

Study on the use of CNN Models for High-pass Filtering in Ground Motion Data Generation for Evaluating Seismic Fragilities

2023.10.26 (Thu.)

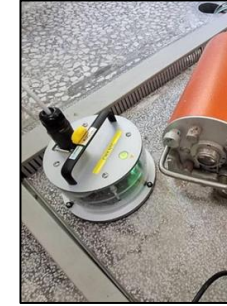
Jin Koo Lee*, JeongBeom Seo, Seong Jin Jeon
KIT Valley co., Ltd.

- Background and Objective
- Dataset and Preprocessing
- Training and Results
- Summary and Reference

BACKGROUND AND OBJECTIVE

빅데이터 기반 원전 부지 입력지진 평가기술 개발

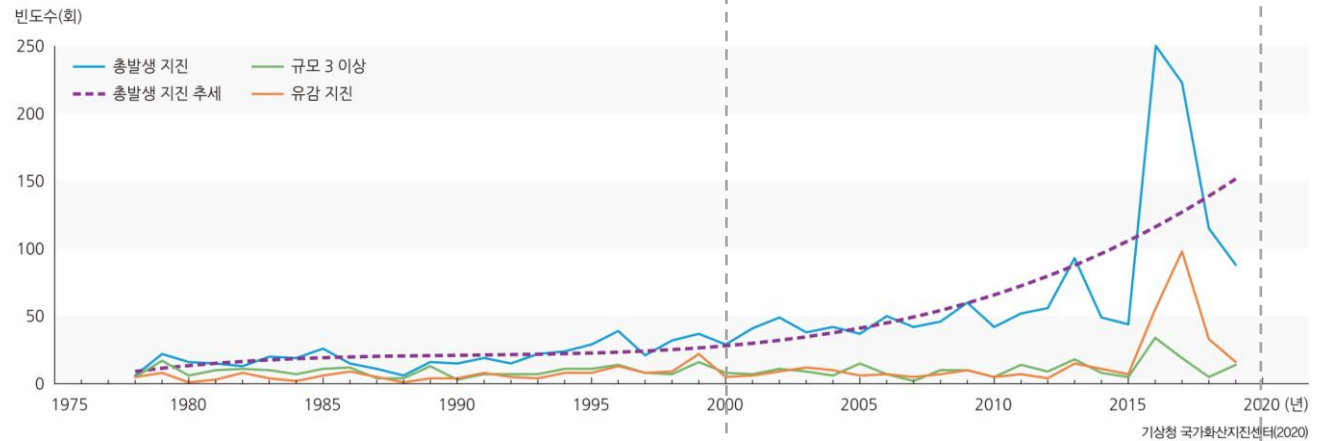
- 스마트 지진 센서 개발
- 지진 데이터, 관측소 메타데이터 수집 및 처리
- 입력지진 평가 기술 개발
 - ✓ 원전 부지내 지반운동의 비상관성을 고려한 확률론적 부지 응답 해석 기술 개발
 - ✓ 내진성능평가를 위한 부지고유 입력지진동 평가 기술 개발



■ 연간 지진 발생 빈도 및 지진 관측소 현황

- 연간 지진 발생 빈도는 증가 추세
- 2023년 기준 300여개 이상의 지진 관측소 운용(기상청), 60여개 지진관측소 운용(지질자원연구소)
- 한국원자력안전기술원은 2022년 9월부터 220대 규모의 원전부지 지진 관측망 운용(김민욱 외, 2022)

