

대한민국 ICT 인재를 키우는 힘, 한이음 ICT 멘토링 프로젝트 참여기 2

딥러닝(GAN) 기반 나만의 아바타 화상회의 서비스 개발

강자원

컴퓨터시스템응용기술사
/ KBS MNC(Media
Network Center)팀
덕성여자대학교 김지윤,
문지연, 박지혜, 안효원

이번 호에서는 2022년 4월부터 11월까지 덕성여대 학생들과 함께 진행한 딥러닝 기술인 GAN 알고리즘을 활용한 프로젝트에 대해 소개하는 시간을 가져보려 한다.

딥러닝(GAN) 기반 나만의 아바타 화상회의 서비스란?

3D 아바타를 사용한 화상회의 플랫폼이다. 코로나19 사태의 장기화로 화상회의 수요가 증가함에 따라 고안되었다. 자신의 모습과 닮은 개성 있는 아바타를 생성하고 아바타가 자신의 모습을 대체하여 화상회의에 참가한다. 딥러닝 알고리즘에 의해 사용자의 표정이 실시간 분석되어 아바타에 반영된다. 또한, 아바타를 이용해 다양한 표정의 이모티콘을 만들어 커뮤니티, 채팅, 화상회의 등 서비스에 사용할 수 있다. 또한, 안드로이드 모바일의 챗봇 기능을 통해 화상회의 일정 관리도 가능하다.



프로젝트 성과

참여한 프로젝트 중 특히 우수 프로젝트로 선정되어 앱 개발 관련 기획의 전반과 아이디어 및 기술에 대해 특허출원을 신청하여 현재 예비 출원을 부여받았다.

- 한이음 공모전 - 입선
- 한국정보처리학회 논문 게재

화상회의 서비스를 위한
GAN 기반 아바타 생성 및 애니메이션 구현 기술
(2022년 한국정보처리학회 학술대회논문집, 29권 2호)



프로젝트 명 : Avatwin - 딥러닝(GAN) 기반 나만의 아바타 화상회의 서비스

팀원	이름	소속	부서/학과	역할	직위/학년
멘토	강자원	KBS	MNC	프로젝트 지도	감독
멘티	팀장	덕성여자대학교	컴퓨터공학과	PM, 백엔드 개발 인프라설계	4학년
	팀원1			프론트엔드 개발	
	팀원2			안드로이드, 백엔드 개발 인공지능 알고리즘 개발	
	팀원3			인공지능 알고리즘 개발	

최종구현 화면

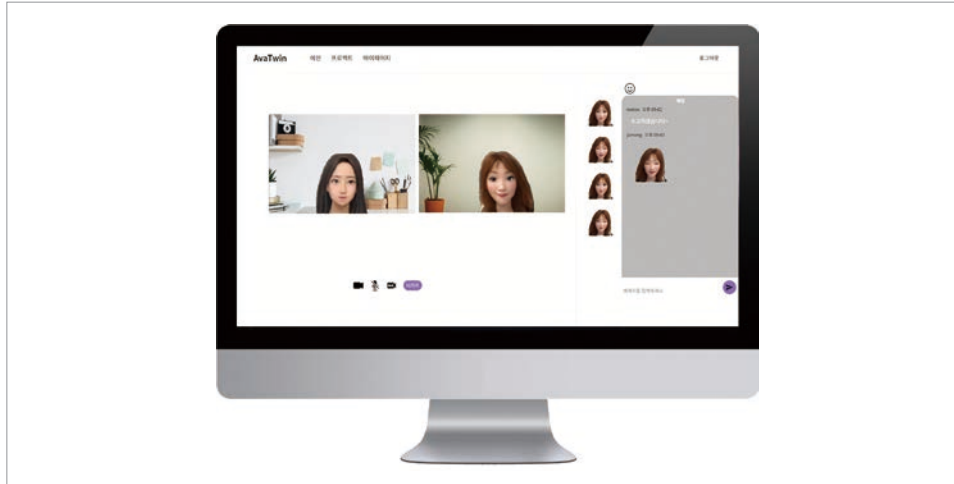
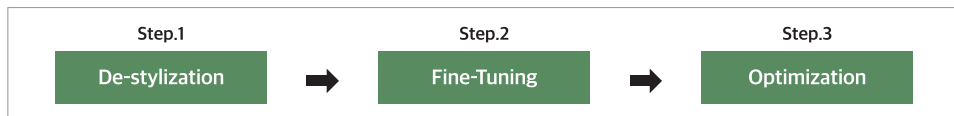


