Transfer Learning for Predicting Thermal Comfort in Office Environments with Climate Similar to Tunisia: Overcoming Data Scarcity with Deep GRU-BiGRU Models

Mohamed Khayri RAHMANI[1,2], Hajer CHTIOUI[1], Jalel BEN HADJ SLAMA[1], Mireille GETTLER SUMMA[2,3], Anouar BEN KHALIFA[1,4]

***Abstract*— Transfer learning is considered an effective technique that enhances model performance by using knowledge acquired from a source dataset to tackle a similar task on a target dataset. This technique is particularly valuable in fields where labeled data is limited, such as thermal comfort prediction. In the Tunisian context, the lack of specific thermal comfort data for office spaces occupied by multiple people represents a major challenge to optimizing workplace environments. To address this, we introduce a transfer learning-based approach using ASHRAE RP-884 data from countries with similar climatic conditions. In fact, these data were purified based on climate conditions and selected for office-type buildings. Three transfer learning methods were evaluated using three models: a Deep GRU-BiGRU model, a Deep GRU model, and a BiGRU model. Our results present that the transfer learning approach based on the Deep GRU-BiGRU model achieves the highest accuracy, reaching 66.15% in thermal comfort prediction, outperforming the other methods.**

***—Keywords*—Transfer learning, Thermal comfort prediction, Deep GRU-BiGRU model, Data scarcity, ASHRAE RP-884**

I. Introduction

The rapid global urbanization process has significantly increased in energy consumption, specifically due to the widespread use of heating, ventilation, and air conditioning (HVAC) systems [1]. This trend is especially pronounced in developing countries, where infrastructural expansion comes with a high energy demand. However, earlier research indicates that more than 80% of individuals spend most of their time indoors, and more than half of the occupants report dissatisfaction with current thermal comfort conditions [2]. Given these concerns, accurately predicting and managing thermal comfort is important for both energy efficiency and occupant satisfaction. Traditional thermal comfort prediction models have long been applied as benchmarks. However, these models have limited real-world applicability, specifically in accounting for person-occupant variations, including metabolism and clothing. In response to this limitation, machine learning and deep learning-based approaches, have emerged as promising solutions. Algorithms including Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and advanced neural network models have demonstrated improved accuracy in predicting thermal comfort [3]. Additionally, the incorporation of personalized comfort models, which consider individual occupant preferences, has proven effective in optimizing both comfort and energy management in buildings [4]. Despite advancements, personalized models rely heavily on extensive environmental and subjective data. In fact, collecting sufficient occupant feedback poses a significant challenge, as surveys are intrusive and require considerable manual effort [5]. To address this limitation, Transfer Learning (TL) proposes a crucial solution enabling models pre-trained on existing datasets to adapt to new contexts with minimal additional data. In a related context, the Deep CNN-BiGRU model [6], which combines multiple data streams in gesture recognition, has also shown good performance in transfer learning for predicting thermal comfort.

Based on this idea, this study proposes a deep learning-based model with knowledge transfer to enhance thermal comfort prediction. The general layout of the proposed approach is illustrated in Fig. 1. The goal of this research is to develop a more generalized and efficient model capable of adapting predictions to various environments and climatic conditions while reducing the reliance on large training datasets.

The key contributions of this paper are outlined below:

- Proposal of a Transfer Learning approach based on Deep GRU - BiGRU for thermal comfort prediction in office environments, seek to improve adaptability across various climatic conditions.
- A thorough comparative evaluation with machine learning models to evaluate the performance of the proposed approach.

The structure of this paper is as follows: Section 2 provides an overview of related works on Transfer Learning for thermal comfort prediction. In Section 3, wediscuss the proposed Deep GRU - BiGRU approach and the comparative methods. Section 4 details the experimental analysis, and Section 5 offers a summary of the conclusions.

[1]Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, LATIS - Laboratory of Advanced Technology and Intelligent Systems, 4023, Sousse, Tunisie.

