

CAM 계열의 시각화 방법 중 어떻게 가장 유의미한가?

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 딥러닝 모델이 어떻게 이런 결과를 내는지 알아야 모델을 개선할 수 있고, 신뢰
2 할 수 있다. 이를 위해 CAM 계열의 시각화 방법이 사용되고 있다. 하지만 이를
3 이용해 얻은 시각화 결과는 방법에 따라 차이가 발생한다. 어떤 방법이 가장
4 유의미한지에 대해 비교해보려고 한다.

1 CAM 계열 시각화 방법이란?

6 기본적인 컨셉은 피쳐맵을 가지고 시각화하는 것이다. 큰 값을 가질수록 더 중요한 부분이라고
7 생각할 수 있다. 피쳐맵들을 어떻게 조합해서 시각화를 할 것인가에 따라 방법들이 달라진다.

1.1 CAM

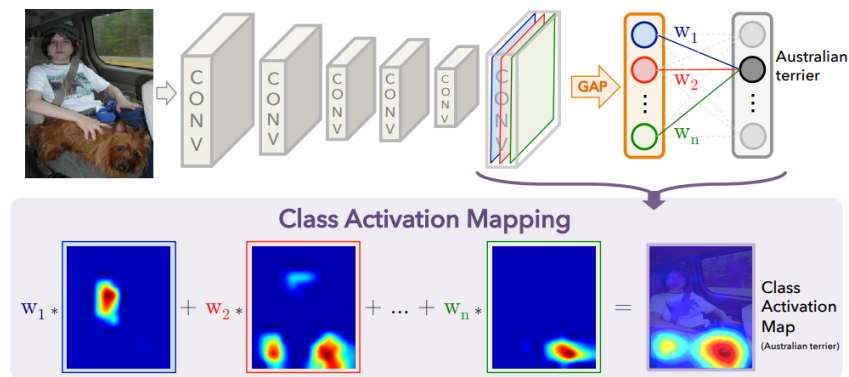


Figure 1: CAM의 모델구조

9 CAM은 figure 1과 같은 구조에서만 사용가능한 모델이다. 최종 출력층의 피쳐맵을 얼마나 가중
10 합해 예측을 하는지 학습하고, 이 가중치를 이용해 시각화하는 방법이다.

11 <https://arxiv.org/abs/1512.04150>

1.2 Grad-CAM

13 Grad-CAM은 figure 2와 같이 사용할 수 있다. 학습이 완료된 후, 클래스에 대한 기울기를 가중치
14 로 사용해 피쳐맵을 가중합하여 사용한다.

15 <https://arxiv.org/abs/1610.02391>

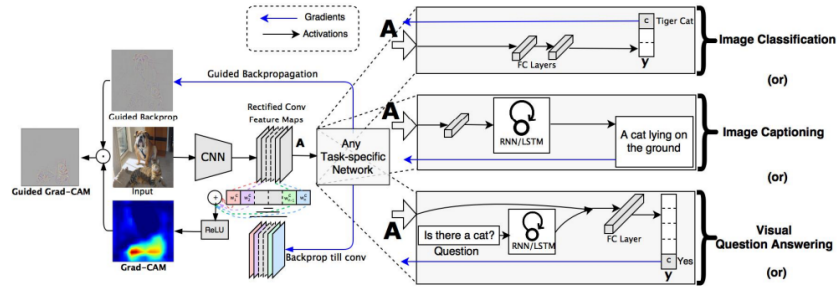


Figure 2: Grad-CAM의 모델구조

16 2 비교방법

17 "stanford dogs" 데이터셋으로 학습한 CNN 모델에 대해 CAM과 Grad-CAM 결과를 비교해본다.

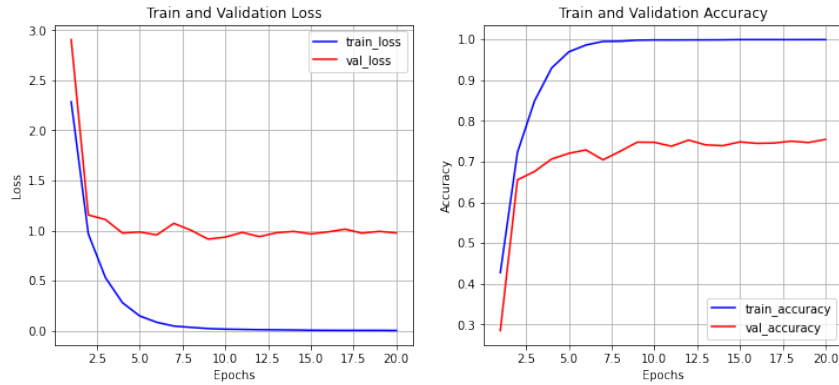


Figure 3: 모델 성능 그래프

18 모델의 성능을 보면 validation set에 대해 0.75 정도의 정확도를 보인다.

19 3 결과 비교

20 좌측이 CAM의 결과, 우측이 Grad-CAM의 결과이다. CAM의 경우 물체와 상관없는 부분에서
21 강조되는 부분이 발생하는 경우가 있었다. 하지만 그 외의 사진에서는 대부분 비슷한 결과를 보
22 였다.

23 4 결론

24 CAM과 Grad-CAM은 유의미한 차이를 보이지 않았다. CAM의 경우 특정 모델 구조를 강요하기
25 때문에, Grad-CAM이 CAM보다 더 좋은 방법이라고 판단된다.

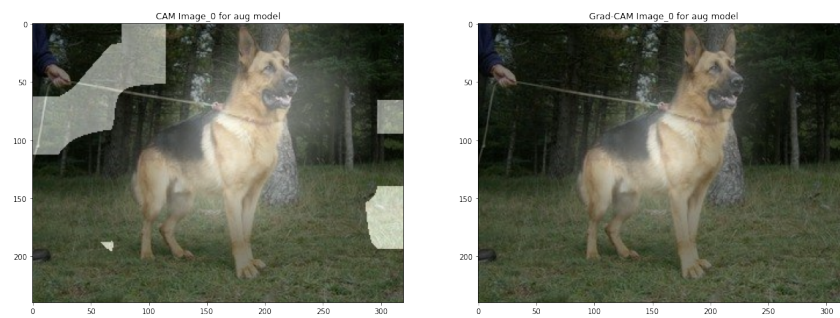


Figure 4: 결과비교

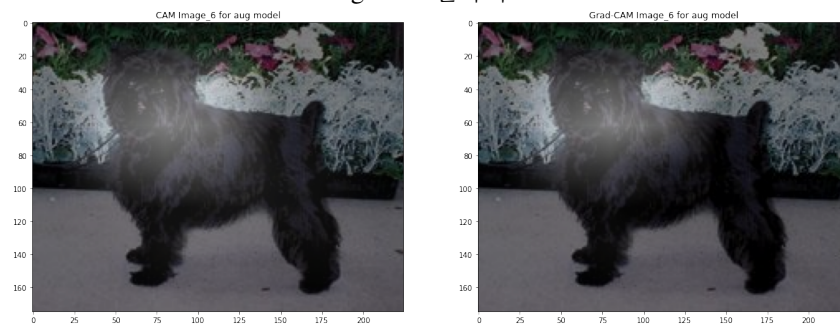


Figure 5: 결과비교 2