

캡스톤디자인 I 계획서

제 목	국문	인공지능을 활용한 일러스트 생성기					
	영문	Illust Generator using A.I.					
프로젝트 목표 (500자 내외)	구상한 프로젝트의 주제는 인공지능을 활용한 일러스트 생성기입니다. 게임 제작이나 개인방송의 화면, 그림의 창작 등에 활용될 수 있는 일러스트를 추가적인 비용과 시간을 소비할 필요 없이 즉석에서 제작할 수 있게 만드는 것이 목적입니다. 사용자가 원하는 그림의 주제와 특징을 설명한 문장을 한글로 입력하면 특징 키워드를 추출하고 학습된 이미지 데이터셋과 연결하여 선택된 이미지들로 심층학습 과정을 진행하여 이미지를 생성 후 출력합니다. 최종적으로는 이를 웹사이트의 형태로 구현하여 관련 서비스의 수요를 가진 사용자에게 대한 접근성을 확보하는 것이 이번 프로젝트의 목표입니다.						
프로젝트 내용	파이썬 언어를 기반으로 자연어처리에는 Konlpy 라이브러리를 응용하여 키워드를 추출하고 데이터셋 준비에는 웹 크롤링을 활용하여 데이터를 확보한 뒤 전처리, 이미지 학습모델 구현에는 PyTorch를 기반으로 심층신경망을 제작해서 학습, 마지막으로 웹사이트 구현에는 Django 프레임워크를 사용하며, 간편한 UI로 사용하기 편한 것을 특징으로 프로젝트의 기본적인 목적을 달성한 후 오류의 최소화와 성능의 향상에 집중할 예정입니다.						
중심어(국문)	딥러닝	컴퓨터 비전	자연어 처리	웹 크롤링			
Keywords (english)	Deep Learning	Computer Vision	Natural Language Processing	Web Crawling			
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;">2021 년 3 월 9 일</p> <p style="text-align: right;">책임자 : 연제원 (인)</p> <p style="text-align: right;">희망 지도교수 : 임경태 교수님</p>							

1. 캡스톤디자인의 배경 및 필요성

인공지능을 활용한 이미지의 분석 및 생성은 최근 심층 신경망 학습 알고리즘의 등장을 통해서 비약적인 발전을 이루었습니다. 특히 적대적 생성 알고리즘(GAN)이 대표적인 사례로 꼽히는데 이 분야에서 저희는 GauGAN을 대표적인 결과물로 생각합니다.

하지만 저희가 GauGAN을 사용해보니, 소재가 자연물/자연적인 풍경으로 한정되고 그림판 위에 직접 드로잉을 해야하니 그림에 정말 조예가 없는 사람은 사용하기가 힘들며 텍스트입력을 사용해도 영어로 입력을 받으니 한국어 화자가 사용하기에 나름의 불편이 따른다는 문제점이 있었습니다.

하지만 GauGAN의 사례를 보고 저희 중 그림에 관심이 있던 조원이 원화를 지망하는 지망생 대부분이 캐릭터를 그려서 배경 분야는 항상 인력이 부족하다는 업계인의 언급과 자신이 인터넷에서 커미션(그림 제작 주문)을 신청하고자 하여도 높은 가격과 신뢰도의 부족으로 망설였던 경험을 바탕으로 일종의 일러스트 제작기를 프로젝트 주제로 제안하였습니다.

그래서 GauGAN을 사용하고 생각했던 문제점을 바탕으로 저희의 관점에서 그것을 해결해보고자 가능한 넓은 주제를 자유롭게 다룰 수 있도록, 또 한국어 문장으로 입력을 받아 해당 조건에 맞는 그림을 이미지로 출력하는 프로그램을 저희가 사용했던 바와 같이 웹사이트에서 편한 접근성에서 사용할 수 있게 제작해보는 것을 이번 캡스톤디자인의 목표로 결정하게 되었습니다.

2. 캡스톤디자인 목표 및 비전

저희가 구상한 프로젝트의 주제는 일러스트 생성기로 게임 제작이나 개인방송 화면, 기타 창작물의 제작과정 등에 활용될 수 있는 가벼운 느낌의 일러스트를 번거롭게 비용을 들이고 다른 곳에 작업을 맡기고서 며칠씩 기다릴 필요가 없도록 비교적 짧은 시간 안에 그 자리에서 필요한 이미지를 제작하는 일종의 AI 페인팅 프로그램을 목표로 하고 있습니다.