[2]Be Excellent Creative Optimist Motivated Energic Sasu (BECOME), Paris, France.

[3]Modyco Laboratory UMR 7114, University of Paris Nanterre, Nanterre, France.

[4]Université de Jendouba, Institut National des Technologies et des Sciences du Kef, 7100, Le Kef, Tunisie.

Email: mohmedkhayri.rahmani@eniso.u-sousse.tn

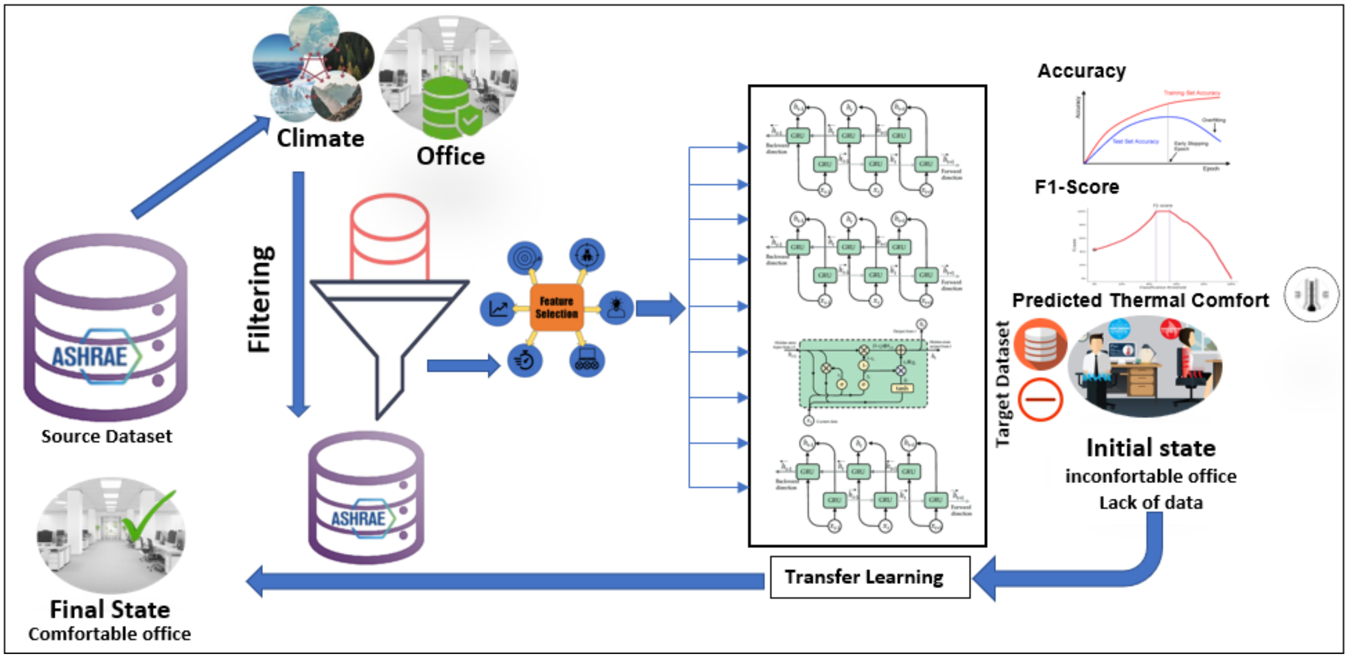


Fig. 1. General layout of the proposed approach

II. Related Works on Transfer Learning for Thermal Comfort Prediction

Thermal comfort prediction is a crucial aspect of energy-efficient building management, and recent advancements in deep learning and transfer learning have significantly enhanced predictive capabilities. Prior research have explored various methodologies to enhance the accuracy and flexibility of thermal comfort models, specifically in environments with limited data. One of the key challenges in predicting thermal comfort is the inconsistency in data availability and environmental conditions across various regions. To deal with this limitation, Jiao and Tan [7] suggests a transfer learning method applied in West Bengal, employing deep learning to improve indoor thermal comfort prediction in tropical climates. Similarly, Zhang and Li [8] Explored the application of a Transformer-based transfer learning model to enhance thermal comfort prediction when only limited data is available, demonstrating the effectiveness of leveraging spatiotemporal relationships for improved model generalization. According to these results, Somu et al. [9] presented a hybrid deep transfer learning approach that combines domain adaptation techniques to enhance thermal comfort predictions in various building environments. In a complementary approach, Park and Park [10] explored a transfer learning approach based on ensembles, which leverages wearable and environmental sensors to enhance individual thermal comfort prediction and enhancing both accuracy and robustness. Recently, Li et al. [11] discussed an instance-based transfer learning approach incorporating Nearest Neighbor Search (NNS) and an improved TrAdaBoost algorithm, optimizing data selection from the source domain to enhance prediction reliability.

