

# 블랙박스 영상 기반 차량 및 배경 대체 영상을 이용한 실시간 MR 콘텐츠의 설계

김성호

상지대학교 컴퓨터공학과 교수

## Design of Real-time MR Contents using Substitute Videos of Vehicles and Background based on Black Box Video

Sung-Ho Kim

Professor, Department of Computer Science and Engineering, Sangji University

**요약** 본 논문에서는 차량용 블랙박스로 촬영된 고속도로 주간 주행 영상을 기반으로 차량을 종류별로 검출하고 추적한다. 그리고 검출된 차량의 종류별 대체 영상을 새로운 배경 영상의 같은 위치에 올려놓음으로써 새롭게 창조될 수 있는 실시간 MR 콘텐츠 제작 방안을 설계한다. 차량을 종류별로 검출하고 추적하기 위해서는 딥러닝의 객체 검출 분야에서 가장 잘 알려지고 유명한 YOLO 알고리즘을 사용한다. 또한, 검출된 차량의 종류별 대체 영상을 위해서는 RGB 색상을 기반으로 하는 Mask 기법을 사용한다. 실시간 MR 콘텐츠를 위해 사용될 차량 대체 영상의 크기는 원본 영상에서 검출된 차량의 영역 크기와 같은 크기로 대체된다. 본 논문에서는 실시간 MR 콘텐츠 설계가 가능함을 실험 및 시뮬레이션으로 확인하였으며 VR 콘텐츠 분야에서 유용하게 활용할 수 있을 것으로 판단한다.

**주제어** : 차량용 블랙박스, 차량 검출, 대체 영상, MR, 딥러닝, YOLO

**Abstract** In this paper, we detect and track vehicles by type based on highway daytime driving videos taken with black boxes for vehicles. In addition, we design a real-time MR contents production method that can be newly created by placing substitute videos of each type of detected vehicles in the same location as the new background video. To detect and track vehicles by type, we use the YOLO algorithm. And we also use the mask technique based on RGB color for substitute videos of each type of vehicles detected. The size of the vehicle substitute videos to be used for MR content are substituted by the same size as the area size of the detected vehicles. In this paper, we confirm that real-time MR contents design is possible as a result of experiments and simulations and believe that It will be usefully utilized in the field of VR contents.

**Key Words** : Vehicle Black Box, Vehicle Detection, Substitute Video, MR, Deep Learning, YOLO

### 1. 서론

최근 4차 산업혁명 시대를 대표하는 IT/ICT 분야의 가장 중요한 핵심 기술이며 키워드인 딥러닝 기술은 다

양한 산업 분야에서 연구 및 사업화되고 있다. 딥러닝은 자율주행 자동차와 무인 자동차 등 미래 자동차 산업을 이끌고 가기 위한 핵심 기술로도 활용되고 있는데, 이때 사용되는 가장 큰 기능 중의 하나는 바로 영상

\*This research was supported by Sangji University Research Fund, 2019.

\*Corresponding Author : Sung-Ho Kim(kimsh1204@sangji.ac.kr)

