<div align="center">

|  |  |  |
| --- | --- | --- |
| **2018** | **Report – 2018/ 12월 3주** | **Seong-gu** |

</div>

**I. Plan for this week**

***Personal Research Subject***

* 학회 제출용 논문 작성  
  (Introduction 완료)

**II. Plan for next week**

* 논문 작성 완료

**III. Response on Prof. Shin's feedback last week**

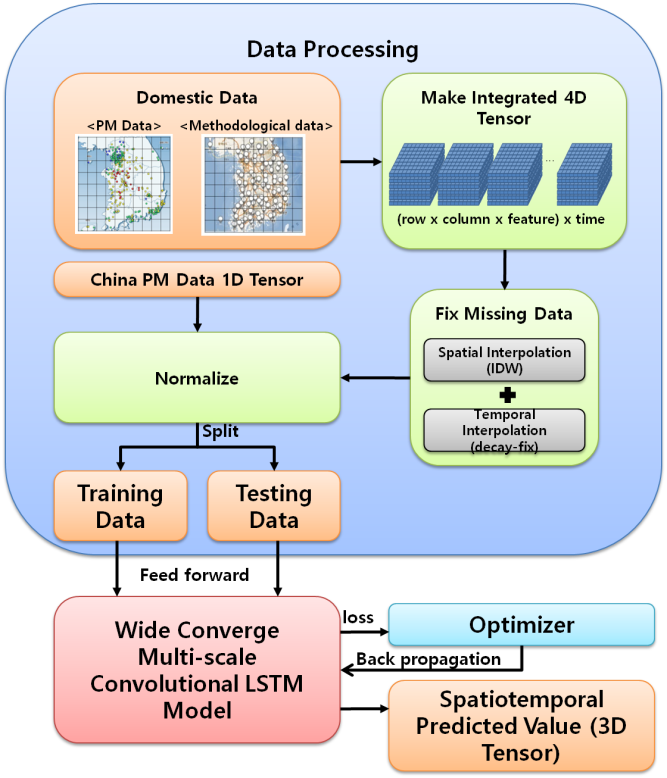
* 모델별로 구조 검토해보고, 가장 미세먼지 농도 예측에 나은 모델링 선정
* 결과로 보여준 median, min, max는 의미가 없어 보임.
* 결과를 잘 보여주는 방법에 대해 검토
* 모델링에 대해 자세히 검토할 것

**IV. Milestone**

1. 실험실 환경에서 미세먼지 농도 예측 시스템 구축   
   1. 머신러닝 기법을 활용한 미세먼지 농도 예측  
        
      **- 한국의 지리학적 특성 및 계절풍의 영향을 고려하여, 활용할 측정소 범위를 국내 전역으로 하고, 중국의 미세먼지 농도를 모델링 데이터로 사용함.**(중국데이터를 확보하는 방법 조사할 것)

**- 결측치는 공간적 보간법인 IDW 기법과 시간적 보간법인 decay-fix 기법을 결합한 시공간 보간법을 사용하여 추정함.**

**- 모델링은 미세먼지 유입의 공간적인 특성을 적극 반영하기 위해서 최근에 연구되고 있는 Convolutional LSTM 모델을 사용.**(모델링 최적화 필요)



**V. Reports**

* 개인연구 진행사항
* **모델링 관련**

현재 진행되고 있는 모델링 단계입니다.

1. **측정소 선별, Convolutional LSTM input matrix에 data mapping**

* 미세먼지 측정소를 52개, 기상측정소를 3개만 선별하고, 공간좌표를 9x9로 줄임.

