

캡스톤디자인 II 계획서

| | | | | |
|--|---|--|---------------|----------------------|
| 제 목 | 국문 | 캡션을 이용한 이미지(일러스트) 생성 구현 | | |
| | 영문 | Image Generating from Caption Sentence | | |
| 프로젝트 목표 | <p>프로젝트의 목표는 한국어로 이루어진 단어나 문장(캡션)을 입력받아 그림과 같은 느낌의 이미지를 생성하는 기능을 갖추고 해당 기능을 웹 페이지에 포함하여 결과물을 보이는 것으로, 이전의 캡스톤디자인 I 에서는 필요한 학습과 기능적 요소의 구현에만 집중하여 기술적인 도전과제를 도외시하였던 단점이 있으며 차후 단계에서는 해당 단점의 보완을 위한 모델의 탐색/개선 또는 스타일의 추가 등 기술적인 탐구와 함께 사용을 위한 배포를 위한 구현을 계획 중이다.</p> | | | |
| 프로젝트 내용 | <p>목표에서 제시한 점을 기반으로 이어질 캡스톤디자인 II에서의 내용은 크게 두 가지로 나눌 수 있으며 하나는 기술적인 도전과제의 탐색 및 진행, 두 번째는 사용자를 위한 접근성을 확보할 웹 페이지 제작이 될 예정으로 전자에서는 컴퓨터 비전적인 요소에 집중해서 학습 데이터의 종류를 더함으로써 접근할 수 있는 스타일을 추가하고 더 나은 모델구조의 탐색이나 파라미터의 변경을 통해 사용된 인공지능 모델의 개선을 추구할 것이며 후자에서는 node.js, Django 등 다양한 웹 구현 관련 프레임워크를 비교하며 더 유용하다고 판단한 것을 취사선택하여 구현에 도전할 예정이다.</p> | | | |
| 기대효과 | <p>국내에는 존재하지 않던 일반 사용자를 대상으로 삼은 인공지능을 이용한 그림제작 서비스를 준비하는 것으로 해당 기술의 인식 확산과 함께 스타일을 추가하는 과정을 통해 추후 다양한 서비스의 확장을 위한 일차적인 기반을 만들고, 관련 시장에 파급효과를 가져올 것을 기대할 수 있다.</p> | | | |
| 중심어(국문) | 웹 | 컴퓨터 비전 | 딥 러닝 | 자연어 처리 |
| Keywords (english) | Web | Computer Vision | Deep Learning | NLP |
| 멘토 | 소속 | - | 이름 | - |
| 팀 구성원 | 학년/반 | 학 번 | 이 름 | 연락처(전화번호/이메일) |
| | 4 | 20161616 | 연제원 | 010-4924-9552 |
| | 4 | 20197124 | 하상욱 | 010-7468-8864 |
| | 4 | 20198028 | 이창현 | 010-8891-7629 |
| <p>컴퓨터공학과와 캡스톤디자인 관리규정과 모든 지시사항을 준수하면서 본 캡스톤디자인을 성실히 수행하고자 아래와 같이 계획서를 제출합니다.</p> <p style="text-align: center;">2022 년 7 월 1 일</p> <p style="text-align: right;">책 임 자 : 연제원 (인) 지도교수 : 임경태 교수님 (인)</p> | | | | |

캡스톤디자인 계획서

1. 캡스톤디자인의 배경 및 필요성

프로젝트의 배경은 해외의 DALL-E, GauGAN의 사례를 보고 저희 중 그림에 관심이 있던 조원이 원화가를 지망하는 지망생 대부분이 캐릭터를 그려서 일러스트 생성분야는 항상 인력이 부족하다는 업계인의 언급과 자신이 인터넷 상에서 커미션(일종의 그림 제작 주문)을 신청하고자 하여도 높은 가격과 신뢰도의 부족으로 망설였던 경험을 바탕으로 일종의 일러스트 제작기를 프로젝트 주제로 제안하였던 것이 기반이 되었다.