Received April 13, 2021

Revised June 5, 2021

Accepted June 20, 2021

Published June 28, 2021

을 이용한 실시간 객체 인식 능력이다. 즉, 컴퓨터가 영상을 실시간으로 분석하여 인간, 차량, 건물, 도로 등을 검출하고 추적함으로써 차량 스스로 안전하게 자율주행이 가능하게 한다. 전국 고속도로에는 수많은 차량이 차량용 블랙박스로 실시간 전후방 촬영과 녹화를 수행하고 있다. 차량용 블랙박스 영상을 기반으로 실시간 객체를 검출하고 추적하는 기술은 VR/AR 및 MR 분야에서도 활용하거나 응용할 수 있다. 일반적으로 대부분의 VR 콘텐츠들은 가상으로 제작한 3D 공간을 게임, 교육, 의료 등 다양한 산업 분야에서 활용하고 있다. 그리고 360도 파노라마 VR 콘텐츠 등과 같은 실사 영상을 이용한 VR 콘텐츠도 있다. 그러나 차량용 블랙박스 영상을 실시간 VR 콘텐츠 제작에 사용한 사례는 없었다. 만약 다수의 승객이 탑승한 승용차나 고속버스와 같은 대형 차량이 장거리 운행을 할 경우, 지루하고 답답한 차량 내부에서 탑승객들을 위해 차량용 블랙박스의 실시간 영상을 토대로 한 실시간 MR 콘텐츠를 제작하여 제공할 수 있으면 VR 콘텐츠 산업 분야에 매우 큰 영향을 줄 것이다. 그러므로 본 논문은 차량용 블랙박스 고속도로 주간 주행 영상을 사용하여 차량 검출 및 추적을 수행하고, 검출된 차량을 기존에 준비한 영상으로 대체한 후 새로운 영상의 동일 위치에 합성함으로써 새롭게 창조 가능한 실시간 MR 콘텐츠 제작을 위한 설계를 목적으로 한다. 본 논문의 2장에서는 차량용 블랙박스 영상 혹은 CCTV 영상을 기반으로 차량 검출 및 추적과 관련된 관련 연구를 확인한다. 특히, 가장 대표적인 알고리즘 2가지에 대한 선행 연구 결과를 기술한다. 3장에서는 차량용 블랙박스 고속도로 주간 주행 영상을 기반으로 차량 검출 및 추적 알고리즘을 수행한 결과를 기술하고, 4장에서는 실시간 MR 콘텐츠 실험 및 시뮬레이션 결과를 기술한다. 마지막 5장에서는 결론 및 향후 연구 과제를 기술하고 마무리한다.

## 2. 관련 연구

본 논문의 주제 및 목적과 관련된 차량용 블랙박스 영상 기반 차량 검출 및 추적 관련 연구 동향을 살펴보면 대표적인 두 가지 알고리즘 즉, Semantic Segmentation 알고리즘과 YOLO 알고리즘으로 구분된다. Semantic Segmentation 알고리즘은 이미지 속에 있는 객체들을 의미가 있는 단위로 분할하는 것으로서 최근 고속도로 혹은 CCTV 영상을 기반으로 차량뿐

만 아니라 도로, 사람, 건물 등 다양한 객체들의 의미 분할을 위한 연구 결과들[1-11]이 있었다. 여기서 [2-4]와 [8,9]는 Convolutional Networks, [5]는 Pyramid scene parsing network, [6]은 Multi-path refinement networks, [7]은 Pyramid attention network, [11]은 Bilateral segmentation network 등에 중점을 두고 Semantic Segmentation을 수행하고 있다. 정지 영상이나 동영상에 있는 객체들을 의미가 있는 단위로 분할하는 Semantic Segmentation을 이용하여 차량뿐만 아니라 다양한 객체들을 검출하는 연구들은 매우 다양하게 수행되어왔으며 Fig. 1은 그 예시이다. 본 논문에서는 Semantic Segmentation 알고리즘이 고속도로 주간 주행 동영상을 기반으로 차량 검출 및 추적에 적합한지를 확인하기 위해 실험과 시뮬레이션에 사용해보았다.



Fig. 1. Example of semantic segmentation[1]

그러나 그 결과는 본 논문의 목적에는 매우 부적합하다는 것을 확인하였다. 왜냐하면 CCTV와 같이 특정 위치에 고정으로 설치된 카메라 영상의 경우에는 도로를 주행하는 차량을 검출해서 같은 크기의 신규 영상으로 대체하고자 하는 경우 기존의 영상 활용으로 가능하지만, 블랙박스 영상을 기반으로 하는 경우 카메라가 계속 이동하기 때문에 일치하는 영상으로 대체하는 것은 해결하지 못하였기 때문이다. Semantic Segmentation 알고리즘과는 대조적으로 우수한 성능을 자랑하고 있는 딥러닝 기반의 YOLO(You Only Look Once) 알고리즘을 이용하여 블랙박스 혹은 CCTV 영상을 한번 보는 것만으로 객체의 종류와 위치를 추측하는 관련 연구 결과들[12-15]이 있었다. 최근 딥러닝 기반의 객체 인식 알고리즘인 YOLO를 이용한 객체 검출 연구들[12-14]도 있었으며, Fig. 2와 같이 매우 우수한 효율성을 나타내고 있다. 특히, [16]은 동

영상을 기반으로 한 객체 검출을 위해서는 Semantic Segmentation 알고리즘보다 YOLO 알고리즘 효율성이 더 우수하다고 결론을 내리고 있다.

