



제 135회 석사 학위 졸업 논문 발표

# 모바일 적응형 딥러닝 구조 탐색을 이용한 인간 행동 인식

CAU Artificial Intelligence LAB

임원선

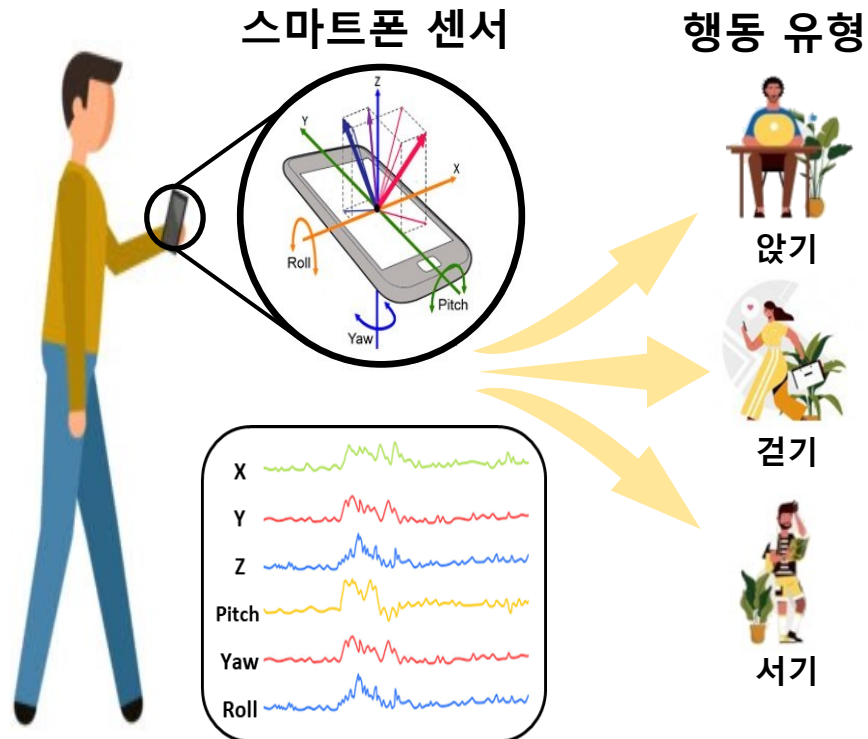
[costor107@gmail.com](mailto:costor107@gmail.com)

12<sup>th</sup> May, 2021

# 스마트폰 기반 인간 행동 인식(Human Activity Recognition, HAR)

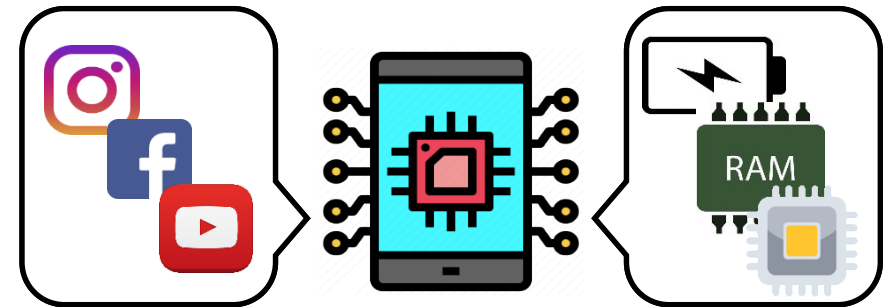
- ✓ 스마트폰 센서의 시계열 데이터를 활용한 행동 예측 연구
- ✓ 모바일 환경은 어플리케이션 동작하여 **계산 자원 활용 제한적**
- ✓ **다양한 하드웨어 사양**에 따라 계산 자원 다름

스마트폰 기반 HAR [Radu et al., 2016,  
Nweke et al., 2018]

