😀😅😭 Emopic

텍스트 감정분석을 통한 이미지 표정 변화 서비스

Level up 조경원 강유경 김태영 이효은

[index]

Emopic

😂❤️👍😭🙏

chapter 1 기획배경

chapter 2 데이터 설명

chapter 3 데이터 전처리

chapter 4 API 및 크롤링

chapter 5 Kobert 모델

chapter 6 Stargan 모델

chapter 7 서비스 설명 및 시연

chapter 8 결론

1

기획배경

한경 예술에 눈뜬 AI...디지털 콘텐츠 시장 '판' 바꾼다

글로벌 리포트

그림 그리는 AI 시대 개막

이미지 생성하는 AI 잇단 등장
美 미술대회서 1등 수상하기도

GPU 발달로 제작 기술 크게 진화
"2028년 디지털 이미지 시장 48억弗"

이미지 저작권 모호해 논란일기도
AI 수용한 셔터스톡 "새 세계 열렸다"

지난 8월 게임 디자이너 제이슨 앨런은 인공지능(AI) 프로그램으로 생성한 그림 '스페이스 오페라극장'(사진)을 콜로라도 주립박람회 미술대회에 출품해 1등상을 받았다. 150년 미술대회 역사상 AI의 그림이 1등을 차지한 것은 처음이었다.

AI 앱' 상반기 글로벌 다운로드 3억건 돌파...챗봇·이미지가 주류

올 상반기 전 세계 모바일 시장에서 인공지능(AI) 애플리케이션이 3억건 이상 다운로드된 것으로 알려졌다. '챗GPT'와 '렌사 AI' 등 유료 앱도 크게...

1주 전

매드타임스

AI 앱 다운로드 수 3억건 돌파, 이미 2022년 넘어서... AI 기술이 가장 많이 사용되는 분야는 '챗봇', '이미지 생성'

[매드타임스 신성수 기자] 글로벌 모바일 시장 데이터 분석 기업 센서타워(Sensor Tower)가 2023년 전 세계 인공지능(AI) 앱의 다운로드와 수익 성과...

1주 전

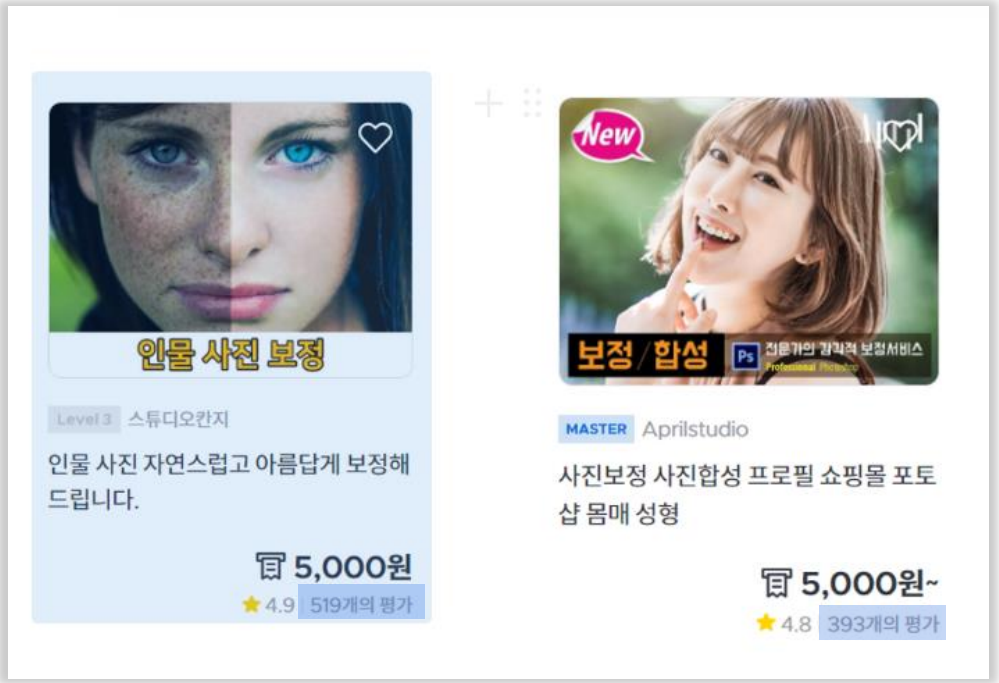
뉴스1

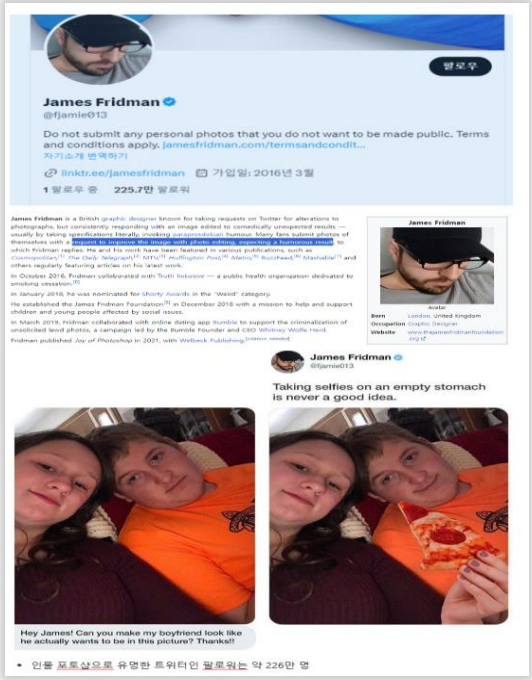
'AI 앱 전성시대' 상반기 다운로드 3억건 돌파...챗봇·이미지 생성 앱 인기 - 뉴스1

(서울=뉴스1) 손엄지 기자 | 다양한 인공지능(AI) 앱이 등장하면서 글로벌 AI 시장이 급속도로 성장하고 있다. 이용자들은 챗봇과 이미지 생성·처리...

