1. 프로젝트 주제 소개

MVC(Mobile Voice Cloning)서비스. 딥러닝을 이용한 voice cloning 오픈소스를 활용해 보편적으로 사용가능한 어플리케이션 구축

1. 주제 배경과 필요성, 중요성, 차별성

우선 보이스 클로닝 서비스가 필요한 사용자의 입장을 생각해 보았다. 더 이상 보지 못하는 가족이나, 좋아하던 사람의 목소리를 재현하여 듣고 싶은 사람들이 서비스의 주 사용자일 것이다.

현재 공개되어 있는 소스는 텐서플로우 기반의 소스로 딥러닝 소스이다. 앞서 말한 이 서비스가 필요한 사용자들에겐 높은 사용장벽 요인이다. 소스코드를 활용하는 것뿐만 아니라 연산에 필요한 컴퓨터 자원을 투입하는 것도 일반적인 상황에 불가능하다.

따라서 사용자는 모바일 어플리케이션 단에서 학습에 필요한 목소리, 그리고 출력하려는 문장만 입력한다. 이 후 서버에서 연산 및 음원을 전송하는 방식으로 이 서비스를 구현하려 한다.

1. 기술, 툴, 언어, 프로그램 소개

따라서 이 서비스를 위해 기술적으로 구현할 부분은 클라이언트와 서버별로 다음과 같다.

클라이언트는 모바일 어플리케이션으로 사용자 식별과 음성 전송과 수신을 담당한다. 서버는 사용자 데이터 대조, 모델 학습, 음원 생성 및 발신을 담당한다.

이를 위해 사용할 딥러닝 소스는 <https://github.com/CorentinJ/Real-Time-Voice-Cloning>이며, 클라이언트는 플러터를 이용한 하이브리드 프로그래밍을 통해 안드로이드와 IOS를 지원할 계획이다. 서버는 Ubuntu에서 호스팅하며 코드 실행뿐 아니라 기존 학습된 데이터 모델 관리, 동시 사용자 접근을 위한 컨테이너 관리 등을 담당할 예정이다.

1. 예상되는 강점과 약점, 그리고 극복하기 위한 방안 강조(아마 차별성과 관련된 이야기일듯)

실행시간 모델임에도 사용자가 한번에 집중될 경우 이를 분산할 시스템이 필요하며, 이를 컨테이너화 하여 해결해 보려 한다.

생성된 결과물에 단어간 공백이 있어 이를 편집해 줘야 좀 더 자연스럽게 들린다. 이를 지원할 방안으로 클라이언트에 다운받은 음원을 사용자가 편집할 수 있는 식으로 기능을 지원하려 한다.