

全方位カメラを用いた横断歩道走行のための 自動車のトラッキングおよび速度予測に関する研究

渡邊, 陸 / WATANABE, Riku

(出版者 / Publisher)

法政大学大学院理工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学研究科編

(巻 / Volume)

63

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2022-03-24

(URL)

<https://doi.org/10.15002/00025404>

全方位カメラを用いた横断歩道走行のための 自動車のトラッキングおよび速度予測に関する研究

A STUDY ON VEHICLE TRACKING AND SPEED PREDICTION
AT PEDESTRIAN CROSSING USING OMNI-DIRECTIONAL CAMERA

渡邊陸

Riku WATANABE

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 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 YOLO-v4 and Deep SORT algorithm. The YOLO-v4 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, YOLO-v4, Deep SORT*

1. はじめに

近年、周囲をすべて撮影できる全方位カメラが、幅広く販売されるようになった。Fig.1に示すように、全方位カメラを中心点に、周囲 360 度を一度に記録することができる。単眼カメラから得られる画像とは異なり、周辺環境を一枚のパノラマ画像として扱うことができる。全方位カメラはその広い画角から、自律移動ロボットの分野においても多く使用されている。屋外の人物検出システム[1,2]や UAV の自己位置推定[3]等、さまざまな用途に対応している。

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コース内には歩行者用信号のある横断歩道と、信号のない横断歩道が存在する。ロボットは、正確に横断歩道の状況を認識して横断をする必要がある。とくに歩行者用信号のない横断歩道では、自動車が停車しない場合がある。また、人間の目ならば遠くから接近してくる自動車を認識することは容易だが、ロボットにとっては困難であり、安全のための対策が必要であると考えられる。

ロボットが安全に歩行者用信号のない横断歩道を渡るためには、遠くから接近してくる自動車を検出しトラッキングをすることで、自動車の速度を予測しておく必要がある。周辺環境を画角の制限なく撮影できる全方位カメラを用いることで、遠くから接近してくる自動車の速度をよりはやくリアルタイムに検出できると考えた。

以上より、本研究の目的はリアルタイムに自動車の検出とトラッキングを行ない、速度を予測するアルゴリズムの作成である。全方位カメラから得たパノラマ動画から周辺の自動車を検出し、検出された物体ごとにトラッキングを行ない、速度を予測する。

物体検出には YOLO-v4(You Only Look Once version 4) アルゴリズム[4-7]を用いる。物体検出アルゴリズムのひとつであり、リアルタイムに物体を検出することができる。検出された物体は、その物体を丁度囲むバウンディングボックスで表される。また、物体をトラッキングする手法として、Deep SORT(Simple Online and Realtime Tracking with a Deep Association Metric)[8-10]を用いる。この手法は、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 ID を紐づける手法である。Deep SORT には YOLO-v4 でフレームごとに検出された物体のバウンディングボックスの中心座標を入力する。

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、予測の範囲内に物体があれば同一物体としてIDを紐づけし、それ以外なら新しいIDを与える。

本論文の構成は、2章で関連研究、3章で本研究における問題の定義、4章で提案するアルゴリズム、5章で検証実験、6章で結論について述べる。



Fig.1 Original omnidirectional camera image

2. 仮定と問題の提起

(1) 仮定の記述

本研究で提案するリアルタイムな自動車の検出とトラッキングを行なうために、必要な仮定を以下に示す。

A1) ロボットは横断歩道に垂直な向きに停止している。

A2) 走行環境は一般的な横断歩道と想定する。

以下に、本研究で想定する横断歩道の詳細を示す。

R1) 長さ 10 m, 幅 3.0 m, 白線の間隔 0.45 m

R2) 平面であり、凹凸は考慮しない。

(2) 問題の記述

本研究で提案するリアルタイムな自動車の検出とトラッキングを行なうにあたって、解決すべき問題点を以下に示す。

P1) 全方位画像から自動車をどのように検出・トラッキングするのか。

P2) 検出した自動車の速度予測はどのように行なうか。

3. 提案手法

Fig.2 に本研究で提案するアルゴリズムを示す。はじめに、全方位カメラから取得した画像をパノラマ画像に変換し、YOLO-v4 に入力する。入力したパノラマ画像から、自動車のクラス、バウンディングボックス、信頼度スコアを出力する。その後、検出結果に対して、重複検出補正をかける。補正されたデータをトラッキングアルゴリズムである Deep SORT に入力し、YOLO-v4 で検出した自動車をトラッキングする。最後に、バウンディングボックスの座標とフレームレートから、自動車の速度を予測する。