연간 지진 발생 빈도



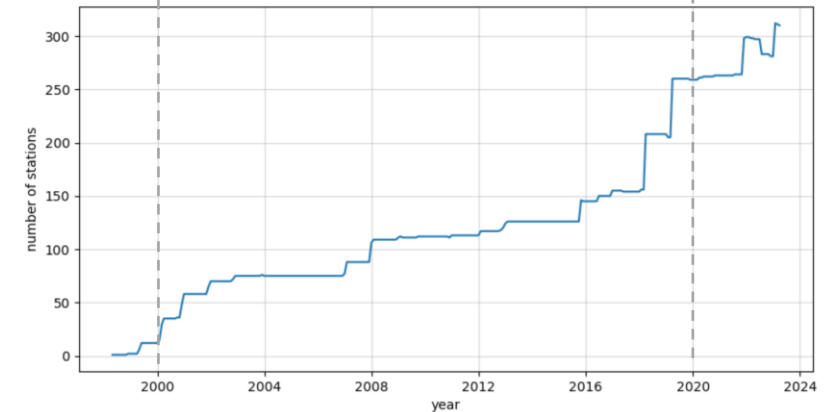
아파트 외벽 금 가...장수군 지진 피해 6건 발생

2023.07.30 21:15 일력 -
 원문로 선암기자 박용팔 광기명 기자

‘규모 3.5’ 올해 지진 중 3위
 모두 균열...인명피해는 없어
 부산 포함 ‘유감 신고’ 52건



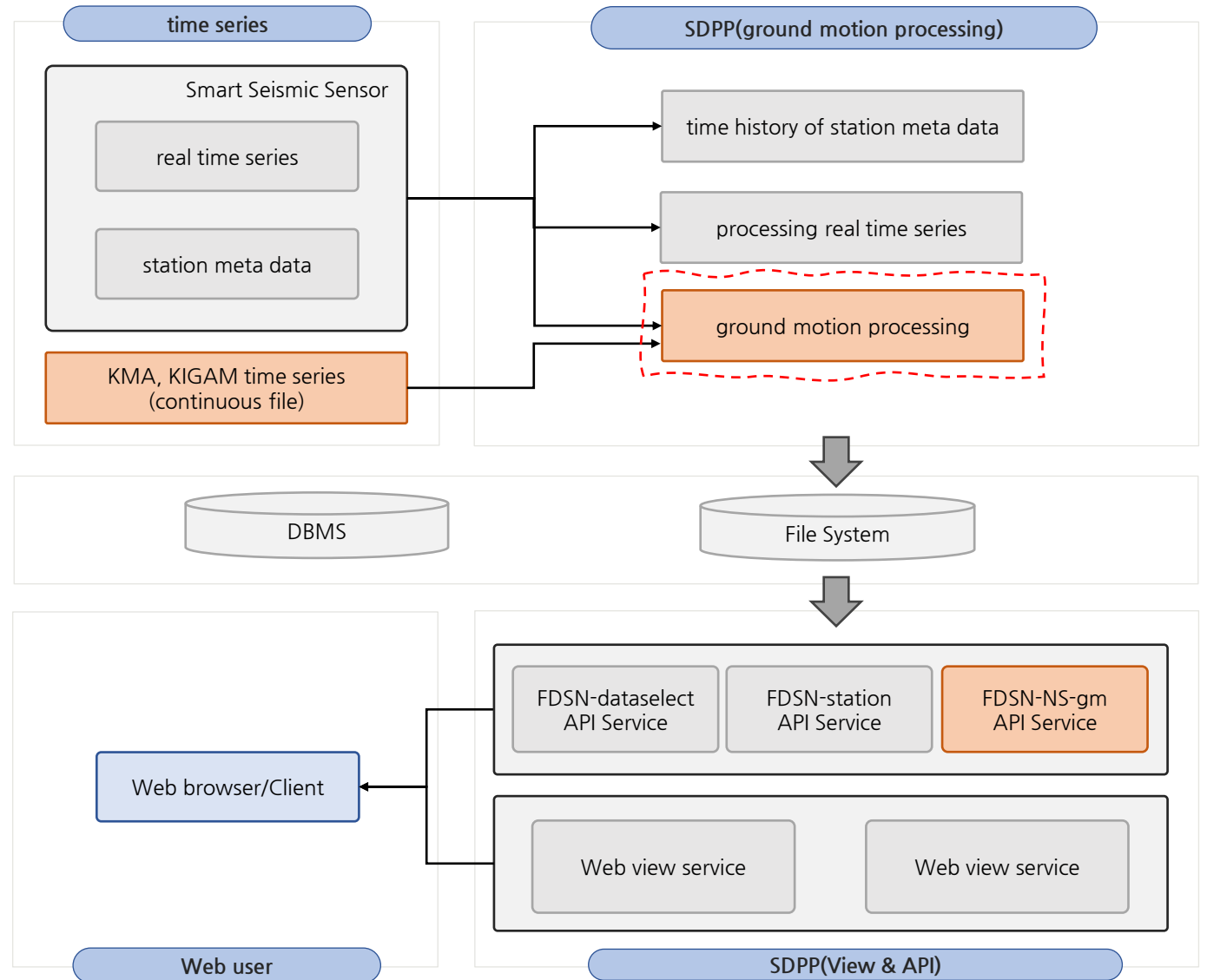
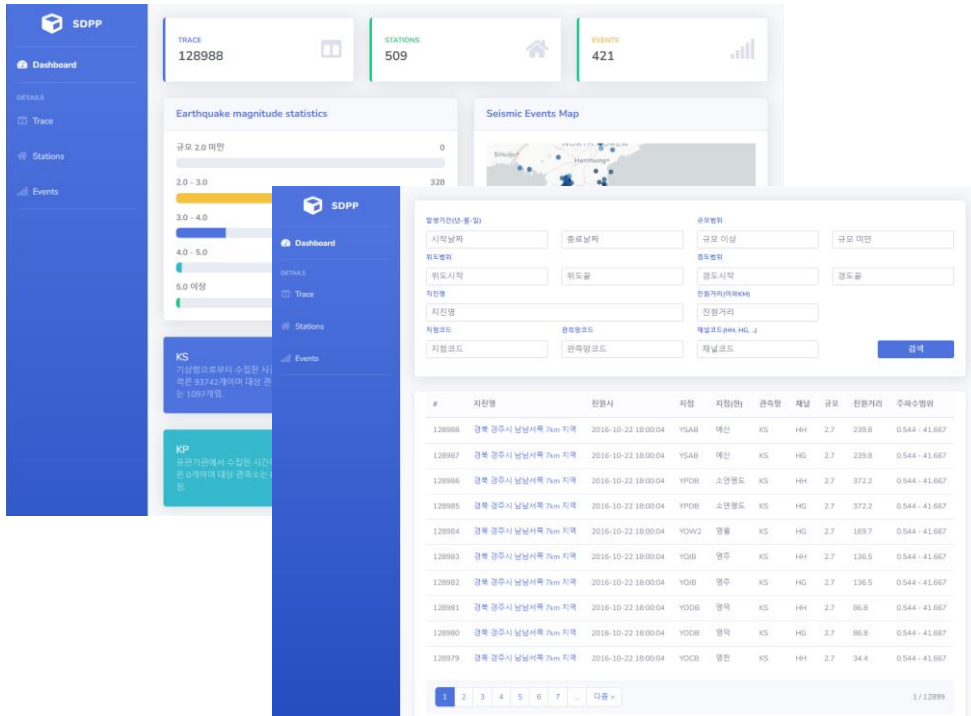
지난 29일 발생한 지진으로 인해 전체 전라군에 한 아파트 외벽이 균열해 있다. 왼쪽도 제공



(The number of station per year, NECIS)

SDPP(Seismic Data Processing Platform)

