

명지대학교 ICT융합대학 정보통신공학과
강환일 교수 연구실 소개

정보 및 전산지능 연구실



주관기관: 명지대
협력기업: (주) 이드웨어

정보 및 전산지능 연구실 (지도교수: 강환일)

1. 정보 및 전산지능 연구실 현재 주요 연구내용

- 음성인식 인식을 개선
- 저전력을 이용한 최소형 컨트롤러 장착연구
- 딥러닝 내용 심화 연구

2. 음성인식 상세 연구내용

- 세부 연구내용 연결 관계 다이어그램

3. 최적화 기법 비교 연구

- Adam, Adagrad, Adadelata 및 RMSProp 등 최적화 기법 비교

4. Adam optimizer의 learning rate 및 성능 비교

- Adam optimizer의 learning rate simulation

5. 음성 신호 속성인 멜주파수 캡스트럼 계수에 의한 성능 비교

- 속성크기에 의한 최적화 기법 비교

6. 링크 정보

- <https://blog.naver.com/hwanilkang>
- <https://github.com/hwankang/deep-learning-lab>
- hwanilkang@naver.com hwan@mju.ac.kr

<https://github.com/hwankang/deep-learning-lab>



명지대학교 ICT융합대학 정보통신공학과
강환일 교수 연구실 소개

정보 및 전산지능 연구실



주관기관: 명지대
협력기업: (주) 이드웨어

https://blog.naver.com/hwanilkang

https://blog.naver.com/hwanilkang

새로 고침 (Ctrl+R)

YouTube 지도 (13) 윈드라이브 계... Softmax Regression... 인프런 - 한시적 무... YouTube Music [딥러닝] 뉴럴 네트... 2020-1학기 대학원...

(hwanilkang)

hwanilkang님의 블로그입니다. [EDIT](#)

[프로필](#)

[글쓰기](#) [관리](#) [통계](#)

카테고리

[전체보기](#) (13) [EDIT](#)

[정보 및 전산지능연구실](#)

[포토로그](#)

[비디오로그](#)

태그 [EDIT](#) [최근 | 인기](#)

사용된 태그가 없습니다.

[모두보기](#)

지도위에 글 0

위치가 첨부된 글이 없습니다.
맛집, 여행, 사진의 추억을 지도위에 남겨보세요.

[글쓰기](#)

전라남도

정보 및 전산지능연구실

음성인식

[이수](#) · 4시간 전

[URL 복사](#) [통계](#)

음성인식 연구내용

- 핵심어 인식 (Keyword spotting)에 관한 연구임
- 알고리즘은 depth separable convolutional neural network (DS-CNN)* 를 이용함
- 데이터 수집, 정련, 훈련을 진행하고 모델을 양자화하여 Raspberry Pi에 장착함

레이블 Raw Data

전처리

음성파일

훈련, 테스트 유효데이터셋

테스트 데이터셋

특징벡터

DNN, DS-CNN 등 학습

모델선택 변수최적화 정확도개선

Model

양자화 Model

Raspberry Pi에 장착

New 데이터

레이블

예측 및 동작

DNN: deep neural network
DS-CNN: depth separable convolutional neural network

음성인식에 관한 다이어그램

사용중인 아이템 보기

TODAY TOTAL 1 447

[RSS 2.0](#) [RSS 1.0](#) [ATOM 0.3](#)

powered by **NAVER blog**

https://blog.naver.com/hwanilkang

https://blog.naver.com/hwanilkang/222193708258

Gmail YouTube 지도 (13) 원드라이브 계... Softmax Regression... 인프런 - 한시적 무... YouTube Music [딥러닝] 뉴럴 네트... 2020-1학기 대학원..

프로로그 | 블로그 | 포토로그 | 비디로그

지도 | 서재 | 메모 | 태그 | 안부

정보 및 전산지능연구실 13개의 글

이수 (hwanilkang)
hwanilkang님의 블로그입니다. [프로필](#)

글쓰기 [후편라-통계](#)

카테고리

- 전체 보기 (13) [EDIT](#)
- 정보 및 전산지능연구실
- 포토로그
- 비디로그

태그 [EDIT](#) 최근 | 인기

사용된 태그가 없습니다. [모두보기](#)

지도위에 글 0 [+](#)

정보가 부족합니다. 위치가 정확하지 않습니다. 맛집, 여행, 사진의 추억을 지도위에 남겨보세요. [글쓰기](#)