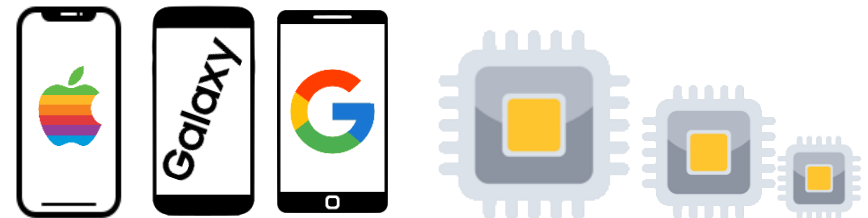


모바일 환경

제한된 계산 자원



다양한 모바일 사양

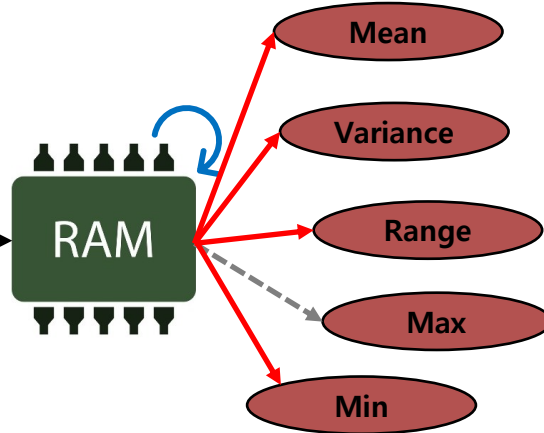
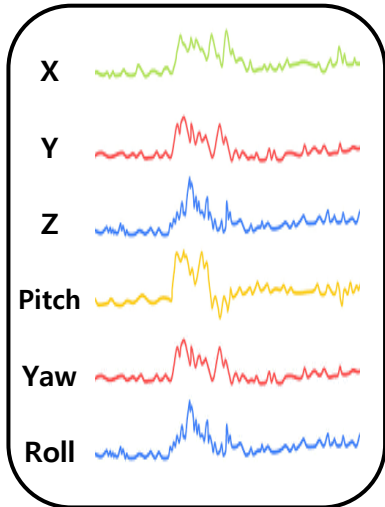


# 기존 연구 : 기계 학습 기반 HAR

- ✓ **특징 추출 과정**이 분류 모델 성능에 크게 기여
- ✓ 효과적인 **특징 발굴 및 판별 작업**으로 인한 **오버헤드**
- ✓ 메모리 접근 유발하여 **에너지, 실시간 처리 측면 비효율**

## 특징 추출 과정

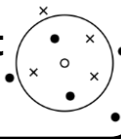
실시간 시계열 데이터



## 분류 모델

[Ignatov, 2018]

K-Nearest Neighbor



[Anguita et al., 2013]

Support Vector Machine



[Feng et al., 2015]

Random Forest

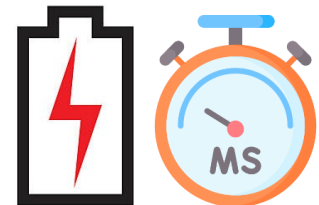


## 한계점

특징 판별 오버헤드



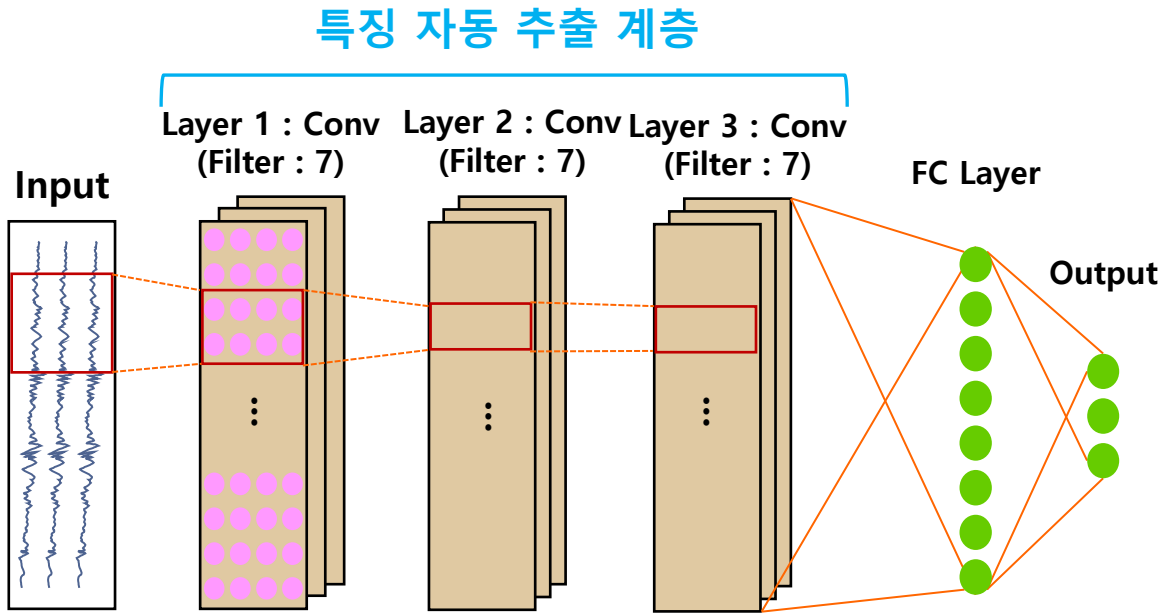
실시간 처리 비효율



# 최근 연구 : 딥러닝 기반 HAR

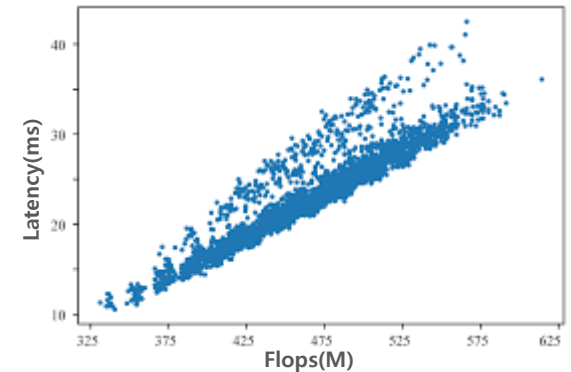
- ✓ **특징 자동 추출**하여 실시간 처리 효율적, 정확도 높음
- ✓ 파라미터 많아 **계산량 및 지연시간 높음**
- ✓ 다양한 모바일에 반해 **단편화 된 모델 구조**

