



دانشگاه صنعتی شریف  
دانشکده‌ی مهندسی کامپیوتر

پایان‌نامه‌ی کارشناسی ارشد  
هوش مصنوعی و رباتیکز

عنوان:

# ردیابی ژرف چندشیء با بازشناسایی پاره-محور در مسابقات فوتبال

نگارش:

امیر محمد منصوریان

استاد راهنما:

دکتر شهره کسائی

مرداد ۱۴۰۲





## اظہارنامہ

(اصالت متن و محتوای پایان نامہ کارشناسی ارشد)

عنوان پایان نامہ: .....

نام استاد راهنما: .....

نام استاد راهنمای همکار: .....

نام استاد مشاور: .....

این جانب .....

اظہار می دارم:

۱- متن و نتایج علمی ارائه شده در این پایان نامہ اصیل بوده و منحصرأً توسط این جانب و زیر نظر استادان (راهنما، همکار و مشاور) نام برده شده در بالا تهیه شده است.

۲- متن پایان نامہ به این صورت در هیچ جای دیگری منتشر نشده است.

۳- متن و نتایج مندرج در این پایان نامہ، حاصل تحقیقات این جانب به عنوان دانشجوی کارشناسی ارشد دانشگاه صنعتی شریف است.

۴- کلیه مطالبی که از منابع دیگر در این پایان نامہ مورد استفاده قرار گرفته، با ذکر مرجع مشخص شده است.

نام دانشجو: .....

تاریخ

امضا

نتایج تحقیقات مندرج در این پایان نامہ و دستاوردهای مادی و معنوی ناشی از آن (شامل فرمول‌ها، توابع کتابخانه‌ای، نرم افزارها، سخت افزارها و مواردی که قابلیت ثبت اختراع دارد) متعلق به دانشگاه صنعتی شریف است. هیچ شخصیت حقیقی یا حقوقی بدون کسب اجازه از دانشگاه صنعتی شریف حق فروش و ادعای مالکیت مادی یا معنوی بر آن یا ثبت اختراع از آن را ندارد. همچنین کلیه حقوق مربوط به چاپ، تکثیر، نسخه برداری، ترجمه، اقتباس و نظائر آن در محیط‌های مختلف اعم از الکترونیکی، مجازی یا فیزیکی برای دانشگاه صنعتی شریف محفوظ است. نقل مطالب با ذکر ماخذ بلامانع است.

نام استادان راهنما: .....

نام دانشجو: .....

تاریخ

تاریخ

امضا

امضا

## تصویب نامه

به نام خدا  
دانشگاه صنعتی شریف  
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد

عنوان:

نگارش:

کمیته ممتحنین

استاد راهنما:

امضاء

استاد ممتحن داخلی:

امضاء

استاد ممتحن خارجی:

امضاء

اول از همه از استاد بزرگواریم دکتر کسائی به خاطر راهنمایی‌ها و کمک‌هایشان، چه در زمینه این تحقیق و چه خارج از آن، کمال تشکر و قدردانی را دارم.

همچنین از دوستان خوبم در آزمایشگاه پردازش تصویر به خاطر راهنمایی‌ها و لحظات خوشی که در طول یک سال اخیر برای بنده به وجود آورده‌اند تشکر بسیار دارم.

در نهایت از خانواده‌ی مهربانم سپاسگزارم که همانند سایر مراحل زندگی‌ام همواره حامی و پشتیبان من بوده‌اند، به خصوص خواهر عزیزم که در دوران تحصیلم همیشه راهنما، مشوق و همراه من بوده‌است.