정보 및 전산지능연구실

정보 및 전산지능 연구실 (Deep learning 기반 음성인식 연구)

이수 5시간 전 URL 복사 [로그인](#)

1. 음성인식율 개선 연구 (MFCC 계수 및 최적화 optimizer)
2. Mobile net와 Resnet의 응용연구
3. Learning rate의 최적화
4. 저전력 최소형 컨트롤러 기반 음성인식장치 구현

최근댓글

이웃 블로거

활동정보

블로그 이웃 0 명
글 보내기 0 회
글 스크랩 0 회

사용중인 아이템 보기

TODAY TOTAL 1 447

RSS 2.0 RSS 1.0 ATOM 0.3

powered by NAVER blog

1. 정보 및 전산지능연구실 현재 주요 연구 내용

◆딥러닝 (deep learning) 기반 음식인식 인식율 개선

- 암호 구호 인식 시스템 → 스마트 팩토리에 적용
- 핵심어 인식 시스템 채용
- 협력기업: (주)이드웨어

◆저전력을 이용한 최소형 컨트롤러 장착연구

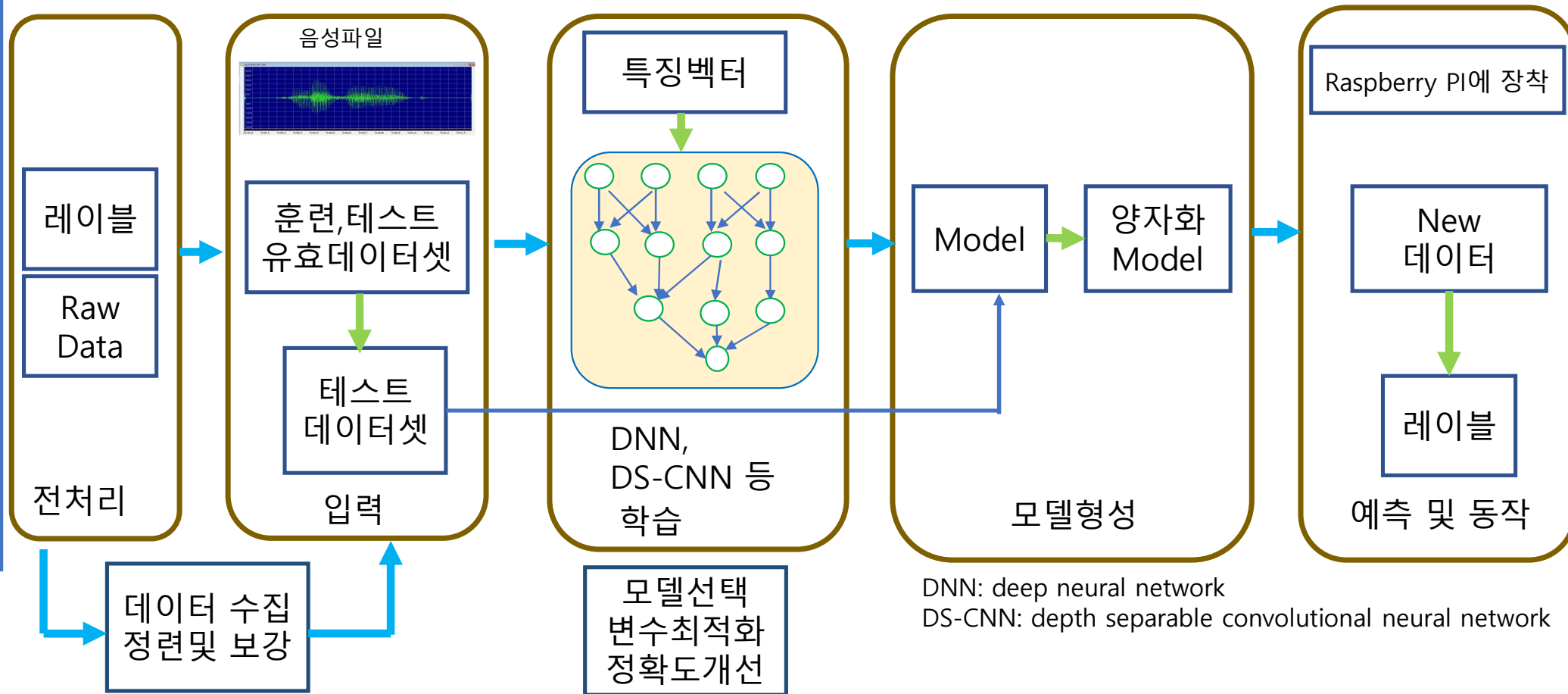
- 암호 구호를 쉽게 변경 가능 (무선으로 암호 구호 변경 가능)

