

YOLOv2 기반의 영상 워핑을 이용한 강인한 오토바이 번호판 검출 및 인식

*당순정 **김응태

한국산업기술대학교

*xtdang93@gmail.com **etkim@kpu.ac.kr

Robust Motorbike License Plate Detection and Recognition using Image Warping based on YOLOv2

*Dang, Xuan Truong **Kim, Eung Tae

Korea Polytechnic University

요약

번호판 자동인식 (ALPR: Automatic License Plate Recognition)은 지능형 교통시스템 및 비디오 감시 시스템 등 많은 응용 분야에서 필요한 기술이다. 대부분의 연구는 자동차를 대상으로 번호판 감지 및 인식을 연구하였고, 오토바이를 대상으로 번호판 감지 및 인식은 매우 적은 편이다. 자동차의 경우 번호판이 차량의 전방 또는 후방 중앙에 위치하며 번호판의 뒷배경은 주로 단색으로 덜 복잡한 편이다. 그러나 오토바이의 경우 키 스탠드를 이용하여 세우기 때문에 주차할 때 오토바이는 다양한 각도로 기울어져 있으므로 번호판의 글자 및 숫자 인식하는 과정이 훨씬 더 복잡하다. 본 논문에서는 다양한 각도로 주차된 오토바이 데이터셋에 대하여 번호판의 문자 인식 정확도를 높이기 위하여 2-스테이지 YOLOv2 알고리즘을 사용하여 오토바이 영역을 선 검출 후 번호판 영역을 검지한다. 인식률을 높이기 위해 앵커박스의 사이즈와 개수를 오토바이 특성에 맞추어 조절하였다. 그 후 기울어진 번호판을 검출한 후 영상 워핑(Image Warping) 알고리즘을 적용하였다. 모의실험 결과, 기존 방식의 인식률이 47.74%에 비해 제안된 방식은 80.23%의 번호판의 인식률을 얻었다. 제안된 방법은 전체적으로 오토바이 번호판 특성에 맞는 앵커박스와 이미지 워핑을 통해서 다양한 기울기의 오토바이 번호판 문자 인식을 높일 수 있었다.

1. 서론

오토바이는 인도, 중국, 태국 베트남 등 많은 나라의 대표 대중 교통 수단이다. 뿐만 아니라 한국, 일본 등 다른 나라에서도 편리성 및 이동성 높기 때문에 오토바이는 지역배달에 많이 사용되고 있다. 또한 몇 년 전부터 계속 성장하고 있는 오토바이를 이용하여 택시 용도로 그랩바이크(Grab Bike)도 등장하였다. 특히 동남아시아에서는 오토바이를 사용하는 인구가 계속 늘어남에 따라 오토바이 번호판 인식에 대한 요구가 증가하고 있다.

최근까지 비디오감시시스템 및 지능형 교통시스템 등 다양한 환경에서 차량용 번호판을 인식하는 연구들이 많이 진행되고 있다 [1]-[3]. 그 연구들의 경우 특정한 배경, 위치, 조명 같은 조건하에서 성능이 개선되어져 왔다. 한편 최근에 다양한 분야에 심층학습(deep learning)기술이 적용되어 더 좋은 성과들을 보이고 있다. 번호판 인식 분야에서도 Faster R-CNN 및 YOLOv2 같은 심층 합성곱 신경망(Deep Convolution Neural Network)을 적용하여 인식률이 계속 높아지고 있다. 그러나 기존 딥러닝 기반의 번호판 검출 및 인식 [3-8] 방식들은 대부분 자동차, 버스, 트럭 대상으로 학습이 되어서 차량 번호판에 대해서는 검출 및 인식이 잘 되나 오토바이를 대상으로 적용시 오류나 검출이 잘 안 되는 현상이 나타난다.

