A Novel Approach to Indoor Environment Assessment: Artificial Intelligence of Things (AIoT) Framework for Improving Occupant Comfort and Health in Educational Facilities

Min Jae Lee

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Building Construction Science and Management

Ruichuan Zhang, Chair

Alireza Shojaei

Nazila Roofigari-Esfahan

May 9th, 2024

Blacksburg, Virginia

Keywords: Deep Learning, Artificial Intelligence of Things (AIoT), Multimodal Data Fusion, Educational Facilities, Occupants Comfort and Health, Transformer, CNN

A Novel Approach to Indoor Environment Assessment: Artificial Intelligence of Things (AIoT) Framework for Improving Occupant Comfort and Health in Educational Facilities

Min Jae Lee

(ABSTRACT)

Maintaining the quality of indoor environments in educational facilities is crucial for student comfort, health, well-being, and learning performance. Amidst the growing recognition of the impact of indoor environmental conditions on occupant comfort, health, and well-being, there has been an increasing focus on the assessment and modeling of Indoor Environmental Quality (IEQ). Despite considerable advancements, current IEQ modeling and assessment methodologies often prioritize and limit to singular comfort metrics, potentially neglecting the comprehensive and holistic factors associated with occupant comfort and health. Furthermore, existing indoor environment maintenance practices and building systems for educational facilities often fail to include feedback from occupants (e.g., students and faculty) and exhibit limited adaptability to their needs. This calls for more inclusive and occupant-centric IEQ assessment models that cover a broader spectrum of environmental parameters and occupant needs. To address the gaps, this thesis proposes a novel Artificial Intelligence of Things (AIoT)-based IEQ assessment framework that bridges gaps by utilizing multimodal data fusion and deep learning-based prediction and classification models. These models are developed to utilize real-time multidimensional IEQ data, non-intrusive occupant feedback (MFCC features from audio recordings, video/thermal features extracted by Vision Transformer (ViT)), and self-reported comfort and health levels, placing a focus on occupant-centric and data-driven decision-making for intelligent educational facilities.

The proposed framework was evaluated and validated at Virginia Tech Blacksburg campus, achieving a 91.9% in R^2 score in predicting future IEQ conditions and 97% and 96% accuracy in comfort and health-based IEQ conditions classifications.

A Novel Approach to Indoor Environment Assessment: Artificial Intelligence of Things (AIoT) Framework for Improving Occupant Comfort and Health in Educational Facilities

Min Jae Lee

(GENERAL AUDIENCE ABSTRACT)

Ensuring optimal Indoor Environmental Quality (IEQ) in educational facilities is crucial for enhancing student comfort, health, and learning performance. Despite the numerous research emphasizing the importance of IEQ conditions and their impact on students' comfort, health, and productivity, existing IEQ modeling and assessment approaches inadequately address the full spectrum of environmental and occupant-specific factors, often disregarding comprehensive and holistic assessments in favor of single isolated metrics. Furthermore, existing indoor environment maintenance practices in educational facilities lack responsiveness to student feedback, necessitating a shift toward more inclusive, occupant-centric IEQ assessment models. To address the gaps, this paper proposes an Artificial Intelligence of Things (AIoT)-based framework. The proposed framework integrates multimodal data fusion with predictive and classificatory deep learning models, utilizing real-time IEQ data and occupant feedback, including non-intrusive occupant feedback and self-reported comfort and health levels. The proposed framework demonstrated high performance in classifying occupant comfort and health levels and predicting future IEQ conditions.

Dedication

This thesis is dedicated to my family and friends, who have been my unwavering support system throughout this 1 year journey. To my parents and grandparents, who instilled in me the values of hard work and perseverance, and to my sister, whose endless encouragement fueled my ambition. To my friends (Peterson, Pourkhodagholi, Zekovic, Agha, Tempone, and Kim), who believed in me even when I doubted myself, and who stood by me, offering their trust and support through every challenge. This achievement is not just a reflection of my efforts, but a testament to the faith you all had in me. Thank you for being my source of motivation and for trusting in my abilities. Your belief in me has made all the difference.

Acknowledgments

First and foremost, I would like to extend my deepest thanks to my advisor, Dr. Ruichuan Zhang, whose guidance, patience, and support made this thesis possible. His expertise and critiques have not only shaped this thesis but also enriched my academic and research abilities. His belief in my potential and his encouragement to pursue new ideas have fostered my growth. I also extend my thanks to Dr. Alireza Shojaei and Dr. Nazila Roofigari-Esfahan for continued guidance on my theis and constructive feedback.

Contents

List of Figures

List of Tables

List of Abbreviations

AIoT Artificial Internet of Things

ANN Artificial Neural Network

CNN Convolutional Neural Network

DNN Deep Neural Network

EEG Electroencephalogram

IAQ Indoor Air Quality

IEQ Indoor Enviromental Quality

IoT Internet of Things

kNN k-nearest neighbors

LSTM Long short-term memory

MFCC Mel-Frequency Cepstral Coefficients

PM Particulate Matter

PSO-SVM Particle Swarm Optimization Support Vector Machine

ResNet Residual Network

SBS Sick Building Syndrome

SVM Support Vector Machine

UI	User Interface
ViT	Vision Transformer
VOC	Volatile Organic Compounds

Chapter 1

Introduction

1.1 Overview

In contemporary society, humans have become predominantly an indoor species, with studies showing that individuals spend as much as 90% of their time indoors (EPA 2023). This shift underscores the importance of Indoor Environmental Quality (IEQ) in our daily lives, especially in educational facilities that play a critical role for over 20 million college students in the United States. IEQ, which encompasses various factors such as indoor air quality (IAQ), lighting levels, thermal conditions, and acoustics, is crucial for the health, comfort, and productivity of occupants (USGBC 2014). Research has consistently shown that well-maintained IEQ conditions in places of learning not only support cognitive function and reduce stress but also improve overall health outcomes and academic performance (Samet and Spengler 2003, Shan et al. 2018, Brink et al. 2021). Conversely, suboptimal IEQ conditions have been linked to health issues ranging from transient discomfort, and sick building syndrome (SBS), to chronic diseases (Cincinelli and Martellini 2017, Jones 1999). Therefore, ensuring that educational facilities provide a comfortable and healthy indoor environment is not just a matter of promoting well-being among students – it is also a strategic approach to enhancing their learning experiences, fostering positive social relationships, encouraging independence, and boosting overall productivity.

Despite the critical role of IEQ conditions in educational facilities, current practices for its assessment and modeling are often inadequate. They frequently fail to include feedback from occupants (e.g., students and faculty), exhibit limited adaptability to their needs and preferences, and lack of exploration of how the integration of IEQ data with occupant feedback might influence the predictions of multidimensional IEQ conditions (Lai and Yik 2007, Frontczak and Wargocki 2011, Gupta et al. 2016). Moreover, current practices are based on code-defined comfort ranges, maximum design occupancy, predefined thresholds, or historical data rather than occupant-centric approach (Virginia Law 2023, Klein et al. 2012). These practices tend to be laborious and reactive, often leading to less-than-ideal indoor environment conditions that do not consider occupant comfort and health levels. Also, existing efforts fall short of integrating IEQ data and occupant feedback through heterogeneous channels. For example, students can show dissatisfaction through their behaviors instinctively or express that they feel uncomfortable with the current IEQ conditions directly. The complex but semantic-rich multimodal environmental and occupant feedback data can provide valuable insights that could optimize IEQ conditions and in turn improve occupants' comfort, health, and well-being (Jazizadeh et al. 2014, Yang et al. 2014, Day et al. 2020). While some studies have begun to explore the integration of various data sources, such as environmental sensors and occupant feedback, there remains an opportunity to leverage advanced data fusion techniques for a more comprehensive and holistic understanding of IEQ conditions (Lahat et al. 2016, Kanezaki et al. 2019). Post-pandemic, these challenges have become even more prominent, as students grapple with the transition back to large-scale college classes following an extended period of remote learning (Zhao and Watterston 2021).

The potential solutions lie in leveraging advancements in deep learning, multimodal data fusion, and the Internet of Things (IoT) to create data-driven, occupant-centric indoor environment modeling and assessment systems. IoT systems have transformed real-time data collection and analysis – it has enabled many concepts such as smart homes/cities and digitally connected healthcare facilities (Gao et al. 2021, Tekler et al. 2022, Wu et al. 2022). Deep learning, capable of discovering and extracting meaningful patterns and representations from large and complex datasets, offers significant predictive and classificatory power. Recent research efforts have shown that deep learning algorithms can effectively fuse heterogeneous data with multimodality and exploit the interrelationships between these modalities (Chen and Jahanshahi 2017, Wang et al. 2019, Zhang and El-Gohary 2023). Artificial Intelligence of Things (AIoT), utilizing both IoT and deep learning algorithms, promises real-time monitoring, enhanced adaptability, and the potential to significantly improve IEQ conditions modeling and assessment by supporting proactively assessing building systems based on the knowledge acquired from both IEQ data and occupant feedback.

In response to the aforementioned challenges, this thesis proposes an occupant-centric AIoT framework for modeling and assessment of indoor environments of educational facilities. The primary goal of the proposed framework is to leverage the capabilities of IoT networks, multimodal fusion approach, and deep learning algorithms to predict and classify multidimensional IEQ conditions at education facilities, incorporating self-reported comfort and health levels and non-intrusive occupant feedback, ensuring a rich combination of subjective and objective feedback. Such framework will support the creation of educational facility assessment systems that proactively respond to the varied needs of their occupants and effectively place occupants at the center of the control progress. This proposed AIoT framework serves as a fundamental component for climate-adaptive, data-driven, and occupant-centric educational facilities. Furthermore, this thesis prioritizes the occupant-centric aspects of building control, emphasizing the health and comfort of individuals. Unlike traditional approaches that predominantly target energy and operational efficiency, this thesis centers on the direct impacts of IEQ on the occupants themselves. Through the development and validation of the proposed AIoT framework, including both the high-resolution IEQ prediction model and the

IEQ classification model, the study aims to enhance the understanding and implementation of building systems that actively respond to and prioritize occupant needs and feedback. This approach represents a significant shift towards creating more adaptive and responsive learning environments that focus on occupant well-being as a primary objective. The development of the proposed framework involves four key steps: (1) multidimensional IEQ data collection through a network of IoT devices, (2) multimodal occupant feedback (non-intrusive occupant feedback and self-reported comfort and health level) collection through sensor networks and web-based user interface (UI), (3) IEQ prediction model development, and (4) IEQ classification model development.

1.2 Literature Review

1.2.1 Influence of IEQ on Comfort and Health

Indoor Environmental Quality (IEQ), as defined by the U.S. Green Building Council (USGBC), encompasses the conditions inside a building – air quality, lighting, thermal conditions, ergonomics – and their effects on occupants or residents (USGBC 2014). Moreover, humans in modern society have become an indoor species, spending 90% of their time indoors: at home, at work, in transportation, and in many other public and private spaces (EPA 2023, Samet and Spengler 2003). The intricacies of IEQ greatly influence everyday experiences and outcomes related to comfort, health, and productivity for individuals spending the majority of their time indoors.

Numerous studies have highlighted the significant impact of various IEQ factors on health, underlying their critical importance. Poorly maintained IEQ conditions have been consistently linked to adverse psychological and physical health effects. Lai and Yik (2007) provided evidence of a direct correlation between IEQ conditions and the health levels of occupants. Samet and Spengler (2003) stated that in environments compromised by indoor air pollutants, surface contaminants, and increased personal contact, there is an elevated risk of various diseases. Shan et al. (2018) identified that inadequately controlled IEQ parameters can significantly impact occupants' cognitive health. Similarly, Mujan et al. (2019) noted the significant effect of indoor environments on the health of individuals in both residential and office settings. Indoor air quality (IAQ), in particular, has been identified as a significant factor affecting occupant health, as shown by Cincinelli and Martellini (2017) and Jones (1999). Beemer et al. (2019) reviewed the impact of the built environment on mental health, focusing on nature connection, personal control, and indoor air quality.

IEQ conditions also influence occupant comfort. Patnaik et al. (2017) have found that IEQ has a direct impact on occupant comfort. Kaushik et al. (2022) investigated how IAQ and lighting directly impact occupant comfort and productivity in office environments. Kamaruzzaman et al. (2017) found that lighting, IAQ, and esthetic perception are crucial for occupant comfort satisfaction. The design and operation of building systems are also crucial to promote well-being and comfort, as noted by Al horr et al. (2016). Campano-Leborda et al. (2020) identified occupants' SBS and their perception of indoor environments as contributing factors to a more comprehensive indoor comfort assessment.

1.2.2 Current Indoor Environment Assessment

A key challenge that research in indoor environment assessment systems continues to face is the issue of limited adaptability. The default operation mode of many existing systems involves the use of predefined thresholds based on historical data to control and manage the indoor environment (Virginia Law 2023). However, these static configurations often disregard the preferences and unique needs of individual occupants, leading to inefficiencies. For example, current building assessment systems frequently prioritize energy efficiency at the cost of occupant comfort and tend to operate according to fixed schedules and maximum design occupancy assumptions, strictly relying on code-defined comfort ranges (Klein et al. 2012). This inflexible approach can result in conditions that do not fully meet the wide-ranging requirements of occupants, a reality that becomes increasingly evident in settings such as educational facilities. In the research conducted by Frontczak et al. (2012) emphasized that the consideration of individual occupant needs is crucial to enhancing occupant satisfaction levels in indoor environments.

Furthermore, limited studies have explored the causal and temporal relationships between multidimensional IEQ data and occupant comfort and health levels. Frontczak and Wargocki (2011) highlighted this issue stating that prior literature reviews examining occupant comfort in indoor environments have mostly focused on individual environmental conditions' effects on humans (e.g., temperature and occupant comfort). The majority of the studies conducted previously focused on the relationship between occupant comfort and thermal/air speed conditions, but not multidimensional IEQ conditions with occupant comfort and health (Cheng et al. 2023; He et al. 2023; Quintana et al. 2023; Zhou et al. 2023). This signals the need for a more holistic and comprehensive approach, one that considers the full range of factors contributing to occupants' comfort and health levels in indoor environments. With such an approach, it has the potential to improve both the efficiency and effectiveness of indoor environment management/assessment systems, ensuring spaces that are not just compliant with codes and regulations, but truly cater to the needs of their occupants.

1.2.3 Occupant-Centric Approaches in IEQ Modeling

Occupant feedback plays a vital role in modeling, calibrating, and weaving occupant needs into building controls, as pointed out by abundant efforts (Lai and Yik 2007, Jazizadeh et al. 2014, Gupta et al. 2016). Findings in neuroscience, psychology, and building science have shown that humans have built-in feedback mechanisms to the indoor environment conditions (Karakas and Yildiz 2020, Sussman and Hollander 2021, Day et al. 2020). Recent developments in occupant-centric approaches for IEQ modeling show a shift toward more interactive and personalized analytics that explicitly leverage data generated through these feedback mechanisms. Song et al. (2013) envisioned an occupant-engaged platform that allows users to write building system control rules dynamically, illustrating the potential for user-driven environmental management. Lam et al. (2014) introduced a model that correlates temperature with occupant comfort, incorporating feedback on a seven-point scale to fine-tune environmental conditions to individual preferences. Building upon the concept of real-time feedback, Gupta et al. (2016) introduced an environmental framework where occupants actively set their preferred temperature ranges through a software application. Similarly, Pritoni et al. (2017) developed a system that captures occupant thermal comfort on a five-level scale, fostering an energy-efficient approach tailored to occupant comfort in university buildings. Jazizadeh et al. (2014) took a decentralized approach, proposing an occupant-centric Heating, Ventilation, and Air Conditioning (HVAC) operation system where users adjust their thermal comfort on a sliding scale, demonstrating a model where occupant comfort directly influences environmental control.

Among all sorts of occupant feedback, self-reported data has become increasingly important in assessing and modeling IEQ conditions, because it encapsulates occupant perceptions that may not be fully captured by objective measures alone. Kallio et al. (2020) leveraged self-reported data to assess perceived IEQ conditions, applying a support vector machine (SVM) to discern patterns and implications on the stress levels of occupants. Ali et al. (2022) utilized machine learning to align hospital patients' satisfaction with IEQ conditions to their self-reported comfort levels. Fritz et al. (2022) utilized a combination of low-cost sensor data and self-reported sleep quality levels to investigate correlations between indoor air quality and sleep quality highlighting the importance of subjective well-being measures. Orman et al. (2023) explored the relationship between IEQ and self-reported productivity, concluding that intelligent building systems contribute to improved occupant productivity. Ortiz and Bluyssen (2022) clustered office workers based on their self-reported preferences for IEQ and psychosocial comfort, particularly in the context of home-based work during the COVID-19 pandemic. This study underscored the diversity of occupant needs in unprecedented circumstances. Fransson et al. (2007) compared objective and subjective indoor comfort indicators, finding that self-reported sensory ratings were more indicative of overall comfort, emphasizing the significance of personal feedback in IEQ assessments.