◆딥러닝(deep learning) 내용연구

- Mobile net, Resnet 의 음식인식 응용
- Adam, Adagrad, Adadelata 및 RMSProp 등 최적화 기법 비교
- GPU 여부에 따른 성능 비교

2. 음성인식 상세 연구내용

- 핵심어 인식 (Keyword spotting)에 관한 연구임
- 알고리즘은 depth separable convolutional neural network (DS-CNN)* 를 이용함
- 데이터 수집, 정련, 훈련를 진행하고 모델을 양자화하여 Raspberry PI에 장착함



3. 최적화 기법 비교 연구

- 4가지 최적화 기법에 따른 핵심어 인식율 비교
- 성능은 Adagrad > Adam > RMSProp > Adadelta 의 순서로 좋다.
- 학습율에 따라 성능은 다르다.

최적화기법 (optimizer)	학습율(learning rate)	Accuracy (%) (train)	Accuracy (%) (validation, best validation)	Accuracy (%) (test)
Adam	0.0005,0.0001,0.00002	96.00	92.94(92.31)	93.50
RMSProp	0.0005,0.0001,0.00002	95.50	93.00(93.32)	93.31
Adadelta	0.5, 0.1, 0.02	94.00	92.24(92.31)	92.31
Adagrad	0.05, 0.01, 0.002	97.00	93.99(94.29)	94.19

Steps (10000, 10000, 10000)

Batch size 200

구글 speech data 이용 12개의 단어 분류

4. Adam optimizer의 learning rate 및 성능 비교

최적화기법	학습율	Accuracy (%) (훈련)	Accuracy (%) (유효, best 유효)	Accuracy (%) (test)
Adam	0.0005,0.0001,0.00002	96.00	92.31(92.94)	93.50
Adam	0.05,0.01,0.002	97.50	94.83(95.19)	94.95
Adam	0.05,0.01,0.001	98.00	94.98(95.28)	94.89
Adam	0.06,0.02,0.004	96.00	95.07(95.32)	94.85
Adam	0.05,0.01,0.004	96.00	94.89(95.03)	94.76
Adam	0.05,0.02,0.001	97.50	94.62(95.03)	94.70
Adam	0.05,0.01,0.004	97.00	94.58(94.98)	94.72
Adam	0.06,0.02,0.002	96.50	95.01(95.10)	94.68
Adam	0.05,0.01,0.004	97.00	94.67(95.07)	94.58
Adam	0.05,0.01,0.003	97.50	94.49(94.76)	94.52
Adam	0.06,0.02,0.005	97.50	94.78 (94.94)	94.23

Steps (10000, 10000, 10000)

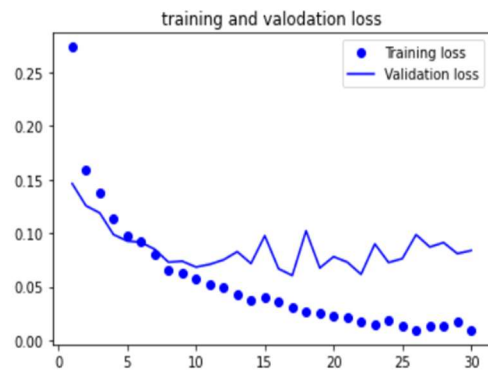
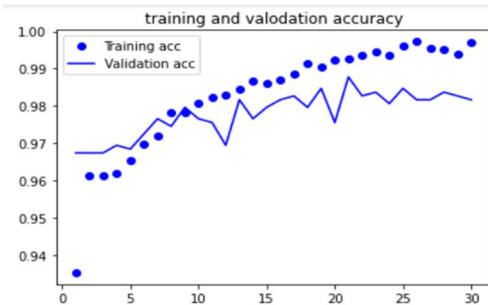
Batch size 200

