

# YOLOv5를 사용한 수영모 착용 여부 판별 알고리즘

고광현, 김민철, 이승환, 유지나\*, 김재현

아주대학교 전자공학과, \*아주대학교 AI 융합네트워크학과

{kokh0128, dktl530, tmdghks7643, \*jina1114, jkim}@ajou.ac.kr

## Algorithm for Determining Detection to Wear a Swimming Cap using YOLOv5

Kwang Hyun Ko, Min Cheol Kim, Seung Hwan Lee, Jina Yu\*, Jae-Hyun Kim

Department of Electrical and Computer Engineering, Ajou Univ.,

\*Department of Artificial Intelligence Convergence Network, Ajou Univ.

### 요 약

대부분의 실내 수영장에서는 수영모 착용을 의무화하고 있다. 본 논문에서는 이미지를 이용하여 수영모 착용 여부를 판별하는 알고리즘을 제안한다. 객체 탐지 기법으로 성능이 좋은 YOLOv5 알고리즘을 사용하고 그 중 실시간 판별을 위해 가장 빠른 모델인 YOLOv5s를 사용한다. 오픈 데이터셋 크롤링을 통해 표본을 수집하여 수영모를 착용한 사람과 착용하지 않은 사람 두 가지로 클래스를 정하고 라벨링을 진행한다. 수집한 데이터셋은 YOLOv5 알고리즘을 통해 이미지 속 사람의 머리를 인식하여 수영모 착용 여부를 판별한다. 수집한 데이터셋을 학습시킨 결과로 96.67%의 높은 정확도를 보이며 수영모 착용 여부를 잘 판별할 수 있다.

### I. 서론

최근 COVID-19로 인한 사회적 거리두기가 완화됨에 따라 야외활동이 늘어나고 있으며 수영장 또한 다시 운영을 재개하고 있다 [1]. 우리나라 대부분의 실내 수영장에서는 수영모 착용을 의무화하고 있고, 일부 야외 수영장에서도 반드시 수영모를 착용해야 이용이 가능하다. 하지만 모든 수영장에 수영모 착용 여부를 관리하는 인력을 추가적으로 배치하기는 현실적으로 쉽지 않다.

본 논문에서는 직접 육안으로 수영모 착용 여부를 확인하는 것이 아닌 you only look once v5 (YOLOv5)를 사용하여 이미지를 자동으로 판별할 수 있는 딥 러닝 알고리즘을 활용하려고 한다. YOLOv5는 속도가 매우 빠르며 한 번에 여러 객체를 탐지할 수 있는 특성을 활용하여 실시간으로 자동 판별이 가능하다. 객체 탐지 딥 러닝 알고리즘을 통해 사람들의 머리를 인식하여 수영모 착용 여부를 판별할 수 있는 알고리즘을 제안한다.



그림 1 데이터셋 라벨링

### II. 본론

#### 1. 객체 탐지

객체 탐지는 주어진 이미지 내에서 사용자가 관심 있는 객체를 탐지하는 기술이다. 객체 탐지는 이미지 분류와는 다르게 객체에 예측 박스를 만들어야 한다. 여기서 이미지 속 객체의 공간적 위치를 표현하기 위한 경계선을 bounding box라고 한다. Bounding box를 통하여 객체를 올바르게 탐지할 수 있다. 또한 딥 러닝 과정에서 bounding box 내의 영역만 탐지 대상이 되기 때문에, 딥 러닝을 효율적으로 수행할 수 있다.

Ground truth란 이미지 속 객체의 위치에 대한 개발자가 지정한 정답을 의미한다. 객체를 잘 감지한 딥 러닝 모델이라면 ground truth와 가장 가까운 bounding box를 표현한 모델이 좋은 성능을 가진 딥 러닝 모델이라고 할 수 있다. 딥 러닝 모델이 예측을 잘했는지를 수치로서 비교하기 위해 intersection over union (IoU) 공식을 사용한다.

