

# 이어폰의 음향 감지를 이용한 구강 악습관 인지

이재승, 강승우<sup>1</sup>  
한국기술교육대학교

jaeseung.lee@misl.koreatech.ac.kr, swkang@koreatech.ac.kr

## Recognizing Oral Habits Using Acoustic Sensing on Earphones

Jaeseung Lee, Seungwoo Kang<sup>1</sup>  
KOREATECH

### 요 약

턱관절 장애 치료의 기본 방향은 턱관절이 더 이상 손상되지 않도록 보호하고 기능을 안정시키는 것이다. 의료기관에서는 환자의 고통을 완화하고 치료 효과를 높이기 위하여 구강 악습관에 대한 교육을 기반으로 치료를 진행한다. 이러한 환자교육에는 환자 스스로 구강 악습관을 파악하고 장기간 추적 및 관리해야 하는 부담이 따른다. 본 논문에서는 귀에 착용하는 이어폰 형태의 웨어러블 디바이스를 이용하여 일상 중 무의식적으로 나타나는 구강 악습관을 인지하는 시스템을 제안한다. 제안하는 시스템은 이어폰에서 수집되는 음향 신호 데이터를 기반으로 만든 인공 신경망 모델을 통하여 사용자의 구강 악습관을 파악한다.

### I. 서 론

턱관절 장애(Temporomandibular joint disorder)는 측두하악관절, 저작근 및 주변 구조물과 관련된 많은 임상적 문제들을 포괄하는 집합적 용어이다. 국민건강보험공단에 따르면 2015년부터 2019년까지 환자의 수는 지속적으로 증가하여 약 40만 명에 이른다.

이 질환의 대표적인 증상은 턱관절 주변(머리, 목, 귀 등)에 통증이 발생하거나 턱관절 운동 제한으로 입이 잘 벌어지지 않는 등 일상생활에 불편함이 생기는 것이다. 이 질환의 원인은 크게 두 가지, 행동적 요인(구강 악습관)과 정신적 요인으로 나눌 수 있다. 구강 악습관은 크게 편측 저작, 딱딱하거나 질긴 음식 섭취, 이갈이, 이 악물기, 입 크게 벌리기 등을 포함하며 정신적 요인은 심리적 혹은 감정적인 불안 상태를 의미한다. 정신적 요인이 직접적으로 턱관절 장애를 발생시키지 않으나 불안한 심리상태는 구강 악습관을 발생시키며 이러한 행동이 지속적으로 반복될 경우 턱관절 장애를 발생시킬 수 있다[1].

일반적으로 구강 악습관을 파악하고 지속적으로 관리하기 위해서는 환자들이 자신의 행동을 되돌아보고 직접 기록해야 할 것이다. 그러나 무의식중에 발생하는 구강 악습관을 기록하는 것은 번거로운 일이며 신뢰도가 낮다. 이 문제를 해결하기 위하여 유비쿼터스/웨어러블 컴퓨팅 분야에서 구강의 움직임에 관한 다양한 연구가 시행되었으나 한계가 명확하였다. Li et al.[2]은 임베디드 가속도계를 이용하여 사람의 구강 행동(저작운동, 음료 마시기, 말하기, 기침)을 인지하고 분류하였으나 센서를 치아 표면에 부착하는 게 일상생활에 불편을 준다는 한계가 있었다. Bondareva et al.[3]는 무선 이어폰 형태의 웨어러블 디바이스에서 얻은 IMU 센서 데이터를 활용하여 구강 악습관 중 이갈이, 이 악물기를 분류하였는데 분류할 수 있는 행동의 개수가 적고, 구강

악습관 외 다른 행동에 대한 데이터가 주어졌을 때 성능이 하락하는 한계가 있었다.

본 논문에서는 환자의 구강 악습관을 인지하기 위하여 음향 감지 기술을 이용하였다. 이어폰 스피커에서 처프 신호를 내보내면 외이도 표면에 반사되어 이어폰 마이크에 들어오는 음향 신호를 이용하는 것이다. 이렇게 수집되는 음향 신호는 턱관절 주변 근육의 움직임이 외이도 형상에 주는 변화에 민감하므로 수집되는 음향 신호 데이터는 특정한 턱관절 운동에 대응시킬 수 있다.