1주 전

디지털 콘텐츠에서 **얼굴 표정의 중요성**이 **증대**되고 있으며
이미지 생성 AI가 각광받고 있음





이미지 **포토샵 작업에 대한 수요**가 상당함

2.1 글로벌 전체 시장 규모

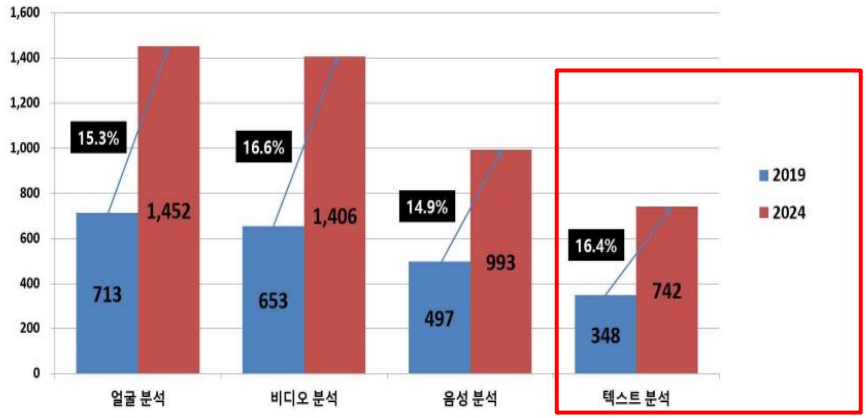
□ 전 세계 감정 분석 시장은 2018년 20억 9,020만 달러에서 연평균 성장률 17.05%로 증가하여, 2023년에는 45억 9,330만 달러에 이를 것으로 전망됨

[그림 2-1] 글로벌 감정 분석 시장 규모 및 전망

※ 출처 : TechNavio, Global Emotion Analytics Market, 2019

[그림 2-3] 글로벌 감정 분석 시장의 종류별 시장 규모 및 전망

(단위: 백만 달러)

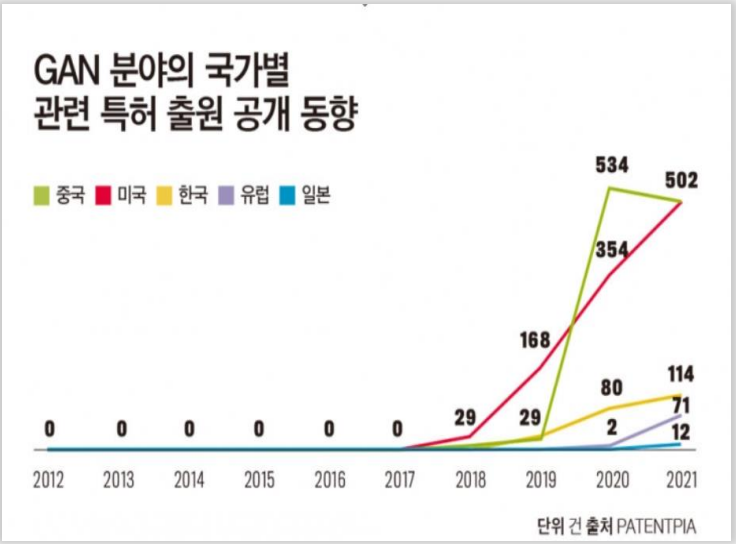


※ 출처 : MarketsandMarkets, Emotion Analytics Market, 2019

텍스트 감정 분석 시장이 급속 성장하고 있음

2020~2021	팬데믹으로 인한 AI 발전 가속화	2021년 한 해 AI특허출원 2015년 대비 30배 증가
오늘날	생성형 AI의 대중화 시작	챗GPT, DALL-E, Midjourney, Stable Diffusion의 등장

그리고 2023년을 맞아 챗GPT를 계기로 GAN 기술 및 생성형 AI가 잠금해제되고 대중화 시대를 맞이했다. 포문을 연 챗GPT는 온라인의 모든 문서를 학습해 질문에 대한 문서형의 최종 맞춤형 출력물을 생성하지만, 이후 그림, 동영상, 음악 등을 생성형 AI가 더욱 고도화할 것으로 전망된다. 생성형 AI 기술은 일반적으로 비용이 많이 들고 전문인력이 필요한 고급 생



출처 : 포브스

생성형 AI인 **GAN 관련 기술의 대중화**로 활용 영역 확대

Emopic

텍스트 감정분석을 통한 이미지 표정 변화 서비스

2

데이터 설명

데이터 종류	활용 데이터	출처
텍스트 데이터	감정 분류를 위한 대화 음성 데이터 셋	AIHUB
이미지 데이터	한국인 감정인식을 위한 복합 영상	AIHUB

총 **7가지 감정**에 대해 **5명이 라벨링** 한 데이터로 발화에 대한 감정 라벨링

총 50만장의 이미지로 구성되어 있는 **감정별 이미지** 데이터

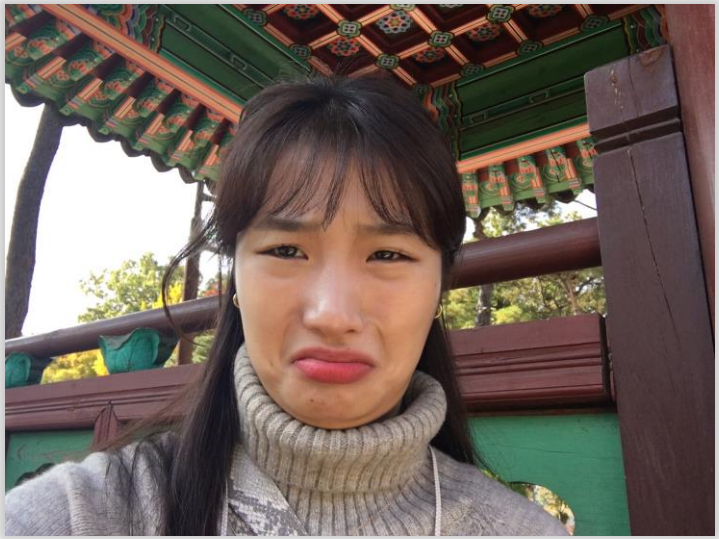
3

데이터 전처리

〈변경 전〉		〈변경 후〉		
이미지 라벨	**텍스트 라벨**	**이미지 라벨**	**텍스트 라벨**	
기쁨	happiness	기쁨	기쁨	
분노	angry	분노	분노	
상처	–	–	–	〉〉 이미지 데이터 삭제
–	disgust	–	–	〉〉 텍스트 데이터 삭제
불안	fear	불안	불안	
중립	neutral	중립	중립	
슬픔	sadness	슬픔	슬픔	
당황	surprise	놀람	놀람	〉〉 이미지 라벨 변경

