International Journal of Heat and Mass Transfer 202 (2023) 123728

Contents lists available at ScienceDirect

International Journal of Heat and Mass Transfer

journal homepage: www.elsevier.com/locate/hmt

Review

A systematic approach to optimization of ANN model parameters to predict flow boiling heat transfer coefficient in mini/micro-channel heatsinks

Yue Qiu[a], Tinh Vo[a], Deepak Garg[b], Hyounsoon Lee[c], Chirag R. Kharangate[a,*]

[a] Department of Mechanical and Aerospace Engineering, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA
[b] Mechanical and Aerospace Engineering Department, University of California, Los Angeles, CA 90095, USA
[c] School of Mechanical Engineering, Chung-Ang University, Seoul, South Korea

ARTICLE INFO

Article history:
Received 16 September 2022
Revised 12 November 2022
Accepted 28 November 2022

Keywords:
Machine learning
Neural networks
Optimization
Heat transfer
Flow boiling

ABSTRACT

Flow boiling in mini/micro-channel heatsinks is often utilized to meet the high heat dissipation requirements of thermal management systems. However, accurate prediction of the heat transfer coefficients in these flow boiling systems is challenging due to complex two-phase fluid behavior compounded by the thermal complexities of these systems. Traditionally, universal correlations, theoretical models, and computational simulations were utilized in predicting heat transfer even though with lower accuracy. Recent work includes use of machine learning tools in more accurately predicting heat transfer behaviors. In this study, Artificial Neural Network (ANN) is utilized and optimized using a comprehensive physics-based and data-driven approach to predict heat transfer coefficients for saturated flow boiling in mini/micro channels. A database of 16,953 data points for flow boiling heat transfer in mini/micro-channels amassed from 50 sources was included in the analysis. The predictive capabilities of the ANN-based method are thoroughly optimized for its input parameters and model architecture parameters something not attempted in any past work. A systematic approach to optimization the model parameters is developed based on correlations from prior literature on flow boiling heat transfer coefficients, Pearson coefficient correlations, and mutual information feature selection method. The final optimized 17 input parameters are trained on the ANN model with network hidden layers (10, 20, 50, 100, 200, 400). This test case achieved an MAE of 8.48% far superior to universal correlations or prior machine learning results for saturated flow boiling heat transfer in mini/micro-channels. These results demonstrate great improvements in accuracy, and a potentially useful framework for optimizing machine learning models for predicting heat transfer coefficients that can be implemented on other two-phase flow configurations and parameters.

1. Introduction

1.1. Demand for two-phase flow boiling heat exchangers

In the past few decades, the rapid development of microelectronics and the rapid reduction in their size has increased their energy consumption per surface area. This along with various inherently energy-intensive applications, like power electronics, lasers, nuclear reactors etc., call for more aggressive cooling methods to effectively dissipate their heat. A well-designed compact cooling heatsink should not only maximize the performance of the devices but also prevent any failures that can be caused by overheating. Two-phase cooling schemes in mini/micro-channel heatsinks can provide significantly higher heat transfer coefficients in comparison to their single-phase cooling counterparts making them a prime candidate for future thermal management systems. Mini/micro-channels boiling can provide multiple benefits, including higher surface-to-volume ratios, higher heat transfer coefficients, more uniform wall temperature distribution, and low fluid inventory requirements [1,2].

Flow boiling in a mini/micro-channel heatsinks transitions across various flow patterns depending on the inlet flow rate and the local flow quality [3]. In addition, the corresponding local heat transfer coefficient variation also depends on if the flow is nucleate boiling dominant or convective boiling dominant [3]. Traditionally, predicting heat transfer coefficients in single-phase flows or two-phase flows corresponding to a specific design was achieved with the use of empirical or semi-empirical correlations and numerical

* Corresponding author.
E-mail address: chirag.kharangate@case.edu (C.R. Kharangate).

