

딥러닝팀

1팀

김예찬

윤지영

채소연

한지원

홍지우

INDEX

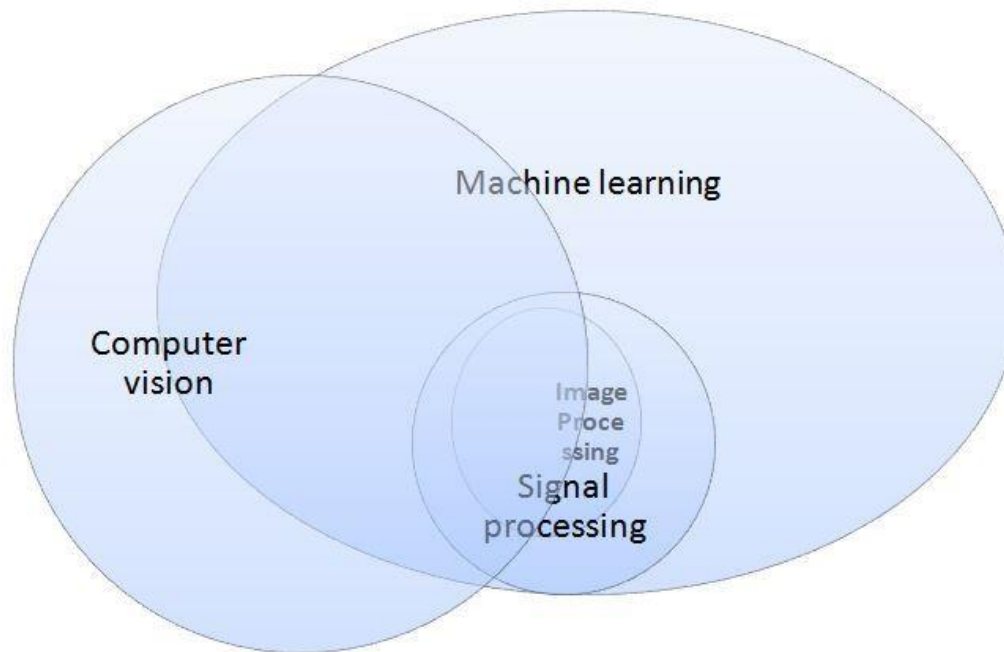
1. 컴퓨터 비전
2. 자연어처리
3. 기타

1 컴퓨터 비전

- 컴퓨터 비전(Computer Vision)

컴퓨터 비전이란?

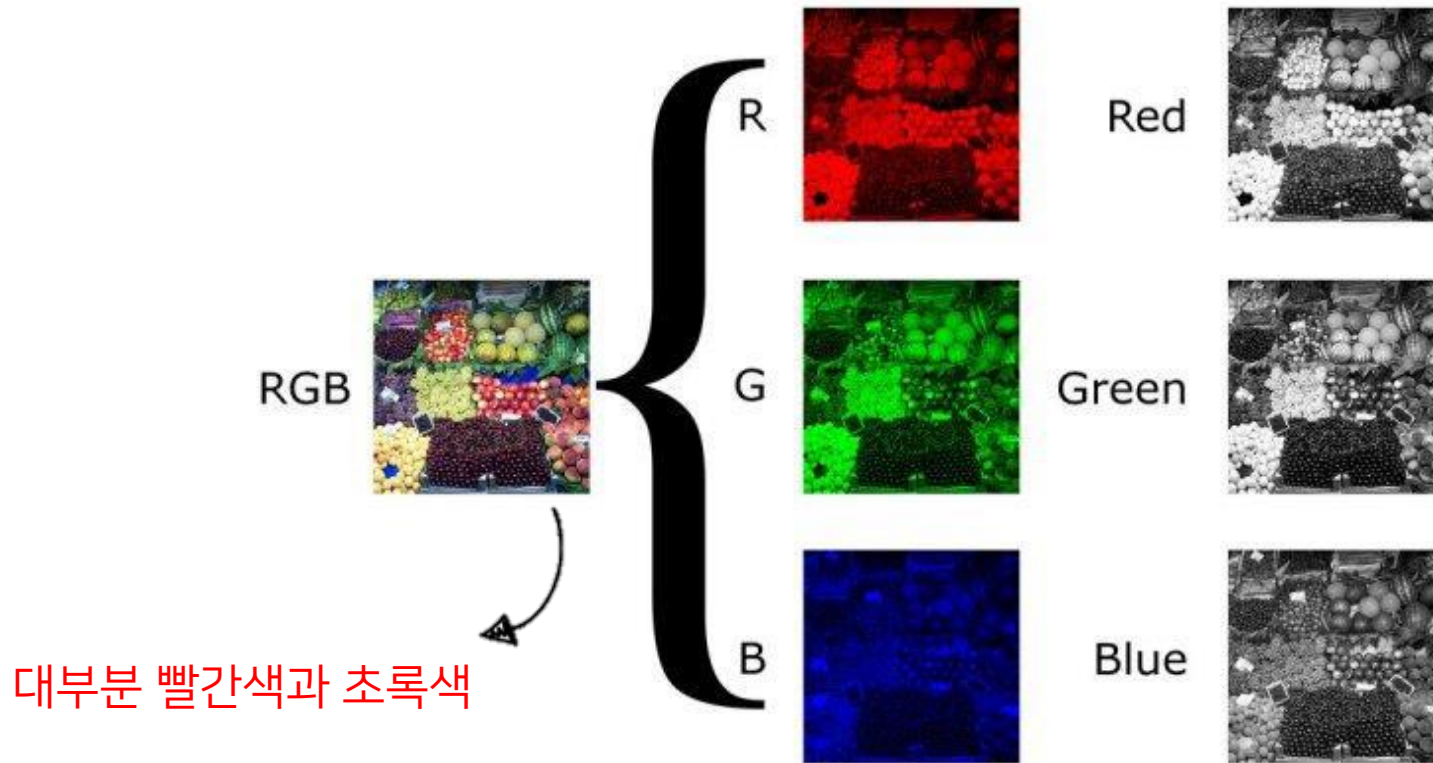
: 컴퓨터가 시각적인 세계를 이해하고 해석할 수 있도록 컴퓨터를 학습시키는 것이 목적인 연구 분야



1 컴퓨터 비전

- 이미지 데이터

이미지 데이터의 구조

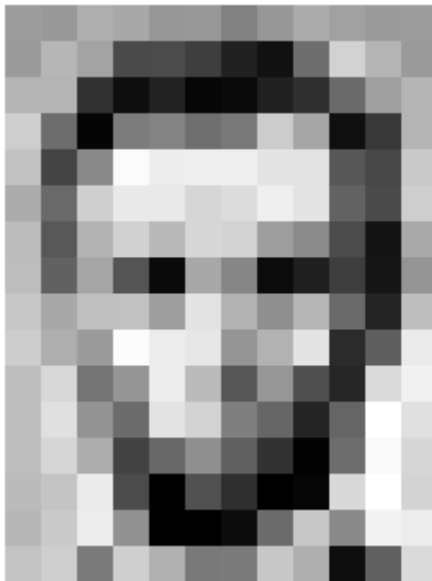


1 컴퓨터 비전

- 이미지 데이터

이미지 데이터의 구조

: 이미지를 이루는 픽셀의 수가 많아지면 해상도가 높아짐



157	153	174	168	150	152	129	151	172	181	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	85	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

8비트 이미지 (0-255)

157	153	174	168	150	152	129	151	172	181	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	85	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

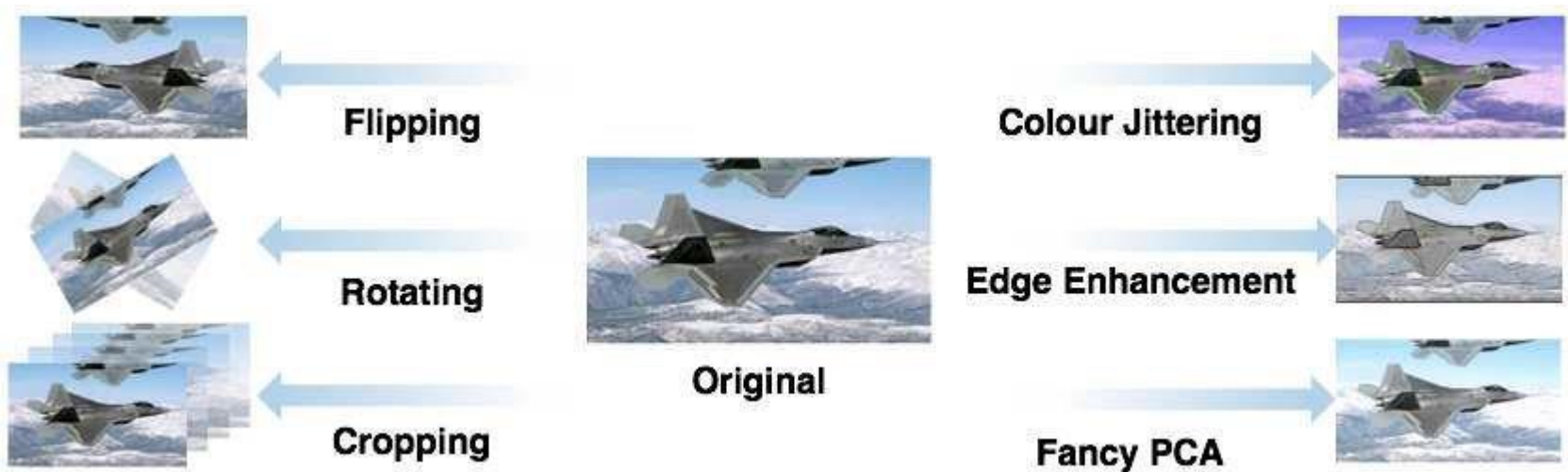
비트 = 정보의 양 ($256 = 2^8$ = 8비트)

1 컴퓨터 비전

- 이미지 데이터

이미지 데이터의 변형 (Data Augmentation)

