

[사회현안 문제 7번]

지상 관측과 레이더 관측 자료를 활용한
주요 지점별 강우량 예측 모델 개발

팀명 : 비전

프로젝트명 : 레인보우



1. 프로젝트 개요

프로젝트명

Rain + 保佑 (保 보호할 보, 佑 도울 우) = 레인보우

1. 프로젝트 개요

문제 정의

- 학습 데이터 종류 : ① 대상 지점의 지상 관측 데이터 (시계열)
② 레이더로 관측한 한반도 영역의 강우분포 데이터 (이미지)
- 학습 기간 : 9개 년도 (2012~2020년)의 6~9월 (1시간 간격)
- 예측 시점 : 2021년 6~9월의 854개 시점
- 예측 장소 : 10개 (서울, 원주, 동해, 대전, 안동, 전주, 대구, 광주, 부산, 여수)
- 예측 값 : 854개 시점에서 10개 장소에 대한 미래 1, 2, 3, 4, 5, 6시간 후의 시간당 강우량
- 성능 평가 : 10개 지점의 평균절대오차 (MAE), 상관계수 (r)의 평균과 분산

2. 활용 데이터

1) 지상 기상 관측자료

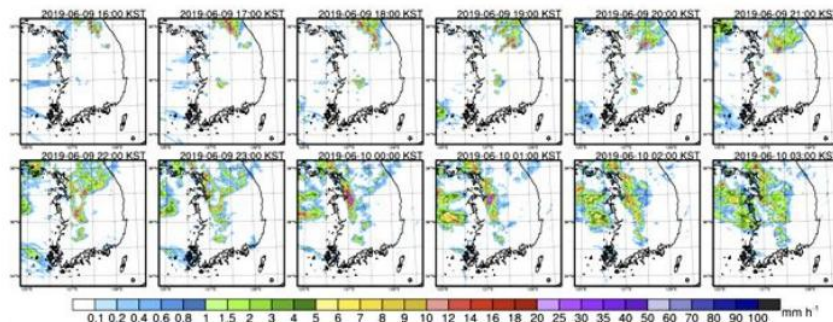
- 파일형식: txt
- 기간/용량: 2012~2020년 6~9월 (1시간 간격) / 약 1.2GB
- 내용: 지점번호 (STN), 기온 (TA), 풍향 (WD), 풍속 (WS), HM (습도), RN_HR1 (시간당 강우량), PA (현지기압), PS (해면기압)

| # | YYMMDDHHMI | STN | TA | WD | WS | RN_DAY | RN_HR1 | HM | PA | PS |
|--------------|------------|------|-------|-----|-----|--------|--------|--------|--------|-----|
| # | KST | ID | C | deg | m/s | mm | mm | % | hPa | hPa |
| 201203010000 | 12 | 2.8 | 41.7 | 2.5 | 0.0 | 0.0 | 63.6 | 1014.0 | 1021.5 | |
| 201203010000 | 13 | 9.0 | 25.3 | 3.1 | 0.0 | 0.0 | 76.6 | 1013.4 | 1019.7 | |
| 201203010000 | 59 | 1.7 | 42.1 | 1.8 | 0.0 | 0.0 | 79.7 | 1005.4 | 1021.4 | |
| 201203010000 | 82 | 4.0 | 358.6 | 1.2 | 0.0 | 0.0 | 47.5 | 1012.7 | 1021.2 | |
| 201203010000 | 83 | 3.3 | 45.5 | 1.9 | 0.0 | 0.0 | 60.0 | 1018.2 | 1021.8 | |
| 201203010000 | 84 | 4.9 | 284.3 | 0.3 | 0.0 | 0.0 | 48.2 | 1013.2 | 1021.8 | |
| 201203010000 | 85 | 2.1 | 117.8 | 0.5 | 0.0 | 0.0 | 77.3 | 1018.7 | 1021.6 | |
| 201203010000 | 86 | 2.3 | 33.1 | 0.7 | 0.0 | 0.0 | 89.7 | 1016.4 | 1021.2 | |
| 201203010000 | 87 | 7.3 | 80.5 | 2.6 | 0.0 | 0.0 | 69.8 | 1013.1 | 1021.1 | |
| 201203010000 | 89 | 11.4 | 45.0 | 3.3 | 0.0 | 0.0 | 72.6 | 1016.9 | 1019.1 | |
| 201203010000 | 90 | 4.4 | 283.3 | 2.0 | 0.0 | 0.0 | 76.0 | 1018.0 | 1020.3 | |
| 201203010000 | 92 | 4.3 | 270.0 | 1.0 | 0.0 | 0.0 | -99.0 | 1011.7 | 1020.7 | |

2) 레이더 합성자료 (Post-processed)

- 파일형식: netCDF
- 기간/용량: 2012~2020년 6~9월 (1시간 간격) / 약 5.0GB
- 격자구성: 128(H) × 128(W), 공간해상도 4 km
- 내용: rain1h (시간당 강우분포, mm/h),

XLAT (위도, degrees), XLON (경도, degrees)

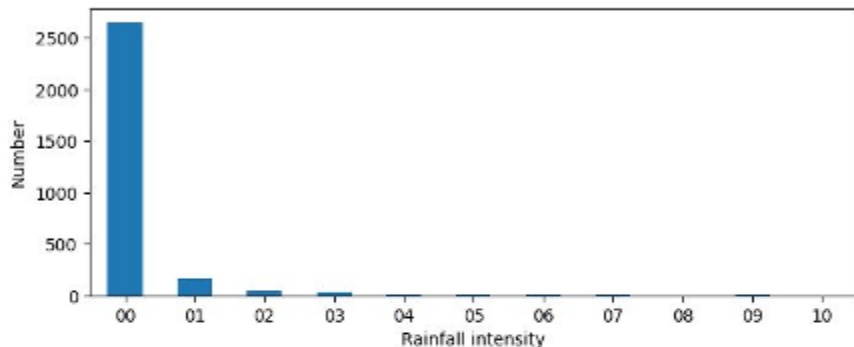


3. 모델 개발 방법

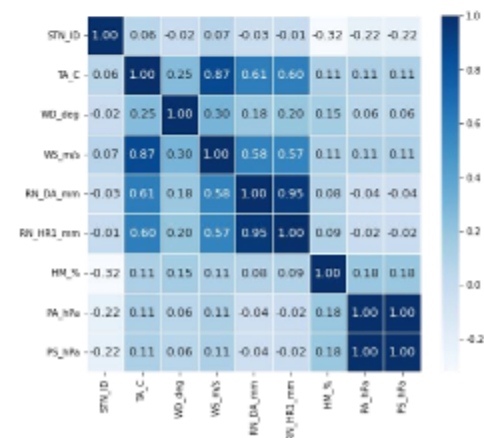
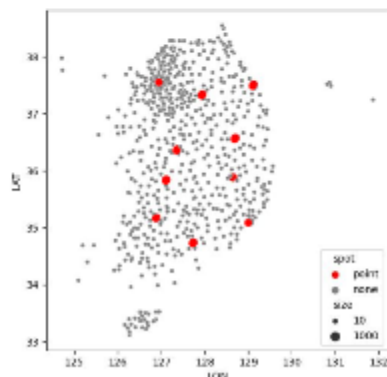
1) 데이터 분석 (EDA)

| | 관측소 수 | 결측치 (-99 value) | 강우량 누락 |
|------|-------|-----------------|--------|
| 2012 | 653 | 3.1% | 0% |
| 2013 | 656 | 3.3% | 8.3% |
| 2014 | 657 | 3.3% | 0% |
| 2015 | 664 | 3.6% | 3.3% |
| 2016 | 676 | 3.2% | 3.4% |
| 2017 | 684 | 3.8% | 3.4% |
| 2018 | 684 | 3.7% | 3.3% |
| 2019 | 700 | 3.7% | 3.3% |
| 2020 | 709 | 3.7% | 3.3% |

