

컴퓨터 비전과 딥러닝

Chapter 01 인간의 시각을 흉내 내는 컴퓨터 비전

차례

1.1 인간의 시각

1.2 왜 컴퓨터 비전인가?

1.3 컴퓨터 비전은 왜 어려운가?

1.4 컴퓨터 비전의 역사

1.5 컴퓨터 비전 체험 서비스

1.6 컴퓨터 비전 만들기

1.7 읽을거리와 볼거리

Preview

■ 놀라운 인간의 시각

- 인간은 아래 영상을 보고 인식, 추론, 예측, 상상 등을 수행함
- 선수가 얻을 점수까지 추정



그림 1-1 인간이 쉽고 정확하게 해석할 수 있는 영상

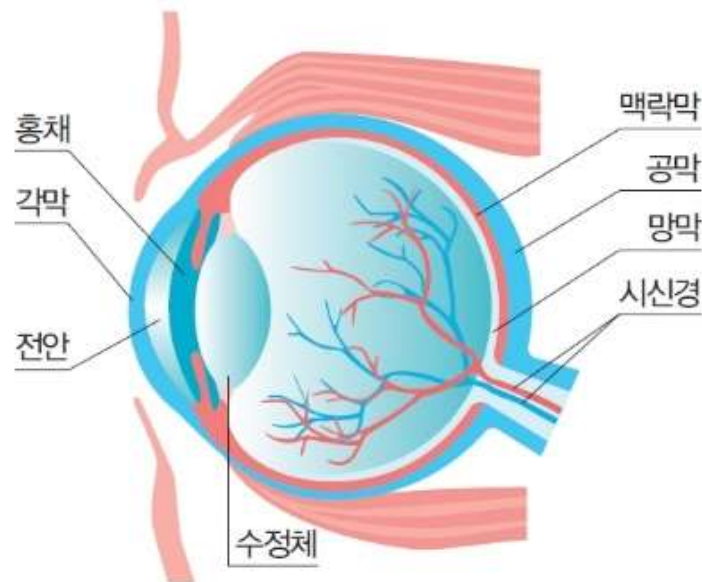
■ 컴퓨터가 인간 시각을 흉내 낼 수 있을까? 컴퓨터 비전 세계로 쏙~~~

