Wiley
Indoor Air
Volume 2026, Article ID 5960599, 15 pages
https://doi.org/10.1155/ina/5960599

WILEY

Research Article

Application of Deep Neural Networks for Leakage Airflow Rate Estimation From Three-Dimensional Thermal Patterns

Diego Tamayo-Alonso, Irene Poza-Casado, and Alberto Meiss

Department of Architectural Construction, University of Valladolid, Valladolid, Spain

Correspondence should be addressed to Diego Tamayo-Alonso; diego.tamayo@uva.es

Received 21 August 2025; Revised 4 November 2025; Accepted 26 November 2025

Academic Editor: Praveen Barmavatu

The employment of deep convolutional neural networks (CNNs) signifies a substantial progression in the domain of image analysis. The application of this method is particularly suitable when the image set represents a spatial structure and predictive analysis can only be performed using Gaussian processes, which are computationally complex. The uncontrolled airflow of air into buildings, known as infiltration, poses a significant challenge in terms of characterisation and quantification. The irregular contours of gaps and cracks through the enclosure create a virtually endless variety of cases, making a generalizable scientific interpretation that can be applied to existing buildings very difficult. This circumstance is always clearly manifested by an irregular, three-dimensional incoming airflow. This study presents an innovative methodology for estimating airflow rates based on three-dimensional thermal patterns captured through infrared thermography. The experimental setup employs a 3D-printed matrix of spheres, facilitating the characterisation of the spatial temperature distribution within the airflow. The resulting thermal images are processed using a CNNs, which integrates the spatial information contained in the thermograms with a scalar input representing the inlet air temperature. The model's performance was assessed under a range of conditions, including reduced image resolutions, varying experimental configurations (involving different flow apertures) and a comparison between full thermographic inputs and thermal difference-based features. The results indicate that the model can accurately infer airflow rates within the same aperture (medium absolute error [MAE] < 2%). While generalisation to new apertures presents a greater challenge, the experiments demonstrate that a sufficiently diverse training dataset can enhance the model's predictive capacity for configurations not included in the training phase. These findings underscore the potential of deep learning as a nonintrusive and efficient tool for estimating airflow in systems where conventional measurement techniques are either difficult to apply or impractical.

Keywords: building airtightness; deep convolutional neural network; infiltration; pressurisation test; thermography

1. Introduction

Buildings are responsible for a substantial portion of global energy consumption, with heating, ventilation and air conditioning (HVAC) systems accounting for nearly half of this usage [1]. Leakages, the unintentional movement of air through gaps and cracks in the building envelope, significantly impact both thermal comfort, ventilation performance and energy efficiency [2]. In this regard, it is estimated to contribute between 30% and 50% of HVAC loads [3].

Minimising air leakage is a key focus in modern building energy codes and standards to reduce overall energy demand [4]. Beyond energy loss, uncontrolled airflow can also

compromise material durability and promote mould growth due to moisture transport, as well as affect indoor air quality with these pollutants [2]. Improving airtightness is therefore a critical aspect of enhancing building performance and reducing greenhouse gas emissions, aligning with targets set by various regulations and directives aimed at achieving more energy-efficient and net-/nearly-zero-energy buildings (nZEBs) [5–8].

Conventional methods are employed to assess building airtightness and to identify air leakage paths. Pressurisation tests, such as the well-known and widely used Blower-Door test, are effective in quantifying the overall air leakage rate through the building envelope [9, 10]. However, one of their main limitations is the inability to pinpoint the exact shapes and flow rates of the leaks [4]. Infrared thermography (IRT) offers a noninvasive approach to visualise surface temperatures and detect thermal anomalies associated with energy gains or losses [7, 11]. IRT can effectively identify leakage points by detecting temperature differences on building surfaces, often highlighting areas around windows, doors and service penetrations, provided there is a temperature differential between the interior and exterior. Nevertheless, these methods offer only a qualitative, rather than quantitative, assessment of the airflow rate associated with infiltration. Quantitative IRT methods aim to correlate surface temperature patterns with leakage characteristics; however, they often necessitate complex heat transfer models and numerous parameters and can be complicated by thermal bridging effects. Early applications of IRT for air leakage detection were not always accurate, requiring precise parameters and environmental conditions [12]. Furthermore, traditional methods may involve extensive data collection.

Recent years have seen advances in both the characterisation of air infiltration and the application of machine learning techniques to thermographic data. Novel methodologies that approach the characterisation of infiltrations from a three-dimensional perspective using IRT have emerged [9, 10]. By capturing temperature data from measurement points and transposing them into a digital model, it is possible to analyse these points in three dimensions, allowing for the extraction of planes to understand airflow behaviour. This approach characterises infiltration using a single thermographic image, providing a "steady-stage" airflow image and significantly reducing data acquisition and processing time compared to previous methods. This characterisation method is one of the baselines of this research.

Convolutional neural networks (CNNs) have emerged as a central tool in image-based diagnostics, renowned for their capacity to learn hierarchical representations from raw data without the need for manual feature engineering [13]. Within the context of IRT, CNNs have shown remarkable performance in thermal image classification [14, 15], anomaly segmentation [12, 16], concealed object detection [17] and internal defect localisation in materials [18]. These capabilities render them particularly promising for building diagnostics, especially in the identification of air leakages, thermal bridges and envelope anomaly factors directly linked to energy losses and occupant discomfort [15, 19].

CNN-based models have proven effective across a range of real-world scenarios. They have been used to detect structural damage postdisaster [20], identify subsurface defects through dual-stream processing of thermal and visible imagery [21] and automate flashpoint recognition in chemical testing using minor thermal signatures [22]. Furthermore, hybrid approaches—such as CNNs coupled with support vector machines (SVMs) and graph-based localisation—have demonstrated reliable leak detection in complex fluid distribution networks [23]. Advanced architectures employing multiscale feature extraction and attention mechanisms have also improved fault detection in HVAC systems where datasets are inherently imbalanced [13].

Recent studies have also examined CNNs under constrained data regimes and portable diagnostic platforms. Notably, the DeepRadiation system demonstrated that energy performance prediction is achievable using only 36 annotated thermal images [19], and passive thermographic learning approaches have been validated through transfer learning in varying operational conditions [14]. These advances, alongside foundational work on deep learning theory [11], highlight CNNs' growing integration in real-time, edge-level applications.

Within the specific domain of air infiltration in buildings, accurate quantification of airflow rates remains a challenge. Although transient IRT has proven that it is possible to identify air leaks without the need for sustained pressurisation [2], no effective IRT quantification method has been developed, and few studies have addressed the use of CNNs to infer volumetric airflow from thermal patterns. Most existing work focuses on the localisation of thermal anomalies [12, 16], not on predictive estimation of airflow magnitude, and generalisation to unseen geometries or openings remains largely unexplored.

Preliminary efforts, such as those by Gertsvolf et al. [12], have prepared data pipelines for CNN-based thermal diagnosis of building envelopes, while Deng et al. [18] validated the applicability of CNNs to internal defect recognition even under low thermal contrast. These findings suggest that with appropriate training protocols, CNNs could reliably infer air leakage rates from thermograms alone. This study contributes to the evolving field by proposing a CNN-based approach trained on passive thermal imagery for the inference of airflow rates through uncharacterised building openings. The method seeks to offer a scalable, accurate and nonintrusive diagnostic framework for real-time energy loss detection in buildings.

Despite progress in generating three-dimensional thermal data for airflow characterisation and applying deep learning to thermographic analysis, a gap remains in directly leveraging 3D thermal patterns with deep neural networks for the specific task of quantifying infiltration airflow rates.

While 3D characterisation methods can produce detailed thermal datasets [10] and neural networks can infer airflow from surface temperatures [24], a method that explicitly takes the volumetric or multiplanar thermal information derived from the 3D characterisation techniques as direct input into a deep learning model for nonintrusive airflow estimation has not been widely demonstrated in the literature.

This paper addresses this knowledge gap by proposing and evaluating a novel methodology for estimating infiltration airflow rates. Our objective is to develop a method based on CNNs trained on 3D thermal patterns, obtained from advanced IRT characterisation techniques, to infer the airflow rate through building envelope discontinuities directly. This approach is aimed at providing a more accurate and nonintrusive means of quantifying air infiltration at specific leakage points, combining the spatial richness of 3D thermal data with the powerful pattern recognition capabilities of deep learning.

