UKAEA-CCFE-PR(25)334

Yadu Krishnan Sukumarapillai, Michelle Baxter, Perumal Nithiarasu

Machine Learning for Data Assimilation/Inverse Modelling in Thermal Problems

Machine Learning for Data Assimilation/Inverse Modelling in Thermal Problems

Yadu Krishnan Sukumarapillai, Michelle Baxter, Perumal Nithiarasu

This is a preprint of a paper submitted for publication in
Annual Review of Heat Transfer

MACHINE LEARNING FOR DATA ASSIMILATION/ INVERSE MODELING IN THERMAL PROBLEMS

Yadu Krishnan Sukumarapillai[1], ***Michelle Baxter***[2], ***& Perumal Nithiarasu***[1,*]

[1]Zienkiewicz Institute for Modelling, Data and AI, Faculty of Science and Engineering, Swansea University, Swansea, SA1 8EN, United Kingdom
[2]Culham Science Centre, United Kingdom Atomic Energy Authority (UKAEA), Abingdon, OX14 3DB, United Kingdom

[*]Address all correspondence to: Perumal Nithiarasu, E-mail: P.Nithiarasu@swansea.ac.uk

Abstract

This study aims to systematically examine various methods leveraging machine learning (ML) to assimilate data for investigating thermal systems. These measured or observed data may include temperature or thermal material properties and could be synthetically (computationally) generated or experimentally obtained. The goal of these ML-augmented methods is to derive the unknown material properties and/or reconstruct the full temperature field by integrating such measured data into physics-based computational models, such as FE models. The present work continues the previously conducted review of ML in heat transfer with a strong focus on inverse modeling techniques. It also attempts to closely incorporate ML into the FE workflow. Data assimilation and inverse modeling are closely linked tasks. While inverse modeling typically focuses on recovering unknown parameters or inputs from given observational data, data assimilation incorporates observations into dynamic models in a sequential manner, often with the goal of improving forecasting performance. In this review, we use the terms interchangeably for simplicity, though they arise from distinct methodological traditions.

KEY WORDS: *Inverse heat transfer problems, Machine learning, data assimilation, deep learning, digital twin, heat transfer, finite element method*

1. INTRODUCTION

Inverse problems in thermal systems are vital to a wide range of engineering and scientific applications. These problems are characterized by identifying the unknown causes, including boundary conditions, internal heat sources, or material properties, from observed effects. Unlike the direct or forward problem, where known system parameters are used to estimate the temperature field, inverse problems require the reconstruction of unknown inputs from limited or noisy data. These measurements often come with challenges when direct measurements are difficult or impractical, and only surface or sparse measurements are available. Another inherent characteristic of inverse heat transfer problems (IHTPs) is ill-posedness. According to Hadamard's criteria[1], well-posed problems must have a solution that exists, is unique, and has continuous dependency between the solution and initial conditions. Inverse problems mostly violate one or more of these conditions, making it difficult to solve. Furthermore, inverse problems are susceptible to minor errors in the measurements. This further increases with noise measurement, limited sensor placements, and cost or technical challenges in obtaining a large amount of high-fidelity data. Due to these facts, direct inversion of the governing equation always leads to an unstable or unrealistic solution.

To address these challenges, various traditional methods and numerical techniques were developed. Regularization methods, such as Tikhonov regularization[2], introduce additional constraints or penalty terms to stabilize the solution and effectively eliminate the effects of noise and ill-conditioning. Sequential approaches, such as Beck's sequential function[3], are efficiently used in time-dependent or dynamic problems. Analytical and semi-analytical methods, such as space marching techniques[4], are employed for inverse heat conduction problems. Optimization-based strategies, such as the Levenberg–Marquardt algorithm[5], and classical gradient-based optimization[6], are widely implemented to estimate unknown parameters or reconstruct a thermal field from limited observations. In conjunction with these traditional methods, data assimilation (DA) has emerged as a vital framework for integrating model predictions and observations to estimate unknown physical states and parameters in thermal systems. Traditional DA techniques such as Kalman filters and their variations[7], variational methods[8], and ensemble-based techniques[9] methodically incorporate new measurements into dynamic models and update the estimates in real-time as new data become available. In the context of IHTPs, data assimilation techniques not only stabilize the solution but also allow real-time monitoring, control, and optimization capabilities. Despite their success, these conventional methods face significant challenges in various aspects. These shortcomings include computational complexity, being prone to local minima, high dimensionality and nonlinearity, handling complex priors, and limitations in scalability due to the increased dimensionality.[10]

The growing complexities of thermal systems and the need for rapid, robust solutions brought attention to the limitations of the traditional methods. To mitigate the limitations, machine learning (ML) techniques are used as an aid to address the inverse problems. In particular, deep learning (DL) using deep neural networks (DNNs) has demonstrated a potential capability in learning patterns from data and extracting complex features.[11] Furthermore, DNNs are efficient in finding complex, nonlinear relations between sparse measurements and unknown parameters or fields in the context of thermal systems. Using a large dataset or experimental data, these models can serve as efficient surrogate models that provide rapid and accurate predictions once trained. Combining data-driven models and physics-based solvers shows greater potential as the governing laws and constraints are incorporated into the learning process, eventually estimating physically meaningful solutions.

The present work systematically examines the diverse machine learning methods currently explored for assimilating data into thermal systems. The primary focus is on inverse thermal problems of unknown parameter estimation or reconstruction of the thermal field, where ML capabilities are efficiently incorporated. This review mainly includes different methods where ML algorithms are used to identify parameters from measured or observed data, whether synthetically generated or experimentally obtained. It also includes techniques for reconstructing the full thermal field from sparse measurements and approaches that demonstrate coupling between machine learning and physics-based models.

2. PRINCIPLES OF DATA ASSIMILATION AND INVERSE PROBLEMS IN THERMAL SYSTEMS

Data Assimilation is the process of integrating measured data, which are typically sparse and noisy, into numerical models like the FE model to generate an improved state or to predict the underlying parameters of a system. These principles are widely used in diverse fields, from optimizing heat exchangers and designing advanced materials to climate modeling and bio-medical engineering.[12] Thermal systems, governed by heat conduction, convection, and radiation, present significant challenges for accurate modeling. In traditional forward problems, the system behavior is predicted based on known inputs such as material properties, initial conditions, and boundary conditions, which often come with uncertainties or are incomplete.[13] In contrast, inverse problems can generally be categorized into two types: (1) determining the system state or unknown parameters from observed causes and effects[14], and (2) finding the causes from limited measurements or observations.[15] These demand the adoption of data assimilation and inverse problem techniques. Objective of the the inverse modeling is to find the unknown parameters or inputs from the given observational data. Since data assimilation is a special class of inverse problems[16], the terms *data assimilation* and *inverse problems* are used interchangeably in this review, though they arise from distinct methodologies. Data Assimilation (DA) technique focuses on sequentially integrating real-time or near real-time measurements into a dynamic model of a physical system to improve the accuracy and best possible estimate of the system's present state.[17,18] This is beneficial for forecasting and active control applications where the system state evolves. In some cases, DA is capable of finding the best description of uncertain state variables, parameters, and model control, or all of them simultaneously.[19] The preliminary background on the DA process as described by Arcucci et al.[20], is provided here.

The system state u is replaced with u^{DA}, which incorporates the observations through the predictor-correction cycle. The forecasting model $\dot{u}$ describes the evolution of the systems given as:

$$\dot{u} = \mathcal{M}(u, t, \theta) \tag{1}$$

where u is the state, $\mathcal{M}$ is a nonlinear function, θ is a state parameter. $t = \in [0, T]$ is time and The observation of the state u is provided as,

$$v = \mathcal{H}(u) + \epsilon \tag{2}$$

where $\mathcal{H}$ is the observation function and ϵ represents the measurement error. In this scenario, data assimilation explores the method of obtaining a possible state of the system as a function of time using observations Eq. (2) and the model Eq. (1). Consider a fixed time step, say t_k, the estimated system state is provided as:

$$u_k = M u_{k-1} \tag{3}$$

where M is a discretization of a first order approximation of $\mathcal{M}$. The corresponding observation of the state at time t_k is v_k, and let $H : u_k \rightarrow v_k$ be the discretization of a first-order approximation of $\mathcal{H}$. The DA method involves determining u^{DA} (referred to as analysis) as an optimal solution between the predictions derived from the estimated system state or background state and measured observations. The prediction-correction cycle process is illustrated in Fig.1. The u^{DA} is computed as an inverse solution as

$$v_k = H u^{DA} + e_{R_k} \tag{4}$$

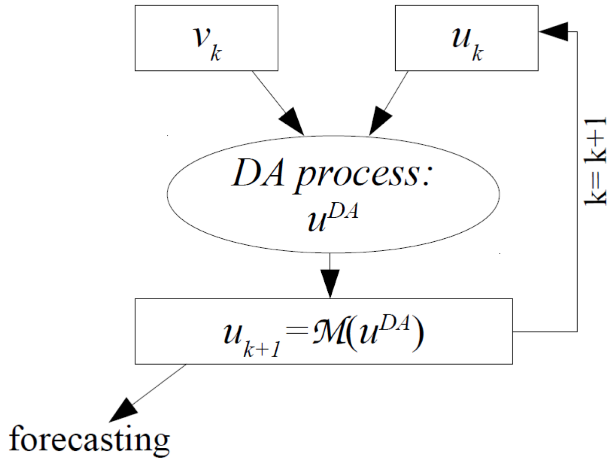


FIG. 1: The prediction-correction cycle (Reprinted under a Creative Commons Attribution 4.0 International license, Copyright 2021).[20]

with constraint

$$u^{DA} = u_k + e_{B_k} \tag{5}$$

where e_{R_k} and e_{B_k} are the observation and model errors respectively. The observation error accounts for both instrumentation and representativity errors[21], and the discretization, approximation, and any other numerical inaccuracies contribute to both errors. The Eq. (4) is usually ill-posed, which means at least one of the conditions, such as (i) the solution exists, (ii) uniqueness, (iii) stability, is not satisfied.[1,22,23,24]This means that H is usually non-square and/or rank deficient; therefore, we cannot invert H to find u^{DA} directly.[25] The Tikhonov regularization mitigates the ill-posedness problem with a minimization of the least-squares cost function[26], and the DA problem can be defined as:

$$u^{DA} = argmin_u\{\|u - u_k\|^2_{B_k^{-1}} + \|v - Hu\|^2_{R_k^{-1}}\} \tag{6}$$

where R_k and B_k are the observation and model error covariance matrices, respectively.

Several methods are used to solve the DA problem specified in Eq. (6). The Variational approach and Kalman Filter (KF) are primarily used for the traditional data assimilation process to improve the system state by integrating observations.[27] In the variational approach, minimization of a functional which estimates the difference between the numerical solution and measurements is carried out by finding a solution which minimizes a cost function.[28,20] On the other hand, the Kalman Filter (KF) finds a solution with minimum variance.[29] However, both methods show limitations due to their mathematical and statistical assumptions. The major issue is their dependence on unrealistic assumptions, such as linearity, multivariate normality of errors, and stationary state-transition functions, along with the assumption of zero error covariances.[30,21,31] These presumptions often are not valid for complex, highly nonlinear systems, where even a minor uncertainty in the input can cause an amplified error in the system behavior.[32] In the case of the variational approach, its performance greatly depends on the accuracy of the background and observation error covariance matrices.[33] A precise definition of these matrices is intricate and computationally challenging, especially in the context of high-dimensional data, as they are critical for effectively communicating between observed and unobserved variables. Moreover, this approach does not naturally account for the temporal evolution during the DA process, and there is a high risk of convergence to the local minima in the cost function rather than a globally optimal solution.[34,35] Conversely, the Kalman Filter (KF) faces different challenges. The primary limitation of the KF is the size of the Kalman gain matrix, which can be very large in the case of high-dimensional problems. Since KF relies on the linear assumption for the system dynamics and observations, it restricts its applications to strongly nonlinear physical systems[36]. Also, the linear assumptions make it less effective in situations where small uncertainties can propagate and amplify dramatically, leading to erroneous system behavior.[37]

Furthermore, computational cost poses a major hurdle for traditional DA methods. They involve significant computational resources for solving, storing and manipulating larger matrices, especially in higher-dimensional problems and are less efficient when dealing with very large data sets.[38,39] Another important limitation of these traditional methods is their inability to learn from past error corrections. While DA aims to improve the forecast, it doesn't improve the error propagation within the forecast model. Consequently, frequent corrections need to be applied at each step without the benefit of previous experience from the assimilation process. Additionally, taking the information in a temporal window into account, these methods often struggle to take the past information into account due to the complexity of the forecasting problem and computational cost.[20]

These limitations exhibit the need for novel approaches that can handle complex, high-dimensional, nonlinear inverse problems more effectively and robustly. Machine learning (ML) techniques, with their ability to learn complex patterns from data and their potential for creating faster surrogate models, offer significant advantages in the context of data assimilation (DA) and inverse problems. Several works explored integrating DA with ML to improve the efficiency and accuracy of numerical simulations. While this integration increases the prediction reliability and reduces errors by incorporating physically significant information from observed data, learning from the past state of a DA process leads to a faster and more accurate model.[30] For higher-dimensional systems, DA can be combined with reduced-order modeling (ROM) techniques, including ML-based autoencoders, to reduce computational cost.[33] ML is also being used to correct model errors in the DA and forecast system state.[40] In the broader concept of inverse problems, ML offers a more efficient and accurate approach by analyzing large data compared to traditional methods. Its main advantage lies in its ability to self-learn and predict trends through algorithms, an opportunity for continuous training with an increasing dataset, and more accurate results over time.[41]

In the field of heat transfer and thermal systems, it often involves complex physics, which can be identified through reduced-order modeling and extracting system information from the measured data. ML techniques can provide computationally efficient models for accurate forecasting and robust optimization in this domain. The applications of ML to thermal problems include building reduced-order models, predicting heat transfer coefficients and pressure drop, analyzing complex experimental data, and optimizing large-scale thermal systems.[42]

3. MACHINE LEARNING PARADIGMS FOR DATA ASSIMILATION AND INVERSE PROBLEMS IN THERMAL SYSTEMS

Machine Learning (ML) techniques are emerging as a powerful tool for data assimilation in thermal systems, overcoming significant challenges of traditional methods outlined in the previous section. Thermal or heat transfer problems are often classified into two categories: direct and inverse problems. The direct problem deals with heat transfer, testing, and forward temperature field construction.[43] The finite element method is often considered the most widely used technique[44], especially for heat conduction; however, for the inverse problems, conventional techniques such as FEM are often paired with optimization algorithms. This creates a bottleneck, as the direct problem must be solved several times using FEM in each iteration of the optimization problem.[45] This repeated process can lead to substantial computational resources and time and may require regularization strategies to ensure stability, especially in the case of ill-posed inverse problems.[46,47,48] In contrast, ML and DL offer a more robust approach that can circumvent these challenges. The following sub-sections discuss the principal ML paradigms for data assimilation/inverse problems in thermal systems.