Beyond improving prediction models within a single environment, other studies have addressed transfer learning across different geographical contexts. Gao et al. [12] tackled the challenge of data scarcity by leveraging transfer learning across multiple cities within the same climate zone. Their proposed transfer learning-based multilayer perceptron model (TL-MLP-C*), trained on datasets from ASHRAE RP-884, the Scales Project, and Medium US Office, showed superior Prediction accuracy in comparison with cutting-edge methods. Furthermore, Hu et al. [13] utilized heterogeneous transfer learning for thermal comfort modeling, addressing the challenge of data heterogeneity between different environments. Recent advancements have also focused on combining multiple transfer learning techniques to further enhance predictive performance. For instance, Ahn and Kim [14] applied a TL-LSTM model on simulation data to predict power consumption with minimal training, achieving high accuracy across climate zones. Yang et al. [15] employed unsupervised domain adaptation to predict personalized thermal comfort with limited labeled data. Their deep adaptation methods improved accuracy by up to 15%, demonstrating the effectiveness of transfer learning in adapting models to new conditions.

In summary, transfer learning has emerged as a crucial technique for enhancing thermal comfort prediction models by addressing challenges related to data scarcity and variability across different environments. By leveraging knowledge from previously trained models, transfer learning enables more efficient learning processes, reducing the need for large labeled datasets while improving predictive accuracy. Various strategies, such as domain adaptation, ensemble learning, and hybrid modeling, have been integrated to further optimize performance.

III. Proposed Approach and Comparative Methods for Transfer Learning

This section discusses the proposed Deep Gated Recurrent Unit - Bidirectional Gated Recurrent Unit (Deep GRU-BiGRU) approach for thermal comfort prediction and compares it with other relevant models.

A. Technical Characteristics of the Proposed Method: Deep GRU-BiGRU

The Deep GRU-BiGRU model is a deep learning architecture that combines GRU and Bi-GRU layers. This architecture is specifically designed to efficiently extract temporal patterns from thermal comfort data. The model is initially trained using source data, and then it is fine-tuned on a target dataset through the application of transfer learning.

For the training process, the categorical cross-entropy loss function is employed, as defined by Equation 1:

$$\mathcal{L} = -\sum_{f=1}^{n}\sum_{k=1}^{K} A_{f,k}\log(\hat{A}_{f,k}) \tag{1}$$

where:

- $A_{f,k}$ denotes the true label for the f-th instance and class k.
- $\hat{A}_{f,k}$ is the predicted probability for the f-th instance and class k.
- n represents the number of instances, and K is the number of possible classes.