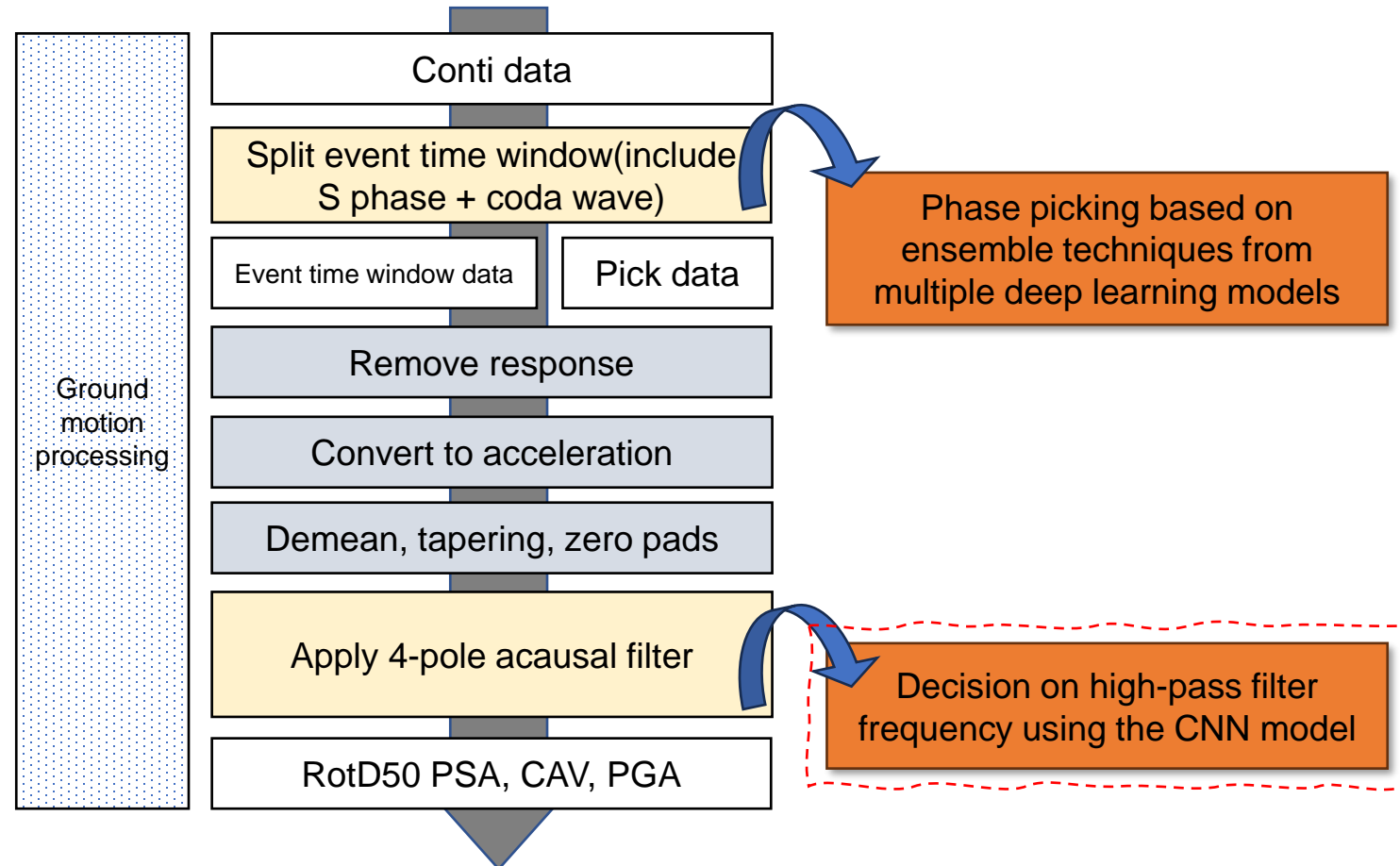
- 시간이력수집 및 지반운동 처리
- 웹기반의 인터페이스 제공(대시보드 등)
- FDSN API기반의 legacy data 제공



※ FDSN: Federation of Digital Seismograph Networks

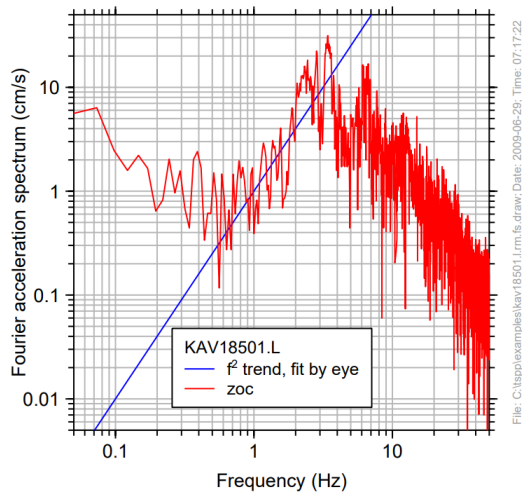
■ 지반운동 처리

- PEER(Pacific Earthquake Engineering Research Center)의 처리과정을 참조
- Deep learning을 이용한 자동화 처리

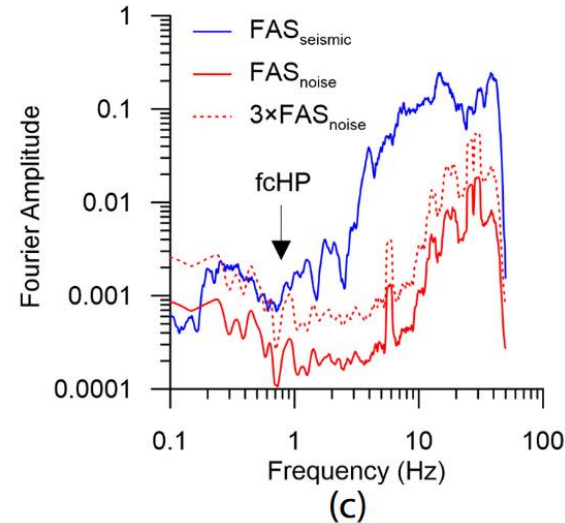


High-pass 주파수 결정을 위한 기존의 방법들

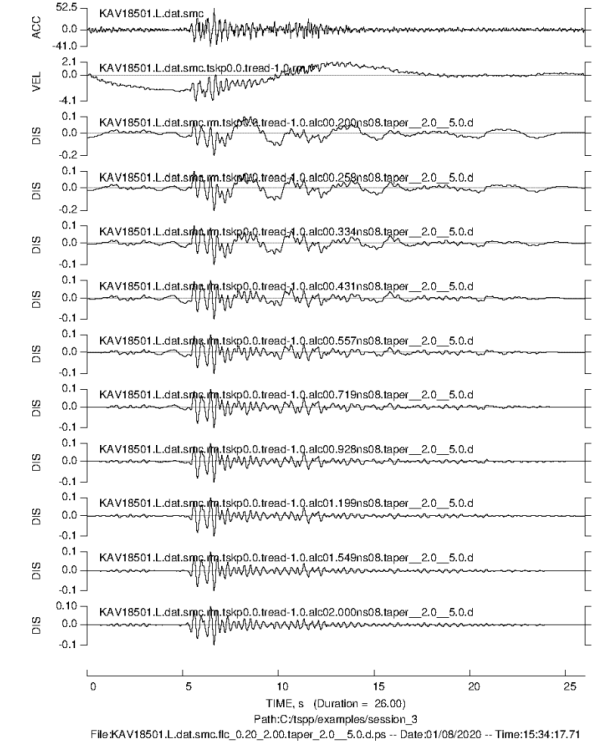
- f^2 trend 곡선과의 비교
- FAS SNR을 비교
- 특정 주파수로 필터링 이후 Displacement 곡선을 확인



