entraineurs. La simplicité du langage naturel supprime l'étape technique qui consiste à apprendre des langages spécifiques de programmation comme SQL, Python,etc.

Les auteurs ont testé leur solution sur un club de football semi-pro en Angleterre et les résultats se sont montrés très positif. En effet le club, Leatherheard FC occupait la 20 ème position dans leur championnat avant l'installation de la plateforme (après 5 matchs) et ensuite le club à terminer à la 8 ème position en fin de saison suite à l'apport de la solution. On ne peut évidemment pas affirmer à 100% que cela est dû à l'apport de la solution mais on peut constater que dès l'installation de celle-ci, les résultats se sont améliorés.

Annotation pour les analystes vidéo

Un autre solution (FootApp) fournissant des statistiques et s'aidant de l'intelligence artificielle a été développée par [12]. Ils ont créé une app ou plutôt un système complet pour annoter des matchs de football. L'architecture se présente de la manière suivante : l'application fournit tout d'abord un lecteur vidéo pour visionner et annoter les matchs. Les utilisateurs vont utiliser cette application pour étiqueter les événements de matchs. Les étiquettes générées seront ensuite stockées dans une « database ». Les joueurs sont également porteurs de capteurs inertiels pendant le match, qui enregistre tous les mouvements. Ensuite, les étiquettes générées automatiquement par les capteurs inertiels sont intégrée aux étiquettes qui ont été annotées manuellement par l'humain. L'intégration des étiquettes manuelle et automatiques se font par un

module d'intégration, créant un ensemble cohérent d'étiquettes hiérarchisées. Une fois cette étape réalisée, c'est maintenant que l'IA va agir. En effet, un module basé sur l'IA va venir vérifier la cohérence entre les étiquettes manuelle et automatique et détecter les erreurs potentielles. Les erreurs sont ensuite signalées à l'annotateur humain qui peut lui vérifier à l'œil et corriger.

Nous allons expliquer de manière plus concrète le fonctionnement de ce système. Prenons l'exemple d'un match de champions play-off opposant la Royale Union Saint-Gilloise au RSC Anderlecht. Chaque joueur de l'Union Saint-Gilloise et d'Anderlecht est équipé de capteurs inertiels Xsens Dot. Ceux-ci vont uniquement mesurer les données physiques. Le match est ensuite enregistré via une caméra grand angle, ça c'est très important car la caméra montre souvent des ralentit, des gros plans sur les coachs, etc et donc loupe de cette façon des actions de jeu. Les capteurs collectent dans le même temps des données sur tous les mouvements des joueurs à une fréquence de 120 Hz. Ce qui est plutôt bon. Une fois le match terminé, les données sont téléchargées et un analyste peut utiliser FootApp pour revoir l'enregistrement vidéo du match. L'analyste va ensuite annoter manuellement les événements du match comme les passes, tirs, tacles, but, etc. Il va par exemple annoter « event ID03- min 00 :10 sec – Passe réussie- number23 (Puertas) to number47(Amoura)-emplacement X :40 Y :20-right foot ». Une fois les annotations terminées par l'analyste vidéo, il va également aller chercher les données physiques. Il va ensuite procéder à une intégration des étiquettes. Les étiquettes générées automatiquement par les capteurs sont intégrées aux annotations manuelles. Revenons sur l'exemple plus haut, si le capteur détecte que le joueur numéro 47, Mohammed Amoura courait à une certaine vitesse au moment de recevoir la passe, cette information est ajoutée à l'annotation de la passe. La dernière étape consiste à détecter les erreurs d'annotation et c'est précisément ici que l'IA va intervenir. L'application se base sur des techniques de « Frequent Itemset Mining » afin de vérifier les incohérences qu'il pourrait y avoir entre le manuelle et l'automatique. Par exemple, si une passe est annotée mais que le capteur indique que le joueur qui a fait la passe était immobile, une alerte est générée car il est impossible de faire une passe en étant immobile, il faut que la jambe et le corps bouge. L'analyste vidéo reçoit l'alerte et va vérifier. En effet, il peut voir que ce n'est pas le numéro 23 qui a fait la passe mais le numéro 13. Il a donc commis une erreur et va donc corriger son annotation.

Position d'un joueur en fonction de ses statistiques

D'autre auteurs [13] ont utilisé des techniques de machine Learning afin savoir si on savait déterminer la position sur le terrain d'un joueur sur base de ses statistiques tactiques et

techniques. Chaque joueur a été évalué selon 52 descripteurs. Les données ont été récoltée à partir d'événements générés par OPTA. Une fois ces données récoltées, elles ont été standardisées afin de pouvoir effectuer des comparaisons. Il a fallu ensuite réduire la dimensionnalité et c'est là que les algorithmes de machine Learning comme t-SNE et UMAP sont utilisés. La réduction de la dimensionnalité sert à réduire le nombre de variables tout en conservant autant que possible les informations importantes. Grâce à l'application du t-SNE, on va pouvoir réduire les dimensions et ainsi obtenir une visualisation en 2D. Cette visualisation 2D nous permettra de placer par exemple notre joueur sur une « carte » qui reprend toutes les positions de terrain possible. Ensuite, l'algorithme RIPPER qui est en réalité un algorithme de machine Learning va à son tour avoir son rôle dans la classification des joueurs. L'algorithme RIPPER génère des règles qui aident à distinguer les différentes positions de jeu en fonction des comportements techniques et tactiques des joueurs. Par exemple, les milieux de terrains défensifs pourront être identifiés par un plus grand nombre d'interceptions et de passes courtes.

Pour illustrer, nous allons vous donner un exemple concret sur un joueur de l'Union Saint-Gilloise. Supposons que nous disposions de statistiques suffisantes comme les passes, tirs, interceptions, etc. Le problème qui pourrait se poser avec les données récoltées, c'est que certaines pourraient être difficile à comparer. Par exemple, le nombre de passes réussies peut être beaucoup plus élevé que le nombre de tirs cadrés. Chaque variable est par conséquent transformée de sorte à obtenir une moyenne de zéro et un écart-type d'un. On réduit ensuite en 2D, on utilise l'algorithme RIPPER et puis on compare. Utilisons le joueur suivant : Cameron Puertas. J'ai envie de déterminer si je dois le placer au poste de milieu défensif, offensif ou même en attaque. Admettons que Puertas cette saison ce soit : 2150 passes réussies,120 tirs cadrés,180 dribles réussis (sur 40 matchs). On va calculer les données standardisées, donc on a besoin des moyennes et écarts-types des autres milieux offensifs de la ligue.

Supposons que les moyennes et écart-types des autres milieux offensifs soient les suivantes : Passes réussies : Moyenne = 2000, Ecart types=200. Les tirs cadrés : Moyenne = 100, Ecarts type= 20. Les dribbles réussis : Moyenne = 200, Ecart-Type=30. Nous allons utiliser ces valeurs pour standardiser les performances de Puertas. Nous n'allons pas aller dans le détail des calculs étant donné que ce n'est pas le but dans cette revue de la littérature mais nous allons tout de même vous montrer les résultats de manière simplifiés. Après standardisation, nous avons respectivement pour les passes réussies, les tirs cadrés et les dribbles réussis les valeurs suivantes : 0.75,1,-0.67. Nous pouvons maintenant réduire la dimensionnalité et utiliser l'algorithme RIPPER. En analysant les résultats, nous pouvons donc conclure que le joueur Cameron Puertas est mieux adapté au poste de milieux offensif. En effet, ses performances de passes réussies et de tirs cadrés sont supérieures à la moyenne, ce qui correspond aux attentes pour un joueur occupant ce poste.

Cependant, ses dribbles réussis sont inférieur à la moyenne. On pourrait donc l'entrainer plus dans ce domaine-là pour qu'il s'améliore. Pour finir avec cet exemple, on peut voir qu'un entraineur peut prendre une décision basée sur des données et optimiser le placement du joueur.

1.3 Les modèles de détections

Les modèles de détections sont nombreux afin de faire ressortir les caractéristiques des images. La vision par ordinateur est la manière de détecter qui va nous intéresser dans ce mémoire. Nous allons d'abord introduire cette vision par ordinateur et démontrer son utilité. Nous nous intéresserons ensuite à la vision par ordinateur par réseaux de neurones convolutif (CNN).

1.3.1 Vision par ordinateur

La vision par ordinateur est une discipline de l'intelligence artificielle qui permet aux machines de percevoir, interpréter et comprendre l'environnement visuel à travers des images et des vidéos. Cette technologie, inspirée de la vision humaine, utilise des algorithmes avancés pour traiter et analyser des données visuelles, facilitant ainsi une variété d'applications dans différents domaines.

Des auteurs [14], ont mis au point une technologie capable de reconnaître intelligemment les mouvements incorrects des athlètes à l'aide de la vision par ordinateur. Cette technologie peut être particulièrement utile dans la gymnastique par exemple, ou encore le football. La technologie commence par acquérir une grosse fréquence d'images à l'aide de caméras équipées de capteurs, qui capturent des images en couleur et en profondeur des athlètes en mouvement. Ces images permettent de créer un modèle tridimensionnel dynamique des mouvements, intégrant les données sur les positions du squelette, les profondeurs, et les couleurs. Ces images sont ensuite transformées en ondelettes. Cette méthode aide à décomposer les images en différents niveaux de fréquence, ce qui facilite l'identification des détails cruciaux tout en supprimant les interférences indésirables. Le système va extraire les contours des images, détectant ainsi les bords et les changements significatifs dans les mouvements des athlètes. La masse d'image extraite est ensuite utilisée pour construire un modèle dynamique tridimensionnel des mouvements des athlètes, qui est comparé à un modèle idéal pour détecter les écarts et les erreurs. Le système utilise un classificateur basé sur l'algorithme de classification bayésienne pour identifier les mouvements incorrects, entraîné sur des exemples de mouvements corrects et incorrects. La technologie

proposée par les auteurs offre plusieurs avantages pour des sports comme la gymnastique ou le football. En gymnastique, où chaque détail technique est essentiel, le système peut détecter des erreurs subtiles telles que des déséquilibres ou des écarts de jambes, fournissant aux juges des informations précises et objectives pour évaluer les performances. Cela contribue à une évaluation plus juste et plus cohérente, tout en fournissant aux athlètes des feedbacks détaillés pour améliorer leurs techniques.

Dans le football, cette technologie peut être utilisée pour analyser les mouvements des joueurs, tels que les techniques de dribble, les passes, et les tirs. En identifiant les mouvements inefficaces ou incorrects, les entraîneurs peuvent adapter les programmes d'entraînement pour améliorer les performances des joueurs. De plus, l'analyse objective fournie par ce système peut aider à prévenir les blessures en identifiant les mouvements à risque.

1.3.2 Modèles de détections par convolution (CNN)

Ces dernières années, les réseaux de neurones profonds, et en particulier les réseaux de neurones à convolution, ont marqué une avancée significative dans le domaine de la vision par ordinateur. Les CNN se sont distingués par leurs performances remarquables en classification d'images et en détection d'objets, devenant ainsi un outil essentiel dans ces domaines. À mesure que ces modèles évoluent, leur capacité à prédire avec précision s'améliore également. Selon [15], les recherches récentes ont confirmé que les CNN profonds sont des solutions extrêmement efficaces pour les tâches de reconnaissance et de détection d'objets.

Dans l'article [15], les auteurs décrive la vision par ordinateurs comme un domaine crucial de l'informatique. Le Machine Learning, qui sous-tend le fonctionnement des CNN, se divise en trois techniques principales :

Apprentissage supervisé : Utilise des données étiquetées (entrées et sorties connues) pour entraîner les modèles. Les CNN sont ajustés pour minimiser les erreurs de prédictions, permettant une classification précise des nouvelles images.

Apprentissage non - supervisé : N'utilise pas de données étiquetées. Les modèles identifient des structures ou des regroupements naturels dans les données, découvrant des motifs cachés utiles pour l'analyse exploratoire.

Apprentissage par renforcement : Apprend par essais et erreurs, en maximisant une récompense cumulative. Bien que moins fréquent pour les CNN, il est crucial dans des domaines comme la navigation autonome et la prise de décision en temps réel.

Dans le cadre de ce mémoire, nous utiliserons un apprentissage supervisé.

