**선행기술조사 보고서**

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트명 | 인공지능을 이용한 웹 소설 플랫폼(NOVAI) | 작성자(조) | 8조 |

1. 프로젝트 개요

본 프로젝트의 목표는 illustration generation model을 자체적으로 구축한 web platform을 만드는 것이다.

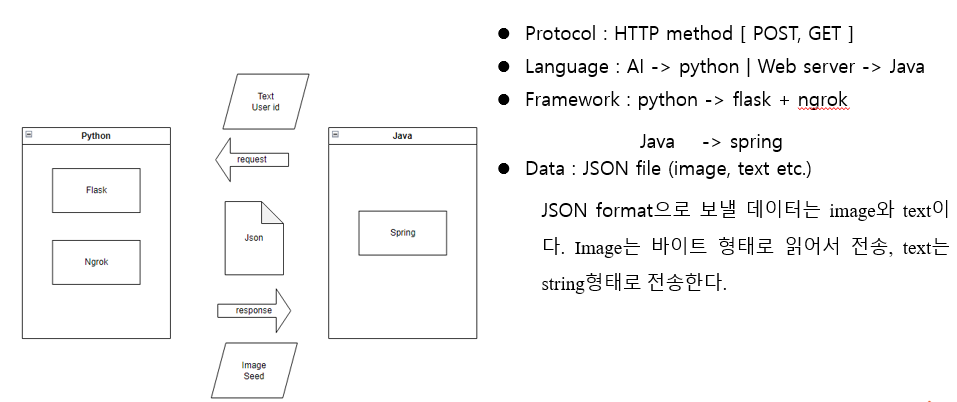
AI 관련 기술 중 image generation 분야의 경우, stability AI에서 stable diffusion model을 open source로 제공함으로써 각광받게 되었다. 해당 model은 latent diffusion model (LDM)의 일종으로써 computation cost가 너무 비싸 local에서는 돌리기 힘들었던 model size를 대폭 낮췄다는 점에서 의미가 있었다. 자체적으로 구축을 한다면 많은 비용을 들이지 않고 image를 제공할 수 있다는 점에서 조사를 진행하게 되었고, 웹 소설 작가의 불필요한 금전적·시간적 요소에 적용시켜 해결해보고자 한다.

본 프로젝트에서는 사용자 요구사항에 대한 정의를 먼저 한 뒤, 요구사항에 맞춰서 시스템을 설계했다. 이에 해당하는 주요 기능은 다음과 같다.

1. 웹 소설 작가에게 웹 소설 플랫폼 제공
2. 다양한 등장인물 이미지를 작가에게 제공 (등장인물 sample)
3. 작가가 선택한 등장인물을 일관성 있게 뽑기 위한 AI model 학습
4. 일관성 있는 등장인물 삽화 제공
5. 작가가 생성한 이미지를 웹소설에 첨부하는 기능 제공

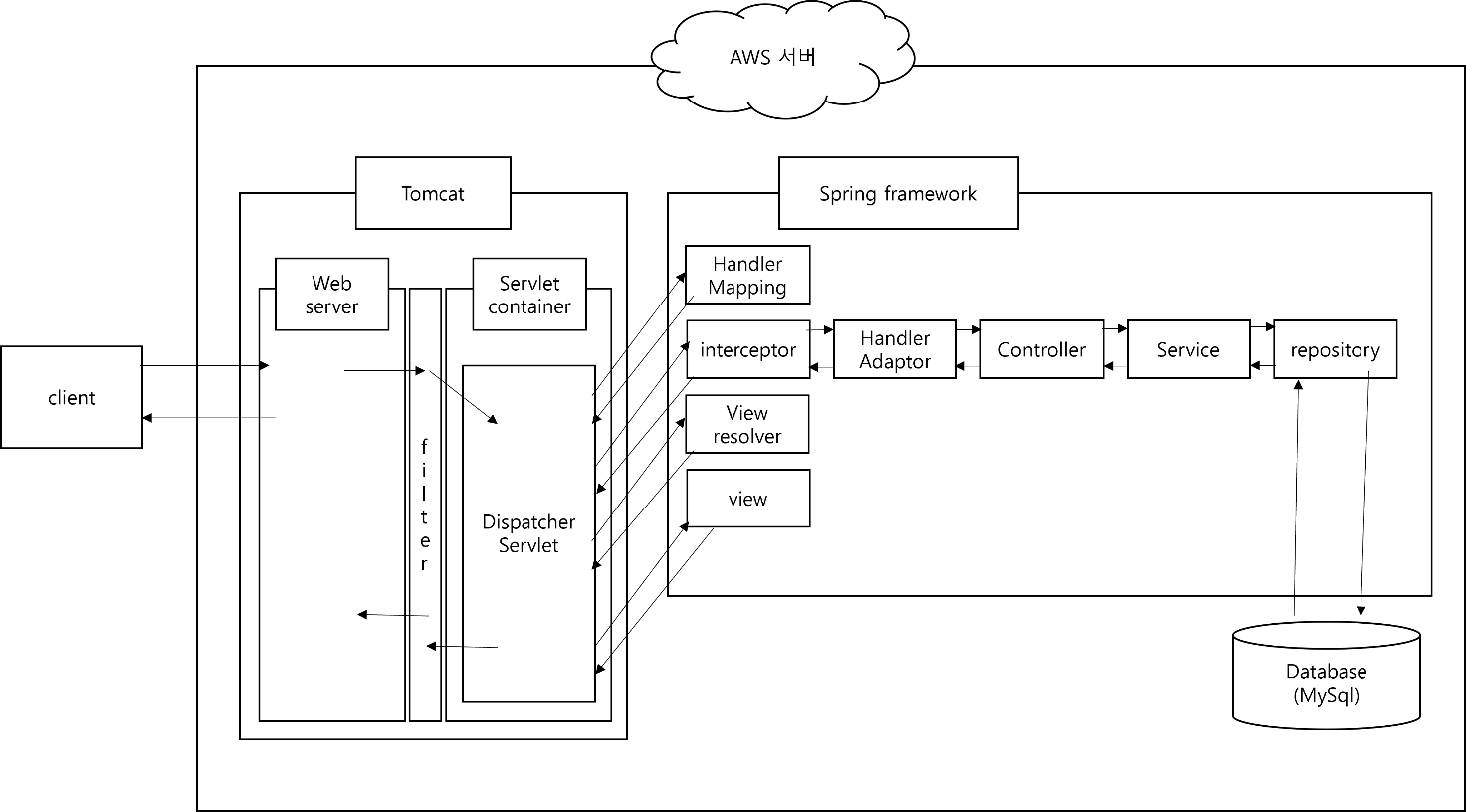
주요 기능을 구현하기 위해 웹 소설 플랫폼과 AI server를 구축해야 한다. 굳이 독립적으로 server를 운영하는 이유는 하드웨어적 제약 사항으로 인해 AI model을 웹 server와 동시에 구축할 수 없기 때문이다. 두 server 간의 data 통신은 HTTP method의 protocol을 따르면서 JSON format으로 통신할 것이다.

