



재실자 정보 기반 실내 CO₂ 농도 예측모델 개발 및 성능 비교

Development of an Indoor CO₂ Prediction Model and Optimal Ventilation Control Algorithm Based on an Occupant Personal Information

배강우* · 김태원** · 변재윤*** · 서민채**** · 문진우*****

Kang Woo Bae* · Tae Won Kim** · Jae Yoon Byun*** · Min Chae Seo**** · Jin Woo Moon*****

* Main author, Graduate Student, Dept. of Architecture Building Science, Chung-Ang Univ., South Korea (rkddn2@gmail.com)

** Coauthor, Graduate Student, Dept. of Architecture Building Science, Chung-Ang Univ., South Korea (xngilsh@naver.com)

*** Coauthor, Graduate Student, Dept. of Architecture Building Science, Chung-Ang Univ., South Korea (byunjaeyoon0320@gmail.com)

**** Coauthor, Graduate Student, Dept. of Architecture Building Science, Chung-Ang Univ., South Korea (alsco0630@gmail.com)

***** Corresponding author, Professor, Dept. of Architecture Building Science, Chung-Ang Univ., South Korea (gilerbert73@cau.ac.kr)

ABSTRACT

Purpose: This study aimed to develop a CO₂ prediction model to enhance indoor air quality (IAQ) management by incorporating real-time occupant-specific characteristics such as activity levels (MET), gender, and BMI. Traditional models typically rely on basic variables like occupant count, often ignoring individual factors that significantly affect CO₂ emissions. By including these variables, the model seeks to improve predictive accuracy and optimize ventilation control systems for enhanced energy efficiency and IAQ management. **Method:** Two predictive models were developed using DNN and GRU machine learning algorithms. One model utilized basic number of occupants and environmental (outdoor CO₂, indoor CO₂, ventilation system flowrate) data, while the enhanced model incorporated MET, gender, and BMI. Both models were trained on mock-up data collected from a controlled environment, including CO₂ concentration, ventilation flow rates, and occupant information. Their performance was evaluated using MAE, CvRMSE, and R². **Result:** The enhanced model, integrating occupant-specific variables, demonstrated significant improvements in prediction accuracy compared to the traditional model, with MAE and CvRMSE values improving by 3.93% and 6.92%, respectively. These results highlight the importance of detailed occupant data for real-time IAQ management and the potential for greater efficiency in ventilation control and energy savings in sustainable buildings.

KEYWORD

실내 공기질
예측모델
재실자 정보

Indoor Air Quality
Prediction Model
Occupants Information

ACCEPTANCE INFO

Received Sep. 20, 2024

Final revision received Oct. 5, 2024

Accepted Oct. 11, 2024

© 2024. KIEAE all rights reserved.

1. 서론

1.1. 연구의 배경 및 목적

현대 사회의 대다수의 사람들은 하루 평균 90%의 시간을 건물 내부에서 보내고 있으며, 이에 따라 재실자 건강을 위한 실내공기질 (indoor air quality, IAQ) 관리의 중요성이 강조되고 있다[1]. 또한, 주요 국가들은 실내 공기질 기준과 규제를 강화하고 있으며, 이를 준수하기 위한 다양한 정책적 노력을 기울이고 있다[2]. 이러한 정책들은 재실자 건강 보호와 더불어 지속 가능한 건물 운영을 가능하게 한다.

실내 공기질 저하의 주요 원인으로는 재실자의 호흡으로 발생하는 CO₂ 축적, 건축 자재에서 방출되는 휘발성 유기 화합물, 외부에서 유입되는 미세먼지 등이 있다[3]. 이러한 오염 물질에 장시간 노출될 경우 재실자의 건강에 악영향을 미칠 수 있으며, 피로, 두통, 집중력 저하 등 다양한 증상을 유발할 수 있다[4]. 실내 CO₂ 농도에 대한 기준은 국가, 건물, 용도별로 최대 5,000ppm 까지 다양하게 제시되고 있는데[5], ASHRAE (American Society of Heating,

Refrigeration and Air Conditioning Engineers)에서는 실외 CO₂ 농도 대비 700ppm 이상 넘어가지 않도록 권장하고 있다. 국내 기준은 실내공기질 관리기준에서 사무공간의 경우 1,000ppm 이하로 유지할 것으로 권고하고 있다. 실내 CO₂ 농도가 높아지게 되면 실내 공기질 저하를 의미하며, 실내 재실자들에게 피로감, 두통, 집중력 저하 등과 같은 증상을 유발할 수 있고, 이는 특히 사무 환경에서 생산성 저하로 이어질 수 있기 때문이다. 또한, 실내 CO₂ 농도는 환기 시스템의 효율성을 평가하는 데도 중요한 변수로 작용하며, 다양한 센서를 통해 실시간으로 모니터링하고 수집하기 용이한 데이터이다. 이러한 맥락에서 본 연구는 실내 CO₂ 농도를 예측모델의 핵심 변수로 선정하였다.

실내 CO₂ 농도를 관리하기 위한 기술 발전으로, 최근 사물인터넷 (Internet of Things, IoT) 기반의 예측 및 제어 시스템이 도입되면서 실내 CO₂ 농도 관리가 더욱 효율적으로 이루어지고 있다. 이러한 시스템들은 실시간 데이터를 활용하여 최적의 환기 전략을 제안함으로써, 에너지 절감과 실내 공기질 개선에 기여하고 있다[6].

이러한 배경에서 본 연구는 실내 공기질 쾌적도를 유지하기 위해 실시간 재실자 수뿐만 아니라 MET, 성별, BMI 개인 정보를 반영한 실내 CO₂ 농도 예측모델을 제안하고자 한다. 이러한 방식은 기존 연

구들이 반영하지 않은 재실자의 개인 특성을 반영하였기 때문에 기존의 예측모델에 비해 더 높은 예측 정확도를 보일 것으로 예상된다. 나아가, 재실자 개인 특성에 맞춘 최적 환기량을 적용함으로써 보다 정교한 실내 공기질 관리와 효율적인 환기시스템 운영을 통해 에너지 낭비를 줄이고, 에너지 효율성을 높여 지속 가능한 건물 관리에 크게 기여할 수 있을 것으로 기대된다.