3.1. Supervised Learning Approaches

The most common type of machine learning technique is supervised learning. Supervised learning algorithms can discover patterns in the given data with existing dependent and independent factors to predict the future of the dependent factors.[49] In other words, these algorithms can learn a mapping between a set of input and output variables and use this mapping to predict the output for unseen data.[50] Supervised learning can be classified into two categories: Regression and Classification. In regression, the output is continuous, and the latter is implemented for categorical output.[51,52] In the context of thermal problems, input might be sparse sensor measurements of temperatures or fluxes at specific locations, or geometric parameters, interfacial temperature and pressure, while outputs could be the unknown material properties, boundary conditions, reconstructed temperature profile or fluxes, and heat transfer coefficient (HTC).[53,54]

3.1.1 Deep Neural Networks (DNNs) for Parameter Identification and Field Reconstruction

Artificial Neural Networks (ANNs) are nonlinear, non-convex regression models with excellent predictive abilities, regarded as universal approximators of continuous functions.[55] ANNs are capable of learning and generalizing underlying information from data.[56] An artificial neuron or perception[57] is the fundamental building block of an ANN. A more sophisticated version of the ANN, with deep architecture, advanced training algorithms for higher feature abstraction and complexity[58] is known as Deep Neural Networks (DNNs). ANNs with more than three layers, including two or more hidden layers, qualify as deep neural networks. If the NN has a single hidden layer, it is usually called a shallow network.[59] Each hidden layer consists of multiple neurons, thus increasing its nonlinearity and feature prediction. A schematic of a DNN is provided in Fig. 2. The architecture comprises an input layer, an output layer, and multiple hidden layers, and each hidden layer can have multiple neurons.

This section explores the application of DNNs in effectively implementing the DA/inverse problems in the context of thermal systems. Standard DNNs, including Feedforward Neural Networks (FNNs), are widely used as flexible function approximations. In thermal data assimilation or inverse problems, they can be used to map sparse observation data directly to the parameter of interest.

The work conducted by Tamaddon-Jahromi et al.[41] uses a database generated from forward FE solutions to train a DNN. In this earlier work, the trained network was used for finding the boundary conditions from sparse temperature measurements in linear/nonlinear heat conduction/convection problems. Various test cases were introduced, and the FE mesh used for this study is shown in Fig. 3. They performed a forward problem with constant tem-

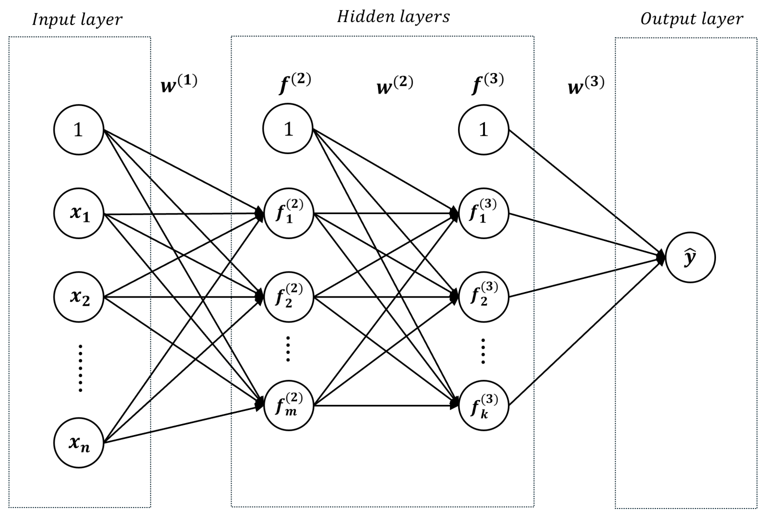


FIG. 2: Schematic of DNN with two layers, where $\{1, x_1, x_2, .., x_n\}$ is the input dataset, $\{w^{(1)}, w^{(2}, w^{(3)}\}$ is the weights for each layer, $\{f^{(2)}, f^{(3)}\}$ is the vector of activation functions and $\hat{y}$ is the predicted output.

perature boundary conditions and obtained the temperature distribution inside the domain. The objective was to use a selected number of these outputs as input for the DNN and to obtain an accurate boundary condition, which was used as input for the forward problem to check the accuracy of the inverse model. Various network architectures based on the number of layers, neurons per layer, and different temperature measurements are established. For example, linear and nonlinear conduction problems with three temperature measurements in the domain were used to train the network to predict the boundary condition. Linear and nonlinear thermal conductivity are considered as $K(T) = k_0$ and $K(T) = k_0 T$, respectively, with $k_0 = 1$. Fig. 4 shows the model accuracy and loss of the 3-64-32-16-4 NN for the linear and nonlinear heat conduction problems. As the number of measured temperatures inside the geometry increases, the performance of the NN improves, and it predicts the solution with higher accuracy. However, another research shows that interpolation methods perform better for linear heat conduction problems, and DNN performs very well for nonlinear heat conduction problems.[60]

In reference[61], ANN is trained with data from both FEM and experiments to invert thermo-physical properties, including density, thermal conductivity, and heat capacity of

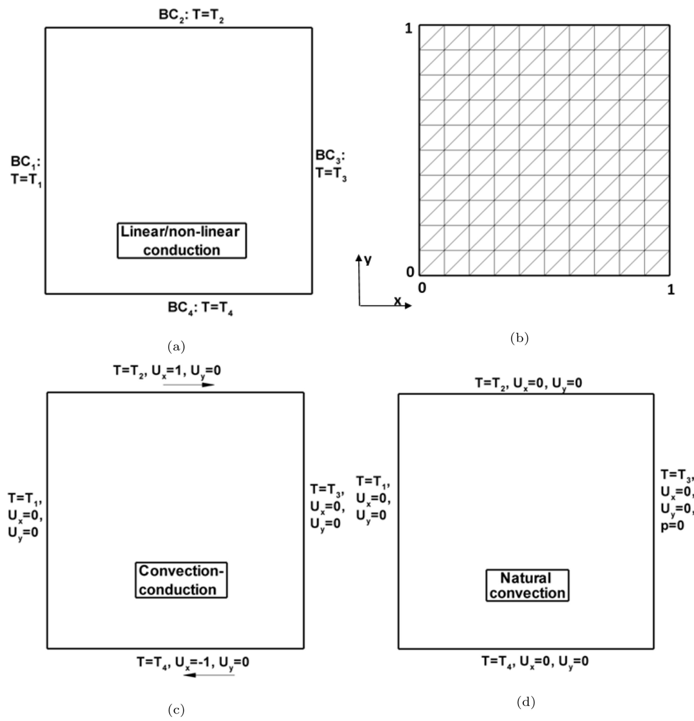


FIG. 3: Problem descriptions and FE mesh.[41]

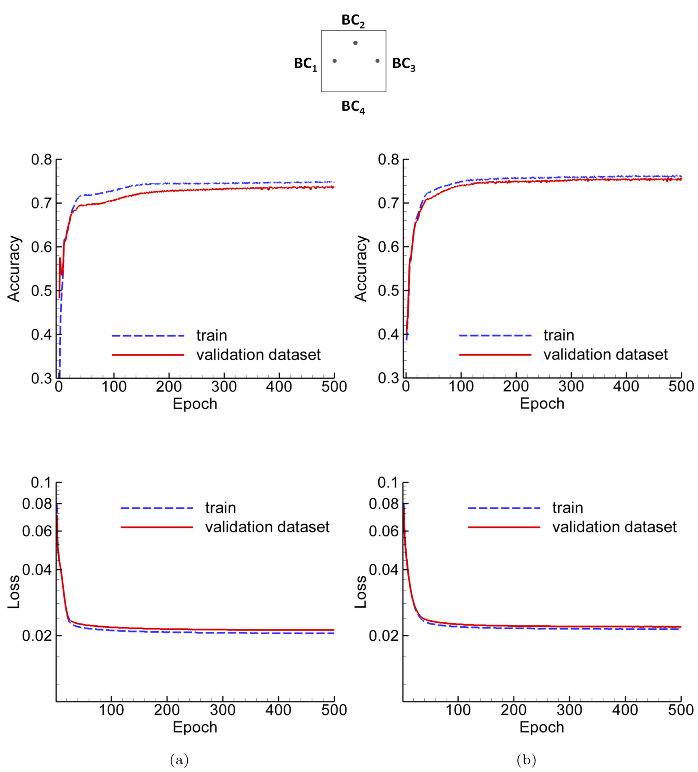


FIG. 4: Model accuracy and loss (Three measured temperatures), (a) Linear heat conduction, (b) Nonlinear heat conduction.[41]

lunar regolith simulant based on thermal probe measurements. They later implemented a multivariate Newton iterative algorithm for further optimization of the ANN. The results show that the ML model was effective in predicting the temperature responses, thus offering a novel approach to measuring the in situ thermal properties of extraterrestrial regolith. Similarly, in reference[62], a machine learning model is introduced to solve the Inverse Heat Conduction Problem (IHCP), such as identifying the inner wall shape of a pipe by estimating the relationship between temperatures and thermal conductivity.

A critical aspect of these supervised methodologies is the generation of a large dataset. This dataset must span the expected range of input measurements and output parameters to ensure the trained model can generalize well with unseen data. Often, the data may not be clean, of good quality, or require manual labeling. Researchers usually rely on small data or computationally expensive simulations. To overcome the difficulties of restricted data and improve machine learning methods for IHCPs, synthetic data can be generated by introducing algorithms like the Gaussian Copula.[63] The deep integration criterion emphasized in this review suggests that parameters identified by DNNs are subsequently used to inform or update an FE model, or the DNN itself is part of the larger FE-based workflow. For example, DNN trained on FEM data can predict temperatures with higher accuracy and speed, and it allows for inverse modeling to find the thermal conductivity from measured temperatures.[64]

Estimating hot-wall surface heat flux from internal temperature measurement presents a challenging IHCP due to its ill-posedness and nonlinearity. To address this, Wang et al.[65] discussed a novel inverse ANN. A one-dimensional heat conduction problem with net heat flux at the front and heat sink or radiation and convection at the back was considered. They formulate the inverse estimation of the incident heat flux as a nonlinear model and introduce a nonlinear ANN due to its ability to approximate nonlinear functions. A single hidden-layer feature with two temperatures as inputs and an ANN output estimated the heat flux. To validate the inverse problem, a new type of hot-wall heat flux sensor was incorporated, and the estimated heat flux agrees with the calibrated values. A feed-forward neural network (FFNN) with a digital filter approach was proposed for one-dimensional IHCPs with and without a moving boundary, along with and without temperature-dependent material properties.[66] The temperature from a limited time window was used as the input of the network, and the current time step heat flux value was the output. They examined the effectiveness of the ANN in terms of prediction accuracy and computational time. The results revealed that the ANN performed well in the robustness test, and it predicted accurate and consistent results when the temperature measurements were near the boundary surface. In another research[67], the ANN was considered as a digital filter to find near-real-time heat flux using temperature measurements. The ANN model was trained with a set of temperatures and heat flux values. This trained network was effectively able to capture various heat flux profiles for constant and temperature-dependent material properties. Another important discovery was that the ANN required only a smaller set of data from the future or previous time steps to obtain the current time step heat flux, compared to the traditional filter-form Tikhonov regularization algorithm. Mirsephai et al.[68] have proposed an ANN for an inverse heat transfer model to identify the absorbed heat in radiation-dominated heat transfer problems. The ANN with three layers was trained with experimental data and proved its applicability to predict the input heat to the system. The real heat applied to the system and the predicted heat from the ANN were accurate and consistent for various test problems.

3.1.2 Convolutional Neural Networks (CNNs) for Spatially Distributed Data

Convolutional neural networks (CNNs) are a powerful machine learning algorithm based on traditional ANNs. The architecture consists of an input layer, an alternating series of convolution and pooling layers, a fully connected layer, activation functions, and an output layer.[69]. A typical CNN architecture is illustrated in Fig.5. The input raw data was converted to an abstract or higher-level feature representation through the convolution and pooling layers, often known as a feature extractor.[70,69] This process typically involves a combination

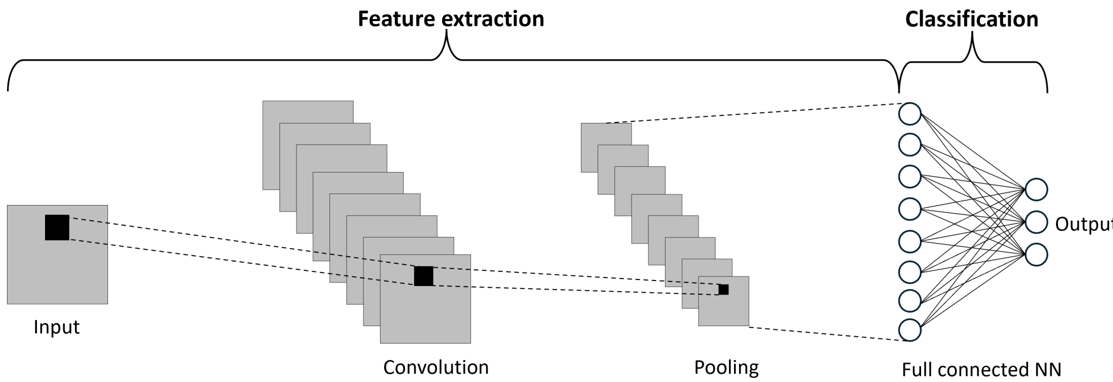


FIG. 5: Schematic of CNN.

of linear and nonlinear processes using convolution and activation functions.[71] CNNs offer significant advantages over traditional neural networks, including strong adaptability, robust feature learning, minimal training parameters, reduced redundant computations, and higher recognition accuracy.[72,70,69] The objective of a CNN is to interpret an array of data through local operations in various parts of the array and send the resultant output to the new layer.[73] Convolution layer consists of kernels or filters to generate an output feature map of the input data[74]. The feature map is then provided to the pooling layers, where the higher spatial dimension of the feature is reduced without losing the dominant features. Thus reducing the size of the data and parameter count.[75] Another core part of the CNNs is the activation function, which maps the input to the output through nonlinear functions.[76] Finally, the fully connected layer, similar to a feed-forward ANN, situated at the end of the NN, acts as a CNN classifier.[74]

Early research shows the benefit and efficiency of CNNs to solve forward problems and reconstructions of thermal fields. Edalatifar et al.[77] modeled a convolutional neural network to learn the physics of 2D heat transfer without prior knowledge of the underlying governing PDE. The input of the neural network consisted of two-channel images of geometry and boundary conditions, and the output was an image of the temperature distribution. CNN was trained with a large dataset (images) created using the Finite Volume Method (FVM). Another study[78] also used CNN to reconstruct the physical field in a thermal problem and obtained the Nusselt number (Nu) and friction factor (f). This study shows the significant potential of the computational efficiency of a well-trained CNN model by comparing GPU-accelerated with typical CFD solvers. CNNs techniques were systematically applied to various thermal problems like boiling heat transfer[79], fast prediction of 2D steady state conduction[80], convection flow of nanofluids[81], heat flux in turbulent flow[82], surrogate modeling of conjugate heat transfer[83], and many more. For brevity, we restrict the discussion of such problems and focus on the use of CNN in DA/inverse problems of heat transfer.

