



## 9주차

### 지난 주차 회의록 요약

1. (제안) 수작업으로 텍스트를 제거 처리한 샘플을 아웃페인팅하여 효과 확인
2. (제안) 분할 확장 시 최적의 최소 사이즈가 있는지 확인
3. (제안) 포토샵 배경 제거 테스트
4. (완료) 구글 Vision API의 성능 우수함 확인
5. (필요) 보더라인 판단 추후 개선
6. (필요) 단색 판단 개선 필요
7. (필요) 보더라인 판단 일반적 성능 개선

### 현재 메인 이슈

1. 새로운 객체나 연장선상이 아닌 새로운 컨셉의 이미지가 추가된다.
2. 원본에 테스트가 존재하는 경우 배경에 텍스트를 생성한다.
3. 마스크의 성능에 따라 사람, 상품이나 상품이 제대로 인식되지 않는다.

### 이번 주 목표

1. 리팩토링
2. 포토샵 배경 제거 적용
3. 텍스트 제거 테스트
4. 분할확장

### 진행사항

#### ▼ 리팩토링 완료

##### ▼ [utils]

- 프로세스 관련

##### ▼ MaskGenerator

- make\_mask(main)
  - make\_mask\_using\_photoshop
  - make\_mask\_using\_rembg
- ▼ BorderTouchChecker
  - check\_mask\_border(main)
- ▼ PaddingProcessor
  - addPadding(main)
  - chop\_top\_and\_bottom(main)
- ▼ DalleExpander
  - download\_image
  - outpainting\_using\_Dalle2
  - outpainting\_using\_Dalle3
  - outpainting(main)
- 전처리(단색 판단)
  - ▼ ForegroundRemover
    - remove\_foreground(main)
  - ▼ SimpleExpander
    - remove\_foreground
    - determine\_foreground\_color(main)
- 전처리(테두리 제거)
  - ▼ BorderRemover
    - remove\_border(main)
  - ▼ BorderRemoverV2
    - 우경창님 알고리즘 클래스
  - ▼ BorderRemoverV1
    - 기존 알고리즘 클래스
- 전처리(테두리 추가 - 보류)
  - ▼ WhiteBorderAdder
    - add\_white\_border(main)
- ▼ [processor]
  - 테스트 모듈

#### ▼ ImageProcessor

- process\_image
- single\_process\_image
- batch\_process\_image

test.py

config.py

#### ▼ 텍스트 제거 후 확장 결과 확인

- 이미지 블러: 텍스트는 거의 제거 되었지만, 블러의 흔적이 확장에 부작용을 주는 것 같음
- 투명 픽셀 대체: 텍스트는 거의 제거 되었지만, 투명 픽셀이 확장에 부작용을 주는 것 같음.
- 특이사항: 로고 같은 경우에 사전에 지우지 않았지만 텍스트와 비슷한 효과를 불러일으키는 것 같음

