

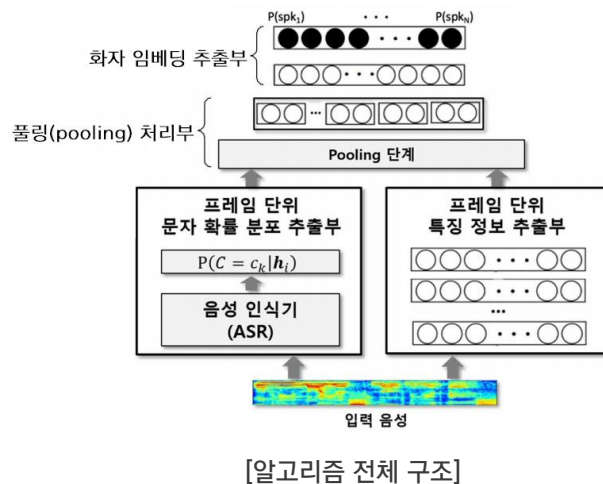
# 다양한 분야에 적용이 가능한 화자 임베딩 추출 방법

서울대학교 공과대학 김남수 교수, 전기정보공학부

## 기술내용

- ASR 모델을 활용하여 음성으로부터 프레임 단위의 문자 단위 확률 분포를 추정하고, 이를 풀링(pooling) 과정에서 활용하여 화자 임베딩 추출 단계에서 화자 정보뿐만 아니라 발화 문장 정보도 함께 추출하도록 하는 화자 인식을 위한 음성인식기 기반 풀링 기법의 화자 임베딩 추출 방법 및 시스템

## 주요도면 및 사진



## 기술개발 배경

- 최근 다양한 딥러닝 기술들이 발전되고, 대용량 데이터 셋이 제공됨에 따라 딥러닝 기반의 화자 임베딩 기법들이 성능 향상을 이뤘으며, 다양한 연구들이 진행되고 있음
- 기존 화자 임베딩 추출 방법은 화자 정보만 고려하는 방식으로, 화자 인식 성능에 제한이 있음
- 기존 기술은 정답 화자에 의존한 지도 학습 방식을 구성하며, 화자 임베딩 추출을 위해 식별 정보만 활용하고 있어 효과적인 화자 인식을 위해서는 중요한 정보는 더 강조하여 처리하는 접근 방식이 요구되고 있음

## 특장점(효과)

- 기존 기술과 달리 프레임 단위의 출력을 집계하는 과정에서 문자 단위의 개별적 처리 과정을 갖기 때문에 추론 단계에서 화자 임베딩간의 유사도 계산 시 특정 발음 간의 특징 비교가 가능해, 이를 통해 화자 인식 시스템에서 화자 정보 및 문장 발화 정보를 동시에 고려하여 비교 분석 가능
- 음성으로부터 프레임 단위의 문자 단위 확률 분포를 추정해, 이를 풀링(pooling) 과정에서 활용하여 화자 임베딩 추출 단계에서 화자 정보 뿐만 아니라 발화 문장 정보도 함께 추출하도록 구성함
- 기존 모델 구조를 적용하지만 목적에 맞도록 적절한 모델 구조로 쉽게 대체가 가능하여 다양한 분야에 활용이 가능

## 기술활용분야

- 화자인식활용 음성 처리 기술

응용분야 및 적용제품	관련 업체
<ul style="list-style-type: none"> <li>• 응용 분야                             <ul style="list-style-type: none"> <li>- 음성인식기(ASR), 화자 인식, 화자 검증, 데이터 및 개인정보 보안, 접근제어, 디지털 포렌식, 음성 데이터 관리, 개인화, 지능형 로봇 제어 등</li> </ul> </li> <li>• 적용제품                             <ul style="list-style-type: none"> <li>- 자동주행, AI 스피커, 음성인식 보안이 필요한 모든 음성처리 기술이 적용된 제품</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Sound처리기술 개발 기업</li> <li>• 음성인증 생체보안 기업</li> </ul>

## 기술개발단계



## 지식재산권 현황

No.	기술명	출원번호	등록번호	국가
1	화자인식을 위한음성인식기반풀링기법의화자임베딩추출방법및시스템,그리고 이를위한기록매체	10-2020-0130575	10-2429656	KR
2	대화자스타일음성합성생성소프트웨어	프로그램	C-2022-024495	-
3	대화자스타일음성합성학습소프트웨어	프로그램	C-2022-024496	-

기술이전상담 및 문의: 서울대학교 산학협력단 신앙일 변리사 ✉ youmei21@snu.ac.kr ☎ 02-880-2026