Facial mask detection 을 활용한 마스크 불량 착용 탐지 및 출입 통제

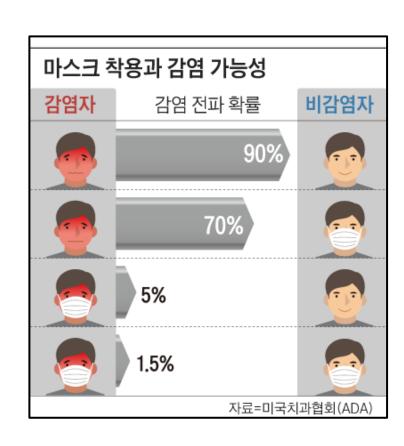


마스크좀쓰시조

박지원 이찬주 김동현 정성훈 이동재

Index

프로젝트 배경 프로젝트 팀 구성 및 역할 02 프로젝트 수행절차 및 방법 03 04 프로젝트 수행 결과



'대구에서의 지난 1차 대유행을 억제하였던 마스크 착용 문화'

깜깜이 환자 통해 전국으로 n차 감염... 2주간 4000명 육박 소미앙

기사입력 2020-08-28 06:00:00 기사수정 2020-08-28 08:02:01



신규확진 33% '경로 깜깜이' 방역당국 통제 대응 역부족

수도권 지방 안 가리고 전방위 확산 사랑제일교회 관련 검사 32% 그쳐 '8·15 집회' 5만명 중 검사자도 16%뿐







'시행 목적이 올바른 착용 유도인만큼 통제 시스템 또한

경고 수준으로 기획'

프로젝트 팀 구성 및 역할

팀원	역할
이찬주	기획안 및 회의록 작성 , 자료 수집, ResNet 설계, 학습 및 성능 평가, ppt 작성
김동현	자료 수집, 마스크 합성하여 데이터셋 만들기, VGGNet 설계, 학습 및 성능 평가, ppt 작성
박지원	Face, eye detection, mask area 추출, AlexNet 모델 학습 및 평가, 예측시각화, 발표
이동재	합성 데이터로부터 mask-area 추출 로직 담당 , thesis reference model 및 InceptionV3 모델 학습 및 평가 , 회의 스케줄링 , 실제 예측 시각화 로직 작성
정성훈	자료 수집 , LeNet-5 모델 설계 , 학습 및 성능 평가 , 예측 시각화, ppt작성

기간	내용
11/12 ~ 11/16	주제 선정
11/17 ~ 11/19	데이터셋 수집 및 학습 데이터 생성
11/19 ~ 11/21	CNN 파생 모델들 학습 및 성능 평가
11/21 ~ 11/23	결과 보고서 작성 및 발표 준비

데이터 정의

조원들의 착용 이미지

전처리 과정에서 이미지 증식을 통한 효과가 상쇄될 것으로 판단.

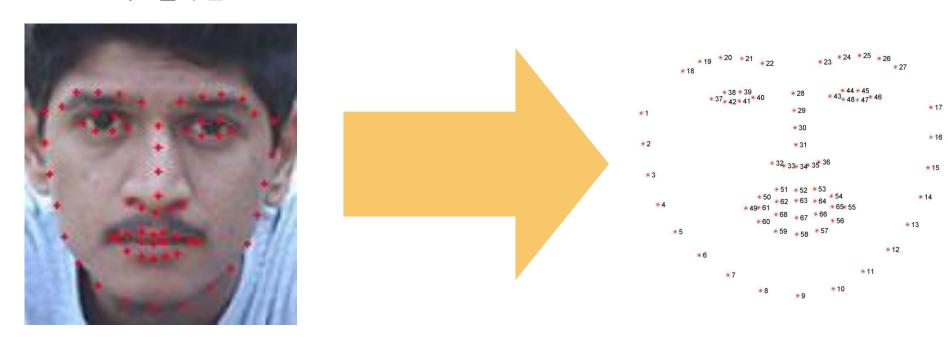
공개 데이터셋으로 생성

Kaggle의 Flicker-Face 데이터셋으로부터 다양한 인종과 각도의 사람 얼굴데이터를 수집, face recognition 패키지로 부터 마스크 합성으로 통하여 프로젝트에 활용할 데이터를 생성

데이터 정의

■ 마스크 합성

Face reconition의 face_landmark를 통해 양쪽 눈썹, 양쪽 눈, 콧대, 코끝, 윗입술, 아랫입술, 턱의 landmark 좌표를 구함.