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}, \quad (1)$$

IoU 공식은 ground truth의 bounding box와 예측된 경계 상자가 겹치는 부분의 너비에 두 경계 상자가 사진에서 차지하는 너비를 나눈 값이다. Area of Overlap이란 모델이 예측한 bounding box와 ground truth bounding box가 겹치는 교집합 영역을 의미하고, Area of Union이란 모델이 예측한 bounding box와 ground truth bounding box를 포함한 합집합 영역을 의미한다. IoU가 높다는 것은 딥 러닝 모델이 bounding box를 정확하게 예측했다는 것을 의미한다 [2].

#### 2. YOLOv5

YOLOv5는 단일 단계 방식의 객체 탐지 딥 러닝 기법으로 매우 빠른 속도와 추론과정을 통한 실시간 물체 인식 기술로 fast region based convolutional neural networks (Fast R-CNN)과 같은 기존의 실시간 객체 탐지 모델들보다 뛰어난 성능을 보인다 [3].

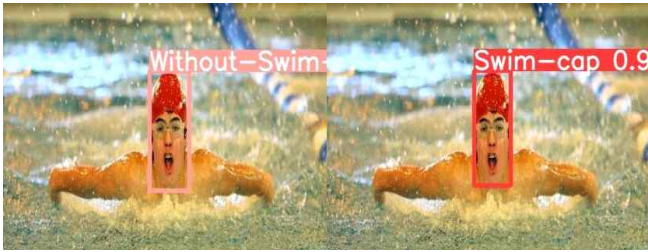


그림 2 오류 개선 결과

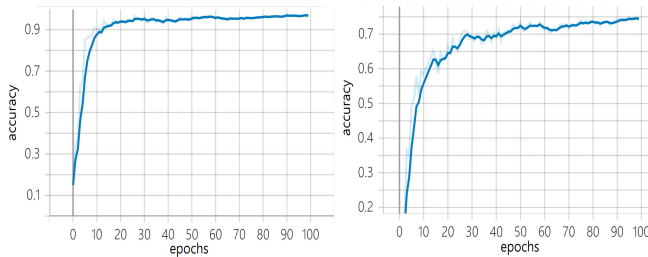


그림 3 (a) mAP\_0.5, (b) mAP\_0.5:0.95 정확도 그래프

YOLO는 동영상을 실시간으로 처리할 수 있으며 이미지 전체를 보고 물체의 일반적인 부분을 학습하기 때문에 단일 대상의 특징뿐만 아니라 이미지 전체의 맥락을 학습하게 된다. 또한, 대상의 일반적인 특징을 학습하여 다른 영역으로의 확장에서도 뛰어난 성능을 보인다. 본 논문에서는 YOLO 알고리즘 중 가장 최신 버전이자, 성능이 좋은 YOLOv5를 사용하며 YOLOv5의 s, m, l, x모델 중 실시간 객체 탐지 속도가 가장 빠른 s모델을 사용한다.

### 3. 데이터셋 수집 및 라벨링

본 논문은 Open Images Dataset V6에서 데이터셋을 사용하고, Pxhere 사이트와 구글 등에서 직접 이미지들을 수집하여 더 많은 데이터셋을 추가한다. 데이터셋은 훈련용 데이터와 검증용 데이터, 테스트용 데이터로 나누어 사용한다.

Ground truth 영역으로 수영모를 포함한 머리 부분을 경계 상자로 설정한다. 그림 1에서와 같이 수영모를 착용한 모든 사람의 머리를 Swim-cap으로 라벨링하고, 수영모를 착용하지 않은 모든 사람의 머리를 Without-Swim-cap으로 라벨링하여 분류한다.