1. **Fixing missing value – 보류**

**3) Convoltional LSTM 모델링 및 파라미터 최적화 –**

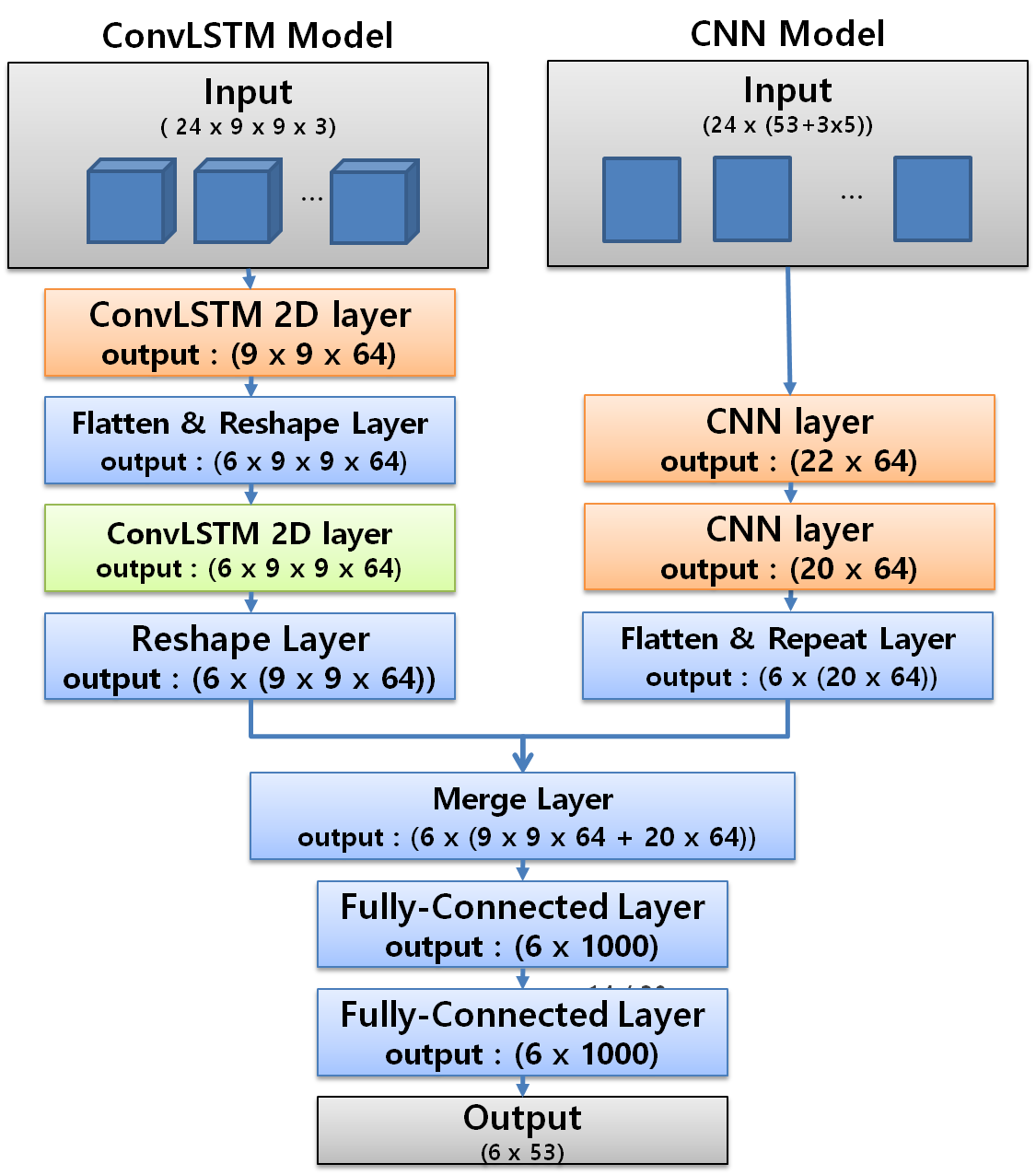
모델링에 바람의 영향을

사용 데이터는

|  |  |  |  |
| --- | --- | --- | --- |
| **Region** | | | |
| Gyeonggi, Incheon, Seoul | | | |
| **Period** | | | |
| 2014.1.1 1:00 – 2015.12.31 24:00 (Training Set) & 2016.1.1 1:00 – 2016.12.31 24:00 (Testing Set) | | | |
| **Input Parameters** | **Unit** | **Source** | **Model Input** |
| Past 24-h PM10 concentration | ug/m3 | 53 of PM10 stations | Convolutional LSTM, CNN |
| Past 24-h Temperature | ℃ | 3 of methodological stations | CNN |
| Past 24-h Humid | % | 3 of methodological stations | CNN |
| Past 24-h Vapor Pressure | hPa | 3 of methodological stations | CNN |
| Past 24-h Wind speed | m/s | 3 of methodological stations | CNN |
| Past 24-h Wind sin component | m/s | 3 of methodological stations | Convolutional LSTM |