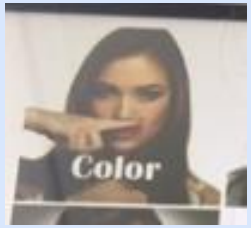
텍스트 라벨 한글로 번역 후 이미지와 텍스트 **라벨 통일**

이미지 데이터 resize 결과



학습 효율을 높이기 위해 opencv 활용하여 **이미지 데이터 resize**

삭제 예시

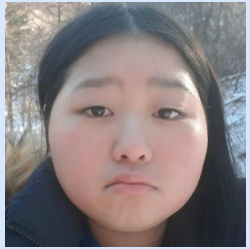


현수막, 간판 등을
얼굴로 인식한 경우

조명이 과하거나
그림자가 진 경우

기쁨_00578.jpg

라벨링이 잘못된 경우

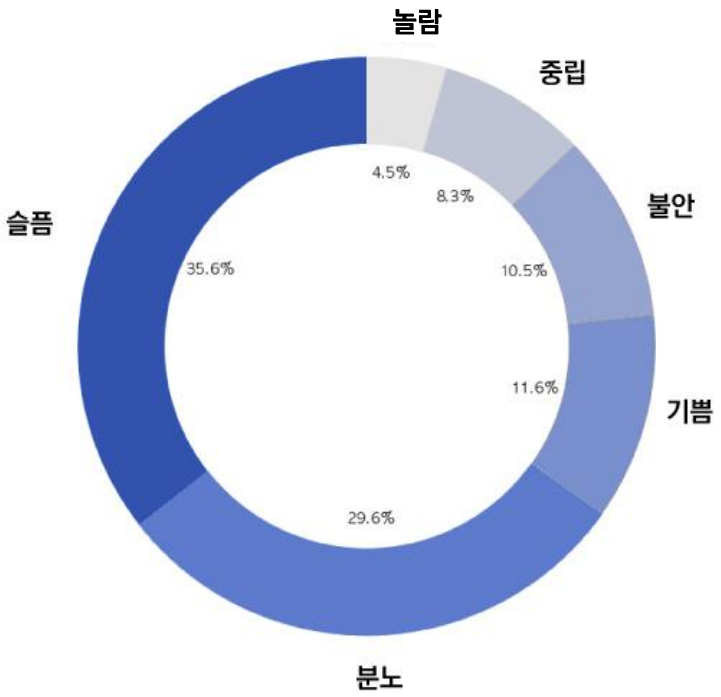


기쁨_07112.jpg

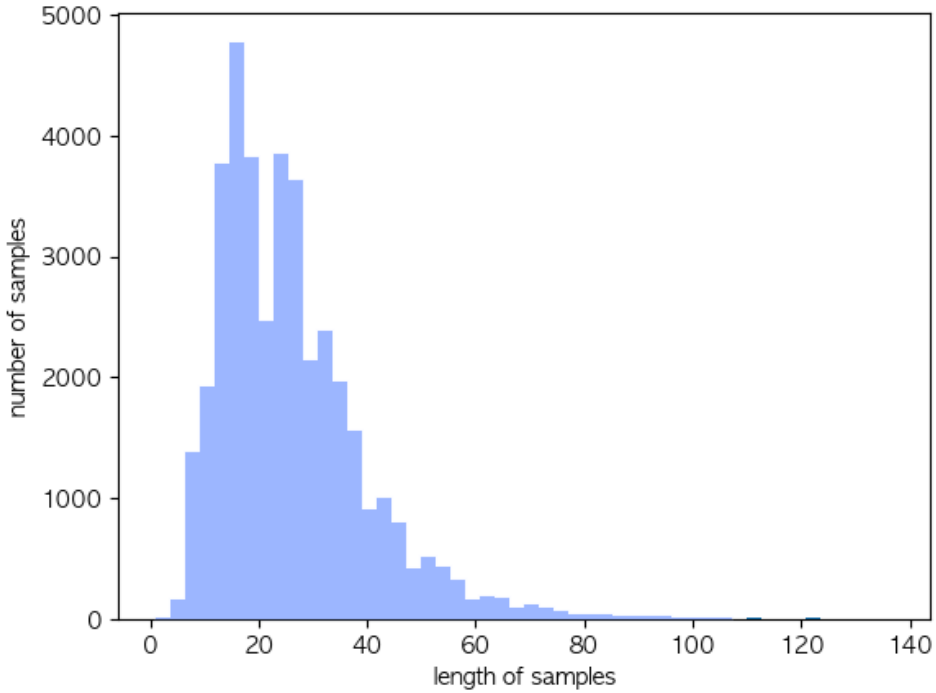
표정이 애매한 경우

사용 불가능한 데이터 삭제

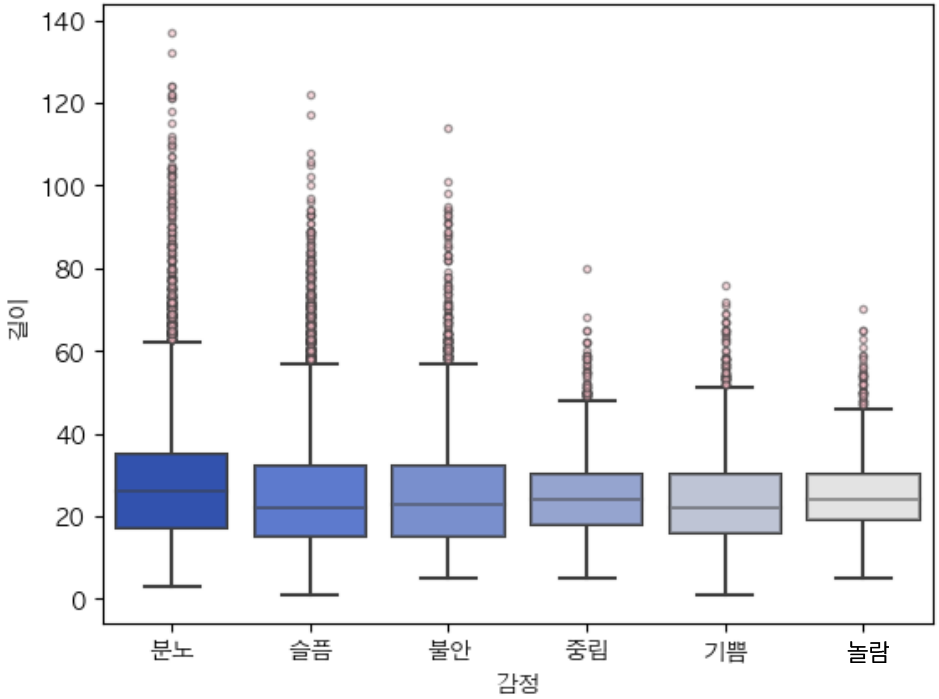
데이터 종류	활용 데이터	최종 수량
텍스트 데이터	감정 분류를 위한 대화 음성 데이터 셋	39,331
이미지 데이터	한국인 감정인식을 위한 복합 영상	115,208



