

Supporting Information

Deep learning-assisted active metamaterials with heat-enhanced thermal transport

Peng Jin, Liujun Xu, Guoqiang Xu, Jiaxin Li, Cheng-Wei Qiu, and Jiping Huang**

Peng Jin, Jiping Huang

Department of Physics, State Key Laboratory of Surface Physics, and Key Laboratory of Micro and Nano Photonic Structures (MOE), Fudan University, Shanghai 200438, China
E-mail: jphuang@fudan.edu.cn

Liujun Xu, Guoqiang Xu, Jiaxin Li, Cheng-Wei Qiu

Department of Electrical and Computer Engineering, National University of Singapore,
Singapore 117583, Singapore
E-mail: chengwei.qiu@nus.edu.sg

Liujun Xu

Graduate School of China Academy of Engineering Physics, Beijing 100193, China

Section S1: Physical mechanism of $|\nabla T_1|/|\nabla T_0|$ as a constant

The temperature gradient in the core and background regions varies with the external temperature gradient, while the ratio of the former to the latter is constant, regardless of the external temperature gradient. Here, the external temperature gradient is defined as $(T_H - T_C)/L$, which is uniform in the background region because the metamaterial works without disturbing the background temperature field. See the schematic in Figure 1b in the main text. To understand the above, we give the analytic expression of the temperature gradient ratio when considering an anisotropic monolayer structure (bilayer structure can also be regarded as the anisotropic monolayer case)

$$\frac{|\nabla T_1|}{|\nabla T_0|} = \frac{-4R_1^{m_1-1}R_2^{m_1+1}m_1\kappa_{rr}\kappa_3}{R_1^{2m_1}(-\kappa_1+m_1\kappa_{rr})(m_1\kappa_{rr}-\kappa_3)-R_2^{2m_1}(\kappa_1+m_1\kappa_{rr})(m_1\kappa_{rr}+\kappa_3)}, \quad (\text{S1})$$

where R_1, R_2 are the inner and outer radii of the anisotropic layer. κ_1, κ_3 are the thermal conductivity of the core and background regions. $\kappa_{rr}, \kappa_{\theta\theta}$ are the radial and tangential thermal conductivity of the anisotropic layer. $m_1 = \sqrt{\kappa_{\theta\theta}/\kappa_{rr}}$. As we can see, the ratio indeed doesn't include the term of $(T_H - T_C)/L$ (derived from the continuity equation of heat flow^[1]). The temperature gradient ratio can reflect the temperature field modulation effect of the metamaterial on the external temperature field, therefore representing thermal functionality (ratio < 1 for cloaking; ratio > 1 for concentration).

Section S2: Dataset collection and training of ANN

For collecting the dataset, we perform finite-element simulations with commercial software “COMSOL MULTIPHYSICS” linking MATLAB. Thanks to the bilayer structure's high performance without disturbing the background's thermal fields, we replace the bilayer structure with a pure background (size: $200 \times 200 \text{ mm}^2$) for convenience of dataset collection. In Section S3, we elaborate the rationality of replacing the bilayer structure with a pure background for dataset collection. After setting the left boundary as a high temperature (T_H) and the right boundary as a low temperature (T_C), we could extract the temperature information $T^{(0)}$ at the circle with a radius of $R_3 = 60 \text{ mm}$ from the finite-element simulation results. Drawing parallels from nonlinear optics^[2]-specifically, the way polarizability or electrical conductivity is influenced by electric intensity-we introduced the concept of a configurable nonlinear thermal material whose effective thermal conductivity [equivalent to rotating velocity $\omega_{1,\text{Target}}(\nabla T)$] being responsive to external temperature gradients ∇T . For a predefined scale, these external temperature gradients can be simplified to the external temperature difference, given a consistent interval L between the hot and cold boundaries. This insight was crucial in

formulating the corresponding labeled $\omega_{1,\text{Target}}$, laying the foundation for our configurable nonlinear thermal material. Therefore, we artificially set the corresponding labeled $\omega_{1,\text{Target}}$ as

$$\omega_{1,\text{Target}}(\Delta T) = \begin{cases} 0, & \Delta T > \frac{4(T_{\text{H,max}} - T_{\text{C,min}}) + \Delta T_0}{5} \\ \omega_{\text{tran}} \times \varphi_1(\Delta T), & \frac{2(T_{\text{H,max}} - T_{\text{C,min}}) + 3\Delta T_0}{5} < \Delta T \leq \frac{4(T_{\text{H,max}} - T_{\text{C,min}}) + \Delta T_0}{5} \\ (\omega_{\text{max}} - \omega_{\text{tran}}) \times \varphi_2(\Delta T) + \omega_{\text{tran}}, & \Delta T \leq \frac{2(T_{\text{H,max}} - T_{\text{C,min}}) + 3\Delta T_0}{5} \end{cases} \quad (\text{S2})$$

where $\Delta T = T_{\text{H}} - T_{\text{C}}$ is the temperature difference between external hot and cold sources, T_{C} went from $T_{\text{C,min}} = 283$ K to $T_{\text{C,max}} = T_{\text{C,min}} + \Delta T_0$ K, and T_{H} went from $T_{\text{H,min}} = T_{\text{C}} + \Delta T_0$ K to $T_{\text{H,max}} = 313$ K. $\varphi_1(\Delta T) = \frac{9 - 9.9 \times \frac{\Delta T - 10}{20}}{5}$ and $\varphi_2(\Delta T) = \frac{4 - 9.9 \times \frac{\Delta T - 10}{20}}{4}$ are the antilinear modulation functions as empirical formulas. Such form ensures that smaller external temperature difference ΔT outputs larger rotating velocity $\omega_{1,\text{Target}}(\Delta T)$, thus leading to smaller target region's temperature gradient. Here, ΔT_0 is the minimum temperature difference between external hot and cold sources (artificially given) when generating the dataset. We take $\Delta T_0 = 10$ K, transitional and maximal rotating velocity as $\omega_{\text{tran}} = 0.001$ rad s⁻¹ and $\omega_{\text{max}} = 0.12$ rad s⁻¹, respectively. After iterating through all the high and low temperature values, the entire structure was rotated 1° around the center. The whole dataset was generated until the entire structure was rotated 360°. During dataset collection, the temperatures at the left and right boundaries are the variables to calculate the labeled $\omega_{1,\text{Target}}$ without using the temperature information $\mathbf{T}^{(0)}$. The reason is that the temperatures at the left (T_{H}) and right (T_{C}) boundaries are indeed vital determinants of $\omega_{1,\text{Target}}(\Delta T)$. However, the external temperature settings (T_{H} and T_{C}), which we utilize in Eq. (S2) are effectively mirror to the surrounding temperature values at R_3 [$\mathbf{T}^{(0)}$]. This means that when we reference the external temperature settings, we are indirectly accounting for the temperature information at R_3 . By substituting these external temperature settings ($\Delta T = T_{\text{H}} - T_{\text{C}}$) into the Eq. (S2), we can directly obtain the target value for $\omega_{1,\text{Target}}(\Delta T)$. This approach allows us to maintain the fidelity of our model without explicitly referencing $\mathbf{T}^{(0)}$ in the equation. While the temperature values at R_3 may not appear explicitly in the equation, its effects are inherently captured through the external temperature settings. This makes the model both efficient and representative of the physical system. The modulation impact of Eq. (S2) on the device's functionality is showcased in Figure 1c of the main text.

Choosing the right hyperparameters for an artificial neural network is crucial for its performance. We decide commonly tuned hyperparameters, including learning rate (determines

the step size when updating the weights and biases), batchsize (number of training examples utilized in one iteration), batches (a batch is a subset of the training dataset, and the number of batches is determined by dividing the total number of training samples by the batch size), number of epochs (defines how many times the algorithm will cycle through the entire training dataset), number of hidden layers, and number of neurons per layer.^[3-5] A popular method for hyperparameter tuning is grid search, where we specify a discrete set of values for each hyperparameter and evaluate every possible combination. Instead of evaluating all combinations, we perform a random search for hyperparameters from predefined distributions and evaluate a fixed number of these combinations. The rationale is that not all hyperparameters are equally important, and random search can efficiently explore the space with fewer trials. After all iterations, we select the hyperparameters that gave the best performance on the testing set, with an accuracy rate of over 90%. We define a cost function C to evaluate the performance of the ANN. We utilize the backpropagation algorithm to minimize the cost of the ANN. The backpropagation algorithm is a fundamental optimization method for training artificial neural networks (ANNs). It's a supervised learning algorithm that allows the network to adjust its weights or biases based on the error of its predictions. To be more understandable, we have provided a flow chart (**Figure S1**) to illustrate the backpropagation process in the ANN. The input layer receives the initial data. A forward pass takes the data through the hidden layers and produces an output in the output layer. The network's output is compared to the actual target to the cost function, which is then used in the backward pass to compute the derivative of the cost function. Based on the derivative and the learning rate (whose mathematical expression is the chain rule), the weights and biases are adjusted to minimize the cost function. This process repeats until convergence or for a specified number of epochs.