2. 캡스톤디자인 목표 및 비전

목표는 한국어 단어, 문장으로 접근할 수 있는 그림 생성 프로젝트를 제작하는 것으로 국내의 사용자로 하여금 더 간단하게 해당 서비스에 접근하고 사용하도록 돕고 그에 더해 계획하고 있는 데이터의 추가를 통한 화풍의 추가가 제대로 진행된다면 그를 기반으로 해서 이후에 필요한 스타일을 지속적으로 추가해가며 특정 화풍의 그림이 특별히 필요한 상황이나 관련 상품의 제작 등에서 유용하게 사용될 수 있을 것으로 관련된 경제적 효과를 기대할 수 있을 것으로 기대 중이다.

3. 캡스톤디자인 내용

주요 기능은 Text to Image Generation으로 사용자의 입력에 맞추어 새로운 이미지를 생성하는 프로그램을 기획함. 새로운 이미지를 생성한다는 목적에 맞게 이미지 생성 모델인 Glide를 도입하였고 기존에 이미 사용한 Sketch-Generation 모델 외에 화풍을 선택할 수 있도록 모델의 수정이나 다른 모델의 추가를 진행하고 웹 프레임워크를 통해 사이트를 제작하여 그를 거친 입력과 결과물 확인이 가능하도록 진행할 예정이다.

결과물의 평가 기준을 놓고서 고민이 있었고 육안을 거친 평가가 유용하다고 판단했으나 정량적 평가 기준이 필요하다면 여러 이미지를 종합하여 사람이 평가하는 것과 비슷한 결과를 보인다는 장점이 있는 FID와 보다 직관적으로 이미지의 품질을 표현한다는 장점이 있는 IS를 병용할 계획이고 웹사이트의 구성에서는 되도록 간결한 UI를 채택하면서 기능을 가진 모듈과의 연결이 원활하게 이루어지도록 제작할 예정이다.

4. 캡스톤디자인 추진전략 및 방법

이후의 계획 추진에서 겪을 것으로 생각되는 문제점으로는 팀 협업 시 정보 및 업무 공유의 부족, 최종적인 이미지 생성의 정확도 문제 등이 있으며 이를 해결하기 위해 고려중인 방안으로는 팀 프로젝트를 위한 별도의 앱을 활용하고 수시로 대화를 가지는 것과 함께 정확도에 대한 문제는 모델을 가능한 만큼 개선해 볼 계획이고 기존의 프로젝트 진행을 통해 준비된 기능 모듈들과 얻게 된 지식을 활용하여 정했던 목표인 인공지능을 활용한 일러스트 생성을 더 폭넓게 구현할 생각이며, 프로젝트의 진행을 위한 각 팀원의 역할 범위는 1) 연제원 : 프로젝트 구상 및 웹 제작, 2) 하상욱 : 모델의 탐구 및 개선, 3) 이창현 : UI 제작과 백엔드 준비이며 이상의 역할을 바탕으로 프로젝트를 진행하되 지난 프로젝트의 진행에서 각자 경험과 지식의 부족이 문제가 되었던 바가 있어 서로 적극적으로 기술과 정보를 공유하고 도움을 주어가며 진행될 예정이다.

5. 캡스톤디자인 결과의 활용방안

현재 이 프로젝트 자체적으로는 생각할 수 있는 경제적 가치가 제한적이지만 모델의 개선 및 스타일 추가를 위한 기반이 제대로 준비된다면 그림의 창작이나 제작을 필요로 하는 다양한 분야에서 소요되는 비용을 절감하여 더 큰 경제적 파급효과를 생각할 수 있고 또한 각 개인의 창작에 대한 참여의 문턱을 낮추고 인공지능과 미술의 결합이라는 생소한 분야에 사례를 더한다는 의의가 있다고 생각한다.

6. 참고문헌

[1] [<https://arxiv.org/abs/1903.07291>]

(****Semantic Image Synthesis with Spatially-Adaptive Normalization****)

[2][<https://github.com/TZYSJTU/Sketch-Generation-with-Drawing-Process-Guided-by-Vector-Flow-and-Grayscale>]

[3] <https://arxiv.org/abs/2112.10741>

(****GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models****)