

# Simposio CEA de Robótica, Bioingeniería, Visión Artificial y Automática Marina 2025



## Estimación del confort térmico mediante imágenes termográficas

Leal, M.<sup>a,b,\*</sup>, Abad-Alcaraz, V.<sup>a,b</sup>, Castilla, M.<sup>a,b</sup>, Álvarez, J. D.<sup>a,b</sup>

<sup>a</sup>CIESOL, Centro de Investigaciones en Energía Solar. Centro mixto UAL-CIEMAT, Universidad de Almería - ceiA3 - Crta. Sacramento, s/n, 04120, La Cañada de San Urbano, Almería, España.

<sup>b</sup>Dpto. de Informática, Universidad de Almería - ceiA3, Crta. Sacramento, s/n, 04120, La Cañada de San Urbano, Almería, España

### Resumen

Este trabajo presenta una metodología para estimar la percepción del confort térmico a partir de imágenes termográficas obtenidas con una cámara de bajo coste. El estudio se desarrolló en el edificio bioclimático CIESOL y constó de seis ensayos experimentales con 18 participantes, realizados en condiciones de invierno y verano. Para abordar la limitación del conjunto de datos, se aplicaron técnicas de segmentación y aumento de imágenes. Se entrenaron modelos EfficientNet y Xception para clasificar las imágenes según la sensación térmica reportada por los sujetos, obteniéndose los mejores resultados con imágenes segmentadas y EfficientNet, alcanzando una exactitud del 84 %. A pesar de la variabilidad inherente a la percepción térmica, los modelos mostraron un rendimiento satisfactorio. Como trabajo futuro, se propone ampliar el conjunto de datos y explorar arquitecturas multimodales que integren datos fisiológicos y ambientales, con el fin de mejorar la personalización y robustez de los sistemas de confort térmico.

Palabras clave: Confort térmico, Visión por Computador, Redes Convolucionales, Inteligencia Artificial

#### Thermal comfort estimation using thermographic images

#### Abstract

This work presents a methodology for estimating thermal comfort perception using thermographic images captured with a lowcost camera. The study was conducted in the bioclimatic CIESOL building and involved six experimental trials with 18 participants under both winter and summer conditions. To address the limitations of a small dataset, image segmentation and augmentation techniques were applied. EfficientNet and Xception models were trained to classify thermal images according to the thermal sensation reported by the participants. The best results were obtained using segmented images and the EfficientNet model, achieving an accuracy of 84 %. Despite the inherent subjectivity and variability of thermal perception, the models demonstrated satisfactory performance. Future work will focus on expanding the dataset and exploring multimodal architectures that integrate physiological and environmental data to improve the personalization and robustness of thermal comfort systems.

Keywords: Thermal comfort, Computer Vision, Convolutional Networks, Artificial Intelligence

#### 1. Introducción

La creciente preocupación por la eficiencia energética y la reducción de las emisiones de gases de efecto invernadero ha situado al sector de la edificación como uno de los principales focos de los esfuerzos para mitigar el cambio climático. En 2023, los edificios fueron responsables del 34 % del consumo energético global y del 34 % de las emisiones de carbono relacionadas con la energía (UNEP, 2025). Dentro de este contexto, los sistemas de Calefacción, Ventilación y Aire acondicionado (HVAC del inglés *Heating, Ventilation and Air Conditioning*)

edificios, el consumo de energía y las emisiones siguen creciendo, superando los niveles previos a la pandemia (UNEP, 2021). Una gestión más adaptativa, centrada en el confort térmico de los ocupantes, es esencial para reducir estos impacto. Sin embargo, el comportamiento de los usuarios sigue siendo un factor clave de ineficiencia (Delzendeh et al., 2017).

representan entre el 40 % y el 60 % del consumo energético de

El confort térmico se define como "la condición mental que expresa satisfacción con el ambiente térmico" (ISO7730, 2023;

un edificio (Goh et al., 2018). A pesar de los avances en la eficiencia energética de los

<sup>\*</sup>Autor para correspondencia: marta.leal@ual.es

Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

ASHRAE, 2009), lo que refleja su naturaleza subjetiva, influida por factores físicos, fisiológicos y psicológicos. Normativas como ISO 7730 y ASHRAE 55 proponen índices para su evaluación, siendo el más común el Voto Medio Previsto (PMV del inglés *Predicted Mean Vote*), calculado a partir de variables ambientales (temperatura, humedad, velocidad del aire, etc.) y del usuario (actividad, vestimenta). Este índice se relaciona con el Porcentaje Estimado de Insatisfechos (PPD del inglés *Predicted Percentage of Dissatisfied*), que estima el grado de disconformidad térmica. No obstante, estos modelos no contemplan factores individuales como el género o la edad, lo que limita su capacidad de personalización (Chaudhuri et al., 2019).

