

法政大学学術機関リポジトリ

HOSEI UNIVERSITY REPOSITORY

PDF issue: 2024-06-30

Deep Learningを用いた横断歩道走行のための 自動車トラッキングに関する研究

AOYAGI, Shuya / 青柳, 秀弥

(出版者 / Publisher)

法政大学大学院理工学・工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学・工学研究科編

(巻 / Volume)

62

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2021-03-24

(URL)

<https://doi.org/10.15002/00023991>

Deep Learning を用いた横断歩道走行のための 自動車トラッキングに関する研究

A STUDY OF VEHICLES TRACKING ACROSS A PEDESTRIAN CROSSWALK
APPLYING DEEP LEARNING TECHNIQUE

青柳秀弥

Shuya AOYAGI

指導教員 小林一行 教授

法政大学大学院理工学研究科システム理工学専攻修士課程

This paper describes a moving vehicle detection and tracking algorithm that is useful for crossing a pedestrian crosswalk using an omnidirectional camera. The omnidirectional camera is suitable for such applications due to its wide viewing angles without blind spots; however, since captured image from the omnidirectional camera is distorted that are not suitable for conventional object tracking algorithm. In order to overcome the omnidirectional camera image distortion problem, we propose a new algorithm that is consisting of YOLOv3 and a deep SORT algorithm. The YOLOv3 is mainly detecting object type and location and the deep SORT is used to track each detected object. The combination of these two algorithms enables robust and stable object tracking even if occlusion phenomena has happened. The validity of the proposed algorithm is confirmed by applying actual captured pedestrian crosswalk movie images.

Key Words : omnidirectional camera, object detection, tracking

1. はじめに

(1) 研究背景

現在の日本では高齢化が進んでいる。高齢者の総人口に占める割合は28.1%である[1]。また、国内の交通事故における死者数を見ると、全年齢死者数に対し、高齢者死者数は55.4%と半数以上になる[2]。その対策として、バリアフリーの推進が国内で行われている。特に2006年に公布された「高齢者、障害者等の移動等の円滑化の促進に関する法律」に基づいて警察庁ではバリアフリー対応型信号機の整備や見やすく分かりやすい道路標識・道路標示等の整備を推進している[3, 4]。しかし、信号が変わってから横断歩道に侵入してくる自動車を検出可能なシステムはまだ整備されていない。

また、自律移動ロボットの技術開発および実証実験の場としてつくばチャレンジが行われている。ここでは、人間が普段生活している環境の下で自律移動ロボットの自律走行技術を進歩させることを目的としている技術チャレンジである。この技術チャレンジの一つに自律移動ロボットが横断歩道を横断する課題が設けられている[5]。この課題は、ロボットが歩行者用信号と自動車の状況を認識して横断することである。

(2) 研究の目的と目標

本研究の目的は、人が交通事故なしに安全に生活する

ことである。高齢者や障害者をはじめ、歩行者が交通事故に遭遇しないために歩行者側が得る周囲の環境を把握することが重要である。将来的に、横断歩道を人が安全に渡るためのシステムを構築したい。

この目的の中で、本研究の目標は、リアルタイムに自動車の検出とトラッキングを行なうシステムの開発である。特に自動車と接近する横断歩道周辺で利用可能なシステムを開発する。また、本研究では全方位カメラを採用し、横断歩道周辺の360°すべての情報を1つのデバイスで取得させる。

2. 関連研究

(1) 全方位カメラにおける動的物体のトラッキング

Marković らの研究について述べる[6]。この研究では、移動ロボットに搭載された全方位カメラを用いて、逐次的に動いている物体を検出し、追跡と追従をする手法が述べられている。全方位カメラからの画像データに対して、移動ロボットのオドメトリ情報を参照しながら、静的物体のベクトルと動的物体のベクトルを単位球上で分類している。動的物体のベクトルをパラメータとしたフォンミーゼスフィッシャー分布に基づくベイズ推定を提案した。これにより、逐次的に物体の位置を推定している。しかし、物体の分類はされていない。

(2) 単眼カメラにおける物体検出とトラッキング

Feichtenhofer らの研究について述べる[7]. この研究では, 単眼カメラからの入力に対し, 物体検出とトラッキングを共同で行なうシステムの提案がされている. フレームベースで物体検出とトラッキングを同時に行なう畳み込みニューラルネットワークを構築した. トラッキングでは特に時間経過に対して, 物体の相関特徴量を算出し, フレームの差で追跡を行なっている.

