

대화 기반 행동 및 음성 분석을 활용한 CBT 보조도구 연구

김수연¹, 신정윤², 양희정², 이유중³, 정은채², 김인권⁴

¹덕성여자대학교 바이오공학전공 학부생

²덕성여자대학교 컴퓨터공학전공(3과) 학부생

⁴아이스큐브랩

kimsy4725@duksung.ac.kr, sju08227@duksung.ac.kr, reveur0yhj@duksung.ac.kr,
iyu20622@gmail.com, 0509alice@naver.com, ingunbi@gmail.com

Research on CBT Support Tools Using Conversation-Based Behavior and Speech Analysis

Su-Youn Kim¹, Jung-Yoon Shin², Hee-Jung Yang², Yu-Jong Lee³, Eun-Che Jeong²,

In-Kwon Kim⁴

¹Dept. of Biotechnology, Duk-sung University

^{2,3}Dept. of Computer Science, Duk-sung University

⁴Icecubelab. Inc

요 약

코로나19 이후 비대면 커뮤니케이션의 확산과 메신저의 발달로 인해 말하기 공포가 증가하여, 진단 및 치료를 위한 임상적 대화 훈련 해결책의 필요성이 높아지고 있다. 인공지능(AI)과의 말하기 시뮬레이션을 통해 사회불안장애 진단 및 솔루션 제공이 가능한 플랫폼 개발을 목표로 한다.

1. 서론

아르바이트 전문 포털 알바천국이 MZ세대(1980년~2000년생) 2735명을 대상으로 조사한 설문에서 응답자 전체의 29.9%는 전화 통화 시 긴장과 불안, 두려움 등을 느끼는 이른바 '콜 포비아(전화 공포증)'를 겪고 있다고 답했다.

이 문제를 해결하기 위해 사회불안장애 환자에게 효과적인 치료법을 탐구한 결과, 인지행동치료(CBT)의 핵심 기법인 노출 치료가 불안 극복에 효과적임을 확인하였다. 특히 두려움을 느끼는 상황에 점진적이고 반복적으로 노출되는 것이 불안 극복의 중요한 요소임을 파악하였다. 이에 따라, 사용자가 심리적 부담을 최소화할 수 있는 가상 환경에서 일상 대화를 연습할 수 있는 대화 훈련 플랫폼을 기획하였다. 더불어, 음성 및 행동 피드백을 제공함으로써 의사소통 패턴을 파악할 수 있도록 설계하였다.

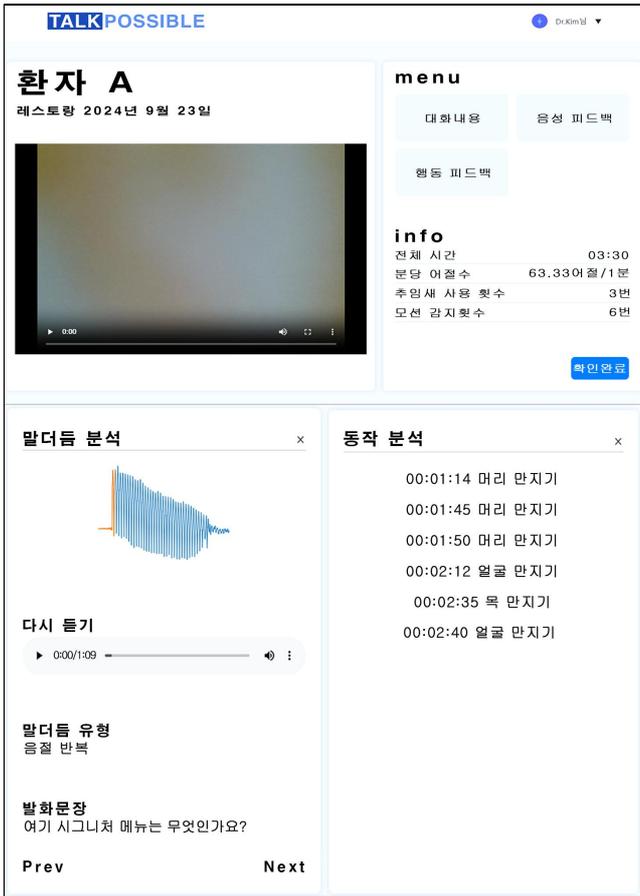
2. CBT 보조도구의 주요 기능

해당 서비스는 본 연구에서 제안하는 CBT 보조도구를 웹 애플리케이션 형태로 구현하였다. 간단한 UI/UX로 설계되었으며, 직관적인 인터페이스를 통해 다양한 사용자층이 쉽게 접근할 수 있도록 개발

되었다. 시스템의 주요 기능은 (1)시뮬레이션 대화, (2)대화 내용 분석, (3)임상 피드백 제공, (4)환자 관리 도구로 구성되어 있다. 시스템의 기능은 진단적 기능과 치료 보조 기능으로 나누어 설명할 수 있다. 진단적 기능은 (1)시뮬레이션 대화와 (2)대화 내용 분석을 포함하고, 치료 보조 기능은 (3)임상 피드백 제공과 (4)환자 관리 도구로 구성되어 있다.