1.2.4 Internet of Things (IoT) in Indoor Environments

In the field of indoor and outdoor environmental monitoring, low-cost sensor devices are widely being used to collect various IEQ data, including temperature, lighting level, humidity, air pressure, and air pollution such as particulate matter (PM) levels. The rapid progression and development of IoT devices have entered a new era, transforming how we interact with smart objects in built environments (Tekler et al. 2022). Sharma et al. (2018) utilized cost-effective environmental monitoring devices to gather data regarding the concentrations of certain chemicals and PM levels in the air. They further processed and analyzed this data through predictive machine learning models. Huang et al. (2017) presented an affordable and versatile sensor node prototype that can support five sensing functions. Anik et al. (2022) emphasized the growing need for IoT devices stating IoT devices are in constant need of environmental monitoring. The use of IoT devices in such contexts paves the way for more efficient, effective, and responsive indoor environmental assessment/management systems and smart buildings, opening possibilities for enhancing the comfort, health, and well-being of occupants.

1.2.5 Multimodal Data Fusion and Deep Learning in IEQ Modeling

The advancement of multimodal data fusion represents an advancement toward a holistic understanding of complex phenomena. Multimodal data encompasses information collected from a diverse array of sources, including various types of instruments and measurement devices, each contributing unique insights into the system of interest (Lahat et al. 2016).

This rich data, when combined, can provide a more comprehensive understanding than any single modality alone could offer. Deep learning, a subset of machine learning with improved abilities to compute, classify, and interpret large datasets, stands at the forefront of this advancement. With strong abilities to compute and classify data, deep learning is highly expected to improve overall performance, especially with multimodal data fusion algorithms (Meng et al. 2020) – compared with a single deep learning model, fusion models can achieve improved performance and smaller errors (Zhen et al. 2023).

The field of IEQ modeling is evolving through the integration of various data types and the application of machine/deep learning. This advancement is crucial in improving the conditions of indoor spaces for comfort and health. Lee and Zhang (2024) utilized feedback from building occupants alongside sensor data to predict multidimensional IEQ conditions, improving the understanding of the relationship between subjective human experience and objective environmental conditions. Pollard et al. (2021) synthesized passively collected high-resolution IEQ data and occupant location data in an office building. Kallio et al. (2021) bridged occupancy patterns with long-term air quality data to predict CO2 levels, with the potential for rapid development in HVAC system applications. The model fusion approach is further advanced by Wang et al. (2018), predicting occupancy using machine learning and data fusion of environmental sensing and Wi-Fi sensing. Guo et al. (2016) designed an indoor environmental monitoring system based on embedded systems and a multi-sensor data fusion algorithm.

1.2.6 Importance of Educational Facilities

Educational facilities bear significance for college students, and thus their impact on academic performance, comfort and health, and well-being demands attention. Previous research has been undertaken to investigate the role of IEQ conditions in educational facilities or academic buildings. Earthman (2004) highlighted the influence of IEQ conditions such as appropriate temperature, indoor air quality (IAQ), lighting, and acoustics on student achievement. This study emphasized the importance of physical conditions in educational facilities and how these conditions impact the learning process. Uline and Tschannen-Moran (2008) conducted a comprehensive study for the Institute of Education Science, aiming to find the relationship between three key aspects: the quality of educational facilities, the overall atmosphere or 'climate' of the school, and student achievement. Their finding suggested a clear positive relationship among these factors: schools with well-maintained, high-quality facilities were found to create a conducive, pleasant climate, which, in turn, supported enhanced student performance. Wang and Zamri (2013) found that the performance of occupants in an educational or work environment correlates directly and indirectly with all conditions of IEQ. This study indicates that maintaining optimal IEQ conditions plays a pivotal role in promoting academic and professional success. Choi et al. (2014) discovered a direct link between the overall IEQ condition of a classroom and students' perceived impact of IEQ on their learning, which in turn, indirectly influenced students' satisfaction rate with their courses. These studies highlight the relationships between IEQ conditions, learning, and student satisfaction within educational facilities, further emphasizing the need for proactive and occupant-centric approaches to managing these spaces. Furthermore, compared to other building types (e.g., office buildings and residential buildings), occupants are not allowed to make changes in indoor environments.

1.3 State-of-the-Art and Knowledge Gap

Advancement in indoor environment assessment, particularly in educational facilities, is being significantly influenced by the intersection of IoT, deep learning, and multimodal data fusion. The widespread adoption of IoT has prompted a shift towards smarter and more intelligent indoor environments (Gao et al. 2021), leveraging continuous data for real-time condition monitoring, thus enhancing IEQ conditions. Alongside the development of IoT devices, the trend towards a more occupant-centric approach is apparent, underscoring that educational facilities should adapt to the diverse comfort and health needs of occupants (e.g., students and faculty). Consequently, the need for integrating occupant feedback into indoor environment assessment/management is gaining increasing attention (Khan et al. 2020, Duarte Roa et al. 2020). For example, Kim et al. (2019) introduced EEG signals to machine learning for IAQ classification. Marino et al. (2012) developed an indoor comfort classification index. Zhao and Yan (2014) utilized a Particle Swarm Optimization Support Vector Machine (PSO-SVM) focusing on thermal and IAQ parameters. Gan et al. (2021) combined Building Information Modeling (BIM) with machine learning for thermal comfort analysis, highlighting a trend toward integrating digital modeling with algorithmic prediction for IEQ assessment. Predictive modeling has also seen significant advancement, with studies such as Gong et al. (2023) and Feng et al. (2023) using ANN and kNN models to predict thermal comfort. Subsequently, the application of advanced machine/deep learning models, including deep neural networks (DNN) and transfer learning, signifies the evolving landscape of predictive IEQ modeling. Deep learning, with its capability to process raw data, detect patterns, and make more accurate predictions, is becoming increasingly pivotal in managing indoor environments (Bucarelli and El-Gohary 2023, Chen et al. 2023) and is being harnessed to process large datasets from IoT devices, demonstrating the potential to model and assess multidimensional conditions for enhancing occupant experiences in indoor environments. Table 1.1 shows a list of State-of-the-Art Studies in the area of intelligent and adaptive indoor environments.

Current methodologies in indoor environment modeling and assessment, while robust, exhibits gaps in knowledge and application.

1. **Limited Personalization and Occupant Feedback Integration**: A gap in the current indoor environment assessment system is their lack of responsiveness to individuals' needs, as they typically rely on generic thresholds or historical data, which fails to account for the diverse preferences of occupants. Additionally, despite the importance of occupant feedback, many systems show limited incorporation of occupant feedback into IEQ modeling and assessments, leading to a misalignment between system operations and actual user experience.

2. **Insufficiency in Addressing Occupant Comfort and Health**: Another gap in current building assessment systems is their focus on operational efficiency and energy conservation at the expense of the occupants' comfort and health. This indicates the need for a transition to occupant-centric approaches that prioritize the enhancement of the indoor environment, with a particular emphasis on the direct impact on occupants' comfort and health

3. **Unidimensional Focus and Inadequate Holistic Assessment**: Previous IEQ assessment methodologies tend to prioritize thermal comfort metrics or singular IEQ parameters, neglecting other factors that contribute to occupant comfort and health. This oversight underscores the necessity for a more holistic, multimodal, and data-driven approach that assimilates multimodal data streams, including environmental parameters, behavioral indicators, and direct occupant feedback, to provide a fuller understanding of indoor environments.

Table 1.1: State-of-the-Art Studies in the Area of Intelligent and Adaptive Indoor Environments

Papers	Space Type	Occupant Feedback Modality	Prediction
Cheng et al. 2023	Commercial	Gender, Age, Clothing Level, Body Temperature	Thermal Comfort Preference
Cho et el. 2023	Academic	Thermophysiological Status	HVAC Control Parameters
Deng et al. 2022	Academic	Occupancy	HVAC Control Parameters
Dikel et al. 2018	Commercial	Occupancy	Lighting Level
Erikson et al. 2013	Academic	Occupancy	HVAC Control Parameters
Gao and Keshav 2013	Commercial	Clothing Level	Thermal Comfort Preference
Jazizadeh et al. 2013	Commercial	Thermal Preference	HVAC Control Parameters
Jazizadeh et al. 2014	Commercial	Thermal Preference	HVAC Control Parameters
Khan et al. 2023	Residential	Thermal Preference	Energy, Temperature, Humidity
Kim et al. 2019	Academic	Chair Usage	HVAC Control Parameters
Nagy et al. 2016	Academic	Occupancy	Lighting Level
Park et al. 2019	Academic	Thermostat Usage, Occupancy	Lighting Level
Peng et al. 2018	Commercial	Occupant Behavior	HVAC Control Parameters
Pritoni et al. 2017	Academic	Occupant Comfort Vote	HVAC Control Parameters
Rubinstein 2010	Commercial	Occupancy	Lighting Level
Tagliabue et al. 2021	Academic	N/A	Indoor Air Quality
Winkler et al. 2016	Academic	Occupant Comfort Vote	HVAC Control Parameters
Yu et al. 2023	Residential	N/A	Indoor Air Temperature
Zhou et al. 2023	Residential	Body Temperature	Indoor Air Speed
Zou et al. 2018	Commercial	Occupancy	Lighting Level

1.4 Research Objectives

1. Assess the impact of IEQ on occupant comfort and health in educational facilities.

    - Investigate the correlation between IEQ conditions and the perceived comfort and health levels of occupants in educational facilities.

2. Evaluate the effectiveness of multimodal data fusion in IEQ modeling.

    - Examine the effectiveness of integrating multimodal occupant feedback, in improving the accuracy and reliability of IEQ modeling and assessment.

3. Assess the prediction and classification power of AIoT framework .

    - Evaluate the accuracy and reliability of AIoT systems in predicting and classifying to changes in IEQ that affect occupant comfort and health in educational facilities. The focus would be on empirically testing the prediction models within the AIoT framework to determine their effectiveness.

1.5 Research Questions

1. How can multidimensional IEQ conditions, self-reported comfort and health levels, and non-intrusive occupant feedback data be effectively integrated into an AIoT framework to enhance the prediction and classification of IEQ conditions in educational facilities? What hardware and computational approaches are needed for such integration?

2. What is the impact of integrating self-reported comfort and health levels and non-intrusive occupant feedback on the accuracy and reliability of AIoT framework in the prediction and classification of IEQ conditions?

3. Which computational model architecture - among statistical, machine learning, or deep learning algorithms - most effectively accommodates the intricacies of multimodal data for the prediction and classification of Indoor Environmental Quality (IEQ) conditions?

Chapter 2

Proposed Framework

The proposed AIoT framework for indoor environment assessment at educational facilities encapsulates two distinct but complementary models: the high-resolution IEQ prediction model and the IEQ classification model based on occupants' self-reported comfort and health levels. The prediction model utilizes a CNN-based multimodal fusion to analyze IEQ data, self-reported occupant comfort and health levels, MFCC features from audio recordings, and ViT features from video/thermal recordings, delivering high-resolution predictive insights. On the other hand, the classification model utilizes a Transformer-based multimodal fusion approach to process the same data for real-time assessment of current conditions based on self-reported comfort and health levels. The scope of this thesis is focused on the development and validation of IEQ prediction and classification models, and potential future applications and extensions of the framework is later discussed in the *Section 4.3 Extension and Downstream Tasks Enabled by the Proposed AIoT Framework*.

The proposed AIoT framework consists of four main components (Figure 2.1): 1) **Multidimensional IEQ Data Collection**: This component involves the systematic capture of multidimensional IEQ data utilizing a network of IoT devices, 2) **Multimodal Occupant Feedback Collection**: Simultaneously, it collects multimodal occupant feedback (self-reported comfort and health levels and non-intrusive occupant feedback) through various sensors and IoT networks, ensuring a rich combination of subjective and objective occupant feedback, 3) **IEQ Prediction Model Development**: Leveraging the collected data, a deep learning-based predictive model is developed to predict future IEQ conditions, and 4) **IEQ Classification Model Development**: Concurrently, a classification model is formulated to deliver immediate assessments of the current IEQ conditions, informed by self-reported comfort and health levels along with multimodal occupant feedback.

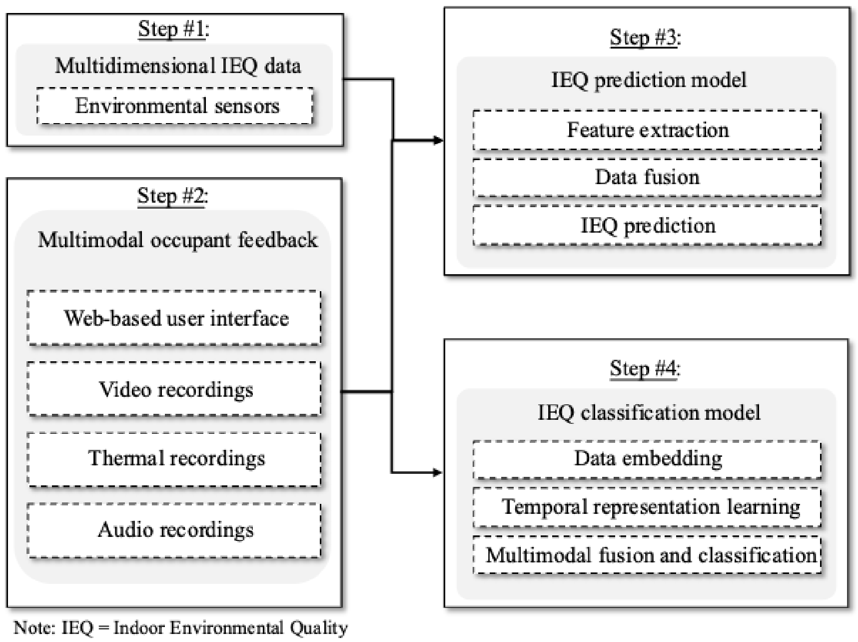


Figure 2.1: Proposed AIoT framework

2.1 Multidimensional IEQ Data Collection

The first component involves deploying a network of cost-efficient and user-friendly IoT devices. The IoT network comprises of low-cost environmental sensors for IEQ data collection, single-board computers, and an user interface (UI) for IEQ condition monitoring and displaying as shown in Figure 2.2.

The cost-effective IEQ sensors and single-board computers collect real-time multidimensional IEQ data. This data collection method facilitates a comprehensive objective understanding

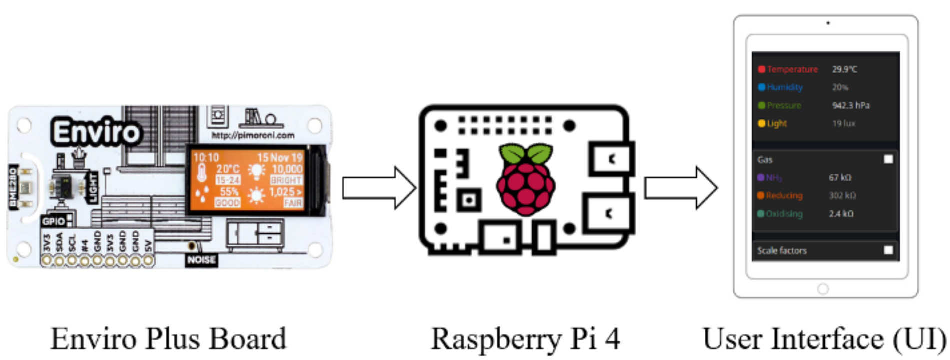


Figure 2.2: IoT device network for multidimensional IEQ data collection

of the indoor environment of education facilities. Collected IEQ data, then, is transferred to the UI through the single-board computer and Wi-Fi. It displays current IEQ conditions, allowing students and other occupants to make informed decisions for indoor environment control if necessary.

To collect IEQ data, a range of sensors and devices were deployed. Specifically, Raspberry Pi 4 with EnviroPlusBoard, equipped with a PM sensor, was utilized. The sensors measured multidimensional IEQ conditions including temperature, pressure, humidity, lighting level, carbon monoxide, nitrogen dioxide, ammonia level, PM1, PM2.5, and PM10.

- Temperature, barometric pressure, and humidity were captured through the BME 280 sensor. BME 280 sensors have a yearly stability rating of ±1.0 hPa (barometric pressure), 0.5% (relative humidity), and ±1.0 1.5 °C (temperature) (Ultra Librarian 2022) and thus suffice the proposed framework.

- To measure light levels, the LTR-599 sensor was used. LTR-599 sensors detect light levels over a wide range from 0.01 lux to 64,000 lux (Pimoroni 2023).

- To gauge the concentration levels of specific volatile organic compounds (VOCs) such as carbon monoxide, nitrogen dioxide, and ammonia, the MICS 6814 sensor, which is an analog gas sensor, was utilized. MICS 6814 sensors were selected because they are equipped with an analog-to-digital converter (ADC), allowing users to work with digitized data and facilitating a more efficient analysis process. PMS 5003 sensor was utilized to collect PM levels from the air.

For the proposed AIoT framework, multidimensional IEQ data was collected every 1 second. This approach was adopted for two reasons: 1) Indoor Air Quailty (IAQ) level, especially the VOC levels, fluctuate rapidly, and capturing these subtle changes is essential to model and assess IEQ conditions accurately and holistically and 2) the high-resolution IEQ data collection supports the data-driven approach, allowing the detailed and nuanced understanding that the proposed AIoT framework require to achieve the best performance.

