

# COVID-19 Risk Estimating : Crowd Density and Face Mask Recognition using DNN

## 요약

최근까지도 이어지고 있는 COVID-19 의 위험에 대비하는 가장 좋은 방법은 예방이다. 이를 위해 한국질병관리청에서 정한 방역 수칙을 기준으로 하여 현 상황에 대한 COVID-19 Risk Estimation 을 진행한다. 위험도를 측정하기 위한 세 가지 기준은 군중 밀집도, 마스크 착용 여부, 공간 개폐 여부로 한다. 각각의 기준에 Deep Learning 알고리즘을 적용하여 값을 산출해내고 질병 전파 모델로 COVID-19 감염 위험도를 계산한다. 연구 결과를 CCTV 와 같은 Live Feed 영상에 적용해 위험도를 알리고 조치를 취해 예방에 기여할 수 있도록 한다.

## 1. 서론

### 1.1. 연구배경

2020 년 3 월 11 일 WHO 는 COVID-19 pandemic 을 선언하였다. 그 이후로 2020 년 10 월이 된 현시점까지 COVID-19 에 대한 위협이 사그라지지 않고 있다. 백신도 개발되지 못한 현 상황에서 예방만이 COVID-19 에 대비한 최선의 방법이다. 국가에서는 이를 예방하기 위해 마스크를 착용하고 손을 자주 씻고 거리 두기를 생활화하라고 얘기하고 있다. 이 중 마스크 착용은 의무화하였고 수도권은 사회적 거리 두기 2 단계가 시행 중이다. 하지만 이러한 노력에도 불구하고 아직도 확진자 수가 안정적인 상태에 접어들지 못하였다.

아직까지는 COVID-19 에 관련하여 확진자의 동선을 확인할 수 있는 서비스를 제외하면 감염을 예방하기 위한 서비스가 제공되고 있지 않다. 확진자가 다녀간 동선을 확인하는 것은 최선의 예방이 아니다. 따라서 사람들의 경각심을 일깨우고 위험도가 높은 장소나 상황을 빠르게 파악하여 대처하기 위해 감염 위험도를 가시화하였다.

COVID-19 의 위험에 얼마나 노출되어 있는지를 수치화하기 위해 세 가지 기준을 정하여 계산하였다. 공간이 개방된 공간인지 혹은 밀폐된 공간인지가 그 첫 번째다. 밀폐된 공간이 개방된 공간보다 위험도가 높아지도록 한다. 두 번째는 얼마나 많은 사람이 밀집되어 있는지이며 많은 사람이 모여있을수록 위험도가 올라간다. 마지막으로 마스크를 착용하고

있는지 아닌지를 판단하는 것이다. 마스크를 착용하지 않은 사람들이 많아질 수록 위험도가 상승하게 된다.

Deep Learning 알고리즘을 통해 각각의 상황을 학습하여 모델들을 만들어 CCTV와 같은 실시간 영상에서 사용될 수 있도록 한다. Real Time으로 적용하여 CCTV 영상에서 각 기준에 대한 결괏값을 보여주고 이 결과들을 이용하여 계산된 위험도도 같이 영상을 통해 볼 수 있도록 서비스를 제공한다.

## 1.2. 연구목표

한국 질병 관리청에서 정한 방역 수칙을 참고해 현재 감염 위험도가 얼마나 있는지 계산하는 방법은 확진자가 다녀간 동선을 파악하는 것과 확진이 된 후 처리하는 것이 아닌 사전에 차단하도록 도와준다는 점에서 차이점이 있다.

만약 그 공간에 있는 사람들에게 위험도를 보여주게 된다면 스스로 경각심을 가지고 거리를 더 유지하도록 노력하고 마스크를 안 쓴 사람들에게도 더 큰 부담감을 느끼게 할 수 있다.

혹은 위험도가 일정 값 이상으로 올라가게 될 시 그곳의 담당자나 관리자가 강제로 위험도를 낮추기 위해 조치를 취할 수 있게 된다. 특히 음식점이나 카페와 같은 곳에서는 지금 상황이 얼마나 위험한지를 수치로 보여주는 것이 더욱 유용하게 사용된다. 그 이외의 공공장소에서도 이를 이용해 감염 예방 수칙을 이행하기가 수월해진다.