특히 오토바이는 키 스탠드 이용하여 주차하기 때문에 주차할 때 오토바이가 많이 기울어져 번호판도 따라서 상당히 기울어진다. 따라서 기울어진 번호판의 인식률이 낮아지는 현상이 발생한다. 또한 다수

의 오토바이들이 주차시에도 빠르게 구별하여 인식할 수 있는 방식이 필요하다. 본 논문에서는 이러한 문제를 해결하기 위해서 2줄형식의 오토바이 번호판 인식을 위해 YOLOv2 기반하에 번호판에 맞는 앵커 박스와 이미지 워핑을 적용하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 오토바이 번호판 검출 관련 연구들을 설명하고, 3장에서는 개선된 YOLOv2 기반 영상워핑을 적용한 번호판 검출 구조를 제안한다. 4장에서는 실험결과 분석을 통해 기존 방식들과 비교한다. 마지막으로 5장에서는 본 연구의 결론을 맺고자 한다.

2. 번호판 인식관련 기존연구

일반적인 번호판 인식 시스템은 특징 추출, 탐식, 분류, 위치 추정과 같은 단계들로 이루어져 있다[9,10]. 특히 Edge, Conner, Harr-like, HOG(Histogram of Oriented Gradient) 와 같이 특징들을 사용하였다. 이런 기존 방법은 객체 데이터를 학습하는 과정이 필요 없으므로 구현이 용이한 장점이 있지만, 다양한 환경, 조명, 크기 등 여러 가지 조건들에 따라서 성능이 저하되는 단점도 있다.

한편 최근에 심층학습기술을 발전하기 때문에 합성곱 신경망을 적용하여 영상처리 분야에서 좋은 결과를 보이고 있다. 번호판 검출 및 인식하는 데도 마찬가지로 응용하는 합성곱 신경망을 이용하여 연구들이 늘어나고 있다. 심층학습 합성곱 신경망은 전통적인 분류기들과는 달리 기존의 특징벡터들을 사용하지 않으며, 훈련 샘플에 대하여 최적화된 벡터들을 스스로 학습하고, 이를 활용하여 특징맵(Feature

Map)을 생성하여 사용하는 것이다. 그래서 객체인식률은 높게 하는 데에 혼련 데이터셋트가 중요한 역할이다. 현재 번호판 인식하기 위한 SSIG, AOLP, CD-HARD, UFPR-ALPR 등 많은 데이터 세트 등장하였는데 안타깝게 모든 데이터셋트는 자동차, 버스, 트럭 위주로 만든 데이터셋트이다. 또한 심층학습 합성곱을 사용하여 객체 인식할 때 핵심 키워드중에 Bounding box는 하나이다. Bounding box는 객체 어떤 칼라 스인지 및 객체의 좌표 정보들을 담는다. Faster R-CNN 및 YOLOv2 모델에 모든 앵커 박스 먼저 설정하여 앵커 박스 통하여 bounding box를 찾는 것이다. 앵커 박스가 더 많을수록 객체 인식률을 높이지나 연산자 늘어지기 때문에 속도가 느려진 것이다. Faster R-CNN[2]는 앵커 박스는 9개 있으며 YOLv2[1]는 5개만 있다. 번호판은 영상에 비하면 상당히 작은 객체라서 Faster R-CNN는 YOLOv2보다 번호판 잘 찾지만 속도가 많이 느리다. 빠른 속도로 정확하게 번호판 인식하기 위해 [5-8]논문은 번호판 검출은 2단계로 나누었다. 첫 번째는 입력 영상에 차량 찾는 단계이며 둘 번째는 찾은 차량 영상에는 번호판 찾는 단계이다. 2 단계 나눠 번호판 인식하면 번호판은 찾은 차량 사진에 비하여 상당히 적당한 사이즈이라 속도 빠르며 인식률을 많이 높아졌다.

3. 제안된 YOLOv2기반 오토바이 번호판인식 시스템

그림 1은 제안된 오토바이 번호판 인식시스템을 나타낸다.