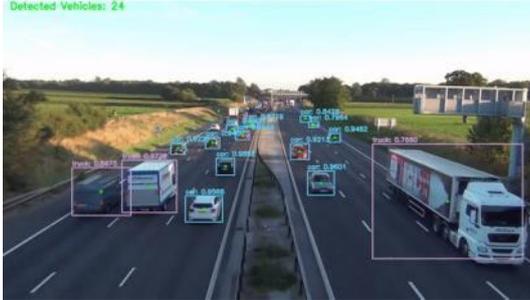


Fig. 2. Example of YOLO vehicle counter [13,14]

그러나 Semantic Segmentation이나 YOLO 등과 같은 알고리즘을 이용하여 영상으로부터 객체 검출과 추적 기능은 있었지만, 검출된 객체를 이용한 활용이나 응용과 관련된 연구 결과는 찾을 수 없었다. 그러므로 본 논문에서는 고속도로 차량용 블랙박스 영상을 기반으로 차량 검출 및 추적을 먼저 수행한 후 실시간 MR 콘텐츠를 설계하고 실험 및 시뮬레이션을 통해 VR 콘텐츠 산업 분야의 적용 가능성을 확인하는 것을 목표로 한다. 또한, 기존 연구 결과와의 대표적인 차별성을 정리해봄으로써 본 연구 결과의 미래 가능성을 확인한다. 차량은 용도와 외형에 따라 승용차, 스포츠카, 상용차, 중장비, 특수차, 이륜차(자전거, 오토바이 등), 농기계, 군용차 등 매우 다양하게 구성되어 있다. 본 논문에서는 다양한 차량의 종류 중에서 고속도로를 주행하는 대표적인 차량인 승용차, 버스, 트럭을 실시간 MR 콘텐츠 설계를 위한 실험 및 시뮬레이션의 주요 대상으로 한다. 물론 시뮬레이션의 확장을 위해 자전거, 오토바이, 기차 등도 포함하여 잠재적인 시뮬레이션의 효율성을 높이도록 설계하였다. 본 논문에서는 차량용 블랙박스의 고속도로 주간 주행 동영상에서 검출된 모든 차량은 유사도 측정으로 계산[12]하였다. 본 논문에서는 YOLO 알고리즘을 이용하여 고속도로 주간 주행 동영상을 기반으로 차량 검출 및 추적을 위한 실험 및 시뮬레이션을 시도해본 결과 매우 우수한 결과를 얻을 수 있었다. 그러므로 본 논문에서는 YOLO 알고리즘 사용하도록 한다.

### 3. MR 콘텐츠 설계 및 시뮬레이션

#### 3.1 시스템 설계 및 모듈 구성

본 논문에서는 차량용 블랙박스의 고속도로 주간 주행 동영상(원본)을 기반으로 차량을 검출 및 추적하고, 검출된 차량의 크기와 위치 정보를 이용하여 원본과 같은 해상도를 가지는 신규 동영상에 원본에서 검출된 차량의 대체 동영상을 배치함으로써 실시간 MR 콘텐츠를 구현하기 위한 설계를 목표로 하고 있다.