Validation data:4445, test data::4890

5. 음성신호 속성인 멜주파수 캡스트럼 계수 기반 성능비교

- 구글 speech dataset에서 stop 핵심어를 인식하는 프로그램에서 최적화 방법의 성능 비교

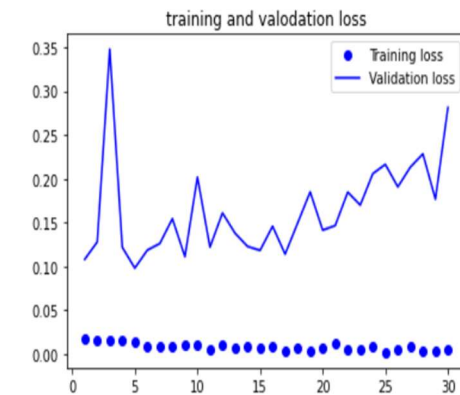
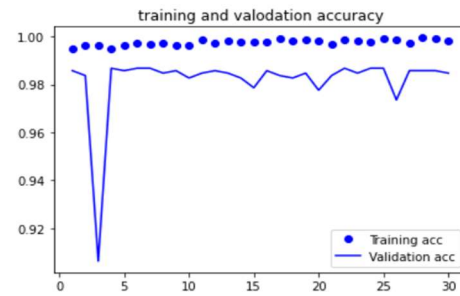
1. Adam



learning_rate=0.0001

Test acc. 0.9852

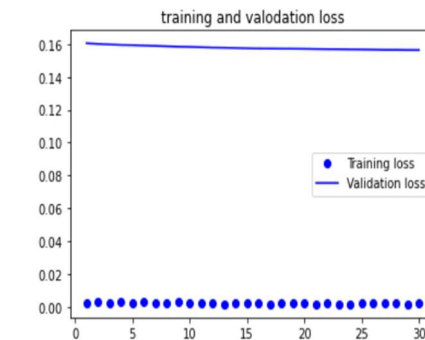
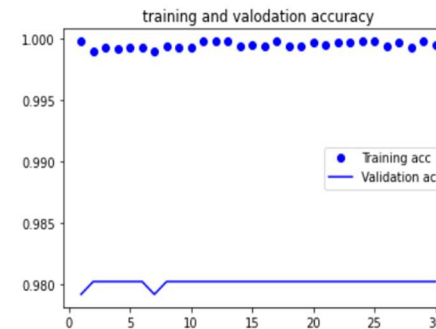
2. RMSProp



learning_rate=0.0001

0.9884

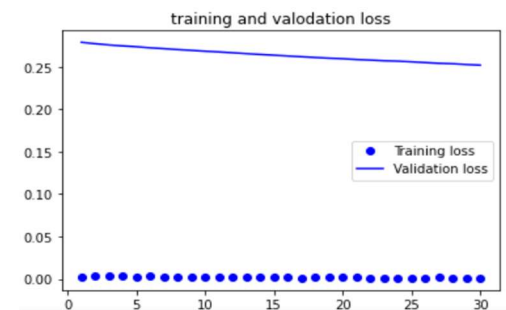
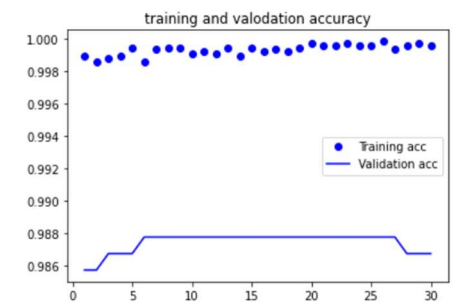
3. Adadelta



learning_rate=0.1

0.9745

4. Adagrad



learning_rate=0.001

0.9894

5. 음성신호 속성인 멜주파수 캡스트럼 계수에 의한 성능비교

Adagrad>Adam>Adadelata>RMSProp

	Adam	RMSProp	Adadelata	Adagrad	
정확도					
	0.0001	0.0001	0.1	0.001	학습율
훈련	0.9969	0.9983	0.9996	1.0000	
유효	0.9817	0.9847	0.9868	0.9812	
테스트	0.9852	0.9844	0.9745	0.9894	

```
model.compile(loss='binary_crossentropy',optimizer='Adam',learning_rate=0.0001,metrics=['acc'])
```

30회 수행 구글 데이터에서 핵심어 stop 인식 시스템
MFCC 계수 (10,25)

6. 링크 정보

◆ <https://github.com/hwankang/deep-learning-lab>

- 음성인식 자료 피피티
- 인공지능 코드

◆ <https://blog.naver.com/hwanilkang>

- 음식인식 다이어아 그램 구성
- 일상의 일

◆ hwanilkang@naver.com hwan@mju.ac.kr

- Email