

컴퓨터 비전 입문

Welcome to computer vision

컴퓨터 비전

사람과 인공지능 시스템의 인지 기능

- 사람은 다섯 가지 인지 기능 시각, 청각, 촉각, 후각, 미각 을 통해 주변 환경을 인식하고 경험함
- 인간을 모사해서 만든 인공지능 시스템은 애플리케이션 인공지능스피커, 자율주행자동차 등 에 따라 시스템이 처한 세계를 가장 잘 이해할 수 있는 감지 장치 눈(카메라), 귀(마이크), 피부(인공 피부), 코(인공 코), 혀(인공 혀)의 역할을 하는 센서 를 선택함

인공지능 시스템(AI system)의 핵심 개념

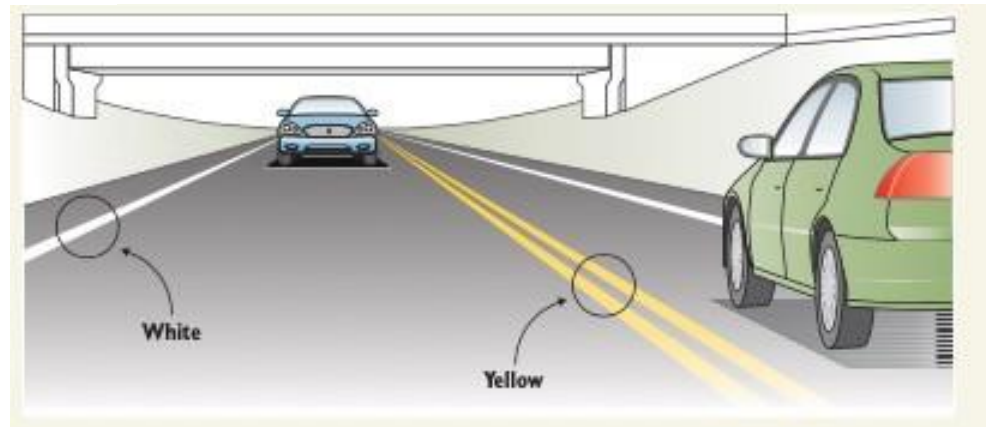
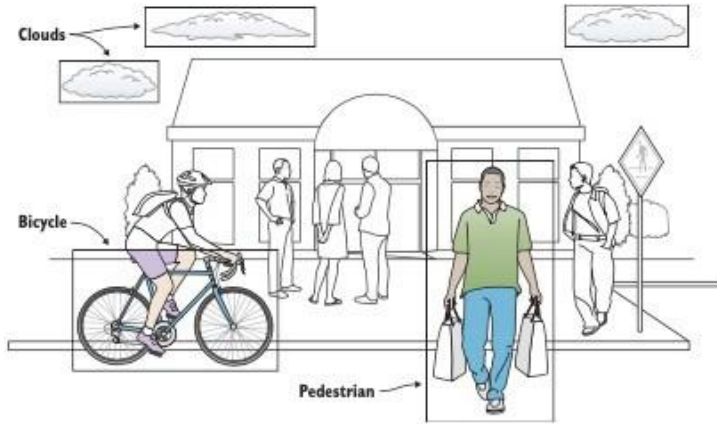
- 자신의 주변 환경을 인식(perception)하고,
그 인식을 기반으로 행동(action)을 취하는 것

컴퓨터 비전(computer vision)의 개념

- 시각은 세계를 이해하기 위한 한가지 수단으로 <컴퓨터 비전>은 시각적 인지(visual perception)를 다루는 분야
- 특히, 영상(images)과 비디오(videos)를 이용해 세계를 인식(perceiving)하고 이해(understanding)하는 분야

컴퓨터 비전

시각적 인지 (Visual Perception) 란 무엇인가



컴퓨터 비전

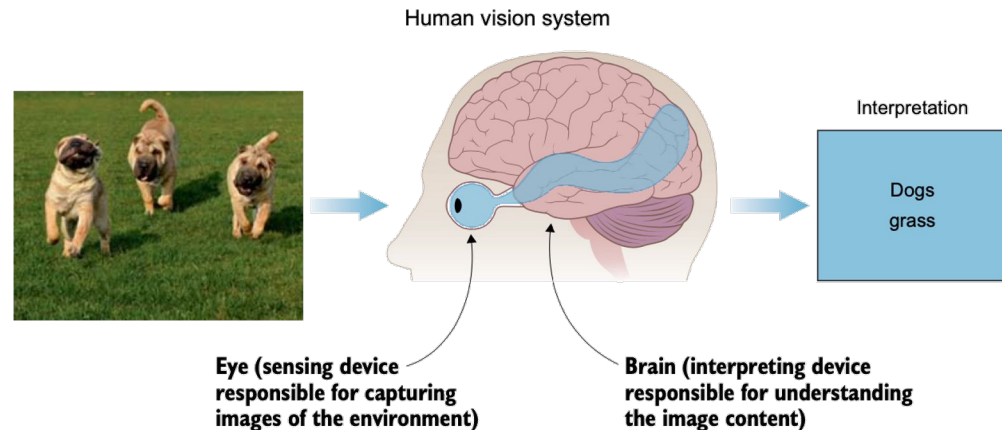
비전 시스템 (Vision Systems)

컴퓨터 비전 시스템의 정의

- 전통적으로 이미지 처리 기법만 컴퓨터 비전 시스템의 전부라 생각
- 최근에는 이미지 처리 기법 뿐만 아니라 **이미지를 통해 어떤 일이 일어났는지 이해하는 것**을 포함하여 컴퓨터 비전 시스템이라 정의함

사람의 시각 시스템(Human Vision Systems)

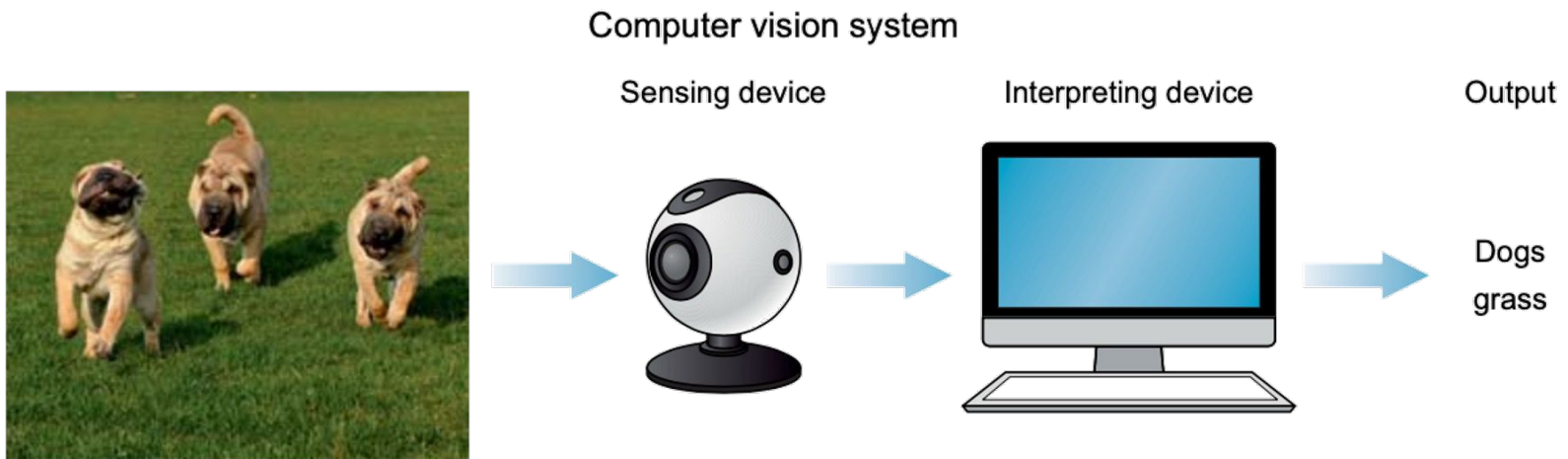
- 시각 시스템은 이미지를 받아들이는 <시각 센서인 눈>과 <이미지를 해석하는 뇌>로 구성됨
- 시스템의 출력은 이미지에서 얻은 데이터를 기반으로 한 이미지 요소에 대한 예측 결과임



컴퓨터 비전

비전 시스템 (Vision Systems)

인공지능의 비전 시스템(AI Vision Systems)



컴퓨터 비전

감지 장치 (Sensing devices)

비전 시스템은 특정 태스크(specific task)를 수행하도록 설계

시스템의 목적에 맞는 감지 장치를 선택하는 것이 중요

- 감지 장치로의 예: 카메라, 레이다, 엑스레이, CT스캔, 라이다 등
- 비전 시스템은 <단일 감지 장치> 혹은 <이들의 조합>을 통해 주변 환경 정보를 얻음

