

패턴 인식

과목 개요

?? 찾기

- 1, 2, 3, 4, 5, ??
- 1, 2, 4, 8, 16, ??
- 2, 10, 9, ??

??

1

2

3

4

5

6

125	134	125	122	127	127	120	130	139	135	139	140	133	127	127	130	133	135	138	133	137	139	134	130	125	121
117	123	114	116	120	122	118	120	122	117	122	126	124	117	106	100	99	102	105	120	118	113	109	105	106	111
109	110	105	102	112	123	130	135	147	171	191	184	183	174	157	139	124	107	90	92	87	88	92	93	88	89
108	105	100	116	117	129	163	195	210	217	205	215	211	198	185	176	167	143	117	91	80	77	88	91	84	79
107	103	102	120	146	173	200	193	172	165	138	141	135	123	118	125	139	143	137	121	99	84	85	88	82	81
104	107	115	134	159	171	170	136	115	129	107	83	83	82	80	83	90	103	113	125	108	93	91	90	86	83
107	120	137	160	150	125	139	150	167	174	115	99	94	93	98	98	89	87	91	104	103	99	97	95	94	95
111	133	155	134	151	157	189	206	216	212	136	114	92	?	97	110	108	100	98	97	101	101	95	92	103	120
130	145	164	165	185	213	219	210	212	196	158	108	123	137	137	123	111	121	134	145	132	130	147	159	163	171
138	151	170	185	195	215	222	211	214	218	209	160	152	151	157	163	166	167	166	159	155	160	180	193	195	193
142	153	171	190	190	204	218	213	207	214	218	213	204	196	192	189	183	178	173	161	159	163	171	183	189	187
141	151	164	188	178	180	197	204	201	197	196	196	193	190	187	176	163	157	156	156	161	163	166	174	186	192
144	151	160	185	183	176	176	187	192	191	188	193	184	178	177	174	165	156	151	148	163	177	182	188	200	203
152	160	168	176	193	193	182	180	180	174	172	164	161	159	154	146	140	143	149	173	184	190	190	193	199	205
159	168	178	178	202	206	197	194	187	175	175	167	172	179	180	176	176	188	203	215	212	206	204	202	204	205
161	171	185	197	210	204	199	211	210	206	212	219	210	206	215	225	226	220	215	214	209	210	214	216	211	200

패턴 인식

- 그림, 음원, 글 등의 데이터 안에서 특정한 패턴을 찾아내는 기술
- 사운드/이미지에서 많이 사용되는 기술
- 초반에는 Heuristic하게 알고리즘을 만들었으나, 기계학습/딥러닝 등을 통해 정확도가 향상됨

영상 처리

- 입력된 영상을 어떤 목적을 위해 처리하는 기술
- 어떤 목적을 위해 “수학적 연산”을 이용해 화소들에 대해 변화를 주는 것
- 아날로그 영상 처리/ 디지털 영상 처리

컴퓨터 비전

- 영상처리 기술에 인공지능 기술(패턴 인식 등)을 이용하여 인간의 시각을 흉내내는 것

인식 (recognition)

- 사람과 동물의 인식 기능-매우 강건^{robust}



(a) 누구인가?

인식 인식 인식
인식 인식 인식
인식 인식 인식
인식 인식 인식

(b) 무슨 글자인가?

(키,몸무게,높은혈압,낮은혈압,
혈당,SGOT,SGPT)
=(169,71,130,80,94,18,26)

(c) 정상인가?

그림 1.1 인식 해보자.

- 컴퓨터라는 기계도 인식할 수 있나?

1.1 왜 패턴 인식인가?

- 사람과 기계의 차이가 연구자들을 끌어들이는 매력으로 작용
 - 인식은 사람에게 극히 쉽다.
 - 인식은 기계에게 극히 어렵다.
- 과학적 접근
 - 뇌의 정보 처리 과정에 대한 어느 정도의 이해를 바탕으로
 - 뇌를 모방한 인식 컴퓨터에 대한 욕구
 - 예, 신경망 연구

1.1 왜 패턴 인식인가?

- 공학적 접근
 - 우리 주위에 널려 있는 인식 기계
 - 기존 제품에 인식 기능 추가하여 부가 가치 높임
 - 우편물 분류기
 - PDA 필기 입력기
 - 동작인식 핸드폰
 - 지문인식 마우스
 - 과속 단속기
 - 청소 로봇



(a) 주차 단속용 번호판 인식기



(b) 청소 로봇

그림 1.2 인식을 하는 지능 기계

1.1 왜 패턴 인식인가?

■ 팽창하는 응용

- 공장 자동화
- 문서 인식
- 음성 인식
- 자연어 처리
- 데이터 마이닝
- 정보 검색
- 사람 컴퓨터 인터페이스^{HCI}
- 생체 인식
- 지능 교통 시스템^{ITS}
- 지능 자동차
- 생물 정보학
- 지능 로봇
- 상황 인식
- ...

1.2 어떻게 인식하나?