강우량 분포 예시



전체 지점 및 10개 지점 위치 기상 정보간 상관성 분석

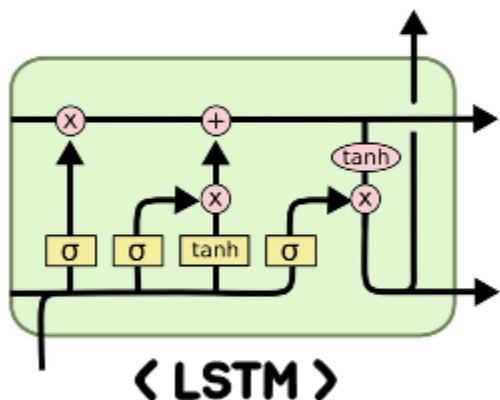


2) 지점 강우량 예측 모델 데이터 전처리 :

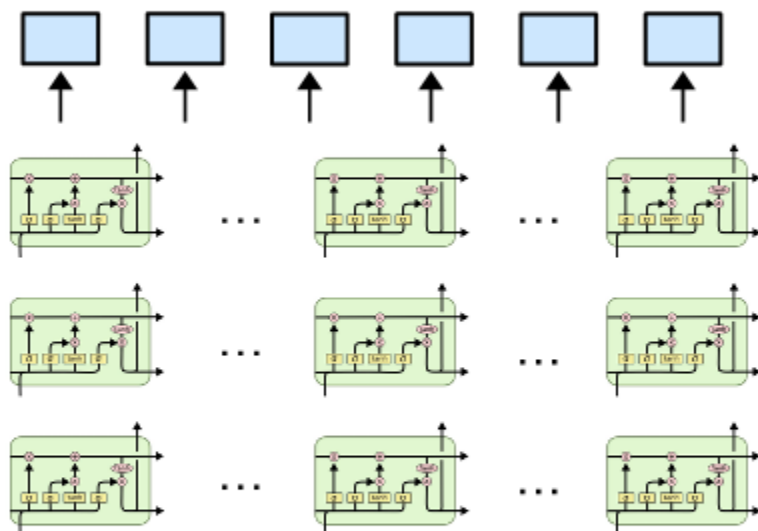


1. 강우량 데이터와 관측소 정보 통합
2. 관측소 정보가 누락된 데이터 제거
3. 결측치 (-99 value)를 0으로 변환
4. 비연속적인 강우량 데이터를 temporal interpolation
5. 9개년 모두에서 존재하는 관측소 데이터 사용

3) 지점 강우량 예측 모델:



미래 6시간 동안의 시간당 강우량



과거 6시간 동안의 시간당 강우량 & 기온

4) 레이더 강우분포 예측 모델을 위한 전처리 :