자율주행자동차

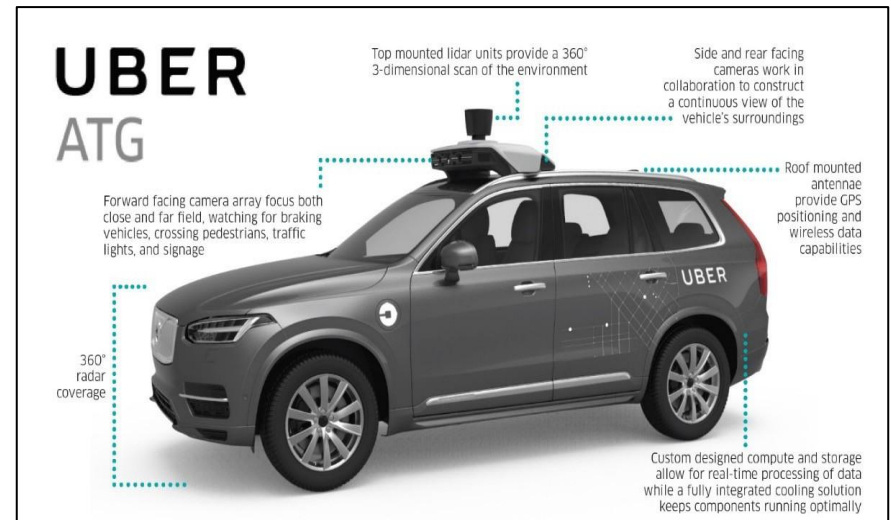
- 라이다 Lidar / 카메라 / 레이다 Radar

의학 진단 영역

- 엑스레이나 방사선 단층 촬영

농업용 시각 시스템

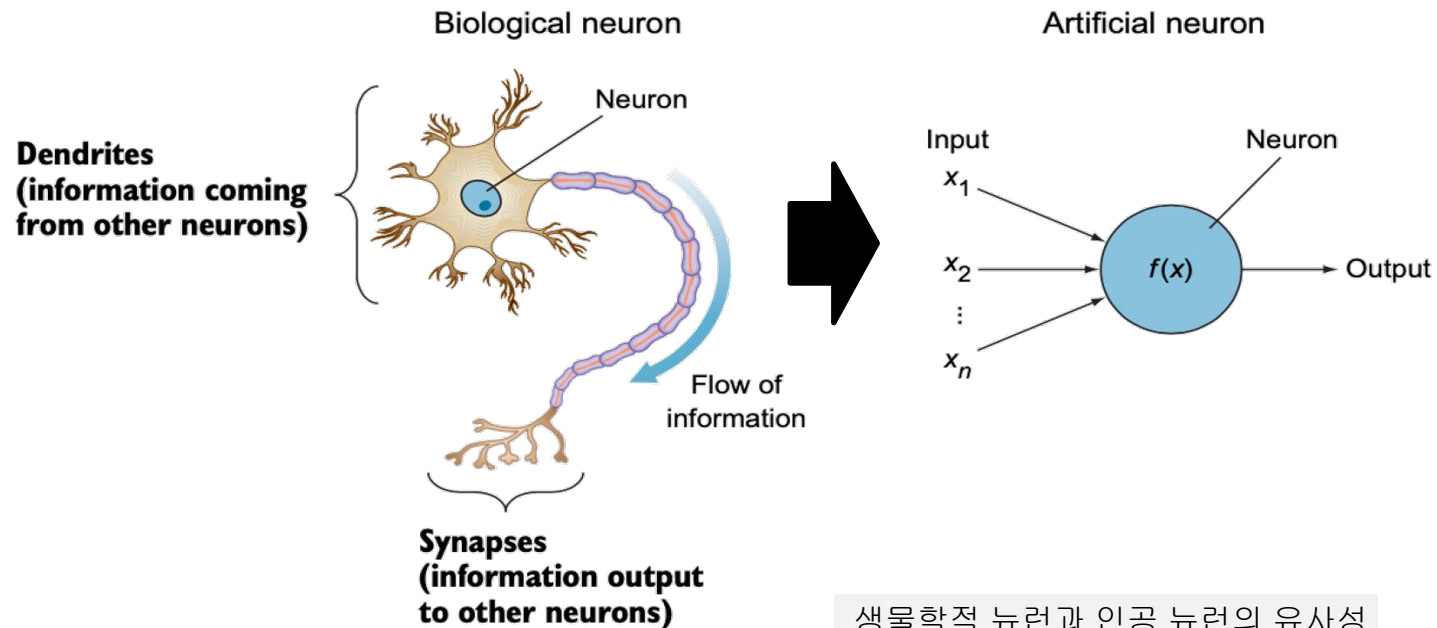
- 또 다른 유형의 레이다



컴퓨터 비전

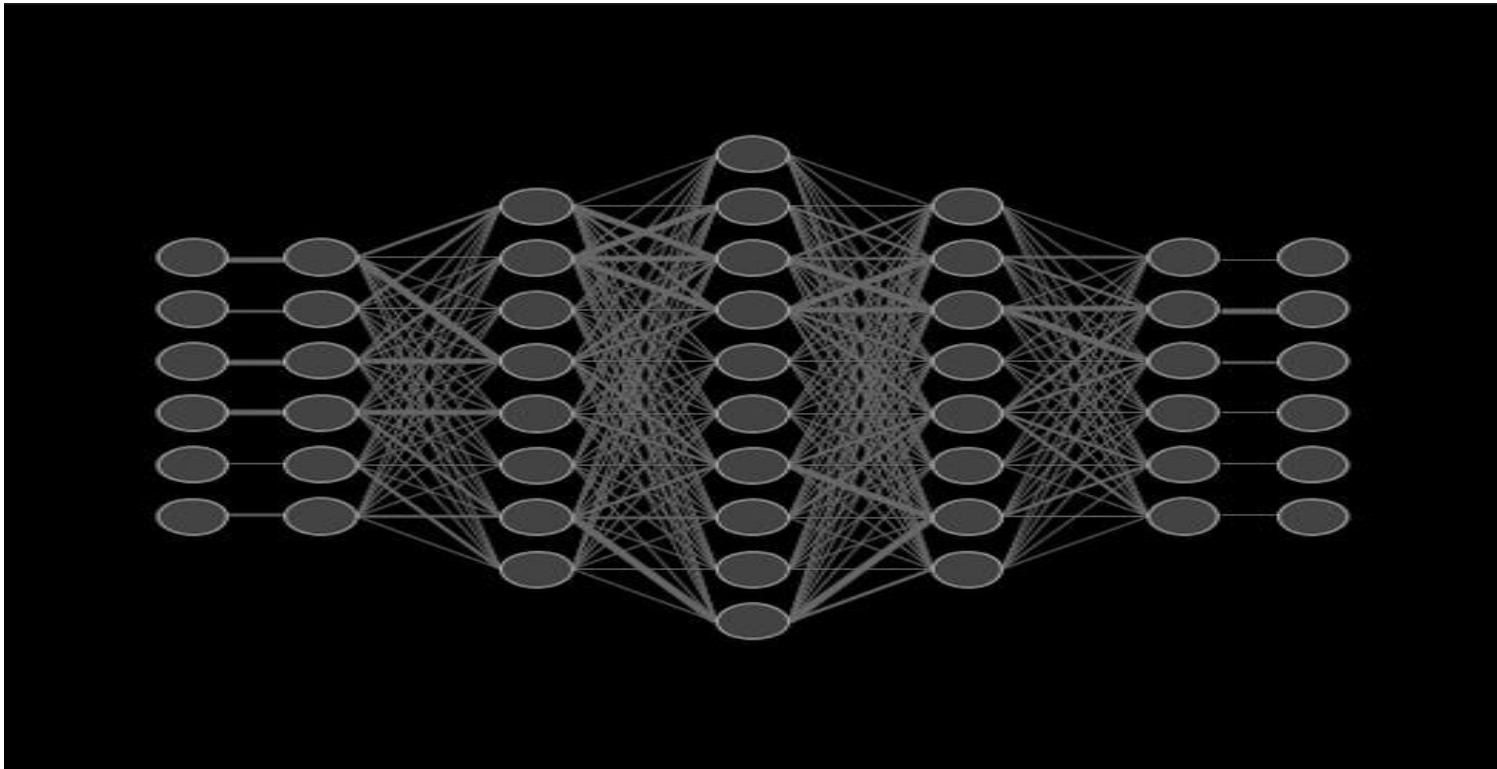
해석 장치

- 일반적으로 <컴퓨터 비전 알고리즘>이 해석 장치의 역할을 담당함
- 최근 우리의 뇌로부터 영감을 받아 뇌의 동작을 분석하는 방법으로 인공 뇌를 만들려는 시도가 이루어 졌고,
그 결과가 바로 **인공 신경망 (ANN: Artificial Neural Network)**에 해당됨