Virupaksha et al.[84] discussed the application of CNNs to transient natural convection in heterogeneous porous media. This study focuses on surrogate modeling, inverse modeling, and time series prediction. In this inverse problem, the goal was to predict the domain heterogeneity from temperature maps. The data was generated using FE solver OGS.[85] Initially, Encoder-Decoder CNN (ED-CNN)[86] was used with a single input image created by merging the transient data, and the output was the Rayleigh-Darcy map. Its performance in inverse modeling was poor compared to surrogate modeling. The authors hypothesized that merging the transient thermal images caused a loss of information about the temporal progress of the heat flow, and ED-CNN was unable to use the time progress information. To address this, authors have developed another technique, 3D-CNN, where a temperature map was provided at a given instant as independent inputs. This allowed the model to leverage the time evolution of heat flow. The 3D-CNN was evaluated for classification and extended to regression in heterogeneous domains. For the inverse problem, the 3D-CNN with transient data achieved greater accuracy and significantly outperformed steady and transient data trained on 2D-CNNs. A multiple regression convolutional neural network (MRCNN) was introduced to estimate various parameters in IHTPs.[87] Traditional CFD and DL are used to generate a dataset, and the proposed method was validated with cubic cavity experiments. The MRCNN technique shows good accuracy for predicting unknown parameters.

Lockwood et al.[88] propose an inverse surrogate model (ISM) using CNNs to evaluate various building parameters such as wall insulation conductivity. Synthetic data was generated using building energy simulation software EnergyPlus[89] by varying parameters to obtain the room temperature series and parameter values. The geometry model was divided into zones analogous to rooms in a building. The modeled CNN was used to predict the building parameters when the time series was given as input. The performance of the CNN model was promising, especially for smaller periods. However, the model struggled to predict attic insulation and glazing conductivity as the number of zones was reduced from 15 to 4. Masrouri & Tahsini[90] discussed a comparative study between a novel one-dimensional convolution network and traditional methods such as the Levenberg–Marquardt and conjugate gradient methods for solving inverse IHTP. The objective was to find the heat transfer coefficient and the generated temperature of an industrial heat gun from temperature measurements. The numerical solution was used as the training data, and measurements were taken from experiments. Results show that one-dimensional convolution network predictions were slightly different from other traditional methods for the estimated parameters. However, the predicted temperature profiles were reasonably well matched with the measurements.

A combination of CNNs and Long Short-Term Memory Networks (LSTMs), CNN-LSTM, was introduced to estimate the thermal boundary parameters for a transient IHTP.[91] The training, testing, and validation data were generated from CFD analysis. The authors outline that CNN-LSTM models were able to extract the temperature features accurately and additionally use the correlation between time series to enhance the predictive accuracy of the parameters. And the performance of the CNN-LSTM model was significantly superior to that of standalone CNNs and LSTM models. In another study[92], taking advantage of the CNN-LSTM model was discussed for the reconstruction of the 3D temperature field at multiple future moments from non-invasive combustion features derived from Light Field (LF) tomography. This type of problem is called the inverse radiative heat transfer problem. To address the challenges of the ill-posed and time-consuming nature of the traditional methods, the authors combined CNNs and LSTMs for a novel CNN-LSTM model for online prediction of instantaneous flame temperature. This model uses CNNs for the extraction of

3D flame temperature features and radiation images, while LSTM predicts accurate future moments of the flame temperature using time series data of extremely dynamic combustion flame profiles. The model learns the complex flame evolution with high accuracy and robustness to the noise. These studies facilitate the greater potential of CNN-LSTM models in IHTPs.

3.1.3 Recurrent Neural Networks (RNNs) for Transient Thermal Problems

Recurrent Neural Networks (RNNs) are a type of ANN to process sequential data.[93] Their fundamental architecture consists of an input layer, a hidden layer, and an output layer. A key difference between RNNs and FNNs is the presence of recurrent cells, as shown in Fig.6, enabling the information to cycle within the networks.[94] This allows the network to maintain a hidden state that captures information from previous inputs. The inherent time series memory makes RNNs best for modeling and processing temporal dependencies within the sequential data.[95]. The mathematical formulation of a standard recurrent cell can be written as[15]

$$\begin{aligned} h_t &= \sigma(W_h h_{t-1} + W_x x_t + b) \\ y_t &= h_t \end{aligned} \tag{7}$$

where x_t, y_t, h_t, and h_{t-1} are the input, output, and hidden state of the recurrent cell at time t, and hidden state at time $(t-1)$. W_h and W_x indicate the weights, b is the bias and σ is the activation(sigmoid) function.

Traditional RNNs, often with sigmoid or tanh activation functions, struggle to learn important features when the input gap is large. This is because of vanishing or exploding gradients during training due to long-term dependencies in the dataset.[96] To overcome these limitations, LSTMs were developed by incorporating gating mechanisms to control the information flow in the network. These gated cells consist of an input gate(i_t), an output gate (o_t), and a forget gate (f_t).[97] These gates regulate how much input data to consider, how much prior state to forget, and how much cell state to send out.[94] Schematic of an LSTM

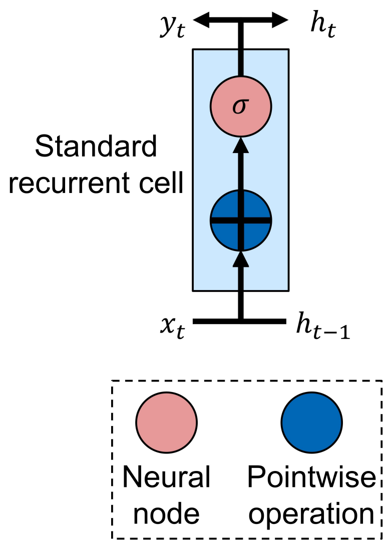


FIG. 6: Schematic of a Recurrent cell.[15]

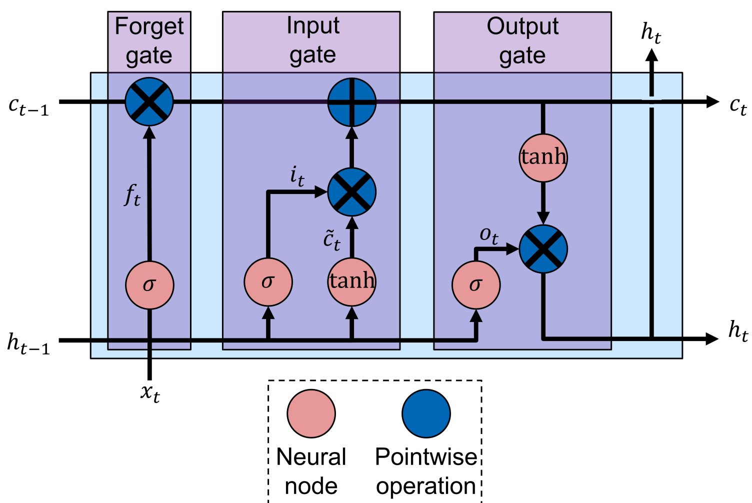


FIG. 7: Schematic of LSTM architecture.[15]

model is provided in Fig.7. The following paragraphs discuss the recent progress in the applications of RNNs to DA and inverse problems in thermal systems.

Zwart et al.[98] proposed a novel integration of DL and DA for a real-time forecasting system of stream water temperature in the Delaware River basin. Observed data were assimilated into the DL model to predict stream temperature for 7 days into the future. The ability to capture temporal relations in the data, an LSTM network was used in their methodology. The performance of the DL model was assessed with and without data assimilation.

The ensemble Kalman Filter (EnKF) was employed for data assimilation to adjust LSTM states according to the observed maximum temperature measurements and reduce bias in predictions. Results demonstrated that the DL-DA approach improved the forecast performance of maximum stream temperature compared to the LSTM model without DA. This study underscores the significance of ML models, especially RNNs like LSTM, for accurate and reliable forecasts to support decision-making.

A nonlinear-autoregressive-exogenous recurrent neural network (NARX RNN or NARX) was proposed to solve IHCPs in one-dimensional domains.[66] The authors discussed two machine learning approaches, NARX and FFNNs based on the digital filter, to estimate near-time heat flux from internal temperature measurements. NARX is discussed in this section, and the latter is discussed earlier in this review. NARX network consists of recurrent connections, and the major difference from the traditional RNNs is the exogenous inputs. To understand the performance of the proposed models, various cases, including those with and without a moving boundary, with and without temperature-dependent material properties, are adopted. Synthetic data generated from COMSOL Multiphysics was used to train and test the proposed networks. The Location of measurement sensors, time consumption, and delays associated with proposed DL models were discussed. For cases with constant and temperature-dependent material properties with fixed boundary, FNNs behave much better than NARX for triangular and parabolic heat flux profiles, and NARX performed well for step function heat flux profiles. In cases with moving boundaries, FFNNs outperformed NARX for almost every type of heat flux profile.

Solving IHCPs for 3D complex structures was investigated by Wang et al.[99] Since the traditional algorithms and iterations for 3D IHCPs are computationally expensive and time consuming, the authors adopted a new framework founded on a convolution-LSTM (ConvLSTM) architecture.[100] The objective was to predict transient heat flux for nonlinear 3D heat conduction problems from measured temperatures. A 3D complex model with nonlinear boundary conditions and temperature-dependent parameters served as the physical model. FEM-based software, COMSOL Multiphysics, was used to generate training and testing datasets. Furthermore, temperature data and geometries at each step were combined to act as the input to the DL network. The proposed model, convolution-LSTM, is the combination of CNNs and LSTM networks. The ConvLSTM primarily consists of three layers of ConvLSTM at the beginning, followed by three LSTM layers and a fully connected network at the end, as shown in Fig.8. Initially, a simple 3D model was tested to assess the potential of the proposed framework, and the predicted heat flux matched the ground truth. In cases of complex geometries, the framework performed reasonably well and exhibited excellent generalization ability to geometric shapes. To reassure confidence in the proposed framework, additional comparisons are made with traditional methods such as congregate (CG) and the sequential function specification method (SFSM). This finally concluded that the fully trained model outperformed the traditional models in terms of time consumption and accuracy.

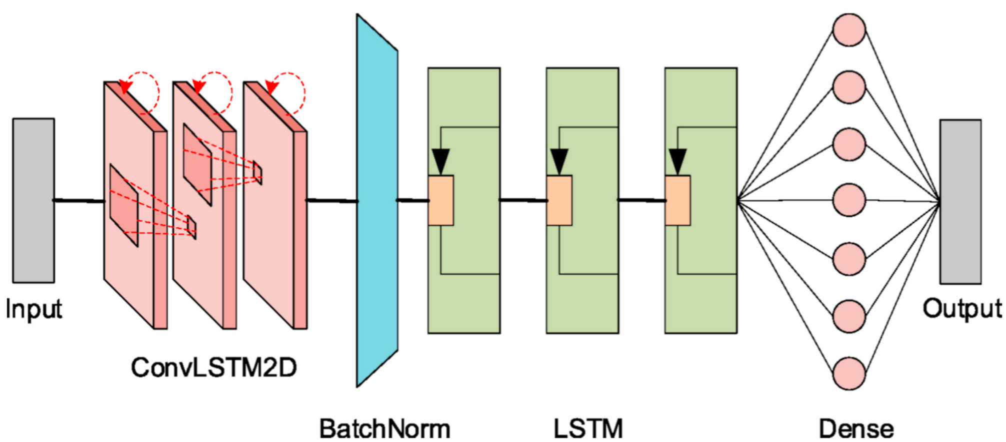


FIG. 8: Proposed Convolution- LSTM architecture for solving transient IHCPs (Reprinted under a Creative Commons Attribution 4.0 International license, Copyright 2022).[99]

Bielajewa et al. conducted a comparison study between the LSTM model and transformers for transient thermal field reconstruction[15]. The authors selected four transformer-based models and an LSTM model for reconstructing the temperature profile of one and two-dimensional heat conduction problems. Transformers[101] represent a class of DNNs that employ a self-attention mechanism to capture the features within the sequential data effectively.

Unlike LSTM models, transformers do not have any recurrent connections and perform well with long dependencies. The absence of recurrent connections helps transformers reduce their training time and facilitate parallel processing.[102]. Training, validation, and testing data were generated using FDM and FEM packages for one and two-dimensional models, respectively. Transient simulation data obtained for 1000s, where 700s of data were used for training, 100s for validation, and the remaining 200s of data were used for testing the models. Temperature values at randomly selected points inside the domain were given to DL networks, and a fully reconstructed temperature profile was generated. For the case of a one-dimensional problem, LSTM exhibits the lowest training time and normalized Root Mean Squared Error (NRMSE) compared to other transformer models. Similarly, in the case of two-dimensional transient heat conduction, the LSTM model consistently performed better than transformer models. Fig.9 illustrates the time-averaged prediction error distribution for the two-dimensional IHTP. The authors emphasized that there was no significant advantage of using transformers over a simple ML model like LSTM for transient field construction problems.

