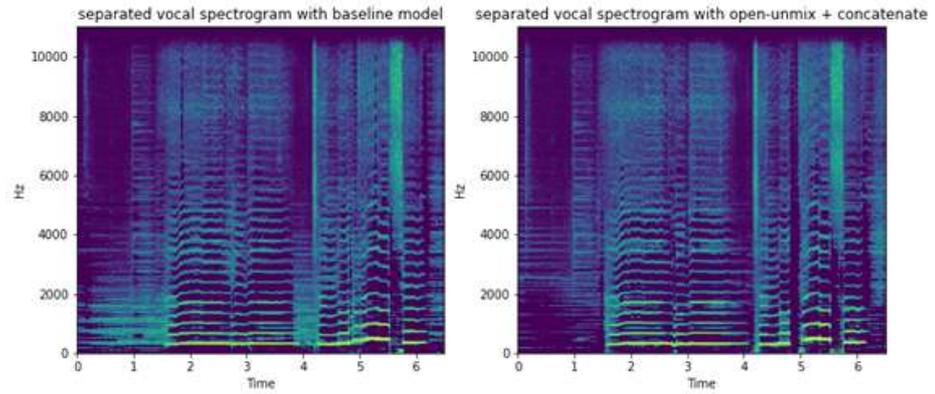


## ■ 연구과제 요약문

<b>과제명(기간)</b>	Factorized low-level acoustic feature 기반 풍부한 표현이 가능한 한국어 TTS 시스템 개발 (2019.04.01.~ 2020.03.31.)
<b>연구책임자</b>	이 교 구 (kglee@snu.ac.kr)
<b>개요</b>	<ul style="list-style-type: none"> <li>- 가사 정보를 활용한 가창 음원 분리(Lyrics-informed Singing Voice Separation)</li> <li>- 본 연구에서는 가사 정보를 활용해 딥러닝 기반의 가창 음원 분리 네트워크의 성능을 향상시키는 기법을 제안한다. 반주와 가창이 섞인 혼합 음원에서 가창을 분리해 내는 가창 음원 분리 문제는 노래방 시스템 등의 상업적 활용도와 자동 악보 채보 등의 다양한 음악정보검색 분야 연구의 전처리 과정에 사용될 수 있기 때문에 오디오 신호 처리 분야에서 가장 활발하게 연구되고 있는 분야 중 하나이다.</li> <li>- 본 연구는 정렬되지 않은 가사 정보를 활용한 음원 분리 네트워크를 제안하기 위한 선행 연구로서 정렬된 가사 정보를 활용한 음원 분리 네트워크를 제안한다. 음원 분리 네트워크에 가사 인코딩 정보를 입력으로 활용하였을 때, 가사 인코딩 정보를 사용하지 않았을 때보다 음원 분리 성능이 향상됨을 확인하였다. 또한, 본 연구에서 제안한 방법을 통해 학습된 가사 인코더 네트워크가 실제 정렬된 가사의 타이밍 정보와 음소 정보를 음원 분리에 잘 활용하는 것을 실험을 통해 확인하였다.</li> </ul> <div style="text-align: center; margin: 20px 0;"> </div> <p style="text-align: center;">&lt;그림 1&gt; 가사정보를 활용한 가창음원 분리 개요</p>
<b>연구개발 결과</b>	<div style="text-align: center;"> </div>



<그림 2> 가창음원 분리 결과 예시. 좌측 위: 혼합음원, 우측 위: 원래의 가창음원, 좌측 아래: 기존 모델의 가창음원 분리 결과, 우측 아래: 제안 모델의 가창음원 분리 결과

- 본 연구에서 제안한 가사 정보를 입력으로 활용했을 때 그렇지 않은 기본 모델보다 가창음원 분리의 성능이 높아짐(<그림 2>).
- 가사는 악보 등과 비교했을 때 비교적 손쉽게 구할 수 있는 정보이므로 활용 가능성이 높음
- 언어와 관계 없이 음향학적 모델을 사용하였기 때문에 다국어 모델로 확장이 가능함
- 본 연구에서는 가사와 음악오디오가 정렬되어 있다는 것을 가정하였지만, 추후 비정렬된 가사를 사용하여 가창음원 분리에 적용할 계획임

**활용분야 및 기대효과**

- 본 연구에서는 정렬된 가사를 이용해 음원 분리 네트워크의 성능 향상이 가능함을 입증하였다.
- 보컬 활동 정보와 가사의 음소 정보를 활용한 가창음원 분리 네트워크의 성능 향상을 확인하였다.
- 가사 인코더를 기존 네트워크와 로컬 컨디셔닝 혹은 Concatenate 방법을 활용해 결합함으로써, 기존 네트워크의 성능 향상에 기여할 수 있을 것으로 기대된다.
- 노래방 등 가창을 제외한 반주 부분을 합성하여 원곡과의 이질감이 심한 형태의 서비스가 아니라, 임의의 곡에 대하여 가창 음원만 제거함으로써 원곡과 똑같은 형태의 새로운 서비스로 활용이 가능하다.