2025 09 11
발표 자료

광운대학교 로봇학과
FAIR Lab

김한서

KWANGWOON UNIVERSITY

2D CNN 모델

FAIR Lab

KWANGWOON UNIVERSITY

이번 주 진행사항

- 2D CNN 모델 구성
- 데이터셋 및 실험 세팅
- 저번 주 모델과 이번 주 모델 비교

  → 1D CNN label scale 1.0과 2D CNN label scale 2.0
- 2D CNN 비교

  → label scale 1.0과 2.0 실험과 시각화 비교

  → label return day 1, 5, 10, 20 각각 실험 비교

Quant data 모델 학습

FAIR Lab

KWANGWOON UNIVERSITY

현재 진행 중인 실험 설명

- Input / Label 전 종목 사용
- Input feature → Open, Close, High, Low, Volume, Vwap, Ticker

  Label feature → 1_day_return_rate / 5_day_return_rate /

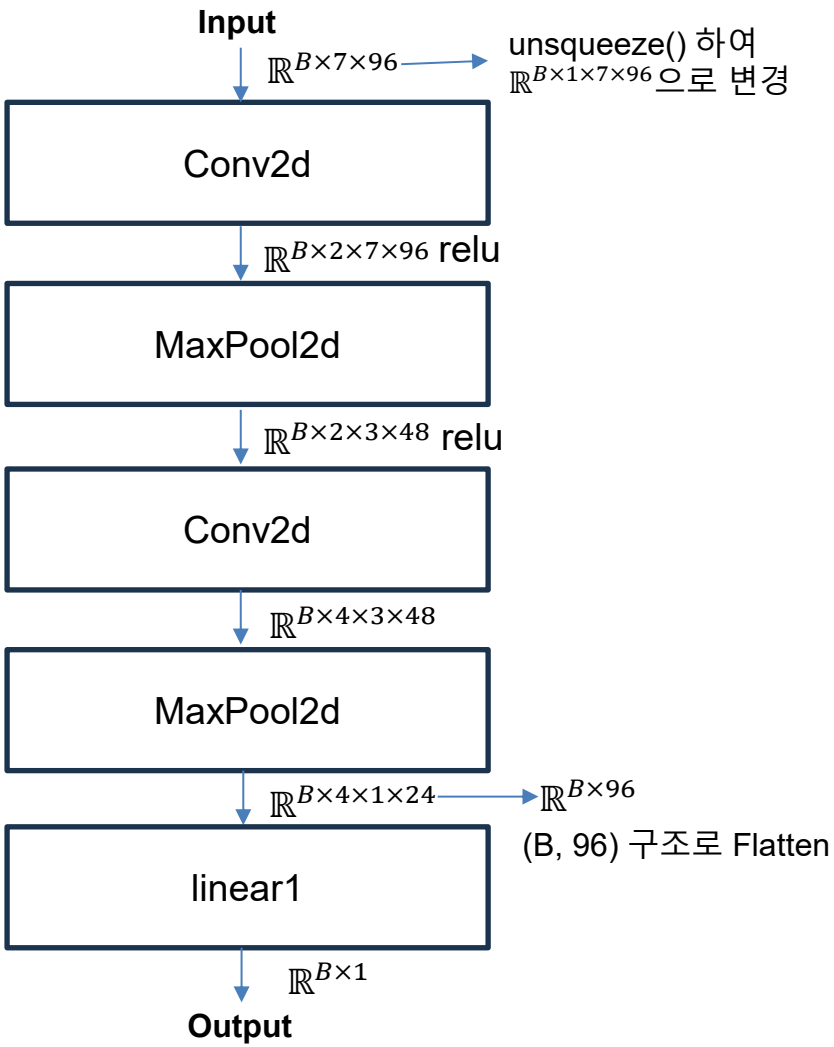
  10_day_return_rate / 20_day_return_rate
- 2D CNN 모델 Label Scale 1.0과 2.0 실험 비교
- 랜덤 종목 2개, 랜덤 샘플 시각화 비교