The real-time estimation of heat flux is crucial in several industrial applications, including cutting processes. During the cutting process, high heat is generated due to the plastic yielding of the material.[103] The heat can be accumulated in the small tool chip interface, leading to irregular temperature distribution and high-temperature gradients. This can affect the machining quality, tool life, and production safety. In this context, Han et al.[104] propose an LSTM-based encoder-decoder (ED) technique to estimate the real-time prediction of heat flux in the tool-chip region during the turning process. The training data was generated using FEM simulations, and the temperature fields are solved using a preconditioned conjugate gradient (PCG) solver. Other networks, including MLP, CNN, and LSTM, were used to compare the performance and efficacy of the LSTM-ED model. The same training data with normalization was used for all the neural networks. Heat flux and temperatures estimated from the FEM simulation, which are not included in the training dataset, were

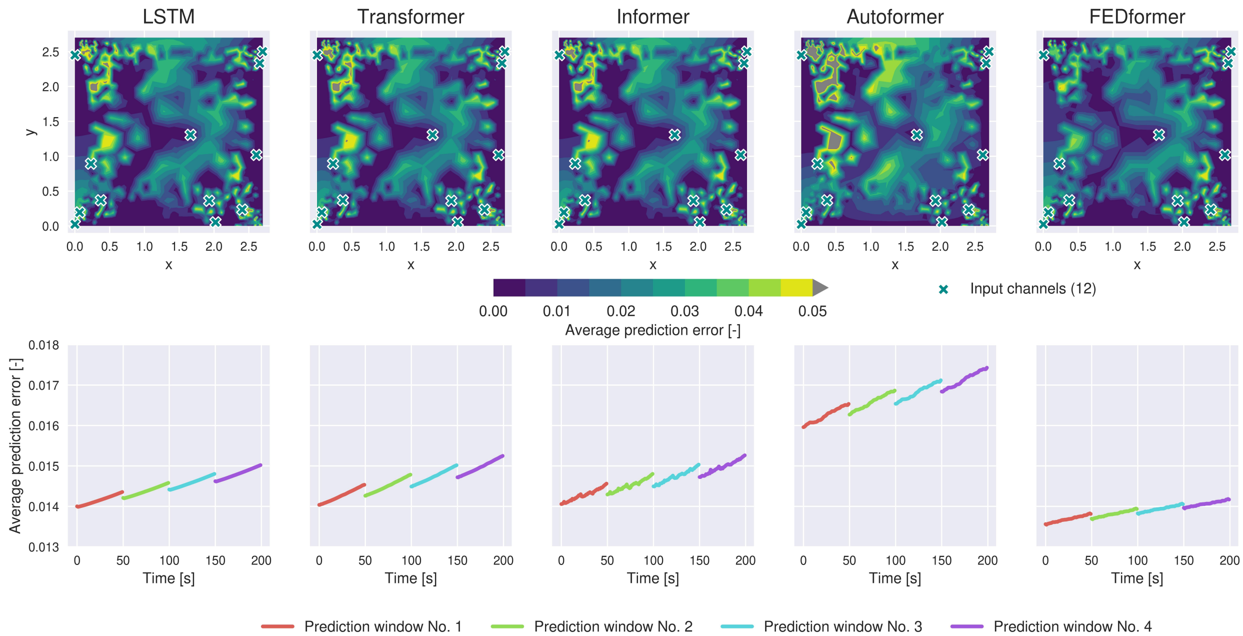


FIG. 9: Time averaged prediction error distribution (top)and prediction error averaged at each time step (bottom) for 2D conduction problem. The green crosses indicate the ML inputs.[15]

used for the testing of the models. Trained models were input with sensor temperatures to predict the heat flux. To further understand the efficiency of the proposed model, different noise levels were introduced into the temperature measurements. The results demonstrated that the LSTM-ED model outperformed other NN models and was highly immune to noisy levels ranging from 1K to 20K with reasonable computation time. Authors have tested the robustness and effectiveness of the proposed LSTM-ED model with numerical analysis and prediction accuracy through experiments.

In another study, a physics-guided LSTM was presented for predicting the dynamic heat load of buildings[105]. In situ measurements and numerical simulations were used to generate the dataset. Extra data on indoor temperature and heat loads were obtained from numerical analysis using TRNSYS. Several inputs, including historical data of outdoor meteorological conditions, indoor-air temperature, heat load, time and date given to the network, and heat loads for the next seven days of the building, are predicted. Results show that the proposed method exhibits greater accuracy and robustness for both measured and numerical data. Since the simulation set has a larger amount of data than the in-situ data, the model behaved much better for the simulated data set. It concluded that as the number of data increased, the accuracy of the proposed model increased.

3.2. Physics-Informed Machine Learning (PIML)

Physics-Informed Neural Networks (PINNs) are a modern scientific approach to solving partial differential equations (PDEs) by combining ML techniques and governing physical laws. The PINNs show greater potential since the traditional numerical methods, such as finite element or finite difference methods, face challenges with complex mesh generation, difficulty with noisy experimental data integration, or higher-dimensional parametric PDEs or ill-posed inverse problems.[106] PINNs offer a mesh-free alternative method for solving PDEs by integrating the physical laws directly into the training process. PINNs approximate PDE solutions by training the NNs to minimize the loss function that includes the governing PDE, initial conditions (ICs), and boundary conditions (BCs). It can be considered as an unsupervised strategy since PINNs do not require labeled data from previous simulations or experiments.[107] Usually, the inputs of the PINNs are the independent variables, such as spatial coordinates and time, and the outputs are the dependent variables, such as temperature, for a heat conduction problem. The fundamental process of PINNs is shown in Fig.10. This PINN has two spatial and a temporal input, with two hidden layers, where $\{w^{(1)}, w^{(2)}, w^{(3)}\}$ indicates the weights/parameters of each layer, $\{a^{(2)}, a^{(3)}\}$ are the vector of activated neurons, and $\hat{u}$ is the predicted output. The output of the NN is given to the Automatic differentiation (AD) component to compute the derivative to obtain the residual and integrated into a loss function to improve the accuracy of the model.[108] Another major difference between PINNs and conventional PDE solvers is that PINNs consider the solution process as a regression, where the number of data points or equations is greater than that of unknown parameters. This is in contrast with conventional methods, which typically have the same number of equations and unknowns.[106] This section discusses the applications of

PINNs to DA and inverse problems in thermal systems.

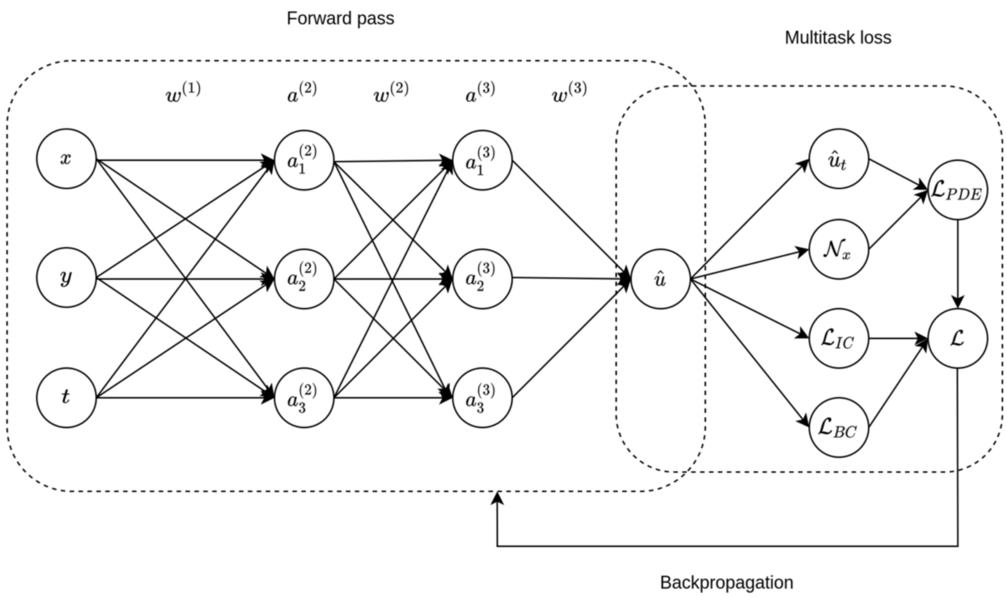
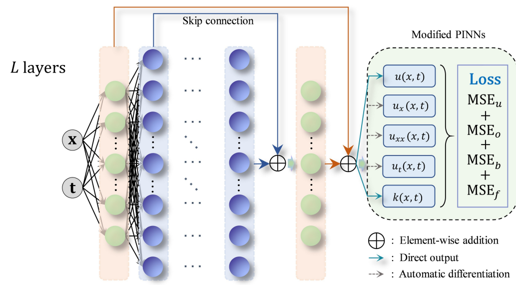


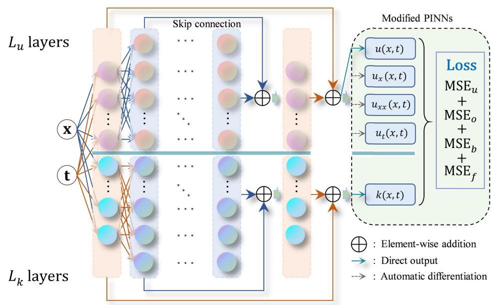
FIG. 10: Schematic of PINN architecture.[106]

Recent research shows that PINNs are gaining popularity among researchers for solving IHTPs. To obtain the heat flux in a 1D IHCP, a PINN was developed.[109] The neural network was trained with physical constraints such as governing laws, boundary conditions, initial conditions, and temperature data. The performance of the PINN was observed by changing the activation function, network architecture, and different forms of heat flux. The results show reasonable performance in predicting different heat flux forms from temperature data with noise. Furthermore, IHCPs in gas turbines rotating cavities were investigated using PINNs.[110] The objective was to predict the surface heat flux value from the experimentally captured radial temperature profile. The PINN was trained with synthetic data without noise generated from 2D FE analysis and noisy experimental data. The predicted heat flux values were in good agreement with the FE solution. Another application of PINNs to solve IHTPs within an aluminum wall subjected to heat flux on one side and thermal insulation on the other side was proposed.[46] The objective was to predict the temperature profile as the forward problem and heat flux and thermal diffusivity as the inverse problem from limited experimental data. The experimental data of temperature at random points was generated using an analytical solution. The network takes spatial and temporal independent variables as input and estimates the dependent variable, temperature, as the output for each iteration. At each iteration, a loss function was calculated and optimized using the Adam optimizer. For the forward problem, the model predicted the temperature with higher accuracy, with an absolute error of 1.01^oC for the no noise case and 1.32^oC for the case of 10% noise. In the case of the inverse problem, the model was able to predict the heat flux and corresponding temperature with higher accuracy for both noise-free and noisy data.

To address the direct and indirect heat conduction problem, a novel ML framework based on PINNs is discussed by He et al.[43] An improved method was proposed, where the loss function takes advantage of training data, mean squared error (MSE) of initial and boundary conditions, and enhances the training convergence. Unlike the normal PINNs, the proposed method considers an inversion algorithm for the variable parameters. The authors introduced two network architectures coupled with unknown parameters. Type A is a single network with two kinds of outputs, temperatures (u) and an unknown parameter (k) of the inverse problem, and Type B consists of two networks, one for u and the other for k. A schematic of the proposed models is shown in Fig11. To reduce the complexity and increase the accuracy in the inversion of the variable parameter, skip connections[111] were implemented in both networks. Skip connections help the model to handle the degradation problem in NN due to the higher depth of the architecture. The result of the study indicated that the proposed method exhibits higher accuracy and convergence rate than the conventional PINNs.[112] Further to analyze the efficacy of the model, random Gaussian noise levels of 5%, 10%, and 15% were added to the raw data, and the model performed with higher accuracy, even when the noise reached 15% for a problem with one constant inverse parameter. However, the inversion of a two-parameter problem shows less accuracy than single-parameter inversion. For the case without noise, both networks achieved similar accuracy, but deteriorated when noise was introduced. The Type A model exhibits less error growth and higher robustness under noise compared to the Type B model. This study shows the potential and limitations of the proposed study for IHCPs.



(a) Type A



(b) Type B

FIG. 11: Proposed architectures of coupled neural networks (Reprinted with permission from Elsevier, Copyright 2021).[43]

Among various approaches to solving the IHCPs, an ensemble physics-informed neural network (E-PINN) with adversarial training (AT) was explored by Jiang et al.[113] This model was primarily focused on solving and quantifying the uncertainties of space-dependent IHCPs. An ensemble of PINNs trained with different parameters was used for the predictions and uncertainty estimation. Fig.12 shows the proposed E-PINN model where (a) the quantity of interest **m** is represented separately in the neural network, (b) the parameter **m** is represented based on the derivatives of the temperature field **d**, (c) training of PINN, and (d) predicted statistics from the ensemble of models. Results show that only five models, $M = 5$, were sufficient to obtain high-quality uncertainty estimates. For the heat flux inversion problem, the proposed method exhibits higher accuracy compared to other classical inversion methods. The model exhibits a relative L2 error of 0.1034 and 0.0810 for a noise level of 0.05 with and without adversarial training, respectively. Similar performance was observed in the material diffusivity inversion problem.

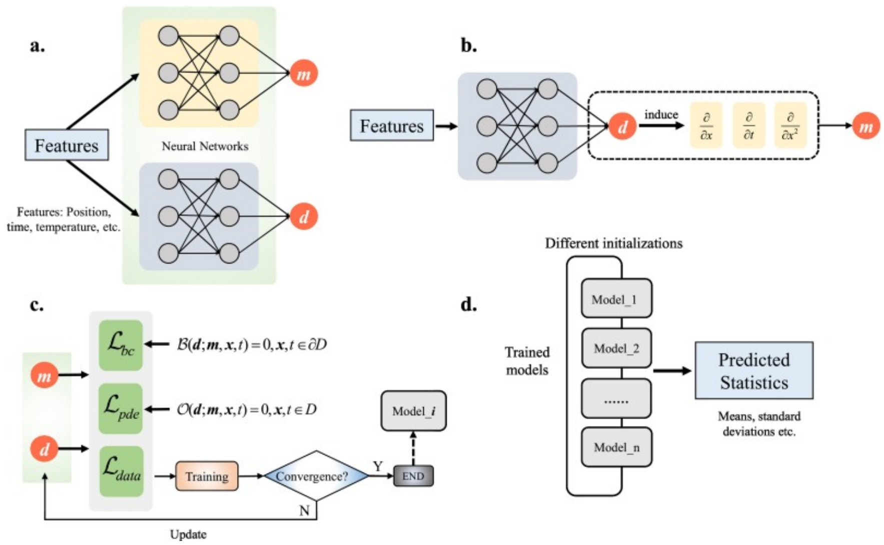
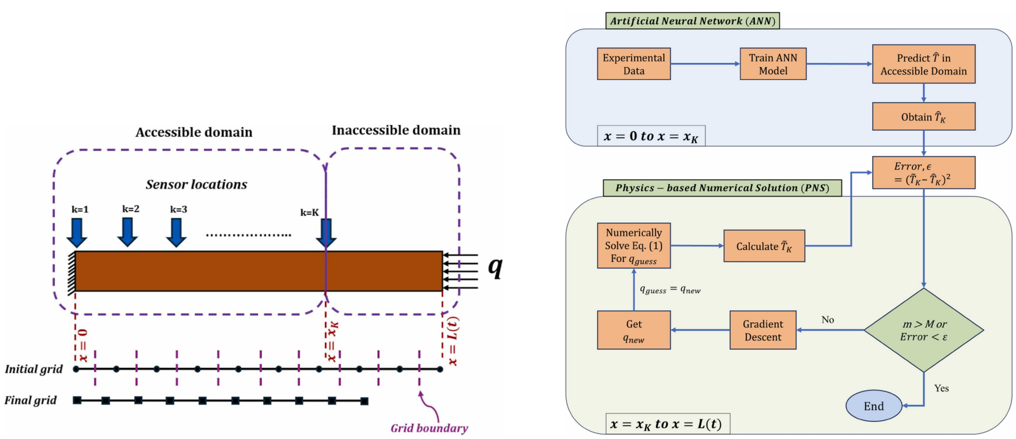