1.1 인간의 시각

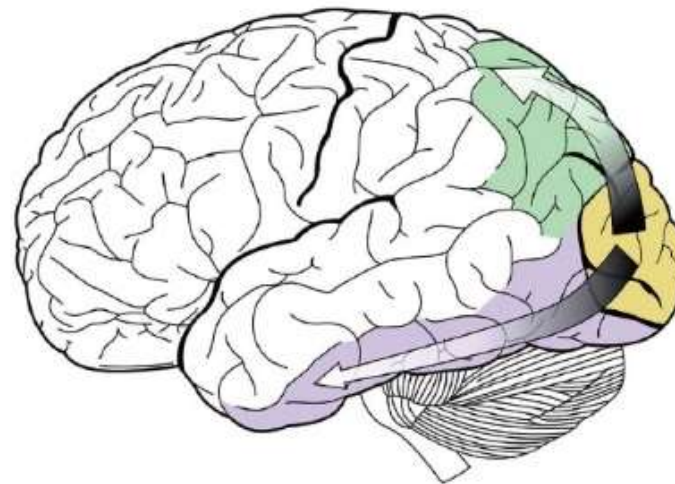
■ 시각은 오감 중에서 가장 뛰어남

■ 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

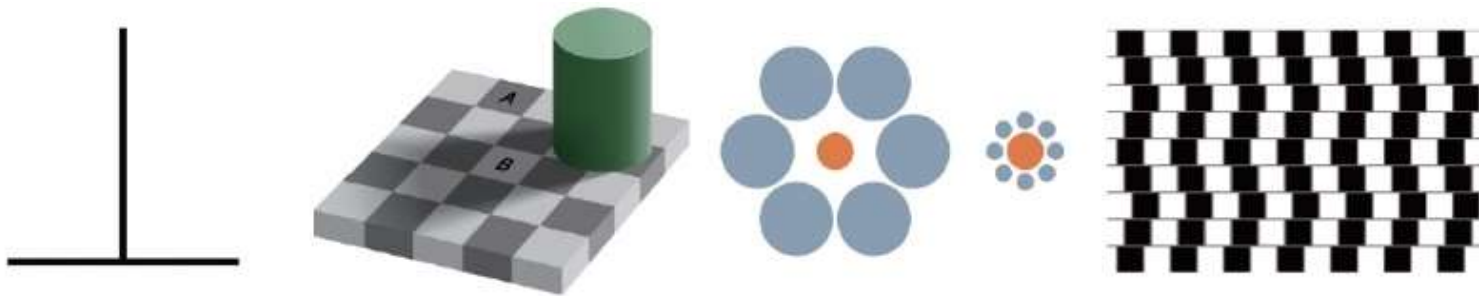


그림 1-3 인간 시각의 착시 현상(출처: 영문 위키피디아 'optical illusion')

- 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

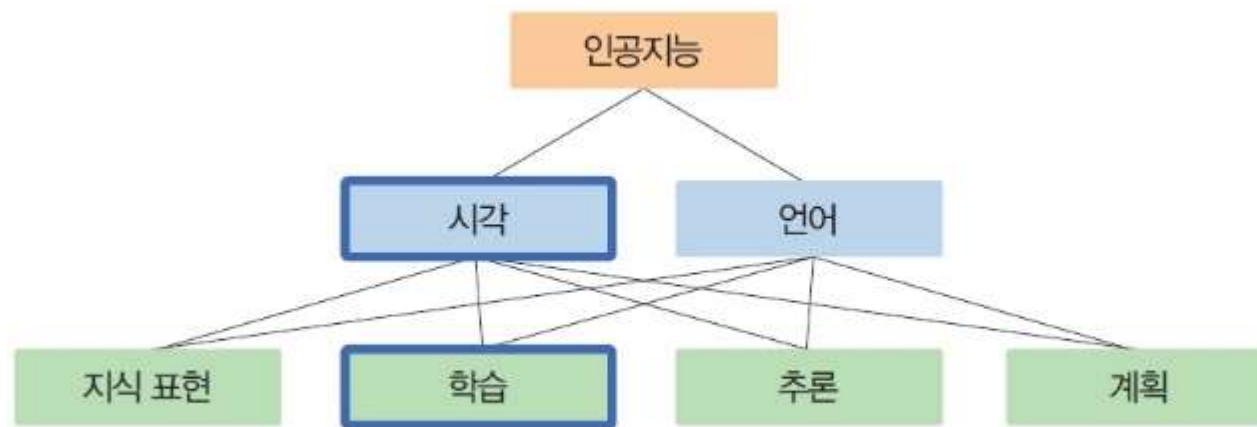


그림 1-4 인공지능의 실현

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

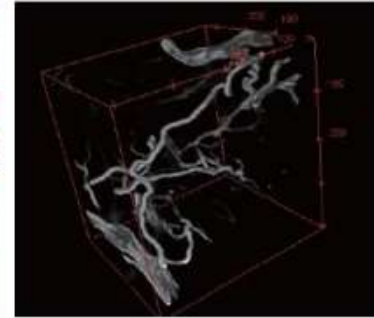
1.2 왜 컴퓨터 비전인가?

■ 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



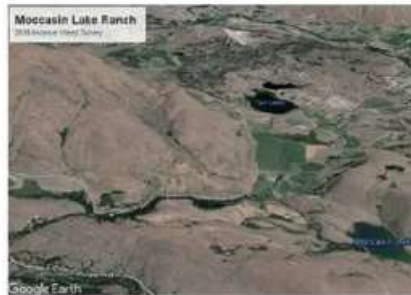
(g) 얼굴 인식 보안



(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 (영상종류의 다양성)
- 컴퓨터 -> 넘버 크런처 (number cruncher)