Les réseaux de neurones CNN appartienne au Deep Learning, une branche du machine Learning. Le Deep Learning est une branche du Machine Learning se distingue par ses capacités uniques. Il est capable de créer de nouvelles caractéristiques sans intervention humaine. L'explosion des données disponibles ces dernières années a permis au Deep Learning d'exploiter cette masse d'informations pour l'analyse et l'apprentissage, menant à des progrès significatifs dans divers domaines. Par exemple, des applications dans les voitures autonomes, la reconnaissance vocale et visuelle a montré le potentiel du Deep Learning à rapprocher le machine Learning de l'intelligence artificielle véritable. Nous expliquerons plus en détail dans la méthodologie le Deep Learning ainsi que tout une série de concepts lié à ce dernier. La figure 1.1 ci-dessous vous montre la place du Deep Learning dans l'intelligence artificielle.

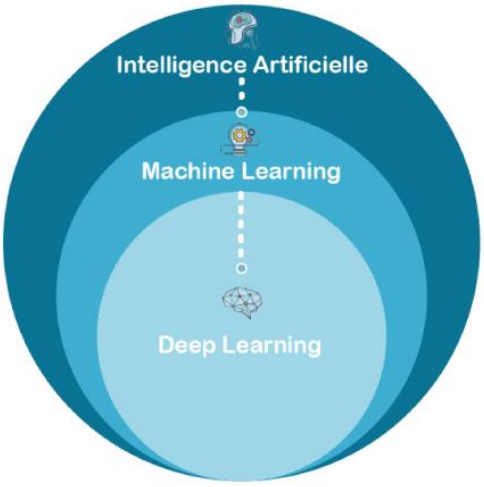

Figure 1.1 : Schéma illustrant la place du Deep Learning dans l'IA [16]

Dans le papier [17], les auteurs ont conçu un modèle de détection fiable capable de reconnaître les joueurs sur n'importe quel match de football. Ce modèle repose sur les technologies avancées d'intelligence artificielle, en particulier sur les réseaux de neurones convolutifs (CNN). Ils ont utilisé un ensemble de données nommé ISSIA-CNR soccer, ainsi que des vidéos de matchs de football pour entraîner et évaluer leur modèle. Cela a permis de tester la généralité de l'algorithme.

Voici dans la figure 1.2 l'architecture du modèle de détection utilisée par [17] :

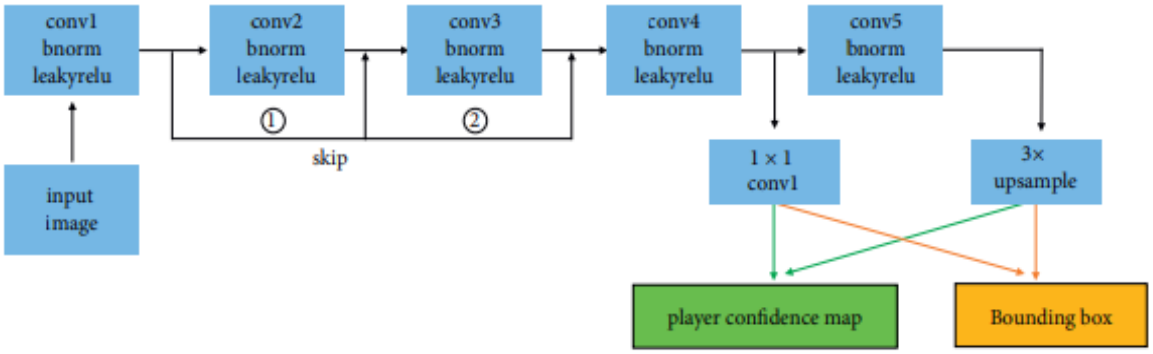

Figure 1.2 : Architecture du modèle de [17]

Le modèle analyse les vidéos de football image par image. D'abord, la vidéo est divisée en plusieurs images par seconde. Les différentes couches de convolution agissent comme des filtres, extrayant des caractéristiques spécifiques telles que les contours, les motifs, et les différences de couleur. Après chaque couche de convolution, les résultats sont normalisés pour stabiliser et accélérer l'apprentissage, en ajustant les activations. Ensuite, une fonction d'activation, Leaky ReLU (Rectified Linear Unit), est appliquée pour introduire de la non-linéarité, aidant le modèle à capturer les variations complexes des caractéristiques de l'image.

À chaque étape, le modèle produit des "feature maps" qui représentent des aspects clés de l'image. Par exemple, une feature map pourrait mettre en évidence les contours des objets, tandis qu'une autre pourrait se concentrer sur les zones de couleurs spécifiques. Certaines de ces cartes de caractéristiques peuvent être agrandies pour retrouver une taille plus grande ou même la taille d'origine de l'image, afin de conserver les détails qui pourraient avoir été perdus dans les étapes précédentes.

Les caractéristiques extraites à différentes résolutions sont ensuite combinées pour fournir une vue d'ensemble plus complète. Cela permet au modèle de mieux comprendre les détails complexes de l'image, comme la différenciation des éléments similaires.

Le modèle génère également une carte de confiance (Confidence Map) qui indique les zones où il y a une forte probabilité de trouver un objet d'intérêt, comme un joueur de football. Enfin, il dessine des boîtes englobantes (Bounding Boxes) autour des zones où il détecte ces objets, permettant de les localiser précisément dans l'image.

Ce modèle de détection par convolution est efficace pour analyser des vidéos en temps réel et peut être adapté à d'autres applications de détection d'objets.

Un autre article [18] qui a également pour objectif de détecter et de traquer les joueurs utilise aussi les réseaux de neurones convolutifs. Le modèle principal utilisé dans cette étude est le Mask R-CNN, un réseau de neurones spécialisé dans la segmentation d'images. Mask R-CNN utilise ce que l'on appelle un RPN pour générer des propositions de régions dans l'image susceptibles de contenir des objets. Ces propositions sont ensuite affinées pour prédire les classes des objets et ajuster les cadres de délimitation. Dans ce Mask R-CNN, deux modèles pré-entrainés ont été utilisés pour extraire les caractéristiques des images : ResNet101 et DenseNet. Dans l'article [18], ils les appellent les backbones. Les backbones sont des composants essentiels dans les réseaux de neurones, agissant comme des extracteurs de caractéristiques. Ces modèles ont été initialement entraînés sur des bases de données d'images générales, et ont ensuite été adaptés aux spécificités des données utilisées par les auteurs dans cette étude. Les backbones jouent un rôle critique en capturant des informations détaillées à partir des images, ce qui est ensuite utilisé par le Mask R-CNN pour la détection et la segmentation.

L'entraînement du modèle a été réalisé sur une période de 80 époques, chaque époque comportant 500 étapes d'entraînement. Pendant cette phase, les images ont été converties au format RGB, compatible avec les exigences de Mask R-CNN. La perte d'entraînement et de validation a été calculée pour chaque modèle (ResNet101 et DenseNet). DenseNet a montré une meilleure réduction de la perte, indiquant une meilleure généralisation sur les données de validation. DenseNet a donc logiquement été choisi pour ses capacités à extraire des cartes de caractéristiques plus détaillées et pour ses performances supérieures par rapport à ResNet101.

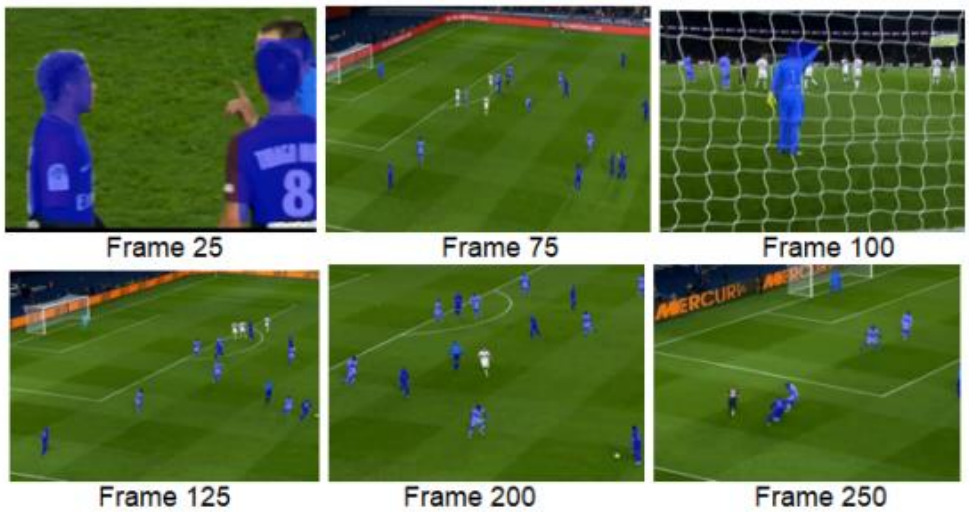

Figure 1.3 : résultats du Mask R-CNN sur une nouvelle vidéo [18]

On peut observer visuellement sur la figure 1.3 les résultats obtenus par les auteurs de l'étude. Leur modèle, basé sur des réseaux de neurones convolutifs (CNN), permet d'identifier efficacement les différents joueurs présents sur les différentes images. Ce qui est particulièrement notable, c'est que cette identification demeure précise même sous différents angles de caméras, et avec des joueurs dans toutes les positions. Que les joueurs soient en mouvement, accroupis, ou en plein saut, le modèle parvient à les reconnaître correctement.

Cela est rendu possible grâce à un entraînement de haute qualité. En effet, le modèle a été soumis à un grand nombre d'images variées, ce qui a permis de renforcer sa capacité à généraliser et à s'adapter à des situations diverses. Cette étude met en évidence comment il est possible d'arriver à des résultats très satisfaisants dans l'identification de joueurs, grâce à l'utilisation des réseaux de neurones.

1.4 Importance des statistiques dans le sport

On peut lire dans l'article [19], que les statistiques occupent une place croissante dans le sport moderne, devenant un outil essentiel pour le management des équipes. L'analyse statistique permet de quantifier les performances individuelles et collectives, fournissant ainsi des données objectives pour les décisions stratégiques. Ici, dans le contexte du basketball français, cette tendance est particulièrement notable, où l'introduction des statistiques a transformé la manière dont les entraîneurs et les joueurs perçoivent et évaluent les performances.

Historiquement, l'utilisation des statistiques dans le sport était limitée, mais l'essor des technologies de l'information a révolutionné ce domaine. Les données collectées permettent de suivre diverses métriques telles que les points marqués, les rebonds, les passes décisives, et les interceptions, offrant une vision détaillée des contributions de chaque joueur. Cette précision aide non seulement à ajuster les stratégies de jeu mais également à identifier les forces et les faiblesses des équipes adverses. Les statistiques sont devenues un outil incontournable pour la préparation des matchs, influençant les choix tactiques des entraîneurs. Cependant, l'accent mis sur les statistiques peut aussi mener à une individualisation excessive des performances. Les joueurs peuvent être tentés de se concentrer sur leurs propres statistiques pour améliorer leur visibilité et leurs perspectives de carrière, parfois au détriment de la cohésion d'équipe. Ce phénomène, observé dans le basketball français, soulève des questions sur l'équilibre entre les intérêts individuels et collectifs. Les « joueurs de stats », qui privilégient leurs performances personnelles, peuvent nuire à la dynamique de l'équipe, bien que ces statistiques soient également cruciales pour le recrutement et la gestion des carrières.

Un autre auteur souligne également l'importance des statistiques dans le sport, mais insiste sur le fait qu'elles ne doivent être qu'un outil pour les entraîneurs et les joueurs, afin de mieux comprendre leurs propres performances. Selon lui, c'est toujours l'humain qui prend la décision finale, et non les chiffres. En côtoyant des entraîneurs ayant une grande expérience internationale, l'auteur a compris que les statistiques et la technologie jouent un rôle crucial dans l'amélioration des performances des équipes de haut niveau. Cette influence s'étend à l'ensemble des sports collectifs. Cependant, les chiffres ne peuvent pas capturer toute l'essence du sport de haut niveau. Les entraîneurs reconnaissent que l'humain occupe une place centrale et indispensable dans le succès d'une équipe. Les statistiques servent simplement de support pour aider à mieux comprendre les forces et les faiblesses des joueurs, permettant ainsi aux équipes de se concentrer sur la stratégie contre leurs adversaires.

Dans les moments critiques d'une compétition, comme lorsque Cristiano Ronaldo tire un penalty crucial, le joueur se retrouve seul face à sa tâche. Les statistiques peuvent fournir un soutien émotionnel et mental, aider à la gestion de l'énergie ou guider le choix technique, mais l'exécution finale reste imprévisible et unique. Le haut niveau est aussi caractérisé par la créativité et l'originalité des sportifs, qui peuvent surprendre et réinventer leur discipline pour mener leur équipe à la victoire.

