



逢甲大學學生報告 *ePaper*

基於深度學習之單目車距量測系統

Vehicle Distance Detection Based on Deep Learning Using Monocular Camera

作者：黃子霖

系級：電機系

學號：D0383251

開課老師：陳冠宏、梁寶芝

課程名稱：專題研究(一)

開課系所：電機系

開課學年：105 學年度 第二學期



中文摘要

摘要

新聞報導因交通事故而造成人員傷亡與破碎的家庭不計其數，又現今交通工具的先進駕駛輔助系統（ADAS）尚未普及，且價格昂貴。本次計畫提出基於單目視覺的駕駛輔助系統，並利用現今已普及的產品如行車記錄器或手機進而研究與開發，希望此次研究與實作能有效應用於實際道路狀況，改善交通安全。

作品研究動機

交通安全一直是受關注的議題，如<<全球道路安全現狀報告>>指出，每年約有125萬人死於道路交通事故。而在台灣根據<<105年內政統計通報>>顯示，歷年來死傷人數亦是居高不下。

年(月)別	機動車輛數 ②	道路交通事故(A1+A2)						肇事率 (件/萬輛)		
		發生件數 (件)	A1類	A2類	死傷人數 (人)	死亡①	受傷	A1類	A2類	
民國95年	20,307,197	160,897	2,999	157,898	214,316	3,140	211,176	80.11	1.49	78.61
民國96年	20,711,754	163,971	2,463	161,508	219,500	2,573	216,927	79.95	1.20	78.75
民國97年	21,092,358	170,127	2,150	167,977	229,647	2,224	227,423	81.39	1.03	80.36
民國98年	21,374,175	184,749	2,016	182,733	249,086	2,092	246,994	87.01	0.95	86.06
民國99年	21,721,447	219,651	1,973	217,678	295,811	2,047	293,764	101.94	0.92	101.02
民國100年	22,226,684	235,776	2,037	233,739	317,318	2,117	315,201	107.30	0.93	106.37
民國101年	22,346,398	249,465	1,964	247,501	336,122	2,040	334,082	111.94	0.88	111.05
民國102年	21,562,645	278,388	1,867	276,521	375,496	1,928	373,568	126.80	0.85	125.95
民國103年	21,290,313	307,842	1,770	306,072	415,048	1,819	413,229	143.67	0.83	142.85
民國104年	21,400,897	305,413	1,639	303,774	411,769	1,696	410,073	143.08	0.77	142.31
民國104年 1-10月	21,386,814	249,272	1,326	247,946	335,688	1,378	334,310	116.82	0.62	116.20
民國105年 1-10月	21,502,605	239,168	1,286	237,882	314,908	1,327	313,581	111.49	0.60	110.89
較104年同期 增減(%)	0.54	-4.05	-3.02	-4.06	-6.19	-3.70	-6.20	③ -5.33	③ -0.02	③ -5.31

資料來源：交通部、本部警政署。

說明：1. A1類係指造成人員當場或24小時內死亡之交通事故；A2類係指造成人員受傷或超過24小時死亡之交通事故。

2. 本表道路交通事故統計範圍包括A1類及A2類。

3. 肇事率係指每萬輛機動車輛肇事事件數。

4. 依新修正之道路交通安全規則第30條第3項規定，若機車未繼續使用，並完全符合下列狀況：(1)出廠逾10年。

(2)最近5年內無使用道路違規，無投保汽車責任強制險及無參加環保排氣檢驗紀錄。經過切結同意後，可由監理機關協助完成報廢登記，以減輕後續汽燃費之負擔，故102年起機車數逐年減少。

附註：①係指A1類死亡人數，不含A2類超過24小時死亡人數。

②機動車輛數係為期底數。

③係指增減數。

正是因為駕駛者的視野及反應速度有限，如常見肇事原因「未注意車前狀態」、「變換車道或方向不當」、「未保持行車安全距離」…等，故有必要尋找解決方案。

關鍵字：深度學習, 車輛辨識, 單目測距

Abstract

Driver inattention and misjudgment are the major causes of vehicle accidents. Extensive research has proved that Advanced Driver Assistance Systems(ADAS) can significantly reduce the number of these accidents. Since all perception sensors have their own weaknesses, Sensor Fusion and Machine Learning are crucial to achieve functional safety. We have built a monocular visual processing system which can estimate vehicle distance.