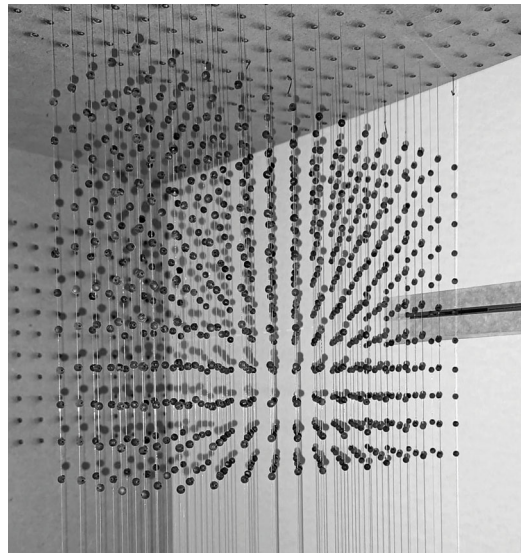

FIGURE 1: Three-dimensional matrix of spheres.

2. Materials and Methods

The experimental setup, described in detail in a previous study by the authors [10], was used in this study to capture three-dimensional thermal patterns using IRT. Unlike the previous work, this investigation employed thermal video sequences instead of individual frames, allowing for a sufficient data volume to train the model.

The main equipment consisted of a FLIR E75 thermal camera for the detection of low-magnitude temperature variations (temperature accuracy ± 2%) with a spatial resolution of 320 × 240 pixels and calibrated optics. The airflow passed through a three-dimensional matrix of spheres, manufactured using resin-based 3D printing. The periodic matrix of the spheres induced local thermal differences without causing interference in the path of the flow field (Figure 1).

Five different inlet configurations (A1–A5) were analysed, each with a different flow area and geometry. Openings A1–A3 had regular dimensions, the same height and different widths, while A4 and A5 had irregular form but the same area as the others (Figure 2). Specifically, the A5 opening has an irregular shape with no parallel sides, resembling more of a crack in the wall. For each configuration, two independent data collections were conducted and grouped under the labels 01–10 (Table 1).

The airflow rate (Qm) was measured using an IST FS5 sensor with an accuracy of ≤ 3%. The inlet air temperature was measured using a TSIC 501F sensor, with an accuracy of ±0.1 K. Environmental conditions were maintained at a constant level and recorded for each test. The primary objective in terms of environmental conditions is to maintain a thermal differential of 5°C as constant as possible. This is achieved thanks to the management of the chamber's climate control system. The temperature of the inlet air and the temperature in the chamber are constantly measured in the measurement area. The logger continually calculates the temperature differential between the test chamber air and the inlet air. At the same time, it controls the ignition of the underfloor heating, thus maintaining a stable temperature differential of 5°C. The inlet air is outdoor air, as it would be in a real house, and is somewhat tempered as it passes through the ducts, similar to what would happen when passing through an architectural enclosure. In this way, the inlet air temperature changes slightly between tests, as the outside temperature fluctuates. These slightly changing conditions enable us to obtain a broader range of temperatures. The humidity conditions remain unchanged, corresponding to the actual humidity of the outside air.

2.1. Image Acquisition and Processing. Thermal images of the airflow were recorded in video format at acquisition frequencies of 10 and 30 fps (Table 1). Each thermogram extracted from the videos had an original resolution of 320 × 240 pixels. To ensure spatial consistency, all images were cropped to a uniform size of 200 × 200 pixels by removing areas outside the sphere matrix.

To evaluate the influence of resolution on model performance, a series of tests was conducted in which the cropped images were resampled using bicubic interpolation to resolutions of 120 × 120, 80 × 80, 40 × 40 and 20 × 20 pixels (Figure 3). For each of these resolutions, the model was trained from scratch using the corresponding thermographic series.

In parallel, given the notable improvement in flow visualisation achieved by computing thermal differences between preexcitation and postexcitation frames—an approach validated in previous studies—a comparative analysis was conducted. For this purpose, thermal difference images were computed by subtracting the thermogram

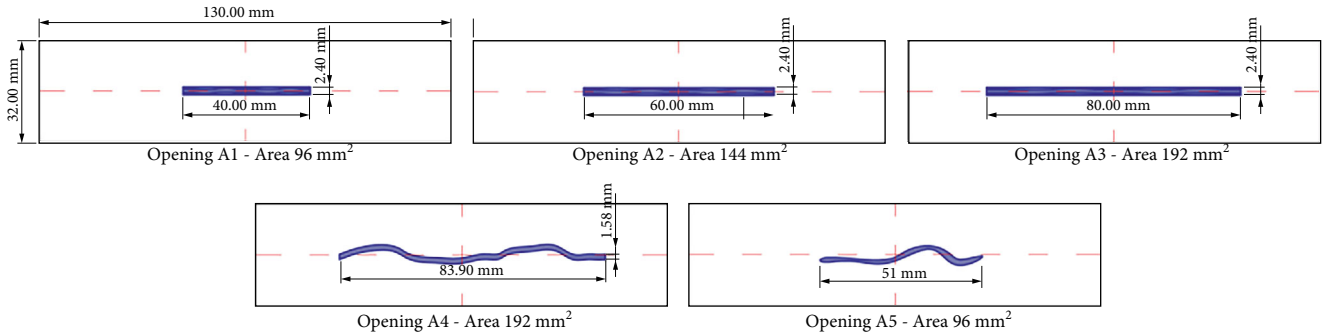

FIGURE 2: Opening configurations.

TABLE 1: Data collection timing and ambient conditions.

Data collection	Opening	Freq. (fps)	Duration	No. of thermography	μ°C chamber	μ°C flow	%HR
01	A1	10	0:11:00	6600	24.84	19.68	45
02	A1	30	0:10:45	19,350	25.05	19.76	44
03	A2	10	0:11:21	6810	24.45	19.87	45
04	A2	30	0:11:00	19,800	24.89	20.10	45
05	A3	10	0:11:40	7000	24.78	20.15	44
06	A3	30	0:11:30	20,700	25.15	20.07	44
07	A4	10	0:11:00	6600	25.28	20.18	44
08	A4	30	0:11:00	19,800	24.99	20.20	45
09	A5	10	0:11:00	6600	25.09	20.15	45
10	A5	30	0:10:50	19,500	25.45	20.24	45

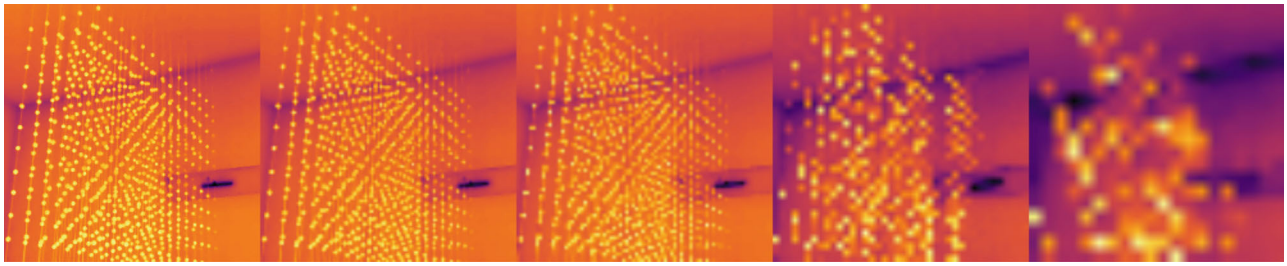

FIGURE 3: Comparison of IRT images at different resolutions.

recorded before depressurisation from each subsequent frame (Figure 4).

All inputs were normalised to the range [0, 1] before training to eliminate the effect of absolute temperature values and focus on relative thermal patterns. Each input was labelled with its corresponding experimentally measured airflow rate (Qm).

2.2. Dataset Preparation. The thermographic images were organised according to the experimental collection (01–10) and their corresponding opening configuration (A1–A5). Each image was labelled with its experimentally measured airflow rate (Qm), thus forming a supervised dataset.

