

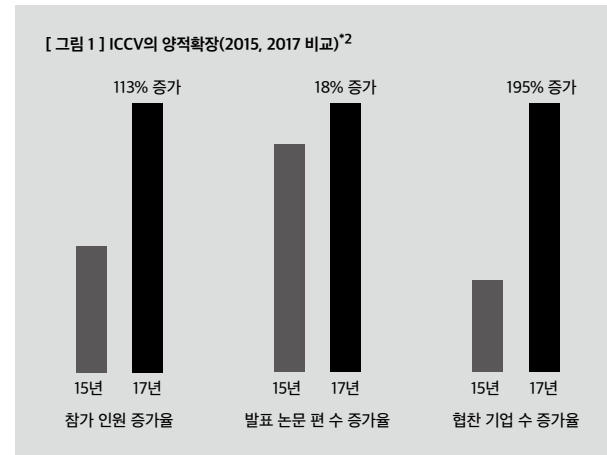
# ICCV 2017 참관기

세계에서 가장 오래된 국제 영화제가 열리는 곳이자, 가장 아름다운 관광 도시 중 하나로 꼽히는 이탈리아 베니스에서 ICCV 2017(국제컴퓨터비전학회, international conference on computer vision, ICCV)이 개최되었다.

ICCV는 CVPR(컴퓨터비전 및 패턴인식 학회, conference on computer vision and pattern recognition)과 함께 컴퓨터 비전 분야에서 가장 권위있는 학회로 꼽힌다. 1987년 영국 런던에서 첫 학회가 열렸고, 1~3년을 주기로 진행되다가 1999년부터 2년에 한 번씩 개최되고 있다. 전기전자공학 분야의 최대 기술 조직인 IEEE(전기 전자 기술자 협회, institute of electrical and electronics engineers)가 주최하는 학회로서, 발표 논문으로 채택되는 것은 매우 어렵다.

## 컴퓨터 비전의 가장 권위있는 학회, 2017 ICCV를 다녀오다

최근 개최된 CVPR<sup>1</sup>과 유사하게 ICCV 역시 딥러닝의 인기로 힘입어 양적 확장이 눈에 띄었다. 수치적으로 비교해 보면, 2009년에서 2015년까지 1100~1400여명이 참석하던 것에 비해 이번에는 무려 3100명이 넘는 많은 사람이 참석을 했다. 그 외 투고되는 논문 편수나 협찬 기업의 수 등 많은 면에서 증가가 있었다.



국내 연구자들의 ICCV에 대한 기여는 큰 것으로 확인되고 있다. 한국은 무려 230여명이 참석해 4번째로 참석 인원이 많은 국가였다. 또한 구체적인 수치가 공개되지는 않았으나, 논문 제출 상위 연구기관에 KAIST(한국과학기술원)와 서울대학교가 이름을 올리는 등 양과 질 두 측면 모두에서 ICCV에 많은 기여를 하고 있음이 확인 가능했다. 이는 국내에서 컴퓨터 비전을 연구하고 있는 대학원 또는 기업의 연구소 수는 상대적으로 적지만, 연구진의 수준은 결코 낮은 편이 아님을 방증한다. ICCV에 다녀온 뒤 느낀 몇 가지 생각과 Best Paper로 선정된 논문들을 정리해 보았다.

## 트렌드를 따라가는데 기본이 되는 튜토리얼(tutorial)

학회 튜토리얼(tutorial)이 상당히 알찬 경우가 많다. ICCV의 경우 GAN(generative adversarial network) 이라던가 Instance-level Visual Recognition 등의 주제로 다양한 튜토리얼이 열렸다. 각 튜토리얼은 가장 기본이 되는 논문부터 본 학회에서 발표될 가장 최근의 논문 소개까지 진행됐다. 이러한 튜토리얼은 개별 주제에 익숙하지 않은 연구자라고 하더라도, 최근의 논문을 이해하는데 큰 도움을 줄 것으로 생각됐다. 특히, 각각의 알고리즘을 제안한 저자의 직강을 듣고 질문할 수 있는 프로그램의 경우 잘 알고 있는 분야라고 하더라도 한번 더 확인하는 계기가 됐다. 다른 사람들에게 가장 추천하고 싶은 시간이다.

## 핵심을 짚어주는 스포트라이트

CVPR 2016에서 처음 접했던 스포트라이트(spotlight)가 ICCV에도 적용이 되었다. 스포트라이트는 포스터 세션(poster session)에 발표될 논문 중 일부에 대해 짧게 핵심만 짚어주는 시간이다. 발표된 논문에 대해 질문을 할 때 알고리즘 설명을 되묻지 않고 바로 질문을 하거나 관심있는 논문을 선별하는데 많은 도움을 주었다. 오랄 세션(oral session)이 1회부터 9회까지의 직관하는 야구 경기라면, 스포트라이트는 스포츠 뉴스처럼 승부처만 요약해서 보는 야구 경기 같은 느낌이었다.