https://doi.org/10.1016/j.ijheatmasstransfer.2022.123728
0017-9310/

Nomenclature

AR	aspect ratio
b	bias
Bd	Bond number, $Bd = g(\rho_f - \rho_g)D_h^2/\sigma$
Bo	Boiling number, $Bo = q''_H/Gh_{fg}$
C	Confinement number
Co	Convection number, $Co = [(1-x)/x]^{0.8}(\rho_g/\rho_f)^{0.5}$
C_p	specific heat at constant pressure
C_v	specific heat at constant volume
D_h	hydraulic diameter of flow channel
E	total error
F	enhancement factor
E_m	loss function
Fr_f	saturated liquid Froude number, $Fr_f = [G(1-x)]^2/(\rho_f^2 gD_h)$
Fr_g	saturated vapor Froude number, $Fr_g = (Gx)^2/(\rho_g^2 gD_h)$
Fr_{fo}	liquid-only Froude number, $Fr_{fo} = G^2/(\rho_f^2 gD_h)$
Fr_{go}	vapor-only Froude number, $Fr_{go} = G^2/(\rho_g^2 gD_h)$
F_f	fluid-dependent parameter
f	activation function
G	mass velocity
g	gravity acceleration
h	heat transfer coefficient
$h_{tp,exp}$	experiment measured two-phase heat transfer coefficient
$h_{tp,pred}$	predicted two-phase heat transfer coefficient
h_{fg}	latent heat of vaporization
i	connecting node i
j	connecting node j
k	liquid conductivity
L	total number of input and output layers / channel length
l	layer
m	training example
M	number of training examples; molecular mass
MAE	mean absolute error
MSE	mean square error
n	number of input parameters
Nu	Nusselt number
P	pressure
P_c	critical pressure
Pe_f	saturated liquid Peclet number, $Pe_f= Re_f\ Pr_f$
Pe_g	saturated vapor Peclet number, $Pe_g= Re_g\ Pr_g$
P_F	wetted perimeter of channel
pH	heated perimeter of channel
P_R	reduced pressure, $P_R= P_c/P$
Pr_f	saturated liquid Prandtl number, $Pr_f = \mu_f c_{pf}/k_f$
Pr_g	saturated vapor Prandtl number, $Pr_g = \mu_g c_{pg}/k_g$
q''	heat flux
q''_H	heat flux based on heated perimeter of channel
R	relative roughness, $R = e/D_h$
R^2	coefficient of determination
Re_f	saturated liquid Reynolds number, Re_f=$G(1\text{-}x)D_h/\mu_f$
Re_g	saturated vapor Reynolds number, Re_g=GxD_h/μ_g
Re_{fo}	liquid-only Reynolds number, Re_{fo}=GD_h/μ_f
Re_{go}	vapor-only Reynolds number, Re_{go}=GD_h/μ_g
S	output, suppression factor
Su_f	saturated liquid Suratman number,
Su_g	saturated vapor Suratman number,
T	temperature
u	velocity at inlet
w	weight of connecting node
We_f	saturated liquid Weber number, $We_f = [G(1-x)]^2 D_h/(\rho_f\sigma)$
We_g	saturated vapor Weber number, $We_g = (Gx)^2 D_h/(\rho_g\sigma)$
We_{fo}	liquid-only Weber number, $We_{fo} = G^2D_h/(\rho_f\sigma)$
We_{go}	vapor-only Weber number, $We_{go} = G^2D_h/(\rho_g\sigma)$
Xtt	lockhart-martinelli parameter, $Xtt = (\mu_f/\mu_g)^{0.1}[(1-x)/x]^{0.9}(\rho_g/\rho_f)^{0.5}$
x	quality, value of the node, input vector
x_{di}	dryout incipience quality

Greek symbols

α	vapor void fraction
α_r	regularization parameter
β	aspect ratio, exponential decay rate
δ	error at the node
ε	percentage data predicted within ±50%,
θ	percentage data predicted within ±30%,
μ	dynamic viscosity
λ	learning rate
ρ	density, Pearson's correlation coefficient
σ	surface tension

Subscripts

c	critical
cb	convective boiling
e	effective value with flow
exp	experimental
f	saturated liquid, fluid
fo	liquid only
g	saturated vapor
go	vapor only
l	liquid
nb	nucleate boiling
$pred$	predicted
$pool$	pool boiling
sp	single-phase
tp	two-phase
w	heated wall; inner wall

simulations [4]. Experiments are first performed over a range of operating and geometric parameters, then correlation form is selected, and finally the coefficients for the correlation form are calculated such that the mean error based on all the data points is minimized. These correlations forms are developed based on physical understanding of the flow behaviors and/or the experimental tests showing specific flow parameters having a direct impact on the heat transfer coefficient data.