1.2. 연구의 범위 및 방법

본 연구는 실내 CO₂ 농도를 중심으로 재실자 정보(재실 인원, MET, BMI, 성별)를 반영한 예측모델을 개발하고 그 성능을 평가하는 것을 목표로 한다. MET와 CO₂ 농도를 연계한 이유는 MET가 개인의 신체 활동 강도를 나타내는 지표로서, 재실자의 에너지 소비량을 반영하기 때문이다. 신체 활동이 활발할수록 CO₂ 배출량이 증가하므로, MET는 정확한 CO₂ 농도 예측에 중요한 변수로 작용한다. 예를 들어, 앉아 있는 상태에서의 CO₂ 배출량과 걷는 상태에서의 배출량은 크게 차이가 나기 때문에, MET 값을 포함함으로써 보다 정밀한 CO₂ 농도 예측이 가능해진다. 다음으로 BMI와 성별 또한 중요한 변수로 고려하였다. BMI는 개인의 신체 크기와 체중을 반영하며, 호흡량과 CO₂ 배출량에 영향을 미친다. 연구 결과에 따르면 BMI가 높은 사람일수록 호흡 활동이 활발해져 더 많은 CO₂를 배출하는 경향이 있다. 따라서 BMI를 예측모델에 포함시킴으로써 개인별로 상이한 CO₂ 배출량을 보다 정밀하게 반영할 수 있다. 성별 역시 호흡량에 차이를 보이는 변수로, 일반적으로 남성의 CO₂ 배출량이 여성보다 높게 나타난다. 이러한 특성을 모델에 반영함으로써 예측 성능을 더욱 높일 수 있을 것이다.

실내 CO₂ 농도 예측 성능을 평가하기 위해 실제 1인 사무공간 크기와 유사한 mock-up을 제작하여 실험을 수행하였다. 원활한 실험 진행을 위해 실내외 CO₂ 농도, MET, 성별, BMI, 환기시스템의 풍량 등을 모니터링 할 수 있는 영상 및 IoT 센서를 설치하여 실시간 데이터를 수집하였다. 여기서 환기시스템 풍량은 4단계(미풍, 약풍, 중풍, 강풍)로 구성된다. 구축된 실험 환경에서 재실자 조건을 다양하게 변화시키며 수집된 데이터를 모니터링하고 분석하였다. 해당 데이터를 바탕으로, 기계 학습 알고리즘을 활용하여 실내 CO₂ 농도 변화 패턴을 학습시킨 재실자 정보가 포함되지 않은 기존 방식의 예측모델과 재실자 정보를 포함한 예측모델 두 가지를 개발하였다. 개발된 예측모델들은 다양한 통계적 기법을 사용하여 성능을 평가하였으며, 각각의 기존 방식 예측모델과 재실자 정보를 포함한 예측모델 중에서 가장 우수한 성능을 보이는 모델을 선정하였다. 이후, 재실자 개인 정보를 반영한 예측모델이 기존 모델 대비 얼마나 성능이 개선되었는지를 실험적으로 검증하였다.

2. 기존 연구 및 이론 고찰

2.1. 기존 연구 분석

최근 진행된 실내 CO₂ 농도 환기시스템 제어 연구에서 Taheri & Razban(2021)[7]은 수요 예측 환기(DCV) 시스템을 위한 실내 CO₂ 농도 예측모델을 개발하였다. 이 모델은 온도, 습도 등의 환경

데이터를 활용하여 실내 CO₂ 농도를 실시간으로 예측하며, 이를 통해 환기시스템이 재실자의 수에 따라 자동으로 조정되었다. 그 결과, 에너지를 절감하면서도 쾌적한 실내 공기질을 유지하는 데 효과를 나타냈다. 또한, Segala & Giacomo(2021)[8]은 제한된 환경 데이터를 활용한 실내 CO₂ 농도 예측을 위한 신경망 모델을 제안하였다. 이 연구는 일주일간의 데이터 수집만으로도 높은 예측 정확도를 달성하였고, HVAC 제어 시스템에 실시간으로 적용될 수 있음을 입증하였다. 시간이 지남에 따라 예측 정확도는 더욱 향상되어 장기적인 실내 공기질 관리에 기여할 수 있음이 확인되었다.

또한, IEA EBC Annex 79 프로젝트에서는 재실자의 행동을 모델링하여 건물의 에너지 소비와 실내 환경의 질을 최적화하기 위한 다양한 연구를 진행해 왔다. 이 프로젝트에서는 재실자의 유무, 인원, 밀도와 같은 정보를 활용하여 환기량을 조절하는 방법이 개발되었으며, 이러한 재실자 중심 제어(occupant-centric control, OCC)의 적용은 건물의 열환경 및 환기 효율성을 향상시키는 데 기여하였다 [9-12]. OCC는 건물 내 재실자들의 행동 패턴을 반영하여 설비 시스템의 제어를 정교하게 하고, 불필요한 에너지 낭비를 줄이는 데 효과적이었다. 그러나 기존 연구들은 재실자 정보의 일부분만 반영하고 있으며, 재실자의 활동량(metabolic equivalent task, MET), 성별, 체질량지수(body mass index, BMI)와 같은 개인 특성들을 충분히 고려하지 않은 한계가 있었다.

2.2. 이론 고찰

BMI와 호흡량 사이에는 선형적인 관계가 존재하는데, BMI가 높을수록 호흡 활동이 활발해지면서 더 많은 CO₂를 배출하게 된다는 연구 결과가 있다[13]. 다양한 연령과 체지방율을 가진 피실험자들을 대상으로 한 CO₂ 배출량 실험에서도, BMI가 높은 인원은 일수록 CO₂ 배출량이 증가하며, 성별 차이로는 남성이 여성보다 더 많은 CO₂를 배출하는 것으로 나타났다[14]. 이는 재실자의 신체적 특성이 CO₂ 농도에 중요한 영향을 미친다는 것을 보여준다.

재실자의 활동량 역시 CO₂ 배출에 큰 영향을 미친다. 활동량이 많을수록 에너지 소비가 증가하고, 이에 따라 호흡 빈도가 높아져 CO₂ 배출량도 증가한다. 예를 들어, 사무실에서 주로 앉아 있는 활동보다, 체육관이나 판매시설과 같이 활동량이 많은 환경에서는 CO₂ 농도가 더 빠르게 증가하는 경향이 있다[15]. 이러한 활동 정보가 반영되지 않으면, 환기시스템의 정밀한 제어가 어려우며, 이는 실내 공기질 악화와 에너지 낭비로 이어질 수 있다.

따라서, 본 연구는 이러한 기존 연구의 한계를 극복하기 위해, 재실자의 개인 특성(BMI, 성별, MET 등)을 반영한 새로운 CO₂ 농도 예측모델을 제안하고자 한다. 이러한 접근은 기존 예측 모델 대비 더 높은 예측 정확도를 기대할 수 있으며, 장기적으로 보다 효율적인 실내 공기질 관리와 최적의 환기 전략을 수립하는 데 기여할 것이다.