: 비정형 데이터 분석 시, 데이터의 수를 인위적으로 늘려주는 기법

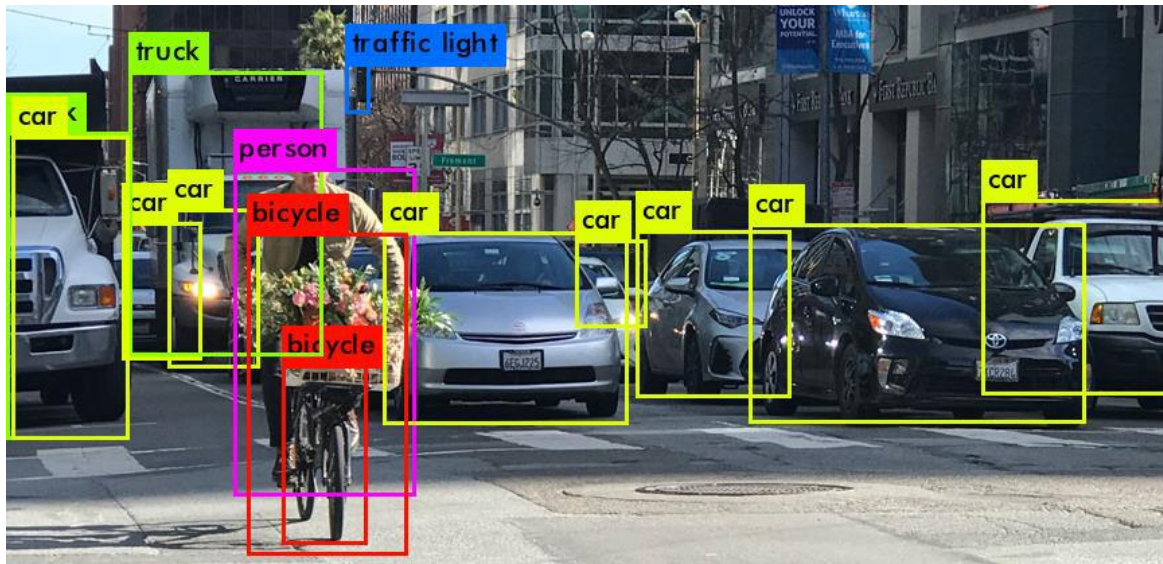


1 컴퓨터 비전

● 활용 분야 - (1) Object Detection

Object Detection

: 이미지 내에 존재하는 객체들을 탐지하는 과제



Bbox

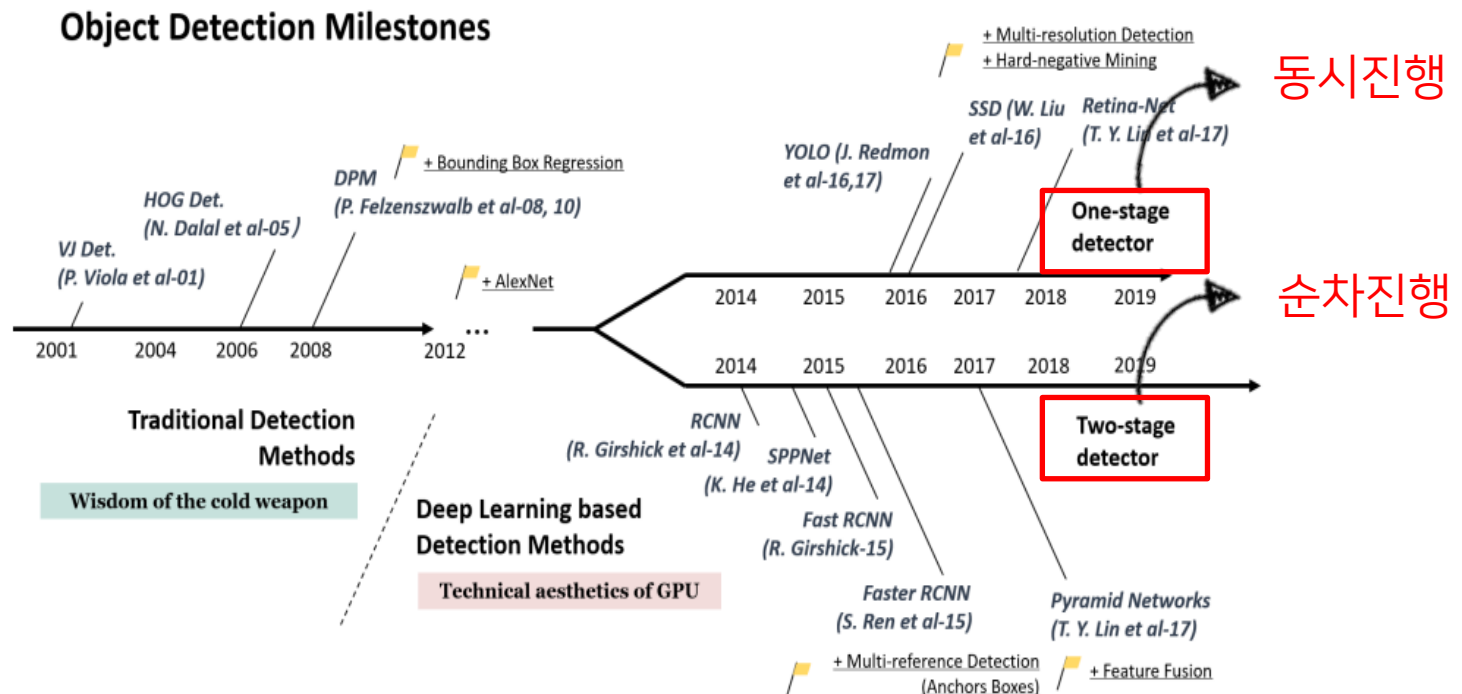
- Localization: 이미지에 맞는 bbox를 찾는 과정
- Classification: 객체를 분류하여 라벨을 예측하는 과정

1 컴퓨터 비전

● 활용 분야 - (1) Object Detection

Object Detection

- Localization과 Classification을 동시에 진행하는지 여부에 따라 두 가지로 나뉨



1 컴퓨터 비전

- 활용 분야 - (1) Object Detection

Object Detection

- One-Stage Detector & Two-Stage Detector

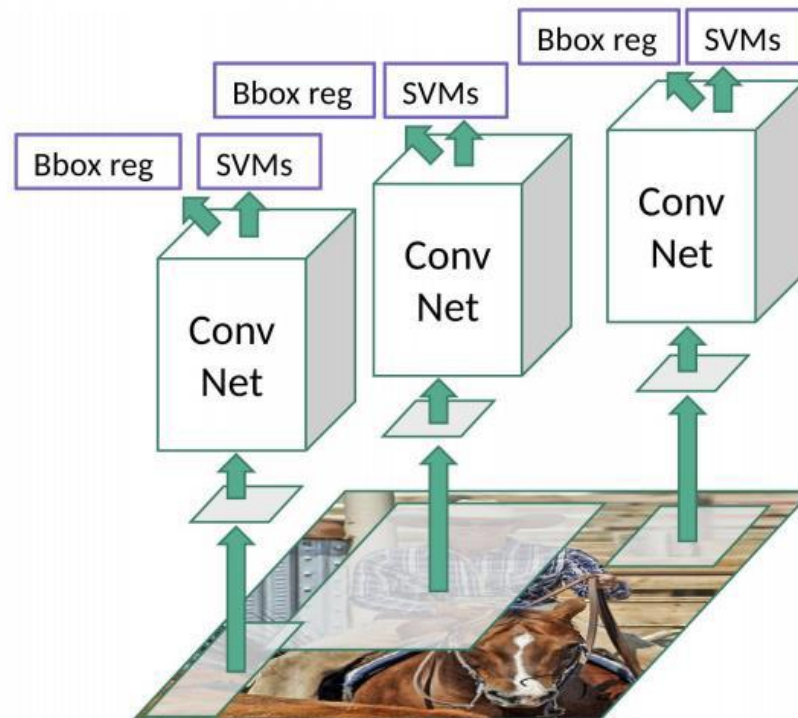
One-Stage Detector	Two-Stage Detector
Localization과 Classification 동시 진행	Localization과 Classification 순차 진행
빠른 작업 수행	느린 작업 수행 & 높은 정확도

1 컴퓨터 비전

- 활용 분야 - (1) Object Detection

RCNN(Region-based CNN)

: 2-Stage Detector의 대표적인 모델

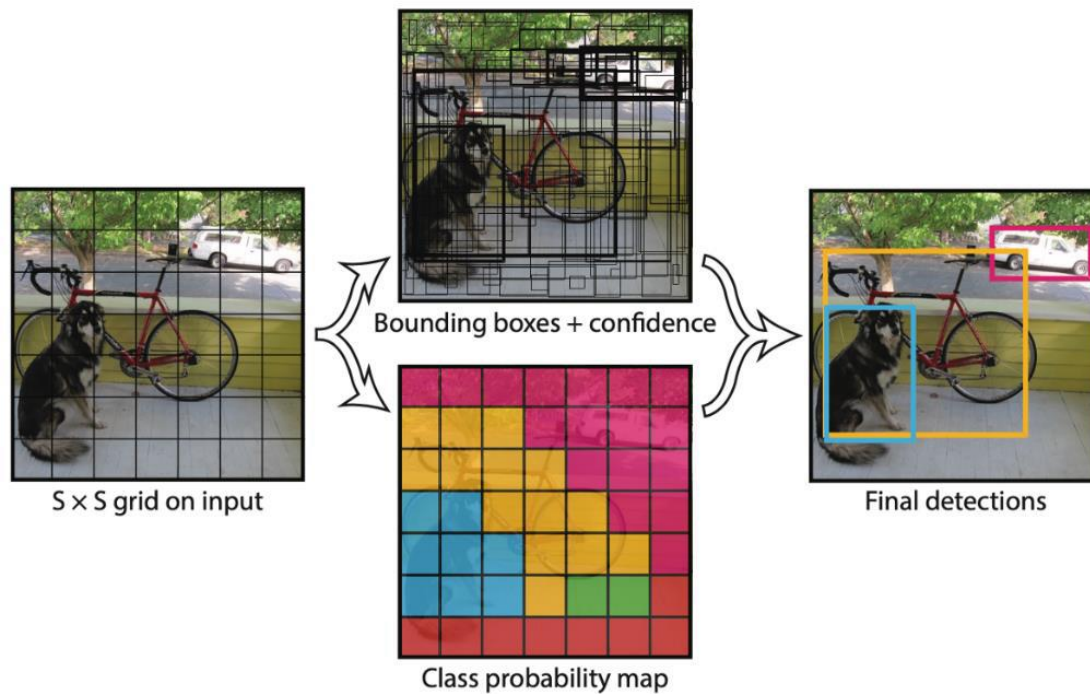


1 컴퓨터 비전

- 활용 분야 - (1) Object Detection

YOLO(You-Only-Look-Once)