레이더 합성자료

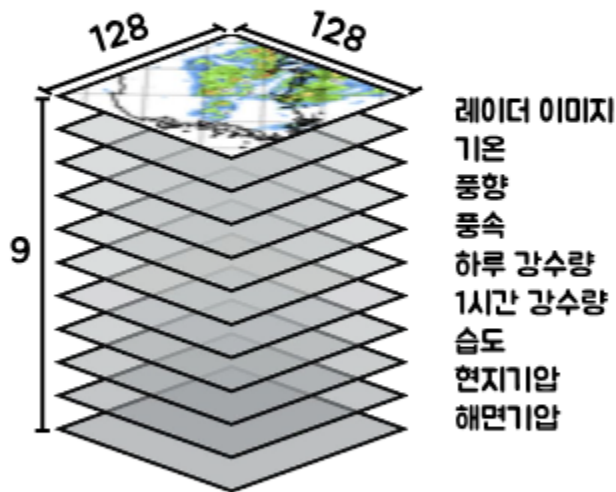
- 누락 값에 대한 temporal interpolation

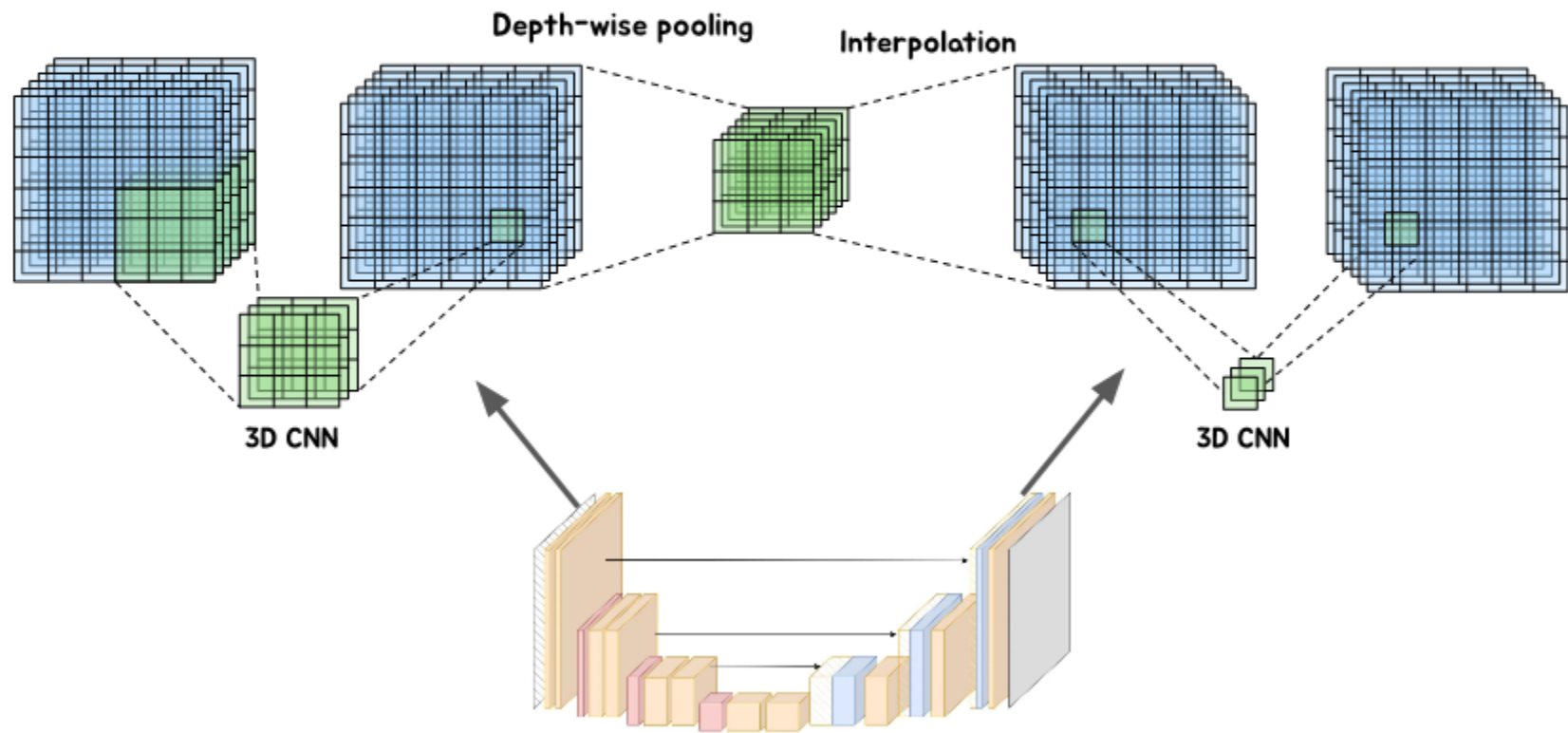
지상 기상 관측자료

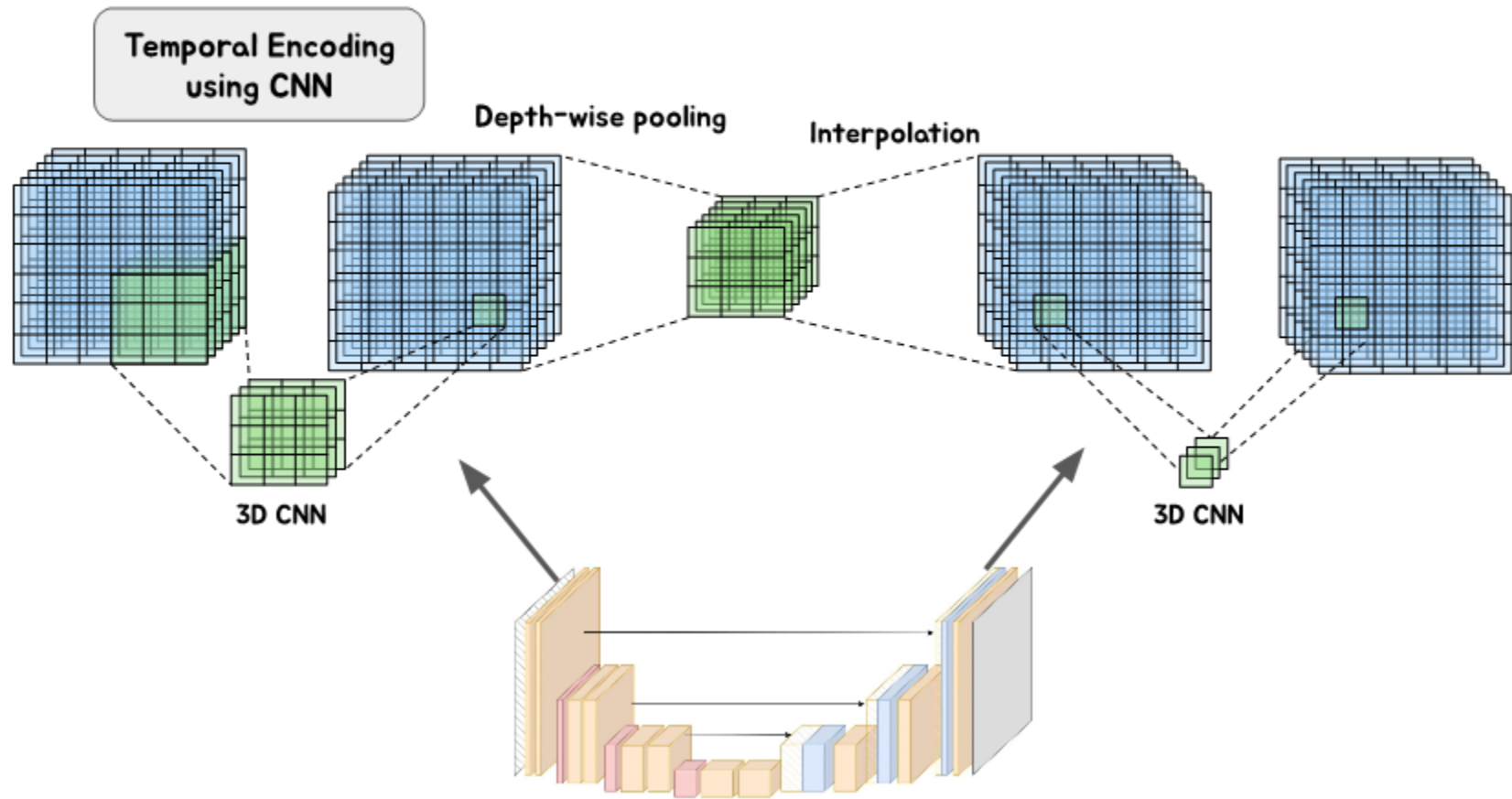
- Spatial interpolation 하여 각 지점의 변수들 이미지화

Multi-modal 통합 데이터

- Sampling: 12시간 평균 강수량 > 0.8mm/hr
- Channel-wise normalization

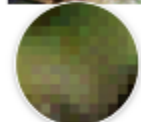
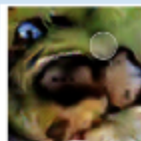
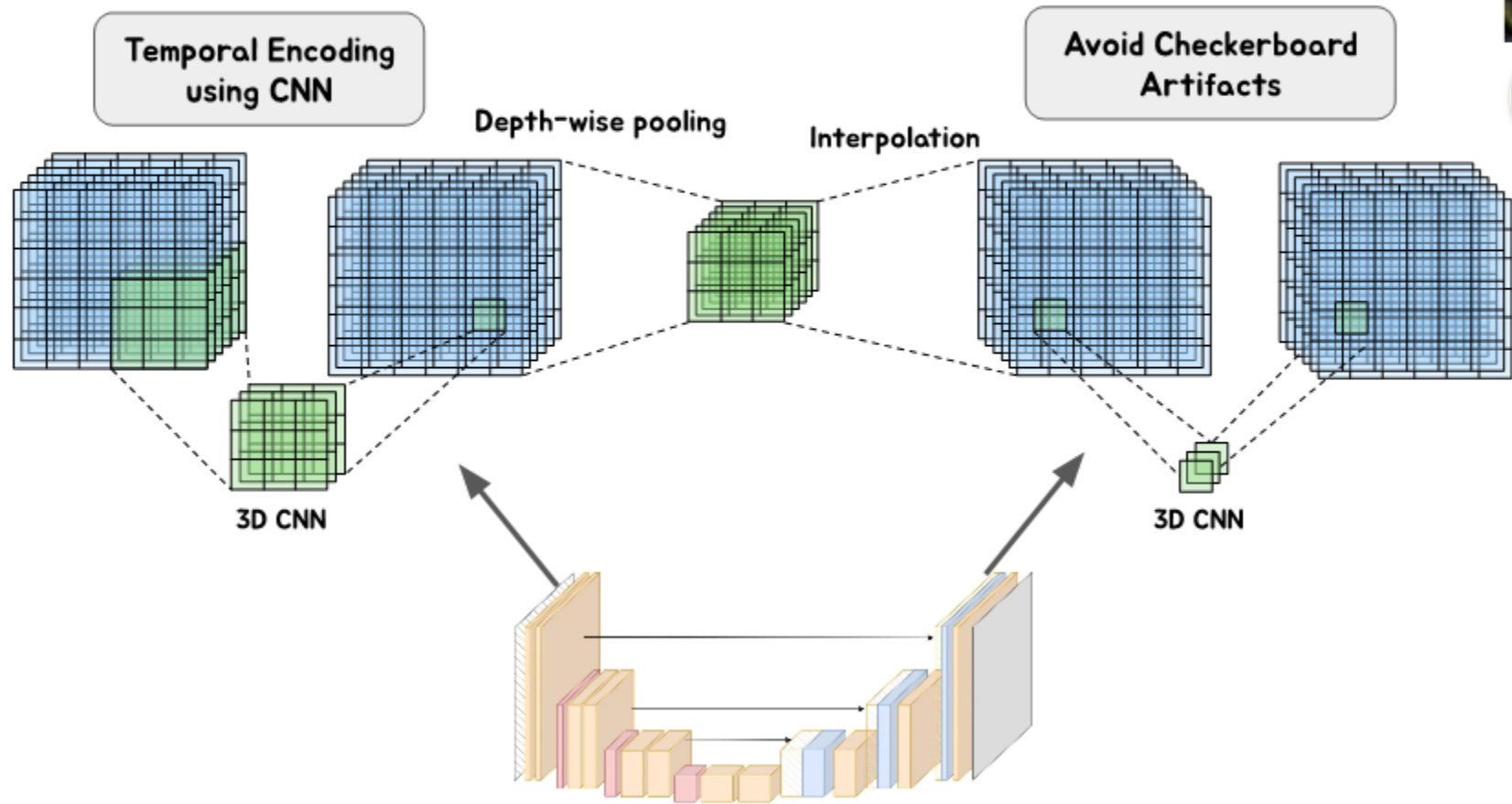






