Thermal Science and Engineering Progress 49 (2024) 102477

Contents lists available at ScienceDirect

Thermal Science and Engineering Progress

journal homepage: www.sciencedirect.com/journal/thermal-science-and-engineering-progress

Deep learning model for rapid temperature map prediction in transient convection process using conditional generative adversarial networks

Munku Kang, Nam Phuong Nguyen, Beomjin Kwon *

School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ 85287, USA

ARTICLE INFO

Keywords:
3D transient mixed convection
Conditional generative adversarial networks

ABSTRACT

This article presents a deep learning model for three-dimensional (3D) transient mixed convection in a horizontal channel with a heated bottom surface. Using Conditional Generative Adversarial Networks (cGAN), we successfully approximate temperature maps at arbitrary channel locations and time steps. The model is specifically designed for mixed convection at Reynolds number 100, Rayleigh number 3.9×10^6, and Richardson number 88.8. To investigate the impact of the discriminator network architecture on model accuracy, we compare Convolutional Neural Network (CNN) based classifiers with PatchGAN classifiers, both with and without strided convolutions. Remarkably, the cGAN with PatchGAN based classifier (without strided convolutions) yields the highest clarity and accuracy in inferring temperature maps. Additionally, other factors such as image contrast, spatiotemporal variation rate of temperature, and the number of channels in the temperature image significantly influence cGAN accuracy. This work highlights the potential of deep learning in efficiently modeling complex transport processes.

1. Introduction

Recently, there has been a surge in interest in utilizing deep learning approaches for modeling transport processes, a domain that has traditionally relied heavily on computational fluid dynamics (CFD). Deep learning algorithms offer the advantage of being able to effectively identify features within unstructured data, such as images, texts, and signals. This eliminates the need for separate feature extraction processes in data-driven modeling, streamlining the overall workflow.

While CFD techniques are well-established and powerful, they often demand substantial computational resources and time to tackle large-scale spatiotemporal dynamic problems. This can be a significant drawback in practical applications. However, with access to simulation and/or experimental data for transport processes, deep learning models can be trained to address complex transport problems without incurring significant computational costs. In essence, these models can act as efficient reduced-order models. By leveraging the data-driven capabilities of deep learning, researchers and practitioners can explore novel avenues for modeling transport processes and potentially unlock new insights. The ability to reduce reliance on extensive computational resources makes deep learning an attractive option for tackling real-world transport challenges.

Numerous deep learning techniques have been employed in modeling heat transfer processes, including the pioneering use of conditional generative adversarial networks (cGAN). In one early application [1], cGANs were utilized to infer steady-state temperature fields within a two-dimensional (2D) domain where conduction served as the sole heat transfer mechanism. The training dataset for this work was prepared using a finite volume method. The cGAN model effectively generated temperature fields based on 2D images representing initial temperature distributions.

Expanding on this success, the cGAN model was further trained to approximate temperature, velocity, and pressure fields for a natural convection process in a 2D square cavity [2]. Here, the model inputs consisted of five-channel 2D images, incorporating initial and boundary conditions. In another recent application [3], the cGAN technique was employed to infer 2D distributions of temperature, Nusselt number, and friction factor within a heated channel, considering a wide range of Reynolds numbers from 100 to 27750. Notably, when combined with a data augmentation technique, the cGAN model demonstrated exceptional predictive capabilities for unseen fluid channel geometries, including narrowed, widened, and rotated channels, highlighting the technique's nonlinear interpolation and extrapolation capabilities. Traditionally, convolutional neural networks (CNN) have been widely used for image processing and analysis. However, CNNs necessitate a

---

* Corresponding author.
E-mail address: kwon@asu.edu (B. Kwon).

https://doi.org/10.1016/j.tsep.2024.102477
Received 10 October 2023; Received in revised form 23 January 2024; Accepted 19 February 2024

Nomenclature

c_p	specific heat, J/kg·K
H	channel height, m
k	thermal conductivity of fluid, W/m·K
L	channel length, m
Pr	Prandtl number, dimensionless
ΔP	pressure drop between the channel inlet and outlet, Pa
R^2	coefficient of determination, dimensionless
Ra	Rayleigh number, dimensionless
Re	Reynolds number based on channel diameter, dimensionless
Ri	Richardson number, dimensionless
t	time, (s)
T_b	base plate temperature, K
T_i	inlet temperature of the fluid, K
T_{min}	minimum temperature of fluid, K
T_{max}	maximum temperature of fluid, K
ΔT	mean temperature difference between the ground truth and generated images, K
ΔT_{max}	maximum temperature difference between the ground truth and generated images, K
W	channel width, m

Greek symbol

α	thermal diffusivity, m^2/s
ρ	density, kg/m^3
λ	weight parameter, dimensionless
μ	dynamic viscosity, Pa·s

Acronyms

ADAM	adaptive moment estimation
AE	autoencoder
cGAN	conditional generative adversarial network
CNN	convolutional neural network
CFD	computational fluid dynamics
DNN	deep neural networks
DN	discriminator network
DCGAN	deep convolutional generative adversarial network
FVM	finite volume model
GAN	generative adversarial network
GN	generator network
MAE	mean absolute error
ReLU	rectified linear unit

carefully designed loss function to achieve satisfactory performance [4]. To address this challenge and enhance accuracy, the concept of Generative Adversarial Networks (GAN) was introduced [5]. GANs automatically adapt the loss function to the data, making them well-suited for deep learning-based image generation tasks. As a result, GANs have gained popularity in the field.

