**선행기술조사 보고서**

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트명 |  | 작성자(조) |  |

1. 프로젝트 개요

본 프로젝트의 목표는 현재 웹소설 사이트들이 다소 웹툰 사이트보다 인기가 적은 이유를 분석해보았다. 그 결과 웹소설은 그림이 있지 않아서 흥미나 몰입도가 줄어드는 점이라는 것이다. 그렇기에 우리 팀은 인공지능의 생성 모델을 이용하여, 삽화를 인공지능이 생성해주는 웹을 만드는 것이다.

시스템 설계는 다음과 같다.

스크린샷, 콜라주, 사람, 텍스트이(가) 표시된 사진

자동 생성된 설명

등장인물을 생성한 후 2. 등장인물에 대해 학습을 시키고 3. 등장인물과 관련된 삽화를 생성하는 것이다. 각 프로세스에 대한 설계는 다음과 같이 진행된다.

1. 등장인물 생성

등장인물 생성과정에서작가는 등장인물을 고르기 위해 다양한 태그를 선택하여 등장인물을 정의한다.작가는 해당 태그를 통해 생성된 등장인물 sample 4장을 얻을 수 있다. Ai model은 CivitAI에서 제공하는 model을 사용하고 이때 이용하는 generation toolkit으로써는 Invokeai를 사용한다.

Character Tag

등장인물을 정의하는 태그이다. 생김새를 정의하는 부분이기 때문에 행동이나 배경 과 같이 동적이고 등장인물과 상관없는 태그는 제외한다. AI model의 input으로 들어간다.

1. 등장인물 학습

작가가 선택한 등장인물의 대표 이미지를 dreambooth 기법을 적용시켜 개인화 시켜주는 과정이다. Base model은 openjourney를 사용. 한 model에 대해서 다양한 instance prompt를 학습한다. User가 선택한 등장인물의 prompt와 seed값을 사용한다.

1. 등장인물 삽화 생성

등장인물 학습 후 학습된 데이터를 바탕으로 등장인물 삽화를 생성한다. 등장인물 삽화 생성 시 다음과 같은 방법을 통해 등장인물 삽화의 활용도를 높일 수 있다.

동작 방식

* 작가

동작 하는 방식은 작가가 글을 작성하는 과정과 독자가 글을 읽는 과정으로 크게 나누었다. 웹에서 작가는 ‘소설 작성’을 하게 된다. 소설 작성을 하기 앞서서 인공지능 모델이 이미지를 생성하는데 거치는 과정이 다소 있기 때문에, 우선 주인공 생성 과정이 필요하다. 주인공 생성은 civitai를 통해서 작가가 원하는 스타일의 주인공을 고를 수 있다. civitai에서 4가지 모델을 채택하여 스타일을 4가지 다른 것으로 만들어 낼 수 있다. 다음으로 주인공의 생김새를 결정하는 방식은 tag 방식을 통해서 얼굴이나, 눈 색, 머리 색, 나이, 성별을 택하여 일차적으로 주인공을 생성할 수 있다. 다음 과정으로는 선택한 주인공을 Dreambooth 모델에 넣어 준다. 그 결과 동일한 인물에 대하여 instance화 되어 얼굴이 지속적으로 계속 나오고 나머지 옷과 같은 다른 형상을 선택하여 다양한 상황의 인물을 생성할 수 있다.

* 독자

독자는 기존 시중에 나와 있는 독서 프로세스와 동일하게 진행된다.

기대효과

이 프로젝트의 목표는 소설을 읽는 독자들에게 글만 있는 것이 아니라 그림을 삽입하여 몰입도를 더욱 이끌어내고 작가는 자신이 구상한 글에 대한 그림을 그리는 행위를 외주를 통해 타인에게 맞기는 과정에서 발생하는 문제점을 해결할 수 있다. 문제점은 외주를 맡기게 되면 발생하는 비용을 인공지능 모델로 생성하여 매우 절감할 수 있고, 그림 그리는 데 걸리는 시간적 문제를 줄일 수 있다. 또한 작가가 외주로 맞긴 그림이 원하는 스타일이 아닐 수 있는데, 인공지능 모델은 작가가 원하는 이미지를 뽑아 낼 수 있도록 수정하고 커스텀 할 수 있다는 장점이 있다.

(프로젝트 목표, 시스템 설계안, 동작 방식, 기대효과 등 기술합니다.)