■ 패턴 인식의 가장 간단한 관점

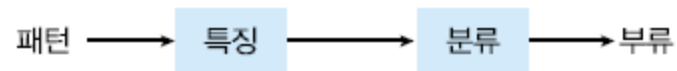


그림 1.3 패턴 인식의 가장 간단한 관점

■ 개념적으로 생각해 보자.

- 얼굴이 작고, 코가 뾰족하고, 눈썹이 짙고, 눈이 작은 샘플이 있다면,

■ 특징

- 얼굴 크기 (x_1), 코의 모양 (x_2), 눈썹의 짙은 정도 (x_3), 눈의 크기 (x_4)

■ 분류

- x_1 =작다, x_2 =뾰족하다, x_3 =짙다, x_4 =작다 라는 패턴이 들어왔을때, 이미 알고있는 지식에 비추어 아무개일 가능성이 높다는 의사 결정 과정

1.2 어떻게 인식하나?

■ 데이터베이스 수집

□ 용어

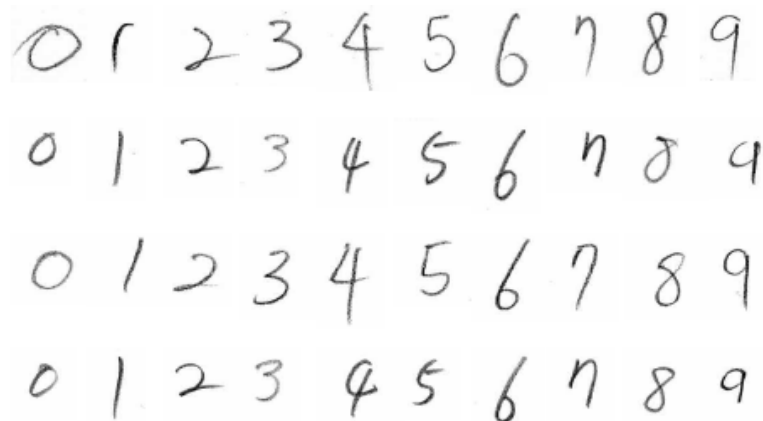
■ 샘플

■ 훈련 집합과 테스트 집합

□ 고성능 인식 시스템을 만들기 위해서는 고품질 데이터베이스 필수

■ 양적 우수성과 질적 우수성

□ 필기 숫자 예



00011110 00001111 01100000 01111100 00011000 01111110 00001100 01111110 00111111 00111111
00110110 00111000 00110000 00001100 00111000 01000000 00010000 11100011 01110001 11000011
01100001 01100000 00010000 00011000 01101000 01000000 00100000 11000001 00110011 10001111
00100001 11100000 00010000 00111100 11101111 01110110 01100000 00000011 00011111 01110011
10100001 11000000 00010000 00000011 00001100 00000001 11000110 00000011 00111100 00000010
01000010 11000000 00111111 00000001 00001100 00000001 11000011 00000110 01100010 00000010
01000110 11000000 01100000 00000011 00000100 10000011 11000001 00000100 11000010 00000010
00111100 01000000 11000000 11111100 00000100 11111100 11111110 00001000 11111100 00000010

00001111 00000111 11110000 11111111 00111100 00111000 00000011 01100111 00011110 00010011
00011111 00000111 10001000 10000001 01101100 01101111 00000110 01111101 00000111 01000011
00110001 00000111 00000100 00011111 11011000 11011000 00011000 01110001 00001110 11000011
01100001 00001110 00000100 00011110 11011011 10000000 00110000 11100001 00110110 10000110
11000001 00011110 00000100 00000011 11111110 11111000 01111000 11000011 01100110 11111010
11000011 00111100 00001100 00000011 00100000 00001000 11100100 00000111 11000110 00000010
11000111 00111100 00001111 00000110 01100000 01011000 11001100 00000110 10001100 00000010
01111100 01111000 01111100 11111000 01100000 01110000 11111000 00001110 11111000 00000010

00111110 00000011 00001110 00001111 00010110 01111111 00000011 00111111 00011111 01111111
01110011 00000011 00000001 00000001 00110100 01100000 00000110 11100001 01100001 11000001
01000001 00000110 00000001 00000110 01100100 01000000 00001100 10000011 01111010 11000011
11000001 00001100 00000001 00011100 11001111 11000000 00011000 00000110 00001100 00101001
10000001 00011000 00000011 00000111 00001000 11001110 00100000 00001100 00110110 00000011
10000001 01110000 00000110 00000001 00001000 01100011 01110000 00011000 01000010 00000011
10000010 11100000 00111100 00000011 00001000 00000010 11001100 00110000 11000010 00000011
01111100 11000000 11110000 11111100 00001000 11111100 11011000 01100000 11101100 00000011

00011110 00000111 01110000 00111111 00011110 00001110 00000011 00111111 00111111 00011110
00011111 00000111 10001000 11000011 01100110 00111001 00000110 00111001 00110001 01100011
00110001 00000111 00001000 00011110 11000100 01111110 00001100 01100001 00010110 11000011
01100001 00001111 00001000 00011110 11001001 01000000 00011000 01000001 00011100 10000011
01100001 00011110 00001000 00000011 11111111 01111000 00110000 11000011 00111100 11111111
11000110 00111100 00011111 00000011 00110000 00001100 01111000 00000010 01101100 00100010
10011100 01111000 01111000 00000110 00100000 00001100 11001100 00000110 11001100 00000010
11111000 11110000 01100000 11111100 01100000 11111000 11111000 00000110 11111000 00000010

1.2 어떻게 인식하나?