We give the cost function as $C = \frac{1}{N} \sum_{i=0}^{N-1} C_i = \frac{1}{N} \sum_{i=0}^{N-1} [\text{ReLU}(\omega_1) - \text{ReLU}(\omega_{1,\text{Target}})]^2$, and train the parameters (weights and biases) of the ANN via the backpropagation algorithm (based on the generated dataset, or say, training samples). Here, C_i is a single cost function corresponding to a single training sample, and N is the total number of training samples. Now let's review the rules of connection between neurons and give the following mathematical symbols:

$$C_i = [\text{ReLU}(\omega_1) - \text{ReLU}(\omega_{1,\text{Target}})]^2, \quad (\text{S3a})$$

$$z_j^{(L)} = \sum_{k=0}^{n_{L-1}-1} \omega_{jk}^{(L)} h_k^{(L-1)} + b_j^{(L)}, \quad (\text{S3b})$$

$$h_j^{(L)} = \text{ReLU}[z_j^{(L)}], \quad (\text{S3c})$$

where $h_k^{(L-1)}$ and $h_j^{(L)}$ are activations of the k-th and j-th neuron in the $(L-1)$ -th and L -th layer of the ANN, respectively. $\omega_{jk}^{(L)}$ is the weight between k-th neuron in the $(L-1)$ -th layer and j-th neuron in the L -th layer. $b_j^{(L)}$ is the bias for the j-th neuron in the L -th layer. n_{L-1} is the total number of neurons in the $(L-1)$ -th layer. In particular, $z_0^{(5)}$ is ω_1 in our case. Next, we calculate the derivative of C_i to the parameters of the ANN, which is the core of backpropagation algorithm. Concrete chain rules are as follows:

$$\frac{\partial C_i}{\partial \omega_{jk}^{(L)}} = \frac{\partial z_j^{(L)}}{\partial \omega_{jk}^{(L)}} \frac{\partial h_j^{(L)}}{\partial z_j^{(L)}} \frac{\partial C_i}{\partial h_j^{(L)}}, \quad (\text{S4a})$$

$$\frac{\partial C_i}{\partial b_j^{(L)}} = \frac{\partial z_j^{(L)}}{\partial b_j^{(L)}} \frac{\partial h_j^{(L)}}{\partial z_j^{(L)}} \frac{\partial C_i}{\partial h_j^{(L)}}, \quad (\text{S4b})$$

$$\frac{\partial C_i}{\partial h_k^{(L-1)}} = \sum_{j=0}^{n_{L-1}} \frac{\partial z_j^{(L)}}{\partial h_k^{(L-1)}} \frac{\partial h_j^{(L)}}{\partial z_j^{(L)}} \frac{\partial C_i}{\partial h_j^{(L)}}, \quad (\text{S4c})$$

$$\frac{\partial C_i}{\partial h_j^{(L)}} = \sum_{j=0}^{n_{L+1}-1} \frac{\partial z_j^{(L+1)}}{\partial h_j^{(L)}} \frac{\partial h_j^{(L+1)}}{\partial z_j^{(L+1)}} \frac{\partial C_i}{\partial h_j^{(L+1)}}. \quad (\text{S4d})$$

Based on this chain rule, weights and biases are updated when all training samples are cycled and cycled. Finally, the cost function is optimized to the minimum based on the backpropagation algorithm, all parameters are iterated to fit, and the performance of the ANN achieves the best.

The proposed ANN with selected hyperparameters (**Figure S2a**) is trained with 114418 labeled samples, and all the weights and biases are adjusted to their most appropriate values based on the backpropagation algorithm. We choose 10% of all samples as testing samples, and the accuracy of ANN reaches up to 92.7% (Figure S2b). Here, if the relative deviation between the output result from ANN and the target result is within 20%, the output result is accurate. Also, a mean square error has a good convergence effect; see Figure S2c.

Section S3: The rationality of replacing the bilayer structure with a pure background for dataset collection

We should prepare a set of input data and their targeted output data for collecting the dataset. In our case, the input data is the temperature values of the bilayer structure's surroundings (circle R_3) along the circumference in a counterclockwise direction from finite-element simulations. Due to the need for a large amount of training samples, we choose a background with simpler parameters as a replacement to fast collect these samples. The replaceable basis is

that this bilayer structure works without disturbing the background thermal field. To elaborate, we have added the comparison of temperature results from the two cases in different external temperature and rotating velocity settings. As shown in **Figure S3**, these temperature values from the background structure are precisely consistent with those of the bilayer structure. If the data is collected by including the bilayer structure, we get the same data with the case of the background.

Section S4: Effect on background thermal fields of the heat-enhanced thermal diffusion metamaterial

In the main text, we present the performance of the heat-enhanced thermal diffusion metamaterial in Figure 2, which shows enlarged-range temperature-gradient distributions in the core region under different ambient temperatures. To further showcase the effect on the background thermal fields, we plot temperature distributions on two lines of the bilayer structure in **Figure S4a** (from finite-element simulation results in Figure 2b). Here, Line 1 is tangent to the outer contour of the outer layer of the bilayer structure, and Line 2 passes through the center of the bilayer structure; see Figure S4a. We can see from Figure S4b that temperature data distributed along x direction on Line 2 has an enlarged-range inclination change in the core region compared to Line 1. Also, background temperature data on Line 1 and Line 2 attenuate uniformly along the direction of heat transfer, indicating undisturbed background thermal fields. We still demonstrate the effect for all three rotating velocities and all four directions of external thermal fields (from finite-element simulation results in Figure 2). **Figure S5** displays temperature-gradient distributions in the background in the above cases, which are uniform under each rotating velocity and each direction of the external thermal field. These results indicate the heat-enhanced thermal diffusion metamaterial works well without affecting the background thermal fields. **Figure S6** exhibits the original temperature distributions in the core region of the bilayer structure in all cases of Figure 2.

Section S5: Response of heat-enhanced thermal diffusion metamaterial under different cold sources and non-uniform external thermal fields

The heat-enhanced thermal diffusion metamaterial also works well in other cold sources. In the following simulations, the hot source ($T_H = 313$ K) is fixed, while the cold source (T_C) is varied. We extract the temperature data ($T_a^{(0)}, T_b^{(0)}, T_c^{(0)}$) of $N = 36$ equally spaced positions in the yellow dashed circle in three cases ($T_C = 303, 293, 283$ K) of a static bilayer structure, as shown in **Figure S7a**. In each case, the first data point corresponds to the temperature at the position

marked 0° in the dashed circle. The temperature of the other positions is taken counterclockwise, forming the input layer of the pre-trained ANN. Hence, the relevant rotating velocities of the PDMS ω_1 are calculated as 0.109, 0.00079, 0 rad s⁻¹, set in finite-element simulations. Finally, the temperature profiles (modulated T) with three ω_1 under three cold sources are shown in Figure S7b-d and the corresponding temperature profiles of static pure background (original T_0) are depicted in Figure S7e-g. To characterize the effect of the metamaterial on the original background thermal fields, we compute the difference between the temperature distributions in Figure S7b-d and those in Figure S7e-g; see Figure S7h-j. As expected, the temperature fields in the core region are modulated effectively. Notably, the temperature-difference distributions in the background are almost zero, and the background thermal fields are not disturbed when the heat-enhanced thermal diffusion metamaterial is operating. We then extract temperature-gradient distributions in the core region from the results of Figure S7b-g. The original temperature-gradient distributions in Figure S7k (from Figure S7e-g) are modulated as results in Figure S7l (from Figure S7b-d), enlarging the temperature-gradient tunability in the core region. These results demonstrate the performance of the heat-enhanced thermal diffusion metamaterial under different cold sources.

We further consider the performance of the heat-enhanced thermal diffusion metamaterial in non-uniform external thermal fields. Keeping the invariant sizes and components of the bilayer structure (Figure S4), we set two circle sources in the left (hot source), and right (cold source) parts, whose radii are both 0.003 m and positions of the center of the circles are at (± 0.09 , 0) m. The cold source is fixed at 283 K, and we vary the hot source to 293, 303, and 313 K to create different external temperature conditions. We simulate the system using COMSOL MULTIPHYSICS, setting the background boundaries as thermal insulation. We collect the temperature data $T_a^{(0)}$, $T_b^{(0)}$, $T_c^{(0)}$ of the static bilayer structure's surroundings under three external temperature conditions; see **Figure S8a**. By inputting $T_a^{(0)}$, $T_b^{(0)}$, $T_c^{(0)}$ into the pre-trained ANN, the relevant rotating velocities of the PDMS ω_1 are calculated as 0.127, 0.00097, 0 rad s⁻¹, used in finite-element simulations. Finally, the temperature profiles (modulated T) with three ω_1 under three hot sources are shown in Figure S8b-d and the corresponding temperature profiles of static pure background (original T_0) are depicted in Figure S8e-g. An intuitive phenomenon is the tunable uniformity of temperature distributions in the core region of Figure S8b-d, compared with relevant results in Figure S8e-g. To examine the effect of the heat-enhanced thermal diffusion metamaterial on the original background thermal fields, we calculate the temperature difference between Figure S8b-d and Figure S8e-g; see Figure S8h-j. Notably, the temperature-difference distributions in the background are

almost zero, making known undisturbed background thermal fields when the heat-enhanced thermal diffusion metamaterial works. Subsequently, we extract temperature-gradient distributions in the core region from the results of Figure S8b-g. The original temperature-gradient distributions shown in Figure S8k (from Figure S8e-g) are modulated in Figure S8l (from Figure S8b-d), enlarging the temperature-gradient tunability in the core region. These results demonstrate that the heat-enhanced thermal diffusion metamaterial is effective in non-uniform external thermal fields. We further investigate the cases of three or four circle heat sources at random locations. These results in **Figure S9** and **Figure S10** demonstrate that the pre-trained ANN maintains robust performance, to some extent, on intelligently modulating the temperature fields, even when working in non-uniform external thermal fields. Specifically, smaller (larger) external temperature differences bring about larger (smaller) rotating velocities, resulting in smaller (larger) temperature gradients in the core region.