* Interface 설계는 다음과 같다.



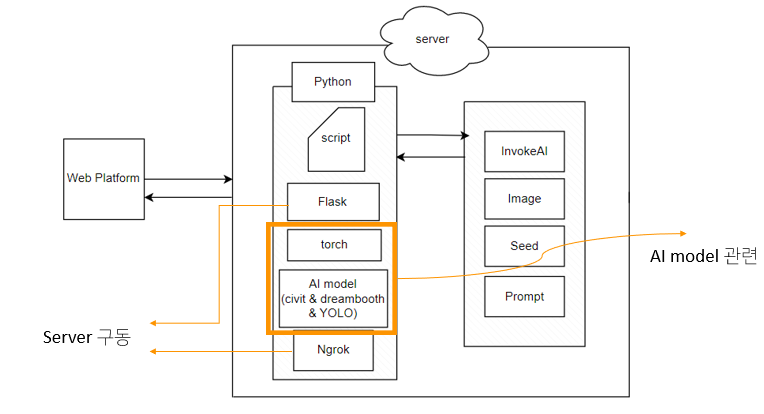
* Web platform 설계는 다음과 같다.

AWS server를 구축하여 tomcat으로 배포를 한다. Java를 기반으로 작동하기 때문에 Spring framework를 사용하여 필요한 기능들을 구현할 것이다.



* AI server의 설계는 다음과 같다.

AI를 다루기 위해 python과 torch를 사용한다. python에서 제공하는 Flask와 Ngrok로 server를 배포하고 python script로 invokeAI라는 image generation toolkit과 image, seed, prompt를 다룬다.



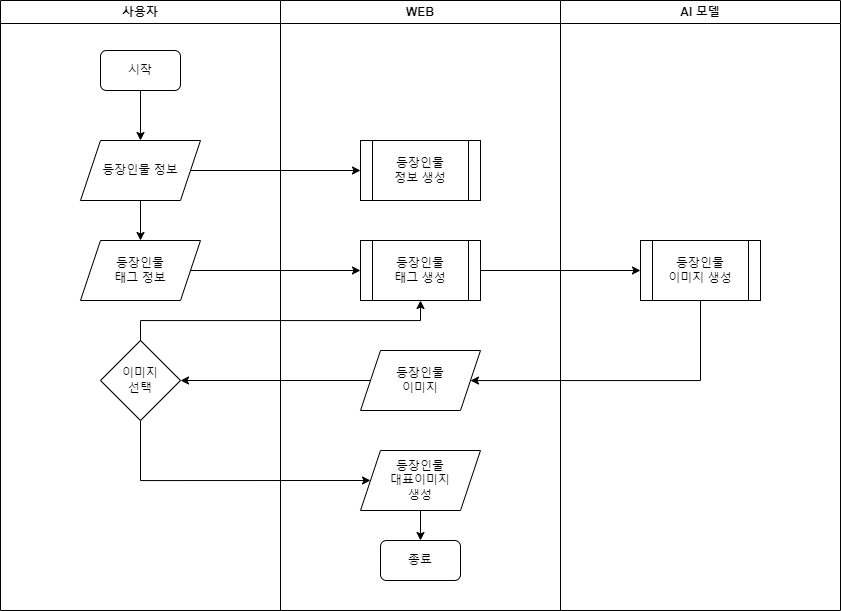
* 프로세스 설계



* 등장인물의 프로세스 설계는 다음과 같다.

1. 등장인물 sample 생성

태그를 input으로 등장인물 sample image를 뽑는 과정이다. 이때 사용되는 AI model은 작가의 선호하는 작화를 고려하여 4개를 선정했다. 각 model의 style은 anime, midjourney 등이 있다. 작가는 원하는 등장인물이 나올 때 까지 sample을 보다가 마음에 드는 등장인물이 나오면 해당 사진을 선택하면 된다.