Recientes estudios han demostrado el potencial de las imágenes termográficas combinadas con aprendizaje automático para predecir el confort térmico personal. En (Huang et al., 2025) fusionaron imágenes visibles e infrarrojas para mejorar la predicción en presencia de oclusiones faciales, alcanzando una exactitud del 92,73 %. Por otro lado, en (Baek et al., 2023) utilizaron una Red Neuronal Convolucional (*CNN*, del inglés *Convolutional Neural Networks*) para estimar la sensación térmica a partir de la distribución de temperatura en el cuerpo, obteniendo una exactitud del 96 %. Finalmente, en (Cosma and Simha, 2019) desarrollaron un enfoque no invasivo en tiempo real basado en información térmica de distintas partes del cuerpo, logrando una exactitud del 80 %. Estos trabajos confirman la viabilidad de estas tecnologías para aplicaciones personalizadas en entornos climatizados.

Este estudio describe la metodología empleada en una serie de ensayos experimentales para estimar el confort térmico mediante imágenes capturadas con una cámara termográfica de bajo coste. Los experimentos se llevaron a cabo en los laboratorios del Centro de Investigación de Energía Solar (CIESOL), situado en el campus de la Universidad de Almería (UAL). Se aplicaron técnicas de preprocesamiento de datos, aumento de imágenes y redes neuronales convolucionales para analizar las imágenes termográficas y estimar la percepción subjetiva del confort térmico de los ocupantes, complementada con la percepción del ocupante en cada instante de captura.

El artículo se organiza de la siguiente manera: la sección 2 describe el edificio CIESOL, el diseño experimental y los métodos de recopilación y procesamiento de datos, incluyendo técnicas de preprocesamiento de imágenes, aumento de datos y redes neuronales convolucionales. En la sección 3 se presentan y discuten los resultados. Finalmente, la sección 4 expone las conclusiones y posibles trabajos futuros.

#### 2. Materiales y métodos

#### 2.1. Recopilación y preparación de datos

El CIESOL<sup>1</sup> es un centro mixto de investigación en energía solar, fruto de la colaboración entre la UAL y el Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT). Sus instalaciones albergan grupos de investigación de ambas instituciones y cuentan con una superficie de 1072 m<sup>2</sup> distribuidos en dos plantas. El centro dispone de ocho laboratorios orientados al norte, seis despachos en la planta baja (más uno principal en la planta superior), así como espacios para el personal incluyendo cocina y aseos.

El centro, que ha sido diseñado bajo principios de arquitectura bioclimática, integra un sistema de climatización basado en refrigeración solar. Este sistema está compuesto por un campo de captadores solares, un sistema de almacenamiento de agua caliente, una caldera de alto rendimiento y una máquina de absorción con torre de refrigeración. Asimismo, el edificio está dotado de una red avanzada de sensores que permiten la monitorización de más de 725 señales a través de un Sistema de Supervisión, Control y Adquisición de Datos (SCADA).

Los ensayos experimentales se han llevado a cabo en los laboratorios 5 y 6, ubicados en la planta superior. El laboratorio 5 tiene una superficie de 27,97 m<sup>2</sup>, mientras que el laboratorio 6 alcanza los 42,11 m<sup>2</sup>. En la Figura 1 se muestra el interior del laboratorio 6 donde se realiza la mayor parte de los ensayos y está dotado de una amplia red de sensores.



Figura 1: Interior del laboratorio 6 y sus principales sensores.