궁극적으로 위험도를 낮춤으로써 COVID-19 감염을 효과적으로 예방하는 것이 목표이다.

## 2. 관련연구

### 2.1. 코로나 위험도 산정 참고 모델

Covid-19의 위험도 산정을 위해서 크게 두 가지의 접근 방식을 사용할 수 있다. 첫번째로는 한국질병관리청에서 정한 방역 수칙을 참고해 해당 수칙이 제대로 지켜지는지에 대한 척도를 위험도 산정에 사용할 수 있다. 두번째로는, Covid-19나 이를 포함하는 이미 알려진 공기중으로 전파가 되는 전염병이 어떤 환경에서 잘 전파가 되는지에 대한 연구를 참고해 실제로 어떤 조건에서 전파가 잘 되는지 참고해 위험도 연산을 하는 방식이 있다. 첫번째 방식은 연산이 단순하고 어떤 수칙이 지켜졌느냐, 지켜지지 않았느냐에 대한 점수로 출력을 한다. 하지만 이 숫자는 그저 군중의 행동을 분석하는 숫자일 가능성이 높고, 숫자에 대한

평가가 다시 이루어져야 한다는 한계점이 있다. 반면, 두번째 방식은 고려해야할 매개변수가 많고, 복잡한 식을 가지므로 연산이 힘들다. 연산 과정이 복잡한 만큼 현재 상황에 대한 분석이 조금 더 잘 되어 출력된 값을 바로 위험도에 대한 척도로 활용할 수 있다.

## 2.2. 군중 밀집도

“COVID-19” 바이러스는 주로 사람 간에 전파 된다고 여겨지고 있으며, 감염자와 밀접 접촉자들 간, 감염자의 호흡기를 통해 나온 비말 등이 주위 사람들의 입이나 코에 닿거나 호흡기를 통해 흡입되면 전파될 수 있다고 알려져 있다.[hj\_ref\_1] 이에 코로나 전파 위험도를 평가할 경우, “군중 밀집도”, “사람들 간의 거리”는 중요한 평가 요소가 된다.

### 2.2.1. Object Detection

“군중 밀집도”를 평가하기 위해서는 먼저 해당 Image 속에 어떤 물체들이 존재하는지 파악해야 한다. 이 경우 현재 CCTV, 자율주행 자동차 및 OCR 등의 분야에서 많이 활용되고 있는 “Object Detection” 기술이 요구된다. Image 속 Object 들이 종류인지 분류(Classification), 해당 Object 의 위치 정보를 파악한다(Localization).

### 2.2.2. Crowd Count(Crowd Density)

“군중 밀집도”는 Image 속에 있는 물체들 중 사람에 의해서만 평가되며, 그 중에서도 사람들의 수, 사람들 간의 위치, 거리가 중요한 평가 요소로 작용한다. Crowd Count(Crowd Density)는 Image 속 사람의 수를 세거나 예측, 해당 Image 속에 단위 면적별 사람들의 밀집도를 추정하는 분야다. Crowd Count(Crowd Density)을 측정하는 방식으로는 Detection-based methods, Regression-based methods, Density estimation-based methods, CNN-based methods 등이 있으며, 위 방식을 input 이 되는 Dataset 의 환경에 맞게 선택한다.[hj\_ref\_2][hj\_ref\_3]

### 2.2.3 Social Distancing

Social Distancing(사회적 거리두기)은 비말에 의해 전파될 수 있는 질병에 경우 강력한 예방책 중 하나로 여겨지고 있다. 서로 물리적으로 거리를 두어 접촉을 줄이는 것만으로도 전염병의 확산을 크게 줄일 수 있다. Social Distancing 분야는 사람 간의 거리를 Image 속 픽셀의 단순 거리로 계산하는 것이 아니라, 2D Image 의 사람들의 위치를 3D 로 인식하여, 사람들의 원근에 따라 다른 계산식(Perspective Mapping)을 적용 및 측정한다.

## 2.3. 마스크 얼굴인식

마스크 착용 여부는 비말 감염으로 주로 전파되는 "COVID-19"예방에 있어서 큰 영향력을 갖는다. 비말 차단 기능을 갖는 마스크의 착용은 바이러스를 직접 차단하거나 손과 얼굴의 접촉을 방지함으로써 전반적으로 바이러스 감염 확률을 대폭 낮출 수 있다. 이에 따라 "마스크 착용 여부"를 코로나 감염 위험도 산출 요소로 포함한다.