La technologie, dans ce contexte, est un outil qui permet aux athlètes de sublimer leurs performances et de continuer à nous émerveiller. Cependant, il est essentiel de préserver l'aspect humain du sport. La magie du sport ne se résume pas aux chiffres ; elle réside dans l'unicité et l'imprévisibilité des grands moments sportifs. Le sport ne doit pas devenir purement technologique ; il doit conserver son essence humaine.

Les statistiques sont par conséquent devenues indispensables dans le sport, offrant des outils puissants pour l'analyse et la gestion des équipes. Cependant, leur utilisation doit être équilibrée et intégrée de manière à ne pas éclipser l'aspect humain du sport, qui reste au cœur de ses moments les plus mémorables et inspirants.

Un autre article de [20] souligne l'importance des statistiques dans le sport contemporain. Le sport d'aujourd'hui fait appel aux statistiques afin de mieux se préparer aux multiples enjeux des compétitions. Des entreprises comme Opta ou Prozone enregistrent des milliers de points de données provenant d'élites compétitions sportives. Même si l'importance des statistiques est évidente, deux approches dominantes persistent.

D'une part, il y a les instrumentalistes qui s'enthousiasment pour la perspective d'améliorer les performances sportives grâce à des statistiques ciblées. Pour eux, les statistiques offrent une méthode objective et scientifique pour optimiser l'entraînement, les tactiques et même les choix

nutritionnels des athlètes. D'autre part, les romantiques déplorent la perte de véritable authenticité des sports, déjà hautement commercialisés. Ils estiment que l'obsession pour les données quantitatives réduit la beauté et l'émotion inhérentes aux performances sportives.

Les auteurs de l'article soutiennent que ces deux points de vue négligent le rôle actif des statistiques dans la constitution des pratiques sportives. Pour remédier à cette lacune, ils proposent que les statistiques sportives soient considérées comme performatives, c'est-à-dire qu'elles contribuent à créer la réalité sportive de la même manière que les modèles économiques ont été théorisés comme performatifs des institutions économiques. Une telle approche pourrait éclairer de nombreuses questions empiriques, telles que l'influence des statistiques sur le développement des talents, la prise de décisions tactiques et la régulation des marchés sportifs.

[20] plaident pour une perspective performative des statistiques sportives, permettant de mieux comprendre comment les données influencent non seulement les performances, mais aussi la manière dont les sports sont perçus et pratiqués à travers des dynamiques socio matérielles complexes.

D'autres auteurs [21] soulignent que de grandes quantités de données et de statistiques sont présentes dans tous les sports, mais que les organisations sportives qui parviennent à extraire des connaissances utiles de ces données obtiennent un avantage sur leurs concurrents. En effet, en transformant les données brutes en informations exploitables, ces organisations peuvent optimiser leurs stratégies de jeu, améliorer les performances individuelles et collectives, et prendre des décisions plus éclairées en matière de recrutement et de développement des talents. La capacité à interpréter correctement les données et à en tirer des conclusions pratiques permet également de mieux préparer les matchs en identifiant les points faibles des adversaires et en élaborant des plans de jeu efficaces pour les contrer. Il est donc important de savoir quelles sont les données intéressantes et celles qui ne le sont pas.

2. Contribution

Le chapitre deux présente la contribution de cette thèse en détaillant la motivation et le cadre théorique derrière le projet. Il introduit les concepts clés du Machine Learning, avec une attention particulière portée au fonctionnement des réseaux de neurones CNN et du modèle de détection YOLO. Le traitement des données, l'entraînement de l'IA, ainsi que les résultats obtenus tant sur le plan technologique que managérial sont également abordés. Enfin, ce chapitre discute des limitations rencontrées et des défis liés à l'application pratique du modèle.

2.1 Motivation de notre contribution

À l'heure actuelle, les statistiques jouent un rôle crucial dans le football. De plus en plus d'équipes basent leurs tactiques sur des analyses statistiques approfondies. Les analystes vidéo et les data analystes passent leur semaine à décortiquer les statistiques de leur propre équipe ainsi que celles de leurs adversaires pour préparer le match du week-end. Cependant, seuls les grands clubs de première division peuvent se permettre d'accéder à des statistiques avancées, souvent fournies par des entreprises spécialisées comme Opta.

À l'heure où nous écrivons cette thèse, nous allons commencer par répondre à la question : Comment sont les statistiques dans le monde professionnel sont-elles créées ? En fait, le match est diffusé en direct à une équipe (souvent sous-traité en Inde) de 5/6 personnes qui a pour rôle de noter ce que l'on appelle des événements. Ils vont en fait annoter chaque événement du match, ce qui créera à la fin une énorme base de données. Une fois le match terminé ou même parfois pendant le match, les personnes en Inde traite les données d'événements afin de les envoyer le plus rapidement possible. Nous développerons plus tard en détail le mode de fonctionnement de ces annotations.

Nous nous sommes donc posé la question suivante : Avec les avancées technologiques récentes en intelligence artificielle, serait-il possible de détecter automatiquement les différents joueurs, le ballon, l'arbitre et le gardien de but, pour qu'ensuite les événements soit notés de manière automatique. Dans cette contribution, nous allons explorer cette possibilité et voir comment il est envisageable de développer une solution capable de reconnaître les différents acteurs et objets d'un match.

Pour détecter nos différents éléments, plusieurs modèles de détection existe. Nous allons utiliser le modèle YOLO que j'expliquerai plus en détail par la suite. Ensuite, nous entrainerons ce modèle sur un dataset afin d'améliorer sa précision et pour finir nous vous présenterons nos résultats.

2.2 Théorie et modèle

Dans cette section, nous allons explorer les fondements théoriques et les modèles utilisés pour développer notre solution de détection automatique des objets. Nous aborderons tout d'abord les concepts essentiels du Machine Learning, suivis d'une partie consacrée à l'apprentissage automatique à travers les réseaux de neurones convolutifs (CNN). Enfin, nous détaillerons le modèle de détection YOLO (You Only Look Once), qui sera au cœur de notre approche. Cette combinaison de techniques avancées nous permettra de créer un système puissant et efficace pour la reconnaissance et la classification des objets.

2.2.1 Introduction au Machine Learning

Pour répondre à la question de recherche, notre solution sera programmée avec du Machine Learning. « Le Machine Learning est le moteur qui contribue au progrès dans le de développement de l'IA » [22] . En effet, le Machine Learning permet aux ordinateurs d'entreprendre des tâches complexes avec un puissance de calcul, un énorme stockage, et une mémoire impressionnante. Le Machine Learning consiste à programmer une machine qui va apprendre à réaliser des tâches sur base de pleins de données afin de définir un modèle. Par exemple, en utilisant une formule comme $Y = Wx + b$. La machine va elle ajuster les paramètres W et b afin d'obtenir le meilleur modèle possible. Initialement, W et b sont choisis aléatoirement et le but est de trouver les paramètres W et b qui donne le meilleur modèle possible, c'est-à-dire le modèle qui s'ajuste le mieux à nos données.

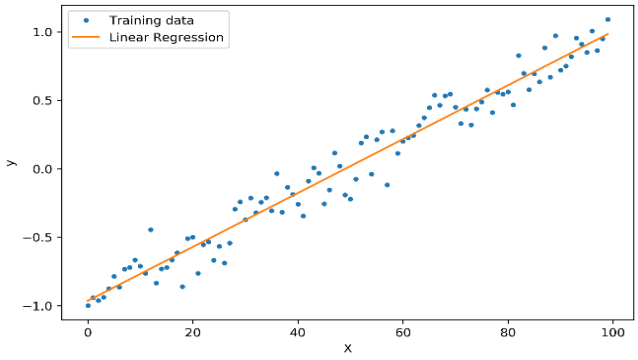

Figure 2.1 : Modèle d'ajustement des paramètres W et b

L'algorithme va donc essayer pleins de valeurs différentes jusqu'à trouver celle qui minimise la différence entre les prédictions et les résultats réels. On obtient donc la « Cost Function ». Nous développerons plus tard ce qu'est une fonction de coût.

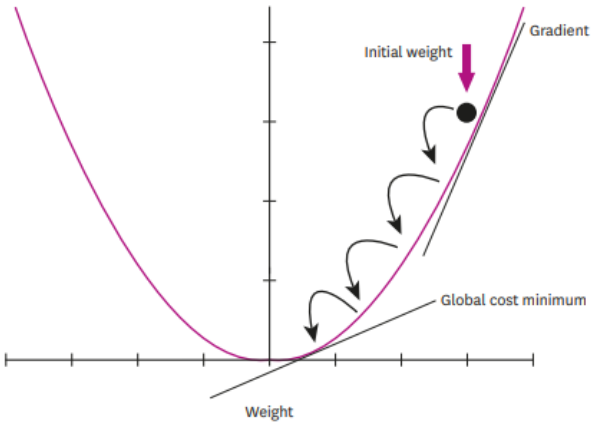

Figure 2.2 : Fonction de coût [23]

C'est ainsi que le Machine Learning permet d'entrainer des modèles pour qu'ils deviennent de plus en plus précis. En ajustant continuellement les paramètres pour minimiser la fonction de coût, la machine affine ses prévisions. Ce processus d'apprentissage permet à la machine de s'améliorer de manière autonome.

Pour notre modèle que nous présenterons plus tard, nous utiliserons du Deep Learning. Le Deep Learning est une intégration du Machine Learning. Le Deep Learning fonctionne de la même manière mais ce dernier utilise des fonctions beaucoup plus complexes. On ne se contente plus de faire fonctionner la machine sur une simple fonction linéaire. Le Deep Learning utilise des réseaux complexes de neurones artificiels pour traiter des tâches encore plus sophistiquées. Ces réseaux de neurones sont en réalité des réseaux de fonctions connectés les unes aux autres. Plus il y a de fonctions, plus les prédictions seront meilleures. Ce système est similaire à celui du système nerveux humain [24]. Le terme Deep Learning va utiliser des réseaux de neurones profonds, c'est-à-dire des réseaux avec de nombreuses couches cachées, d'où l'appellation Deep Learning.

2.2.2 Approche d'apprentissage automatique

Fonctionnement d'un neurone artificiel

Nous allons dès à présent vous expliquer le fonctionnement d'un neurone en illustrant avec des images. Un neurone est représenté en 3 couches. Une couche d'entrée, une autre cachée et une couche de sortie. Dans notre cas, nous avons ceci :

- **Entrée** : Pixel correspondant aux images d'un match de football.

- **Cachées** : La couches cachée effectue des calculs complexes pour détecter des motifs comme les bords, textures, formes et objets complets (joueurs, ballons, etc.). Ces couches utilisent des fonctions d'activation pour transformer et extraire les caractéristiques importantes des données d'entrée.

- **Sortie** : La couche de sortie produit les résultats finaux sous forme de boîtes englobantes et de classes pour les objets détectés. Pour chaque grille de l'image, le modèle prédit les coordonnées des boîtes englobantes, les scores de confiance, et les classes des objets. Par exemple, il reconnait un ballon avec une confiance de 0.8.

Un neurone reçoit donc plusieurs entrées $(x1, x2, .. , xn)$ et les multiplient par des poids $(W1, W2, .. Wn)$. Un biais est également ajouté. Le neurone va s'activer si la somme pondérée des entrées dépasse un certain seuil.

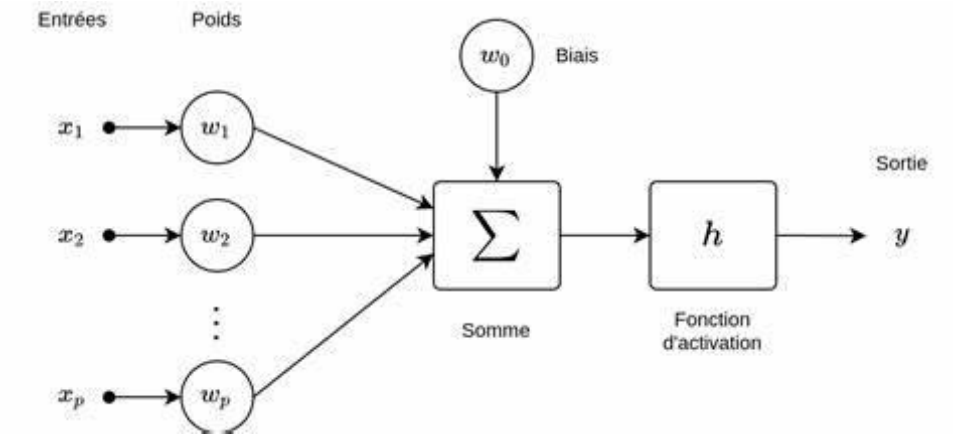

Figure 2.3 : Visualisation d'un neurone artificielle [25]

On pourrait imaginer chaque entrée $(x1, x2, .., xn)$ comme des signaux envoyés au neurone et ensuite le neurone va multiplier les signaux $(x1, x2, .., xn)$ par des poids $(w1, w2, .. wn)$ qui vont venir ajuster chaque signal. La fonction d'agrégation se présente sous la forme suivante :

$$f(x) = \sigma(W_1 x_1 + W_2 x_2 + W_3 x_3 + b)$$

Où:

- Les différents x sont les entrées.
- Les différents W sont les poids associés à chaque entrée.
- b est le biais.
- σ sigma est la fonction d'activation.