La fase experimental se desarrolló a lo largo de seis ensayos realizados en días diferentes, tres bajo condiciones de funcionamiento del sistema HVAC en modo verano y los tres restantes en modo invierno. Participaron un total de 18 voluntarios residentes en Almería, con edades entre 22 y 44 años y un Índice de Masa Corporal (IMC) entre 18,72 y 25,08. En cada ensayo participaron seis personas (tres hombres y tres mujeres). En los experimentos en modo invierno, se empleó como vestimenta estándar una camiseta de manga larga, ropa interior, pantalones vaqueros largos y zapatos, con un valor de aislamiento térmico estimado en 0,64 clo. En modo verano, la camiseta se sustituyó por una de manga corta, reduciendo el aislamiento a 0,4 clo, según (ASHRAE, 2023)). En todos los casos, la actividad metabólica de los participantes se mantuvo constante, de acuerdo a una situación de reposo con actividades ligeras como lectura y escritura (1,2 met), y la velocidad del aire se mantuvo por debajo de 0,2 m/s, de acuerdo con las recomendaciones de (ASH-RAE, 2023). Se indicaron pautas previas a seguir por los participantes antes de cada ensayo, como evitar alcohol, tabaco o ejercicio intenso, y se rotó la ubicación de los participantes para minimizar el efecto de posibles asimetrías en el flujo de aire.

<sup>&</sup>lt;sup>1</sup>https://ciesol.com/

Cada ensayo tuvo una duración de 210 minutos. El protocolo experimental se desarrolló como se muestra en la Figura 2. Durante los primeros 15 minutos, los participantes permanecieron en un laboratorio ajeno al del ensayo, el laboratorio 5, acondicionado a una temperatura neutra de 24 °C. A continuación, se trasladaron al laboratorio del ensayo, el laboratorio 6, donde permanecieron 195 minutos. Este espacio, acondicionado inicialmente a 20,1 ± 3,2 °C, alcanzó progresivamente los 27,0 ± 2,5 °C.



Figura 2: Protocolo experimental.

En el transcurso de los ensayos se obtuvo información sobre la percepción subjetiva del confort térmico de los ocupantes mediante encuestas cada 20 minutos estructuradas como se muestran en la Figura 3(a), considerando únicamente la escala A en este estudio. También se recopilaron datos fisiológicos, con una frecuencia de 40 minutos, como son la frecuencia cardíaca, la saturación de oxígeno, la presión arterial y la temperatura de la piel. Para medir la temperatura de la piel se ha utilizado la cámara termográfica FLIR ONE®. Es una cámara asequible que se puede acoplar a teléfonos móviles y tabletas y puede medir temperaturas en un rango de -20 °C a 120 °C con una precisión de ±3 %. Se pretendía registrar la temperatura en la frente, el cuello, ambas muñecas (cara dorsal y palmar) y los brazos. Para ello, los participantes se sentaron en una posición estandarizada, colocando una mano con la palma hacia arriba y la otra con la palma hacia abajo durante la toma de cada imagen.

Se obtuvieron un total de 194 imágenes termográficas, normalizadas utilizando la misma escala térmica en el rango de de 20–40 °C para garantizar la uniformidad entre muestras. Para su representación visual, se emplearon dos paletas de color termográficas comúnmente utilizadas: *Ironbow* y *Rainbow*, tal como se muestra en la Figura 3(b). La paleta *Ironbow* ofrece una transición progresiva entre tonos fríos (negros y azules) y cálidos (naranjas y blancos), lo cual facilita la identificación de gradientes térmicos en regiones cutáneas. Por su parte, la paleta *Rainbow* incorpora una gama cromática más amplia, lo que permite distinguir con mayor claridad pequeñas variaciones de temperatura (FLIR Systems, 2024). Estas diferencias visuales pueden influir en la capacidad de las redes neuronales para aprender patrones térmicos discriminativos, por lo que ambas configuraciones fueron consideradas en los experimentos.

#### 2.2. Preprocesamiento de imágenes

El tratamiento adecuado de las imágenes térmográficas constituye un paso esencial para garantizar un rendimiento ro-

busto de los modelos de redes neuronales convolucionales empleados en este estudio. Desde la elección de la paleta termográfica hasta las técnicas de segmentación y aumento de datos, se aplicaron distintos procesos para maximizar la calidad del aprendizaje automático en un contexto de datos limitados.