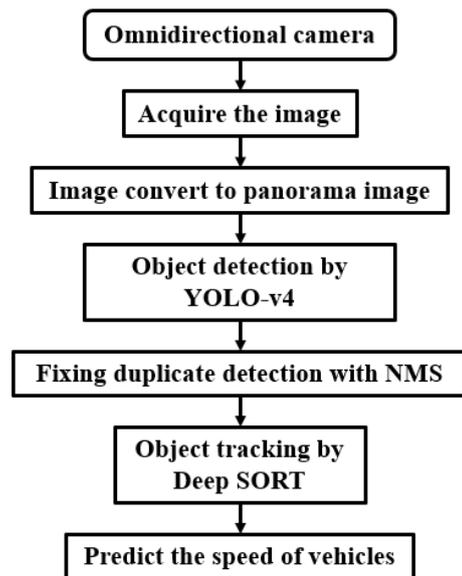


Fig.2 Proposed algorithm

(1) 全方位カメラ画像を用いた自動車検出

本研究では、全方位カメラ画像をパノラマ状に展開した画像において自動車の検出を行なう。

a) 全方位カメラ画像のパノラマ変換処理

本研究で扱う全方位カメラ画像は歪みを含んでいる。そのため、自動車をより正しく検出するために歪み補正を行ない、画像をパノラマ状に変換する。はじめに、全方位カメラのパラメータを事前に取得する。そのパラメータから、画像処理のためのライブラリである OpenCV を使用して歪み補正を行ない、パノラマ状に変換する。以下の Fig.3 に、入力する全方位カメラ画像、Fig.4 に、パノラマ状に変換した画像を示す。

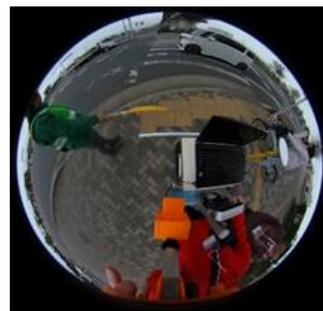


Fig.3 Omnidirectional camera image (land)



Fig.4 Panorama Image

b) YOLO-v4によるパノラマ画像上の自動車検出

自動車検出には、YOLO-v4 アルゴリズムを用いる。リアルタイムに物体を検出できるという特徴があり、検出した物体を、その物体を丁度囲むバウンディングボックスという矩形で出力する。

また、検出の評価指標として IOU 値と平均適合率 AP を考える。IOU 値とは、YOLO-v4 による画像上の予測領域と、実際に物体が存在する正解領域との重なりによって表される。平均適合率 AP は、検出したバウンディングボックスが正しい比率 p 、検出すべきバウンディングボックスを検出できている比率 r によって表される。IOU 閾値を決めることで、予測したバウンディングボックスの正誤を判定し、平均適合率 AP によって検出結果を評価する指標を算出する。以下の式(1)(2)(3)に、その平均適合率 AP を求める式を示す。尚、TP (True Positive) は、画像中に実際に存在する物体を検出できたことを示す。FP (False Positive) は、画像中に物体が存在しないはずが検出をしてしまったことを示す。FN (False Negative) は、画像中の物体を検出できなかったことを示す。Table.1 に、IOU 値を変化させたときの平均適合率 AP を示す。

$$p = \frac{TP}{TP+FP} \tag{1}$$

$$r = \frac{TP}{TP+FN} \tag{2}$$

$$AP = \int_0^1 p(r)dr \tag{3}$$

Table.1 AP results for varying IOU thresholds

IOU Threshold	0.3	0.4	0.5	0.6	0.7
AP	0.821	0.812	0.767	0.649	0.408

Fig.5, Fig.6に YOLO-v4 による実際の検出結果を示す。パノラマ画像に対して、自動車と予測した部分のバウンディングボックスを正しく出力できていることが確認できる。