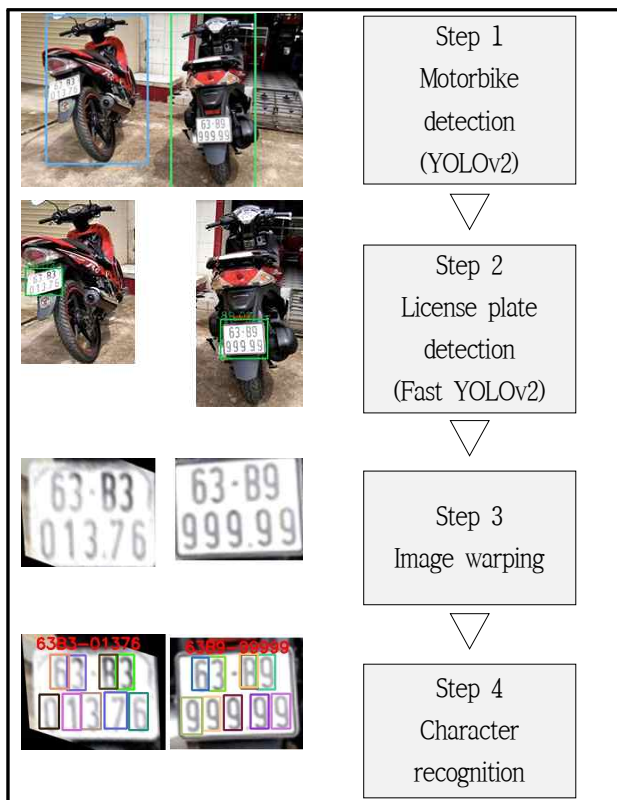


그림 1. 제안된 번호판 인식 시스템

3.1 오토바이 검출 및 번호판 검출

제안된 오토바이 번호판을 검출하기 위해서 2개의 CNN 구조를 가진다. 첫번째는 영상내 오토바이를 인식하기 위한 YOLOv2 모델이며, 두 번째는 검출된 오토바이내 번호판을 인식하기 위해 Fast-YOLOv2 모델이다. 수행 속도 및 인식률을 고려해서 복잡한 배경을 갖는 입력 영상에서 오토바이 인식률이 높이기 위한 YOLOv2 방식을 사용한 후

간단한 배경 가진 오토바이 번호판의 인식 속도를 높이기 위해 Fast-YOLOv2 사용하였다. 이는 Fast-YOLOv2는 보다 간단한 배경에서 객체를 빨리 인식을 할 수 있으나 복잡한 환경에서는 인식률이 떨어지기 때문이다.

YOLOv2와 Fast-YOLOv2 model 의 각각 모델에는 클래스 수를 맞추기 위한 마지막 콘볼루션 레이어의 필터 수는 바뀌어 한다. YOLOv2의 필터의 개수를 계산 하는 식 : $filters = (C + 5) \times A$

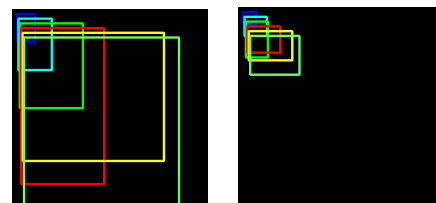
A는 앵커 개수이며, 상수 5는 각 bounding box의 4가지 좌표 (x, y, w, h)와 confidence를 포함하고, C는 클래스의 개수이다.

오토바이 인식 및 번호판 인식은 모두 하나의 클래스를 인식하므로 클래스의 개수는 1이고 객체의 다양한 크기를 인식하기 위한 앵커 개수는 6이라면 필터 개수는 36이 된다. 본 논문에서 쓰는 YOLOv2 구조는 표 1에 표현한다.