#### ▼ 포토샵 배경 제거 적용

자동화 완료

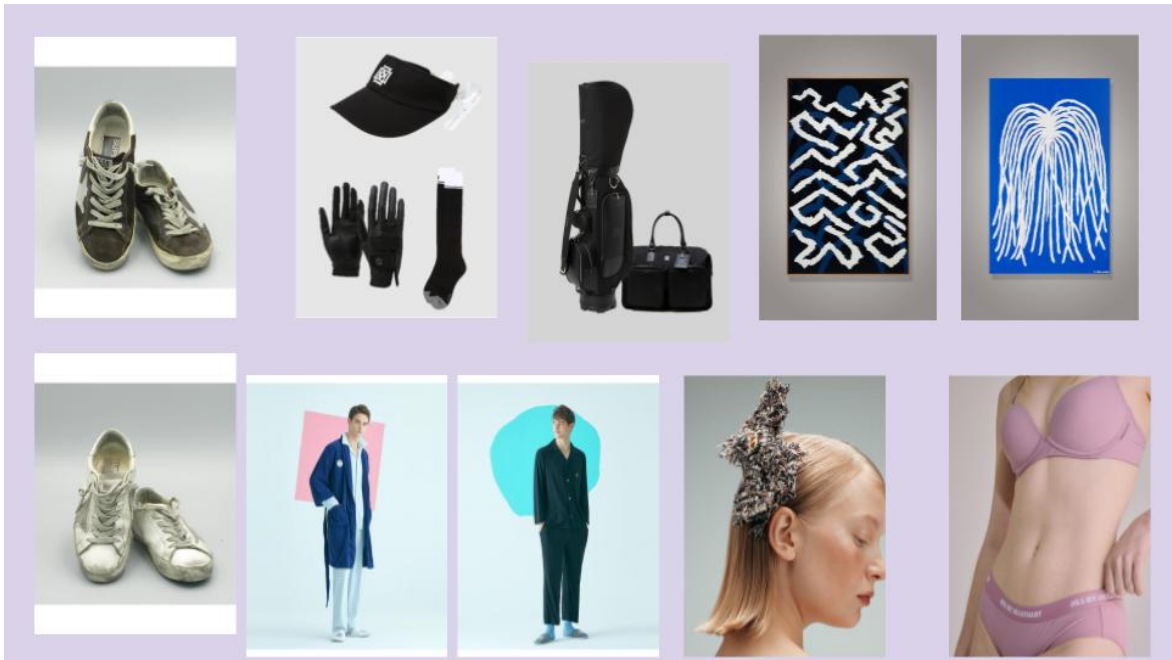
#### ▼ 분할 확장 결과 확인

- 샘플 추출 조건
  1. 배경 확장에 실패했던 샘플
  2. 텍스트나 마스크 이슈가 없는 샘플
  3. 이론상 배경확장에 무리가 없어야할 샘플