1.2. Previous flow boiling correlations

There have been several past efforts focused on developing heat transfer coefficient correlations in the context of flow boiling. Shah [5], Cooper [6], Gungor and Winterton [7], Liu and Winterton [8] developed correlations for macro-sized channels, while the mini/micro-channels correlations were developed by Oh and Son [9], Lazarek and Black [10], Tran et al. [11], Warrier et al. [12], Yu et al. [13], Agostini and Bontemps [14], Bertsch et al. [15], Li and Wu [16], Ducoulombier [17], and Kim and Mudawar [3]. Among these authors, Lazarek and Black [10], Cooper [6], Tran et al. [11], Warrier et al. [12], Yu et al. [13], Ducoulombier et al. [17], and Oh and Son [9] interpolated correlations from their own experimental data; while Shah [5], Gungor and Winterton [7], Liu and Winterton [8], Bertsch et al. [15], Li and Wu [16], and Kim and Mudawar

[3] consolidated databases to derive their correlations. It is useful to understand how each of the authors select relationships for predicting heat transfer coefficient.

Table 1 summarizes the past correlations proposed for predicting flow boiling heat transfer. In these correlations, some of them calculated the two-phase heat transfer coefficient directly from input parameters. Examples of such correlations are those by Lazarek and Black [10], Cooper [6], Tran et al. [11], Yu et al. [13], Oh and Son [9] and Agostini and Bontemps [14]. Other correlations use some form of correlation between the single-phase and nucleate boiling and employ the use of suppression (S) and enhancement factor (E) as weights for each of the terms. These include correlations by Shah [5], Gungor and Winterton [7], and Liu and Winterton [8], with Warrier et al. [12] only correlates the two-phase heat transfer coefficient as a function of the single-phase boiling term with the enhancement factor E as a weight. Similarly, correlations by Bertsch et al. [15] and Kim and Mudawar [3] correlate between convective and nucleate boiling terms with suppression (S) and enhancement factor (E) as weights, while Ducoulombier et al. [17] determines the heat transfer coefficient as the maximum between nucleate and convective boiling terms.

The individual studies see various correlation forms being used for the nucleate boiling (also called two-phase term in some correlations) and convective boiling terms. In some early work, Chen [18] assumed an interaction between micro-convective and macro-convective mechanisms on net flow boiling heat transfer behaviors. The full form of the original Chen [18] correlation is as follows:

$$h = h_{mic} + h_{mac} \tag{1}$$

$$h_{mic} = 0.00122\left(\frac{k_{fo}^{0.79} C_{p,l}^{0.45} \rho_{fo}^{0.49} g^{0.25}}{\sigma^{0.5} \mu_f^{0.29} h_{fg}^{0.24} \rho_{go}^{0.24}}\right)(\Delta T_e)^{0.24}(\Delta P_e)^{0.75} S \tag{2}$$

$$h_{mac} = 0.023 Re_{fo}^{0.8} Pr_{fo}^{0.4}\left(\frac{k_f}{D_h}\right) F \tag{3}$$

Where

$$S = \left(\frac{\Delta T_e}{\Delta T}\right)^{0.24}\left(\frac{\Delta P_e}{\Delta P}\right)^{0.75} \tag{4}$$

$$F = \left(\frac{Re}{Re_{fo}}\right)^{0.8} \tag{5}$$

Here, the macro-convective term was derived from the Dittus-Boetler [19] correlation, with a modification factor, F, being the ratio of two-phase Reynolds number to the liquid-only Reynolds number while the micro-convective term was derived from the Foster and Zuber [20] correlation form. This correlation form is popular and used in later studies conducted by Bertsch et al. [15], Gungor and Winterton [7], Liu and Winterton [8], Warrier et al. [12], Ducoulombier et al. [17]. Enhancement weight E in Liu and Winterton's correlation [8] is calculated from Chen's Reynolds analogy [18] in a homogenous model, excluding the Martinelli parameter, and coupling with the Prandtl number in its final form. Ducoulombier [17] separated the boiling data into 3 regimes. For type I, with low quality (x<0.1), the Cheng et al. [21] version of the Cooper correlation is used. For higher quality, the flow is then divided into a high Boiling number regime (type II) and low Boiling number regime (type III). Type III uses interpolated data with the inverse of Martinelli parameter. Yu et al. [13] modified Chen [18] correlation form with the Martinelli parameter.