A noted concern with this configuration of devices involves the heat emitted by the single-board computer's CPU, which could potentially interfere with the accuracy of data collection, particularly temperature and humidity readings. To counteract this problem, the authors have employed a mitigation strategy by utilizing a 40-pin general purpose input-output (GPIO) cable extension, approximately one foot in length, to physically distance the sensors from the single-board computer, thus minimizing the effect of heat dissipation on the accuracy of the sensors and the quality of the data collected.

2.2 Multimodal Occupant Feedback Collection

2.2.1 Self-reported Comfort and Health Level Collection

Central to our framework is the component that captures self-reported comfort and health levels from the building occupants. This component draws from the concept of self-reported health (SRH) levels, predominantly utilized in the medical field and supported by the National Institute of Health (NIH). It takes a holistic approach, considering the multifaceted nature of health and comfort by correlating with multidimensional IEQ conditions and other occupant feedback. Moreover, this self-assessment tool is designed with the foresight that self-reported comfort and health levels can serve as a predictive marker for future health/-comfort needs. By offering a means to collect self-reported comfort and health levels directly from occupants, this component allows the framework to dynamically adapt and accurately model and assess IEQ conditions, placing a strong emphasis on the experiences and needs of the occupants.

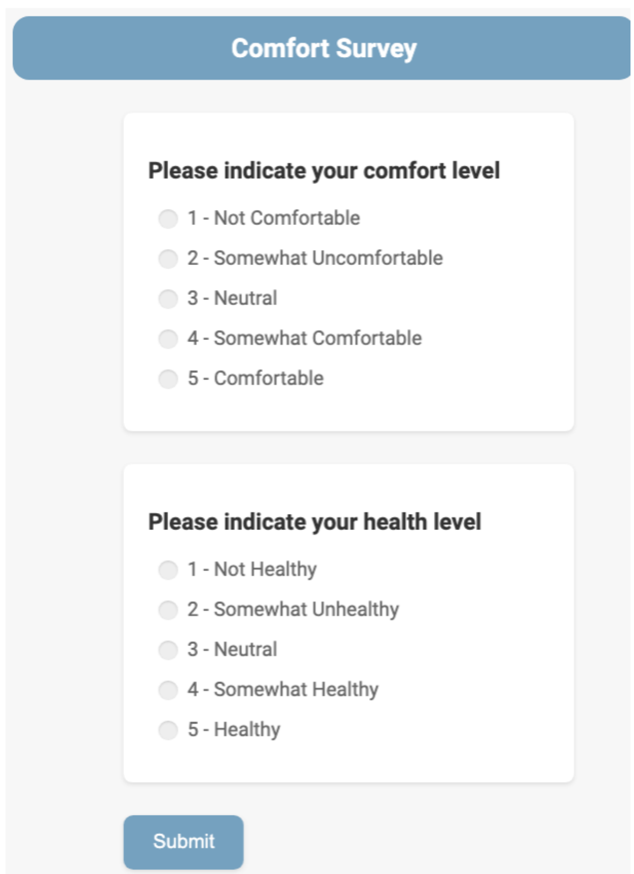


Figure 2.3: Web-based UI for self-reported comfort and health levels collection

The collection process is facilitated through a web-based UI (Figure 2.3) designed by authors based on two criteria: (1) the interface is accessible from anywhere with an internet connection, ensuring convenience for occupants, and (2) the design is simple to encourage user engagement, allowing occupants to promptly express their perceptions of comfort and health within the indoor environment.

Within this component, occupants can explicitly indicate their perceived comfort and health levels, utilizing the developed web-based UI to articulate how comfortable and healthy they feel in the current setting in Likert scale, offering a direct channel for expressing personal experiences of the indoor environment's impact.

The self-reported comfort and health levels collected through the platform are measured on a Likert scale, which is then integrated with the IEQ data and non-intrusive occupant feedback (i.e., MFCC features and video/thermal features), creating a unified dataset that fuses the measurable aspects of the environment with the personal experiences of the occupants. In a case where different comfort/health levels were reported at the same time, an average was taken to represent the collective representation of the occupant feedback.

2.2.2 Non-Intrusive Occupant Feedback Collection

This component involves collecting video, thermal video, and audio data while minimizing disruptions to occupants in the indoor environment. (1) *Video and thermal video data.* Through the strategic placement of cameras, video/thermal recordings were captured from different angles, offering diverse perspectives that reflect varied experiences. Further enhancing our framework's sensitivity to the subtleties of occupant behaviors, Vision Transformer (ViT) was applied to analyze video and thermal recordings to understand underlying patterns that contribute to the assessment of occupant comfort and health, extending beyond the direct interpretation or labeling of videos. Pre-trained ViT models were utilized to extract a rich set of features from the recordings instead of specific labeling of each frame from the video or thermal recordings. (2) *Audio data.* In parallel, audio recordings are processed to extract Mel-frequency cepstral coefficients (MFCCs), transforming complex sound landscapes into structured features. These features present meaningful patterns from background noise and occupant interactions within the space. MFCCs provide an understanding of the auditory environment, which, when combined with other visual features, enhances the framework's ability.

This proposed method of collecting video/thermal recordings and audio recordings showcases the non-intrusive nature of the proposed occupant feedback mechanism, setting it apart from more intrusive assessment methods that require occupants to do certain tasks (e.g., wearing sensors, interacting with robots, using headsets, etc.). This method allows occupants to contribute to indoor environment assessment effortlessly, without the need for active participation or alteration from their normal activities.

2.3 IEQ Prediction Model Development

The third component of the proposed framework includes the development of a deep learning model architecture for IEQ and occupant data analytics. The proposed model consumes IEQ data with self-reported comfort and health levels, features extracted from video and thermal recordings using pretrained Vision Transformer (ViT) or ResNet50, and MFCCs from audio recordings and predicts multidimensional IEQ conditions that satisfy the students' comfort and health needs. The model architecture (Figure 2.4) consists of three main parts: feature extraction, data fusion, and IEQ condition prediction.

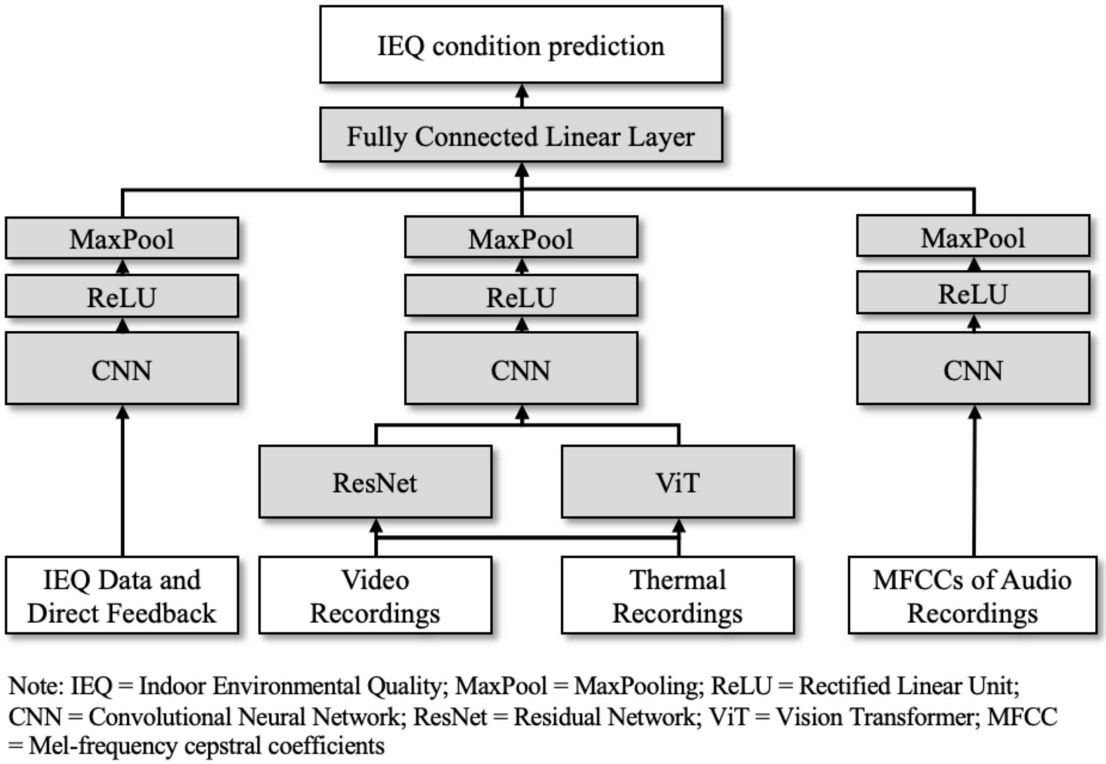


Figure 2.4: Proposed IEQ prediction model architecture

2.3.1 Feature Extraction Model

The authors employed and adapted both residual network (ResNet) and Vision Transformer (ViT) to build the feature extractor to evaluate and compare their respective performances within the proposed framework. The objective was to evaluate which model architecture could provide a more useful and discriminative feature representation. ResNet was chosen due to its strong performance in extracting features robust and discriminative for downstream tasks from images (Allen-Zhu and Li 2020). ResNet incorporates the concept of skip connections and shortcut paths, which prevent the problem of vanishing gradients during training. The model used in this framework is pretrained ResNet50 (ResNet composed of 50 layers) and its model architecture is modified by removing the original final fully connected layer. The layers prior to the final fully connected layer capture high-level features and representations of the input image. Removing the last layer allows us to access and use these features. The pretrained ViT was also utilized as a feature extractor due to its outstanding ability in analyzing image data through the application of transformer architecture (Dosovitskiy et al. 2020). ViT, similar to ResNet, has demonstrated high performance in extracting salient features from images which are valuable for various downstream tasks. Unlike traditional convolutional neural networks (CNNs), ViT applies Transformers, initially designed for natural language processing (NLP) tasks, to understand global, long-range dependencies within an image. The model deployed in this framework is a pretrained version of ViT, which means it has already learned rich, discriminative feature representation from extensive image data sets. The feature extractor takes in individual frames from each video, each of which has been reshaped and normalized to fit the model's input requirements. ResNet50 then outputs a feature vector with 2,048 dimensions and ViT outputs a feature vector with 768 dimensions for each extracted frame. These high-dimensional feature vectors serve as a compact but rich representation of the original visual content of the frame. Each dimension in this space represents an abstracted feature extracted by the model during its pretraining on a vast amount of labeled image data. Therefore, despite the abstraction, each dimension can capture a unique characteristic of each frame, allowing a holistic understanding of the visual and thermal conditions. While thermal comfort is subjective and may vary between individuals, the feature extraction capabilities of ViT or ResNet50 can help identify subtle visual cues that correlates with general comfort level. It also reduces computational requirements by focusing subsequent analysis on the most important patterns learned by models.

2.3.2 Data Fusion Model

In the proposed model architecture, a multi-layer convolutional neural network (CNN) is then used for fusing multimodal data. This fusion model is designed to learn from various features including data, self-reported comfort and health levels, MFCC features from audio recordings, as well as video and thermal features generated from ViT feature extractors. This architecture is based on the idea of model fusion, which combines three distinct processing units, each dedicated to a specific type of input data. The model architecture consists of three CNNs that process different types of input data. Each CNN processes a unique type of data: IEQ data collected from IoT devices with direct feedback, MFCC data derived from audio recordings, and concatenated feature vectors from video and thermal recordings. This design allows each data stream to be processed independently within its respective CNN. Each CNN is composed of a convolutional layer, a ReLU activation function, and a max-pooling layer. The convolutional layer is used for the extraction of significant features from the input data. Following the convolutional layer, a ReLU activation introduces non-linearity into the model, setting all negative values to zero and leaving positive values unchanged, allowing the model to learn and capture complex patterns in the data. Subsequently, a max-pooling layer reduces the spatial dimensions of the output from the preceding layer, decreasing the computational complexity of the model. The outcomes of all three CNNs — each representing the processed IEQ data, MFCC data, and video and thermal feature vectors—are concatenated. This concatenation process is pivotal as it creates a combined representation that encapsulates the multidimensional spectrum of prior IEQ conditions and corresponding occupant behaviors. The fused representation is then fed into the IEQ prediction model.

2.3.3 IEQ Prediction Model

The proposed model architecture uses a fully connected layer for prediction. This layer leverages the fused data to predict the final multidimensional IEQ conditions. This prediction reflects the model's interpretation of the multidimensional IEQ conditions, taking occupant health and comfort into consideration. These factors are considered alongside a comprehensive analysis of both previous IEQ conditions and direct and indirect feedback from occupants. For example, at timestamp i, the model takes in the fused representations learned from the data of previous l timestamps (denoted by D_{i-l}, D_{i-l+1},..., D_{i-2}, D_{i-1}) and predict the IEQ conditions of the current timestamp T_i.

2.4 IEQ Classification Model Development

The fourth component of the proposed framework involves developing a multimodal data fusion-based Transformer model for classifying IEQ in terms of occupant self-reported comfort and health levels based on collected multimodal data (i.e., IEQ conditions, video and thermal features, MFCCs). Similar to the IEQ prediction model, the importance of the proposed model architecture lies in its model fusion approach, where different data types are individually processed and later integrated to create a comprehensive understanding of indoor environments and occupants. The Transformer model consumes IEQ data, MFCC features generated from audio recordings, and concatenated features from video and thermal recordings and classifies IEQ conditions. The model consists of three main components, as shown in Figure 2.5: (1) data embedding, (2) temporal representation learning, and (3) multimodal data fusions and classifications.

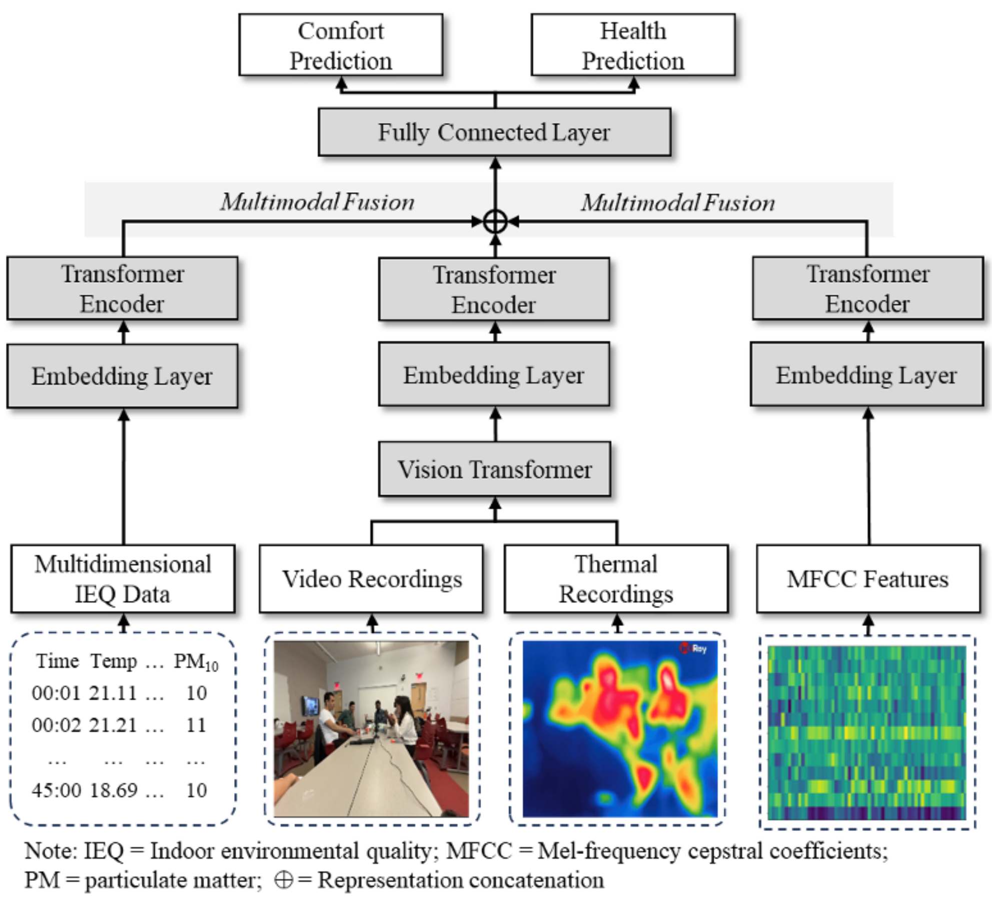


Figure 2.5: Proposed IEQ classification model architecture

2.4.1 Data Embedding

This segment of the model architecture is dedicated to the initial processing phase of each data type: IEQ data, MFCC data, and video/thermal features. During preprocessing, each data point, comprising IEQ, MFCCs, and video/thermal features for each time step, is labeled based on self-reported comfort and health levels. This labeling is pivotal as it directly ties the raw data to occupant-centric subjective parameters. The model initiates the processing of IEQ data, audio, video, and thermal features by passing them through an embedding layer, which maps the original input size to a new size for subsequent Transformer encoder.

2.4.2 Temporal Representation Learning

Each of these embedded inputs is then processed through their respective Transformer encoders. These encoders consist of multiple layers of Transformer encoder layers, which are adept at handling sequential data and capturing intricate dependencies within the input sequences through self-attention mechanisms. Self-attention mechanisms, a key component of Transformer, allow each element in the input sequence to interact with and weigh the importance of all the other elements, effectively enabling the model to consider the entire sequence. Each timestamp in the input sequence attends to all other timestamps, producing a set of weights. These weights are subsequently aggregated to generate a contextual temporal representation for each timestamp. The use of distinct encoders for each data modality allows for specialized handling of different data modalities, ensuring that the unique characteristics of each are effectively captured.