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
107 120 137 160 150 125 138 150 167 174 115 98 94 93 98 98 89 87 91 104 103 99 97 95 94 95
111 133 156 134 151 157 189 206 216 212 136 114 92 83 97 110 108 100 98 97 101 101 95 92 103 120
130 145 164 165 185 213 219 210 212 196 158 108 123 137 137 123 111 121 134 145 132 130 147 159 163 171
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
142 153 171 190 190 204 218 213 207 214 218 213 204 195 192 189 183 178 173 161 159 163 171 183 189 187
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
159 168 178 178 202 206 197 194 187 175 175 167 172 179 180 176 176 188 203 215 212 208 204 202 204 205
161 171 185 197 210 204 199 211 210 208 212 219 210 206 215 225 226 220 215 214 209 210 214 216 211 200



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 어쩌면 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

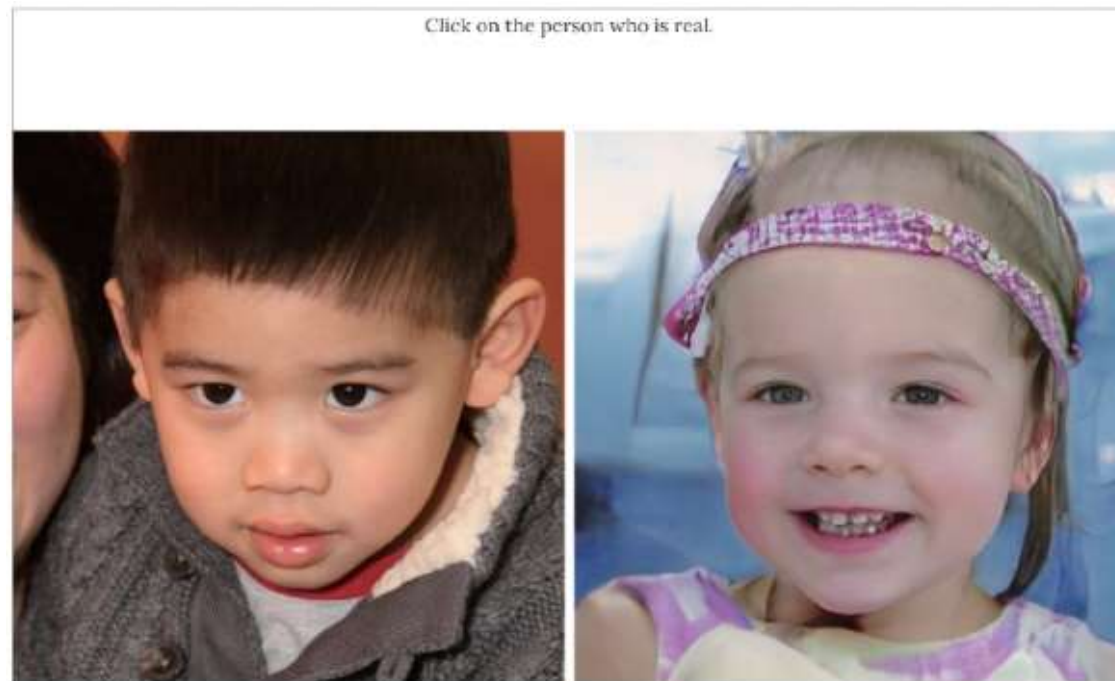


그림 1-9 Which face is real?(왼쪽이 진짜)