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 |  |
| 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 | 1021283990000 | 등록일  2020.06.24 | ㈜자이언트스텝 | AI 기반의 얼굴 애니메이션 구현을 위한 학습데이터 생성 방법, AI 기반의 얼굴 애니메이션 구현 방법 및 컴퓨터 판독 가능한 저장매체 | Y |
| 2 | 한국 | 1020230054211 | 2023.04.24 | 김동준 | 인공지능 기반 대화 이미지 생성 방법 및 이를 수행하는 장치 | A |
| 3 | 한국 | 1020210097314 | 2021.08.09 | 박경미 | 인공지능 기반 이미지 자동 생성 시스템 | Y |
| 4 |  |  |  |  |  |  |

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

1. 선행기술과의 대비
2. 선행문헌 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 발명의 명칭 | | AI 기반의 얼굴 애니메이션 구현을 위한 학습데이터 생성 방법, AI 기반의 얼굴 애니메이션 구현 방법 및 컴퓨터 판독 가능한 저장매체 | | |
| 문헌번호 | | 1021283990000 | 관련성 | Y |
| 등록일 | | 2020.06.24 | 출원인 | ㈜자이언트스텝 |
| 기술  요지 | 목적 | 본 발명의 일 실시예에 따른 AI 기반의 얼굴 애니메이션 구현을 위한 학습데이터 생성 방법은, 제1 카메라에 의해 얼굴에 부착된 마커의 3차원 위치 정보를 포함하는 제1 데이터를 획득하는 단계; 제2 카메라에 의해 상기 마커가 부착된 얼굴의 영상을 포함하는 제2 데이터를 획득하는 단계; 상기 제2 데이터를 상기 얼굴의 영상에서 상기 마커의 이미지만을 제거하여 제3 데이터로 변환하는 단계; 및 상기 제1 데이터 및 상기 제3 데이터를 학습 모델의 학습을 위한 학습데이터로 사용하는 단계를 포함할 수 있다. | | |
| 구성 | 상기 과제를 해결하기 위해서, 본 발명의 일 실시예는 AI 기반의 얼굴 애니메이션 구현을 위한 학습데이터 생성 방법을 제공한다.  상기 AI 기반의 얼굴 애니메이션 구현을 위한 학습데이터 생성 방법은, 제1 카메라에 의해 얼굴에 부착된 마커의 3차원 위치 정보를 포함하는 제1 데이터를 획득하는 단계; 제2 카메라에 의해 상기 마커가 부착된 얼굴의 영 상을 포함하는 제2 데이터를 획득하는 단계; 상기 제2 데이터를 상기 얼굴의 영상에서 상기 마커의 이미지만을 제거하여 제3 데이터로 변환하는 단계; 및 상기 제1 데이터 및 상기 제3 데이터를 학습 모델의 학습을 위한 학습데이터로 사용하는 단계를 포함할 수 있다.  한편, 본 발명의 다른 실시예는 AI 기반의 얼굴 애니메이션 구현 방법을 제공한다.  상기 AI 기반의 얼굴 애니메이션 구현 방법은, 얼굴에 부착된 마커의 3차원 위치 정보를 포함하는 데이터 및 상기 마커가 부착된 얼굴의 영상에서 상기 마커의 이미지만을 제거한 데이터를 포함하는 학습 데이터를 생성하는 단계; 학습 모델에서 상기 학습 데이터를 이용하여 상기 마커의 3차원 위치 정보를 학습하는 단계; 및 상기 학 습 모델에서 마커가 미부착된 얼굴의 영상을 입력 받아서 가상의 마커의 3차원 위치 정보를 출력하는 단계를 포함할 수 있다.  덧붙여 상기한 과제의 해결수단은, 본 발명의 특징을 모두 열거한 것이 아니다. 본 발명의 다양한 특징과 그에 따른 장점과 효과는 아래의 구체적인 실시형태를 참조하여 보다 상세하게 이해될 수 있을 것이다. | | |
| 효과 | 본 발명의 일 실시예에 따르면, 얼굴 애니메이션 구현을 위한 작업 과정에서 기존의 마커 부착 및 정렬 과정을 생략할 수 있어 보다 빠르고 정확하게 얼굴 애니메이션을 구현할 수 있다. | | |
| 도면 |  | | |
| 검토  의견 | 유사점 | **실제 사진을 이용하여 생성 모델에 넣어서 비슷한 얼굴이 나올수 있도록 생성 모델을 학습한다는 점이다.** | | |
| 차이점 | **우리 기술에는 데이터 증강에 대한 기술을 부여하지 않아서 과정 중 목적에 동일하게 부합하는 것이 없다.** | | |