■ 특징 (필기 숫자 예)

- 개별적인 화소를 특징으로 한다면,
 - 64개의 특징
 - 64-차원 특징 벡터 (feature vector) $\mathbf{x}=(x_1, \dots, x_{64})$
 - 이들 특징의 분별력은?
- 검은 화소의 비율을 특징으로 한다면,

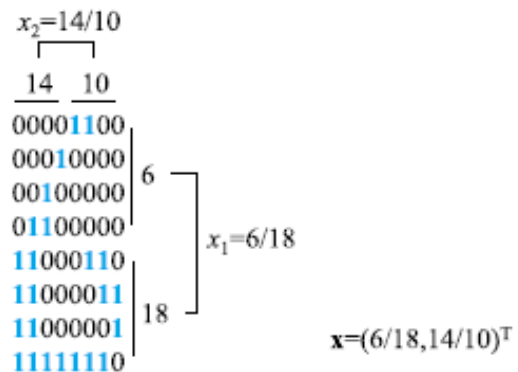


그림 1.5 필기 숫자에서 특징 추출 예

□ 중요한 문제

- 분별력 discriminating power
- 차원의 저주 curse of dimensionality

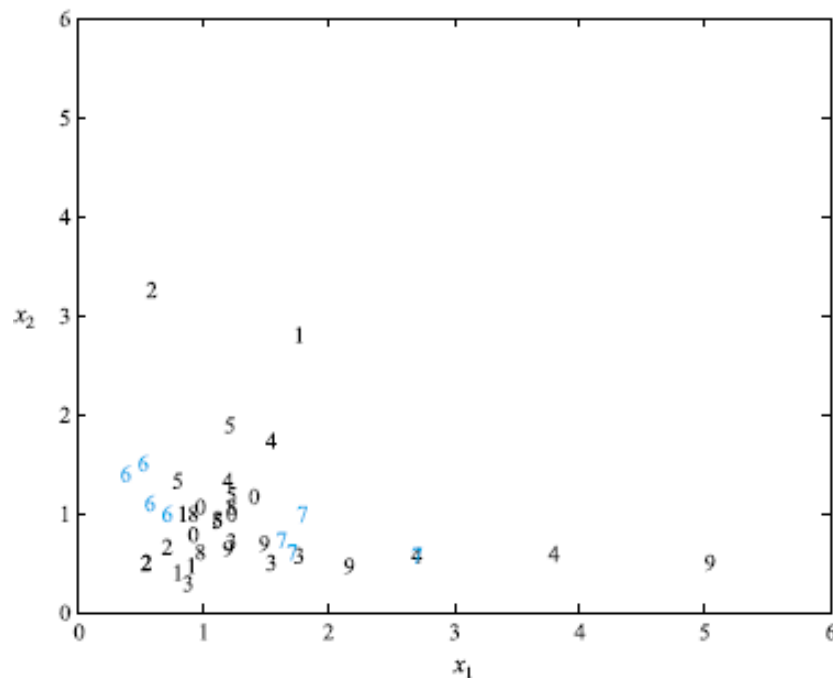


그림 1.6 특징 공간에서 샘플의 분포

1.2 어떻게 인식하나?

- 분류

- 두 단계의 설계

- 모델 선택과 학습^{learning}

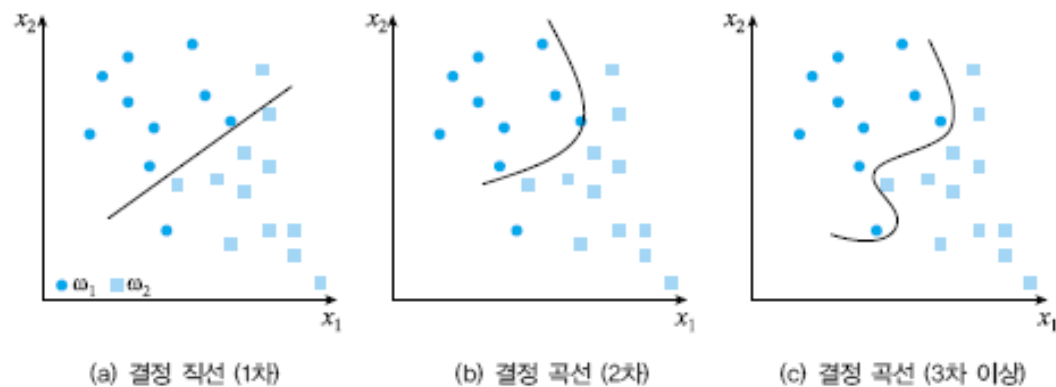


그림 1.7 분류기 설계 (모델 선택과 학습의 두 단계를 거침)

