

임베디드 환경에서 딥러닝을 이용한 얼굴 검출

박명숙, 김상훈*
 한경대학교 전기전자제어공학과
 e-mail:kimsh@hknu.ac.kr

Face Detection using Deep Learning in Embedded Platform

Myeong-Suk Pak, Sang-Hoon Kim*
 Dept of Electrical, Electronic and Control Engineering, Hankyong National University

요 약

최근 몇 년 동안 딥러닝 기법을 이용한 객체 검출이 뛰어난 성능을 보여주었다. 얼굴 검출은 도전적인 문제로 많은 연구가 되고 있다. 본 논문에서는 임베디드 환경에서 적용할 수 있는 객체 검출을 위한 딥러닝 방법을 살펴보고, 얼굴 데이터 셋을 이용하여 훈련시켜 얼굴 검출에 적용한다. 훈련된 모델의 크기는 임베디드 환경에 적합한 메모리 요구량을 보여준다.

1. 서론

얼굴 검출은 얼굴인식에서 먼저 수행되는 중요한 단계이며 생체인식, 감시, 인간 컴퓨터 상호작용(HCI) 등에 적용되고 있다. 컴퓨터 비전에서 영상 내에 얼굴이 있는지 결정하고 그 위치를 찾는 것은 매우 어려운 과제이며 많은 연구가 있었다. 얼굴은 조명, 표정, 방향의 변화와 선글라스, 콧수염 등으로 인한 가림과 같은 요소들 [1, 2]로 인해 검출이 어려워진다.

얼굴 검출 기법은 사전 정의된 규칙을 사용하는 지식 기반 방법, 자세 및 조명 변화에 강인한 얼굴 구조 특징을 찾는 특징 불변 방법, 미리 저장된 얼굴 템플릿을 사용하는 템플릿 매칭 방법, 얼굴 데이터 셋으로부터 모델을 학습하는 외형 기반 방법의 네 가지로 분류할 수 있다 [1]. 최근 몇 년 동안 Faster R-CNN [3], R-FCN [4], SSD [5], YOLO [6] 와 같은 딥러닝 기법을 이용한 객체 검출이 뛰어난 성능을 보여주어 얼굴 검출 성능도 향상되었다.

생체인식이나 인간 컴퓨터 상호작용은 로봇을 포함하여 다양한 모바일 장치에 적용될 수 있어야 하는데 기존의 딥러닝 객체 검출 기법은 임베디드 환경에 적합한 성능을 보여주지 못한다. 이를 위해서는 검출 정확도를 유지하면서도 속도가 빠르고 모델 크기가 작아야 한다. 최근에 관련 연구가 이루어지고 있으며, 본 연구에서는 이들 기법에 대해 설명하고 얼굴 검출에 적용해 본다.

2. 임베디드 환경을 위한 객체 검출 방법

2.1 Tiny YOLO

YOLO는 최첨단의 실시간 객체 검출 시스템으로 최근 YOLOv3[6]가 제시되었다. 이 방법은 Faster R-CNN[3]과 같은 2단계 검출기와 다르게 단일 컨볼루션 신경망을 통

해 경계상자와 클래스 확률을 동시에 예측하는 1단계 검출기이다. 네트워크는 영상을 그리드 셀(grid cell)로 분할하고, 해당 객체의 중심점이 위치한 그리드에서 경계상자를 예측하고 객체가 그리드 셀에 존재하는 지 알려주는 신뢰도 점수(confidence score)를 계산한다.

YOLOv3에서는 Darknet-53이라는 새로운 네트워크를 사용하여 이전 버전보다 강력해졌다. 320 x 320 YOLOv3는 22.2 ms에서 28.2 mAP로 실행되고 SSD[5]만큼 정확하면서 3배 빠르다. Tiny YOLO[7]는 YOLO의 각 버전마다 포함되어 있으며 제약이 있는 환경에서도 사용할 수 있는 작은 모델이다. 성능이 개선된 Tiny YOLOv3도 제시되었다.

2.2 Tiny SSD

Tiny SSD[8]는 임베디드 장치를 위해 객체 검출 성능을 유지하면서 모델 크기를 최소화하도록 설계된 컨볼루션 신경망이다. 이 방법은 SqueezeNet[9]과 SSD[5]에 영감을 받아 고도로 최적화된 비균일 Fire 서브 네트워크 스택과 고도로 최적화된 SSD 기반 보조 컨볼루션 피쳐 레이어(auxiliary convolutional feature layer)의 비균일 서브 네트워크 스택으로 구성된다.

VOC2007에서 61.3 mAP를 달성하면서 모델 크기가 2.3 MB로 Tiny YOLO보다 현저히 작아 필요한 메모리 요구량을 줄이는 효과를 보여준다.