컴퓨터 비전

해석 장치



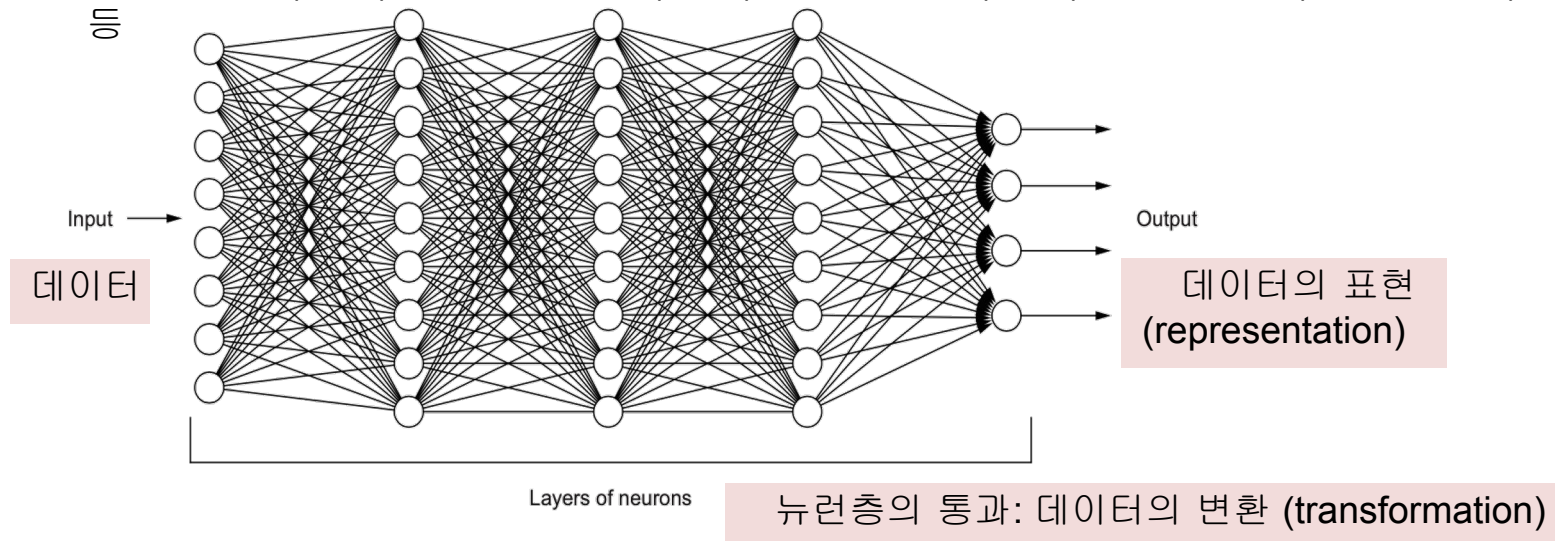
1.1 컴퓨터 비전

해석 장치

- 여러 층을 가진 신경망을 만드는 기법을 딥러닝(deep learning)
- 딥러닝 기법은 뉴런으로 구성된 층에 데이터를 통과시켜 데이터를 여러 번 변환하는 방법으로 데이터의 표현(representation)을 학습함

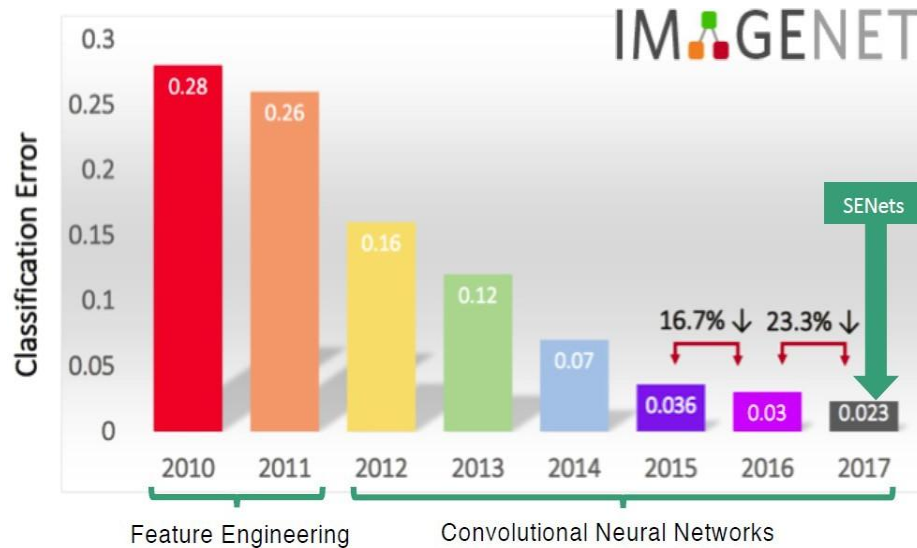
딥러닝 구조의 종류

- 인공 신경망(ANN), 합성곱 신경망(CNN), 순환 신경망(RNN), 트랜스포머(Transformer) 등



1.1 컴퓨터 비전

머신러닝의 성능이 사람을 능가할 수 있을까



영상 인식분야에서 사람의 능력을 압도한 딥러닝 기법



AI 프로그램 미드저니가 생성한 “스페이스 오페라 극장”

컴퓨터 비전 응용 분야

- 컴퓨터 비전 응용 분야의 예
 - 이미지 분류
 - 사물 인식/물체 검출
 - 화풍 모방하기
 - 새로운 이미지 창조하기
 - 안면 인식
 - 이미지 검색
 - 기타

1.2 컴퓨터 비전 응용 - 이미지 분류

이미지 분류(image classification)는

미리 정의한 유한한 수의 레이블을 이미지에 부여하는 태스크

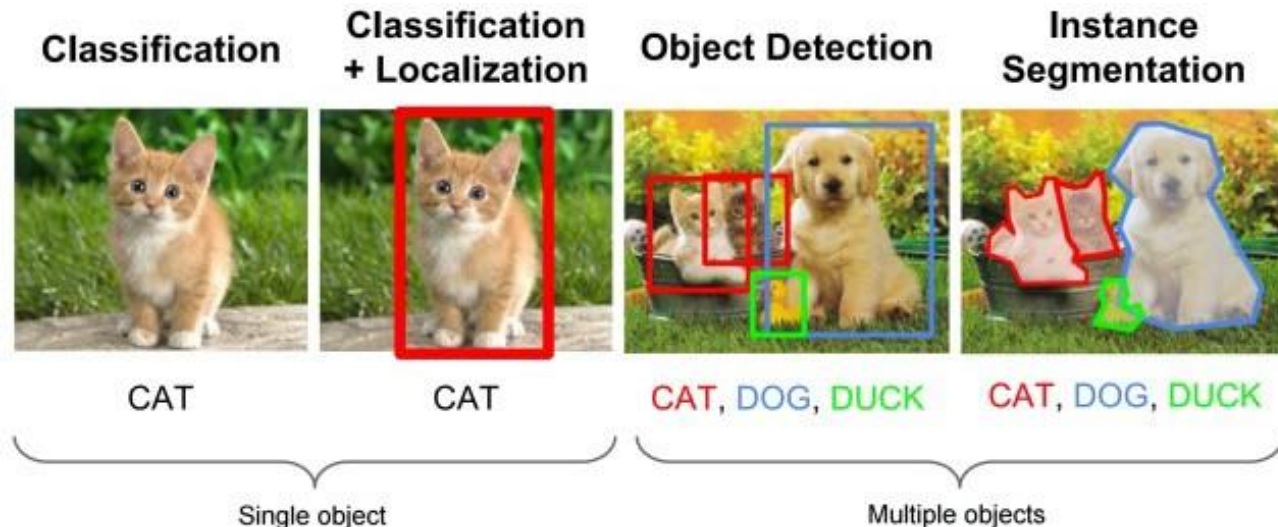


CIFAR10 데이터셋 예시

컴퓨터 비전 응용 - 사물 인식/물체 검출

분류(classification)는 단일 물체가 찍힌 이미지를 인식(=분류)하는 태스크

- 분류 및 위치 인식(classification & **localization**)은 단일 물체를 인식하고 위치까지 파악하는 태스크
- 물체 검출(object detection)은 이미지 안에서 다중 물체를 인식하고 위치까지 파악하는 태스크



컴퓨터 비전 응용

화풍 모방하기

- 신경망을 이용한 화풍 모방(Neural Style Transfer) 기법은 원본 이미지에 특정 스타일을 입혀 새로운 형태의 이미지를 생성하는 것