Section S6: Weak direction-dependence of the metamaterial's performance

In this section, we discuss the weak direction-dependence of the metamaterial's performance. The designed metamaterial comprises isotropic materials with isotropic geometry. Notably, every time the micro infrared camera operates, it starts capturing temperature data from a designated initial position (marked 0°), regardless of the direction from which external thermal energy originates. The data is read sequentially in a counterclockwise direction and then fed into the ANN as input data, symbolized as $T^{(0)}$. While the external thermal flow's direction changes, the extracted temperature values as input data are different for neurons of the input layer of the ANN, according to the given measurement rules. Even though there's inconsistency in the $N=36$ temperature data points for the input layer when varying directions of external thermal flow, training the ANN with the dataset ensures that the metamaterial exhibits stable thermal responses to external thermal fields, regardless of their directionality. Specifically, the deviation between the output results under different external heat flow's direction and the target result is less than 20%, ensuring a similar temperature field modulation effect (please see results in **Figure S11**). Here, if the relative deviation between the output result from ANN and the target result is within 20%, the output result is accurate. For clarity, we have provided a flow chart to illustrate this "direction-dependency" (Figure S11), reflected in the differences among input temperature values under different external heat flow's directions.

Section S7: Principle of the thermal signal modulator

In this section, we introduce the principle of the thermal signal modulator. Suppose Alice and Bob transmitted information to each other using a heat communication technology. They agreed on a rule for encoding information, where they prepare signal waveforms by presenting the temperature gradient distributions in the encoding zones of arranged thermal metadevices. For clarity, we discuss the mechanism of how one single thermal signal modulator works. See the schematic in **Figure S12a**. A common metadevice works under external temperature settings hot source T_A and cold source T_B , preparing an original temperature gradient signal in the encoding zone (the dashed box). Due to some thermal noise, original temperature gradient signals are sometimes enhanced or suppressed, whose oscillation amplitude is suppressed. For the accuracy of original signals, thermal signal modulators take effect, making thermal signals oscillate within a larger temperature gradient amplitude. For example, original temperature gradient signals oscillate within the range of 50 to 150 K m⁻¹ in our case. The thermal signal modulator reads the surrounding temperature data at the yellow dashed circle. Then, the pre-trained ANN outputs the rotating velocity ω_1 , modulating the temperature gradient in the core region of the bilayer structure (modulated zone in Figure S12a). According to the pre-trained ANN, smaller (larger) external temperature differences bring about larger (smaller) rotating velocities, resulting in smaller (larger) temperature gradients in the modulated zone. In our case, the original temperature gradient 50 (150) K m⁻¹ is prepared for the ANN, outputting the $\omega_1 = 0.12$ (0) rad s⁻¹. Finally, the modulated temperature gradient is 0 (238.5) K m⁻¹. We have realized the modulation of original thermal signals, making these signals oscillate within a larger range of 0 to 238.5 K m⁻¹.

Section S8: Effect of outer shape on the metamaterial's performance

Recently, researchers have designed thermal functional devices with an outer contour arbitrary shape, including the bilayer scheme (this requires non-uniform/anisotropic material distribution in the outer-layer metamaterial region while easy to prepare through 3D printing technology), based on topology optimization methods,^[6-7] transformation thermotics,^[8] or the conservation equation of heat flow.^[9] Our work can be extended to scenes requiring complex shapes, while maintaining highly tunable performance.

Specifically, we consider a bilayer structure whose outer and middle contours are replaced with square shape, and the side length of the outer and middle square are denoted by L_{outer} and L_{inner} , respectively; see **Figure S13a**. Minor changes are made to the size parameters, where we set $L_{outer} = 0.12$ m, $L_{outer} - L_{inner} = 2d = 0.042$ m. Here, d is the thickness of the outer layer. Other parameters are the same as cases in the main text. We first

divide the outer-layer metamaterial region into two parts with thermal conductivity of $\kappa_1 = \text{diag}(\kappa_{1,xx}, \kappa_{yy})$ and $\kappa_2 = \text{diag}(\kappa_{2,xx}, \kappa_{yy})$, respectively; see Figure S13a. According to the conservation equation of heat flow,^[9] we get the thermal conductivity of the outer-layer metamaterial as $\kappa_1 = \text{diag}(9.8, 300)$, $\kappa_2 = \text{diag}(9.8 \times \frac{L_{\text{outer}}}{L_{\text{outer}} - L_{\text{inner}}}, 300) \text{ W m}^{-1} \text{ K}^{-1}$, respectively. Then, we perform finite-element simulations to show the tunable performance of the designed device. Figure S13b-d show that the central temperature gradient can be tuned within a large range, reflected by the uniformity of temperature distributions. Meanwhile, the background thermal fields are not disturbed.

Section S9: Effect of length scale on the metamaterial's performance

When discussing the effect of length scale on the metamaterial's performance, we'd better ensure that all other conditions remain unchanged, especially the external temperature gradient settings. Under this condition, when the dimensions are increased or decreased, surrounding temperature data becomes further (surrounding temperature data varies within a larger range than the current situation, corresponding to the larger external temperature difference case), or closer with each other (this corresponds to the smaller external temperature difference case), respectively. Therefore, if the dimensions of the bilayer regions are increased (or decreased), these conditions can be equivalently regarded as increasing the external temperature difference (or decreasing) of the bilayer regions. According to the pre-trained artificial neural network, smaller (larger) external temperature differences bring about larger (smaller) rotating velocities, resulting in smaller (larger) temperature gradients in the core region.

We further characterize the performance and verify the above analysis upon adjusting the systems' dimension based on the current pre-trained artificial neural network. In our simulations, the dimensions of the bilayer regions are equally decreased (1/10 times), unchanged, and increased (10 times) with length scale $L = 0.02, 0.2, 2 \text{ m}$, respectively. Further, the external temperature gradient is maintained at 100 K m^{-1} in three cases. External temperature settings are displayed in **Figure S14c**. Then, we extracted the temperature data ($T_a^{(0)}, T_b^{(0)}, T_c^{(0)}$) of $N = 36$ equally spaced positions in the yellow dashed circle in three cases ($L = 0.02, 0.2, 2 \text{ m}$) of a bilayer structure. As expected, surrounding temperature data $T_a^{(0)}$ [$T_c^{(0)}$] gets closer (further) with each other than the unchanged case $T_b^{(0)}$, as shown in Figure S14b. The relevant rotating velocities of the PDMS ω_1 are calculated as $0.18, 0.00066, 0 \text{ rad s}^{-1}$ with respective to $L = 0.02, 0.2, 2 \text{ m}$, set in finite-element simulations. Figure S14c exhibits the modulated temperature profiles of the bilayer structure with different sizes $L = 0.02, 0.2,$

2 m, respectively. It is noted that larger (smaller) size induces the larger (smaller) external temperature difference, making these temperature data from $T_c^{(0)}$ [$T_a^{(0)}$] further (closer) with each other. Finally, the pre-trained ANN outputs a smaller (larger) rotating velocity, generating a larger (smaller) temperature gradient in the core region.

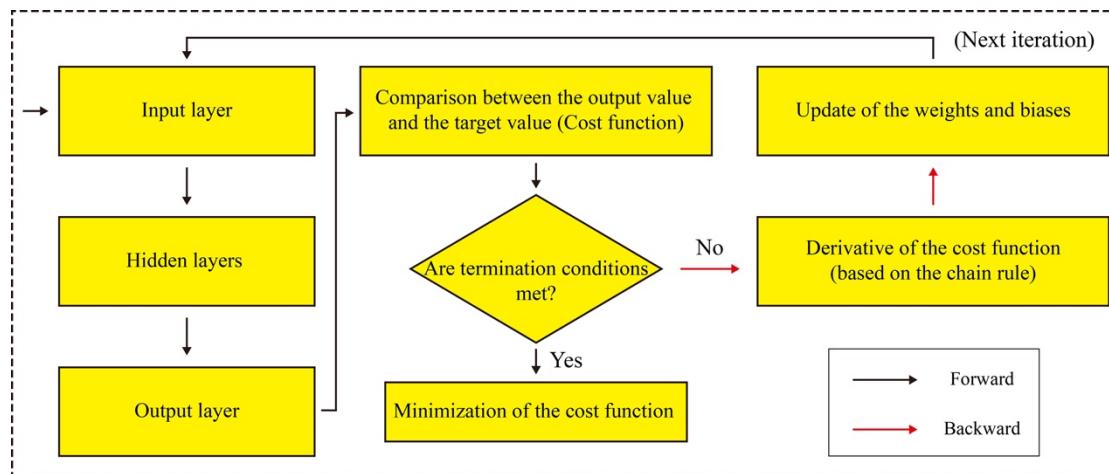


Figure S1. Flowchart of a backpropagation algorithm for minimizing the cost of the ANN.

a **Table 1:** Selection of hyperparameters.