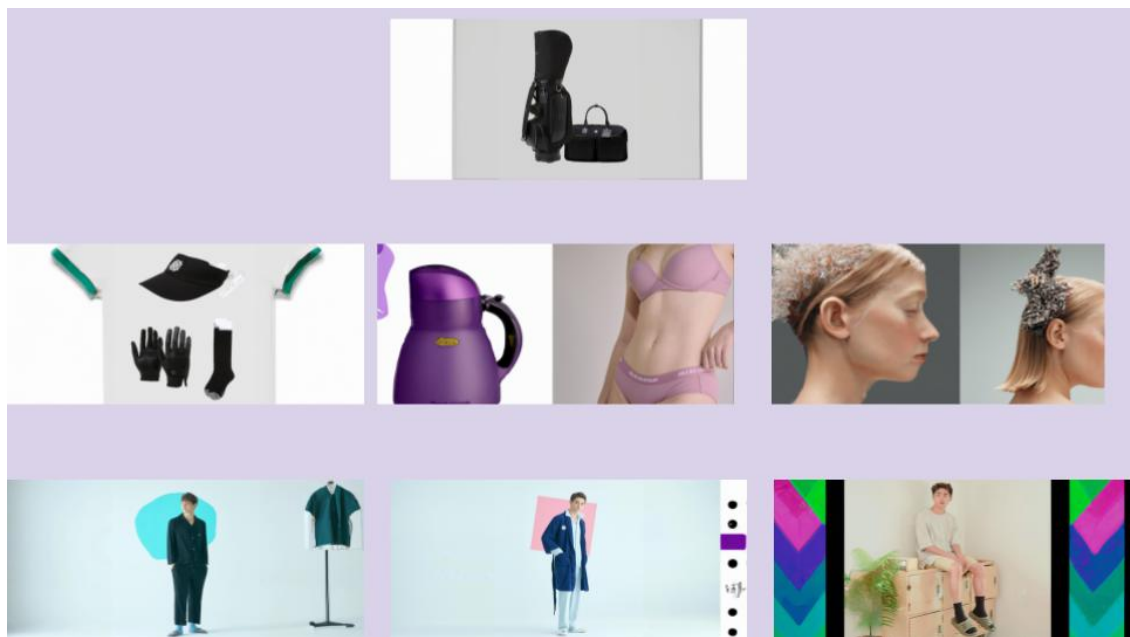
### 테스트 샘플 선택

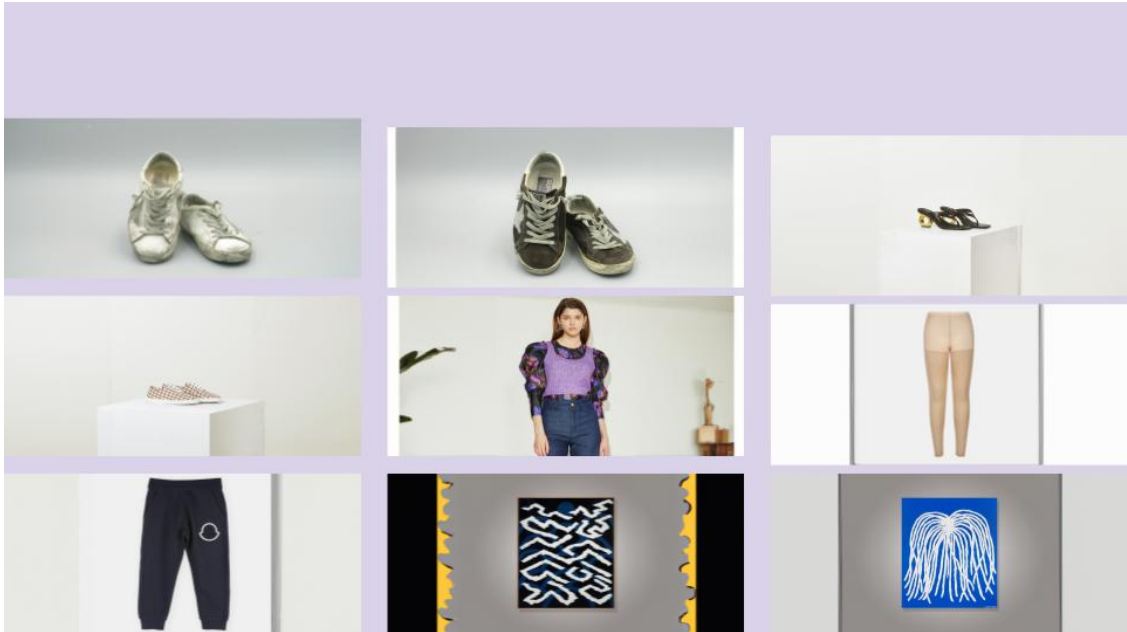
- 텍스트가 없으면서, 마스크 잘 인식되는 이미지
- 그럼에도 안 좋은 결과를 도출했던 이미지





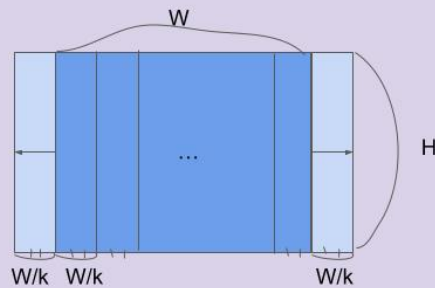
- 기존 알고리즘으로 재시도 결과





- 분할 원리

- 2분할 확장
- 원본 이미지의 가로 길이의  $1/k$ 에 해당하는 길이만큼 양옆으로 확장
- 이후, 1:2 비율로 확장



1. 이미지를  $k$ 등분한 만큼의 길이만큼 양 옆에 각각 확장
2. 나머지 영역 확장

- 중간 과정 확장 결과

k	원본	3	4	5	7
good	x	8	9	<b>12</b>	<b>11</b>
bad	x	8	7	4	5
success rate	x	50%	56.25%	<b>75%</b>	68.75%

- 가설: 잘게 자를 수록 좋아야 한다.
- 결론: 그러한 경향을 보이지만, 최대한 작게 잘라도 나쁠 때가 꽤 많다.

• 최종 확장 결과(2분할)

k	원본	3	4	5	7
good	6(3)	5(2)	3(2)	9(4)	x
bad	10	11	10	6	x
success rate	37.5%	31.25%	18.75%	56.25%	x

평가 기준: 전체적으로 어색하지 않고 쓸만한 경우 good, 괄호 내부에는 완벽한 이미지 개수

- 가설: 특정 k에서 최적의 결과를 낼 것이다.
- 결론: k=5에서 가장 좋은 결과를 얻었지만 가장 가까운 k=4와 차이가 많이나므로 추가적인 검증이 필요하다.