: 1-Stage Detector의 대표적인 모델



1 컴퓨터 비전

- 활용 분야 - (2) Image Segmentation

Image Segmentation

: 각 픽셀이 어떤 분류에 해당하는지 라벨을 예측함으로써 객체 간 경계 탐지



Semantic
Segmentation

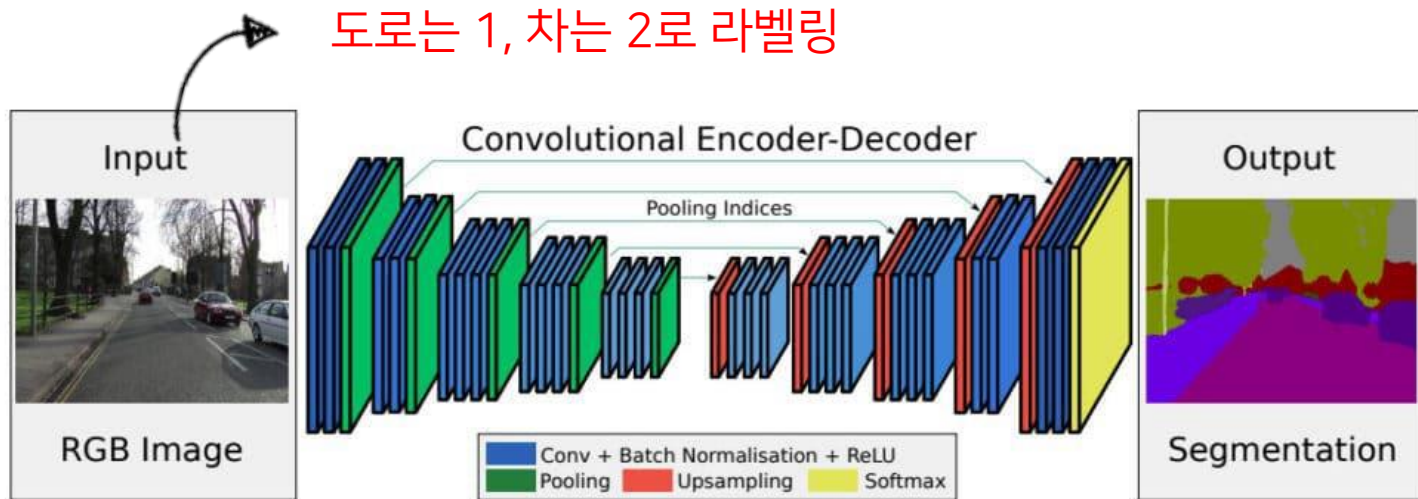
Instance
Segmentation

1 컴퓨터 비전

● 활용 분야 - (2) Image Segmentation

Semantic Segmentation

- 이미지 내의 모든 픽셀에 라벨을 부여하는 것이 기본 아이디어

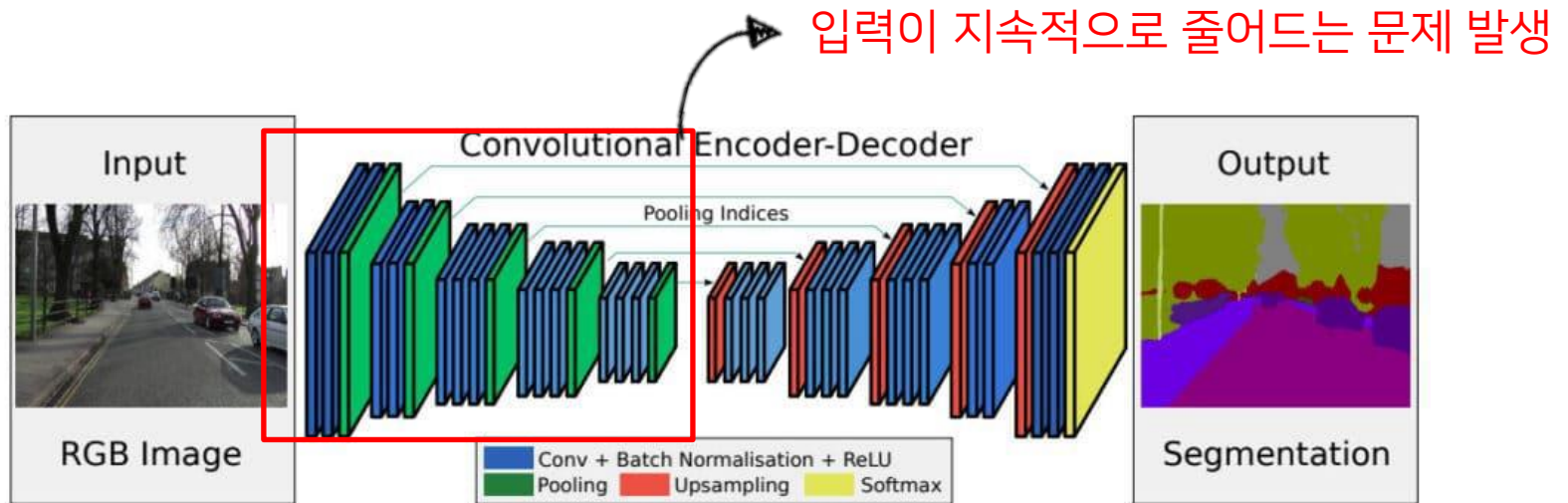


1 컴퓨터 비전

● 활용 분야 - (2) Image Segmentation

Semantic Segmantation

- 이미지 내의 모든 픽셀에 라벨을 부여하는 것이 기본 아이디어

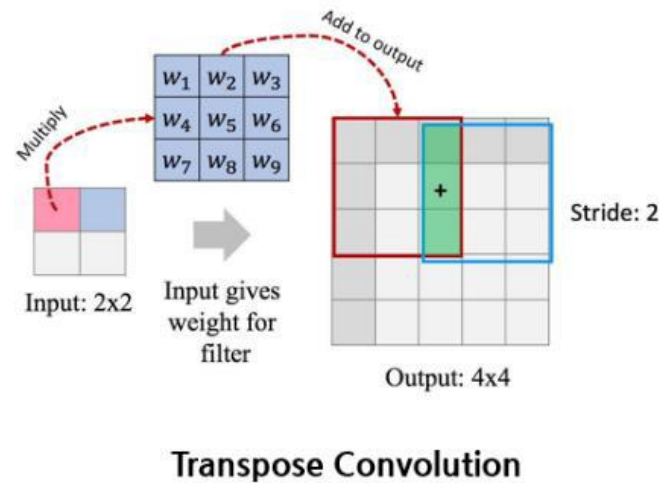
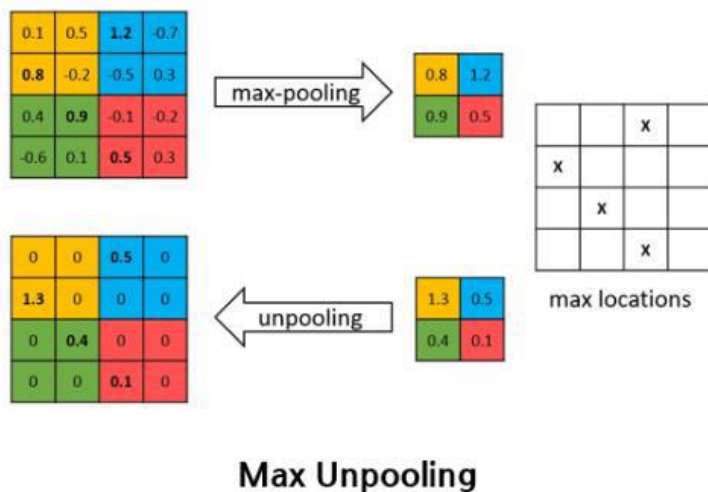