## 딥러닝 모델 [Wang et al., 2019 ,Wan et al., 2020]

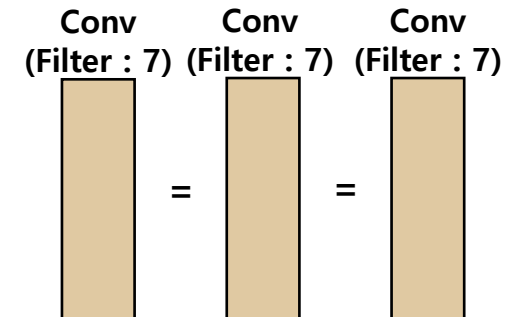


## 한계점

### 계산량 & 지연시간



### 단편화 된 모델



# 제안하는 아이디어 소개

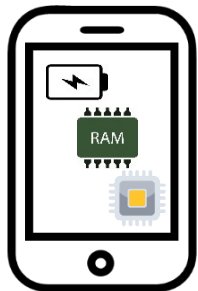
- ✓ **모바일 적응형 인간 행동 인식 딥러닝 구조 자동 탐색**을 제안
- ✓ 컴퓨터 비전의 딥러닝 구조 자동 탐색 기술 활용
- ✓ **설계 이슈** 도출 및 **재설계**

## 모바일 환경

다양한 하드웨어 사양



제한된 계산 자원



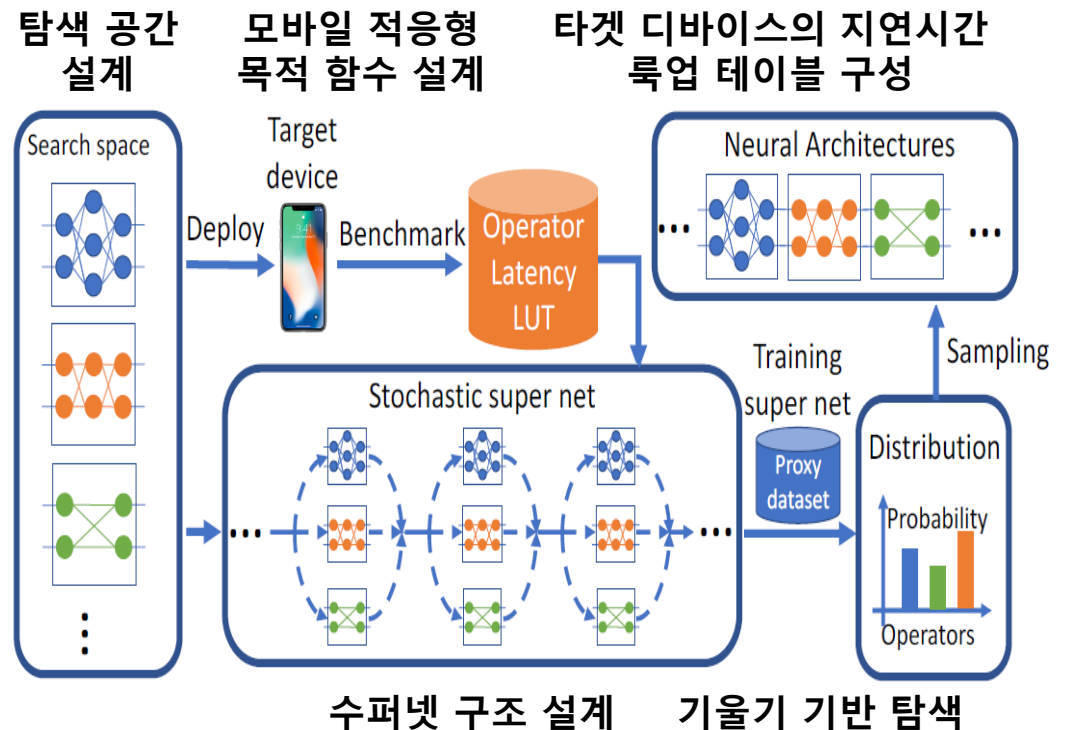
**자동**

**다변화 된  
모델 구조 설계**

**+**

**타겟 모바일의  
지연시간 고려**

## FBNet의 딥러닝 구조 자동 탐색 [Wu et al., 2019]



# 제안하는 아이디어 소개

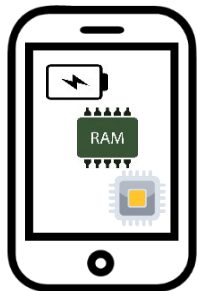
- ✓ **모바일 적응형 인간 행동 인식 딥러닝 구조 자동 탐색**을 제안
- ✓ 컴퓨터 비전의 딥러닝 구조 자동 탐색 기술 활용
- ✓ **설계 이슈** 도출 및 **재설계**