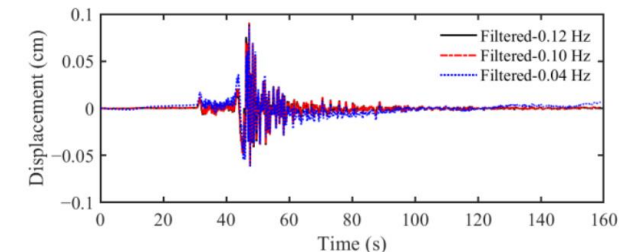
(FAS for the sample record. The blue line has a slope of 2, Boore, 2009)



(The comparison of 3XFAS-noise and FAS-seismic, Choi, 2019)



(Displacement curve after filtering, Boore, 2008)



(Displacement curve after filtering with certain frequency, Liu, B. 2022)

DATASET AND PREPROCESSING

■ 데이터 기본 전처리

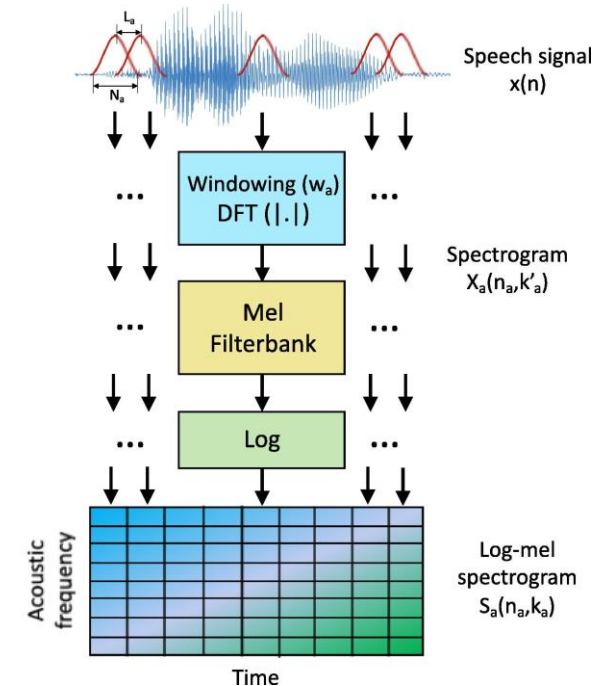
- 100Hz 샘플링 레이트로 조정
- 속도데이터는 가속도 데이터로 변환
- Demean, 기기응답 제거
- 물리적 단위로 변환

■ 멜-스펙트로그램에 의한 특징 추출

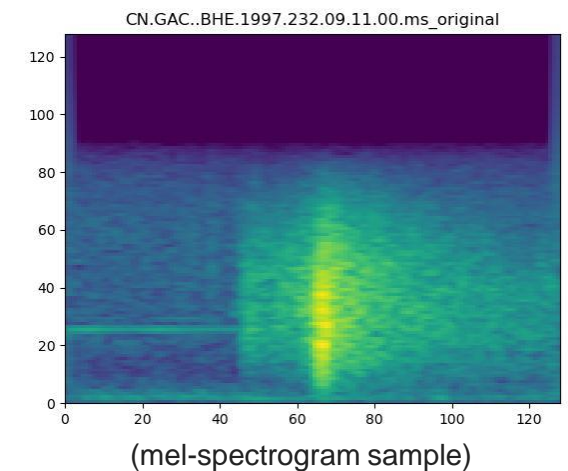
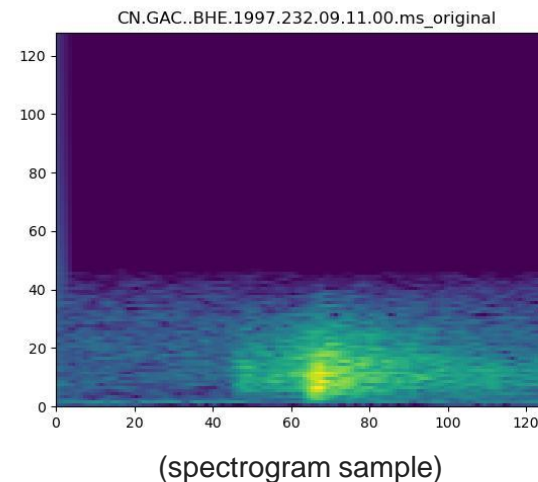
- 음성 인식, 음성 신호 추출 분야에 사용
- 오탐지 분류(Seo, 2021), 지진 탐지(Shakeel, 2021)
- 저주파수 대역에 대한 민감도 증가
- 오버래핑-윈도우 단위로 분할
- 각 윈도우에 FFT변환
- Mel스케일, Db스케일 변환
- 시간축에 따라 결과물을 배치

■ 입력데이터 형식 통일

- 128 x 128 사이즈로 변환



(Gallardo-Antolín, A., & Montero, J. M., 2021)



■ 데이터셋

- SK: Database of response history for historical earthquake records on the Korean peninsula (Choi, et al., 2020) + NECIS(국가지진종합정보시스템)
- CENA: PEER NGA database (Chiou, et al., 2008) + IRIS DMC

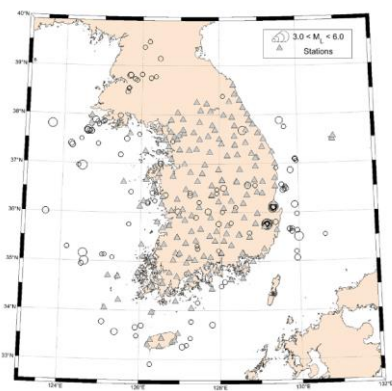
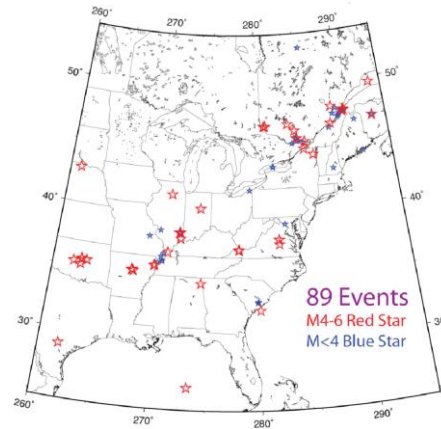
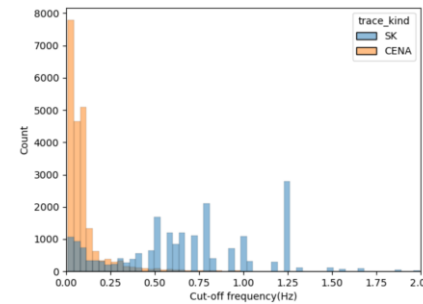