### 2.3.1 Object detection

마스크 착용 여부를 탐지하기 위해 우선적으로 Object detection 모델을 정한다. 기본적으로 실시간 탐지를 목표로 하기 때문에 One-shot-detection 기법을 사용하는 알고리즘을 사용해야 할 것이다. SSD 와 Yolo 는 One-shot-detection 기반의 대표적 알고리즘 모델이다.

### 2.3.2 Yolo(You Only Look Once)

최초로 One-shot-detection 기법을 사용한 알고리즘으로, Object Localization 단계와 수많은 bounding box 를 모두 CNN 으로 처리하여 Object Classification 을 별도의 과정을 통해 처리하는 Two-shot-detection 기법과 달리 위의 과정을 한 번에 처리한다.

Yolo 는 feature map 을  $S \times S$  grid cell 로 나누고 각 cell 에 지정된 개수만큼의 Anchor Box 를 포함시키는데, 각각의 Anchor Box 는 상자 안에 물체가 포함될 확률, box 의 위치( $x, y, w, h$ ), detecting 하고자 하는 class 들의 일치 가능성을 파라미터로 갖게 된다. 이후 학습을 통해 물체가 포함될 확률이 낮은 box 들을 없애고 최종적으로 Object 들을 탐지, 분류한다.

### 2.3.3 SSD(Single Shot Multibox Detector)

임의로 정해진 셀의 크기에 따라 detection 가능한 물체의 크기가 한정되는 Yolo 와 달리 convolution 과정의 초기부터 classification 을 실행한다. 따라서 셀의 개수가 많은 초기의 Feature Map 에서는 작은 크기의 물체를, convNet 을 거치면서 점점 더 큰 물체를 대상으로 탐지하는 것을 목표로 하며 학습을 통해 각각의 Classification 들을 종합하여 최종적인 결과를 도출한다.

## 2.4. 공간 개폐

Covid-19 의 위험도 산출에 있어서 환경은 큰 factor 를 차지한다. 밀폐된 실내의 경우 실내 공기를 다수의 인파가 계속해서 호흡하고, 이 때 감염자가 있을 경우 감염될 가능성이 매우 높다. 따라서, Covid-19 Risk Estimator 에서 산출하는 이미지의 공간이 어떤 유형인지 감지하는 것은 꼭 필요하다고 할 수 있다. 이미지가 어떤 공간인지 감지하는 모델은 옛날부터 많은 연구가 진행되었다[sm\_ref\_5]. 옛날부터 연구가 진행된 만큼 모델도 많은 수가 있고 Dataset 도 많다[sm\_ref\_6][sm\_ref\_7]. 대체로 이미지를 input 으로 받는 만큼 Convolution 연산을 사용하는 CNN 모델 구성이 많고, 모델 별 각기 다른 Dataset 을 사용하다가 근래 들어 Places Dataset 을 사용하는 경향이 있다. Places Dataset 은 각 Scene 별 특징도 label 되어있어 Indoor/Outdoor 분석뿐만이 아니라 Scene Detection 에도 적합한 Dataset 이다. 따라서, 공간 개폐의 경우, 모델의 성능 및 연산 속도에 대해 비교해보고 적합도가 가장 높은 모델을 사용할 수 있다.