## چکیده

ردیابی و بازشناسایی موثر افراد برای تجزیه و تحلیل ویدیوهای ورزش تیمی امری ضروری است. این کار اما به دلیل حرکت غیرخطی بازیکنان در مقایسه با عابر پیاده، شباهت ظاهری بازیکنان یک تیم، فاصله دوربین از افراد حاضر در زمین و انسدادهای مکرر، کاملاً چالش برانگیز است. بنابراین، توانایی استخراج بازنمایی‌های معنادار برای نمایش افراد در توسعه یک سیستم ردیابی و بازشناسایی موثر بسیار مهم است. در ورزش های تیمی اما، اطلاعات دیگری وجود دارد که می تواند در بازشناسایی افراد استفاده شود، مانند وابستگی تیم، اطلاعات نقش هر فرد و شماره پیراهن. با ظهور شبکه های عصبی پیچشی ژرف و پیشرفت آن ها در وظایف ورزشی، در حال حاضر شبکه های خوبی با دقت بالا برای وظایف گفته شده وجود دارد. با این حال، روش های موجود معمولاً از دو مشکل رنج می برند: اول، آموزش شبکه های مجزا برای هر یک از آن وظایف با هزینه های محاسباتی بالایی همراه است، و دوم، انسدادهای سنگین و ظاهر مشابه در ویدیوهای ورزشی، که دقت راه حل های موجود برای این کارها را محدود می کند. در این پژوهش، یک روش بازنمایی فرد پاره-محور چند منظوره به نام *PRTreID* پیشنهاد شده است که سه وظیفه دسته بندی نقش، وابستگی تیم و بازشناسایی را به طور همزمان انجام می دهد. برخلاف کارهای موجود، یک شبکه واحد با نظارت چند وظیفه ای برای حل هر سه کار به طور مشترک آموزش داده می شود که از نظر محاسباتی موثر است. روش پیشنهادی پاره-محور است و از اطلاعات مبتنی بر قسمت های مختلف بدن برای هر فرد بهره می برد که می تواند به طور قابل توجهی در برنامه های انسداد مفید باشد. همانطور که توسط نتایج کمی و کیفی نشان داده شده است، یادگیری چند وظیفه ای منجر به بازنمایی های غنی تر و تمایز دهنده تر می شود. علاوه بر این، مدل *PRTreID* پیشنهادی با یک روش ردیابی مبتنی بر بازشناسایی یکپارچه شده و یک الگوریتم پس پردازش پاره-محور برای مدیریت ردیابی بلندمدت پیشنهاد شده است. با مدل چندمنظوره بازشناسایی پیشنهادی، روش ردیابی به دست آمده، *PRT - Track*، قادر است به طور موثر افراد را حتی در برنامه های چالش برانگیز ردیابی کند، و از همه روش های ردیابی اخیر در مجموعه داده چالش برانگیز *SoccerNet - Tracking* بهتر عمل کند. روش پیشنهادی معیارهای *HOTA* و *AssA* را نسبت به بهترین روش موجود، به ترتیب، به مقدار ۱/۵۷ و ۲/۵۳ درصد بهبود می دهد.

**کلیدواژه ها:** بینایی کامپیوتر، یادگیری ژرف، ویدیوهای ورزش تیمی، بازشناسایی، ردیابی چندشیء، بازشناسایی پاره-محور، وابستگی تیم، یادگیری چندوظیفه ای، یادگیری بازنمایی

# فهرست مطالب

۱	مقدمه	۱
۱-۱	مقدمه	۱
۲-۱	تعریف و اهمیت مسئله و کاربردهای آن	۱
۳-۱	چالش‌ها	۱
۴-۱	هدف پژوهش	۱
۵-۱	دستاوردهای پژوهش	۲
۶-۱	ساختار پایان‌نامه	۲
۲	ادبیات پژوهش	۳
۱-۲	مقدمه	۳
۲-۲	ردیابی چندشیء	۳
۳-۲	جمع‌بندی	۳
۳	روش پیشنهادی	۵
۱-۳	مقدمه و شرح روش به صورت کلی	۵
۲-۳	بازنمایی چندمنظوره فرد پاره-محور	۵
۳-۳	جمع‌بندی	۵

۶	۴ ارزیابی
۶	۱-۴ مقدمه
۶	۲-۴ مجموعه داده
۶	۳-۴ معیارهای ارزیابی
۶	۴-۴ جزئیات پیاده سازی
۶	۵-۴ نتایج تجربی
۶	۱-۵-۴ مقایسه با کارهای پیشین
۶	۲-۵-۴ مطالعات فرسایشی
۷	۶-۴ جمع بندی
۸	۵ جمع بندی و کارهای آتی
۸	۱-۵ جمع بندی و نتیجه گیری
۸	۲-۵ کارهای آتی
۹	آ مطالب و نتایج تکمیلی



# فهرست تصاویر

۴	۱-۲ نمای کلی یک الگوریتم ردیابی .....
---	---------------------------------------

# فهرست جداول

۷	۱-۴ نتایج ردیابی .....
---	------------------------

# فصل ۱

## مقدمه

۱-۱ مقدمه

۲-۱ تعریف و اهمیت مسئله و کاربردهای آن

۳-۱ چالش‌ها

۴-۱ هدف پژوهش

## ۵-۱ دستاوردهای پژوهش

مشارکت‌های<sup>۱</sup> این پژوهش به صورت زیر خلاصه می‌شود:

- ارائه یک روش ردیابی افراد با استفاده از ویژگی‌های مبتنی به بخش و یک روش پس‌پردازش پاره-محور برای حل مسئله ردیابی بلند مدت.