The architecture of the model includes several recurrent layers structured as shown in Table. I:

TABLE I
ARCHITECTURE OF THE PROPOSED MODEL

Layer	Type	Details
1	Bi-GRU	256 units, BatchNorm, Dropout 40%
2	Bi-GRU	128 units, BatchNorm, Dropout 30%
3	GRU	64 units, BatchNorm, Dropout 30%
4	Bi-GRU	32 units, BatchNorm, Dropout 20%
5	Dense	64 neurons, ReLU activation
Final	Dense	5 output units, Softmax activation

The model is trained with the Adam optimizer and a learning rate of $\eta = 0.0005$ across 50 epochs, using a batch size of 16. To adapt the model to a new dataset, transfer learning is used, which involves freezing the early layers of the pre-trained model and fine-tuning the final layers with the target dataset. This adaptation process is mathematically expressed as (Equation 2):

$$\Theta^* = \arg\min_{\Theta}\sum_{f=1}^{n} l(A_f, f(q_f;\Theta)) \tag{2}$$

where Θ^* represents the parameters after fine-tuning.

B. Technical Characteristics of Compared Models

- GRU Model:

The Deep GRU model leverages stacked layers to model temporal dependencies, mitigating the issue of vanishing gradients. It processes information selectively using gating mechanisms, which allow it to retain relevant data and discard irrelevant parts. Dropout regularization further advances its generalization ability. The underlying equations are provided in Equations 3-6.

$$p_k = \sigma(E_p q_k + U_p S_{k-1} + b_p) \tag{3}$$

The update gate p_k (Equation 3) determines the portion of the previous hidden state S_{k-1} to retain. It depends on:

- q_k: The input at time step k.
- S_{k-1}: The hidden state from the previous step.
- E_p, U_p, and b_p: The weight matrices and bias term for the update gate, where E_p and U_p represent the weights for both the current input and the preceding hidden state.
- σ: The sigmoid activation function, which produces output values ranging from 0 to 1.

$$G_k = \sigma(E_G q_k + U_G S_{k-1} + b_G) \tag{4}$$

The reset gate G_k (Equation 4) controls the extent to which the previous hidden state is ignored. Similar to the update gate, it involves parameters related to the input and the previous hidden state.

$$\tilde{S}_k = \tanh(E_S q_k + U_S(G_k \odot S_{k-1}) + b_S) \tag{5}$$

The candidate hidden state $\tilde{S}_k$ (Equation 5) is computed with the reset gate's influence. The inputs include:

- $G_k \odot S_{k-1}$: The reset gate modifies the previous hidden state.
- E_S, U_S, and b_S: Weight matrices and bias term for the candidate hidden state.
- $\tanh$: The hyperbolic tangent activation function that outputs values between -1 and 1.

$$S_k = (1 - p_k) \odot S_{k-1} + p_k \odot \tilde{S}_k \tag{6}$$

The final hidden state S_k (Equation 6) is derived from a combination of the previous hidden state S_{k-1} and the candidate hidden state $\tilde{S}_k$, with the update gate p_k dictating the proportion.

- Bi-GRU Model:

The Bi-GRU architecture enhances the GRU by incorporating bidirectional layers, Allowing the network to process sequences in a bidirectional mannerboth previous and subsequent contexts, leading to a richer representation. The fusion of outputs from both directions is mathematically expressed in Equations 7-9.

$$\overrightarrow{S_k} = GRU(q_k, \overrightarrow{S_{k-1}}) \tag{7}$$

$$\overleftarrow{S_k} = GRU(q_k, \overleftarrow{S_{k+1}}) \tag{8}$$

$$S_k = [\overrightarrow{S_k}, \overleftarrow{S_k}] \quad (9)$$

$\overrightarrow{S_k}$, $\overleftarrow{S_k}$, $\overrightarrow{S_{k-1}}$ and $\overleftarrow{S_{k+1}}$: Hidden states in both directions at times k, $k-1$, and $k+1$.

Table II summarizes the key characteristics of the Deep GRU and Bi-GRU models.