웹사이트에서 사용자가 원하는 그림의 주제와 특성을 설명하는 문장을 입력하면 핵심이 되는 특징 키워드를 추출하고, 그 키워드에 맞춘 이미지 데이터셋으로 학습한 모델이 GAN 알고리즘을 거쳐 새 이미지를 생성 및 출력하는 구조의 프로그램으로 구상하였습니다.

해당 프로젝트의 구현에 성공한다면 이전에 있었던 엔비디아의 GauGAN, 카카오브레인의 minDALL-E 등의 다른 예시와 비슷한 목적을 가지면서도, 다르게 ‘사실적인’ 풍경만이 아닌 보다 가벼운 느낌의 게임이나 애니메이션 풍의 일러스트를 구현하여 원화 및 일러스트 제작 업계에서 해당 작업을 지원하거나 관련 작업에 대한 필요를 가진 개개인에게 웹사이트를 이용한 접근성을 바탕으로 서비스를 제공하는 데 사용될 수 있을 것으로 기대하고 있습니다.

그 기대에 더해 저희로서도 기존에 있던 훌륭한 프로그램들의 주제를 약간 비틀어서 저희의 구상대로 구현하고 그에 따른 어려움을 해결해가는 과정에서 데이터 사이언스와 머신러닝이라는 이전까지는 난해하게 느껴지던 분야를 직접 다루며 한층 더 깊이 이해하고 거기에서 파생된 GAN이라는 알고리즘이 더 사실적인 이미지만을 구현하는 게 아닌 인간이 생각하는 미술과 비슷한 방향으로도 나아갈 수 있음을 확인해보는 계기가 될 것으로 생각하고 있습니다.

3. 캡스톤디자인 내용

이 프로젝트의 주요 기능은 Text-to-image Generation으로, 배경 일러스트의 생성이 목적입니다. 한글 텍스트를 입력받아 내부에서 Text와 Image Pair를 학습시켜서 비교하며 GAN 과정을 거쳐 조건에 부합하는 새로운 이미지를 만들어내는 구조의 프로그램입니다.

저희는 한글 입력으로 새로운 이미지를 만들어낸다는 목적에 발맞추어 결과물의 정확도에 중점을 두는 평가 기준을 자체적으로 채택하고자 합니다. 정량적 평가 기준이 필요할지, 필요하다면 어떤 것으로 준비할지가 저희 사이에서도 고민되는 주제였습니다. 고민한 결과, 수많은 결과를 육안만으로 평가하는 것에는 한계가 있으며 설문조사 방식으론 불명확한 결과가 나올 가능성이 크다는 이유를 들어 필요할 것이라는 결론을 내렸습니다. 평가할 기준으로는 교수님께서 제시해주신 FID가 여러 이미지를 종합하여 사람이 평가하는 것과 비슷한 결과를 보여준다는 장점이, IS는 이해하기에 더 직관적이고 이미지의 품질과 다양성이라는 이해하기에 명료한 기준이 있다는 점에서 필요에 따라 둘을 병용할 예정입니다.

그 외에 고려할 사항으로는 우선 데이터의 수집으로, 웹 크롤링으로 특정 단어와 연관된 이미지들을 모아 데이터셋으로 변환하고자 생각합니다만 데이터의 수집에서 단순히 많은 데이터를 긁어모으기보다는 법적/도의적 문제가 발생할 수 있는 데이터의 수집을 지양할 필요가 있음을 인지하고 있으며 이 문제에 대처하고자 일단 robots.txt를 확인하고 크롤링하도록 해당 프로그램을 제작하는 방식으로 해당 요구사항을 부족하게나마 충족시키고 그 뒤에 다른 대안을 더 구하려고 합니다.

웹사이트의 형식들 사용한다는 특성상 생길 수 있는 보안에 대한 필요는 사이트로 들어오는 요청들에 XSS 필터를 적용하여 안전한 값만 전달되도록 하는 방식으로 악성 스크립트의 실행을 막아 보안을 갖추도록 할 것이며, 마지막으로 유지보수 측면에선 제작에 착수하며 프로젝트를 본격적으로 디자인할 때부터 객체 지향적으로 설계하고 도중에는 프로젝트의 형상 관리를 철저히 하여 애초에 이후의 유지보수 필요를 최소화해야 한다고 생각 중입니다.