Three main data partitioning strategies were defined:

- Intratest (IT-XX): The model was trained and validated using data from the same collection.
- Intraopening (AX-XX): The model was trained on one collection and validated on the other collection corresponding to the same aperture.

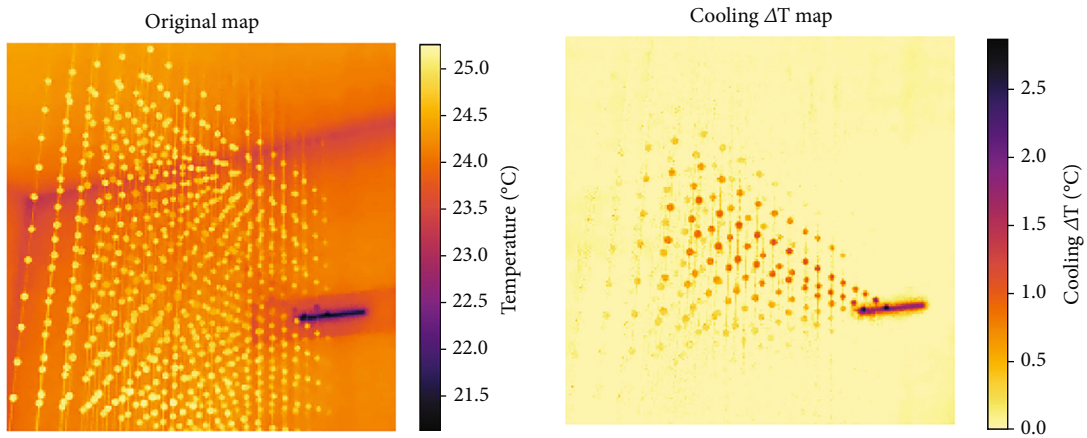

FIGURE 4: Comparison of the original thermography with the differential temperatures.

TABLE 2: Test configurations.

Intratest								
No. of test	**Training data collection (80%)**	**No. of thermography**	**Inferred data collection (10%)**	**No. of thermography**	**MAE**	**MSE**	**MAE (%)**	
IT-01	01	5280	01	660	0.014	0.003	0.2	
IT-02	02	15,480	02	1935	0.030	0.001	0.5	
IT-03	03	5448	03	681	0.054	0.005	0.9	
IT-04	04	15,840	04	1980	0.044	0.003	0.7	
IT-05	05	5600	05	700	0.078	0.011	1.3	
IT-06	06	16,560	06	2070	0.049	0.004	0.8	
IT-07	07	5280	07	660	0.043	0.004	0.7	
IT-08	08	15,840	08	1980	0.030	0.002	0.5	
IT-09	09	5280	09	660	0.085	0.012	1.4	
IT-10	10	15,600	10	1950	0.045	0.003	0.7	
Intraopening								
No. of test	**Training data collection (80%)**	**No. of thermography**	**Inferred data collection (100%)**	**No. of thermography**	**MAE**	**MSE**	**MAE (%)**	**Comments**
A2-01	03	5448	04	19,800	0.263	0.154	4.4	
A2-02	03	5448	04	19,800	1.356	2.682	22.6	Resizing 20 × 20
A2-03	03	5448	04	19,800	0.737	0.859	12.3	Resizing 40 × 40
A2-04	03	5448	04	19,800	0.724	0.801	12.1	Resizing 80 × 80
A2-05	03	5448	04	19,800	0.715	0.683	11.9	Resizing 120 × 120
A2-06	03	5448	04	19,800	0.745	1.113	12.4	Temperature difference
Interopening								
No. of test	**Training data collection (60%)**	**No. of thermography**	**Inferred data collection (100%)**	**No. of thermography**	**MAE**	**MSE**	**MAE (%)**	
IO-01	01 and 05	8160	03	6810	0.578	0.511	9.6	
IO-02	01, 03 and 05	12,246	03	6810	0.048	0.004	0.8	
IO-03	01, 03, 05 and 07	16,206	09	6600	0.652	0.617	10.9	
IO-04	01, 03, 05, 07 and 09	20,166	09	6600	0.098	0.013	1.6	

TABLE 3: Test to estimate the influence of thermal resolution.

No. of test	Training data collection (80%)	Inferred data collection (100%)	MAE	MSE	MAE (%)	Comments
A2-01	03	04	0.263	0.154	4.4	
A2-02	03	04	1.356	2.682	22.6	Resizing 20 × 20
A2-03	03	04	0.737	0.859	12.3	Resizing 40 × 40
A2-04	03	04	0.724	0.801	12.1	Resizing 80 × 80
A2-05	03	04	0.715	0.683	11.9	Resizing 120 × 120

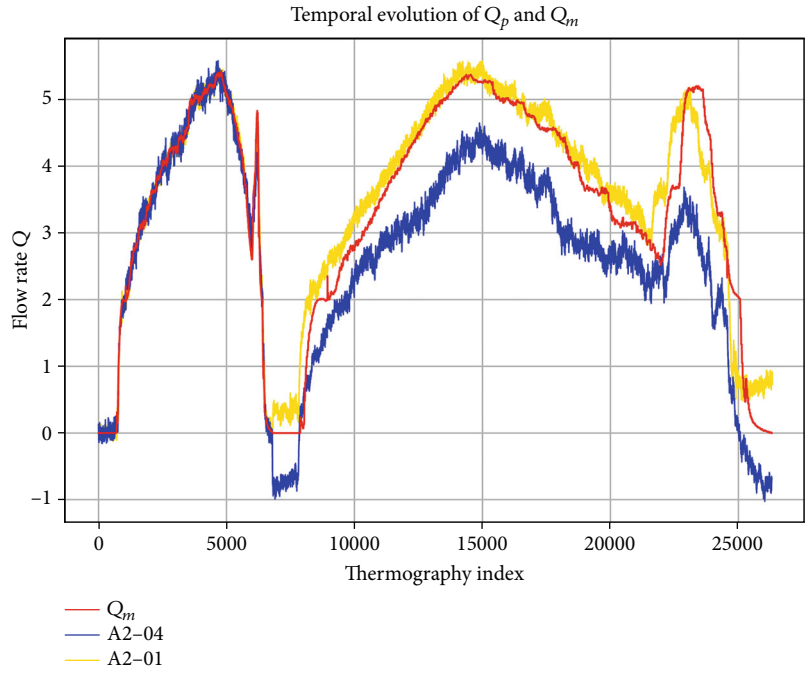

FIGURE 5: Comparison of *Qp* between full (A2-01) and reduced 80 × 80 px (A2-04) thermograms.

TABLE 4: Comparison test of complete thermography and thermal differences.

No. of test	Training data collection (80%)	Inferred data collection (100%)	MAE	MSE	MAE (%)	Comments
A2-01	03	04	0.263	0.154	4.4	
A2-06	03	04	0.745	1.113	12.4	Difference temperature

- Interopening (IO-XX): The model was trained using data from multiple apertures and validated on an aperture not included in training.

These combinations (Table 2) made it possible to evaluate the model performance under low variability conditions (within the same collection), moderate generalisation (within the same aperture) and high generalisation complexity (across different apertures).

In the process of creating the dataset for each test, the data is shuffled before being divided into training, validation and test sets. This ensures that all flows are represented in each dataset.