FIG. 12: Components of E-PINN for space-dependent IHCPs (Reprinted with permission from Elsevier, Copyright 2023).[113]

Generally, due to extreme thermal conditions and moving boundaries, it is challenging to predict the temperature and heat flux in ablative materials. To tackle such an IHTP to find heat flux in ablative materials, a hybrid physics-based method was proposed.[114] For this study, a computational domain that changes with time was considered. The left end of the domain was insulated, and the right boundary was subjected to an unknown heat flux. And the entire domain was divided into accessible and inaccessible domains as shown in Fig.13a. The accessible domain corresponds to having fixed sensor data, and the domain to the right of the starting point from the last sensor was considered as an inaccessible domain where the temperature data was unknown. Initially, the PINN-based model was used to predict the temperature, and it behaved well when the boundary data were available. However, when a lack of boundary data in the training set, the PINN and ANN were unable to predict the unknown heat flux at the right boundary. It was noticed that ANN underpredicted and PINN overpredicted the temperature value at the right boundary. Especially when the heat flux was very large, the PINN model was unable to converge. To address these shortcomings, a hybrid model was introduced with an ANN model used in the accessible region and a physics-based numerical solution (PNS) for the inaccessible domain, as illustrated in Fig.13b. The trained ANN model was only used to predict the temperature at the accessible domain ($\hat{T}_k$) and to predict the temperature at the last sensor location. On the other hand, PNS iteratively obtained the temperature ($\tilde{T}_k$) in the inaccessible region and hence predicted the heat flux on the right boundary until $\tilde{T}_k$ approaches $\hat{T}_k$. For the direct problem of predicting temperature distribution with constant thermal properties and heat flux, the hybrid model demonstrates high accuracy. For the inverse problem, the hybrid model exhibited higher accuracy in predicting heat flux at the right boundary. This study demonstrates the potential of integrating numerical results into a data-driven model to enhance the accuracy and efficiency of IHTPs.



(a) **(b)**

FIG. 13: (a) Computational domain and (b) Schematic of ANN-PNS algorithm flowchart (Reprinted under a Creative Commons Attribution 4.0 International license).[114]

Reconstruction of the temperature field is often challenging in various engineering applications. Insufficiency of measured data or sparse measurements may result from challenging in-situ environments, limitations in instrument resolution, cost, and sensor availability. A recent study by Wang et al.[45] addressed this issue and proposed a novel method, the analytical solution-embedded physics-informed neural network (A-PINN), to reconstruct the temperature field from sparse data. Initially, the Fourier transform was used to simplify and reduce the degrees of freedom of the ill-posed problem. This helped to reduce the 3D-continuous temperature field into a series of 2D frequency-domain cross-section distributions that vary along the z-axis. The proposed model A-PINN, has similar features to a CNN. The input has sparse measurements, spatial coordinates, and the corresponding temperature and initial layer with two channels, which calculates weight functions related to the frequency and spatial variable. Hidden layers with four channels estimate the solution of temperature and heat flux in the frequency domain, and later combine these two in two channels; finally, combining these two channels, the inverse Fourier transform was applied to obtain the temperature field and the heat flux. The results indicated that the A-PINN model achieved a higher accuracy in 3D temperature field reconstruction with less than 1% of average error in maximum temperature.

Another study[115] addressed a hybrid physics-based DL thermal model of additive manufacturing (AM) using PINNs for estimating the temperature field and unknown material and process parameters from limited observed temperature data. The main difference between the PINNs and the proposed method is the addition of a data-based loss term in the loss function. The authors outlined that the advantages of the added data loss are a) the labeled data can act as auxiliary data to guide and accelerate the training to solve forward problems,b) solve inverse problems, and c) create a hybrid-physics-informed model using arbitrary experimental data in governing laws. One numerical and one experimental example were discussed to show the accuracy and usability of the proposed model. In the numerical example, the model predicted laser absorptivity, heat capacity, and thermal conductivity with less than 5% of error from partially observed temperature data. A full-field temperature was obtained with a root mean squared error (RMSE) of 47.28K from partially observed data from experiments. These results show a promising avenue for this proposed network to model thermal problems in additive manufacturing.

The study by Go et al.[116] introduced a PINN-based surrogate model for a virtual thermal sensor (VTS) with real-time simulations. The proposed surrogate model takes limited sensor measurements from the real simulations and estimates the thermal profile or heat flux in the areas without the sensors. Initially, a self-adaptive PINN (SA-PINN) was discussed, where the loss weight during the training was adjusted adaptively. This strategy offers a flexible loss weight to each training point rather than the entire loss component. Physical sensor temperature measurements, along with spatial and temporal data, were used as input to the PINN-based surrogate model to predict the complete thermal profile. Losses were calculated similarly to SA-PINN, and when calculating the PDE loss, the gradient based on the input sensor data was also included. Several examples were discussed, along with a 2D plate with an unknown body heat flux. In this example, the proposed model was used to estimate the temperature profile and unknown heat flux from the limited sensor data. Training with no-noise data, the model predicts the temperature and heat flux with higher accuracy when compared with the ground truth. However, with 3% noise, the predicted solution on heat flux obeys the ground truth trend with some oscillations. The proposed model predicted the solution in 158.746ms compared to 71s by the commercial software Abaqus. However, 152,984s were utilized for the training of the model.

An alternative model was proposed for the reconstruction of temperature response using a transfer machine learning framework.[117] A pre-model was obtained from the simulation and experimental models and then provided to the transfer learning techniques along with the experimental model to reconstruct the temperature field and obtain thermophysical parameters. Initially, generated data from the matched simulation model from the experimental model was fed to the DNN, and temperature reconstruction was achieved through forward propagation. The gradient descent optimization algorithm was used to update the network parameters by constraining the loss function. Through transfer learning, the final PINN model was achieved from the pre-model with limited experimental data. A schematic of the proposed transfer machine learning technique is illustrated in Fig14. Moreover, the authors introduced three fine-tuning strategies by freezing 0, 2, and 4 network layers of the DNN consisting of one input, six fully connected, and one output layer, respectively. Training time for these fine-tuning strategies shows that strategy 2 exhibits higher performance with optimal neuron parameters compared to the other strategies. From several examples discussed, it was concluded that the proposed model has significant advantages, such as training efficiency and accuracy over typical DNNs, due to the physics-driven adaptation of the pre-model and transfer learning approach.

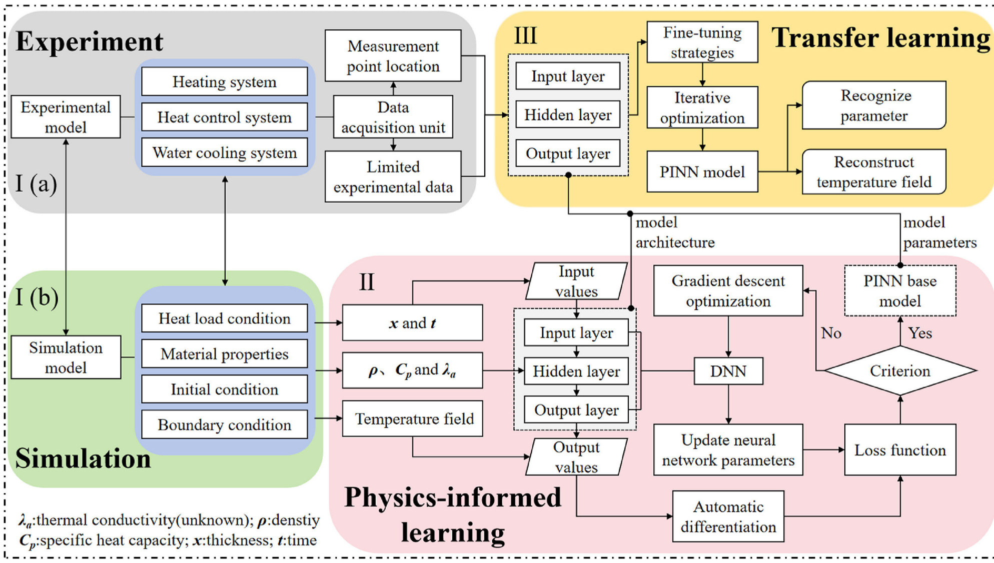


FIG. 14: Schematic of Transfer Machine learning model (Reprinted with permission from Elsevier, Copyright 2025).[117]

A modified PINN is explored for estimating thermal conductivity from noisy and limited experimental data.[118] The proposed model contains multiple dedicated NNs, each with a specific purpose. The architecture was divided into three segments. In the first step, normalization of the spatial and temporal data was performed using the maximum scaling technique. In step 2, the limited dataset was expanded from the initial measured data using physical equations. In this step, three different NN segments were introduced, named Pre-NN-A1, Pre-NN-A2, and Pre-NN-B, based on the conventional PINN architecture. The purpose of Pre-NN-A1 was to augment the labeled data and estimate $(\frac{dT}{dt})$ as a parameter for every time step, and this data was used to train Pre-NN-A2. Next, the temperature data at every time step obtained from Pre-NN-A2 was subjected to a reality test by checking whether the predicted temperature follows the 1D heat transfer law. Later, Pre-NN-B was trained with data that passed the reality test. A significantly increased amount of labeled data, compared to the initial set, is obtained from this process and later provided to the Main-NN, where the unknown physical parameter is predicted. Results show that the modified PINN significantly outperformed the conventional PINN in terms of accuracy and computational time for all the validation cases. Even with only three data points for each time step, the modified PINN produces excellent results consistently.

Tanaka and Nagai[119] extended the application of PINN to the thermal management of spacecraft systems. The objective was to obtain the complete temperature distribution from limited sensor data and model uncertainties. They proposed a physics-informed machine learning (PIML) thermal analysis with limited temperature measurements, with Thermal mathematical model (TMM) constraints. Initially, FFNN was built with nodal coordinates as inputs and nodal temperature as the outputs. Next, residual loss, observation loss, and boundary loss were calculated, and the solution was optimized to satisfy the governing law and observed data (pseudo-observation). This study demonstrates that PIML substantially improved temperature prediction with higher accuracy. The authors outlined that by using PIML, it is possible to predict the temperature of an in-orbit spacecraft from limited sensor data and incomplete TMM. Table 1 presents a detailed summary of the reviewed papers, highlighting the types of machine learning methods used, problem objectives and dimensionality, data requirements, and accuracy.

3.3. Digital Twinning of Thermal Systems

In recent years, the Digital Twins (DT) concept has gained significant attention in various fields, including thermal systems[120], healthcare[121], automotive industry[122], medical field[123], and agriculture[124]. In simple terms, DT is a real-time virtual replica of a physical system that enables real-time monitoring and optimization.[125] Any model that accurately represents the behavior of the physical system can be used as a digital twin.[126] For example, the digital twin can be generated based on CAD (computer-aided design) models or CAE (computer-aided engineering) simulation. However, the traditional 2D/3D simulations are far from being able to exchange real-time data with their physical counterpart.[127] To mitigate these challenges, a surrogate model, typically data-driven, can be developed, which processes information more quickly than any physics-based model and

TABLE 1: Summary of ML-based inverse heat transfer studies

Reference	ML Method	Inverse Problem	Data	Dimension	Accuracy
Tamaddon-Jahromi et al.[41]	ANN	Boundary conditions from sparse temperature measurements in linear and nonlinear heat conduction/convection problems.	FEM	2D	Linear heat conduction problem, the model accuracies of 75%, 98%, and 98% were achieved to predict the boundary temperatures given three, four, and five locations inside the domain geometry, respectively.
Zhang et al.[61]	ANN	Thermo-physical properties (density, thermal conductivity, heat capacity) from thermal probe measurements.	FEM & Experiment.	2D-Axisymmetric	Total 12 validation parameters, the errors of ten cases are less than 10%, and the other two cases are 12.23% and 13.17%.
Haolong and Zhanli[62]	ANN	Identifying the thermal conductivity and inner wall shape of a pipe.	Stochastic model & FEM	2D	As measurement error decreases, the inversion results become more accurate.
Wang et al.[65]	Nonlinear ANN	Estimating hot-wall surface heat flux from internal temperature measurements.	Experiment	1D	Estimated heat flux agrees with calibrated values.
Allard and Najafi[66]	FFNN and NARX	Surface heat flux using internal temperature measurements (with and without a moving boundary, and with and without temperature-dependent material properties).	FEM	1D	Performed well in robustness tests and predicted accurate, consistent results, especially when measurements were near the boundary surface.

TABLE 1: *Continued*

Reference	ML Method	Inverse problem	Data	Dimension	Accuracy
Najafi and Woodbury[67]	ANN	Near-real-time heat flux using temperature measurements for constant and temperature-dependent material properties.	Analytical	1D	Able to capture various heat flux profiles for constant and temperature-dependent material properties.
Mirsephai et al.[68]	ANN	Identifying absorbed heat in radiation-dominated problems.	Experimental data	2D	Predicted heat was accurate and consistent with experimental values.
Virupaksha et al.[84]	CNN	Predicting domain heterogeneity from temperature maps.	FEM	2D	3D-CNN with transient data achieved greater accuracy and outperformed 2D-CNNs.
Lockwood et al.[88]	CNN	Evaluating building parameters (e.g., wall insulation conductivity) from time series data.	FDM	3D	Promising performance, especially for smaller periods.
Masrouri and Tahsini[90]	CNN	finding heat transfer coefficient and generated temperature of an industrial heat gun	DNS	1D	Parameter predictions were slightly different from traditional methods, but predicted temperature profiles matched with measurements.
Zhu et al.[91]	CNN-LSTM	Thermal boundary parameters for a transient IHTP.	CFD	3D	Significantly superior to standalone CNNs and LSTM models.
Niu et al.[92]	CNN-LSTM	Reconstruction of the 3D flame temperature field.	Time series data from light field tomography.	3D	High accuracy.
Zwart et al.[98]	LSTM	Real-time forecasting of stream water temperature.	Observed data	3D	Improved forecast performance with DA-LSTM compared to a standalone LSTM model.
Wang et al.[99]	ConvLSTM	Transient heat flux from measured temperatures.	FEM	3D	Higher accuracy with error of 1.07% for regular adat set and 6.06% for complex data set.