• 중간확장 성공 후 성공할 확률

k	원본	3	4	5	7
good	x	5	3	9	x
bad	x	8	6	3	x
success rate	x	62.5%	33.33%	75%	x

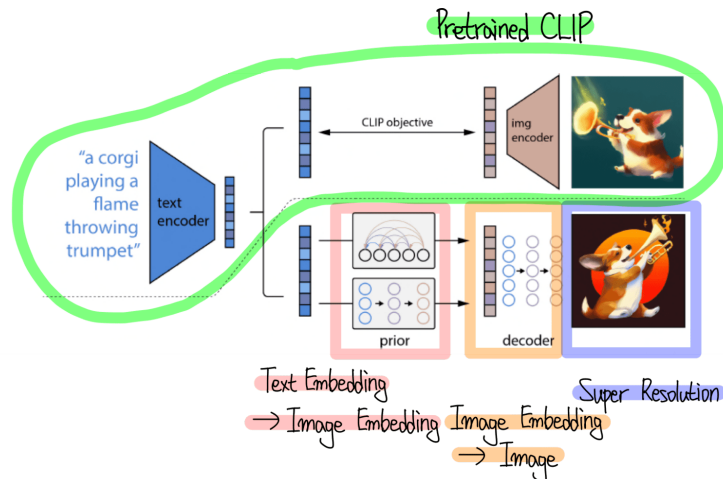
- 가설: 확장 영역이 넓을 수록 결과가 안좋을 것이다.
- 결론: 가설의 예측이 완전히 일치하지 않는다.

• 결론

- 예상외로 유의미한 결론이 없었으며, 운의 요소가 너무 크게 작용하는 것으로 보인다.
- 확실한 결론을 얻기 위해서는 다양한 샘플에 대해 횟수를 늘려서 검증해야 할 것 같다.
- 추가적인 검증 방안으로는 6등분 확장 3번 반복이 있을 것이다.
- $((640/813) * (8/6))^3 = 1.86 \rightarrow$  마지막 확장은 2:1에 맞춰서

▼ 프롬프트 개선방안

## 논문 요약



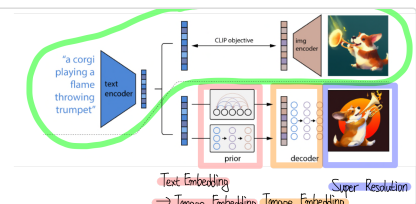
기본적으로 Dall-E는 Open AI에서 개발한 CLIP이라는 성공적인 이미지 학습 모델을 사용한다. CLIP은 자연어 또는 이미지를 의미론적으로 해석하여 임베딩을 생성할 수 있다.

Dall-E는 CLIP을 통해 인코딩된 이미지와 텍스트 간에 연관성을 파악하여 공동 표현 공간(joint representation space)이 학습된 모델이다. CLIP에 의해 인코딩된 텍스트 임베딩은 prior에서 새로운 이미지 임베딩으로 생성된다. 이것이 diffusion decoder 과정을 거치고, 해상도를 회복해서 최종 이미지가 생성되는 것이다.

### DALLE2 - Diffusion Model 논문 리뷰

DALLE2 논문의 핵심 내용을 리뷰합니다. 먼저 기존 방법의 문제점을 살펴봅니다. 이어서 이를 해결하기 위한 DALLE2의 제안 방법을 살펴봅니다. 마지막으로 성능 비교 실험을 통해 DALLE2의 효과를 확인합니다.

<https://fighting.net/deep-learning-paper-review/diffusion-model/dalle2/>



## 프롬프트 테스트

첫 시도: "Simply extend background without introducing any new objects or texts."

### ▼ 테스트 프롬프트 9가지

1. 공백 (5점)
2. Simply extend background(9점)
3. Simply extend background widely without any new objects or texts(10점)
4. Simply extend monochromatic tone background widely without any new objects or texts.(10점)
5. high-quality banner image.(7점)
6. Minimalism style background.(12점)
7. Minimalism style background without introducing any new objects or texts.(10점)
8. Wide-angle shot with ample whitespace(9점)

## 9. Extend background of Minimalist style whitespace photography.(13점)

### ▼ 테스트 코드

```
prompt_list = [
    " ",
    "Simply extend background.",
    "Simply extend background widely without any new objects or texts",
    "Simply extend monochromatic tone background widely without any new objects or texts.",
    "high-quality banner image.",
    "Minimalism style background.",
    "Minimalism style background without introducing any new objects or texts.",
    "wide-angle shot with ample whitespace",
    "Extend background of Minimalist style whitespace photography."
]
for idx, prompt_text in enumerate(prompt_list):
    ImageProcessor.batch_process_images(target_folder, output_folder_name= "promptTest"+str(idx),
                                       mask_folder=mask_folder, percentage = BATCH_PERCENTAGE,
                                       key=DALLE_KEY, prompt_text=prompt_text)
```