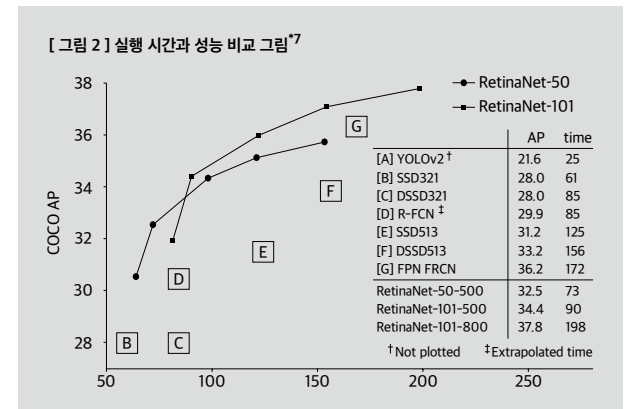
## 단 한가지 아쉬운 점

아쉬운 점을 딱 한가지 뽑자면, 학회에서 발표되는 논문이 이미 arXiv 를 통해 수 개월, 길게는 수 년 전에 공개되다 보니, 종종 본인 발표 마지막에 추가 연구를 진행한 새 논문의 arXiv 주소를 공개하는 경우가 있었다는 점이다. 최신 논문이 더 이상 최신 논문이 아니라는 생각이 들었고, 이미 알고 있는 내용이다 보니 흥미가 떨어지는 순간도 있었다. 이러한 문제는 ICCV 뿐만 아니라 CVPR이나 타 학회 발표 역시 갖고 있다. arXiv를 통한 선공개는 분명 장점이 더 많지만, 이러한 단점에 대한 보완책이 미비하다는 점 또한 아쉬웠다.

## 최우수 논문 소개

### 1) Best Student Paper : Focal Loss for Dense Object Detection

딥러닝 기반의 검출기는 크게 Fast/Faster R-CNN<sup>3,4</sup>처럼 후보 영역을 검출하고 후보 영역에서 정확한 bounding box와 카테고리를 구별하는 2-stage 검출기와, YoLo<sup>5</sup>, SSD<sup>6</sup>처럼 한 번에 bounding box와 카테고리를 구별을 하는 1-stage 검출기로 구별된다. [그림 2]를 보면 2-stage 검출기(A, B, C, D, E, F)는 상대적으로 느리지만, 더 정확한 성능을 보이고 1-stage 검출기(G)는 반대로 매우 빠르지만 다소 부족한 성능을 보인다.



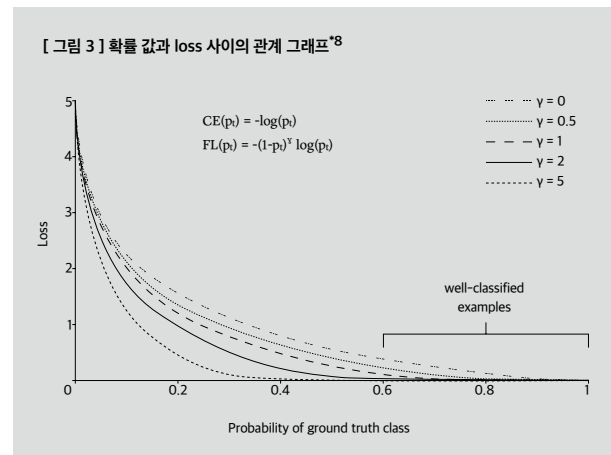
글 | 이주영 michael.lee@kakaocorp.com

대학원을 졸업하고 9년여의 시간을 컴퓨터비전 관련해서 회사에서 연구/개발 업무를 해왔지만, AI 리포트에 기고할 때마다, 실수한건 없을까 싶어 항상 긴장하게 되는 것 같습니다. 다음에 기고할 기회가 주어진다면, 참관기가 아닌 제 논문에 대한 이야기를 다룰 수 있게 되기를 기대합니다.

글 | 노명철 joshua.roh@kakaocorp.com

오랜만에 참석하는 권위있고 큰 학회였습니다. 많은 논문들이 발표돼 부지런히 돌아 다녔는데, 시차 때문인지 체력적으로는 너무 힘든 학회였습니다. 대부분의 논문 내용들은 인터넷을 통해 검색할 수 있는 시대이지만 전체적인 추세나 분위기의 경우 학회장에 참석하지 않고는 알기 힘들음을 다시 한번 느끼며, 견문을 넓힐 수 있었던 좋은 기회 였다 생각합니다.