Another common correlation form used for the nucleate boiling term is the Cooper correlation [6] form developed based on pool boiling heat transfer coefficient. The Cooper correlation form is as follows:

$$h_{tp} = 55 P_R^{0.12}(-log_{10}(P_R))^{-0.55} M^{-0.5} q''^{0.67}_H \tag{6}$$

This form can be seen in studies by Gungor and Winterton [7], Liu and Winterton [7], Bertsch et al. [15], and Ducoulombier et al. [17]. This correlation was originally developed to be applicable for pool boiling but was later extended to be used for flow boiling. The parameters include Prandtl number, molecular mass, and heat flux as their significance is easily captured on the authors' working data set.

Ducoulombier et al. [17], and Oh and Son [9] used the Schrock and Grossman [22] correlation form. The Schrock and Grossman [22] correlation form is as follows:

$$h_{tp} = 170\left(Bo + 1.5 \times 10^{-4} Xtt^{-2/3}\right) h_{fo} \tag{7}$$

Schrock and Grossman [22] correlated the heat transfer with the Boiling number and an inverse power of Martinelli parameter. Ducoulombier et al. [17] used this form to calculate the convective boiling term, while Oh and Son [9] modified the form to get the final two-phase heat transfer coefficient. Warrier et al. [12] proposed a correlation with enhancement factor based only on the Boiling number and the vapor quality. Bertsch et al. [15] used the Hausen correlation form [23] for the pure liquid and pure vapor convective heat transfer terms, noting low Reynolds number associated with mini/micro-channels, and considering the bubble confinement in small channels with enhancement and suppression weighting terms. Liu and Winterton [8] correlated the nucleate and convective boiling terms according to the Kutateladze [24] correlation form. Kutateladze [24] recognized the effect of flow rates at low pressure, and combined the effects of pool boiling and forced convective boiling as follows:

$$h_{tp}^2 = \left(F h_{fo}\right)^2 + \left(S h_{pool}\right)^2 \tag{8}$$

Ducoulombier et al. [17] used the Hihara and Tanaka [25] correlation form for type II (Bo $>1.1 \times 10^{-4}$) of their correlation, which is an empirical interpolation of the Schrock and Grossman [22] correlation form on their own data and includes effect of heat flux and agrees well with low mass fluxes heat transfer. Tran et al. [11] made modifications to the Tran et al. [11] form to calculate the two-phase heat transfer coefficient. Tran et al. defined their heat transfer coefficient as follows:

$$h_{tp} = 8.4 \times 10^5 \left(Bo^2 We_{fo}\right)^{0.3}\left(\frac{\rho_g}{\rho_f}\right)^{0.4} \tag{9}$$

Kim and Mudawar [3] used the Churchill and Usagi [26] general form of correlation interpolation between limiting solutions. This interpolation form has no theoretical reasoning behind it but showed good performance in interpolating experimental data and solutions for correlating rates of transfer for uniformly varying processes. Churchill and Usagi's [26] general correlation form is used to correlate convective and nucleate boiling terms by Kim and Mudawar [3] in the form:

$$h_{tp} = \left(h_{nb}^n + h_{cb}^n\right)^{1/n} \tag{10}$$

Kim and Mudawar [3] follow the functional form of Schrock and Grossman [22] equation for the convective boiling term and the functional form of Dittus and Boetler [19] correlation for the nucleate boiling term based on Reynolds number in the pure liquid phase. Martinelli parameter is deemed insignificant and is omitted in nucleate boiling, while a term is introduced that is a function of quality to account for nucleate boiling suppression. Reduced pressure is included for both boiling modes to adapt to different working fluid properties and working flow conditions; wetted and heated perimeter ratio is included to account for 3-sided wall heating. They also included Weber's number for the pure fluid phase We_{fo} to account for inertia and surface tension effects encountered due to small geometry sizes in convective boiling regime.

Lazarek and Black [10] calculated Nusselt number as a weighted function of Boiling number and Reynolds number. A strong heat

Table 1
Flow boiling heat transfer correlations.