For each configuration, the metrics mean absolute error (MAE) and mean squared error (MSE) were calculated,

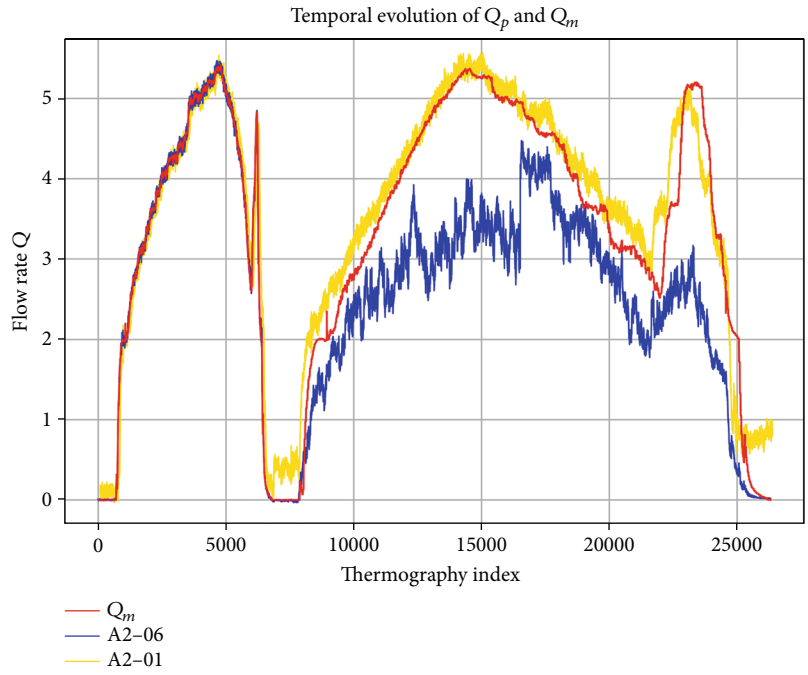

Figure 6: Comparison of *Qp* between full (A2-01) and thermal differences (A2-06) thermograms.

expressed in absolute units (cubic meters per hour) and as a percentage of the operating range (0–6 m^3/h).

2.3. Neural Network Model. In a previous study by Royuela-del-Val [24], four neural network architectures were evaluated for the inference of airflow rate during an infiltration process. The study concluded that the most balanced option in terms of accuracy and generalisation capability was a deep CNN with a multimodal architecture. This model incorporates two inputs: a thermal image (single-channel matrix) and a scalar value corresponding to the inlet air temperature. In the present work, this previously validated model is adopted.

The convolutional branch of the model consists of two consecutive Conv2D layers (20 filters, 5×5 and 3×3, respectively), both using ReLU activation and "same" padding. The convolutional output is then flattened using a flatten layer.

The scalar value is fed through a second independent input. Both inputs are concatenated and passed to a dense layer with 100 neurons and ReLU activation. The final output of the model is a single linear neuron that predicts the airflow rate (*Qp*).

The model was compiled using the Adam optimiser [25] (learning rate = 1^{e-3}, clipnorm = 1.0), with MSE as the loss function and MAE as the evaluation metric. All graphs showing the progress of the training sessions are located in Appendix 1.

3. Results

3.1. Influence of IR Resolution. To assess the impact of thermal image resolution on the model performance, independent neural networks were trained using resampled subsets at resolutions of 120×120, 80×80, 40×40 and 20×20 pixels. In all cases, training and inference were performed on data from the same aperture (Table 3).

The network was trained with the data corresponding to Data Collection 3 using 80% of the data for training, 10% for validation and 10% for testing. Posttraining test results showed that the neural network remained stable in terms of accuracy up to a resolution of 80×80, with a loss of only 2% accuracy. Below this resolution, the accuracy worsens significantly.

Subsequently, each trained network was used to infer the second dataset of the same aperture, each resampled to match the training resolution of the network. In these cases, a marked decrease in accuracy was observed as the resolution decreased, as illustrated in Figure 5.

The results indicate that lowering the resolution degrades the model's accuracy and significantly reduces its generalizability, as the MAE increases from 4.4% using the full resolution to 11.9% at 120×120 pixels. As we continue to reduce the resolution, accuracy and generalisation continue to deteriorate, reaching an MAE of 22.6% in the case of a resolution of 20×20 pixels.

3.2. Comparison Between Full Thermographic Images and Thermal Differences. The test was performed using Aperture 2 (Data Collections 03 and 04). Two models were trained: one using full thermographic images and another using thermal differences. Both models were trained on data from Collection 03 and evaluated on Collection 04, corresponding to the same aperture (Table 4).

The results favoured the use of full thermographic images (view in Figure 6), with significantly lower errors

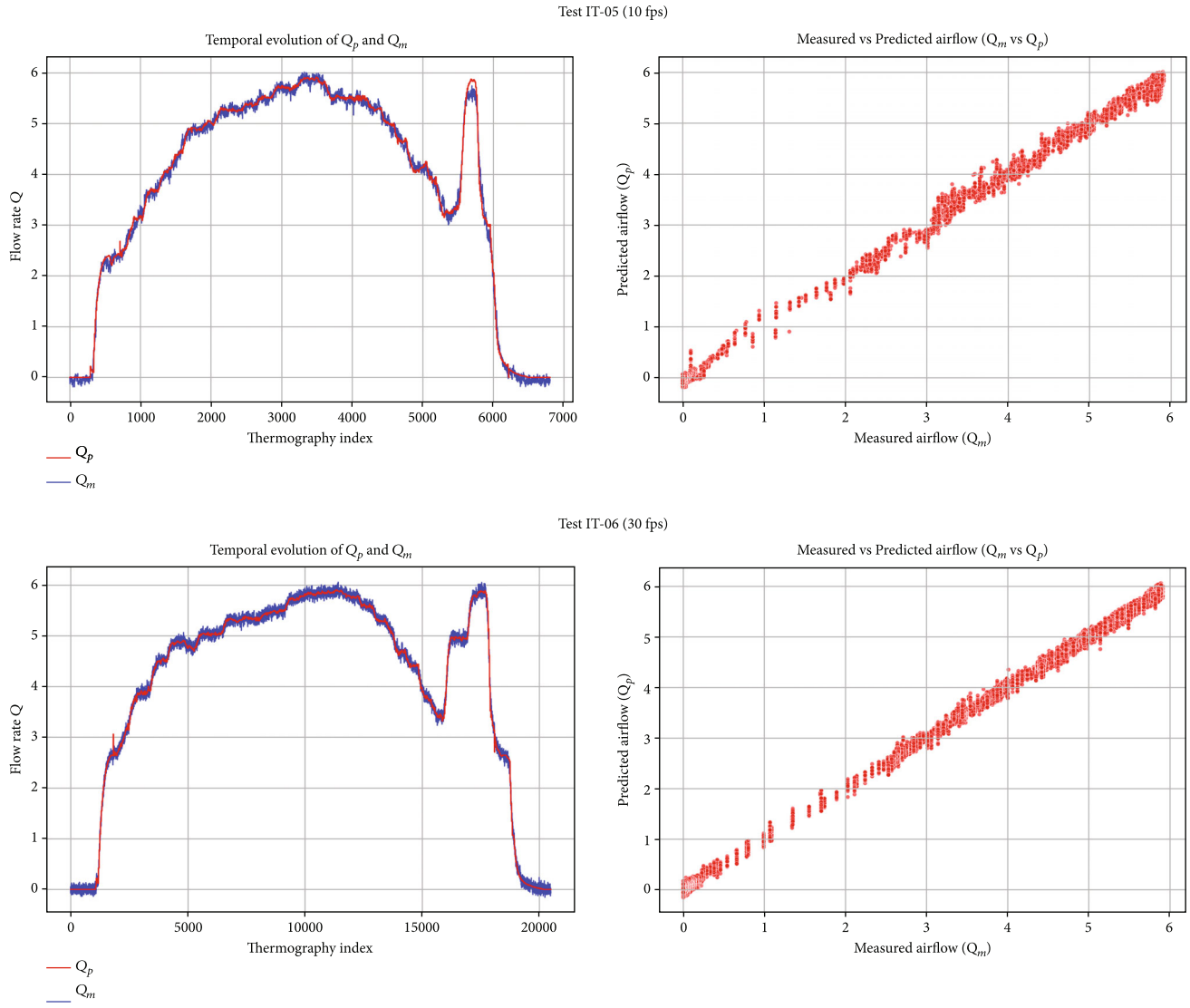

Figure 7: Comparison of *Qp* versus *Qm* between data collection at 10 and 30 fps.

across all evaluated metrics. The model based on thermal differences showed a reduced generalisation capacity when switching the dataset, suggesting a loss of relevant spatial information.

This comparison supports the use of full images as input to the model, especially when aiming to extend its applicability to new conditions within the same geometric configuration.

3.3. Intratest Inference. This section evaluates the model's accuracy when it is both trained and tested using data from the same experimental collection, without changing the aperture configuration or flow conditions. This scenario represents the lowest possible variability and serves as a baseline reference for model performance. Every data collection was tested and corresponds with tests IT-01 to IT-10.