슬픔, 분노, 기쁨, 불안, 중립, 놀람 순으로 많이 분포



전체 발화문 중 **길이가 71 이하**인 샘플 비율: 98.98%



감정별 발화문 길이는
분노, 슬픔, 불안 데이터에서 예외적으로 길이가 긴 발화문이 많음

〈감정별 워드클라우드〉

분노

슬픔

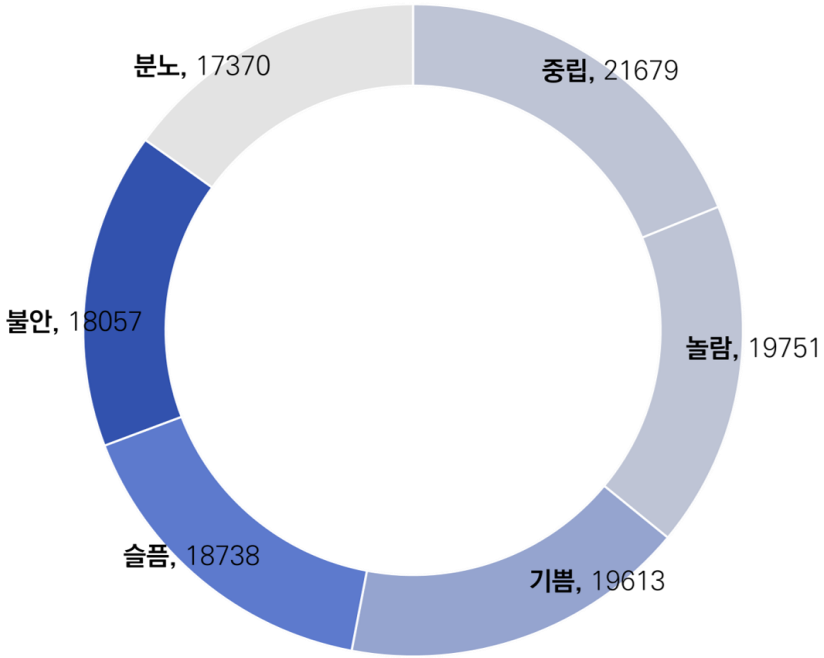
불안

〈감정별 워드클라우드〉

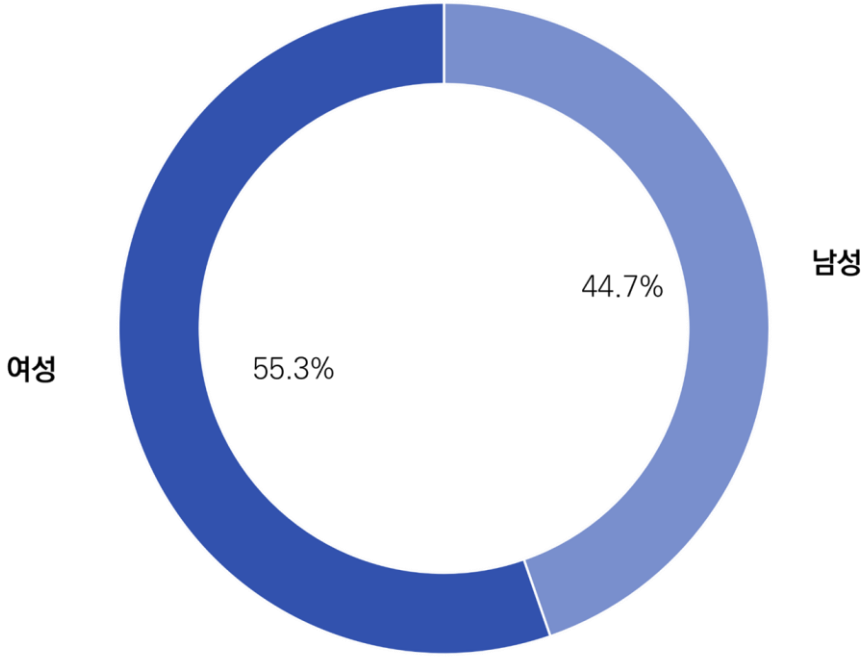
중립

기쁨

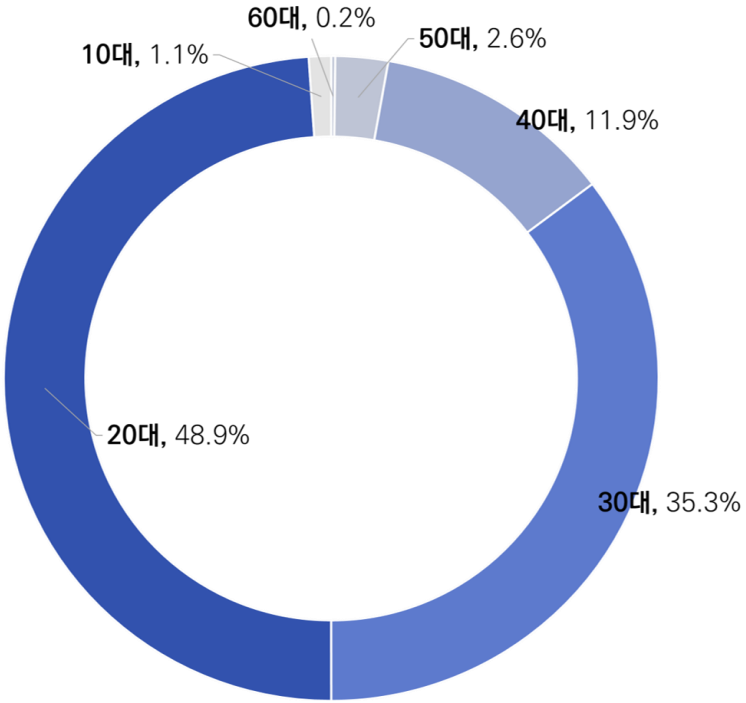
놀람



이미지 데이터의 **감정별 분포**는 고르게 분포되어 있음



성별 분포는 **여성**이 더 많음

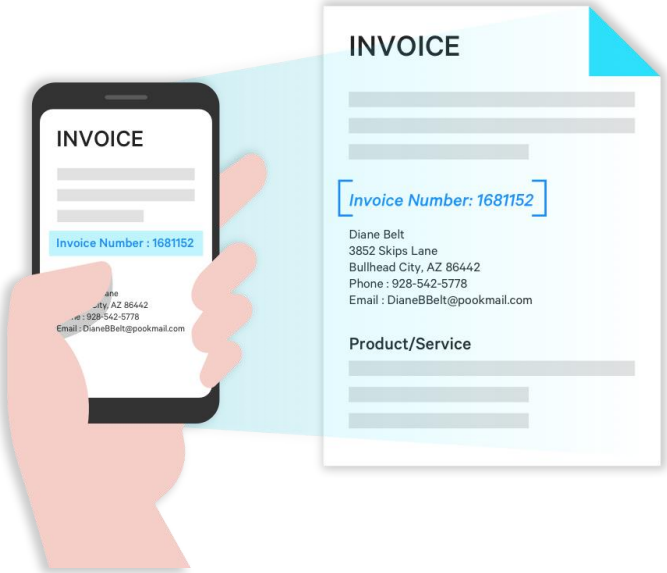


20대, 30대가 84.2%를 차지

4

API 및 크롤링

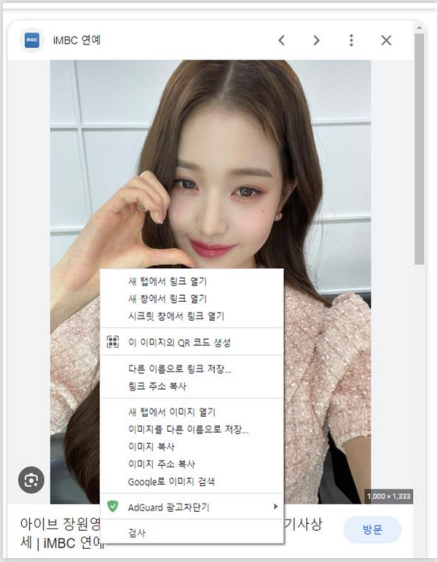
NAVER CLOUD
CLOVA OCR API



〉〉 이미지 속 문자 추출하여 컴퓨터 데이터로 변환해주는 기능

자필로 쓴 글씨 이미지를 입력 데이터로 변환하는데 활용

screenshotlayer API



〉〉 URL 입력하면 화면 캡쳐해주는 기능

이미지 URL 입력하여 **URL의 인물 이미지** 변형하는데 활용

Remove.bg API

〉〉 이미지의 배경을 제거해주는 기능

이미지 저장 시, **배경 제거 옵션** 추가 가능

욕설/한국어 나무위키

사용자가 **욕설**을 입력하면 **재입력**하도록 안내하는 기능에 활용

5

Kobert 모델

