

사용한 학습 사진



산타 모자를 쓴 제제(1살, 강패)

1번 모델 출력 결과

---

there is a cat wearing a santa hat and holding a pair of scissors

---

고양이가 산타 모자를 쓴 것까지 정확히 맞췄지만 아래의 매듭을 가위를 들고 있는 것으로 착각

2번 모델 출력 결과

['a black and white cat wearing a red hat']

텍시도 고양이를 맞췄지만 빨간모자로 출력됨

모델 비교:

첫 모델이 비교적 세세한 부분을 캐치했다고 생각한다 (산타 모자) 하지만 잘못된 사실을 반영함 (가위 든 고양이?)

두번째 모델은 디테일 한 것은 캐치하지 못하였으나 (근데 텍시도 고양이를 맞춘걸 봐서 또 어느 정도는 예리함) 사실과 외곡된 내용은 없음 (산타 모자지만 빨간모자로 반영)

따라서 두번째 모델의 성능이 좀 더 좋다고 생각한다.