Cette fonction $f(x)$ va nous donner la sortie Y en fonction des entrées reçues.

Les poids

Afin de bien comprendre la fonction d'agrégation, nous allons vous expliquer brièvement ce qu'est un poids et son rôle.

Les poids sont en fait les coefficients W qui vont venir déterminer l'importance de chaque entrée dans chaque neurone. Lorsqu'une entrée est multipliée par un poids, cela permet au modèle d'ajuster l'influence de chaque entrée sur le résultat final. Chaque poids représenté par les traits sur la figure 2.4 est ajusté au cours de l'apprentissage pour réduire l'erreur entre la prédiction du modèle et la réalité.

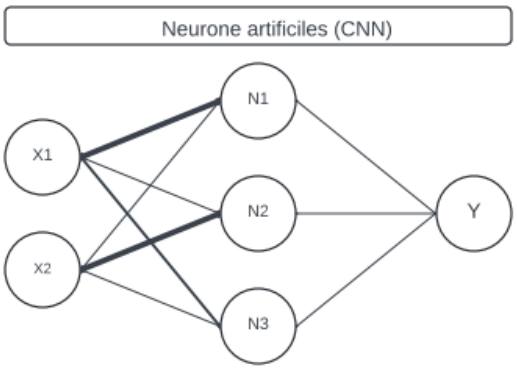

Figure 2.4 : Représentation de différents poids

Alors comment les poids sont réajustés ? On va utiliser ici la descente de gradient. L'idée est de calculer le gradient de la fonction de coût par rapport aux poids et de mettre à jour les poids en remontant couche par couche le réseau mais cette fois ci dans l'autre sens. Cette démarche appelé rétropropagation permet de réduire progressivement l'erreur de prédiction du modèle.

Fonction d'activation

Passons maintenant au signe σ sigma qui sur la fonction d'agrégation représente la fonction d'activation. Cette fonction va transformer la somme pondérée des entrées et du biais en une sortie qui sera utilisée pour les couches suivantes du réseau ou comme sortie finale.

La fonction d'activation introduit de la non-linéarité dans le réseau de neurones, ce qui permet au modèle d'apprendre et de modéliser des relations complexes dans les données. Sans fonction d'activation, le réseau serait simplement une combinaison linéaire des entrées, limitant ainsi sa capacité à capturer des motifs complexes. Il existe pleins de type de fonctions de d'activation. Dans cette thèse, nous utiliserons la fonction Rectified Linear Unit, pour sa simplicité.

La fonction ReLu est très utilisée car elle est très facile. Selon La fonction est définie de la manière suivante :

$$f(x) = \max(0, x)$$

On peut observer que cette fonction va considérer toutes les valeurs d'entrée négative ou nulle comme égale à zéro et chaque valeur d'entrée positive gardera sa valeur. Cette fonction est non-linéaire et c'est cette non-linéarité qui va faire en sorte que le réseau de neurone sera puissant. ReLu estrès utilisée car elle résout le problème de la disparition des gradients connu sous le nom « vanishing gradients »[26].

On peut donc observer ici graphiquement ce que nous avons explicité au-dessus.

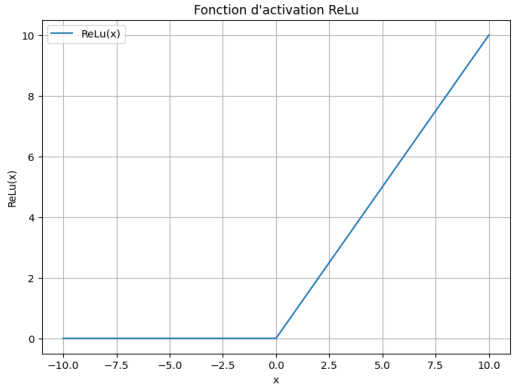

Figure 2.5 : Fonction d'activation ReLu

Fonction de perte

La fonction de perte que vous pouvez observer sur la figure 2.6 joue un rôle important dans la construction du modèle. En effet, comme expliqué plus haut, les paramètres W et b vont s'ajuster jusqu'à minimiser l'écart la valeur prédite et la valeur réelle. L'objectif est de minimiser cette perte pour améliorer la précision du modèle. Lors de l'entraînement, le modèle utilise cette mesure pour ajuster ses paramètres afin de réduire cette différence. Plus la perte est faible, plus les prédictions du modèle sont proches des valeurs réelles.

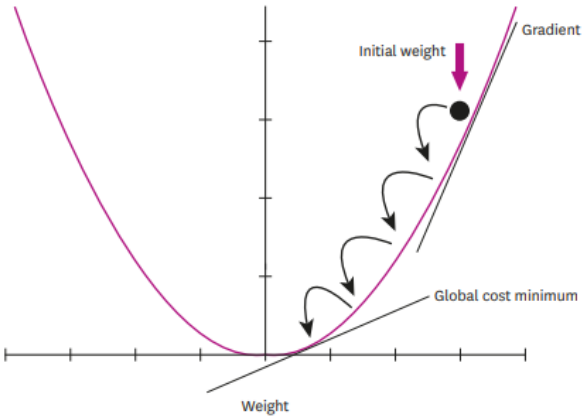

Figure 2.6 : Fonction de perte [23]

Réseaux de neurones convolutif (CNN)

Dans cette partie, nous allons tenter de vous expliquer au mieux les réseaux de neurones convolutifs.

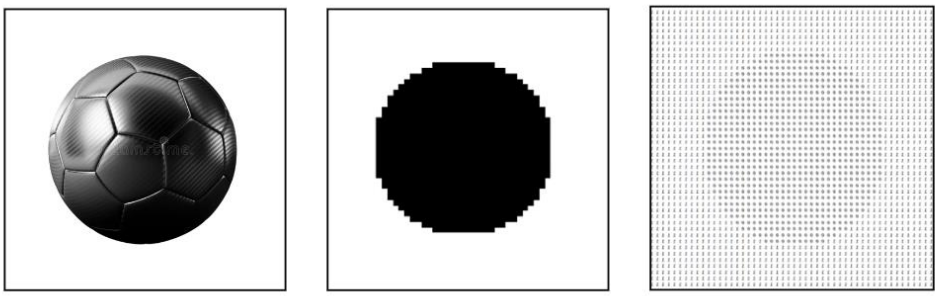

Figure 2.7 : Lecture d'un ballon par la machine

Les différentes images que nous allons faire passer à notre machine (voir figure 2.7) seront lues sous forme de matrices. Chaque pixel prendra une valeur spécifique. Pour simplifier nos

explications, nous allons d'abord utiliser l'exemple d'une image en noir et blanc. Prenons l'exemple d'un ballon de foot en noir et blanc. Si nous observons cette série d'images, nous pouvons voir d'abord une image ordinaire du ballon, puis le ballon est pixelisé, et finalement, est représentée sous forme de matrice avec des valeurs 0 et 1. Zéro représente le noir et 1 le blanc.

Pour les images en couleur, nous utilisons trois matrices (bleu, vert, rouge) superposées. Chaque matrice prend des valeurs de 0 à 255, où 0 représente le noir et 255 le blanc. La superposition de ces trois matrices permet de reconstituer l'image en couleur.

Nous allons appliquer différentes matrices de convolution sur une image pour mettre en évidence certaines caractéristiques, ce qui nous donnera une nouvelle image. Par exemple, certaines matrices, comme la matrice identité, produiront la même image, tandis que d'autres matrices mettront en évidence les contours, etc. On peut voir cela comme un filtre appliqué à l'image. Concrètement, notre matrice de convolution passe sur l'image de gauche à droite et de haut en bas.

Pour la reconnaissance d'objets, nous utiliserons les réseaux de neurones convolutifs. La structure du réseau de convolution est la suivante :

1. **Convolution** : Une première couche de convolution prend en entrée une image et applique des filtres. Ces filtres mettent en évidence certaines caractéristiques, comme les pixels horizontaux ou verticaux. Le nombre de pixels de l'image reste le même après la convolution, ce qui permet de faire ressortir des couleurs, des contours, etc.

2. **Activation** : La fonction d'activation ReLU (Rectified Linear Unit) est ensuite appliquée. ReLU transforme chaque élément de la carte de caractéristiques en mettant à zéro toutes les valeurs négatives et en laissant les valeurs positives inchangées, introduisant ainsi de la non-linéarité.

3. **Pooling** : L'image passe ensuite par une couche de pooling. Le pooling réduit la taille de l'image tout en gardant les informations les plus importantes. Le pooling le plus courant est le max-pooling, qui prend le maximum d'une région de la carte de caractéristiques.

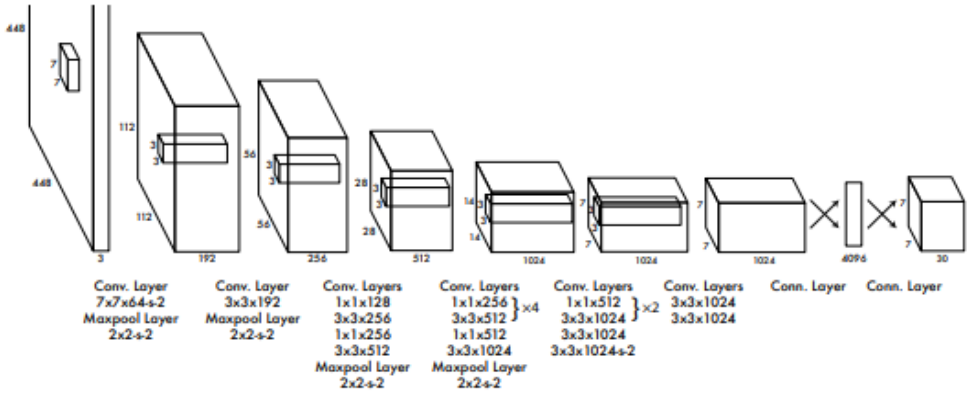

Figure 2.8 : Architecture CNN [27]

Ces étapes de convolution, activation et pooling (voir figure 2.8) sont répétées plusieurs fois dans le réseau. Chaque répétition permet d'extraire des caractéristiques de plus en plus complexes de l'image, rendant le modèle de plus en plus capable de reconnaître des motifs spécifiques.

Après plusieurs couches de convolution, d'activation et de pooling, les différentes matrices remplies de caractéristiques sont réduites petit à petit et sont transformées en un vecteur plat (flattening). Ce vecteur passe ensuite à travers des couches cachées entièrement connectées pour finaliser la reconnaissance d'objet. Voilà donc concrètement, comment la dernière phase de reconnaissance fonctionne. Après avoir aplati nos données, nous obtenons de nouvelles entrées $x1, x2, x3, \dots, xn$.

Pour illustrer cela plus en détail, prenons l'exemple de la reconnaissance du ballon de football noir que vous avez vu juste avant :

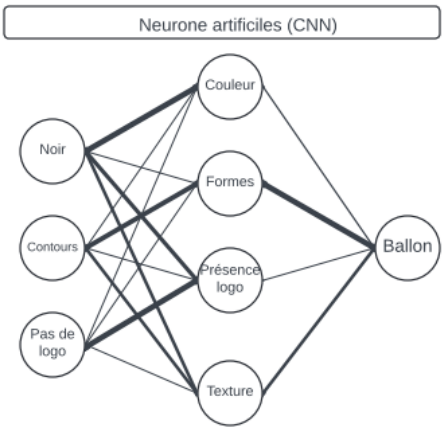

Figure 2.8 : Composition d'un CNN du ballon de football

Comme montré dans la figure 2.8, nous avons plusieurs entrées ($x1, x2, x3$) qui sont reliées à des neurones ($n1, n2, n3, n4$), lesquels sont en réalité des fonctions. Un poids W ajuste l'importance de chaque entrée pour chaque neurone. Par exemple, l'entrée $x1$, qui représente la couleur noir détecté sur l'image, est très importante pour le neurone $n1$, dont la fonction est de détecter l'intensité des couleurs. De même, $x2$, l'entrée qui représente les contours ovales de la balle, a un poids W important pour le neurone $n2$, qui détecte les différents contours et formes. A l'inverse l'entrée $x2$ n'a évidemment aucune importance pour le neurone $n1$, qui lui s'occupe des couleurs.