In addition to the cGAN technique, various other deep learning methods have been employed for heat transfer modeling. CNN [6] and autoencoder (AE) [7] were utilized to predict steady-state temperature distributions within a 2D domain with conduction as the sole heat transfer mechanism. Similar to GAN, both models effectively reconstructed output images from input images containing boundary conditions. In these studies, the CNN model utilized a 2D heat equation as its loss function [6], while the autoencoder model employed mean squared errors or mean of maximum squared errors as its loss function [7]. In another demonstration, CNN was trained to predict the 2D distribution of local heat flux in a fully developed turbulent channel flow, using 2D images of wall-shear stresses and pressure fluctuations as inputs [8]. The loss function in this work was defined as the sum of mean squared error and regularization loss. Recently, physics-informed neural networks (PINN) were reported for one-dimensional (1D) and 2D heat conduction problems [9]. By defining the loss function to incorporate errors in the conduction heat transfer equation, boundary and initial conditions, deep neural networks (DNN) were able to predict temperature distributions in a 2D domain without requiring a pre-generated training dataset. DNNs are also commonly used for regressing discrete numeric outputs from discrete numeric inputs. In a previous work [10], two-layer artificial neural networks learned to predict the fouling resistances of a cross-flow heat exchanger when six operating condition parameters were provided.

Despite the considerable progress made in deep learning techniques for modeling heat transfer processes, there remains uncertainty about their effectiveness in inferring complex phenomena, such as transient, multi-mode conjugate, or various turbulent-flow heat transfer processes. The question arises whether deep learning models can outperform numerical simulations in solving these complex heat transfer processes, particularly in terms of speed. If deep learning models can achieve faster results while maintaining accuracy, they would be invaluable in designing and optimizing 2D or three-dimensional (3D) transport systems. This study aims to explore the potential of deep learning by investigating deep conditional generative adversarial networks capable of predicting 3D transient laminar mixed convection phenomena.

Mixed convection involves comparable effects of free and forced convection, resulting in buoyancy-driven secondary flows and unstable, inhomogeneous property fields. The flow characteristics of mixed convection are known to enhance heat transfer up to 4 to 5 times compared to pure forced convection [11–15], making it an area of substantial interest in engineering and scientific applications, such as heat exchangers [16–18], radiation energy collectors [19,20], plants [21], and groundwater and geothermal systems [22,23]. By studying deep learning models in the context of mixed convection, this research seeks to explore their applicability and potential benefits in tackling such challenging and important heat transfer scenarios.

2. Mixed convection problem and data generation

We investigate a mixed convection flow in a three-dimensional (3D) channel with dimensions: width (W) of 60 mm, height (H) of 15 mm, and length (L) of 130 mm, as depicted in Fig. 1. The flow is driven by water entering the channel at 35 °C (T_i) through the inlet with a uniform velocity distribution, resulting in a Reynolds number (Re) of 100, based on the channel hydraulic diameter. The bottom wall temperature (T_b) is set to 43 °C, while the other walls are thermally insulated. The prescribed boundary conditions yield a Rayleigh number (Ra) of 3.9×10^6 and a Richardson number (Ri) of 88.8.

To analyze the transient flow and temperature fields, we employ a finite volume model (FVM) implemented in commercial CFD software (ANSYS Fluent). To reduce computational cost, the simulation considers only half-section of the channel, utilizing the yz-plane at $x = 0$ as a symmetric plane. The convergence criteria for the continuity equation are set at 10^{-4}, and for the momentum and energy equations, it is set at 10^{-5}. The time step size for the simulation model is 0.05 s, and the total number of time steps is 7200. In our model, we determine the density (ρ) and dynamic viscosity (μ) of water as functions of temperature without employing the Boussinesq approximation. Other water properties, such as specific heat (C_p), Prandtl number (Pr), and thermal conductivity (k), are assumed to be constant at the mean temperature of the water inlet and bottom wall.