(1) 시뮬레이션 대화는 사용자가 선택한 주제에 따라 생성형 AI와의 대화를 통해 데이터 수집을 진행하며, 이는 치료와 진단을 위한 기초 자료로 활용된다. (2) 대화 내용 분석은 음성 및 행동 데이터를 실시간으로 처리하여, 대화 속도, 말더듬, 행동 패턴을 분석하고, 이를 통해 임상적으로 유의미한 정보를 추출한다. 이 분석은 치료의 중요한 참고 자료로 사용된다. (3) 임상 피드백 제공 기능을 통해 사용자는 대화 종료 후 시스템이 제공하는 치료 경과 피드백을 확인할 수 있다. 피드백은 사용자 행동 및 대화 패턴에 기반한 분석 결과로 제공되며, 이는 치료 과정에서 사용자가 개선해야 할 부분을 제시하는데 활용된다. (4) 환자 관리 도구는 병원 관계자(의사)가 마이페이지에서 환자별로 시뮬레이션 결과를 확인하고 관리할 수 있는 기능이다. 이를 통해 의사는

각 환자의 상태를 모니터링하고, 필요한 조치를 취할 수 있다.



<그림 1> 피드백 화면 - 음성, 행동

3. CBT 보조도구의 핵심 기술

(1) 행동 인식 및 특정 행동 검출

시뮬레이션은 사용자가 버튼을 눌러 생성형 AI와 대화를 시작하면, 웹캠을 통해 참여자의 행동이 감지된다. 이때 MediaPipe를 통해 특정 제스처(얼굴만지기, 머리 만지기, 목 만지기)가 인식되었을 경우 동작 지속 시간을 계산하여 3초 이상인 경우에만 동작 감지 내역을 세션 스토리지에 저장한다. 동작 감지는 최근 5개 프레임의 평균값을 관절 검출 민감도 기준으로 한다. 상황 종료 시 세션 스토리지의 정보를 DB에 저장할 수 있도록 백엔드로 전송한다.

(2) 말더듬 및 억제 행동 분석

사용자의 음성은 실시간 Azure API를 통한 텍스트 변환과 Media Recorder를 통한 음성 녹음 및 Azure Storage로의 파일 업로드가 이루어진다. 변환된 텍스트는 프롬프트 엔지니어링을 거친

ChatGPT-3.5-turbo API에 전달되어 적절한 응답 값을 추출한다. 생성된 응답 값은 Google Cloud API를 통해 음성으로 변환되어 사용자에게 제공된다. 사용자의 음성 녹음이 종료된 후 python 스크립트를 호출하여 Azure Storage의 음성파일에 접근하고 전처리 과정을 수행하게 된다. 어절 단위의 음성파일을 mfcc 처리하고 학습시킨 KNN 모델을 통해 말더듬기 판별을 수행한다. 결과값이 일정 이상이면 말더듬으로 판단하고 학습된 XGBoost 모델을 통해 말더듬 유형(Prolongation, SoundRep)을 분류한다. 이후 해당 어절의 파장 이미지, 오디오, 말더듬 유형, 텍스트를 DB에 저장한다. 추가적으로 말더듬 억제 행동으로 추임새를 사용했는지 판별하기 위해 학습시킨 CNN 모델을 사용한다. 총 추임새 사용 횟수를 카운트해 최종 결과를 DB에 저장한다.

4. 결론

코로나19 이후 비대면 소통의 확산으로 대면 대화에 대한 불안이 증가하고 있으며, 이는 사회불안장애로 이어지는 경향이 있다. 본 연구에서는 사회불안장애를 겪는 사람들이 AI 기반 말하기 시뮬레이션을 통해 음성 및 행동 피드백을 받아 불안을 진단하고 개선하는 방법을 제안한다. GPT API를 사용한 실시간 대화와 행동 검출, 말더듬 여부 및 유형 판단, 말더듬 억제 행동(추임새) 횟수 기록과 관련된 피드백을 제공한다. 이 시스템은 의료, 상담, 교육 및 훈련 분야에서 다양하게 활용될 수 있을 것으로 기대된다.

※ 본 논문은 과학기술정보통신부 대학디지털교육역량강화 사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.

참고문헌

- [1] 이재윤, "통화가 두려운 MZ세대, 콜포비아 60% '생각할 틈이 없어'", 머니투데이, 2022.09.20.
- [2] 윤혜영, 박기우, 조상현, "사회불안 경향자를 위한 가상현실노출치료 프로그램 개발 및 효과검증", 연구방법논총, 4(3), pp. 63-88, 2019.
- [3] 조용래, "비디오 피드백이 사회불안장애 집단의 사회적 자기상의 개선에 미치는 효과", 한국인지행동치료학회, 인지행동치료, 6(1), pp. 23-39, 2006.