Keyword : Deep Learning, Vehicle Detection, Monocular Camera, Distance Estimation



目次

摘要	1
一、系統設計說明	3
1.1 軟硬體平台	3
1.2 系統流程設計	3
二、系統實作內容	3
2.1 車輛感測-前置程序	3
2.2 車輛感測-偏移程序	6
2.3 車輛感測-距離程序	7
三、實際測試畫面	7
3.1 Server (Remote Desktop)	7
3.2 Client (iPhone)	8
3.3 Lane Curvature + Vehicle Distance Detection	8
四、結論	8
完成功能說明	8
五、參考文獻	9



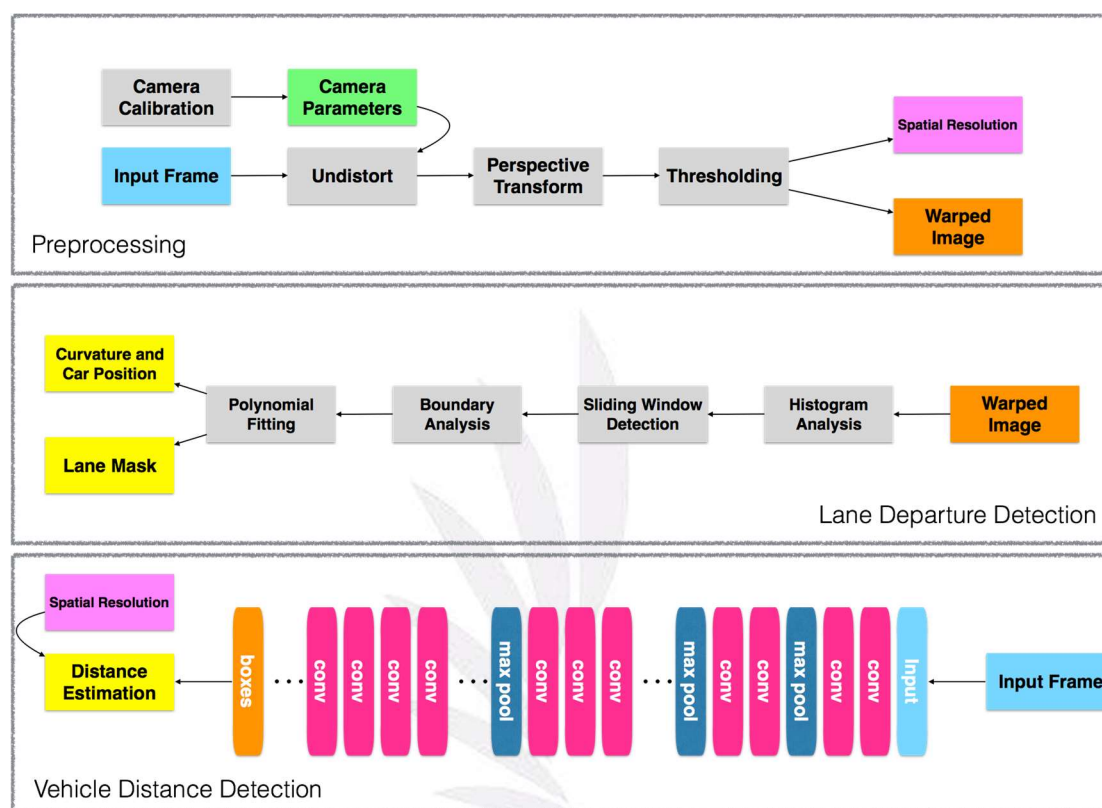
一、系統設計說明

1.1 軟硬體平台

軟體：OpenCV、Python、TensorFlow、Keras

硬體：iPhone、JetsonTX2

1.2 系統流程設計



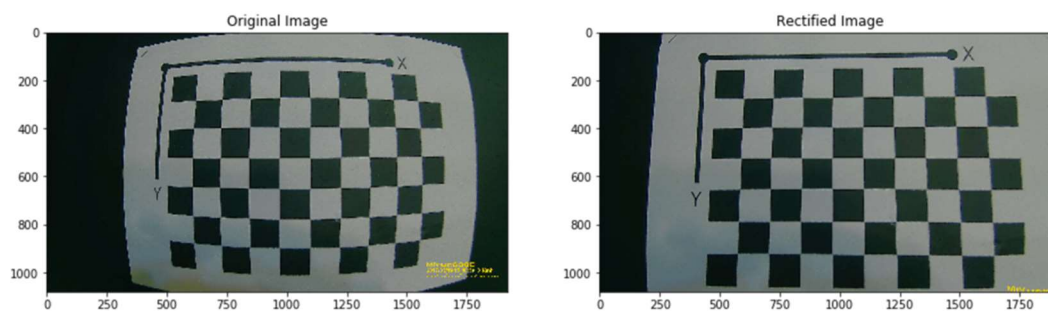
二. 系統實作內容

2.1 車輛感測-前置程序