Todas las imágenes, en total 194, han sido redimensionadas a un tamaño estándar de entrada de 224×224×3 píxeles en consonancia con las arquitecturas de redes neuronales convolucionales preentrenadas seleccionadas. En cuanto al formato de las imágenes, se analizaron tres variantes: (i) imágenes completas sin procesar, (ii) imágenes recortadas centradas en la persona, y (iii) imágenes segmentadas con máscara, donde se suprime el fondo para conservar únicamente la silueta corporal. Para el segundo enfoque se aplica una umbralización para generar una máscara binaria, luego se realiza un análisis de componentes conectados para aislar la región de mayor área y, a partir de ella, se extrae la caja delimitadora que garantiza el centrado de la persona en la imagen. El último enfoque, la segmentación con máscara, se realiza aplicando técnicas de procesamiento como umbralización binaria (Seelaboyina and Vishwakarma, 2023) e interpolación mediante contornos activos (Chan and Vese, 2001) a partir de la imagen recortada centrada en la persona. En primer lugar, se transformaron las imágenes al espacio de color Lab, el cual permite separar la luminosidad de los componentes cromáticos y facilita la segmentación de regiones homogéneas, especialmente en entornos con variaciones suaves de color (Gonzalez and Woods, 2008). Luego, se aplica una umbralización global sobre la versión en escala de grises de la imagen para obtener una máscara binaria inicial, la cual fue invertida y erosionada con un elemento estructurante circular para eliminar ruido y definir mejor la región de interés. Posteriormente, se utilizó el método de contornos activos, que refina la segmentación mediante la evolución de la máscara binaria inicial hasta ajustarse a los bordes relevantes de la silueta corporal. Este proceso permite generar máscaras que conservan las zonas cutáneas más significativas (frente, mejillas, cuello, muñeca, etc.), eliminando información contextual irrelevante como el fondo de la escena.

Debido a la limitada cantidad de imágenes disponibles, se aplicaron técnicas de aumento de datos durante la fase de entrenamiento, con el objetivo de mitigar el sobreajuste. Entre las transformaciones utilizadas se incluyen: rotaciones aleatorias, escalado, traslaciones, reflexiones horizontales, deformaciones por cizalla y suavizado gaussiano. Estas operaciones se aplican aleatoriamente a un 33 % del conjunto de entrenamiento, generando nuevas muestras sintéticas que enriquecen la variabilidad de dicho conjunto.

Por último, se evaluaron dos configuraciones de clasificación: una con 7 clases correspondientes a la escala ASHRAE, y otra reducida a 3 clases (fresco, neutro y cálido), agrupando categorías poco representadas para aumentar la estabilidad del entrenamiento, como se observa en la Tabla 1. Este enfoque ha sido empleado también en trabajos previos de la literatura (Baek et al., 2023).

Tabla 1: Correspondencia entre la escala térmica de 7 clases y la de 3 clases										
Escala 7 clases	Frío	Fresco	Ligeramente frío	Neutro	Ligeramente cálido	Cálido	Sofocante			
Escala 3 clases	Fresco	Fresco	Fresco	Neutro	Cálido	Cálido	Cálido			





(b) Imágenes tomadas con la cámara térmográfica: Izquierda paleta *Ironbow*, derecha *Rainbow* 

Figura 3: Toma de datos mediante encuestas e imágenes tomadas con cámaras termográficas.

#### 2.3. Arquitectura de redes neuronales convolucionales

Dado el carácter limitado del conjunto de imágenes disponibles, se consideró fundamental estructurar los subconjuntos de entrenamiento y prueba para evitar desequilibrios de clase que pudieran comprometer el aprendizaje del modelo. Este aspecto resulta crítico en categorías minoritarias como *Fresco*, donde la escasez de muestras podría dificultar la obtención de un modelo robusto si no se aplican mecanismos de compensación adecuados. Para abordarlo, se ha aplicado un balanceo estratificado de clases durante el entrenamiento y la fase de evaluación.

Tras realizar varios ensayos, el 90 % de las imágenes se ha destinado al entrenamiento del modelo mediante una validación cruzada estratificada de tipo *k-fold*, con el objetivo de reducir la varianza asociada a la aleatoriedad en la partición de los datos. Esta técnica divide el conjunto de entrenamiento en cinco bloques manteniendo la proporción de clases en cada uno de ellos. El 10 % restante se ha reservado como conjunto de prueba independiente, seleccionando también las muestras de forma estratificada para preservar la proporción original entre las clases *Fresco, Neutro* y *Cálido*.