Author	*Correlation*	*Notes*
Lazarek and Black [10]	$h_{tp} = (30Re_{fo}^{0.857}Bo^{0.714})(\frac{k_f}{D_h}),\ Re_{fo} = \frac{GD_h}{\mu_f},\ \mathrm{Bo} = \frac{q''_H}{Gh_{fg}}$	D = 3.15 mm, R113, nucleate boiling dominant
Shah [5]	$h_{tp} = Eh_{sp} + Sh_{nb},\ h_{sp} = 0.023\mathrm{Re}_f^{0.8}\,\mathrm{Pr}_f^{0.4}\,\frac{k_f}{D_h}$ for $\mathrm{N} > 1.0, \mathrm{S} = \frac{1.8}{\mathrm{N}^{0.8}}, \mathrm{E} = 230\mathrm{Bo}^{0.5}$ for $\mathrm{Bo} > 3 \times 10^{-5}$, or $\mathrm{E} = 1 + 46\mathrm{Bo}^{(0.5)}$ for $\mathrm{Bo} < 3 \times 10^{-5}$, *for* $0.1 < \mathrm{N} \leqslant 1.0, S = \frac{1.8}{N^{0.8}}, E = FBo^{0.5}\exp(2.74N^{-0.1})$ *for* $\mathrm{N} \leqslant 0.1, S = \frac{1.8}{N^{0.8}}, E = FBo^{0.5}\exp(2.47N^{-0.15})$ $F = 14.7$ for $\mathrm{Bo} \geqslant 11 \times 10^{-4}$, or $F = 15.43$ for $\mathrm{Bo} < 11 \times 10^{-4}$ $N = Co$ *for vertical tube,* $N = Co$ *for horizontal tube with* $Fr_f \geqslant 0.04$ $N = 0.38Fr_f^{-0.3}\mathrm{Co}$, for horizontal tube with $\mathrm{Fr_f} < 0.04$, $Re_f = \frac{G(1-x)D_h}{\mu_f}, Co = (\frac{1-x}{x})^{0.8}(\frac{\rho_g}{\rho_f})^{0.5} Fr_f = \frac{G^2}{\rho_f^2 gD_h}$	D = 6–25.4 mm, water, R11, R12, R22, R113, cyclohexane, 780 data points
Cooper [6]	$h_{tp} = 55P_R^{0.12}(-log_{10}(P_R))^{-0.55}M^{-0.5}q_H''^{0.67}$	6000 data points, developed for nucleate pool boiling
Gungor and Winterton [7]	$h_{tp} = Eh_{sp} + Sh_{nb},\ h_{sp} = 0.023Re_f^{0.8}Pr_f^{0.4}\frac{k_f}{D_h}$ $E = 1 + 24000Bo^{1.16} + 1.37(\frac{1}{X_{tt}})^{0.86}$, $E = 1 + 24000Bo^{1.16} + 1.37(\frac{1}{X_{tt}})^{0.86}$, $h_{nb} = \mathrm{h_{tp,Cooper}},\ S = (1 + 1.15 \times 10^{-6}\mathrm{E}^2 Re_{\mathrm{f}}^{1.17})^{-1}$, *for horizontal tube with*$Fr_f \leqslant 0.05$,*replace E and S with,* $EFr_f^{(0.1-2Fr_f)}$ and $SFr_f^{0.5}$, respectively	D = 2.95–32.0 mm, water, R11, R12, R113, R114, R22, ethylene glycol, 4300 data points
Liu and Winterton [8]	$h_{tp} = [(\mathrm{E}h_{sp})^2 + (\mathrm{Sh_{nb}})^2]^{0.5},\ h_{sp} = 0.023Re_{fo}^{0.8}Pr_f^{0.4}\frac{k_f}{D_h}$ $\mathrm{E} = [1 + xPr_f(\frac{\rho_f}{\rho_g} - 1)]^{0.35}$, $h_{nb} = \mathrm{h_{tp,Cooper}},\ \mathrm{S} = (1 + 0.055E^{0.1}Re_{fo}^{0.16})^{-1}$, for horizontal tube with $Fr_f \leqslant 0.05$, replace E and S with $EFr_f^{(0.1-2Fr_f)}$ and $SFr_f^{0.5}$, respectively	Same data as Gungor and Winterton's [7]
Tran et al. [11]	$h_{tp} = 8.410^5 \times (Bo^2We_{fo})^{0.3}(\frac{\rho_g}{\rho_f})^{0.4},\ We_{fo} = \frac{G^2D_h}{\rho_f\sigma}$	D = 2.46, 2.92 mm, D_h = 2.40 mm, R12, R113, nucleate boiling dominant