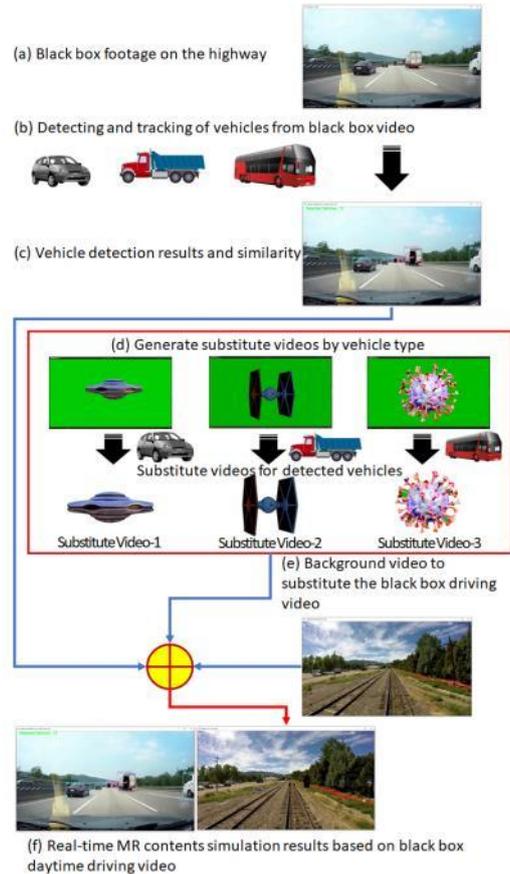


Fig. 3. System Configuration and Flow Chart

Fig 3은 본 논문에서 제안한 연구 방법과 절차, 시뮬레이션을 위한 모듈 구성과 기능 등을 포함하는 전체적인 시스템의 흐름도를 나타낸 것이다. 본 논문의 목적을 위한 첫째 단계는 Fig 3의 (a)와 같이 차량용 블랙박스 고속도로 주간 주행 영상 파일을 준비하는 것이다. 둘째 단계는 (b)와 같이 (a)의 영상에서 Yolo 알고리즘

을 이용하여 차량을 종류별로 검출하고 추적한다. 이때 (c)와 같이 (b)에서 유사도 측정[12]을 통해 검출 및 추적된 차량의 종류를 확인한다. 셋째 단계는 (d)와 같이 (c)에서 검출된 대표적인 차량(승용차, 트럭, 버스)의 대체 영상을 준비하고 교체하도록 한다. 여기서 (d)에 대한 자세한 설명은 3.2에서 기술하도록 한다. 넷째 단계는 (a)와 동일 해상도를 가지는 신규 영상을 (e)와 같이 준비한다. 마지막 다섯째 단계는 (c)에서 유사도 측정으로 검출 및 추적된 차량의 종류별 대체 영상을 (d)와 같이 생성하고 (e)와 같은 신규 동영상의 동일 위치에 합성함으로써 (f)와 같은 실시간 MR 콘텐츠가 완성될 수 있음을 실험 및 시뮬레이션을 통해 확인한다.

### 3.2 검출된 차량을 대체 영상으로 교체

본 논문에서 실험 및 시뮬레이션에 사용하기 위한 차량용 블랙박스의 고속도로 주간 주행 동영상 속에 등장하는 차량의 종류는 고속도로에서 쉽게 볼 수 있는 승용차, 버스, 트럭 등 3가지로 제한한다. 특히, 본 논문에서는 YOLO 알고리즘으로 검출한 3가지 차량에 대해서 서로 다른 색상의 직사각형으로 구분[12]한다. 실시간 MR 콘텐츠 설계 및 시뮬레이션을 위해서는 고속도로 주간 주행 동영상 속에 등장하는 차량을 Fig. 3의 (b) 및 (c)와 같이 검출한 후 (d)와 같이 동일 해상도를 가지는 영상으로 대체한다. 즉, 본 논문에서는 원본 영상에서 검출된 차량을 대체하기 위한 영상으로 Fig. 4와 같이 배경이 녹색으로 채워진 3가지 동영상 파일 (Video 1, 2, 3)을 사용한다. 먼저 동영상에서 검출한 차량을 포함하는 직사각형 크기와 같은 동일 해상도로 배경이 녹색으로 채워진 대체 영상의 해상도 크기를 조정한다. 배경이 녹색으로 채워진 대체 영상은 모두 RGB 색상을 기본적으로 가지고 있는데, 본 논문에서는 녹색 배경을 제거한 후 사용하기 위해 Mask 기법을 사용한다. 본 논문에서는 동영상의 RGB 색상을 HSV 색상으로 변경한 후 녹색 배경을 Mask로 지정하는데, 녹색 배경 색상의 범위는 실험을 거쳐 최솟값은 HSV = (50, 100, 100), 최댓값은 HSV = (100, 255, 255)로 설정하였다.