Learning rate	Batchsize	Batches	Epochs	Hidden layers	Neurons per layer
0.005	200	572	100	4	50

b **Table 2:** Performance of ANN.

Training Samples	Testing Samples	Correct Samples	Accuracy
114418	12721	11792	92.7%

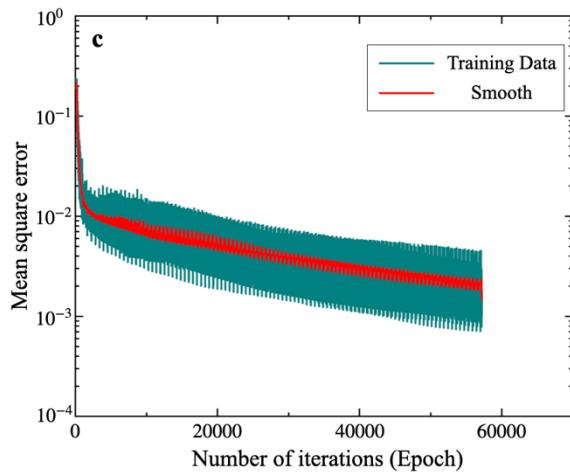


Figure S2. Hyperparameters and performance of ANN. a) Selection for hyperparameters of the ANN. b) The accuracy of the pre-trained ANN. c) Mean square error of the training data over the epoch.

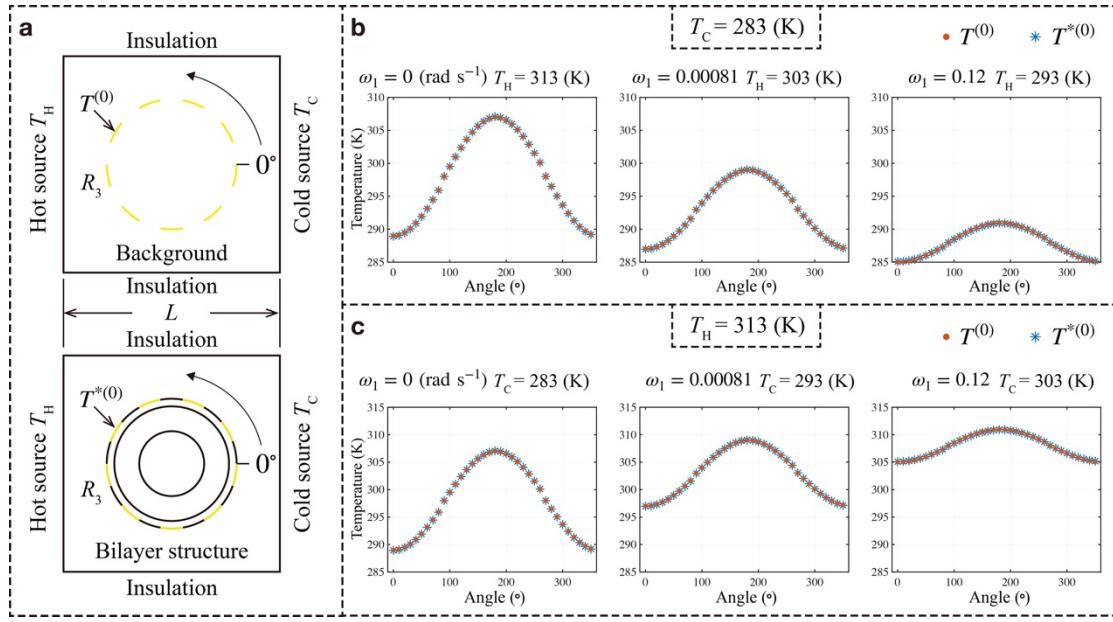


Figure S3. Comparison between $T^{(0)}$ from the pure background and $T^{*(0)}$ from bilayer structure. a) Schematic. b,c) $N = 36$ equally spaced temperature data $T^{*(0)}$ [$T^{(0)}$] at the yellow dashed circle in bilayer structures (pure background) under different external temperature conditions and rotating velocities.

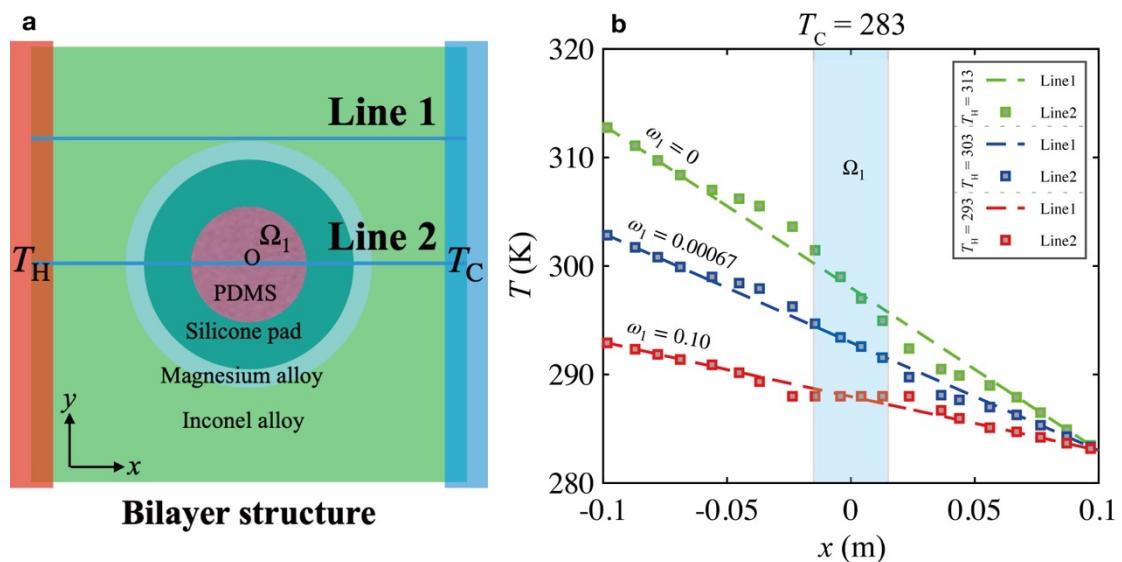


Figure S4. Temperature distributions on Line 1 and Line 2 of the bilayer structure (from finite-element simulation results in Figure 2b in the main text).

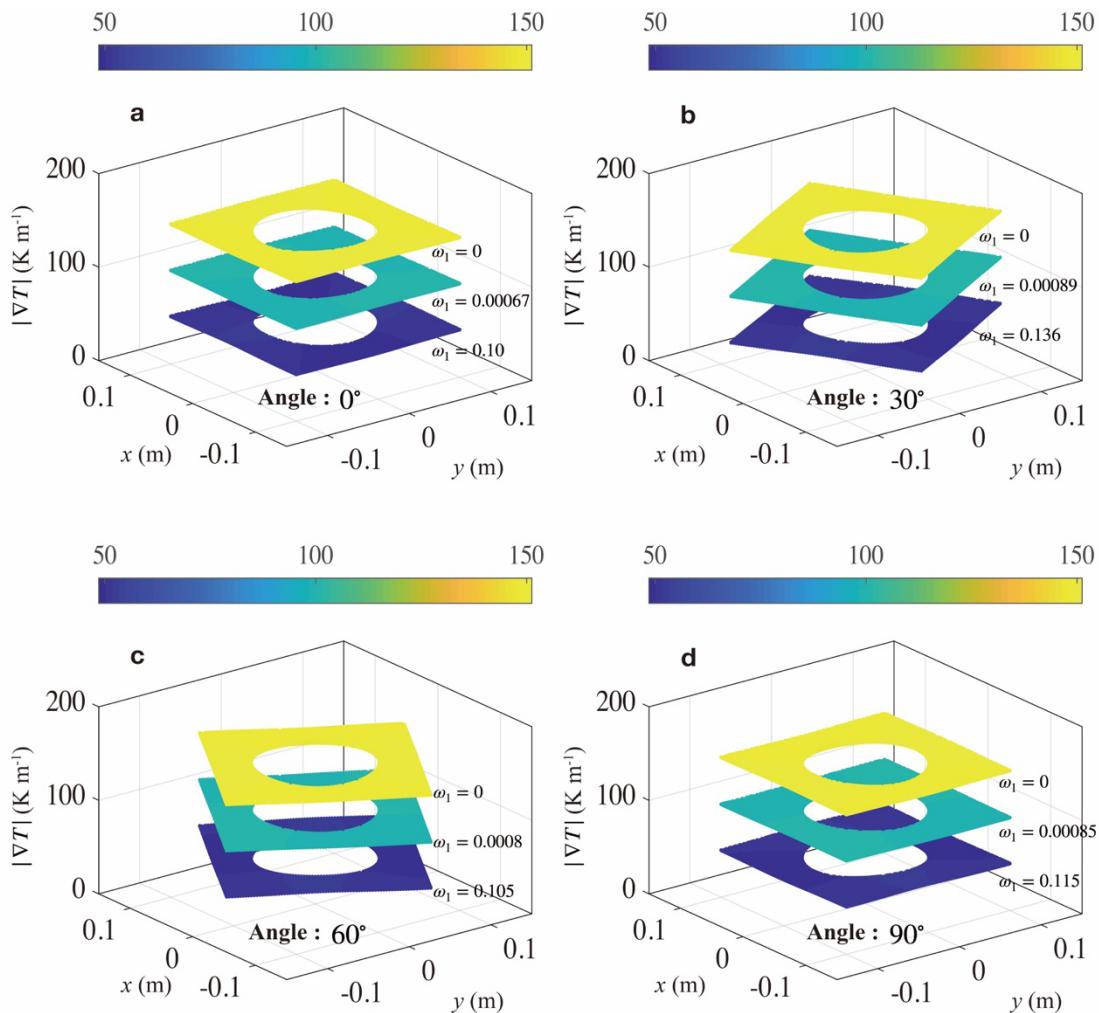


Figure S5. Temperature-gradient distributions in the background (Inconel alloy) of the bilayer structure (from finite-element simulation results in Figure 2b,e,h,k in the main text).