한국어 언어 모델

- Encoder-Centric Models: BERT 계열
- Decoder-Centric Models: GPT 계열
- Encoder-Decoder Models: Seq2seq 계열

Encoder-Centric Models: BERT 계열

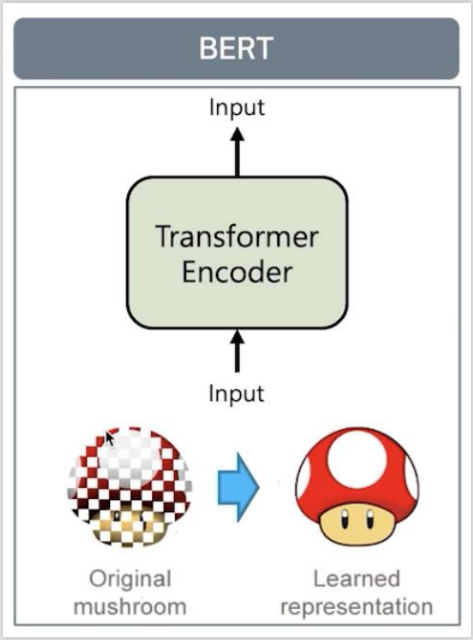
모델	개발	학습 데이터	Tokenizer	Vocab	Params
KorBERT	ETRI	뉴스/백과사전 23GB	Morphology, WordPiece	30,349(Morphology) 30,797(WordPiece)	110M
KoBERT	SKT	위키피디아 50M	Sentence-Piece	8,002	92M
HanBERT	투블럭 AI	일반/특허문서 70GB	Moran	54,000	128M

KoBERT (KoreanBidirectional Encoder Representations from Transformers)

SKTBrain에서 공개한 기계번역 모델로

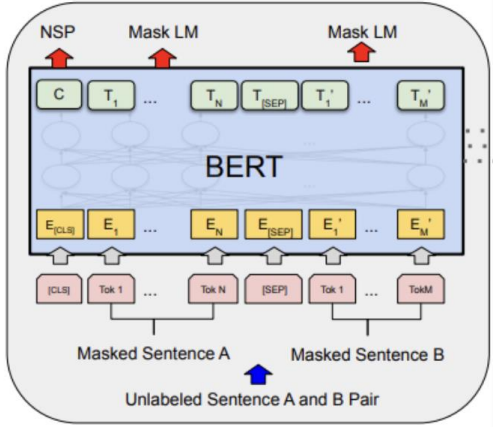
구글 BERT base multiligual cased의 한국어 성능 한계를 극복하기 위해 개발된 모델

Model	Accuracy
BERT base multilingual cased	0.875
KoBERT	0.901
KoGPT2	0.899



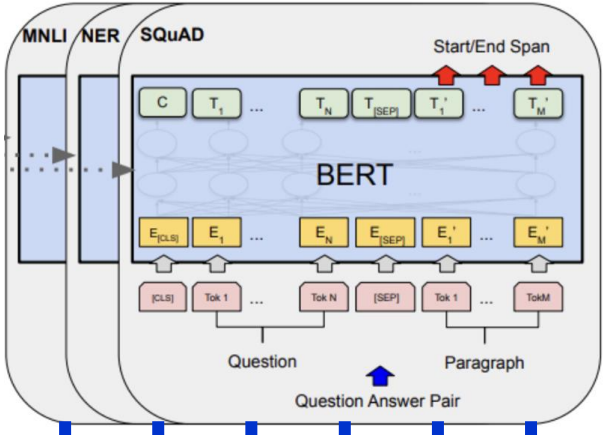
양방향성을 지니고 **마스킹 기법**을 사용하는 Bert 모델

pre – trained KoBERT



sentiment_talk
(감성발화문)

