

■ 연구과제 요약문2

<p>과제명(기간)</p>	<p>E2E형 음성 인식/합성 및 DRL 기반 음악 생성 기술 연구 (2017-04-01~2019-03-31)</p>
<p>연구책임자</p>	<p>박 종 헌</p>
<p>개요</p>	<ul style="list-style-type: none"> - 본 연구과제는 다음과 같은 3가지 목표를 가지고 진행되었음 - Phone-based language model에 기반하여, 다국어 처리가 가능한 음성 인식/합성 기술 개발 - Speaker recognition 분야의 여러 문제에 활용할 수 있도록 deep learning 기반의 representation learning 방법으로 speaker의 음성적 특징을 얻는 기법을 개발 - Deep learning 기반의 generative model을 이용하여 자동으로 음악을 생성하는 모형을 개발
<p>연구개발 결과</p>	<ul style="list-style-type: none"> - 음성 인식/합성에 많이 활용되는 sequence-to-sequence 모형 기반의 새로운 decoding 기법 및 speaker diarization 방법을 제안하고, neural machine translation 및 image 도메인에서 활용되어온 unsupervised learning 방식을 음성 인식/합성 분야에 적용할 수 있도록 새로운 학습 알고리즘을 개발하였음 - Text를 이해할 수 있는 AI를 구현하기 위하여, sentence completion 문제와 후속 sentence 등장 확률 등을 예측할 수 있는 다양한 deep neural network 학습 기법을 개발하였음 - Hierarchical VAE, SeqGAN, deep reinforcement learning 등을 이용하여, chord sequence가 주어졌을 때, 이에 어울리는 음악을 자동으로 생성하는 모형을 개발하였으며, 아울러 음원이 주어졌을 때, 음악의 chord 진행을 자동으로 파악하는 transformer 기반의 neural network 모형을 개발하였음
<p>활용분야 및 기대효과</p>	<ul style="list-style-type: none"> - 본 연구에서 개발된 모형 및 기법들은 음성인식, 음성합성, 자연어 이해, 자동 작곡 등 다양한 응용 서비스에서 활용될 수 있을 것으로 기대되며, 구축된 연구 프레임워크를 이용하여, 향후 다양한 deep learning 모형들의 개발, 개선, 평가 및 검증을 수행할 수 있을 것으로 기대됨