4. 캡스톤디자인 추진전략 및 방법

우선 프로그램의 구현을 위해서는 Python을 기반으로 한 여러 후술할 프레임워크를 활용할 예정이며 개발 및 테스트환경으로 GPU를 사용한 머신러닝 테스트를 클라우드에 무료로 제공한다는 점에서 구글 Colab을 선택하였지만 시간적 한계나 처리할 수 있는 데이터양에 한계가 있는 상황을 고려하여 초반엔 주로 사용되 더 나은 성능이 필요해지면 다른 IDE를 준비할 계획입니다. 거기에 더해 팀의 협업을 위해 Github Collaborate Project도 새로 개설하여 작업물과 현황을 공유하면서 팀의 의사소통이 엇나갈 것에 대비해 Notion 프로그램을 통해 관련 정보와 향후 계획을 자세히 주고받으며 프로젝트를 추진할 예정입니다. 마지막으로, 프로젝트를 수월하게 진행하기 위해 각 팀원의 역할 및 책임은 다음과 같으며.

****조장**** : 연제원(팀 총괄 및 GAN 구현 담당)

****조원**** : 하상욱(크롤링 및 데이터셋 담당), 이창현(UI, 백엔드 연결 담당)

아래의 목표들을 하나씩 월간목표로 정해서 단계적으로 달성하고자 생각하고 있습니다. 현재 계획 중인 단계는 대략 다음과 같습니다. 먼저 입력된 문장을 분석하기 위해 konlpy 라이브러리에서 명사 분석 기능을 사용하여 문장의 단어들을 추출하는 기능을 구현합니다. 단, 명사 추출만으로는 특징 키워드를 제대로 잡아내지 못할 가능성이 크기에 이 부분은 여유가 된다면 형태소 분석으로 바꾸는 것을 계획중입니다.

그 뒤에 추출된 단어들을 미리 시험 주제로 선정한 단어들과 비교해서 정확히 비교되는 것을 확인하면, 해당하는 단어들과 연관되어 심층학습 과정에 사용할 이미지 데이터셋을 준비합니다. 이 때, 데이터셋의 준비에는 처음에는 실험목적으로 인터넷상에 준비된 데이터셋을 주제에 맞춰 끌어와서 활용하고 먼저 시험이 끝난 뒤에 네트워크 프로그래밍에서 학습한 지식을 활용하여 웹 크롤링으로 본격적인 데이터 수집에 들어갈 예정입니다.

단, 크롤링 시에 저작권으로 문제가 발생할 수 있음을 인지하고 있으며 크롤링 거부의를 밝힌 페이지를 피해가는 방법, 역시 저작권 침해 콘텐츠를 피해가는 방법이 있는지를 추가로 조사하여 적용할 예정입니다.

실험용 데이터셋의 준비가 1차적으로 완료된 뒤에는 PyTorch를 활용하여 데이터셋을 사용할 수 있는 GAN 과정을 구현합니다. 그렇게 데이터셋을 거쳐서 새 이미지가 생성되는 것을 확인한 이후 다음 단계로 프로그램 및 데이터셋의 최적화에 들어가 결과에서 오류의 최소화, 학습이 제대로 이루어지는지에 대한 검토 등과 함께 해당 프로그램을 웹사이트로 구현하는 단계로 진행합니다.

웹사이트로의 구현에는 파이썬 웹 프로그래밍용 프레임워크인 Django의 활용을 고려하는 중이며 프론트엔드의 작업으로는 가장 간결한 수준의 디자인을 채용한 UI의 구현과 백엔드 프로그램과의 연결을 수행하고, 백엔드의 작업으로는 데이터셋을 저장하고 그 접근을 관리할 DB의 준비에 더해 Django 프로그램의 구조에 맞추어 해당 DB를 model에 넣고 앞서 작성된 프로그램을 Views에 넣어 실행될 수 있도록 하는 것까지가 계획에 있으며 이 단계까지 완료된다면 이후부터는 기한이 남은 만큼 더 나은 성능을 가지도록 만드는데 주력할 예정입니다.

5. 참고문헌

[1] Soctt Reed et al. Generative Adversarial Text to Image Synthesis

[2]<https://blogs.nvidia.com/blog/2021/11/22/gaugan2-ai-art-demo/>

[3]<https://github.com/kakaobrain/minDALL-E>

[4] Esser et al. ImageBART: Bidirectional Context with Multinomial Diffusion for Autoregressive Image Synthesis, NeurIPS 2021.

[5]<https://paperswithcode.com/task/text-to-image-generation#task-libraries>