Fig. 1 depicts the temperature distribution at select yz-planes. Near the channel entrance ($z < 2H$), only the fluid near the bottom wall is heated, while most of the fluid maintains a temperature close to the inlet

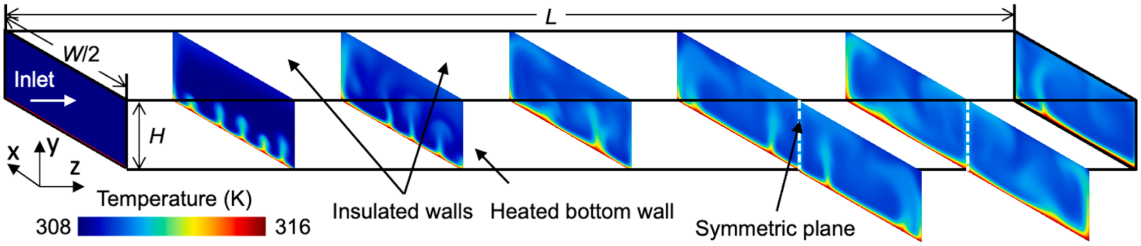

Fig. 1. Temperature distribution at selected cross-sections in a rectangular channel induced by mixed convection heat transfer.

temperature. Along the x-direction, thermal plumes develop at regular intervals, representing concentrated hot fluid penetrating into the colder region above. At this stage, there is a high temperature gradient from the heated wall toward the unheated fluid region. As z increases along the channel, the thermal plumes grow in size, reach the top wall, and circulate around the channel, contributing to the homogenization of fluid temperature. Beyond the middle region of the channel ($z > L/2$), the temperature gradient in the fluid significantly reduces due to hydraulic and thermal mixing.

The solutions obtained from the FVM represent temperature distributions in the simulation domain at discrete time steps. These results are then used to create training and testing datasets for the cGAN model. The datasets consist of 4320 two-dimensional (2D) temperature images with a resolution of 64 × 128, obtained from xy-planes at 360 different time steps. Two approaches are considered for generating the temperature images: (1) RGB images with three channels and (2) grayscale images with a single channel. For this study, we opted to use 16-bit grayscale images due to their advantages in achieving higher accuracy with significantly fewer training parameters. Each pixel of a 16-bit grayscale image can contain a value between 0 and 65535. The grayscale pixel intensity corresponds to the absolute temperature multiplied by a scaling factor, which is 100 in this work. Due to the temperature range relevant to our problem, the pixel values range from 30,800 to 31600. For numerical stability during training, we offset the pixel values by half of 65,535 and normalize them so that each pixel may have a value between -1 and 1. However, for visualization purposes, the subsequent temperature maps obtained from the cGAN model are presented in RGB scales accompanied by color bars. The single-channel image-based training required only a few hundred parameters for the generator training, whereas the three-channel image-based training demanded thousands of parameters. For our problem, there was no significant difference in training duration between the single-channel and three-channel image-based training.

In the training process, 90 % of the images from the dataset are selected to train the cGAN model, while the remaining 10 % are reserved for testing the model's performance. Specifically, we sampled the test data with a uniform distribution in both the spatial and temporal domains. The temperature images were collected at a constant time step interval of 10 s and a constant step distance along the z-direction, i.e., a 10 mm interval. Uniform data sampling enabled our model to be trained and tested without specific biases at particular time steps or channel locations.

3. Deep learning methodology

We utilize the cGAN [24] to achieve rapid prediction of flow temperature distribution. Among various deep learning algorithms for image processing, such as CNN [25] and variational autoencoder (VAE) [26], we specifically select cGAN due to its superior capability in image generation. CNN is a popular class of artificial neural networks used for tasks like classifying, recognizing, analyzing, or segmenting images. However, generating high-quality target images using CNN alone without designing sophisticated loss functions can be challenging. On the other hand, VAE is an artificial neural network that maps input image characteristics into a latent space and reconstructs images from vectors in that space. While VAE training is stable, the generated outputs may not be as close to the ground truth images as those produced by cGAN.

The generator network (GN) of the cGAN learns to create images that closely resemble the ground truth images, tricking the discriminator network (DN). In Fig. 2(a), we present the architecture of the GN. The input to GN is a 1 × 8 conditional vector that describes the channel geometry, thermal and hydraulic boundary conditions, fluid properties, time (t), and the z location in the channel, i.e., input (c) = [channel aspect ratio (AR), Re, Pr, Ri, T_i, T_b, t, z]. The fully connected layer (G1) feeds and reshapes the input vector into an input layer with a shape of 1 × 1 × 512 array. Subsequently, the input layer goes through seven blocks (from G2 to G8), each comprising three layers, including transpose convolutional, batch normalization, and ReLU (Rectified Linear Unit) [27] activation layers. The transpose convolutional layer performs an inverse convolution operation, generating an output feature map with greater dimensions than the input. The batch normalization [28] layer transforms the means and variances of the layer inputs to ensure a stable and fast learning process. In this work, G1 does not include a batch normalization layer, as GN performs well without it. Finally, the output layer (G9) uses the Tanh activation function, generating a flow temperature image with a desired size of 64 × 128.