<별이 빛나는 밤> 화풍을 다른 사진에 적용해서 원래 그림과 비슷한 새로운 그림을 만드는 화풍 모양

Original image



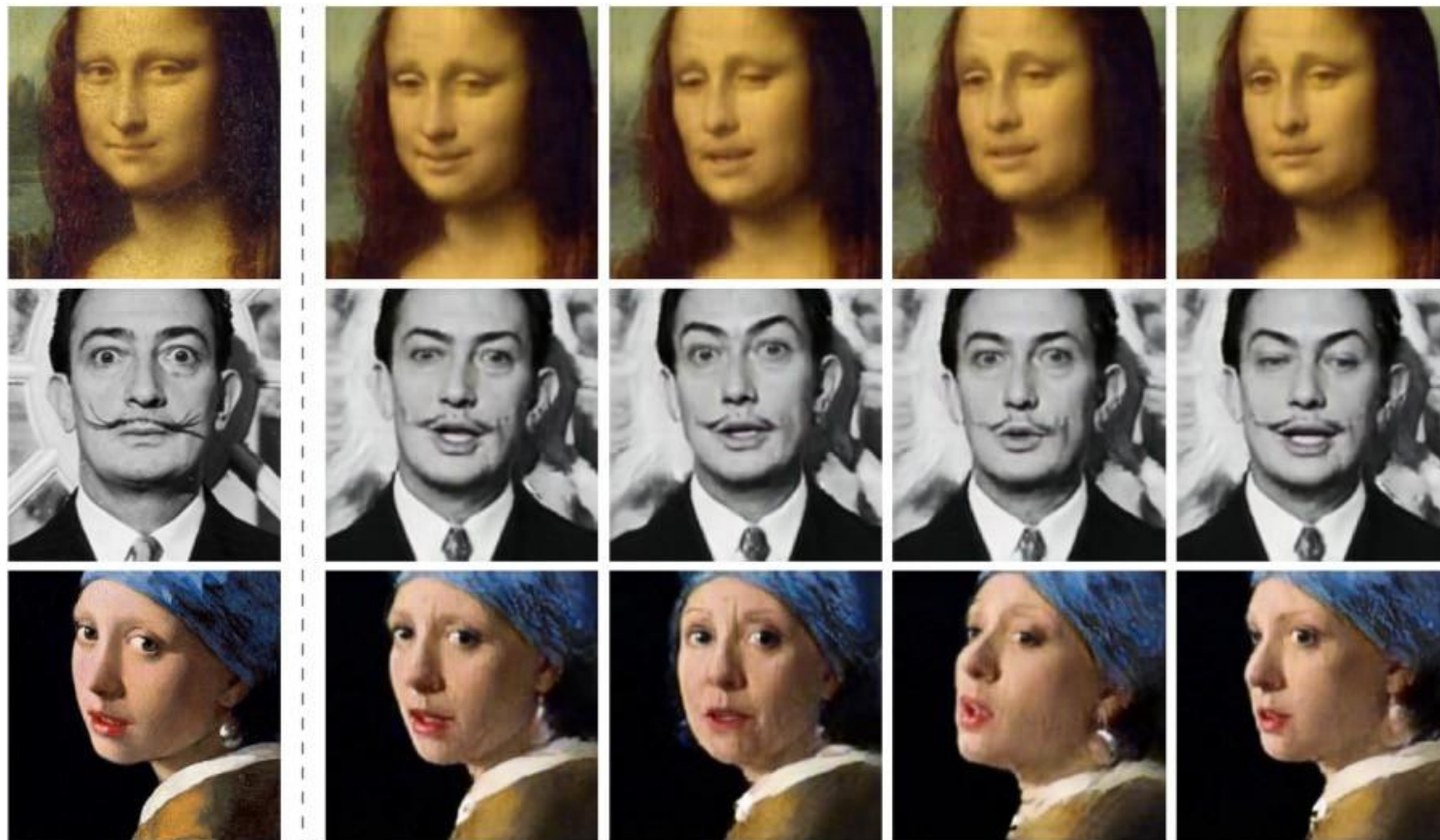
Style



Generated art



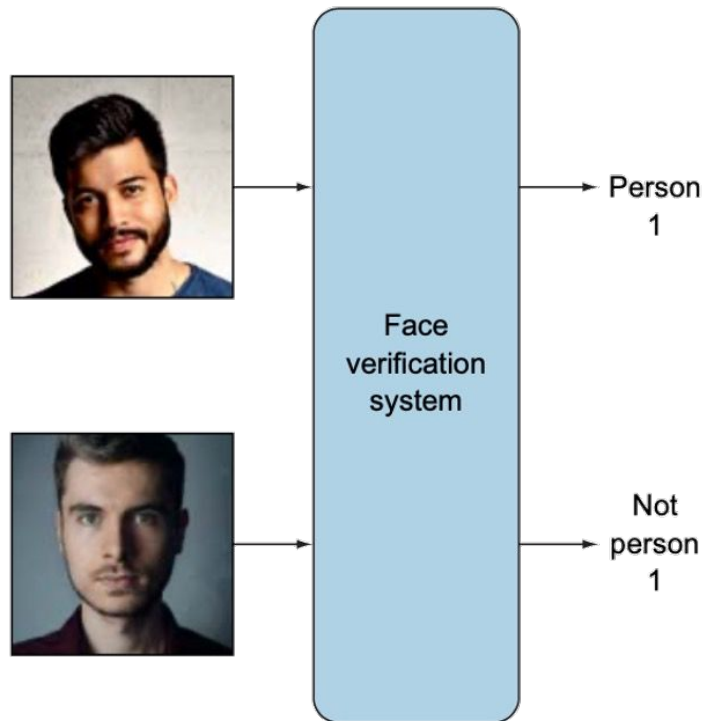
컴퓨터 비전 응용



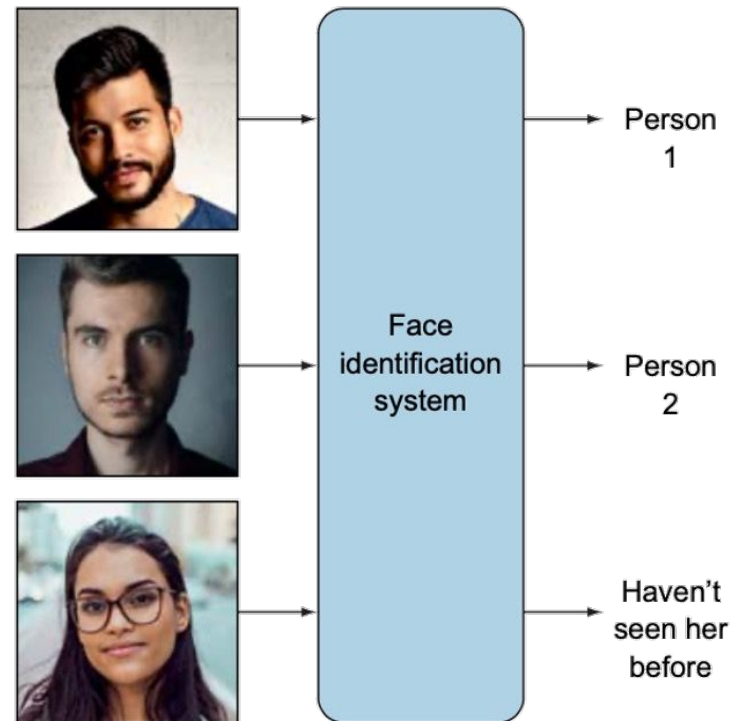
[Deepfake Videos: GAN Synthesizes a Video From a Single Photo](#)

컴퓨터 비전 응용

- 안면 대조 face verification



- 안면 식별 face identification

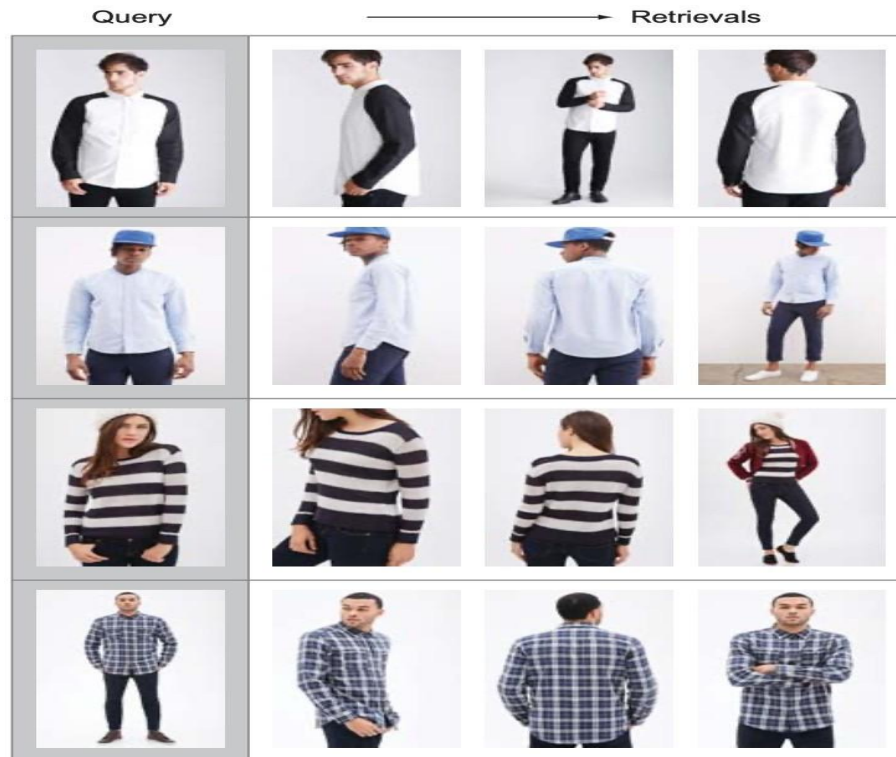


컴퓨터 비전 응용 분야

이미지 검색

- 이미지 검색 image retrieval 은 질의 영상과 유사한 영상을 찾는 기술
- 전자상거래 사이트에서 이전에 선택한 상품과 생김새가 비슷한 상품을 찾는데 활용

