

# 2013년과 2017년의 CVPR을 비교하다

출발은 한두 주 전쯤으로 기억한다. 지인의 초대로 페이스북 그룹채팅방에 들어가보니 직접적으로 혹은 한다리 건너 알만한 대학원 연구실 학생들이며 회사에 소속된 연구원들 20~30 명이 모여 있었다. 채팅방 이름은 'CVPR2017\_kor 수다방' 제목 그대로 CVPR에 참석하는 한국인들의 수다방이었다. 학회 전까지는 자기소개며, 어떤 비행기를 타고 가는지 등의 잡담을, 학회 중에는 실시간 발표자에 대한 평이나 저녁 시간 번개를 모집하는 글 등 자유로운 글들이 여과없이 올라왔다. 직감적으로 이번 학회에 상당히 많은 한국인들이 참석하게 될 것을 느꼈고 예상대로 인천공항과 기내에서 여러 명의 지인과 마주쳤다

컴퓨터 비전 및 패턴 인식 분야에는 '컴퓨터 비전 및 패턴 인식 컨퍼런스(CVPR)', '유럽컴퓨터비전학회(ECCV)', '국제컴퓨터비전학회(ICCV)' 등 외에도 다양한 학회가 있다. 그 중 CVPR은 규모나 영향력 지수(impact factor) 측면에서 가장 인기 있는 학회이다. 특히나 최근에는 딥러닝의 인기로 모든 지표에서 학회의 성장을 확인할 수 있다. 당장 작년과 비교해 봐도 논문 투고 수 40% 증가, 등록 인원 37% 증가, 스폰서 펀딩 금액 79% 증가 등 가파른 성장세를 보이고 있다. 개인적으로 가장 인상 깊었던 것은 기업체 지원(sponsor) 부분이었다. 2010년 전후만 해도 시그그래프(SIGGRAPH) 학회에 참석하게 되면 스폰서 리스트에 표시된 수십 개의 기업체 마크와 화려한 전시 부스를 보며 내심 부러웠는데, 이제 CVPR은 시그그래프보다 많은 130여개 업체가 스폰서십으로 참여할 정도로 기업체로부터 대단히 큰 관심과 지원을 받는 학회가 되었다. 이렇듯 CVPR은 학계와 산업계 모두에서 크게 주목받으며 성장하는 학회이다.

본 학회에 마지막으로 참석했던 2013년 비교하여 올해 내가 받은 인상을 세가지만 뽑아서 정리해 보았다.

글 | 천영재 yeongjae.cheon@kakaobrain.com

지난 10년 동안 컴퓨터비전 분야(얼굴/사물 인식) 한 우물만 파왔다. 방법보다는 문제에 집중하며, 효율적이면서 효과적인 기술이다 싶으면 어느새 빠져드는 전형적인 엔지니어이다. 사람과 시간을 소중히 하며 끝까지 매진하여 좀 더 크고 가치 있는 결과물을 만들어내고 싶다.

## 모든 면에서 폭발적인 성장

먼저 워크숍이 열린 홀의 크기와 참석 인원에 놀라지 않을 수 없었다. 메인 학회에 앞선 워크숍 첫날 동시에 10여개의 워크숍들이 진행되었는데, 필자가 참석한 자율주행 관련 워크숍은 메인 학회에 조인트로 열리는 워크숍이라고 보기 어려울 정도로 상당히 큰 홀에서 진행되었다. 더군다나 워크숍 첫날은 보통 민망할 정도로 사람이 없기 마련인데 이번 워크숍에서는 첫날부터 발표장이 상당히 붐볐다. 메인 컨퍼런스는 3개의 세션으로 나뉘어 각자 다른 대형 홀에서 진행되었는데 등록자가 많다 보니 어느 세션에서도 빈자리를 찾기가 쉽지 않았다. 이러한 열기와 참여는 메인 학회와 워크숍 마지막 날까지 이어졌다.

한국에서도 많은 학생과 연구원들이 참석한 것을 보고 한번 더 놀랐다. 필자가 대학원 석사과정이었다던 10년 전에는 논문을 발표하는 학생만 학회에 참석하는 것을 당연하게 생각했었는데, 이번 학회에는 한국의 복수의 대학원 연구실에서 논문이 없는 학생들에게도 학회에 참관할 수 있는 기회가 주어졌다. 또한 국책연구소나 회사에서도 많게는 십여 명의 연구원들이 오는 것을 볼 수 있었다. 덕분에 본인도 한국에서는 보기 힘들던 대학원 지도 교수님 및 선배님들과의 모임을 가지기까지 하였다.