The discriminator network (DN) learns to distinguish ground truth images from the fake images generated by GN. Several common types of DN have been developed for GANs. In the original GAN [5], DN is defined by multilayer perceptrons, providing a classifier with a scalar probability that the input image belongs to the ground truth group rather than being from GN. In recent GANs, multilayer perceptrons have been replaced by CNNs to enhance training stability. In the deep convolutional GAN (DCGAN) [29], DN uses a modified CNN that replaces pooling layers with strided convolutions and incorporates batch normalization. In PatchGAN [4], the classifier receives small patches of images and classifies them, requiring a smaller number of parameters in DN compared to the traditional GAN classifier.

Fig. 2(b) illustrates the DN architecture, which reduces the input image into an output array. Our DN is based on a PatchGAN, which produces a 64 × 64 image (D7) with each pixel representing the prediction for a 64 × 128 patch of the input image. The DN does not employ strided convolutions to maintain the feature map size after D1. Our selection of DN architecture is through the comparison between CNN and PatchGAN architectures, which is described in our supplementary material.

The objective of a cGAN combines the goals of both the GN and the DN. The GN aims to generate a flow temperature map (Y) from a given input vector (c) that closely resembles the CFD simulation result (X) while trying to deceive the DN. Conversely, the DN's objective is to distinguish between the real CFD simulation result (X) and the generated flow temperature map (Y). The cGAN's objectives are formulated as follows:

$$L_{cGAN}(GN,DN) = E_{c,X}[\log DN(c,X)] + E_{c,Y}[\log\{1\text{-}DN(c,Y)\}] \quad (1)$$

where $DN(c, X)$ represents the probability that the DN receives input c and real image X and classifies it as X, and $DN(c, Y)$ represents the

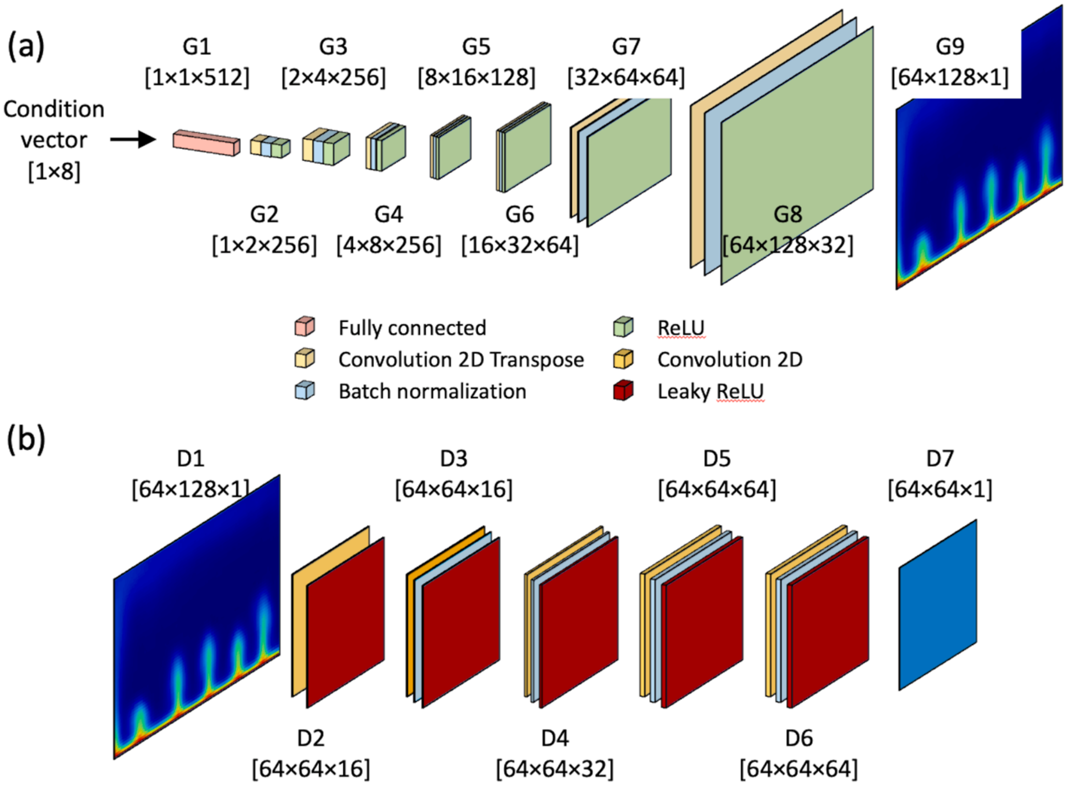

Fig. 2. cGAN architectures: (a) generator and (b) discriminator.

probability that the DN receives input *c* and generated image *Y*, and classifies it as *X*. The DN aims to maximize the expectation values of *DN* (*c*, *X*) and 1 - *DN*(*c*, *Y*), while the GN seeks to minimize these values as per Equation (1).

