

YOLO를 이용한 객체인식 최적환경 구성

김선호

mksj869@naver.com

Organizing the optimal environment for object recognition using YOLO

Seonho Kim

School of Computer Science and Engineering, Kyung Hee University

요 약

YOLO(You only look once)는 기존 Object Detection 모델들이 극복하지 못하였던 실시간 처리에 가까운 속도를 구현한 모델로 빠른 속도뿐만 아니라 높은 정확도를 구현하는 것이 중요하다. 이에 본 논문은 YOLO를 이용한 객체인식의 정확도를 높이는 환경 및 방법들을 연구하려 한다. 연구 실험에 사용한 방법으로는 학습 이미지 변형 및 테스트 이미지들의 특성에 따른 정확도 변화를 관찰하여 분석하고 YOLO모델을 변형하는 등의 방법을 연구한다.

1. 서 론

기존 YOLO를 이용한 불건전 객체 모자이크 서비스에서 프로젝트 과정 중 첫 학습파일의 정확도와 인식능력이 높게 나온 후에 학습한 가중치 파일들의 정확도가 높아지지 않거나 인식능력이 떨어지는 현상이 발생했었다. 학습에 사용한 이미지의 양이 2배의 차이가 있음에도 정확도와 인식능력이 오히려 떨어지는 현상을 해결하는 과정이 필요하다고 보이며 이에 YOLO를 이용한 불건전 객체 모자이크 서비스 프로젝트에서 부족했던 인식능력을 개선하고 추가적인 불건전 객체를 학습 및 물체 인식 능력을 발전시키기 위한 프로젝트를 기획하였다. 학습이 상대적으로 어려운 문신을 추가적으로 학습시킬 계획이며 기존 담배와 함께 어떤 학습 환경에서 물체 인식 능력이 가장 높아지는지 학습 및 인식결과 비교를 진행하여 가장 효율적인 학습 환경을 찾아 인식능력을 향상시키는 것을 목표로 한다. 실험할 계획인 방법들은 학습 이미지 형태별 학습 정확도 비교, YOLO 학습 모델별 정확도 비교, YOLO 학습 모델 변형 후 비교 등이 있다.

2. 기존 연구

기존 연구인 YOLO를 이용한 불건전 객체 모자이크 서비스에서 학습에 사용한 이미지는 3000장의 담배 이미지로 학습 이미지 양적으로 문제가 없는 충분한 양의 이미지를 학습시켰다. 학습 결과로 가장 좋은 인식률과 정확도를 보인 모델은 1200장의 이미지를 학습한 모델이었으며 지속적인 정확도 개선을 위해 이미지를 늘리고 모델을 변경하는 등 변화를 주어 학습시켰지만 정확도가 오히려 떨어지는 등의 문제가 발생하였다. 학습과정에서 Best 모델에서 달라진 것은 학습에 사용한 학습모델과

train 이미지, test 이미지로 학습모델에 의한 정확도 변화를 실험한 결과 서로 다른 학습모델들의 학습 정확도 결과가 기존에 존재하는 모델의 성능값과 유사하게 변화하는 것을 관찰하였기에 성능저하에 큰 영향을 준 원인이 아님을 확인하였다.

3. 문제 정의 및 연구 방향

학습에 사용한 이미지들의 품질 상태와 정확한 YOLO 마킹 규칙을 정하지 않고 이미지 전처리를 시행하여 학습시킨 것과 테스트 이미지의 품질상태가 불균등하고 좋지 않은 것이 학습에 사용한 이미지의 양이 2배 이상 늘어났음에도 정확도를 저하시킨 원인으로 예상하고 있다.

학습에 사용할 이미지를 수집하는 과정에서 이미지를 선별하는 과정을 수행하여 저품질의 이미지 또는 학습할 물체가 명확하게 나오지 않은 이미지 등을 제외하여 수집하는 것으로 저품질 이미지가 학습 정확도를 저하시키는 것을 해결 할 수 있을 것으로 보인다. 또한 자체적인 마킹 규칙을 정하여 일정한 이미지 마킹을 시행하고 학습에 사용할 테스트 이미지의 품질변화에 따른 정확도 변화를 확인하는 과정을 통해서 더 효율적으로 학습 가능한 환경을 구성하는 것을 목표로 한다.

4. YOLO 학습

4.1. YOLO 학습 모델 선정

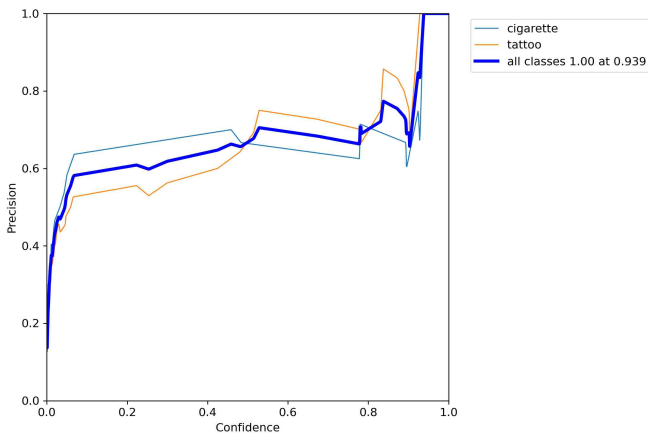
YOLO는 기존 객체 탐지 모델이 학습하는 방식과는 다

르게 과정을 단순화하여 기존 CNN모델 등이 처리하지 못했던 실시간 처리의 속도를 구현한 모델이다.