2.4.3 Multimodal Data Fusion and Classification

Post-encoding, the model performs a crucial step of combining the encoded features from all three data types. This is achieved by concatenating the output of the IEQ, MFCC, and video/thermal encoders. This concatenated feature set represents a comprehensive understanding of the indoor environment, integrating various aspects such as physical parameters, audio cues, visual/thermal dynamics, and corresponding occupant comfort and health levels. The combined and labeled features are then passed through a final fully connected layer, which acts as the central classification unit of the model. This layer translates the rich, fused features into final predictions, classifying IEQ conditions based on self-reported comfort and health levels. The average of cross-entropy loss of comfort and health levels was used as the objective function for training the model. Cross-entropy loss (Eq. (2.1)) measures the performance of a classification model whose output is a probability value between 0 and 1, where N denotes the number of observed IEQ conditions, y_{ik} denotes the binary indicator of whether class k is the correct classification for observation i, and $\hat{y}_{ik}$ is the predicted probability that observation i belongs to class k. Cross-entropy loss increases as the predicted probability diverges from the actual label, indicating worse performance, and thus is minimized during the training process.

$$\text{Cross Entropy Loss} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{ik}\log(\hat{y}_{ik}) \tag{2.1}$$

Chapter 3

Experiments and Results

3.1 Data Collection

To collect data and test the proposed AIoT framework, a set of experiments was conducted across three uniquely characterized educational facilities on the Virginia Tech Blacksburg campus: Bishop-Favrao Hall, Pamplin Hall, and New Classroom Building (Figure 3.1). Each of these buildings represents a distinct and unique indoor environment, allowing for a comprehensive assessment of the proposed AIoT framework's capabilities and adaptability.

Figure 3.1: Experiment educational facilities (Top: Bishop-Favrao Hall, New Classroom Building, Pamplin Hall; Bottom: Example Spaces)

The experimental data collection process was conducted over six sessions, distributed across three different educational facilities to assess how occupants respond to various types and styles of buildings. In each building, the sessions were held at two different times: one in the early morning and another in the afternoon, to ensure a comprehensive analysis of occupant reactions within varied temporal contexts and potential fluctuations in student engagement levels due to the time of the day.

- Bishop-Favrao Hall was selected for the first two sessions, representing contemporary and the most widely used educational settings. These sessions aimed to assess the effectiveness of the AIoT framework within an environment that closely mirrors the majority of the current educational facilities.

- New Classroom Building was selected for the third and fourth sessions. It was built in the late 2010s, representing a technologically advanced educational setting. The goal of these sessions was to examine how the AIoT network interacts and adapts to a space equipped with more recent technologies.

- The fifth and sixth sessions were conducted in Pamplin Hall. As one of the oldest buildings on the Virginia Tech Blacksburg campus, Pamplin Hall represents a contrasting environment to assess the adaptability of the AIoT network. These sessions aimed to evaluate how effectively the proposed network could be integrated into more outdated indoor environments, which represent unique challenges compared to contemporary or technologically advanced buildings.

The experimental design also considered the different spatial configurations in educational facilities. Each session was conducted in various locations within the selected buildings, ranging from small conference rooms to larger lecture halls, with each classroom size ranging from 800 square feet to 1,200 square feet (74 m^2 to 112 m^2). This approach provided a thorough assessment of how the AIoT framework adapts to different space sizes and layouts. One EnviroPlusBoard with Raspberry Pi 4, located near the occupants, was deployed per session to collect IEQ data.

The authors also manipulated key environmental conditions within the selected room for each session to simulate different and changing indoor environments. In each session, the lighting levels were adjusted to mimic both optimal and sub-optimal light and visual conditions, ranging from bright illumination to softer and dimmer settings. The airspeed from the fans was also altered, creating different levels of ventilation and air movement across the classrooms, which also impacted the temperature. Additionally, the participating students' seating arrangements were altered to further diversify the spatial configurations and the spatial distribution of the occupants. As a result of IEQ conditions manipulation, some participants showed signs of discomfort by changing their body position and changing lighting levels in the classroom.

In order to simulate a typical class time period, each experimental session lasted for 45 minutes. Within these sessions, a variety of activities were conducted, including self-studying, group discussions, lectures, and other student-led tasks, to replicate realistic classroom scenarios. Participants were prompted to behave as naturally as they would in their regular classrooms, including performing routine actions such as talking, listening, and walking.

3.2 Data Preprocessing

To preserve the distinct characteristics in each individual session, the processing of each data modality and each session's data was executed separately in this step. The separate, parallel preprocessing approach was adopted because each session and data modality represent a unique set of indoor environments, with its own specific environmental variables, spatial and temporal distributions, student behaviors, and student-building interactions. Furthermore, a slightly different approach was taken for each model (e.g., prediction model and classification model), tailoring it to better fit the unique requirements and objectives inherent to their respective analytical functions.

3.2.1 Prediction Model Data Preprocessing

Three key steps were adopted for data preprocessing for the IEQ prediction model: (1) data sampling and synchronization, (2) data normalization, and (3) data sequence generation and splitting (Figure 3.2).

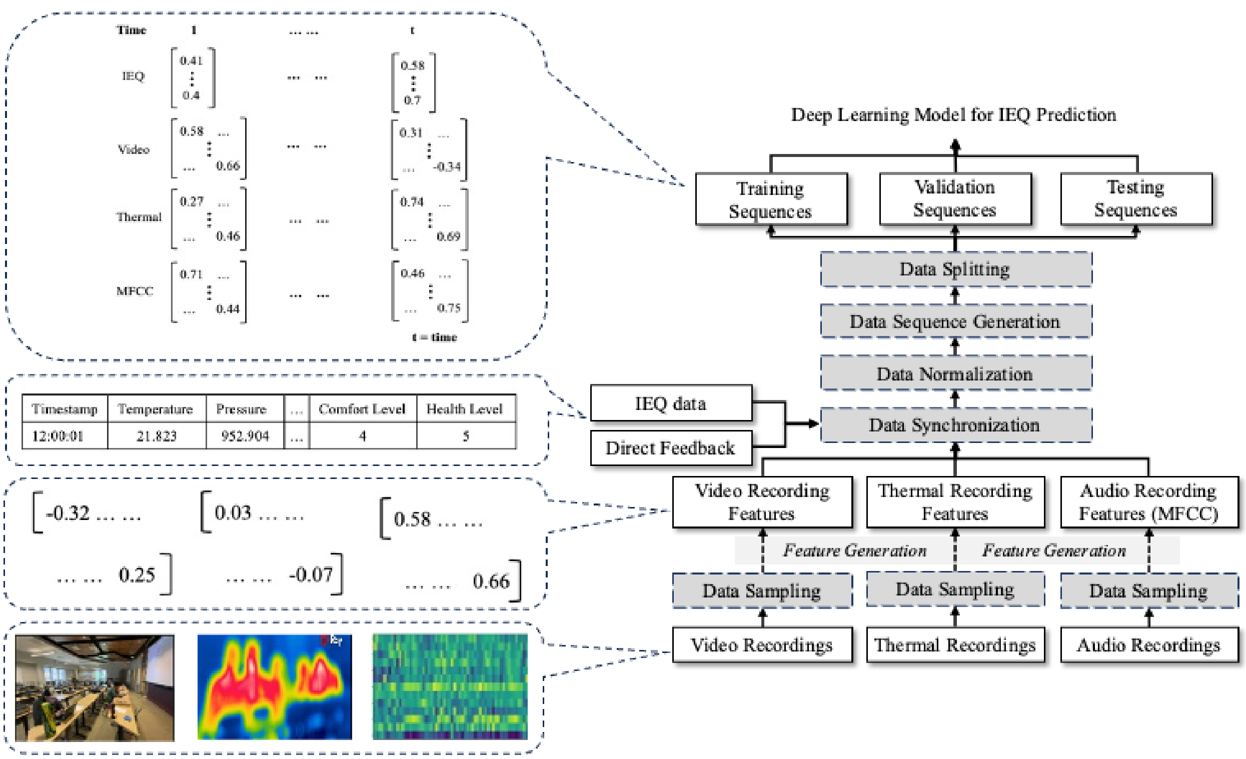


Figure 3.2: IEQ prediction model data preprocessing step for each experimental session

Data Sampling and Synchronization

Each data modality and session were processed individually due to their distinct nature – each modality and session encapsulates unique indoor environments, characterized by specific environmental variables, spatial-temporal distributions, and student behaviors.

- IEQ data: the IEQ data were collected every 1 second. They were represented in single-precision float numbers and stored in CSV files along with the timestamp.

- Self-reported comfort and health levels: the students' self-reported comfort and health levels were collected through the web-based UI with the response time. Occupants were instructed to use the web-based UI every 5 minutes or whenever they felt significant changes in their comfort and health levels. They were represented in five Likert scales and stored in CSV files along with the timestamp.

- Video and thermal recording: the video and thermal recordings were initially recorded in MOV format. Frames were then extracted at a rate of 1 frame per second to synchronize with the IEQ data, from each video and thermal file using the Python library OpenCV. The frames later served as input data for the feature extraction models, ResNet50 and ViT. Feature extractions for each type of model (i.e., ResNet and ViT) were done separately. After all, the extracted features were single-precision float numbers and were temporarily saved in NPY format.

- Audio data: the audio data were initially recorded in WAV files. They were trimmed and resampled for synchronizing with the IEQ data. The audios were then used for MFCC extraction with the Python library Librosa. The extracted MFCC features were single-precision float numbers stored in NPY format.

Data Normalization

All modalities of representations (IEQ data, self-reported comfort and health levels, ViT features extracted from video and thermal video representations, and MFCC features extracted from audio recordings) were then normalized. Min-max scaling was adopted to transform all representations to a range between 0 and 1. The normalization step (1) ensured a standard scale across all features and reduced variance; (2) mitigated the risk of any features disproportionately affecting the model and thus making the IEQ condition prediction model less biased; and (3) stabilized the gradient descent during the training of the model and thus allowed a faster convergence for real-time IEQ condition prediction and indoor environment assessment.

Data Sequence Generation and Splitting

Sequences were created from the normalized data following a sliding window method. This method organized the data into sequences appropriate for prediction based on predefined parameters of sequence length and forecast horizon. These parameters were chosen to best reflect the temporal structure of the data. For this experiment, sequence length was set to 60, and forecast horizon was set to 1, meaning that the proposed model learns from the past 60 data points and predicts the next 1 data point. A total of 9,126 sequences were generated. The decision to use past 60 seconds data to predict next 1 second was adopted from results found from previous studies: using high-resolution data for short-term prediction results in improved model performance. Recent studies have compared different prediction intervals and found that high-resolution data with shorter prediction intervals result in increased prediction performance (Afroz et al. 2018, Park and Kim 2023, Xing et al. 2023, Jiang et al. 2022).

The generated sequences were then split into three subsets: training, validation, and testing sets, following the 80-10-10 split for each session's collected data. The training set was used to train the deep learning IEQ condition prediction model, the validation set was employed during the model tuning process to monitor model performance and prevent any overfitting or underfitting in a cross-validation manner, and the testing set was used for the final evaluation of the model's performance.

3.2.2 Classification Model Data Preprocessing

Four key steps were adopted for data preprocessing for the IEQ classification model: (1) data normalization, (2) discretization and labeling of self-reported comfort and health levels, (3) data sequence generation, and (4) data splitting for model training, validation, and testing.

Data Normalization

Initially, raw IEQ data along with MFCC features extracted from audio recordings, and ViT features extracted from video/thermal recordings were normalized to ensure uniformity in scale across all input types. MinMax Scaler was used, which scales the data to a default range of 0 to 1. This scaling is crucial for managing the disparity in magnitude among different features, making them comparable and suitable for analysis by deep learning models.

Discretization and labeling of self-reported comfort and health levels

Following data normalization, the self-reported comfort and health levels collected from occupants are discretized into categorial labels. This step is vital for converting subjective occupant feedback into quantifiable targets for the proposed model. By assigning numerical labels to different levels of comfort and health, the data becomes actionable for the training classification model. Each IEQ data point, along with corresponding MFCC features, and video/thermal features, is then associated with these labeled comfort and health levels, effectively labeling IEQ conditions with the subjective assessments of occupants for every timestamp.

Data Sequence Generation

Similar to the prediction model data preprocessing step, the sliding window method was applied to generate data sequences from the normalized and labeled data for subsequence processing and analysis. The sliding window method involves moving a fixed-size window across the time series data, one-time step at a time, to capture and analyze subsets or segments of the time series comprising IEQ conditions, their corresponding comfort and health levels, and non-intrusive occupant feedback features. Creating sequences post-labeling ensures that each segment used for model training encapsulates both environmental conditions and the associated occupant feedback, crucial for learning the relationship between IEQ and occupant perceptions over time. Each type of data goes through separate preprocessing to maintain its unique characteristics before being integrated. This approach enables a comprehensive analysis, leveraging the strengths of each data type to improve the model's understanding of indoor environments.

Data Splitting

The labeled sequences are then divided into distinct sets for training, validation, and testing, with a specified ratio of 8:1:1. Ensuring a comprehensive evaluation process: the training set is used to fit the model, the validation set allows for an iterative process of model tuning and refinement, and the testing set evaluates the model's ability to generalize to new, unseen data. Finally, the structured and labeled sequences are loaded into iterators, which facilitate efficient batching and shuffling of data for model training. These iterators are instrumental in optimizing the training process, providing a streamlined mechanism for feeding data into the deep learning model in manageable batches and enhancing computational efficiency and model performance. After splitting is completed, all data are loaded into the Data Loader.

3.3 Performance of Prediction and Classification Models

3.3.1 Evaluation Metrics for Prediction Model

The performance of the prediction model was assessed using four metrics: mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination, denoted as R^2. MSE, represented by Eq (3.1), quantifies the average of the squared differences between the predicted and actual values of IEQ at time stamp i ($\hat{T}_i$ and T_i). A higher MSE indicates a larger discrepancy between the predicted and actual values and thus lower model performance (similarly for MAE and RMSE).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(T_i - \hat{T}_i)^2 \tag{3.1}$$

MAE, defined in Eq (3.2), is another metric used to measure the average magnitude of errors in predictions, calculated as the average absolute differences between the predicted and actual values of IEQ at time stamp i. Compared to MSE, MAE is less sensitive to outliers.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|T_i - \hat{T}_i\right| \tag{3.2}$$

RMSE, as shown in Eq (3.3), is the square root of the MSE. It offers a more interpretable evaluation metric by representing the average magnitude of the prediction errors in the same units as the target IEQ variables.

$$RMSE = \sqrt{MSE} \tag{3.3}$$

R^2, defined in Eq (3.4), is a statistical measure that represents the proportion of the variance for a dependent variable that is explained by an independent variable in a regression model, where R^2 of 1 indicates the regression prediction is perfect. SS_{res} denotes the sum of squares of residuals and SS_{tot} denotes the total sum of squares.

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \tag{3.4}$$

3.3.2 Prediction Model Comparative and Ablation Analysis

The proposed prediction model architecture was tested through an extensive ablation and comparative analysis involving different model variants. The primary objective was to evaluate each model's performance and understand (1) the impact of self-reported comfort and health levels and non-intrusive occupant feedback; (2) the impact of various modalities of input data; and (3) the impact of the model architecture for feature extraction. Additionally, through ablation analysis, the significance of specific feedback types related to occupants' perceived comfort and health levels was assessed. For *each* of the three analyses, a comparison between the proposed CNN model and an LSTM model for multimodal data fusion was also conducted. Given that LSTM models are recognized for their proficiency in sequential data prediction, it is worth comparing their performance with the proposed CNN architecture.

All models shared the same hyperparameters, with variations only in the input dimensions reflecting the different types of data they consume and the number of epochs to best optimize the training of each model variation. All models were trained on Python 3 using PyTorch and ran on NVIDIA A100.

Tables 3.1 and 3.2 present an analysis of the performance of the proposed CNN model and the LSTM model across several variants. In Table 3.1, the full CNN model demonstrates the best performance with an MSE of 0.0162, RMSE of 0.0995, MAE of 0.0782, and R^2 of 0.919. When self-reported comfort and health levels were removed, there was a decrease in the model performance as MSE, RMSE, and MAE increased by 17.3%, 17.1%, and 18%, and R^2 decreased by 3.3%. When the non-intrusive occupant feedback (MFCC and video/thermal features) were removed, the decrease in the model's performance was more significant. Compared to the full model, MSE increased by 62.3%, RMSE by 47.4%, MAE by 47.8%, and R^2 score decreased by 10.2%. The results suggest that non-intrusive occupant feedback has a substantial contribution to the model's performance. When only IEQ data were used, which implies the removal of both self-reported comfort and health levels and non-intrusive occupant feedback, showed an increase in MSE, RMSE, and MAE by 22.8%, 21.5%, and 22.4% and R^2 decreased by 4.2% The results highlight the combined impact of self-reported comfort and health levels and non-intrusive occupant feedback.