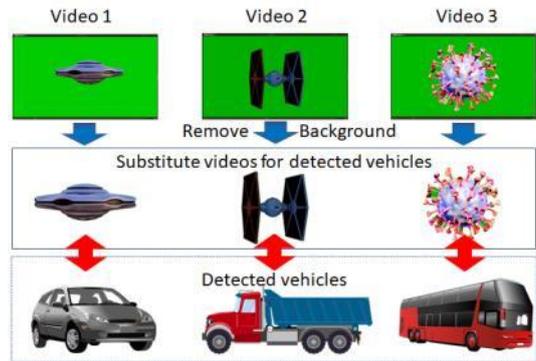


Fig. 4. Substitute videos according to vehicle type

Mask로 선택된 녹색 배경은 Bitwise\_not 연산을 이용하여 녹색 배경을 제거한 남은 객체 부분을 선택하게 하였다. 마지막으로 녹색 배경을 제거한 객체 부분만 보이는 동영상을 고속도로 주간 주행 동영상과 같은 해상도를 가진 새로운 동영상 위에 배치한다. 본 논문에서는 실험을 위하여 차량용 블랙박스의 고속도로 주간 주행 동영상에서 검출된 차량의 종류에 따라 서로 다른 색상의 직사각형(x, y, w, h)으로 나타내는데, 같은 위치의 직사각형 시작점(x, y)을 기준으로 같은 크기의 해상도(w, h)로 새로운 동영상에 배치한다.

### 3.3 MR 콘텐츠 실험 및 시뮬레이션 결과

본 논문에서는 고속도로 주간 주행 동영상을 기반으로 차량을 검출 및 추적하면서 차량의 종류에 따른 대체 동영상을 신규 동영상과 혼합함으로써 실시간 MR 콘텐츠 설계를 실험 및 시뮬레이션으로 확인하였다. 원본을 대체할 새로운 동영상은 철로를 달리는 기차의 조종석에서 촬영한 철로 전방 동영상을 기반으로 하고 있으며, 고속도로 주간 주행 동영상에서 검출한 차량을 대체할 대체 동영상은 UFO, 우주선, 코로나바이러스 모형 등을 사용하였다. Fig. 5의 (a)는 고속도로 주간 주행 동영상, (b)는 (a)를 기반으로 YOLO 알고리즘을 이용하여 검출한 차량과 그 종류 및 유사도를 나타내고 있다. 그리고 (c)는 (a)를 대체할 새로운 동영상이고, (d)는 (b)에서 검출 및 추적한 차량의 종류, 위치 및 크기에 대응하는 대체 동영상을 (c)에 배치한 실시간 MR 콘텐츠 설계를 위한 실험 및 시뮬레이션의 결과물이다.