TABLE 1: *Continued*

Reference	ML Method	Inverse problem	Data	Dimension	Accuracy
Bielajewa et al.[15]	LSTM and Transformers	Transient thermal field reconstruction.	FDM & FEM	1D and 2D	Consistently performed better than transformer models for 1D and 2D problems.
Han et al.[104]	LSTM-based Encoder-Decoder (ED)	Real-time prediction of heat flux in the tool-chip region during the turning process.	FEM	3D	Outperformed other NN models (MLP, CNN, LSTM).
Wang et al.[105]	Physics-guided LSTM	Predicting the dynamic heat load of buildings.	In-situ measurements and numerical simulations (TRNSYS).	3D	Greater accuracy, especially with a larger dataset from simulations.
Qian et al.[109]	PINN	Prediction of heat flux.	Sampled temperature data.	1D.	Reasonable performance in predicting different heat flux forms.
Puttock-Brown et al.[110]	PINN	Predict surface heat flux from radial temperature profiles.	FEM	2D	Predicted heat flux values were in good agreement with the FE solution.
Billah et al.[46]	PINN	Predict thermo-physical parameters such as the material's thermal diffusivity and heat flux.	Analytical solution	1D	model predicted the temperature with higher accuracy, with an absolute error of 1.01^{o}C for the no noise case and 1.32^{o}C for the case of 10% noise"
Jiang et al.[113]	E-PINN with adversarial training (AT)	Solving and quantifying the uncertainties of space-dependent IHCPs (e.g., heat flux and material diffusivity inversion).	FEM	2D	Exhibits higher accuracy compared to other classical inversion methods.

TABLE 1: *Continued*

Reference	ML Method	Inverse problem	Data	Dimension	Accuracy
He et al.[43]	PINN with skip connections	Solving direct and indirect heat conduction problems, and inverting variable parameters.	Training data	1D	Higher accuracy and convergence rate than conventional PINNs.
Islam and Dutta[114]	ANN + Physics-based Numerical Solution	Finding heat flux in ablative materials with moving boundaries.	FVM	1D	High accuracy in predicting heat flux at the boundary. Outperformed standalone ANN and PINN models.
Wang et al.[45]	Analytical solution-embedded PINN (A-PINN)	Reconstructing a 3D temperature field from sparse measurements.	Sparse measurements	3D	Achieved higher accuracy with less than 1% average error in maximum temperature.
Liao et al.[115]	PINN	Estimating temperature fields and unknown material/process parameters in additive manufacturing.	FEM	3D	Predicted parameters with less than 5% error.
Go et al.[116]	Self-Adaptive PINN (SA-PINN)	Surrogate model for a virtual thermal sensor (VTS) to estimate thermal profiles or heat flux from limited sensor measurements.	FEM	1D and 2D	High accuracy compared with ground truth, even with measurement noise.
Chen et al.[117]	PINN and Transfer Learning techniques	Reconstructing the temperature field and obtaining thermophysical parameters.	FEM and experiments	1D	High accuracy over typical DNNs.
Jo et al.[118]	PINN	Estimating thermal conductivity from noisy and limited experimental data.	FVM and experimental	1D and 2D	Significantly outperformed conventional PINN in accuracy. Produced excellent results consistently.
Tanaka and Nagai[119]	PINN	Temperature distribution.	Thermal Mathematical Model (TMM) data	3D	Substantially improved temperature prediction with higher accuracy.

efficiently updates the physical state.[128]

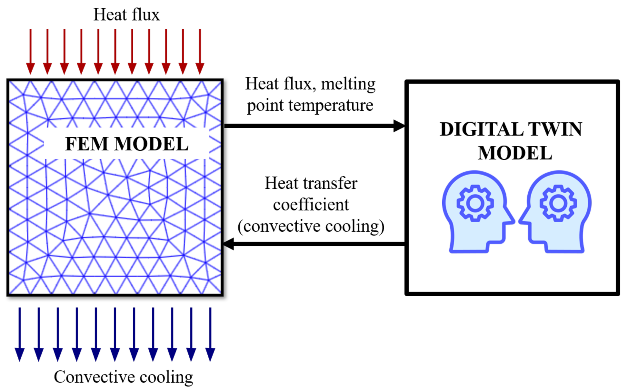
Yang et al.[129] proposed a DT framework for optimal control of the continuous casting process. This framework consists of a high-accuracy fundamental model, coordinate optimization, and dynamic control of cooling and stirring. A high-fidelity heat transfer model was used as the surrogate model and calibrated with a combination of offline and online data with NNs. Metallurgical laws were incorporated for optimization and solved using the particle swarm optimization algorithm. The results demonstrated a significant reduction in micro-segregation and improvement in strand quality using this DT framework. Integration of proper orthogonal decomposition(POD) and ANNs as digital twins was proposed for thermal stratification in pressurizer surge lines.[128] A data-driven snapshot-based method was used to extract the dominant lower modes of the fluid field and project the complete field onto these modes to obtain the reduced state coefficients using POD. For rapid prediction of the system behavior, an ANN is trained to learn a mapping between the inputs and the POD-reduced state coefficients. The performance of the proposed DT was monitored by its determination coefficient (R^2) and relative error cloud plots. The proposed digital twinning method successfully reconstructs the temperature and stress variables. Moreover, the computational time required for the proposed method was significantly lower than that of traditional solvers.

Ma et al.[130] discussed a digital twin-assisted deep transfer learning method (DTL) for thermal error modeling of electric spindles. This method was introduced to address the limited or unavailable thermal error data due to complications in data measurement and variable working conditions. DT-generated thermal data accounts for the training data in the source domain, and limited or unavailable data from the physical system corresponds to the transfer learning of the target domain. They developed a distance-guided domain adversarial network (DGDAN) to reduce the inconsistency by establishing a connection between the virtual and real data. The performance of the proposed methods was validated using simulation data and real data under variable operating conditions from DT models and a physical system. The results show that the combination of digital twins (DT) and transfer learning is a viable tool to mitigate the challenges in thermal error modeling, especially when thermal error samples are limited or unavailable.

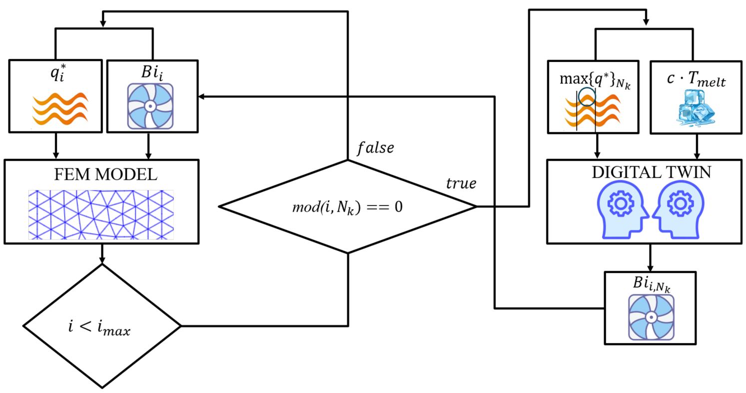
A digital twinning of a transient heat transfer system in a solid medium was proposed by Di et al.[127] The objective was to identify the heat transfer coefficient to dynamically control the maximum temperature below a threshold when transient heat flux is provided to the thermal system. A DNN was used as the digital counterpart of the physical system, while an FEM model served as the surrogate physical system. A 2D FE model, as a physical system, with transient heat flux was applied to the top surface, and a convective heat transfer boundary with an unknown heat transfer coefficient was applied to the bottom. This unknown heat transfer coefficient was calculated by the DT (trained model). The two functions of the FEM model was to generate data to train the neural network and test the DT for its performance. Once the NN was trained, validated and tested, a coupling between the physical system (FEM) and NN was established as shown in Fig15a. For the training of the NN, only steady-state data were used, which were sufficient to map the relationship between the heat flux, heat transfer coefficient, and maximum temperature. In this proposed method, at certain time intervals, the heat flux and maximum allowed temperature data are fed to the DT, and the time-dependent heat transfer coefficient is estimated. The flow chart of the proposed digital twinning is illustrated in Fig.15b. To test the performance of the DT, various heat flux distributions are provided. Results show that for all cases, DT estimated the appropriate convective heat transfer coefficient and limited the maximum temperature below the threshold to keep the system safe. These observations are extended to 3D cases.

3.4. Challenges and Future Outlook

The review of machine learning applications in the inverse problems of heat transfer shows great potential but also comes with significant challenges and research gaps. Few of these issues are addressed in this section. A primary issue is the handling of the data, which is an important factor in all machine learning models. Discussed applications, whether DNNs, CNNs, or RNNs, rely on large, clean, and labeled datasets, often synthetically generated from computationally expensive simulations. However, this scenario does not represent the reality, since the experimental data are inherently sparse and corrupted by noise. Studies show that even small levels of noise could introduce oscillations in their predictions, and machine learning models such as CNN-LSTM or modified PINN strategies show some degree of robustness to noise and sparsity. This necessitates the development of more advanced data augmentation and pre-processing techniques. Future focus must be on the research of various models that can learn efficiently from heterogeneous data with seamless



(a)



(b)

FIG. 15: (a) Schematic of FEM-DT coupling (b) Flowchart of ML-based DT.[127]

integration of sparse experimental measurements. Another critical challenge of machine learning models is their black-box nature, which can severely impact their interpretability and explainability. For most engineering problems, it is important for engineers and scientists to understand the underlying physical mechanisms and reasons for the model predictions rather than just providing accurate results. The review discusses some problems where DNN informs the FE model, but the understanding of how the ML model reaches its conclusion is opaque. However, PINN-based models offer better transparency and are more interpretable by embedding the governing equations than purely data-driven models. Future research needs to study methods revealing the black-box nature of the models for better understanding, transparency, and insights into the physical systems being modeled.

Another major hurdle observed from this review is the scalability of ML methods for higher dimensions and complex problems. While some of the ML methods are efficient in 3D applications, they often come with higher computational cost during the training phases. The training time can be very high for problems like iterative design-optimization and real-time inverse problems, making it difficult to solve without massive computational resources. Reviewed studies show that the greater computational efficiency during inference is a significant advantage over traditional models; however, there is very little discussion on the computational resources used during training and testing. Without the proper discussion on the computational resources at every stage of these ML models, it is difficult for widespread adoption of these for complex and novel problems. This challenge is also linked to the problem of generalization. A model trained for specific geometry or a set of boundary conditions may perform poorly on a new, slightly different problem. Although the ML models are used to reduce the computational burden, due to a lack of generalization, they require a computationally intensive training cycle for each new problem. Future efforts should focus on developing better methodologies that can generalize more effectively across different problems. Another important observation made from this review is the lack of standardized benchmarks and comprehensive comparison studies between different ML models. Even though the review provided a list of different methods and their applications, an in-depth comparison of their performance on a standard problem is rare. This scenario makes it difficult to make an informed decision about the true applicability and efficiency of various methodologies and to guide the engineers and researchers in selecting the most appropriate strategy. Suggestion from the authors to the community is to agree on a set of various benchmark problems, from linear to non-linear, 1D to 3D, and different levels of noise and sparsity, to enable an effective comparison. Furthermore, there is very limited research on the coupling of various methods presents a significant opportunity. Some of the discussed coupled models show great potential in solving IHTPs. Combining the best features of the different ML models, for example, using a CNN to extract spatial features and PINN to enforce governing law and solve the inverse problem, could lead to robust hybrid models. These next-generation models need to be well documented and assessed against standardized problems to provide reliable evaluations and comparisons that can lead to widespread adoption in engineering and scientific applications.

4. CONCLUSION

This present study provides a detailed review of significant progress and ongoing development in the field of machine learning techniques in solving inverse problems of thermal systems. Traditionally, solution strategies for the inverse heat transfer problem have heavily depended on mathematical and numerical methods such as regularization, sequential estimation, optimization algorithms, or semi-analytic approaches. These methods have shown great potential to tackle inverse problems for decades; however, they have often struggled with their sensitivity to noisy data, sparse or limited data, and unstable or non-unique solutions. Data assimilation techniques such as the Kalman filter or variational methods further improved the ability to integrate real-time measurements into the model for enhanced predictions, improved robustness in parameter estimation, or field reconstruction. Recent advances in machine learning frameworks and deep learning models have provided a new direction for addressing the limitations of traditional methods. Deep neural networks have shown strong potential for solving inverse heat transfer problems, especially when a physics-based model is integrated into the machine learning framework. Hybrid models, such as DNNs for accessible domains and numerical methods for the inaccessible domains, can outperform standalone machine learning models. The incorporation of transfer learning and data-based loss functions to DNNs shows significant improvement in prediction accuracy under limited experimental data. Furthermore, CNN-inspired architecture has demonstrated exceptional accuracy in temperature field reconstruction from sparse and noisy data. These models leverage the ability to process spatially distributed data and extract features, allowing efficient transformation of the sparse data into accurate thermal and flux distributions. Incorporating the governing laws into the deep learning model further provides a new way to solve inverse problems efficiently. In the absence of large training datasets, PINN-based architecture often helps by using physical laws to minimize loss and generalization.

Furthermore, research shows that various hybrid methodologies and modifications of vanilla PINNs can further enhance the prediction accuracy. The synergy between inverse problem-solving techniques, data assimilation, and a machine-learning framework leads to digital twins. These digital twins can assimilate the data in real time, update predictions, and optimize the system dynamically. These digital twin frameworks represent a transformative step in the practical application of inverse and data assimilation techniques. Although these machine learning techniques show great potential, it is important to address the current limitations and future research directions. The discussed ML techniques used for the IHTPs are often considered as a surrogate model. For higher accuracy and prediction capability, it is important to train these models with a large amount of data, which is usually complex and computationally challenging. The research studies demonstrated the performance of the proposed models in terms of computational time and provided little information on the training time. Many data-driven methods or hybrid models have shown high accuracy for a given dataset, and they may not generalize well to new or different scenarios. Most articles only explored a single or a couple of methods, and this restricts the adoption of a suitable method for IHTPs. Furthermore, there is no clarity on whether the proposed methods work for different types of problems. Some of these challenges and future directions are also discussed.

Acknowledgments

This work has been funded by the Fusion Futures Programme. As announced by the UK Government in October 2023, Fusion Futures aims to provide holistic support for the development of the fusion sector.

References

1. Hadamard, J., Lectures on Cauchy's Problem in Linear Partial Differential Equations, Courier Corporation, 2014.

2. Tikhonov, A. N., and Arsenin, V. I. A. K., Solutions of Ill-Posed Problems, 1977.

3. Beck, J. V., Litkouhi, B., and Clair Jr, C. R. S., Efficient Sequential Solution of the Nonlinear Inverse Heat Conduction Problem, *Numerical Heat Transfer, Part A Applications*, vol. **5**, no. 3, pp. 275–286, 1982.

