

딥러닝 영상 분할의 정확도 향상을 위한 처리방법 연구

최동규 · 김민영 · 장종욱*

동의대학교

A Study on the Processing Method for Improving Accuracy of Deep Learning Image Segmentation

Donggyu Choi · Minyoung Kim · Jongwook Jang*

Dong-eui University

E-mail : dgchoi@office.deu.ac.kr / kmyco@deu.ac.kr / jwjang@deu.ac.kr

요 약

자율주행, CCTV, 휴대폰 보안, 주차시설 등 카메라를 통한 이미지 처리는 실생활의 많은 문제를 해결하기 위해 사용되고 있다. 간단한 구분의 경우는 이미지 처리를 통해 해결하지만, 복잡하게 섞인 물체의 이미지 또는 이미지 내 특징을 찾아내기 어렵다. 이런 특징점 해결을 위해 사람에 가깝게 생각하고 판단할 수 있도록 영상데이터에 분류, 탐지, 분할에서 딥러닝 기술을 도입하고 있다. 물론 이미지 처리만 수행하는 것보다 결과가 좋지만, 딥러닝을 사용한 영상 분할의 방법에서 판단된 결과물이 실제 객체와 편차가 있는 것을 확인하였다. 본 논문에서는 영상 분할의 정밀도를 높이기 위해 딥러닝 영상 분할의 결과물을 출력하기 직전 간단한 이미지 처리를 통하여 정확도 향상을 수행하는 방법에 관해 연구하였다.

ABSTRACT

Image processing through cameras such as self-driving, CCTV, mobile phone security, and parking facilities is being used to solve many real-life problems. Simple classification is solved through image processing, but it is difficult to find images or in-image features of complexly mixed objects. To solve this feature point, we utilize deep learning techniques in classification, detection, and segmentation of image data so that we can think and judge closely. Of course, the results are better than just image processing, but we confirm that the results judged by the method of image segmentation using deep learning have deviations from the real object. In this paper, we study how to perform accuracy improvement through simple image processing just before outputting the output of deep learning image segmentation to increase the precision of image segmentation.

키워드

Deep Learning, Image Segmentation, Image Classification, Computer Vision

1. 서 론

기존의 해결하지 못한 문제들의 해결을 위해서 다양한 분야에 인공지능 기술이 활용되고 있다. 인공지능의 대표적인 수행 능력은 추론 및 추측에 대해 결과물을 보여주는 것이며 제일 많이 사용되고 있는 방법은 많은 데이터에 근거하여 사람의

뇌와 비슷한 구조를 알고리즘 모델을 수행하고 전체 데이터의 가중치를 연산한다. 이 가중치에 대해 추측할 데이터를 입력하고 결과물을 확인하는 것이 일반적이다[1].

대표적으로 영상에서 주로 객체에 관하여 사용되는 기술은 Detection, Segmentation, Classification 이다. 딥러닝을 활용한 결과물을 확인하면 기존에 활용되던 과거의 일반적인 영상처리를 통한 방법보다 흥미로운 결과를 얻을 수 있다. 하지만, 데이

* corresponding author

터 세트의 정제 정확도와 인공지능 모델의 구조에 따라 결과물의 정확도의 편차가 있다[2].

본 논문에서는 인공지능 기술 중 딥러닝 활용에 있어서 사용되는 탐지 기술의 정확도를 높이는 방법에 대해 기술한다. 학습 이후 사용된 모델과 가중치에는 데이터 세트가 다르지 않으면 정확도의 변동은 없으나, 객체의 탐지 이후의 출력에 대해 영상처리를 통하여 검출된 결과의 정확도를 높일 수 있는 보정 방법을 제안한다.