To enhance the ability of the GN further, the *L1* loss is often included as an additional objective. The *L1* loss measures the mean absolute pixel difference between the generated image *Y* and the real image *X*, encouraging the GN to produce clearer results. The *L1* loss is formulated as follows:

$$L_{L1}(GN) = E_{c,X,Y}[\|Y - X\|] \tag{2}$$

Thus, the final objective of the cGAN is described as

$$\mathrm{L_{Total}} = \mathrm{min_{GN}max_{DN}L_{cGAN}(GN,DN)} + \lambda \mathrm{L_{L1}(GN)} \tag{3}$$

where the hyperparameter λ is a weight that determines the relative importance of the *L1* loss in the overall objective.

We carefully optimized the hyperparameters, such as training epochs and λ, to strike a balance between training quality and computation duration. A sufficient number of training epochs is essential to reduce the training losses to desirable levels. We compared the results obtained with epochs set to 500, 1000, 2000, and 3000. As the number of epochs increased, the training duration also increased from 1.5 h to 8.8 h, and the mean temperature difference (ΔT) between the ground truth and generated images decreased from 0.22 K to 0.093 K. Note, ΔT is equivalent to the mean absolute error (MAE) in this work. Considering the training duration and ΔT, we conducted all subsequent model training with an epoch value of 3000. For the model training, we employ the ADAM (adaptive moment estimation) optimizer [30] with a learning rate of 0.0002 and β_1 of 0.5. Detailed information regarding hyperparameter optimization is provided in the supplementary material.

The hyperparameter λ plays a crucial role in balancing the two objectives during generator training. As λ increases, the GN focuses more on minimizing the mean absolute difference between the real image (*X*) and the generated image (*Y*), thus improving low-frequency accuracy. Given that many temperature images in our convection problem include low-frequency features, such as small temperature gradients across unit image pixels, we raised λ to prioritize the *L1* loss over L_{cGAN} and minimize model errors effectively. We identified an optimal value of λ as 5×10^5, since this λ made the mean prediction error ΔT, averaged over all channel locations and time steps, less than 0.1 K.

4. Model testing

The trained cGAN serves as reduced-order models for mixed convection, enabling the inference of temperature maps at arbitrary channel locations (z) and time points (t). Fig. 3(a) shows the temperature maps at an early stage (t = 12 s) of the mixed convection flow for selected z locations that were not seen by the model during training. The cGAN model successfully generates the characteristic development of convection transverse rolls in mixed convection flow. The temperature maps exhibit similar features arising from buoyancy-driven secondary flows at this early stage. At t = 12 s, the prediction error ΔT, averaged over all channel locations, ranges from merely 0.15 K to 0.24 K.

The accuracy of the models is affected by the image contrast determined by the temperature gradient. Fig. 3(b) shows the model predictions during a stable stage (t = 342 s). The trained models can infer elongated transverse convection rolls and the secondary flow development across the entire channel region. However, during this stable stage, the spatial temperature variation is noticeably reduced compared to the initial stage due to flow mixing induced by the secondary flows. The Michelson image contrast, calculated as (T_{max} - T_{min})/(T_{max} + T_{min}) where T_{max} and T_{min} are the maximum and minimum temperatures, decreases from 0.0128 to 0.0113 when comparing the ground truths at t = 12 s and 342 s at z = 60 mm. We find that models tend to generate blurred features, particularly indistinct boundaries of convection rolls at z = 60 mm when the Michelson image contrast is approximately below 0.011. At t = 342 s, the prediction error ΔT, averaged over all channel locations, is 0.132 K.

