

<https://doi.org/10.7236/JIIBC.2024.24.6.135>

JIIBC 2024-6-20

딥러닝 및 패치 기반 커널 PCA를 이용한 미세먼지 추정

A Particulate Matter Estimation Using Deep Learning and Patch-Based Kernel PCA

이재원, 인치호*

Jae-Won Lee, Chi-Ho Lin*

요약 본 연구는 이미지 기반 미세먼지(PM_{2.5} 및 PM₁₀) 농도 예측을 위해 패치 기반 커널 PCA와 딥러닝을 이용한 미세먼지 추정 방법론을 제안한다. 커널 PCA는 이미지의 비선형적 특징을 효과적으로 추출하고, CNN 모델은 이를 학습하여 높은 예측 성능을 달성한다. 강화 학습 기반 마스크 최적화와 다중 커널 통합을 통해 예측 정확도와 학습 효율성을 향상했으며, 다양한 환경에서 일관된 성능을 유지했다. 실험 결과, 제안된 방법론은 기존 센서 기반 방법보다 넓은 범위에서 실시간 모니터링이 가능하며, 저사양 하드웨어에서도 우수한 실시간 성능을 보여준다. 향후 연구는 다양한 환경에서의 데이터 수집과 모델 최적화를 통해 실시간 응용 가능성을 더욱 높이는 데 주력할 것이다.

Abstract In this paper, we propose a methodology that combines patch-based kernel PCA and deep learning for image-based particulate matter (PM_{2.5} and PM₁₀) concentration prediction. Kernel PCA effectively extracts the nonlinear features of images, and a CNN model is trained to achieve high prediction performance. Through reinforcement learning-based mask optimization and multi-kernel integration, we improve the prediction accuracy and learning efficiency, while maintaining consistent performance in different environments. Experimental results show that the proposed methodology is capable of real-time monitoring in a wider range than existing sensor-based methods, and shows good real-time performance even on low-end hardware. Future work will focus on data collection in different environments and model optimization to further enhance its real-time applicability.

Key Words : Patch-based Kernel PCA, Deep Learning, Particulate Matter Prediction

1. 서 론

미세먼지(PM_{2.5} 및 PM₁₀) 오염은 전 세계적으로 주요한 공중 보건 문제로 대두되고 있습니다. PM_{2.5}와 같은 초미세먼지는 매우 작은 입자 크기(2.5 마이크로미터 이하)로 인해 호흡기를 통해 체내에 쉽게 침투하여 폐포

에 도달할 수 있으며, 이는 다양한 호흡기 질환, 심혈관 질환, 그리고 조기 사망과 같은 심각한 건강 문제를 유발할 수 있습니다^[1, 2]. 특히, PM_{2.5}는 대기 중에서 오래 머무를 수 있어 넓은 지역에 걸쳐 오염을 확산시키며, 인체에 미치는 해로운 영향을 더욱 증가시킵니다. 이러한 위험성 때문에, 전 세계 많은 도시는 미세먼지 농도를 실

*정회원, 세명대학교 컴퓨터학부
접수일자 2024년 8월 28일, 수정완료 2024년 11월 8일
게재확정일자 2024년 12월 6일

Received: 28 August, 2024 / Revised: 8 November, 2024 /
Accepted: 6 December, 2024

*Corresponding Author: ich410@semyung.ac.kr
School of Computer Science, Semyung University, Korea

시간으로 모니터링하고 이에 따라 대기질 관리 정책을 시행하고 있습니다^[3].

현재의 미세먼지 농도 측정 방법은 주로 물리적 센서를 기반으로 합니다^[4]. 이러한 센서들은 높은 정확도를 제공할 수 있지만, 설치 및 유지 비용이 많이 들고, 물리적 위치에 따라 제약을 받기 때문에 광범위한 지역에서의 실시간 모니터링에는 한계가 있습니다. 예를 들어, 개발도상국이나 인프라가 부족한 지역에서는 이러한 센서를 광범위하게 배치하기 어렵습니다. 또한, 센서 기반 측정은 특정 지점에서만 데이터를 수집할 수 있으므로, 도시 전체의 공기질을 반영하는 데는 한계가 있을 수 있습니다.

