

# 미세먼지 보간 모델을 위한 Local BatchNormalization

윤진용\*, 안석호\*\*, 서영덕\*\*

## Local Batch Normalization for Air Pollution Interpolation Models

JinYong Yun\*, Seokho Ahn\*\*, and Young-Duk Seo\*\*

### 요 약

배치 정규화(Batch Normalization, BN) 기법을 도입하여 미세먼지 보간 모델을 개선한다. 기존 보간 연구는 배치 정규화를 활용한 공간적 특징의 안정화를 고려하지 못했다. 본 연구에서는 미세먼지 보간 모델에 글로벌 배치 정규화와 로컬 배치 정규화 계층을 적용하여 이러한 문제를 해결하고자 한다. 실험 결과, 다양한 수준에서 배치 정규화 계층을 적용하면 보간 정확도가 개선됨을 확인했다.

### Abstract

We propose a method to enhance spatial dependencies in air pollution interpolation models by applying Batch Normalization (BN) layers. Global BN and Local BN are applied to all stations and individual stations, respectively, to improve spatial feature stability. Experiments using PM2.5 data from Antwerp demonstrated that models with BN layers achieved higher interpolation accuracy compared to conventional models. This confirms that BN layers can effectively enhance the performance of air pollution interpolation models.

### Key words

Air pollution, Interpolation, ConvLSTM, Batch Normalization, PM2.5

## 1. 서 론

도시 내 미세먼지 농도를 정확히 측정하기 위해 곳곳에 미세먼지 측정소가 설치되고 있지만, 모든 지역을 커버할 만큼 충분한 미세먼지 측정소를 설치하는 것은 상당한 비용과 공간적인 제약이 따른다. 따라서 미세먼지 측정소가 없는 지역의 농도 추

정을 위해 다양한 보간 연구가 활발히 진행되고 있다 [1-3]. 이러한 연구에서는 주로 시간과 공간의 상호작용을 고려한 시공간(spatiotemporal) 모델이 사용되며, 특히 시계열적 성능 개선에 초점을 맞춰 보간 성능을 향상시켜왔다.

시계열적 성능 개선 방식으로 배치 정규화를 적용할 경우, 각 시점에서 개별적으로 평균과 분산을

\* 인하대학교 컴퓨터공학과, 12191632Y@inha.edu,

\*\* 인하대학교 전기컴퓨터공학과 sokho0514@inha.edu, mysid88@inha.ackr(교신저자)

※ 본 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(NRF-2022R1C1C1012408, 우수신진연구)과 정보통신기획평가원의 지원(No.2022-0-00448/RS-2022-II220448, 인간처럼 회상이 가능한 인공 신경망 지속학습 플랫폼 개발, No.RS-2022-00155915, 인공지능융합혁신인재양성(인하대학교))을 받아 수행한 연구임.

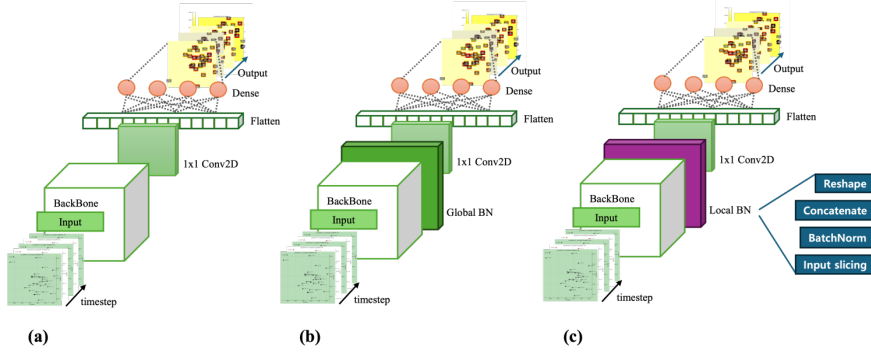


그림 1. 배치 정규화 적용 모델 구조도

Fig. 1. Model Architecture with Batch normalization

계산해 정규화하기 때문에 시간 축에 따른 데이터의 일관성이 깨질 위험이 있다 [4]. 반면, 공간적 성능 개선을 위해 배치 정규화를 적용하면 각 측정소의 미세먼지 농도의 범위와 변화 패턴을 일정하게 조정하여 미세먼지의 공간적 분포가 왜곡되지 않도록 방지하며, 공간적 특성을 안정화시킬 수 있다.

본 논문에서는 지역적으로 보간 모델에 배치 정규화를 진행하여 미세먼지 농도 보간 모델의 성능을 향상시킨다. 또한 본 방식은 데이터의 양이나 모델의 크기를 증가시키지 않으면서도 보간 성능을 비용 효율적으로 향상시킬 수 있는 방식이다. ConvLSTM, CNN 보간 모델을 백본 모델로 실험한 결과, 보간 정확도가 지역 배치 정규화, 전역 배치 정규화, 기본 모델 순으로 지역 배치 정규화가 가장 높은 보간 성능을 보였다.

## II. 제안 방법

본 논문에서 제안하는 방식은 그림 1(c)와 같다. 구체적으로, 그림 1(a)와 같은 기존의 미세먼지 보간 모델에 지역적으로 배치 정규화를 수행하여 공간적인 특성을 반영한다. 기존 전역적인 배치 정규화 방식인 그림 1(b)와의 차이점은 각 공간적 위치에 대해서 배치 정규화를 독립적으로 수행한다는 점이다. 이는 측정소별 공간적 특성을 반영하지 못하는 전역적 배치 정규화 방식의 문제를 해결할 수 있다.