데이터 정의

■ 마스크 합성 결과 예시





Full mask





Nose mask





Chin mask

03

프로젝트 수행절차 및 방법

논문 소개

Journal of the Korea Academia-Industrial cooperation Society Vol. 21, No. 8 pp. 395-400, 2020 https://doi.org/10.5762/KAIS.2020.21.8.395 ISSN 1975-4701 / eISSN 2288-4688

출입 통제에 활용 가능한 딥러닝 기반 마스크 착용 판별

이승호 한국기술교육대학교 융합학과

Deep learning based face mask recognition for access control

Seung Ho Lee

Department of Future Technologies, Korea University of Technology and Education

요 약 전 세계적으로 유행하며 수많은 확진자와 사망자를 발생시킨 코로나바이러스-19(COVID-19)는 일상에서 사람 간 전염이 가능하여 국민들을 불안과 공포에 떨게 하고 있다. 감염을 최소화하기 위해서는 건물 출입시 마스크 착용이 필수적이지만 일부 사람들은 여천히 마스크 없이 얼굴을 노출시킨 채 건물에 출입하고 있다. 본 논문에서는 효율적인 출입 통제를 위해 얼굴에 마스크를 착용했는지 여부를 자동으로 판별하는 방법을 제안한다. 제안 방법은 양쪽 눈 영역을 검출하고 눈 위치를 참조하여 마스크 착용 영역(양쪽 눈 아래 얼굴 영역)을 예측한다. 이 때 마스크 착용 영역을 보다 착확히 예측하기 위해 양쪽 눈 위치가 수평이 되도록 얼굴 영역을 화전하여 정렬한다. 정렬된 얼굴 영역에서 추출된 마스크 착용 영역은 이미지 분석에 특화된 답러닝 기법인 CNN(Convolutional neural network)을 통해 마스크 착용 여부(착용 또는 미착용)를 최종 판별한다. 총 186장의 테스트 이미지에 대해 실험한 결과, 98.4%의 판별 정확도를 보였다.

Abstract Coronavirus disease 2019 (COVID-19) was identified in December 2019 in China and has spread globally, resulting in an ongoing pandemic. Because COVID-19 is spread mainly from person to person, every person is required to wear a facemask in public. On the other hand, many people are still not wearing facemasks despite official advice. This paper proposes a method to predict whether a human subject is wearing a facemask or not. In the proposed method, two eye regions are detected, and the mask region (i.e., face regions below two eyes) is predicted and extracted based on the two eye locations. For more accurate extraction of the mask region, the facial region was aligned by rotating it such that the line connecting the two eye centers was horizontal. The mask region extracted from the aligned face was fed into a convolutional neural network (CNN), producing the classification result (with or without a mask). The experimental result on 186 test images showed that the proposed method achieves a very high accuracy of 98.4%.

Keywords: Face Mask Detection, Deep Learning, Convolutional Neural Network, Eye Detection, OpenCV

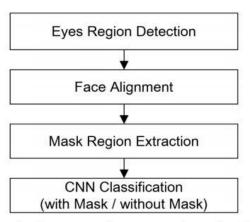
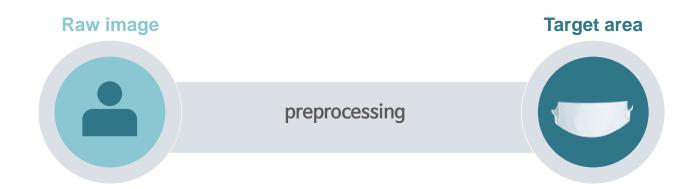


Fig. 1. Block diagram of proposed method.