최근 몇 년간 인공지능, 특히 딥러닝 기술의 발전은 이미지 분석 분야에서 상당한 성과를 보여주었습니다^[5]. 이러한 기술은 의료 영상 분석^[6], 자율주행차^[7], 보안 감시 시스템^[8] 등 다양한 응용 분야에서 사용되고 있으며, 이미지 기반 미세먼지 농도 추정에도 적용될 수 있습니다. 예를 들어, CNN(Convolutional Neural Network)과 같은 딥러닝 모델은 이미지에서 다양한 특징을 자동으로 추출하고 이를 바탕으로 높은 정확도의 예측을 수행할 수 있습니다^[9, 10].

이미지 기반 분석은 저비용으로 광범위한 지역의 공기질을 모니터링할 가능성을 제공합니다^[11, 12]. 스마트폰, 드론, CCTV 카메라와 같은 다양한 이미지 캡처 장치를 활용하면 물리적 센서를 설치하지 않아도 많은 양의 데이터를 수집할 수 있습니다. 이러한 데이터는 도시 전체의 공기질을 실시간으로 모니터링하고, 미세먼지 농도를 추정하는 데 매우 유용하게 사용될 수 있습니다. 또한, 이미지 기반 접근법은 시간적, 공간적 해상도가 높아 특정 지점이 아닌 넓은 지역에서의 공기질 변화를 효과적으로 감지할 수 있습니다^[13, 14].

본 연구의 목적은 패치 기반 커널 주성분 분석(kernel PCA)과 딥러닝을 결합한 새로운 방법론을 통해 이미지 데이터를 활용하여 미세먼지 농도를 효과적으로 추정하는 것입니다^[15, 16]. 기존 연구에서는 패치 기반 커널 PCA를 사용하여 이미지의 복잡한 비선형적 특징을 효과적으로 추출하고 이를 딥러닝 모델의 입력으로 사용함으로써 모델의 성능을 향상시키려는 시도가 있었습니다^[17]. 커널 PCA는 비선형 차원 축소 기법으로, 데이터의 비선형 구조를 반영하여 복잡한 데이터의 특징을 더 효과적으로 추출할 수 있습니다. 이를 통해 이미지의 다양한 지역적 변화를 포착하고, 미세먼지 농도 추정의 정확도를 높일 수 있습니다^[18].

딥러닝 모델은 이러한 커널 PCA로부터 추출된 특징을 학습하여 이미지에서 미세먼지 농도를 추정할 수 있습니다. 예를 들어, CNN은 이미지의 시각적 특징을 학습하는 데 매우 강력한 도구로, 다양한 크기의 필터를 사용하여 이미지의 공간적 계층 구조를 파악합니다^[19]. 패치 기반 커널 PCA로 생성된 초기 마스크는 네트워크 학습 초기 단계에서 중요한 역할을 하며, 네트워크가 중요한 시각적 패턴에 집중하도록 도와줍니다. 이로 인해 네트워크는 빠르게 수렴하고 높은 예측 성능을 달성할 수 있습니다^[7, 8].

본 연구는 이미지 기반 미세먼지 농도 추정을 위한 새로운 방법론으로 패치 기반 커널 PCA와 딥러닝을 결합하였습니다. 제안된 방법론이 기존의 센서 기반 방법보다 더 넓은 공간적 구역을 제공하고, 실시간 공기질 모니터링의 가능성을 제시합니다. 다양한 환경에서의 실험을 통해 제안된 방법론의 효과를 검증하고, 그 결과를 기존 방법론과 비교 분석합니다.

본 논문의 나머지 부분은 다음과 같이 구성됩니다. II장에서 제안된 방법론의 세부 구현을 설명합니다. III장에서는 실험 결과와 그에 대한 분석을 제시하며, 마지막으로 IV장에서는 연구의 결론과 향후 연구 방향을 논의합니다.