3. 연구 방법

3.1. Mock-up 구축

1) Mock-up 개요

Mock-up은 다수의 인원이 재실하고 있는 공간에서 공기질 관련 실험을 위해 환기시스템을 동의없이 강제로 제어하는 것은 재실자 건강에 악영향을 미칠 수 있다. 따라서, 본 연구에서는 통제된 환경에서의 데이터 취득과 성능평가를 위해 환경센서, 영상센서, 환기시스템이 설치된 mock-up을 구축하였다. Mock-up의 개요는 Table 1.과 같고, 구축 사진은 Fig. 1.과 같다. 설치 위치는 서울 동작구에 위치한 'C' 대학교 2층이다. 내부 면적은 정부청사관리규정 시행규칙 제274호 제2조 2항 '청사 취득 및 배경면적' 기준에 따라 업무 전용공간인 3.54㎡의 기준과 유사한 2.6m(D) * 1.35m(W) * 2.4m

Table 1. Mock-up information

Information	Details
Plan	
Location	'C' University, Seoul, South Korea
Scale	Floor area : 3.51㎡
Outshell	Uninterruptible clear panel
Ocupants	1 person
Ventilation System	Type 1 ventilation system
Variables	<ul style="list-style-type: none"> Outdoor & Indoor sensors <ul style="list-style-type: none"> - Drybulb temperature [°C] - Relative humidity [%] - PM₁₀ [μg/m³] - PM_{2.5} [μg/m³] - CO₂ [ppm] Ventilation system <ul style="list-style-type: none"> - Flowrate [CMH] CCTV <ul style="list-style-type: none"> - Number of occupants [person] - Occupant activity level [met] - BMI [kg/m²] - Gender [0,1,2]



Fig. 1. Mock-up setup

(H) 규모이다. 바닥을 제외한 모든 외피는 재실자 정보 관측이 가능한 투명한 무정전 패널이며, 바닥은 MDF로 구성하였다. Mock-up은 실내에 위치하며, 실과 mock-up 사이 공간에는 서버 및 관리자가 위치한다.

2) Mock-up 실험 개요 및 피실험자 정보

2월 1일부터 2월19일까지 mock-up을 통한 데이터 수집을 진행하였으며, 남성 5명과 여성 5명, 총 10명의 피실험자를 선정하였다. 다만 피실험자 E, J의 경우 학습용 데이터 취득이 아닌 예측모델 성능평가를 위한 데이터를 취득하였다. 실험 기간 및 피실험자 정보는 Table 2.와 같다.

피실험자는 표준화된 데이터를 확보하여 실험 결과의 신뢰성을 높이고자 세계보건기구 기준 BMI 18.5~25의 정상 범위에 속하는 20, 30대 인원으로 선정하였다. 선정된 인원들은 Fig. 2.와 같이 mock-up 내부에서 환기시스템 가동 상태(미가동, 미풍, 약풍, 중풍, 강풍)별로 앉기 (met 1.0), 서기 (met 1.2), 걷기 (met 1.7) 3가지 행동을 각각 40분씩 총 15회 수행하였다.

3.2. 환기시스템 구성 및 제어

환기시스템은 협소한 mock-up 내부공간으로 인해 기성 제품을 설치가 불가하여 직접 제작하였으며, 구성은 Fig. 3.과 같다. 급배기를 위한 2개의 팬과 제어 모드(off, 외기도입, 내부순환) 구현을 위한 4개의 댐퍼를 설치하였으며 개별 제어가 가능하도록 구성하였다.

②, ③번 위치에 설치된 급배기 팬(BFM17051B)은 24V DC 팬으로 크기는 170mm (D) * 170mm (W) * 51mm (H)이다. 정적 압력이 13.54 mmH₂O 일 때, 최대 377.4 CMH의 풍량을 발생시킨다. 제어 모드를 결정하는 댐퍼는 총 4대로 ①-1~4에 해당하며, 제어 모드별 댐퍼작동은 다음과 같다.

Table 2. Participant information and experiment duration

ID	Sex	Height (cm)	Weight (kg)	Age	BMI (kg/m ²)	Participant period
A	M	181	75	32	22.89	2/12
B	M	171	55	28	18.81	2/13, 15
C	M	173	72	26	24.06	2/1, 2
D	M	176	70	24	22.6	2/8
E	M	168	63	33	22.31	3/13, 18, 19
F	F	157	45	28	18.26	2/14, 16
G	F	158	47	25	18.1	2/18, 19
H	F	159	48.5	28	19.18	2/3, 12
I	F	162	50	28	19.04	2/6, 7
J	F	161	52	25	20.06	3/17, 20, 24

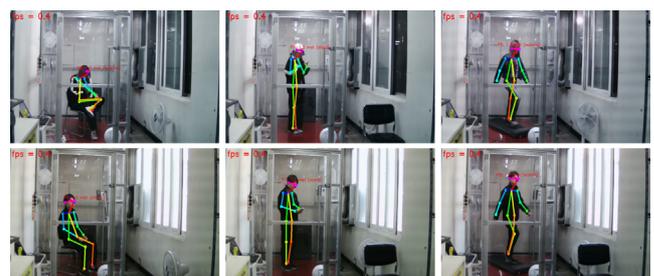


Fig. 2. Occupant activity level experiment

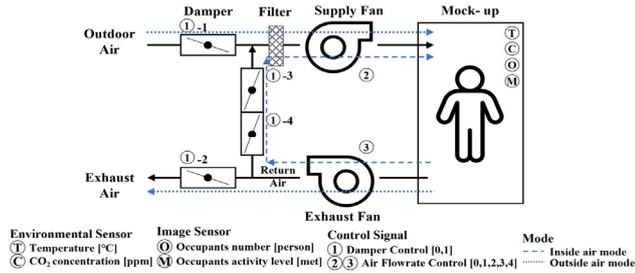


Fig. 3. Ventilation system configuration diagram

- 시스템 미작동(off): ①-1~4 모든 댐퍼 off
- 외기 도입모드: ①-1, 2 댐퍼 on, ①-3, 4 댐퍼 off
- 내부 순환모드: ①-1, 2 댐퍼 off, ①-3, 4 댐퍼 on