Para abordar la tarea de clasificación del confort térmico se ha hecho uso de redes convolucionales profundas mediante aprendizaje por transferencia, utilizando arquitecturas preentrenadas. En concreto, se evaluaron dos modelos: EfficientNet-B0 y Xception, ampliamente utilizados en tareas de visión por computador debido a su capacidad para capturar patrones espaciales complejos. Ambos ofrecen una buena eficiencia computacional, destacando EfficientNet-B0 por ser más eficiente que Xception. Estas redes han sido incorporadas como parte del modelo propuesto, cuya estructura general se ilustra en la Figura 4.

Ambas arquitecturas reciben como entrada imágenes redimensionadas a 224×224×3 píxeles, las cuales son normalizadas mediante centrado en cero (*zero-centering*), lo que mejora la estabilidad numérica del entrenamiento. Sobre la red base preentrenada se incorpora la capa totalmente conectada (*fully connected*) con activación de la función Unidad Lineal Rectificada (*ReLU*, por sus siglas en inglés de *Rectified Linear Unit*), seguida de una capa de abandono (*dropout*) empleada como técnica de regularización. Durante la fase de entrenamiento, se aplica un ajuste fino parcial (*fine-tuning*) descongelando de forma inicial únicamente las diez últimas capas convolucionales del modelo. Esta estrategia permite que los pesos de dichas capas puedan ajustarse al nuevo dominio de imágenes termográficas, aprovechando la capacidad de las capas iniciales para captar patrones básicos. En la Tabla 2 se muestran con más detalle los hiperparámetros y configuraciones adicionales empleadas durante el proceso de entrenamiento.

Tabla 2: Parámetros utilizados durante el entrenamiento de las redes

Parámetro	Valor
Optimizador	Adam
Tasa de aprendizaje inicial	$1 \times 10^{-4}$
Tamaño de mini lote (minibatch)	64
Número máximo de épocas	80
Frecuencia de validación	Cada 20 iteraciones
Regularización L2	$1 \times 10^{-5}$
Plan de reducción de tasa ( <i>LR schedule</i> )	Tipo escalonado (piecewise)
Capa de abandono	0,02

#### 3. Resultados

Para evaluar el rendimiento de los modelos de clasificación multiclase se emplearon diversas métricas, entre ellas la exactitud (accuracy), la precisión (precision), la recuperación (recall), la puntuación F1 y la matriz de confusión. Esta última permite analizar la relación entre las etiquetas reales y las predichas. A partir de la matriz se derivan las principales métricas: la exactitud indica el porcentaje de aciertos sobre el total de muestras, mientras que en contextos con desequilibrio entre clases resulta más representativo analizar la precisión, la recuperación y la puntuación F1 por clase. La precisión mide la proporción de aciertos entre todas las predicciones positivas, la recuperación refleja la proporción de casos positivos correctamente identificados, y la puntuación F1 combina ambas mediante una media armónica. Los resultados obtenidos con la configuración de 3 clases para ambas paletas se presentan en la Tabla 3.

Los mejores resultados se obtuvieron utilizando la red EfficientNet con imágenes segmentadas mediante máscaras, tanto para la paleta *Ironbow* como para *Rainbow*. La segmentación permitió eliminar información de fondo irrelevante, centrando el aprendizaje en las regiones corporales más representativas



Figura 4: Arquitectura de la red

Tabla 3: Métricas de clasificación sobre el conjunto de prueba para redes CNN Xception y EfficientNet, con distintas imágenes y paletas								
Paleta	Red CNN	Tipo Imagen	Exactitud	Precisión	Recuperación	Puntuación F1	Tiempo (min)	
1	EfficientNet	Original	0,58	0,57	0,56	0,53	14,35	
1	EfficientNet	Recortada	0,58	0,71	0,56	0,58	14,30	
1	EfficientNet	Con máscara	0,79	0,81	0,79	0,80	14,10	
1	Xception	Original	0,58	0,61	0,57	0,57	47,20	
1	Xception	Recortada	0,53	0,67	0,59	0,48	46,95	
1	Xception	Con máscara	0,74	0,75	0,73	0,73	43,92	
2	EfficientNet	Original	0,53	0,53	0,53	0,53	23,35	
2	EfficientNet	Recortada	0,53	0,57	0,55	0,53	13,78	
2	EfficientNet	Con máscara	0,84	0,88	0,83	0,82	13,87	
2	Xception	Original	0,58	0,55	0,56	0,55	49,92	
2	Xception	Recortada	0,58	0,58	0,58	0,56	47,10	
2	Xception	Con máscara	0,84	0,87	0,83	0,84	46,77	