fine – tuning



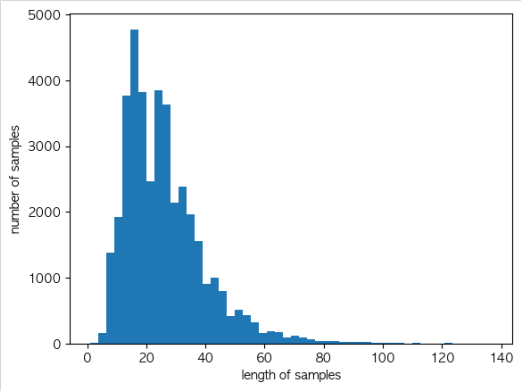
0	1	2	3	4	5
기쁨	놀람	슬픔	중립	분노	불안

```
tokenizer = KoBERTTokenizer.from_pretrained('skt/kobert-base-v1')
bertmodel = BertModel.from_pretrained('skt/kobert-base-v1', return_dict=False)
vocab = nlp.vocab.BERTVocab.from_sentencepiece(tokenizer.vocab_file, padding_token='[PAD]')
```

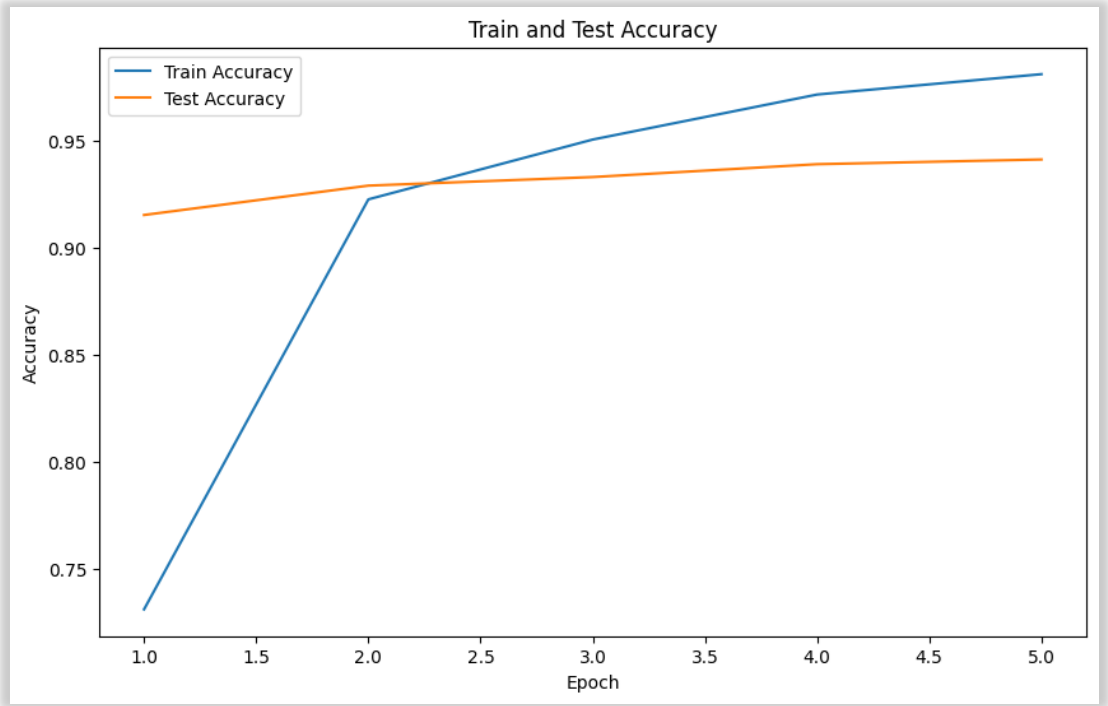
```
text = "3조 텍스트 감정분석을 통한 이미지 표정 변화"
tokens = tokenizer.tokenize(text)
print(tokens)
```

```
['_3', '조', '_', '텍', '스트', '_감정', '분석', '을', '_통한', '_이미지', '_표정', '_변화']
```

```
#setting parameters
max_len = 71 #padding
batch_size = 64
warmup_ratio = 0.1
num_epochs = 5
max_grad_norm = 1
log_interval = 200
learning_rate = 5e-5
```



전체 발화문 중 길이가 71 이하인 샘플 비율: 98.98%



epoch 1

train accuracy 0.731

test accuracy 0.915

epoch 5

train accuracy 0.980

test accuracy 0.941

확인하고싶은 문장을 넣어보세요 : 아니 벌써 시간이! 너랑 있으면 시간가는지 모르고 놀아 너무 재밌어
기쁨이 느껴집니다.

확인하고싶은 문장을 넣어보세요 : 지금 수정 요청하면 어쩌라는거야...? 나 내일 휴가야 말이야 방구야
분노이 느껴집니다.

확인하고싶은 문장을 넣어보세요 : 왜 나한테만 이래라저래라 간섭하고 난리야?? 지금 싸우자는 거야?
분노이 느껴집니다.

확인하고싶은 문장을 넣어보세요 : 악!! 주먹만한 바퀴벌레!! 웩...
놀람이 느껴집니다.

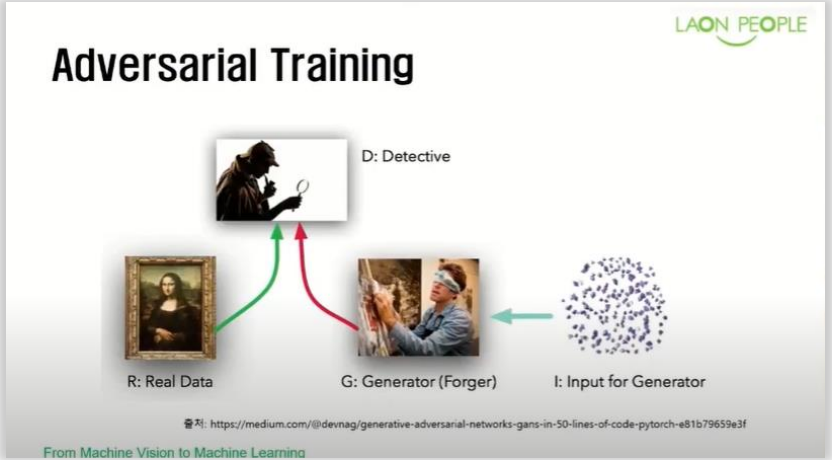
확인하고싶은 문장을 넣어보세요 :