1. 선행문헌 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 발명의 명칭 | | 인공지능 기반 대화 이미지 생성 방법 및 이를 수행하는 장치 | | |
| 문헌번호 | | 1020230054211 | 관련성 | A |
| 공개일 | | 2023.04.24 | 출원인 | 김동준 |
| 기술  요지 | 목적 | 본 발명은 딥러닝을 이용한 감정인식을 통해 대화형 이미지를 생성하는 방법에 관한 것으로 본 발명에 따른 인공지능 기반 대화 이미지 생성 방법은 입력된 이미지의 객체를 인식하는 단계; 상기 인식된 객체의 객체 별 감정을 학습된 신경망을 통해 인식하는 단계; 상기 인식된 감정의 조합 정보와 상기 매칭 후보 텍스트 내 문장 단위의 감정 조합 정보를 비교하여 추출된 매칭 텍스트를 수신하는 단계; 및 상기 객체의 적어도 일부와 상기 텍스트를 매핑하여 대화 이미지를 생성하는 단계를 포함한다. 본 발명에 따르면, 임베딩 기반의 판단 모델과 신경망 모델을 이용함으로써, 감정 정보를 보다 정확하게 판단 및 분류할 수 있다. | | |
| 구성 | 본 발명은 모바일 기반의 감정 인식 모델을 이용하면서 이미지에 따른 감정과 매칭되는 감정 정보를 추출하는 것을 목적으로 한다.  보다 구체적으로, 본 발명은 딥러닝 기반의 자연어 처리를 통해 텍스트의 감정을 분류하고, 이미지의 감정에 매칭되는 텍스트로 대화 이미지를 생성하는 것을 목적으로 한다.  또한, 본 발명은 개인화된 이미지의 장소나 배경 자체의 특징을 반영하고, 개인정보의 이용을 최소화한 대화 이 미지 생성 방법을 제안하는 것을 목적으로 한다.  상기 기술적 과제를 해결하기 위한 본 발명에 따른 인공지능 기반 대화 이미지 생성 방법은 입력된 이미지의 객 체를 인식하는 단계; 상기 인식된 객체의 객체 별 감정을 학습된 신경망을 통해 인식하는 단계; 상기 인식된 감 정의 조합 정보와 상기 매칭 후보 텍스트 내 문장 단위의 감정 조합 정보를 비교하여 추출된 매칭 텍스트를 수 신하는 단계; 및 상기 객체의 적어도 일부와 상기 텍스트를 매핑하여 대화 이미지를 생성하는 단계를 포함한다. | | |
| 효과 | 본 발명에 따르면, 임베딩 기반의 판단 모델과 신경망 모델을 이용함으로써, 감정 정보를 보다 정확하게 판단 및 분류할 수 있다.  또한, 본 발명은 임베딩 기반의 판단 모델을 이용함으로써, 저사양 모바일에서도 용이하게 이용할 수 있다.  또한, 사용자의 이미지를 신경망을 학습시킴으로써 사용자의 특성에 따른 감정 정보에 대한 적응성을 높임과 동 시에 학습을 위한 개인정보의 이용을 최소화할 수 있다 | | |
| 도면 |  | | |
| 검토  의견 | 유사점 | **이미지 정보를 학습 시키는 점이 동일한 것 같다.** | | |
| 차이점 | **하지만 과정이 imageto text가 아니라 text to image모델이여서 모델의 난이도 쉽고, 우리가 하고자 하는 과정을 담을 수 없을 것 같다.** | | |