The cGAN demonstrates higher accuracy in predicting the simpler

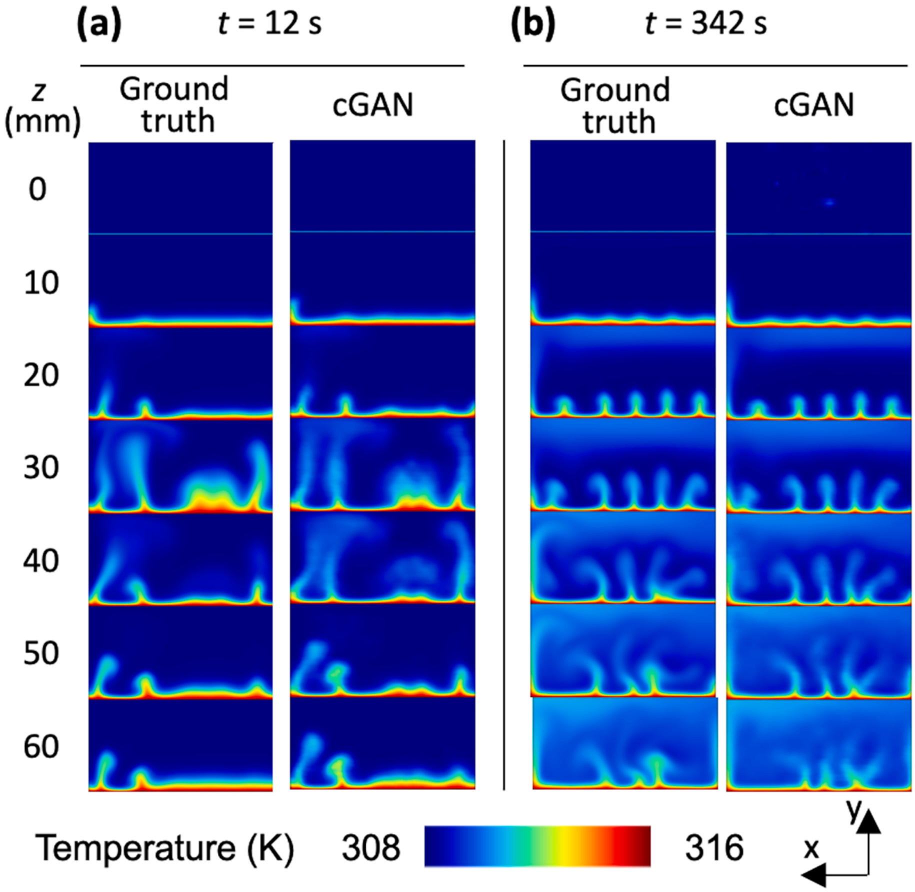

Fig. 3. Temperature distributions at selected cross-sections in a heated channel predicted at (a) $t = 12$ s and (b) $t = 342$ s by cGAN model and FVM (denoted as ground truth).

temperature maps occurring during the stable stage. The coefficient of determination (R^2), calculated between the predicted and ground truth temperatures, tends to be greater at smaller z values, particularly near the channel entrance region, where the image features are simpler and have higher contrast compared to locations at greater z values. R^2 also tends to be greater with a large t, where the time-dependent temperature variation is less compared to the initial flow development stage at small t. The time-averaged R^2 for all time steps at various z locations ranged between 0.834 and 0.999 as shown in Figure S7. For Fig. 3, Table 1 presents ΔT and the maximum temperature difference between ground truth and generated images (ΔT_{max}), which are also consistent with the dependence of R^2 on z and t.

Fig. 4 displays the temperature maps predicted by the cGAN at unseen channel locations ranging from $z = 5$ mm to $z = 55$ mm. Fig. 4(a) and 4(b) show the model predictions during the early unstable stage at $t = 12$ s and the stable stage at $t = 342$ s, respectively. Remarkably, even for the unseen channel locations, the trained cGAN successfully approximates the development of convection rolls at all time steps. Consistent with the previous model test conducted for various time points, the model performs better when predicting temperature maps with higher contrast. We observe that the Michelson image contrast in the ground truth, averaged for all selected channel locations, is 0.0114 at $t = 12$ s and 0.0109 at $t = 342$ s, indicating that the image contrast diminishes when the temperature field becomes mixed and homogenized during the stable stage. Consequently, the cGAN exhibits a greater location-averaged R^2 at $t = 12$ s ($R^2 = 0.907$) compared to $t = 342$ s ($R^2 = 0.862$). Based on our observations regarding the cGAN's accuracy dependence on the image contrast, we can expect improved performance if the color range of the temperature maps is adjusted to suit different stages of the convection process. This adjustment may enhance the cGAN's ability to predict temperature variations during the stable stage when the image contrast is relatively low.

Table 1
ΔT and ΔT_{max} at various values of t and z in Fig. 3.

	$t = 12$ s		$t = 342$ s	
z (mm)	ΔT (K)	ΔT_{max} (K)	ΔT (K)	ΔT_{max} (K)
10	0.033	1.30	0.030	0.49
20	0.087	1.68	0.13	1.84
30	0.43	3.67	0.16	2.44
40	0.51	3.66	0.17	4.44
50	0.34	4.60	0.20	1.79
60	0.26	3.74	0.21	2.49

5. Conclusions

This study investigates the application of a cGAN model for rapidly approximating temperature maps in a 3D transient mixed convection process. The trained cGAN effectively infers temperature distributions at any channel location and time, provided an input conditional vector containing numerical information about the channel geometry, thermal and hydraulic boundary conditions, location, and time.

During cGAN training, we compared four architectures of the discriminator network: the PatchGAN classifier and CNN classifier with or without strided convolutions. The comparison suggests that the PatchGAN classifier without strided convolutions is well-suited for generating complex temperature features in unstable mixed convection processes.