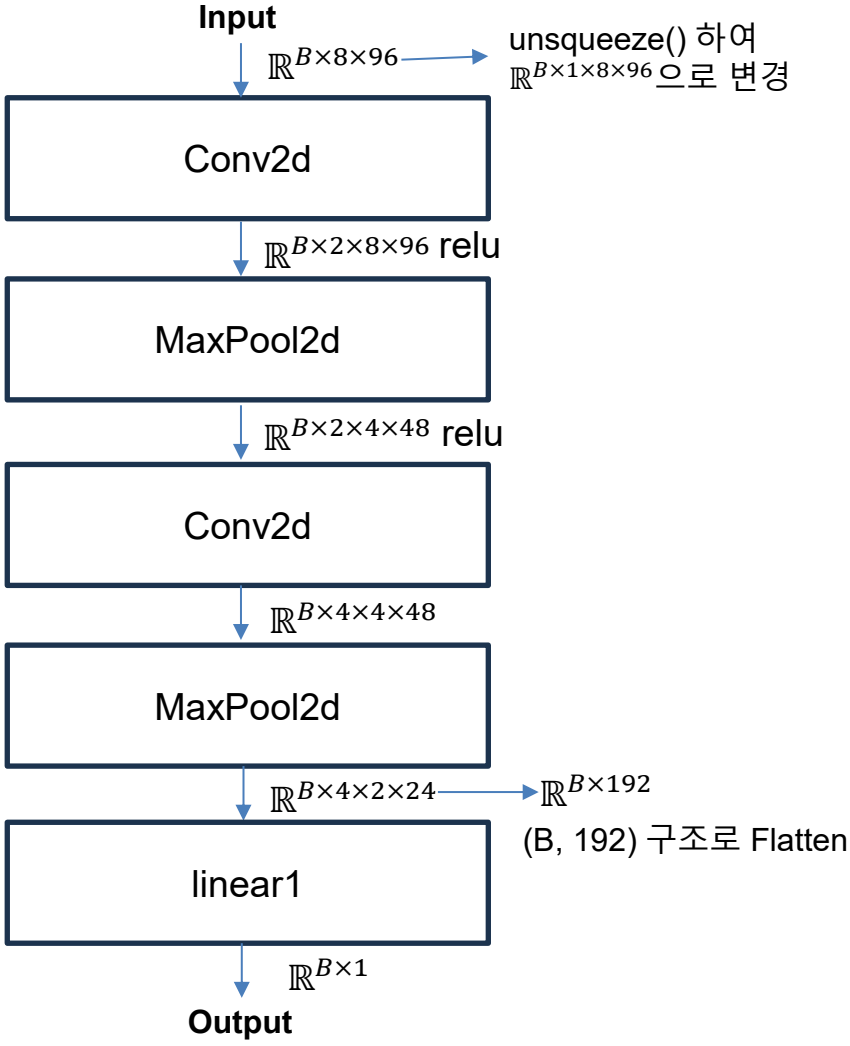
모델 구조

FAIR Lab

KWANGWOON UNIVERSITY

입력 차원에 따른 모델(7개, 8개 모델)





데이터셋 및 실험 세팅

- 사용한 종목: 전체 종목
- 데이터 기간 : 2009-12-31 ~ 2023-12-31
  데이터 분할: Train, Valid, Test 6:2:2
- 전처리: 결측치 제거 및 np.inf 삭제
- 정규화: StandardScaler

*B : Batch Size

Learning rate	0.00001
Epoch	100 / 300
Batch size	64
Loss function	MSE Loss
Sequence Length	96
input_feature	7 / 8
Output_window	1