3. 모델 개발 방법 - 레이더 강우 분포

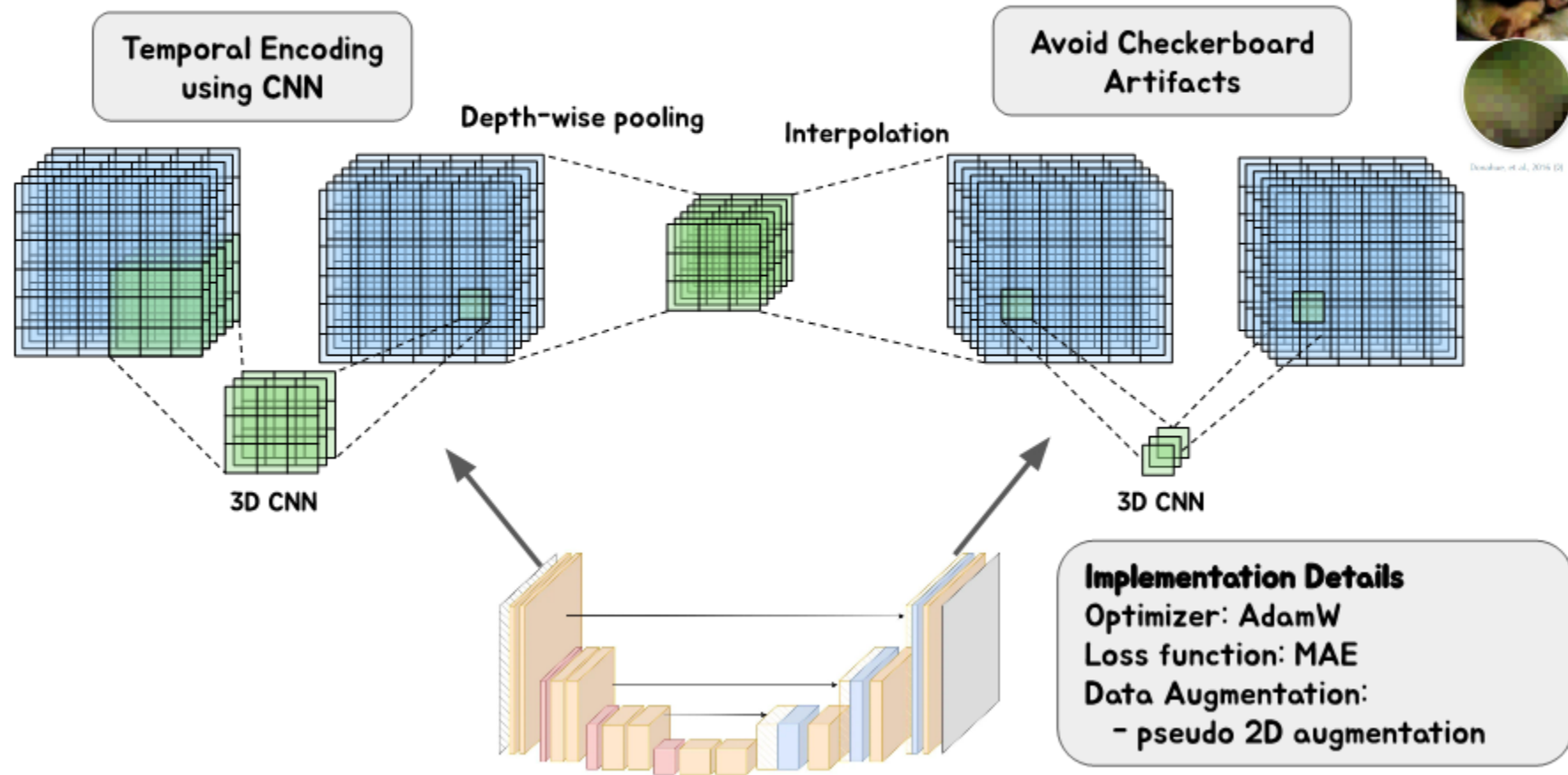
개발 방법 및 시스템 구조 설명



Deledan et al., 2016 01

3. 모델 개발 방법 - 레이더 강우 분포

개발 방법 및 시스템 구조 설명

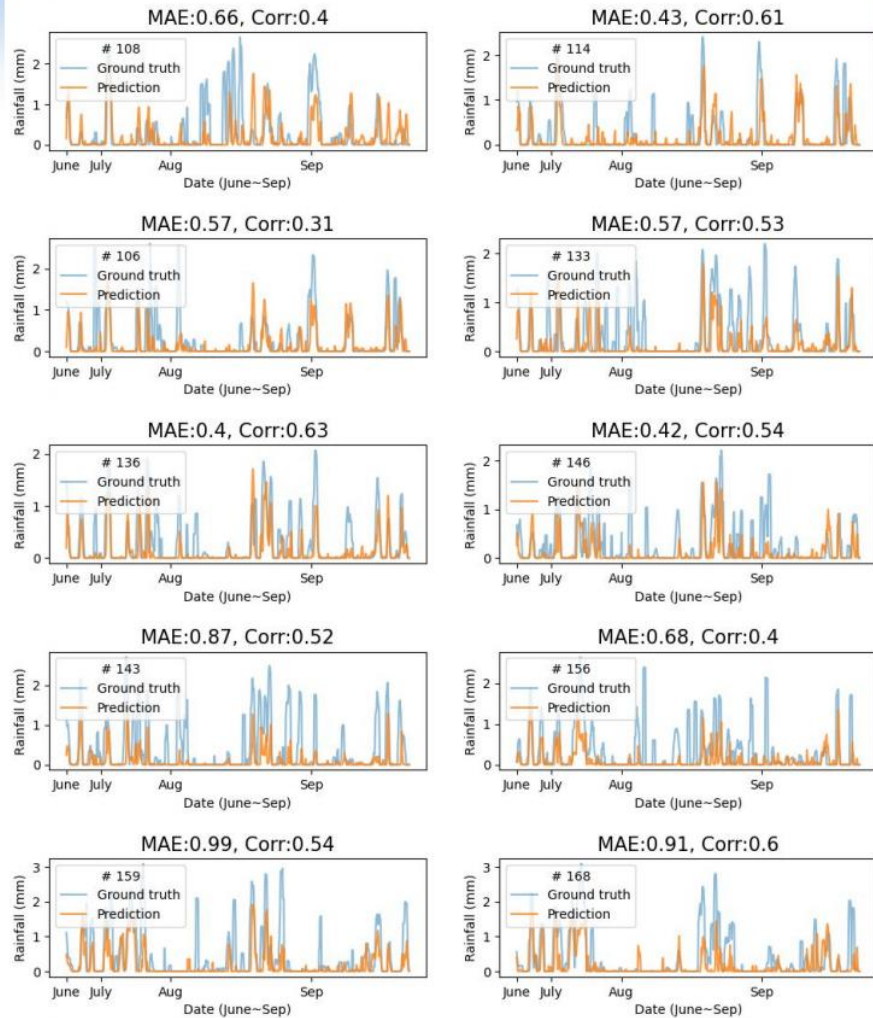


4. 실험 및 평가 - 지점 강우량

| | 평균절대오차 MAE (분산) | 상관계수 (분산) |
|----------|--------------------|--------------|
| Test set | 0.65 (0.04) | 0.51 (0.01) |

| 지점 | 평균절대오차 MAE | 상관계수 |
|----------|------------|------|
| 108 (서울) | 0.66 | 0.40 |
| 106 (동해) | 0.57 | 0.31 |
| 136 (안동) | 0.40 | 0.63 |
| 143 (대구) | 0.87 | 0.52 |
| 159 (부산) | 0.99 | 0.54 |
| 114 (원주) | 0.43 | 0.61 |
| 133 (대전) | 0.57 | 0.53 |
| 146 (전주) | 0.42 | 0.54 |
| 156 (광주) | 0.68 | 0.40 |
| 168 (여수) | 0.91 | 0.60 |