3. 실험결과

이 장에서는 Tiny YOLOv3와 Tiny SSD를 사용하여 임베디드 환경에서 손 검출 실험에 대하여 설명한다. 실험은 NVIDIA Jetson TX2를 사용하였다. 훈련(training)에

사용된 Titan Xp는 3,840개의 CUDA core를 가지고 있고 메모리가 12GB인데 비해 테스트에 사용된 Jetson TX2는 256개의 CUDA core를 가지고 있고 메모리는 8GB로 제한된 자원을 가지고 있다. Tiny YOLOv3는 [7]의 소스코드, Tiny SSD는 [10]의 파일을 이용하여 구현하였다.

데이터 셋은 WIDER FACE[11]를 이용하여 훈련하였고 FDDB[12]를 이용하여 테스트하였다. WIDER FACE는 총 32,203 영상과 스케일, 자세 및 가림의 가변성이 높은 주석이 달린 393,703 얼굴을 포함하며, 12,880 영상이 트레이닝 셋에서 제공된다.

그림 1은 얼굴 검출 결과를 보여준다. 왼쪽은 Tiny YOLOv3의 결과이고 오른쪽은 Tiny SSD의 결과이다. 두 알고리즘 모두 얼굴 방향, 조명, 스케일 등에 대해 강건함을 보여주지만 일부 가려진 얼굴에 대해 Tiny YOLOv3가 잘 찾는 경우(그림 1의 3행)가 있고, Tiny SSD가 잘 찾는 경우(그림 1의 4행)가 있다. 그림 1의 마지막 행은 가려진 얼굴에 대해 두 방법이 검출하는 결과가 다르게 나타났다. 또한 훈련된 모델의 크기의 경우 Tiny YOLOv3는 34.7 MB로 원래의 YOLOv3 보다 훨씬 작고, Tiny SSD는 4.6 MB로 Tiny YOLOv3 보다 작아 두 방법 모두 임베디드 보드에 적용하기 적합한 메모리 공간을 차지한다는 것을 알 수 있다.



(그림 1) FDDB 데이터 셋의 얼굴 검출 결과

4. 결론

얼굴 검출은 얼굴인식에서 먼저 수행되는 중요한 단계이며 딥러닝 기법을 이용한 객체 검출의 큰 성공으로 얼굴 검출도 성능이 향상되었다. 본 논문에서는 임베디드 환경에서 최신 딥러닝 기법을 이용하여 얼굴 검출을 적용하였다. 실험에서 두 방법 모두 얼굴 방향, 조명, 스케일 변화에 대해 얼굴 검출한 결과를 보였고, Tiny YOLOv3와 Tiny SSD는 각각 34.7 MB와 4.6 MB의 모델 크기를 가져서 임베디드 환경에 적합하다고 볼 수 있다. 본 연구의 결과를 통하여 임베디드 환경에 적합한 모델을 위해서 필요한 요소를 이해하고 이어지는 연구에서는 얼굴 검출 성능을 유지하면서 속도가 빠른 임베디드 장치를 위한 모델을 개발하고자 한다.

참고문헌

[1] Zhang C. and Zhang Z. "A survey of recent advances in face detection" Technical Report MSR-TR-2010-66, 2010

[2] Chihouei M., Elkefi A., Bellil W. and Amar C.B. "A Survey of 2D face recognition techniques" Computers, Vol. 5, No. 4, 2016

[3] Ren S., He K., Girshick R., Sun J. "Faster r-cnn: Towards real-time object detection with region proposal networks", Proceedings of the 28th International Conference on Neural Information Processing Systems, 2015, pp.91-99

[4] Dai J., Li Y., He K. and Sun J. "R-FCN: Object Detection via Region-based Fully Convolutional Networks", arXiv:1605.06409, 2016

[5] Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.Y., Berg A.C., "SSD: Single Shot MultiBox Detector", European Conference on Computer Vision, 2016, pp.21-37

[6] Redmon J. and Farhadi A. "YOLOv3: An incremental improvement" arXiv preprint arXiv:1804.02767, 2018

[7] YOLO: Real-Time Object Detection, <https://pjreddie.com/darknet/yolo/>

[8] Wong A., Shafiee M.J., Li F., Chwyl B. "Tiny ssd: A tiny single-shot detection deep convolutional neural



network for real-time embedded object detection" arXiv preprint arXiv:1802.06488, 2018

[9] Iandola F.N., Han S., Moskewicz M.W., Ashraf K., Dally W.J. and Keutzer K. "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5 mb model size" arXiv preprint arXiv:1602.07360, 2016

[10] object detection of tiny ssd, <https://github.com/lampsonSong/tinySSD>

[11] Yang S., Luo P., Chen C.L. and Tang X. "WIDER FACE: A Face Detection Benchmark" IEEE Conference on Computer Vision and Pattern Recognition, 2016

[12] Jain V. and Learned-Miller E.G. "Fddb: A benchmark for face detection in unconstrained settings" Umass Amherst Technical Report, 2010