- 선형 분류기와 비선형 분류기

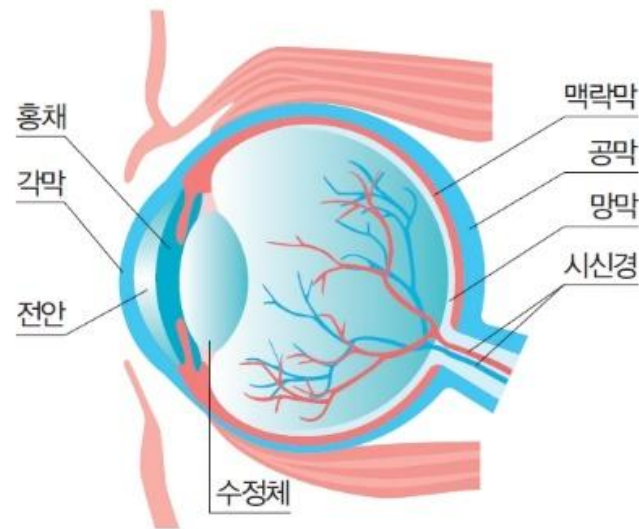
- 결정 직선, 결정 평면, 결정 초평면
 - 결정 곡선, 결정 곡면, 결정 초곡면

1.1 인간의 시각

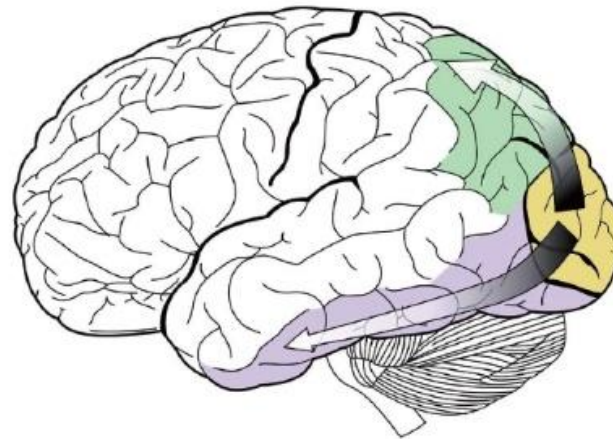
■ 시각은 오감 중에서 가장 뛰어남

■ 인간의 눈의 구조와 동작

- 등쪽 경로(녹색)는 주로 물체의 움직임, 배쪽 경로(보라색)는 주로 물체의 부류를 알아냄
- 매 순간 빠르고 정확하게 그리고 아주 손쉽게 인식



(a) 눈의 구조



(b) 시각 정보 처리를 담당하는 시각 피질

그림 1-2 인간의 시각 시스템

1.1 인간의 시각

■ 인간 시각의 강점

- 분류, 검출, 분할, 추적, 행동 분석에 능숙
- 3차원 복원 능력
- 빠르고 강건
- 다른 지능 요소인 지식 표현, 추론, 계획과 협동
- 사전 행동_{proactive}에 능숙
- 과업 전환이 매끄럽고 유기적이고 빠름
- 비주얼 서보잉이 뛰어남

1.1 인간의 시각

■ 인간 시각의 한계

- 착시가 있음

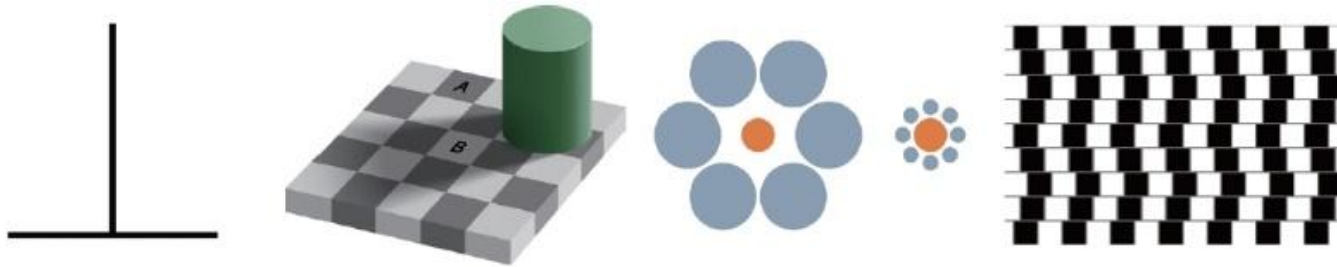


그림 1-3 인간 시각의 착시 현상(출처: 영문 위키피디아 'optical illusion')

- 정밀 측정에 오차
- 시야가 한정됨
- 피로해지고 퇴화

1.2 왜 컴퓨터 비전인가?