출처: 한국산학기술학회 논문지 21(8), 2020.8, 395-400(6pages)

데이터 전처리

합성된 이미지로부터 전처리 과정을 거친 후 모델의 입력으로 활용



데이터 전처리 과정

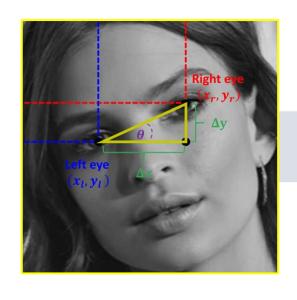


haarcascade_eye.xml (일반 검출)

haarcascade_eye_tree_eyeglasses.xml (안경 눈 검출)

haarcascade_lefteye_2splits.xml (측면 각도 검출)

▶ 데이터 전처리 과정

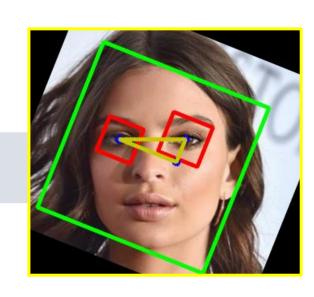


Face Alignment

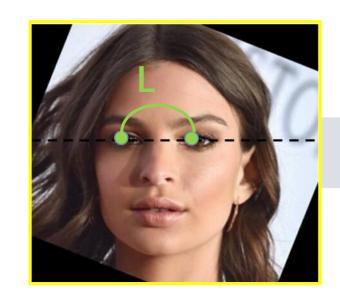
$$\Delta x = x_r - x_l$$

$$\Delta y = y_r - y_l$$

$$\boldsymbol{\theta} = \arctan \frac{\Delta y}{\Delta x}$$



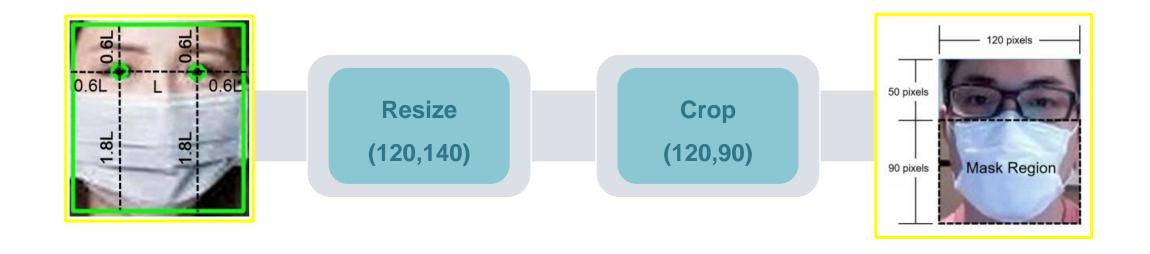
데이터 전처리 과정



Face area



데이터 전처리 과정



데이터 전처리 결과 예시















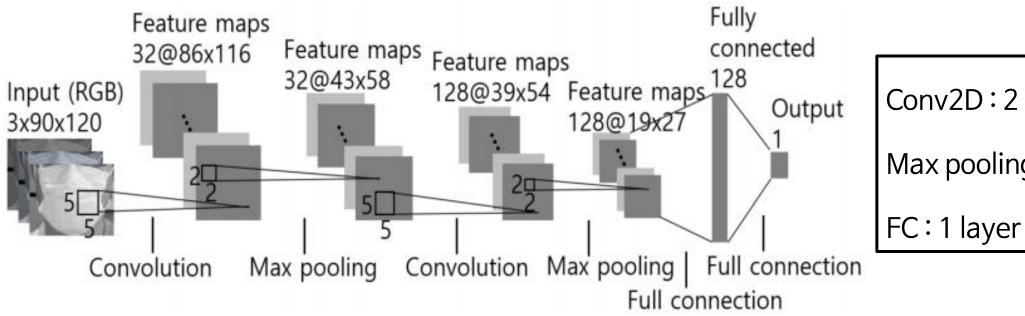


모델 생성 및 학습

후보 모델	Epoch	Train:Val:Test
Basic CNN (thesis reference)		
InceptionV3		
ResNet	50	7:2:1
LeNet	30	/ •∠• 1
AlexNet		
VGG16		

Basic CNN (thesis reference)

참고논문의 모델 아키텍쳐를 바탕으로 basic CNN 모델을 구성하였음



Conv2D: 2 layers

Max pooling: 2 times

Basic CNN (thesis reference)



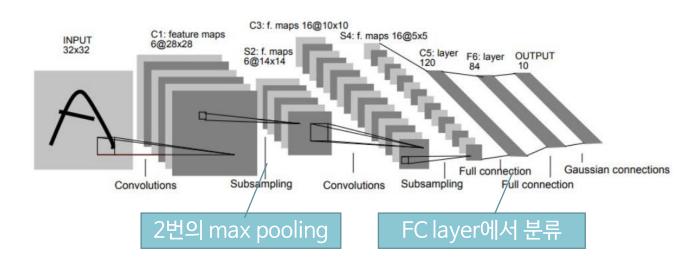
Train ACC	Validate ACC	Test ACC
0.9696	0.9613	0.9600