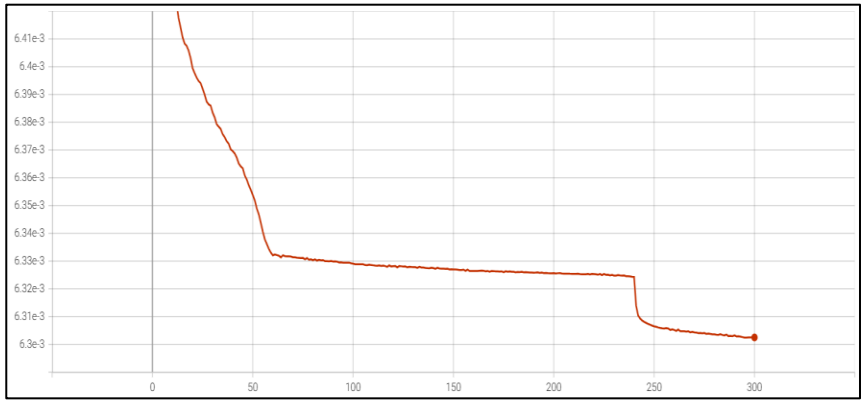
제안 모델

FAIR Lab

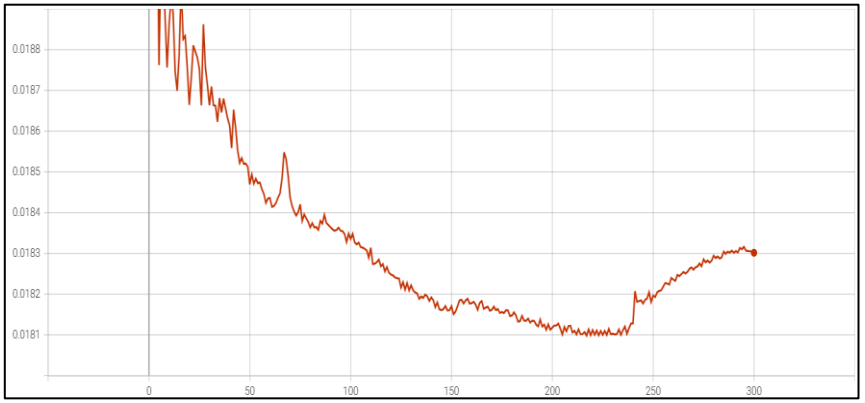
KWANGWOON UNIVERSITY

저번 주 모델 1D CNN

train loss



validation loss



x: epoch y: loss

result

Test MSE
0.437039

epoch	Valid Loss
205	0.0181

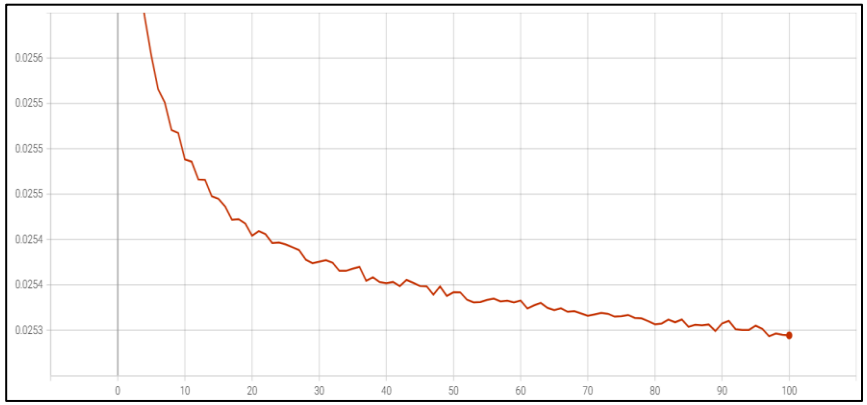
제안 모델

FAIR Lab

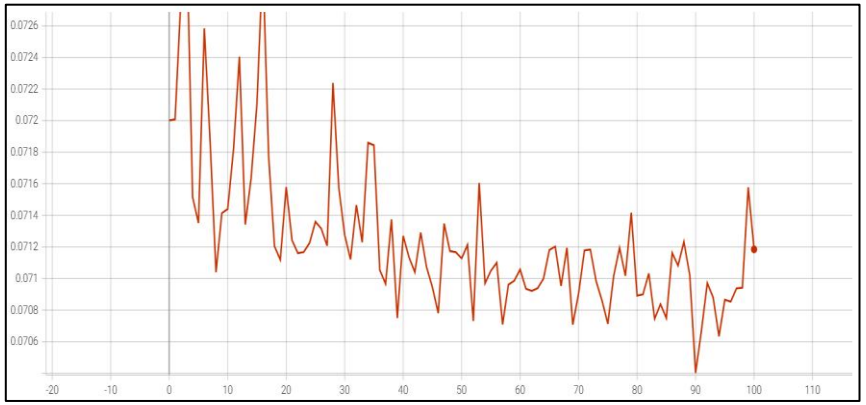
KWANGWOON UNIVERSITY

이번 주 모델 2D CNN + label scale 2.0

train loss