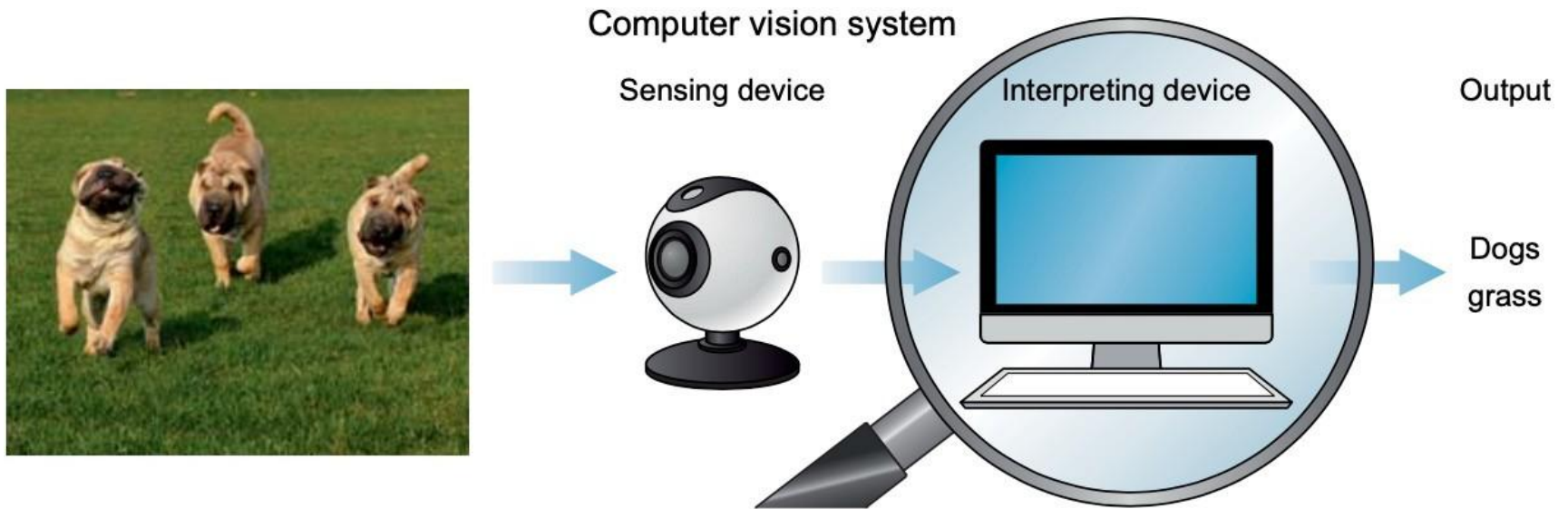
의류상품 검색의 예



컴퓨터 비전 파이프라인 전체 처리 과정

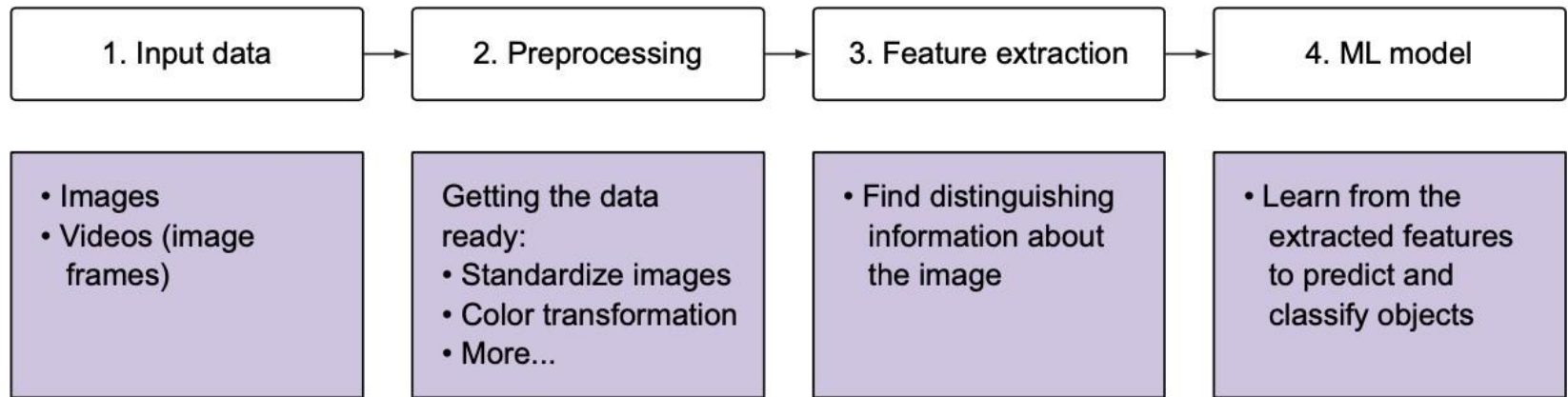
컴퓨터 비전 시스템

- 비전 시스템은 감지 장치와 해석 장치로 구성됨

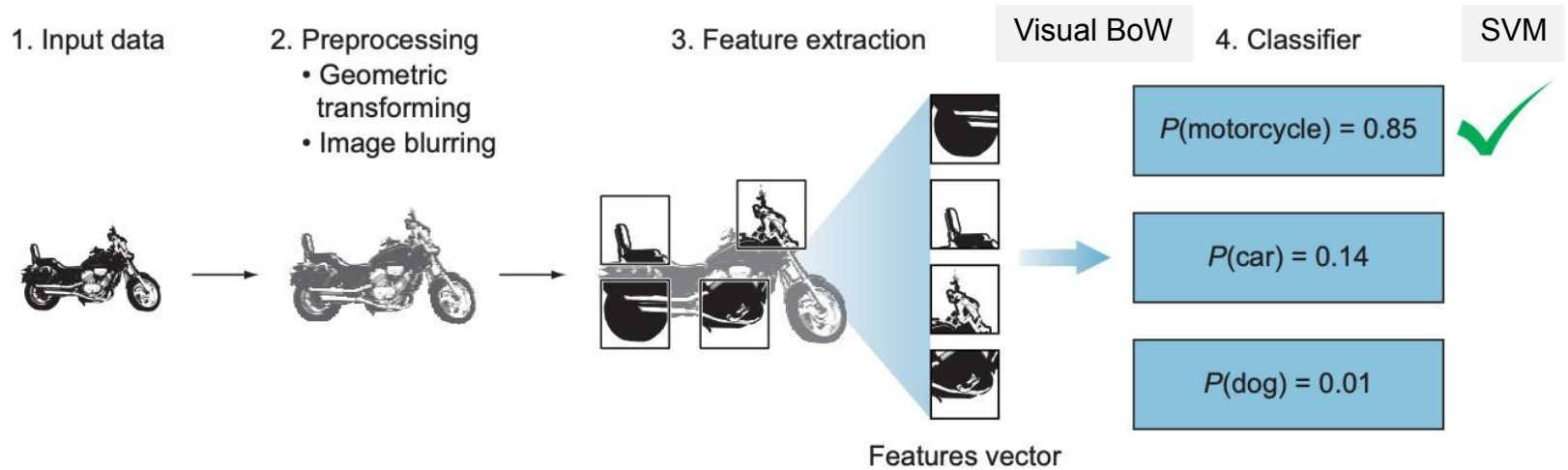


컴퓨터 비전 파이프라인 전체 처리 과정

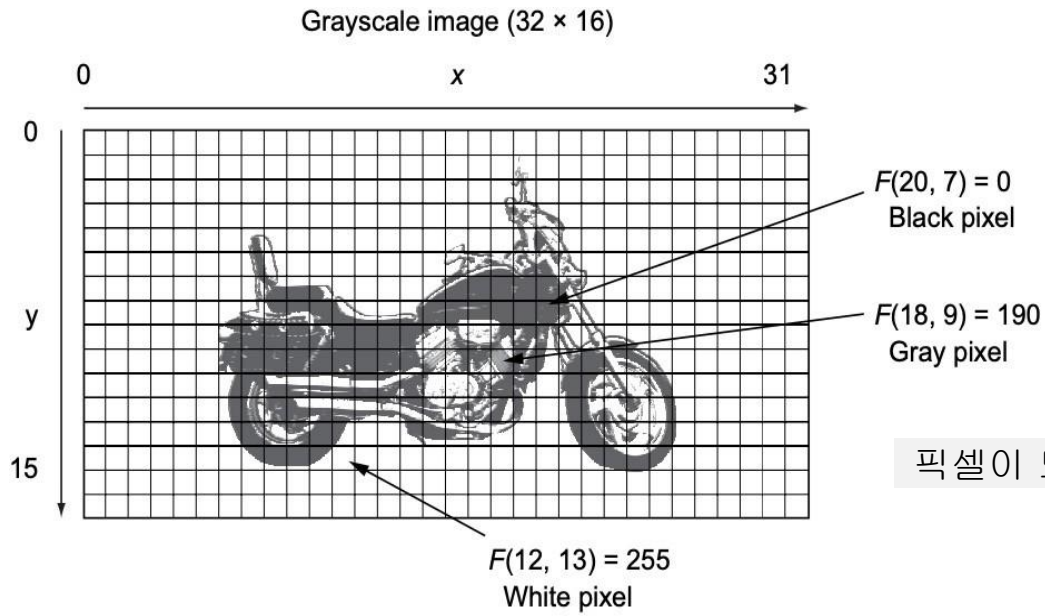
컴퓨터비전 파이프라인 : 데이터를 입력 받고, 처리하고, 정보를 추출한 다음 머신러닝 모델에 전달해 학습