간단한 구조임에도 좋은 성능을 보임

이미지 데이터의 특징이 뚜렷

LeNet

LeNet - 5 Architecture:



Conv2D: 2 layers

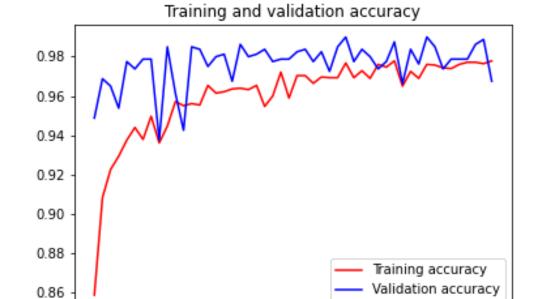
Max pooling: 2 times

FC: 3 layer

CNN의 개념을 처음으로 도입한 Yann LeCun이 만든 고전적인 network으로, 기존 딥러닝

모델의 기울기 손실 문제를 극복하고 기하급수적인 연산량을 해결

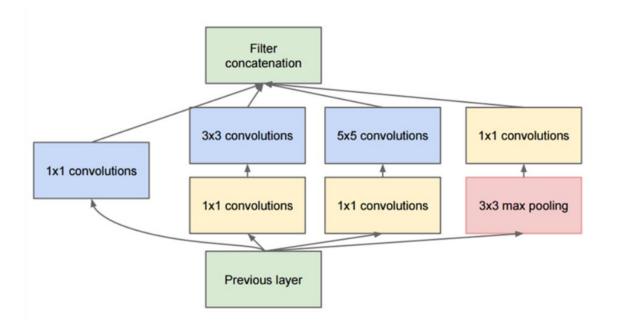
LeNet



Train ACC	Validate ACC	Test ACC
0.9779	0.9675	0.9725

Evaluate 결과에서 약 97.2%의 정확도를 보인다.

InceptionV3

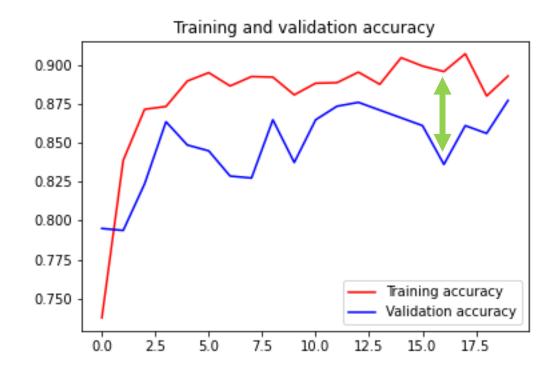


2014년 처음 제안되어 현재 V4 까지 개량된 구글의 CNN 모델

다른 기존 CNN모델들과 다르게 한번에 여러 크기의 필터를 동시에

사용하여 다음 층으로 정보를 전달할 때 concat 과정이 존재

InceptionV3



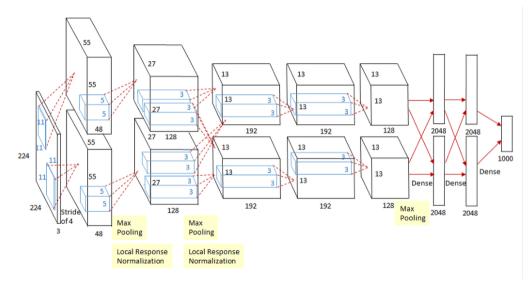
Train ACC	Validate ACC	Test ACC
0.8932	0.8775	0.9150

환경상 학습을 온전히 마치지 못함.