그림 1. 최종 구현된 앱 및 기능 소개

아바타 생성을 위한 딥러닝(GAN) 기반 적용 모델 및 기술

캐릭터 생성 적용모델 - DualStyleGan

DualStyleGan은 실제 얼굴에 새로운 캐릭터 얼굴의 스타일을 특성화하여 자연스러운 얼굴 합성을 만들어내는 모델이다. 실제 인물과 캐릭터를 매칭하여 나만의 아바타를 생성한다. 캐릭터 생성을 위한 학습 단계는 아래와 같다.



Step 1. De-stylization : 예술적 초상화에서 사실적인 얼굴을 복구하는 것으로 고정된 얼굴-초상화를 한 쌍으로 형성하는 것을 목표로 한다.



그림 2. 얼굴 디스타일화 결과

Step 2. Fine-Tuning : 점진적 미세 조정 단계로, 목표 도메인의 색상, 얼굴 구조, 얼굴 스타일을 차례대로 기존 도메인에 전송한다.



그림 3. Fine-Tuning 결과

Step 3. Optimization : 잠재 최적화와 샘플링으로, 조정된 모델로 얻은 이미지는 훨씬 스타일 이미지에 가까워진다. ‘여신강림’의 만화 캐릭터를 이용하여 사용자와 닮은 최종 학습된 캐릭터 이미지다. 왼쪽 이미지는 사용자, 가운데 이미지는 점진적 전이학습을 실행한 결과, 오른쪽 이미지는 스타일이 전승되어 최종 학습된 결과이다.