Une fois les entrées multipliées par leurs poids et le biais ajouté, une fonction d'activation transforme la somme pondérée en une sortie non-linéaire. Dans notre cas, nous utilisons la fonction d'activation ReLU (Rectified Linear Unit), qui met à zéro toutes les valeurs négatives et garde les valeurs positives inchangées.

Ensuite, les couches cachées, qui sont des fonctions, interprètent ces entrées afin de produire une sortie. Chaque fonction a des poids différents, permettant ainsi au réseau de neurones de reconnaître et de classifier l'objet avec précision.

2.2.3 Modèle de détection

Modèle YOLO

Afin de développer notre solution d'intelligence artificielle dans l'apprentissage automatique pour ensuite dans le futur sortir des statistiques automatisés, nous avons utilisé le modèle de détection You only look once (YOLO). Ce modèle a été développé en 2015 par [27]. YOLO utilise une technique de détection très innovante par rapport à d'autres. YOLO traite la détection d'objets comme un problème de régression. Cela signifie que le modèle prédit directement les coordonnées des boîtes englobantes des objets ainsi que les probabilités des classes associées. Un seul réseau de neurones convolutif (CNN) prédit ces boîtes et les probabilités de classes en une seule évaluation. La première version de ce modèle permettait de traiter 45 images par seconde, la dernière version est capable d'en traiter 155 par seconde sur du matériel haut de gamme. Cette architecture YOLO fait de lui un modèle extrêmement rapide.

Processus de détection du modèle YOLO

1. Division de l'Image en Grille

Le modèle prend une image d'entrée et la divise en une grille de S × S cellules (voir figure 2.9). La cellule qui est chargée de détecter l'objet est la cellule qui contient le centre de l'objet. Chaque cellule va donc prédire B boîtes englobantes. Cela signifie que chaque cellule évalue plusieurs hypothèses pour la présence d'objet. Dans notre code, j'utiliserai la version Yolov8, qui prédira jusqu'à 3 boites englobantes par cellule.

Figure 2.9 : Exemple de division d'une image en grille [28]

2. Prédiction des Boîtes Englobantes et des Scores de Confiance

Pour chaque cellule de la grille, YOLO prédit plusieurs boîtes englobantes (bounding boxes). Chaque boîte englobante est définie par 5 éléments :

- Les coordonnées du centre de la boîte (bx, by)
- La largeur (bw) et la hauteur (bh) de la boîte.
- Un score de confiance.

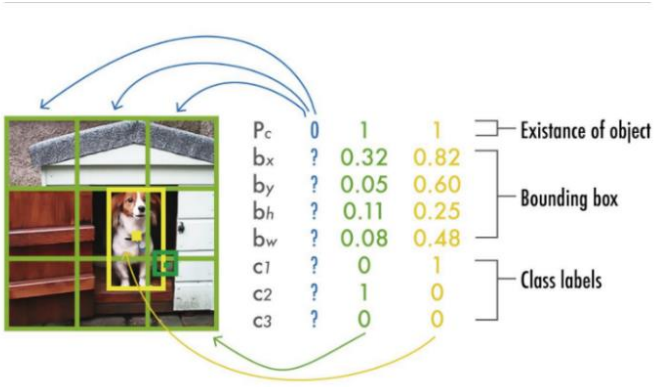

Figure 2.10 : Vecteur de prédiction [28]

3. Calcul du Score de Confiance

Le score de confiance reflète deux choses :

- La probabilité qu'il y ait un objet dans cette boîte. (Pc)
- La précision de cette boîte par rapport à la boîte de vérité (ground truth).

Le score de confiance est calculé comme : **Confiance = Pr (Object) * IOU** où l'IOU (Intersection over Union) mesure combien la boîte prédite chevauche la boîte de vérité.

3.1 Fonctionnement de IoU :

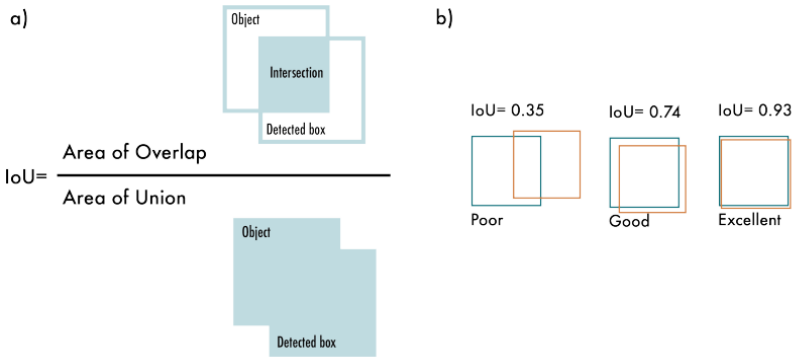

Figure 2.11 : Illustration de l'IoU [28]

L'IOU, ou Intersection over Union, est la métrique utilisée pour évaluer la précision des prédictions des modèles de détection d'objets. En fait, elle mesure le chevauchement entre la boîte englobante prédite par le modèle et la boîte englobante réelle (vérité terrain) de l'objet. L'IOU est calculée comme suit :

$$\frac{\textbf{Aire de l'intersection}}{\textbf{Aire de l'union}}$$

1. **Aire de l'intersection** : La région commune entre la boîte prédite et la boîte réelle.
2. **Aire de l'union** : La somme des aires des deux boîtes, moins l'aire de leur intersection.

4. Prédiction des Classes d'Objets

En plus des boîtes englobantes, chaque cellule prédit aussi la classe d'objets pour chaque boites. Cela signifie que si une cellule prédit qu'il y a un objet, elle la classe possible de cet objet.

5. Calcul des Scores de Confiance Spécifiques à la Classe

À l'étape de test, pour chaque boîte englobante prédite, YOLO calcule les scores de confiance spécifiques à chaque classe en multipliant les probabilités conditionnelles des classes par les prédictions de confiance des boîtes basé sur l'IOU. La formule est :

$$\Pr(\text{Class}_i|\text{Object}) * \Pr(\text{Object}) * \text{IOU}_{\text{pred}}^{\text{truth}} = \Pr(\text{Class}_i) * \text{IOU}_{\text{pred}}^{\text{truth}}$$

Cela donne des scores de confiance spécifiques à chaque classe pour chaque boîte. Ces scores montrent à quel point le modèle est sûr que l'objet dans la boîte appartient à une certaine classe.

6. Suppression Non-Maximale

Parfois, plusieurs boîtes peuvent détecter le même objet. La technique de suppression non-maximale est utilisée pour éliminer les boîtes les moins probables et ne garder que la meilleure boîte pour chaque objet. La NMS fonctionne en sélectionnant la boîte englobante avec le score de confiance le plus élevé et en supprimant toutes les autres boîtes qui ont une forte superposition (IOU élevé) avec cette boîte sélectionnée.

2.3 Traitement de données

Afin d'avoir une solution performante, il était essentiel de l'entrainer sur des données spécifiques. Nous avons donc été cherché un dataset sur la plateforme « roboflow » reprenant un tas d'images de pleins de matchs différents. Nous aurions pu créé nous même notre propre dataset mais cela aurait prit un temps considérable et peu pertinent pour cette recherche qui a pour but démontre qu'il est possible de détecter des acteurs d'un match.

2.3.1 Base de données

Afin d'entraîner notre modèle pour qu'il soit plus performant, nous avons limité sa reconnaissance à 4 classes : les joueurs, le ballon, l'arbitre et le gardien. Sur le modèle YOLO pré-entrainé de base, plus de 80 classes sont créées. Mais le fait d'avoir trop de classes aurait perturbé la reconnaissance spécifique. Nous avons donc utilisé un dataset avec plusieurs centaines d'images annotées contenant seulement ces 4 classes.

Nous allons illustrer directement une image de notre échantillon afin de baser nos explications sur cette image.

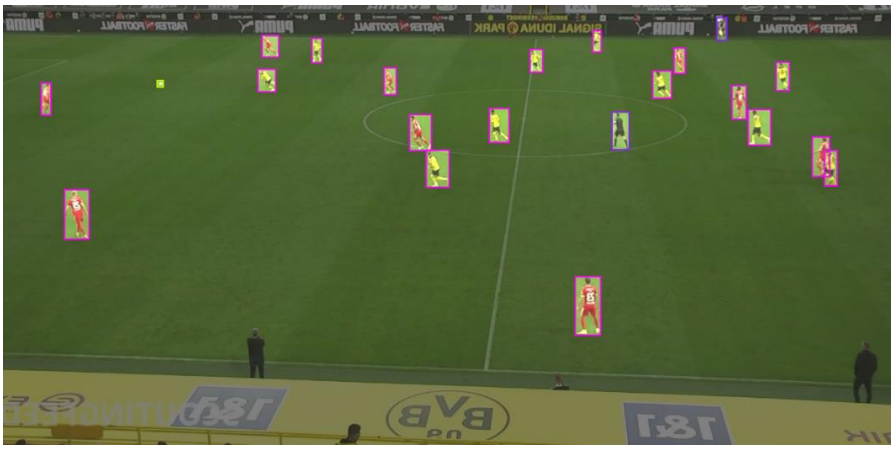

Figure 2.11 : Annotations des différents acteurs/objets d'un match

Vous pouvez voir qu'un humain à lui-même encadré différentes choses (voir figure 2.11). Il a encadré l'arbitre, les joueurs, la balle et on ne le voit pas sur ce plan mais dans d'autres images, il a également encadré le gardien. Un code s'est ensuite créé derrière avec des différents « labels » (player,ball,etc), les positions et d'autres paramètres que nous ne développerons pas ici. C'est avec ces données que l'on va entrainer notre intelligence artificielle.

L'utilisation d'un ensemble de données conséquent, comprenant des centaines d'images, est cruciale pour assurer la précision de l'apprentissage de notre intelligence artificielle. Chaque image supplémentaire contribue à enrichir la variété des scénarios rencontrés par le modèle lors de son entraînement, améliorant ainsi sa capacité à généraliser et à détecter avec précision les objets cibles dans de nouvelles situations.

2.3.2 Entrainement de l'IA

Entrainer notre IA a été la clé de la réussite de notre projet. Sans entrainement, un réseau de neurones est incapable d'être performant et de reconnaitre des objets. Il a donc fallu sur base de notre dataset entrainer notre modèle à reconnaitre et classifier les différentes classes. Pour cela, le dataset a été organisé de manière à séparer l'ensemble des images en trois types de données. Les données de train, valid et test.

Les données de train servent à entrainer le modèle. C'est grâce à cet ensemble de données que le modèle apprend à reconnaitre les motifs et ajuste ses paramètres pour minimiser l'erreur de prédiction. Pendant l'entrainement, à intervalles régulier, le modèle est évalué sur les données de validation. Cette phase sert à évaluer la performance du modèle pendant l'entraînement pour

ajuster les hyperparamètres et surveiller le surapprentissage. Cela permet de mesurer la capacité du modèle à généraliser sur des données qu'il n'a pas vues pendant l'entraînement. Une fois l'entraînement terminé, le modèle est évalué sur un ensemble de données totalement nouveau, séparé des ensembles de train et de validation. On l'appelle la phase de test. Les performances sur les données de test donnent une estimation de la capacité du modèle à généraliser sur des données inconnues. Cette phase fournit une estimation honnête de la performance du modèle en conditions réelles, sans aucun biais d'entraînement.

Pour notre première analyse vidéo avec notre modèle YOLO non-entraîné, nous avons utilisé le modèle YOLOv8x de base pré-entraîné. Ce modèle initial reconnaissait seulement les classes standards parmi les 80 disponibles dans YOLO. Voici une capture d'écran du résultat de cette première tentative (figure 2.12). La vidéo d'une durée de 5 secondes a été coupée en 439 images. Nous avons pris une de ces images pour illustrer nos explications.