1. 등장인물 학습

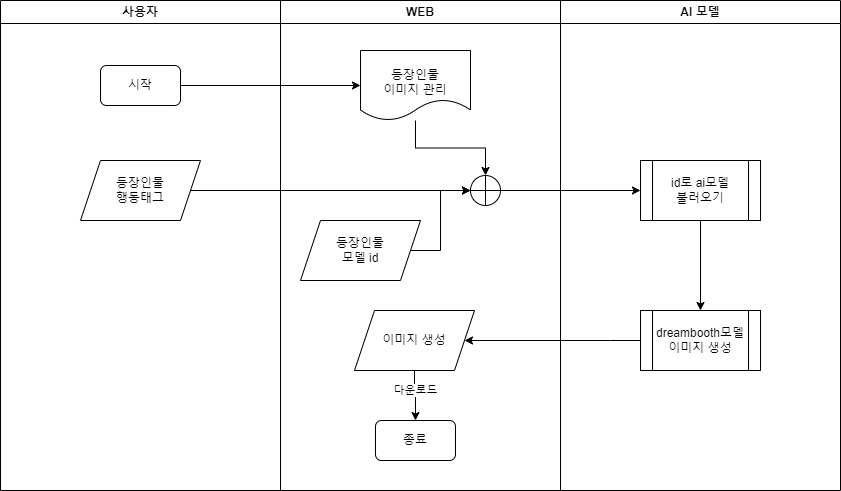
작가가 선택한 등장인물의 사진에 대한 user id, tag, seed값을 전달받는다. 해당 정보를 바탕으로 등장인물을 다양한 각도, 배경, 표정에 대해 복제한 뒤, dreambooth라는 학습 기법을 적용시킨다. 이때 복제하는 이유는 dataset을 늘리기 위함이고 dreambooth는 일관성 있는 삽화를 뽑기 위한 기법이다.

도표이(가) 표시된 사진

자동 생성된 설명

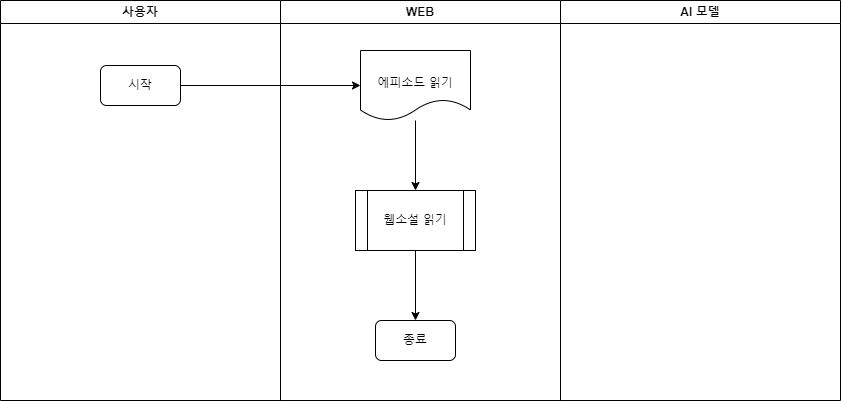
1. 등장인물 삽화 생성

학습이 완료된 AI model로 일관성 있는 삽화를 생성하는 기능이다. 등장인물을 특정 id에 mapping 시킨 후, 해당 등장인물을 뽑을 때 id로 AI model을 불러온다.

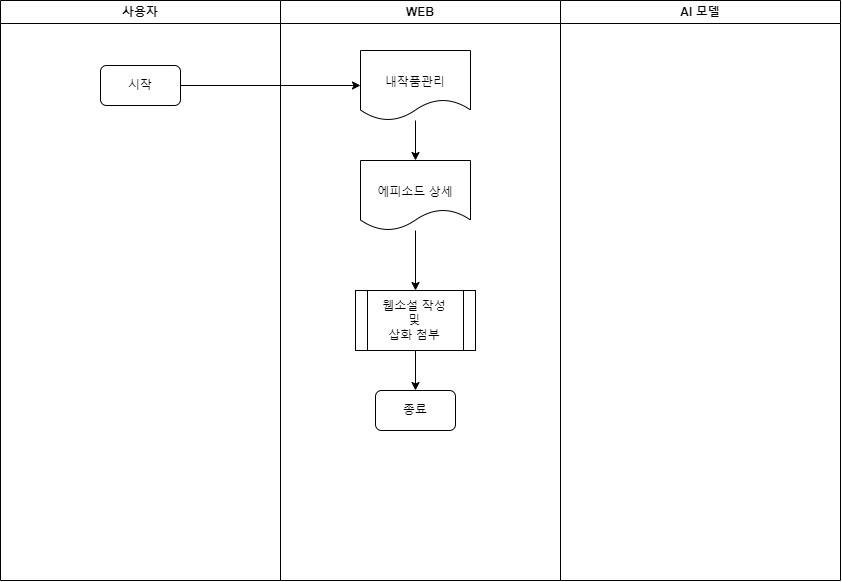


* 웹 소설 읽기, 쓰기의 프로세스는 다음과 같다.

1. 웹 소설 읽기



1. 웹 소설 쓰기



작가는 글 쓰는 분야에서 전문가라고 볼 수 있지만, 그림을 그리는 분야에 대해선 잘 알지 못하는 경향이 존재한다. 작가가 직접 그림을 그릴 수 없으니 결국 외주를 맡겨야 하고 자연스럽게 불편 사항과 불만들이 생길 수 있다. 자신이 상상하는 등장인물을 표현하고 싶은데 어려움을 겪는 경우, 경제적 여건, illustrator와의 소통에서 오는 정신적 고통 및 욕구 불만 등을 예시로 들 수 있다. 작가의 의견 즉시 반영. 독자의 흥미도 증가. (매 회차마다 삽화를 집어넣을 수 있는 환경이 조성되기 때문에)

이러한 불편 및 불만 사항을 최대한 해결하고자 했고 우리가 생각하는 기대효과는 다음과 같다.

웹 소설 플랫폼의 경우, AI model을 자체적으로 구축하여 값싸고 안정적인 system 제공받는다. 최근 존재하는 image generation model을 구축한 web site들을 살펴보자면, 다양한 사진 생성에 초점이 맞춰져 있다. 방향 자체가 다르기도 하고 외부 site에서 AI model을 불러오면 결국 이전과 비교해서 효율적으로 바뀌지 않는다. 이에 반해, 본 프로젝트의 핵심은 등장인물의 사진을 일관성 있게 제공하는 것이다. 작가의 문맥과 story에 맞는 삽화를 제공함으로써 이전에 말한 불편사항들을 어느정도 해소할 수 있다.