## 모바일 환경

다양한 하드웨어 사양



제한된 계산 자원



자동  
다변화 된  
모델 구조 설계

+

타겟 모바일의  
지연시간 고려

## FBNet의 딥러닝 구조 자동 탐색 [Wu et al., 2019]

모바일 적응형 인간 행동 인식  
딥러닝 구조 자동 탐색

- 1 HAR 딥러닝 구조 탐색 공간 및 수퍼넷 재설계
- 2 목적 함수 설계 및 지연시간 록업 테이블 구성
- 3 기울기 기반 최적 탐색 알고리즘

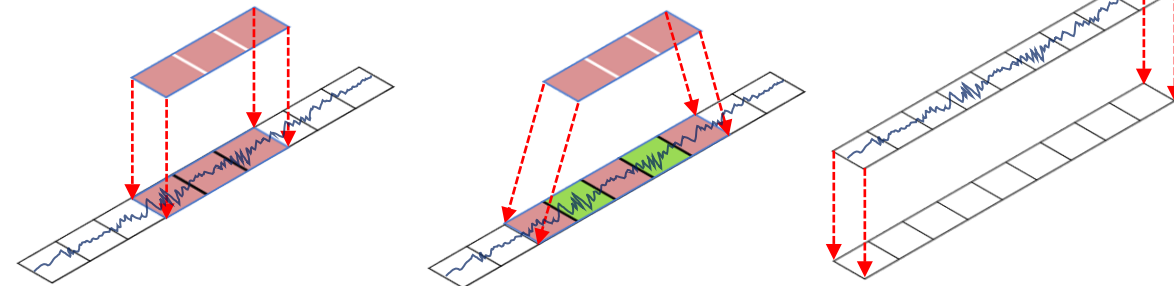
- ✓ 인간 행동 인식에 적합한 연산자로 탐색 공간 재설계
- ✓ 1차원 컨볼루션 연산자와 항등(Identity) 연산자로 탐색 공간 구성
- ✓ 가변적(Variable) 계층과 고정적(Fixed) 계층으로 수퍼넷 구성

## 탐색 공간 설계

Normal Convolution  
1 x 3

Dilated Convolution  
1 x 3, d=2

Identity Operator



필터 사이즈 : {3, 5, 7, 9}

ConvOp 3

ConvOp 5

DiConvOp 3

DiConvOp 5

IdentityOp

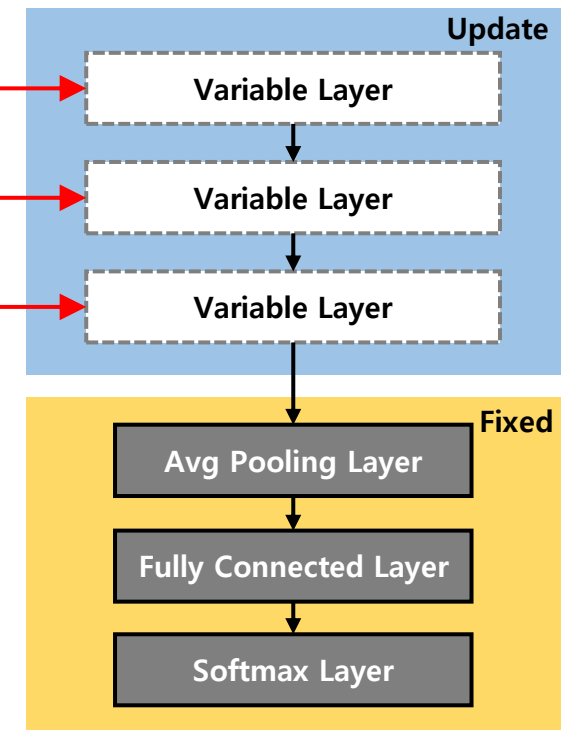
ConvOp 7

ConvOp 9

DiConvOp 7

DiConvOp 9

## 수퍼넷 설계

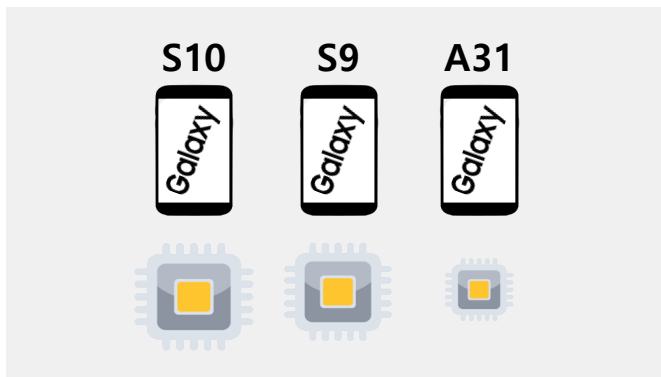


- ✓ 목적 함수는 크로스엔트로피 손실과 모델 지연시간 추정값
- ✓ 다양한 타겟 디바이스 A31, S9, S10 선정
- ✓ 디바이스에서 연산자의 지연시간 측정 후 룩업 테이블 구성