제안하는 시스템의 목표는 사용자의 동작에 따라 변화하는 음향 신호의 특징을 기반으로 사용자의 구강 악습관을 모니터링 하는 것이다.

### II. 본론

본 논문은 귀에 착용하는 이어폰 형태의 웨어러블 디바이스에서 수집되는 음향신호 데이터를 기반으로 사용자의 구강 악습관을 인지하는 것을 목표로 한다.

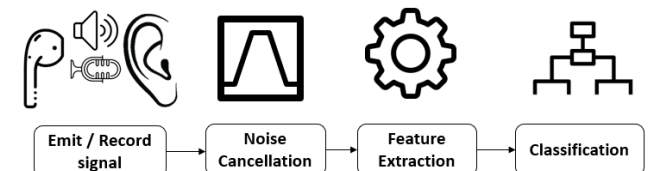


그림 1. 데이터 처리 파이프라인

데이터 수집부터 최종 분류 과정까지 데이터 처리 파이프라인은 그림 1과 같다. 실험에 사용된 이어폰은 왼쪽과 오른쪽 모두 각자의 스피커와 마이크를 이용하여 데이터를 수집할 수 있도록 구현하였다. 첫번째 단계, 능동 소나 시스템은 이어폰 스피커를 통하여 17~18kHz 범위의 처프 신호를 발생시킨다. 이 신호는 20대 이상 성인에게 잘 들리지 않는 주파수 대역의 신호이므로 사용자에게 불편을 주지 않는다. 구강 악습관 혹은

<sup>1</sup> 교신 저자

일상의 움직임에 의하여 외이도에서 반사된 음향 신호는 이어폰 마이크를 통하여 48kHz 로 샘플링되어 수집된다. 이렇게 수집되는 신호에는 잡음이 있는 경우가 흔하다. 이 문제를 해결하기 위하여, 원신호에 Band-pass filter 를 적용하여 17~18kHz 범위 밖에 존재하는 잡음을 제거하는 단계를 거친다. 세번째 단계에서는 잡음이 제거된 신호로부터 FFT, Envelope detection, Channel response 의 특징을 계산하여 총 21 가지의 특징을 추출한다. 이렇게 추출된 특징은 2-layer CNN 과 3-layer CNN 이 결합된 Multi-view classifier 을 통하여 학습된다. 최종적으로, 표 1 과 같이 Unilateral Chewing(UC), Crunchy-chewy Foods(CF), Grind(GD), Clench(CH), Wide mouth(WD), Others(OT) 로 분류된다.

표 1. 수집 데이터 분류

| 이름 | 동작             | 구분      | 시간   |
|----|----------------|---------|------|
| UC | 편측 저작          | 푸딩 젤리   | 20 초 |
|    |                | 부드러운 빵  |      |
| CF | 딱딱하거나 질긴 음식 먹기 | 질긴 건어물  |      |
|    |                | 딱딱한 견과류 |      |
|    |                | 껌       |      |
| GD | 이갈이            | 앞/뒤 방향  |      |
|    |                | 좌/우 방향  |      |
| CH | 이 악물기          |         |      |
| WD | 입 크게 벌리기       |         |      |
| OT | Others         | 말하기     |      |
|    |                | 물 마시기   |      |
|    |                | 걷기      |      |

## 2-2. 데이터 수집

실험 참가자는 총 5 명(남: 4, 여: 1, 나이: 평균: 27.6)으로 모두 턱관절 건강에 이상이 없는 사람으로 구성하였다. 실험은 참가자가 자리에 앉아 이어폰을 착용하는 것으로 시작한다. 실험 중 양쪽 귀에서 발생하는 모든 데이터는 왼쪽 귀 데이터와 오른쪽 귀 데이터로 분리되어 스마트폰에 저장된다. 실험 순서는 가만히 있기, 일반적인 음식(푸딩 젤리, 부드러운 빵) 먹기, 딱딱하거나 질긴 음식(오징어, 브라질 너트, 껌) 먹기, 이갈이(앞/뒤, 좌/우), 이 악물기, 입 크게 벌리기, 말하기, 물 마시기, 걷기 순서로 동작마다 20 초 동안 실시하였다. 본 연구진은 한 사람당 420 초 길이의 데이터를 수집하여 총 5 명의 실험 참가자에게 2100 초 길이의 데이터를 수집하였다.

