

확장강의계획서

(2024 년도 2 학기)

과목명	오디오인식 및 합성변환	과목번호	CESG109/CSE5109 AIEG109/AIE5109
구분(학점)	3학점	수강대상	석사과정/학부4학년
수업시간	월/수 12:00~13:15PM	강의실	

담당교수 (사진)	성명: 김지환	과목홈페이지: eclass.sogang.ac.kr
	E-mail: kimjihwan@sogang.ac.kr	연락처: 이메일
	장소: AS 713 면담시간: 이메일로 사전 예약	

I . 교과목 개요(Course Overview)

1. 수업개요							
본 과목은 교과서 대신 논문을 교재로 사용할 예정임 (논문을 미리 읽었는지를 확인하기 위해, 논문의 핵심단어(수치)에 대해 단답형으로 묻는 퀴즈를 수시 시행할 예정임)							
본과목의 목표는 대화형 사용자 인터페이스 시스템을 구성하고 있는 주요 요소인 음성합성, 오디오 이벤트 인식등의 기본 개념을 이해하는 것이다. 오디오 저장 방법 및 스펙트럼 생성 방법을 공부하고, 딥러닝 이전의 음성합성 방법 (concatenative 방식등)을 살펴본다. 딥러닝 기반의 방법론을 이해하기 위해 RNN, Seq2Seq, Transformer 중심으로 살펴보고, Wavenet, Tacotron, Deepvoice를 중심으로 딥러닝 기반 음성합성 방법에 대해서 공부한다.							
2. 선수학습내용							
없음							
3. 수업방법 (%)							
강의	토의/토론	실험/실습	현장학습	개별/팀 별 발표	기타		
100 %	%	%	%	%	%		
4. 평가방법 (%)							
중간고사	기말고사	퀴즈	발표	프로젝트	과제물	참여도	기타
45 %	45 %	10 %	%	%	%	%	%

II . 교과목표(Course Objectives)

본과목의 목표는 대화형 사용자 인터페이스 시스템을 구성하고 있는 주요 요소인 음성합성, 오디오인식등의 기본 개념을 이해하는 것이다. 오디오 저장 방법 및 스펙트럼 생성 방법을 공부하고, 딥러닝 이전의 음성합성 방법 (concatenative 방식등)을 살펴본다. 딥러닝 기반의 방법론을 이해하기 위해 RNN, Seq2Seq, Transformer 중심으로 살펴보고, Wavenet, Tacotron, Deepvoice를 중심으로 딥러닝 기반 음성합성 방법에 대해서 공부한다.

III. 수업운영방식(Course Format)

(* I -3의 수업방법의 구체적 설명)

IV. 학습 및 평가활동(Course Requirements and Grading Criteria)

-본 과목은 중간고사 (45%), 기말고사 (45%), 퀴즈 (10%)로 평가한다.

1 중간고사

- 강의 시간에 배운 핵심 알고리즘의 이해도를 측정한다

2 기말고사

- 강의 시간에 배운 핵심 알고리즘의 이해도를 측정한다

3. 퀴즈

- 논문을 수업전에 미리 읽었는지를 확인하기 위함

V. 수업규정(Course Policies)

출석은 성적에 반영하지 않음. 비대면 강의에 대한 학교의 FA 방침에 따라 FA는 적용함

VI. 교재 및 참고문헌(Materials and References)

[WaveNet] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. W. Senior, and K. Kavukcuoglu, "WaveNet: A generative model for raw

audio", arXiv:1609.03499, 2016

[**Tacotron1**] Y. Wang, R. Skerry-Ryan, D. Stanton, Y. Wu, R. J. Weiss, N. Jaitly, Z. Yang, Y. Xiao, Z. Chen, S. Bengio, Q. Le, Y. Agiomyriannakis, R. Clark, and R. A. Saurous, "Tacotron: Towards end-to-end speech synthesis", *Proc. Interspeech*, pp. 4006–4010, 2017