1. 相機校正 (Camera Calibration)

在相機內參矩陣 (Intrinsic Matrix) 當中，因其假設為針孔相機模型 (Pinhole Camera Model)，然而於實際相機的成像模型中，尚須考慮鏡頭形變 (Lens Distortion)，此部份採用棋盤格 (6x9 checkerboard) 的方法校正及計算相機內參矩陣，圖為校正前後的狀況：

基於深度學習之單目車距量測系統



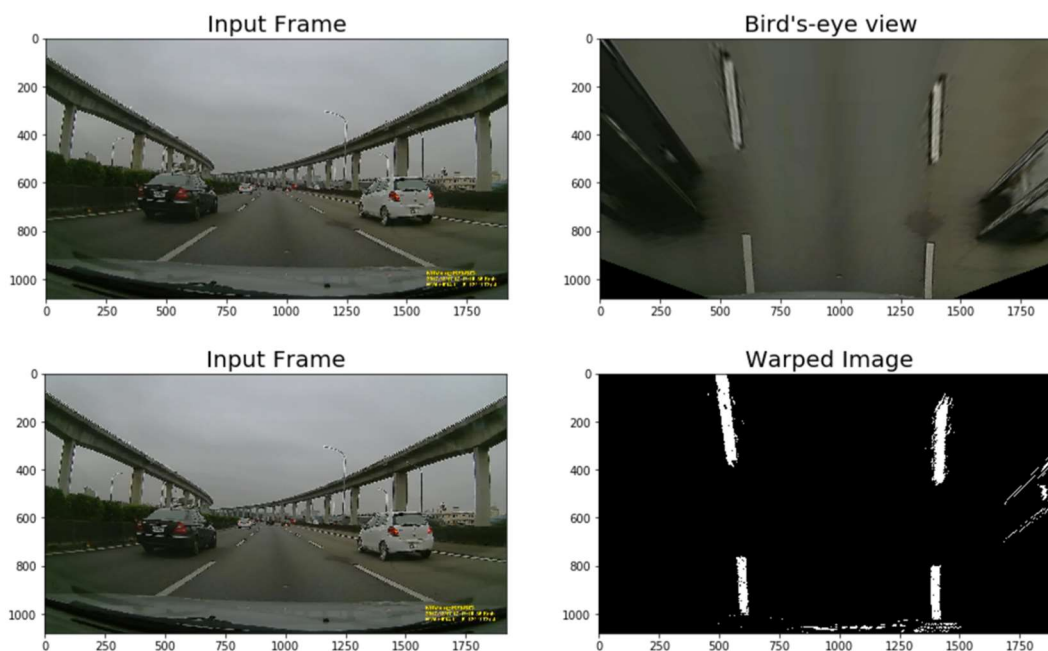
Camera intrinsic matrix:

```
[1.28153230e+03  0.00000000e+00  9.94173047e+02]
[0.00000000e+00  1.28405374e+03  4.62061719e+02]
[0.00000000e+00  0.00000000e+00  1.00000000e+00]
```

Camera distortion parameters:

```
[-0.47005176  0.26970871  0.00178749  0.00142575  0.0222831]
```