para la percepción térmica, lo que contribuyó a mejorar significativamente el rendimiento de los modelos. En particular, la paleta *Rainbow* alcanzó una exactitud, expresada en porcentaje, del 84 % y una puntuación F1 de 0,82, reflejando una mejor capacidad del modelo para distinguir entre las distintas clases de sensación térmica. Aunque esta exactitud es ligeramente inferior a la reportada en estudios como (Huang et al., 2025; Baek et al., 2023) y superior a la obtenida en (Cosma and Simha, 2019), se considera muy competitiva dado el uso de cámaras de bajo coste, la heterogeneidad de las imágenes (diferentes estaciones, rango amplio de temperaturas y gran número de sujetos) y un conjunto de datos limitado.

En este mejor caso, con la paleta *Rainbow*, las métricas de rendimiento fueron similares entre EfficientNet y Xception. Sin embargo, se seleccionó EfficientNet como modelo final debido a su significativa ventaja computacional. EfficientNet completó el procesamiento en 13,87 minutos, frente a los 46,77 minutos requeridos por Xception, lo que representa una reducción del 70% en el tiempo de ejecución. Esta diferencia es especialmente relevante en aplicaciones en tiempo real o con recursos limitados, lo que justifica la elección de EfficientNet por su rendimiento competitivo y su alta eficiencia.

Las matrices de confusión correspondientes a los mejores casos de clasificación se muestran en la Figura 5. En ambos casos, se utilizó la red EfficientNet con imágenes preprocesadas mediante máscaras, aplicadas sobre dos paletas térmicas distintas: *Ironbow* (Paleta 1) y *Rainbow* (Paleta 2).



(a) Paleta Ironbow con máscara



(b) Paleta Rainbow con máscara

Figura 5: Matrices de confusión con EfficientNet y segmentación, usando las paletas térmicas *Ironbow* (a) y *Rainbow* (b)

La matriz de confusión de la paleta *Ironbow* muestra un buen desempeño general, aunque con cierta confusión entre las clases *Neutro* y *Cálido*. Esta confusión puede deberse a la variabilidad subjetiva en la percepción térmica de los individuos, especialmente en condiciones próximas al confort térmico neutro. Por otro lado, la matriz de la paleta *Rainbow* evidencia un rendimiento superior en la clasificación de la clase *Cálido*, con 7 aciertos de 7 posibles, y una mejor distribución en las predicciones de las clases *Neutro* y *Fresco*, lo que sugiere una mayor precisión y consistencia del modelo con esta paleta. La Figura 6 muestra ejemplos aleatorios del conjunto de prueba que ilustran aciertos y errores del modelo.



Figura 6: Ejemplos de predicciones del modelo con imágenes segmentadas

#### 4. Conclusiones

Este estudio ha demostrado el potencial del uso de imágenes termográficas combinadas con técnicas de aprendizaje profundo para estimar la percepción subjetiva del confort térmico, una tarea compleja debido a su naturaleza influida por factores fisiológicos y psicológicos individuales. A pesar de contar con un conjunto de datos limitado, se aplicaron estrategias como el aumento de datos y la segmentación de imágenes para maximizar su aprovechamiento. Estas medidas permitieron entrenar modelos de redes neuronales convolucionales con una capacidad de generalización razonable, obteniendo resultados aceptables en tareas de clasificación multiclase del confort térmico.

En cuanto al rendimiento de los modelos, los mejores resultados se obtuvieron con la arquitectura EfficientNet sobre imágenes segmentadas, destacando su eficiencia en comparación con otras redes como Xception, que, aunque presentaron métricas similares, requirieron tiempos de cómputo significativamente superiores. Esta ventaja convierte a EfficientNet en una opción más adecuada para aplicaciones en tiempo real o entornos con recursos computacionales limitados.