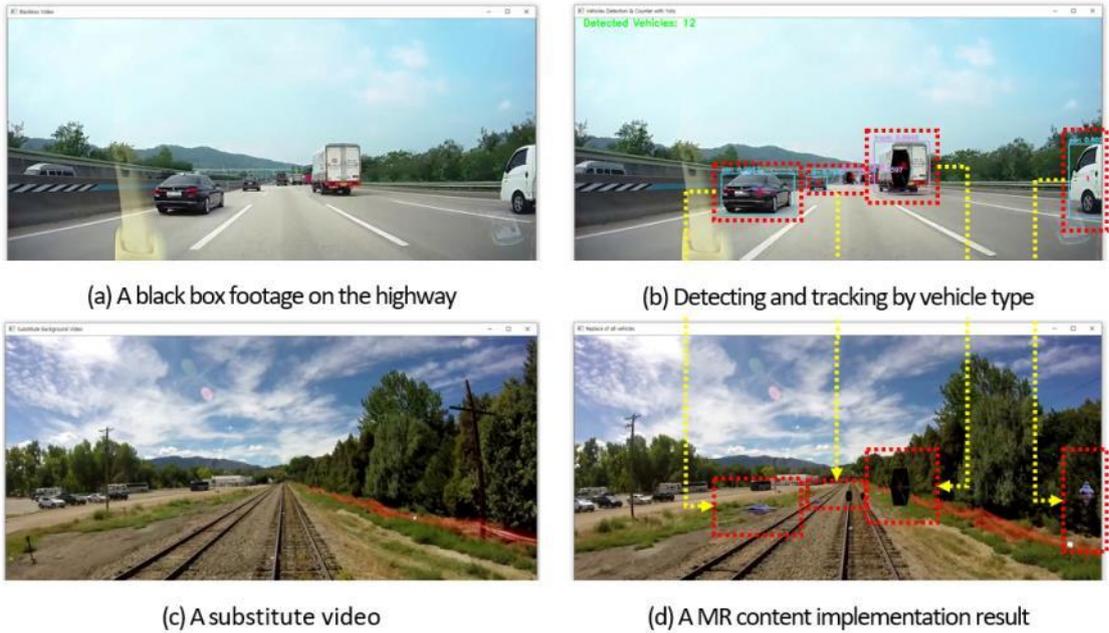


Fig. 5. A real time MR contents simulation result

즉, Fig. 5의 (b)에서 검출 및 추적된 승용차와 트럭은 (d)에서 UFO와 우주선으로 대체되었다. 본 논문에서 실험 및 시뮬레이션을 위해 사용한 영상 (a)는 1280×720 크기의 해상도와 30fps로 설정한 차량용 블랙박스(아이나비 QXD1000 모델) 주간 주행 영상이다. (c)는 Royal Gorge Route Railroad를 운전사의 시각에서 촬영한 주간 영상으로서 (a)와 같은 해상도를 가지는 새로운 동영상이다. 본 논문의 실험 및 시뮬레이션을 위한 시스템 개발 환경으로는 운영체제 Windows 10에서 Anaconda 3 기반의 Python v3.7 과 MS Visual Code를 사용하였으며, Numpy, imutils, OpenCV2 등의 라이브러리를 추가로 사용하였다. 본 논문에서는 다양한 모델의 차량용 블랙박스 고속도로 전후방 주간 주행 영상 수 개를 모두 실험 및 시뮬레이션에 사용하였다. 그 결과 차량 검출 및 추적이 정상적으로 이루어질 경우, 실시간 MR 콘텐츠 제작을 위한 과정에서 나타난 문제점은 없었음을 확인하였다. 그러므로 본 논문에서 제안한 블랙박스 영상 기반 차량 검출 및 대체 영상을 이용한 실시간 MR 콘텐츠 설계는 정지 영상이 아닌 동영상을 사용한 실험 및 시뮬레이션 결과에서 알 수 있듯이 MR 콘텐츠에 있어서 새로운 분야를 개척하였다고 자부할 수 있게 되었다. 결국 본 논문에서는 Table 1과 같이 차량 검출이 목표

인 기존 연구에서 더 나아가 실시간 MR 콘텐츠 제작으로 연구 영역을 확대함으로써 기존 연구와의 차별성을 나타낼 수 있었다.

Table 1. Differentiation from existing research results

Division	Existing Researches	This Research
Front and rear vehicle detection	○	○
Detected Vehicle Replacement	×	○
MR Content Design	×	○

## 5. 결론 및 향후 연구

본 논문에서는 차량용 블랙박스의 고속도로 주간 주행 영상을 기반으로 차량을 검출 및 추적한 후 원본과 같은 해상도를 가진 새로운 동영상에 검출된 차량을 대체 동영상으로 대체하고 혼합함으로써 실시간 MR 콘텐츠 제작이 가능함을 설계하고 실험 및 시뮬레이션으로 확인하였다. 차량용 블랙박스의 고속도로 주간 주행 동영상에서 검출된 차량을 대체할 대체 동영상을 위해서는 검출된 차량의 크기와 같은 녹색 배경을 가진 동영상을 사용하였으며, 배경을 제외한 객체만을 사용하기 위해서는 Mask 기법이 유용하다는 것을 확인하였