■ 컴퓨터 비전은 인간의 시각을 흉내 내는 컴퓨터 프로그램

- 인공지능의 중요한 구성 요소, 예) 시각 기능이 없는 로봇은 낮은 성능
- 파란 테두리 상자는 이 책의 범위

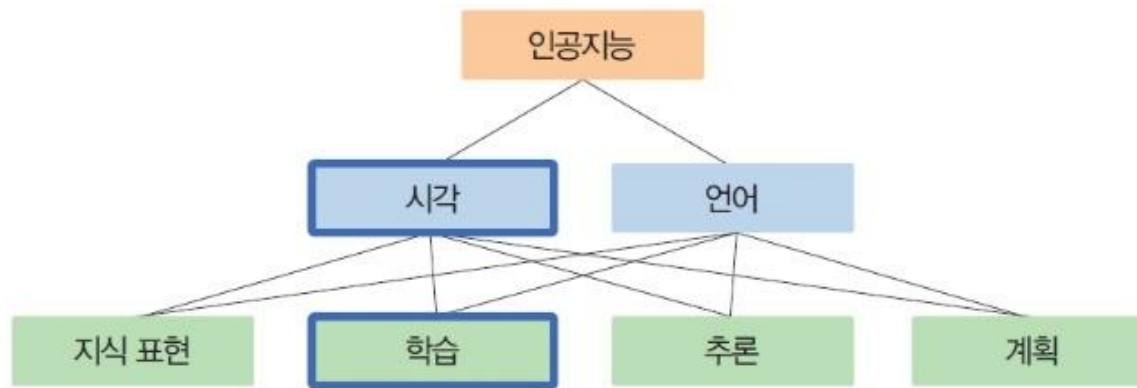


그림 1-4 인공지능의 실현

- 현재 컴퓨터 비전 기술로 인간에 필적하는 시각 구현은 불가능
- 과업을 한정하면 인간 성능에 가깝거나 뛰어넘는 응용이 무궁무진

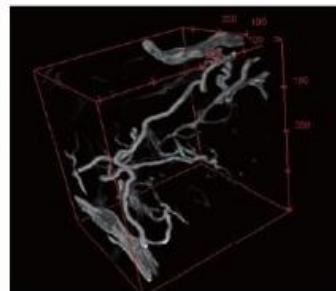
1.2 왜 컴퓨터 비전인가?

■ 몇 가지 대표적인 응용 사례

- 농업
- 의료
- 교통
- 스마트 공장
- 스포츠
- 유통



(a) 과일 수확 드론



(b) 혈관 분할



(c) 자율주행



(d) 불량 검사



(e) 선수의 행동 분석



(f) 고객의 동선 분석

1.2 왜 컴퓨터 비전인가?

■ 몇 가지 대표적인 응용 사례(...계속)

- 보안
- 에너지
- 엔터테인먼트
- 환경
- 우주과학
- 감시
- 예술
- 가사
- 휴머노이드 로봇



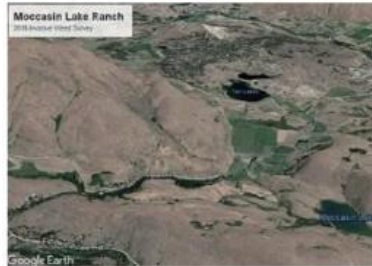
(g) 얼굴 인식 보안



(h) 태양광 모니터링



(i) 게임 플레이(알파스타)



(j) 지형 모니터링



(k) 화성 탐사선



(l) 광장 감시



(m) 에드몽 벨라미



(n) 청소 로봇



(o) 휴머노이드 로봇

그림 1-5 컴퓨터 비전의 응용 사례

1.3 컴퓨터 비전은 왜 어려운가?

■ 컴퓨터 비전이 어려운 이유는 명확

■ 세상의 변화무쌍함

- 환경 (낮밤, 날씨 등) 변화, 보는 위치와 방향의 변화, 강체와 연성 물체
- 원자부터 우주까지 긴 스펙트럼에서 영상 수집

■ 컴퓨터는 넘버 크런처

125	134	125	122	127	127	120	130	139	135	139	140	133	127	127	130	133	135	138	133	137	139	134	130	125	121
117	123	114	116	120	122	118	120	122	117	122	126	124	117	106	100	99	102	105	120	118	113	109	105	106	111
109	110	105	102	112	123	130	135	147	171	191	184	183	174	157	139	124	107	90	92	87	88	92	93	88	88
108	105	100	116	117	129	163	195	210	217	205	215	211	198	185	176	167	143	117	91	80	77	88	91	84	79
107	103	102	120	146	173	200	193	172	165	138	141	135	123	118	125	139	143	137	121	99	84	85	88	82	81
104	107	115	134	159	171	170	136	115	129	107	83	83	82	80	83	90	103	113	125	108	93	91	90	86	83
107	120	137	160	150	125	139	150	167	174	115	99	94	93	98	98	89	87	91	104	103	99	97	95	94	95
111	133	155	134	151	157	189	205	216	212	136	114	92	83	97	110	108	100	98	97	101	101	95	92	103	120
130	145	164	165	185	213	219	210	212	195	158	108	123	137	137	123	111	121	134	145	132	130	147	159	163	171
138	151	170	185	195	215	222	211	214	218	209	160	152	151	157	163	166	167	166	159	155	160	180	193	195	193
142	153	171	190	190	204	218	213	207	214	218	213	204	195	192	189	183	178	173	161	159	163	171	183	189	187
141	151	164	188	178	180	197	204	201	197	196	196	193	190	187	176	163	157	156	156	161	163	166	174	186	192
144	151	160	185	183	176	176	187	192	191	188	193	184	178	177	174	165	156	151	148	163	177	182	188	200	203
152	160	168	176	193	193	182	180	180	174	172	164	161	159	154	146	140	143	149	173	184	190	190	193	199	205
159	168	178	178	202	206	197	194	187	175	175	167	172	179	180	176	176	188	203	215	212	206	204	202	204	205
161	171	185	197	210	204	199	211	210	208	212	219	210	206	215	225	228	220	215	214	209	210	214	216	211	200