그림 4. 네이버 웹툰 ‘여신 강림’ 최종 학습 결과

아바타의 표정 생성 및 애니메이션 - Talking head anime model & Image Animation

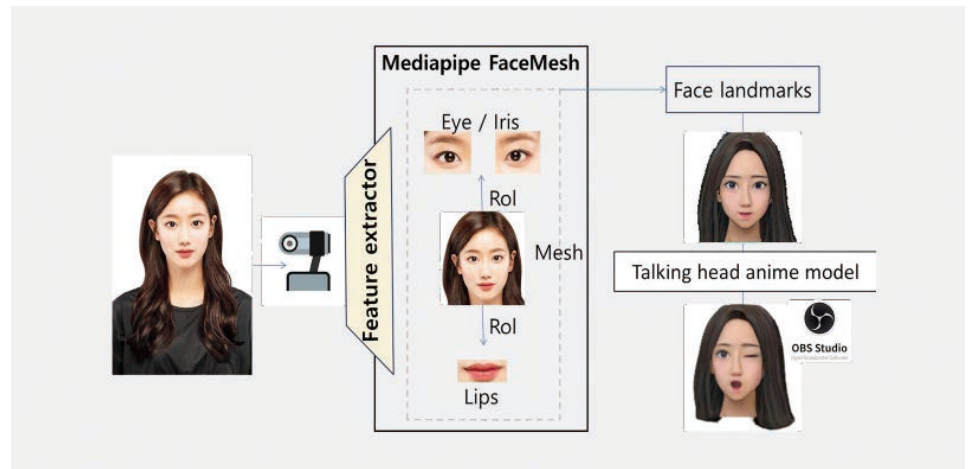


그림 5. 이미지 애니메이션 알고리즘

DualStyleGan을 통해 만들어진 고정적인 2D 캐릭터 이미지를 화상회의 안에서 사용하기 위해서는 사용자의 움직임에 따라 실시간 표정 변화와 얼굴의 움직임이 필요했다. 우선 웹캠에 비친 사용자의 움직임에 대한 얼굴 좌표값을 얻어내기 위해 Mediapipe를 이용하였다. Mediapipe의 FaceMesh는 실시간으로 사용자의 얼굴 특징을 추출할 수 있는 솔루션으로, 눈과 눈동자, 눈썹, 코, 입술, 그리고 얼굴 윤곽의 랜드마크를 추출하여 위치를 얻을 수 있다. 이렇게 추출한 얼굴 랜드마크 값을 Talking head anime model에 보낸다. Talking head anime model은 실시간으로 2D Single Image의 표정을 바꾸고 3D와 같은 움직임을 얻을 수 있는 GAN 모델이다. Talking head anime model은 눈썹, 눈, 입, 눈동자 크기 및 위치, 얼굴 회전의 값을 가진 6차원 포즈(pose) 벡터를 Face Morpher와 Face Rotator 네트워크로 전달해서 아바타 이미지를 변경한다. 여기서 6차원 포즈 벡터로 Mediapipe를 이용하여 얻은 랜드마크 값을 사용하였다. Face Morpher는 눈과 입이 열리는 정도의 표정을 변화시킨다.

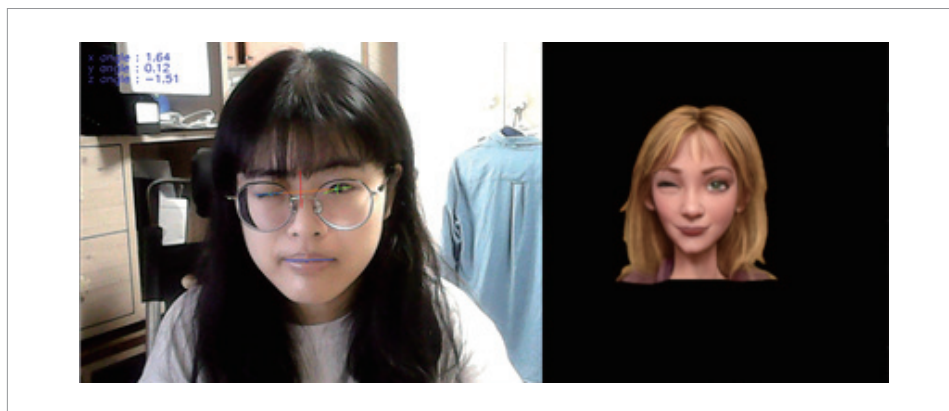


그림 6. 이미지 애니메이션 최종 결과물

이모티콘

현재 이미지 애니메이션에서 사용하는 Talking head anime model이 실시간으로 눈, 코, 입, 귀의 좌표값을 가져오기 때문에 원본 이미지에서 각 좌표값을 조절하여 다양한 표정의 이모티콘을 만들 수 있다. 고정된 표정의 2D 캐릭터를 다시 [기쁨, 슬픔, 화남, 윙크] 총 4가지의 표정을 가진 캐릭터로 변경, 생성시킬 수 있다.

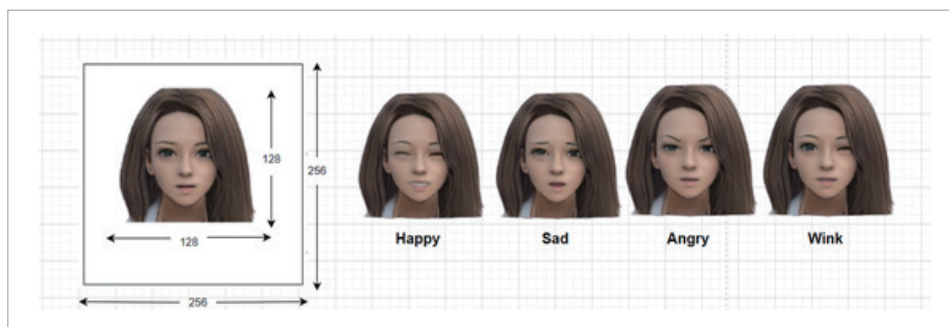


그림 7. 이모티콘 표정 구성도

먼저, 2D 캐릭터 이미지는 256×256의 크기로 RGBA 형식을 가지며 배경이 투명해야 한다. 이를 위해 이미지의 크기를 재조정하고, 컬러에 알파 채널을 추가한다. 얼굴 포즈 값에 각 표정에 따른 고정된 수치값을 적용한다. 포즈 값은 0~1 사이의 실수로 1의 값과 가까워질수록 뚜렷한 표정 변화가 나타난다. 보다 정확한 표정을 보여주기 위하여 최댓값인 1을 적용한다. 예를 들어, [기쁨] 표정의 경우 눈과 입에 해당하는 포즈 벡터값을 조절하여 웃고 있는 눈과 올라간 입 모양을 표현하였다.