II. 딥러닝 및 패치 기반 커널 PCA를 사용하여 미세먼지 추정 강화

본 연구에서는 고해상도 이미지에서 객체 검출의 정확도를 높이고 연산 효율성을 극대화하기 위해 패치 기반 커널 PCA, 강화 학습 기반 마스크 최적화, 주목 모듈을 결합한 딥러닝 네트워크 학습 전략을 제안합니다. 이는 기존 PCA 기반 방법이 전체 이미지 연산에 드는 높은 계산 비용과 성능 한계를 극복하고, 더 정밀한 객체 경계를 검출하도록 설계되었습니다. 본 절에서는 제안한 방법론의 세부 절차와 각 단계에서의 주요 특징을 설명하며, 기존 모델과의 아키텍처 비교를 통해 본 연구에서 추가된 개선 사항을 논의합니다.

고해상도 이미지에서 객체 검출의 정확도를 높이는 것은 컴퓨터 비전 분야에서 매우 중요한 과제입니다. 기존 연구에서 전체 이미지 단위의 PCA 적용은 연산 비용이 많이 들고, 특히 복잡한 객체와 배경이 혼재된 이미지에서 객체 경계를 명확하게 구분하는 데 한계를 가집니다.

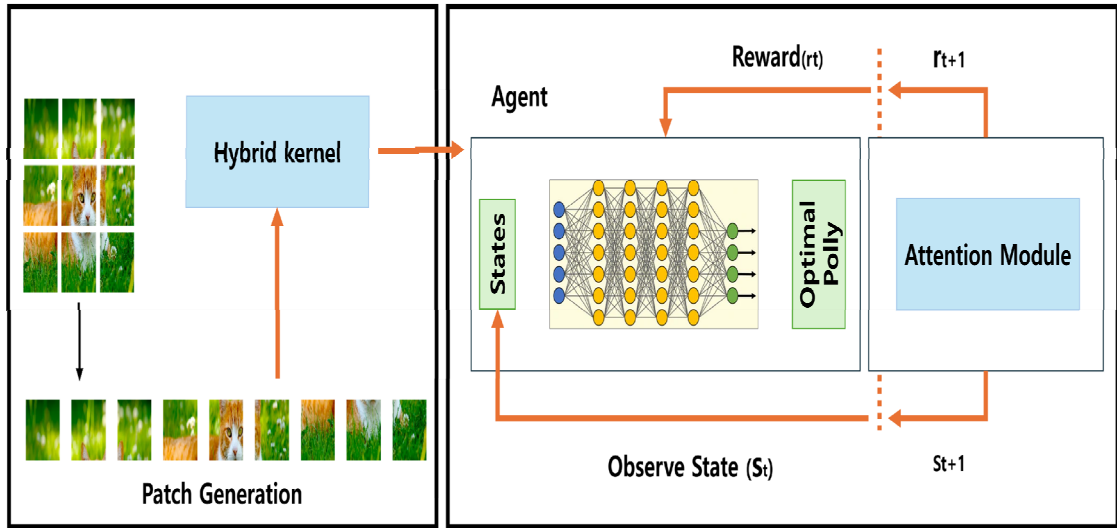


그림 1. 딥러닝 및 패치 기반 커널 PCA를 이용한 미세먼지 추정 강화 알고리즘
 Fig. 1. An Algorithm for Reinforcing Fine Dust Estimation Using Deep Learning and Patch-Based Kernel PCA

본 연구에서는 이러한 문제를 해결하기 위해 이미지를 작은 패치로 분할하고, 각 패치에 대해 독립적으로 커널 PCA를 적용하는 방식을 도입하였습니다.

각 패치에 대해 커널 PCA를 독립적으로 수행하며, 이때 패치 내 주요 특징을 추출하고, 고차원 공간에서의 데이터를 저차원 공간으로 변환하여 객체와 배경을 구분하는 데 유리한 정보를 얻습니다. 본 연구에서는 RBF(Radial Basis Function) 커널을 기본으로 사용하되, 다양한 커널을 통합한 하이브리드 커널 PCA를 도입하여 더 다양한 데이터 분포를 포착할 수 있도록 하였습니다.

이 하이브리드 커널 방식에서는 단일 커널에 의존하지 않고, 각 패치의 데이터 특성에 따라 여러 커널을 통합하여 사용합니다. 적응적 커널 선택 알고리즘은 패치 내 데이터의 밀도와 분포에 따라 가장 적합한 커널을 가중치로 선택하도록 설계되었습니다.