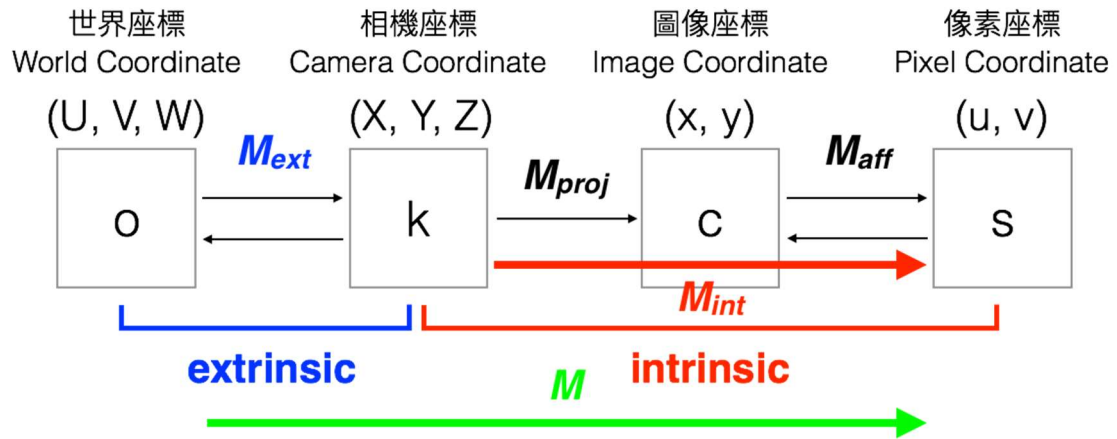
2.1.2 閾值處理 (Thresholding) & 透視轉換 (Perspective Transform)



為了方便計算車輛偏移以及車輛距離，我們利用透視轉換來得到道路的鳥瞰圖 (bird's-eye view)，然而為影像中並非所有的像素皆屬於車道線的一部份，故需經過閾值處理。

3. 空間解像力 (Spatial Resolution)

為了方便分析與描述數位影像與真實三維空間座標的狀況，故採用針孔相機模型 (Pinhole Camera Model) 以及齊次座標 (Homogeneous Coordinates) 表示，圖示為相關座標間的轉換關係：



世界座標(U, V, W)與像素座標(u, v)之關係式：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = sM \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} = sM_{int}M_{ext} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix}$$

$$M = M_{int}M_{ext}, \quad M_{int} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \quad M_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix}$$

s : scalar factor, M_{int} : Intrinsic matrix, M_{ext} : Extrinsic matrix

經透視變換 (perspective transform) 呈現鳥瞰圖 (bird's-eye view) ($W = 0$) 後，其關係式轉變如下：

$$\begin{bmatrix} u_w \\ v_w \\ 1 \end{bmatrix} = sH \begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = s \begin{bmatrix} fs_x & 0 & c_x \\ 0 & fs_y & c_y \\ 0 & 0 & 1 \end{bmatrix} [R_1, R_2, T] \begin{bmatrix} U \\ V \\ 1 \end{bmatrix}$$

$$H = [h_1, h_2, h_3] = \begin{bmatrix} fs_x & 0 & c_x \\ 0 & fs_y & c_y \\ 0 & 0 & 1 \end{bmatrix} [R_1, R_2, T]$$