(3) 全方位カメラにおける物体検出とトラッキング

井上らの研究[8]では, 8面キューブマップと呼ばれる投影方法を提案し, 物体検出を行なっている. またトラッキングには, 従来の投影手法である正距円筒図法を用いて8面キューブマップで検出した物体と合わせている. 物体検出には, YOLOv2[10]のアルゴリズムを用いている. ここで使用された学習モデルは, 単眼カメラの画像で学習されたモデルを利用しているため, 全方位カメラからのデータを直接検出することはできない. 一度2種類の投影方法に変換してから物体検出とトラッキングを行なっている.

3. 事前実験

関連研究から, 物体検出は全方位画像を展開した画像に対して単眼カメラの画像で学習されている重みは適していないとなっている. ここでは, 全方位カメラの画像をパノラマ展開した画像で物体検出の精度を検証し, その妥当性を確認する.

(1) 検証とシステムの概要

全方位カメラと物体検出アルゴリズム YOLOv3 [9, 10, 11]を用いた自動車の検出の実験を述べる. Fig.1 に実験の模式図を示す. 図の左下に示す全方位カメラは横断歩道正面を撮影する. 前方道路には, 自動車が行っており横断歩道の全長は約 11m である.

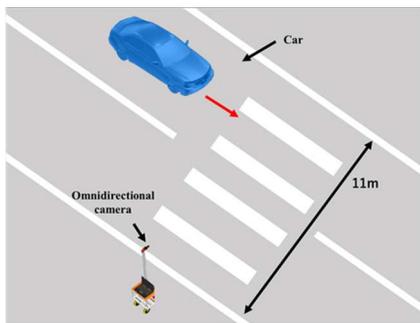


Fig.1 Schematic of vehicle recognition verification

Fig.2 に実験で用いた全方位カメラとその取り付け図を示す. 全方位カメラは地面と平行に高さ 1.4m の位置に取り付け横断歩道全体を見渡せるものとする.



Fig.2 Experimental equipment

Fig.3 に全方位カメラで取得した画像を示す. 左が地面側, 右が空側の画像となっている.



Fig.3 Omnidirectional camera image

Fig.4 に地面側をパノラマ変換した画像を示す. 展開された画像は横軸が角度, 縦軸が対応する角度の距離である. 正面を 180° とした. パノラマ展開した画像に対し, YOLOv3 で自動車を検出する.

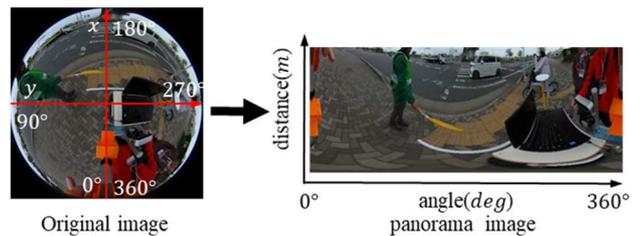


Fig.4 Panorama converted image (lower lens image)

Table.1 に実験に使用した PC のスペックを記載する.

Table.1 Computer specifications	
Topics	Spec
OS	Ubuntu 18.04 LTS
CPU	Intel Core™ i7-9700 CPU @ 3.00GHz
GPU	NVIDIA GeForce GTX 1650
RAM	16.0 GB
NVIDIA-SMI	450.80.02
CUDA Version	10.1

(2) 検証内容と結果

単眼カメラで学習済みの重み[12]で, 横断歩道前で撮影したパノラマ画像の自動車の適合率と, 信頼度スコアを

調べ、パノラマ画像に適用する場合の問題点を洗い出す。

a) パノラマ画像の自動車の適合率

自動車が存在する 2000 枚のパノラマ画像をそれぞれ自身の手でクラス”Car”のラベルを付けて、アノテーションを行なった。そのデータに対して、学習済みモデルで平均適合率 (AP) の検証を行なった。Table.2 に IOU(Intersection over union)の閾値を変化させたときの AP の結果を示す。IOU 値は、”Car”が存在すると予測した領域の割合である。

Table.2 AP results for varying IOU thresholds

IOU Threshold	0.3	0.4	0.5	0.6	0.7
AP (Average precision)	0.821	0.812	0.767	0.649	0.408