Table 3.1: Impact of self-reported comfort and health levels and non-intrusive occupant feedback – Proposed CNN Model

Model Variant	MSE	RMSE	MAE	R^2
Full Model	**0.0162**	**0.0995**	**0.0782**	**0.919**
Without Self-reported Comfort and Health Levels	0.0190	0.1165	0.0923	0.889
Without Non-intrusive Occupant Feedback	0.0263	0.1467	0.1156	0.825
Only IEQ	0.0199	0.1209	0.0957	0.880

Note: **Bold** = best performance

In Table 3.2, the full LSTM model achieved an MSE of 0.0512, RMSE of 0.2186, MAE of 0.1759, and R^2 of 0.611. Similar to the proposed CNN model, the LSTM model benefits from both self-reported comfort and health levels and non-intrusive occupant feedback. However, despite the LSTM model's known ability for handling sequential data, the proposed model displayed better performance across all metrics and all model variants. For example, for full models, the proposed CNN model outperformed the LSTM model with an increase of 215.4% in MSE, 119.6% in RMSE, and 124.9% in MAE, and a decrease of 33.5% in R^2.

Table 3.2: Impact of self-reported comfort and health levels and non-intrusive occupant feedback – LSTM Model

Model Variant	MSE	RMSE	MAE	R^2
LSTM Model	**0.0512**	**0.2186**	**0.1759**	**0.611**
Without Self-reported Comfort and Health Levels	0.0542	0.2265	0.1861	0.583
Without Non-intrusive Occupant Feedback	0.0967	0.3101	0.2470	0.218
Only IEQ	0.0598	0.2387	0.1924	0.536

Note: **Bold** = best performance

In Table 3.3, the full CNN model performed the best with an MSE of 0.0162, RMSE of 0.0995, MAE of 0.0782, and an R^2 score of 0.919. When the video and thermal layer were removed there was a notable decrease in performance, with an increase of 49.4% in MSE, 36% in RMSE, and 37.9% in MAE, and a decrease of 7.4% in R^2. The model variant without self-reported comfort and health levels also showed a decrease in performance as MSE, RMSE, and MAE increased by 17.3%, 17.1%, and 18 and R^2 decreased by 3.3%. When audio feedback was removed, the model also showed a decrease in performance as MSE, RMSE, and MAE increased by 38.9%, 28.8%, and 32.4%.

Table 3.3: Impact of Feedback Data Modality – Proposed CNN Model

Model Variant	MSE	RMSE	MAE	R^2
Full Model	**0.0162**	**0.0995**	**0.0782**	**0.919**
Without Video and Thermal	0.0242	0.1353	0.1078	0.851
Without Audio	0.0225	0.1281	0.1035	0.866
Without Self-reported Comfort and Health Levels	0.0190	0.1165	0.0923	0.889

Note: **Bold** = best performance

Table 3.4 displays the LSTM model performance. For the full LSTM model, the MSE was 0.0512, RMSE was 0.2186, MAE was 0.1759, and the R^2 score was 0.611. Similar to the proposed CNN model, the LSTM model benefits from additional modalities of feedback data, including video, thermal, audio, and self-reported comfort and health levels. However, overall, the proposed model consistently demonstrated better performance than the LSTM model across all variants, underscoring the effectiveness of the proposed CNN model despite the removal of certain input data and their corresponding layers.

Table 3.4: Impact of Feedback Data Modality – LSTM Model

Model Variant	MSE	RMSE	MAE	R^2
Full Model	**0.0512**	**0.2186**	**0.1759**	**0.611**
Without Video and Thermal	0.0930	0.3046	0.2623	0.245
Without Audio	0.0915	0.3016	0.2533	0.260
Without Self-reported Comfort and Health Levels	0.0542	0.2265	0.1861	0.583

Note: **Bold** = best performance

The impact of the feature extraction model architecture on the performance of both the

CNN and LSTM models was also tested, as shown in Tables 3.5 and 3.6. Two pre-trained models were tested, ViT and ResNet. The CNN model, when paired with ViT as the feature extractor, achieved the best performance – it yielded an MSE of 0.0162, MAE of 0.0782, RMSE of 0.0995, and a R^2 score of 0.919. However, when the feature extractor was switched to ResNet, there was a decrease in model performance as MSE, RMSE, and MAE increased by 1,130%, 10.9%, and 15.6%, and R^2 decreased by 2%.

For the LSTM model, when ViT was used, the MSE was 0.0512, MAE was 0.1759, RMSE was 0.2186, and the R^2 was 0.611. When ResNet was used as the feature extractor, the performance further decreased, with the R^2 score declining by 11.3%. The choice of feature extractor has a notable impact, with the ViT leading to better performance compared to ResNet for both CNN and LSTM models.

Table 3.5: Impact of Feature Extraction Model – Proposed CNN Model

Feature Extractor	MSE	RMSE	MAE	R^2
ViT	**0.0162**	**0.0995**	**0.0782**	**0.919**
ResNet	0.1993	0.1103	0.0904	0.901

Note: **Bold** = best performance

Table 3.6: Impact of Feature Extraction Model – LSTM Model

Feature Extractor	MSE	RMSE	MAE	R^2
ViT	**0.0512**	**0.2186**	**0.1759**	**0.611**
ResNet	0.0598	0.2371	0.1902	0.542

Note: **Bold** = best performance

3.3.3 Predicted IEQ Conditions

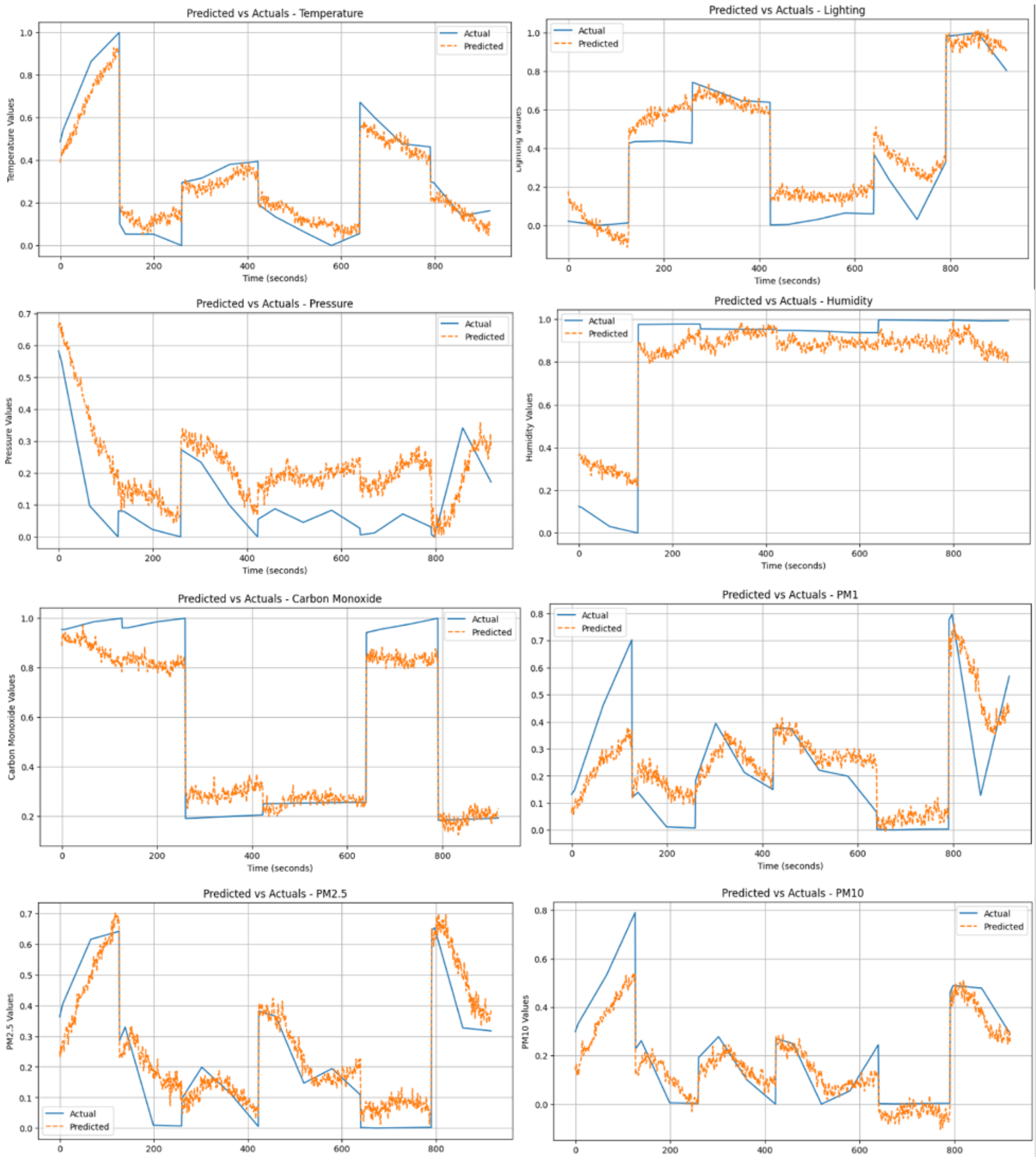


Figure 3.3: Predicted and Actual IEQ Values (from Left to Right, Top to Bottom: Temperature, Light, Air Pressure, Humidity, Carbon Monoxide, PM_1, $PM_{2.5}$, and PM_{10})

Figure 3.3 shows our proposed prediction model's performance on the testing dataset, which is a combination of every 10% of the data collected from each session, following our 80-10-10 data split. The line graph for humidity, lighting level, pressure, and temperature demonstrates a close alignment between the predicted and actual values over the time series. This correspondence further validates the performance of the model, underscoring the proposed model's effectiveness in predicting these IEQ conditions along with other parameters such as VOC and PM concentration levels.

3.3.4 Evaluation Metrics for Classification Model

The proposed classification model's performance is assessed using four different metrics in comparative and ablation analysis: accuracy, precision, recall, and F1 score. Accuracy (Eq. (3.5)) measures the proportion of true results, including both true positives and true negatives, among the total number of cases examined. Precision (Eq. (3.6)) is the number of true positives for the class divided by the total number of cases that were predicted as belonging to that class. Recall (Eq. (3.7)) measures the proportion of true positives that are correctly identified as such for a particular class. Here, we use TP, TN, FP, and FN to denote the number of true positives, true negatives, false positives, and false negatives in Eqs. (3.5) to (3.7).

$$\text{Accuracy} = \frac{TP + TN}{\text{Total number of samples}} \tag{3.5}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.6}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.7}$$

F1 score (Eq (3.8)) is the single metric that combines recall and precision using harmonic mean, which is more sensitive to low values. The F1 score will only be high if both recall and precision are high. F1 score ranges from 0 to 1, where 0 means the worst possible performance and 1 indicates the best possible performance.

$$F1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.8}$$

3.3.5 Classification Model Comparative and Ablation Analysis

The proposed model fusion-based Transformer model was also put through comprehensive comparative and ablation analysis to validate the model performance. The goal of this analysis was to understand the model's performance compared to other deep learning-based models and the impact of occupant feedback data modalities. All models shared hyperparameters (e.g., batch size, optimization algorithm, and learning rate). The number of epochs was adjusted for each model to better train and optimize. All models were trained and tested on Python 3 using PyTorch and ran on NVIDIA V100.

An in-depth analysis of the proposed model's performance on comfort levels (Table 3.7 and Figure 3.4) was conducted across four key metrics: accuracy, precision, recall, and F1 score. The proposed Transformer performed the best across all metrics. The Transformer model performed on average 2.74% better than LSTM-CNN across all metrics. When compared with the LSTM model, the Transformer model's average improvement was 14.65%. Furthermore, the model's predictive performance was 15.98% superior on average to the CNN model and held a 13.5% higher average over the ANN model. The comparative analysis evidently shows that the Transformer model is much better at classifying comfort levels.

Table 3.7: Proposed model's performance on comfort classification

Model	Accuracy	Precision	Recall	F1 score
Transformer (Proposed)	**0.9776**	**0.9783**	**0.9776**	**0.9776**
LSTM-CNN	0.9536	0.9487	0.9536	0.9509
LSTM	0.8546	0.8505	0.8546	0.8518
CNN	0.8448	0.8405	0.8448	0.8422
ANN	0.8667	0.8610	0.8667	0.8623

Note: **Bold** = best performance

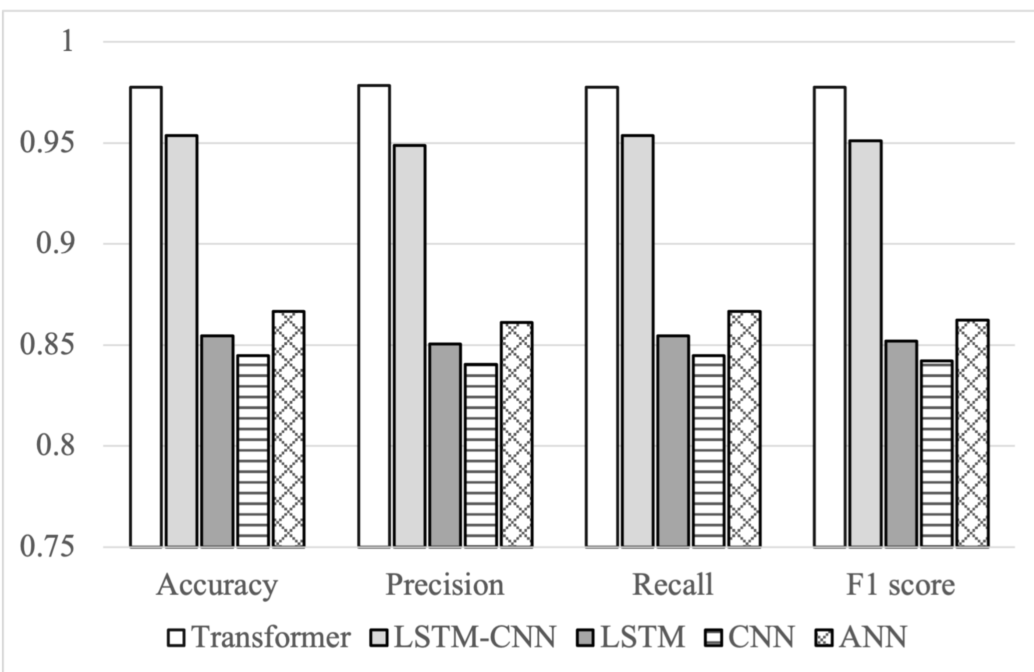


Figure 3.4: Comparative analysis: comfort classification of IEQ conditions

Table 3.8 and Figure 3.5 provide a comprehensive overview of how well each model can classify health levels. Similar to the comfort analysis, the proposed Transformer model displayed the best performance across all metrics when compared to other deep learning-based models. Specifically, the Transformer model performed on average 3.55% better than the LSTM-CNN model, 79.84% better than the LSTM model, 13.20% better than the CNN model, and 4.95% better than the ANN model. These improvements in accuracy, precision, recall, and F1 score show the robustness and reliability of the Transformer model in classifying health levels with high precision.

Table 3.8: Proposed model's performance on health classification

Model	Accuracy	Precision	Recall	F1 score
Transformer (Proposed)	**0.9727**	**0.9734**	**0.9727**	**0.9719**
LSTM-CNN	0.9426	0.9339	0.9426	0.9382
LSTM	0.5792	0.4992	0.5792	0.5156
CNN	0.8585	0.8614	0.8585	0.8585
ANN	0.9268	0.9288	0.9268	0.9249

Note: **Bold** = best performance

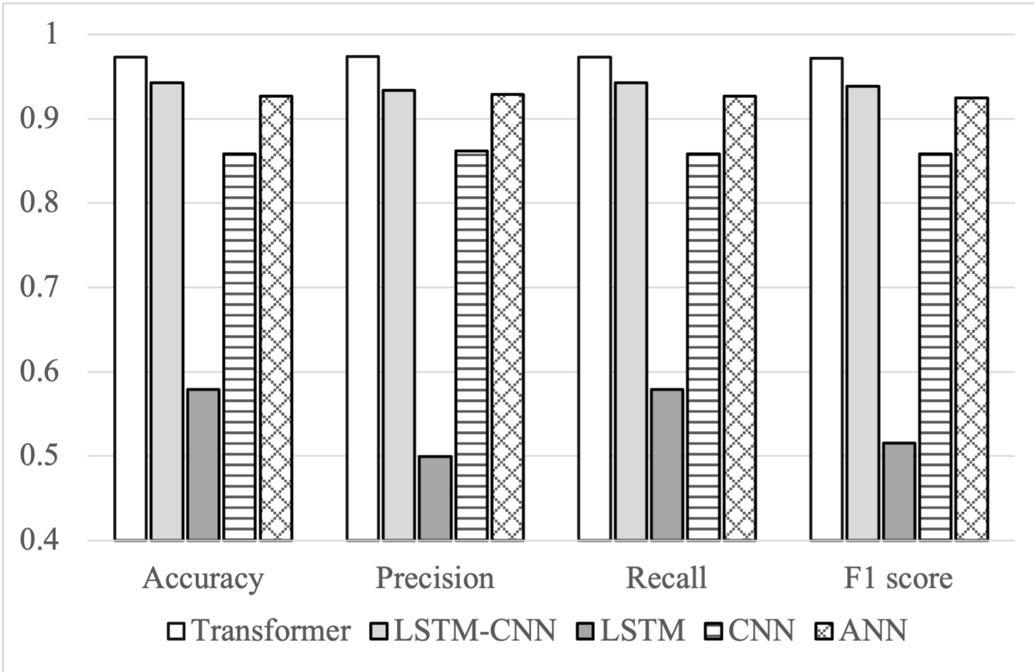


Figure 3.5: Comparative analysis: health classification of IEQ conditions

Table 3.9 and Figure 3.6 show the comfort ablation analysis of the proposed model's performance depending on the different data modalities. The proposed model with all data demonstrates the best performance across all metrics. Comparatively, the model without MFCC features resulted in an average performance decrease of approximately 6.42%. Additionally, the exclusion of video/thermal data resulted in an average decrease of approximately 3.57% in the model's performance, highlighting the benefit of integrating these modalities. These results further emphasize the necessity of each component in achieving the highest model performance.