## ۶-۱ ساختار پایان‌نامه

---

<sup>1</sup>contributions

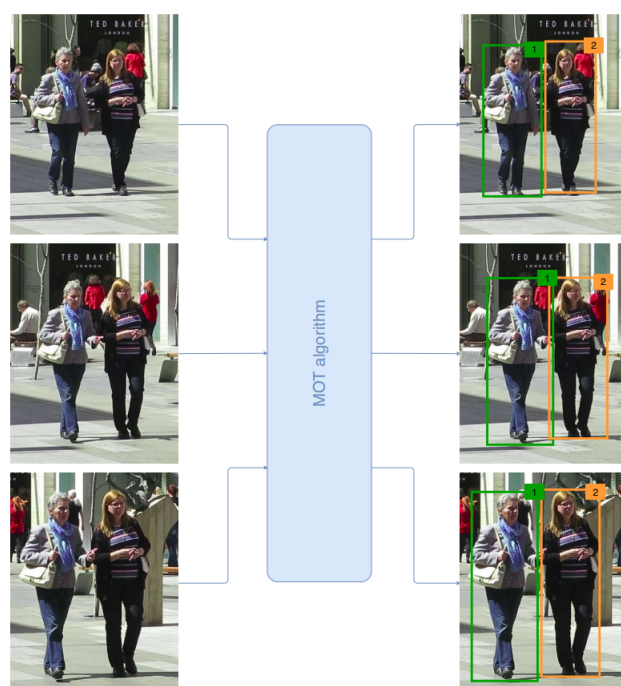
## فصل ۲

### ادبیات پژوهش

۱-۲ مقدمه

۲-۲ ردیابی چندشیء

۳-۲ جمع بندی



شکل ۲-۱: نمای کلی یک الگوریتم ردیابی. ورودی شامل یک ویدیو بوده که از تعدادی قاب تشکیل شده است. در خروجی در هر قاب برای هر شیء (فرد) یک جعبه مرزی و یک شماره به عنوان شناسه آن در طول ویدیو بدست می آید [۱].

## فصل ۳

### روش پیشنهادی

۳-۱ مقدمه و شرح روش به صورت کلی

۳-۲ بازنمایی چندمنظوره فرد پاره-محور

۳-۳ جمع‌بندی

## فصل ۴

### ارزیابی

۴-۱ مقدمه

۴-۲ مجموعه داده

۴-۳ معیارهای ارزیابی

۴-۴ جزئیات پیاده سازی

۴-۵ نتایج تجربی

۴-۵-۱ مقایسه با کارهای پیشین

۴-۵-۲ مطالعات فرسایشی



## ۴-۶ جمع‌بندی

جدول ۴-۱: نتایج ردیابی. مقایسه عملکرد روش پیشنهادی *PRT-Track* و روش‌های ردیابی اخیر بر روی مجموعه تست مجموعه داده ردیابی *Soccernet-Tracking*. (نماد † به معنای این است که نتایج گزارش شده از [۲] است.)

Oracle detections using ground truth						
↓ IDs	↑ IDF۱	↑ MOTA	↑ AssA	↑ DetA	↑ HOTA	Method
–	–	۹۴/۸۴	۵۸/۶۶	۸۲/۶۲	۶۹/۵۲	†DeepSORT
–	–	۹۴/۵۷	۶۰/۷۱	۸۴/۳۴	۷۱/۵	†ByteTrack
۶۰۷۹	۷۴/۷۹	۹۶/۷۶	۶۶/۹۸	۹۷/۸۱	۸۰/۹۴	OC-SORT
۲۸۱۵	۷۹/۱۳	۹۴/۶۷	۷۳/۷۸	۹۵/۰۸	۸۳/۷۵	StrongSORT
۲۶۱۹	۷۹/۷۶	۹۴/۶۲	۷۴/۳۶	۹۵/۰۷	۸۴/۰۸	StrongSORT++
–	۸۶/۱۰	۹۹/۴۰	۸۰/۰۰	۹۹/۴۰	۸۹/۲۰	CBIOU
۳۳۵۵	۸۸/۴۷	۹۸/۶۶	۸۲/۵۳	۹۹/۸۳	۹۰/۷۷	<b>PRT-Track</b>