1 컴퓨터 비전

● 활용 분야 - (2) Image Segmentation

Semantic Segmentation

- Upsampling으로 입력의 사이즈를 키움으로서 문제를 해결

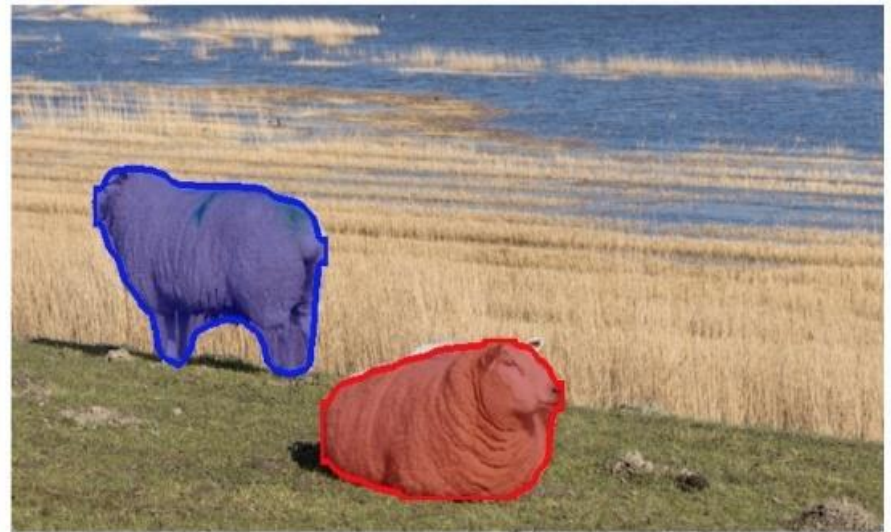


1 컴퓨터 비전

- 활용 분야 - (2) Image Segmentation

Instance Segmentation

- 같은 양이라도 서로 다른 객체로 분리하여 그 경계를 확인함

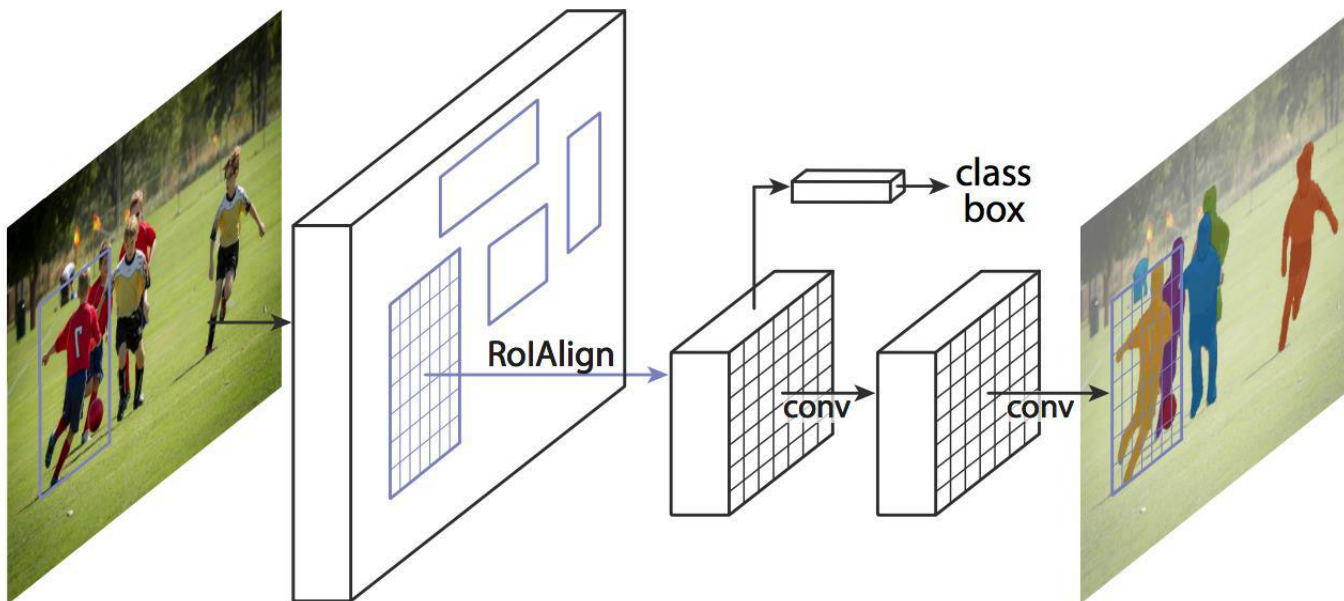


1 컴퓨터 비전

- 활용 분야 - (2) Image Segmentation

Instance Segmentation - Mask RCNN

: Instance Segmentation을 수행하는 대표적인 모델

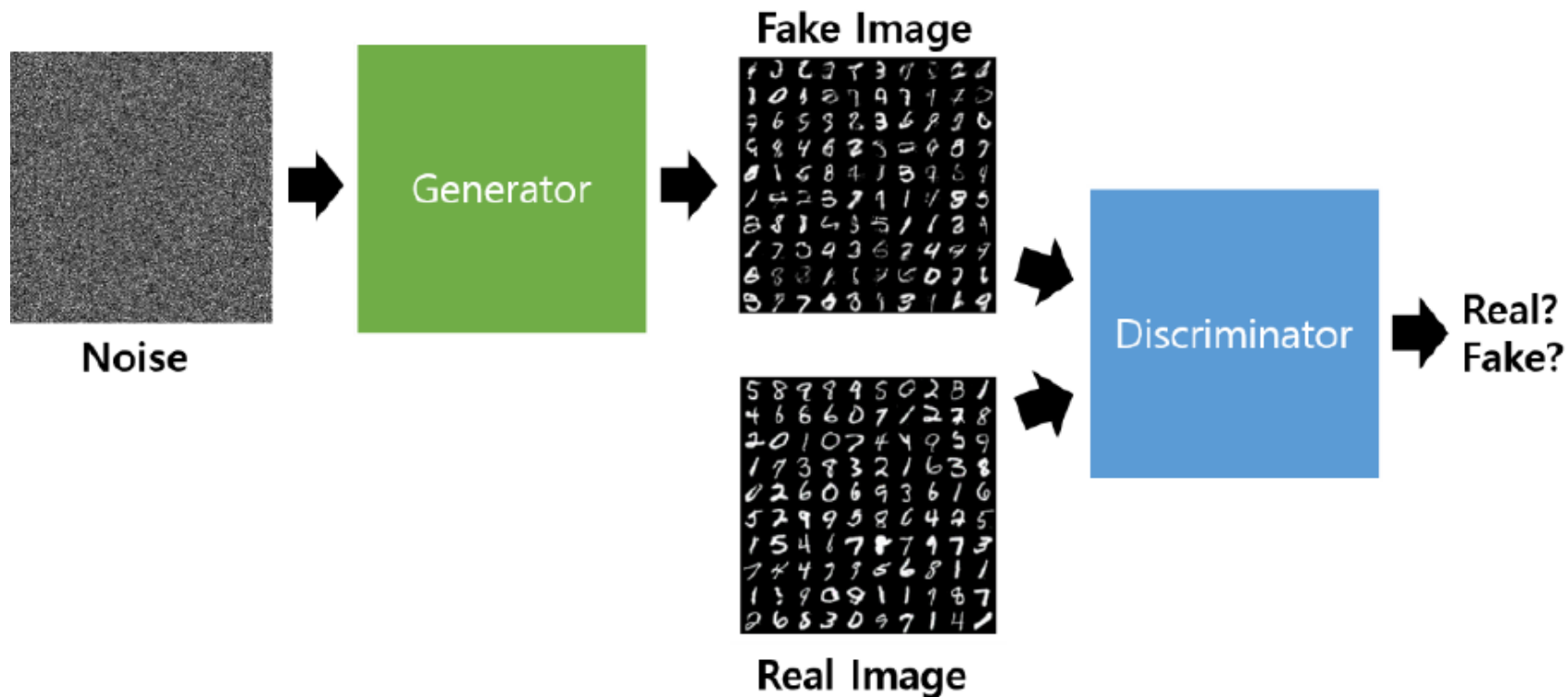


1 컴퓨터 비전

- 활용 분야 - (3) Image Generation

Image Generation

: 주어진 이미지를 바탕으로 새로운 이미지 생성

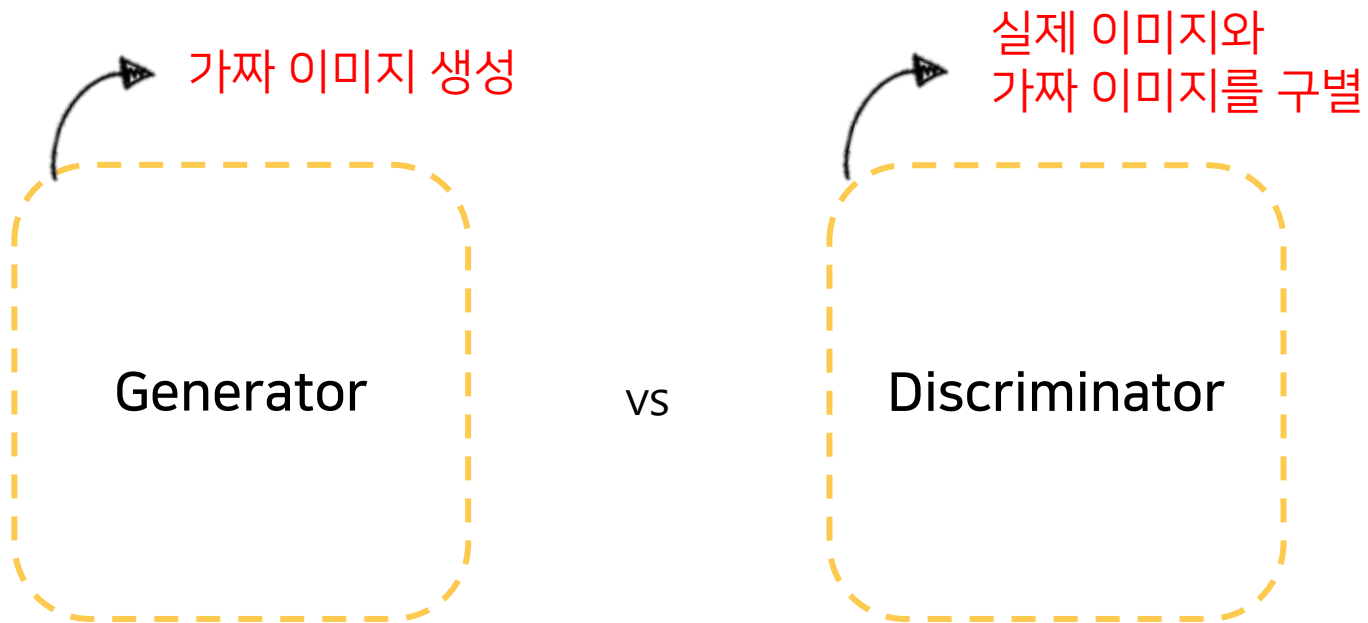


1 컴퓨터 비전

- 활용 분야 - (3) Image Generation

GAN (Generative Adversarial Network)

: 주어진 이미지를 바탕으로 새로운 이미지를 만드는 모델

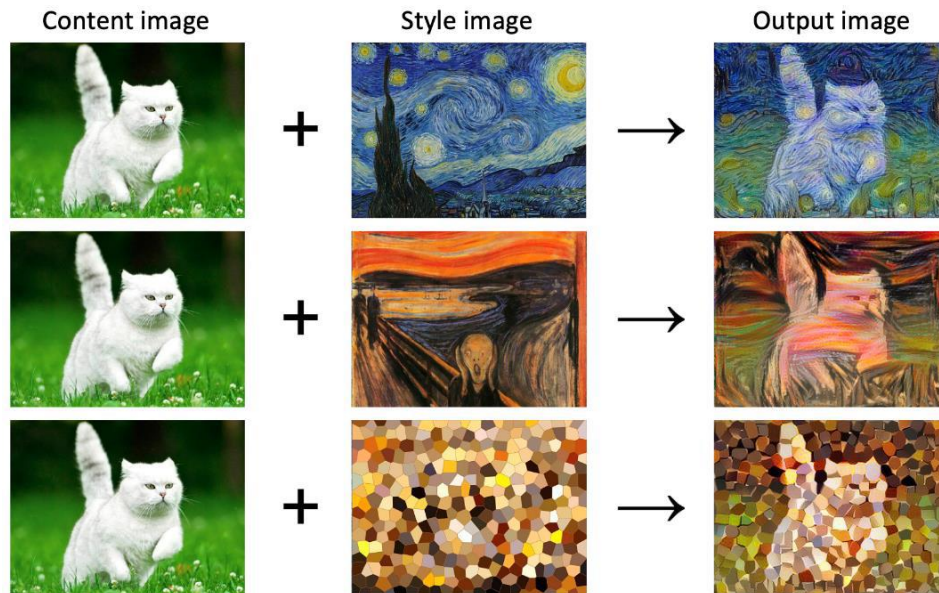


1 컴퓨터 비전

- 활용 분야 - (3) Image Generation

Style Transfer

: Style image의 특징을 추출하여 Content image에 적용하는 것을 목적으로 하는 알고리즘



Content Difference, Style Difference를 줄이는 것을 목적으로 손실함수 구성

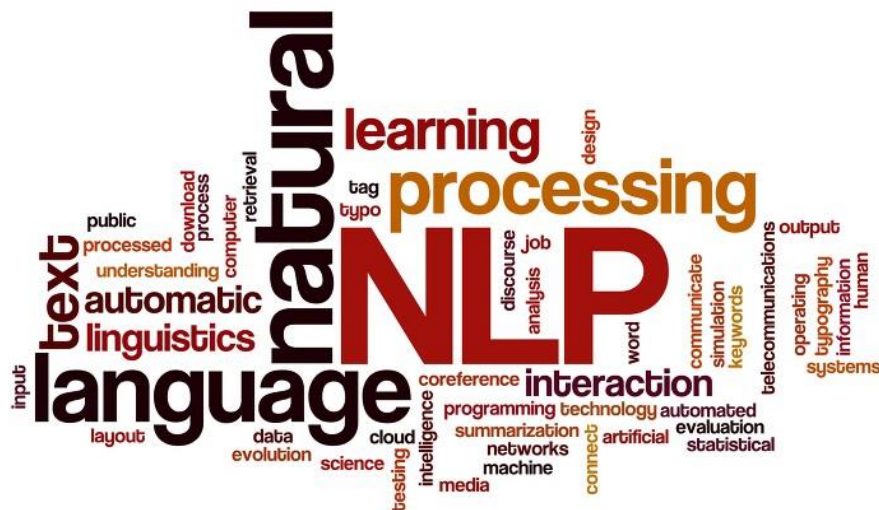
2

자연어처리

2 자연어처리 (Natural Language Processing)

- 자연어처리

자연어처리란?



- 사람의 언어 → 컴퓨터가 이해하는 언어
- 텍스트 데이터의 순서 정보 (RNN 모델 사용)

2 자연어처리 (Natural Language Processing)

- 텍스트 데이터의 변환

자연어의 속성 1: 모호성

원문	나는 철수를 안 때렸다.
해석 1	철수는 맞았지만, 때린 사람이 나는 아니다.
해석 2	나는 누군가를 때렸지만, 그게 철수는 아니다.
해석 3	나는 누군가를 때린 적도 없고, 철수도 맞은 적이 없다.

- 중의적 문장
- 지역 방언 (같은 단어, 다른 의미)

2 자연어처리 (Natural Language Processing)

- 텍스트 데이터의 변환

자연어의 속성 2: 불연속성

Color	Color Name	HEX Code	RGB Code
	Black	#000000	rgb(0, 0, 0)
	Maroon	#800000	rgb(128, 0, 0)
	DarkRed	#8b0000	rgb(139, 0, 0)
	Red	#ff0000	rgb(255, 0, 0)
	FireBrick	#b22222	rgb(178, 34, 34)
	Brown	#a52a2a	rgb(165, 42, 42)
	IndianRed	#cd5c5c	rgb(205, 92, 92)
	RosyBrown	#bc8f8f	rgb(188, 143, 143)
	LightCoral	#f08080	rgb(240, 128, 128)

VS



연속성

불연속성

2 자연어처리 (Natural Language Processing)

- 자연어 전처리

1. 정제(Cleaning) 단계

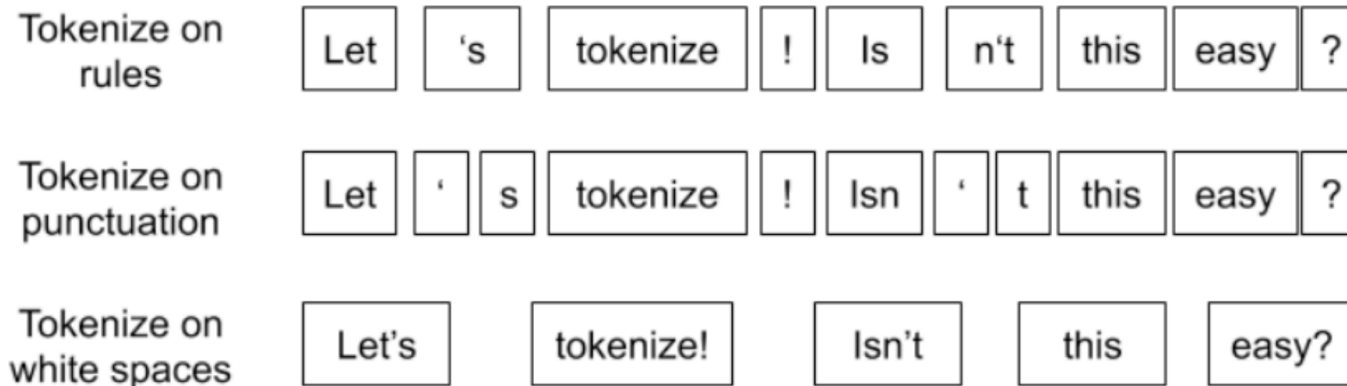
There is an apple
=
there is an apple

- 의미 분석에 필요 없는 변수 제거
- 코퍼스(Corpus): 정제 대상이 되는 전체 문장 데이터 집합

2 자연어처리 (Natural Language Processing)

- 자연어 전처리

2. 토큰화(Tokenize) 단계



Let's tokenize! Isn't this easy?