Warrier et al. [12]	$h_{tp} = Eh_{sp},\ h_{sp} = 0.023Re_{fo}^{0.8}Pr_f^{0.4}\frac{k_f}{D_h}$ $\mathrm{E} = 1.0 + 6.0\mathrm{Bo}^{1/16} - 5.3(1 - 855\mathrm{Bo})\mathrm{x}^{0.65}$	D_h = 0.75 mm, five parallel, FC84
Yu et al. [13]	$h_{tp} = 6.410^6 \times (Bo^2We_{fo})^{0.27}(\frac{\rho_g}{\rho_f})^{0.2}$	D = 2.98 mm, water, ethylene glycol, nucleate boiling dominant
Agostini and Bontemps [14]	$h_{tp} = 28q_H''^{2/3}G^{-0.26}x^{-0.10}$ for $\mathrm{x} < 0.43$, $h_{tp} = 28q_H''^{2/3}G^{-0.64}x^{-2.08}$ for $\mathrm{x} > 0.43$	D_h = 2.01 mm, 11 parallel, R134a
Bertsch et al. [15]	$h_{tp} = \mathrm{E}h_{cb} + Sh_{nb},\ h_{cb} = \mathrm{h_{sp,fo}}(1 - \mathrm{x}) + \mathrm{h_{sp,go}}\mathrm{x}$, $\mathrm{E} = 1 + 80(\mathrm{x}^2 - \mathrm{x}^6)\exp(-0.6\mathrm{N_{conf}}),\ \mathrm{S} = 1 - \mathrm{x}$, $h_{nb} = 55P_R^{0.12-0.2log_{10}(R_p)}(-log_{10}(P_R))^{-0.55}M^{-0.5}q_H''^{0.67}$, $N_{conf} = \sqrt{\frac{\sigma}{g(\rho_f-\rho_g)D_h^2}}$, $h_{sp,fo} = (3.66 + \frac{0.0668\frac{D_h}{L}Re_{fo}Pr_f}{1+0.04[\frac{D_h}{L}Re_{fo}Pr_f]^{2/3}})\frac{k_f}{D_h}$, $h_{sp,go} = (3.66 + \frac{0.0668\frac{D_h}{L}Re_{go}Pr_g}{1+0.04[\frac{D_h}{L}Re_{go}Pr_g]^{2/3}})\frac{k_g}{D_h}\ ,Re_{fo} = \frac{GD_h}{\mu_f},\ Re_{go} = \frac{GD_h}{\mu_g}$	D_h = 0.16–2.92 mm, water, refrigerants, FC-77, nitrogen, 3899 data points
Li and Wu [16]	$h_{tp} = 334 \times Bo^{0.3} = (BdRe_f^{0.36})^{0.4}\frac{k_f}{D_h},\ \mathrm{Bd} = \frac{g(\rho_f-\rho_g)D_h^2}{\sigma}$	$\mathrm{D_h}$ = 0.16–3.1 mm, water, refrigerants, FC-77, ethanol, propane, CO_2, 3744 data points
Ducoulombier et al. [17]	$h_{tp} = \max(h_{nb}, h_{cb})$, $h_{nb} = 131P_R^{-0.0063}(-log_{10}(P_R))^{-0.55}M^{-0.5}q_H''^{0.58}$, *if Bo > 1.1 × 10–4,* $h_{cb} = [1.47 \times 10^4 Bo + 0.93(\frac{1}{X_{tt}})^{2/3}](0.023Re_{fo}^{0.8}Pr_f^{1/3}\frac{k_f}{D_h})$, *if Bo < 1.1 × 10–4,* $h_{cb} = [1 + 1.80(\frac{1}{X_{tt}})^{0.986}](0.023Re_f^{0.8}Pr_f^{0.4}\frac{k_f}{D_h})$	D = 0.529 mm, CO_2
Oh and Son [9]	$h_{tp} = 0.034Re_f^{0.8}Pr_f^{0.3}[1.58(\frac{1}{X_{tt}})^{0.87}](\frac{k_f}{D_h})$	D = 1.77, 3.36, 5.35 mm, R134a, R22
Kim and Mudawar [3]	$h_{tp} = (h_{nb}^2 + h_{cb}^2)^{0.5,}$ $h_{nb} = [2345(Bo\frac{P_H}{P_F})^{0.70}P_R^{0.38}(1-x)^{-0.51}](0.023Re_f^{0.8}Pr_f^{0.4}\frac{k_f}{D_h})$, $h_{cb} = [5.2(Bo\frac{P_H}{P_F})^{0.08}We_{fo}^{-0.54} + 3.5(\frac{1}{X_{tt}})^{0.94}(\frac{\rho_g}{\rho_f})^{0.25}](0.023Re_f^{0.8}Pr_f^{0.4}\frac{k_f}{D_h})$,	10,805 data points D = 0.19–6.5 mm, R113, R134a, R22, R11, R123, Water, CO_2, FC72, R236fa, R245fa, R404A, R407C, R410A, R417A, R152a, R1234yf, R32, R1234ze, Circular/rectangular, horizontal, vertical upflow, G = 19–1608 kg/m2 s, x = 0–1, 12,974 data points from 31 sources