検証結果より、IOU の閾値は、小さくすれば適合率は、良くなるが、あげすぎると過剰適合になる。適合率を 0.75 程度にするには、IOU の閾値が 0.5 にすると良いことがわかった。

b) パノラマ画像の自動車の検出と信頼度スコアの検証

つぎに、IOU の閾値を 0.5 に設定した場合の、自動車の検出と信頼度スコアの検証を行った。

検出できた例

Fig.5 に入力画像、Fig.6 に検出したバウンディングボックスと番号、Table.3 にはクラスと信頼度スコア(0 ≤ p ≤ 1)を示す。



Fig.5 Input data

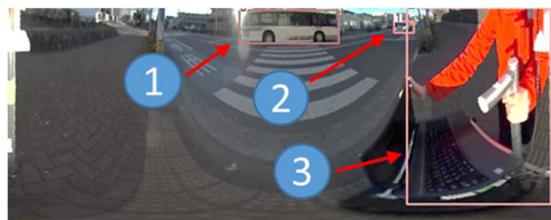


Fig.6 Output image with bounding boxes

Table. 3 Detection result

Number	Class	Confidence score
1	Bus	0.870
2	Car	0.459
3	Person	0.206

①の”Bus”は 0.87 で検出できている。②の”Car”は小さく写っているが、検出されている。

重複検出された例

Fig.7, 8 に入出力画像を示す。Fig.8 下に重複検出された部分の拡大画像を示す。Table.4 にクラスと信頼度スコアを示す。



Fig. 7 Input data

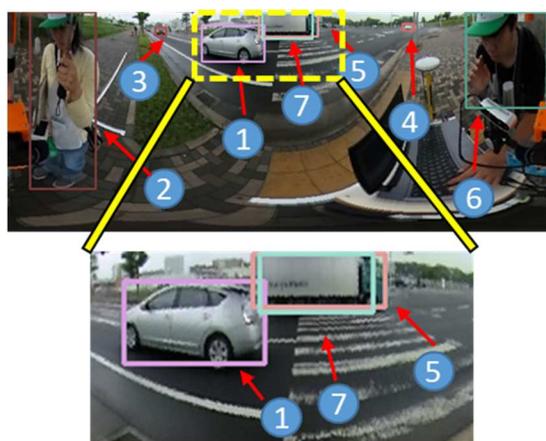


Fig.8 Output image and zoomed image

Table.4 Detection result

Number	Class	Confidence score
1	Car	0.987
2	Person	0.601
3	Car	0.591
4	Car	0.399
5	Truck	0.300
6	Person	0.127
7	Car	0.112

Fig.8 では同一の物体が、2種類のクラス(⑤, ⑦)として検出された。

IOU 値を 0.5 に設定することで、単眼カメラの画像で学習されている重みでもパノラマ画像の物体検出が可能であるが、重複検出の問題が生じやすくなることがわかった。

4. 問題と条件の定義

関連研究と事前検証の結果を前提に、本研究の目標で

あるリアルタイムに自動車の検出とトラッキングを行なうシステムを提案する。提案する上で解決しなければならない問題を以下に示す。

- P1: 自動車の重複検出の補正
- P2: 自動車のトラッキング
- P3: 自動車が進んでいるかの判定方法

5. 提案手法

Fig.9 に提案するシステムの流れを示す。最初に全方位カメラ画像をパノラマ画像に変換し、YOLOv3 に入力する。YOLOv3 で自動車(“Car”, “Bus”, “Truck”)のクラス、バウンディングボックス、信頼度スコアを出力する。検出した自動車から、重複検出補正をする。そして自動車のバウンディングボックスデータをトラッキングアルゴリズムである Deep SORT[13, 14, 15]に代入し、YOLOv3 で検出した自動車が続くフレームで同一のものであるかを識別する。同一な自動車が識別されているとき、連続する自動車の中心座標からベクトルを計算し、自動車が止まっているかを判定する。以下に詳細を述べる。

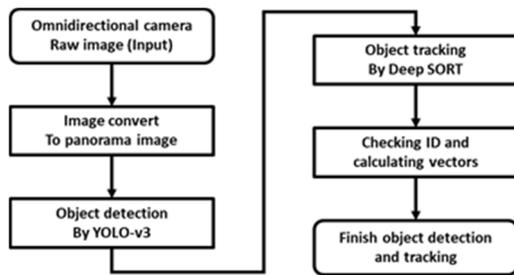


Fig.9 Flow of the developed system

(1) 自動車の重複検出の補正

重複検出は、Non-maximum suppression[16]を YOLOv3 の後処理として行なう。この処理は、検出した物体のバウンディングボックスの重なりを IOU 値で判定する。もし重なりが大きければ、信頼度スコアの小さい方を取り除く処理である。