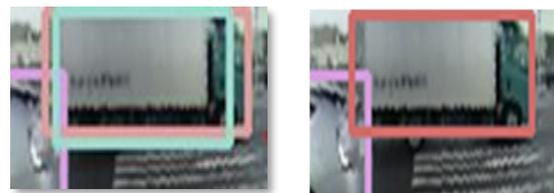
Fig.5 Input image



Fig.6 Output image with bounding boxes

c) NMSによる重複検出補正処理

YOLO-v4 を用いて検出した自動車の出力結果に、重複して同一物体を示すものが発生することがある。その補正処理に NMS (Non-maximum suppression) アルゴリズム [11]を用いる。2つの同じクラスとして検出された結果の IOU 値が閾値以上の場合、同一の物体を示していると認識し、信頼度の低い出力を抑制する。以下の Fig.7 に、NMS アルゴリズムによる補正結果を示す。同一の自動車を示しているバウンディングボックスの出力が、1つに抑制されていることが確認できる。



(a) Before

(b) After

Fig.7 Comparison of before and after correction processing

(2) 全方位カメラ画像を用いた自動車トラッキング

トラッキングには、Deep SORT(Simple Online and Real time Tracking with a Deep Association Metric)を用いる。物体のトラッキングを行なう機械学習モデルであり、物体に ID を付与してトラッキングする。この手法は、カルマンフィルタを用いたトラッキングアルゴリズム SORT と、ReID アルゴリズムの組み合わせである。YOLO-v4 によって出力されたバウンディングボックスの中心座標とアスペクト比を 1 フレーム毎に入力する。前フレームにおける情報から、現在フレームの入力を予測し、予測の範囲内にトラッキング対象が存在した場合、同一の ID を継続して付与する。対象の位置が予測から大きく外れている場合、新たな ID を付与してトラッキングを開始する。Fig.8 に、トラッキングを行なっている画像を示す。同一の車両に ID を付与し、トラッキングを行なえていることが確認できる。

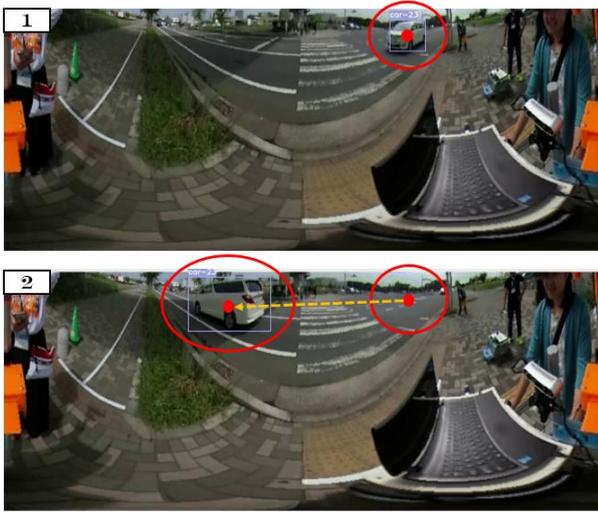


Fig.8 Object tracking in successive frames

(3) 被検出自動車の速度予測方法

ロボットが安全に横断歩道を渡るため、トラッキングした自動車の速度予測を行なう。Fig.9 に入力する画像を示す。入力する全方位カメラ画像は 740×1846 [pixel] であり、縦方向に仰角($^\circ$)、横方向に方位角($^\circ$)を表している。したがって 1 pixel あたり、仰角 0.122° 、方位角 0.195° と求まる。Fig.10 にロボット、全方位カメラ、検出対象の自動車における座標系を示す。本研究では、全方位カメラを地表からの高さ 1.5 メートル部分に取り付ける。全方位カメラの高さ h と、検出した自動車を表すピクセルから、1 フレーム毎にロボットから自動車の座標 P 、距離 D を求める。動画のフレームレート F_{ps} から、1 フレームあたりに自動車の進む距離、速度 S を算出する。以下、(4) (5) (6) (7) (8) に式を示す。