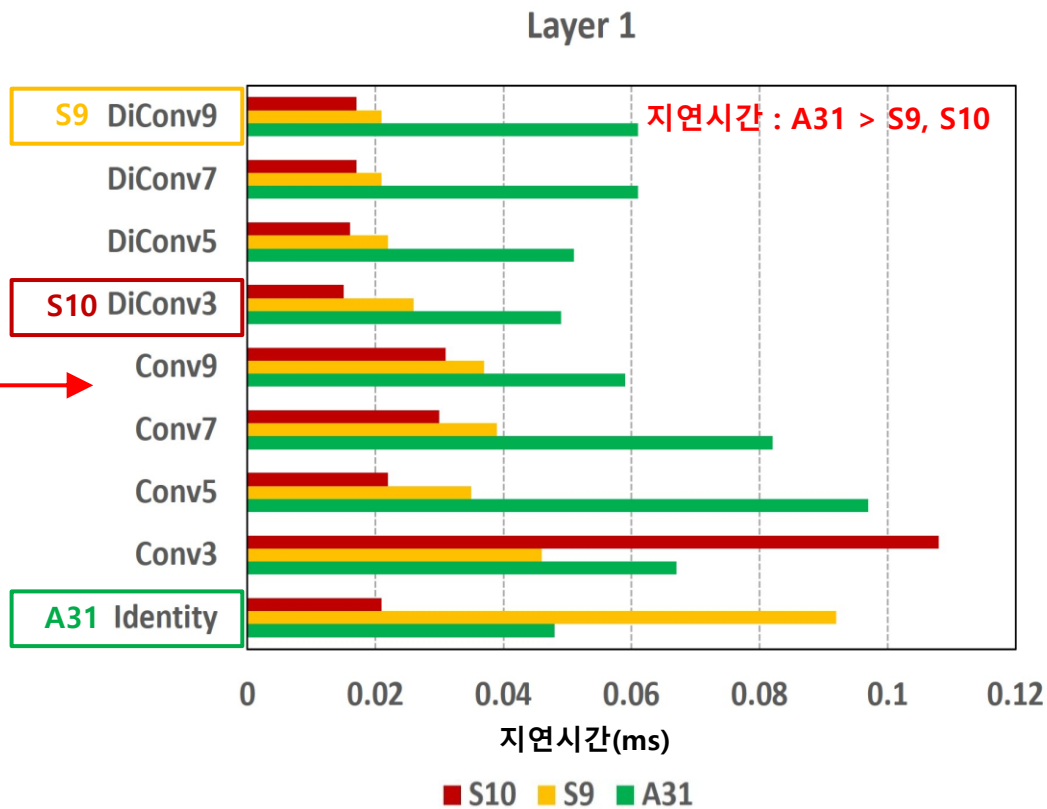
## 목적 함수 설계

$$Loss = Loss(CE) + \lambda \mathbb{E}[Latency]$$

$$\begin{aligned} \mathbb{E}[Latency] &= \sum_i \mathbb{E}[Latency_i] \\ &= p_{i,1} \times F(IdentityOp) + \\ &\quad p_{i,2} \times F(ConvOp\ 3) + \\ &\quad \dots + \\ &\quad p_{i,9} \times F(DilatedConvOp\ 9) \end{aligned}$$