각 패치에서 데이터 특성을 반영하기 위해, 여러 커널 $k_i(x, x')$ 에 대해 최적의 가중치 α_i 를 학습하여 결합 커널을 생성합니다. 결합 커널 $k^{*}_{hybrid}(x, x')$ 은 다음과 같이 정의됩니다:

$$k^{*}_{hybrid}(x, x') = \sum_i \alpha_i k_i(x, x') \quad (1)$$

각 가중치 α_i 는 패치의 데이터 밀도나 분포에 따라 조정되며, 이때 가중치는 다음과 같이 최적화 기준을 통

해 업데이트됩니다. 특히 α_i 는 강화 학습 기반 최적화와 결합하여 패치 내의 주요 특징을 반영하며 업데이트됩니다:

1. 밀도 기반 가중치 조정: 패치 내 데이터의 밀도를 분석하여, 고밀도 구역에서는 비선형 특성을 잘 반영할 수 있는 RBF 커널에 가중치를 높게 부여합니다. 밀도가 낮은 경우에는 선형적 분포를 잘 반영할 수 있는 Polynomial 커널이나 Sigmoid 커널에 더 높은 가중치를 부여합니다.
2. 분산 분석에 따른 커널 선택: 각 패치 내 데이터의 분산을 측정하여, 분산이 높은 경우에는 비선형 특성이 강하게 나타나는 영역으로 판단하고 RBF와 Polynomial 커널의 가중치를 증가시킵니다. 반대로 분산이 낮은 경우, 선형에 가까운 데이터를 반영할 수 있도록 Sigmoid 커널에 가중치를 부여합니다. 이 가중치 α_i 는 패치 내 분포 통계값(밀도와 분산)을 기반으로 주어진 적응적 최적화 함수로 업데이트됩니다. 또한, 강화 학습의 보상 함수에서 학습된 정보를 활용하여 주목 모듈과 결합하여 각 패치의 가중치를 최적화합니다. 이 과정에서 선택된 결합 커널 $k^{*}_{hybrid}(x, x')$ 는 각 패치에 최적화된 커널로서의 역할을 수행합니다.

강화 학습 기반 최적화는 마스크의 정확도를 높이기

위해 네트워크가 학습 과정에서 점진적으로 개선된 마스크를 생성하도록 합니다. 강화 학습의 보상 함수는 각 학습 단계에서 마스크가 주요 시각적 특징을 잘 반영하도록 도와줍니다. 구체적으로, 에이전트는 아래의 보상 함수를 통해 마스크의 정확도를 높이며 업데이트합니다:

$$R(s_t, a_t) = \lambda \cdot IOU_t - \beta \cdot Reg_t \quad (2)$$

여기서 IOU_t 는 현재 마스크와 객체 간의 교집합 비율을 나타내며, Reg_t 는 과도한 경계 왜곡을 최소화하는 정규항입니다. 이 최적화 과정에서 학습된 가중치 α 는 각 패치의 중요한 특징을 반영하며 주목 모듈과의 연계를 통해 강화 학습 기반 최적화에 반영됩니다.

주목 모듈은 객체 검출 과정에서 특정 시각적 특징에 더 집중하도록 설계되었습니다. 본 연구에서는 최적화된 주목 모듈을 통해 중요한 특징에 높은 가중치를 부여하여, 복잡한 배경에서의 객체 경계 검출이 보다 쉬워졌습니다.

강화 학습 기반 최적화와 연계된 주목 모듈은 소프트 맥스 함수를 통해 주목할 영역을 선택하며, 아래와 같은 가중치 업데이트 방식을 따릅니다:

$$\alpha_{ji} = \frac{\exp(e_{ij})}{\sum_k \exp(e_{ik})} \quad (3)$$

여기서 e_{ij} 는 패치 j 의 특징이 특정 객체와 얼마나 유사한지를 나타내며, 이를 통해 객체 경계에서의 중요한 특징이 강화됩니다. 이 가중치는 강화 학습 기반의 보상 함수 $R(s_t, a_t)$ 와 연계되어 주목 모듈의 성능 향상하며, 객체 경계의 중요도를 강조하여 모델이 이미지 내 중요한 특징에 집중하도록 합니다. 본 연구에서 제안한 아키텍처는 기존 모델의 한계를 보완하고 성능을 향상하기 위해 다양한 구성 요소들을 개선하여 최적화하였습니다. 표 1은 기존 모델과 제안 모델의 아키텍처를 비교하여, 본 연구의 차별화된 요소와 개선 사항을 요약한 것입니다.