- 토큰(Token): 의미를 가지는 최소 단위
- 여러 기준에 따라 토큰화 가능

2 자연어처리 (Natural Language Processing)

- 자연어 전처리

2. 토큰화(Tokenize): nltk - word_tokenize()

```
from nltk.tokenize import word_tokenize
```

```
print(word_tokenize("Don't be fooled by the dark sounding name, Mr. Jone's Orphanage is as  
cheery as cheery goes for a pastry shop."))
```

```
['Do', 'n't', 'be', 'fooled', 'by', 'the', 'dark', 'sounding', 'name', ',', 'Mr.', 'Jone', 's', 'Orphanage',  
'is', 'as', 'cheery', 'as', 'cheery', 'goes', 'for', 'a', 'pastry', 'shop', '.']
```

- 가장 기본적인 토큰화 함수
- 띄어쓰기 & 구두점 기준

2 자연어처리 (Natural Language Processing)

- 자연어 전처리

2. 토큰화(Tokenize): nltk - WordPunctTokenizer()

```
from nltk.tokenize import WordPunctTokenizer
```

```
print(WordPunctTokenizer().tokenize("Don't be fooled by the dark sounding name, Mr. Jone's  
Orphanage is as cheery as cheery goes for a pastry shop."))
```

```
['Don', "'", 't', 'be', 'fooled', 'by', 'the', 'dark', 'sounding', 'name', ',', 'Mr', '.', 'Jone', "'", 's',  
'Orphanage', 'is', 'as', 'cheery', 'as', 'cheery', 'goes', 'for', 'a', 'pastry', 'shop', '.']
```

모든 구두점 기준

2 자연어처리 (Natural Language Processing)

- 자연어 전처리

2. 토큰화(Tokenize): Keras – text_to_word_sequence()

```
from tensorflow.keras.preprocessing.text import text_to_word_sequence
```

```
print(text_to_word_sequence("Don't be fooled by the dark sounding name, Mr. Jone's  
Orphanage is as cheery as cheery goes for a pastry shop."))
```

```
["don't", 'be', 'fooled', 'by', 'the', 'dark', 'sounding', 'name', 'mr', "jone's", 'orphanage', 'is', 'as',  
'cheery', 'as', 'cheery', 'goes', 'for', 'a', 'pastry', 'shop']
```

- 대문자 → 소문자
- 쉼표, 마침표 제거

2 자연어처리 (Natural Language Processing)

● 자연어 전처리

2. 토큰화(Tokenize): 한국어의 경우

```
from konlpy.tag import Hannanum  
  
nanum = Hannanum()  
  
nanum.morphs("드디어 교안 다 썼다!")  
  
['드디어', '교안', '다', '쓰', '었다', '!']
```

```
from konlpy.tag import Kkma  
  
kkma = Kkma()  
  
kkma.morphs("드디어 교안 다 썼다!")  
  
['드디어', '교안', '다', '쓰', '었', '다', '!']
```

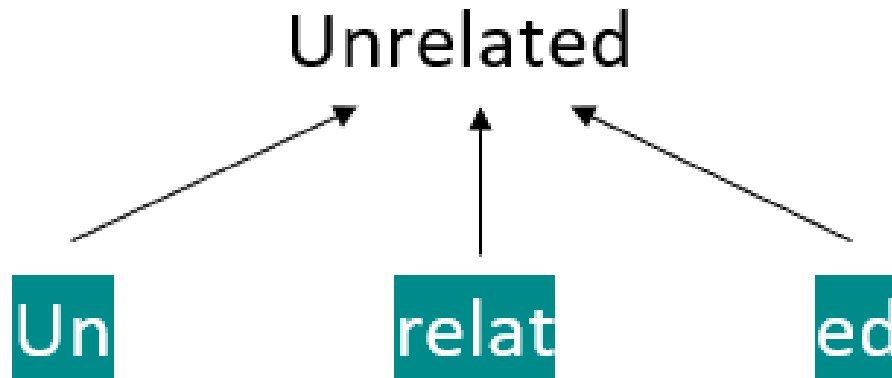
► konlpy

- 교착어: 어간 + 접사 Ex) '가다' = '가-' + '-쓰다'
- 형태소 단위

2 자연어처리 (Natural Language Processing)

- 자연어 전처리

3. Subword Segmentation

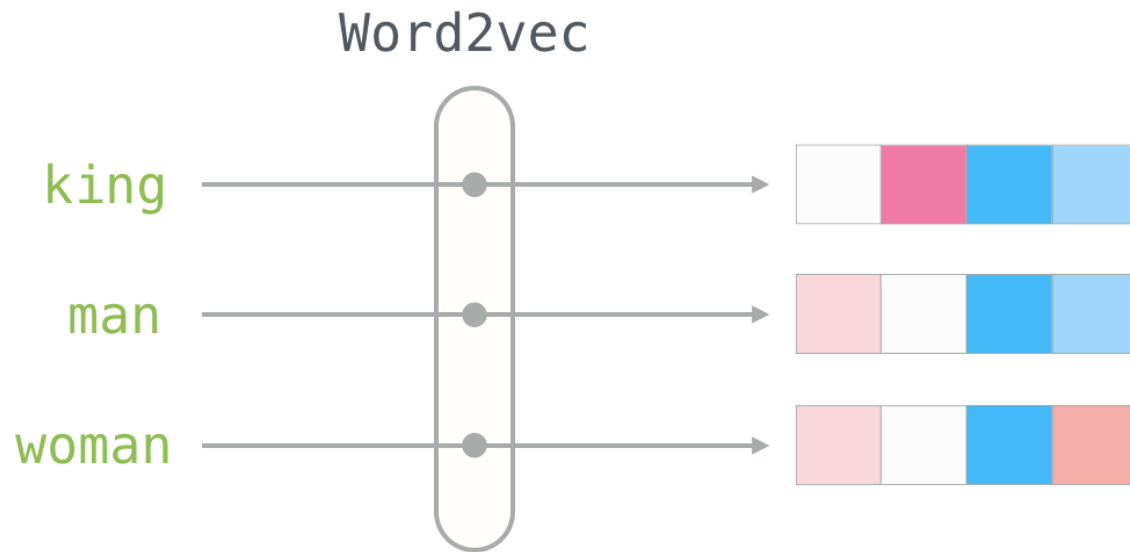


- 단어를 더 작은 단위로 분해
- OOV(Out-of-Vocabulary) 문제 해결
- 구글 Sentencepiece 라이브러리로 이용 가능

2 자연어처리 (Natural Language Processing)

● Word2Vec

Word2Vec이란?

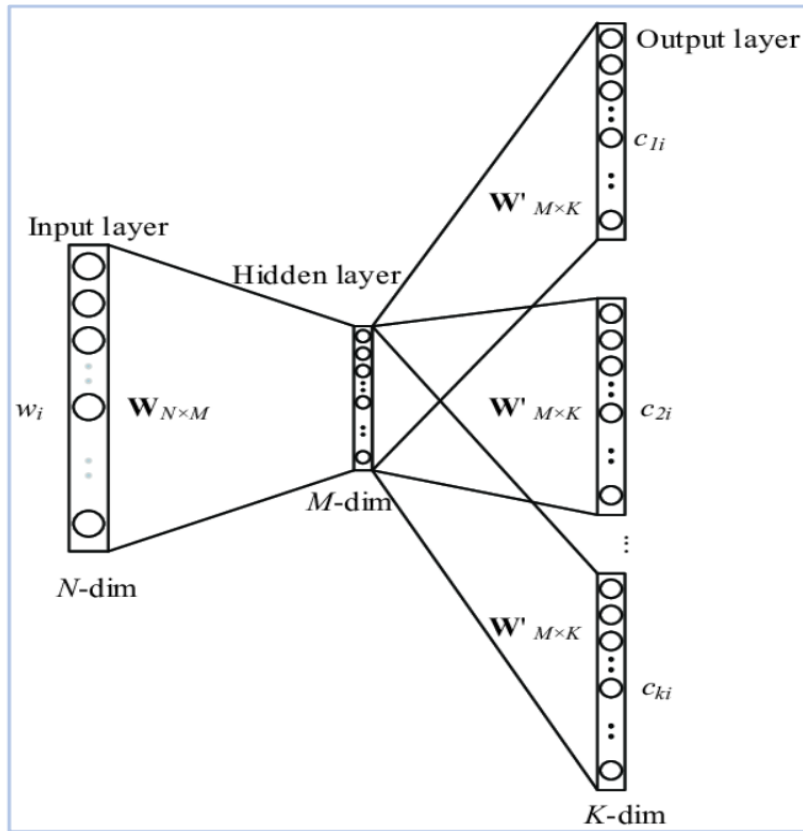


- 가정: 비슷한 위치에 등장하는 단어는 비슷한 의미를 가진다
- CBOW / Skip-gram

2 자연어처리 (Natural Language Processing)

Word2Vec

Skip-gram

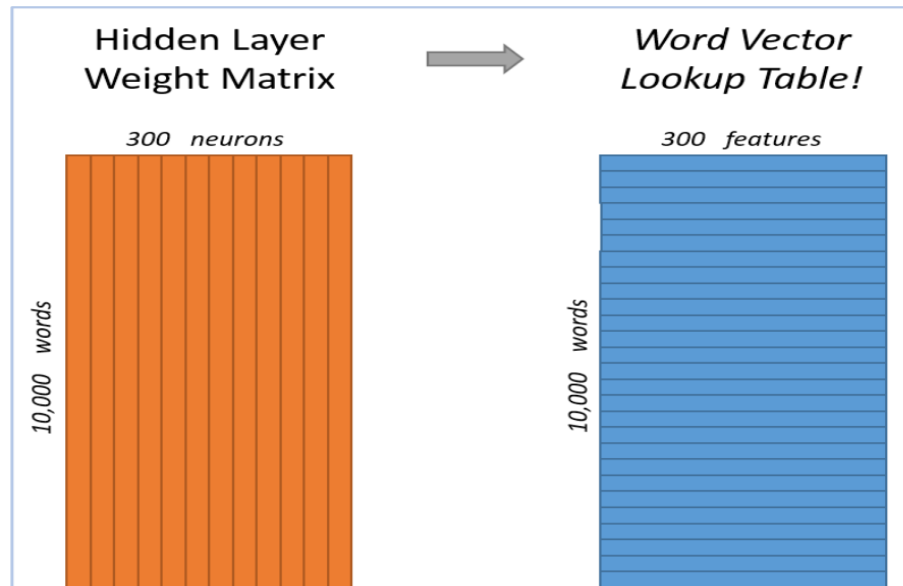


- 가중치 행렬: W, W'
- 벡터차원 N
= 문장에 서로 다른 단어 총 N 개
- 은닉층 차원 M

2 자연어처리 (Natural Language Processing)

- Word2Vec

Skip-gram - 임베딩(Embedding)



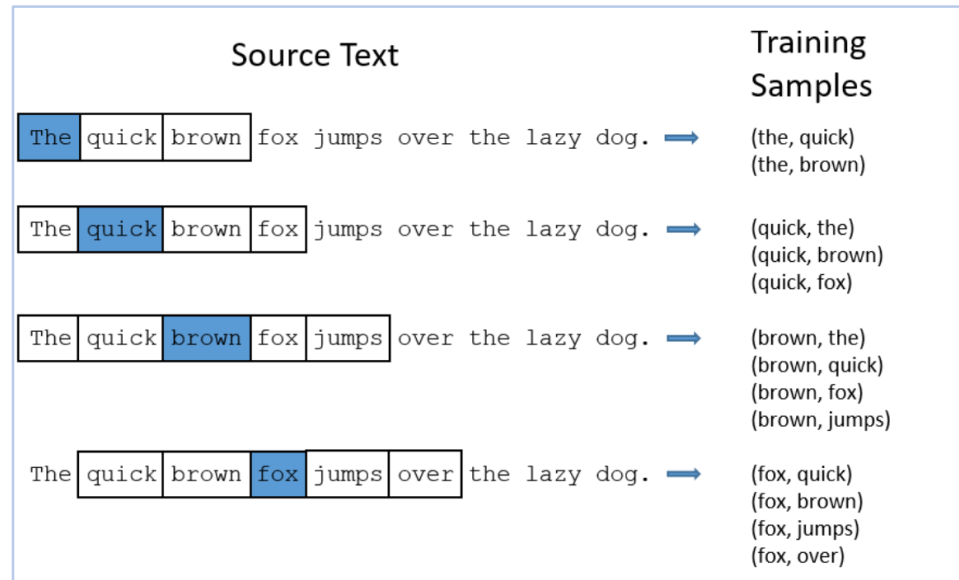
$N=10000, M=300$ 은닉층의 가중치
 \times
One-Hot Encoding 변환 벡터