Fig. 1. Location of the selected events and stations used for ground motion database

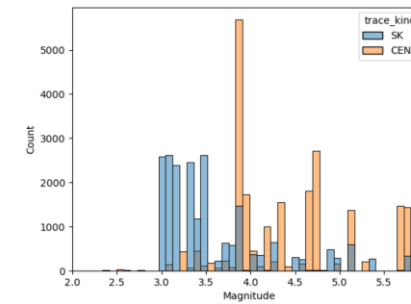
(SK, Choi et al., 2020)



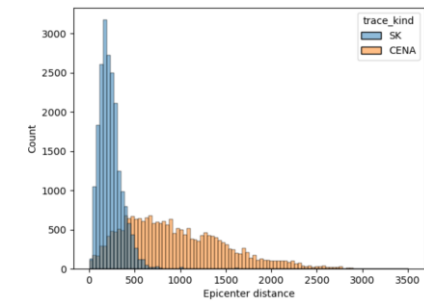
(CENA, Chiou et al., 2008)



(High-pass 주파수값 분포)



(규모 분포)

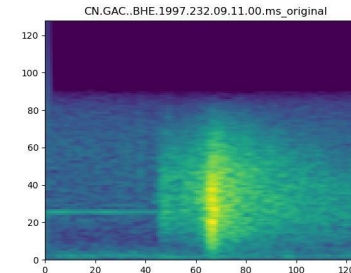


(진앙거리 분포)