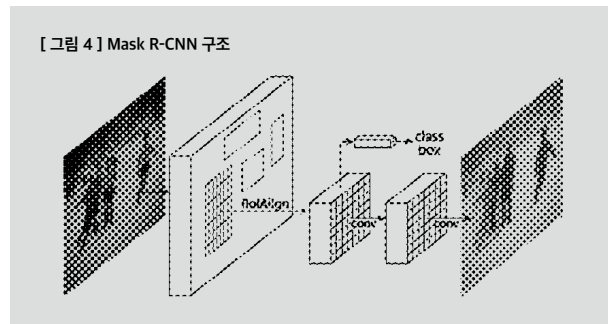
이러한 결과는 쉽게 구별되는 객체가 아닌 영역(background)이 객체 영역보다 매우 많이 학습에 참여하게 되어 발생하는 것으로 간주된다. 이를 해결하기 위해 새로운 손실 함수(loss function)를 제안했다. 새로 제안하는 손실 함수는 잘 분류되는 예제는 손실(loss)이 적게 발생해 [그림 3]과 같은 특징을 가지도록 디자인 되었고, 검출 결과에서 좋은 성능을 보였다. 문제로 가정했던 것이 맞았음이 실험적으로 증명된 것이다.



흥미로웠던 점은 객체 검출에서 성능을 높이기 위해 현재 알고리즘이 구별하기 어려워하는 대상을 학습에 더 많이 참여시키는 방법(online hard example mining, OHEM<sup>9</sup>) 보다도 더 높은 정확도를 보였다는 점이다.

## 2) 최우수 논문 : Mask R-CNN

최우수 논문은 객체 검출(object detection)과 instance segmentation을 다룬 ResNet을 처음 제안한 카이밍 허(Kaiming He), 사람 검출 알고리즘의 전문가 피오트 달러(Piotr Dollar), Fast(er) R-CNN 을 제안한 로스 걸식(Ross Girshick) 등이 참여한 Mask R-CNN 이었다. 쟁쟁한 저자들이 공동작업한 논문이기에 저자들의 명성만으로도 충분히 좋은 논문이 나올 것 같았다. 이 논문은 특히 쉽고 단순한 방법으로 instance segmentation 알고리즘을 제안했다. 기존에 객체 검출에서 높은 성능을 보인 Faster R-CNN<sup>10</sup> 의 네트워크 구조 [그림 4]에 instance segmentation task를 담당하는 분기(branch) 를 추가하였고, 이렇게 새롭게 제안한 네트워크는 COCO 2016 challenge 에서 1등을 하여, 우수함을 증명하였다.



"Simple Is the Best"라는 표현이 더 잘 어울리는 논문은 없을 것 이란 생각이 들만큼, 단순하고 정확한 결과를 보인 알고리즘이라 개인적으로 매우 관심이 가는 논문이다. [그림 5] 참고



## 정리하며

학회에 참석하고 나면, 하고 있는 연구와 업무에 더 집중할 수 있고 스스로를 되짚어 보게 되어 항상 좋은 느낌을 받았던 것 같다. 더군다나 2년 뒤에 열리는 ICCV는 대한민국 서울에서 개최될 예정이라고 하니 마음 한구석으로 학회장 어느 한 곳에서 내가 발표를 하게 된다면 어떨까? 라는 상상까지 하게 되었다.

마지막으로, 인공지능 관련 인재 영입을 위해 학회를 찾아온 패션 기업이나 VR, AR 체험 부스 등 이전의 머신러닝 학회에서는 볼 수 없었던 기업들의 참여 모습도 인상적이었다. 또 딥러닝의 시대로 넘어오면서 데이터의 양이 이전 보다 훨씬 중요하게 되었는데, 이러한 데이터의 분류, 마킹(marking)을 대신 해주는 회사들도 부스를 차리고 나와 있었다. 딥러닝이 얼마나 학회와 산업계에 큰 영향을 미치고 있는지를 알 수 있는 부분이었다. 이 딥러닝 기술이 언제까지 지속될지, 또 다른 새로운 기술은 언제 나타날지 벌써부터 궁금해진다.

<sup>\*1</sup> 참고 | <https://brunch.co.kr/@kacao-it/143> <sup>\*2</sup> 참고 | <http://iccv2017.thecvf.com/files/OpeningICCV17.pdf> <sup>\*3</sup> 논문 | Ren, S., He, K. & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, doi : arXiv:1506.01497. <sup>\*4</sup> 논문 | Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection, doi : arXiv:1506.02640 <sup>\*5</sup> 논문 | Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection, doi : arXiv:1506.02640 <sup>\*6</sup> 논문 | Liu, W. et. al. (2016). SSD: Single Shot MultiBox Detector, doi : arXiv:1512.02325 <sup>\*7</sup> 논문 | Lin, T., Goyal, P., Girshick, R., He, K. & Dollar, P. (2017). Focal Loss for Dense Object Detection, doi : arXiv:1708.02002 <sup>\*8</sup> 논문 | Lin, T., Goyal, P., Girshick, R., He, K. & Dollar, P. (2017). Focal Loss for Dense Object Detection, doi : arXiv:1708.02002 <sup>\*9</sup> 논문 | Shrivastava, A., Gupta, A., & Girshick, R. (2016). Training Region-based Object Detectors with Online Hard Example Mining, doi : arXiv:1604.03540 <sup>\*10</sup> 논문 | Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection, doi : arXiv:1506.02640