Table 3.9: Comfort ablation analysis on different data modalities

Data modality	Accuracy	Precision	Recall	F1 score
All data	**0.9776**	**0.9783**	**0.9776**	**0.9776**
Without MFCC	0.9142	0.9212	0.9142	0.9102
Without video/thermal	0.9459	0.9377	0.9459	0.9415
Without any feedback	0.9656	0.9587	0.9656	0.9620

Note: **Bold** = best performance

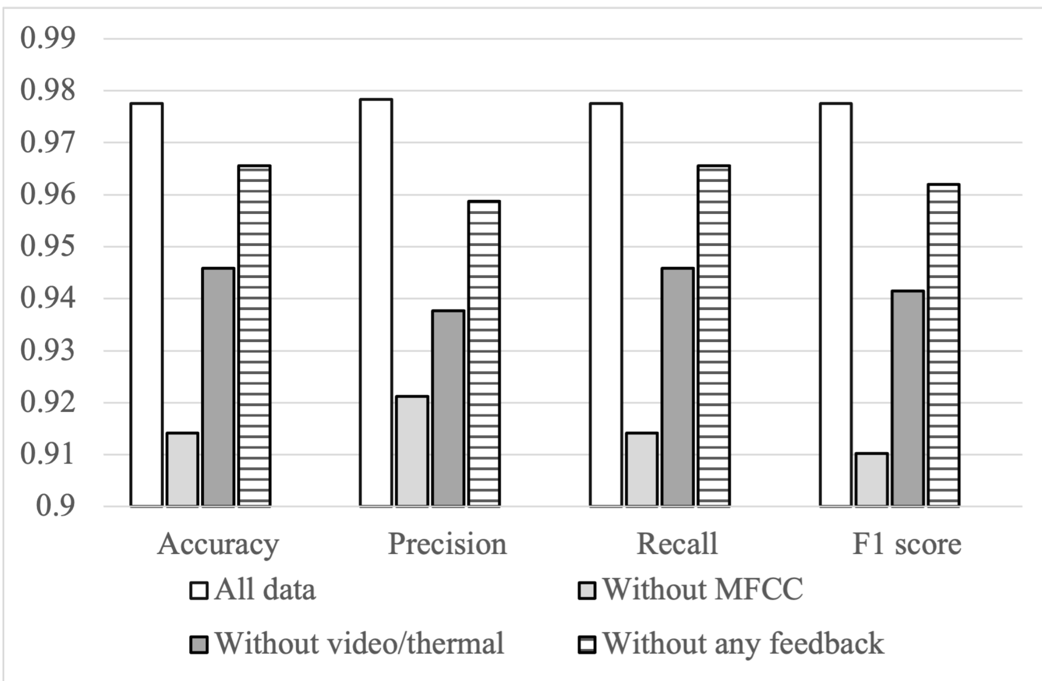


Figure 3.6: Ablation analysis: comfort classification of IEQ conditions

Table 3.10 and Figure 3.7 show the ablation analysis on different data modalities for the model's performance in classifying the health levels. Following the same trend as comfort, the proposed model with all data outperforms all of the other deep learning-based models across all performance metrics. When MFCC features are excluded, the model's performance is decreased by an average of 3.15%. The absence of video/thermal features results in a 3.59% average reduction in all performance metrics. Furthermore, the model's effectiveness diminishes by an average of 1.81% without any form of feedback, which highlights the importance of multimodal occupant feedback.

Table 3.10: Health ablation analysis on different data modalities

Data Modality	Accuracy	Precision	Recall	F1 score
All data	**0.9727**	**0.9734**	**0.9727**	**0.9719**
Without MFCC	0.9415	0.9466	0.9415	0.9384
Without video/thermal	0.9426	0.9301	0.9426	0.9357
Without any feedback	0.9590	0.9488	0.9590	0.9534

Note: **Bold** = best performance

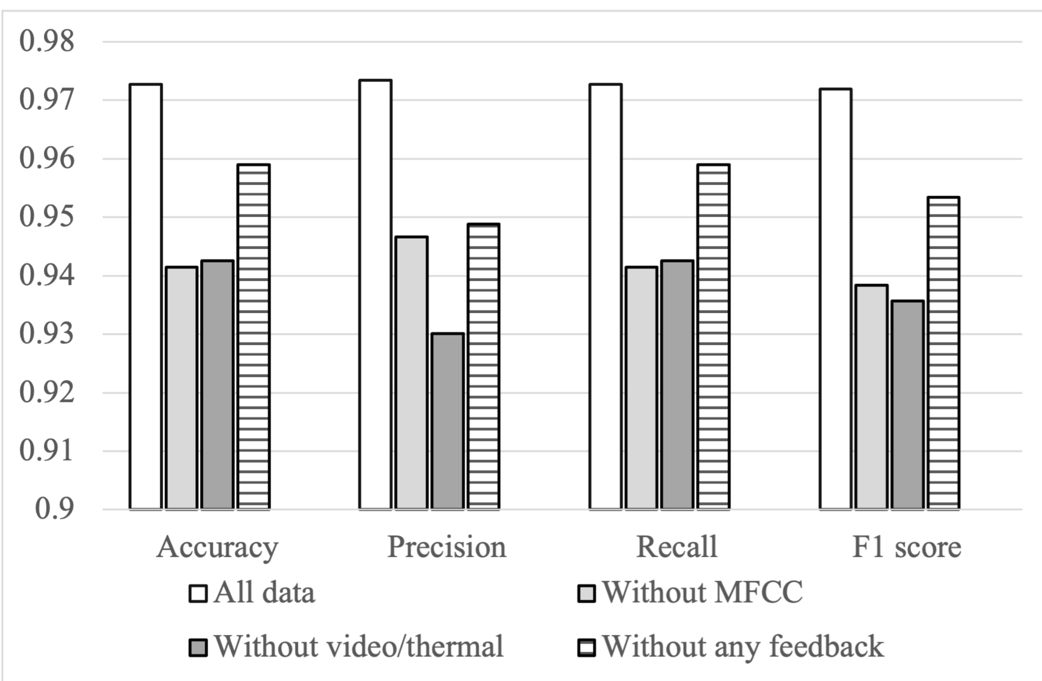


Figure 3.7: Ablation analysis: health classification of IEQ conditions

3.3.6 Confusion Matrix of Classification Model

Two confusion matrices (Figure 3.8) are provided to demonstrate the performance of the proposed model in classifying health and comfort levels. For health-level classification, the model displays a robust capability, as shown by the high concentration of true positives in the middle of the matrix. Similarly, the model's performance in classifying comfort levels is also promising, reflecting a strong classification accuracy. Overall, the two confusion matrices indicate a high level of accuracy in the model's classification abilities.

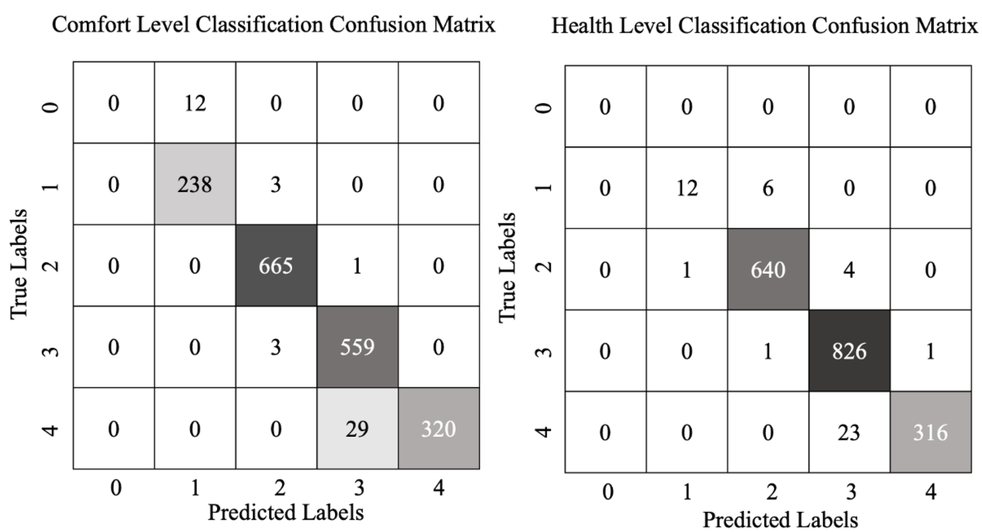


Figure 3.8: Confusion matrices of comfort and health-based IEQ classification

3.3.7 Findings from Comparative and Ablation Analysis

Prediction Model

The results from comprehensive comparative and ablation analysis of both the proposed CNN and LSTM models show key insights in predicting IEQ conditions.

The significant drop in performance (acrosss all four metrics) can be observed when ViT features from video and thermal recordings is removed, emphasizing the importance of ViT features. The second significant performance drop was observed when MFCC features from audio recordings were removed, highlighting the importance of including auditory features into IEQ predictions. Moreover, both models (Proposed CNN and LSTM) that only had IEQ conditions faced the third largest performance drop in model performance, underscoring the importance of including both non-intrusive occupant feedback and self-reported comfort and health levels. The removal of self-reported comfort and health levels, while not as impactful as the first three input modalities, still noticeably affected model performance. In summary, the findings collectively highlights the importance of a comprehensive and holistic approach to IEQ modeling, where the combination of multidimensional IEQ data and multimodal occupant feedback contributes to more accurate and reliable prediction model. Furthermore, the comparative analysis between ViT and ResNet50 reveals that ViT is more suitable for discerning and identifying patterns from video and thermal recordings and well-suited to the multifaceted nature of indoor environment assessment.

Classification Model

The integration of multimodal occupant feedback into the proposed Transformer-based classification model is important, as shown by its superior performance over other deep learning models (LSTM, CNN, LSTM-CNN, and ANN). By utilizing a rich combination of auditory, visual, and thermal data, the proposed Transformer model captures a more holistic understanding of occupant comfort and health, which is reflected in its consistently higher accuracy, precision, recall, and F1 score. While the proposed Transformer model achieves the best performance, the ablation analysis shows varying degrees of influence each data modality exerts on the classification model's performance. MFCC features had the most impact on the model's performance for both comfort and health classification as shown in our ablation analysis. Removing ViT features from video/thermal recordings had the second biggest drop in model performance, however, the results were not as impactful as removing MFCC features. In summary, this multimodal approach enables the model to recognize subtle patterns that single-mode data might miss, resulting in a more robust and reliable system for assessing occupant comfort and health.

Chapter 4

Conclusions

This thesis presents an improvement toward more intelligent and occupant-centric assessment of indoor environments, particularly in educational facilities. By leveraging the power of deep learning, IoT devices, and multimodal data fusion, it offers a novel approach to IEQ modeling and assessment based on both environmental (multidimensional IEQ data) and multimodal occupant data (self-reported comfort and health levels and non-intrusive occupant feedback), which can support the future development of data-driven and occupant-centric educational facilities. The proposed AIoT framework consists of five main components: (1) smart and connected devices for collecting multidimensional IEQ data, (2) multimodal methods for collecting non-intrusive occupant feedback data (MFCC features from audio recordings and ViT features from video and thermal recordings), (3) CNN-based multimodal data fusion model for IEQ prediction, (4) Transformer-based multimodal data fusion model for IEQ classification based on self-reported comfort and health levels. Furthermore, dual-model integration system for proactive and adaptive assessment of indoor environment is discussed. The proposed deep learning model architectures (both predictive and classificatory models) utilize pre-trained feature extraction models, ViT, and a CNN and Transformer-based model with a multimodal data fusion approach. The learned features from video, thermal, audio, and IEQ data, as well as the self-reported comfort and health levels from occupants (e.g., students and faculty), are then fused by each model, allowing the framework to predict and classify multidimensional IEQ conditions. This thesis stands out for its emphasis on occupant feedback and its integration of multiple data modalities, which enable a more adaptive and personalized response to individual needs. The proposed framework classifies IEQ conditions with promising performance, achieving 97% and 96% (comfort and health) accuracy, and predicts IEQ conditions with a 91.9% R^2 score.

4.1 Limitations

Even though the results collected from the experiments are quite promising, the limitations of this study have been recognized for furthering the context and setting directions for future research.

Outdoor environmental factors. The proposed framework focuses on indoor environment parameters and multimodal occupant feedback, possibly underestimating the impact of external environmental conditions on IEQ and occupants. Factors such as outdoor air quality, meteorological conditions, and urban heat island effects can significantly influence indoor environments. The lack of integration of these external factors into the IEQ assessment model may limit its ability to provide a holistic assessment of the factors affecting indoor comfort and health, potentially compromising the accuracy and relevance of the predictions in real-world settings.

Subjective occupant experience. The framework's performance is partly reliant on the accuracy and reliability of self-reported comfort and health levels. While this approach provides valuable insights into the subjective experiences of occupants, it is subject to individual biases and variations in perception. This subjective nature of self-reported data can lead to inconsistencies, where different individuals may interpret or respond to the same environmental conditions in diverse ways. This variability poses a challenge in accurately modeling and assessing IEQ conditions based solely on self-reported metrics, as it introduces a layer of subjectivity that may not always align with objective measures of IEQ.

Data diversity. The model was trained on data collected with varying spatial, locational, and temporal configurations, and specific environmental conditions were intentionally modified to enhance data diversity and model performance. However, it's important to acknowledge that the collected datasets might not fully encapsulate the vast range of real-world environments and scenarios. Variations in data, including diverse contexts, indoor climates, and occupant behaviors, may not be fully represented, which could affect the model's capacity to generalize its predictions to unseen situations. Additionally, the model's reliance on self-reported comfort and health levels to improve performance will need to be content with the practical challenges of consistently obtaining accurate, timely, and consistent feedback in real-world settings.

Quality of non-intrusive occupant feedback. While extensive preprocessing and normalization steps were implemented, noise or errors inherent to video and thermal images could affect the performance of the model's efficiency. Improvements in data cleaning and higher-resolution cameras for enhanced quality assurance can further optimize the model performance. Additionally, while thermal comfort is inherently subjective and can differ among individuals, it is hoped that the feature extraction power of ViT could discern subtle visuals correlating with general comfort levels.

Temporal alignment and data synchronization. The assumptions of perfect temporal alignment in the data, with all data types synchronized at one-second intervals, may not hold true in practice. Slight inaccuracies or discrepancies in the alignment could impact the predictive power of the model.

Focus on Occupant-Centric Building Control This thesis is predominantly focused on improving occupant comfort and health by prioritizing occupant factors within the indoor environment. While this focus brings considerable advancements in understanding and improving IEQ conditions, it may overlook the broader operational efficiencies of building operation such as energy consumption and system-wide optimization.

4.2 Contribution to the Body of Knowledge

The proposed AIoT framework for intelligent educational facilities contributes to the evolving body of knowledge at the intersection of IoT, deep learning, multimodal data fusion, and indoor environment monitoring and assessment, especially within the context of educational facilities. The framework applies deep learning and multimodal occupant feedback collected from various sensors (IoT) and monitoring devices, creating a data-driven and occupant-centric model for modeling and predicting IEQ conditions of educational facilities, and addressing the diverse needs and experiences of students. The contributions of the proposed AIoT framework are fivefold.

Use of self-reported occupant comfort and health levels and non-intrusive feedback. Key to this framework is the inclusion of various occupant feedback, which current building control and indoor environment systems often neglect. The emphasis on occupant feedback, both self-reported levels and inferred through non-intrusive means, for IEQ assessment, is another contribution. By focusing on the subjective experiences of occupants regarding their comfort and health levels, the framework aligns closely with the principles of occupant-centric control (OCC). This alignment naturally facilitates downstream tasks such as indoor environment management and building system control. This research underscores the importance of considering the human dimension in building assessment/management systems, advocating for a shift towards environments that adapt to the needs and preferences of their users.

Advancement in multimodal data fusion. The proposed framework utilizes multimodal data fusion, which brings together diverse data sources such as multidimensional IEQ data from IoT devices, self-reported comfort and health levels through a web-based UI, MFCC features from audio recordings, and deep learning-extracted features from video and thermal recordings. This approach allows for a richer, more comprehensive understanding of the indoor environment, surpassing traditional methods that often rely on singular data sources.