X: (128x128) 2D array

y: float number

Total	Training	Validation	Test
42,980	34,384	4,298	4,298

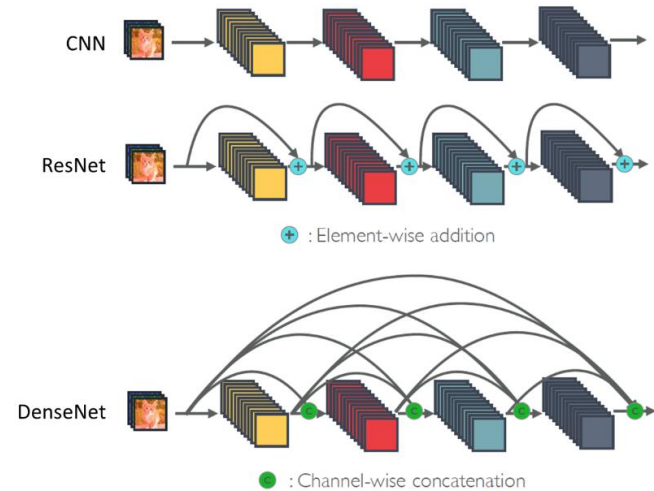


0.423

TRAINING AND RESULTS

■ 대상 모델

- ResNet(Residual Net)
- WRN(Wide ResNet)
- DenseNet
- EfficientNet

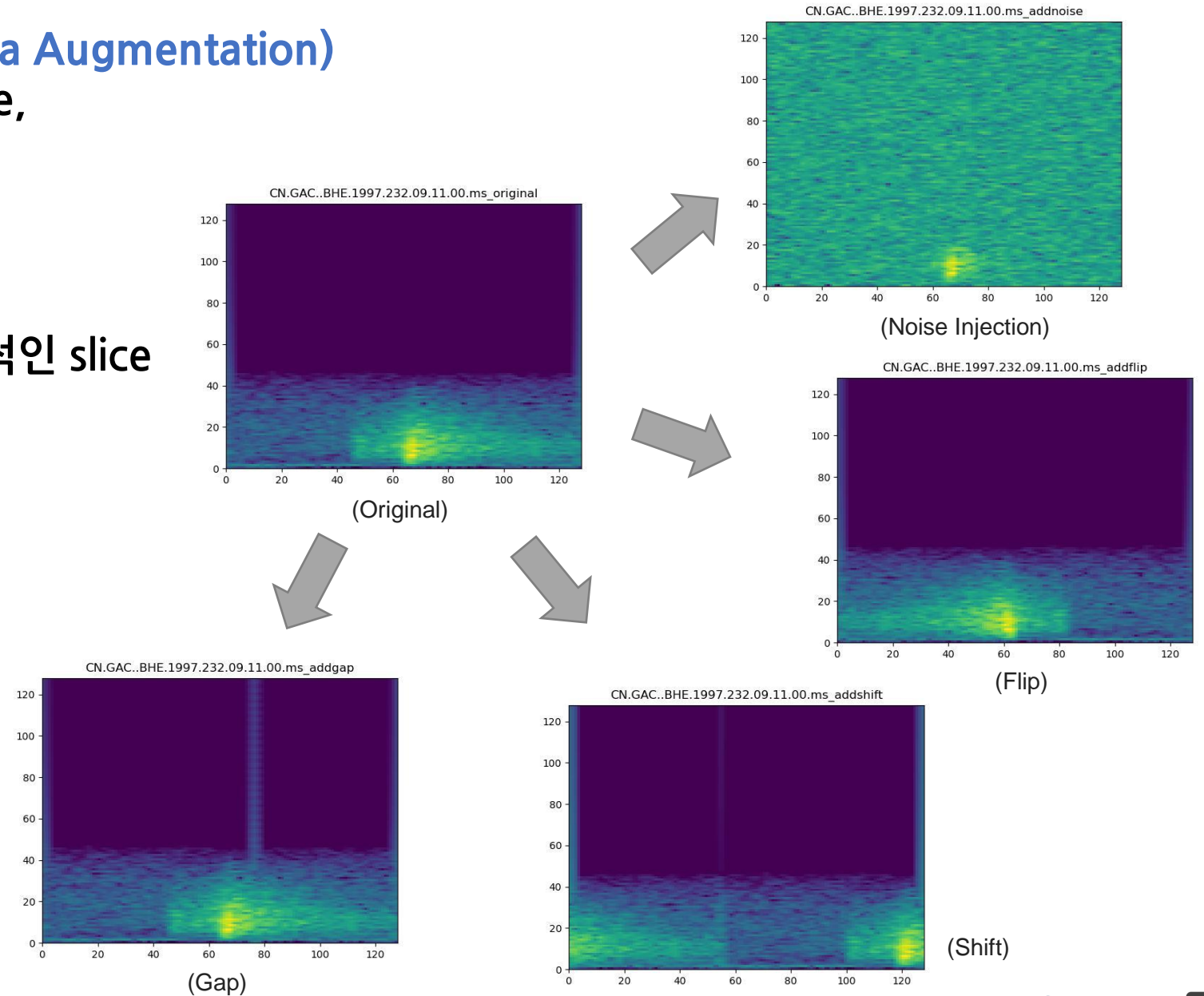


	ResNet	WRN	DenseNet	EfficientNet
Publication year	2015	2016	2017	2019
Params(M)	25.6	69	8.06	5.33
G-FLOPs*	8.2	22.8	5.8	0.8
Features	ReLU, 7x7 conv with pooling, 1x1 conv shortcut downsample	ReLU, 7x7 conv with pooling, 1x1 conv shortcut downsample	ReLU, (1x1 conv/3x3 conv) dense block, 1x1 conv shortcut with average pool	Compound scaling
Detailed name	ResNet-50	WRN-50	DenseNet-121	EfficientNet-B0
Author	Kaiming He	Sergey Zagoruyko	Gao Huang	Mingxing Tan

*G-FLOPs: GPU Floating point Operations Per Second

■ 시계열 데이터에서의 데이터 증강(Data Augmentation)

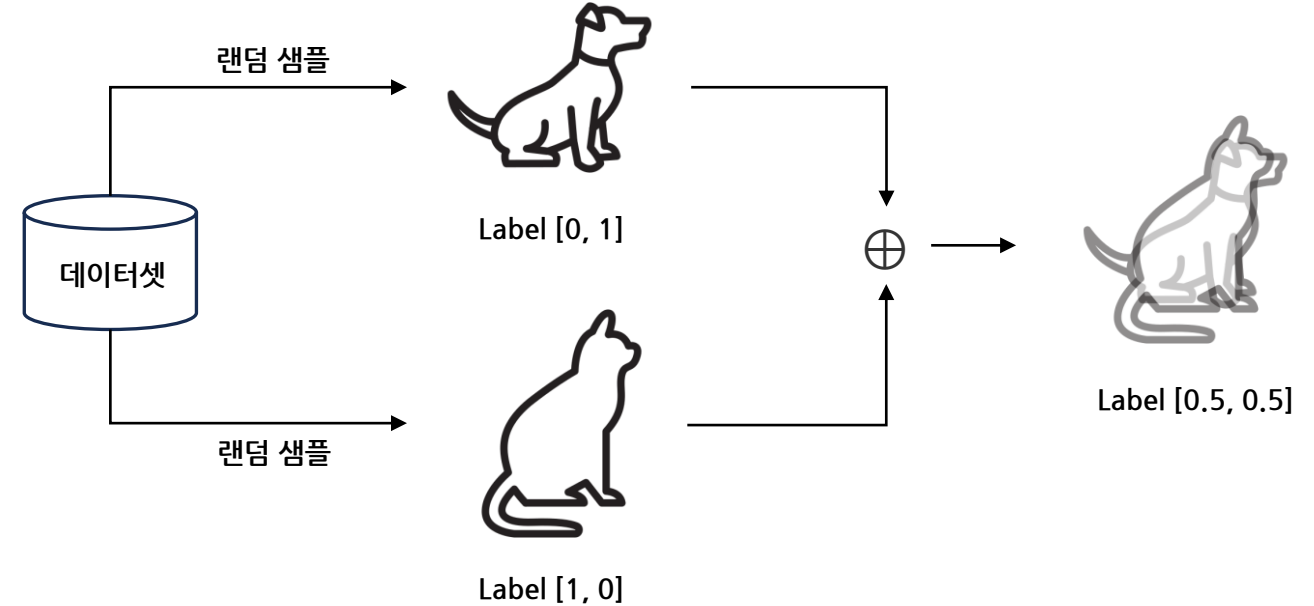
- Noise Injection: Gaussian, Spike, Random 노이즈 추가
- Gap: 공백 추가
- Flip: 좌우, 상하 뒤집기
- Shift: 좌우를 이동
- Cropping, Slicing: 랜덤하게 연속적인 slice를 추출
- Wrapping: 압축 또는 확장



■ Mixup(Zhang, 2017)

- 랜덤 샘플에 대한 선형적 보간
- $X_{new} = \lambda X_i + (1 - \lambda)X_j$
- $Y_{new} = \lambda Y_i + (1 - \lambda)Y_j$

- 오버피팅 방지와 데이터 증강 효과



(Mixup)

■ C-Mixup(Yao., et al., 2022)

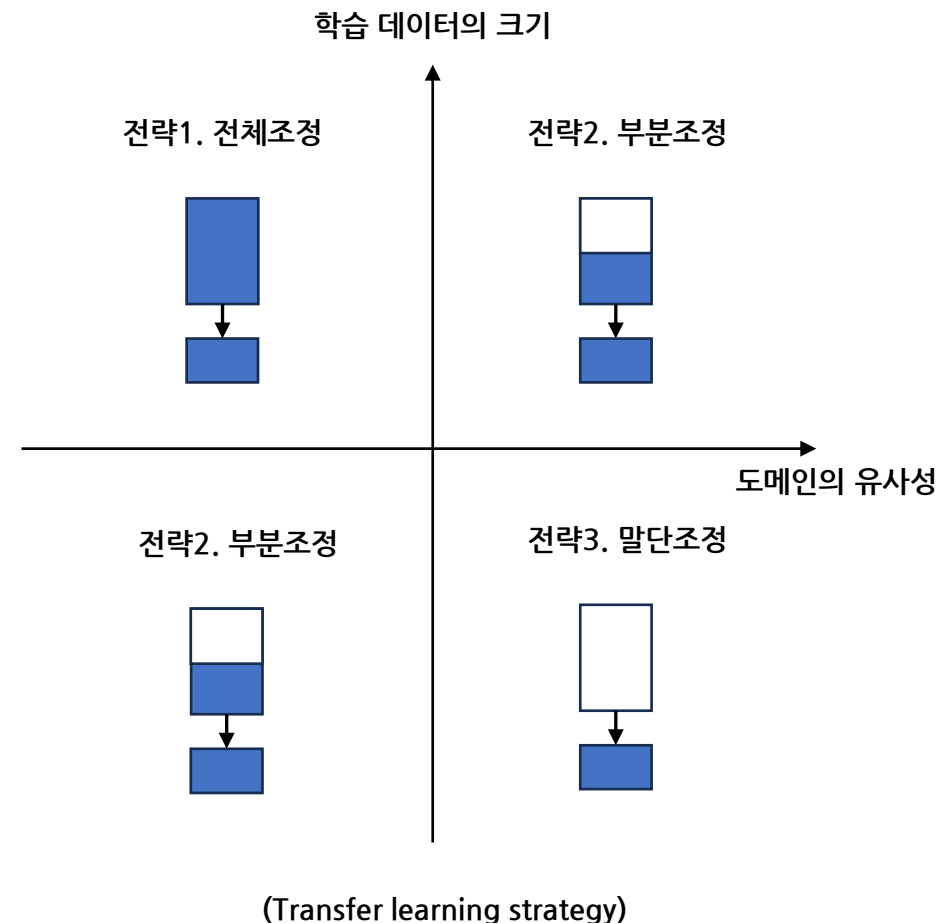
- Regression을 위한 Mixup
- 랜덤이 아닌 유사도가 높은 샘플들에 대해 짝을 지음