작가의 경우, illustrator에게 외주를 맡기지 않아도 된다. 외주를 맡기지 않으니 시간적으로나 금전적으로나 이전보다 여유가 생기고 즉각적인 의견 반영으로 인해 사용자의 요구사항을 충분히 들어줄 수 있다. 다양한 작화를 고려하여 여러 style의 sample들이 나오므로 성능 면에서도 효과를 보여준다. 추가적으로 AI에 대한 공부를 하지 않아도 사용할 수 있다는 장점이 있다.

illustrator뿐만 아니라 자신만의 style을 가진 모든 user에게 idea를 제공할 수 있다. 그림을 그리는 사람의 경우, 자신만의 style을 구축한 사람이 많이 존재한다. 하지만, 인간의 특성 상 상상력의 한계나 시간적 소모, 정신적 소모로 인해 한계에 부딪히는 경우가 많다. AI를 활용한다면 style을 유지하면서 새로운 길이나 방향을 제시할 수도 있다.

본 프로젝트의 핵심은 웹 소설 플랫폼에 집중하기 보단, 자체적으로 AI model을 구축함에 있어서 의의가 있다.

1. 선행기술조사 범위 및 방법

|  |  |  |
| --- | --- | --- |
| 조사범위 | 조사국가 | ■한국 □미국 □유럽 □일본 □중국 □PCT □기타 |
| 조사대상 | 특허/실용신안 공개 및 등록공보 |
| 사용 DB | KIPRIS |
| 조사 기준일 | 2023.04.27 |
| 키워드 | AI, 인공지능, 이미지, 삽화, 사진, 생성 | |

|  |  |  |
| --- | --- | --- |
| 국가 | 검색식 | 검색 건수 |
| KR | (AI+인공지능)\*(이미지+삽화)\*생성 | 107,290 |
| US | AI\*(image+picture+illustration)\*create | 146,308 |
| JP | 1. (AI+인공지능)\*(이미지+삽화)\*생성 2. AI\*(image+picture+illustration)\*create | 1. 30,023 2. 1,290 |
| EP | AI\*(image+picture+illustration)\*create | 18,174 |
| CN | - | - |
| PCT | - | - |

1. 선행기술 목록

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 순번 | 국가 | 문헌번호 | 공개일 | 출원인 | 발명의 명칭 | 관련성 |
| 1 | 한국 | 1020230009577 | 2023.01.17 | 마인드로직 주식회사 | 인공지능을 이용하는 고해상도 이미지 생성 방법 및 그 시스템 | A |
| 2 | 한국 | 1020220151142 | 2022.11.14 | 주식회사 디자이노블 | AI를 이용한 디자인 생성 방법 및 그 장치 | Y |
| 3 | 한국 | 1022009090000 | 등록일  2021.01.11 | ㈜ 에스프레소미디어 | AI 기반 이미지 스타일 학습 및 공유 시스템 | Y |
| 4 |  |  |  |  |  |  |

\* 관련성 X, Y, A는 각각 관련높음, 관련있음, 참고할만함을 의미함.