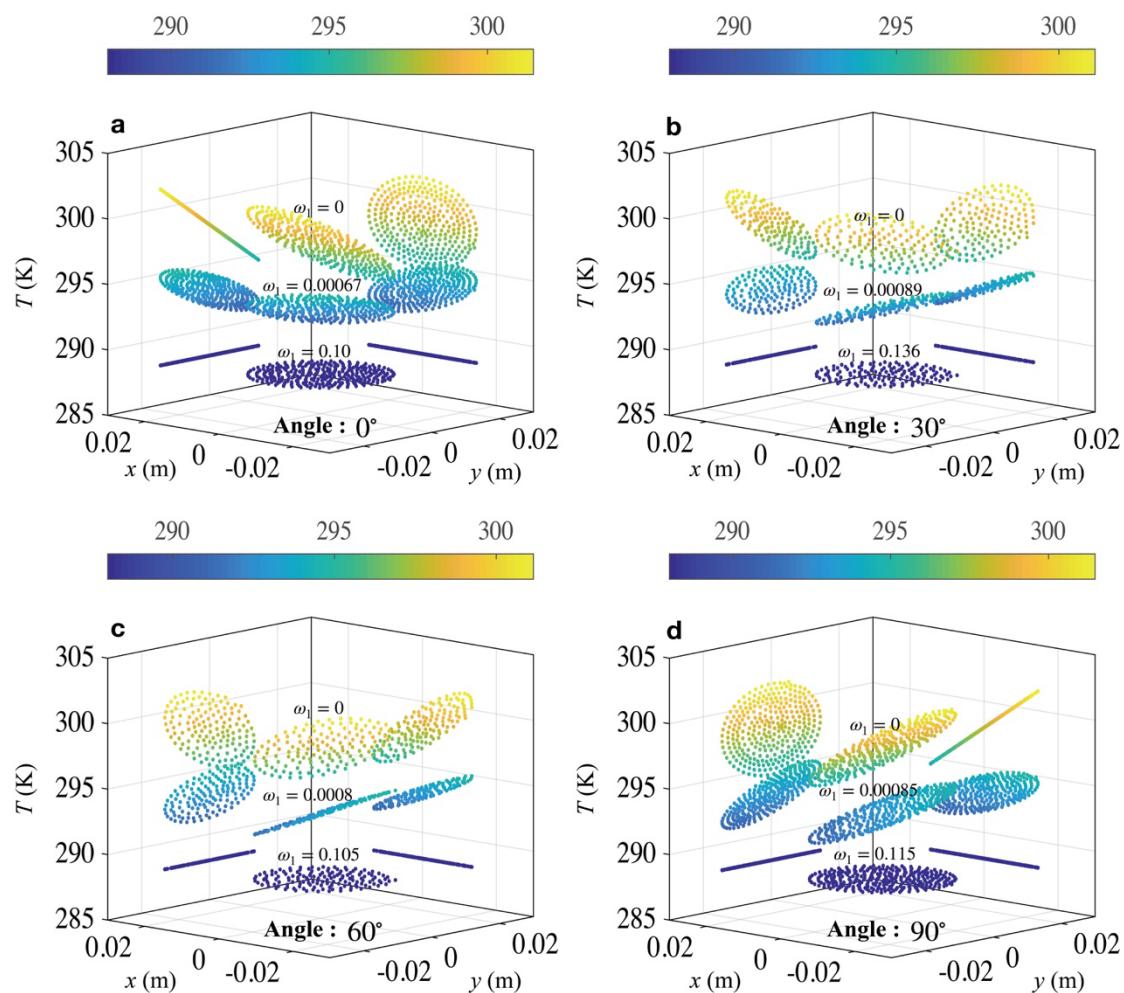


Figure S6. Temperature distributions in the core region (PDMS) of the bilayer structure (from finite-element simulation results in Figure 2b,e,h,k in the main text).

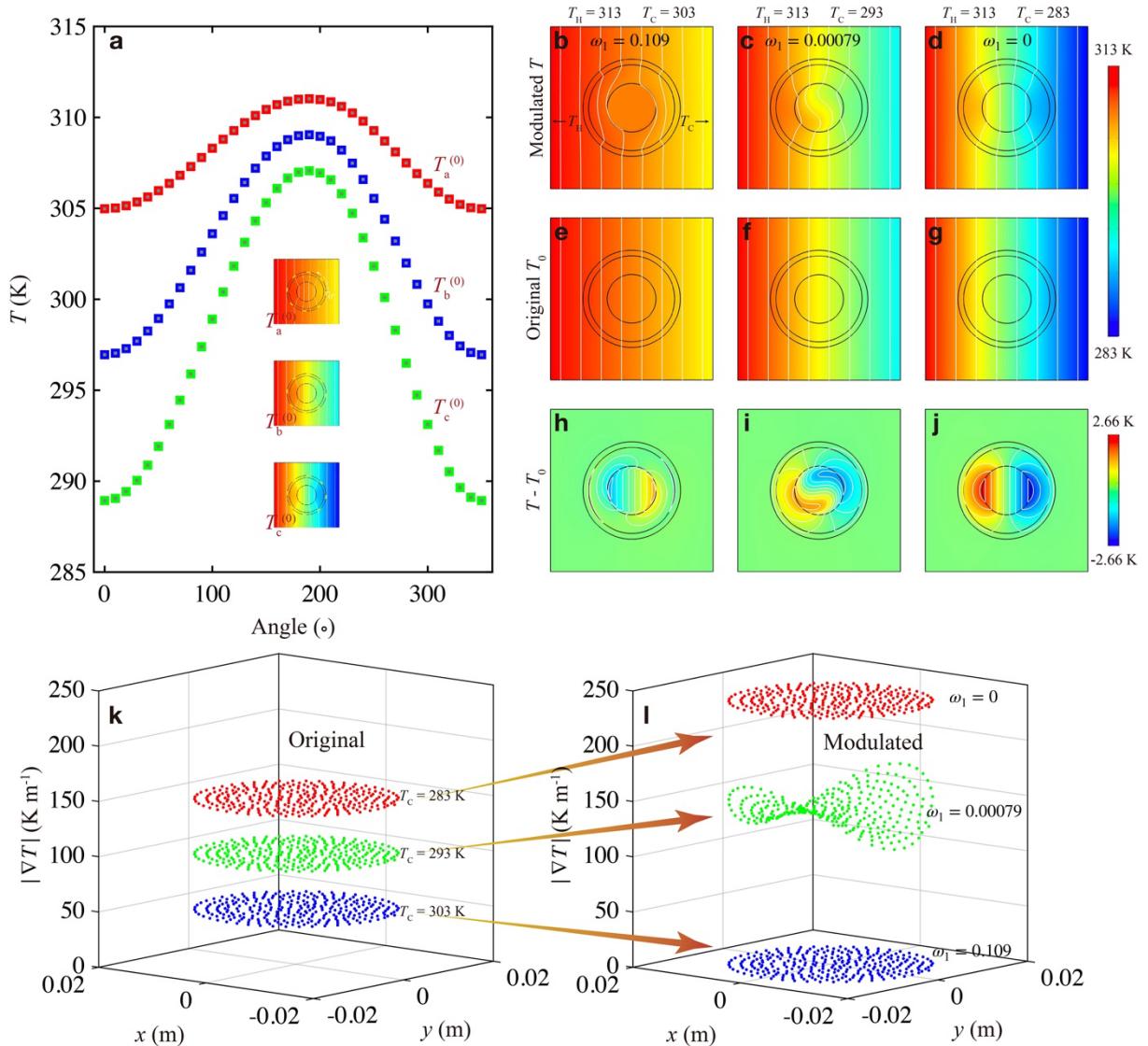


Figure S7. Response of heat-enhanced thermal diffusion metamaterial under different cold sources. The cold source is set as 303, 293, and 283 K, respectively. The hot source is fixed at 313 K. a) N = 36 equally spaced temperature data $T_a^{(0)}, T_b^{(0)}, T_c^{(0)}$ in the yellow dashed circle in static bilayer structures under three external temperature conditions. b-d) Temperature profiles of the bilayer structure with rotating velocity $\omega_1 = 0.109, 0.00079, 0$ rad s⁻¹, respectively. e-g) Temperature profiles of the static pure background with the three cold sources (the whole bilayer structure is made of Inconel alloy). h-j) Temperature-difference profiles between results in (b-d) and results in (e-g). k) Original temperature-gradient distributions in core region Ω_1 of the static pure background with the three cold sources. l) Modulated temperature-gradient distributions in core region Ω_1 of the bilayer structure with the three cold sources and $\omega_1 = 0.109, 0.00079, 0$ rad s⁻¹, respectively.