스폰서 기업과 이 기업들의 전시 부스도 이제는 학회 행사의 중요한 한 축을 담당하게 되었다. 수년 전 아기자기 하던 전시 부스들은 어디로 가고 이제는 상업 전시회를 방불케 할 정도로 큰 공간에서 다수의 글로벌 IT 기업들을 포함한 130여 기업의 부스가 차려졌다.

## arXiv를 통한 출판 전 논문 공개와 오픈 소스 활성화

한두 해 전부터 이미 활성화 된 arXiv를 통한 논문의 출판 전 공개와 깃허브(github)를 통한 소스 공개 경향은 더욱 두드러져 보였다. 베스트 페이퍼를 받은 논문들은 말할 것도 없거니와 다수의 구두 발표 논문들이 이미 arXiv로 통해서 공개 된 것들이었다. 결과적으로 최신 논문들로 가득해야 할 학회 발표장에는 이미 많이들 읽어본 심지어는 피인용수가 50건을 넘긴 논문들을 찾아 볼 수 있었다. 이러한 이유에서인지 몇 해 전부터 본 학회는 발표하는 모든 논문을 온라인 상에 무료로 공개하고 있으며<sup>1</sup>, 작년부터는 튜토리얼과 구두발표, 스팟라이트 발표 동영상까지도 유튜브를 통해서 공개<sup>2</sup>하고 있다.

저자가 소스를 깃허브(github)를 통해서 소스를 공개하는 모습도 쉽게 찾아 볼 수 있었다. 게다가 arXiv로 미리 공개된 유명 논문들은 저자가 소스코드를 공개하지 않더라도 독자들이 앞다투어 각자 다른 버전의 코드를 작성해 공개하고 있다. 한 예로, 최우수

논문상을 받은 DenseNet<sup>3</sup>은 저자가 공개한 구현<sup>4</sup> 외에도 15개 이상의 서로 다른 구현을 깃허브를 통해서 내려 받을 수 있다.

## 리크루팅(recruiting) 전쟁터

전시 부스의 가장 큰 목적은 리크루팅이라고 느껴졌다. 단순히 기술을 전시하고 홍보하는 것이 아니라, 적극적으로 학생들과 연구원들의 연락처를 수집하고 경우에 따라서는 현장에서 회사의 연구원과 학생 사이에 리크루팅 관련 질의 응답이 이루어지는 경우도 있었다.

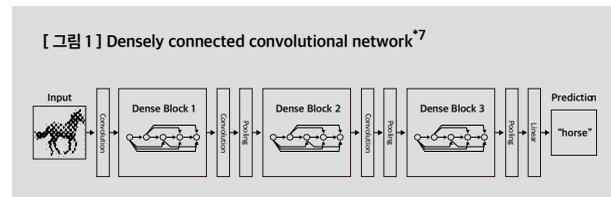
저녁에는 우버, 구글, 트위터, 마이크로소프트 와 같이 주로 북미 회사들이 네트워킹 파티를 열었다. 나는 스냅챗(SnapChat) 파티에 초대 받아 참석하였는데 스냅챗 연구원들 뿐 아니라 실리콘 벨리의 벤처 창업자들이나 다른 회사 연구원들이 연구 주제와 현재의 소속 앞으로의 계획 등에 대하여 물어왔다. 게다가 일부 회사들은 비공개 소규모 리셉션을 한다거나 그룹 단위로 학생들을 초대하여 식사를 하는 경우도 많았던 것으로 알고 있다.

덕분에 학생들도 바쁜 일정을 소화해야 했을 것이다. 봐야할 포스터들도 많았을 뿐더러, 졸업을 앞둔 학생들은 여러 팀들과 인터뷰를 본다고, 또 저녁 시간에는 파티나 식사 모임에도 참석해야 했으니 말이다.