(N,M) 은닉층 가중치 행렬의
행벡터

2 자연어처리 (Natural Language Processing)

● Word2Vec

Skip-gram



- 파란색 입력 → 주변 2n개의 단어 예측
- n: window size
- 문장 여러 번 학습 → 임베딩(Embedding) 벡터

2 자연어처리 (Natural Language Processing)

- Word2Vec

임베딩 벡터 표현의 장점

$$\overrightarrow{King} - \overrightarrow{Man} + \overrightarrow{Woman} = \overrightarrow{Queen}$$

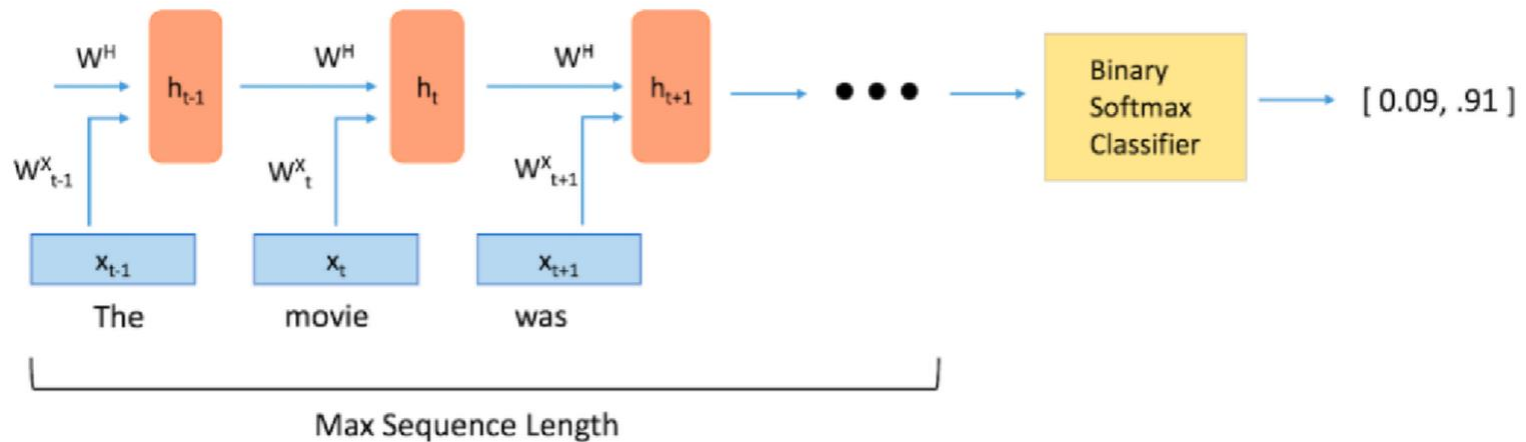
각 단어가 고차원 벡터공간의 한 점으로 표시

→ 벡터간 연산 가능

2 자연어처리 (Natural Language Processing)

- 활용 분야 - (1) Sentiment Analysis

감성 분석(Sentiment Analysis)

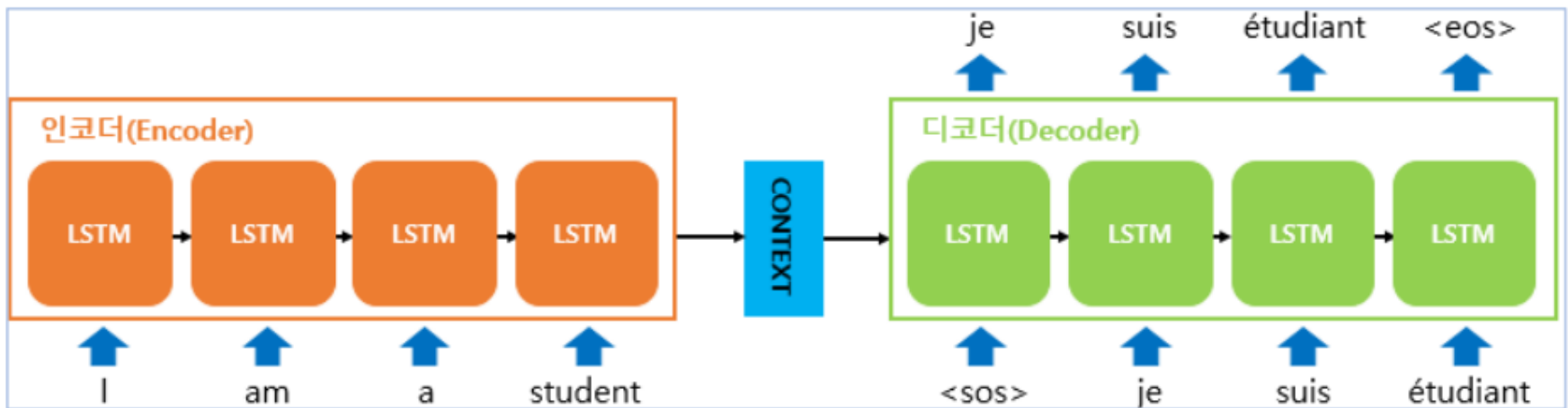


- 문장에 드러나는 감성을 긍정/부정 등으로 분류
- RNN 모델로 구현 가능
- Many-to-One

2 자연어처리 (Natural Language Processing)

● 활용 분야 - (2) Machine Translation

기계번역(Machine Translation)

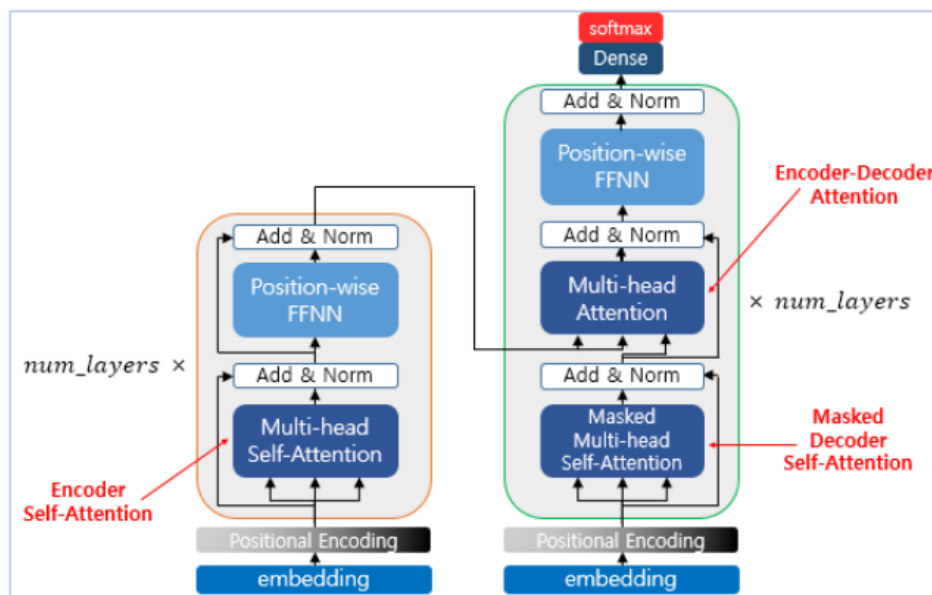


- 입력 문장을 다른 언어로 번역하여 출력하는 것을 목표로 하는 과제
- Encoder-Decoder가 대표적으로 사용되는 모델 구조
- Seq2Seq

2 자연어처리 (Natural Language Processing)

● 활용 분야 - (2) Machine Translation

Transformer



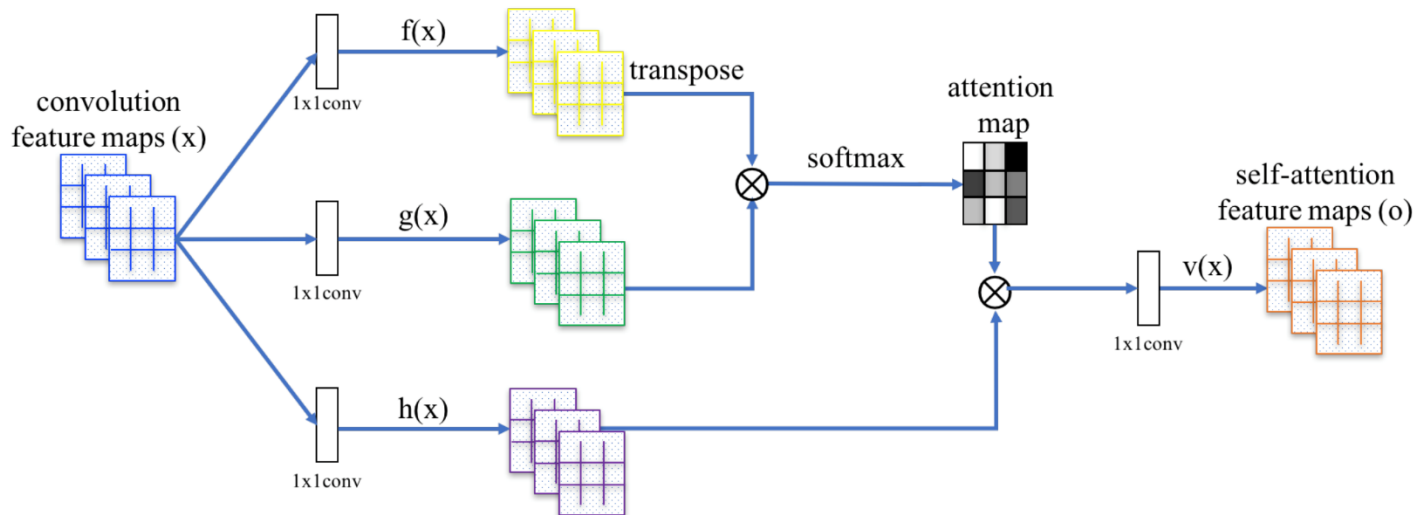
- RNN을 활용한 Encoder-Decoder 구조의 단점을 보완하기 위해 제안된 모델
- 2017 년 구글의 “Attention Is All You Need”라는 논문에 소개

- Encoder-Decoder 구조를 유지하면서 **RNN 을 사용하지 않음**
- RNN 의 순차적 입출력 구조를 배제하는 동시에 높은 정확도를 유지

2 자연어처리 (Natural Language Processing)

