

OpenAI DALL-E API를 활용한 이미지 생성 및 편집 튜토리얼 (2025년 3월 기준)

이 Jupyter Notebook에서는 OpenAI의 DALL-E 모델 API를 활용하여 텍스트로부터 이미지 생성 및 기존 이미지 편집(Inpainting/Outpainting) 기능을 구현하는 방법을 소개합니다.

주요 학습 내용:

1. DALL-E 모델 개요 및 환경 설정
2. OpenAI API 클라이언트 생성
3. 기본 이미지 생성
4. 이미지 편집

1. DALL-E 모델 개요

DALL-E는 OpenAI가 개발한 최첨단 이미지 생성 AI 모델로, 텍스트로 설명한 내용을 바탕으로 새로운 이미지를 만들어냅니다. 예를 들어 `"우주비행사가 말을 타고 있는 사진"`이라는 프롬프트를 주면, 우주비행사가 말 위에 올라탄 독창적이고 현실감 있는 이미지를 생성할 수 있습니다. DALL-E 2 모델은 자연어 설명만으로 사실적인 이미지와 예술 작품을 생성할 수 있으며, 여러 개념과 색상, 스타일을 조합하여 창의적인 결과물을 만들어냅니다.

2023년에 공개된 DALL-E 3 모델은 한층 발전된 기능을 제공합니다. 이 모델은 더 복잡한 프롬프트를 이해하고 이미지 안에 텍스트(예: 간판의 글자 등)까지 생성할 수 있으며, 가로 또는 세로로 긴 다양한 비율의 이미지도 지원합니다. 또한 이전보다 상세하고 정교한 이미지를 만들어내며, 프롬프트를 자동으로 보완하여 최적의 결과를 얻도록 도와줍니다. 이러한 DALL-E 시리즈 모델들은 디자인, 예술 창작, 시각화 등 다양한 분야에서 활용되고 있습니다 (예: 광고 시안 제작, 게임 컨셉 아트, 시나리오 장면 묘사 등).

2. OpenAI API 클라이언트 생성

위 코드에서는 .env에서 불러온 API 키로 client 객체를 생성했습니다. 이제 이 client를 통해

DALL-E를 비롯한 OpenAI API 요청을 보낼 수 있습니다. (참고로, `api_key`를 생략하면 `OPENAI_API_KEY` 환경 변수를 자동으로 참조합니다.)

3. 기본 이미지 생성

텍스트 프롬프트를 사용한 이미지 생성은 DALL-E API의 가장 기본적인 기능입니다. 앞서 생성한 OpenAI client를 사용하여 원하는 내용을 설명하는 프롬프트를 전달하면, 해당 설명에 부합하는 이미지를 얻을 수 있습니다.

이미지 생성 요청하기

OpenAI Python SDK에서는 이미지 생성을 위해 `client.images.generate()` 메서드를 제공합니다. 가장 간단한 형태로 이 메서드에 프롬프트 문자열만 전달하면 한 장의 이미지를 생성합니다. 예를 들어 다음 코드는 "숲이 보이는 밝은 자연 이미지를 그려줘"이라는 간단한 프롬프트로 이미지를 생성하는 요청을 만듭니다:

위 코드에서는 `prompt`에 설명 텍스트를 넣고, `n=1`로 이미지 한 장을 요청했습니다. `size`는 원하는 해상도를 지정하며, "256x256"로 설정하여 256픽셀 정사각형 이미지를 얻도록 했습니다. DALL-E 2 모델의 경우 256x256, 512x512, 1024x1024의 해상도를 지원하며 필요한 경우 선택할 수 있습니다. `client.images.generate()` 호출 시 응답 객체(`response`)에는 생성된 이미지에 대한 정보가 담기며, `response.data` 리스트 내부에 각 이미지 결과가 포함됩니다. 각 결과 항목에서 `.url` 속성을 통해 생성된 이미지의 URL을 얻을 수 있습니다.

참고: OpenAI 이미지 생성 API의 기본 출력은 URL 형태입니다. 즉, 생성된 이미지는 일시적으로 접근 가능한 클라우드 URL로 제공되며, 위 예시처럼 `response.data[0].url`로 URL을 받아 사용할 수 있습니다. 이 URL을 웹 브라우저에서 열거나 Python 코드로 다운로드하여 이미지를 확인할 수 있습니다. 만약 바이너리 이미지 데이터(base64 인코딩 형태)로 직접 받고 싶다면 `response_format="b64_json"` 옵션을 추가할 수 있습니다. 그러면 응답에서 `url` 대신 `b64_json` 필드로 이미지 데이터가 제공되며, 이를 디코드하여 이미지로 저장하거나 표시할 수 있습니다.

위 코드에서는 `prompt`에 설명 텍스트를 넣고, `n=1`로 이미지 한 장을 요청했습니다. `size`는 원하는 해상도를 지정하며, "512x512"로 설정하여 512픽셀 정사각형 이미지를 얻도록 했습니다. DALL-E 2

모델의 경우 256x256, 512x512, 1024x1024의 해상도를 지원하며 필요한 경우 선택할 수 있습니다. `.client.images.generate()` 호출 시 응답 객체(response)에는 생성된 이미지에 대한 정보가 담기며, `response.data` 리스트 내부에 각 이미지 결과가 포함됩니다. 각 결과 항목에서 `.url` 속성을 통해 생성된 이미지의 URL을 얻을 수 있습니다.