표 1. 오토바이 인식 단계에서 사용하는 YOLOv2 구조

Layer	Filters	Size	Input	Output	
0	conv	32	3 x 3 /1	416x416 x 3	416x416 x 32
1	max		2 x 2 /2	416x416 x 32	208x208 x 32
2	conv	64	3 x 3 /1	208x208 x 32	208x208 x 64
3	max		2 x 2 /2	208x208 x 64	104x104 x 64
4	conv	128	3 x 3 /1	104x104 x 64	104x104 x 128
5	conv	64	1 x 1 /1	104x104 x 128	104x104 x 64
6	conv	128	3 x 3 /1	104x104 x 64	104x104 x 128
7	max		2 x 2 /2	104x104 x 128	52x52 x 128
8	conv	256	3 x 3 /1	52x52 x 128	52x52 x 256
9	conv	128	1 x 1 /1	52x52 x 256	52x52 x 128
10	conv	256	3 x 3 /1	52x52 x 128	52x52 x 256
11	max		2 x 2 /2	52x52 x 256	26x26 x 256
12	conv	512	3 x 3 /1	26x26 x 256	26x26 x 512
13	conv	256	1 x 1 /1	26x26 x 512	26x26 x 256
14	conv	512	3 x 3 /1	26x26 x 256	26x26 x 512
15	conv	256	1 x 1 /1	26x26 x 512	26x26 x 256
16	conv	512	3 x 3 /1	26x26 x 256	26x26 x 512
17	max		2 x 2 /2	26x26 x 512	13x13 x 512
18	conv	1024	3 x 3 /1	13x13 x 512	13x13 x 1024
19	conv	512	1 x 1 /1	13x13 x 1024	13x13 x 512
20	conv	1024	3 x 3 /1	13x13 x 512	13x13 x 1024
21	conv	512	1 x 1 /1	13x13 x 1024	13x13 x 512
22	conv	1024	3 x 3 /1	13x13 x 512	13x13 x 1024
23	conv	1024	3 x 3 /1	13x13 x 1024	13x13 x 1024
24	conv	1024	3 x 3 /1	13x13 x 1024	13x13 x 1024
25	route	16			
26	conv	64	1 x 1 /1	26x26 x 512	26x26 x 64
27	reorg			26x26 x 64	13x13 x 256
28	route	27 24	/2		
29	conv	1024	3 x 3 /1	13x13 x 1280	13x13 x 1024
30	conv	36	1 x 1 /1	13x13 x 1024	13x13 x 36
31	detection				

또한, 오토바이 특성에 맞는 번호판 인식을 위해 그림 2와 같은 형태의 앵커박스를 사용하였다. 오토바이 및 번호판 인식률을 높이기 위해 앵커 박스는 6개 쓰고 사이즈도 k-means 알고리즘 통하여 기존 앵커 박스보다 오토바이 모양에 더 적합하게 만들었다.



(a) 오토바이용 (b) 번호판용

그림 2. 앵커 박스 모양

3.2 영상 워핑

그림 1에서 보는 바와 같이 오토바이 주차할 때 번호판이 항상 기울어짐을 알 수 있다. 이 상황에서 기울어진 문자의 인식이 떨어질 수 있으므로 번호판 문자 인식 단계 전에 번호판을 이미지 워핑하여 보정해 줄 필요가 있다.

본 논문에서 속도 측면을 고려하여 이미지 워핑하기 위하여 번호판의 4개 꼭짓점을 찾아서 투영변환(Perspective Transformation)을 사용한 기하학적 변형 방법을 사용한다. Hough Transform를 사용하여 영상의 모든 직선 성분을 찾아서 Voting 방법으로 가장 많이 나타난 각도를 가진 직선들을 찾는다. 찾은 직선들을 바로 번호판의 수직 축 테두리나 수평 축 테두리의 후보들이다. 번호판이 많이 기울어져도 수평 축 테두리와 수직 축 테두리의 각도 차이 범위는 [45,115]에 따라 찾은 직선들을 기반으로 남은 대응한 직선들을 찾는다. 따라서 수직 축 및 수평 축 테두리의 후보들을 모두 찾을 수 있다. Voting방법으로 후보들 중에 수직 축 테두리와 수평 축 테두리의 각각 2개의 제일 정확한 직선을 선택하여 이제 찾은 직선 만나는 4점은 바로 번호판의 4개 꼭짓점이다. 그림 3은 영상 워핑 과정을 나타낸다.