III. 실험 및 결과

본 연구에서는 제안된 패치 기반 커널 PCA와 딥러닝 네트워크 방법론을 활용하여 이미지 데이터를 기반으로 미세먼지(PM2.5 및 PM10) 농도를 예측하는 성능을 평가하기 위해 다양한 실험을 수행했습니다. 실험은 여러 환경에서 촬영된 이미지 데이터셋을 활용하여 모델의 예측 정확도와 효율성을 평가하는 방식으로 진행되었습니다.

다양한 기상 조건 및 환경에서 수집된 이미지를 사용하여 실험을 진행했습니다. 본 연구에 사용된 데이터셋은 여러 도시에서 맑음, 흐림, 비, 안개 등의 다양한 기상 조건과 도심, 교외, 공업 지역 등 여러 환경에서 촬영된

표 1. 기존 모델과 제안 모델의 아키텍처를 비교

Table 1. Compare architecture between traditional and proposed models

구성 요소	기존 모델	제안 모델	설명
입력 이미지 전처리	기본 이미지 처리	패치 기반 커널 PCA 적용	이미지 분할을 통해 지역적 특징 강조
커널 적용	단일 커널 또는 다중 커널	다중 커널 통합 및 최적화	다양한 커널 통합을 통해 비선형 특징 반영
초기 마스크 최적화	없음	강화 학습 기반 마스크 최적화	주요 특징 강조 및 학습 효율성 증대
주목 모듈 (Attention Module)	포함된 경우도 있음	최적화된 주목 모듈 사용	특정 시각적 특징 강조
다중 스케일 접근법	포함된 경우도 있음	개선된 다중 스케일 접근법 적용	다양한 크기와 해상도에 맞춘 일관된 검출 성능

표 2. 미세먼지 농도 예측 모델의 성능 비교

Table 2. Performance Comparison of Airborne Particulate Matter Prediction Models

모델	평균 절대 오차 (MAE)	평균 제곱근 오차 (RMSE)	결정 계수 (R²)	인퍼런스 속도
전통적인 CNN 모델	9.2 $\mu\text{g}/\text{m}^3$	11.4 $\mu\text{g}/\text{m}^3$	0.78	20 FPS
패치 기반 커널 PCA 적용 모델	7.8 $\mu\text{g}/\text{m}^3$	9.5 $\mu\text{g}/\text{m}^3$	0.87	25 FPS
단일 커널 적용 모델	6.6 $\mu\text{g}/\text{m}^3$	8.8 $\mu\text{g}/\text{m}^3$	0.89	28 FPS
다중 커널 통합 모델	5.9 $\mu\text{g}/\text{m}^3$	7.8 $\mu\text{g}/\text{m}^3$	0.91	30 FPS
강화 학습 마스크 최적화 모델	6.2 $\mu\text{g}/\text{m}^3$	8.1 $\mu\text{g}/\text{m}^3$	0.91	29 FPS

이미지들로 구성되어 있습니다. 각 이미지에는 해당 시
간과 장소에서 측정된 PM2.5 및 PM10 농도 정보가 포
함되어 있습니다.

이미지의 일관성을 유지하기 위해 모든 이미지를 동일
한 해상도(256x256픽셀)로 조정하고, 색상 보정 및 노이즈
제거 등의 전처리 작업을 수행했습니다. 이러한 전처
리 과정을 통해 이미지 내 불필요한 변수를 최소화하여,
모델이 미세먼지 관련 특징에 집중할 수 있도록 했습니다.