TABLE II
COMPARISON OF DEEP GRU AND BI-GRU MODELS

Characteristic	**Deep GRU**	**Bi-GRU**
Number of layers	3 GRU layers	2 Bi-GRU layers
Neurons per layer	128, 64, 32	64, 32
Activation function	ReLU	Tanh
Regularization	Dropout (0.2)	Dropout (0.3)
Optimizer	Adam	Adam
Learning rate	0.001	0.0005
Loss function	Categorical Crossentropy	Categorical Crossentropy
Batch size	32	64
Number of epochs	50	50

Both models are well-suited for sequence classification tasks, but Bi-GRU, due to its bidirectional nature, is particularly beneficial when full contextual information is needed. Deep GRU, on the other hand, offers strong feature extraction capabilities through its multiple stacked layers. The choice of the hyperparameters and the architecture of the Deep GRU - BiGRU model was optimized after several trials, ensuring the best performance for thermal comfort prediction, as well as for the compared models (Deep GRU and Bi-GRU). In addition, other existing transfer learning approaches, including Transfer Learning Multilayer Perceptron (TL-MLP) [12], Transfer Learning Long Short-Term Memory (TL-LSTM) [14], Domain-Adversarial Adaptive Network (DAAN) and Domain-Specific Adaptation Network (DSAN) [15], were also evaluated and compared with our proposed approach to assess their effectiveness in different thermal comfort prediction scenarios.

IV. EXPERIMENTAL ANALYSIS

A. Datasets

The selected dataset is the ASHRAE RP-884 database [16], which contains more than 25,000 data points from 52 studies conducted across 26 cities. This dataset was initially compiled to develop the adaptive model proposed by De Dear [17]. It has been extensively utilized in thermal comfort research [18], [19].

to guarantee the relevance and quality of the data utilized, filtering was applied to the *Climate* and *Building type* columns, as outlined in Table III. This process narrowed the dataset to include only those matching the climate types and building categories relevant to our study, enabling transfer learning from offices in climates similar to Tunisia to an office in Tunisia, where offices are workplace environments typically occupied by multiple occupants, leading to diverse thermal comfort preferences. Thus, we selected from the filtered dataset the portion relevant to Tunisia as the target dataset, while the rest of the data formed the source dataset. Additionally, extreme thermal sensations, such as the values +3 and +2, -3 and -2, were grouped into broader classes, where the two classes +3 (Hot) and +2 (Warm) were combined into a single class +2 (Warm/Hot), while the two classes -3 (Cold) and -2 (Cool) were merged into a single class -2 (Cool/Cold). This approach aims to create a more balanced data distribution, facilitating the analysis. Finally, the parameters selected for both the *source dataset* and the *target dataset* are presented in Table III, to clarify the criteria used in our model.

TABLE III
COMPARISON OF SOURCE AND TARGET DATASETS

Feature	**Source Dataset**	**Target Dataset**
Climate	Hot semi-arid, Hot-summer Mediterranean	Hot semi-arid, Hot-summer Mediterranean
Building type	Office	Office
Cooling strategy (building level)	Naturally Ventilated, Air Conditioned	Naturally Ventilated, Air Conditioned
Clothing	0.31 - 2.19	0.41 - 1.96
Metabolism	0.8 - 2.9	1 - 2.8
Air temperature (°C)	15.4 - 39.8	16 - 31.8
Air velocity (m/s)	0 - 2.54	0 - 1.59
Relative humidity (%)	10.3 - 86.8	48 - 79
Outdoor monthly air temperature (°C)	16 - 40.9	16 - 27
Thermal Sensation	5 classes (-2, -1, 0, +1, +2)	5 classes (-2, -1, 0, +1, +2)
Number of instances	18725	324

B. Results and Discussion

The results explains in Table IV highlight the success of the Deep GRU - BiGRU model in predicting thermal comfort through transfer learning. Our method attains an accuracy of 66.15% and an F1-score of 58.60%, surpassing traditional deep learning models. These metrics confirm the model's robustness in adapting knowledge from source datasets to the target environment. Examining the confusion matrices in Figure 2, we observe notable improvements in classifying intermediate thermal sensation levels (-1, 0, +1), which are typically more challenging to predict. This indicates that the combination of GRU and BiGRU architectures enhances feature extraction, allowing the model to learn from both past and future contexts, leading to better generalization across diverse climatic conditions. Additionally, the results highlight the advantages of transfer learning from regions with similar climates to Tunisia. By leveraging knowledge from offices in these regions, the model improves classification accuracy and better adapts to the thermal comfort patterns in Tunisian environments. This approach illustrates the potential of using existing datasets to improve predictions in cases with limited local data.