validation loss



x: epoch y: loss

result

Test MSE
1.741656

Epoch	Valid Loss
90	0.0704

제안 모델

FAIR Lab

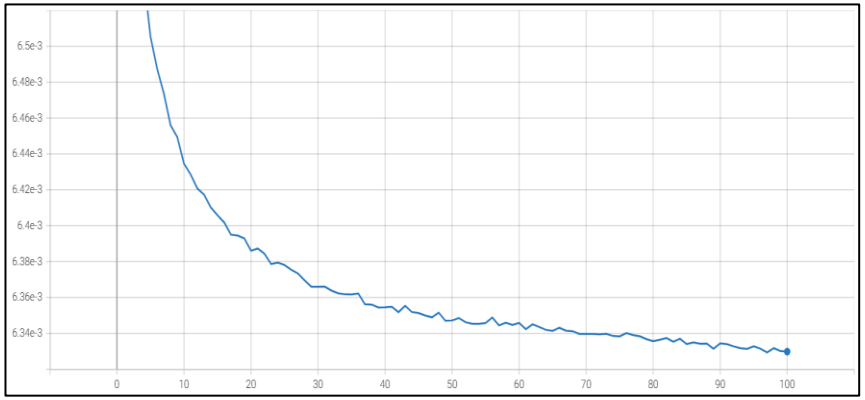
KWANGWOON UNIVERSITY

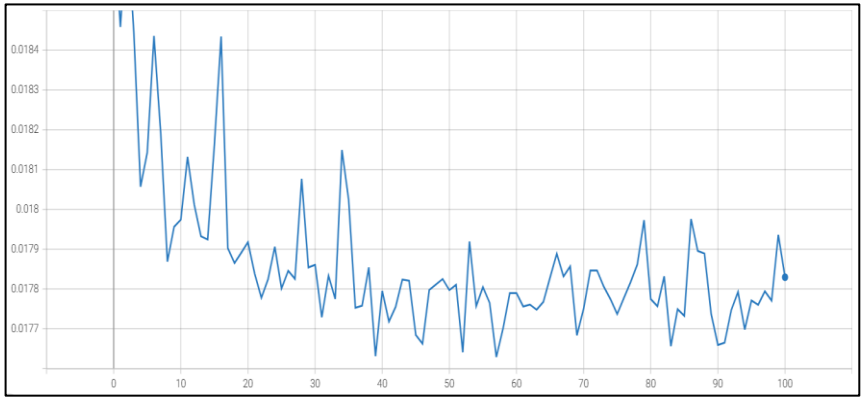
이번 주 모델 2D CNN label scale 실험 결과

train loss

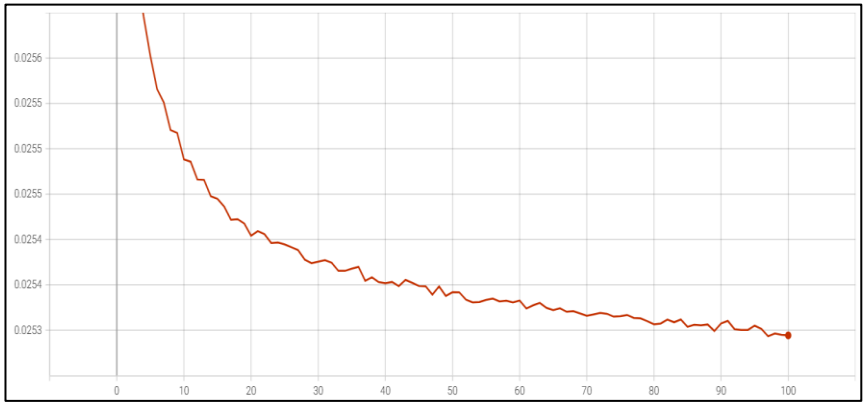
validation loss

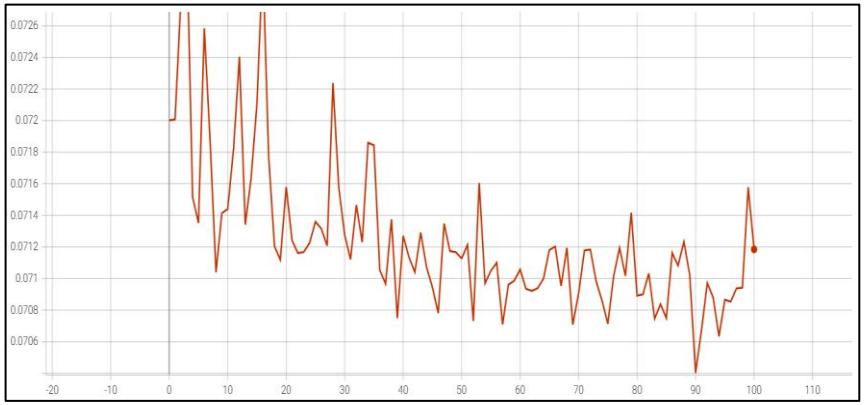
①





②





Test MSE ①	Test MSE ②
0.435675	1.741656