| Past 24-h Wind cos component | m/s | 3 of methodological stations | Convolutional LSTM |
| **Output Parameters** |  |  | |
| Next 6-h PM10 concentration | ug/m3 | 53 of PM10 stations | |

**개선된 모델링**

개선된 모델링은 바람에 의한 미세먼지 유입을 모델링한 ConvLSTM과, 시간 feature를 뽑아낸 CNN Model의 output을 병합하여 최종 결과를 얻게 됩니다.

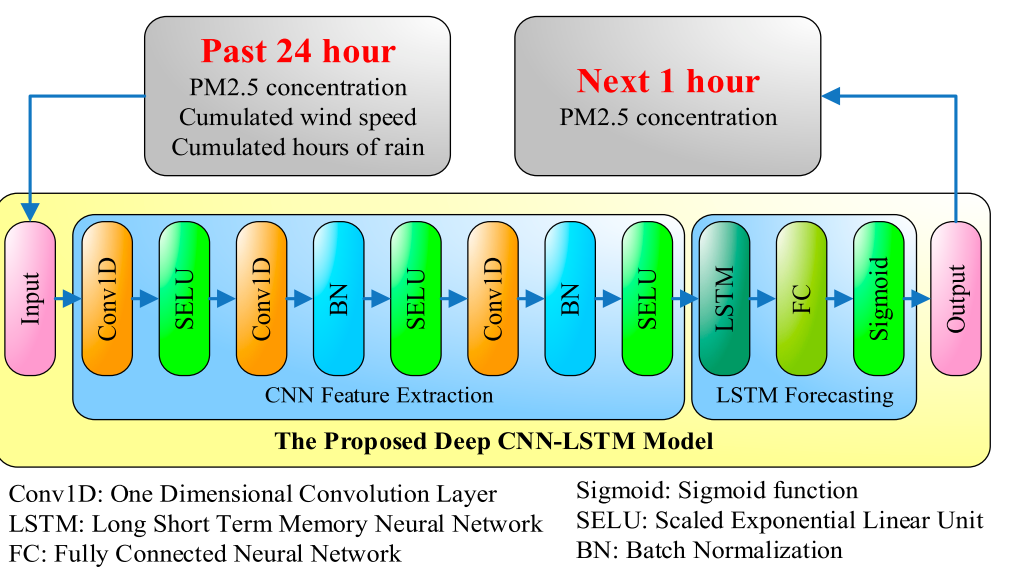
****

**4) 다른 모델링(LSTM, CNN, CNN-LSTM 등)과 결과값 비교- – 진행중**

아래 표는 6시간 뒤의 모든 측정소의 RMSE값 평균입니다.

|  |  |
| --- | --- |
|  | RMSE(average) |
| ConvLSTM 1x1 | 25.87 |
| ConvLSTM 3x3 | 24.53 |
| ConvLSTM 5x5 | 26.85 |
| LSTM | 25.84 |
| CNN-LSTM | 27.43 |

\*CNN-LSTM은 다른 논문의 모델링 과 파라미터 및 LSTM이 Encode-Decode 구조로 변경된 것을 제외하고, 동일하게 제작 되었습니다.



**결과값 정리 관련 추가 예정 작업**

1) 시간대별로 RSME, MAE, MAPE값 표로 추가.

2) 비교 모델에 MLP, 기타 모델들 추가

3) 하나의 측정소에 대한, 모델별 실제값 비교 그래프 추가

4) 미세먼지 농도 기준에 대한 labeling 정확도 비교 표 추가 (option)