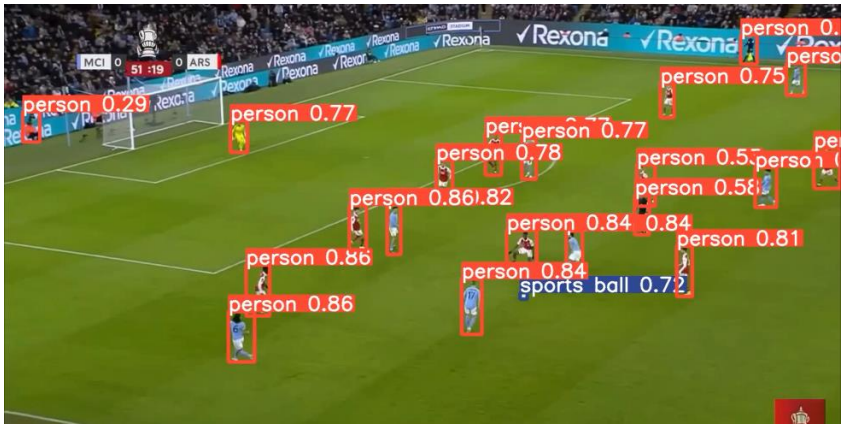

Figure 2.12 : Détection sans entrainement (seulement le modèle pré-entrainé YOLO)

Sur cette image initiale, le modèle a réussi à identifier seulement deux classes : des personnes et une balle de sport. Les chiffres que vous observez représentent la confiance de la prédiction. Ainsi, lorsque la boîte englobante affiche « person 0.85 », cela signifie que le modèle a identifié la classe « Personne » avec une prédiction fiable à 85%.

Plusieurs problèmes se posent ici et peuvent donc être amélioré.

1er problème : La détection des joueurs

Dans le football, sur le terrain, il y a 3 catégories personnes. Il y a les joueurs, les gardiens et les arbitres. Le problème ici c'est qu'il identifie les différents acteurs du match comme une seule catégorie, la catégorie « personne ». Il ne fait pas la différence entre les 3 catégories. De plus, il

identifie des stewards sur le côté également comme des personnes. L'objectif ici est de juste prendre en compte les acteurs du terrain.

2ème problème : La fiabilité

La fiabilité est déjà assez bonne mais peut être améliorée.

Pour améliorer notre intelligence artificielle, nous allons maintenant expliquer comment nous avons procédé.

Dans un premier temps, sur base du dataset avec les 3 jeux de données (train,valid,test), nous avons fait tourner ces dernières dans un premier temps sur 10 époques. Les époques sont quelque chose de fondamental dans l'entrainement de notre modèle. Une époque correspond à une itération complète à travers l'ensemble des données d'entraînement. Cela signifie que toutes les images dans l'ensemble de données sont utilisées pour ajuster les poids du modèle. L'entraînement en plusieurs époques permet au modèle d'apprendre progressivement à partir des données, en les observant dans un ordre aléatoire pour aider à généraliser et à éviter le surapprentissage. Augmenter le nombre d'époques est souvent crucial pour améliorer la précision du modèle en lui donnant plus de temps pour ajuster ses paramètres et améliorer sa capacité à détecter correctement les objets. Voici les résultats sur 10,25, et 75 époques.

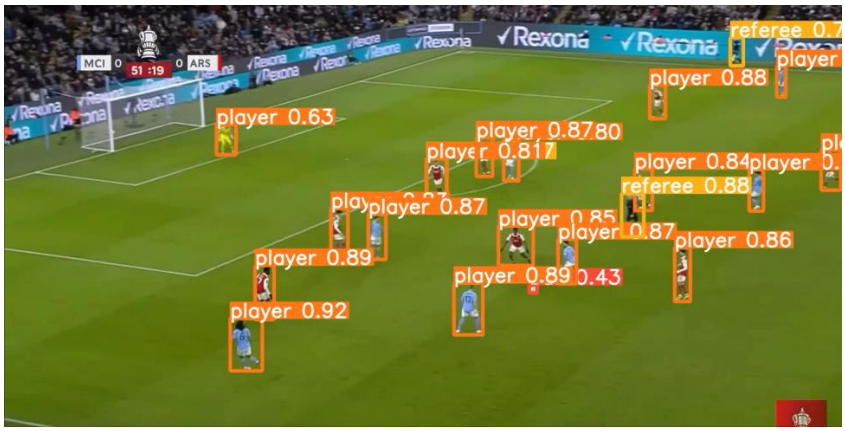

Figure 2.13 : Détection après un entrainement de 10 époques

On peut observer qu'avec un entrainement de 10 époques, la reconnaissance est déjà améliorée (voir figure 2.13). Le modèle a été suffisamment entrainé pour reconnaitre les différents joueurs avec un indice de confiance moyen de 86% pour les joueurs. C'est une bonne moyenne pour un début. La balle est quant à elle également identifié avec une confiance de 43%. Les arbitres que l'on voit ici en tunique noir sont également classifier à 88%. Le problème qui persiste est que le gardien est considéré comme un joueur. De plus on voit bien que le modèle n'est pas sûr car il identifie le gardien comme un joueur mais avec une confiance de 63%, ce qui est largement

inférieur à la moyenne des bonnes prédictions de joueurs. L'indice de confiance toute classification confondue est estimé à 78%.

Figure 2.14 : Détection après un entrainement de 25 époques

Après un entrainement de 25 époques (figure 2.14), on voit que les indices de confiances ont légèrement augmenté pour les joueurs. Elle passe à 88% en moyenne. Pour l'arbitre elle se situe cette fois à 84%, une petite baisse qui n'est pas inquiétante et quoi doit être due au fait que les modèle à ajusté ses poids différemment. La moyenne de la reconnaissance de l'arbitre sur les 439 images est supérieure par rapport à avant donc pas de souci. La balle est identifiée à 76 %. Cependant, cette fois le gardien est identifié comme un gardien avec un indice de confiance à 88%. On voit déjà une amélioration dans la classification des différents acteurs du jeu. Un indice de confiance général (toute classification confondue) qui passe de 78% avec 10 époques à 87 % avec 25 époques.

Figure 2.15 : Détection après un entrainement de 75 époques

Sur cette image (figure 2.15), on peut voir que la moyenne des prédictions des joueurs est restée à 88% dû à la tendance logarithme qu'à l'évolution d'un modèle, on commence à stagner. Le gardien est lui maintenant prédit avec une confiance de 95%. La balle est également restée à 76%. L'arbitre lui a pris un peu plus de confiance est estimé à 91%.

Nous tenons à préciser que nous aurions pu aller encore plus loin dans nos époques d'entrainement mais il faut savoir qu'entrainer trop le modèle peut avoir un effet de surapprentissage et donc le modèle aura plus de mal à généraliser sur de données non vues. Pendant l'entrainement de notre modèle, une fois les 75 époques atteintes, le modèle n'évoluait plus et nous avons donc décidé d'arrêter l'entrainement après 125 époques étant donné que cela faisait 50 époques qu'il ne progressait plus. Nous n'avons volontairement pas pris le modèle après 125 époques étant donné que le meilleur modèle était identifié (avoir le meilleur ajustement de poids W) était à 75 époques. Pour augmenter la confiance, un dataset beaucoup plus conséquent aurait fait la différence.

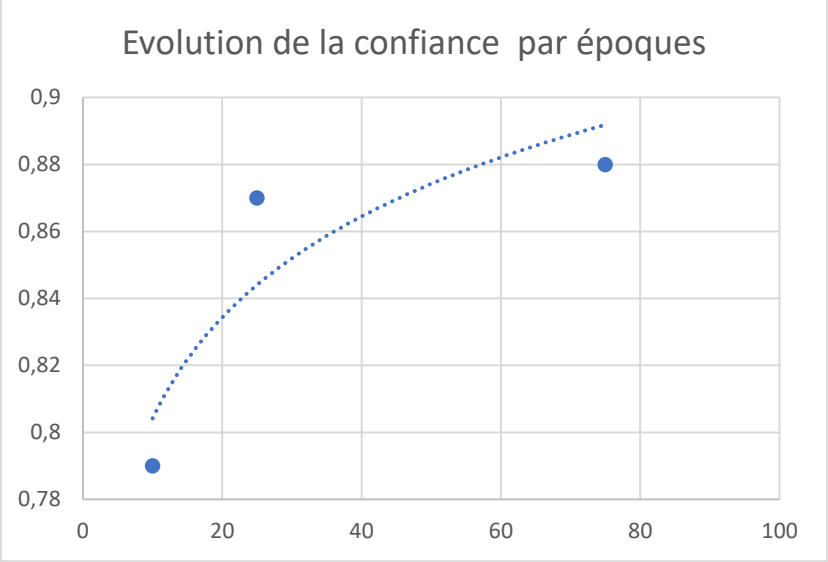

Figure 2.16 : Graphique montrant l'évolution de la confiance par époques

On peut voir sur ce graphique (figure 2.16), l'évolution de l'indice de confiance général en fonction du nombre d'époques. On observe naturellement une relation positive entre le nombre d'entrainement et l'indice de confiance mais déjà avec 3 points de données, on peut observer que la tendance suit celle d'une tendance logarithme. En effet au début les performances augmentent mais à la fin on finit par stagner et ici une fois arrivé autour des 75 époques, on stagne. Cela s'appelle le point de saturation [29].

2.4 Tests et résultats

Dans cette section, vous retrouvez les différents résultats de notre contribution.

2.4.1 Résultats d'un point de vue technologique

Comme précité au début de notre partie contribution, à l'heure d'aujourd'hui, les statistiques sont notés par des humains. Nous allons expliquer plus en détail le process et par la suite valoriser et montrer ce que peut apporter notre solution d'un point de vue technologique.

Lors d'un match, prenons un match X se déroulant à 18h. Une équipe d'annotateurs va se mettre en place au même moment. Une fois que le match débute, ce dernier va être retransmis en direct à l'équipe d'annotateurs qui se trouve généralement à l'autre bout de la planète pour des raisons budgétaires. Cette équipe va noter ce que l'on appelle des événements grâce à des boutons et des raccourcis. Derrière cela, va sortir un dataset brut avec tous les événements. À chaque passe, contrôle, tir, etc., les annotateurs vont créer un événement avec la minute et la seconde précises de l'événement, la position exacte sous forme de coordonnées x,y, l'id du joueur, le type d'événement, etc. Voici un exemple :

Figure 2.17 : Action de match (Chaine Eleven Dazn)

On peut voir ici sur les images (Figure 2.17) le joueur n°47 Amoura faire une passe au n°23 Puertas. Ce dernier tir et marque. Voici Comment l'annotateur a enregistré cette action sous forme de données.

Figure 2.18 : Données Brutes récoltées par les annotateurs

Une fois ces événements notés sous forme de données brutes (figure 2.18), le data analyste intervient afin de traiter ces données pour les rendre lisible pour un entraineur ou la TV.

Comme vous l'avez vu, pour annoter un événement, il faut d'abord savoir reconnaître les joueurs, la balle, etc. L'annotateur doit être capable de différencier un gardien, un joueur, un arbitre, une balle, etc. Notre solution vise à éliminer toute cette main-d'œuvre en mobilisant des annotateurs. En effet, notre intelligence artificielle est capable de reconnaître les joueurs, la balle, le gardien et l'arbitre. Ces quatre éléments sont fondamentaux dans la construction de ce projet. C'est une première étape vers les annotations automatiques. Notre solution a été construite sur un dataset de seulement 350 images annotées, entraîné sur 75 époques. Le modèle a obtenu des résultats concluants et est capable de reconnaître les différents acteurs d'un match avec une fiabilité moyenne de 88 %. Évidemment, il est possible de tendre vers les 100 %, mais l'objectif de cette thèse était de démontrer la faisabilité d'un modèle de détection. Pour augmenter cette fiabilité, il suffit d'augmenter le nombre d'images dans nos données d'entraînement.

De plus, les annotateurs humains peuvent être sujets à des erreurs, des biais et de la fatigue, ce qui peut affecter la précision et la cohérence des données collectées. Une IA bien entraînée peut surpasser ces limitations humaines en offrant une reconnaissance des événements plus précise et cohérente. Dans notre cas, notre modèle a atteint une fiabilité moyenne de 88 % pour la reconnaissance des différents acteurs d'un match, ce qui est déjà très prometteur avec un dataset limité de 350 images annotées et un entraînement sur 75 époques.