Fig.9 Panorama Image in pixels

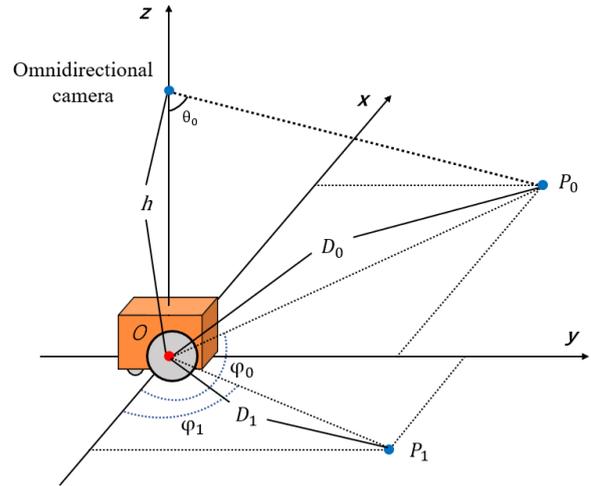


Fig.10 Coordinate system including omnidirectional camera

$$D_0 = \frac{h}{\tan(90^\circ - \theta_0)} \quad (4)$$

$$D_1 = \frac{h}{\tan(90^\circ - \theta_1)} \quad (5)$$

$$P_0 : (x_0, y_0) = (D_0 \sin \varphi_0, D_0 \cos \varphi_0) \quad (6)$$

$$P_1 : (x_1, y_1) = (D_1 \sin \varphi_1, D_1 \cos \varphi_1) \quad (7)$$

$$S_{01} = \frac{\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}}{1/F_{ps}} \quad (8)$$

提案システムでは、速度予測方法に YOLO-v4 から出力されたバウンディングボックスを用いている。しかし、出力はフレーム毎に誤差が生じる。そのため、本研究では、予測したフレーム毎の速度から近似を求めることで、ばらつきのあるデータから速度を予測する。

4. 検証

検証は、実際の片側一車線道路の交通状況を全方位カメラで記録した動画上で行なった。Fig.11 に、本研究で使用するロボットを示す。ロボットの幅の中心部分、高さ 1.5 メートルの位置に、全方位カメラを横に倒して設置する。Fig.12 に検証環境を示す。仮定より、ロボットは横断歩道に対して垂直な向きで停止している。