## 최우수 논문상을 받은 4편의 논문

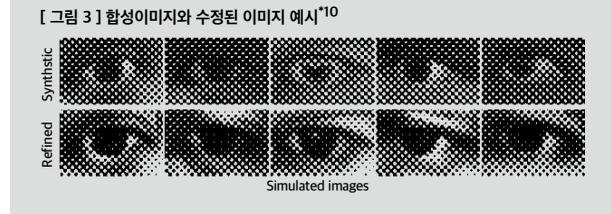
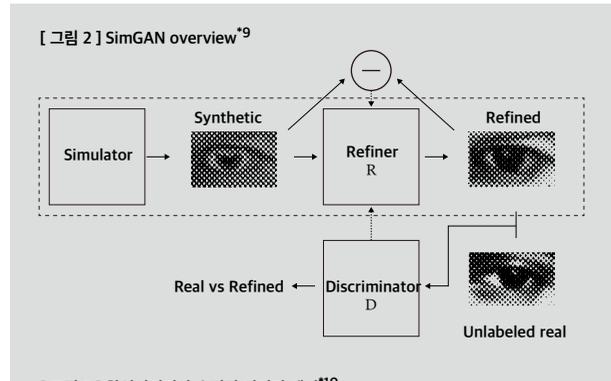
### 1) 2편의 최우수 논문상(Best Paper Awards).

첫번째는 페이스북인공지능연구소(Facebook AI Research, FAIR)에서 작년 12월 발표한 '성기게 연결된 컨볼루션 네트워크(Dense Connected Convolutional Networks, DenseNet)<sup>5</sup>'이었다. 본 논문은 ResNet<sup>6</sup> 등과 같은 최신 컨볼루션 네트워크(convolutional networks) 연구에서 입력 레이어와 출력 레이어 사이에 짧은 연결들이 포함되면 더욱 깊은 구조의 모델을 효과적으로 학습하여 높은 정확도를 얻을 수 있다는 직관으로부터 시작한다. 논문에서 저자들은 더욱 성기게 연결되는(densely connected) 단순한 구조의 모델을 제안하고 있다. 결과적으로 DenseNet은 물체 인식(object recognition) 벤치마크 태스크에서 발표 당시 빠르면서도 가장 좋은 성능을 보였다.



두번째는 작년 12월 애플이 비밀주의를 깨고 최초로 논문을 공개하여 화제가 되었던 '적대적 훈련을 통해 시뮬레이션된 이미지와 감독되지 않은 이미지로부터의 학습(Learning from Simulated and Unsupervised Images through Adversarial Training)\*8이다. 지도 학습(supervised learning)은 불가피하게 대량의 학습 이미지와 함께 정답(ground truth)을 필요로 하는데 이는 많은 경우 시간과 돈의 문제로 제한을 받는다. 이를 해결하기 위한 방법 중 최근 가장 각광을 받고 있는 방법이 본 논문에서와 같이 실제 데이터가 아닌 그래픽스 엔진을 이용하여 데이터를 합성하여 사용하는 방법이다.

하지만 이러한 방법은 필연적으로 실제 데이터와 합성된 데이터 사이의 차이로 인한 성능 저하 문제를 가지게 된다. 본 논문은 [그림 2]에서와 같이 합성된 이미지를 실제 이미지들과 구분하기 어려우면서도 합성이미지와도 유사한 이미지로 수정하는 Refiner를 적대적 훈련(adversarial training) 방법으로 학습함으로 이 문제를 해결하고 있다. [그림3] 에서 합성이미지가 보다 사실적으로(realizm) 개선되는 것을 확인 할 수 있다. 이 방법을 시선 추정(gaze estimation) 문제와 손동작 추정(hand pose estimation)문제에 적용하여 기존 방법 대비 성능 개선을 얻을 수 있었다.



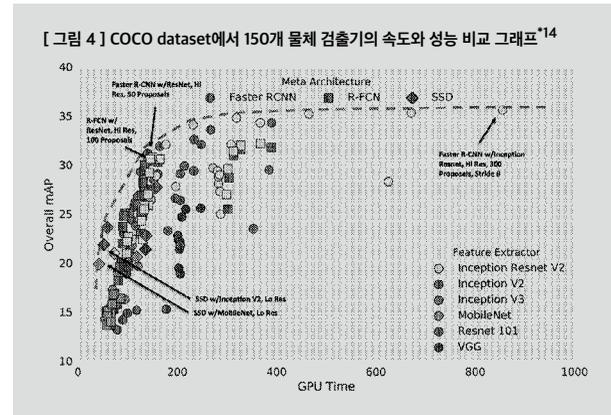
2) 2편의 최우수 영애 멘션상(Best Paper Honorable Mention Awards)