TABLE IV
EVALUATION OF DIFFERENT TRANSFER LEARNING MODELS FOR THERMAL COMFORT PREDICTION

Model	Accuracy (%)	F1-Score (%)
Deep GRU	60.00	53.61
BiGRU	61.54	55.29
TL-MLP [12]	50.76	53.60
TL-MLP-C* [12]	54.50	55.12
TL-LSTM [14]	53.20	51.90
DAAN [15]	60.35	57.00
DSAN [15]	60.88	57.50
Proposed Deep GRU - BiGRU	**66.15**	**58.60**

C. Ablation Study

To underscore the critical role of our transfer learning approach, we conducted an ablation study where the same models were trained without utilizing transfer learning on a combined dataset that includes both the source and target datasets. Table IV displays the accuracy and F1-scores for these models, and the results are telling. The performance drops significantly compared to our transfer learning approach, highlighting the substantial impact of pre-trained knowledge in improving The model's capability to accurately predict thermal comfort. This demonstrates that transfer learning is not merely a supplementary technique, but a crucial factor that substantially enhances model performance, especially when dealing with complex environmental variables that may vary across different contexts. Furthermore, Figure 3 illustrates the confusion matrices for the three models: *Deep GRU*, *BiGRU*, and *Deep GRU-BiGRU* without transfer learning. A more pronounced distribution of misclassifications can be observed, confirming the models' difficulty in generalizing effectively without knowledge transfer.

TABLE V
EVALUATION OF DIFFERENT MODELS FOR THERMAL COMFORT PREDICTION WITHOUT TRANSFER LEARNING

Model	Accuracy (%)	F1-Score (%)
Deep GRU	42.97	32.95
BiGRU	42.91	33.04
Proposed Deep GRU - BiGRU	43.08	35

V. CONCLUSION

This study introduces a transfer learning method using Deep GRU-BiGRU models to predict thermal comfort in office environments with climates similar to Tunisia's, addressing the challenge of limited local data. Our results show that this method outperforms others, achieving 66.15% accuracy. The study highlights The impact of transfer learning in enhancing model accuracy when data is limited and provides