오히려 간단한 이미지라 복잡한 구조의

CNN 모델에서는 학습이 잘 되지 않음

AlexNet

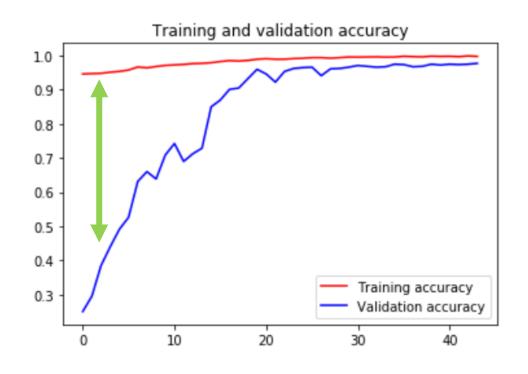


AlexNet은 2012년도 ImageNet 우승을 차지한 CNN 모델

LeNet-5와 비슷한 기본구조를 가지며 2개의 병렬적인 모델 구조로 설계되었다는 점이 특징

- * Tanh 함수 대신에 ReLU 함수 사용
- * dropout 적용

AlexNet

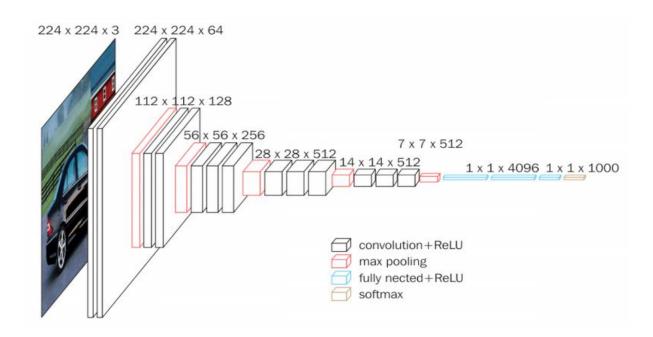


Train ACC	Validate ACC	Test ACC
0.9979	0.9775	0.9675

loss: 0.0098 - accuracy: 0.9979 - val_loss: 0.1112 - val_accuracy: 0.9775

에폭이 증가할수록 점차적으로 validation 정확도가 증가하는 것을 확인할 수 있음 test set에 대해 96.75%의 정확도를 보임.

VGGNet

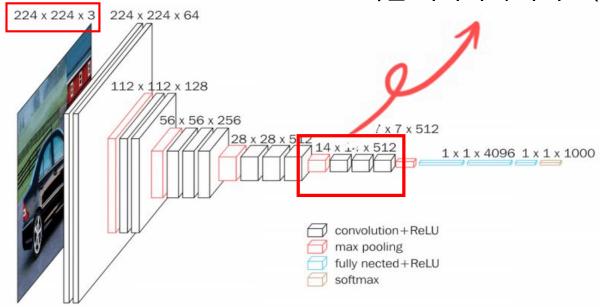


Conv kernel size = 3, padding = 1 Pooling kernel size = 2, stride = 2 네트워크의 깊이(depth)가 어떤 영향을 주는지 연구를 하기 위해 설계된 CNN model

구조가 간단하고 단일 네트워크에서는 좋은 성능을 보이지만 최종단의 세 FC 층의 파라미터 수가 많다는 것이 단점

VGGNet

학습 이미지의 사이즈(90x120x3)를 고려하여 모델의 복잡도를 줄여 사용

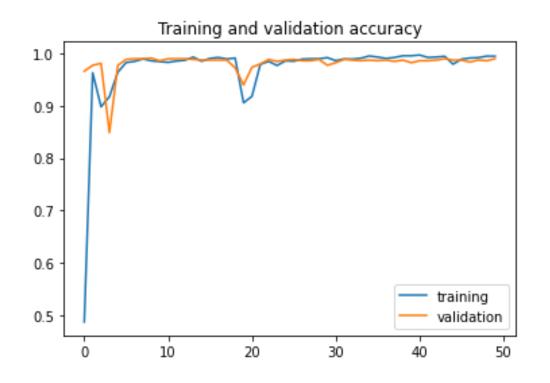


Conv2D(512) layer 3개: 제거

pooling:5->4번

FC 유닛: 4096 - > 512

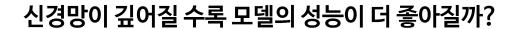
VGGNet

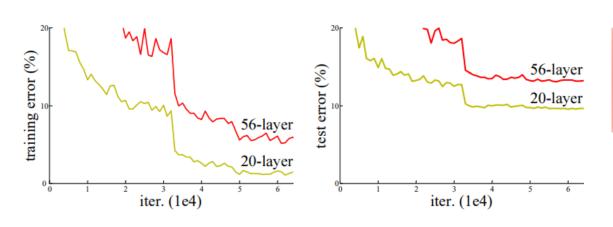


Train ACC	Validate ACC	Test ACC
0.995	0.99	0.985

Evaluate 결과에서 약 98.5%의 정확도를 보임. 학습 데이터가 단순하여 좋은 성능을 보이는 것 같음.