■ 전이학습(Transfer learning)

- 모델의 파라미터 수는 5백 만개 이상
- 학습 데이터의 크기(4만)와 비교해 볼 때 학습이 쉽지 않음

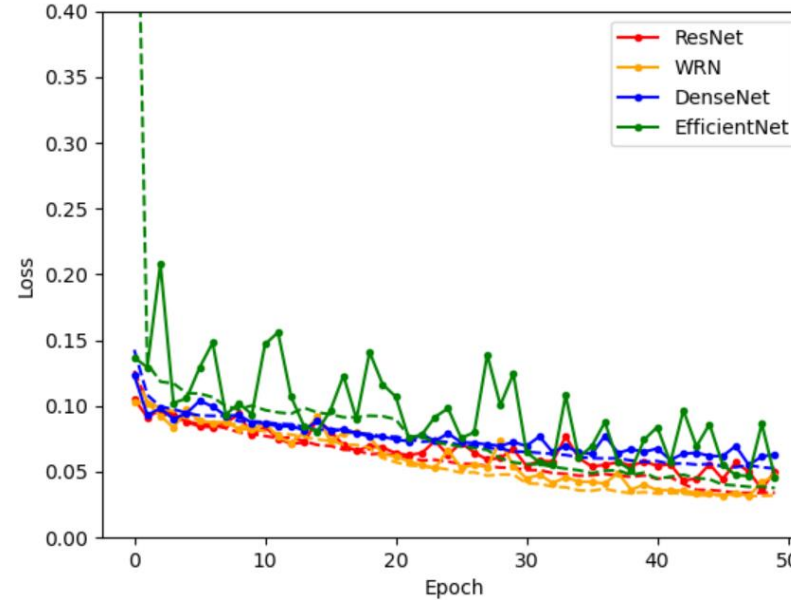
■ 대상 및 전략

- Imagenet-1K(1,281,167 training, 50,000 validation, 100,000 test images)
- 1,000개 이미지에 대한 분류
- huggingface의 학습모델
- 구하려는 문제가 바뀔때 따라 모델 출력 레이어 변경 (Classification → Regression)

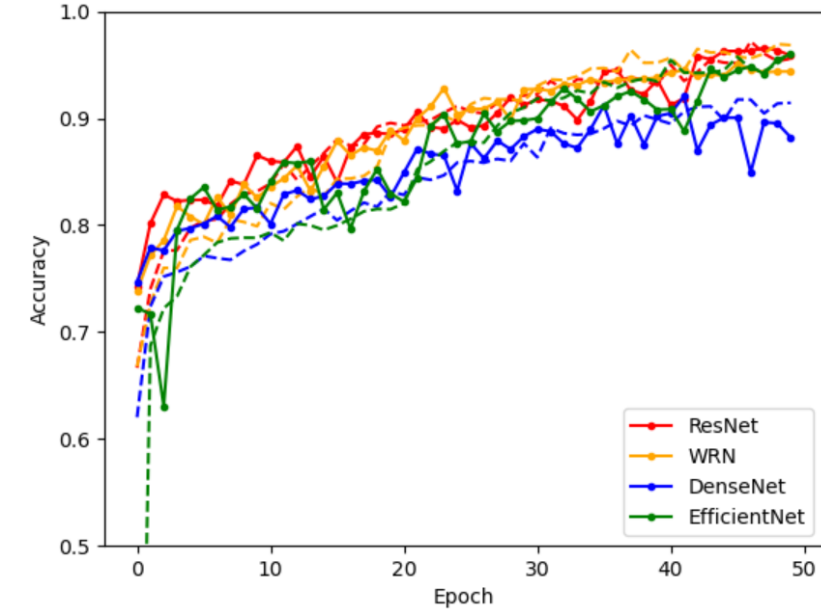


■ 학습 결과

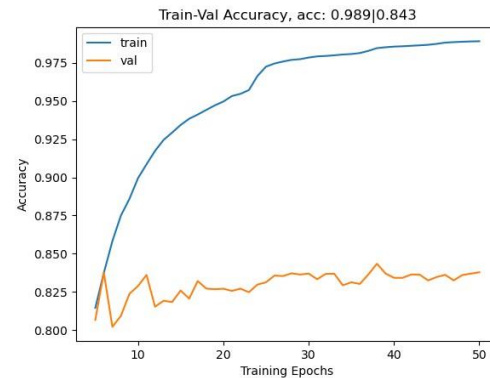
- 각 epoch에 따른 loss 및 accuracy 그래프
- 점선: 훈련 데이터셋
- 실선: 검증 데이터셋



(The loss of CNN models in Train and Validation)



(The accuracy of CNN models in Train and Validation)



(The loss/accuracy of CNN models in Train and Validation, No data augmentation)

	ResNet	WRN	DenseNet	Efficient Net
Best Accuracy	0.97	0.95	0.92	0.96
Best loss	0.04	0.03	0.06	0.05