그림 1-6 컴퓨터 비전이 인식해야 하는 영상은 아주 큰 숫자 배열

■ 인공지능의 미숙함

- 지식 표현, 추론, 계획, 학습이 유기적으로 동작할 때만 강한 인공지능 가능
- 강한 인공지능은 먼 미래의 일 또는 영영 불가능

1.4 컴퓨터 비전의 역사

■ 신문 산업에서 태동한 디지털 영상

- 1920년 유럽과 북미 간 케이블을 통해 사진 전송하는 Bartlane 시스템 개통

■ 1946년 세계 최초의 범용 전자식 컴퓨터인 에니악 탄생

- 빠른 계산이 주목적 (에니악은 초당 3000개 가량 덧셈)

■ 1957년 스캐너를 통해 디지털 영상을 컴퓨터에 저장

- 5cmx5cm 사진에서 획득한 176x176 디지털 영상 ← 컴퓨터 비전의 태동



Bartlane 시스템(유럽-북미 해저 케이블로 사진을 전송하는 시스템)으로 전송된 세계 최초의 디지털 영상(1920년)

→
37년



세계 최초로 컴퓨터에 저장된 디지털 영상 (1957년)

→
65년



자율주행차(현재)

그림 1-7 컴퓨터 비전의 발전

1.4 컴퓨터 비전의 역사

표 1-1 컴퓨터 비전의 역사

연도	사건
1920	• Bartlane 영상 전송 케이블 시스템 구축 [McFarlane1972]
1946	• 세계 최초 전자식 범용 디지털 컴퓨터인 에니악 탄생
1957	• 커쉬가 세계 최초로 디지털 영상을 컴퓨터에 저장
1958	• 로젠블라트의 퍼셉트론 제안(이후 Mark 1 Perceptron에서 문자 인식 실험)
1968	• 소벨의 소벨 에지 연산자 제안
1979	• IEEE Transactions on Pattern Analysis and Machine Intelligence 창간 • ACRONYM 시스템 발표 [Brooks1979]
1980	• 후쿠시마의 네오코그니트론 논문 발표 [Fukushima1980]
1983	• 제1회 CVPR(Computer Vision and Pattern Recognition)이 미국 알링턴에서 개최
1986	• 캐니의 캐니 에지 연산자 논문 발표 [Canny1986] • 루멜하트의 『Parallel Distributed Processing』 출간(다층 퍼셉트론 제안) [Rumelhart1986]

1.4 컴퓨터 비전의 역사

1987	<ul style="list-style-type: none">• International Journal of Computer Vision 창간• 런던에서 제1회 ICCV(International Conference on Computer Vision) 개최(홀수 연도)• Marr상 제정(ICCV에서 시상)• 덴버에서 제1회 NIPS(Neural Information Processing Systems) 개최(2018년에 NeurIPS로 개명)
1990	<ul style="list-style-type: none">• 프랑스 안티베에서 제1회 ECCV(European Conference on Computer Vision) 개최(짝수 연도)
1991	<ul style="list-style-type: none">• Eigenface 얼굴 인식 논문 발표 [Turk1991]
1998	<ul style="list-style-type: none">• 르쿤의 컨볼루션 신경망 논문 발표 [LeCun1998]
1999	<ul style="list-style-type: none">• 로우의 SIFT 논문 발표 [Lowe1999]• 엔비디아에서 GPU 발표
2000	<ul style="list-style-type: none">• CVPR에서 OpenCV 알파 버전 공개
2001	<ul style="list-style-type: none">• Viola-Jones 물체 검출 논문 발표 [Viola2001]
2004	<ul style="list-style-type: none">• 그랜드 챌린지(고속도로 자율주행)
2005	<ul style="list-style-type: none">• PASCAL VOC 대회 시작