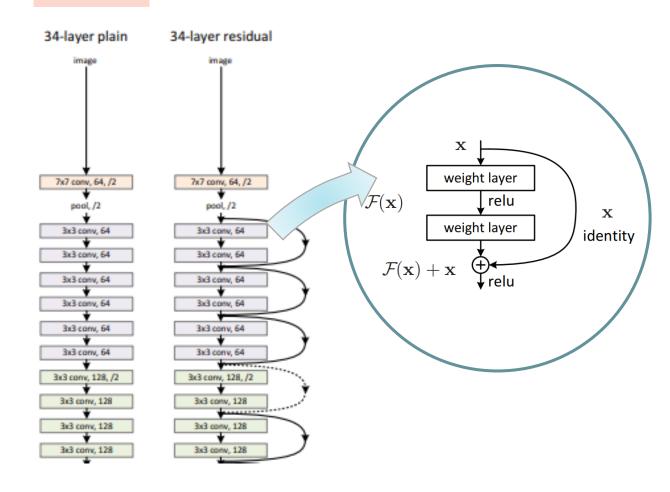
ResNet





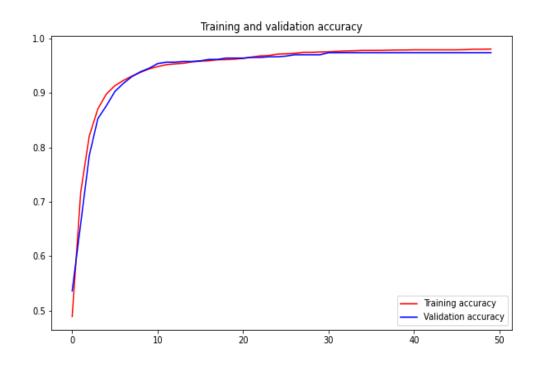
Layer가 깊어질수록 기울기 손실발생 =〉 성능 저하 발생

ResNet



입력을 layer의 출력에 바로 연결시키는
"skip connection"을 사용하여 나머지를
학습하게 하는 특징을 가지고 있음

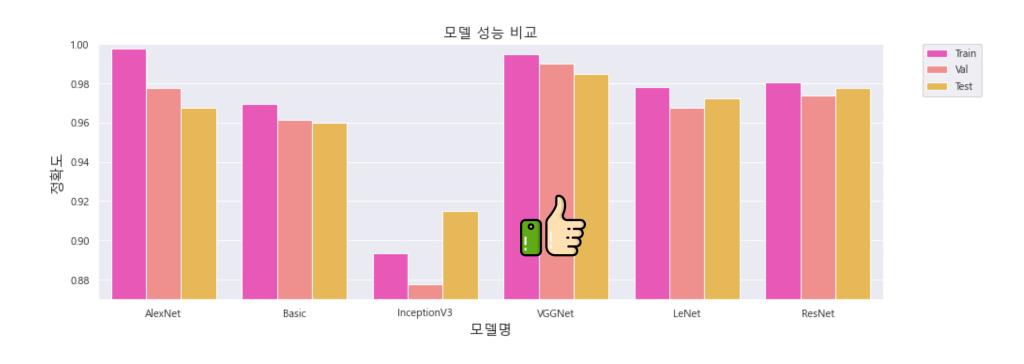
ResNet



Train ACC	Validate ACC	Test ACC
0.9804	0.9737	0.9775

Evaluate에서 약 0.97 정도의 정확도를 보였음. 에폭이 늘수록 Loss가 꾸준히 줄어드는 점이 인상깊었음

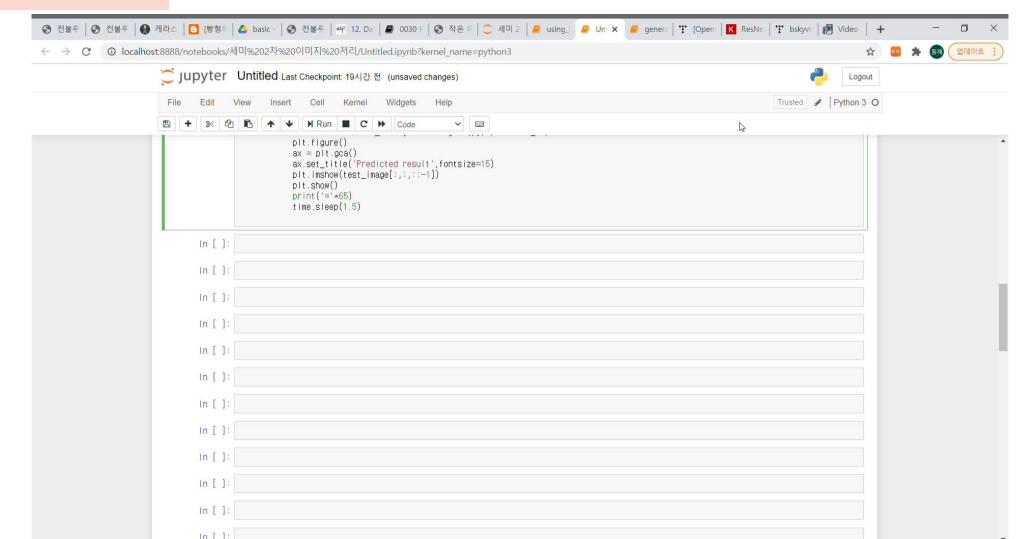
모델 성능 비교



최종적으로 VGG16 개량 모델이 가장 이상적인 모델로 확인

<u>프로젝트</u> 수행 결과

실제 예측 영상



프로젝트 개선 방안

- 1.동영상이나 웹캠을 통해 실시간 마스크 착용 여부 판단
- 2. 흰 마스크 뿐만 아니라 다른 색의 마스크들 까지 구별할 수 있도록 구현
- 3. 데이터셋을 좀 더 늘려서 학습해보기 : 전체 사진의 개수, 동양인 추가
- 4. 실제 사람이 마스크를 쓰고 있는 데이터셋 구축
- 5. 얼굴 측면 사진의 마스크 착용 여부 판단

느낀점









데이터 수집이 문제였는데, 팀원이 마스크 합성을 제안하였고, face landmark를 이용하여 마스크 합성에 성공, 원하는 데이터셋을 만들어 낼 수 있었기에 모델 학습, 성능 평가까지 큰 걸림돌 없이 잘 수행할 수 있었던 것 같다.