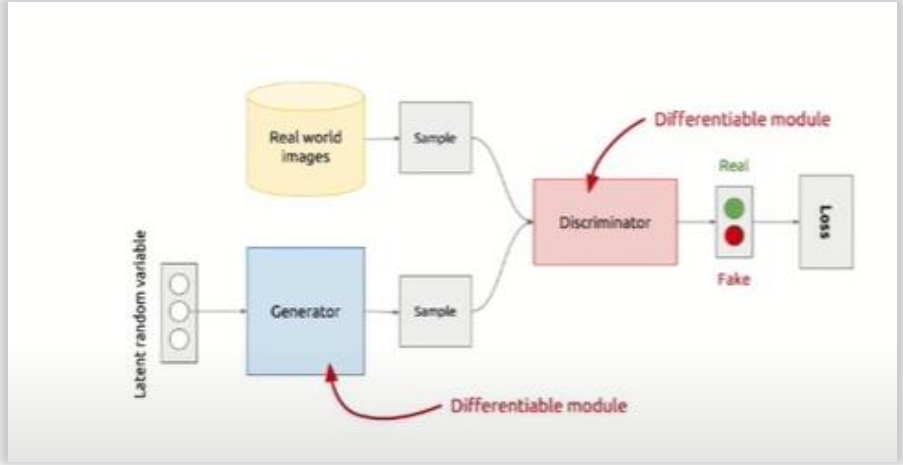
6

StarGan 모델

GAN이란?

생성적 적대 신경망(Generative Adverserial Network)





생성 모델과 판별모델이 경쟁하면서 적대적으로 학습

Output Obama Video

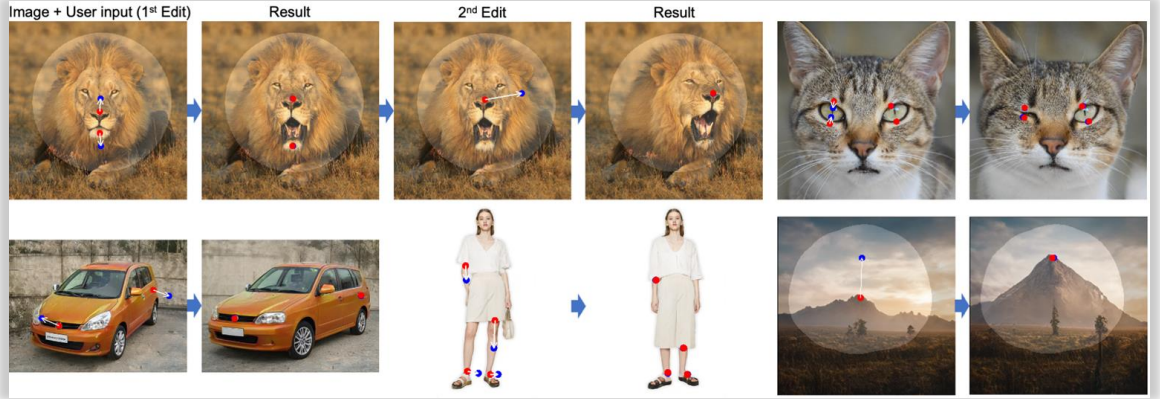
- 그림 1-10 Fake Obama Video (출처: Supasorn Suwajanakorn. 2017. Synthesizing Obama: Learning Lip Sync from Audio)