일반적으로 학회 조직 위원회(committee)에 의하여 선정되는 최우수 영애 멘션상은 Polygon-RNN을 이용하여 미지에서 물체에 주석(annotation)을 만드는 작업을 적은 수의 수작업으로 가능하게 하는 방법을 제안한 '다각형-RNN을 이용한 물체에 주석달기(Annotating Object Instances with a Polygon-RNN)\*11와

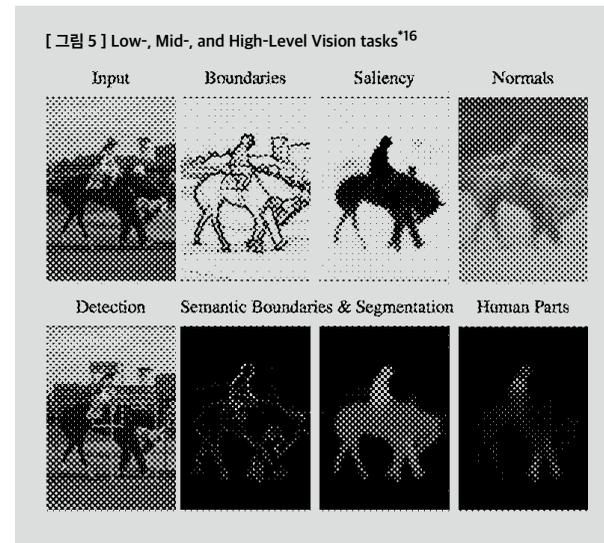
모델의 크기를 조절하여 속도와 성능을 상대적으로 다양하게 조정 할 수 있으면서도 어떠한 경우에도 기존의 다른 알고리즘보다 빠르면서 좋은 성능을 보장하는 '요로9000(YOLO9000: Better, faster, stronger)\*12이 수상하였다.

3) 그밖에 인상 깊었던 논문들

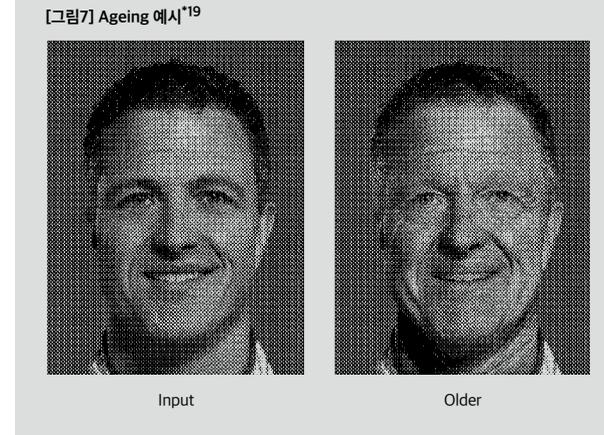
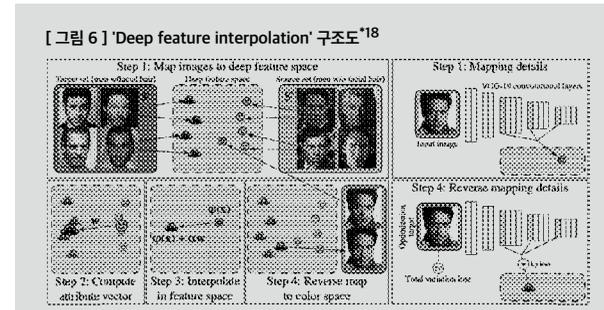
(1) '최신 컨볼루션 물체 검출기들의 속도와 정확도 사이의 트레이드오프(Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors)\*13 (Google) : 이미지에서 물체를 검출하는 문제를 풀고자 할 때, 주어진 어플리케이션이나 플랫폼에 따라서 속도, 메모리, 성능 등의 제약 조건이 다를 수 있다. 이때 널리 알려진 다양한 딥러닝 모델 구조 중 옳은 선택을 하기 위한 가이드를 제공하고 있다. 뿐만 아니라 제안하고 있는 딥러닝 모델구조와 하이퍼파라미터를 쉽게 조정할 수 있는 구현 코드 모두를 공개하고 있다.