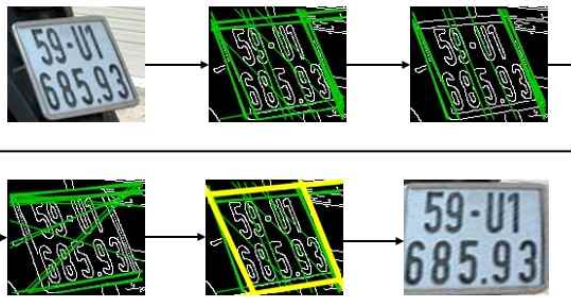


그림 3 영상 워핑 하는 과정

3.3 문자 인식

최근에 각 문자는 하나의 객체처럼 CNN 통하여 인식하는 연구 있다. 그래서 문자 분할 및 인식을 따로 실행하지 않고 문자 분할 및 인식을 동시에 실행하는 것이다. 특히 [7]논문예 YOLO 네트워크 약한 수정하여 OCR-net 만들었다. 다양한 나라 번호판 문자 인식 잘 하기 위하여 OCR-net는 많은 나라의 번호판 하고 인공 번호판같이 트러닝 하였. 본 논문에도 OCR-net 기반하여 문자 인식하였다. 다만 [7]은 3:1 비율 가진 자동차 번호판 인식하기 때문에 OCR-net의 입력 영상 사이즈는 240 x 80인데 본 논문에는 3:2 비율 가진 오토바이 번호판 인식하기 때문에 입력 영상 사이즈는 200 x 160으로 바꿨다. 본 논문에서 사용하는 OCR-net 구조는 표 2.으로 나타난다.

특히 동남아 오토바이는 번호판내 2 라인 있는데 첫 번째 라인은 4 가지 문자 있는데 그 중에 첫 번째, 둘 번째, 넷 번째 위치는 숫자이며 셋 번째 자리는 글자이고 둘 번째 라인은 문자 4개나 5개 문자 있는데 모든 숫자이다. 만일 숫자이어야 되는 위치에 문자를 인식하면 그 문자 비슷한 숫자 찾아서 대입하는 것이다. 구체적으로 밑처럼 바꾸는 것이다.

+ Line 1, 3rd position(글자): 0 -> D, 1 -> I, 2 -> Z, 4 -> L, 5 -> S, 8 -> B, 6 -> G

+ 2 lines, other position(숫자): D -> 0, I -> 1, Z -> 2, L -> 4, S -> 5, B -> 8, G -> 6

표 2. 오토바이 인식 단계 사용하는 OCR-net 구조

Layer	Filters	Size	Input	Output
0 conv	32	3 x 3 /1	200x160 x 3	200x160 x 32
1 max		2 x 2 /2	200x160 x 32	100x80 x 32
2 conv	64	3 x 3 /1	100x80 x 32	100x80 x 64
3 max		2 x 2 /2	100x80 x 64	50x40 x64
4 conv	128	3 x 3 /1	50x40 x64	50x40 x 128
5 conv	64	1 x 1 /1	50x40 x 128	50x40x 64
6 conv	128	3 x 3 /1	50x40x 64	50x40 x128
7 max		2 x 2 /2	50x40 x128	25x20 x126
8 conv	256	3 x 3 /1	25x20 x126	25x20 x256
9 conv	128	1 x 1 /1	25x20 x256	25x20 x512
10 conv	256	3 x 3 /1	25x20 x512	25x20 x256
11 conv	512	3 x 3 /1	25x20 x256	25x20 x512
12 conv	256	3 x 3 /1	25x20 x512	25x20 x256
13 conv	512	3 x 3 /1	25x20 x256	25x20 x512
14 conv	80	1 x 1 /1	25x20 x512	25x20 x80
15 detection				