Use of advanced deep learning models and comprehensive assessment of IEQ conditions. By developing a framework capable of classifying and predicting multidimensional IEQ conditions based on a holistic set of parameters (e.g., self-reported comfort and health levels and non-intrusive occupant feedback), the proposed framework addresses gaps in existing IEQ modeling and assessment methods. Leveraging and adapting Transformer models for classification tasks, traditionally used in NLP and CV tasks, and CNN-based model IEQ prediction marks a novel application of these advanced deep learning models in the field of time series and behavioral data analytics. The ability of Transformer models, especially following the multimodal data fusion approach, to handle temporal dynamics and relationships within IEQ data and occupant feedback signifies a leap forward in IEQ modeling and assessment capabilities.

Impact of occupant feedback in IEQ modeling and assessment. This thesis explores the impact and intricate relationship of different types of occupant feedback data in predicting multidimensional IEQ conditions through comprehensive ablation and comparative analysis of model variations.

Generalizability and adaptability of the proposed framework. The proposed framework's design facilitates generalizability to different experimental settings. Researchers and practitioners can input their datasets, including CSV files for IEQ data, self-reported comfort and health levels, video, and thermal recordings, and adjust both model and training hyperparameters to cater to their specific needs. Furthermore, the overall cost to implement the proposed framework is considerably cheaper than the existing practices.

4.3 Extension and Downstream Tasks Enabled by the Proposed AIoT Framework

Integration of a proposed prediction and classification models system brings about a shift in the approaches used in indoor environment assessment within educational facilities. The proposed dual-model integration consists of two distinct models: the IEQ prediction model and the IEQ classification model, each serving a unique role within the integrated framework to enhance the overall quality of indoor spaces within educational facilities.

The proposed predictive model serves in the forward-looking perspective to allow proactive adjustments and inform long-run strategic planning. In other words, this model can predict what state the indoor environment is going to take on at a time in the future in such a manner that the maintenance task or scheduling of HVAC operation by facilities management occurs in an anticipatory manner (any downstream tasks) rather than being carried out in a reaction to not-maintained IEQ or equipment state.

The proposed classification model, on the other hand, provides an immediate indoor environment condition assessment (based on self-reported comfort and health levels) that can be easily and almost immediately diagnosed against the optimal IEQ conditions. This classification model is designed to aggregate data from environmental sensors as well as occupant feedback to evaluate the present conditions of comfort and health, thereby enabling the prompt execution of appropriate actions based on how occupants feel in the indoor environment.

Based on the outcomes from the prediction and classification models, buildings by themselves can take informed, rule-based actions to maintain and optimize the indoor environment. For instance, if the prediction model forecasts that the temperature will rise above an optimal threshold - determined by standards like ASHRAE and research on ideal learning conditions - and the classification model simultaneously indicates occupant comfort and health levels are below neutral, the system can initiate a change. Specifically, it could adjust the HVAC system's temperature set point to ensure the environment remains conducive to comfort and effective learning. Such rule-based systems enable the system to automatically respond to changing IEQ conditions, ensuring that occupants' comfort and health are actively maintained without the need for manual intervention.

Furthermore, the classification model acts as a checking mechanism for the predicted multidimensional IEQ conditions, verifying that the predicted IEQ conditions are well in line with occupants' perceived comfort and health levels, strengthening the credibility and validity of the system. Leveraging a data-driven approach, the proposed framework predicts and classifies multidimensional IEQ conditions, enabling buildings (or facility managers) to selectively extract specific IEQ data pertinent to targeted tasks, thus optimizing the indoor environment in alignment with the evolving needs of the building occupants. Through double strengths of predictions and real-time classifications, this proposed framework aligns the assessment of educational facilities to the dynamic needs of the occupants and hence promises an intelligent, adaptive, and occupant-centric AIoT framework for intelligent educational facilities.

4.4 Future Works

Despite the promising results from the proposed AIoT framework, there is room for improvement for future work. First, future work will incorporate additional data modalities such as occupancy level and outdoor environmental data (e.g., meteorological conditions and outdoor air quality). This could lead to a system where the building can sense both the indoor environment, outdoor environment, and occupants for improved model performance. Second, human factor analysis will be conducted. The authors plan to investigate how different people express different levels of comfort or health levels. Our future framework will include exploring alternative consensus-building occupant feedback methods. Third, longitudinal studies will be conducted to understand the long-term effects of optimized IEQ on comfort and health. Fourth, future research will be conducted to compare the effectiveness of proposed framework when compared to the traditional operational methods in terms of energy usage/efficiency. Our ultimate goal is to leverage the insights from IEQ conditions to create ubiquitous sensing and computing-supported intelligent and occupant-centric indoor environments that enhance occupants' health, comfort, and well-being.

References

Afroz, Z., Urmee, T., Shafiullah, G. M., & Higgins, G. (2018). Real-time prediction model for indoor temperature in a commercial building. *Applied Energy*, *231*, 29–53. https://doi.org/10.1016/j.apenergy.2018.09.052

Al horr, Y., Arif, M., Katafygiotou, M., Mazroei, A., Kaushik, A., & Elsarrag, E. (2016). Impact of indoor environmental quality on occupant well-being and comfort: A review of the literature. *International Journal of Sustainable Built Environment*, *5*(1), 1–11. https://doi.org/10.1016/j.ijsbe.2016.03.006

Ali, H. H., Abdullah, M., & Wedyan, M. (2022). Application of machine learning techniques to predict patient's satisfaction of indoor environmental quality in Jordanian hospitals. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-022-04021-6

Allen-Zhu, Z., & Li, Y. (2019). *What Can ResNet Learn Efficiently, Going Beyond Kernels? **. https://proceedings.neurips.cc/paper_files/paper/2019/file/5857d68cd9280bc98d079fa912fd6740-Paper.pdf

Anik, S. M. H., Gao, X., Meng, N., Agee, P. R., & McCoy, A. P. (2022). A cost-effective, scalable, and portable IoT data infrastructure for indoor environment sensing. Journal of Building Engineering, 49, 104027. https://doi.org/10.1016/j.jobe.2022.104027

Beemer, C. J., Stearns-Yoder, K. A., Schuldt, S. J., Kinney, K. A., Lowry, C. A., Postolache, T. T., Brenner, L. A., & Hoisington, A. J. (2019). A brief review on the mental health for select elements of the built environment. *Indoor and Built Environment*, 1420326X1988965. https://doi.org/10.1177/1420326x19889653

Brink, H.W., Loomans, M.G., Mobach, M.P. and Kort, H.S., 2021. Classrooms' indoor environmental conditions affecting the academic achievement of students and teachers in higher education: A systematic literature review. Indoor air, 31(2), pp.405-425.

Bucarelli, N., & El-Gohary, N. (2023). Deep learning approach for recognizing cold and warm thermal discomfort cues from videos. *Building and Environment*, 110277. https://doi.org/10.1016/j.buildenv.2023.110277

Campano-Laborda, M. Á., Domínguez-Amarillo, S., Fernández-Agüera, J., & Acosta, I. (2020). Indoor Comfort and Symptomatology in Non-University Educational Buildings: Occupants' Perception. *Atmosphere*, *11*(4), 357. https://doi.org/10.3390/atmos11040357

Chen, F.C. and Jahanshahi, M.R., 2017. NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion. IEEE Transactions on Industrial Electronics, 65(5), pp.4392-4400.

Chen, K., Chen, S., Zhu, X., Jin, X., & Du, Z. (2023). Interpretable mechanism mining enhanced deep learning for fault diagnosis of heating, ventilation and air conditioning systems. *Building and Environment*, *237*, 110328. https://doi.org/10.1016/j.buildenv.2023.110328

Cheng, C., Deng, X., Zhao, X., Xiong, Y., & Zhang, Y. (2023). Multi-occupant dynamic thermal comfort monitoring robot system. *Building and Environment*, *234*, 110137. https://doi.org/10.1016/j.buildenv.2023.110137

Cho, S., Hong Jae Nam, Shi, C., Choong Yeon Kim, Byun, S., Agno, K.-C., Byung Chul Lee, Xiao, J., Joo Yong Sim, & Jeong, J.-W. (2023). Wireless, AI-enabled wearable thermal comfort sensor for energy-efficient, human-in-the-loop control of indoor temperature. *Biosensors and Bioelectronics*, *223*, 115018–115018. https://doi.org/10.1016/j.bios.2022.115018

Choi, S., Guerin, D. A., Kim, H.-Y., Brigham, J. K., & Bauer, T. (2014). Indoor Environmental Quality of Classrooms and Student Outcomes: A Path Analysis Approach. *Journal of Learning Spaces*, *2*(2). https://libjournal.uncg.edu/jls/article/view/506/485

Cincinelli, A., & Martellini, T. (2017). Indoor Air Quality and Health. *International Journal of Environmental Research and Public Health*, *14*(11), 1286. https://doi.org/10.3390/ijerph14111286

Day, J. K., McIlvennie, C., Brackley, C., Tarantini, M., Piselli, C., Hahn, J., O'Brien, W., Rajus, V. S., De Simone, M., Kjærgaard, M. B., Pritoni, M., Schlüter, A., Peng, Y., Schweiker, M., Fajilla, G., Becchio, C., Fabi, V., Spigliantini, G., Derbas, G., & Pisello, A. L. (2020). A review of select human-building interfaces and their relationship to human behavior, energy use and occupant comfort. *Building and Environment*, *178*, 106920. https://doi.org/10.1016/j.buildenv.2020.106920

Deng, M., Wang, X., Li, D., & Menassa, C. C. (2022). Digital ID framework for human-centric monitoring and control of smart buildings. *Building Simulation*. https://doi.org/10.1007/s12273-022-0902-3

Dikel, E. E., Newsham, G. R., Xue, H., & Valdés, J. J. (2018). Potential energy savings from high-resolution sensor controls for LED lighting. *Energy and Buildings*, *158*, 43–53. https://doi.org/10.1016/j.enbuild.2017.09.048

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S. and Uszkoreit, J., 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

Duarte Roa, C., Schiavon, S., & Parkinson, T. (2020). Targeted occupant surveys: A novel method to effectively relate occupant feedback with environmental conditions. *Building and Environment*, *184*, 107129. https://doi.org/10.1016/j.buildenv.2020.107129

Earthman, G.I. (2004) Prioritization of 31 Criteria for School Building Adequacy, American Civil Liberties Union Foundation of Maryland. http://www.aclumd.org/aTop%20Issues/Education%20Reform/EarthmanFinal10504.pdf

EPA. (2018, July 16). *Indoor Air Quality | US EPA*. US EPA. https://www.epa.gov/report-environment/indoor-air-quality

Erickson, V. L., Achleitner, S., & Cerpa, A. E. (2013). *POEM: power-efficient occupancy-based energy management system*. https://doi.org/10.1145/2461381.2461407

Feng, X., Zainudin, E. B., Wong, H. W., & Tseng, K. J. (2023). A hybrid ensemble learning approach for indoor thermal comfort predictions utilizing the ASHRAE RP-884 database. *Energy and Buildings*, *290*, 113083. https://doi.org/10.1016/j.enbuild.2023.113083

Fransson, N., Västfjäll, D., & Skoog, J. (2007). In search of the comfortable indoor environment: A comparison of the utility of objective and subjective indicators of indoor comfort. *Building and Environment*, *42*(5), 1886–1890. https://doi.org/10.1016/j.buildenv.2006.02.021

Fritz, H., Kinney, K. A., Wu, C., Schnyer, D. M., & Nagy, Z. (2022). Data fusion of mobile and environmental sensing devices to understand the effect of the indoor environment on measured and self-reported sleep quality. *Building and Environment*, *214*, 108835. https://doi.org/10.1016/j.buildenv.2022.108835

Frontczak, M., & Wargocki, P. (2011). Literature survey on how different factors influence human comfort in indoor environments. *Building and Environment*, *46*, 922–937. https://doi.org/10.1016/j.buildenv.2010.10.021

Frontczak, M., Schiavon, S., Goins, J., Arens, E., Zhang, H., & Wargocki, P. (2011). Quantitative relationships between occupant satisfaction and satisfaction aspects of indoor environmental quality and building design. *Indoor Air*, *22*, 119–131. https://doi.org/10.1111/j.1600-0668.2011.00745.x

Gan, V. J. L., Luo, H., Tan, Y., Deng, M., & Kwok, H. L. (2021). BIM and Data-Driven Predictive Analysis of Optimum Thermal Comfort for Indoor Environment. *Sensors*, *21*(13), 4401. https://doi.org/10.3390/s21134401

Gao, P. X., & Keshav, S. (2013). SPOT: a smart personalized office thermal control system. *Proceedings of the the Fourth International Conference on Future Energy Systems - E-Energy '13*. https://doi.org/10.1145/2487166.2487193

Gao, X., Pishdad-Bozorgi, P., Shelden, D. R., & Tang, S. (2021). Internet of Things Enabled Data Acquisition Framework for Smart Building Applications. *Journal of Construction Engineering and Management*, *147*(2), 04020169. https://doi.org/10.1061/(asce)co.1943-7862.0001983

Gong, P., Cai, Y., Chen, B., Zhang, C., Stravoravdis, S., Sharples, S., Ban, Q., & Yu, Y. (2023). An Artificial Neural Network-based model that can predict inpatients' personal thermal sensation in rehabilitation wards. *Journal of Building Engineering*, *80*, 108033–108033. https://doi.org/10.1016/j.jobe.2023.108033

Guo, L., Wang, G., & Yu, X. (2016). Design for Indoor Environment Monitoring System based on Embedded System and Multi-sensor Data Fusion Algorithm. *International Journal of Smart Home*, *10*(1), 31–40. https://doi.org/10.14257/ijsh.2016.10.1.04

Gupta, S. K., Atkinson, S., O'Boyle, I., Drogo, J., Kar, K., Mishra, S., & Wen, J. T. (2016). BEES: Real-time occupant feedback and environmental learning framework for collaborative thermal management in multi-zone, multi-occupant buildings. *Energy and Buildings*, *125*, 142–152. https://doi.org/10.1016/j.enbuild.2016.04.084

He, X., Zhang, X., Zhang, R., Liu, J., Huang, X., Pei, J., Cai, J., Guo, F., & Wang, Y. (2023). More intelligent and efficient thermal environment management: A hybrid model for occupant-centric thermal comfort monitoring in vehicle cabins. *Building and Environment*, *228*, 109866. https://doi.org/10.1016/j.buildenv.2022.109866

Huang, Q., Mao, C. and Chen, Y., 2017. A compact and versatile wireless sensor prototype for affordable intelligent sensing and monitoring in smart buildings. In *Computing in Civil Engineering 2017* (pp. 155-161).

Jazizadeh, F., Ghahramani, A., Becerik-Gerber, B., Kichkaylo, T., & Orosz, M. (2014). User-led decentralized thermal comfort driven HVAC operations for improved efficiency in office buildings. *Energy and Buildings*, *70*, 398–410. https://doi.org/10.1016/j.enbuild.2013.11.066

Jiang, B., Gong, H., Qin, H., & Zhu, M. (2022). Attention-LSTM architecture combined with Bayesian hyperparameter optimization for indoor temperature prediction. *Building and Environment*, *224*, 109536. https://doi.org/10.1016/j.buildenv.2022.109536

Jones, A. P. (1999). Indoor air quality and health. *Atmospheric Environment*, *33*(28), 4535–4564. https://doi.org/10.1016/s1352-2310(99)00272-1

Kallio, J., Tervonen, J., Räsänen, P., Mäkynen, R., Koivusaari, J., & Peltola, J. (2021). Forecasting office indoor CO2 concentration using machine learning with a one-year dataset. *Building and Environment*, *187*, 107409. https://doi.org/10.1016/j.buildenv.2020.107409

Kallio, J., Vildjiounaite, E., Koivusaari, J., Räsänen, P., Similä, H., Kyllönen, V., Muuraiskangas, S., Ronkainen, J., Rehu, J., & Vehmas, K. (2020). Assessment of perceived indoor environmental quality, stress and productivity based on environmental sensor data and personality categorization. *Building and Environment*, *175*, 106787. https://doi.org/10.1016/j.buildenv.2020.106787

Kamaruzzaman, S. N., Egbu, C. O., Mahyuddin, N., Ahmad Zawawi, E. M., Chua, S. J. L., & Azmi, N. F. (2017). The impact of IEQ on occupants' satisfaction in Malaysian buildings. *Indoor and Built Environment*, *27*(5), 715–725. https://doi.org/10.1177/1420326x16689493

Kanezaki, A., Kuga, R., Sugano, Y., & Matsushita, Y. (2019). Deep Learning for Multimodal Data Fusion. *Elsevier EBooks*, 9–39. https://doi.org/10.1016/b978-0-12-817358-9.00008-1

Karakas, T., & Yildiz, D. (2020). Exploring the influence of the built environment on human experience through a neuroscience approach: A systematic review. *Frontiers of Architectural Research*. https://doi.org/10.1016/j.foar.2019.10.005

Kaushik, A. K., Arif, M., Syal, M. M. G., Rana, M. Q., Oladinrin, O. T., Sharif, A. A., & Alshdiefat, A. S. (2022). Effect of Indoor Environment on Occupant Air Comfort and Productivity in Office Buildings: A Response Surface Analysis Approach. *Sustainability*, *14*(23), 15719. https://doi.org/10.3390/su142315719

Khan, A. N., Rizwan, A., Ahmad, R., & Kim, D. H. (2023). An OCF-IoTivity enabled smart-home optimal indoor environment control system for energy and comfort optimization. *Internet of Things*, *22*, 100712. https://doi.org/10.1016/j.iot.2023.100712