## 2-3. 평가

본 논문에서는 구강 악습관 인지 시스템의 성능을 검증하기 위하여 계층별 10 겹 교차검증을 이용하였다. 실험 참가자 5 명에게 수집한 전체 데이터 세트는 무작위로 섞인 후 모든 클래스의 비율이 균일한 10 개의 데이터 세트로 나뉜다. 1 개의 데이터 세트를 평가할 때마다 나머지 9 개의 데이터 세트로 학습된 분류 모델이 만들어진다. 이 과정을 10 회 반복하여 각 데이터 세트에 대한 평가지표(Precision, Recall, F1-score)의 평균을 계산하여 구강 악습관 분류 모델에 대한 성능을 평가하였다.

평가 결과, 한쪽 이어폰의 데이터를 이용하였을 때 평가지표는 모두 0.91, 양쪽 이어폰의 데이터를 모두 이용하였을 때 평가지표는 모두 0.96 을 달성하였다.

결론적으로, 한쪽 이어폰을 사용하였을 때와 양쪽 이어폰을 이용하였을 때 모두 높은 성능을 보였으며 한쪽 이어폰을 사용하였을 때보다 양쪽 이어폰을 사용하였을 때 성능이 향상된 것을 확인하였다. 양쪽 이어폰을 이용하였을 때 구강 악습관에 대한 분류결과는 그림 2 와 같다.

표 2. 성능 평가

| 평가지표   | Precision | Recall | F1-score |
|--------|-----------|--------|----------|
| 한쪽 이어폰 | 0.91      | 0.91   | 0.91     |
| 양쪽 이어폰 | 0.96      | 0.96   | 0.96     |

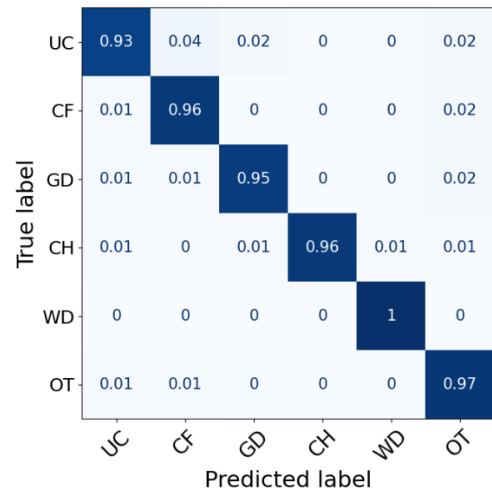


그림 2. 구강 악습관에 대한 혼동행렬

## III. 결론

본 논문에서는 이어폰 형태의 웨어러블 디바이스를 이용하여 일상 중 나타나는 구강 악습관을 인지하는 시스템을 제안하였다. 제안하는 시스템은 사용자의 동작에 따라 변화하는 음향 신호의 특징을 기반으로 사용자의 행동을 6 가지 클래스로 분류한다. 총 5 명의 실험 참가자를 모집하여 실험을 진행하였고 계층별 10 겹 교차검증 방식으로 평가한 결과 모든 평가지표에 대하여 충분한 성능을 보임을 확인하였다. 향후 연구에서는 모바일 애플리케이션을 개발하여 구강 악습관 개선이 필요한 환자에게 적용해보면 제안하는 시스템이 실제 환경에서 얼마나 효과적인지 검증할 수 있을 것이다.

## ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업의 연구결과로 수행되었음 (IITP-2020-0-01778)

## 참 고 문 헌

- [1] E. G. Seo, "Prevalence of Temporomandibular disorders and its factors among office workers, service workers and teachers", Master's degree, Korea Univ, 2010.
- [2] Li et al. Sensor-embedded teeth for oral activity recognition. In *ISWC'13*
- [3] Bondareva et al. Earables for Detection of Bruxism: a Feasibility Study. In *EarComp '21*