YOLOv۸ detections						
↓ IDs	↑ IDF۱	↑ MOTA	↑ AssA	↑ DetA	↑ HOTA	Method
–	–	۳۳/۹۱	۳۳/۷۶	۴۰/۰۲	۳۶/۶۳	†DeepSORT
–	–	۵۰/۶۹	۴۱/۷۷	۴۶/۳۱	۴۳/۹۱	†FairMOT
–	–	۳۱/۷۴	۵۰/۲۵	۴۴/۴۹	۴۷/۲۲	†ByteTrack
۳۵۹۳	۶۲/۵۲	۷۶/۱۸	۴۷/۰۷	۶۳/۴۷	۵۴/۶۰	OC-SORT
۲۱۷۸	۶۵/۱	۷۴/۵۲	۴۸/۷۹	۶۲/۱۹	۵۴/۸۶	StrongSORT
۲۱۰۶	۶۶/۵۳	۷۵/۰۲	۵۰/۲۷	۶۲/۸۹	۵۶/۲۱	StrongSORT++
۱۴۲۸	۷۴/۴۴	۷۳/۰۷	۵۸/۵۵	۶۱/۰۹	۵۹/۷۷	<b>PRT-Track</b>

## فصل ۵

### جمع‌بندی و کارهای آتی

در این بخش ابتدا یک جمع‌بندی از کارهای انجام‌شده در این پژوهش ارائه می‌شود. در ادامه به کارهایی که می‌تواند در آینده در ادامه پژوهش فعلی انجام‌شود بررسی می‌شود.

#### ۵-۱ جمع‌بندی و نتیجه‌گیری

#### ۵-۲ کارهای آتی

پیوست آ

مطالب و نتایج تکمیلی

# Bibliography

- [1] G. Ciaparrone, F. L. Sánchez, S. Tabik, L. Troiano, R. Tagliaferri, and F. Herrera. Deep learning in video multi-object tracking: A survey. *Neurocomputing*, 381:61–88, 2020.
- [2] A. Cioppa, S. Giancola, A. Deliege, L. Kang, X. Zhou, Z. Cheng, B. Ghanem, and M. Van Droogenbroeck. Soccernet-tracking: Multiple object tracking dataset and benchmark in soccer videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3491–3502, 2022.

# واژه‌نامه

## الف

انسداد ..... occlusion  
آشکارسازی ..... detection  
اطمینان ..... confidence  
الحاق ..... concatenate  
امیدوارکننده ..... promising  
اتلاف سه‌گانه ..... triplet loss  
استنتاج ..... inference  
آستانه ..... threshold  
انتها به انتها ..... ent-to-end

## ب

بازشناسی ..... recognition  
برون خط ..... offline  
بازیابی ..... retrieval  
برچسب ..... label  
بازنمایی ..... embedding  
برافزایی ..... augmentation

## Abstract

Effective tracking and re-identification of persons is essential for analyzing team sport videos. However, this task is challenging due to the nonlinear motion of players, the similarity in appearance of players from the same team, the distance of the camera from the persons on the pitch, and frequent occlusions. Therefore, the ability to extract meaningful embeddings to represent persons is crucial in developing an effective tracking and re-identification system. In team sports, there is other information that can be used for re-identification of persons, such as team affiliation, role information, and jersey number. However, existing methods usually suffer from two problems: first, training separate networks for each of those tasks comes with high computational costs, and second, heavy occlusions and similar appearances in sports videos limit the solutions for these tasks. In this research, a multi-purpose part-based person representation method, called PRTreID, is proposed that performs three tasks of Role Classification, Team Affiliation, and Re-Identification, simultaneously. In contrast to available literature, a single network is trained with multi-task supervision to solve all three tasks jointly, which is computationally effective. The proposed method is part-based, leveraging body part-based information for each person, which can be significantly helpful in occlusion scenarios. The multi-task learning leads to richer and more discriminative representations, as demonstrated by both quantitative and qualitative results. Furthermore, the proposed PRTreID model is integrated with a re-identification based tracking method and a part-based post-processing module to handle long-term tracking is proposed. With the powerful re-identification model, the resulting tracking method, named PRT-Track, is capable of effectively re-identifying persons even in challenging scenarios, outperforming all recent tracking methods on the challenging SoccerNet-Tracking dataset. The proposed method improves the best existing method in terms of HOTA and AssA by 1.57 and 2.53 percentage points, respectively.

**Keywords:** Computer Vision, Deep Learning, Team Sports Videos, Re-Identification, Multi-Object Tracking, Part-based Re-Identification, Team Affiliation, Multi-task Learning, Representation Learning



Sharif University of Technology  
Department of Computer Engineering

M.Sc. Thesis

# **Deep Multi-Object Tracking by Part-Based Re-Identification in Soccer Matches**

By:

**Amir Mohammad Mansourian**

Supervisor:

**Dr. Kasaei**

August 2023