화상회의 구현 및 그 외 기타 기능

화상회의 - Openvidu

Openvidu는 WebRTC 미디어 서버인 Kurento 기반의 중개 서버를 애플리케이션에 쉽게 추가할 수 있도록 완전한 기술 스택을 제공하는 플랫폼으로, 웹과 모바일 환경에서 영상 통화 기능을 쉽게 추가할 수 있도록 한다. 또한, WebRTC 미디어 서버를 구현하는 데 소모되는 리소스를 절약해서 빠르게 실시간 통신을 추가할 수 있다.

채팅 - reids & webSocket

채팅을 위해 WebSocket과 redis의 pub/sub을 이용했다. 이벤트(메시지)를 발행하는 Publisher가 존재하며, Publisher는 특정 Topic에 이벤트를 전송한다. 특정 Topic을 구독하는 Subscriber가 존재하며, Publisher에 관계 없이 발행된 이벤트를 받을 수 있다.

챗봇 - Google Dialogflow

Google Dialogflow는 대화식 사용자 인터페이스를 모바일 앱, 웹 등에 설계하고 통합할 수 있는 플랫폼으로 NLP(자연어처리)를 기반으로 구성된 챗봇이다. Google Dialogflow를 사용한 이유는 한글에서의 높은 정확도를 보이고 확장성이 뛰어나다.

알림 - FCM(Firebase Cloud Messaging)

알람 메시지 전송 과정을 위하여 FCM과 앱을 서로 연동한다. FCM에서 토큰을 발급받아 앱에서 백엔드로 전송한다. 백엔드는 토큰을 유저 정보와 연동하여 DB에 보관한다. 그리고 채팅 메시지나 일정 예약 알람 등과 같은 알림이 필요할 때 푸시 알림 내역을 FCM에 요청하게 된다. FCM은 앱에 알람 메시지를 전송하게 된다.

프로젝트 완료 후기

문지연 : 생소한 주제에 익숙하지 않은 기술을 사용하면서 많은 문제를 겪었습니다. 그럴 때마다 적극적으로 의견을 내고 맡은 일을 열심히 수행해준 팀원들과 조언을 아끼지 않고 주신 멘토님께 감사드립니다.

김지윤 : 이번 프로젝트를 통해 새로운 시도를 할 수 있었습니다. 처음에는 혼자 구현해야 하는 부분에 있어서 막막함도 느꼈습니다. 하지만 맡은 일을 열심히 해나가는 팀원들을 보며 해결해나갈 수 있었습니다. 새로운 기술, 팀워크 등 많은 것을 배울 수 있는 뜻깊은 프로젝트였습니다. 프로젝트를 잘 마무리할 수 있게 이끌어준 팀원들과 도움 주신 멘토님 정말 감사합니다.

박지혜 : 중간에 주제를 바꾸기도 하고 정한 주제의 개발 방향을 바꾸는 일도 생겼었습니다. 그럴 때마다 팀원들이 적극적으로 자신의 의견을 공유하고 연구하며 점차 진행을 할 수 있었습니다. 인공지능을 웹과 앱에 접목하는 프로젝트가 처음이라 뜻깊었고 사용자가 접근하기 편하고 사용하기 쉬운 UI와 기능을 고안해내는 것이 중요하다는 것을 깨달았습니다. 이번 프로젝트를 함께한 팀원과 이끌어주신 멘토님께 감사드립니다.

안효원 : 많은 시행착오가 있었지만 앞으로 또 있을 새로운 프로젝트에 좋은 밑거름이 될 거라 믿습니다. 기술뿐

만 아니라 의사소통, 협동심 등 많은 걸 배울 수 있었습니다. 1년 동안 잘 버텨준 팀원들, 늘 신경 써 주신 멘토님께 감사드립니다. 🙏