1.4 컴퓨터 비전의 역사

2006	<ul style="list-style-type: none">• OpenCV 1.0 공개
2007	<ul style="list-style-type: none">• 어번 챌린지(도심 자율주행)• Azriel Rosenfeld Lifetime Achievement상 제정
2009	<ul style="list-style-type: none">• 페이페이 리가 CVPR에서 ImageNet 데이터셋 발표• OpenCV 2.0 공개
2010	<ul style="list-style-type: none">• Xbox 360을 위한 Kinect 카메라 시판• 제1회 ILSVRC 대회 개최• MS COCO 데이터셋 발표
2012	<ul style="list-style-type: none">• ILSVRC 대회에서 AlexNet 우승 [Krizhevsky2012]• 시각 장애인을 태운 자율주행차의 시범 운행 성공
2013	<ul style="list-style-type: none">• 아타리 비디오 게임에서 사람 성능 추월 [Mnih2013]• 스코츠데일에서 제1회 ICLR(International Conference on Learning Representations) 개최
2014	<ul style="list-style-type: none">• RCNN 논문 발표 [Girshick2014]• 생성 모델인 GAN 발표 [Goodfellow2014]• ILSVRC에서 GoogLeNet이 우승, VGGNet이 준우승
2015	<ul style="list-style-type: none">• 텐서플로 서비스 시작• ILSVRC에서 ResNet이 우승

1.4 컴퓨터 비전의 역사

2016	<ul style="list-style-type: none">• 파이토치 서비스 시작• YOLO 논문 발표 [Redmon2016]
2017	<ul style="list-style-type: none">• 트랜스포머 논문 발표 [Vaswani2017]• Open Images 데이터셋 공개• 구글 렌즈 서비스 시작
2018	<ul style="list-style-type: none">• 인공지능이 그린 에드몽 벨라미가 경매에서 5억 원에 낙찰• 벤지오, 힌튼, 르쿤 교수가 딥러닝으로 튜링상 수상
2019	<ul style="list-style-type: none">• 알파스타가 스타크래프트에서 그랜드마스터 수준 달성• 트랜스포머를 위한 파이썬 라이브러리 transformers 2.0 공개
2020	<ul style="list-style-type: none">• OpenAI 재단의 GPT-3 발표• iPad Pro에 라이다 센서 장착
2021	<ul style="list-style-type: none">• 비전 트랜스포머 발표 [Dosovitskiy2021]• OpenAI 재단의 DALL·E 발표 [Ramesh2021]
2022	<ul style="list-style-type: none">• 구글의 Imagen 발표 [Saharia2022]

1.5 컴퓨터 비전 체험 서비스

■ 컴퓨터 비전 커뮤니티의 공개 문화

- SOTA 달성한 연구자는 논문 발표와 더불어 깃허브에 소스 코드와 데이터 공개하는 문화
- 이를 활용한 웹/앱 서비스 활성화



그림 1-8 Google 앱

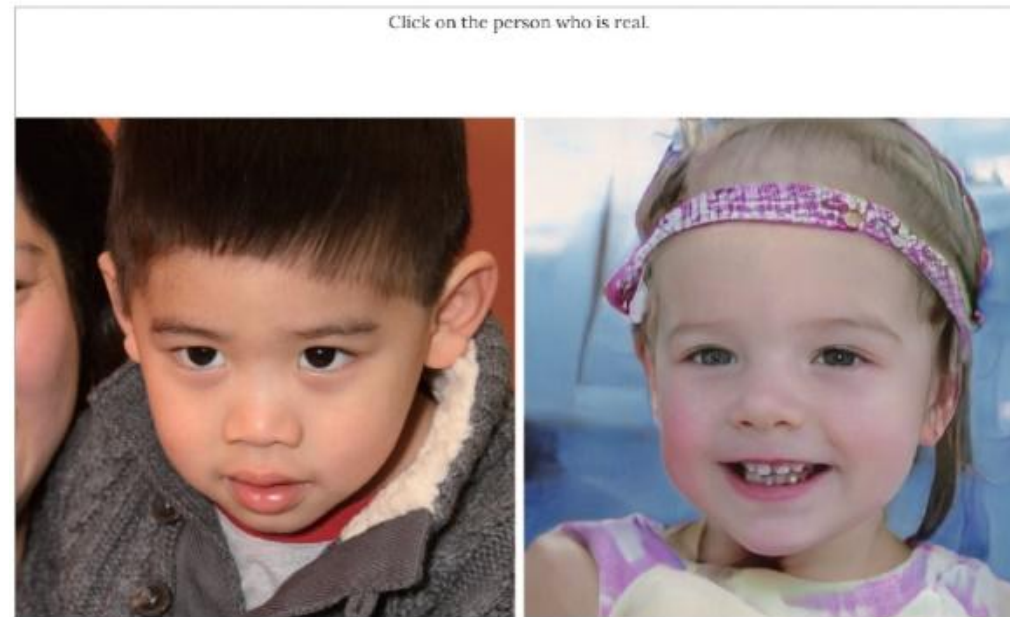


그림 1-9 Which face is real?(왼쪽이 진짜)