(2) 検出した自動車のトラッキング

トラッキング手法は、Deep SORT (Simple Online and Realtime Tracking with a Deep Association Metric)[12, 13, 14]を用いる。これは、再帰的カルマンフィルタと連続するフレームごとのデータを参照して、検出された物体に Track-ID を紐づける手法である。Deep SORT には YOLOv3 でフレームごとに検出された自動車のデータを入力する。自動車のデータはバウンディングボックスの中心位置(u, v)とアスペクト比(γ)と高さ(h)である。カルマンフィルタを用いて前のフレームのデータから現在の入力を予測し、予測の範囲内であれば同一物体として Track ID を紐付けし、それ以外なら新しい Track ID を与える。Fig.10 に連続するフレームの切り抜き画像を示す。フレームが変化したときに同じ Track ID が別フレームの自動車にも割り当てられた。

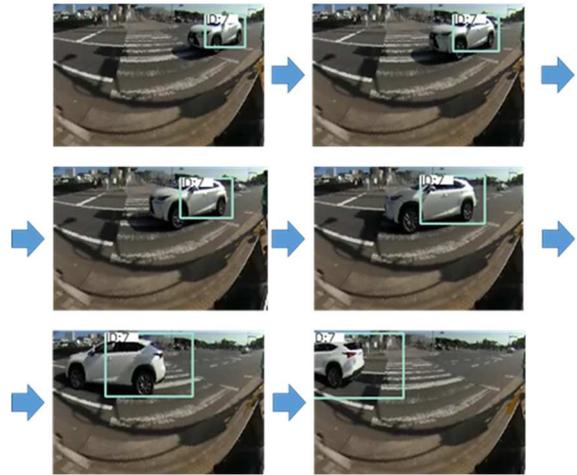


Fig. 10 Object tracking in successive frames

(3) 物体が進んでいるかの判定方法

物体が進んでいるかの判定としては、現在のフレームに存在する自動車のベクトルを計算する。トラッキングで紐づけた Track ID から、直前の 1 秒間のバウンディングボックスの中心座標(u, v)を抽出する。抽出した座標のうち座標 u のフレーム間の差分を取り、1 秒間のフレーム数で平均を計算する。計算した平均値の絶対値が 1 未満のとき止まっている、それ以外を進んでいると判定する。

Fig.11 に連続するフレームでのトラッキング画像を示す。このとき、画像右に存在する自動車は停止していることがわかる。

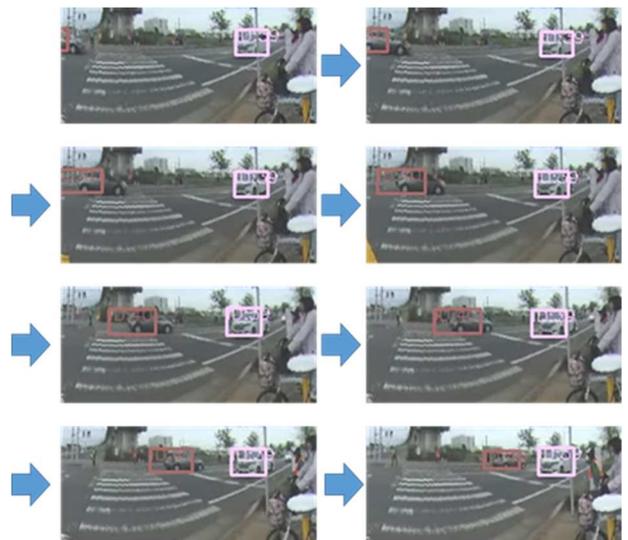


Fig.11 Object tracking in successive frames

6. 検証実験

提案した手法をそれぞれ検証していく。

(1) 自動車の重複検出の補正

重複検出補正に関して検証する。事前実験で検証した Fig.7 と同じ画像で行なう。Fig.12 に出力画像と事前実験で重複検出された部分の拡大画像を、Table.6 にクラスと