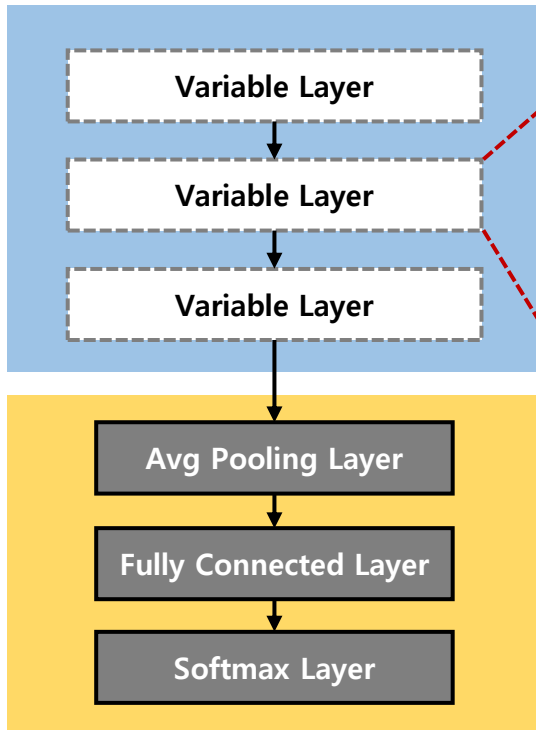


## 지연시간 룩업 테이블 구성

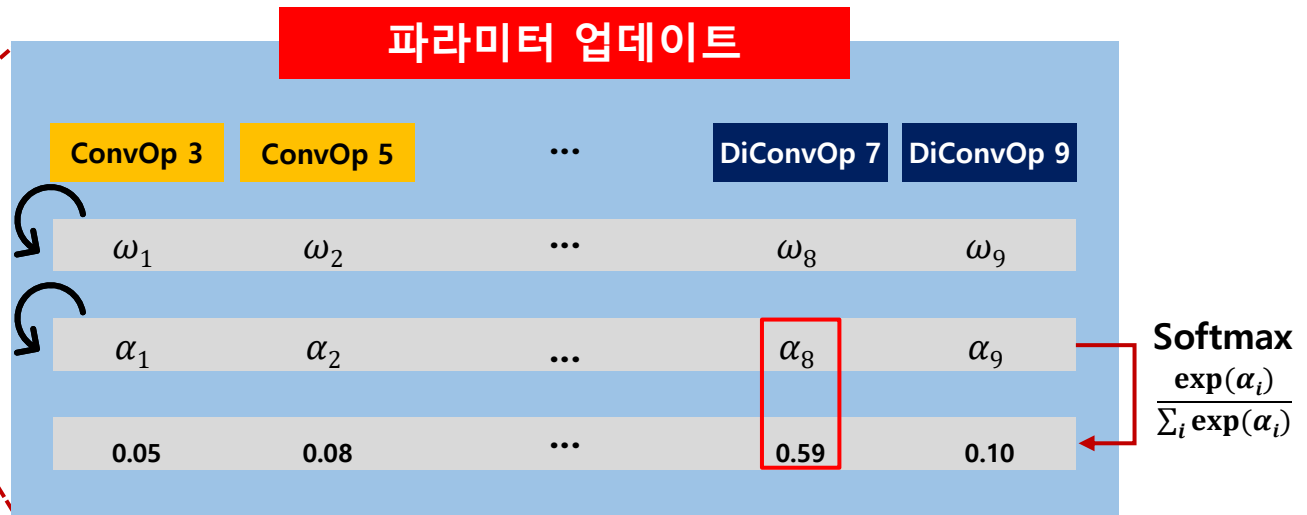


- ✓ 이산적 탐색 공간을 연속적 공간으로 완화
- ✓ 가중치와 구조 파라미터를 2단계로 학습
- ✓ 가중치 → 정확도 향상 & 구조 → 연산자 선별

## 수퍼넷



## 기울기 기반 최적 탐색



훈련 데이터 80%  
 $\omega$  업데이트



훈련 데이터 20%  
 $\alpha$  업데이트

## 실험 설계

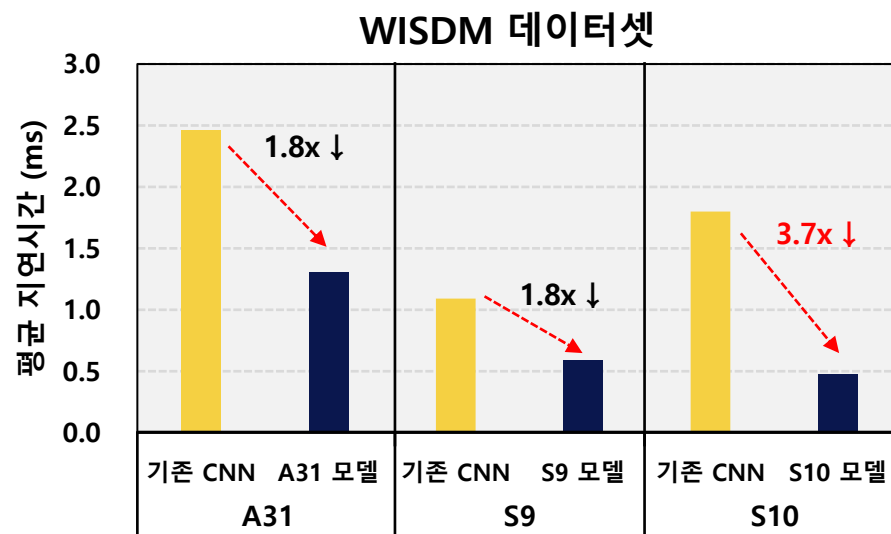
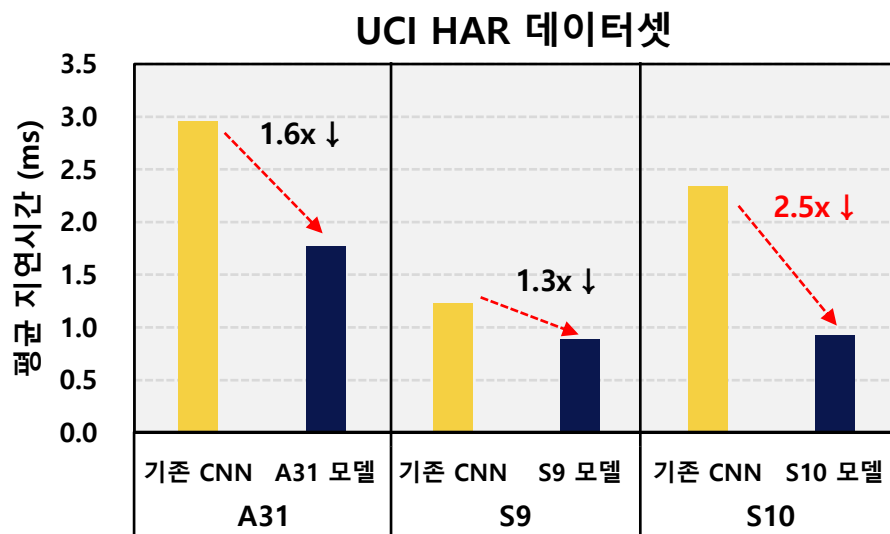
- ✓ 실험 검증 : 제안 방법 설계 모델의 **모바일 환경 효율성 증명**
- ✓ 검증 척도 : 지연시간 & F1-Score
- ✓ 비교 대상 알고리즘 : 기존 대표 CNN 모델 [Wan et al., 2020]
- ✓ 실험 데이터셋 : UCI HAR, WISDM 데이터셋 [Anguita et al., 2013, Kwapisz et al., 2011]
- ✓ 실험 : 훈련용, 검증용 데이터 (7 : 3)