Rainfall Prediction at 2021

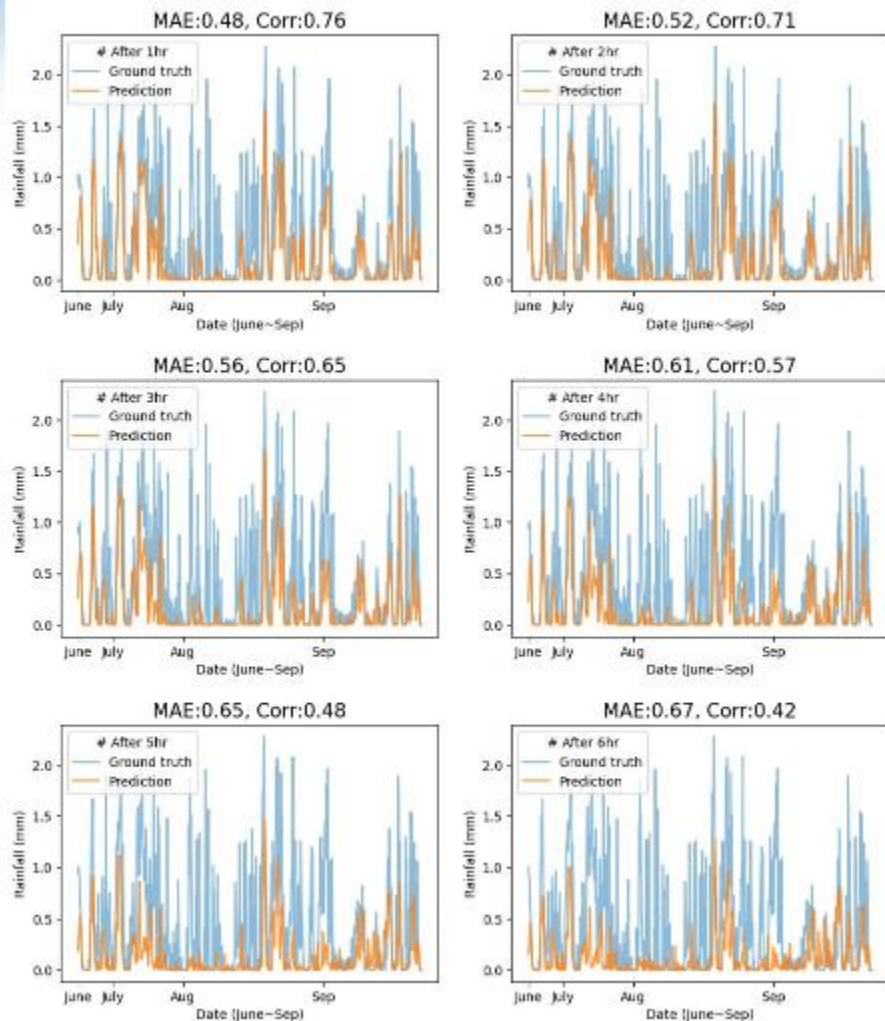


4. 실험 및 평가 - 지점 강우량

| | 평균절대오차 MAE (분산) | 상관계수 (분산) |
|----------|--------------------|--------------|
| Test set | 0.65 (0.04) | 0.51 (0.01) |

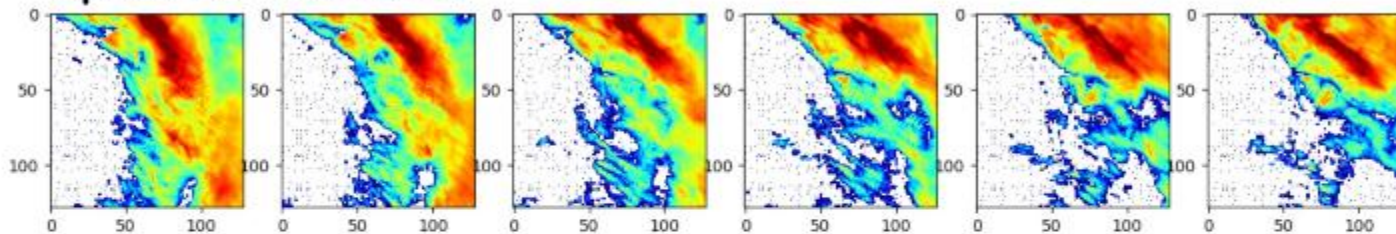
| 예측 시간 | 평균절대오차 MAE | 상관계수 |
|-------|------------|------|
| 1시간 후 | 0.48 | 0.76 |
| 2시간 후 | 0.52 | 0.71 |
| 3시간 후 | 0.56 | 0.65 |
| 4시간 후 | 0.61 | 0.57 |
| 5시간 후 | 0.65 | 0.48 |
| 6시간 후 | 0.67 | 0.42 |

Rainfall Prediction at 2021

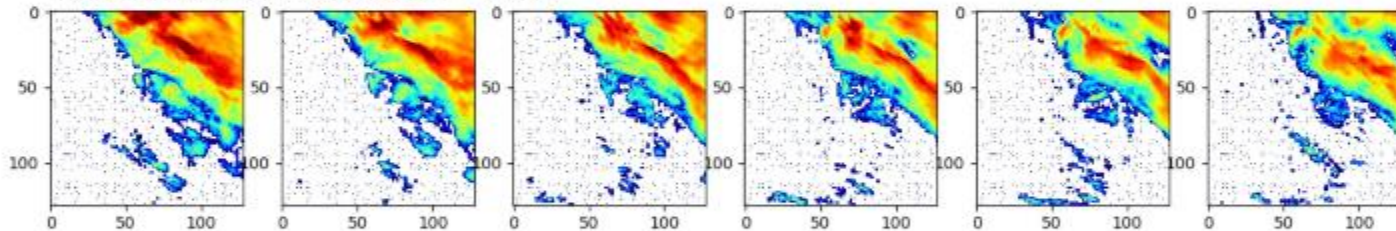


4. 실험 및 평가 - 레이더 강우 분포

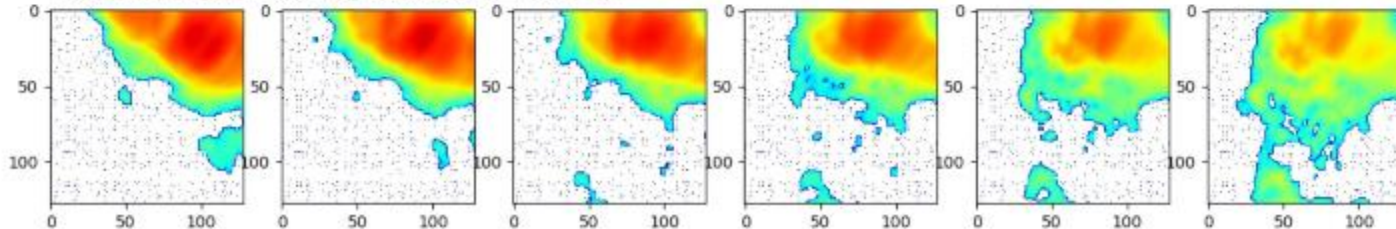
Input data (과거 6시간)



Actual data (미래 6시간 실제)



Predicted data (미래 6시간 예측)



5. 활용 계획 및 기대효과

2023년

[폭우 피해]사망 41명·실종 9명...오송 공평지하차도서 14명 사망

A 기자실 기자 | 안영희 2023.07.10 07:50 | 보도자료

기상청장 "가뭄·홍수 동시 걱정...기후변화로 책임 커져"

등록 2023.10.16 10:22:12 수정 2023.10.16 11:12:44

태그 [이](#) [스](#) [이](#) [가](#)

매년 반복되는 침수 사고, 오송 지하차도 참사 막을 수 없었나?

전주 C 실 기자, 오송 C 기자

2023.08.29

올 여름, 장마철 강수량 역대 '1위'

●기상청 기록본의 결과 발표
장우-전남 평균 강수량 765.5mm
기온 2도 뽕뽕보다 0.8도 높아

2023.09.11 16:26

전주 C 기자 jubkang@jilbo.com

한 달 치 비가 한꺼번에...평년 강수량 이미 넘었다

등록 2023.07.16 16:14 수정 2023.07.16 16:17

2022년

‘극한 집중호우’ 올해만 30차례... “기후변화 맞춰 방재대책 재수립해야”

이미지 기자

입력 2022-06-16 03:00 | 업데이트 2022-06-16 03:00

올 여름 집중호우로 대규모 산사태 위험 우려

저기압-대기 불안정으로 집중호우 발생 확률 높아
산림과학원 14일부터 ‘산사태 예측분석센터’ 가동
(대전·울진=뉴스1) 박찬수 기자 | 2021-05-14 09:33 송고

2021년

11일째 이어진 폭우...