실험은 NVIDIA RTX 3090 GPU를 사용하여
TensorFlow와 PyTorch 딥러닝 프레임워크로 진행되
었습니다. 데이터 증강 기법을 활용해 데이터셋의 다양
성을 높였으며, 이를 통해 모델의 일반화 능력을 향상했
습니다. 모델 학습은 Adam 옵티마이저를 사용하여 학습
률 0.0001에서 시작했으며, 학습이 진행됨에 따라 학습
률이 점차 감소하도록 설정했습니다. 배치 크기는 32로
설정하여 학습의 안정성을 유지했습니다.

본 연구에서는 미세먼지 농도 예측의 정확도와 모델의
성능을 평가하기 위해 세 가지 주요 지표를 사용했습니다:

평균 절대 오차(MAE)는 예측된 미세먼지 농도와 실제
농도 간의 평균 절대 차이를 측정하여 모델의 예측 정확
도를 평가합니다. 낮은 MAE 값은 모델이 실제 값에 더
가까운 예측을 수행했음을 의미합니다.

평균 제곱근 오차(RMSE)는 예측값과 실제값 사이의
차이 제곱 평균의 제곱근을 계산하며, 큰 오차에 더 민감
하게 반응합니다. RMSE는 모델의 전반적인 예측 성능을
평가하는 데 사용되며, 낮을수록 성능이 우수함을 나타
냅니다. 결정 계수(R^2)는 모델이 데이터를 얼마나 잘 설
명하는지를 나타내며, 1에 가까울수록 모델의 설명력이
높음을 의미합니다. 이 지표는 모델의 예측이 얼마나 정
확한지를 평가하는 데 중요한 역할을 합니다.

패치 기반 커널 PCA를 적용한 모델은 이미지에서 미
세먼지 농도와 관련된 중요한 특징들을 효과적으로 추출
했습니다. 이러한 접근법은 특히 다양한 환경과 기상 조
건에서 미세먼지 농도를 예측할 때 유용했습니다.

고해상도 이미지 데이터셋 결과, 고해상도 이미지 데
이터셋을 사용한 실험에서는 패치 기반 커널 PCA를 적
용한 딥러닝 모델이 평균 제곱근 오차(RMSE) 9.5
 $\mu\text{g}/\text{m}^3$, 결정 계수(R^2) 0.87을 기록했습니다. 이는 고해
상도 이미지에서 미세먼지 농도를 예측하는 데 있어 제
안된 방법론이 매우 효과적임을 나타냅니다.

강화 학습 기반 마스크 최적화 기법을 적용한 결과,
네트워크가 이미지 내 미세먼지와 관련된 주요 특징을
더 잘 학습할 수 있었습니다. 이는 모델의 예측 정확도를

높이는 데 중요한 역할을 했습니다. 학습 초기 단계: 강
화 학습 기반 마스크 최적화를 적용한 네트워크는 학습
초기 단계에서 빠르게 수렴하였으며, 10 에포크 내에 R^2
값이 0.80에 도달했습니다. 이는 기존 방법론보다 약
30% 더 빠른 수렴 속도를 나타내며, 마스크 최적화가 네
트워크 학습의 효율성을 크게 개선했음을 보여줍니다.

다중 커널 통합 기법을 통해 다양한 데이터 특징을 효
과적으로 학습할 수 있었습니다. 실험 결과, 다중 커널을
사용한 모델은 미세먼지 농도를 보다 정확하게 예측할
수 있었으며, 다양한 환경에서 일관된 성능을 유지했습
니다. 다중 커널을 사용한 모델의 평균 절대 오차(MAE)
는 5.9 $\mu\text{g}/\text{m}^3$, 평균 제곱근 오차(RMSE)는 7.8 $\mu\text{g}/\text{m}^3$
로, 단일 커널을 사용한 모델 대비 약 10% 성능 향상을
보였습니다. 이는 다중 커널이 다양한 환경에서 미세먼
지 농도를 예측하는 데 있어 더 높은 정확도를 제공함을
나타냅니다.

본 연구에서 제안한 방법론은 실시간 처리 성능에서도
뛰어난 결과를 나타냈습니다. 인퍼런스 속도는 저사양
하드웨어에서도 초당 30프레임 이상을 유지할 수 있었
으며, 이는 실시간 미세먼지 모니터링 시스템에 적용할 가
능성을 높였습니다.