학습에 사용한 모델은 기존의 YOLO_v3이 아닌 YOLO_v5로 기존 YOLO_v3의 성능이 FPS(Frame per seconds)는 높지만 mAP(mean Average Precision)이 낮은 모델인 반면 YOLO_v5는 FPS와 mAP 모두에서 좋은 성능을 보이는 모델이다. 또한 모델들이 크기별로 다양하게 존재하여 학습에 적합한 모델을 선정하는 것이 중요하다. v5의 x모델은 크기가 큰 모델은 정확도는 높지만 속도가 느린 반면, s모델은 크기가 작아 속도는 빠르지만 정확도가 상대적으로 낮은 모델로 학습에 사용한 모델에 따라 정확도와 처리속도가 매우 다르게 나타난다. 해당 논문의 실험에서는 YOLO_v5의 s모델을 학습에 사용하였다.

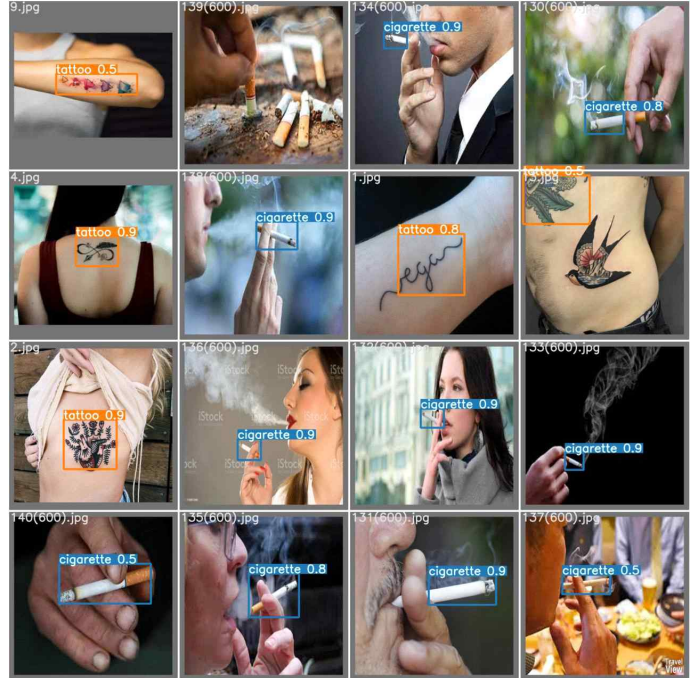
4.2. YOLO 학습 및 결과

담배 학습에 1200장의 학습 이미지를 사용하였고 YOLO를 이용한 불건전 객체 서비스에서 최고 성능을 보였던 모델과 동일한 환경으로 구성하여 동일 성능을 보임을 그래프를 통해 확인하였다.



[그림 1] Cigarette& tatto P_Curve Graph

문신 학습에는 200장의 학습 이미지를 사용하였고 테스트 정확도와 인식 영역이 50%보다 낮게 나왔으며 학습에 사용한 이미지의 양이 적은 것으로 판단하여 학습 이미지를 추가할 예정이다. 또한 문신 학습 후 테스트 과정에서 문신 학습에서 발생하는 오탐지와 미탐지를 발견하였다. 전신 문신과 같은 뾰뾰한 문신들의 경우 인식을 못하거나 일정 부분만 인식하는 문제가 발생하였으며 문신 그림과 비슷한 그림이나 배경을 탐지하는 문제가 발생하였다.



[그림 2] 학습 후 테스트 이미지

위 그림에서 발생한 미탐지와 오탐지 문제를 해결하기 위해서 학습에 사용할 이미지에 학습에 사용하지 않은 전신문신을 추가하고 그림을 오탐지하는 문제는 학습에 사용한 이미지의 양이 충분한 상황에서 다시 발생하는지 확인할 필요가 있다고 판단하였다.

5. 결론 및 향후 연구

YOLO를 이용한 불건전 객체 서비스 연구에서 최고 성능을 보였던 모델을 그대로 다시 구현하였고 이에 문신을 추가하여 학습하여 첫 학습 실험을 시행하였다. 결과적으로 문신의 학습 정확도가 아직 낮아 추가적인 이미지 수집과 학습과정을 진행하면서 정확도를 올릴 예정이다. 또한 담배의 경우 학습에 사용할 이미지의 양을 무작정 늘리는 것이 아니라 이미지의 품질과 변형 상태 등을 기준을 정하여 수집하고 변형하여 학습에 사용하면서 정확도 변화를 관찰할 예정이다.

6. 참고 문헌

- [1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. You only look once: Unified, real-time object detection. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016, pages 779-788. 2016.
- [2] Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., & Yeh, I. H. CSPNet: A new backbone that can enhance learning capability of CNN. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops 2020, pages

390-391. 2020.