1.5 컴퓨터 비전 체험 서비스

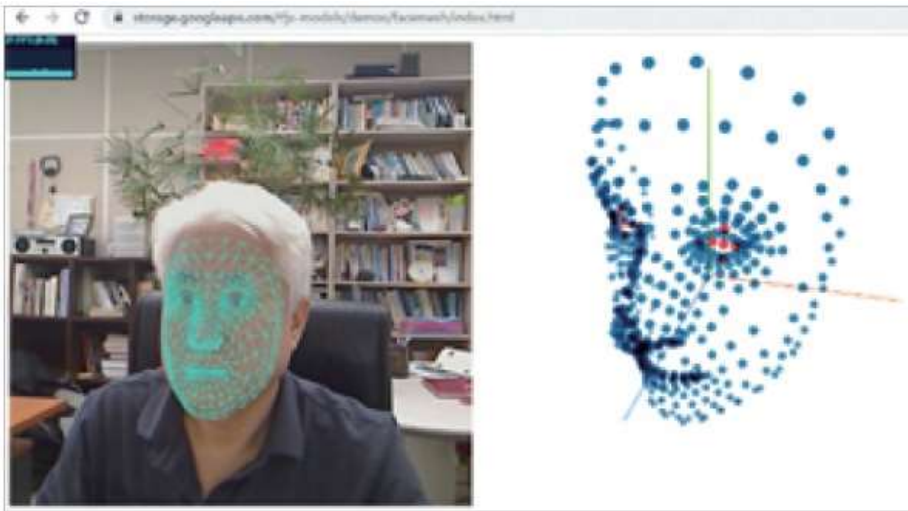


그림 1-10 얼굴 랜드마크 검출



그림 1-11 영상 설명

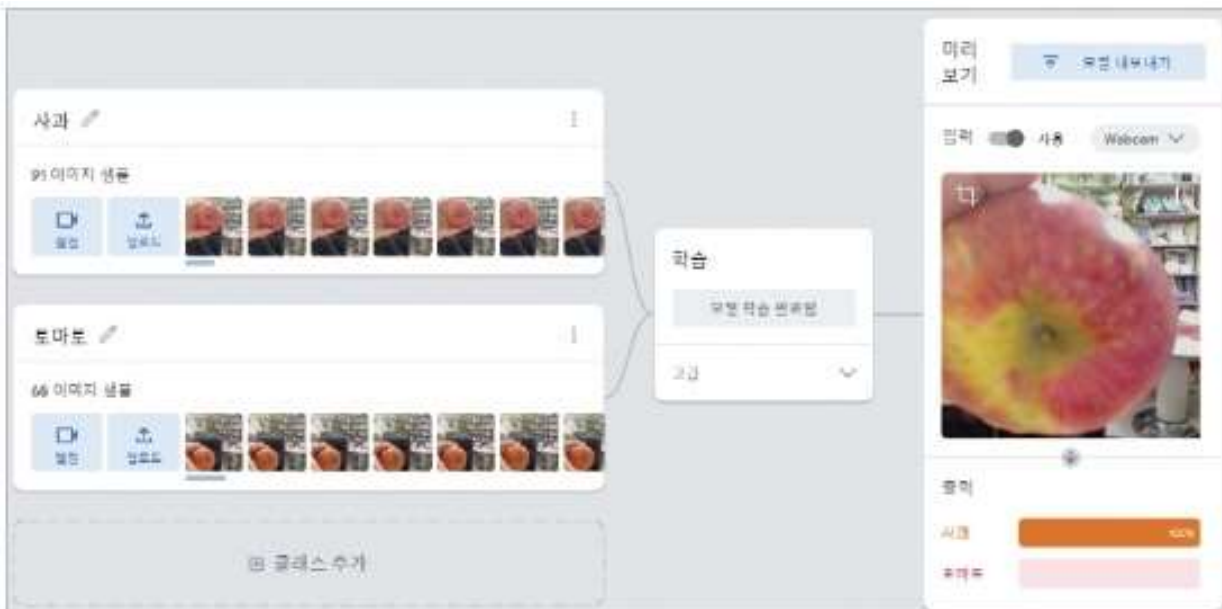


그림 1-12 티처블 머신

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

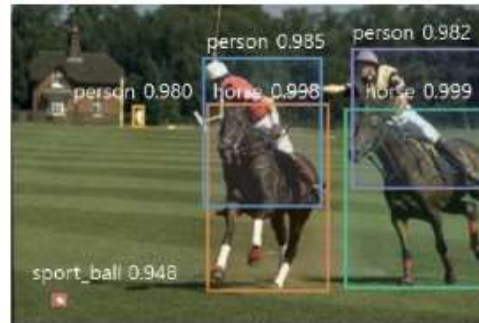
컴퓨터 비전이 풀어야 할 문제

■ 기본 문제

- 분류
- 검출
- 분할
- 추적
- 행동 분석
- ...