### 3. 프로젝트 내용

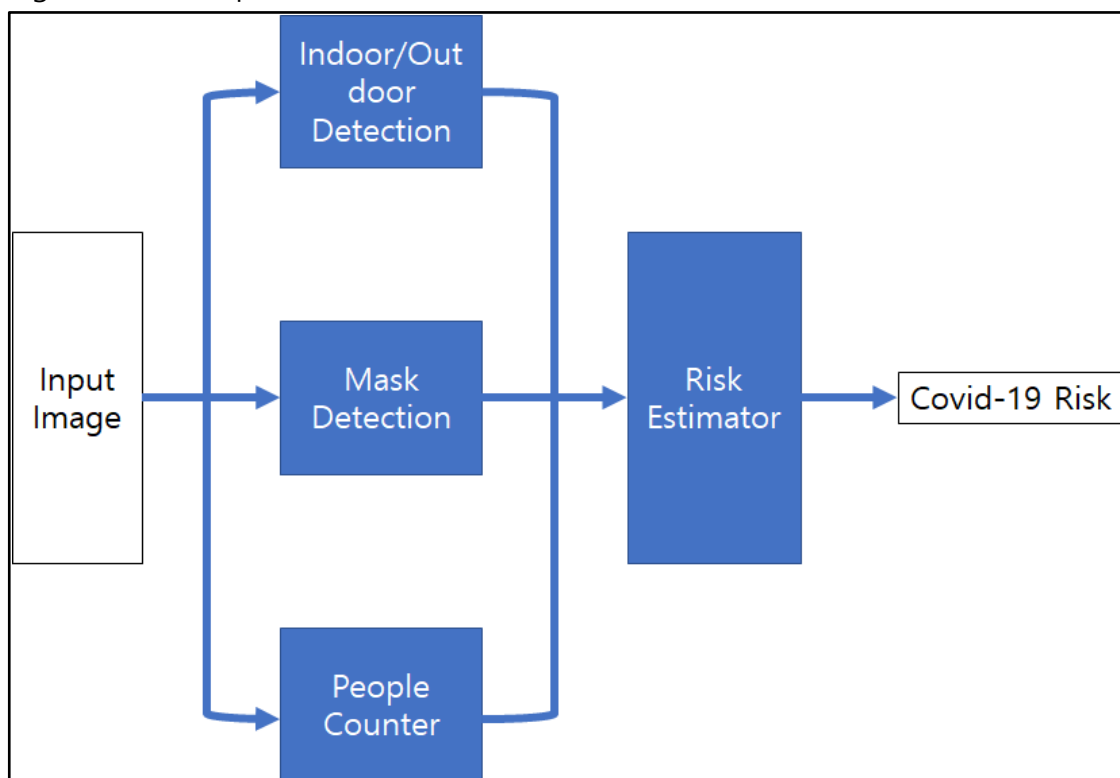
#### 3.1.시나리오(코로나 위험도 산출)

##### 3.1.1 Input

모델의 Input 으로는 Monocular RGB 이미지를 사용한다. 모델이 작동되는 연산장치의 성능과 실제 구현이 어떻게 되는지에 따라 Realtime 적용의 유무가 결정될 예정이다. 모델 설계 과정에서 실시간 처리가 가능한 가벼운 모델을 목표로 설계 예정이다. Realtime 적용이 가능한 상황에서는 CCTV 나 Drone 에서 받아오는 Live Feed 영상을 Input 으로 하여 응용이 가능하다.

##### 3.1.2 산출 과정 및 output

산출 과정은 아직 모델 설계 과정에 있기 때문에, 변동 가능성이 있음을 미리 고지한다. Image 가 모델의 input 으로 들어오면 3 개의 모델이 이미지를 처리한다.



##### 3.1.3 Indoor/Outdoor Detection

모델의 Input 이미지에 대해 이미지가 실내의 이미지인지, 실외의 이미지인지 연산한다. Places365 Dataset 으로 학습된 모델을 사용해 이미지의 공간적 특징을 추출하게 된다. 현재

Places365 Dataset 으로 학습된 Wideresnet 을 사용해 구현 되어있는 상태이다. 해당 모델은 Indoor/Outdoor 외의 해당 Image 의 Scene 의 분석 값(카페, 식당 등)을 출력하는데, 현재는 Indoor/Outdoor 값 만을 활용하도록 작성되었다. 실제 실행 결과, GPU 가속을 활용했을 때 Image 당 0.2 초 미만의 연산 속도를 가져 Realtime 으로 충분히 운용이 가능하나, Camera 가 고정되어 있음을 감안할 때 Image Feed 의 첫 몇 초만 모델의 연산을 진행하고 Image Feed 가 Indoor 인지 Outdoor 인지 출력하고, 해당 결과값을 작동하는 Session 동안 유지하여 연산 장치의 부하를 최소화해 다른 모델이 연산하는데 필요한 Computing Power 를 보장한다.

