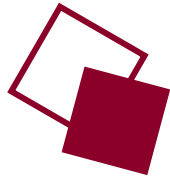


Machine Learning

# 7장 CNN의 응용-자율자동차

---

고려대학교 통계학과  
박유성



# Contents

**01** 교통표식의 인식

**02** 객체 분할(instance segmentation)

**03** 객체의 의미분할(semantic segmentation  
of objects)

# CNN의 응용-자율자동차

- 교통표식 인식, 여러 이미지 중에서 특정 이미지(예를 들어 자동차)를 분리인식, 그리고 여러 물체를 종류별로 동시에 구별해서 인식하는 것은 자율자동차를 운행하기 위한 필수적인 사항이다.
- 자동차는 낮과 밤에 관계없이 운행되므로 교통표식이나 이미지의 밝기는 낮과 밤이 뚜렷한 차이가 있으며 이미지 크기 또한 다를 경우가 많다.
- 이는 이미지자료의 사전자료정리과정이 필수적임을 의미하며 색의 밝기나 이미지자료의 크기에 대한 일종의 표준화작업이 필요하다.

# 01 교통표식의 식별

- 총 43개 종류의 교통표식 이미지를 식별하기 위해 CNN 아키텍처를 사용하고자 한다.
- 교통표식 컬러이미지 자료가 낮과 밤에 찍힌 것을 포함하여 다양한 색의 밝기를 가진 자료로 제공되어 있으므로
- 교통표식을 색의 질, 색의 선명도, 색의 밝기로 분리한 후, 색의 밝기를 표준화할 것이다.

## 02 객체 분할(instance segmentation)

- 객체분할은 다음 절에서 논의할 객체의 의미분할(semantic segmentation of objects)과 구별된다.
- 객체분할은 이미지에서 하나의 객체만을 분할·인식하는 데에 목적이 있지만, 객체의 의미분할은 이미지에서 2개 이상의 객체를 의미론적 관점에서 분할·인식하는 데에 목적이 있다.
- 여기에서 의미론적 관점이란, 사람, 건물, 자동차, 나무 등의 일반명사적인 관점에서의 분할을 의미한다.
- 객체분할은 사람과 같이 일반명사적인 관점에서 하나의 객체를 인식하기도 하지만 특정인이라는 고유명사적 관점에서 하나의 객체를 인식하는 것을 목적으로 할 수 있다. 이 절에서 다루는 객체분할은 일반명사적인 관점에서 자동차를 인식하는 데에 목적을 두고 있다.

## 02 객체 분할(instance segmentation)

- 객체분할의 정밀도를 높이기 위해 이미 제 5장에서 다룬 바 있는 사전 학습된 VGG16 모형의 이전학습과 딥러닝모형의 성능향상을 위한 잔차연결(residual connection)을 이용할 것이다.
- 딥러닝의 은닉층이 너무 많을 때, 잔차연결(제 11장에서 자세하게 다룰 것임)은 초기 은닉층의 정보손실 방지를 위해 초기 은닉층을 후반부 은닉층에 직접 연결하는 기법이다.
- 이를 위해 keras의 function API (제 9장에서 자세하게 논의할 것임)를 사용하여 객체분할을 실행할 것이다.

## 02 객체 분할(instance segmentation)

- `UpSampling2D(size=(2,2))`

- $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \\ 3 & 3 & 4 & 4 \end{pmatrix}$

- `UpSampling2D(size=(2,3))`

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 2 & 2 & 2 \\ 3 & 3 & 3 & 4 & 4 & 4 \\ 4 & 4 & 3 & 4 & 4 & 4 \end{pmatrix}$$

## 03 객체의 의미분할(semantic segmentation of objects)

- 객체의 의미분할은 객체분할(instance segmentation)의 일반화로 생각하면 쉽게 이해할 수 있다.
- 즉, 논리적으로 2개 이상의 객체분할이 객체의 의미분할로 해석할 수 있다. 객체 분할에서 자동차 label인 8이면 1, 그렇지 않으면 0을 부여한 (224,224,1)인 channel이 하나인 목적변수  $y_2$ 를 생성한 바 있다.
- 만약 객체가 12개 있다면 channel을 12개로 늘려서 (224,224,12)로 목적변수를 정의하면 된다.
- 물론 각 channel은 해당 객체에 따라 1 또는 0값을 갖도록 한다. 이러한 목적변수의 구성으로 손실함수는 categorical crossentropy가 되며,



## 03 객체의 의미분할(semantic segmentation of objects)

- $224 \times 224$ 의 각 셀에는 12개의 channel 존재하므로
- categorical crossentropy에 대응하는 softmax 함수에 의해 각 셀이 12개 channel 각각에 포함될 확률을 부여하게 된다.
- 그러므로 손실함수는  $224 \times 224$  개의 categorical crossentropy의 합이 된다.

Q & A