3.3. 학습 모델 선정

다양한 인공신경망 종류가 있지만 현재 건축 및 환경 분야에서 주로 사용되는 모델은 심층 신경망 모델(Deep Neural Network, DNN)과 게이트 매커니즘이 적용된 GRU(Gated Recurrent Units) 모델이다[16]. DNN은 복잡한 비선형 함수 처리에 용이하며, 입·출력 변수 간 연속성, 범주형 변수에 국한되지 않고 분석이 가능하기 때문에 비선형적 특징을 나타내는 데이터 학습에 유리하다[17]. GRU는 시계열 데이터 학습에 유리한 것으로 보고되고 있으며, 건물 환경 및 에너지 관련 데이터는 대부분 시계열 데이터로 표현된다[18]. GRU는 학습 및 예측 과정에서 과거 데이터(historical data)를 효과적으로 반영하여 시퀀스 데이터에서 발생하는 패턴 파악에 강점을 나타낸다[19]. 학습 시간 측면에서 DNN이 GRU보다 빠르지만, GRU는 시계열 데이터 처리에 우수한 성능을 보이기 때문에, 본 연구에서는 학습 시간이 가장 빠른 DNN과 시계열 데이터 처리에 강점이 있는 GRU를 학습 모델로 선정하였다. LSTM은 학습 시간이 GRU보다 길기 때문에 본 연구에서는 사용하지 않았다.

3.4. 입출력 데이터 선정

예측모델 성능확보를 위해서는 적절한 입력변수가 정의되어야 한다. 이를 위해 본 연구에서는 통계적 기법에 기반한 상관관계분석을 실시하였다. 상관관계분석에서 주로 사용되는 방법은 피어슨 상관관계수(r)로 -1과 1사이의 상관계수를 가진다[20]. -1에 가까울수록 음의 상관관계를, 1에 가까울수록 양의 상관관계를 의미하며, 0은 변수 간 상관관계가 없음을 나타낸다.

취득한 원본 데이터에 대한 상관관계분석 결과는 Fig. 4., 이에 따라 최종적으로 선정된 입력변수 및 출력변수 목록은 Table 3.과 같다. Fig. 4.에 의하면 0.5 이상의 상관계수를 가진 변수는 재실자 수, MET, BMI, gender가 있으며, 상관계수가 0.3 이상인 변수는 SA (supply air flowrate), RA (return air flowrate)로 나타났다. SA와 RA의 상관계수는 0.35로 낮은 편이지만, 기존 연구된 예측모델의 입력변수로 사용되어서 본 연구에서도 동일하게 선정하였다. 실외 CO₂ 농도는 상관계수가 0에 가깝지만 환기시스템이 외기를 도입하기 때문에 영향성을 고려하여 포함하였다. 따라서, 입력변수는 실내외 CO₂ 농도, SA, RA, 재실자 수, MET, BMI, 성별로 총 8가지이며,

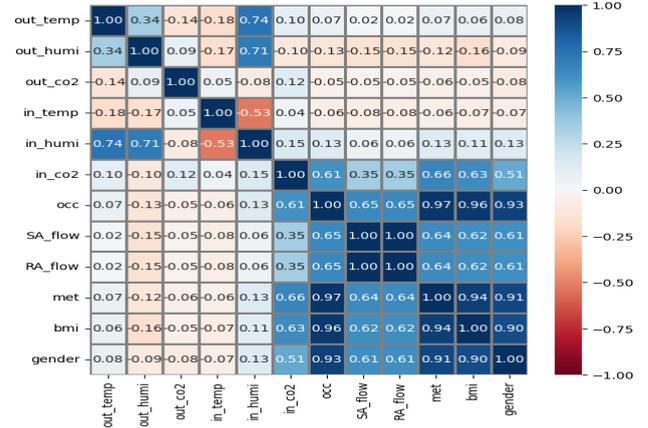


Fig. 4. Pearson correlation analysis

Table 3. Input and output variables for indoor CO₂ concentration prediction model

Type	Variables (abbreviation)
Input Variables	• Outdoor CO ₂ Concentration [ppm] (out_co2)
	• Indoor CO ₂ Concentration [ppm] (in_co2)
	• SA (Supply Airflow) [CMH] (SA_flow)
	• RA (Return Airflow) [CMH] (RA_flow)
	• Number of occupants [person] (occ)
	• MET [met] (met)
Output Variable	• BMI [kg/m ²] (bmi)
	• Gender [0, 1, 2] (gender)
	• Indoor CO ₂ Concentration After 5 Minutes [ppm]

출력변수는 제어주기인 5분 후 실내 CO₂ 농도이다.

본 연구에서는 두 가지 예측모델을 개발하였다. 기존 방식 모델의 입력변수는 실내외 CO₂ 농도, SA, RA, 재실자 수로 총 5가지이며, 성능 개선 모델은 기존 방식 모델의 입력변수에 더해 MET, BMI, 성별을 추가하여 총 8가지 입력변수를 사용하였다. 두 모델의 출력변수는 동일하게 제어주기인 5분 후 실내 CO₂ 농도로 설정하였다. 기존 방식 모델과 성능 개선 모델의 차이는 추가적인 재실자 정보 (MET, BMI, 성별)의 포함 여부로, 이들 변수들이 예측 성능에 미치는 영향을 평가하기 위해 두 모델을 비교 분석하였다.

3.5. 학습데이터 수집

실내 CO₂ 농도 예측모델을 개발 및 성능평가를 위해서 실내외 환경변수, 재실자 정보, 그리고 환기시스템 가동 상태에 대한 실시간 데이터 취득이 필수적이다. 이를 위해 mock-up 내부 벽면 1.2m 높이에 환경센서(AW-1008KP) 3개를 설치하였다. 센서를 통해 온도, 습도, CO₂, PM₁₀, PM_{2.5}가 취득되고 본 실험에서는 CO₂ 데이터만 사용하였으며, 순시값 및 5분 평균값을 DB에 저장하였다.

실내 재실자 데이터 수집을 위해 mock-up 내외부에 웹캠 (STWC-500, HW20)을 설치하여 재실자 수, MET, BMI, 성별 정보를 취득하였다. 재실자 검출에는 기존 연구에서 제안된 모델을 사용하였으며, 이 모델은 영상센서를 활용해 재실자를 검출하고, 활동인지 방법론으로 MET 값을 산출한다[21]. MET의 평균값을 사용할 경우 오차가 발생하였으나, 최빈값을 적용할 때 100% 정확한 대

뜻값을 얻을 수 있었다. 이 모델은 시뮬레이션 및 mock-up 실험을 통해 검증되었다[22, 23]. 본 연구에서는 이 모델을 서버 컴퓨터에 적용하고, 영상센서를 설치하여 재실자 정보를 실시간으로 감지하고, 30초 뒤 환기시스템과 연계해 제어되도록 설정하였다. 재실자 수는 10~15초마다 저장되며, MET 값은 5분 동안의 최빈값이 기록된다.