$$\|R_1\| = \|R_2\| = 1$$

$$fs_y = fs_x \frac{\|h_1\|}{\|h_2\|}$$

(u_w, v_w) : Warped Image Pixel Coordinate

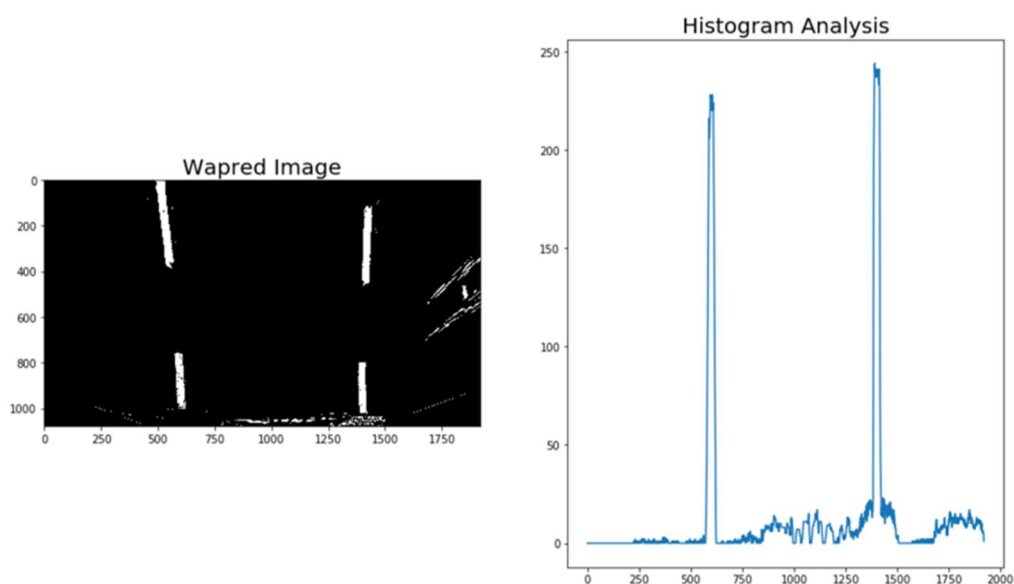
H : Homography Matrix

由於相機成像會將三維空間資訊投影至感光元件上，故只能判斷方向，並無法直接得到距離的資訊，透過上面的關係式即可求得

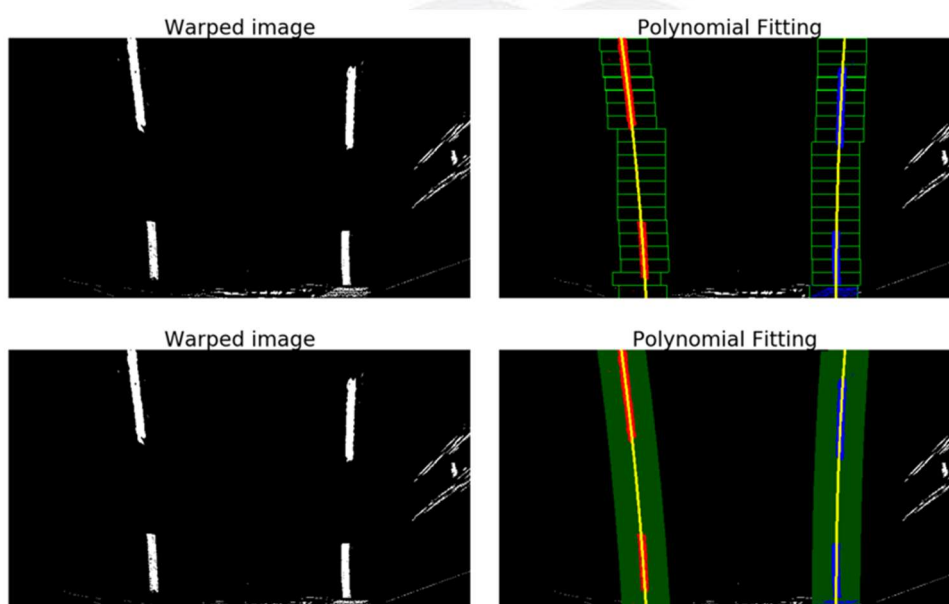
方向	空間解像力 (Spatial Resolution) (pixels/meter)
X	229.7297297297297
Y	61.951245567

2.2 車輛感測-偏移程序

2.2.1 直方圖分析 (Histogram Analysis)



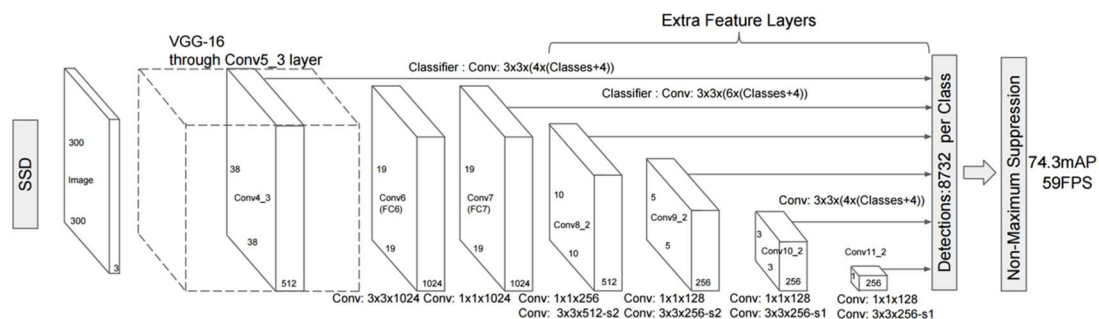
2.2.2 滑動窗口 (Sliding Window)