컴퓨터 비전 파이프라인 전체 처리 과정



이미지 입력



픽셀이 모여 이미지를 만듦

이미지 입력

함수로 나타낸 이미지

회색조 이미지의 픽셀이 숫자 하나로 표현된 것과 달리 컬러 이미지는 픽셀을 세 가지

색의 강도로 나타냄

<RGB 색상 체계>에서는 빨간색의 강도, 녹색의 강도, 파란색의 강도 이렇게 3개의 숫자를 이용하여 픽셀을 나타냄

<HSV>나 <Lab> 같은 다른 색상 체계도 있지만 픽셀 값을 표현하는 방법은 모두 동일

RGB 색상 체계의 컬러 이미지 => $F(x, y) = [\text{red}(x, y), \text{green}(x, y), \text{blue}(x, y)]$

컬러 이미지의 구조

픽셀 (Pixel) : 이미지의 가장 작은 단위

- 컬러 이미지에서 각 픽셀은 여러 색상 값의 조합으로 표현

RGB 색상 체계

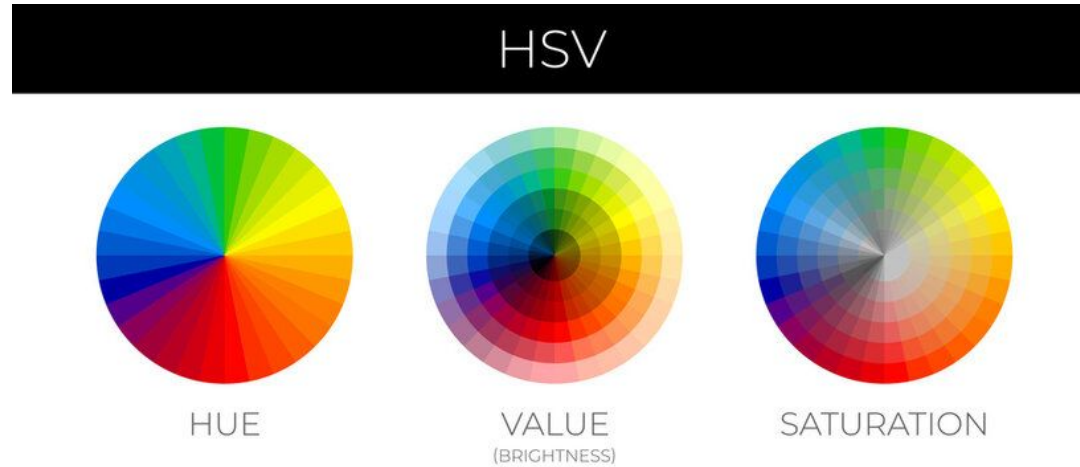
- 가장 흔히 사용되는 컬러 표현 방식입니다.
- 각 픽셀은 세 가지 색상의 강도로 표현됩니다:
 - R (Red): 빨간색의 강도
 - G (Green): 녹색의 강도
 - B (Blue): 파란색의 강도
- 각 색상 강도는 보통 0-255 사이의 값을 가집니다 (8비트 표현)
- 여기서 F는 이미지 함수, R, G, B는 각각 빨강, 초록, 파랑의 강도 값입니다.

예시

순수한 빨간색: [255, 0, 0]

- 순수한 녹색: [0, 255, 0]
- 순수한 파란색: [0, 0, 255]
- 흰색: [255, 255, 255]
- 검은색: [0, 0, 0]

컬러 이미지의 구조



1.

이미지 입력

함수로 나타낸 이미지

- 이미지를 함수로 나타내면 이미지 처리에 유용한 점이 많음
- 이미지를 함수 $F(x, y)$ 로 생각할 수 있고,

수학적 연산을 통해 새로운 이미지 $G(x, y)$ 로 변환할 수 있음

이미지 변환 함수의 예시

Application	Transformation
Darken the image.	$G(x, y) = 0.5 * F(x, y)$
Brighten the image.	$G(x, y) = 2 * F(x, y)$
Move an object down 150 pixels.	$G(x, y) = F(x, y + 150)$
Remove the gray in an image to transform the image into black and white.	$G(x, y) = \{ 0 \text{ if } F(x, y) < 130, 255 \text{ otherwise } \}$

이미지 입력

컴퓨터가 보는 이미지

What we see



What computers see

(HxW) = 24x24

08	02	22	97	38	15	00	40	00	75	04	05	07	78	52	12	50	77	91	08
49	49	99	40	17	81	18	57	60	87	17	40	98	43	69	46	04	56	62	00
81	49	31	73	55	79	14	29	93	71	40	67	53	99	30	03	49	13	36	65
52	90	95	23	04	60	11	42	69	24	48	56	01	32	54	71	37	02	34	91
22	31	14	71	51	67	43	59	41	92	34	54	22	40	40	28	44	33	13	80
24	47	32	60	99	03	45	02	44	75	33	53	78	36	64	20	35	09	12	80
32	98	81	28	64	23	67	10	26	38	40	67	59	54	70	66	18	38	64	70
47	24	20	68	02	62	12	20	95	63	94	39	63	04	49	91	44	49	94	21
24	55	58	05	66	73	99	26	97	17	78	78	94	83	14	88	34	89	63	72
21	36	23	09	75	00	74	44	20	45	35	14	00	41	33	97	34	31	33	95
78	17	53	28	22	75	31	67	15	94	03	80	04	42	16	14	09	53	56	92
16	39	05	42	96	35	31	47	55	58	88	24	00	17	54	24	34	29	85	57
84	56	00	48	35	71	89	07	05	44	44	37	44	40	21	58	51	54	17	58
19	80	61	68	05	94	47	49	28	73	92	13	86	52	17	77	04	09	55	40
04	52	08	83	97	35	99	14	07	97	57	32	16	26	26	79	33	27	98	44
04	36	68	81	57	62	20	72	03	16	33	67	46	55	12	32	43	93	53	69
04	42	14	73	38	25	39	11	24	94	72	18	06	46	29	32	40	62	74	36
20	49	34	41	72	30	23	88	34	62	99	69	82	47	59	85	74	04	34	24
20	23	35	29	78	31	90	01	74	31	49	71	48	86	81	14	23	57	05	54
01	70	54	71	83	51	54	49	16	92	33	48	61	43	51	01	89	19	67	48

컴퓨터가 보는 회색조 이미지의 2차원
행렬

이미지 입력

컬러 이미지

- 표준 RGB 색상 체계를 사용하는 컬러 이미지에는 채널이 3개 (빨간색, 녹색, 파란색)가 존재함
- 즉, 컬러 이미지를 나타내려면 빨간색의 강도를 나타내는 행렬, 녹색의 강도를 나타내는 행렬, 파란색의 강도를 나타내는 행렬 등 총 3개의 행렬이 필요함

Color image



$F(0, 0) = [11, 102, 35]$



Channel 1
Red intensity
values

Channel 2
Green intensity
values

Channel 3
Blue intensity
values

11	158	156	158	158	...
159	159	157	159	159	...
149	151	155	158	159	...
146	146	149	153	158	...
145	143	143	148	158	...
...

102	169	167	169	169	...
170	170	168	170	170	...
160	162	166	169	170	...
...

35	165	163	165	158	...
166	166	164	166	159	...
156	158	162	165	159	...
...

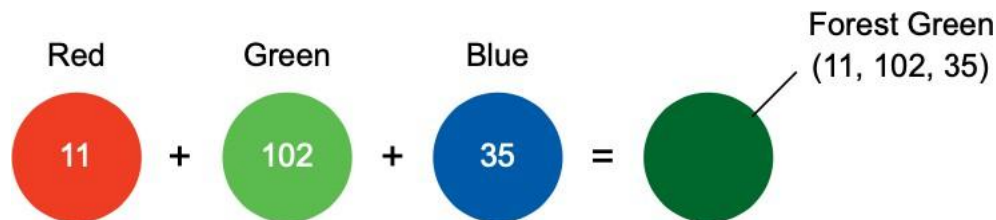
RGB channels

(HxWxC) = 24x24x3

이미지 입력

컬러 이미지

- 컴퓨터의 컬러는 서로 다른 강도의 색상 세 가지를 혼합한 것
- 세가지 색상을 서로 다른 강도로 합쳐 다양한 색을 만들 수 있음



Forest
HEX #0B6623
RGB 11 102 35

Forest green
Codes:
HEX #0B6623
RGB 11 102 35

Mint
HEX #98FB98
RGB 152 251 152

Mint green
Codes:
HEX #98FB98
RGB 152 251 152

Olive
HEX #708238
RGB 112 130 56

Olive green
Codes:
HEX #708238
RGB 112 130 56

Lime
HEX #C7EA46
RGB 199 234 70

Lime green
Codes:
HEX #C7EA46
RGB 199 234 70

Jungle
HEX #29AB87
RGB 41 171 135

Jungle green
Codes:
HEX #29AB87
RGB 41 171 135

Jade
HEX #00A86B
RGB 0 168 107

Jade green
Codes:
HEX #00A86B
RGB 0 168 107

이미지 전처리

컬러 이미지를 회색조 이미지로 변환하기

비전 시스템에서 계산 복잡도를 낮추기 위해 불필요한 정보를 제거해야 하는 경우가 존재함

이미지에서 대상을 인식 사물의 모양이나 특징을 활용 하는 것이 목적이라면 컬러 이미지를 회색조 이미지로 변환하여 사용할 경우 계산 복잡도를 낮출 수 있음