## III. 실험 및 결과

### 3-1. 데이터셋

본 실험에서 사용한 데이터셋은 Antwerp 지역 32개의 미세먼지 측정소에서 OPCN3PM25 센서를 이용해 1분 단위로 측정한 PM2.5 데이터이다 [5]. 모든 측정소에서 동시에 측정된 시간대의 데이터만을 사용했다. 데이터의 결측치는 IDW (Inverse Distance Weighting) 기법으로 보간했다. 각 측정소의 위도와 경도는 정규화하여 32x32 그리드에 미세먼지 측정값과 함께 저장된다. 본 실험에서 사용된 학습, 검증, 평가 데이터의 비는 8:2:2이다.

### 3-2. 실험 환경

본 논문에서는 실험을 위해 미세먼지 보간을 위한 백본 모델로 ConvLSTM [1]과 CNN [6]을 사용하였다. 각 백본 모델에 대해서 전역 정규화와 지역 정규화를 수행하고 그 결과를 비교하였다. 본 연구에서 보간 정확도를 측정하기 위해 사용한 평가 지표는 spRMSE로, 이는 측정소별로 보간 값과 실제 미세먼지 값의 RMSE값의 평균으로 계산된다. 나머지 실험 환경에 대한 세부 설정은 [1]과 [6]에서 제안한 방법을 따랐다.

### 3-3. 실험 결과

표 1. 배치정규화 실험 결과 (spRMSE)

Table 1. Batch normalization experiment result (spRMSE)

Backbone	ConvLSTM [1]	CNN [6]
Baseline	16.6354	18.6550
Global BatchNorm	13.9992	15.9571
Local BatchNorm	13.7041	14.1488

표 1은 두 백본 모델에 대해서 배치 정규화를 수행하지 않았을 때 (Baseline), 전역 배치 정규화를 수행했을 때 (Global BatchNorm), 그리고 지역 배치 정규화 (Local BatchNorm)를 수행했을 때 정량적인 비교실험 결과를 나타낸다. 표 1에서 볼 수 있듯이 두 백본 모델에 대해서 보간 정확도는 로컬 배치 정규화 모델, 전역 배치 정규화 모델, 기본 모델 순으로 높았다. 따라서 미세먼지 보간 모델에 전역 배치 정규화를 적용하는 것이 성능에 유의미한 영향을 끼칠 뿐만 아니라, 측정소별 지역 배치 정규화를 적용하는 방식이 더 효과적임을 보였다.

### 3-4. 사례 연구

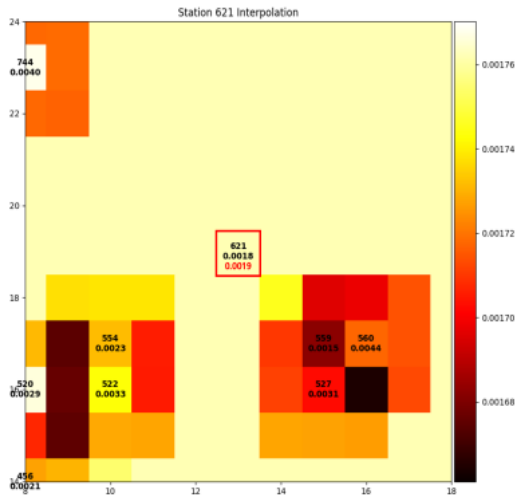


그림 2. 621 측정소 보간 결과 그리드 이미지

Fig. 2. 621 station's Interpolation result grid image

그림 2는 특정 시점에서 621번 측정소의 주변 측정소를 기반으로 지역 배치 정규화를 적용한 보간 결과를 히트맵으로 나타낸 이미지이다. 그림에서의

색은 주변 측정소의 미세먼지 농도를 나타내며, 색이 짙을수록 미세먼지 농도가 높다. 텍스트는 측정소의 번호와 실제 미세먼지 농도를 나타낸다. 621번에서의 텍스트는 미세먼지 보간값이다. 621번 측정소의 보간 값은 0.0018로 해당 시점의 실제 미세먼지 농도 값인 0.0019와 매우 근소한 차이를 보였다.

## IV. 결 론

본 논문에서는 배치 정규화를 미세먼지 보간 모델에 활용했을 때 보간 성능이 올라간다는 가설을 입증한다. 특히 각 측정소별로 배치 정규화를 적용하면 각 측정소별 공간적 특성에 맞게 정규화가 일어나서 더욱 정밀한 보간 결과를 얻을 수 있다.

## 참 고 문 헌

- [1] V. D. Le and T. C. Bui, "Spatiotemporal Deep Learning Model for CityWide Air Pollution Interpolation and Prediction", IEEE BIGCOMP, pp. 55-62, February 2020.
- [2] D. Wong, L. Yuan and S. Perlin, "Comparison of spatial interpolation methods for the estimation of air quality data", nature journal of Exposure Science & Environmental Epidemiology vol. 14, pp. 404-415, September 2004.
- [3] S. H. Kim, et al. "Explainable AI-driven high-fidelity IAQ prediction model for subway stations", Building and Environment, vol. 263, pp. 112220, July 2024
- [4] C. S. Laurent, et al. "Batch Normalized Recurrent Neural Networks", in Proc. IEEE ICASSP, pp. 2657-2661, October 2015.
- [5] M. V. Poppel, et al. "SensEURCity: A multi-city air quality dataset collected for 2020/2021 using open low-cost sensor systems", nature scientific data, vol. 10, pp. 322, May 2023.
- [6] V. D. Le and S. K. Cha, "Real-time Air Pollution prediction model based on Spatiotemporal Big data", The International Conference on Big data, IoT, and Cloud Computing (BIC), April 2018