● 활용 분야 - (2) Machine Translation

Self-Attention

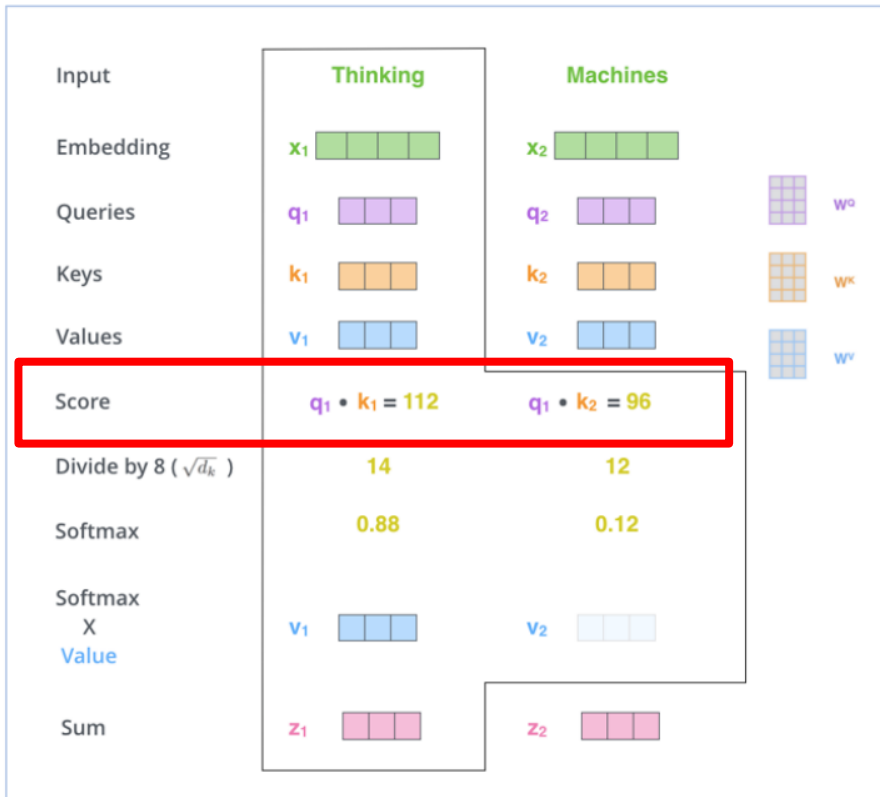


- Attention 기법과 아이디어는 동일
- 입력 문장 전체에 대해 한 번에 계산이 가능하도록 개선

2 자연어처리 (Natural Language Processing)

● 활용 분야 - (2) Machine Translation

Self-Attention 계산

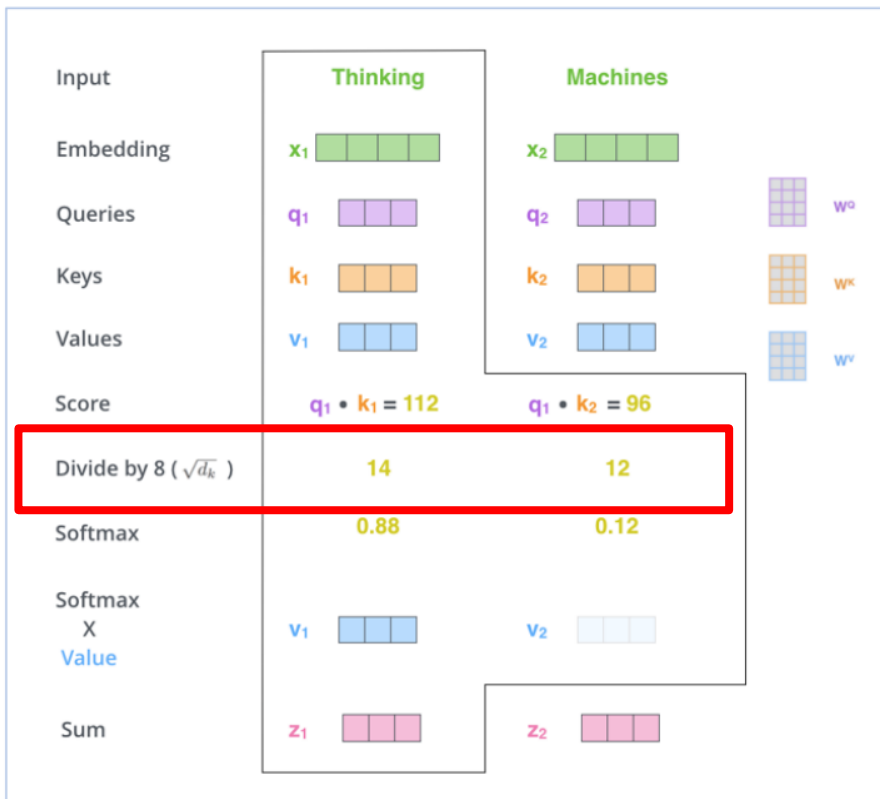


- Attention score 구하기
- Attention Score 를 벡터의 차원의 제곱근으로 나눠줌
- 이를 Softmax 함수에 통과시킨 후 Value 벡터와 가중합

2 자연어처리 (Natural Language Processing)

● 활용 분야 - (2) Machine Translation

Self-Attention 계산

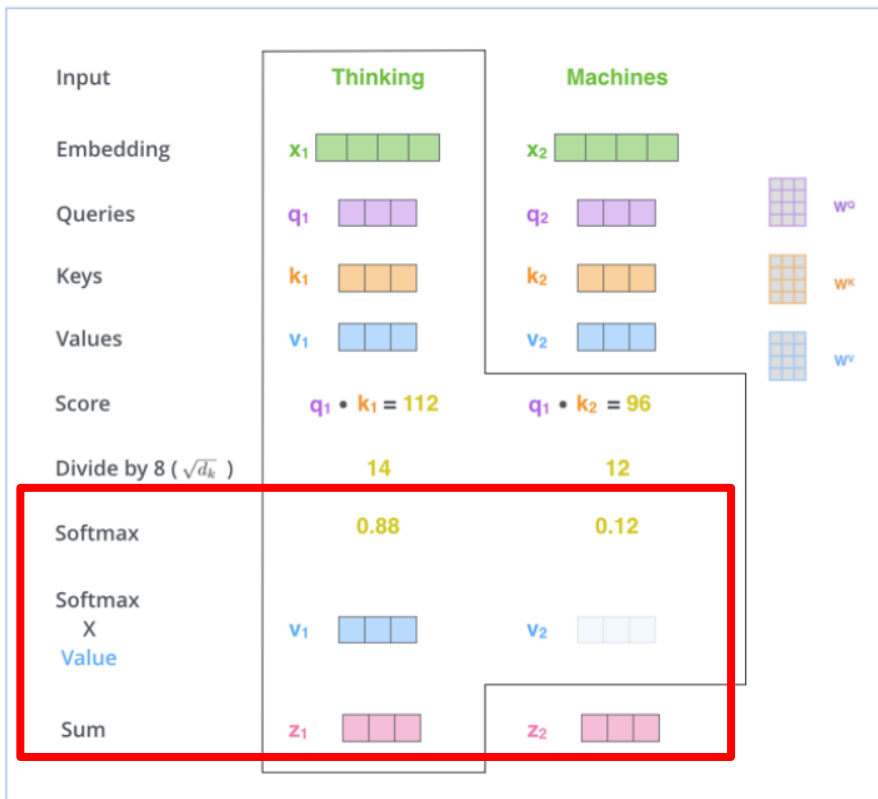


- Attention score 구하기
- Attention Score 를 벡터의 차원의 제곱근으로 나눠줌
- 이를 Softmax 함수에 통과시킨 후 Value 벡터와 가중합

2 자연어처리 (Natural Language Processing)

● 활용 분야 - (2) Machine Translation

Self-Attention 계산



- Attention score 구하기
- Attention Score 를 벡터의 차원의 제곱근으로 나눠줌
- 이를 Softmax 함수에 통과시킨 후 Value 벡터와 가중합

3

기 타

3 기타

- 기타 연구 분야 - (1) 음성

음성 딥러닝



Siri



Bixby

- 음성인식, 음성 합성
- 음의 높낮이, 소리의 크기 등의 특징을 잘 살려 행렬의 형태로 변환

3 기타

- 기타 연구 분야 - (1) 음성

활용 사례 - 음성인식



눈으로 보며 듣는 음성 기록

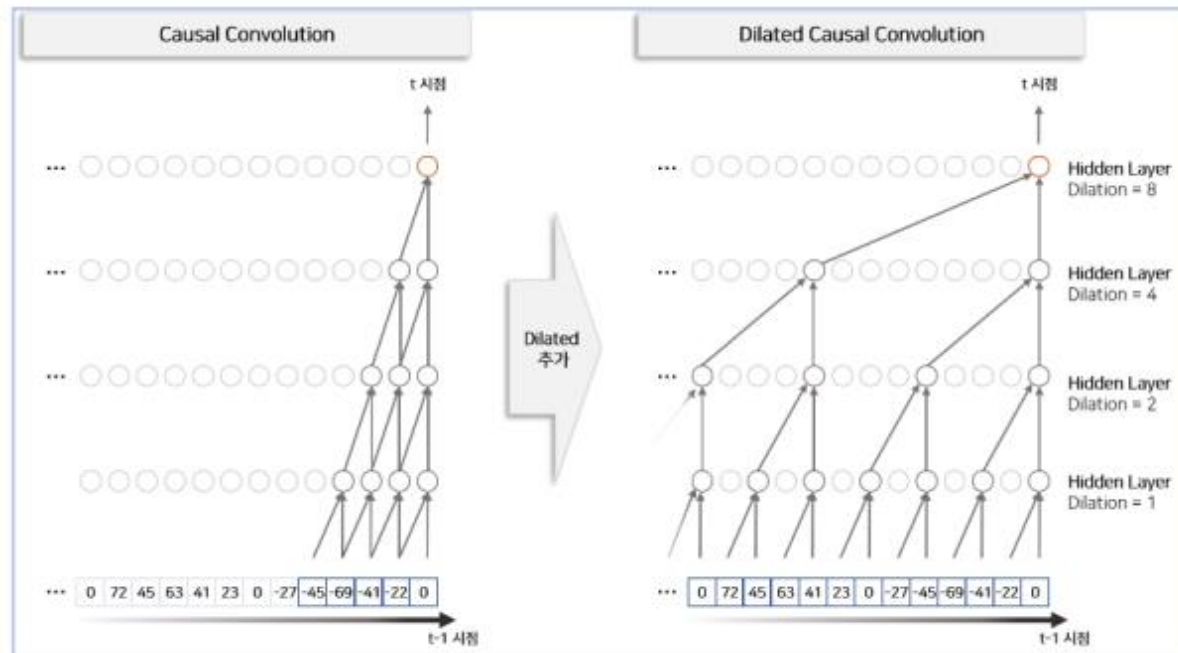
CLOVA Note^β

- 음성 파일, 혹은 진행되고 있는 음성을 입력하면 실시간으로 음성을 텍스트로 변환 (STT)
- 화자를 분별할 수 있도록 하는 화자 인식 기술 또한 포함

3 기타

- 기타 연구 분야 - (1) 음성

활용 사례 - 음성합성



- Residual Block 과 시간 순서를 고려해주는 특수한 Conv. Layer 를 적용
- 2017 년부터 구글 어시스턴트에 적용되어 TTS(Text-to-Speech) 서비스에 사용