x: epoch y: loss

① → 2D CNN label scale 1.0
② → 2D CNN label scale 2.0

제안 모델

FAIR Lab

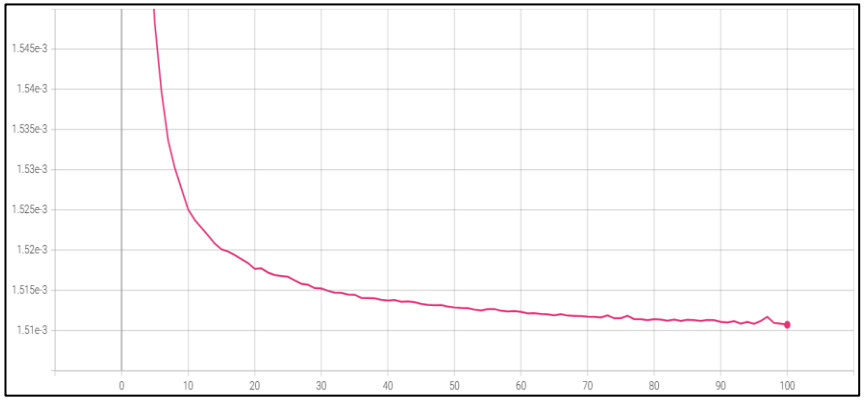
KWANGWOON UNIVERSITY

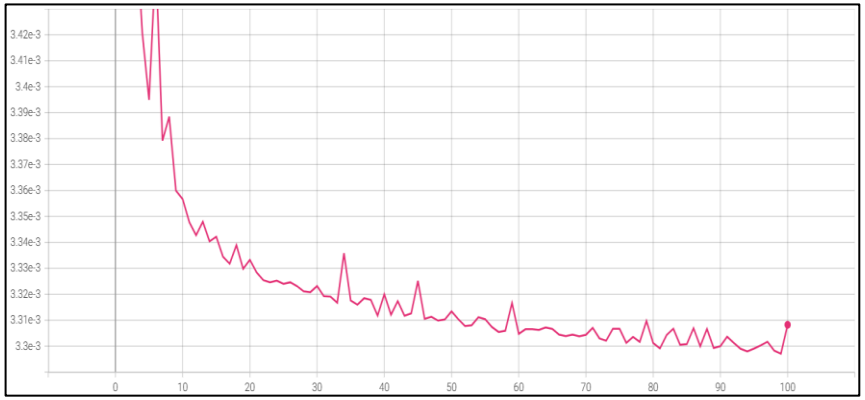
Label return day 1, 5 실험 결과

train loss

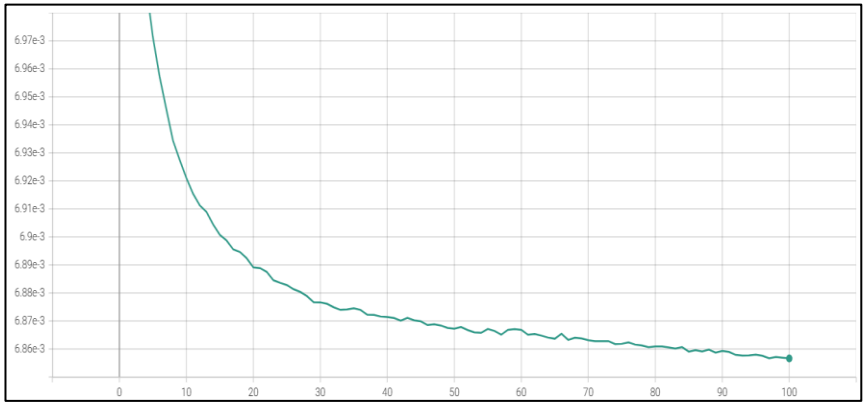
validation loss

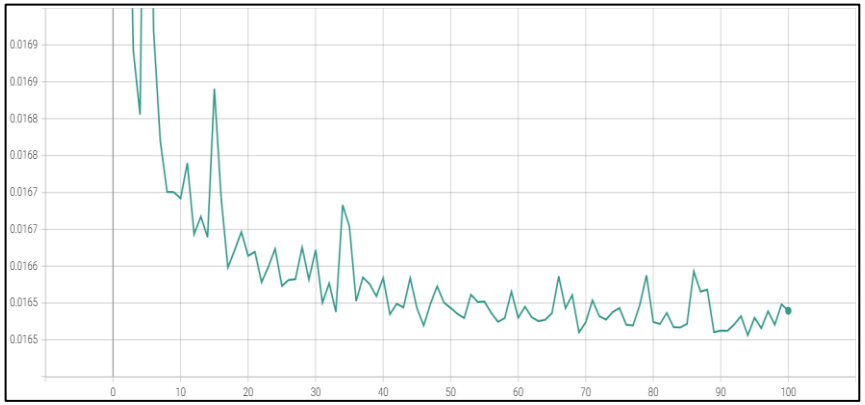
①





②





Test MSE ①	Test MSE ②
0.002655	0.018971

x: epoch y: loss