II. 관련 연구

2.1 Deep learning image processing

딥러닝을 통한 이미지 분석은 크게 분류(Classification), 검출(Detection), 분할(Segmentation)로 분류된다. 분류는 이미지에서 특정한 객체나 배경, 분위기 등의 주제를 잡고 각 이미지에 대해 1개의 답을 가지고 지정하는 것을 의미한다. 검출은 보통 이미지에서 필요한 객체를 찾고 그 위치(Localization)를 Binding Box 형식으로 표기를 해주는 것이 일반적인 형태이다. 이것을 넘어서 분할에서는 이미지에서 필요한 객체를 찾는 검출에 추가로 정확한 객체의 모양을 Polygon 형식으로 Box의 형식보다 더 정밀하게 표기한다[1].

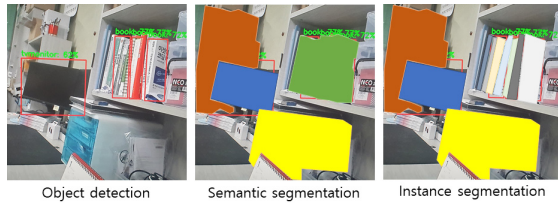


그림 1. 딥러닝을 통한 객체 검출 방법

각 처리방법에 따라 데이터 정제 방식도 다르다. 보통 원하는 결과와 같이 사람이 직접 정제하여 학습 데이터로 사용하게 된다. 이 유사 학습 데이터를 토대로 학습하여 나온 가중치를 토대로 결과에 반영시키기 때문이다.

2.2 이미지 분할(Image segmentation)

딥러닝의 Image Segmentation에는 2가지 방법으로 의미론적 분할(Semantic segmentation)과 의미 있는 객체의 분할(Instance segmentation)이 있다. 의미론적 분할은 객체의 분할을 수행하되 같은 클래스에 속하는 객체들은 같은 영역과 색으로 분할된 것을 의미하며, 의미 있는 객체의 분할은 같은 클래스에 속하더라도 객체들을 모두 구분해 주는 것을 의미한다. 이것은 여러 개의 같은 물체가 있

라도 다른 색으로 분할 한다.

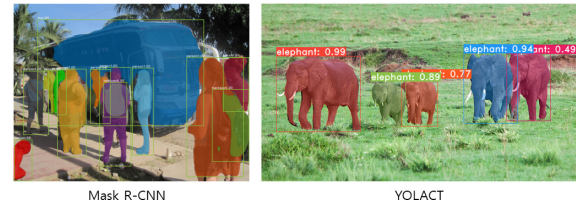


그림 2. 딥러닝을 이용한 이미지 분할 모델

Segmentation도 여러 코드가 있지만 시작을 알린 Mask R-CNN부터 시작하여, 빠른 실시간 처리를 가능하게 한 YOLACT를 주도로 하여 진행하였다 [3,4].

III. 설계 및 수행

수행될 방법은 딥러닝을 통하여 영상 분할이 먼저 진행된 데이터에 후처리를 진행하고자 한다. 딥러닝이 수행된 이미지는 경계선이 완벽히 정돈되지 않은 상태로 본다. 1차로 수행된 이미지 또한 부족하지 않으나 정확도를 올리기 위하여 후처리를 수행한다.

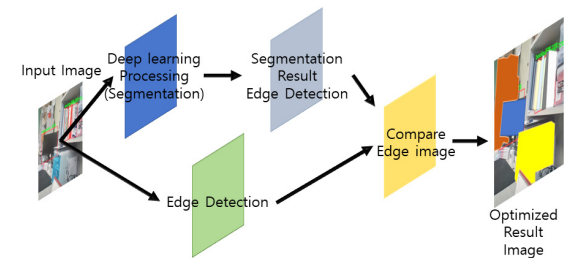


그림 3. 제안하는 방법의 순서도

후처리는 영상처리 방법 중 에지 검출을 진행한다. 에지 검출이란, 이미지 내의 픽셀이 가지고 있는 값에 대해 대체로 색상정보 변화를 감지하여 이미지에 있는 물체의 외곽선을 나타내는 것이다[5]. 에지 검출을 수행한 이미지는 영상에서 외곽선을 나타내게 되고 영상 분할로 수행된 외곽선과 일치하는지 비교하게 된다. 두 과정 이후 나온 데이터에서 유사한 위치에 있는 외곽선이라면 추론으로 나온 딥러닝보다 에지 검출로 나온 것이 실제 데이터로 나온 것이기 때문에 좀 더 정확하다. 만약 여기서 위치 오차가 크다면 데이터 보정을 하지 않는다.