The errors obtained in this configuration were consistently low, with MAE values below 1% of the full range in most cases. This consistency indicates that the model is capable of learning and reproducing the thermal patterns of a specific dataset with high precision.

Additionally, as shown in Figure 7, there is no difference in accuracy and fit between the 10 and 30 fps tests. Therefore, in other test phases where more simultaneous data frames will be collected, the 10 fps data frames will be used, thereby reducing the size of the data frame.

This result is anticipated, as the model does not encounter any variations in geometry or flow distribution. Nonetheless, it provides a useful indicator of the maximum achievable performance under controlled conditions.

3.4. Intra-Aperture Inference. This section analyses the model's performance when trained on one dataset and evaluated

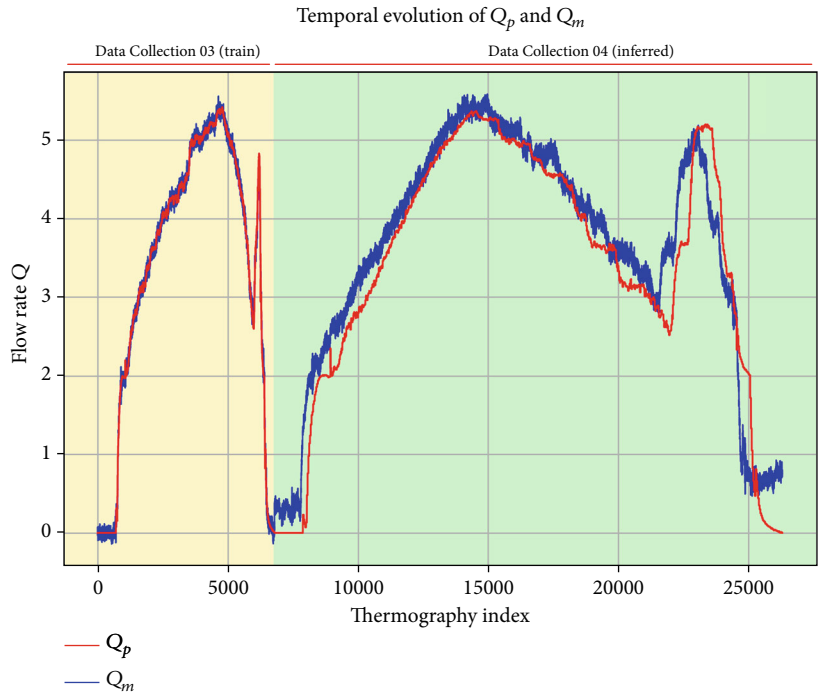

FIGURE 8: Comparison of *Qp* between full and thermal differences of thermograms.

on another dataset corresponding to the same aperture. Although the system's geometry remains unchanged, slight variations in the flow may occur due to boundary or experimental conditions.

The results showed a slight degradation in performance compared to the intratest cases, but generally remained within acceptable margins.

Figure 8 shows the A2-01 test data, which was trained using data from Data Acquisition Number 3 and inferred from Data Acquisition Number 4, both corresponding to the A2 aperture.

The error values remained below 2%, indicating the model's good ability to generalise across datasets corresponding to the same aperture.

This behaviour demonstrates that the model is not limited to memorising training data, but can recognise thermal patterns that represent a specific aperture and apply them to new cases within the same geometrical configuration.

3.5. Interaperture Inference. Finally, the model's ability to generalise across different apertures was evaluated. To do so, we trained with several datasets corresponding to different apertures and tested with another aperture that was not included in the training set (see Table 5).

Two test sets were created, one with rectangular apertures, training the model with the smallest and largest apertures (A1 and A3, respectively) and then inferring on the median aperture (A2) (test IO-01). The network was then trained with all three apertures to observe the differences (IO-02). Subsequently, another model was trained with the apertures A1–A4. This test (IO-03) includes the irregular aperture A4 and another irregular aperture A5, which was inferred. As in the previous case, the model was retrained with the data of this aperture (A5) to test the potential for learning improvement (IO-04).

TABLE 5: Configuration tests interopening.

No. of test	Training data collection (60%)	Inferred data collection (100%)	MAE	MSE	MAE (%)
IO-01	01 and 05	03	0.578	0.511	9.6
IO-02	01, 03 and 05	03	0.048	0.004	0.8
IO-03	01, 03, 05 and 07	09	0.652	0.617	10.9
IO-04	01, 03, 05, 07 and 09	09	0.098	0.013	1.6

The results indicated a considerable increase in errors compared to the intraopening cases. The MAE surpassed 9% (Figures 9 and 10), highlighting a significant loss of predictive accuracy when the model was exposed to unseen geometries. Nonetheless, a gradual improvement was observed as the diversity of the training set increased, suggesting that the model improves from greater variability in thermal patterns (Figures 11 and 12).

These results emphasise the model's sensitivity to geometric changes while also demonstrating its potential for generalisation when provided with a sufficiently representative dataset. This is particularly crucial when the model is intended for application in real-world situations.

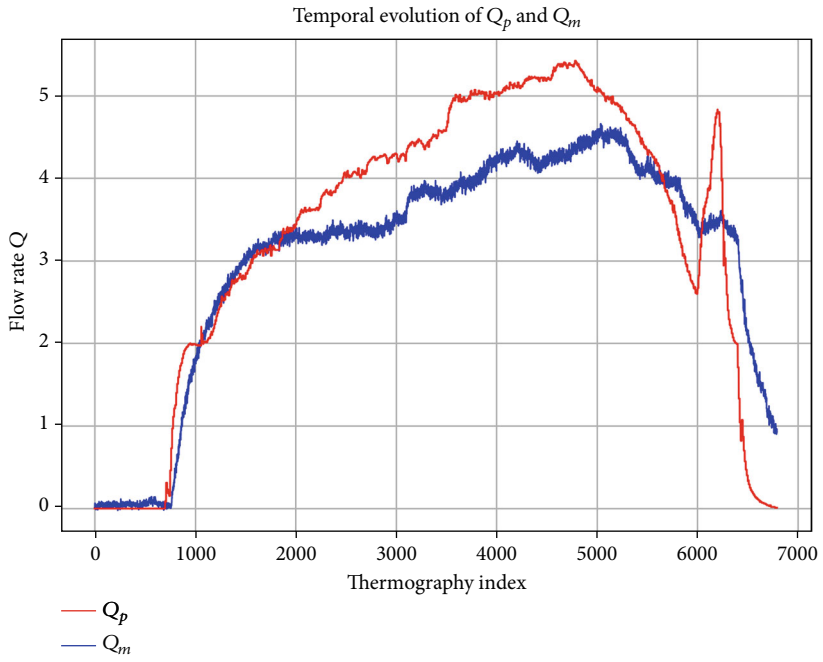

FIGURE 9: *Qp* versus *Qm* in test IO-01 for Data Collection 03.

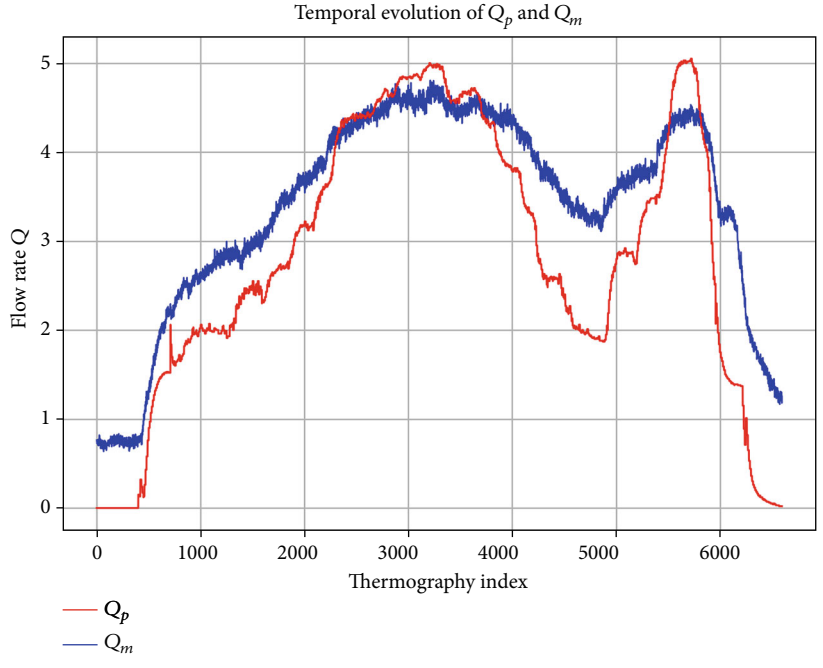

FIGURE 10: *Qp* versus *Qm* in test IO-03 for Data Collection 09.