3 기타

- 기타 연구 분야 - (1) 음성

활용 사례 - 음성합성

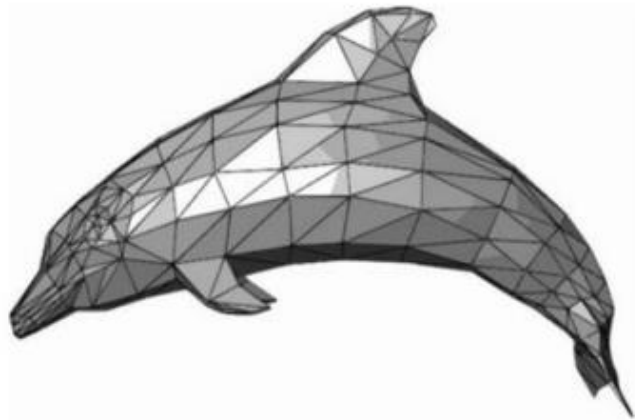


- 그 가수가 부르지 않은 노래를 그 가수가 부른 것처럼 구현
- 모션 인식 기술 또한 딥러닝으로 구현

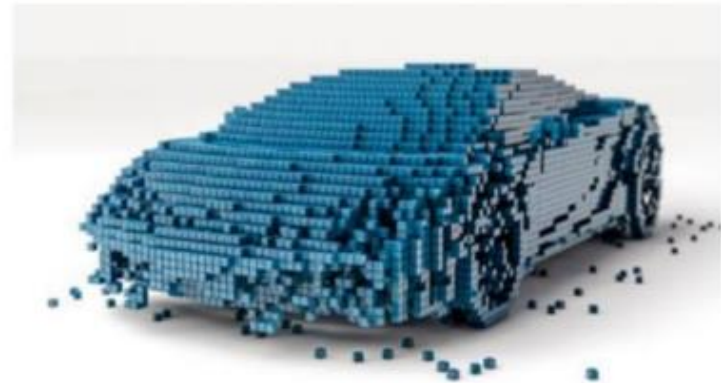
3 기타

● 기타 연구 분야 - (2) 3D

3D



Polygonal Mesh Data



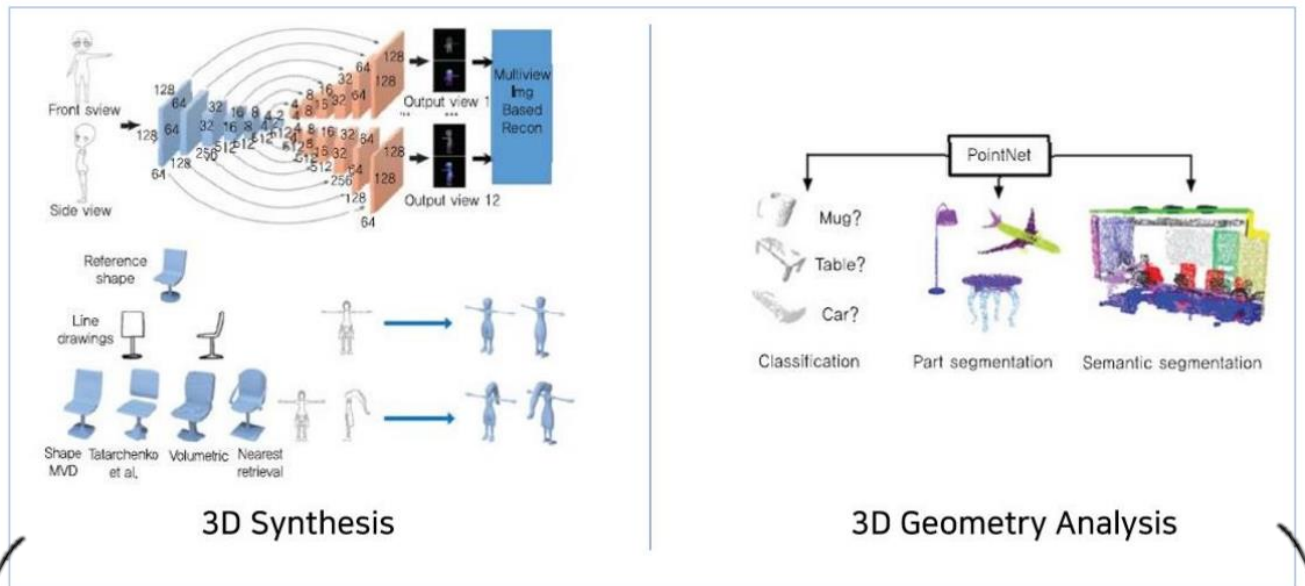
Volume Data

- 이미지, 텍스트, 음성 등의 분야에 비해 발전이 늦게 시작
- 최근 활발히 연구

3 기타

- 기타 연구 분야 - (2) 3D

3D - 활용 사례



3D 객체의 외형 추정

3D 객체에 대한 인식

3 기타

- 발전 중인 연구 분야 - (1) XAI

딥러닝의 단점



- 해석력의 부족 (Black Box)
- 결과에 대한 타당한 근거를 제시할 수 없음



신뢰도의 하락

3 기타

- 발전 중인 연구 분야 - (1) XAI

설명 가능한 인공지능(XAI)

: 인공지능의 판단을 사람이 이해 가능하도록 하는 연구

모델 특정 기법
(Model-Specific)

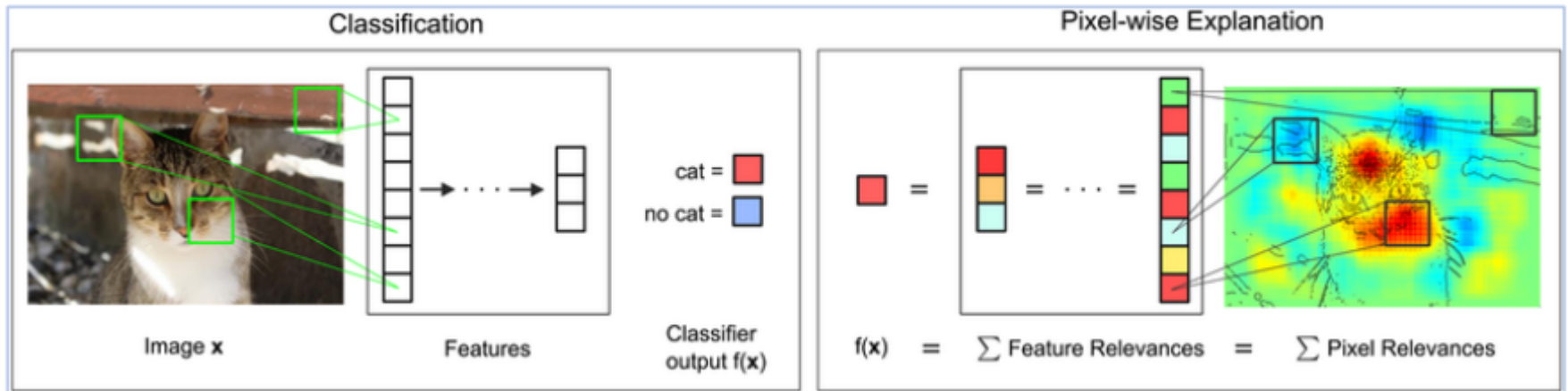
모델 불특정 기법
(Model-Agnostic)

3 기타

- 발전 중인 연구 분야 - (1) XAI

Layer-wise Relevance Propagation(LRP)

: 순전파의 역순으로 Layer 별 Feature 의 기여도를 분해하여 입력의 어떤 영역이 분류 과제에 많이 기여했는지 확인하는 방법

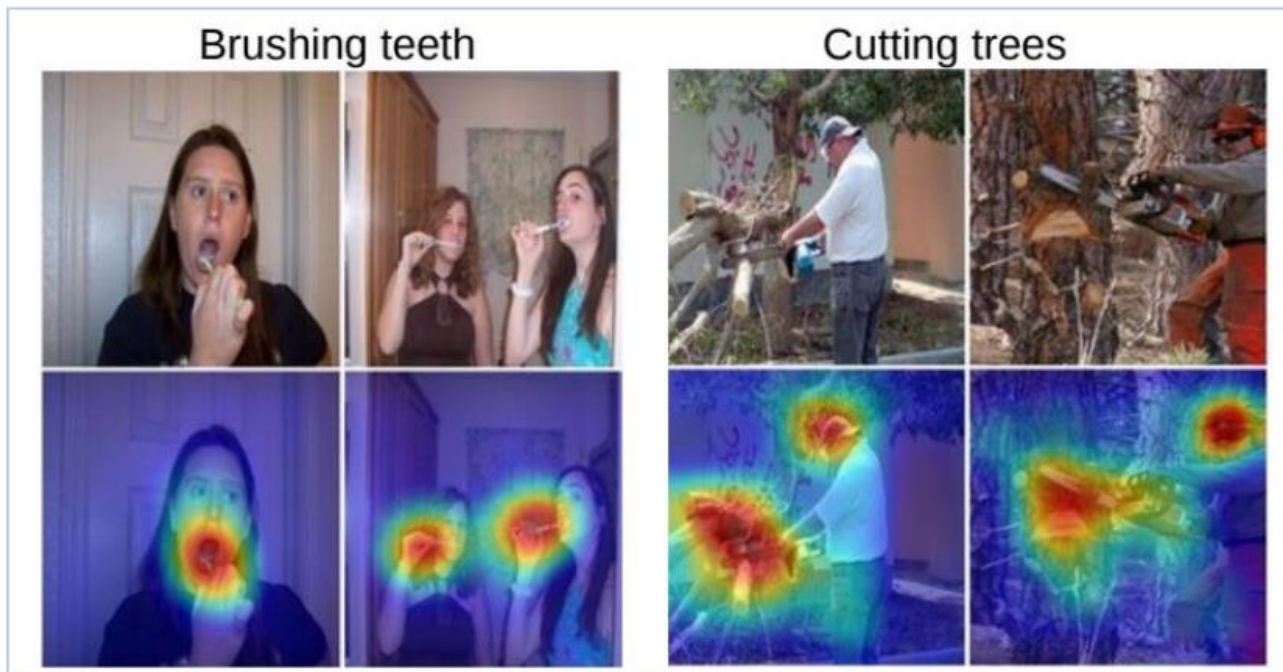


3 기타

- 발전 중인 연구 분야 - (1) XAI

Class Activation Map(CAM)

: 어떤 이미지가 특정 라벨로 예측되었을 때 Feature Map 에서 활성화되는 부분을 정량화하여 나타내는 기법

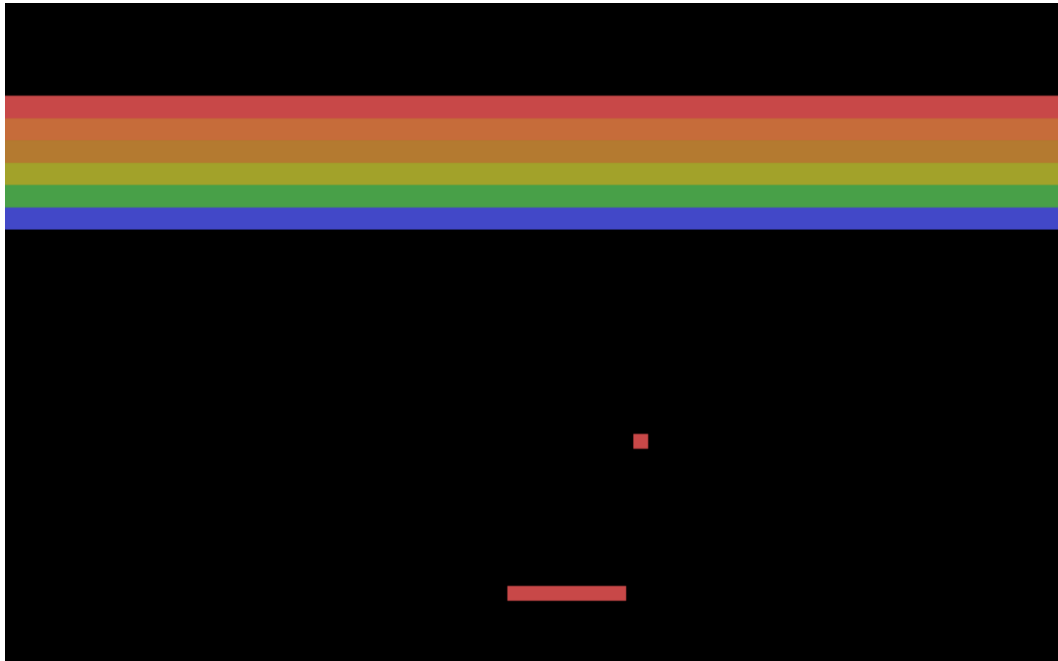


3 기타

- 발전 중인 연구 분야 - (2) 딥러닝 & 강화학습

DQN

- 딥러닝과 강화학습을 접목하는 데 처음으로 성공한 알고리즘
- CNN 모델을 Q-Network에 접목하여 Atari 게임을 클리어



3 기타

- 발전 중인 연구 분야 - (2) 딥러닝 & 강화학습

딥러닝과 강화학습

- 범용성 - 강화학습 알고리즘이 해결할 수 있는 과제의 범위가 넓어짐
- MADDPG- 스타크래프트 2 성공적으로 플레이

✓ 정답이 존재하지 않아도 컴퓨터가 스스로 해답을 찾을 수 있는 범위가 점점 넓어짐





THANK YOU