① → Label return day 1
② → Label return day 5

제안 모델

FAIR Lab

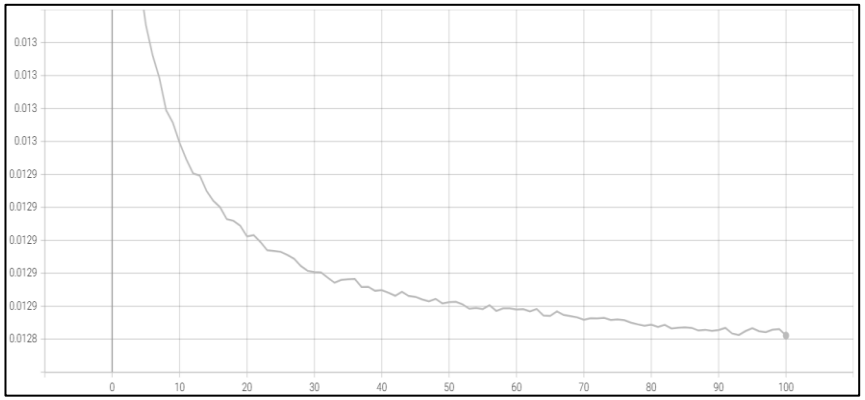
KWANGWOON UNIVERSITY

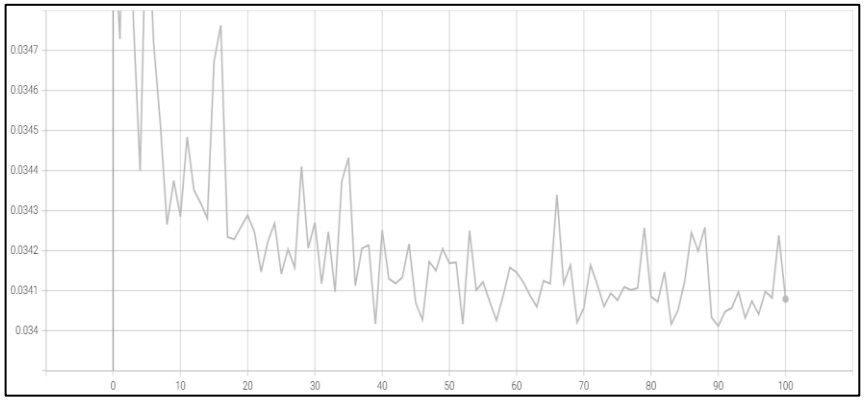
Label return day 10, 20 실험 결과

train loss

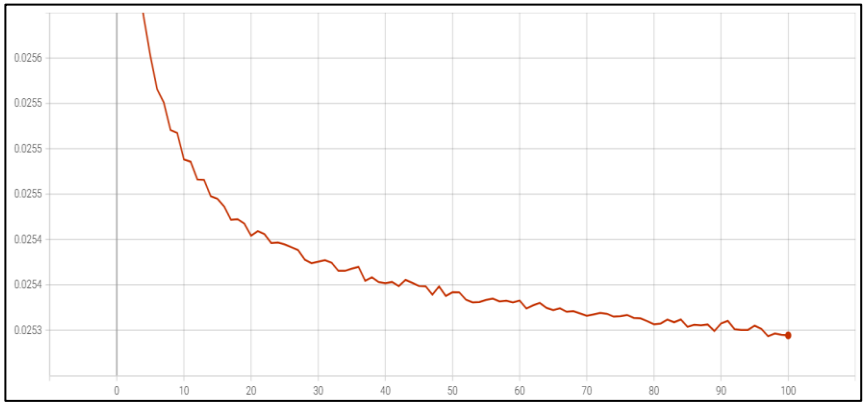
validation loss

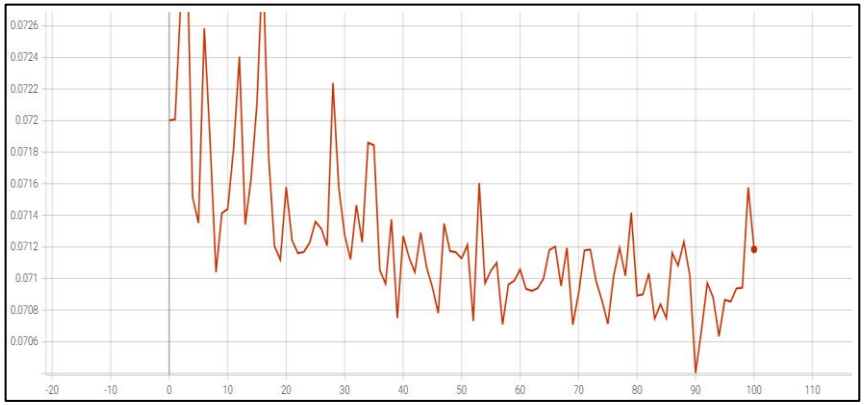
①





②





Test MSE ①	Test MSE ②
0.079227	1.741656

x: epoch y: loss

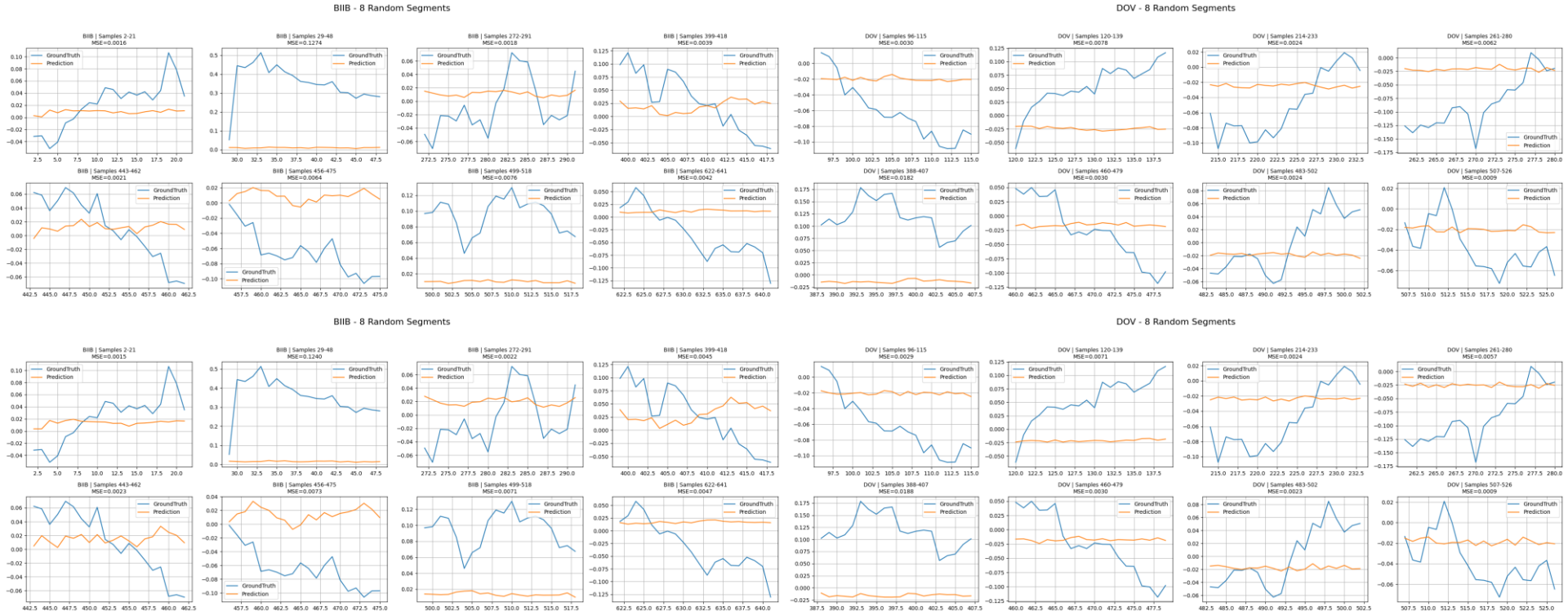