flux influence with minor vapor quality dependence was seen. Tran et al. [11] based the heat transfer coefficient on Lazarek and Black [10], noticing a dependence on Boiling number and Reynolds number. Reynolds number then was replaced by Weber number to reflect the dominance of nucleate boiling to convection boiling. Agostini and Bontemps [14] developed a correlation for the nucleate boiling regime as a function of mass flux, heat flux and vapor quality where the vapor quality of 0.43 was identified as the critical quality where dry-out occurs. Li and Wu [16] accounted for surface tension, body force, viscous and inertia force by including boiling number, Bond number Reynolds number in their correlation.

1.3. Previous machine learning efforts and techniques for thermal analysis

With a plethora of correlations available, each dependent on different set of parameters, the development of a consensual correlation form for heat transfer is elusive. In practice, the factors influencing heat transfer coefficient include groups of dimensionless and dimensional numbers, each valid over a finite range of values. The functions derived from these parameters are then interpolated using the experimental data to develop the correlation. A general trend to be seen from these correlations is that while they suffice in narrow dimensional ranges of flow boiling configuration, whose corresponding experimental data they are cross-referenced against, they deviate significantly in terms of accuracy when applied to configurations out of their tested range.

An approach to better understand and selecting parameters with the highest impact on the heat transfer is the implementation of soft-computing techniques such as Artificial Neural Networks (ANNs), Decision Trees, Random Forest, Gradient Boosting, Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Support Vector Machines (SVM) [27]. Amongst these techniques, ANN is one of the supervised learning techniques capable of recognizing hidden relations and patterns in a large datasets [28]. There have been several efforts to integrate neural networks in the analysis of thermal systems. Goudarzi et al. [29] used a backpropagation trained neural network to the heat transfer rate in the exhaust valve in internal combustion engines, concluding that the Levenberge-Marquardt algorithm provides the best performance when used to train the neural network. Huang et al. [30] compared results from training a neural network to a 1D turbocharger model in the GT-power software in predicting heat transfer in a turbocharger simulation, noting better performance resulting from the neural network. Peng and Ling [31] predicted the Colburn factor and the friction factor in plate-fin heat exchangers, achieving mean square error as low as 1.5% and 1% for the Colburn factor and the friction factor, respectively. Chang et al. [32] predicted heat transfer of supercritical water using neural network over a dataset of 5280 experimental datapoints from various publications, noting the neural network's possible dependence on being trained by a single geometry or working fluid. Tan et al. [33] predicted thermal performance of a compact fin-tube heat exchanger with neural networks and conventional non-linear regression models, resulting in better performance from the neural network compared to the latter. Nie et al. [34] combined the machine learning models with universal correlations for flow condensation. They implemented parameter selection on heat transfer coefficients based on a trained Extreme Gradient Boosting machine learning model (XGBoost) [35] and showed that the new universal correlation provided better predictive performances.

None of the previous studies employed any parametric optimization which can map the relationship between input parameters and the heat transfer. One attempt at preliminary optimization was conducted by Qiu et al. [27], who consolidated a large database of saturated flow boiling heat transfer in mini/micro-channels. However, their final model used the complete set of non-dimensionless input parameters with some parametric investigation of the network architecture to reduce mean error. The model predicted heat transfer coefficient of the test data with an MAE of 14.3%, and the percentage of datapoints lying within ±30% (θ) and ±50% (ε) were 92.0% and 97.4%, respectively. While this indicates reasonably good predicting capability of the heat transfer coefficient, the selected input parameters and the hyperparameters in their study were not thoroughly optimized specific to the problem under investigation.