1. 선행기술과의 대비
2. 선행문헌 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 발명의 명칭 | | 인공지능을 이용하는 고해상도 이미지 생성 방법 및 그 시스템 | | |
| 문헌번호 | | 1020230009577 | 관련성 | A |
| 공개일 | | 2023.01.17 | 출원인 | 마인드로직 주식회사 |
| 기술  요지 | 목적 | 본 발명은 상태 조건(condition)을 적용하여 원하는 형태의 고해상도 이미지를 생성하는 인공지능을 이용하는 고 해상도 이미지 생성 방법 및 그 시스템에 관한 것으로서, 일부분을 나타내는 조각 이미지를 수신하는 단계, 인공 지능 모델 및 이미지 변환 모델을 사용하여 상태 조건(condition)을 적용하는 단계 및 인공지능 모델을 사용하여 상기 조각 이미지를 유지하면서 상태 조건(condition)에 따라 상기 조각 이미지에 적합한 부분 이미지를 자연스럽게 연결하여 고해상도의 전신 이미지를 생성하는 단계를 포함한다. | | |
| 구성 | 본 발명의 실시예에 따른 인공지능(Artificial Intelligence; AI)을 이용하여 상태 조건에 부합하는 고해상도 (High Resolution) 이미지 생성 방법에 있어서, 일부분을 나타내는 조각 이미지를 수신하는 단계, 인공지능 모 델 및 이미지 변환 모델을 사용하여 상태 조건(condition)을 적용하는 단계 및 인공지능 모델을 사용하여 상기 조각 이미지를 유지하면서 상태 조건(condition)에 따라 상기 조각 이미지에 적합한 부분 이미지를 자연스럽게 연결하여 고해상도의 전신 이미지를 생성하는 단계를 포함한다.  상기 수신하는 단계는 얼굴 이미지와 같은 전신의 일부분을 나타내는 상기 조각 이미지를 수신할 수 있다.  상기 적용하는 단계는 Conditional GAN의 인공지능 모델과 Pix2Pix의 이미지 변환 모델을 모두 사용하며, 남녀의 도메인 상태 조건(domain condition)과 얼굴 이미지의 상태 조건, 옷의 상태 조건, 몸의 상태 조건 및 포즈 (pose)의 상태 조건을 적용할 수 있다.  상기 생성하는 단계는 Stylegan2의 인공지능 모델에서 인코더(encoder)와 디코더(decoder) 중간에 스킵 연결 (skip connection)을 넣어 상기 상태 조건에 따라 얼굴 이미지인 상기 조각 이미지에 적합한 몸과 옷의 크기, 위치 및 비율을 나타내는 상기 부분 이미지를 생성한 후, 상기 조각 이미지와 상기 부분 이미지를 자연스럽게 연결하여 고해상도(high resolution)의 전신 이미지를 생성할 수 있다.  상기 생성하는 단계는 L1 손실(L1 loss)를 적용하여 상기 조각 이미지를 유지하는 것을 특징이라 할 수 있다.  상기 생성하는 단계는 상기 조각 이미지를 세분화하여 추출된 얼굴색과 상기 전신 이미지의 일부분에서의 피부 색이 일치되도록 스킨 매칭 알고리즘(skin matching algorithm)을 적용할 수 있다.  상기 생성하는 단계는 옷의 상태 조건 및 포즈(pose)의 상태 조건에 대해 일정한 로스(loss)를 적용하여 상기 조각 이미지에 자연스럽게 연결되는 고해상도의 상기 전신 이미지를 생성할 수 있다.  본 발명의 실시예에 따른 인공지능(Artificial Intelligence; AI)을 이용하여 상태 조건에 부합하는 고해상도 (High Resolution) 이미지 생성 시스템에 있어서, 일부분을 나타내는 조각 이미지를 수신하는 수신부, 인공지능 모델 및 이미지 변환 모델을 사용하여 상태 조건(condition)을 적용하는 적용부 및 인공지능 모델을 사용하여 상기 조각 이미지를 유지하면서 상태 조건(condition)에 따라 상기 조각 이미지에 적합한 부분 이미지를 자연스럽게 연결하여 고해상도의 전신 이미지를 생성하는 생성부를 포함한다.  상기 수신부는 얼굴 이미지와 같은 전신의 일부분을 나타내는 상기 조각 이미지를 수신할 수 있다.  상기 적용부는 Conditional GAN의 인공지능 모델과 Pix2Pix의 이미지 변환 모델을 모두 사용하며, 남녀의 도메인 상태 조건(domain condition)과 얼굴 이미지의 상태 조건, 옷의 상태 조건, 몸의 상태 조건 및 포즈(pose)의 상태 조건을 적용할 수 있다. [0016] 상기 생성부는 Stylegan2의 인공지능 모델에서 인코더(encoder)와 디코더(decoder) 중간에 스킵 연결(skip connection)을 넣어 상기 상태 조건에 따라 얼굴 이미지인 상기 조각 이미지에 적합한 몸과 옷의 크기, 위치 및 비율을 나타내는 상기 부분 이미지를 생성한 후, 상기 조각 이미지와 상기 부분 이미지를 자연스럽게 연결하여 고해상도(high resolution)의 전신 이미지를 생성할 수 있다.  상기 생성부는 L1 손실(L1 loss)를 적용하여 상기 조각 이미지를 유지하는 것을 특징이라 할 수 있다.  상기 생성부는 상기 조각 이미지를 세분화하여 추출된 얼굴색과 상기 전신 이미지의 일부분에서의 피부색이 일치되도록 스킨 매칭 알고리즘(skin matching algorithm)을 적용할 수 있다.  상기 생성부는 옷의 상태 조건 및 포즈(pose)의 상태 조건에 대해 일정한 로스(loss)를 적용하여 상기 조각 이 미지에 자연스럽게 연결되는 고해상도의 상기 전신 이미지를 생성할 수 있다. | | |
| 효과 | 본 발명의 실시예에 따르면, 일부분의 조각 이미지를 기반으로, 상태 조건(condition)을 적용하여 원하는 형태의 고해상도(high resolution) 이미지를 생성하는 인공지능 기술을 제공함으로써, 한 장의 얼굴 이미지를 보유하는 경우에 특정 디자인의 옷을 선택하면, 해당 얼굴 이미지에 어울리는 형태로 해당 디자인의 옷을 입고 있는 몸을 생성할 수 있으며, 해당 얼굴에 합성된 이미지를 높은 해상도로 생성할 수 있다.  또한, 본 발명의 실시예에 따르면, 얼굴 이미지의 변형을 최소화하여 전신 이미지의 합성 전후에도 동일 인물로 인식되도록 제공할 수 있다.  또한, 본 발명의 실시예에 따르면, 얼굴 이미지에 따라 몸과 옷의 크기, 위치, 비율 등을 자연스럽게 연결시켜 얼굴 이미지와 어울리는 형태로 조합하여 전신 이미지를 생성할 수 있다.  또한, 본 발명의 실시예에 따르면, 큰 사이즈의 조각 이미지(1024×1024 pixel 이상)를 사용하여 상태 조건에 따른 전신 이미지를 제공할 수 있다. | | |
| 도면 |  | | |
| 검토  의견 | 유사점 | **AI model에서의 구성요소로 보자면 encoder, decoder, skip connection의 경우 본 프로젝트의 AI model과 동일한 부분이 존재한다.**  **Output이 완성된 image라는 부분에 대해선 동일하다.** | | |
| 차이점 | **구성요소가 같은 부분이 있어도 목적 자체가 다르기 때문에 model의 architecture는 다르다.**  **본 프로젝트에서 사용하는 AI model은 input이 tag로 들어가는 게 필수적인 요소이다. 즉 string의 형태로 text가 들어간다고 생각하면 되는데, 이 system의 경우 image의 일부분을 제공하여 나머지 부분까지 생성해내는 outpainting이다. 즉 input이 image의 일부분임** | | |