Après chaque annotation humaine, l'équipe d'annotateurs regarde à nouveau le match pour vérifier s'ils n'ont pas fait d'erreurs, ce qui prend à nouveau du temps. Une IA peut traiter les données à une vitesse bien supérieure à celle des humains, offrant ainsi des résultats beaucoup

plus rapidement. Un humain va analyser le match à vitesse réelle, alors que l'IA pourra aller beaucoup plus vite et parcourir le match en 5 minutes. Évidemment, cela dépend du matériel de la machine, une bonne RAM est nécessaire

2.4.2 Résultats d'un point de vue management

Ce mode de fonctionnement coûte cher et seuls les clubs professionnels peuvent se le payer. Les clubs de second rang (D2, D3, D4…) ont des budgets restreint et ne savent pas investir autant d'argent que les pros, il faut bien comprendre qu'entre les clubs de 1 ère division et le reste, il y a un monde de différence. C'est pour cette raison qu'avec la solution par intelligence artificielle, cela pourrait drastiquement réduire les coûts des humains qui note ces annotations d'événement.

Les équipes de deuxième division ont des statistiques, elles ne sont pas aussi poussées qu'en D1A mais il y a déjà nombre de statistique différentes. Mais par contre, dès que l'on arrive en D3, plus aucune statistique n'est fournie. Les seuls statistiques que les clubs ont ce sont les données physique donnée par des ceintures cardiaque. Par contre, les matchs sont filmés jusqu'en 6 ème division. C'est une obligation de la fédération belge de football. Toutes les équipes évoluant de la D1 à la D6 ont l'obligation de filmer les matchs avec une caméra et le programme de sportcode « HUDL » et poster la vidéo sur la plateforme « HUDL ».

D'un point de vue managérial, les statistiques peuvent révolutionner la gestion des équipes en influençant divers aspects :

Niveau tactique

Les statistiques permettent aux entraîneurs de prendre des décisions tactiques éclairées. « L'analyse statistique est probablement ces dernières années, la frontière qui a le plus progressé dans le football » [30]. Par exemple, si les données montrent qu'une équipe adverse marque souvent sur des contre-attaques, l'entraîneur pourrait choisir une formation plus défensive pour minimiser ce risque. De même, les statistiques peuvent révéler les points forts et les faiblesses de l'équipe, permettant ainsi de maximiser les points forts et de travailler sur les faiblesses à l'entraînement. Par exemple, si les statistiques montrent que l'équipe est faible dans les phases de jeu arrêtées comme les corners et les coups francs défensifs, l'entraîneur peut planifier des sessions spécifiques pour améliorer ces aspects.

Choix des joueurs

Les statistiques sont cruciales pour prendre des décisions objectives concernant les choix des joueurs. Les données peuvent montrer quels joueurs performent le mieux dans quelles situations. Par exemple, un joueur peut être excellent en attaque mais moins efficace en défense. En analysant

les statistiques, l'entraîneur peut choisir les joueurs les mieux adaptés à une stratégie spécifique pour un match donné. Si les statistiques montrent qu'un joueur réussit un nombre élevé de passes décisives mais qu'un autre a un meilleur taux de conversion des tirs en buts, l'entraîneur peut ajuster la composition de l'équipe en conséquence.

Justification des choix

Les statistiques fournissent des preuves objectives pour justifier les décisions de l'entraîneur. Selon [31], les statistiques améliore la crédibilité. Si un joueur est mis sur le banc, les statistiques peuvent expliquer pourquoi cette décision a été prise. Par exemple, si un joueur a un taux de réussite de passes très bas ou s'il a commis plusieurs erreurs défensives lors des derniers matchs, il sera plus facile pour l'entraîneur d'expliquer cette décision au joueur. Cela permet non seulement de justifier les choix auprès des joueurs, mais aussi de maintenir la transparence et la confiance au sein de l'équipe.

Acceptation des décisions par les joueurs

Lorsqu'un joueur sait que ses statistiques ne sont pas bonnes, il est plus susceptible d'accepter la décision de l'entraîneur de le mettre sur le banc ou de le faire jouer à un poste différent. Par exemple, si les données montrent qu'un joueur a un taux de passes réussies de seulement 60% comparé à la moyenne de l'équipe qui est de 85%, ce joueur comprendra mieux pourquoi il n'est pas titulaire. Cela peut également servir de motivation pour les joueurs, les incitant à s'améliorer dans les domaines où ils sont faibles.

Optimisation des entraînements

Les statistiques peuvent aussi aider à personnaliser les entraînements en fonction des performances individuelles. Par exemple, si les données montrent qu'un joueur a du mal à récupérer le ballon dans certaines situations, des exercices spécifiques peuvent être intégrés à son programme d'entraînement pour améliorer cette compétence. Cela permet aux joueurs de progresser plus rapidement et de manière plus ciblée

Avec notre solution d'intelligence artificielle, cela pourrait avoir un impact énorme sur le management des équipes dans les divisions inférieures.

Interview d'un club

Pour évaluer l'intérêt des clubs de second rang pour un logiciel basé sur l'IA, nous avons mené plusieurs entretiens avec des professionnels et semi-professionnel du football belge. Parmi eux, nous avons eu la chance d'interrogé Hubert Stas, directeur technique et ancien entraîneur du

Rupel Boom FC, club évoluant en quatrième division mais qui était il y a 2 ans en 3 ème division. Hubert Stas était donc la personne idéale étant donné sa connaissance des divisions inférieurs.

Voici un résumé de ses réponses à trois questions fondamentales :

1. **Intérêt pour un logiciel statistique :**

- *Question : Est-ce qu'un logiciel statistique intéresserait l'équipe de Rupel Boom FC ? Serait-ce utile à ce niveau-là ?*

- *Réponse : « Oui, un logiciel statistique serait évidemment très utile pour mon équipe. Je vais te donner un exemple : si je sais que l'équipe contre qui on va jouer ce week-end a tendance à marquer beaucoup sur corner, je vais donc cette semaine plus travailler les phases de corner défensives à l'entraînement. Donc oui, évidemment que ce serait utile, rien que pour savoir également les performances techniques de mes joueurs. »*

2. **Impact sur le management:**

- *Question : Qu'est-ce que l'ajout d'un logiciel comme celui-ci pourrait changer dans le management de votre équipe ?*

- *Réponse : « Actuellement, les décisions du coach sont prises sur la base de ce que l'on voit à l'entraînement, des données physiques et des performances sur le terrain. Mais il y a évidemment une part de subjectivité là-dedans ; l'entraîneur ne peut analyser le match de chaque joueur à tout instant. Ce logiciel permettrait de visualiser les performances techniques de chaque joueur individuellement sur l'ensemble d'un match, puis de prendre des décisions basées sur ces données, comme changer le joueur de poste ou le mettre sur le banc. Ces statistiques permettraient aussi de concevoir des entraînements plus personnalisés pour chaque joueur, ce qui favoriserait leur progression. »*

3. **Budget alloué:**

- *Question : Combien serais-tu disposé à mettre par an dans un outil comme celui-ci ?*

- *Réponse : « Je ne vais pas parler en termes de montant précis, mais plutôt en termes de pourcentage. Je pense que je serais prêt à consacrer 10 % de mon budget à un tel outil. »*

Après avoir sondé plusieurs clubs, nous nous sommes rendu compte de l'intérêt que portent les clubs à un logiciel comme celui-ci. D'après les personnes que nous avons interrogées et selon nos estimations, les clubs seraient prêts à investir entre 10 000 et 20 000 € par an.

2.4.3 Résultats du modèle sur un match amateur

Figure 2.19 : Modèle appliqué sur un match amateur

Appliquons notre modèle sur un match amateur filmé, on peut observer que le modèle basé sur l'IA arrive à détecter les joueurs et la balle (voir figure 2.19). Les scores de confiance sont évidemment plus faibles. Cela s'explique par le fait que le modèle a été entraîné sur des images de matchs professionnels.

Dans les matchs professionnels, les uniformes, le terrain et les conditions de jeu sont très standardisés, ce qui aide le modèle à faire des prédictions précises. En revanche, dans les matchs amateurs, il y a plus de variations dans les uniformes, les couleurs et les mouvements des joueurs, ainsi que dans les conditions de jeu. Ces différences rendent la détection plus difficile pour le modèle, ce qui fait baisser les scores de confiance d'environ 88% à 60%.

Pour améliorer les performances du modèle sur les matchs amateurs, il serait important d'ajouter des images de matchs amateurs à l'ensemble de données d'entraînement. En incluant ces nouvelles données, le modèle pourra mieux reconnaître et s'adapter aux spécificités des matchs amateurs, augmentant ainsi sa précision et sa fiabilité.

2.*5* Discussion

L'objectif principal du modèle est de détecter de manière fiable les différents acteurs d'un match de football, notamment les joueurs, le ballon, le gardien et l'arbitre. Ce modèle a été entraîné principalement sur des images de matchs professionnels, ce qui a conduit à certaines limites significatives, surtout lorsqu'il est appliqué à des contextes de matchs semi-professionnels ou amateurs.

Contrainte liée au Dataset

La première limite notable réside dans le choix du dataset. Le modèle a été entraîné sur des images provenant de matchs professionnels en raison des contraintes de temps et de disponibilité des données. Trouver et annoter des images issues de nombreux matchs semi-professionnels aurait été extrêmement chronophage. De plus, les images de matchs professionnels sont plus accessibles publiquement, ce qui a facilité l'acquisition des données. Cependant, cette approche a entraîné une diminution de la performance du modèle sur des matchs amateurs, comme l'a démontré la comparaison des résultats.

Pour remédier à cette faiblesse, il sera essentiel, à l'avenir, de diversifier les données d'entraînement en incluant des images de matchs de niveaux inférieurs, tels que les divisions D3, D4 et D5. Cela permettra au modèle de généraliser ses capacités de détection à une plus grande variété de contextes de jeu.

Identification du Ballon

Une autre limite importante concerne l'identification du ballon. Le modèle éprouve des difficultés à reconnaître le ballon lorsqu'il est en mouvement rapide, en raison de la déformation visuelle qui rend le ballon moins distinct. De plus, lorsqu'il est caché derrière un joueur ou un groupe de joueurs, l'identification devient encore plus complexe. Un autre [32]étude confirme les même difficultés éprouvées dans l'identification du ballon.

Afin d'améliorer cette situation, il serait nécessaire d'augmenter le volume de données d'entraînement en incluant des images du ballon en mouvement rapide et déformé. L'ajout de plusieurs caméras (2 à 3) sur le terrain pourrait également offrir une vue plus complète et continue du ballon, réduisant ainsi les cas où il est temporairement caché.

Limites liées à l'Entraînement et aux Compétences

Le temps et les ressources limitées ont également influencé l'avancement de l'entraînement du modèle. Alors qu'une fiabilité proche des 100% est théoriquement atteignable avec plusieurs milliers de données et une recherche approfondie, un mémoire de niveau master n'a pas permis d'atteindre ce niveau de perfection. Un projet de cette envergure nécessiterait des ressources considérables, potentiellement dans le cadre d'une thèse de doctorat.

De plus, nos connaissances limitées en codage et en intelligence artificielle ont constitué une autre contrainte. Étant donné mon parcours en sciences de gestion, la création de statistiques à partir du modèle de reconnaissance s'est avérée complexe. Ce travail serait plus adapté à un ingénieur en informatique, spécialisé dans le traitement des données et la création de statistiques.

Limites de l'IA par rapport à la Gestion Humaine

Il est également important de souligner que, bien que les statistiques générées par l'IA puissent améliorer la gestion d'une équipe, elles ne peuvent pas remplacer certains aspects intangibles du management humain. Le sentiment intuitif du potentiel d'un joueur, la gestion des émotions et des dynamiques d'équipe sont des domaines où l'IA ne peut pas rivaliser avec l'expertise et l'expérience humaines.

Conclusion

Cette thèse avait pour but principal de répondre à la question suivante : **« est-il possible de créer une solution capable de détecter automatiquement les différents acteurs d'un match grâce à l'IA ? ».** Nous pouvons répondre oui. En effet, le modèle de détection présenté tout au long de cette thèse montre clairement des résultats positifs et encourageants. Les réseaux de neurones CNN sont capables de détecter les différents acteurs d'un match, notamment les joueurs, l'arbitre, le gardien de but et le ballon. L'utilisation du modèle de détection d'objets YOLO (You Only Look Once), que nous avons entraîné sur un ensemble de données bien précis composé d'images annotées manuellement, a permis d'atteindre une fiabilité de 88 %. Les résultats obtenus montrent qu'il est possible de développer des systèmes de détection automatiques fiables. Cela met également en lumière le potentiel transformateur de cette technologie pour le monde du sport. À mesure que les techniques de Machine Learning continuent de progresser, il est envisageable que des modèles encore plus performants puissent être développés, augmentant ainsi l'efficacité et l'application des systèmes de détection automatique dans divers contextes sportifs.