3.6. 데이터 전처리

데이터 전처리는 학습데이터에 대한 결측값 및 노이즈 현상 제거부터 학습 알고리즘에 입력시키기 위한 구조화 및 스케일링까지 일련의 과정을 의미한다. 적절한 데이터 전처리가 시행되지 않는 경우, 예측모델의 성능이 좋지 않거나, 학습이 진행되지 않는 예외가 발생할 수 있다[24]. 본 연구에서는 mock-up에서 취득한 학습데이터에 결측값 및 노이즈 현상이 발생하지 않았기 때문에 데이터 구조화와 스케일링 처리를 진행하였다. 데이터 구조화는 학습모델인 DNN과 GRU에 맞추어 변환하였다. DNN 모델은 학습데이터와 동일한 2차원 행렬 구조를 입력받으며 n번째의 입력변수로 n+1번째 출력변수를 예측하는 구조이다.

다음은 GRU 학습을 위한 3차원 데이터 배열 형태로 데이터를 구조화하는 과정이다. 먼저, n+1번째 출력변수를 예측하기 위해 n - (window size - 1)부터 n번째 입력변수를 사용하고, horizon 값을 1로 지정하여 1개의 출력변수를 생성하였다. 이후 입력 데이터를 나누기 위해 window size를 2, 3, 4로 설정하여 데이터 구조화를 진행하였다.

데이터 스케일링 및 분할 과정도 진행되었다. 일반적으로 정규화(normalization)와 표준화(standardization) 방법이 사용되며, 본 연구에서는 각 입력 변수의 최솟값과 최댓값을 0과 1로 변환하는 min-max normalization을 적용하였다. 이후, 예측모델의 학습(train), 검증(validation), 테스트(test) 단계를 위해 정규화된 데이터를 8:1:1 비율로 분할하여 각각 2,670개, 333개, 333개로 구분하였다.

3.7. 실내 CO₂ 예측모델 개발 및 평가 방법

실내 CO₂ 예측모델 개발 및 평가를 위해, 먼저 하이퍼파라미터 최적화 과정이 진행되었다. 하이퍼파라미터는 머신러닝 모델의 구조와 학습 방법을 결정하는 중요한 변수로, 뉴런 수, 은닉층, 드롭아웃 비율, 활성화 함수 등이 포함된다[25]. 예측모델의 성능은 하이퍼파라미터의 조합에 따라 크게 달라지며, 최적의 성능을 얻기 위해서는 하이퍼파라미터 튜닝이 필수적이다[26]. 본 연구에서는 탐색 효율성과 시간을 고려하여 베이지안 최적화 방법을 채택하였으며, 하이퍼파라미터 변수는 Table 4.와 같다. Keras-Tuner에서 제공하는 Bayesian Optimization 라이브러리를 활용하였다. 학습에 사용된 손실 함수는 평균제곱오차(Mean Squared Error, MSE), 최적화 함수로는 Adam optimization이 사용되었다. 학습횟수는 100회(epoch)로 설정되었으며, 과적합을 방지하기 위해 5회 이상 성능 개선이 없을 시 학습을 종료하는 early stopping 기능을 적용하였다. GRU 모델의 경우 window size도 주요 하이퍼파라미터로 지정될

Table 4. Input and output variables for indoor CO₂ concentration prediction model

Hyperparameter Variables	Range	
DNN	Neuron	• 32~512, step = 32
	Hidden layer	• 0~5
	Dropout	• 0~0.5, step = 0.1
	Activation function	• ReLU, sigmoid, tanh
	Loss function	• MSE
	Optimization function	• Adam optimization
	Max trial	• 20
	Epoch	• 100
GRU	Training option	• Early stopping = 5
	Window size	• 2, 3, 4

수 있으나, 본 연구에서는 고정값을 적용하였다.

예측모델 성능평가는 MAE(Mean Absolute Error), CvRMSE (Coefficient of Variation of the Root Mean Squared Error), R²(Coefficient of Determination) 지표를 사용하여 이루어졌다. MAE는 예측값과 실제값의 절대 오차의 평균을 나타내어 모델의 예측 오차의 크기를 평가하고, CvRMSE는 RMSE를 실제값의 평균으로 나누어 변동성을 평가하는 지표이다. R²는 모델이 실제값의 변동성을 얼마나 잘 설명하는지를 나타내며, 값이 1에 가까울수록 모델의 성능이 우수함을 의미한다. 이를 통해 예측모델의 정확성과 신뢰성을 검증하였다(Eq. 1~3).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - t_i| \tag{Eq. 1}$$

$$CvRMSE = \frac{1}{\bar{Y}_i} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - t_i)^2} \tag{Eq. 2}$$

$$R^2 = 1 - \sum_{i=1}^n \frac{(y_i - t_i)^2}{(y_i - \bar{Y}_i)^2} \tag{Eq. 3}$$

where n: the number of data point (i), t_i: estimated value of model for data point (i) and y_i: actual value for data point (i) and \bar{Y}_i : mean value of the actual values

4. 예측모델 성능 비교분석

4.1. 예측모델 학습 결과

예측모델의 학습 결과는 학습데이터 중 학습에 사용되지 않은 테스트(test) 데이터를 사용하여 평가하였다. 기존 방식 모델 및 성능 개선 모델의 하이퍼파라미터 최적화 변수 및 예측모델 성능은 Table 5.와 같이 나타났다.

먼저 기존 방식 모델의 DNN 모델은 320, 32, 144개의 뉴런으로 구성된 총 3개의 은닉층 가진다. 학습률(learning rate)은 0.001이

며, 활성화 함수는 ReLU를 사용하였다. 성능은 MAE 48.11ppm, CvRMSE 15.47%, R² 0.7970으로 나타났다. GRU 모델은 가장 성능이 우수한 모델은 window size가 2일 때 2개의 은닉층에 288, 448개의 뉴런을 가지며 학습률은 0.001, 활성화 함수 Tanh가 결정된 모델이다. 성능은 MAE, CvRMSE, R² 각각 36.42ppm, 23.40%, 0.5356으로 나타났다. 가장 성능이 낮은 모델은 window size가 4일 때 MAE, CvRMSE, R²가 48.19ppm, 25.59%, 0.4447로 나타났다. window size가 늘어남에 따라 성능이 저하되었으며, 이는 예측모델이 과거 데이터의 영향을 거의 받지 않음을 의미한다.