### ▼ 채점결과


1. xxx/xxx/xxx/xxx/xxx/xxo/ooo/xxx/xox/xxx = 5
2. xoo/xxo/oxo/oxo/xxx/oxo/xoo/xxx/xxx/oxo = 9
3. Xxo/xxx/oxo/ooo/xxx/xxx/ooo/xxo/xxx/xxo = 11
4. xxo/xxx/xxx/ooo/xxo/oxo/ooo/xxx/xxx/xxx = 10
5. oxx/xxx/xox/oxo/xxx/xxx/xox/xxx/oxo/xxo = 7
6. xxo/oxo/ooo/ooo/xxx/xxo/oox/xxx/xxo/xxx = 12
7. xox/oxo/oxo/oxo/xox/oox/xoo/xxx/xxx/oxo = 10
8. Xox/xxx/ooo/ooo/xxo/oxo/ooo/xxx/xxx/xxx = 9
9. Xoo/oxo/xoo/ooo/oxo/xox/ooo/xxx/xxx/oxo = 13

## 참고자료

### 1. 일정한 이미지를 배출하는 프롬프트(OpenAI community)

#### Ways to prompt consistent image generation for DALL-E API

Hey, I was wondering if there are any tips or tricks for prompting the DALL-E image api to create consistent image prompts that are not too far out there for generating course images that are not too abstract and crazy for business use.

 <https://community.openai.com/t/ways-to-prompt-consistent-image-generation-for-dall-e-api/354196>

#### OpenAI Developer Forum



Currently DALL-E results are not good as others like MidJourney or BlueWillow may in future they may improve but for as a user we have to provide clear, specific, and detailed instructions. Use adjectives, nouns, and references (both other prompts and images) when needed, and iterate as necessary to refine your results.

Example prompt:

```
Generate a high-resolution, 1920x1080 pixel image of a contemporary office boardroom with large windows, a long mahogany conference table surrounded by ergonomic chairs, modern artwork on the walls, and a view of a city skyline through the windows. The room should be well-lit with a professional, minimalist design
```



n, and there should be a laptop, notepads, and a pitcher of water with glasses on the conference table. The overall ambiance should convey a sense of professionalism and corporate sophistication.

## 2. 텍스트 추가를 막기위한 방안(OpenAI community)

### Keep Dalle from including text

Hey all, I'm trying to make some potential logos, but text is always included (and as we know Dalle text tends to be gibberish). As much as I say, "The picture has no text" or "there are no words in the logo" Dalle always seems to want them in

 <https://community.openai.com/t/keep-dalle-from-including-text/32556>

OpenAI Developer Forum



I used someones suggestion to avoid negative prompting and gave a bit more instruction on what I actually wanted - even though it is abstract. I'm also trying to do it programmatically so having some good prompt ideas would be helpful.

"create a square business logo containing simple colours and shapes inspired by artificial intelligence"

I didn't get any text

## 메모

#아이디어1

배경만을 남긴 상태에서 DALL-E 모델 확장하기

#아이디어2

리사이징을 전혀 하지 않고 DALL-E를 양옆으로 각각 확장하면 어떨까? (동영상 참고)

예시) Patch-E

이미지 확장 온라인 툴

## 이번 주 회의록

### 1. 진행중인 과제

#### a. 원본사진에서 텍스트 인식 및 제거 후 확장

i. 테스트 완료, 결과 좋지 않아 사용하지 않기로 함

ii. 블러, 투명 처리한 결과도 인지해서 확장할때 반영

#### b. 분할해서 확장 한 경우 품질 개선 여부 테스트 추가 테스트

• 작게 먼저 확장해서 2회 확장하기 -> 도움이 되지 안됨

• 잘게 여러번 추가로 해보기로 함(2)

#### c. 포토샵 배경제거 테스트

i. 달리 확장 테스트시 최종 결과 생성 가공시 품질개선에 도움될것으로 예상

ii. 추가 테스트 필요

iii. 달리2 프롬프트 개선 제안 - **google vision** 등으로 자동 판단해서 결과 보기로 함(1)

d. 마스크 이용한 배경만으로

i. 확장해서 테스트해보기(3)

ii. 단색판단 해보기(4)

2. 향후과제

a. 보더라인 판단 개선

b. 단색 판단 개선

c. 배경 제거후 배경교체( 전체 생성) 품질 리서치 - dalle, photoshop 등 적용

3. 소스 코드 및 산출물 - git으로 관리 후 주간 미팅때 git 자료로 공유