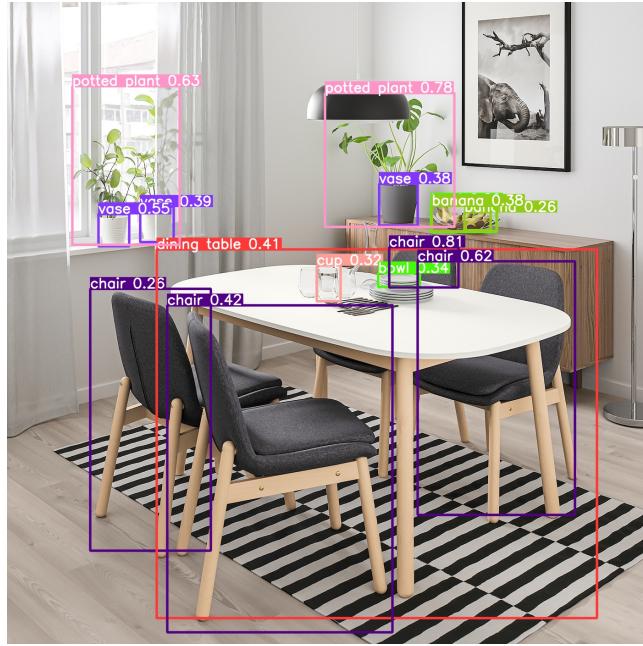


## YoloDiffusion 개선 현황

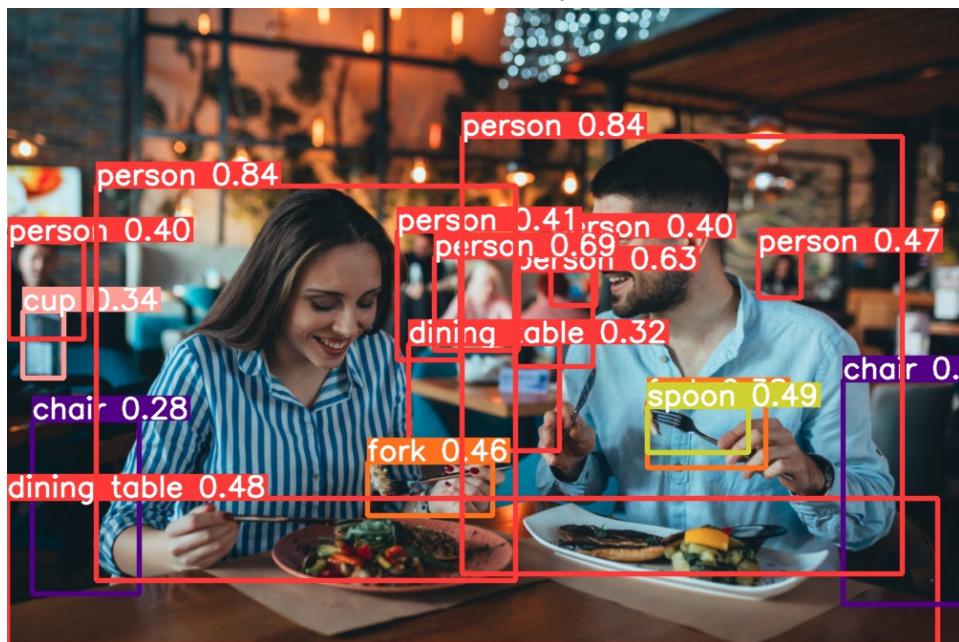
1	COCO 데이터셋 중 핵심 사물로 잘 쓰이지 않는 사물은 제외하기 ex) wheel, traffic light, parking meter, bench, hat, shoe, fork, knife, spoon etc....	무슨 사물을 뺄지 기준이 모호함. 6번 항목이 커버해줌
2	수정된 프롬프트를 GPT를 이용해 img2img 맞춘 유효한 프롬프트로 다듬기 $User\_prompt + yolo\_prompt = concat\_prompt$ $Final\_prompt = GPT(concat\_prompt)$	프롬프트가 훨씬 자연스러워짐, 이미지도 퀄리티도 향상
3	더 좋은 GPU ex) V100, A100	월 \$1, 월 \$5 속도 차이일뿐 굳이 필요없음
4	더 길어진 iterations	실험결과: 큰 의미없음
5	사용자가 strength 조절 가능(+사용자 입력 변수가 하나 더 생기는 case)	0.7~0.8 사이에서 결과물이 좋음
6	Yolo로 인식된 사물들의 크기에 따라 핵심 요소인지 아닌지 판단하여 아닌 것들은 제외, 혹은 위치를 계산하여 끝 쪽에 인식된 사물들 또한 핵심 요소에서 제외	사물의 크기만 판단, 사물의 가로와 세로가 화면의 40% 이상일 경우만 핵심 요소
7	Yolo 인식된 사물이 하나일 때는 무조건 핵심 요소로 지정	