スコアを示す。事前実験で重複検出されていた自動車のうちスコアが低い方を取り除くことができた。

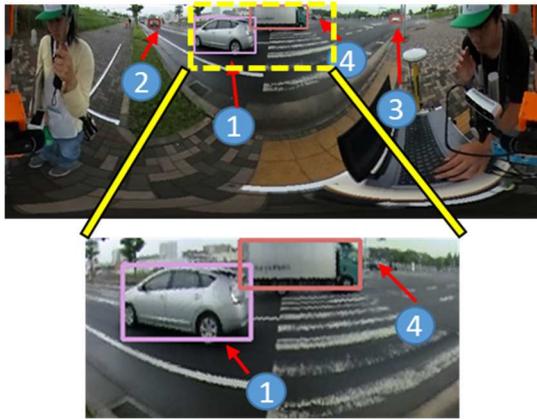


Fig.12 Output image and zoomed image

Table.6 Detection result

Number	Class	Confidence score
1	Car	0.987
2	Car	0.591
3	Car	0.399
4	Truck	0.300

(2) Deep SORT を用いたトラッキング

Fig.12 にトラッキング結果のグラフを示す。グラフの横軸に時間、縦軸に角度を取る。角度は、バウンディングボックスの中心座標(u, v)の u に対応している。

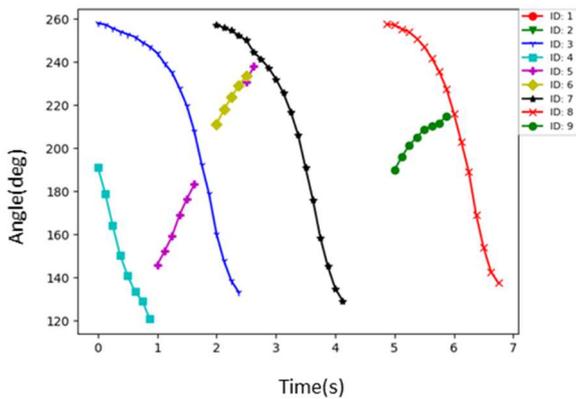


Fig.12 Result of object tracking

グラフより、縦軸の角度が減少する物体は右から左へ移動している自動車であり、実環境ではカメラに近い車線を走る自動車であることがわかる。角度が増加しているのは、左から右へ移動するカメラから遠い車線を走行する自動車である。近くの子線の自動車に隠れてしまう場合 Track ID がずれる。その様子を Fig.13 に示す。

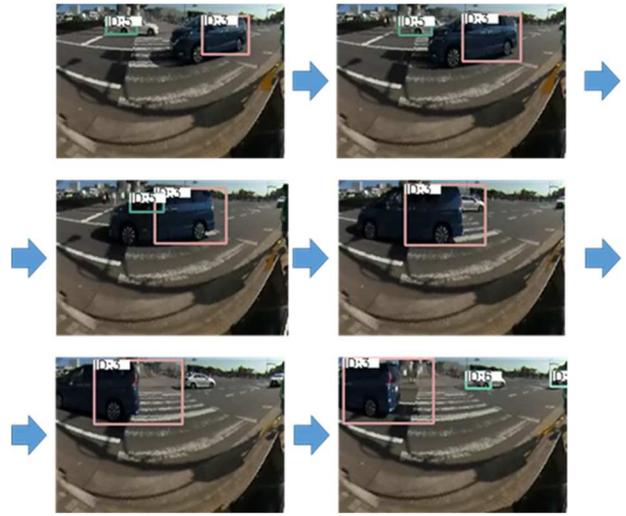


Fig.13 Object tracking in successive frames

(3) 物体のベクトル

Fig.14 にトラッキング結果のグラフの一部を示す。

この結果は実際に停止した自動車があったときのトラッキングである。

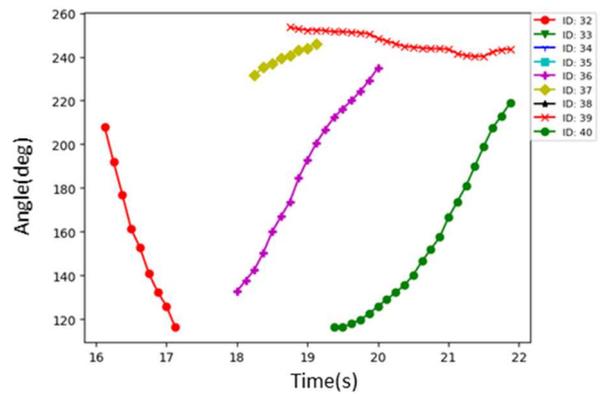


Fig.14 Result of object tracking

Table.7 にトラッキングされた 5 つの自動車がどの程度動いたかを示す。Track ID がそれぞれ整数秒のとき、バウンディングボックスの中心(u, v)の座標 u が平均でどれくらい動いたかを計算した。ID:39 の自動車が停止していることがわかる。