Kim, J., Bauman, F., Raftery, P., Arens, E., Zhang, H., Fierro, G., Andersen, M., & Culler, D. (2019). Occupant comfort and behavior: High-resolution data from a 6-month field study of personal comfort systems with 37 real office workers. *Building and Environment*, *148*, 348–360. https://doi.org/10.1016/j.buildenv.2018.11.012

Kim, J., Kim, H., & Hong, T. (2019). Automated classification of indoor environmental quality control using stacked ensembles based on electroencephalograms. *Computer-Aided Civil and Infrastructure Engineering*, *35*(5), 448–464. https://doi.org/10.1111/mice.12515

Klein, L., Kwak, J., Kavulya, G., Jazizadeh, F., Becerik-Gerber, B., Varakantham, P., & Tambe, M. (2012). Coordinating occupant behavior for building energy and comfort management using multi-agent systems. *Automation in Construction*, *22*, 525–536. https://doi.org/10.1016/j.autcon.2011.11.012

Lahat, D., Adali, T., & Jutten, C. (2015). Multimodal Data Fusion: An Overview of Methods, Challenges, and Prospects. *Proceedings of the IEEE*, *103*(9), 1449–1477. https://doi.org/10.1109/jproc.2015.2460697

Lai, J. H. K., & Yik, F. W. H. (2007). Perceived Importance of the Quality of the Indoor Environment in Commercial Buildings. *Indoor and Built Environment*, *16*(4), 311–321. https://doi.org/10.1177/1420326x07080463

Lam, A. H., Yuan, Y., & Wang, D. (2014). An occupant-participatory approach for thermal comfort enhancement and energy conservation in buildings. *CiteSeer X (the Pennsylvania State University)*. https://doi.org/10.1145/2602044.2602067

Lee, M. J., & Zhang, R. (2024). Human-Centric Artificial Intelligence of Things–Based Indoor Environment Quality Modeling Framework for Supporting Student Well-Being in Educational Facilities. *Journal of Computing in Civil Engineering*, *38*(2). https://doi.org/10.1061/jccee5.cpeng-5632

Marino, C., Nucara, A., & Pietrafesa, M. (2012). Proposal of comfort classification indexes suitable for both single environments and whole buildings. *Building and Environment*, *57*, 58–67. https://doi.org/10.1016/j.buildenv.2012.04.012

Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). A survey on machine learning for data fusion. *Information Fusion*, *57*, 115–129. https://doi.org/10.1016/j.inffus.2019.12.001

Mujan, I., Anđelković, A. S., Munćan, V., Kljajić, M., & Ružić, D. (2019). Influence of indoor environmental quality on human health and productivity - A review. *Journal of Cleaner Production*, *217*, 646–657. https://doi.org/10.1016/j.jclepro.2019.01.307

Nagy, Z., Yong, F. Y., & Schlueter, A. (2016). Occupant centered lighting control: A user study on balancing comfort, acceptance, and energy consumption. *Energy and Buildings*, *126*, 310–322. https://doi.org/10.1016/j.enbuild.2016.05.075

Orman, Ł.J.; Krawczyk, N.; Radek, N.; Honus, S.; Pietraszek, J.; Dębska, L.; Dudek, A.; Kalinowski, A. Comparative Analysis of Indoor Environmental Quality and Self-Reported Productivity in

Intelligent and Traditional Buildings. *Energies* 2023, *16*, 6663. https://doi.org/10.3390/en16186663

Ortiz, M. A., & Bluyssen, P. M. (2022). Profiling office workers based on their self-reported preferences of indoor environmental quality and psychosocial comfort at their workplace during COVID-19. *Building and Environment*, *211*, 108742. https://doi.org/10.1016/j.buildenv.2021.108742

Park, J. Y., Dougherty, T., Fritz, H., & Nagy, Z. (2019). LightLearn: An adaptive and occupant centered controller for lighting based on reinforcement learning. *Building and Environment*, *147*, 397–414. https://doi.org/10.1016/j.buildenv.2018.10.028

Park, B. K., & Kim, C.-J. (2023). Short-Term Prediction for Indoor Temperature Control Using Artificial Neural Network. *Energies*, *16*(23), 7724–7724. https://doi.org/10.3390/en16237724

Patnaik, A., Kumar, V., & Saha, P. (2017). Importance of Indoor Environmental Quality in Green Buildings. *Environmental Pollution*, 53–64. https://doi.org/10.1007/978-981-10-5792-2_5

Peng, Y., Rysanek, A., Nagy, Z., & Schlüter, A. (2018). Using machine learning techniques for occupancy-prediction-based cooling control in office buildings. *Applied Energy*, *211*, 1343–1358. https://doi.org/10.1016/j.apenergy.2017.12.002

Pimoroni, LTR-559 Light & Proximity Sensor Breakout. (n.d.). Shop.pimoroni.com. Retrieved July 10, 2023, from https://shop.pimoroni.com/products/ltr-559-light-proximity-sensor-breakout?variant=13571092611155

Pollard, B., Held, F., Engelen, L., Powell, L., & de Dear, R. (2021). Data fusion in buildings: Synthesis of high-resolution IEQ and occupant tracking data. *Science of the Total Environment*, *776*, 146047. https://doi.org/10.1016/j.scitotenv.2021.146047

Pritoni, M., Salmon, K., Sanguinetti, A., Morejohn, J., & Modera, M. (2017). Occupant thermal feedback for improved efficiency in university buildings. *Energy and Buildings*, *144*, 241–250. https://doi.org/10.1016/j.enbuild.2017.03.048

Quintana, M., Schiavon, S., Tartarini, F., Kim, J., & Miller, C. (2023). Cohort comfort models — Using occupant's similarity to predict personal thermal preference with less data. *Building and Environment*, *227*, 109685. https://doi.org/10.1016/j.buildenv.2022.109685

Rubinstein, F., & Enscoe, A. (2010). Saving Energy with Highly-Controlled Lighting in an Open-Plan Office. *LEUKOS*, *7*(1), 21–36. https://doi.org/10.1582/leukos.2010.07.01002

Samet, J. M., & Spengler, J. D. (2003). Indoor Environments and Health: Moving Into the 21st Century. *American Journal of Public Health*, *93*(9), 1489–1493. https://doi.org/10.2105/ajph.93.9.1489

Shan, X., Melina, A. N., & Yang, E.-H. (2018). Impact of indoor environmental quality on students' wellbeing and performance in educational building through life cycle costing perspective. *Journal of Cleaner Production*, *204*, 298–309. https://doi.org/10.1016/j.jclepro.2018.09.002

Sharma, P. K., De, T., & Saha, S. (2018). IoT based indoor environment data modelling and prediction. IEEE Xplore. https://doi.org/10.1109/COMSNETS.2018.8328266

Song, Z., Zhen, S., Ji, K., & Lu, Y. (2013). Collaborative building control to optimize energy saving and improve occupants' experience. *ASHRAE Transactions*, *119*(1). https://link.gale.com/apps/doc/A331928011/AONE?u=anon~bb4f7c8d&sid=googleScholar&xid=be8a3307

Sussman, A., & Hollander, J. (2021). *Cognitive Architecture: Designing for How We Respond to the Built Environment*. Routledge & CRC Press. https://www.routledge.com/Cognitive-Architecture-Designing-for-How-We-Respond-to-the-Built-Environment/Sussman-Hollander/p/book/9780367468590

Tagliabue, L. C., Re Cecconi, F., Rinaldi, S., & Ciribini, A. L. C. (2021). Data driven indoor air quality prediction in educational facilities based on IoT network. *Energy and Buildings*, *236*, 110782. https://doi.org/10.1016/j.enbuild.2021.110782

Tekler, Z. D., Low, R., Yuen, C., & Blessing, L. (2022). Plug-Mate: An IoT-based occupancy-driven plug load management system in smart buildings. *Building and Environment*, *223*, 109472. https://doi.org/10.1016/j.buildenv.2022.109472

Uline, C., & Tschannen-Moran, M. (2008). The walls speak: the interplay of quality facilities, school climate, and student achievement. *Journal of Educational Administration*, *46*(1), 55–73. https://doi.org/10.1108/09578230810849817

Ultra Librarian, BME280 Datasheet: Temperature, Pressure & Humidity Sensor for Wearable Technology. (2022, February 15). Free Online PCB CAD Library. https://www.ultralibrarian.com/2022/02/15/bme280-datasheet-temperature-pressure-humidity-sensor-for-wearable-technology-ulc#:~:text=As%20shown%20above%2C%20the%20BME280

USGBC. (2014). *Green Building 101: What is indoor environmental quality? | U.S. Green Building Council*. Usgbc.org. https://www.usgbc.org/articles/green-building-101-what-indoor-environmental-quality

Virginia Law. (2023). *22VAC40-73-880. Heating, ventilation, and cooling.* Virginia.gov. https://law.lis.virginia.gov/admincode/title22/agency40/chapter73/section880/#:~:text=A%20temperature%20of%20at%20least

Wang, C. C., & Zamri, Mohd. A. (2013). Effect of IEQ on Occupant Satisfaction and Study/Work Performance in a Green Educational Building: A Case Study. *ICCREM 2013*. https://doi.org/10.1061/9780784413135.022

Wang, W., Chen, J., & Hong, T. (2018). Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings. *Automation in Construction*, *94*, 233–243. https://doi.org/10.1016/j.autcon.2018.07.007

Wang, Z., Hong, T. and Piette, M.A., 2019. Data fusion in predicting internal heat gains for office buildings through a deep learning approach. Applied energy, 240, pp.386-398.

Winkler, D. A., Beltran, A., Esfahani, N. P., Maglio, P. P., & Cerpa, A. E. (2016). *FORCES*. https://doi.org/10.1145/2971648.2971700

Xing, T., Sun, K., & Zhao, Q. (2023). MITP-Net: A deep-learning framework for short-term indoor temperature predictions in multi-zone buildings. *Building and Environment*, *239*, 110388–110388. https://doi.org/10.1016/j.buildenv.2023.110388

Yang, Z., Li, N., Becerik-Gerber, B. and Orosz, M., 2014. A systematic approach to occupancy modeling in ambient sensor-rich buildings. Simulation, 90(8), pp.960-977.

Yu, W., Nakisa, B., Ali, E., Loke, S. W., Stevanovic, S., & Guo, Y. (2023). Sensor-based indoor air temperature prediction using deep ensemble machine learning: An Australian urban environment case study. *Urban Climate*, *51*, 101599. https://doi.org/10.1016/j.uclim.2023.101599

Zhang, R. and El-Gohary, N., 2023. Transformer-based approach for automated context-aware IFC-regulation semantic information alignment. Automation in Construction, 145, p.104540.

Zhang, W., Cui, G., Wang, Y., Zheng, C., & Zhu, Q. (2023). A human comfort prediction method for indoor personnel based on time-series analysis. *Building Simulation*, *16*(7), 1187–1201. https://doi.org/10.1007/s12273-023-1010-8

Zhao, M. H., & Yan, T. (2014). PMV Index Forecasting Algorithm Based on PSO-SVR. *Advanced Materials Research*, *912-914*, 1313–1317. https://doi.org/10.4028/www.scientific.net/amr.912-914.1313

Zhao, Y. and Watterston, J., 2021. The changes we need: Education post COVID-19. Journal of Educational Change, 22(1), pp.3-12.

Zhen, M., Yi, M., Luo, T., Wang, F., Yang, K., Ma, X., Cui, S., & Li, X. (2023). Application of a Fusion Model Based on Machine Learning in Visibility Prediction. *Remote Sensing*, *15*(5), 1450–1450. https://doi.org/10.3390/rs15051450

Zhou, J., Zhang, X., Xie, J., & Liu, J. (2023). Occupant's preferred indoor air speed in hot-humid climate and its influence on thermal comfort. *Building and Environment*, *229*, 109933. https://doi.org/10.1016/j.buildenv.2022.109933

Zou, H., Zhou, Y., Jiang, H., Chien, S.-C., Xie, L., & Spanos, C. J. (2018). WinLight: A WiFi-based occupancy-driven lighting control system for smart building. *Energy and Buildings*, *158*, 924–938. https://doi.org/10.1016/j.enbuild.2017.09.001

Division of Scholarly Integrity and
Research Compliance
Institutional Review Board
North End Center, Suite 4120 (MC 0497)
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-3732
irb@vt.edu
http://www.research.vt.edu/sirc/hrpp

MEMORANDUM

DATE: March 28, 2023

TO: Ruichuan Zhang, Antonio Lee

FROM: Virginia Tech Institutional Review Board (FWA00000572)

PROTOCOL TITLE: Adaptive Academic Buildings for Improving Comfort and Well-Being of College Students Using Artificial Intelligence of Things

IRB NUMBER: **23-205**

Effective March 28, 2023, the Virginia Tech Human Research Protection Program (HRPP) determined that this protocol meets the criteria for exemption from IRB review under 45 CFR 46.104(d) category (ies) 2(i),2(ii).

Ongoing IRB review and approval by this organization is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these activities impact the exempt determination, please submit an amendment to the HRPP for a determination.

This exempt determination does not apply to any collaborating institution(s). The Virginia Tech HRPP and IRB cannot provide an exemption that overrides the jurisdiction of a local IRB or other institutional mechanism for determining exemptions.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

https://secure.research.vt.edu/external/irb/responsibilities.htm

(Please review responsibilities before beginning your research.)

PROTOCOL INFORMATION:

Determined As: **Exempt, under 45 CFR 46.104(d) category(ies) 2(i),2(ii)**
Protocol Determination Date: **March 28, 2023**

ASSOCIATED FUNDING:

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.

SPECIAL INSTRUCTIONS:

Please note: The HRPP office has stopped stamping documents for Exempt protocols. It is your responsibility to maintain these documents and make current versions available on request.

Date*	OSP Number	Sponsor	Grant Comparison Conducted?

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.

Participant Consent Form

You are invited to participate in a research design to investigate *Artificial Intelligence of Things-based Interactive Educational Facilities for Improving Students' Comfort and Health*. This research aims to develop a better building automation system utilizing user-friendly Internet of Things (IoT) devices and a deep learning model that incorporates building occupants' direct and indirect feedback to provide a better indoor environment at academic buildings.

Participants will engage in a series of individual and group classroom activities, including self-study, lecture and discussion, and group discussion. During these activities, IoT devices will be deployed in the test location to collect Indoor Environmental Quality (IEQ) data and participants will be able to monitor current IEQ conditions. Participants will be asked to answer brief survey questions throughout the participation period. Participants will be asked to answer the questions every 5 minutes or whenever they feel discomfort in the built environment. The classroom will be monitored by video and thermal equipment with a microphone. The main goal of recording videos is to analyze occupants' behavior based on IEQ

There are no risks associated with participation in this study. Any collected data throughout the participation period will be stored securely in the cloud database and will be made available only to the research team in this study. Your participation in this study is voluntary. You can decline to participate or withdraw from the study at any time. If you have any questions, please feel free to contact Antonio Lee at minjae0624@vt.edu or Dr. Ruichuan Zhang at ruichuanz@vt.edu.

If you have any concerns related to this study, please contact the Human Research Protection Program office, North End Center, Suite 4120, Virginia Tech, 300 Turner Street NW, Blacksburg, Virginia, 24061, Phone: 540-231-3732, or at Email: irb@vt.edu. IRB protocol number: #23-205

Name and Signature:

Date:

By signing this, you are agreeing to participate in this experiment.

Recruitment email

Dear MLSoC students,

We would like to invite you to participate in data collection for ***Artificial Intelligence of Things-based Interactive Educational Facilities for Improving Students' Comfort and Health***. We would greatly appreciate your participation in this research!

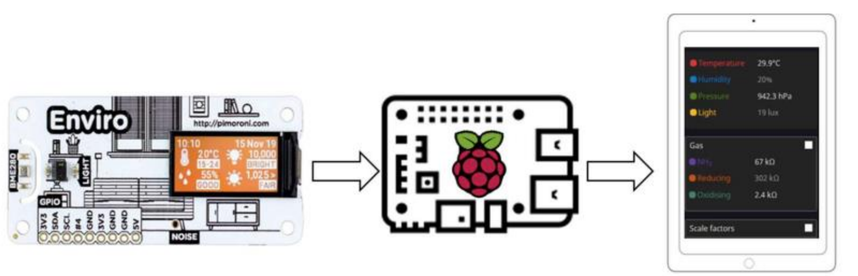
IRB protocol number: #23-205

Eligibility

- Must be 18 years old or older
- Must be a current Virginia Tech student

Descriptions

- Experiment duration: 35 minutes
- You will participate in a series of individual and group class activities and complete a short website-based survey to report your self-comfort and health level every 5 minutes and whenever you feel necessary.
- Internet of Things (IoT) devices will be deployed to monitor Indoor Environmental Quality (IEQ) data
- Test location will be monitored using video and thermal recording equipment



Enviro Plus Board Raspberry Pi 4 User Interface (UI)

The experiment will be conducted after Spring 2023 semester concludes. Free **DONUTS** and **DRINKS** will be provided at each session! If you are interested, please feel free to contact Antonio Lee at minjae0624@vt.edu to schedule the experiment.