4. Discussion

The results obtained in this study provide valuable insights into the capabilities, limitations and future potential of deep learning techniques for nonintrusive airflow estimation based on thermographic data. The proposed approach demonstrates that CNNs can effectively learn complex spatial features embedded in thermal patterns, enabling the quantification of airflow rate with high accuracy, particularly under consistent geometric conditions.

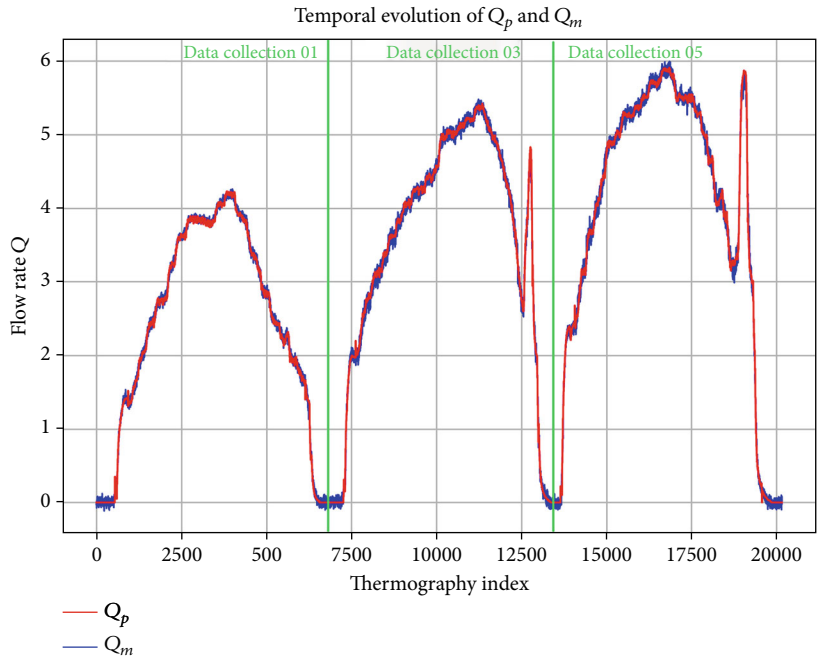

FIGURE 11: *Qp* versus *Qm* in test IO-02 for Data Collection 01–03.

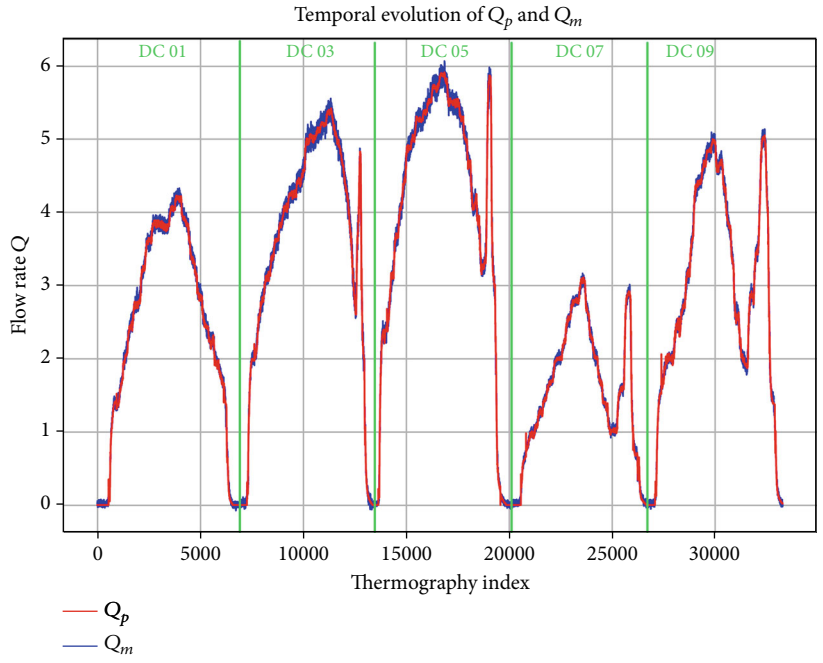

FIGURE 12: *Qp* versus *Qm* in test IO-04 for Data Collection 01–09.

In intratest and intra-aperture scenarios, where the aperture geometry and boundary conditions remain constant, the model consistently demonstrated low error rates, typically below 2% of the full range. This underscores the network's strong capacity for pattern recognition and validates the feasibility of employing thermography combined with

CNNs as a rapid and efficient alternative to traditional measurement methods such as anemometry or hot-wire probes, which may be invasive or unsuitable in specific applications.

A key finding has been the significance of the spatial resolution of the thermal images. Reduced resolution has been shown to diminish the generalisability between different tests. This occurs due to the reduction of thermal gradients and the loss of spatial characteristics of the flow.

Another relevant contribution is the comparison between the use of full thermographic images and differential thermograms (i.e., temperature difference images) as model inputs. While thermal difference images can enhance visual contrast and facilitate qualitative interpretation of flow behaviour, the experiments demonstrated that full images retain more robust spatial information, enabling better generalisation when switching between datasets. The differential preprocessing may inadvertently eliminate subtle but important gradients and local variations that are informative for the network. Therefore, for machine learning applications where predictive accuracy is paramount, complete thermal data is preferable, especially in scenarios with low signal-to-noise ratios or fine-scale temperature structures.

The study also reveals important insights into the generalisability of the model when exposed to previously unknown aperture geometries. Inference in untrained apertures yielded significantly higher error rates. This finding highlights a well-known challenge in supervised learning: the tendency of deep models to perform worse when extrapolating beyond their training domain. However, a progressive performance improvement was observed as more varied training data was incorporated. This suggests that increasing the variability of thermal patterns, either by including additional aperture types or by simulating them, may be an effective strategy to enhance model generalisability.

In this context, transfer learning techniques could be investigated in future work. By fitting a pretrained CNN to a small dataset of a new aperture configuration, it may be possible to adapt the model to new scenarios with minimal additional training. Similarly, domain adaptation strategies could assist in bridging the gap between synthetic data (e.g., from CFD simulations) and real-world measurements, providing a scalable path to wider deployment.

Beyond the technical evaluation, this work opens a promising avenue of research for integrating deep learning models into building diagnostic workflows. For instance, the methodology could be applied in permeability testing by swiftly identifying the most significant leakage paths. Such systems could be particularly valuable during commissioning phases or building energy audits, where time and intrusion represent critical constraints.

However, several limitations remain to be addressed. The current experimental setup is highly controlled, featuring stable environmental conditions. In contrast, real-world environments often present additional challenges, including a diversity of materials and dynamic boundary conditions, among others. These factors can impact the model's performance and must be considered before it is implemented. In any case, the test setup will be determined by the conditions needed to maintain a constant pressure difference and a minimal temperature gradient, which will be dependent on the equipment used.

5. Conclusions

This study demonstrates the feasibility and effectiveness of using deep CNNs to estimate leakage airflow rates from three-dimensional thermal patterns. By combining thermographic input with a scalar inlet temperature in a multimodal architecture, the model achieved high predictive accuracy in intratest and intra-aperture scenarios, with MAEs typically below 2%. These results validate deep learning as a powerful tool for analysing thermographic data and extracting quantitative flow information under controlled conditions.

Among the most significant findings, it was confirmed that reducing the resolution of thermal images results in a decline in the model's generalisation capability. Similarly, full thermographic images outperformed temperature difference images, likely due to their enhanced spatial richness and the preservation of subtle gradients that are crucial for accurate inference.

The results obtained from the interaperture tests have also revealed the challenges the model faces when dealing with previously unseen geometric aperture configurations. Although a significant increase in prediction error was observed in these cases, a gradual improvement became evident as the training dataset incorporated greater diversity. This indicates that, with a sufficiently representative dataset, the model's predictive capacity can be extended to more general scenarios. This aspect becomes particularly relevant when aiming for real-world implementation, where precise reproduction of aperture conditions is not always possible.