성능 개선 모델의 DNN은 512, 128, 176개의 뉴런으로 구성된 3개의 은닉층을 포함하고 있다. 학습률은 0.001이며, 활성화 함수로는 ReLU를 사용하였다. 해당 모델의 성능은 MAE 46.22ppm, CvRMSE 14.40%, R² 0.8242로 나타났다. GRU 모델의 경우, 성능이 가장 우수한 모델은 window size가 2일 때로, 384개와 512개의 뉴런을 가진 두 개의 은닉층을 포함하며, 학습률은 0.001, 활성화 함수로는 Tanh가 결정된 모델이다. 이 모델의 성능은 MAE 32.97 ppm, CvRMSE 21.49%, R² 0.6083로 나타났다. 반면, 성능이 가장 낮은 모델은 window size가 4일 때로, MAE 38.03ppm, CvRMSE 22.27%, R² 0.5793의 성능이 나타났다. 이를 종합하면, 성능 개선 모델이 기존 방식 모델에 비해 전반적으로 높은 예측 정확도를 보였으며, 특히 DNN 기반 모델의 경우 개선된 성능이 두드러졌다.

4.2. 예측모델간 성능 비교 분석

두 모델 모두 DNN을 기반으로 개발된 예측모델이 가장 우수한 성능을 보였다. 따라서, 본 연구에서는 기존 방식 모델과 성능 개선 모델 간의 성능 비교를 위해 DNN 모델을 사용하였다. 예측값 산출을 위해 3월 13, 17, 18, 19일 총 4일간 피실험자 E, J를 통해 데이터를 취득하였으며, 모델간 비교 그래프는 Fig. 5. (a), (b)이며, 성능 비교는 Table 6.과 같다.

전체 기간의 모델간 성능평가 결과 기존 방식 모델의 MAE는 62.33ppm, CvRMSE는 19.83%, R²는 0.8709이며, 성능 개선 모델의 MAE, CvRMSE, R²는 각각 45.30ppm, 18.73%, 0.8762로 나타났다. 하지만 전체 기간의 경우 실내 CO₂ 농도가 400~500ppm을 유지하고 있는 구간이 대부분을 차지하고 있어 실내 CO₂ 농도가 급격히 변화하는 구간의 정확한 성능평가 결과를 판단하기 어렵다. 따

라서 예측모델의 정확한 성능 파악을 위해 timestep에서 재실자 행동이 발생한 [100-200], [240-340], [580-630]의 세가지 구간으로 구분하였고, 각각 구간에 대해 성능평가를 실시하였다.

1) [100-200] 구간

먼저, [100-200] 구간은 2명의 피실험자 모두 환기시스템이 미가동인 상태에서 걷기, 앉기, 서기 행동을 40분씩 진행하였다. 그 결과, 기존 방식 모델의 경우 MAE 143.98ppm, CvRMSE 24.38%, R² 0.7979로 나타났으며, 성능 개선 모델은 MAE, CvRMSE, R² 각각 130.89ppm, CvRMSE 24.22%, 0.7967로 성능 개선 모델이 좀 더 우수한 예측 정확도를 나타냈다. 걷기 행동시 피실험자 E의 BMI 대비 더 많은 CO₂를 배출하였기 때문에 최대치까지 예측하지 못한 것으로 확인되며, 이는 이후 다양한 재실자 데이터를 학습시킬시 충분히 개선될 수 있을 것으로 판단된다. 피실험자 J의 경우, 기존 방식 모델은 재실자 수만을 활용하여 각 행동마다 비슷한 수준의 실내 CO₂ 농도를 예측했으며, 예측값이 1,200~1,300ppm 범위 내에서만 산출된 것을 확인할 수 있다. 성능 개선 모델의 경우 MET, BMI, 성별 등의 입력 변수가 추가되었기 때문에 실측값의 추세를 잘 따라가는 것을 확인할 수 있으며, 서기 행동 시 발생한 오차는 데이터 부족에 기인한 것으로, 추가적인 학습을 통해 성능 개선이 가능할 것으로 판단된다.

2) [240-340] 구간

[240-340] 구간은 환기시스템을 가동한 상태에서 피실험자 E의 실험 결과를 나타낸다. 피실험자 E의 걷기 예측 정확도가 상대적으로 낮았기 때문에, 환기시스템 적용시 예측 정확도 확인을 위해 두 차례 추가 실험이 진행되었다. 그 결과 기존 방식 모델은 MAE 137.59ppm, CvRMSE 27.38%를 나타냈고, 성능 개선 모델은 그래프 상에서 실제값과 잘 부합하는 형태를 보였으나, 실험 시작 시점과 종료 시점에서 실내 CO₂ 농도 예측값이 300ppm보다 낮게 예측되어 정확도가 떨어진 것을 확인했다. 이 문제는 이후 실험에서 최소 CO₂ 농도 설정을 통해 방지할 예정이다. 추가 실험을 포함한 전체 구간의 예측 정확도를 비교한 결과, 기존 방식 모델의 경우 MAE는 122.07ppm, CvRMSE 23.90%, R² 0.6848로 나타났으며, 성능 개선 모델은 각각 MAE 93.44ppm, CvRMSE 22.53%, R² 0.7699로, 성능 개선 모델이 더 우수한 예측 정확도를 보였다.

Table 5. Input and output variables for indoor CO₂ concentration prediction model

Model		Structure	Learning rate	Activation function	MAE	CvRMSE	R ²	
Baseline	DNN	8-320-32-144-1	0.001	ReLU	48.11 ppm	15.47%	0.7970	
	GRU	Window size = 2	8-288-448-1	0.001	Tanh	36.42 ppm	23.40%	0.5356
		Window size = 3	8-32-512-1	0.001	Tanh	49.58 ppm	23.89%	0.5160
		Window size = 4	8-256-416-1	0.001	Tanh	48.19 ppm	25.59%	0.4447
Improved	DNN	8-512-128-176-1	0.001	ReLU	46.22 ppm	14.40%	0.8242	
	GRU	Window size = 2	8-384-512-1	0.001	Tanh	32.97 ppm	21.49%	0.6083
		Window size = 3	8-320-512-1	0.001	Tanh	33.45 ppm	21.58%	0.6051
		Window size = 4	8-256-512-1	0.001	Tanh	38.03 ppm	22.27%	0.5793