그러나 의료 이미지에서 붉은 발진을 확인해야 하는 진단 시스템 또는 자율주행자동차의 차선 감지 시스템은 컬러 정보가 문제 해결에 큰 역할을 하기때문에 계산 복잡도를 낮추기 위해 회색조 이미지로 변환하여 정보를 처리하는 것은 적절한 솔루션이 아님

이미지 전처리

기타

이미지 표준화

- 일부 머신러닝 알고리즘 ^{CNN}은 데이터 셋의 이미지를 모두 같은 크기로 조정해야 한다는 제약이 존재하며, 이런 알고리즘에 이미지를 입력하려면 배율을 조정해서 이미지의 크기를 같게 만들어야 함

데이터 강화

- 기존 데이터셋에 포함된 이미지를 조금씩 변형하는 방식으로 새로운 데이터를 추가하는 것
- 이미지를 변형하는 방법은 배율 조정, 회전, 아핀 변환 등을 주로 사용
- 추가한 데이터로 학습한 모델은 더 다양한 형태와 모양을 가진 대상도 인식할 수 있음

그 외 기법

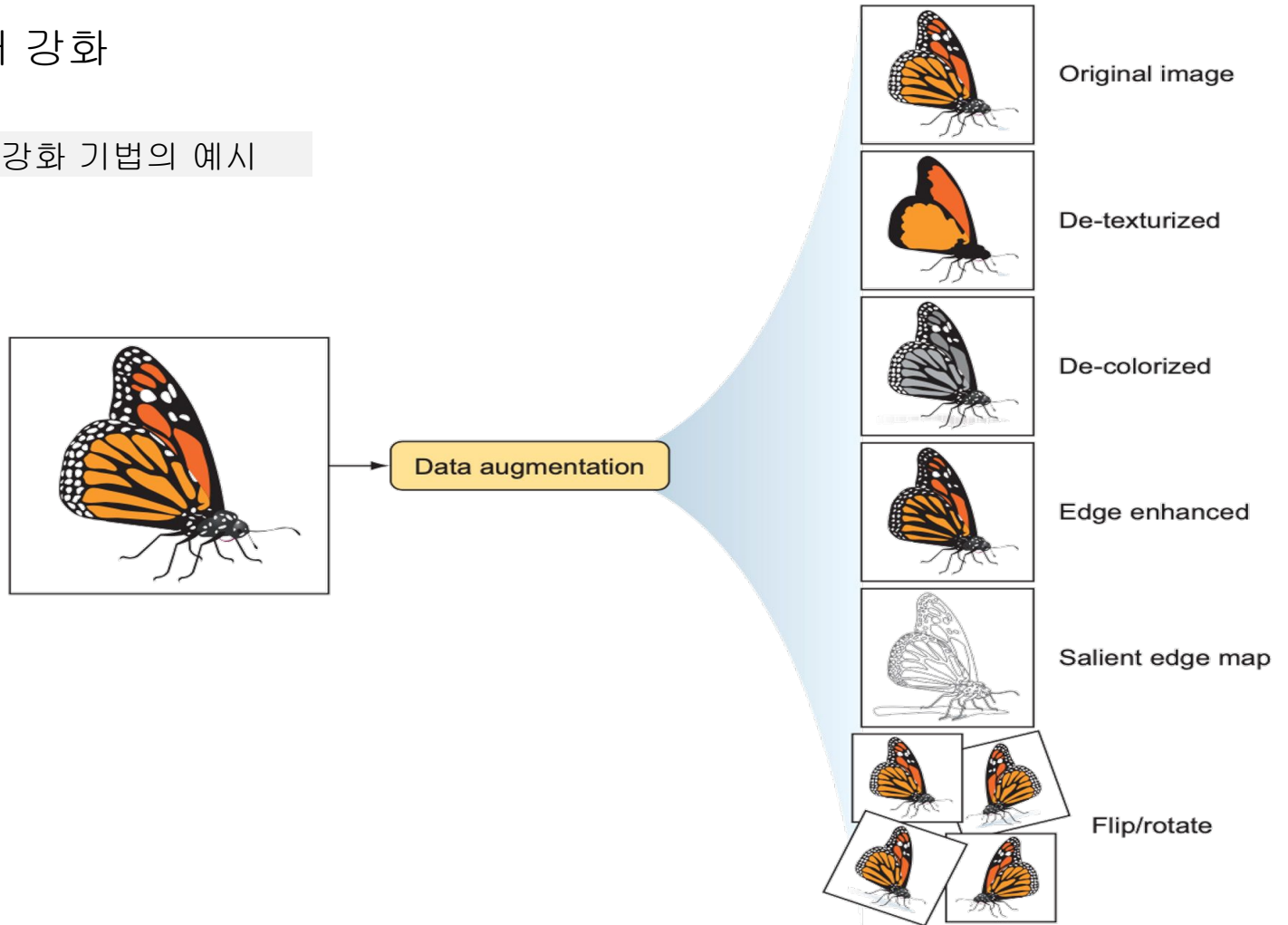
- 이미지에 포함된 노이즈를 줄이기 위해 배경색을 제거
- 필요에 따라 이미지의 밝기를 조절

이미지 전처리

기타

- 데이터 강화

데이터 강화 기법의 예시

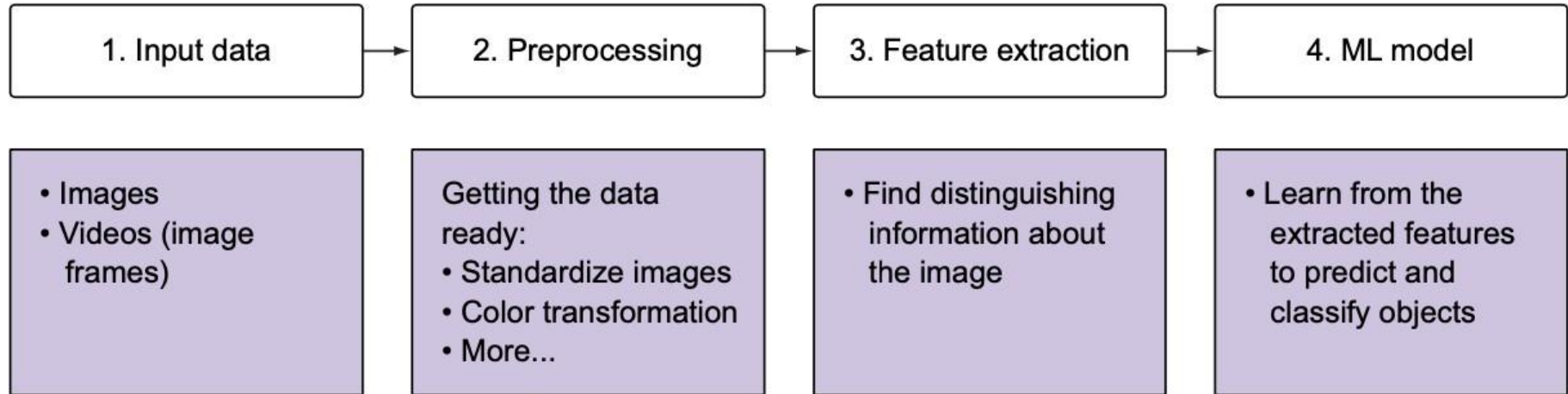


이미지 전처리

이미지 전처리 적용 방법

- 어떤 기법을 적용 할지는 데이터 셋과 해결하려는 문제가 무엇이냐에 따라 달라짐
- 실제 프로젝트 상황에 따라 필요한 기법을 선택하는 방법을 익히는 것이 좋음

특징 추출



특징 추출

컴퓨터 비전에서 사용되는 특징

- 이미지에서 특정 대상 모터사이클, 자전거, 개에만 해당하며 측정 가능한 데이터를 뜻함
- 선, 모서리, 이미지 조각 등 특이한 모양도 특징일 수 있고, 다른 것과 구분되는 색상도 특징이 될 수 있음
- 단, **좋은 특징은 대상을 다른 대상과 쉽게 구분 짓게 해주는 속성을 의미함**
- 좋은 특징 vs 나쁜 특징 예시