In this context, the proposed methodology provides a solid foundation for an alternative, nonintrusive approach to estimating airflow resulting from infiltration. Compared to conventional techniques, this deep learning-based strategy enhances automation and accelerates analysis, potentially contributing to the verification and characterisation of the building envelope. However, for practical deployment, further research is required to assess the model's robustness under dynamic and uncontrolled environmental conditions.

Future research will focus on extending the proposed methodology by investigating the use of pretrained CNNs and fine-tuning strategies. It is expected that these approaches will exploit the representational capabilities of large-scale visual models to improve the accuracy and generalisation of air infiltration rate inference from thermographic imagery. Additionally, future work will explore image transformation and augmentation techniques to enhance model robustness and ensure consistent performance across various building types and environmental conditions.

Appendix 1

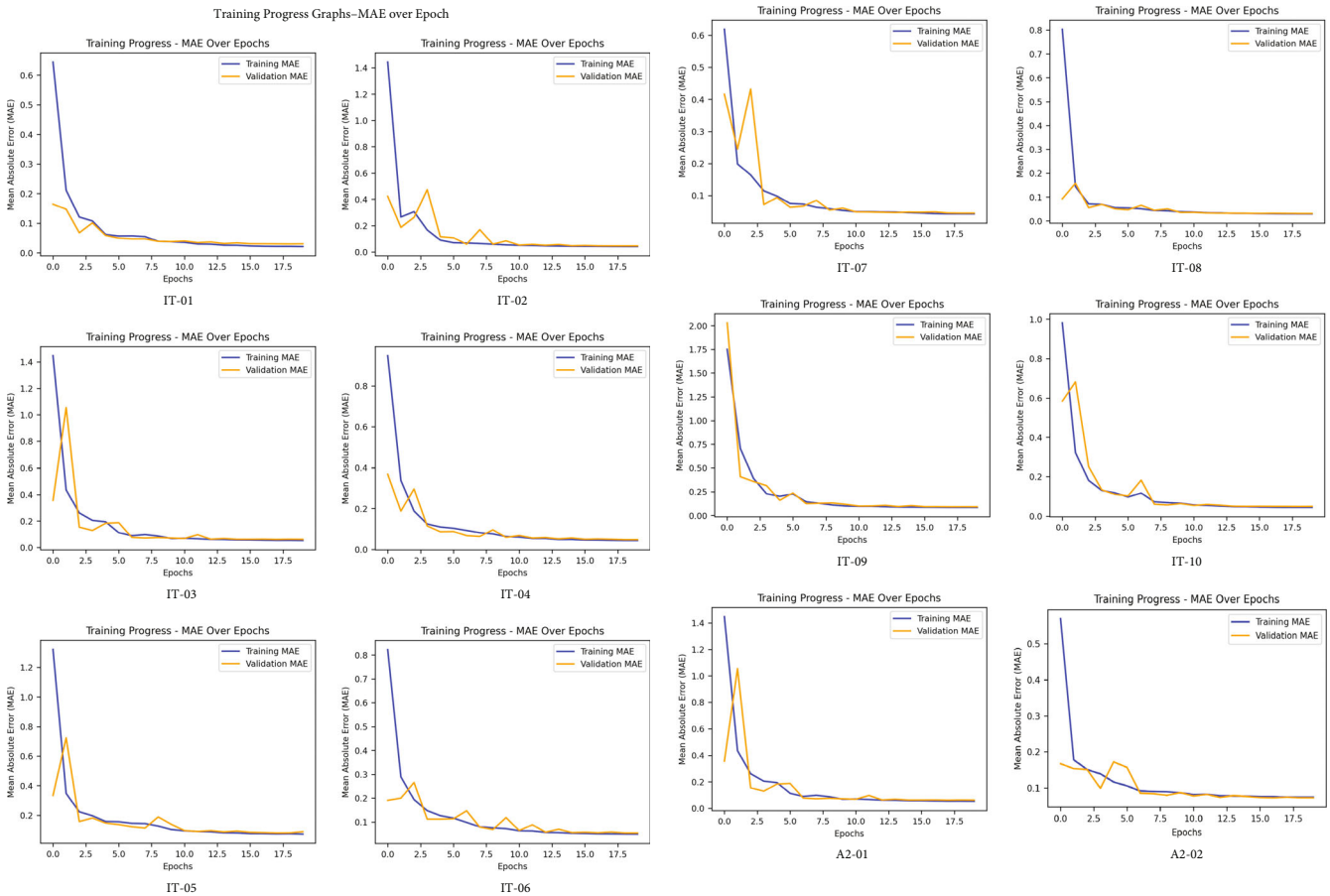

FIGURE A1: Continued.

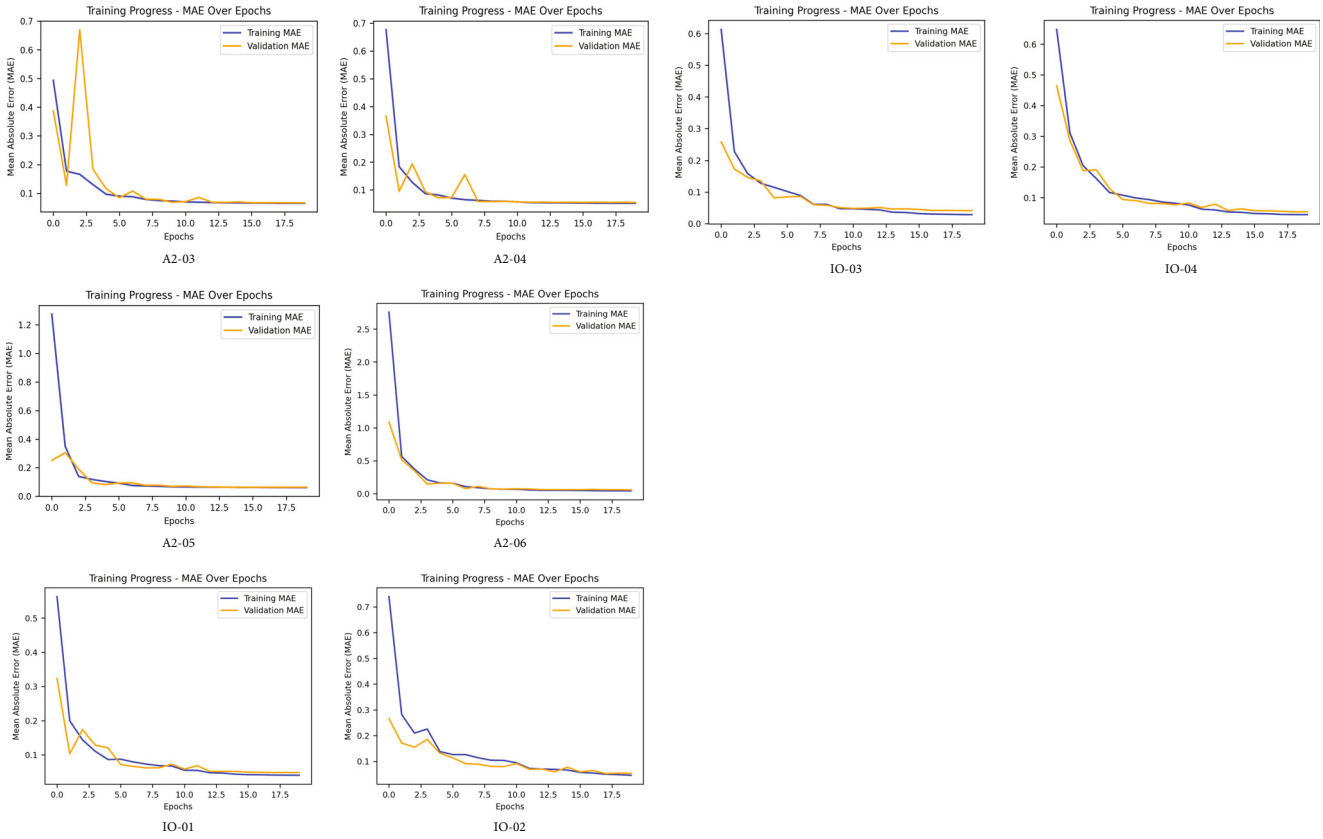

FIGURE A1: Training progress graphs—MAE over epoch.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Author Contributions

Diego Tamayo-Alonso: conceptualization, data curation, formal analysis, investigation, methodology, software, writing – original draft. **Alberto Meiss:** conceptualization, funding acquisition, project administration, resources, supervision, writing – review and editing. **Irene Poza-Casado:** conceptualization, formal analysis, investigation, resources, writing – review and editing.