4. Taler, J., and Zima, W., Space-Marching Method for Two-Dimensional Inverse Heat Conduction Problems, *Encyclopedia of Thermal Stresses*, pp. 4473–4489, 2014.

5. Marquardt, D. W., An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *Journal of the Society for Industrial and Applied Mathematics*, vol. **11**, no. 2, pp. 431–441, 1963.

6. Hào, D. N., and Reinhardt, H.-J., Gradient Methods for Inverse Heat Conduction Problems, *Inverse Problems in Engineering*, vol. **6**, no. 3, pp. 177–211, 1998.

7. Wang, B., Sun, Z., Jiang, X., Zeng, J., and Liu, R., Kalman Filter and Its Application in Data Assimilation, *Atmosphere*, vol. **14**, no. 8, pp. 1319, 2023.

8. Asch, M., Bocquet, M., and Nodet, M., Data Assimilation: Methods, Algorithms, and Applications, SIAM, 2016.

9. Jung, S., Lee, K., Park, C., and Choe, J., Ensemble-Based Data Assimilation in Reservoir Characterization: A Review, *Energies*, vol. **11**, no. 2, pp. 445, 2018.

10. Arridge, S., Maass, P., Öktem, O., and Schönlieb, C.-B., Solving Inverse Problems Using Data-Driven Models, *Acta Numerica*, vol. **28**, pp. 1–174, 2019.

11. Ongsulee, P., Artificial Intelligence, Machine Learning and Deep Learning, *2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE)*, pp. 1–6, 2017.

12. Tikhonov, A. N., Goncharsky, A. V., Stepanov, V. V., Yagola, A. G., Tikhonov, A. N., Goncharsky, A. V., Stepanov, V. V., and Yagola, A. G., Numerical Methods for the Approximate Solution of Ill-Posed Problems on Compact Sets, Springer, 1995.

13. Freitag, M. A., and Potthast, R. W. E., Synergy of Inverse Problems and Data Assimilation Techniques, *Large Scale Inverse Problems—Computational Methods and Applications In the Earth Sciences*, vol. **13**, 2013.

14. Tarantola, A., Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, 2005.

15. Bielajewa, W., Tindall, M., and Nithiarasu, P., Comparative Study of Transformer-and LSTM-Based Machine Learning Methods for Transient Thermal Field Reconstruction, *Comput. Thermal Sci.: An Int. J.*, vol. **16**, no. 3, 2024.

16. Sanz-Alonso, D., Stuart, A. M., and Taeb, A., Inverse Problems and Data Assimilation, *arXiv preprint arXiv:1810.06191*, 2018.

17. Wang, B., Zou, X., and Zhu, J., Data Assimilation and Its Applications, *Proc. Natl. Acad. Sci.*, vol. **97**, no. 21, pp. 11143–11144, 2000.

18. Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G., Data Assimilation in the Geosciences: An Overview of Methods, Issues, and Perspectives, *Wiley Interdiscip. Rev. Climate Change*, vol. **9**, no. 5, pp. e535, 2018.

19. Evensen, G., Vossepoel, F. C., and Van Leeuwen, P. J., Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem, Springer Nature, 2022.

20. Arcucci, R., Zhu, J., Hu, S., and Guo, Y.-K., Deep Data Assimilation: Integrating Deep Learning with Data Assimilation, *Appl. Sci.*, vol. **11**, no. 3, pp. 1114, 2021.

21. Nichols, N. K., Mathematical Concepts of Data Assimilation, *Data Assimilation: Making Sense of Observations*, pp. 13–39, 2010.

22. Kirsch, A., and others, An Introduction to the Mathematical Theory of Inverse Problems, vol. **120**, Springer, 2011.

23. Vogel, C. R., Computational Methods for Inverse Problems, SIAM, 2002.

24. Hansen, P. C., Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, SIAM, 1998.

25. Levin, E., and Meltzer, A. Y., Estimation of the Regularization Parameter in Linear Discrete Ill-Posed Problems Using the Picard Parameter, *SIAM J. Sci. Comput.*, vol. **39**, no. 6, pp. A2741–A2762, 2017.

26. Karl, W. C., Regularization in Image Restoration and Reconstruction, *Handbook of Image and Video Processing*, pp. 183–V, 2005.

27. Zhang, Z., and Moore, J. C., Mathematical and Physical Fundamentals of Climate Change, Elsevier, 2014.

28. Le Dimet, F.-X., and Talagrand, O., Variational Algorithms for Analysis and Assimilation of Meteorological Observations: Theoretical Aspects, *Tellus A: Dyn. Meteorol. Oceanogr.*, vol. **38**, no. 2, pp. 97–110, 1986.

29. Ponomareva, K., and Date, P., An Exact Minimum Variance Filter for a Class of Discrete Time Systems with Random Parameter Perturbations, *Appl. Math. Model.*, vol. **38**, no. 9-10, pp. 2422–2434, 2014.

30. Buizza, C., Casas, C. Q., Nadler, P., Mack, J., Marrone, S., Titus, Z., Le Cornec, C., Heylen, E., Dur, T., Ruiz, L. B., and others, Data Learning: Integrating Data Assimilation and Machine Learning, *J. Comput. Sci.*, vol. **58**, pp. 101525, 2022.

31. Gilbert, R. C., Richman, M. B., Trafalis, T. B., and Leslie, L. M., Machine Learning Methods for Data Assimilation, *Computational Intelligence in Architecturing Complex Engineering Systems*, pp. 105–112, 2010.

32. Brajard, J., Carrassi, A., Bocquet, M., and Bertino, L., Combining Data Assimilation and Machine Learning to Infer Unresolved Scale Parametrization, *Philos. Trans. R. Soc. A*, vol. **379**, no. 2194, pp. 20200086, 2021.

33. Cheng, S., Quilodrán-Casas, C., Ouala, S., Farchi, A., Liu, C., Tandeo, P., Fablet, R., Lucor, D., Iooss, B., Brajard, J., and others, Machine Learning with Data Assimilation and Uncertainty Quantification for Dynamical Systems: A Review, *IEEE/CAA J. Autom. Sin.*, vol. **10**, no. 6, pp. 1361–1387, 2023.

34. Korn, P., Strong Solvability of a Variational Data Assimilation Problem for the Primitive Equations of Large-Scale Atmosphere and Ocean Dynamics, *J. Nonlinear Sci.*, vol. **31**, no. 3, pp. 56, 2021.

35. Chelil, S., Oubanas, H., Henine, H., Gejadze, I., Malaterre, P. O., and Tournebize, J., Variational Data Assimilation to Improve Subsurface Drainage Model Parameters, *J. Hydrol.*, vol. **610**, pp. 128006, 2022.

36. Jia, X., Zhou, W., Li, S., and Chen, X., Combined Prediction of Short-Term Travel Time of Expressway Based on CEEMDAN Decomposition, *IEEE Access*, vol. **10**, pp. 96873–96885, 2022.

37. Ding, L., and Wen, C., High-Order Extended Kalman Filter for State Estimation of Nonlinear Systems, *Symmetry*, vol. **16**, no. 5, pp. 617, 2024.

38. Wang, Y., and Yan, B., On the Feasibility of an Ensemble Multi-Fidelity Neural Network for Fast Data Assimilation for Subsurface Flow in Porous Media, *Available at SSRN 4293917*, 2022.

39. Hou, E., Lawrence, E., and Hero, A. O., Penalized Ensemble Kalman Filters for High Dimensional Non-Linear Systems, *PLoS One*, vol. **16**, no. 3, pp. e0248046, 2021.

40. Farchi, A., Laloyaux, P., Bonavita, M., and Bocquet, M., Using Machine Learning to Correct Model Error in Data Assimilation and Forecast Applications, *Q. J. R. Meteorol. Soc.*, vol. **147**, no. 739, pp. 3067–3084, 2021.

41. Tamaddon-Jahromi, H. R., Chakshu, N. K., Sazonov, I., Evans, L. M., Thomas, H., and Nithiarasu, P., Data-Driven Inverse Modelling Through Neural Network (Deep Learning) and Computational Heat Transfer, *Comput. Methods Appl. Mech. Eng.*, vol. **369**, pp. 113217, 2020.

42. Hughes, M. T., Kini, G., and Garimella, S., Status, Challenges, and Potential for Machine Learning in Understanding and Applying Heat Transfer Phenomena, *J. Heat Transfer*, vol. **143**, no. 12, pp. 120802, 2021.

43. He, Z., Ni, F., Wang, W., and Zhang, J., A Physics-Informed Deep Learning Method for Solving Direct and Inverse Heat Conduction Problems of Materials, *Mater. Today Commun.*, vol. **28**, pp. 102719, 2021.

44. Zeng, Y., Wang, H., Zhang, S., Cai, Y., and Li, E., A Novel Adaptive Approximate Bayesian Computation Method for Inverse Heat Conduction Problem, *Int. J. Heat Mass Transfer*, vol. **134**, pp. 185–197, 2019.

45. Wang, W., Yu, B., Ai, Q., Liu, M., Shuai, Y., Wang, X., and Zhang, X., Reconstructing 3D Temperature Fields from Sparse Discrete Data by Analytical Solution-Embedded Neural Network, *Appl. Therm. Eng.*, vol. **264**, pp. 125448, 2025.

46. Billah, M. M., Khan, A. I., Liu, J., and Dutta, P., Physics-Informed Deep Neural Network for Inverse Heat Transfer Problems in Materials, *Mater. Today Communications*, vol. **35**, pp. 106336, 2023.

47. Qian, Z., and Fu, C.-L., Regularization Strategies for a Two-Dimensional Inverse Heat Conduction Problem, *Inverse Probl.*, vol. **23**, no. 3, pp. 1053, 2007.

48. Li, Y., Wan, D., Wang, Z., and Hu, D., Physics-Constrained Deep Learning Approach for Solving Inverse Problems in Composite Laminated Plates, *Composite Struct.*, vol. **348**, pp. 118514, 2024.

49. Tiwari, A., Supervised Learning: From Theory to Applications, *Artificial Intelligence and Machine Learning for EDGE Computing*, pp. 23–32, 2022.

50. Cord, M., and Cunningham, P., Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval, Springer Science & Business Media, 2008.

51. Valkenborg, D., Geubbelmans, M., Rousseau, A.-J., and Burzykowski, T., Supervised Learning, *Am. J. Orthod. Dentofacial Orthop.*, vol. **164**, no. 1, pp. 146–149, 2023.

52. Alnuaimi, A. F. A. H., and Albaldawi, T. H. K., An Overview of Machine Learning Classification Techniques, *BIO Web of Conferences*, vol. **97**, pp. 00133, 2024.878

53. Vu, A. T., Gulati, S., Vogel, P.-A., Grunwald, T., and Bergs, T., Machine Learning-Based Predictive Modeling of Contact Heat Transfer, *Int. J. Heat Mass Transfer*, vol. **174**, pp. 121300, 2021.

54. Chu, H., Ji, T., Yu, X., Liu, Z., Rui, Z., and Xu, N., Advances in the Application of Machine Learning to Boiling Heat Transfer: A Review, *Int. J. Heat Fluid Flow*, vol. **108**, pp. 109477, 2024.

55. Wu, Y.-C., and Feng, J.-W., Development and Application of Artificial Neural Network, *Wireless Pers. Commun.*, vol. **102**, pp. 1645–1656, 2018.

56. Chen, T., and Chen, H., Universal Approximation to Nonlinear Operators by Neural Networks with Arbitrary Activation Functions and Its Application to Dynamical Systems, *IEEE Trans. Neural Netw.*, vol. **6**, no. 4, pp. 911–917, 1995.

57. Kleyko, D., Rosato, A., Frady, E. P., Panella, M., and Sommer, F. T., Perceptron Theory Can Predict the Accuracy of Neural Networks, *IEEE Trans. Neural Netw. Learning Syst.*, 2023.

58. Masood, A., and Ahmad, K., A Review on Emerging Artificial Intelligence (AI) Techniques for Air Pollution Forecasting: Fundamentals, Application and Performance, *J. Cleaner Prod.*, vol. **322**, pp. 129072, 2021.

59. Kim, D. E., and Gofman, M., Comparison of Shallow and Deep Neural Networks for Network Intrusion Detection, *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 204–208, 2018.

60. Löhner, R., Antil, H., Tamaddon-Jahromi, H., Chakshu, N. K., and Nithiarasu, P., Deep Learning or Interpolation for Inverse Modelling of Heat and Fluid Flow Problems?, *Int. J. Numer. Methods Heat Fluid Flow*, vol. **31**, no. 9, pp. 3036–3046, 2021.

61. Zhang, W., Xu, Z., Zhang, F., Wang, Z., Tang, M., Zhao, H., Zhang, L., and Wang, K., Development of a Data-Driven Subsurface Thermal Probing Technique for Lunar Exploration, *Adv. Space Res.*, vol. **73**, no. 4, pp. 2261–2276, 2024.

62. Haolong, C., and Zhanli, L., Solving the Inverse Heat Conduction Problem Based on Data Driven Model, *Chinese J. Comput. Mech.*, 2021.

63. Biczó, Z., Szénási, S., and Felde, I., A Novel Machine Learning Solution for the Inverse Heat Conduction Problem with Synthetic Datasets, *2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, pp. 000117–000122, 2023.

64. Kumar, R., Heat Transfer in Material Having Random Thermal Conductivity Using Monte Carlo Simulation and Deep Neural Network, *Multiscale Multidiscip. Model. Exp. Des.*, vol. **7**, no. 4, pp. 3173–3186, 2024.

65. Wang, H., Zhu, T., Zhu, X., Yang, K., Ge, Q., Wang, M., and Yang, Q., Inverse Estimation of Hot-Wall Heat Flux Using Nonlinear Artificial Neural Networks, *Measurement*, vol. **181**, pp. 109648, 2021.

66. Allard, D., and Najafi, H., A Systematic Study of Two Machine Learning-Based Approaches for Solving Non-Linear Inverse Heat Conduction Problems in One-Dimensional Domains, *Int. Commun. Heat Mass Transfer*, vol. **155**, pp. 107494, 2024.

67. Najafi, H., and Woodbury, K. A., Online Heat Flux Estimation Using Artificial Neural Network as a Digital Filter Approach, *Int. J. Heat Mass Transfer*, vol. **91**, pp. 808–817, 2015.

68. Mirsephai, A., Mohammadzaheri, M., Chen, L., and O'Neill, B., An Artificial Intelligence Approach to Inverse Heat Transfer Modeling of an Irradiative Dryer, *Int. Commun. Heat Mass Transfer*, vol. **39**, no. 1, pp. 40–45, 2012.

69. Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., and Parmar, M., A Review of Convolutional Neural Networks in Computer Vision, *Artif. Intell. Rev.*, vol. **57**, no. 4, pp. 99, 2024.

70. Cong, S., and Zhou, Y., A Review of Convolutional Neural Network Architectures and Their Optimizations, *Artif. Intell. Rev.*, vol. **56**, no. 3, pp. 1905–1969, 2023.

71. Yamashita, R., Nishio, M., Do, R. K. G., and Togashi, K., Convolutional Neural Networks: An Overview and Application in Radiology, *Insights into Imaging*, vol. **9**, pp. 611–629, 2018.

72. Wang, R., Li, J., Duan, Y., Cao, H., and Zhao, Y., Study on the Combined Application of CFAR and Deep Learning in Ship Detection, *J. Indian Soc. Remote Sens.*, vol. **46**, pp. 1413–1421, 2018.

73. Tamaddon-Jahromi, H. R., Chakshu, N. K., Thomas, H., and Nithiarasu, P., Deep Learning in Heat Transfer, *Annu. Rev. Heat Transfer*, vol. **24**, 2021.

74. Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L., Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions, *J. Big Data*, vol. **8**, pp. 1–74, 2021.

75. Zhao, L., and Zhang, Z., An Improved Pooling Method for Convolutional Neural Networks, *Sci. Rep.*, vol. **14**, no. 1, pp. 1589, 2024.

76. Wang, Y., Li, Y., Song, Y., and Rong, X., The Influence of the Activation Function in a Convolution Neural Network Model of Facial Expression Recognition, *Appl. Sci.*, vol. **10**, no. 5, pp. 1897, 2020.

77. Edalatifar, M., Tavakoli, M. B., Ghalambaz, M., and Setoudeh, F., Using Deep Learning to Learn Physics of Conduction Heat Transfer, *J. Therm. Anal. Calorim.*, vol. **146**, pp. 1435–1452, 2021.

78. Liu, T., Li, Y., Jing, Q., Xie, Y., and Zhang, D., Supervised Learning Method for the Physical Field Reconstruction in a Nanofluid Heat Transfer Problem, *Int. J. Heat Mass Transfer*, vol. **165**, pp. 120684, 2021.

79. Suh, Y., Bostanabad, R., and Won, Y., Deep Learning Predicts Boiling Heat Transfer, *Sci. Rep.*, vol. **11**, no. 1, pp. 5622, 2021.

80. Peng, J.-Z., Liu, X., Aubry, N., Chen, Z., and Wu, W.-T., Data-Driven Modeling of Geometry-Adaptive Steady Heat Transfer Based on Convolutional Neural Networks: Heat Conduction, *arXiv preprint arXiv:2010.03854*, 2020.

81. Hua, Y., Peng, J.-Z., Zhou, Z.-F., Wu, W.-T., He, Y., and Massoudi, M., Thermal Performance in Convection Flow of Nanofluids Using a Deep Convolutional Neural Network, *Energies*, vol. **15**, no. 21, pp. 8195, 2022.

82. Kim, J., and Lee, C., Prediction of Turbulent Heat Transfer Using Convolutional Neural Networks, *J. Fluid Mech.*, vol. **882**, pp. A18, 2020.

83. Ebbs-Picken, T., Romero, D. A., Da Silva, C. M., and Amon, C. H., Deep Encoder–Decoder Hierarchical Convolutional Neural Networks for Conjugate Heat Transfer Surrogate Modeling, *Appl. Energy*, vol. **372**, pp. 123723, 2024.

84. Virupaksha, A. G., Nagel, T., Lehmann, F., Rajabi, M. M., Hoteit, H., Fahs, M., and Le Ber, F., Modeling Transient Natural Convection in Heterogeneous Porous Media with Convolutional Neural Networks, *Int. J. Heat Mass Transfer*, vol. **222**, pp. 125149, 2024.

85. Bilke, L., Flemisch, B., Kalbacher, T., Kolditz, O., Helmig, R., and Nagel, T., Development of Open-Source Porous Media Simulators: Principles and Experiences, *Transport in Porous Media*, vol. **130**, pp. 337–361, 2019.

86. Rajabi, M. M., Javaran, M. R. H., Bah, A.-O., Frey, G., Le Ber, F., Lehmann, F., and Fahs, M., Analyzing the Efficiency and Robustness of Deep Convolutional Neural Networks for Modeling Natural Convection in Heterogeneous Porous Media, *Int. J. Heat Mass Transfer*, vol. **183**, pp. 122131, 2022.

87. Zhu, F., Chen, J., and Han, Y., A Multiple Regression Convolutional Neural Network for Estimating Multi-Parameters Based on Overall Data in the Inverse Heat Transfer Problem, *J. Thermal Sci. Eng. Appl.*, vol. **14**, no. 5, pp. 051003, 2022.

88. Jowett-Lockwood, L. A., and Evins, R., Inverse Surrogate Modelling to Determine Thermal Characteristics of Buildings, *Building Simulation 2023*, vol. **18**, pp. 932–939, 2023.

89. Hopkins, A. S., Simulating a Nationally Representative Housing Sample Using EnergyPlus, 2011.

90. Masrouri, M., and Tahsini, A. M., A Comparative Experimental Study on Inverse Identification of an Industrial Heat Gun Using Deep Learning and Two Inverse Heat Transfer Techniques, *J. Thermal Anal. Calorim.*, vol. **148**, no. 18, pp. 9597–9608, 2023.

91. Zhu, F., Chen, J., Han, Y., and Ren, D., A Deep Learning Method for Estimating Thermal Boundary Condition Parameters in Transient Inverse Heat Transfer Problem, *Int. J. Heat Mass Transfer*, vol. **194**, pp. 123089, 2022.

92. Niu, Z., Qi, H., Sun, A., Ren, Y., He, M., and Gao, B., Efficient and Robust CNN-LSTM Prediction of Flame Temperature Aided Light Field Online Tomography, *Sci. China Technol. Sci.*, vol. **67**, no. 1, pp. 271–284, 2024.

93. Iosifidis, A., and Tefas, A., Deep Learning for Robot Perception and Cognition, Academic Press, 2022.

94. Mienye, I. D., Swart, T. G., and Obaido, G., Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications, *Information*, vol. **15**, no. 9, pp. 517, 2024.

95. Mienye, I. D., and Jere, N., Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions, *IEEE Access*, 2024.

96. Bengio, Y., Simard, P., and Frasconi, P., Learning Long-Term Dependencies with Gradient Descent Is Difficult, *IEEE Trans. Neural Netw.*, vol. **5**, no. 2, pp. 157–166, 1994.

97. Schmidt, R. M., Recurrent Neural Networks (RNNs): A Gentle Introduction and Overview, Nov. 2019, *URL http://arxiv. org/abs/1912.05911*, 2023.

98. Zwart, J. A., Oliver, S. K., Watkins, W. D., Sadler, J. M., Appling, A. P., Corson-Dosch, H. R., Jia, X., Kumar, V., and Read, J. S., Near-Term Forecasts of Stream Temperature Using Deep Learning and Data Assimilation in Support of Management Decisions, *JAWRA J. Am. Water Resour. Assoc.*, vol. **59**, no. 2, pp. 317–337, 2023.

99. Wang, Y., Wang, N., and Ren, Q., Predicting Surface Heat Flux on Complex Systems Via Conv-LSTM, *Case Studies in Thermal Engineering*, vol. **33**, pp. 101927, 2022.

100. Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-C., Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, *Advances in Neural Information Processing Systems*, vol. **28**, 2015.

101. Turner, R. E., An Introduction to Transformers, *arXiv preprint arXiv:2304.10557*, 2023.

102. Islam, S., Elmekki, H., Elsebai, A., Bentahar, J., Drawel, N., Rjoub, G., and Pedrycz, W., A Comprehensive Survey on Applications of Transformers for Deep Learning Tasks, *Expert Systems with Applications*, vol. **241**, pp. 122666, 2024.

103. Jaspers, S. P. F. C., Dautzenberg, J. H., and Taminiau, D. A., Temperature Measurement in Orthogonal Metal Cutting, *Int. J. Adv. Manuf. Technol.*, vol. **14**, pp. 7–12, 1998.

104. Han, J., Xu, L., Cao, K., Li, T., Tan, X., Tang, Z., Shi, T., and Liao, G., Online Estimation of the Heat Flux During Turning Using Long Short-Term Memory Based Encoder-Decoder, *Case Studies in Thermal Engineering*, vol. **26**, pp. 101002, 2021.

105. Wang, Y., Zhan, C., Li, G., Zhang, D., and Han, X., Physics-Guided LSTM Model for Heat Load Prediction of Buildings, *Energy and Buildings*, vol. **294**, pp. 113169, 2023.

106. Sharma, P., Evans, L., Tindall, M., and Nithiarasu, P., Stiff-PDEs and Physics-Informed Neural Networks, *Archives of Computational Methods in Engineering*, vol. **30**, no. 5, pp. 2929–2958, 2023.

107. Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F., Scientific Machine Learning Through Physics–Informed Neural Networks: Where We Are and What's Next, *Journal of Scientific Computing*, vol. **92**, no. 3, pp. 88, 2022.

108. Farea, A., Yli-Harja, O., and Emmert-Streib, F., Understanding Physics-Informed Neural Networks: Techniques, Applications, Trends, and Challenges, *AI*, vol. **5**, no. 3, pp. 1534–1557, 2024.

109. Qian, W., Hui, Xin, Wang, B., Zhang, Z., Lin, Y., and Yang, S., Physics-Informed Neural Network for Inverse Heat Conduction Problem, *Heat Transfer Res.*, vol. **54**, no. 4, 2023.

110. Puttock-Brown, M. R., Bindhu, G. K. M., and Ashby, C. E., A Physics-Informed Neural Network for Solving the Inverse Heat Transfer Problem in Gas Turbine Rotating Cavities, *J. Turbomach.*, vol. **147**, no. 7, 2025.

111. He, K., Zhang, X., Ren, S., and Sun, J., Deep Residual Learning for Image Recognition, *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 770–778, 2016.

112. Raissi, M., and Karniadakis, G. E., Hidden Physics Models: Machine Learning of Nonlinear Partial Differential Equations, *J. Comput. Phys.*, vol. **357**, pp. 125–141, 2018.

113. Jiang, X., Wang, X., Wen, Z., Li, E., and Wang, H., Practical Uncertainty Quantification for Space-Dependent Inverse Heat Conduction Problem Via Ensemble Physics-Informed Neural Networks, *Int. Commun. Heat Mass Transfer*, vol. **147**, pp. 106940, 2023.

114. Islam, M. S., and Dutta, P., Machine Learning Assisted Inverse Heat Transfer Problem to Find Heat Flux in Ablative Materials, *Mater. Today Commun.*, vol. **45**, no. C, 2025.

115. Liao, S., Xue, T., Jeong, J., Webster, S., Ehmann, K., and Cao, J., Hybrid Thermal Modeling of Additive Manufacturing Processes Using Physics-Informed Neural Networks for Temperature Prediction and Parameter Identification, *Comput. Mech.*, vol. **72**, no. 3, pp. 499–512, 2023.

116. Go, M.-S., Lim, J. H., and Lee, S., Physics-Informed Neural Network-Based Surrogate Model for a Virtual Thermal Sensor with Real-Time Simulation, *Int. J. Heat Mass Transfer*, vol. **214**, pp. 124392, 2023.

117. Chen, Y., Chen, Q., Ma, H., Chen, S., and Fei, Q., Transfer Machine Learning Framework for Efficient Full-Field Temperature Response Reconstruction of Thermal Protection Structures with Limited Measurement Data, *Int. J. Heat Mass Transfer*, vol. **242**, pp. 126785, 2025.

118. Jo, J., Jeong, Y., Kim, J., and Yoo, J., Thermal Conductivity Estimation Using Physics-Informed Neural Networks with Limited Data, *Eng. Appl. Artif. Intell.*, vol. **137**, pp. 109079, 2024.

119. Tanaka, H., and Nagai, H., Data-Driven Thermal State Estimation for In-Orbit Systems Via Physics-Informed Machine Learning, *Acta Astronaut.*, vol. **212**, pp. 316–328, 2023.

120. Di Meglio, A., Massarotti, N., and Nithiarasu, P., Digital Twins of Thermal Systems: A Comparison Between Supervised and Reinforcement Learning, *Computational Thermal Sciences: An International Journal*, vol. **17**, no. 3, 2025.

121. Vallée, A., Digital Twin for Healthcare Systems, *Frontiers in Digital Health*, vol. **5**, pp. 1253050, 2023.

122. Caputo, F., Greco, A., Fera, M., and Macchiaroli, R., Digital Twins to Enhance the Integration of Ergonomics in the Workplace Design, *International Journal of Industrial Ergonomics*, vol. **71**, pp. 20–31, 2019.

123. Laubenbacher, R., Mehrad, B., Shmulevich, I., and Trayanova, N., Digital Twins in Medicine, *Nature Computational Science*, vol. **4**, no. 3, pp. 184–191, 2024.

124. Purcell, W., and Neubauer, T., Digital Twins in Agriculture: A State-of-the-Art Review, *Smart Agricultural Technology*, vol. **3**, pp. 100094, 2023.

125. Subramani, T., and Bartscher, S., Predictive Digital Twins for Thermal Management Using Machine Learning and Reduced-Order Models, *arXiv preprint arXiv:2505.06849*, 2025.

126. Wright, L., and Davidson, S., How to Tell the Difference Between a Model and a Digital Twin, *Advanced Modeling and Simulation in Engineering Sciences*, vol. **7**, pp. 1–13, 2020.

127. Di Meglio, A., Massarotti, N., and Nithiarasu, P., A Physics-Driven and Machine Learning-Based Digital Twinning Approach to Transient Thermal Systems, *Int. J. Numer. Methods Heat & Fluid Flow*, vol. **34**, no. 6, pp. 2229–2256, 2024.

128. Yang, Y., Zhao, X., Cheng, Q., Guo, R., Li, M., and Zhou, J., POD–ANN as Digital Twins for Surge Line Thermal Stratification, *Nuclear Engineering and Design*, vol. **428**, pp. 113487, 2024.

129. Yang, J., Ji, Z., Liu, W., and Xie, Z., Digital-Twin-Based Coordinated Optimal Control for Steel Continuous Casting Process, *Metals*, vol. **13**, no. 4, pp. 816, 2023.

130. Ma, S., Leng, J., Zheng, P., Chen, Z., Li, B., Li, W., Liu, Q., and Chen, X., A Digital Twin-Assisted Deep Transfer Learning Method Towards Intelligent Thermal Error Modeling of Electric Spindles, *Journal of Intelligent Manufacturing*, vol. **36**, no. 3, pp. 1659–1688, 2025.