IV. 결 론

본 연구에서는 이미지 기반으로 미세먼지 농도(PM2.5
및 PM10)를 예측하기 위한 새로운 접근법으로, 패치 기
반 커널 PCA와 딥러닝 네트워크를 결합한 방법론을 제
안하고 그 성능을 평가했다. 기존의 물리적 센서 기반 방
법이 가지는 공간적 제약과 비용 문제를 해결하기 위해,
본 연구는 이미지 데이터를 활용하여 미세먼지 농도를
보다 정확하게 추정할 가능성을 탐색했다. 실험 결과, 제
안된 방법론은 미세먼지 농도 예측에서 우수한 성능을
나타냈으며, 특히 다양한 기상 조건과 환경에서 일관된
예측 정확도를 유지할 수 있음을 확인했다.

본 연구의 주요 기여는 다음과 같다. 첫째, 패치 기반
커널 PCA와 딥러닝의 효과적인 결합이다. 본 연구는 패
치 기반 커널 PCA를 사용하여 이미지의 비선형적 특징
을 효과적으로 추출하고, 이를 딥러닝 모델의 학습에 활
용하여 미세먼지 농도 예측의 정확도를 크게 향상했다.
커널 PCA는 고차원 공간에서의 데이터 분포를 더 잘 반
영할 수 있어, 이미지 내의 복잡한 패턴을 효과적으로 포
착하고, 이를 통해 모델의 성능을 극대화할 수 있음을 확

인했다. 둘째, 강화 학습 기반 마스크 최적화 기법을 도입했다. 제안된 강화 학습 기반 마스크 최적화 기법은 네트워크가 이미지 내의 중요한 특징을 더욱 잘 학습하도록 돕고, 초기 학습 단계에서 빠르게 수렴하도록 하여 학습 효율성을 크게 향상했다. 이러한 최적화 기법은 특히 복잡한 환경에서의 이미지 데이터 처리에 매우 효과적이었다. 셋째, 다중 커널 통합의 활용이다. 다중 커널 통합을 통해 다양한 데이터 특징을 학습할 수 있는 모델을 개발하였으며, 이를 통해 미세먼지 농도를 더욱 정확하게 예측할 수 있었다. 다중 커널을 사용한 모델은 단일 커널 모델보다 모든 환경에서 일관된 성능을 보여, 제안된 방법론의 유연성과 적응성을 강조했다. 넷째, 실시간 예측 능력이다. 본 연구의 방법론은 실시간 예측 성능에서도 뛰어난 결과를 나타내었으며, 특히 저사양 하드웨어에서도 높은 인퍼런스 속도를 유지하여 실시간 모니터링 시스템에 적합함을 보여주었다. 이는 제안된 접근법이 실시간 미세먼지 모니터링을 위한 실용적인 솔루션이 될 수 있음을 시사한다.

그러나 본 연구에는 몇 가지 한계가 존재한다. 첫째, 데이터의 다양성 부족이다. 본 연구에서 사용된 이미지 데이터셋은 특정 도시와 환경에 한정되어 있어, 다양한 환경에서의 모델 성능을 충분히 평가하지 못한 한계가 있다. 향후 연구에서는 더 다양한 환경과 조건에서 데이터를 수집하고, 이를 바탕으로 모델의 일반화 능력을 평가할 필요가 있다. 둘째, 모델의 복잡도와 연산 비용이다. 패치 기반 커널 PCA와 딥러닝을 결합한 모델은 높은 예측 성능을 보였지만, 연산 비용이 많이 들어 실시간 응용에 제약이 있을 수 있다. 향후 연구에서는 모델의 경량화를 통해 연산 효율성을 개선하고, 실시간 응용에서의 적용 가능성을 더욱 높이는 방안을 모색할 필요가 있다. 셋째, 미세먼지 농도 예측의 정확도는 이미지의 품질에 크게 의존한다는 점이다. 낮은 해상도나 노이즈가 많은 이미지는 모델의 성능에 부정적인 영향을 미칠 수 있으며, 따라서 이미지 전처리 기법의 개선이 필요하다. 향후 연구에서는 이미지 전처리 기법을 최적화하여 낮은 품질의 이미지에서도 높은 예측 성능을 유지할 방안을 모색해야 한다.