Ce projet ouvre la voie à de nouvelles opportunités pour les équipes de football semi-professionnelles, leur offrant la possibilité d'améliorer leurs performances, la gestion du recrutement et les stratégies tactiques grâce à des analyses de données sophistiquées. Le travail pour commercialiser cette solution est encore long, car la prochaine étape sera de partir de l'identification des différents acteurs réalisée ici pour ensuite construire toute une série de statistiques. Nous avons également remarqué un engouement de la part des clubs rencontrés, avec certains clubs prêts à investir 10 à 20 000 € par an dans ce genre d'outil si le programme voit le jour.

Comme nous l'avons longuement souligné tout au long de cette thèse, il est évident que la technologie améliore la gestion d'une équipe, mais elle ne peut pas remplacer certains aspects intangibles du management humain. En effet, l'humain, fort de son expérience, possède une flexibilité et une capacité d'adaptation supérieures face à de nouvelles situations. Il est donc crucial de favoriser une bonne collaboration entre les experts en intelligence artificielle, les analystes sportifs et les décideurs des clubs pour affiner et adapter ces technologies aux besoins spécifiques des équipes.

Pour aller plus loin, des recherches supplémentaires pourraient être faites afin d'appliquer le modèle à d'autres sports.

En conclusion, cette thèse a démontré la faisabilité de créer une solution capable de détecter automatiquement les différents acteurs d'un match grâce à l'IA, en utilisant des réseaux de neurones CNN et le modèle YOLO. Nous espérons que, dans le futur, les clubs professionnels pourront bénéficier de ces perspectives prometteuses.

References

[1] V. Beygmohammadloo, D. Gharakhani, et M. Naderinasab, « Development of Football Tactics and Strategies Driven by Artificial Intelligence », *AI and Tech in Behavioral and Social Sciences*, vol. 2, nº 2, p. 1-6, 2024. [En ligne]. Disponible sur: https://journals.kmanpub.com/index.php/aitechbesosci/article/view/2611

[2] R. Rein et D. Memmert, « Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science », *SpringerPlus*, vol. 5, nº 1, p. 1410, déc. 2016, doi: 10.1186/s40064-016-3108-2.

[3] V. R. A. Cossich, D. Carlgren, R. J. Holash, et L. Katz, « Technological Breakthroughs in Sport: Current Practice and Future Potential of Artificial Intelligence, Virtual Reality, Augmented Reality, and Modern Data Visualization in Performance Analysis », *Applied Sciences*, vol. 13, nº 23, Art. nº 23, janv. 2023, doi: 10.3390/app132312965.

[4] G. Fialho, A. Manhães, et J. P. Teixeira, « Predicting sports results with artificial intelligence–a proposal framework for soccer games », *Procedia Computer Science*, vol. 164, p. 131-136, 2019. [En ligne]. Disponible sur: https://www.sciencedirect.com/science/article/pii/S1877050919322033

[5] R. P. Bunker et F. Thabtah, « A machine learning framework for sport result prediction », *Applied computing and informatics*, vol. 15, nº 1, p. 27-33, 2019. [En ligne]. Disponible sur: https://www.sciencedirect.com/science/article/pii/S2210832717301485

[6] N. Razali, A. Mustapha, S. Utama, et R. Din, « A review on football match outcome prediction using bayesian networks », in *Journal of Physics: Conference Series*, IOP Publishing, 2018, p. 012004. Disponible sur: https://iopscience.iop.org/article/10.1088/1742-6596/1020/1/012004/meta

[7] S. Moustakidis, S. Plakias, C. Kokkotis, T. Tsatalas, et D. Tsaopoulos, « Predicting football team performance with explainable ai: Leveraging shap to identify key team-level performance metrics », *Future Internet*, vol. 15, nº 5, p. 174, 2023. [En ligne]. Disponible sur: https://www.mdpi.com/1999-5903/15/5/174

[8] C. Gottschalk, S. Tewes, et B. Niestroj, « The innovation of refereeing in football Through AI », *International Journal of Innovation and Economic Development*, vol. 6, nº 2, p. 35-54, 2020,. Disponible sur: https://ideas.repec.org/a/mgs/ijoied/v6y2020i2p35-54.html

[9] T. Joris et À. F. Université de Liège > Master droit, « L'impact de l'intelligence artificielle sur l'arbitrage et sa procédure », janv. 2020, [En ligne]. Disponible sur: https://matheo.uliege.be/handle/2268.2/9283

[10] J. Pavitt, D. Braines, et R. Tomsett, « Cognitive analysis in sports: Supporting match analysis and scouting through artificial intelligence », *Applied AI Letters*, vol. 2, nº 1, p. e21, mars 2021, doi: 10.1002/ail2.21.

[11] « Opta Stats Hub 2023-24 | The Analyst ». [En ligne]. Disponible sur: https://theanalyst.com/2023/08/2023-24-football-stats-opta-data/

[12] S. Barra, S. M. Carta, A. Giuliani, A. Pisu, A. S. Podda, et D. Riboni, « FootApp: An AI-powered system for football match annotation », *Multimed Tools Appl*, vol. 82, nº 4, p. 5547-5567, févr. 2023, doi: 10.1007/s11042-022-13359-0.

[13] A. García-Aliaga, M. Marquina, J. Coterón, A. Rodríguez-González, et S. Luengo-Sánchez, « In-game behaviour analysis of football players using machine learning techniques based on player statistics », *International Journal of Sports Science & Coaching*, vol. 16, n° 1, p. 148-157, févr. 2021, doi: 10.1177/1747954120959762.

[14] W. Lu et J. Hou, « Intelligent Recognition Method of Athlete Wrong Movement Based on Image Vision », *Scientific Programming*, vol. 2021, n° 1, p. 8467906, 2021, doi: 10.1155/2021/8467906.

[15] H. Hakim et A. Fadhil, « Survey: Convolution neural networks in object detection », in *Journal of Physics: Conference Series*, IOP Publishing, 2021, p. 012095. [En ligne]. Disponible sur: https://iopscience.iop.org/article/10.1088/1742-6596/1804/1/012095/meta

[16] « schéma deep learning bulle », Bing.. [En ligne]. Disponible sur: https://www.bing.com/images/search?q=schéma+deep+learning+bulle&FORM=HDRSC3

[17] T. Wang et T. Li, « Deep Learning-Based Football Player Detection in Videos », *Computational Intelligence and Neuroscience*, vol. 2022, n° 1, p. 3540642, 2022, doi: 10.1155/2022/3540642.

[18] C. A. M. Husein, « Detecting and Tracking Player in Football Videos Using Two-Stage Mask R-CNN Approach », *Conference Series*, 2023, doi: 10.34306/conferenceseries.v4i1.643.

[19] P. Broda, « La mesure dans le sport et ses limites: le cas des statistiques dans le basketball professionnel français », 2013, [En ligne]. Disponible sur: https://mpra.ub.uni-muenchen.de/id/eprint/50767

[20] D. Yarrow et M. Kranke, « The performativity of sports statistics: towards a research agenda », *Journal of Cultural Economy*, vol. 9, n° 5, p. 445-457, sept. 2016, doi: 10.1080/17530350.2016.1202856.

[21] « Rátgéber et al., 2013Comparative Review of Statistical... - Google Scholar »[En ligne]. Disponible sur: https://scholar.google.com/scholar?hl=fr&as_sdt=0%2C5&q=R%C3%A1tg%C3%A9ber+et+al.%2C+2013Comparative+Review+of+Statistical+Parameters+for+Men%27s+and+Women%27s+Basketball+Leagues+in+Serbia&btnG=

[22] J. A. M. Sidey-Gibbons et C. J. Sidey-Gibbons, « Machine learning in medicine: a practical introduction », *BMC Med Res Methodol*, vol. 19, n° 1, p. 64, déc. 2019, doi: 10.1186/s12874-019-0681-4.

[23] S.-H. Han, K. W. Kim, S. Kim, et Y. C. Youn, « Artificial neural network: understanding the basic concepts without mathematics », *Dementia and neurocognitive disorders*, vol. 17, n° 3, p. 83-89, 2018, [En ligne]. Disponible sur: https://synapse.koreamed.org/articles/1120136

[24] A. Singh, K. K. Srivastava, et H. Murugan, « Speech Emotion Recognition Using Convolutional Neural Network (CNN) », vol. 24, n° 08, 2020.

[25] « Réseau de neurones artificiels - Modèle », Techno-Science.net. [En ligne]. Disponible sur: https://www.techno-science.net/glossaire-definition/Reseau-de-neurones-artificiels-page-4.html

[26] « RELU-Function and Derived Function Review | SHS Web of Conferences ».. [En ligne]. Disponible sur: https://www.shs-conferences.org/articles/shsconf/abs/2022/14/shsconf_stehf2022_02006/shsconf_stehf2022_02006.html

[27] J. Redmon, S. Divvala, R. Girshick, et A. Farhadi, « You only look once: Unified, real-time object detection », in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, p. 779-788. [En ligne]. Disponible sur: https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html

[28] J. Terven, D.-M. Córdova-Esparza, et J.-A. Romero-González, « A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas », *Machine Learning and Knowledge Extraction*, vol. 5, n° 4, p. 1680-1716, 2023,. [En ligne]. Disponible sur: https://www.mdpi.com/2504-4990/5/4/83

[29] O. G. Ajayi et J. Ashi, « Effect of varying training epochs of a Faster Region-Based Convolutional Neural Network on the Accuracy of an Automatic Weed Classification Scheme », *Smart Agricultural Technology*, vol. 3, p. 100128, févr. 2023, doi: 10.1016/j.atech.2022.100128.

[30] A. Drivet et A. Drivet, « From the Data to the Statistical Analysis of Football: The Case of the Italian Serie A League », https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-7998-3473-1.ch051. [En ligne]. Disponible sur: https://www.igi-global.com/gateway/chapter/www.igi-global.com/gateway/chapter/263575

[31] D. Hahn, « The Effect of Statistics on Enjoyment and Perceived Credibility in Sports Media », *Communication & Sport*, vol. 11, n° 1, p. 53-71, févr. 2023, doi: 10.1177/2167479521998395.

[32] « Applied Sciences | Free Full-Text | A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions ». [En ligne]. Disponible sur: https://www.mdpi.com/2076-3417/12/9/4429

[33] « YOLO Object Detection Explained: A Beginner’s Guide ». [En ligne]. Disponible sur: https://www.datacamp.com/blog/yolo-object-detection-explained

Annexes

1) Code python pour la détection :

code_detection.py > ...

```python
import os
from ultralytics import YOLO

# Définir les chemins du modèle et de la vidéo
chemin_modèle = 'modèle_entrainé/epoch75.pt'
chemin_video = 'vidéo/raal.mp4'

# Vérifier si le fichier du modèle existe
if not os.path.exists(chemin_modèle):
    raise FileNotFoundError(f"Le fichier de modèle {chemin_modèle} est introuvable.")

# Charger le modèle YOLO pré-entraîné pour la détection des joueurs de football
model = YOLO(chemin_modèle)

# Vérifier si le fichier vidéo existe
if not os.path.exists(chemin_video):
    raise FileNotFoundError(f"Le fichier vidéo {chemin_video} est introuvable.")

# Fonction pour charger et prédire les objets dans la vidéo
def yolo_predict(model, chemin_video):
    # Prédire les objets dans la vidéo et sauvegarder les résultats
    predictions = model.predict(chemin_video, save=True)
    return predictions

# Fonction pour afficher les résultats complets de la première image
def results(predictions):
    print("Résultats complets pour la première image :")
    print(predictions[0])
    print('====================================')

# Fonction pour afficher chaque boîte détectée
def print_boxes_détectées(predictions):
    for box in predictions[0].boxes:
        print(box)

# Fonction principale pour exécuter toutes les étapes de détection
def detection():
    # Charger et prédire les objets dans la vidéo
    predictions = yolo_predict(model, chemin_video)

    # Afficher les résultats complets de la première image
    results(predictions)

    # Afficher chaque boîte détectée
    print_boxes_détectées(predictions)

if __name__ == "__main__":
    detection()

```