VGGNet을 학습하는데 시간이 너무 오래 걸려서 답답함을 많이 느꼈지만, 수행결과를 보니 왜 VGGNet이 많은 network에서 응용되고 있는지 알 것 같다.

LeNet 모델은 CNN모델 중 가장 기본적인 모델이지만 LeNet-5 모델까지 발전했습니다. 다양하게 발전된 모형들이 등장하고 있지만 feature의 갯수가 적고 확실하다면 고전적 모델이 보다 효과적이라는 것을 알았습니다. 학습layer가 많이 깊지 않았지만 학습시키는데 시간이 생각보다 오래 걸리는 것을 보면서 깊은 layer의 모델의 학습에는 좋은 성능의 GPU가 필수적 이라는 것을 다시 한번 느끼게 되었습니다.

수업을 들을 때는 이미 만들어진 모델에 데이터 학습만 시키면 예측과 분류가 가능하구나, 쉽다고 생각했습니다. 실제로 해보니 역시 세상엔 쉬운 게 없구나 다시한번 느낄 수 있는 프로젝트였습니다. 이미지넷 대회에서 우승하였던 여러 CNN모델들을 활용하여 학습하고 서로 성능을 비교하는 과정이 뜻 깊었습니다. 기간이 짧고 또 여건 상 더 좋은 환경에서 학습을 진행할 수 없어 아쉬운 점이 있지만 다들 고생하였습니다. 프로젝트를 진행하면서 컴퓨터 비전 분야에 정말 많은 연구가 진행되었고, 응용이 많이 되고 있구나를 체감할 수 있었다. ResNet을 이 프로젝트에 적용시키기가 까다로웠어서 많이 헤멨다. 사전 학습 모델에 input shape를 맞추고, 마스크 분류를 위한 layer를 추가하는 것이 어려웠다. 그리고 train, validation 결과가 이상하게 나와서 조정하는 것이 어려웠다. 하지만 나중에 ResNet을 적용하고 보니, 생각보다 모델 학습 속도가 빠르고 정확도가 빨리 나와서 신기했다.

Q&A