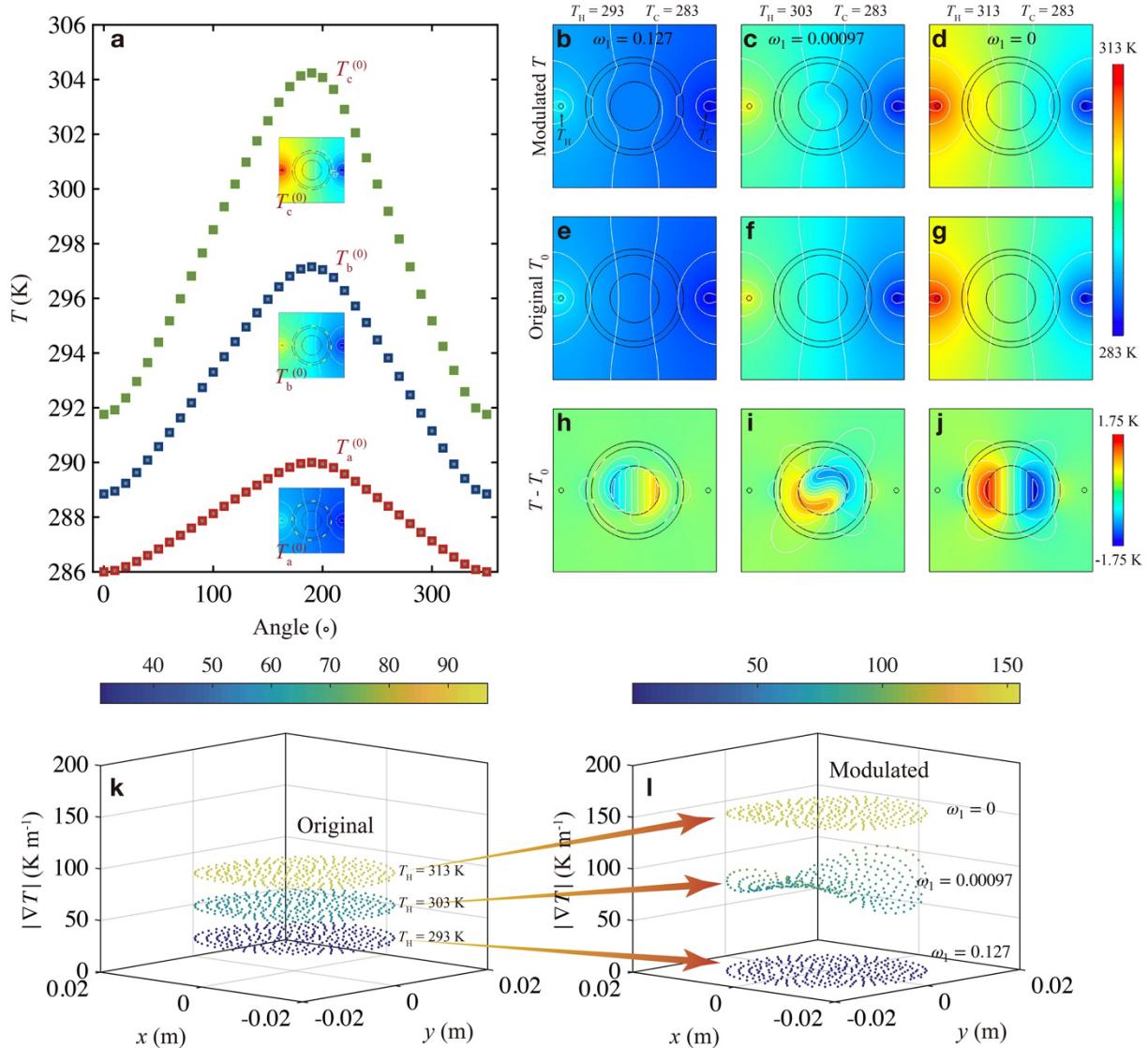


Figure S8. Response of heat-enhanced thermal diffusion metamaterial under non-uniform external thermal fields (circle sources are set in finite-element simulations). The hot source is set as 293, 303, and 313 K, respectively. The cold source is fixed at 283 K. a) $N = 36$ equally spaced temperature data $T_a^{(0)}$, $T_b^{(0)}$, $T_c^{(0)}$ in the yellow dashed circle in static bilayer structures under three external temperature conditions. b-d) Temperature profiles of the bilayer structure with rotating velocity $\omega_1 = 0.127$, 0.00097 , 0 rad s^{-1} , respectively. e-g) Temperature profiles of the static pure background with the three hot sources (the whole bilayer structure is made of Inconel alloy). h-j) Temperature-difference profiles between results in (b-d) and results in (e-g). k) Original temperature-gradient distributions in core region Ω_1 of the static pure background with the three hot sources. l) Modulated temperature-gradient distributions in core region Ω_1 of the bilayer structure with the three hot sources and $\omega_1 = 0.127$, 0.00097 , 0 rad s^{-1} , respectively.

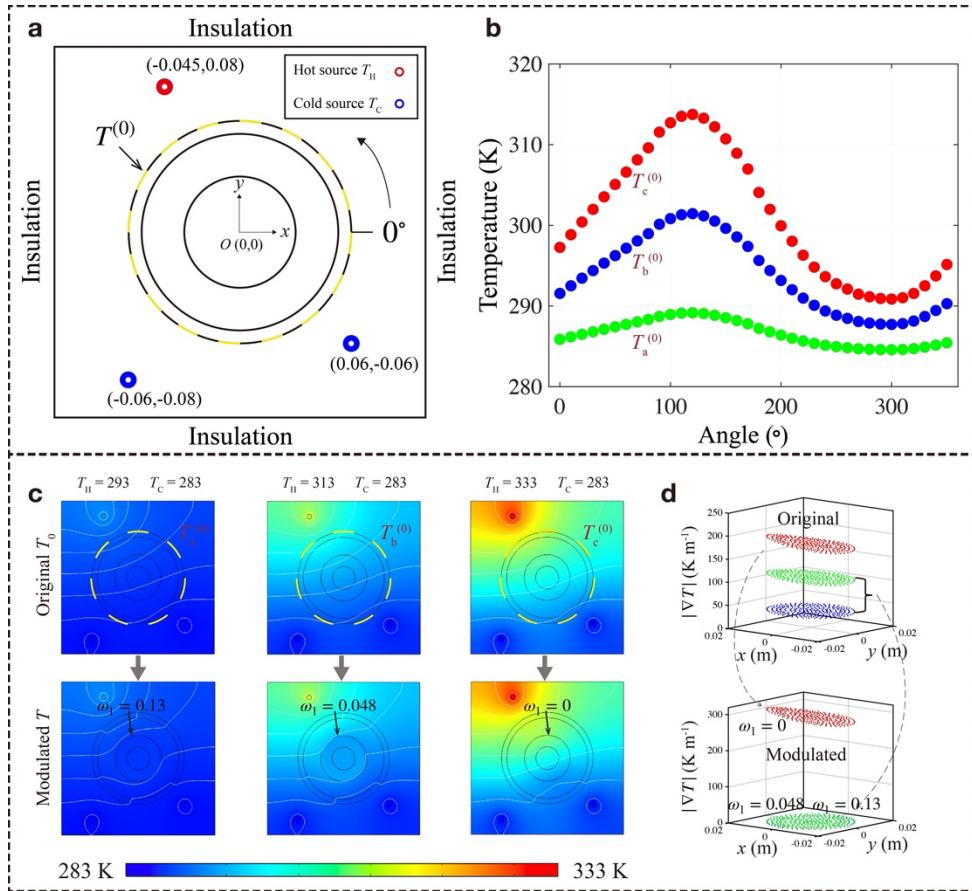


Figure S9. Effect of non-uniform external thermal fields on the metamaterial's performance (three circle sources are set randomly, with the same radii of 0.003 m). One hot source [whose position is at (-0.045, 0.08) m] is set as 293, 313, and 333 K, respectively. Two cold sources [whose positions are at (-0.06, -0.08) and (0.06, -0.06) m] are fixed at 283 K. a) Schematic. b) $N = 36$ equally spaced temperature data $T_a^{(0)}$, $T_b^{(0)}$, $T_c^{(0)}$ at the yellow dashed circle in bilayer structures under three external temperature conditions. c) Original (the whole bilayer structure is made of Inconel alloy) and modulated (rotating velocity $\omega_1 = 0.13$, 0.048, 0 rad s⁻¹, respectively) temperature profiles of the bilayer structure with three hot source settings. d) Original and modulated ($\omega_1 = 0.13$, 0.048, 0 rad s⁻¹, respectively) temperature-gradient distributions in the core region Ω_1 of the bilayer structure with the three hot sources.