(2) '우버넷(UberNet: Training a Universal Convolutional Neural Network for Low-, Mid-, and High-Level Vision Using Diverse Datasets and Limited Memory)\*15 (FAIR) : 마치 스위스칼(swiss knife)과 같이 컴퓨터 비전 분야에서의 '현저한 물체 추정(saliency estimation)', '외각선 검출(boundary detection)', '물체 검출(object detection)', '시맨틱 분할(semantic segmentation)', '물체 파트 검출(object part detection)' 등의 다양한 레벨의 문제를 하나의 모델로 풀 수 있도록 학습하는 방법을 제안하고 있다.



(3) '이미지 콘텐츠 변화를 위한 깊은 특징점 보간법(Deep Feature Interpolation for Image Content Changes)\*17 : [그림6] 에서와 같이 깊은 컨볼루션 피쳐(deep convolutional feature) 공간에서 보간법(Interpolation)을 이용하여 이미지 콘텐츠(나이 혹은 표정)를 변경하는 방법을 제안하고 있다. [그림7]은 입력 얼굴을 나이든 얼굴로 변형한 예시이다.



글을 마치며

arXiv 덕분에 유명 논문들은 학회 전에 이미 읽어 볼 수 있고 나머지 논문들과 심지어 발표 동영상까지 올라오는데 왜 굳이 많은 비용과 시간을 들여서 학회에 가는지를 누군가 물을 수 있을 것이다. 하지만 '백문이 불여일견'이라는 속담처럼 가서 보고 느끼고 네트워킹하며 받는 자극은 분명 단순히 논문을 읽는 것과는 비교할 수 없는 가치가 있을 것이다. 본인은 앞으로도 머신러닝(특히 딥러닝)과 컴퓨터 비전 모두를 동시에 다루는 가장 큰 학회인 CVPR은 꼭 참석하고 싶다. 이 자리를 빌려 학회를 참관할 수 있는 기회를 준 회사와 동료들에게 감사의 마음을 전한다.

\*1 참고 | <http://openaccess.thecvf.com/CVPR2017.py> \*2 참고 | [https://www.youtube.com/channel/UCOn76gicarsN\\_Y9YShWwhw/videos](https://www.youtube.com/channel/UCOn76gicarsN_Y9YShWwhw/videos) \*3 논문 | Huang, G., Liu, Z., Weinberger, K. & Maaten, L. (2017). Densely connected convolutional networks. CVPR. \*4 참고 | <https://github.com/liuzhuang13/DenseNet> \*5 논문 | Huang, G., Liu, Z., Weinberger, K. & Maaten, L. (2017). Densely connected convolutional networks. CVPR. \*6 논문 | He, K., Zhang, X., Ren S. & Sun, J. (2016). Deep Residual Learning for Image Recognition. ECCV. \*7 논문 | Huang, G., Liu, Z., Weinberger, K. & Maaten, L. (2017). Densely connected convolutional networks (p. 3). CVPR. \*8 논문 | Shrivastava, A. et al. (2017). Learning from simulated and unsupervised images through adversarial training. CVPR. \*9 논문 | Shrivastava, A. et al. (2017). Learning from simulated and unsupervised images through adversarial training (p. 2). CVPR. \*10 논문 | Shrivastava, A. et al. (2017). Learning from simulated and unsupervised images through adversarial training (p. 5). CVPR. \*11 논문 | Castrejon, L., Kundu, K., Urtasun, R. & Fidler, S. (2017). Annotating object instances with a Polygon-RNN. CVPR. \*12 논문 | Redmon, J. & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. CVPR. \*13 논문 | Huang, J. et al. (2017). Speed/Accuracy trade-offs for modern convolutional object detectors. CVPR. \*14 논문 | Huang, J. et al. (2017). Speed/Accuracy trade-offs for modern convolutional object detectors (p. 9). CVPR. \*15 논문 | Kokkinos, I. (2017). UberNet: Training a universal convolutional neural network for low-, mid-, and high-level vision using diverse datasets and limited memory. CVPR. \*16 논문 | Iasonas Kokkinos. UberNet: Training a Universal Convolutional Neural Network for Low-, Mid-, and High-Level Vision Using Diverse Datasets and Limited Memory. In CVPR, 2017. \*17 논문 | Upchurch, P. et al. (2017). Deep Feature Interpolation for Image Content Changes. CVPR. \*18 논문 | Upchurch, P. et al. (2017). Deep Feature Interpolation for Image Content Changes (p.2). CVPR. \*19 논문 | Upchurch, P. et al. (2017). Deep Feature Interpolation for Image Content Changes (p.1). CVPR.