참고: OpenAI 이미지 생성 API의 기본 출력은 URL 형태입니다. 즉, 생성된 이미지는 일시적으로 접근 가능한 클라우드 URL로 제공되며, 위 예시처럼 `response.data[0].url`로 URL을 받아 사용할 수 있습니다. 이 URL을 웹 브라우저에서 열거나 Python 코드로 다운로드하여 이미지를 확인할 수 있습니다. 만약 바이너리 이미지 데이터(base64 인코딩 형태)로 직접 받고 싶다면 `response_format="b64_json"` 옵션을 추가할 수 있습니다. 그러면 응답에서 url 대신 `b64_json` 필드로 이미지 데이터가 제공되며, 이를 디코드하여 이미지로 저장하거나 표시할 수 있습니다.

DALL-E 3 모델 사용 (선택 사항)

앞서 언급했듯, OpenAI는 2025년 기준 두 가지 DALL-E 모델을 API로 제공합니다: DALL-E 2(기본값)와 DALL-E 3. 별도 모델 지정을 하지 않으면 DALL-E 2가 사용되는데, DALL-E 3의 향상된 기능을 활용하려면 `model` 파라미터를 지정해야 합니다. 예를 들어 DALL-E 3로 이미지를 생성하려면 위 요청을 다음과 같이 변경합니다:

- `quality`: DALL-E 3 전용 옵션으로, "standard" (기본값) 또는 "hd" 값을 가집니다. "hd"는 더 높은 이미지 품질을 제공하지만 생성 비용과 시간이 증가합니다. "standard"는 빠르고 비용이 저렴하지만 품질이 다소 낮습니다.

- `style`: 마찬가지로 DALL-E 3의 추가 옵션입니다. "vivid"(선명하고 풍부한 색감, 기본값) 또는 "natural"(더 자연스러운 색감과 톤) 중 선택하여 이미지의 스타일 분위기를 조정할 수 있습니다. 원하는 결과에 따라 어떤 스타일이 적합한지 실험해볼 수 있습니다.

주의: DALL-E 3 모델은 현 시점에서 API 호출 한 번에 `n=1` (이미지 1장)만 생성을 지원합니다. 즉, 한 요청으로 여러 장의 이미지를 받고 싶다면 DALL-E 2를 사용하거나, DALL-E 3의 경우 반복 호출을 통해 이미지를 여러 장 받아야 합니다. 또한 DALL-E 3는 출력 이미지 해상도로 1024x1024 (정사각형), 1024x1792 (세로형 직사각형), 1792x1024 (가로형 직사각형)만 지원합니다. 작은 해상도(256 또는 512 크기)가 필요하면 DALL-E 2 모델을 사용해야 합니다. DALL-E 2의 경우 `n` 파라미터로 한 번에 최대 10개까지 이미지를 생성할 수 있으므로, 여러 시안을 비교하고 싶을 때 활용하면 유용합니다.

4. 이미지 편집

이제 기존 이미지를 편집하는 방법을 알아보겠습니다. DALL-E API는 이미지 편집(edit) 기능을 통해 ****인페인팅(Inpainting)****과 ****아웃페인팅(Outpainting)****을 지원합니다. 인페인팅은 이미지의 일부 영역을 다른 내용으로 대체하거나 채우는 것이고, 아웃페인팅은 이미지 바깥의 빈 공간을 주변과 어울리도록 확장 생성하는 기술입니다. 예를 들어, 사진 속 하늘을 밤하늘로 바꾸거나(인페인팅), 그림의 캔버스를 넓혀 배경을 추가로 그려넣는 것(아웃페인팅)이 가능합니다.

Inpainting - 이미지 부분 수정하기

인페인팅을 위해서는 기본적으로 두 개의 이미지 입력이 필요합니다: 하나는 편집의 기준이 될 원본 이미지, 다른 하나는 편집할 영역을 지정하는 마스크 이미지입니다. 마스크(mask)는 수정하고자 하는 영역을 나타내는 투명한 부분과, 유지할 부분을 나타내는 불투명한 부분으로 구성된 PNG 이미지를 사용합니다. DALL-E API에 원본 이미지와 마스크 이미지를 함께 제공하고, 어떤 내용을 추가/변경할지 프롬프트로 설명하면, 모델이 해당 부분을 자연스럽게 채워주는 방식입니다.

예를 들어, 하늘이 맑은 풍경 사진에서 하늘 부분에 구름을 추가하고 싶다면, 원본 사진과 함께 하늘 영역만 투명하게 처리된 마스크 이미지를 업로드하고 프롬프트에 "하늘에 구름 추가"와 같이 지시합니다. 그러면 모델이 투명한 부분(하늘 영역)을 인식하여 그 부분에만 구름을 그려넣고, 나머지 영역은 원본 그대로 유지합니다.

OpenAI Python SDK에서는 `client.images.edit()` 메서드를 사용하여 이미지 편집 요청을 보낼 수 있습니다. 사용법은 다음과 같습니다:

위 코드에서 `original.png`는 편집 대상이 되는 원본 이미지이고, `mask.png`는 편집할 부분(하늘 영역)이 투명 처리된 마스크 이미지입니다. `client.images.edit()` 호출 시 `prompt`에 원하는 편집 내용을 서술하면 모델이 해당 부분을 반영한 이미지를 생성합니다. 결과는 이미지 생성 때와 마찬가지로 `response.data[0].url`에 편집된 이미지의 URL로 반환됩니다.

원본 이미지와 마스크 이미지에 대한 DALL-E API 요구사항

- RGB로 변환 (RGBA ✕ → RGB ✔)
- 정확한 1024x1024 크기 조정
- PNG 최적화 적용 (파일 크기 4MB 미만으로 유지)
- 마스크 이미지를 정확한 1비트(흑백)로 변환 (L 모드 ✕ → 1 모드 ✔)