1. 선행문헌 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 발명의 명칭 | | AI를 이용한 디자인 생성 방법 및 그 장치 | | |
| 문헌번호 | | 1020220151142 | 관련성 | Y |
| 공개일 | | 2022.11.14 | 출원인 | 주식회사 디자이노블 |
| 기술  요지 | 목적 | 본 발명의 실시예에 따르면, 상기 방법은, 이미지 및 텍스트를 포함하는 정보로부터 상기 이미지를 획득하는 단계와 상기 획득된 이미지의 특징을 학습하는 단계와 상기 정보 상의 텍스트를 추출하고, 상기 추출된 텍스트를 상기 학습된 이미지의 특징에 매칭하는 단계와 상기 매칭을 통해, 상기 이미지를 바탕으로 생성하고자 하는 디자인의 조건을 학습하는 단계와 사용자로부터 디자인 이미지 생성을 위한 사용자 텍스트를 입력 받는 단계와 상기 사용자 텍스트에 대응하는 조건을 식별하는 단계와 상기 식별된 조건 기반의 디자인 이미지를 생성하는 단계를 포함할 수 있다. | | |
| 구성 | 본 발명이 해결하고자 하는 기술적 과제는, 획득된 이미지 및 이에 매칭되는 텍스트를 학습함으로써, 디자인의 자동 생성을 위한 조건을 생성하는 방법 및 그 장치를 제공하는 것이다. 구체적으로 본 발명이 해결하고자 하는 기술적 과제는, 이미지에 매칭된 텍스트의 특성을 학습함에 따라, 텍스트에 반영된 트렌드 정보를 기초로 디자인을 자동 생성하는 방법 및 그 장치를 제공하는 것이다.  본 발명이 해결하고자 하는 다른 기술적 과제는, 상기 생성된 조건에 대응하는 텍스트의 입력에 응답하여 디자인을 자동 생성하는 방법 및 그 장치를 제공하는 것이다. 구체적으로 본 발명이 해결하고자 하는 다른 기술적 과제는, 제품의 카테고리 선정 및 사용자에 의해 입력된 텍스트를 디자인 생성의 조건으로 인식하고, 인식된 조건 하에 디자인을 생성하는 방법 및 그 장치를 제공하는 것이다. 보다 자세히, 본 발명이 해결하고자 하는 기술적 과제는, 상기 입력된 텍스트를 상기 선정된 카테고리 디테일한 특징으로 해석하고, 상기 카테고리의 제품에 대한 상기 특징이 반영된 디자인을 생성하는 방법 및 그 장치를 제공하는 것이다.  본 발명이 해결하고자 하는 또 다른 기술적 과제는, 입력되는 텍스트에 의해 생성되는 디자인 이미지의 해상도를 높이는 방법 및 그 장치를 제공하는 것이다. 구체적으로, 본 발명이 해결하고자 하는 또 다른 기술적 과제는, 원본 이미지 및 생성 디자인 이미지 상의 픽셀 기반 손실이 아닌, 특징량 기반의 손실을 고려하여 디자인 이미지의 해상도를 극대화할 수 있는 방법 및 그 장치를 제공하는 것이다.  본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명의 기술분야에서의 통상의 기술자에게 명확하게 이해될 수 있을 것이다.  상기 기술적 과제를 해결하기 위한 학습된 조건 기반의 디자인 생성 방법은 학습된 조건 기반의 디자인 생성 장 치에 의해 수행되고, 상기 방법은, 이미지 및 텍스트를 포함하는 정보로부터 상기 이미지를 획득하는 단계와 상기 획득된 이미지의 특징을 학습하는 단계와 상기 정보 상의 텍스트를 추출하고, 상기 추출된 텍스트를 상기 학습된 이미지의 특징에 매칭하는 단계와 상기 매칭을 통해, 상기 이미지를 바탕으로 생성하고자 하는 디자인의 조건을 학습하는 단계와 사용자로부터 디자인 이미지 생성을 위한 사용자 텍스트를 입력 받는 단계와 상기 사용 자 텍스트에 대응하는 조건을 식별하는 단계와 상기 식별된 조건 기반의 디자인 이미지를 생성하는 단계를 포함 할 수 있다.  일 실시예에 따르면, 상기 방법은, 상기 생성된 디자인 이미지에 대한 해상도 변환을 수행하는 단계와 상기 해 상도 변환을 통해 최종 디자인 이미지를 생성하는 단계를 포함할 수 있으며, 상기 해상도 변환을 수행하는 단계는, 상기 생성된 디자인 이미지의 특징을 추출하는 단계와 상기 학습된 이미지의 특징을 상기 생성된 디자인 이 미지의 추출된 특징과 비교하는 단계와 상기 비교 결과, 차이 값 정보를 기초로, 상기 생성된 디자인의 해상도 변환을 수행하는 단계를 포함할 수 있다.  일 실시예에 따르면 상기 이미지를 바탕으로 생성하고자 하는 디자인의 조건을 학습하는 단계는, 상기 매칭을 통해, 상기 추출된 텍스트 중, 상기 학습된 이미지의 특징에 대응하는 텍스트를 상기 조건으로 학습하는 단계와 상기 조건으로 학습된 텍스트를 상기 이미지 특징을 포함하는 제품의 카테고리로 결정하는 단계를 포함할 수 있다.  일 실시예에 따르면, 상기 사용자 텍스트에 대응하는 조건을 식별하는 단계는, 상기 사용자 텍스트를 벡터화 하는 단계와 상기 벡터화 결과 상기 사용자 텍스트의 벡터 값을 생성하고, 상기 카테고리에 상기 벡터 값이 매칭 되는지 판단하는 단계와 상기 카테고리에 상기 벡터 값이 매칭되는 경우, 상기 카테고리 및 상기 벡터 값을 상기 조건으로 식별하는 단계를 포함할 수 있다.  일 실시예에 따르면, 상기 카테고리에 상기 벡터 값이 매칭되지 않는 경우, 상기 카테고리 및 상기 벡터 값에 스타일 변수를 반영하여, 상기 카테고리에 매칭되는 변환 벡터 값을 추출하는 단계와 상기 변환 벡터 값 및 상 기 카테고리를 상기 조건으로 식별하는 단계를 포함하는, 상기 조건으로 식별하는 단계를 포함할 수 있다. | | |
| 효과 | 본 발명의 일 실시예에 따르면, 제품 디자인에 대하여 사용자마다 사용하는 단어가 다르더라도, 사용자가 입력한 단어를 기초로 사용자의 의도에 부합하는 디자인 이미지를 생성할 수 있는 장점이 있다.  또한, 본 발명의 다른 실시예에 따르면, 사용자에게 최신 디자인 트렌드 정보가 반영된 디자인 이미지가 제공되는 장점이 있다. 즉, 상기 실시예에 따르면, 사용자가 파악하기 어려운 최신 트렌드가 디자인 이미지 생성 모델에 반영되어, 사용자의 입력 텍스트를 최신 트렌드로 해석하고, 최신 트렌드가 반영된 디자인 이미지를 생성할 수 있는 장점이 있다.  또한, 본 발명에 또 다른 실시예에 따르면, 카테고리 및 입력 텍스트로 디자인의 생성 대상을 제한함으로써 디자인 이미지 생성의 연산량을 줄일 수 있는 장점이 있다.  또한, 본 발명의 또 다른 실시예에 따르면, 생성되는 디자인 이미지의 해상도를 향상시킴으로써, 디자인의 제품 화의 효율성이 높아지는 장점이 있다.  본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 본 발명의 기술분야에서의 통상의 기술자에게 명확하게 이해될 수 있을 것이다. | | |
| 도면 |  | | |
| 검토  의견 | 유사점 | **Text를 input으로 image를 생성하는 점이 유사하다. Generation model을 사용한 것이 유사.**  **Text와 image를 매칭하여 embedding을 하는 것이 같고, text에 반영된 정보를 추출하여 이미지를 자동 생성하는 것이 유사하다.**  **Dreambooth 기법을 이용하여 특정 text를 style or object와 mapping 시켜 학습을 하는 것과 비슷하다.** | | |
| 차이점 | **크게 보면 generation model이지만 세부적으로 나누자면 GAN과 Diffusion model로 나눠지므로 둘은 엄밀히 말하자면 서로 다른 model이긴 하다. (작동 방식이 다름)** | | |