사망 32명·실종 10명·이재민 7600명

기사승인 2020-08-12 01:30:02

2020년

115년 관측 사상 기록적 폭우...사망·실종 잇따라

허정연 기자 | 작성 2022-06-04 07:16 | 조회 11,459

폭우 ‘극한 한반도’, 코로나19보다 더 무섭다

기재여 다른기사 보기 [가](#) [가](#) [f](#) [t](#) [v](#)

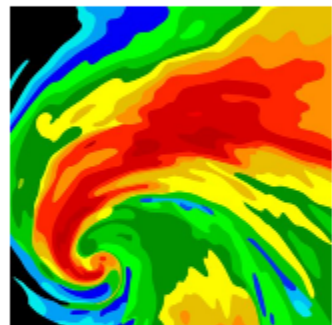
입력 2021.08.01 19:44:00 수정 2022-07-07 19:45:00

"이번 폭우는 시작에 불과" 열받은 한반도 '기후 역습'

이니투데이 | 서울=박광달 기자, 김지은 기자, 기상청 기자

VIEW 14,722 | 2020.08.16 09:00

5. 활용 계획 및 기대효과



레이더 데이터

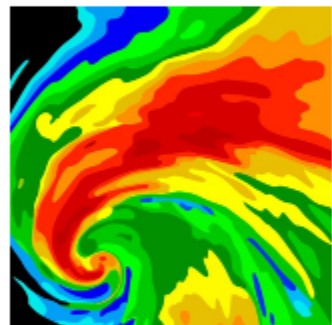


지상 관측 데이터



초단기 강수 예측

5. 활용 계획 및 기대효과



레이더 데이터



지상 관측 데이터



레인보우 어플리케이션

5. 활용 계획 및 기대효과



레인보우 어플리케이션

5. 활용 계획 및 기대효과



공공데이터

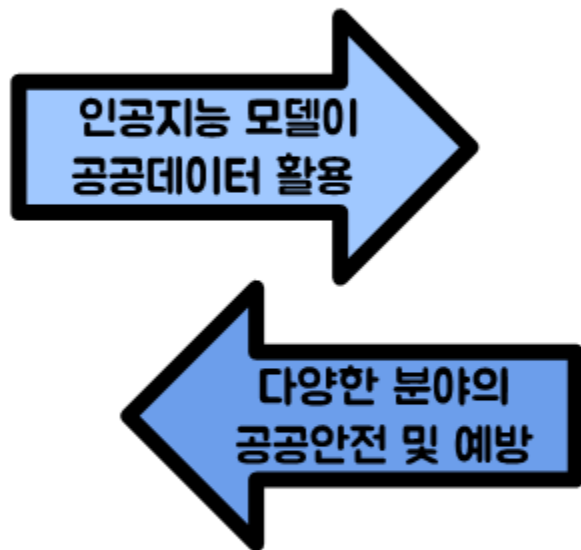


레인보우 어플리케이션

5. 활용 계획 및 기대효과



공공데이터



레인보우 어플리케이션

5. 활용 계획 및 기대효과



5. 활용 계획 및 기대효과



국토교통부

지하안전정보시스템



지반정보

과학기술정보통신부

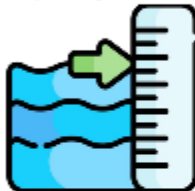
지오빅데이터 오픈플랫폼
Big Data Open Platform



지질정보

환경부

홍수통제소



수위정보

재난안전

교통안전

인공강우

5. 활용 계획 및 기대효과



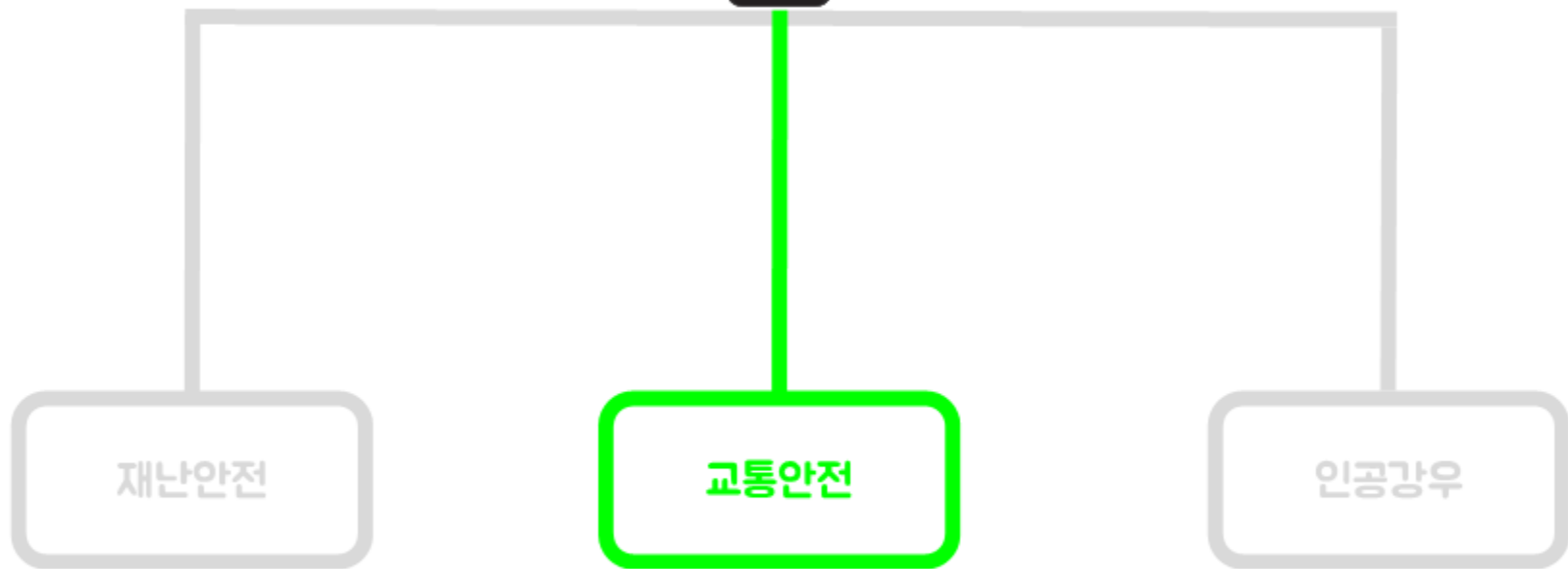
재난안전

교통안전

인공강우



5. 활용 계획 및 기대효과



5. 활용 계획 및 기대효과



5. 활용 계획 및 기대효과



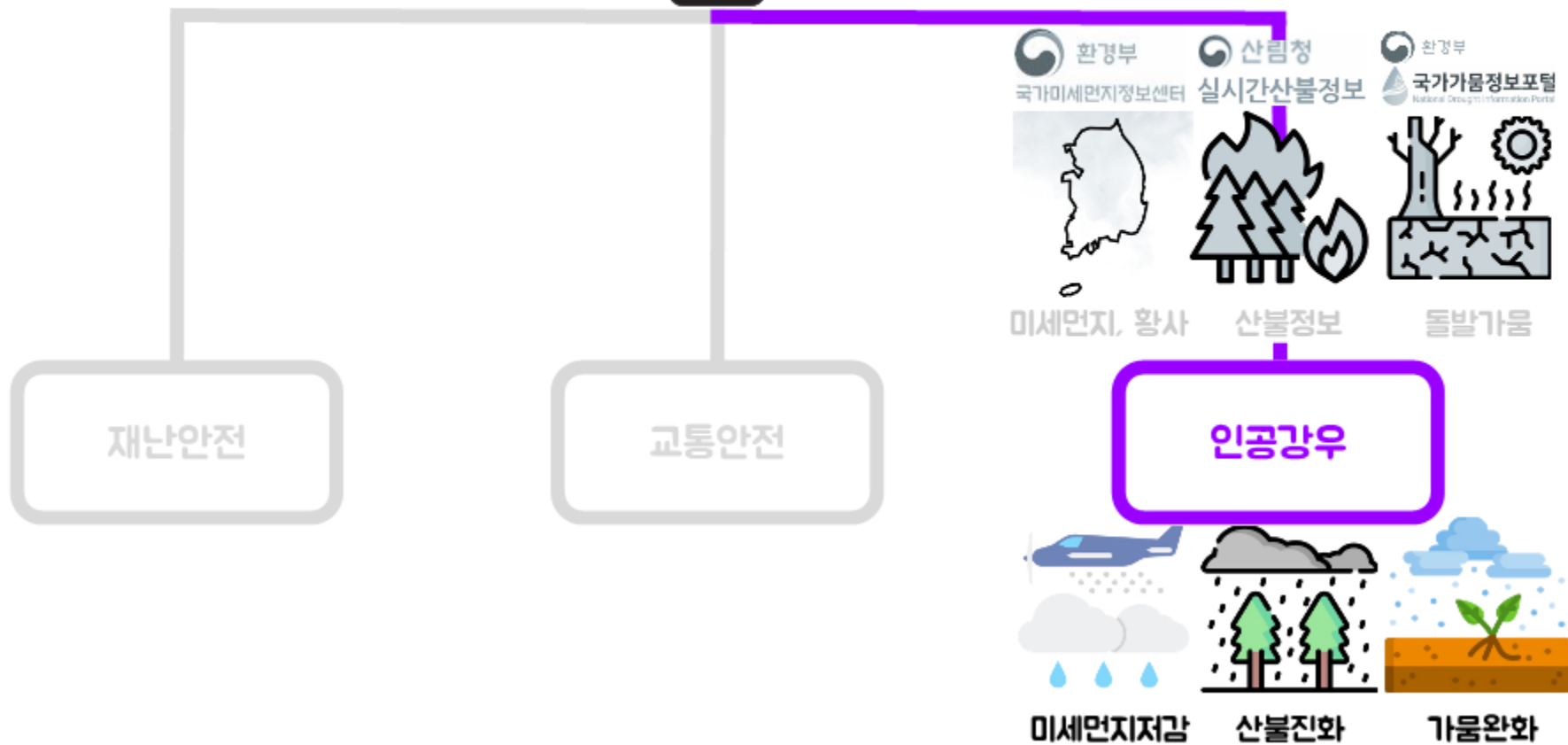
5. 활용 계획 및 기대효과



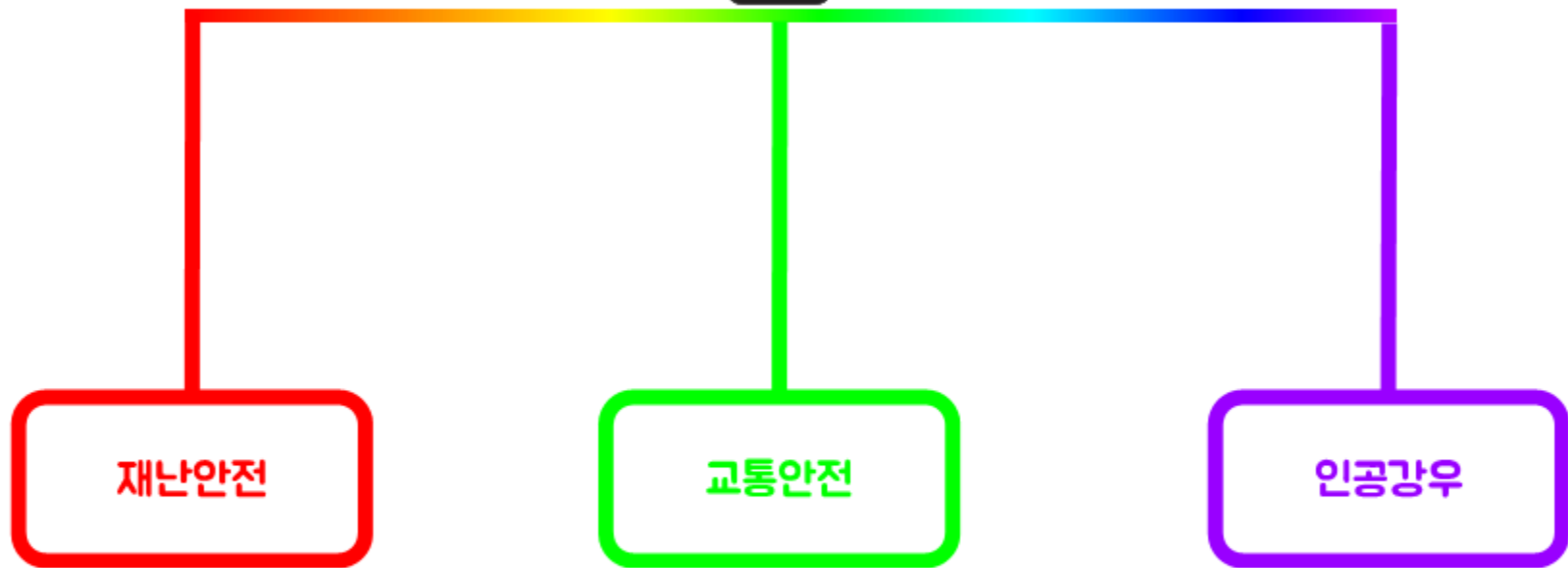
5. 활용 계획 및 기대효과



5. 활용 계획 및 기대효과



5. 활용 계획 및 기대효과



5. 활용 계획 및 기대효과



5. 활용 계획 및 기대효과



재난안전

교통안전

인공강우



도시침수예방



산사태대비



홍수방지



도로안전



항공안전



해양안전



미세먼지저감



산불진화



가뭄완화

5. 활용 계획 및 기대효과



재난안전

교통안전

인공강우



도시침수예방

산사태대비

홍수방지

도로안전

항공안전

해양안전

미세먼지저감

산불진화

가뭄완화

6. 시연

```
[2]: import os
import os.path as osp

from tqdm.auto import tqdm
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn import model_selection

import torch
from torch.cuda import amp
from torch import nn
import torch.nn.functional as F
