Computer Vision Paper Gestalt

# **계기:**

학술 연구 결과를 학계에 전파하는 데 peer review가 필수다. Peer review는 해당 분야의 다른 전문가들이 연구 논문을 철저히 검토하는 절차이다. 그러나 최상위 컴퓨터 비전 학회에 제출되는 논문의 수가 사상 최고치를 기록하는 반면, 그에 비해 숙련된 리뷰어의 수는 부족하여 peer review가 점점 더 어려워졌다. 학회 조직위원회는 결국 경험이 적은 학생들이 peer reviewer가 되어, 그 결과로 수개월 또는 수년 동안 노력해서 작성한 논문의 평가 근거가 부족하거나 불공정한 평가를 받게 된다.

본 논문에서는 이전 년도의 학회의 논문 자료(ICCV/CVPR 학회 및 워크숍 논문)을 활용해 논문의 시각적 외형(Paper Gestalt)을 기반으로 논문의 품질을 판단할 수 있도록 CNN 기반 모델을 개발했다.

# **데이터셋/Train:**

Computer Vision Foundation, CVF 웹사이트에서 데이터 크롤링을 통해 2013-2017년 학회에 제출된 논문 중 Accepted된 논문을 좋은 논문, Rejected된 논문들을 안좋은 논문으로 분류하여 Train 시킨다. 논문의 편파적이지 않고 공정한 분류를 위해 7페이지 이상의 논문들로 선정하였다. 또한 다운받은 논문 pdf를 pdf2image를 사용하여 8페이지를 4행 2열로 구성된 하나의 이미지로 총 2200 x 3400 pixel 사이즈로 변환했다. Data Pre-processing은 논문의 머릿글을 제거하여 논문 분류의 과적합성을 방지했다.

# **Test:**

2018년 CVPR의 accepted/rejected 논문들과 random sampling한 이전년도 논믄들을 test data로 사용했다. Classification method: 기존 imageNet으로 학습된 1000개의 클래스를 Resent-18을 사용하여 두개의 출력 노드(좋은 논문/안좋은 논문)로 fine tuning 했다. 이를 바탕으로 CVPG(Computer Vision Paper Gestalt) 데이터셋(github에 제공됨)을 SGD를 적용하여:

Momentum=0.9

Epoch=50

초기 learning rate=0.001

Decay Learning Rate/ per 10 Epoch= 0.1 (10epoch당 학습률 0.1배 감소)

Loss function = Weighted Cross Entropy loss

Train/test image = 224 x 224 픽셀

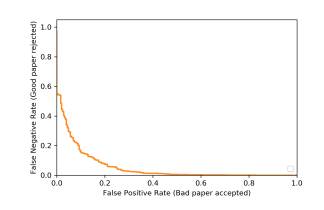
논문의 원래 시각적 내용과 레이아웃을 유지하기 위해, 무작위 자르기(random cropping), 수평 뒤집기(horizontal flipping), 색상 변화(photometric transformation) 등의 일반적인 데이터 증강 기법은 적용하지 않는다.

# **성능:**

2013-2017년도 ICCV/CVPR 학회 및 워크숍 논문인 Trained Data를 학습시켜 Test Data인 2018년도의 논문들에 대해 92%의 정확도를 가진다.

# **Error:**

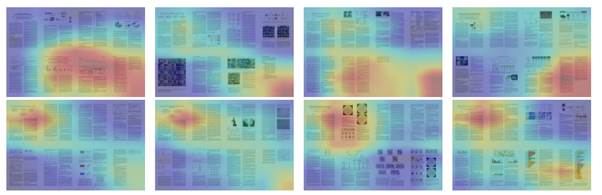
Softmax layer 이후의 임계값들을 변화시키는 것에 대해 모델의 성능을 나타낸 그래프:



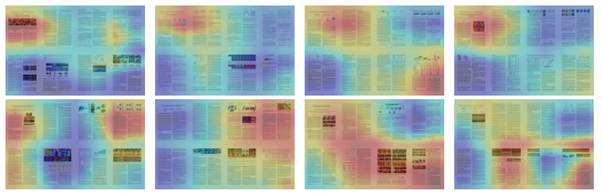
False Positive Rate (FPR) vs False Negative Rate (FNR) 여기서 TPR 대신해서 FNR을 사용하였다. 학회에서 공정하지 않은 평가 시 좋은 논문이 rejected되는 error와 안좋은 논문이 accepted되는 error간의 상충관계를 더 잘 설명하기 위해 사용했다. 전체 dataset은 총 3309개 중 979개의 좋은 논문, 2230개의 안좋은 논문이 있다. 이 논문에서는 0.4% FPR, 50% FNR로 설정하여 많아도 겨우 4개의 좋은 논문(전체 데이타셋의 0.4%) rejected되도록 수용했다. (이러면 안전하게/확실하게 50%의 안좋은 논문을 reject 할수있음)

# **Class-specific activation map:**

Classification 모델이 논문 이미지의 어느부분을 근거로 의사결정을 했는지 보여주는 visualization tech.



모델이 평가한 안좋은 논문 = 8페이지 꽉 채워져있지 않다, 첫페이지에 논문의 전체 흐름을 꿰뚫는 그림이 없다 등 안좋은 논문으로 평가한 이유가 있다



모델이 평가한 좋은 논문-첫페이지에 그림O, 정확한 실험 결과 모형과 설명 O, colorful image O, 이들의 조화로운 밸런스

# **평가 및 한계:**

**• 데이터상, 코딩상 한계:**

모델의 학습 결과 accepted된 좋은 논문들은 전체 주제를 아우르는 그림이 첫페이지에 있으며, 다채로운 이미지나 그림이 있으며, 8페이지 꽉채운 논문들이 좋은 논문을 구성하는 요소로 판단했다는 것이다. 하지만, 예를들어 좋은 논문의 요소를 가졌지만 7페이지 밖에 도지 않을때 97.4%로 안좋은 논문으로 분류한다는 한계가 존재한다. 또한 논문에 사용된 그림이나 장표가 흐릿하거나 불투명하면 안좋은 논문이라 판단하는 경우도 있다. 더 나아가 앞서 설명했듯, 안좋은 논문을 50%확률로 확실히 reject 하기 위해 FPR 0.4%로 설정하여 좋은 논문을 reject했다는 점도 포함된다. But, 본 논문에서는 확실히 좋고 안좋은 논문을 결정하여 accept & reject해야될 논문을 걸러네 peer reviews의 업무량을 줄이는 것을 목표로 하기 때문에 완전성보다 확실히 분류 가능하다는 것에 초점을 두었기에 과감히 감내한다.

**• 시스템적 한계:**

분류를 할때 내용적인 측면이 아닌 논문의 시각적 형태(visual appearance)에만 초점을 맞췄다는 점이다. 또한, 논문의 텍스트 세팅 스타일(CVPR, ICCV등) 일정하다는 가정하에 분류했기 때문에 다른 학회 논문들에 적용하기 힘들다는 점도 있다. 추가로, 논문 pdf가 원본이 아니고 네트워크에서 다운받은 것이기 때문에 특정 논문들은 noise가 굉장히 클수있고 그에 따라 정확한 분류가 어려워진다. 마지막으로, workshop과 conference 모두를 데이터셋으로 설정했지만, workshop에서 conference로 제출된 논문 중 rejected된 논문들에 대한 정보를 얻을 접근 권한이 없다.

양식의 맨 위

양식의 맨 아래