We found that cGAN accuracy was influenced by image contrast,

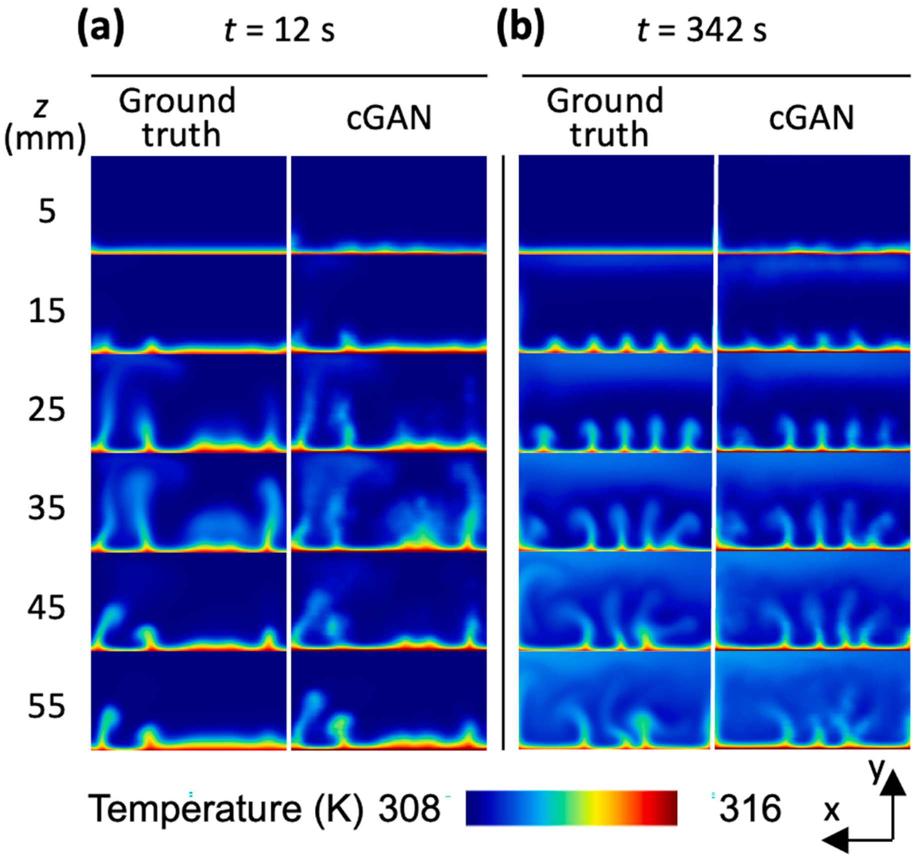

Fig. 4. Temperature distributions at unseen cross-sections in a heated channel predicted by cGAN model and FVM (denoted as ground truth).

spatiotemporal variation rate of temperature, and the number of channels. When the Michelson image contrast was below 0.011, or when temperature features rapidly changed between time steps, the model tended to generate blurred features. Training the model with single-channel images improved the ability to capture complex image features, even for temperature maps with low contrasts.

This work demonstrates the potential of deep learning as a rapid reduced-order model for complex transport processes. The developed cGAN model can be extended to cover a wider range of flow conditions and channel geometries by providing additional training data. This cGAN model will prove valuable for researchers seeking to predict the spatiotemporal variation of fields in heat transfer systems without relying on expensive and computationally intensive numerical simulations. Enhancing the generalization of deep learning models to untrained conditions, such as various flow regimes or geometries, without necessitating extensive training data poses a significant challenge. Exploring the development of models with increased adaptability beyond the training zone emerges as a focal point for future endeavors. One promising opportunity is the incorporation of physics laws into neural networks, as exemplified by the physics-informed neural network framework. This approach holds potential for addressing the need for models that can seamlessly navigate diverse conditions outside their initial training parameters.

CRediT authorship contribution statement

Munku Kang: Data curation, Formal analysis, Investigation, Writing – original draft. **Nam Phuong Nguyen:** Validation, Writing – review & editing. **Beomjin Kwon:** Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

This work was supported by a National Science Foundation grant under Grant No. 2053413.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tsep.2024.102477.

References

[1] A.B. Farimani, J. Gomes, and V.S. Pande, “Deep Learning the Physics of Transport Phenomena,” ArXiv:1709.02432, (2017).
[2] C. Jiang, and A.B. Farimani, “Deep Learning Convective Flow Using Conditional Generative Adversarial Networks,” ArXiv:2005.06422, (2020).
[3] M. Kang, B. Kwon, Deep learning of forced convection heat transfer, J Heat Transfer 144 (2) (2022) 021801.
[4] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, and B.A. Research, “Image-to-Image Translation with Conditional Adversarial Networks,” ArXiv:1611.07004, (2016).
[5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative Adversarial Nets,” ArXiv:2406.2661, (2014).
[6] R. Sharma, A. Barati, F. Bioengineering, J. Gomes Bioengineering, P.E. Bioengineering, and V.P. Bioengineering, “Weakly-Supervised Deep Learning of Heat Transport via Physics Informed Loss,” ArXiv:1807.11374, (2018).