### 인간 행동 인식 데이터셋에 대한 요약

| UCI HAR |       |           | WISDM   |        |         |
|---------|-------|-----------|---------|--------|---------|
| 행동 유형   | 데이터 수 | 특징        | 행동 유형   | 데이터 수  | 특징      |
| 서기      | 1,906 | X축-가속도계   | 서기      | 2,303  | X축-가속도계 |
| 앉기      | 1,777 | Y축-가속도계   | 앉기      | 2,810  | Y축-가속도계 |
| 눅기      | 1,944 | Z축-가속도계   | 조깅      | 15,074 | Z축-가속도계 |
| 걷기      | 1,722 | X축-자이로스코프 | 걷기      | 20,281 |         |
| 계단 내려가기 | 1,406 | Y축-자이로스코프 | 계단 내려가기 | 4,786  |         |
| 계단 오르기  | 1,544 | Z축-자이로스코프 | 계단 오르기  | 5,863  |         |

## 실험 검증 : 지연시간 및 정확도 비교 결과

- ✓ 탐색된 모델과 기존 CNN 모델의 지연시간 약 1.3x~3.7x 배 차이
- ✓ 고사양 기종인 **S10에서 지연시간** 차이 극명
- ✓ 정확도 성능은 비슷하거나 탐색된 모델이 우월



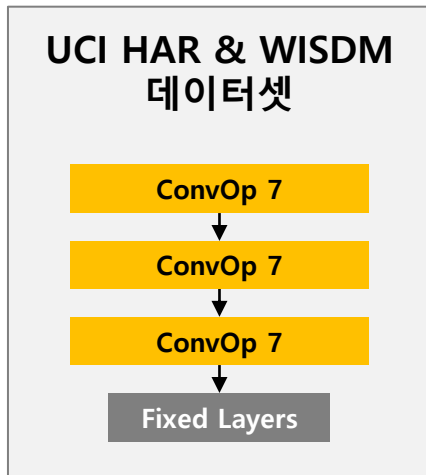
UCI HAR, WISDM 데이터셋에 대한 평균 F1-Score 비교 결과

| 데이터셋    | 기존 CNN | A31 모델 | S9 모델         | S10 모델        |
|---------|--------|--------|---------------|---------------|
| UCI HAR | 0.9392 | 0.9536 | <b>0.9558</b> | 0.9551        |
| WISDM   | 0.8612 | 0.8564 | 0.8617        | <b>0.8710</b> |

# 실험 검증 : 탐색된 모델 구조 비교 결과

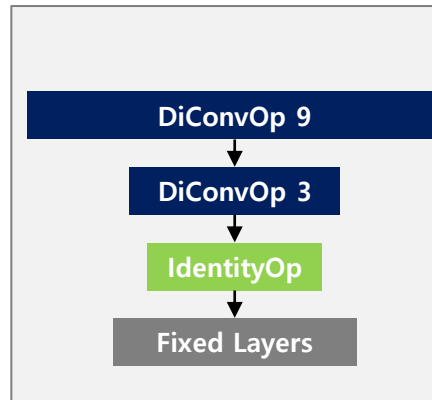
- ✓ 자동 탐색된 모델은 확장 컨볼루션 사용
- ✓ 필터 사이즈를 줄이는 방식으로 모델 설계
- ✓ 각 디바이스별로 서로 다른 모델 탐색