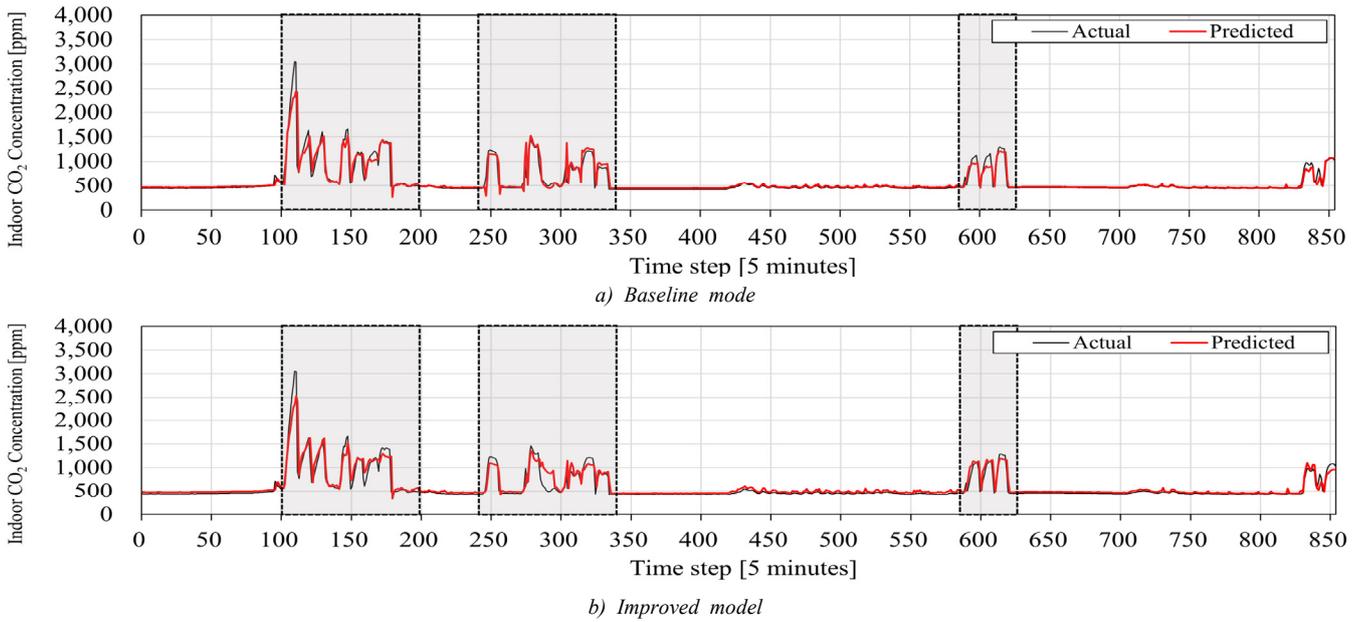


Fig. 5. Model comparison graph

Table 6. Prediction model results

Model	Data segment											
	All			100-200			240-340			580-630		
	MAE [ppm]	CvRMSE	R ²	MAE [ppm]	CvRMSE	R ²	MAE [ppm]	CvRMSE	R ²	MAE [ppm]	CvRMSE	R ²
Baseline	62.33	19.83%	0.8710	143.98	24.38%	0.7979	122.07	23.90%	0.6848	92.39	17.37%	0.8664
Imporved	45.30	18.73%	0.8762	130.89	24.22%	0.7967	95.44	22.53%	0.7699	80.69	14.99%	0.8787

3) [580-630] 구간

[580-630] 구간은 피실험자 J의 실험결과이다. 피실험자 J는 앉기, 서기, 걷기순으로 실험을 진행하였으며, 예측 정확도는 기존 방식 모델의 MAE, CvRMSE, R²는 각각 92.39ppm, 17.37%, R² 0.8664로 나타났고, 성능 개선 모델은 MAE는 80.69ppm, CvRMSE는 14.99%, R²는 0.8787로 세 구간중 가장 우수한 성능을 나타냈다. 비록 성능 개선 모델의 정확도 수치가 더 우수하다고 할 수 있으나, 피실험자 J의 BMI 수치가 학습된 모델의 18.1에서 19.18의 스케일러 범위를 벗어났기 때문에, 학습 범위를 벗어난 입력값에 대해서는 실제 값보다 현저히 낮은 예측값을 보였다. 특히 서기 행동에서 CO₂ 배출량이 상대적으로 많아, 예측값이 실제 CO₂ 농도에 미치지 못한 것으로 나타났다. 이는 모델이 학습되지 않은 범위의 데이터를 처리하는 데 있어 한계를 가지고 있음을 보여준다. 이와 같은 한계는 모델의 일반화 가능성에도 중요한 시사점을 제공한다. 다양한 사용자 또는 재실자의 생리적 변화를 반영하기 위해서는, BMI와 같은 변수의 학습 범위를 확장할 필요가 있다. 추후 연구에서는 다양한 BMI와 MET 조합을 반영한 학습 데이터셋을 구축하여 모델의 예측 범위를 넓히고, 이를 통해 일반화 성능을 개선하고자 한다. 추후 다양한 범주 데이터 수집을 통해 모델 적용성을 높힐 예정이다.

5. 결론

지속 가능한 건물 운영과 재실자 건강 보호를 위해 환기시스템의 효율적인 제어에 대한 관심이 높아지면서, 재실자 중심 제어가 주목받고 있다. 기존의 실시간 재실자 정보는 주로 재실자 유무와 수만을 활용하여 개인 특성을 충분히 반영하지 못했다. 본 연구에서는 인공지능 기계 학습 모델인 DNN과 GRU를 활용해 두 가지 실내 CO₂ 농도 예측 모델을 개발하였다.

기존 방식 모델은 재실자 수만을 변수로 사용한 예측 모델이며, 성능 개선 모델은 MET, 성별, BMI와 같은 재실자 개인 특성을 추가한 모델이다. 두 모델 모두 동일한 실내외 환경 변수와 환기시스템 데이터를 사용했으며, 출력 변수는 5분 후의 실내 CO₂ 농도이다.

예측모델 성능 평가 결과, 기존 방식 및 성능 개선 모델 모두 GRU 다 DNN 기반 모델이 우수한 성능을 보였다. 기존 방식 모델은 MAE 48.11ppm, CvRMSE 15.47%, R² 0.7970, 성능 개선 모델은 MAE 46.22ppm, CvRMSE 14.40%, R² 0.8242로, 성능 개선 모델이 기존 방식 모델 대비 각각 3.93%, 6.92% 더 나은 성능을 보였다. 또한, 학습에 사용되지 않은 두 명의 피실험자 데이터를 추가 분석한 결과, 성능 개선 모델은 MAE에서 최소 9.09%, 최대 21.82%, CvRMSE에서 최소 0.66%, 최대 13.70%의 성능 개선을 보였다.