2.3 車輛感測-距離程序

車輛偵測 Histogram 中文直方圖

選擇使用 SSD: Single Shot MultiBox Detector 神經網路模型



特色：別於 Region Proposal 的方式，亦能保有同時能保有精準度以及速度：

於 PASCAL VOC 2007 上所測試的結果

Method	mAP	FPS
Faster R-CNN (VGG16)	73.2	7
Fast YOLO	52.7	155
SSD300	72.1	58

三、實際測試畫面

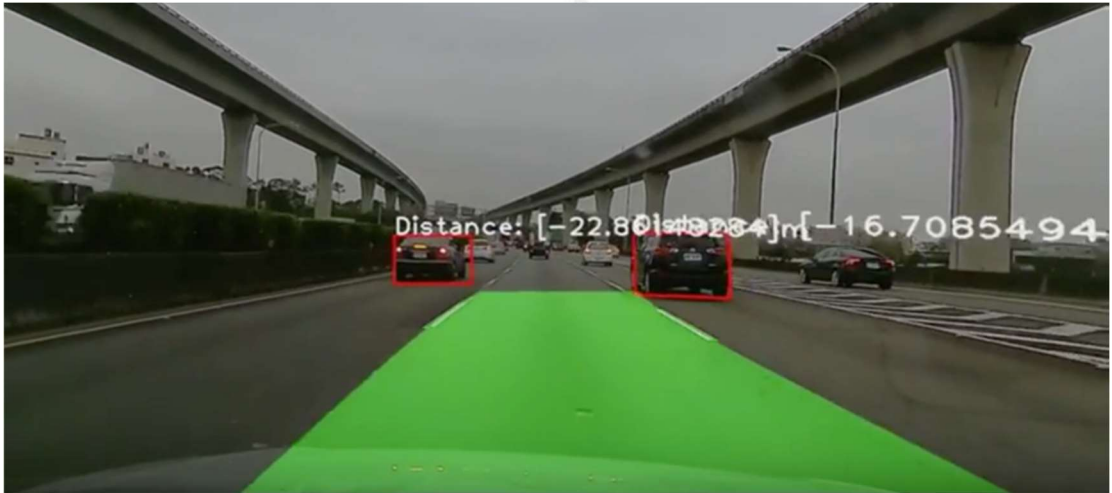
Server (Remote Desktop)



Client (iPhone)



Lane Curvature + Vehicle Distance Detection



Youtube : https://www.youtube.com/watch?v=2_pOgbHJbF8

四、結論

一、完成一個技術模擬系統，可以利用 iPhone 搭配一個 server，在實際開車上路的時候 demo 出如何即時的辨識車子以及車子間的距離

二、僅用單目相機達到車輛辨識與量測距離

三、影像量測汽車，辨識車子的辨識演算法，主要參考 SSD: Single Shot MultiBox Detector、Yolo v2 實現之

四、影像擷取個人化以及與影響處理的設備分離，我們將影像擷取用 client 隨身設備 iPhone 來擷取影像，以及利用一個 server 表示 GPU 處理單元

五、調教影像辨識效能，利用 GPU 之多執行緒處理架構達成之

- GPU 處理設備的局限性：Desktop Server 作為是 Client 的影像辨識的加強設備，後考慮便攜性採用 Nvidia Jetson TX2

- 網路速度的局限性：在實作測試環境 4G/3G 手機網路

六、此次報告實驗於 Nvidia Jetson TX2 可達到即時運算(Logitech Webcam

C920 @30fps)，詳情細節內容可 Email 至 h.tzulin@gmail.com 詢問。

五、參考文獻

- [1] Liu, Wei et al. "SSD: Single Shot Multibox Detector". *Computer Vision – ECCV 2016* (2016): 21-37.
- [2] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection
- [3] Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time