[7] M. Edalatifar, M.B. Tavakoli, M. Ghalambaz, F. Setoudeh, Using deep learning to learn physics of conduction heat transfer, J. Thermal Anal. Calorim. 146 (3) (2020) 1435–1452.

[8] J. Kim, C. Lee, Prediction of turbulent heat transfer using convolutional neural networks, J Fluid Mech 882 (A18) (2020).

[9] N. Zobeiry, K.D. Humfeld, A physics-informed machine learning approach for solving heat transfer equation in advanced manufacturing and engineering applications, Eng Appl Artif Intell 101 (2021) 104232.

[10] S. Sundar, M.C. Rajagopal, H. Zhao, G. Kuntumalla, Y. Meng, H.C. Chang, C. Shao, P. Ferreira, N. Miljkovic, S. Sinha, S. Salapaka, Fouling modeling and prediction approach for heat exchangers using deep learning, Int J Heat Mass Transf 159 (2020) 120112.

[11] D.G. Osborne, F.P. Incropera, Laminar, mixed convection heat transfer for flow between horizontal parallel plates with asymmetric heating, Int J Heat Mass Transf 28 (1) (1985) 207–217.

[12] S. Ostrach, Y. Kamotani, Heat transfer augmentation in laminar fully developed channel flow by means of heating from below, J Heat Transfer 97 (2) (1975) 220–225.

[13] M. Yasuo, U. Yutaka, Forced convective heat transfer between horizontal flat plates, Int J Heat Mass Transf 9 (8) (1966) 803–817.

[14] A.E. Bergles, R.R. Simonds, Combined forced and free convection for laminar flow in horizontal tubes with uniform heat flux, Int J Heat Mass Transf 14 (12) (1971) 1989–2000.

[15] S.T. McComas, E.R.G. Eckert, Combined free and forced convection in a horizontal circular tube, J Heat Transfer 88 (2) (1966) 147–152.

[16] H.F. Oztop, K. Al-Salem, A review on entropy generation in natural and mixed convection heat transfer for energy systems, Renew Sustain Energy Rev 16 (1) (2012) 911–920.

[17] R. Andrzejczyk, T. Muszynski, Thermodynamic and geometrical characteristics of mixed convection heat transfer in the shell and coil tube heat exchanger with baffles, Appl Therm Eng 121 (5) (2017) 115–125.

[18] N. Ghorbani, H. Taherian, M. Gorji, H. Mirgolbabaei, Experimental study of mixed convection heat transfer in vertical helically coiled tube heat exchangers, Exp Therm Fluid Sci 34 (7) (2010) 900–905.

[19] M. Selmi, M.J. Al-Khawaja, A. Marafia, Validation of CFD simulation for flat plate solar energy collector, Renew Energy 33 (3) (2008) 383–387.

[20] M. Edalatpour, J.P. Solano, Thermal-hydraulic characteristics and exergy performance in tube-on-sheet flat plate solar collectors: effects of nanofluids and mixed convection, Internat J Thermal Sci 118 (2017) 397–409.

[21] M. Banna, L. Pietri, B. Zeghmati, Turbulent mixed convection of heat and water vapor transfers in a two-dimensional vegetation canopy, Heat Mass Transfer/ waerme- Und Stoffuebertragung 40 (10) (2004) 757–768.

[22] A.J. Smith, Mixed convection and density-dependent seawater circulation in coastal aquifers, Water Resour Res 40 (8) (2004) 8309.

[23] M. Kumari, Variable viscosity effects on free and mixed convection boundary-layer flow from a horizontal surface in a saturated porous medium - variable heat flux, Mech Res Commun 28 (3) (2001) 339–348.

[24] M. Mirza, and S. Osindero, “Conditional Generative Adversarial Nets,” ArXiv: 1411.1784, (2014).

[25] J. Dai, Y. Lu, and Y.N. Wu, “Generative Modeling of Convolutional Neural Networks,” ArXiv:1412.6296, (2014).

[26] D.P. Kingma, and M. Welling, “Auto-Encoding Variational Bayes,” ArXiv: 1312.6114, (2013).

[27] A. Krizhevsky, I. Sutskever, and G.E. Hinton, in *Advances in Neural Information Processing Systems 25 (NIPS 2012)* (2012).

[28] S. Ioffe, and C. Szegedy, in *Proceedings of the 32nd International Conference on Machine Learning* (2015), pp. 448–456.

[29] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” ArXiv:1511.06434, (2016).

[30] D.P. Kingma, and J.L. Ba, “Adam: A Method for Stochastic Optimization,” ArXiv: 1412.6980, (2014).