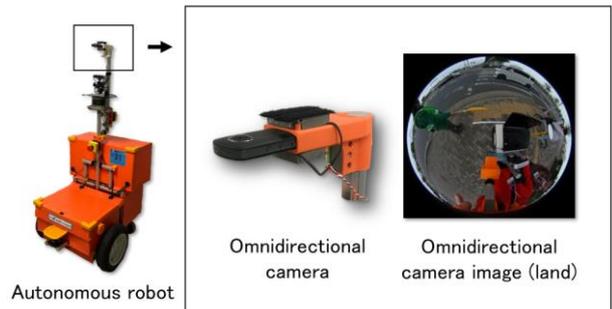


Fig.11 Robot configuration

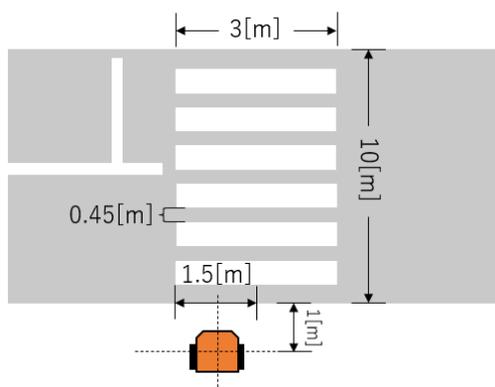
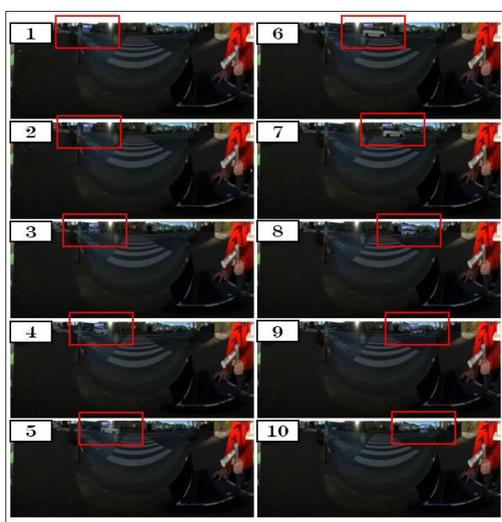
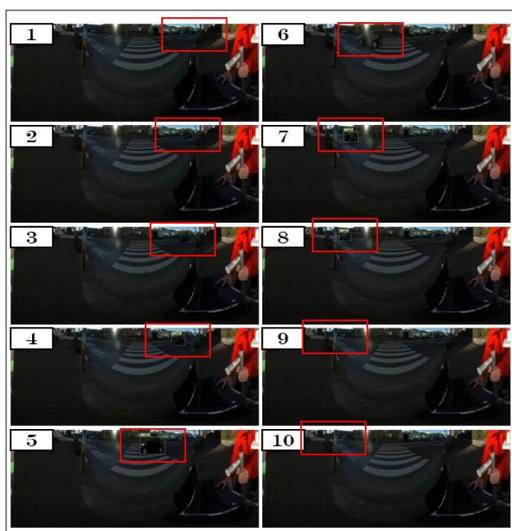


Fig.12 Verification environment

Fig.13, Fig.14, Fig.15 に、検証環境にて自動車の検出およびトラッキング、速度予測を行なった結果を示す。全方位カメラ画像から、自動車を検出しトラッキングが安定して行なえていることを確認した。



(a)



(b)