1. 선행문헌 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 발명의 명칭 | | AI 기반 이미지 스타일 학습 및 공유 시스템 | | |
| 문헌번호 | | 1022009090000 | 관련성 | Y |
| 등록일 | | 2021.01.05 | 출원인 | ㈜에스프레소미디어 |
| 기술  요지 | 목적 | 본 발명은 AI 기반 이미지 스타일 학습 및 공유 시스템에 관한 것이다. 본 발명은, 복수의 스마트 디바이스(100)로 이루어진 스마트 디바이스 그룹, 네트워크(200) 및 이미지 퍼스널 커스텀 서버(300)로 이루어진 AI 기반 이미지 스타일 학습 및 공유 시스템(1)에 있어서, 이미지 퍼스널 커스텀 서버(300)는, 학습 정보 생성 모듈(322); 및 이미지 파일 정보 및 학습 정보를 수집하며, 수집된 학습 파일을 학습 정보 생성 모듈(322)로 전송하는 학습 정보 수집 모듈(321); 을 포함하며, 학습 정보 수집 모듈(321)은 하나의 스마트 디바이스(100)에 설치된 어플(300a) 상에서 스타일을 학습할 파일, 복수의 경우 복수의 파일에 대한 선택이 수행되면, 네트워크(200)를 통해 하나의 스마트 디바이스(100)의 어플(300a)을 통해 수신한 실제 파일을 학습 정보 생성 모듈(322)로 제공하며, 파일이 입력 스토리지(300a)로 전송이 완료되는 경우, 유저 스타일 크리에이터(300c) 상으로 파일정보 및 스타일정보, 기타 옵션정보에 대한 전송을 수행하는 것을 특징으로 할 수 있다. 이에 의해, 유저(User)가 임의의 이미지를 가지고 스타일(style)을 학습시켜 향후 상용 스타일처럼 사용하도록 하기 위한 퍼스널 유저 스타일(Personal User Style) 학습 알고리즘을 제공할 하는 효과가 있다. | | |
| 구성 | 상기의 목적을 달성하기 위해 본 발명의 실시예에 따른 AI 기반 이미지 스타일 학습 및 공유 시스템은, 복수의 스마트 디바이스(100)로 이루어진 스마트 디바이스 그룹, 네트워크(200) 및 이미지 퍼스널 커스텀 서버(300)로 이루어진 AI 기반 이미지 스타일 학습 및 공유 시스템(1)에 있어서, 이미지 퍼스널 커스텀 서버(300)는, 학습 정보 생성 모듈(322); 및 이미지 파일 정보 및 학습 정보를 수집하며, 수집된 학습 파일을 학습 정보 생성 모듈 (322)로 전송하는 학습 정보 수집 모듈(321); 을 포함하며, 학습 정보 수집 모듈(321)은 하나의 스마트 디바이스(100)에 설치된 어플(300a) 상에서 스타일을 학습할 파일, 복수의 경우 복수의 파일에 대한 선택이 수행되면, 네트워크(200)를 통해 하나의 스마트 디바이스(100)의 어플(300a)을 통해 수신한 실제 파일을 학습 정보 생성 모듈(322)로 제공하며, 파일이 입력 스토리지(300a)로 전송이 완료되는 경우, 유저 스타일 크리에이터(300c) 상으로 파일정보 및 스타일정보, 기타 옵션정보에 대한 전송을 수행하는 것을 특징으로 할 수 있다.  이때, 유저 스타일 크리에이터(300c)는, 어플(300a)에서 제공한 파일정보, 그 밖에 학습정보에 해당하는 스타일 정보, 기타 옵션정보에 대한 수집을 수행하며, 입력된 정보의 실제 파일이 있는지 체크 후 학습 정보 생성 모듈 (322)에 의한 학습정보 생성에 대한 진행을 요청하는 것을 특징으로 할 수 있다.  또한, 학습 정보 생성 모듈(322)은, DB(300d)에 학습정보를 입력할 뿐만 아니라, 학습 파일을 작업 경로로 이동하는 기능을 수행하며, DB(300d)에 학습정보 입력에 있어서 신규 학습 시 새로운 Customized Style을 등록하며, 기존 학습 시 기존 Customized Style에 추가 학습을 등록하는 것을 특징으로 할 수 있다.  또한, 학습 정보 생성 모듈(322)은, 유저 스타일 크리에이터(300c)에 의해 DB(300d)에 학습정보가 입력되는데 등록특허 10-2200909 - 5 - 신규 스타일 학습인 경우 유저 스타일 크리에이터(300c)에 의해 생성된 신규 학습정보가 DB(300d) 상에 입력되며, 기존 스타일 학습인 경우 유저 스타일 크리에이터(300c)에 의해 추가 학습정보가 DB(300d) 상에 입력되도록 제어하되, 메타데이터에 해당하는 Key information은 style name으로 지정하는 것을 특징으로 할 수 있다.  또한, 학습 정보 생성 모듈(322)은, 유저 스타일 크리에이터(300c)에 대한 제어를 통해 DB(300d)에 학습정보에 대한 입력이 성공하면 working status=-1(not ready)으로 세팅하는 것을 특징으로 할 수 있다.  또한, 학습 정보 생성 모듈(322)은, 유저 스타일 모니터(300d)에 대한 제어를 통해 입력 스토리지(300b)에 저장된 학습 파일을 실제 작업 경로에 해당하는 워크 스토리지(300f)로 이동하되, DB(300d)에 입력된 학습정보를 보 고 학습 파일을 워크 스토리지(300f)에 해당하는 작업 경로로 이동하면서 학습 파일을 처리할 유저 스타일 워커 (300g)에 대한 지정을 수행하는 것을 특징으로 할 수 있다.  또한, 학습 정보 생성 모듈(322)은, 유저 스타일 모니터(300d)에 대한 제어를 통해 지정된 유저 스타일 워커 (300g)의 워커 정보를 DB(300d)에 기재하도록 하며, DB(300d)에 워커 정보가 기재된 유저 스타일 워커(300g)로 message notify를 전송하도록 유저 스타일 모니터(300d)를 제어하는 것을 특징으로 할 수 있다.  또한, 이미지 퍼스널 커스텀 서버(300)는, Customized Style 명으로 구분하여 학습 작업을 수행하며, 학습 작업 과정 DB에 갱신을 수행하는 학습 수행 모듈(323); 을 더 포함하는 것을 특징으로 할 수 있다.  또한, 학습 수행 모듈(323)은, 유저 스타일 워커(300g)에 대한 제어를 통해 신규 및 기존 스타일 학습 처리를 수행하도록 하는 것을 특징으로 할 수 있다.  또한, 유저 스타일 워커(300g)는, Queue에서 수신된 message notify에 해당하는 message를 하나씩 꺼내어 작업을 수행하되, DB(300d)에서 학습정보를 읽어와서 학습 파일에 대해 학습을 수행하는 것을 특징으로 할 수 있다. | | |
| 효과 | 본 발명의 실시예에 따른 AI 기반 이미지 스타일 학습 및 공유 시스템은, 유저(User)가 임의의 이미지를 가지고 스타일(style)을 학습시켜 향후 상용 스타일처럼 사용하도록 하기 위한 퍼스널 유저 스타일(Personal User Style) 학습 알고리즘을 제공할 수 있는 효과가 있다.  또한, 본 발명의 다른 실시예에 따른 AI 기반 이미지 스타일 학습 및 공유 시스템은, 원본 이미지에 대한 스타일 변환 서비스를 통해, 개인의 프로필 사진을 개성 있게 꾸미거나, 개인화된 이모티콘 제작 및 활용이 가능하며, 스타일 변환 Open API 서비스를 통해, 자신의 그림 스타일을 등록하고, 다양한 사진을 자신만의 스타일로 변환 가능한 효과를 제공한다. | | |
| 도면 |  | | |
| 검토  의견 | 유사점 | **원본 이미지의 스타일 변화를 통한 개인화가 유사하다. 본 프로젝트에서도 image를 일관성있게 제공하기 위해 dreambooth라는 기법을 적용한다. 또한 개인화된 style이나 object를 저장하여 등록 및 재사용이 가능하다는 점에서 같다.** | | |
| 차이점 | **본 프로젝트에서의 중요 기능은 결국 등장인물을 학습하고 생성하는 기능이지만 해당 발명의 기능은 프로필 사진이나 이모티콘 제작으로써 방향성이 다르다.** | | |

