|  |
| --- |
| **1. 주제**  청각 장애인을 위한 AI 배리어 프리 자막 서비스  **분반, 팀, 학번, 이름**  나반, 1팀, 김진우 |

|  |  |
| --- | --- |
| **2. 요약**  청각 장애인을 위해 영상 매체에서 폐쇄형 자막을 자동으로 제공하고자 합니다. 언어적 음성 뿐만 아니라 비언어적인 음성에 대해서도 자막을 제공하는 것입니다. 음성데이터를 학습된 인공지능 모델을 통해 분류한후 텍스트로 변환합니다. 이렇게 자막을 제공함으로써 장애인들이 느끼는 문화적 소외감을 낮춤과 동시에 비장애인들 또한 영상을 놓치는 부분없이 이용할 수 있게 될 것입니다. | **3. 대표 그림**    그림 1. AI 폐쇄형 자막 자동 제공 서비스 |

|  |
| --- |
| **4. 서론**  기술의 발전으로 IT기기 뿐만 아니라 이런 기기에서 사용되는 서비스, 영상 매체 등도 함께 발달하게 되었습니다. 하지만 청각장애인들은 이러한 영상 매체를 제대로 이용하고, 즐기지 못하는 경우가 많습니다. 방송통신위원회에서 발표한 소외계층 미디어 포용 계획을 보면 한국수어방송 의무편성비율을 7%로 늘린다고 하였습니다. 이를 보면 청각장애인들의 환경 개선이 시급한 상황임을 알 수 있습니다.  영화, 드라마, 다큐멘터리 등 수많은 영상 자료들이 여러 ott등을 통해 이용가능하지만 청각장애인들은 이것을 이용할 때 큰 제약이 따르게 됩니다. 자막이 없는 경우도 존재할 것이고, 자막이 있더라도 사람사이의 대화만 자막으로 표시하고 개가 짖는 소리, 바람소리 등의 비언어적인 요소는 자막으로 표시되지 않는 경우가 많기 때문입니다.  이러한 문제는 영상에 담겨있는 모든 소리를 자막으로 표현해주는 폐쇄형 자막을 제공해줌으로써 해결할 수 있습니다. 이러한 폐쇄형 자막은 장애인들에게 매우 큰 도움이 될 뿐만 아니라 영상에 담긴 요소를 하나도 놓치지 않고 즐기기 원하는 비장애인들에게도 도움이 됩니다. 문제는 폐쇄형 자막을 제공하는 것은 큰 시간과 노력을 필요로 한다는 것인데, 저는 이 폐쇄형 자막을 AI를 통해 자동으로 제공해주는 서비스를 제안하고자 합니다. |

|  |
| --- |
| **5. 본론**    우선 학습 데이터를 수집합니다. 이후 음성 데이터를 학습하기 위해 아날로그 데이터인 음성 데이터를 디지털 신호로 변환해줍니다. 파이썬에서 제공하고 있는 librosa 라이브러리를 이용할 것이며, 음성 데이터를 특징 벡터화 해주는 알고리즘인 MFCC로 특징을 추출합니다. 이 결과를 시각화하여 특징을 살펴봅니다. 이후 CNN 모델을 학습시킵니다. Conv2D, MaxPooling2D, Flatten, Dense 등 모델의 층을 쌓고 학습시켜 줍니다. 이후 샘플 데이터를 넣어서 정확히 판별이 되는지를 확인해 줍니다. 이 모델을 통해 비언어적 소리를 구분해 줍니다. 이렇게 구분한 음성, 비언어 음성들을 Google Cloud Speech API를 이용하여 텍스트로 변환해 주고, 이를 사용자에게 제공합니다. |

|  |
| --- |
| **6. 결론**  AI 배리어 프리 서비스는 기존에 많은 시간과 노력이 필요했던 폐쇄형 자막을 자동으로 제공해줌으로써 여러 자원을 절약할 수 있음과 동시에 장애인들이 영상 매체를 즐기고, 문화적 소외감을 줄이는데에 큰 도움을 줄 것입니다.  하지만 여기서 더 나아가서 비언어적인 음성을 자막으로 나타낼 뿐만 아니라 추임새 등 또한 자막으로 나타낼 수 있게 AI를 학습시킨다면 더 풍부한 정보를 제공할 수 있을 것입니다. |

**7. 출처**

오디오 전처리 방법에 따른 콘벌루션 신경망의 환경음 분류 성능 비교- 오원근(p1 – p4)