Fig.13 The results of the tracking

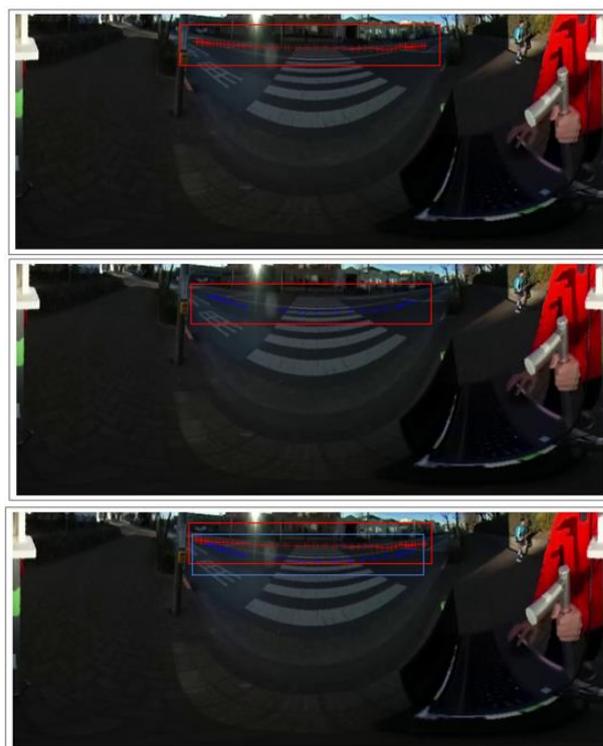


Fig.14 Tracking results for each oncoming lane

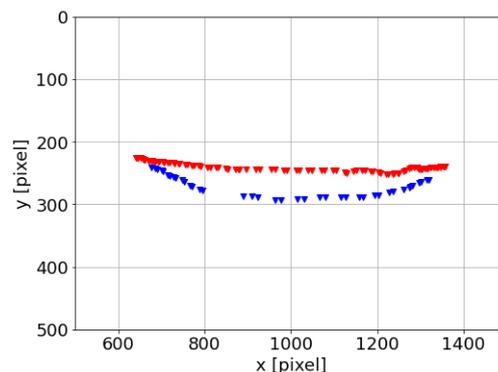


Fig.15 Tracking results for each oncoming lane in pixels

また、Fig.16 と Fig.17 に被検出自動車の速度予測の結果を示す。ロボットの前を通過する自動車、ロボットの前で完全停止する自動車の写った動画を用いて、速度予測の検証を行なう。今回、検証使用する動画は、実際の自動車の速度を測ることができていないため、自動車がロボットの前を通過するのか、完全停車するのかを判断する。結果から、自動車の停止を予測できていることが確認できる。



Fig.16 Completely stopped car

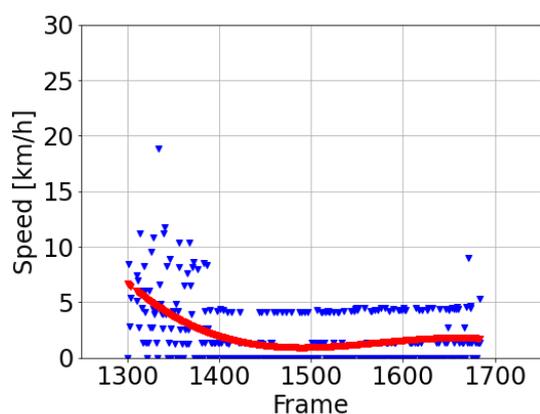


Fig.17 Results of prediction of vehicle speed

5. 結論

本研究では、全方位カメラを用いた自動車の検出・トラッキング、速度予測を行なうシステムを開発した。実際の横断歩道のような、全方位カメラにより撮影した動画を用いて検証を行なった。結果から、リアルタイムに自動車を検出し、車両毎にトラッキングと速度予測が可能であることを確認し、提案手法の有効性が証明された。

6. 今後の展望

本研究で提案したシステムでは、仮定より、ロボットが横断歩道に対して垂直な向きで停止している前提がある。実機に実装する際には、横断歩道手前におけるロボットのランダムな停止位置を考慮した、柔軟な横断歩道の検出方法および自動車の速度予測方法を追加する必要があると考える。

参考文献

- 1) Shunya Tanaka and Yuki Inoue, "Outdoor Human Detection with Stereo Omnidirectional Cameras," J.Robot. Mechatron., Vol.32, No.6, pp. 1193-1199, 2020.
- 2) Shingo Nakamura, Tadahiro Hasegawa, Tsubasa Hiraoka, Yoshinori Ochiai, and Shin'ichi Yuta, "Person Searching Through an Omnidirectional Camera Using CNN in the Tsukuba Challenge," J. Robot. Mechatron., Vol.30, No.4, pp.

540-551, 2018.

- 3) Naoya Hatakeyama, Tohru Sasaki, Kenji Terabayashi, Masahiro Funato, and Mitsuru Jindai, "Position and Posture Measurement Method of the Omnidirectional Camera Using Identification Markers," J. Robot. Mechatron., Vol.30, No.3, pp. 354-362, 2018.
- 4) 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
- 5) Joseph Redmon, Ali Farhadi, "YOLO9000: Better, Faster, Stronger", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6517 – 6525, 2017
- 6) Joseph Redmon, Ali Farhadi, "YOLOv3: An Incremental Improvement", arXiv preprint arXiv: 1804.02767, 2018.
- 7) 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.
- 8) Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, Ben Upcroft, "Simple Online and Realtime Tracking" 2016 IEEE International Conference on Image Processing (ICIP), pp. 3464 - 3468, 2016
- 9) 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
- 10) 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
- 11) 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)