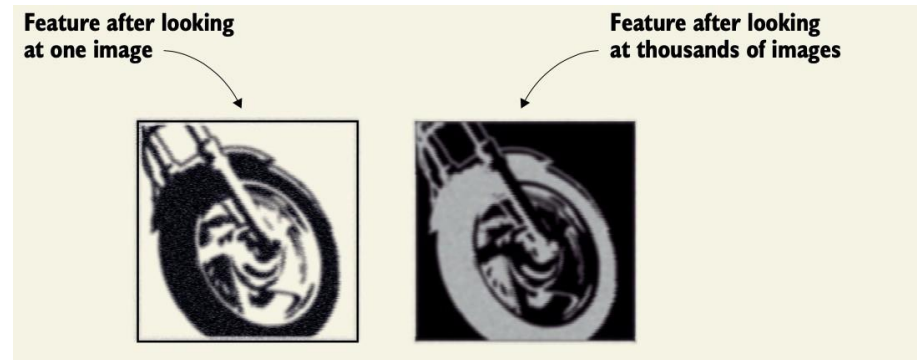


특징 추출

좋은 특징이란 무엇인가

일반화 가능한 특징

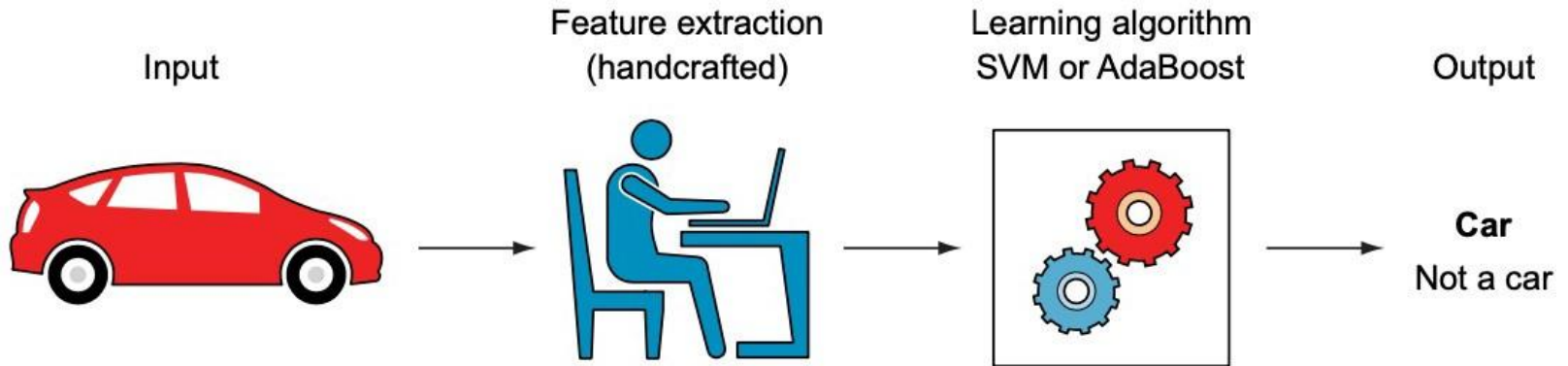
- 특징은 동일 대상에서 모두 찾아낼 수 있는 성질^{반복성, repeatability} 이어야 함
- 예를 들어, 특징은 모든 모터사이클에서 찾아낼 수 있는 성질 이어야 함.
- <바퀴>라는 특성의 경우 입력 이미지의 일부를 그대로 복사한 것을 특징으로 사용할 수 없음.



사물 인식에서 유용한 특징

- 다른 것과 구별될 것
- 추적과 비교가 쉬울 것
- 배율, 밝기, 각도가 달라도 일관적 일 것
- 노이즈가 많은 이미지 또는 대상의 일부만 찍힌 이미지에서도 관찰 가능할 것

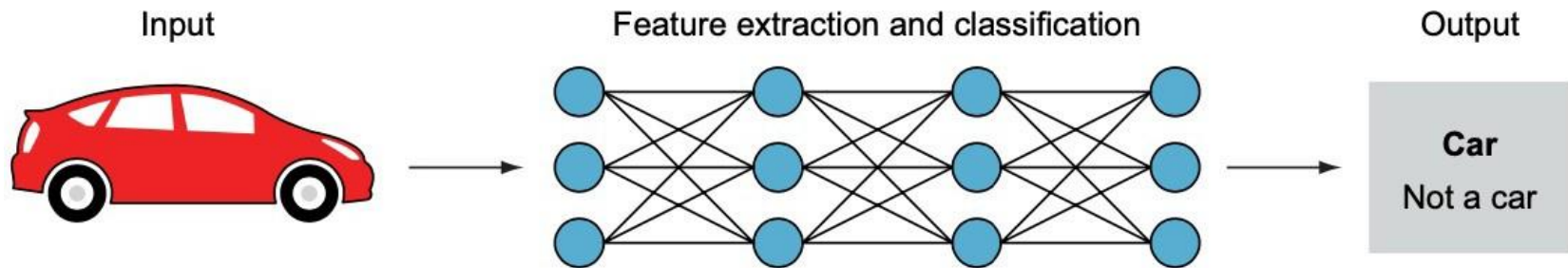
특징 추출



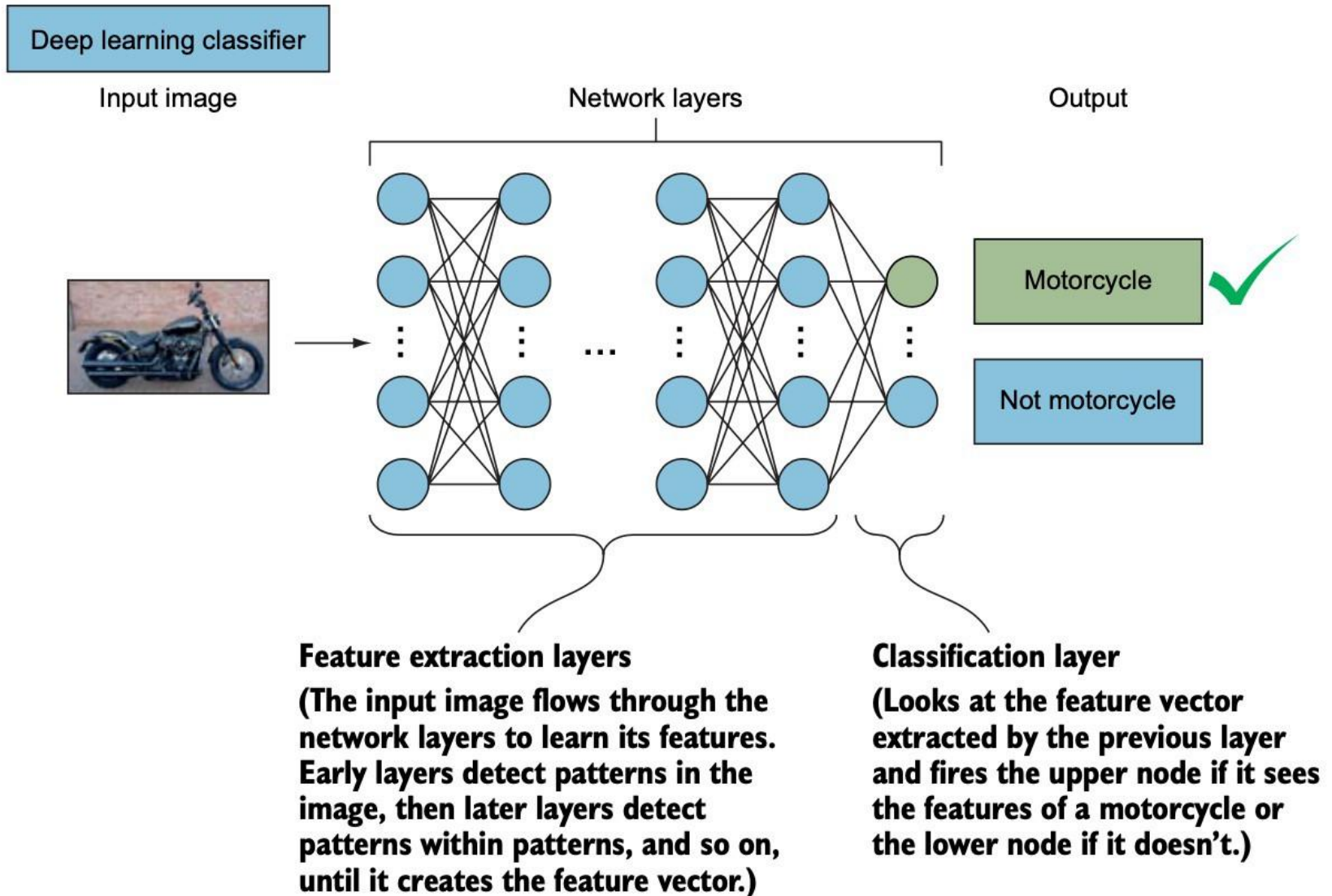
수동 추출된 특징의 예시

- 하르 캐스케이드 Haar Cascades
- 크기 불변 특징 변환 SIFT
- 고속 강인한 특징 추출 SURF
- 기울기 방향성 히스토그램 HOG

특징 추출



분류 학습 알고리즘



마치며

- 사람의 시각 시스템과 컴퓨터의 비전 시스템은 모두 감지 장치와 해석 장치로 구성된다.
- 해석 과정은 데이터 입력, 데이터 전처리, 특징 추출, 모델 생성으로 구성된다.
- 이미지를 x 와 y 의 함수 형태로 나타낼 수 있다. 컴퓨터는 픽셀 값의 행렬 형태로 이미지를 이해한다. 회색조 이미지는 채널 1개, 컬러 이미지는 채널 3개다.
- 이미지 처리 기법은 데이터 셋과 해결하려는 문제의 종류에 따라 달라진다. 상화에 맞춰 컬러 이미지를 회색조로 변환하거나, 신경망 입력 크기에 맞게 이미지 크기를 조정하거나, 데이터를 강화하는 방법이 사용된다.
- 특징은 이미지에 실린 대상을 분류하기 위해 사용되는 속성이다. 고전적인 머신러닝 알고리즘에는 다양한 특징 추출 기법이 쓰인다.