본 연구의 한계는 재실자 데이터 다양성 부족과 최적 풍량 제어가 충분히 구현되지 못한 점, 그리고 평가 기간이 짧았다는 점이다. 그

럼에도 불구하고 성능 개선 모델은 기존 방식 모델보다 예측 성능이 우수했다. 이를 통해 성능 개선 모델은 기존 방식보다 재실자 특성을 반영하여 더욱 정교한 실내 CO₂ 농도 예측이 가능함을 보여주며 향후 연구에서는 다양한 재실자 특성 및 풍량 조건을 반영한 최적 환기 제어 알고리즘 개발과 기존 방식과의 비교 실험을 통해 더 효율적인 환기시스템 제어 방식을 제안할 계획이다.

Acknowledgement

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다(No. RS-2021-KP002461).

References

- [1] A. Cincinelli, T. Martellini, Indoor air quality and health, *International Journal of Environmental Research and Public Health*, 14(11), 2017.10, 1286.
- [2] K.E. Paleologos, M.Y.E. Selim, A.M.O. Mohamed, Indoor air quality: Pollutants, health effects, and regulations, In: Mohamed AMO, Paleologos EK, Howari FM, editors, *Pollution assessment for sustainable practices in applied sciences and engineering*, Oxford, UK: Butterworth-Heinemann, 2021, pp.405-489.
- [3] K. Prashant, B. Imam, Footprints of air pollution and changing environment on the sustainability of built infrastructure, *Science of the Total Environment*, 444, 2013.02, pp.85-101.
- [4] A.J. Koivisto et al., Source specific exposure and risk assessment for indoor aerosols, *Science of the Total Environment*, 668, 2019.06, pp.13-24.
- [5] S.D. Lowther et al., Low level carbon dioxide indoors—A pollution indicator or a pollutant? A health-based perspective, *Environments*, 8(11), 2021.11, 125.
- [6] G. Marques, C.R. Ferreira, R. Pitarma, Indoor air quality assessment using a CO₂ monitoring system based on internet of things, *Journal of Medical Systems*, 43, 2019, 67.
- [7] S. Taheri, A. Razban, Learning-based CO₂ concentration prediction: Application to indoor air quality control using demand-controlled ventilation, *Building and Environment*, 205, 2021.11, 108164.
- [8] G. Segala et al., A practical and adaptive approach to predicting indoor CO₂, *Applied Sciences*, 11(22), 2021.11, 10771.
- [9] 최은지 외 3인, 재실자 활동량 산출을 위한 Pose 분류 모델 개발, 한국생태환경건축학회지, 제18권 제5호, 2018.10, pp.93-98. // (E.J. Choi et al., Development of a human Pose classifying model to estimate the metabolic rate of occupant, *Journal of KIEAE*, 18(5), 2018.10, pp.93-98.)
- [10] 최은지 외 5인, 실내 열환경 쾌적 조성 및 에너지 효율 극대화를 위한 지능형 재실자 활동량 산출모델 개발, 대한설비공학회 하계학술발표대회논문집, 2019.06, pp.704-707. // (E.J. Choi et al., Development of an intelligent MET estimation model for indoor thermal comfort and energy efficiency, *Conference Journal of SAREK*, 2019.06, pp.704-707.)
- [11] B.R. Park et al., Accuracy analysis of DNN-based pose-categorization model and activity-decision algorithm, *Energies*, 13(4), 2020.02, 839.
- [12] E.J. Choi et al., Development of occupant pose classification model using deep neural network for personalized thermal conditioning, *Energies*, 13(1), 2019.12, 45.
- [13] C.K. Lin, C.C. Lin, Work of breathing and respiratory drive in obesity, *Respirology*, 17(3), 2012.04, pp.402-411.
- [14] L. Yang et al., Carbon dioxide generation rates of different age and gender under various activity levels, *Building and Environment*, 186, 2020.12, 107317.
- [15] 최영재 외 4인, 실내 이산화탄소 농도 예측 기반 적응형 환기 제어 알고리즘 개발, 한국생태환경건축학회지, 제23권 제2호, 2023.04, pp.23-29. // (Y.J. Choi et al., Development of adaptive ventilation control algorithm based on indoor carbon dioxide concentration prediction, *Journal of KIEAE*, 23(2), 2023.04, pp.23-29.)
- [16] S.J. Jo et al., A comparative study on the performance of air quality prediction model based on DNN and LSTM, *Proceedings of the Korea Information Processing Society Conference*, 2020.05, pp.577-579.
- [17] M. Du et al., Towards explanation of dnn-based prediction with guided feature inversion, *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018.08, pp.1358-1367.
- [18] R. Dey, F.M. Salem, Gate-variants of Gated Recurrent Unit (GRU) neural networks, 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), 2017.08, pp. 1597-1600.
- [19] K. Irie et al., LSTM, GRU, highway and a bit of attention: An empirical overview for language modeling in speech recognition, *INTERSPEECH 2016*, 2016.09, pp.3519-3523.
- [20] N.J. Gogtay, U.M. Thatte, Principles of correlation analysis, *Journal of the Association of Physicians of India*, 65(3), 2017.03, pp.78-81.
- [21] 최은지 외 3인, 재실자 착의량 산출을 위한 선행 연구 및 기술 분석, 한국생태환경건축학회지, 제19권 제6호, 2019.12, pp.101-106. // (E.J. Choi et al., Analysis of preceding researches and technologies for estimating occupants clothing insulation, *Journal of KIEAE*, 19(6), 2019.12, pp.101-106.)
- [22] E.J. Choi et al., Development of occupant pose classification model using deep neural network for personalized thermal conditioning, *Energies*, 13(1), 2019.12, 45.
- [23] E.J. Choi et al., Comparative analysis of indoor thermal comfort in a residential building by applying dynamic clothing insulation, *Journal of the Architectural Institute of Korea*, 37(4), 2021.04, pp.205-212.
- [24] S.B. Kotsiantis, Supervised machine learning: A review of classification techniques, In: Maglogiannis IG, editor, *Emerging artificial intelligence applications in computer engineering*, Amsterdam, Netherlands: IOS Press, 2007, pp.3-24.
- [25] L. Yang, A. Shami, On hyperparameter optimization of machine learning algorithms: Theory and practice, *Neurocomputing*, 415, 2020.11, pp.295-316.
- [26] R. Andonie, Hyperparameter optimization in learning systems, *Journal of Membrane Computing*, 1(4), 2019.12, pp.279-291.