Table.7 Average of coordinate displacements

ID	17s	18s	19s	20s	21s	22s
32	12.25	----	----	----	----	----
36	----	----	7.40	5.60	----	----
37	----	----	1.78	----	----	----
39	----	----	0.38	0.28	0.84	0.03
40	----	----	----	1.25	4.38	7.69

7. 結論

本研究では、全方位カメラを用いた自動車の検出とトラッキングを行なうシステムを提案した。事前実験から、単眼カメラの画像で学習されている重みで、全方位画像をパノラマ展開した画像で物体検出が可能であることが確かめられ、問題点として重複検出が生じた。提案から、重複検出の補正、自動車のトラッキング、動作判定を解決し、検証を行なった。検証結果から、パノラマ画像で自動車の検出とトラッキングの有効性が確認された。

8. 今後の展望

本研究では、単眼カメラの画像で学習されている重みで自動車の検出を行なったが、パノラマ画像のデータで学習をさせることで検出精度が高まるだろう。

さらに、自動車だけでなく、自転車やバイクなどのクラスも検出できるようにしていく。

参考文献

- 1) 内閣府. 高齢化の状況. 令和元年度版高齢社会白書
https://www8.cao.go.jp/kourei/whitepaper/w-2019/zenbun/pdf/1s1s_01.pdf
- 2) e-Stat 道路の交通に関する統計/交通事故死者数について
<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&lid=000001241709>
- 3) e-Gov. 高齢者、障害者等の移動等の円滑化の促進に関する法律
https://elaws.e-gov.go.jp/search/elawsSearch/elaws_search/lsg0500/detail?lawId=418AC0000000091_20181101_430AC0000000032&openerCode=1
- 4) 警視庁. 安全かつ快適な交通の確保. 令和元年度版 警察白
https://www.npa.go.jp/hakusyo/r01/pdf/09_dai5sho.pdf
- 5) つくばチャレンジ 2019 課題:
<https://tsukubachallenge.jp/2019/regulations/tasks>
- 6) I. Marković, F. Chaumette and I. Petrović, "Moving object detection, tracking and following using an omnidirectional camera on a mobile robot," 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 5630 - 5635, 2014
- 7) C. Feichtenhofer, A. Pinz and A. Zisserman, "Detect to Track and Track to Detect," 2017 IEEE International Conference on Computer Vision (ICCV), pp. 3057-3065, 2017
- 8) 井上慶彦, 岩村雅一, 黄瀬浩一, "全方位カメラを用いた物体検出とトラッキング-視覚障害者支援システムの実現に向けて-" 情報処理学会研究報告(Web) (IPSIJ Technical Report (Web)), Vol.2018-CVIM-212, No.20, 1-6 (WEB ONLY), 2018
- 9) Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779 - 788, 2016
- 10) Joseph Redmon, Ali Farhadi, "YOLO9000: Better, Faster, Stronger", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6517 - 6525, 2017
- 11) Joseph Redmon, Ali Farhadi, "YOLOv3: An Incremental Improvement", arXiv preprint arXiv: 1804.02767, 2018.
- 12) Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár, "Microsoft COCO: Common Objects in Context" In European conference on computer vision, pages 740-755. Springer, 2014.
- 13) Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, Ben Uprocft, "Simple Online and Realtime Tracking" 2016 IEEE International Conference on Image Processing (ICIP), pp. 3464 - 3468, 2016
- 14) Nicolai Wojke, Alex Bewley, Dietrich Paulus, "Simple online and realtime tracking with a deep association metric", 2017 IEEE International Conference on Image Processing (ICIP), pp. 3645 - 3649, 2017
- 15) Nicolai Wojke, Alex Bewley, "Deep Cosine Metric Learning for Person Re-identification", 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 748 - 756, 2018
- 16) Pedro F. Felzenszwalb, Ross B. Girshick, David McAllester, Deva Ramanan "Object Detection with Discriminatively Trained Part-Based Models" IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 32, Issue: 9, Sept. 2010)