from torch.utils.data import Dataset, DataLoader
import torchvision.utils as vutils
from torchvision.transforms import CenterCrop

import datetime
from torch.optim.lr_scheduler import (
    CosineAnnealingWarmRestarts,
    ReduceLROnPlateau,
)
from matplotlib.colors import LogNorm

[3]: # 3D CNN expects: McCaDabW
class MultimodalDataset_3D:
    def __init__(self, npy_path_list, time_step=6):
        self.npy_path_list = npy_path_list
        self.time_step = time_step

    def __len__(self):
        return len(self.npy_path_list)

    def __getitem__(self, idx):
        # T x C x W x H
        day_npy = np.load(self.npy_path_list[idx])

        input_npy = day_npy[self.time_step:,:,:]
        label_npy = day_npy[self.time_step:self.time_step+2,0,:,:]
        label_npy = np.expand_dims(label_npy, axis=1)

        max_rain = 100
        input_npy[:,0,0,:] = (np.clip(input_npy[:,0,0,:], a_min=0, a_max=max_rain))/max_rain

        input_npy = np.transpose(input_npy, (1, 0, 2, 3))
        label_npy = np.transpose(label_npy, (1, 0, 2, 3))
        # CaDabW
        input_npy = torch.tensor(input_npy, dtype = torch.float32)
        label_npy = torch.tensor(label_npy, dtype = torch.float32)
        return input_npy, label_npy

class RadarDataset_3D:
    def __init__(self, npy_path_list, time_step=6):
        self.npy_path_list = npy_path_list
        self.time_step = time_step

    def __len__(self):
        return len(self.npy_path_list)

    def __getitem__(self, idx):
        # T x W x H
        day_npy = np.load(self.npy_path_list[idx])

        input_npy = day_npy[self.time_step:,:]
        label_npy = day_npy[self.time_step:self.time_step+2,0,:]
```

감사합니다.