### 3.1.4. Mask Detection

모델의 Input 이미지에 대해 이미지에서 마스크를 착용한 사람과 그렇지 않은 사람들을 감지한다. Realtime 을 목표로 하고 있기 때문에, 연산 속도가 빠른 YOLO 모델을 사용한다. 마스크를 쓴 얼굴과 그렇지 않은 얼굴의 이미지들을 Train 한 YOLO 모델은 마스크를 착용한 얼굴과 착용하지 않은 얼굴을 감지하고 그 수와 위치를 연산한다. 해당 출력을 다음 단계로 전달하고 계속해서 이미지의 처리를 진행한다.

### 3.1.5. People Counter

모델의 Input 이미지에 대해 이미지에서 사람들의 위치를 감지한다. 역시 Realtime 구현을 감안, YOLO 모델을 사용한다. 사람을 감지하는 모델은 이미 MS COCO Dataset 으로 학습되어 공개된 모델을 사용할 수 있다. 감지된 Object 들 중 Class 가 Person 인 결과값만 사용해 위치와 개수, 크기를 구하고, 크기 값과 위치 값을 사용해 사람이 어느 거리를 두고 있는지를 연산 할 수 있다. 역시 해당 출력들을 다음 단계로 전달하고 계속해서 이미지 처리를 진행한다.

### 3.1.6. Risk Estimator & Output

앞의 3 개의 모델에서 출력한 결과들을 사용해 Covid-19 의 감염 Risk 를 연산한다. 질병 전파 모델로 유명한 SIR 모델을 베이스로 하여, 마스크 착용 및 거리를 추가한 우리만의 연산 척도를 개발 예정이다. SIR 모델은 감염가능한 사람 수인 S, 감염자의 수인 I, 치료가 완료된 사람 수인 R, 이렇게 세 개 값의 인과를 감염률인  $\beta$ , 회복률인  $\gamma$  을 사용해 설명하는 모델이다. 이 모델을 활용해, 많은 논문들이 Covid-19 에 대한 감염률  $\beta$  를 연산했다.[sm\_ref\_9] 우리는 위의 세 모델을 사용해 사람과 사람이 접촉했을 경우를 판단해 어떤 사람이 감염되었을 경우

해당 이미지에 얼마의 사람들이 Covid-19 에 감염되었을지를 예측하려고 한다. 모델은 사람과의 거리, 마스크의 유무, 공간의 형태에 따라  $\beta$  값을 조정할 수 있고, 이  $\beta$  값에 따라 감염률을 계산한다. 물론 질병관리청에서 정한 사회적 거리두기가 얼마나 지켜지고 있는지에 대한 정량적 평가만을 결과로 도출하는 모델 역시 검토 대상이다.

## 3.2. 요구사항

### 3.2.1. 군중 밀집도에 대한 요구사항

군중 밀집도를 측정하기 위해서 사람들이 밀집되어 있거나, 흩어져 있는 경우 또는 사람이 있지 않는 등 다양한 상황의 Image Dataset 을 학습시킨다. 학습에 사용할 DNN 알고리즘 모델은 처리 속도가 우수하여, Object Detection 분야에서 많이 사용되고 있는 Yolo 를 기반으로 진행한다.

학습이 완료되면 해당 모델을 기반으로 위험도를 측정할 Input Data 를 넣고, 해당 Data 를 분석한다. 화면에 존재하는 물체들 중 사람들을 선별하고, 사람들의 수, 각 위치 및 위치별 원근에 따라 사람들 간의 거리를 계산하여, 해당 화면의 밀집도를 결과 값으로 반환한다.

### 3.2.2. 마스크 얼굴인식에 대한 요구사항

마스크 착용 여부를 탐지하기 위해 마스크를 착용한 사람의 얼굴과 마스크를 착용하지 않은 사람의 얼굴에 대한 data 를 각각 학습시킨다. 이후에 학습된 모델을 Object detection 알고리즘에 적용하여 화면에 나타난 모든 사람을 output 상에 with Mask, with no Mask 로 분류하여 labeling 되어 나타내도록 한다.

### 3.2.3. 공간개폐에 대한 요구사항

공간 개폐를 감지하기 위해 Places365 Dataset 으로 학습된 WideResnet Weight 를 활용한다. 해당 Weight 를 사용한 모델로 Input 의 Scene Detection 을 수행하고, 결과값으로 나온 label 들 중 Indoor/Outdoor Label 만을 추출해 Input 이미지가 실내인지 실외인지 출력하도록 한다.