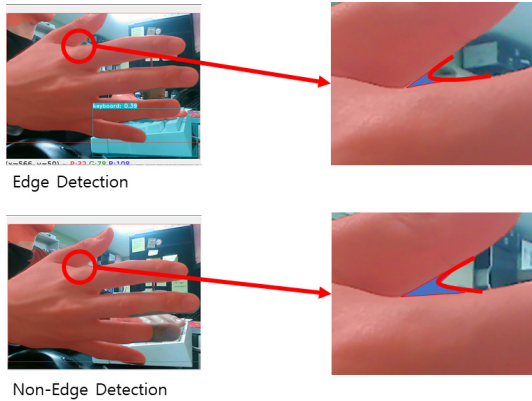


그림 4. 기존과 처리 후 방법의 특정 결과 범위

후처리가 되기 전에도 이미 기존의 Mask R-CNN보다도 검출 영역의 정밀도가 많이 개선되었으나, 후처리 진행 후 필요한 부분에 대해 확연한 개선은 아니었지만 조금은 개선된 모습을 확인할 수 있다. 이미지 출력 프레임마다 차이는 있지만 대체로 영역 안으로 수렴하려는 경향이 있었다.

IV. 결 론

기존 영상처리만으로 해결하지 못한 문제들을 딥러닝이 화두가 된 이후 하드웨어와 소프트웨어의 지속적인 기술 성장으로 더 빠르고 정확한 수행을 해내고 있다. 영상 분할의 경우에도 Mask R-CNN이 처음 나올 당시 실시간으로 하기에는 매우 무거운 작업이었으며, 고성능의 하드웨어를 사용하기 이전에 정확도 개선까지 필요했다. 하지만, 이를 다양한 알고리즘과 연산에 적용하여 YOLACT가 나오게 되었으며, 딥러닝을 통한 해결은 거듭 발전하고 있다.

하지만, 인공지능이 하드웨어에 가중되는 작업이지만 명확한 결과물 때문에 무리하게 사용되기도 한다. 이 외에 정확도에 대한 문제도 입력시킬 데이터의 전처리 정밀도와 다른 환경을 지닌 영상에 따라 결과물은 달라질 수 있다. 본 논문에서 나타낸 것과 같이 인공지능 처리에서의 불필요하게 가중되는 연산 작업을 줄이기 위하여 연구하였으며, 기존의 영상처리 방법을 결과 데이터에 추가함에 따라 더 나은 결과물을 얻고 필요한 환경에 따라 사용할 수 있음을 확인하였다.

Acknowledgement

이 논문은 2020년도 BB21+사업에 의하여 지원되었음. 본 연구는 과학기술정보통신부와 정보통신

기획평가원의 지역지능화혁신인재양성(Grand ICT 연구센터) 사업의 연구결과로 수행되었음(IITP-2021-2020-0-01791).

References

- [1] Brunch, AI Story. Neural Network [Internet]. Available : <https://brunch.co.kr/@gdhan/6>
- [2] Ganghee Lee Tistory. Semantic segmentation and Instance segmentation [Internet]. Available : <https://ganghee-lee.tistory.com/44>
- [3] K. He, G. Gkioxari, P. Dollar and R. Girshick, "Mask R-CNN," in *Proceeding of the International Conference on Computer Vision, Italy* pp. 2980-2988, Oct. 2017.
- [4] D, Bolya, C. Zhou, F. Xiao and Y. J. Lee, "YOLACT : Real-Time Instance Segmentation," in *Proceeding of the International Conference on Computer Vision, Seoul*, pp. 2980-2988, Oct. 2019.
- [5] J. Canny, "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp. 679-698, Nov. 1986.