〉〉 오바마 딥페이크 연설 영상

Original showing Alison Brie

Deepfake showing Jim Carrey instead of Brie

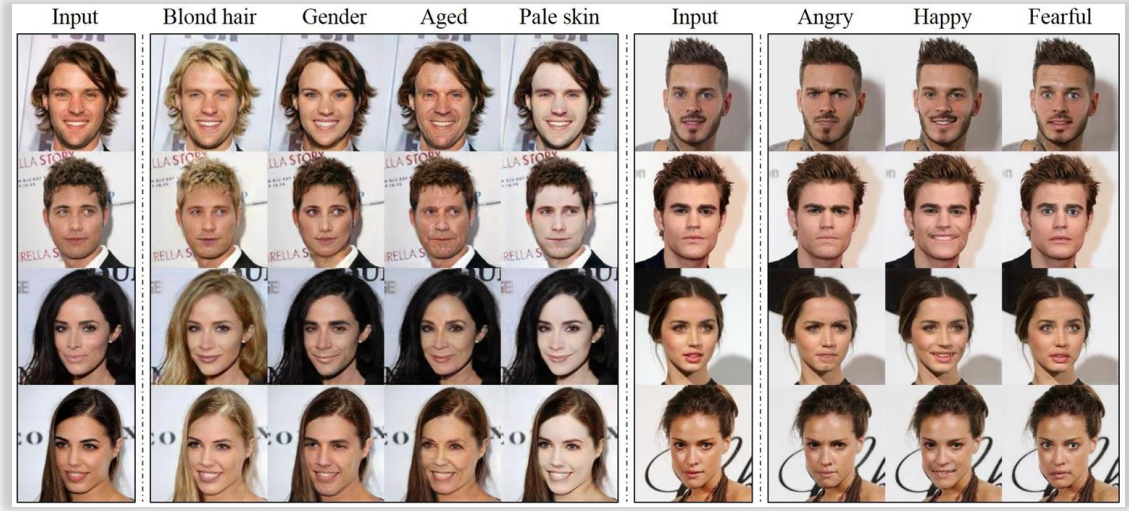
〉〉 앨리슨 브리와 짐 캐리 합성



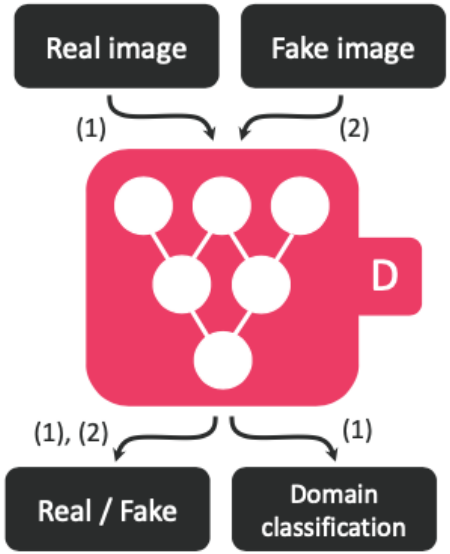
〉〉 마우스로 드래그하면 그에 맞게 사진이 수정되는 DragGAN

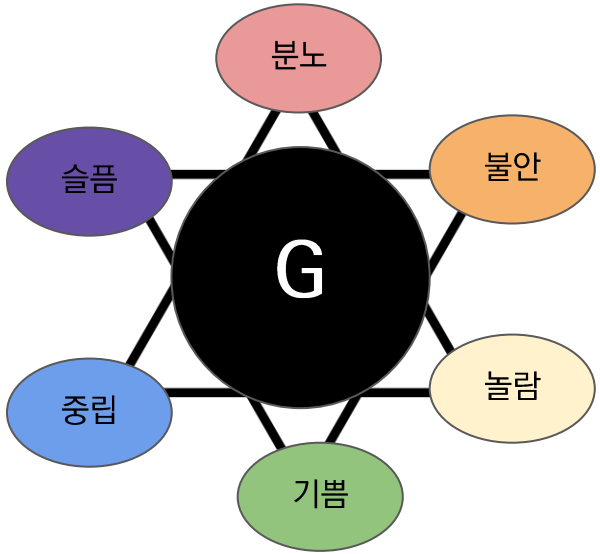
StarGan 모델 선정 이유

StyleGAN



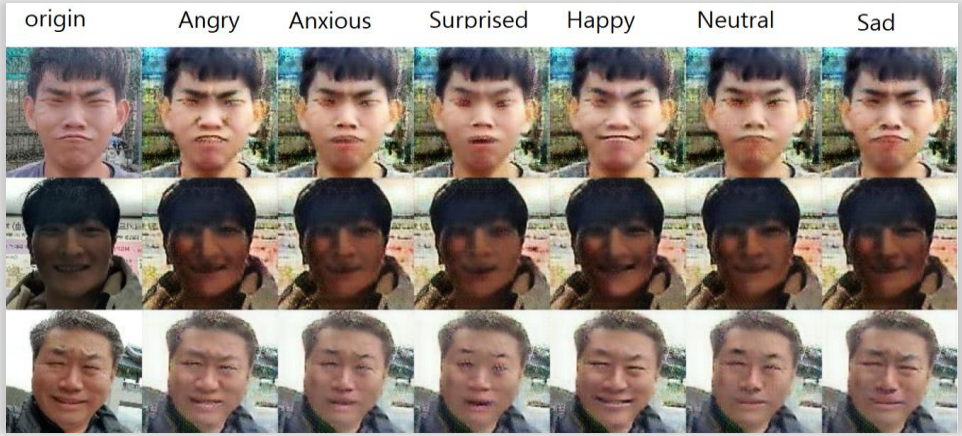
StarGAN


Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real / Fake
Domain classification


분노
슬픔
불안
G
중립
놀람
기쁨

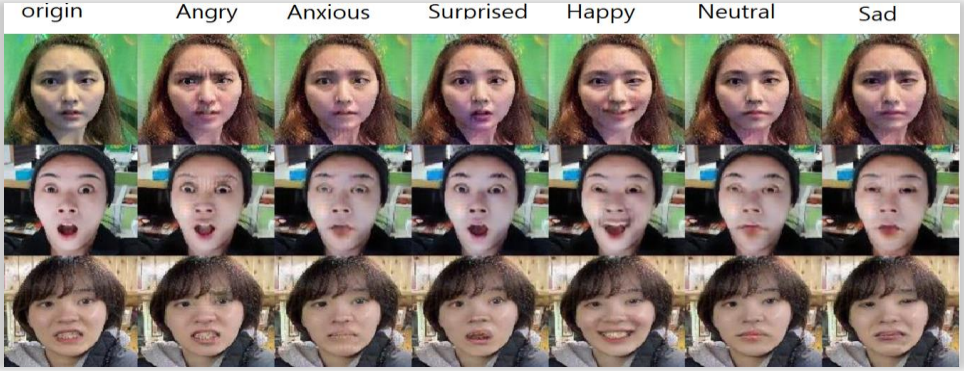
G: Generator

StarGan 모델 학습 결과



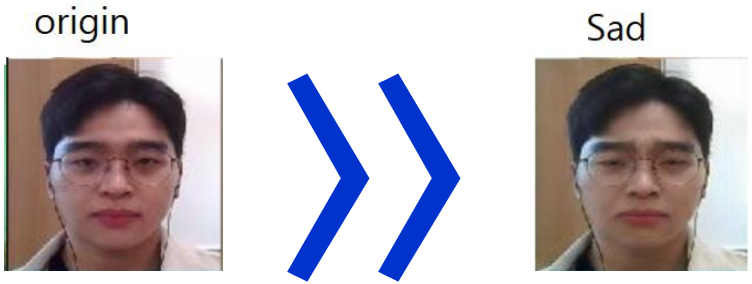
batch_size : 16 iteration : 1만 번

batch_size : 16 iteration : 10만 번



batch_size : 16 iteration : 10만 번

batch_size : 48 iteration : 20만 번



한 가지 도메인에 대해서만 이미지를 생성하게 변형

7

서비스 설명 및 시연

Emopic

텍스트 감정분석을 통한 이미지 표정 변화 서비스

8

결론

POINT 01

감정 인식 정확도 향상

더 많은 다양한 데이터를 수집해 세분화 하게 감정을 나누고 라벨링 품질을 향상 시켜
라벨들간의 일관성을 유지시키고 라벨링 오류를 최소화 한다

POINT 02

다양한 표정 및 감정 처리

해당 서비스는 6가지 감정을 분류 하지만 더 많은 감정과 표정을 인식하고 처리할 수 있도록
다양한 데이터셋을 사용하여 서비스의 다양성을 높이도록 한다.

POINT 03

이미지 사이즈 문제 해결

초기 기획에 맞게 이미지를 임의로 Resize 하지 않고
원본 그대로의 이미지에 표정만 바꿔서 반환힐 수 있도록 연구가 필요하다.

POINT 01

게임 및 엔터테인먼트

- 텍스트 입력을 통해 감정분석해 게임캐릭터의 실시간 표정 변경가능
- 디지털 컨텐츠 산업에서 해당 모델의 캐릭터의 감정을 자연스럽게 표현하거나 특정 장면에 대한 감정을 표현하는 데 사용

POINT 02

의료 및 심리치료

감정 인식 훈련 도구 – 텍스트 감정 인식을 통해 환자들의 자체 관찰 및 피드백으로 감정을 관리하고
환자들의 감정변화를 시각적으로 확인하고 심리 변화를 관찰 가능

Emopic

Q&A

감사합니다