Green : 수행함, Red : 수행 불가능, Yellow : 수행 가능하지만 일부로 뺄, No Color: 아직 안 함



원본 Prompt: bright watercolor image

원본 Prompt: fantasy mood, colorful



Watercolor image, trending on artstation

### Case 1: Softmax()를 통한 인식된 모든 사물 가중치 부여

**Initial\_prompt:** There is (banana:0.38). There is (chair:0.646). There is (bowl:0.423). There is (cup:0.414). There is (bowl:0.42). There is (banana:0.461). There is (vase:0.456). There is (vase:0.45). There is (dining table:1.457). There is (chair:0.984). There is (vase:0.525). There is (chair:0.967). There is (potted plant:0.782). There is (potted plant:0.902). There is (chair:0.734). bright watercolor image

**Gpt\_prompt:** Create a vibrant watercolor image featuring a composition of various objects. Include a (banana:0.38), a (chair:0.646), a (bowl:0.423), a (cup:0.414), a (bowl:0.42), a (banana:0.461), a (vase:0.456), a (vase:0.45), a (dining table:1.457), a (chair:0.984), a (vase:0.525), a (chair:0.967), a (potted plant:0.782), a (potted plant:0.902), and a (chair:0.734). Let the overall tone of the image be bright and cheerful.

예시 인식된 사물이 많을 경우:



다수의 사물, 사람이 중요



**Case 2: Softmax()와 Threshold를 통해 제일 가중치 높은 2~3개 사물만 명시**

**Initial\_prompt:** There is (dining table:1.457). bright watercolor image

**Gpt\_prompt:** Create a bright and vibrant watercolor painting featuring a (dining table:1.457) as the focal point, offering a visually stunning image that captures the essence of elegance and warmth.

예시



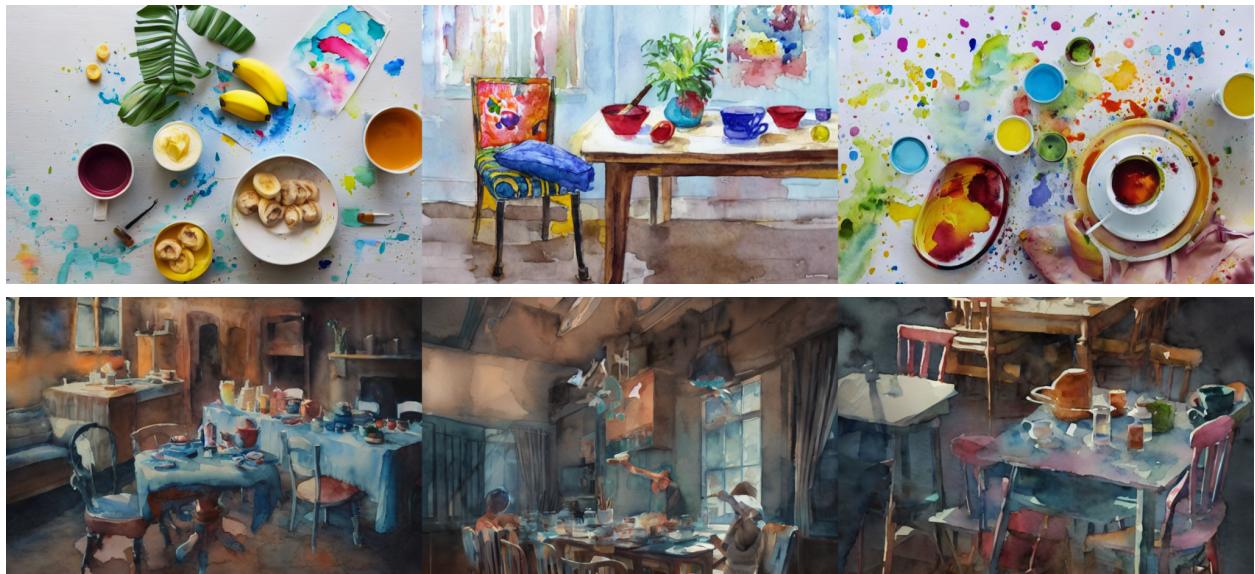


**Case 3: Softmax()**한 가중치 중 **Threshold: 0.1**이하는 가중치 **0.001**로 수정, 가중치 제일 높은 사물만 유지

**Initial\_prompt:** There is (banana:0.001). There is (chair:0.001). There is (bowl:0.001). There is (cup:0.001). There is (bowl:0.001). There is (banana:0.001). There is (vase:0.001). There is (vase:0.001). There is (dining table:1.457). There is (chair:0.001). There is (vase:0.001). There is (chair:0.001). There is (potted plant:0.001). There is (potted plant:0.001). There is (chair:0.001). bright watercolor image

**Gpt\_prompt:** Create a vibrant, bright watercolor image featuring a composition of various objects. Arrange a (banana:0.001), a (chair:0.001), a (bowl:0.001), a (cup:0.001), another (bowl:0.001), another (banana:0.001), two (vases:0.001), a (dining table:1.457), another (chair:0.001), a (vase:0.001), another (chair:0.001), two (potted plants:0.001), and one more (chair:0.001).

예시



**Case 4: ANN**을 이용해 핵심 요소 학습한 모델로 핵심 요소 판단해서 명시



### Case 5: GPT 제외

Softmax로 한정된 사울 제시



ann 으로 한정된 사울 제시