4. 모의실험 결과

4.1 오토바이 인식 및 번호판 인식 성능

모의실험을 위해 그림 4와 같이 다양한 각도의 오토바이 1,230장 포함된 2-lines motorbike license plate (2L_MLP) dataset을 만들어 사용하였다.



그림 4. 2L_MLP Dataset의 사진들

2L_MLP dataset를 사용하여 2-stage YOLOv2기반의 제안된 방식에 대하여 오토바이 및 번호판 인식을 수행한 결과, 오토바이 인식률은 90.91%이며 번호판 인식률은 99.87%인 결과를 얻었다. 또한 오토바이 및 번호판 인식시간은 각각 16.6ms 및 9.5ms가 걸렸다. 표 3의 결과에서 보듯이 번호판 인식의 경우, 간단한 구조로 더 빠르게 인식할 수 있으며 인식율도 엄청 높은 것을 확인 할 수 있었다.

표 3. 오토바이 및 번호판 인식률 및 인식 시간

	인식률	시간
오토바이	90.91%	16.6ms
번호판	99.87%	9.5ms

4.2 영상 워핑에 대한 효과

본 논문에서 제안한 영상 워핑 알고리즘은 그림 5에서 보는 바와 같이 다양한 사이즈 및 각도를 가진 번호판 영상들을 잘 워핑하고 있

음을 알 수 있다. 특히 그림 5의 마지막 영상처럼 문자가 거의 잘 안 보이더라도 워핑을 통해서 인식을 개선할 수 있게 됨을 알 수 있다.

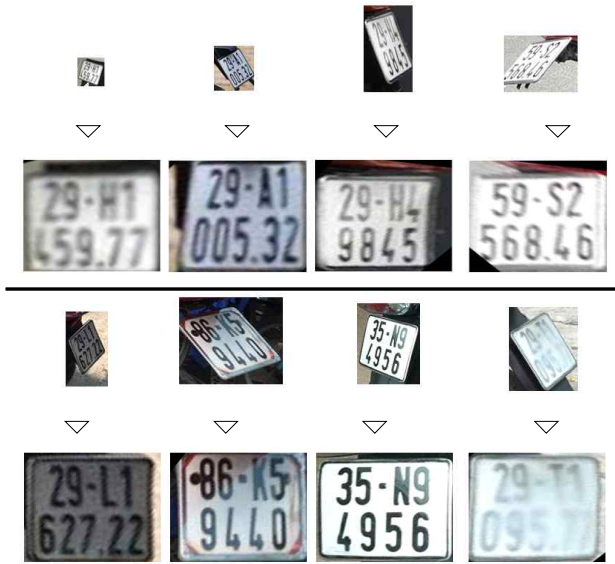


그림 5. 다양한 사이즈 및 각도 가진 번호판 워핑

4.3 문자인식 실험 결과

제안된 방식에서는 Fast-YOLOv2를 수정한 OCR-net를 사용하므로 문자 인식을 위한 경계박스(bounding box)를 찾게 된다. 그림 6에서 보는 바와 같이 영상 워핑 없이 경계박스를 찾게 되면 번호판은 약간 기울어져 있어도 한 경계박스 안에 한 문자아니라 옆에 있는 문자도 일부 들어가는 현상이 나타나기 때문에 인식 잘 안 됨을 알 수 있다. 그러나 번호판을 워핑한 후 문자 인식 잘 됨을 알 수 있다.