마지막으로, 본 연구는 이미지 기반 미세먼지 농도 예측의 가능성을 제시하고, 이를 통해 다양한 응용 분야에서의 활용 가능성을 제시했다. 향후 연구에서는 다양한 환경에서의 데이터 수집을 통해 모델의 일반화 능력을 더욱 강화하고, 실시간 응용을 위한 추가적인 최적화를 진행할 계획이다. 이를 통해 제안된 방법론이 미세먼지

모니터링 및 공기질 관리에서 실질적인 도구로 활용될 수 있기를 기대한다.

References

- [1] C. A. Pope and D. W. Dockery, "Health effects of fine particulate air pollution: Lines that connect," *Journal of the Air & Waste Management Association*, vol. 56, no. 6, pp. 709-742, 2006. DOI: 10.1080/10473289.2006.10464485
- [2] R. D. Brook, S. Rajagopalan, C. A. Pope, et al., "Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association," *Circulation*, vol. 121, no. 21, pp. 2331-2378, 2010. DOI: 10.1161/CIR.0b013e3181d8e1
- [3] C. Zhang, X. Pan, H. Li, et al., "A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 140, pp. 133-144, 2018. DOI: 10.1016/j.isprsjprs.2017.07.014
- [4] T.-Y. Lin, P. Dollár, R. Girshick, et al., "Feature pyramid networks for object detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2117-2125, 2017. DOI: 10.1109/CVPR.2017.106
- [5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015. DOI: 10.1038/nature14539
- [6] G. Litjens, T. Kooi, B. E. Bejnordi, et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, 2017. DOI: 10.1016/j.media.2017.07.005
- [7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 580-587, 2014. DOI: 10.1109/CVPR.2014.81
- [8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, vol. 28, pp. 91-99, 2015. DOI: 10.5555/2969239.2969250
- [9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097-1105, 2012. DOI: 10.5555/2999134.2999257
- [10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations (ICLR)*, 2015. DOI: 10.48550/arXiv.1409.1556

- [11] E. Süel, A. Züfle, and E. Erdem, "Air quality prediction using machine learning algorithms," *Environmental Science & Pollution Research*, vol. 29, no. 22, pp. 33014-33024, 2022.
DOI: 10.1007/s11356-021-17945-0
- [12] A. Mahajan, S. Karkhanis, A. Chitre, et al., "Predicting air quality using deep learning methods," *Applied Sciences*, vol. 14, no. 3, p. 556, 2024.
DOI: 10.3390/app14030556
- [13] Y. Chen, S. Zhang, and S. Xiang, "Machine learning models for PM2.5 prediction using satellite data," *Atmospheric Pollution Research*, vol. 10, no. 1, pp. 44-52, 2019. DOI: 10.1016/j.apr.2018.06.014
- [14] Y. Zhou and X. Xu, "Fine particulate matter (PM2.5) forecasting using a deep learning model," *Scientific Reports*, vol. 7, no. 1, p. 6868, 2017.
DOI: 10.1038/s41598-017-07085-9
- [15] B. Schölkopf, A. J. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299-1319, 1998. DOI: 10.1162/089976698300017467
- [16] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
DOI: 10.1007/978-0-387-45528-0
- [17] B. Han, Y. Zhao, J. Li, et al., "Deep learning based PM2.5 prediction using LSTM network," *Applied Sciences*, vol. 8, no. 5, p. 688, 2018.
DOI: 10.3390/app8050688
- [18] A. Esteva, B. Kuprel, R. A. Novoa, et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115-118, 2017. DOI: 10.1038/nature21056
- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.
DOI: 10.1109/CVPR.2016.90

저 자 소 개

이 재 원(학생회원)



- 2016년 : 세명대학교 컴퓨터학부 졸업
- 2016년 ~ 2019년 : 세명대학교 일반대학원석사과정(컴퓨터학과 전공)
- 2019년 ~ 현재 : 세명대학교 일반대학원박사과정(컴퓨터학과 전공)
- 관심분야 : SoC CAD, 임베디드 시스템 신경, 머신러닝

인 치 호(정회원)



- 2024년도 제24권 1월 호 참고