팀명 : 비전

프로젝트명 : 레인보우

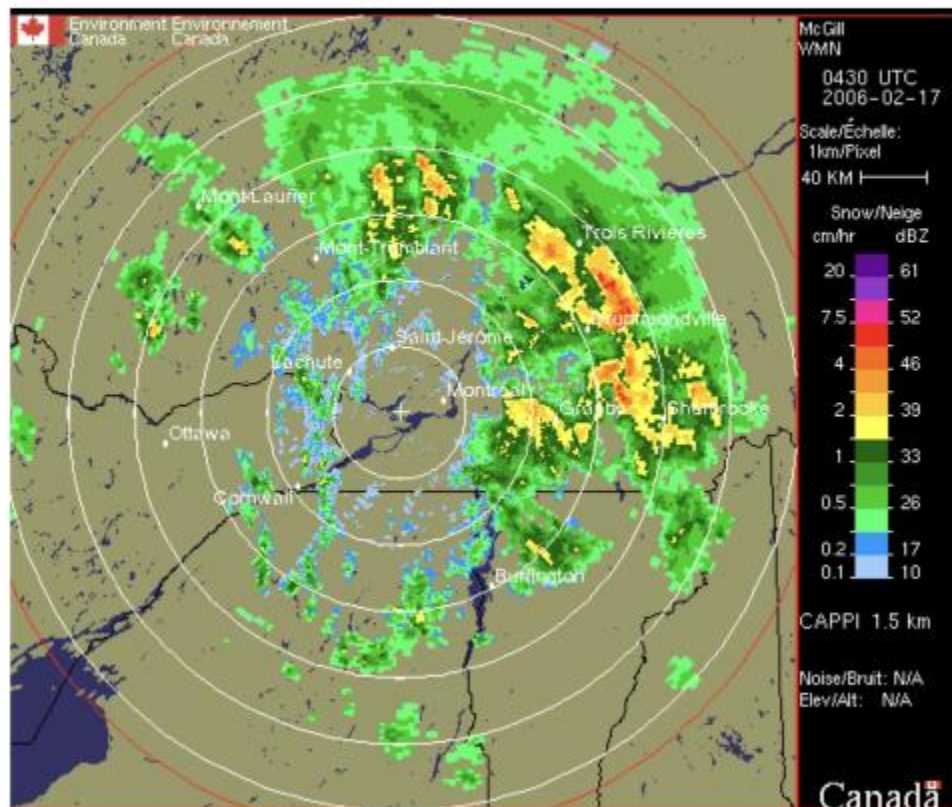


질의응답용 슬라이드

문제 정의

- 대상 지점의 지상 관측 데이터(시계열)와 레이더로 관측한 한반도 영역의 강우분포 데이터(이미지)를 함께 활용하여, **각 지점에 대한 미래 강우량 (F01~F06)을 예측**
- 2012~2020년 6, 7, 8, 9월 기간에 대해서 학습한 모델을 이용하여, 2021년 6, 7, 8, 9월 기간의 강우량을 예측 (상세 예측 대상 기간은 추후 제시)
- 예측 대상 지점: 서울(108), 원주(114), 동해(106), 대전(133), 안동(136), 전주(146), 대구(143), 광주(156), 부산(159), 여수(168), 이상 10개 지점 (괄호는 station ID)

질의응답용 슬라이드



$$R = \left(\frac{10^{\frac{Z}{10}}}{200} \right)^{0.625}$$

R: Estimated rainfall rate (mm/h)

Z: Reflectivity (dBZ)

질의응답용 슬라이드

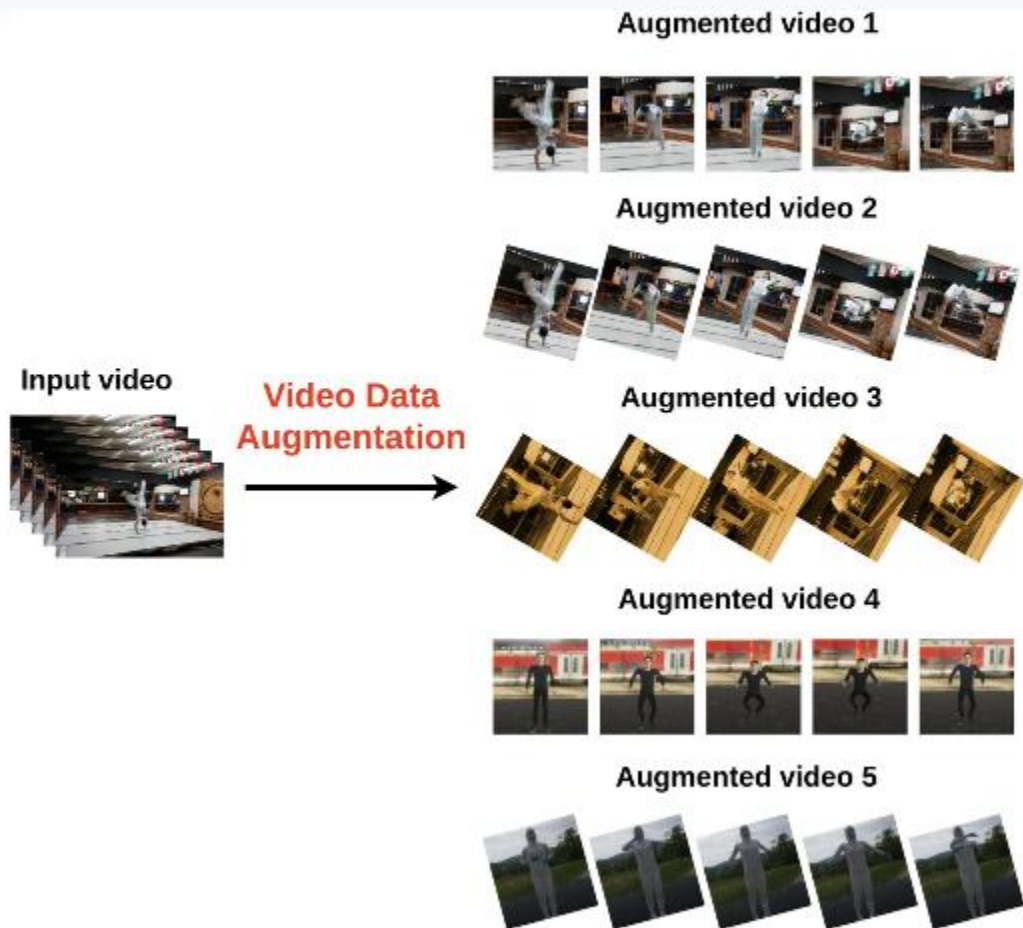


Table 2. Variables provided by the constructed dataset

| Data type | | Variables (unit) | Vertical levels |
|------------|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Radar | | 1-hour integrated rainrate (mm) | Ground |
| MAPLE | | Estimated rainrate (mm h ⁻¹) | Ground |
| AWS & ASOS | | 10-min accumulated rainfall (mm), sea-level pressure (hPa), surface pressure (hPa), temperature (°C), relative humidity (%) | Ground |
| WRF | 2-D | 1-hour accumulated rainfall (mm), sea-level pressure (hPa), terrain height (m) | Ground |
| | | u-wind (m s ⁻¹), v-wind (m s ⁻¹) | 10 m |
| | | temperature (K), water vapor mixing ratio (kg kg ⁻¹) | 2 m |
| | 3-D | pressure (hPa), u-wind (m s ⁻¹), v-wind (m s ⁻¹), w-wind (m s ⁻¹), temperature (K), water vapor mixing ratio (kg kg ⁻¹) | 1, 2, 3, 5, and 10 km |

MAPLE, McGill algorithm for precipitation nowcasting by Lagrangian extrapolation; AWS, automatic weather station; ASOS, Automated Synoptic Observing System; WRF, Weather Research and Forecasting.

질의응답용 슬라이드



Size of the input patch containing satellite and radar images (large, 1024 x 1024 km square) and of the output predicted radar image (small, 64 x 64 km square).

질의응답용 슬라이드