### 4. 구현 결과

일반적으로 객체 탐지 연구에서 성능을 확인하는 지표는 IoU와 mean average precision (mAP)를 도출하여 모델의 정확도를 평가한다. mAP는 IoU를 기준으로 재현율(recall) 값들에 대응하는 정밀도(precision) 값들의 평균을 의미하는 average precision (AP)를 한 개의 객체별로 구하고, 여러 개의 객체 탐지기에 대해서 평균을 구한 것을 뜻한다 [4]. 재현율이란 알고리즘이 이미지에서 찾아야 할 것 중에 실제로 찾은 비율이고 정밀도란 검출된 것 중에 정답의 비율이다. IoU가 0.5를 넘는 경우를 mAP\_0.5라 하고, 0.5에서 0.95 사이의 경우를 mAP\_0.5:0.95라 한다 [5].

학습시킨 결과 중에 그림 2와 같이 노란색 조명이 배경이 되는 특수한 경우 판별이 정확하게 이루어지지 않아 오류가 발생하는 경우가 있다. 이러한 오류를 해결하기 위해 노란색 조명의 배경이 있는 표본들을 추가로 수집하고 반복 학습 횟수를 증가시킨 결과, 정상적으로 판별이 가능해진



그림 4 수영모 이미지 판별 결과

것을 확인할 수 있다.

수집한 데이터셋을 YOLOv5s 모델로 학습한 결과, 그림 3과 같이 mAP\_0.5, mAP\_0.5:0.95는 모두 반복 학습 횟수를 증가시킴에 따라 수영모 착용을 판별할 수 있는 정확도가 높아짐을 확인할 수 있으며 각각 96.67%와 74.23%로 높은 정확도를 보여준다. 그림 4의 경우 이미지를 판별한 결과로 IoU 값이 모두 0.95 이상으로 높은 탐지율을 보이면서 정확하게 수영모를 착용한 사람과 수영모를 착용하지 않은 사람을 구분하는 것을 확인할 수 있다.

### III. 결론

본 논문에서는 딥 러닝 알고리즘인 YOLOv5로 수영모 착용 여부를 판별할 수 있는 방안을 제안하였다. 빠르고 정확하게 수영모 착용 여부를 판별해주는 알고리즘으로 YOLOv5s 모델을 사용하여 여러 번 반복 학습을 진행하여 정확도를 높일 수 있었다. 학습시킨 결과 mAP\_0.5는 96.67%, mAP\_0.5:0.95는 74.23%의 높은 정확도를 확인할 수 있었고 테스트 이미지의 판별 또한 잘 되는 것을 확인할 수 있었다.

향후 이 논문을 토대로 이미지에서 수영모를 착용한 사람과 착용하지 않은 사람을 판별하는 알고리즘을 실생활에 적용할 수 있다면, 수영모 착용 여부를 직접 육안으로 확인할 필요가 없기 때문에 안전성 및 인력감소에 대해 유용하게 사용될 것으로 기대할 수 있다.

### 참 고 문 헌

- [1] e수원뉴스, (2022), “시원한 물속으로 풍덩! 도심 속 실내수영장 ‘수원 청소년문화센터 새천년수영장’”(Online), <https://news.suwon.go.kr/?p=40&viewMode=view&reqIdx=202205162242078738>.
- [2] 진소연, 박종화, 윤상병, 김영수, 이용성, 전지혜, “딥러닝 기반 영상 분석 알고리즘을 이용한 실시간 작업자 안전관리 시스템 개발”, 스마트미디어저널, vol 9, no. 3, pp. 25-30, Sep. 2020.
- [3] 강성우, 송훈, 신영학, “YOLOv5를 이용한 딥러닝 기반 실시간 대장용 종검출 연구”, 대한전자공학회 하계학술대회 논문집, Jun. 2022.
- [4] 박윤환, 송민준, 하란, “YOLO 알고리즘을 이용한 마스크 착용감지 시스템”, 한국소프트웨어종합학술대회 논문집, Dec. 2021.
- [5] 김시현, 이창민, 김효진, 박성용, 김병만, “YOLO-v5 모델을 이용한 실시간 도난 감지 시스템”, 한국컴퓨터종합학술대회 논문집, Jun. 2021.