1.5 컴퓨터 비전 체험 서비스

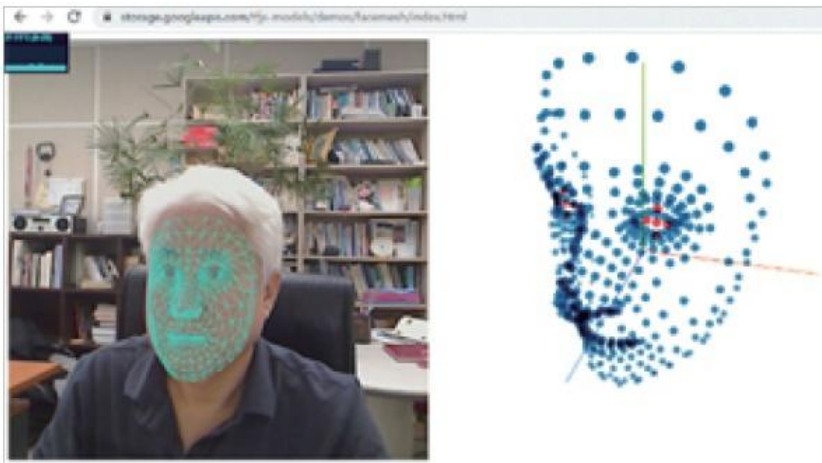


그림 1-10 얼굴 랜드마크 검출



그림 1-11 영상 설명

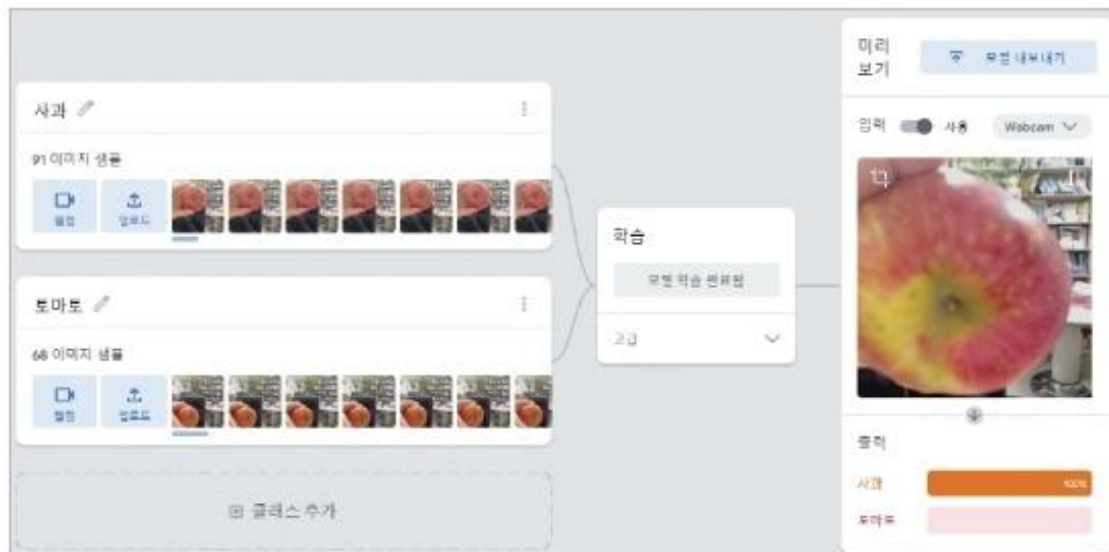


그림 1-12 티처블 머신

1.6 컴퓨터 비전 만들기

■ 궁극적인 목표

- 일반적인 상황에서 잘 작동하는 인간과 같은 시각 (강한 인공지능)
- 영영 불가능하거나 먼 미래에 실현

■ 현실적인 목표

- 제한된 환경에서 특정 과업을 높은 성능으로 달성 (약한 인공지능)
- 컴퓨터 비전 문제를 여러 세부 문제로 구분하고 세부 문제별로 알고리즘 구상

컴퓨터 비전이 풀어야 할 문제

■ 기본 문제

- 분류
- 검출
- 분할
- 추적
- 행동 분석
- ...



(a) 분류



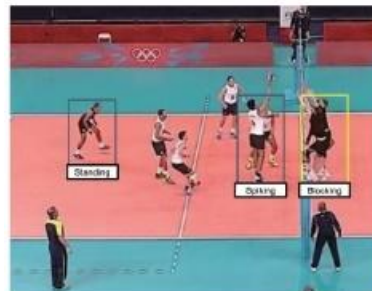
(b) 검출



(c) 분할



(d) 추적(<https://motchallenge.net/vis/MOT17-09-SDP>)



(e) 행동 분석(<https://github.com/mostafa-saad/deep-activity-rec#dataset>)

그림 1-13 컴퓨터 비전이 풀어야 할 문제

컴퓨터 비전이 풀어야 할 문제

■ 특정 상황에 따라 다양하게 변형

- 예) 사과 따는 로봇 비전 → 사과 검출에만 집중. 로봇 손을 위해 정확한 위치가 중요

■ 다른 지능 요소와 협업

- 가장 활발한 협업 분야는 자연어 처리, 예) 영상 설명하기
- 로봇과 협업은 활발, 예) 비주얼 서보잉을 통한 눈-손 협업