기존 CNN 모델

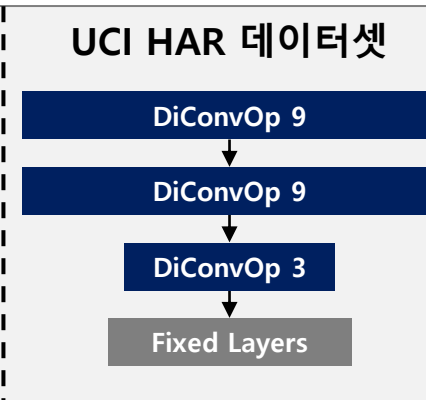


VS

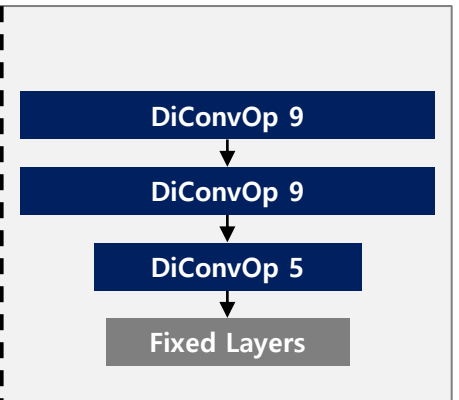
A31 모델



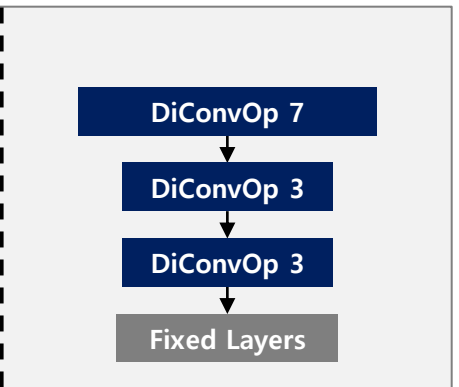
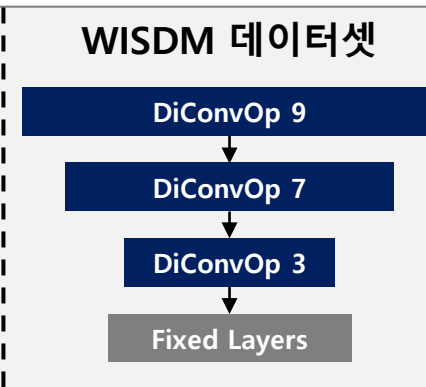
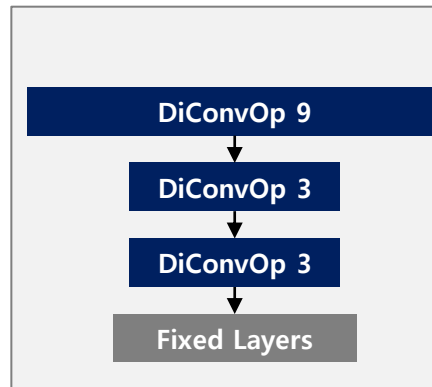
S9 모델



S10 모델



WISDM 데이터셋



## 결론

- ✓ 딥러닝 구조 자동 탐색을 활용한 모바일 인간 행동 인식 딥러닝 모델 자동 설계
- ✓ 기존 대표 CNN 모델보다 낮은 지연시간 성능과 좋은 정확도 성능 결과
- ✓ 인간 행동 인식에서 모바일 환경을 고려한 모델 설계의 중요성

## 향후 연구

1. 탐색 공간 설계하기 위한 해당 분야 지식과 딥러닝 설계 지식 필요  
→ 응용 분야에 독립적으로 활용할 수 있는 연산자에 대한 연구
2. 지연시간 룩업 테이블 구성에 따른 시간적 비용 발생  
→ 지연시간 룩업 테이블 구성하지 않고 계산할 수 있는 지연시간 예측 모델 연구

## 참고 문헌

- [Radu et al., 2016] Radu, V., Lane, N. D., Bhattacharya, S., Mascolo, C., Marina, M. K., and Kawsar, F. (2016). Towards multimodal deep learning for activity recognition on mobile devices. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 185–188.
- [Nweke et al., 2018] Nweke, H. F., Teh, Y. W., Al-Garadi, M. A., and Alo, U. R. (2018). Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Systems with Applications*, 105:233–261.
- [Anguita et al., 2013] Anguita, D., Ghio, A., Oneto, L., Parra, X., and ReyesOrtiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. In *Esann*, volume 3, page 3.
- [Feng et al., 2015] Feng, Z., Mo, L., and Li, M. (2015). A random forest-based ensemble method for activity recognition. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5074–5077. IEEE.
- [Ignatov, 2018] Ignatov, A. (2018). Real-time human activity recognition from accelerometer data using convolutional neural networks. *Applied Soft Computing*, 62:915–922.
- [Wan et al., 2020] Wan, S., Qi, L., Xu, X., Tong, C., and Gu, Z. (2020). Deep learning models for real-time human activity recognition with smartphones. *Mobile Networks and Applications*, 25(2):743–755.
- [Wang et al., 2019] Wang, J., Chen, Y., Hao, S., Peng, X., and Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119:3–11.
- [Wu et al., 2019] Wu, B., Dai, X., Zhang, P., Wang, Y., Sun, F., Wu, Y., Tian, Y., Vajda, P., Jia, Y., and Keutzer, K. (2019). FBnet: Hardware-aware efficient convnet design via differentiable neural architecture search. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10734–10742.
- [Kwapisz et al., 2011] Kwapisz, J. R., Weiss, G. M., and Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2):74–82.

A vertical bar on the left side of the slide, consisting of a red upper half and a grey lower half.

# Thank you

## Appendix : Example

- ✓ 수퍼넷의 가변적 계층에서 연산자들을 혼합하여 계산하는 과정