(a) 분류



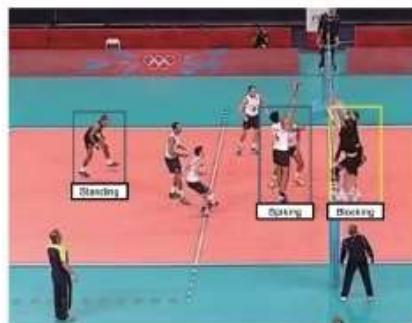
(b) 검출



(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 ([그림 1-11])
- [그림 1-4]의 지식 표현, 추론, 계획과 협업은 매우 소강 상태
- 로봇과 협업은 활발, 예) 비주얼 서보잉을 통한 눈-손 협업

컴퓨터 비전 알고리즘과 프로그래밍

■ 고전 컴퓨터 비전과 딥러닝 컴퓨터 비전

- 대략 2010년을 기점으로 방법론의 대전환. 고전 방법은 규칙 기반, 딥러닝은 데이터 중심
- 이 책의 4~6장은 고전 알고리즘, 7~13장은 딥러닝 위주
- 둘 다 이해하는 것이 중요

■ 프로그래밍 언어

- C/C++, 자바, 파이썬을 주로 사용. 이 책은 파이썬을 사용
- 주로 C/C++는 알고리즘 구현, 파이썬은 인터페이스 언어로 활용됨
- 컴퓨터 비전 지원하는 OpenCV와 딥러닝 지원하는 텐서플로 라이브러리 사용

■ 파이썬

- 배열 처리에 유리
- 파이썬 기초를 다지고 이 책을 공부해야 함

[C 언어]

```
for(i=0; i<n; i++) z[i]=x[i]+y[i];
```

[파이썬 언어]

```
z=x+y
```

1.7 읽을거리와 볼거리

■ 컴퓨터 비전과 딥러닝

- Richard Szeliski 『computer Vision: Algorithms and Applications』 (2nd Edition, 2022)
 - 최신 컴퓨터 비전 내용을 방대하게 담음
 - 무료 pdf 제공 (<https://szeliski.org/Book>)
- Aston Zhang 『Dive into Deep Learning』(2022)
 - 딥러닝에 대한 최신 내용을 충실하게 전달
 - 무료 온라인 교과서. <https://d21.ai>

■ 파이썬 프로그래밍 기초

- 이 책의 부록 A
- 온라인 학습
 - <https://learnpython.org>
 - 코세라 <https://kr.courser.org/courses?query=python>

1.7 읽을거리와 볼거리

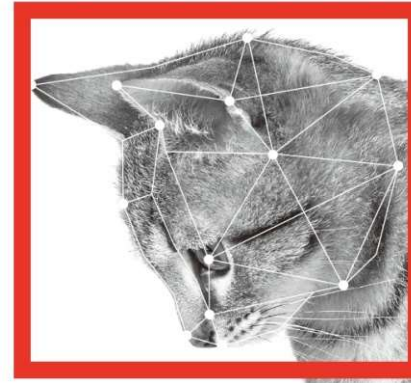
■ TED 강연

- 컴퓨터는 어떻게 사진을 이해하는가? How we're teaching computers to understand pictures?(by FeiFei)
- 컴퓨터는 어떻게 순식간에 물체를 인식할까? How computers learn to recognize objects instantly?(by Redmon)
- 구글 자율주행차 Google's driverless car(by Thrun)
- 자율주행 로봇택시가 거의 왔다 Your self-driving robotaxi is almost here(by Evans)

■ 다큐멘터리

- The Age of AI (시즌 1의 7편은 컴퓨터 비전 활용 이야기)

COMPUTER VISION



DEEP
LEARNING



컴퓨터 비전과 딥러닝

감사합니다.