#### 4. 향후 일정 및 역할분담

배성호 교수 : 프로젝트 면담 진행

신은섭 멘토 : 멘토링 진행

공재호 : PM, 군중 밀집도 분석

이민호 : 마스크 얼굴인식 분석

이승민 : 공간 개폐 분석

지호진 : 군중 밀집도 분석

최성원 : 마스크 얼굴인식 분석

Task	진행일자
팀원 구성	9/11 ~ 9/18
주제 선정	9/18 ~ 9/21
참고자료 조사	9/22 ~ 9/28
기초조사서 작성	9/29 ~ 10/3
오픈소스 코드 분석	10/3 ~ 10/5
설계	10/6 ~ 10/12
논문 작성	10/13 ~ 10/19
설계서 작성	10/20 ~ 10/26
상세 설계	10/27 ~ 11/2
상세 설계서 작성	11/3 ~ 11/9
구현	11/10 ~ 11/24
검증	11/24 ~ 11/30
결과보고서 작성	12/1 ~ 12/7
데모 시연 및 발표	12/8 ~ 12/14

## 5. 결론 및 기대효과

영상 정보를 통해 해당 공간의 COVID-19 위험도를 평가할 수 있다. 영상 정보 내에서 세 가지 정보를 수집한다. 첫째, 군중 밀집도를 수집한다. 군중 밀집도는 해당 공간에 사람들이 얼마나 밀집되어 있는지를 계산한다. 둘째, 사람들의 마스크 착용 여부 정보를 수집한다. 현재 우리나라에서 마스크 착용은 권장사항이 아닌 의무사항이며 감염자로부터 자신을 지킬 수 있는 수단이므로 중요하다. 셋째, 해당 공간이 카페, 도서관과 같은 내부 환경인지 공원, 인도와 같은 외부 환경인지에 대한 정보를 수집한다. 한국질병관리청에서 제시한 방역 수칙에 의하면 외부 환경보다 내부 환경이 더 위험한 환경이므로 이는 중요한 정보가 된다. 수집한 군중 밀집도와 마스크 착용 여부 정보에 대한 위험률을 점수화한다. 만약 내부 환경이라면 위험도 가중치를 올린다. 각 점수화된 내용을 종합하여 위험도를 산출한다. 산출된 위험도를 사용자에게 알려 해당 공간의 위험 수준을 파악할 수 있다.

COVID-19 사태에서 사람들 간의 사회적 거리두기는 중요한 사항이 되었다. 따라서 본 연구는 해당 공간의 위험도를 통해 사회적 거리두기가 잘 이행되고 있는지 확인할 수 있다. 또한, 확인된 위험도를 통해 필요할 조치들을 취할 수 있다. CCTV와 같은 다른 영상 정보에 활용하여 해당 공간 사람들에게 위험도를 알릴 수 있다. 이는 코로나 감염 예방에 도움이 된다. 다른 코로나 관련 개발 또는 연구에도 활용 가능하여 코로나 사태 예방에 기여할 수 있다.

## 6. 참고문헌

- [1] CDC(Centers for Disease Control and Prevention) : 질병 예방 - 전파 경로
- [2] Yuhong Li, Xiaofan Zhang, Deming Chen 저, CSRNet: Dilated Convolutional Neural Networks for Understanding the Highly Congested Scenes
- [3] PULKIT SHARMA
- [4] Mahdi Rezaei, Mohsen Azarmi 저, DeepSOCIAL: Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic
- [5] KCDC, 사회적 거리 두기
- [6] US CDC, Social Distancing
- [7] Goscé, L., Barton, D. & Johansson, A. "Analytical modelling of the Spread of Disease in Confined and Crowded Spaces."
- [8] Cooper, Ian et al. "A SIR model assumption for the spread of COVID-19 in different communities."
- [9] Martin Szummer and Rosalind W Picard. "IndoorOutdoor Image Classification"
- [10] SUN397 Scene benchmark
- [11] The INDECS Database
- [12] Analysis of COVID-19 spread in South Korea using the SIR model with time-dependent parameters and deep learning. Hyeontae Jo, Hwijae Son, Se Young Jung, Hyung Ju Hwang.