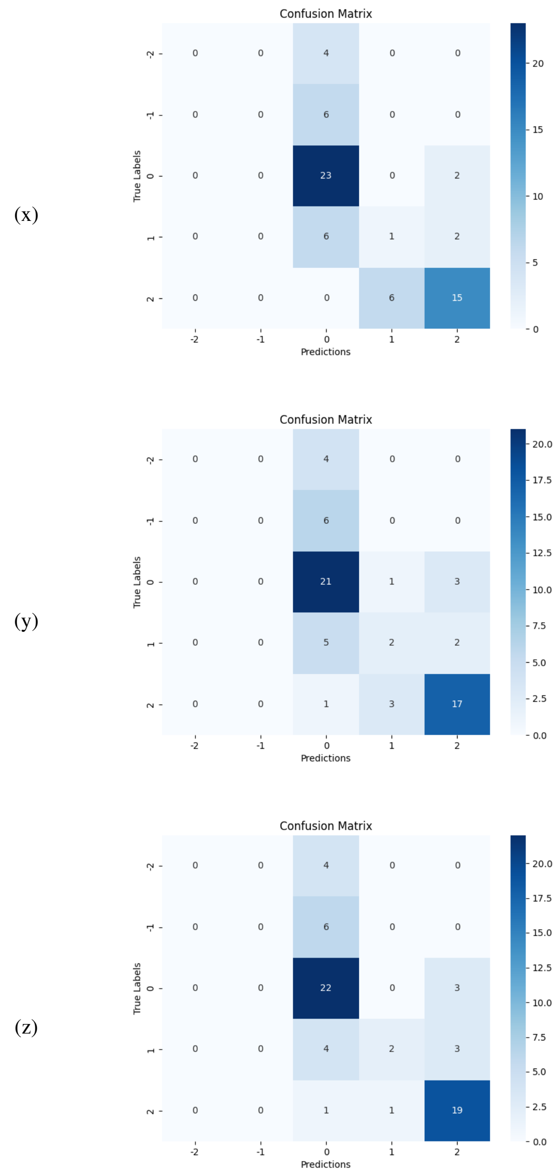
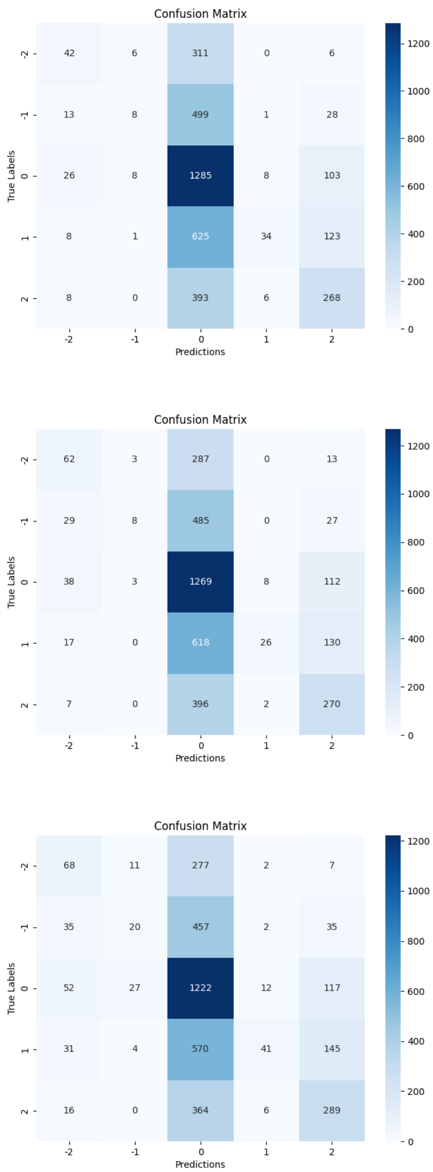


Fig. 2. Confusion matrices for the three transfer learning models: *Deep GRU* (x), *BiGRU* (y), and *Deep GRU - BiGRU* (z)

a framework adaptable to similar climates. The ablation study confirms the importance of pre-trained knowledge, as models without it perform significantly worse. The results pave the way for future studies aimed at enhancing the performance of transfer learning models, particularly by incorporating domain adaptation strategies and active learning techniques. In addition, expanding this approach to include a broader range of environments and real-time data could enable more accurate and Optimizing thermal comfort management in buildings for energy efficiency.

ACKNOWLEDGMENTS

The authors would like to express their deep gratitude to *the Tunisian Ministry of Higher Education and Scientific Research* for its unwavering support. Special thanks are



(x)

(y)

(z)

Fig. 3. Confusion matrices for the three models: (x) *Deep GRU*, (y) *BiGRU*, and (z) *Deep GRU - BiGRU* without transfer learning

extended to *the University of Sousse* for academic guidance. Sincere thanks go to *Bimatech company* for its valuable collaboration and technical support throughout this project. Their contribution has been essential to the completion of this work.

References

[1] H. Niculita-Hirzel, "Latest trends in pollutant accumulations at threatening levels in energy-efficient residential buildings with and without mechanical ventilation: A review," *International Journal of Environmental Research and Public Health*, vol. 19, no. 6, p. 3538, 2022.