그림 6. 문자인식 (위-기존 방법/ 밑 이미지 워핑 적용시)

최종적으로 문자 인식 결과를 표 4에 나타내었다. 번호판내 모든 문자에 대한 인식률을 비교하면, 기존 워핑이 없는 방법은 인식률이 47.74%로 낮은 편이다. 이에 반해 제안한 영상 워핑을 적용한 방식의 인식률은 80.23%를 보였다. 이 결과를 통하여 제안된 YOLOv2기반의 오토바이 번호판 인식 시스템은 전체적으로 오토바이 번호판 특성에 맞는 앵커박스와 이미지 워핑을 통해서 다양한 기울기의 번호판 문자 인식률을 높일 수 있었다.

표 4. 번호판내 모든 문자들의 인식 성공률

	기존 방법	제안된 방법
인식률	47.74%	80.23%

5. 결론

본 논문에는 다양한 각도로 주차된 오토바이내 번호판 인식을 위하여 2-stage YOLOv2 기반의 2-라인 오토바이 번호판 인식 시스템을 구현하였다. 오토바이 및 번호판 영역을 우선 잘 검출하기 위하여 기존 YOLOv2의 앵커 박스 사이즈 및 개수를 조절하였다. 또한 문자 인식할 때 기울어진 번호판 문자를 제대로 인식하지 못하는 점을 개선하기 위하여 번호판 영역내 이미지 워핑 알고리즘 이용하였다. 모의실험 결과, 기존 방법의 인식률은 47.74%에 비해 제안된 오토바이 번호판 인식 방법은 80.23%의 인식률을 얻었다.

향후에 오토바이 뿐 만 아니라 자동차, 버스, 트럭 대상도 인식할 수 있게 적응적인 번호판 인식 방식을 연구하려고 한다.

참고문헌

- [1] J. Redmon and A. Farh.: YOLO9000: Better, Faster, Stronger, Computer Vision and Pattern Recognition(CVPR), 2017 IEEE Conference on, pages 6517-6525
- [2] S. Ren, K. He, R. Girshick and J. Sun.: Faster R-CNN: Towards real-time object detection with region proposal networks, TPAMI, vol. 39, no. 6, pp. 1137-1149, 2017.
- [3] L. Dongsuk, Sook Yoon and L. Jaehwan.: Real-Time License Plate Detection Based on Faster R-CNN, KIPS Tr. Software and Data Eng. Vol.5, No.11, pp.511-520
- [4] S. Zain Masood, G. Shu, A. Dehghan and E. G. Ortiz.: License Plate Detection and Recognition Using Deeply Learned Convolutional Neural Networks, arXiv:1703.07330v2 [cs.CV] 28 Mar 2017
- [5] S. Montazzolli and C. Jung.: Real-Time Brazilian License Plate Detection and Recognition Using Deep Convolution Neural Networks, 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images
- [6] S. Yonetsu, Y. Iwanamoto and Y. Wei Chen.: Two-Stage YOLOv2 for Accurate License-Plate Detection in Complex Scenes
- [7] S. Montazzolli and C. Jung.: License Plate Detection and Recognition in Unconstrained Scenarios, 2018 European Conference on Computer Vision (ECCV)
- [8] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira.: A Robust Real-Time Automatic License Plate Recognition Based on the YOLO Detector, 2018 International Joint Conference on Neural Networks(IJCNN)
- [9] K. Jin-Ho.: Distortion Invariant Vehicle License Plate Extraction and Recognition Algorithm, DOI: 10.5392/JKCA.2011.11.3.001
- [10] L. Hsi-Jian, C. Si-Yuan, W. Shen-Zheng.: Extraction and recognition of license plates of motorcycles and vehicle on highways, Pattern Recognition, 2004. ICPR 2004
- [11] G. Hsu, S. Zeng, C. Chiu and S. Chung.: A Comparison study on motorcycle license plate detection, 2015 IEEE International Conference on Multimedia & Expo Workshops(ICMEW)