

제19회 임베디드SW경진대회 개발요약서

[자유공모]

팀 명	VVS
작품명	VVS(Voice Visualization Service)
작품설명 (3줄 요약)	청각장애인들을 위한 실시간 음성통화 시스템이다. 통화 환경에서 자체 개발한 STT와 TTS 서버를 통해 청각장애인도 실시간으로 음성을 상대방에게 보낼 수 있고, 텍스트 감성분석을 거쳐 본인의 감정을 음성으로도 표현할 수 있다.
소스코드	https://github.com/kusw1006/2021ESWContest_free_1049
시연동영상	https://www.youtube.com/watch?v=hyqGVYRIW48

**작품
기능설명**

0. 개요

현재 배달 관련 시장이 급격히 성장하고 있다. 이제 배달 앱을 이용해 청각 장애인들도 배달음식을 손쉽게 접할 수 있지만, 배달 변경이나 불만 사항 접수 등 추가적 문의에는 아직 음성통화가 주로 사용되어 청각장애인들은 이용하기 어려운 실정이다. 이에 청각장애인들의 통화 접근성 향상을 위해 채팅 기반 음성통화 시스템을 개발하였다.

상대방의 음성이 사용자에게 텍스트로 변환되어 전달되고, 사용자가 작성한 답변 텍스트는 상대방에게 음성으로 변환되어 다시 전달된다. 이 과정들은 모두 실시간으로 통화 환경에서 진행되며, 사용자 측에서 전달할 목소리의 성별 구분이나 사용자가 입력한 문장에서 감정을 분석해 이를 표현한다.

VVS는 청각장애인들이 이처럼 음성통화가 가능하다는 가장 큰 기능을 가진다. 현재 청각장애인들을 위해 시행되고 있는 서비스들이 대부분 영상통화 또는 문자안내 서비스인 데 비하면 VVS는 그들이 실시간으로 음성을 주고받을 수 있는 특색을 가진다. 이 외에도 사용자의 유리한 접근성을 위해 편리한 User Interface 구현과 간소화한 Hardware를 구성했으며, 다양한 딥러닝 모델들을 통해 청각장애인들도 텍스트를 통해 본인의 감정이나 요구사항을 음성으로 드러낼 수 있다.

1. Speech-To-Text (STT)

청각장애인이 VVS를 이용한 통화를 진행할 때 상대방의 음성을 텍스트로 변환해 보여준다.

Kaldi를 활용한 한국어, 배달 및 서비스 제공 특화 음성인식 시스템을 구축하였다. TCP/IP 통신을 활용한 서버 및 클라이언트 네트워크를 구현해 실시간 번역이 가능하고, 사용하는 Domain에 맞게 빠르게 특화하는 코드를 개발하였다. 시뮬레이션 결과, 최대 24시간 이내에 기존 시스템에 수백만 개의 단어 추가 및 적용 가능함을 확인했다. 따라서 배달업계에서 끊임없이 추가되는 신메뉴나 신조어에 대한 빠른 적응이 가능하며, 다른 환경으로 확장하는 것 또한 용이하다.

추가로 가독성 향상을 위해 딥러닝 모델을 통해 띄어쓰기를 수행하여, 실시간으로 상대방에게 전달해 음성합성 기술과 함께 상대방과 청각장애인 간의 양방향 통신이 가능하다.

2. Text-To-Speech (TTS)

청각장애인이 텍스트를 음성으로 변환해 상대방에게 보내주며, 추가적인 기능은 다음과 같다.

사용자가 입력한 텍스트를 통해 감성을 분석하여, 해당 감성에 맞게 음성을 합성하여 발화 상황에 맞는 감정 표현이 가능하다. 또한 감정의 민감도를 설정하여 세밀한 감정표현이 가능하며, 다양한 화자에 대해 학습시켜, 사용자에게 맞게 여성/남성의 목소리로 발화도 가능하다.