Funding

This study was funded by the Ministerio de Ciencia e Innovación (10.13039/100014440) (PID2022-142104OB-I00).

References

[1] A. Kirimtat and O. Krejcar, "A Review of Infrared Thermography for the Investigation of Building Envelopes: Advances and Prospects," *Energy and Buildings* 176 (2018): 390–406, https://doi.org/10.1016/j.enbuild.2018.07.052.

[2] T. Feng, Z. Shen, S. S. Shrestha, and D. E. Hun, "A Novel Transient Infrared Imaging Method for Non-Intrusive, Low-Cost, Fast, and Accurate Air Leakage Detection in Building Envelopes," *Journal of Building Engineering* 91 (2024): 109699, https://doi.org/10.1016/j.jobe.2024.109699.

[3] J. Feijó-Muñoz, C. Pardal, V. Echarri, et al., "Energy Impact of the Air Infiltration in Residential Buildings in the Mediterranean Area of Spain and the Canary Islands," *Energy and Buildings* 188-189 (2019): 226–238, https://doi.org/10.1016/j.enbuild.2019.02.023.

[4] M. Mahmoodzadeh, V. Gretka, S. Wong, T. Froese, and P. Mukhopadhyaya, "Evaluating Patterns of Building Envelope Air Leakage With Infrared Thermography," *Energies (Basel)* 13, no. 14 (2020): 3545, https://doi.org/10.3390/en13143545.

[5] A. Al-Habaibeh, A. Sen, and J. Chilton, "Evaluation Tool for the Thermal Performance of Retrofitted Buildings Using an Integrated Approach of Deep Learning Artificial Neural Networks and Infrared Thermography," *Environment* 2, no. 4 (2021): 345–365, https://doi.org/10.1016/j.enbenv.2020.06.004.

[6] C. Lerma, E. Barreira, and R. M. S. F. Almeida, "A Discussion Concerning Active Infrared Thermography in the Evaluation of Buildings Air Infiltration," *Energy and Buildings* 168 (2018): 56–66, https://doi.org/10.1016/j.enbuild.2018.02.050.

[7] M. Fox, D. Coley, S. Goodhew, and P. de Wilde, "Thermography Methodologies for Detecting Energy Related Building

Defects," *Renewable and Sustainable Energy Reviews* 40 (2014): 296–310, https://doi.org/10.1016/j.rser.2014.07.188.

[8] EPBD, "Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the Energy Performance of Buildings," *Official Journal of the European Union* 1275 (2024): 1–68, http://data.europa.eu/eli/dir/2024/1275/oj (accessed July 17, 2025).

[9] R. Gil-Valverde, D. Tamayo-Alonso, A. Royuela-del-Val, I. Poza-Casado, A. Meiss, and M. Á. Padilla-Marcos, "Three-Dimensional Characterization of Air Infiltration Using Infrared Thermography," *Energy and Buildings* 233 (2021): 110656, https://doi.org/10.1016/j.enbuild.2020.110656.

[10] D. Tamayo-Alonso and A. Meiss, "A Novel Methodology for the Characterisation of Airflow Using Infrared Thermography and Pressurisation Test," *Building and Environment* 272 (2025): 112672, https://doi.org/10.1016/j.buildenv.2025.112672.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature* 521, no. 7553 (2015): 436–444, https://doi.org/10.1038/nature14539.

[12] D. Gertsvolf, M. Horvat, A. Khademi, D. Aslam, and U. Berardi, "Image Processing for Future Machine Learning Algorithm Applications on Infrared Thermography of Building Envelope Systems," in *Environmental Science and Engineering* (Springer Science and Business Media Deutschland GmbH, 2023), 549–556, https://doi.org/10.1007/978-981-19-9822-5_58.

[13] R. Wu, Y. Ren, M. Tan, and L. Nie, "Fault Diagnosis of HVAC System With Imbalanced Data Using Multi-Scale Convolution Composite Neural Network," *History* 17, no. 3 (2024): 371–386, https://doi.org/10.1007/s12273-023-1086-1.

[14] E. A. Setiawan and M. Fathurrahman, "Exploring Insights on Deep Learning-Based Photovoltaic Fault Detection for Monofacial and Bifacial Modules Using Thermography," *International Journal of Cognitive Computing in Engineering* 6 (2025): 495–507, https://doi.org/10.1016/j.ijcce.2025.04.001.

[15] A. Choudhary, T. Mian, and S. Fatima, "Convolutional Neural Network Based Bearing Fault Diagnosis of Rotating Machine Using Thermal Images," *Measurement* 176 (2021): 109196, https://doi.org/10.1016/j.measurement.2021.109196.

[16] G. Park, M. Lee, H. Jang, and C. Kim, "Thermal Anomaly Detection in Walls via CNN-Based Segmentation," *Automation in Construction* 125 (2021): 103627, https://doi.org/10.1016/j.autcon.2021.103627.

[17] W. Khor, Y. K. Chen, M. Roberts, and F. Ciampa, "Automated Detection and Classification of Concealed Objects Using Infrared Thermography and Convolutional Neural Networks," *Scientific Reports* 14, no. 1 (2024): 8353, https://doi.org/10.1038/s41598-024-56636-8.

[18] L. Deng, H. Zuo, W. Wang, C. Xiang, and H. Chu, "Internal Defect Detection of Structures Based on Infrared Thermography and Deep Learning," *KSCE Journal of Civil Engineering* 27, no. 3 (2023): 1136–1149, https://doi.org/10.1007/S12205-023-0391-7.

[19] A. Nakhaee and A. Paydar, "DeepRadiation: An Intelligent Augmented Reality Platform for Predicting Urban Energy Performance Just Through 360 Panoramic Streetscape Images Utilizing Various Deep Learning Models," *Building Simulation* 16, no. 3 (2023): 499–510, https://doi.org/10.1007/s12273-022-0953-5.

[20] L. Zhang and Y. Pan, "Information Fusion for Automated Post-Disaster Building Damage Evaluation Using Deep Neural Network," *Sustainable Cities and Society* 77 (2022): 103574, https://doi.org/10.1016/j.scs.2021.103574.

[21] Y. Cao, Y. Dong, Y. Cao, J. Yang, and M. Y. Yang, "Two-Stream Convolutional Neural Network for Non-Destructive Subsurface Defect Detection via Similarity Comparison of Lock-In Thermography Signals," *NDT and E International* 112 (2020): 102246, https://doi.org/10.1016/J.NDTEINT.2020.102246.

[22] C. St-Antoine, M. C. M. Paradis, P. Gonnel, et al., "Novel Method to Automatize Flash Point Detection in Small Volumes of Liquid by Computer Vision Using Thermal Images," *Measurement* 253 (2025): 117629, https://doi.org/10.1016/j.measurement.2025.117629.

[23] J. Kang, Y. J. Park, J. Lee, S. H. Wang, and D. S. Eom, "Novel Leakage Detection by Ensemble CNN-SVM and Graph-Based Localization in Water Distribution Systems," *IEEE Transactions on Industrial Electronics* 65, no. 5 (2018): 4279–4289, https://doi.org/10.1109/TIE.2017.2764861.

[24] A. Royuela-del-Val, M. Á. Padilla-Marcos, A. Meiss, P. Casaseca-de-la-Higuera, and J. Feijó-Muñoz, "Air Infiltration Monitoring Using Thermography and Neural Networks," *Energy and Buildings* 191 (2019): 187–199, https://doi.org/10.1016/J.ENBUILD.2019.03.019.

[25] D. P. Kingma and J. Lei Ba, "Adam: A Method for Stochastic Optimization): https://doi.org/10.48550/arXiv.1412.6980.