[**Tacotron2**] Jonathan Shen, Ruoming Pang, Ron J Weiss, Mike Schuster, Navdeep Jaitly, Zongheng Yang, Zhifeng Chen, Yu Zhang, Yuxuan Wang, Rj Skerry-Ryan, et al, "Natural tts synthesis by conditioning wavenet on mel spectrogram predictions", *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4779–4783, 2018

[**FastSpeech**] Yi Ren, Yangjun Ruan, Xu Tan, Tao Qin, Sheng Zhao, Zhou Zhao, and Tie-Yan Liu, "FastSpeech: Fast, robust and controllable text to speech", *Proc. Conference on Neural Information Processing Systems*, pp. 3165–3174, 2019

[**Transformer TTS**] Naihan Li, Shujie Liu, Yanqing Liu, Sheng Zhao, Ming Liu, and Ming Zhou. "Close to human quality tts with transformer", arXiv:1809.08895, 2018.

[**Deep Voice**] S. Ö. Arık et al., "Deep voice: Real-time neural text-to-speech", *Proc. Int. Conf. Mach. Learn.*, pp. 195–204, 2017

VII. 주차별 강의계획(Course Schedule)

(* 추후 변경될 수 있음)

1 주차	학습목표	오리엔테이션 및 음성 저장 방법
	주요학습내용	OT & 음성 저장 방법
	수업방법	강의
	수업자료	강의노트
	과제	
2 주차	학습목표	Spectrogram, Pratt
	주요학습내용	Spectrogram, Pratt RNN-T, Seq2Seq, Transformer
	수업방법	강의
	수업자료	강의노트
	과제	
3 주차	학습목표	Concatenative TTS
	주요학습내용	Concatenative TTS
	수업방법	강의
	수업자료	강의노트

	과제	
4 주차	학습목표	딥러닝 (Perceptron, MLP, Back Propagation)
	주요학습내용	딥러닝 (Perceptron, MLP, Back Propagation)
	수업방법	강의
	수업자료	강의노트
	과제	
5 주차	학습목표	딥러닝 (RNN, Seq2Seq, Attention, Transformer)
	주요학습내용	딥러닝 (RNN, Seq2Seq, Attention, Transformer)
	수업방법	강의
	수업자료	강의노트
	과제	
6 주차	학습목표	GriffinLim
	주요학습내용	GriffinLim
	수업방법	강의
	수업자료	논문, 강의노트
	과제	
7 주차	학습목표	Wavenet
	주요학습내용	Wavenet
	수업방법	강의
	수업자료	논문, 강의노트
	과제	
8 주차	학습목표	
	주요학습내용	중간고사
	수업방법	
	수업자료	

	과제	
9 주차	학습목표	오디오 이벤트 인식
	주요학습내용	오디오 이벤트 인식
	수업방법	강의
	수업자료	논문, 강의노트
	과제	
10 주차	학습목표	DCASE 챌린지
	주요학습내용	DCASE 챌린지
	수업방법	강의
	수업자료	논문, 강의노트
	과제	
11 주차	학습목표	DeepVoice1
	주요학습내용	DeepVoice1
	수업방법	강의
	수업자료	논문, 강의노트
	과제	
12 주차	학습목표	DeepVoice2
	주요학습내용	DeepVoice2
	수업방법	강의
	수업자료	논문, 강의노트
	과제	
13 주차	학습목표	Tacotron
	주요학습내용	Tacotron
	수업방법	강의

	수업자료	논문, 강의노트
	과제	
14 주차	학습목표	Tacotron2
	주요학습내용	Tacotron2
	수업방법	강의
	수업자료	논문, 강의노트
	과제	
15 주차	학습목표	FastSpeech
	주요학습내용	FastSpeech
	수업방법	강의
	수업자료	논문, 강의노트
	과제	
16 주차	학습목표	
	주요학습내용	기말고사
	수업방법	
	수업자료	
	과제	

VIII. 참고사항(Special Accommodations)