■ 검증

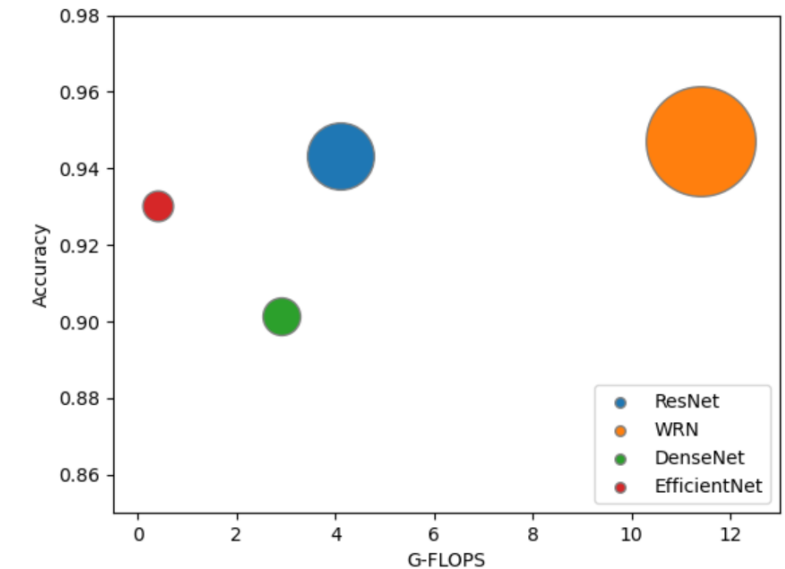
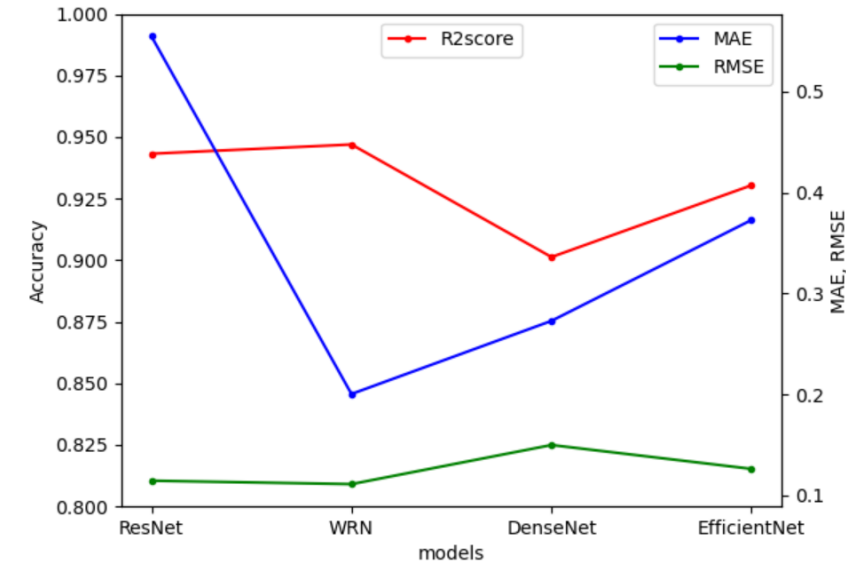
- R^2 : 1=prediction is perfect, 0=prediction is dice, \uparrow good
- MAE(Mean Absolute Error): \downarrow good
- RMSE(Root Mean Square Error): \downarrow good
- G-FLOPS: GPU FLoating point Operations Per Second: \downarrow good

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

	ResNet	WRN	DenseNet	Efficient Net
R^2	0.94	0.95	0.90	0.93
MAE	0.55	0.20	0.27	0.37
RMSE	0.11	0.11	0.15	0.13



SUMMARY AND REFERENCE

■ 요약 및 결론

- 지반운동을 도출하기 위한 high-pass 필터 주파수를 결정하기 위해 딥러닝-CNN기반의 모델을 적용하고 테스트하였음
- 학습데이터가 부족한 문제를 해결하기 위해 전이학습을 사용하였으며 도메인의 유사성이 다름에도 좋은 성능을 나타내었음
- 데이터 증강을 위해 mixup기법을 사용하였으며 이는 다른 기법들 보다 더 높은 성능으로 모델의 오버피팅을 개선할 수 있었음
- 이상의 학습된 딥러닝 모델을 사용함으로써 지반운동을 자동으로 효과적으로 처리할 수 있을 것을 기대함

■ 이후 작업들

- 실제 시스템에 적용하여 성능(처리속도 등)에 대한 평가를 수행해야 함
- CNN모델 외에 어텐션 계열의 모델에 대한 성능 비교

- C. A. Goulet et al., "PEER NGA-East database," *Earthq. Spectra*, vol. 37, no. 1_suppl, pp. 1331-1353, Oct. 2014, doi: 10.1177/87552930211015695.
- J. Seo et al., "Deep Learning-Based, Real-Time, False-Pick Filter for an Onsite Earthquake Early Warning (EEW) System," *J. Earthq. Eng. Soc. Korea*, vol. 25, no. 2, pp. 71-81, Mar. 2021, doi:10.5000/EESK.2021.25.2.071.
- M. Shakeel, K. Itoyama, K. Nishida, and K. Nakadai, "EMC: Earthquake Magnitudes Classification on Seismic Signals via Convolutional Recurrent Networks," 2021 IEEE SICE Int. Symp. Syst. Integr. SII, pp. 388-393, Jan. 2021, doi: 10.1109/IEEECONF49454.2021.9382696.
- S.-W. Choi, J. Rhie, S. H. Lee, and T.-S. Kang, "A Study on Development of an Earthquake Ground-motion Database Based on the Korean National Seismic Transactions of the Korean Nuclear Society Autumn Meeting Gyeongju, Korea, October 26-27, 2023 Network," *J. Earthq. Eng. Soc. Korea*, vol. 24, no. 6, pp. 277-283, Nov. 2020, doi: 10.5000/EESK.2020.24.6.277.
- K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770-778. doi: 10.1109/CVPR.2016.90.
- S. Zagoruyko and N. Komodakis, "Wide Residual Networks." *arXiv*, Jun. 14, 2017. Accessed: Jul. 19, 2023. [Online]. Available: <http://arxiv.org/abs/1605.07146>
- G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks." *arXiv*, Jan. 28, 2018. Accessed: Jul. 19, 2023.
- M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks". International conference on machine learning. PMLR, 2019.
- H. Yao, Y. Wang, L. Zhang, J. Zou, and C. Finn, "C-Mixup: Improving Generalization in Regression." *arXiv*, Oct. 11, 2022. Accessed: Jun. 13, 2023.

THANK YOU FOR YOUR ATTENTION !!