다. 본 논문에서 제시하고 실험 및 시뮬레이션으로 확인한 MR 콘텐츠는 차량을 이용하는 모든 여행객을 대상으로 실시간 블랙박스 영상 기반 MR 콘텐츠 제작 및 응용 분야에 활용할 수 있다. 향후 연구에서는 차량용 블랙박스 고속도로 주간 주행 동영상을 새로운 동영상으로 교체하지 않고 고속도로의 바닥 면과 검출된 차량만 새로운 객체 영상으로 대체하고자 한다. 또한, 최근 MR 콘텐츠 제작 분야에서 유행하고 있는 실사 기반의 360도 파노라마 VR 및 360도 카메라로 촬영한 360도 동영상을 이용한 실시간 MR 콘텐츠의 설계도 추가적인 연구와 실험 및 시뮬레이션을 통해 수행하고자 한다.

## REFERENCES

- [1] [https://docs.openvinotoolkit.org/latest/omz\\_models\\_intel\\_semantic\\_segmentation\\_adas\\_0001\\_description\\_semantic\\_segmentation\\_adas\\_0001.html](https://docs.openvinotoolkit.org/latest/omz_models_intel_semantic_segmentation_adas_0001_description_semantic_segmentation_adas_0001.html)
- [2] O. Ronneberger, P. Fischer & T. Brox. (2015, October). U-net: Convolutional networks for biomedical image segmentation. *In International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
- [3] V. Badrinarayanan, A. Kendall & R. Cipolla. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12), 2481-2495. DOI : 10.1109/TPAMI.2016.2644615
- [4] A. Kendall, V. Badrinarayanan & R. Cipolla. (2018). SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding. *arXiv preprint arXiv:1511.02680*
- [5] H. Zhao, J. Shi, X. Qi, X. Wang & J. Jia. (2017). Pyramid scene parsing network. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2881-2890).
- [6] G. Lin, A. Milan, C. Shen & I. Reid. (2017). Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1925-1934).
- [7] H. LI et al. (2018). Pyramid attention network for semantic segmentation. *arXiv preprint arXiv:1805.10180*.
- [8] L. C. Chen, G. Papandreou, F. Schroff & H. Adam. (2017). Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*.
- [9] L. C. Chen et al. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. *In Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818).
- [10] M. Yang, K. Yu, C. Zhang, Z. Li & K. Yang. (2018). Densnaspp for semantic segmentation in street scenes. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3684-3692).
- [11] C. Yu et al. (2018). Bisenet: Bilateral segmentation network for real-time semantic segmentation. *In Proceedings of the European conference on computer vision (ECCV)* (pp. 325-341).
- [12] <https://github.com/guptavasul213/Yolo-Vehicle-Counter>
- [13] J. Redmon & A. Farhadi. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- [14] T. N. Doan & M. T. Truong. (2020, November). Real-time vehicle detection and counting based on YOLO and DeepSORT. *In 2020 12th International Conference on Knowledge and Systems Engineering (KSE)* (pp. 67-72). IEEE.. DOI : 10.1109/KSE50997.2020.9287483
- [15] D. Bolya, C. Zhou, F. Xiao & Y. J. Lee. (2019). Yolact: Real-time instance segmentation. *In Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9157-9166).
- [16] Y. H. Lee & Y. Kim. (2020). Comparison of CNN and YOLO for Object Detection. *Journal of the semiconductor & display technology*, 19(1), 85-92.

김 성 호(Sung-Ho Kim)

[정회원]



- 1998년 8월 : 송실대학교 컴퓨터학과(공학석사)
- 2005년 2월 : 송실대학교 컴퓨터학과(공학박사)
- 2006년 3월 : 상지대학교 컴퓨터정보공학부 교수
- 2018년 3월~현재 : 상지대학교 컴퓨터공학과 교수
- 관심 분야 : 컴퓨터그래픽스, 컴퓨터 애니메이션, 모션 캡처, 가상현실/증강현실, 드론, 컴퓨터게임, Web3D, 멀티미디어, 지리정보시스템
- E-Mail : kimsh1204@sangji.ac.kr