1. 선행문헌 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 발명의 명칭 | | 인공지능 기반 이미지 자동 생성 시스템 | | |
| 문헌번호 | | 1020210097314 | 관련성 | Y |
| 공개일 | | 2021.08.09 | 출원인 | 박경미 |
| 기술  요지 | 목적 | 인공지능 기반 이미지 자동 생성 시스템이 개시된다. 이 시스템은 영상 콘텐츠로부터 복수의 클립 영상을 생성하는 클립 생성부, 각각의 클립 영상을 분석하여 하나 이상의 감정 항목을 매핑하는 클립 감정 매핑부, 사용자가 선호하는 하나 이상의 영상 콘텐츠를 구성하는 클립 영상들의 감정 항목에 근거하여 사용자 선호 감정 정보를 생성하는 사용자 선호 정보 생성부, 및 타겟 동영상의 클립 영상들 중에서 사용자 선호 감정 정보에 속하는 감정 항목이 매핑된 클립 영상들을 이용하여 하이라이트 콘텐츠를 생성하는 하이라이트 생성부를 포함한다. | | |
| 구성 | 일 양상에 따른 인공지능 기반 이미지 자동 생성 시스템은 영상 콘텐츠로부터 복수의 클립 영상을 생성하는 클 립 생성부, 각각의 클립 영상을 분석하여 하나 이상의 감정 항목을 매핑하는 클립 감정 매핑부, 사용자가 선호하는 하나 이상의 영상 콘텐츠를 구성하는 클립 영상들의 감정 항목에 근거하여 사용자 선호 감정 정보를 생성 하는 사용자 선호 정보 생성부, 및 타겟 동영상의 클립 영상들 중에서 사용자 선호 감정 정보에 속하는 감정 항 목이 매핑된 클립 영상들을 이용하여 하이라이트 콘텐츠를 생성하는 하이라이트 생성부를 포함할 수 있다.  하이라이트 생성부는 각각의 클립 영상별로 프레임을 랜덤 추출하며, 랜덤 추출된 프레임들을 합성하여 단일 이 미지로 이루어진 하이라이트 콘텐츠를 생성할 수 있다.  하이라이트 생성부는 사용자 선호 감정 정보에 속하는 감정 항목이 매핑된 클립 영상들을 랜덤하게 선택하고 조 합하여 하이라이트 콘텐츠를 생성할 수 있다.  클립 감정 매핑부는 클립 영상을 분석하여 클립 정보를 생성하는 클립 정보 생성부, 및 클립 정보에 근거하여 클립 영상별 하나 이상의 감정 항목을 매핑하는 감정 매핑부를 포함할 수 있다.  감정 매핑부는 클립 정보를 다차원 벡터로 변환하는 벡터 생성부, 다차원 벡터를 클러스터링하여 그룹화하는 벡 터 그룹화부, 및 그룹별 고유 감정 항목에 따라 해당 클립 영상에 하나 이상의 감정 항목을 매핑하는 매핑부를 포함할 수 있다.  한편, 일 양상에 따른 인공지능 기반 하이라이트 콘텐츠 생성 방법은 타겟 동영상으로부터 복수의 클립 영상을 생성하는 클립 생성 단계, 각각의 클립 영상을 분석하여 하나 이상의 감정 항목을 매핑하는 클립 감정 매핑 단 계, 및 타겟 동영상의 클립 영상들 중에서 사용자 선호 감정 정보에 속하는 감정 항목이 매핑된 클립 영상들을 이용하여 하이라이트 콘텐츠를 생성하는 하이라이트 생성 단계를 포함할 수 있다. | | |
| 효과 | 본 발명에 따르면, 사용자가 선호하는 감정에 적합한 맞춤형 하이라이트 콘텐츠를 생성하여 제공하는 것이 가능 해진다. | | |
| 도면 |  | | |
| 검토  의견 | 유사점 | **모델의 전치리 단계와 공간 벡터로 생성하고 이를 매핑하는 과정이 stabel diffusion의 모델과 유사하다** | | |
| 차이점 | **모든 과정이 동영상 처리를 위한 모델이여서 입출력 구조가 다르다** | | |

1. 결론

|  |
| --- |
| 유사한 모델을 이용하여 특허를 만든 것을 여러 개 발견할 수 있었다. 특히 오히려 생성 모델을 이용하여 데이터 셋을 증강시켜서 작동하는 특허는 지금 우리 조가 제시하고 있는 모델의 성능을 더욱 뛰어나게 해줄 수 있을 것이라고 생각되는 바이다. 하지만 특허가 공개되었다는 말은 기간이 지났다는 말인데, 저희가 제시하는 모델들은 만들어 진지 3-4개월 된 모델이여서 경쟁력이 있다는 판단이다. 더욱이 우리와 같이 2가지 stable diffusion 모델을 이용하여 원하는 사람을 선정하고 그 사람이 계속 등장하지만 변화하는 것을 만든 메커니즘은 없을 것으로 판단하는 바이다. 만약 우리의 모델을 특허로 등제한다면, 웹소설 시장에 상당한 전환점이 될 것이며, 웹소설 시장의 제 2의 전성기를 불러 올 것이라고 판단한다. |