① → Label return day 10
② → Label return day 20

제안 모델

FAIR Lab

KWANGWOON UNIVERSITY

Label scale 1.0, 2.0 시각화 결과



x: 각 샘플의 index y: 해당 샘플의 값
종목 2가지 (BIIB.csv, DOV.csv)

정리

FAIR Lab

KWANGWOON UNIVERSITY

실험 결과 정리

- 2D CNN Label scale 1.0 / 2.0 비교 결과
  - Label scale 2.0은 1.0과 비교하기 위해 4로 나누어 1.0일 때의 성능으로 환산 → 0.435414

    1.0 Test Loss < 2.0 Test Loss이기 때문에 이번주 실험은 2.0으로 진행하였습니다.
- Label return day 1, 5, 10, 20 비교 결과
  - 기간이 길어질수록 예측이 어려워져 Test Loss 성능이 점점 안 좋게 나오고 있습니다.
- 샘플 시각화 결과
  - 예측 값의 경우 대부분이 flat하게 나와 정답 값을 거의 따라가지 못하는 모습입니다.

Label scale	Test MSE	학습 소요 시간
1.0	0.435675	6시간 45분 30초
2.0	1.741656	6시간 44분 50초

Day1 Test MSE	Day5 Test MSE	Day10 Test MSE	Day20 Test MSE
0.002655	0.018971	0.079227	1.741656

이후 계획

- 하이퍼파라미터 변경 후 실험 진행
- 모델 구조 변경

  → BatchNorm2d 추가 또는 Dropout 추가 후 실험 진행