De cara a futuros trabajos, se plantea como objetivo principal la ampliación del conjunto de datos, lo cual permitirá entrenar modelos más robustos y reducir el impacto de la variabilidad individual. Esto implicará no solo un aumento en el número de participantes, sino también la inclusión de imágenes capturadas en diversas posturas y desde distintos ángulos de visión. También se plantea explorar arquitecturas avanzadas como los *Vision Transformers* y desarrollar modelos multimodales que integren información visual, climática, fisiológica y subjetiva, con el objetivo de lograr una estimación más precisa y personalizada del confort térmico.

En conjunto, los resultados obtenidos son prometedores, especialmente considerando las limitaciones del dataset y la complejidad asociada a la alta variabilidad individual del confort térmico. Este trabajo sienta una base sólida para el desarrollo de sistemas inteligentes de monitorización basados en visión por computador y aprendizaje profundo.

#### Agradecimientos

Este trabajo ha sido realizado como parte del proyecto de I+D+i PID2021-126889OB-I00 financiado por MCI-N/AEI/10.13039/501100011033 y por la "Unión Europea" y del proyecto de I+D+i TED2021-131655B-I00, financiado por MCIN/AEI/10.13039/501100011033 y "Unión Europea Next-GenerationEU".

#### Referencias

- ASHRAE, 2009. ASHRAE Handbook Fundamentals. Refrigerating American Society of Heating and Air-Conditioning Engineers.
- ASHRAE, 2023. Ansi/ashrae standard 55-2023 thermal environmental conditions for human occupancy. Standard.
- Baek, J., Park, D. Y., Park, H., Le, D. M., Chang, S., 2023. Vision-based personal thermal comfort prediction based on half-body thermal distribution. Building and Environment 228, 109877.
  - DOI: 10.1016/j.buildenv.2022.109877
- Chan, T., Vese, L., 2001. Active contours without edges. IEEE Transactions on Image Processing 10 (2), 266–277. DOI: 10.1109/83.902291
- Chaudhuri, T., Soh, Y. C., Li, H., Xie, L., 2019. A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings. Applied Energy 248, 44–53. DOI: 10.1016/j.apenergy.2019.04.065
- Cosma, A. C., Simha, R., 2019. Machine learning method for real-time non-invasive prediction of individual thermal preference in transient conditions. Building and Environment 148, 372–383.
  DOI: 10.1016/j.buildenv.2018.11.017
- Delzendeh, E., Wu, S., Lee, A., Zhou, Y., 2017. The impact of occupants' behaviours on building energy analysis: A research review. Renewable and Sustainable Energy Reviews 80, 1061–1071. DOI: 10.1016/j.rser.2017.05.264
- FLIR Systems, 2024. Picking a thermal color palette. https://www.flir. es/discover/industrial/picking-a-thermal-color-palette/, Último acceso: abril de 2025.
- Goh, T., Ang, B., Xu, X., 2018. Quantifying drivers of co2 emissions from electricity generation – current practices and future extensions. Applied Energy 231, 1191–1204.
- DOI: 10.1016/j.apenergy.2018.09.174
- Gonzalez, R. C., Woods, R. E., 2008. Digital Image Processing, 3rd Edition. Pearson Education.
- Huang, G., Li, D., Ng, S., Wang, L., Zhang, Y., 2025. Vision-based personal thermal comfort modeling under facial occlusion scenarios. Energy and Buildings 335, 115566.

DOI: 10.1016/j.enbuild.2025.115566

- ISO7730, 2023. Moderate thermal environments. determination of the PMV and PPD indices and specification of the conditions for thermal comfort.
- Seelaboyina, R., Vishwakarma, R., 2023. Different thresholding techniques in image processing : A review. In: Kumar, A., Senatore, S., Gunjan, V. K. (Eds.), ICDSMLA 2021. Springer Nature Singapore, Singapore, pp. 23–29.
- UNEP, 2021. 2021 global status report for buildings and construction: Towards a zero-emission, efficient and resilient buildings and construction sector. Tech. rep., United Nations Environment Programme, Nairobi, disponible en www.globalabc.org. Último acceso: abril de 2025.
- UNEP, 2025. Not just another brick in the wall: The solutions exist scaling them will build on progress and cut emissions fast. Tech. rep., United Nations Environment Programme, Nairobi, global Status Report for Buildings and Construction 2024/2025. Disponible en https://wedocs.unep.org/ 20.500.11822/47214. Último acceso: abril de 2025.