[2] W. Song and J. K. Calautit, "Inclusive comfort: A review of techniques for monitoring thermal comfort among individuals with the inability to provide accurate subjective feedback," *Building and Environment*, p. 111463, 2024.

[3] J. Xie, H. Li, C. Li, J. Zhang, and M. Luo, "Review on occupant-centric thermal comfort sensing, predicting, and controlling," *Energy and Buildings*, vol. 226, p. 110392, 2020.

[4] J. Kim, S. Schiavon, and G. Brager, "Personal comfort models–a new paradigm in thermal comfort for occupant-centric environmental control," *Building and Environment*, vol. 132, pp. 114–124, 2018.

[5] X. Xu, H. Yu, Q. Sun, and V. W. Tam, "A critical review of occupant energy consumption behavior in buildings: How we got here, where we are, and where we are headed," *Renewable and Sustainable Energy Reviews*, vol. 182, p. 113396, 2023.

[6] N. M. Bhiri, S. Ameur, I. Jegham, I. Alouani, and A. B. Khalifa, "A deep cnn-bigru network for multi-stream hand gesture recognition framework," in *2024 10th International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, 2024, pp. 893–898.

[7] Y. Jiao and Z. Tan, "Enhancing indoor thermal comfort prediction in tropical regions: A transfer learning strategy in west bengal," *Journal of Building Engineering*, vol. 98, p. 111142, 2024.

[8] X. Zhang and P. Li, "Transfer learning in the transformer model for thermal comfort prediction: a case of limited data," *Energies*, vol. 16, no. 20, p. 7137, 2023.

[9] N. Somu, A. Sriram, A. Kowli, and K. Ramamritham, "A hybrid deep transfer learning strategy for thermal comfort prediction in buildings," *Building and Environment*, vol. 204, p. 108133, 2021.

[10] H. Park and D. Y. Park, "Prediction of individual thermal comfort based on ensemble transfer learning method using wearable and environmental sensors," *Building and Environment*, vol. 207, p. 108492, 2022.

[11] K. Li, Y. Liu, L. Chen, and W. Xue, "Data efficient indoor thermal comfort prediction using instance based transfer learning method," *Energy and Buildings*, vol. 306, p. 113920, 2024.

[12] N. Gao, W. Shao, M. S. Rahaman, J. Zhai, K. David, and F. D. Salim, "Transfer learning for thermal comfort prediction in multiple cities," *Building and Environment*, vol. 195, p. 107725, 2021.

[13] W. Hu, Y. Luo, Z. Lu, and Y. Wen, "Heterogeneous transfer learning for thermal comfort modeling," in *Proceedings of the 6th ACM international conference on systems for energy-efficient buildings, cities, and transportation*, 2019, pp. 61–70.

[14] Y. Ahn and B. S. Kim, "Prediction of building power consumption using transfer learning-based reference building and simulation dataset," *Energy and Buildings*, vol. 258, p. 111717, 2022.

[15] C. Yang, K. Taniguchi, and Y. Akashi, "Transfer learning with unsupervised domain adaptation for personal thermal comfort prediction," *Energy and Buildings*, p. 115449, 2025.

[16] A.-C. E. American Society of Heating, Refrigerating and A. N. S. Institute, *Thermal environmental conditions for human occupancy*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2004, vol. 55.

[17] R. de Dear and G. S. Brager, "Developing an adaptive model of thermal comfort and preference," 1998.

[18] J. Langevin, J. Wen *et al.*, "Relating occupant perceived control and thermal comfort: Statistical analysis on the ashrae rp-884 database," *HVACR Research*, vol. 18, no. 1-2, pp. 179–194, 2012.

[19] W. Liu, Z. Lian *et al.*, "Heart rate variability at different thermal comfort levels," *European Journal of Applied Physiology*, vol. 103, pp. 361–366, 2008.