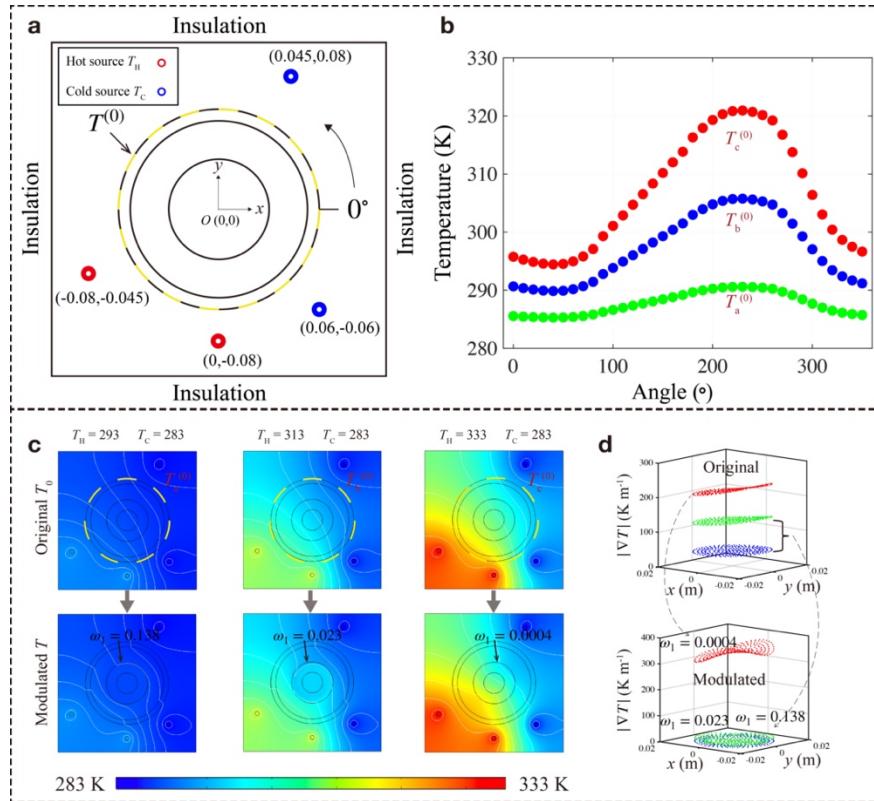


Figure S10. Effect of non-uniform external thermal fields on the metamaterial's performance (four circle sources are set randomly, with the same radii of 0.003 m). Two hot sources [whose positions are at (-0.08, -0.045) and (0, -0.08) m] are set as 293, 313, and 333 K, respectively. Two cold sources [whose positions are at (0.045, 0.08) and (0.06, -0.06) m] are fixed at 283 K. a) Schematic. b) $N = 36$ equally spaced temperature data $T_a^{(0)}$, $T_b^{(0)}$, $T_c^{(0)}$ at the yellow dashed circle in bilayer structures under three external temperature conditions. c) Original (the whole bilayer structure is made of Inconel alloy) and modulated (rotating velocity $\omega_1 = 0.138, 0.023, 0.0004$ rad s⁻¹, respectively) temperature profiles of the bilayer structure with three hot source settings. Here, the circle in the core region is the modulated zone. d) Original and modulated ($\omega_1 = 0.138, 0.023, 0.0004$ rad s⁻¹, respectively) temperature-gradient distributions in the core region Ω_1 of the bilayer structure with the three hot sources.

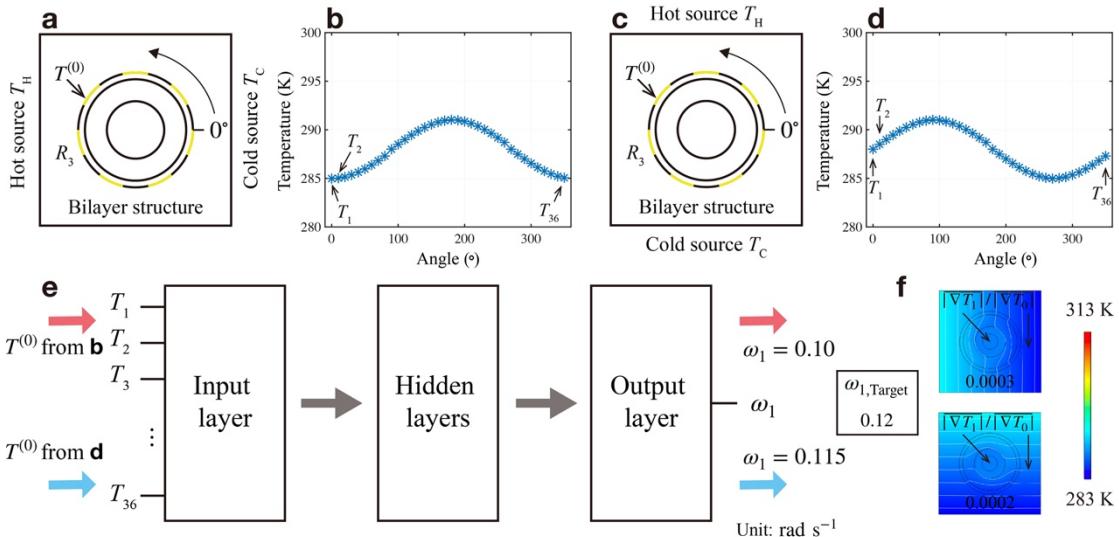


Figure S11. Effect of external thermal flow's direction on the metamaterial's performance. a,c) Schematic of the heat-enhanced thermal diffusion metamaterial working under external temperature settings $T_H = 293$ and $T_C = 283$ K, while with orthogonal external temperature flow's direction for (a) and (c). b,d) $N = 36$ equally spaced temperature data $T^{(0)}$ at the yellow dashed circle in bilayer structures under external temperature conditions of (a) and (c). Flow chart of e) Different temperature flow's direction gives rise to the rotation of the temperature data $T^{(0)}$ for the ANN, making the output ω_1 slightly perturbed. f) Comparison between modulated temperature fields when rotating velocity $\omega_1 = 0.10$ and 0.115 rad s^{-1} . The ratio of the temperature gradient in the core region to the external counterpart $|\nabla T_1|/|\nabla T_0|$ in two cases of (f) are almost the same.

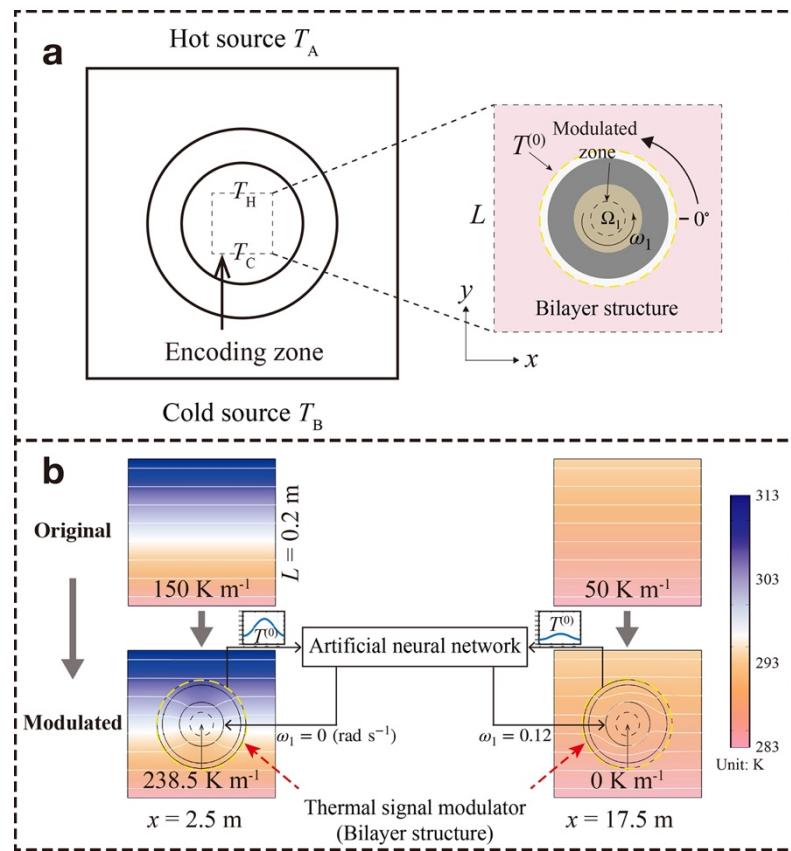


Figure S12. Principle of the thermal signal modulator. a) Schematic of one thermal signal modulator. The dashed box is selected as the encoding zone, and its original temperature gradient is determined by external temperature settings T_A and T_B . The right part is the designed metamaterial discussed in the main text. b) Two original and their modulated temperature profiles of the encoding zones at position $x = 2.5$ and 17.5 m. Color represents temperature distributions, and white lines represent the isotherms.