1. 결론

|  |
| --- |
| 3가지의 발명에 대해서 살펴보았다. Image generation 분야는 최근 각광받고 있는 분야이기도 하지만 application에 바로 적용시켜 볼만큼 성능이 안정적이거나 조절이 쉽지는 않다. 그럼에도 불구하고, 최신 트렌드인 기술을 이용해서 이미 특허를 냈다는 점에서 놀라웠다. 다행인 점은 본 프로젝트에 관련해 이미 특허를 내거나 유사도가 높은 발명은 나오지 않았다는 점이다.  각자 추구하는 방향성과 분야는 다르지만 모두 AI generation model을 사용한다는 점에서 유익하게 읽어볼 수 있었다. 선행문헌 (1)을 읽었을 때, 해당 기술을 본 프로젝트에 적용시켜 세부 기능으로써 사용하고 싶다는 생각이 들었다. Outpainting을 이용하는 기술로써, image의 일부분만 제공하고 나머지를 생성해내는 기법인데 등장인물의 얼굴을 고정시키고 옷이나 배경을 그리는 등 다양하게 활용할 수 있을 거 같다.  특정 object나 style을 개인화 시켜서 사용자에게 제공한다는 점에선 선행문헌 (2), (3)과 상당부분 유사했다. 결국 본 프로젝트의 핵심도 특정 등장인물을 학습하여 일관성 있는 삽화를 제공하는 것으로써, 작가의 등장인물을 특정 embedding에 mapping시켜 개인화를 시켜주는 것이다. 개인화 기법으로 dreambooth를 사용하는데 해당 기법보다 개인화 성능이 뛰어나다면 기능의 일부분을 대체하는 것도 고려해 볼만하다. |