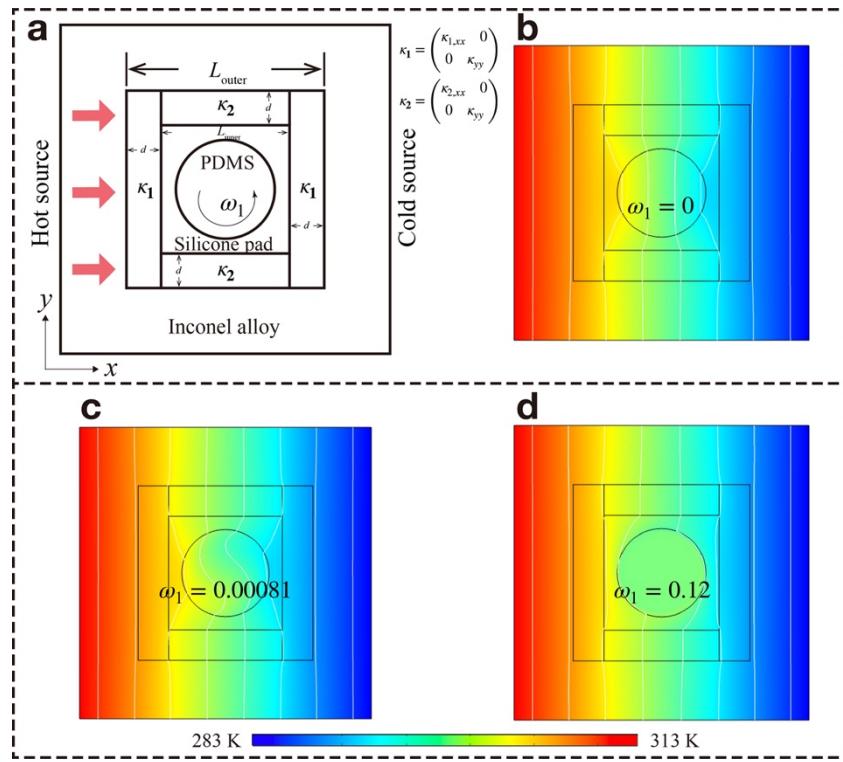


Figure S13. Effect of outer shape on the metamaterial's performance. a) Schematic of a square bilayer case. b-d) Temperature profiles of the bilayer structure with rotating velocity $\omega_1 = 0$, 0.00081, 0.12 rad s⁻¹, respectively. The hot source and the cold source are set as 313 and 283 K, respectively. Color represents temperature distributions, and white lines represent the isotherms.

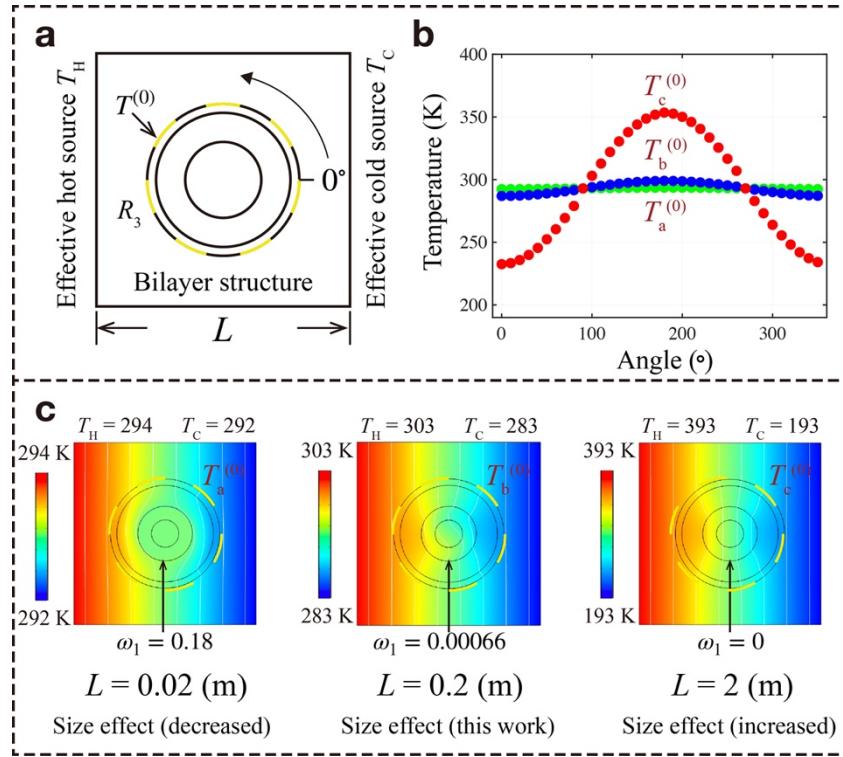


Figure S14. Effect of length scale on the metamaterial's performance. The external temperature gradient is maintained at 100 K m^{-1} in three cases. a) Schematic of the bilayer structure. b) $N = 36$ equally spaced temperature data $T_a^{(0)}, T_b^{(0)}, T_c^{(0)}$ in the yellow dashed circle for three sizes of bilayer structures ($L = 0.02, 0.2, 2 \text{ m}$, respectively). c) Temperature profiles of the bilayer structure with different sizes $L = 0.02, 0.2, 2 \text{ m}$, respectively. The corresponding rotating velocity are $\omega_1 = 0.18, 0.00066, 0 \text{ rad s}^{-1}$, respectively. Color represents temperature distributions, and white lines represent the isotherms. Here, these dimensions of the bilayer regions are equally increased or decreased by 10 times, and L is like a scale bar.

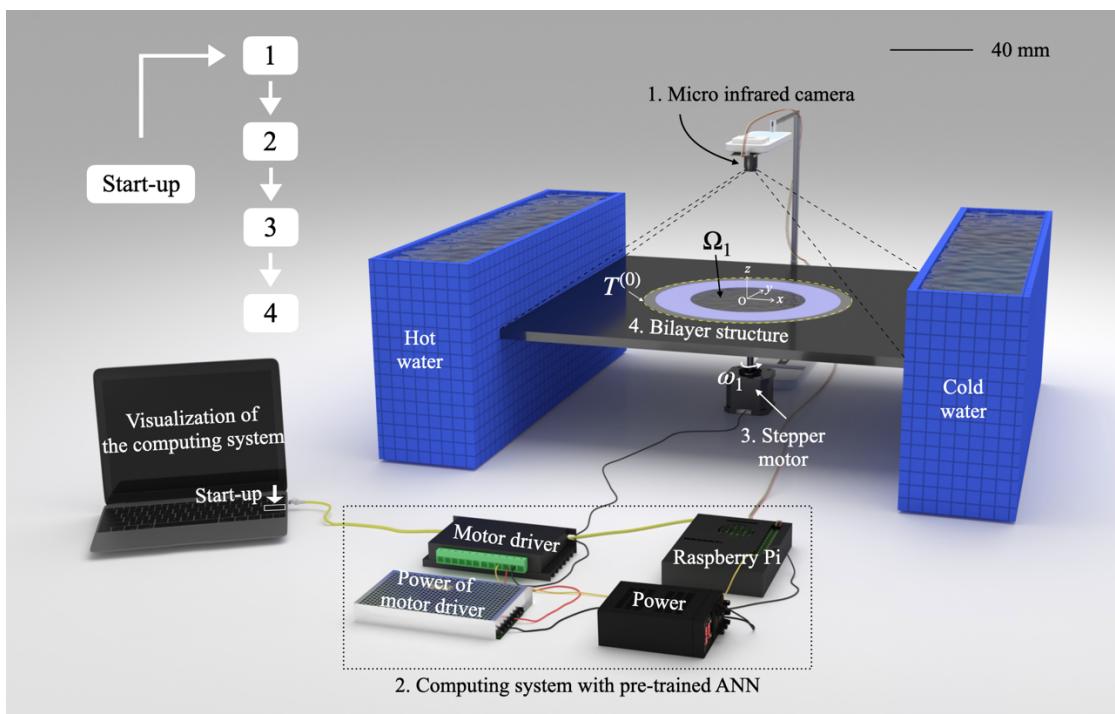


Figure S15. Experimental setup.

References

- [1] L. J. Xu, S. Yang, J. P. Huang, *Phys. Rev. Appl.* **2019**, *11*, 054071.
- [2] Y. R. Shen, *Principles of nonlinear optics*, United States, **1984**.
- [3] C. Qian, B. Zheng, Y. C. Shen, L. Jing, E. P. Li, L. Shen, H. S. Chen, *Nat. Photon.* **2020**, *14*, 383-390.
- [4] J. Bergstra, Y. Bengio, *J. Mach. Learn. Res.* **2012**, *13*, 281-305.
- [5] I. Goodfellow, Y. Bengio, A. Courville, *Deep learning*, MIT press, **2016**.
- [6] W. Sha, M. Xiao, J. H. Zhang, X. C. Ren, Z. Zhu, Y. Zhang, G. Q. Xu, H. G. Li, X. L. Liu, X. Chen, L. Gao, C. W. Qiu, R. Hu, *Nat. Commun.* **2021**, *12*, 7228.
- [7] W. Sha, R. Hu, M. Xiao, S. Chu, Z. Zhu, C. W. Qiu, L. Gao, *Npj Comput. Mater.* **2022**, *8*, 179.
- [8] P. Jin, J. Liu, L. Xu, J. Wang, X. Ouyang, J.-H. Jiang, J. Huang, *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120*, e2217068120.
- [9] C. Q. Wang, L. J. Xu, T. Jiang, L. Zhang, J. P. Huang, *Europhys. Lett.* **2021**, *133*, 20009.