1.4. Objective of study

The objective of the present study is to develop a systematic approach to optimize the Artificial Neural Network (ANN) model for predicting flow boiling heat transfer coefficient in mini/micro-channel heat sinks. The model input parameters will be determined by conjunction of physics-based correlation parameters and data-driven parameters. The model will utilize a consolidated dataset of 16,953 data points, collected from experimental studies published in 50 sources, spanning 16 different working fluids and a large range of operating and geometric parameters from the study by Qiu et al. [27]. The dataset has the following coverage and complete information on the individual databases can be found in [27]:

-Working fluids: CO_2, FC72, Propane, R11, R113, R123, R1234yf, R1234ze, R134a, R152a, R22, R236fa, R245fa, R32, R600a and water; Reduced pressures: $0.0046 < P_R < 0.77$; Hydraulic diameter: 0.15 mm $< D_h <$ 6.5 mm; Mass velocity: 19 kg/m^{2}s $< G <$ 1608 kg/m^{2}s; Liquid-only Reynolds number: $27 < Re_{fo} = GD_h / \mu_f < 55{,}270$; Flow quality: $0 < x < 1$.

There is a lot of randomness associated with parameter selections techniques used for developing machine learning tools. In this study, a systematic approach is suggested to develop an optimized ANN model for predicting flow boiling heat transfer coefficient. The initial basis for the selection of input parameters for the model are prior known correlations from the literature. With the aid of effect size statistics (such as Pearson correlation coefficient) and mutual information between variables, additional input parameters will be added to correlation parameters to enhance the prediction accuracy of the developed ANN model. A comprehensive grid search with a large search space will then be used to tune the model hyperparameters.

2. ANN

2.1. ANN background

The idea of simulating the human neural networks as a predictive algorithm was first introduced by McCullough and Pitts in 1943 [36]. The network architecture consists of nodes, emulating neurons, which are organized in several layers. Typically, an artificial neural network (ANN) consists of an input layer that takes a vector $\vec{x} = (x_1, x_2, \ldots x_n)$ of fixed length n, where n is the number of input parameters. The hidden layers consist of interconnected nodes taking in linear combination of the input x and applying an appropriate function $f(x)$, known as activation function, resulting in an output $f(w^T x + b)$, where 'w' and 'b' are the weights and biases, respectively. which in turn becomes the input for the next layer.

The universal approximation theorem states that a neural network with a single hidden layer consisting of finite number of neurons can approximate any continuous function, provided the activation function is smooth [43]. Despite this, the problem of deciding on optimal width and depth of ANN is mostly experimental. The most widely used learning algorithm applied in ANNs is backpropagation, where the transfer function is transferred forward, and the error is propagated backwards, as demonstrated by Rumelhart and Hinton [37].

2.2. ANN model development

The whole dataset comprising of 16,953 datapoints are split into the training set and the test set, comprising 12,624 and 4239 datapoints, respectively. The metrics used to measure model performance are mean absolute error (MAE) and R^2, defined as:

$$MAE = \frac{1}{N}\sum \frac{\left|h_{tp,pred} - h_{tp,exp}\right|}{h_{tp,exp}} \times 100\% \tag{11}$$

$$R^2 = 1 - \frac{\sum\left(h_{tp,pred} - h_{tp,exp}\right)^2}{\sum\left(h_{tp,pred} - \bar{h}_{tp,exp}\right)^2} \tag{12}$$

To improve the predictive capability and reliability of ANN model, its architecture must be optimized. There are several aspects critical to the performance of an ANN, including data distribution, the number of input parameters, activation function, network width and depth, etc. [38]. Activation functions are chosen based on the problem and training considerations (like exploding gradients or vanishing gradients). Early ANNs used a step function which have derivative zero except at the origin, where it is undefined. Consequently, several other activation functions have also been proposed, but for the present study rectified linear unit has been selected as it is fast to compute and although only one-sided derivatives exist at zero, unlike step function does not create numerical instability while training [40,39].

Since the input parameters vary on widely ranging scales, two of the most common ways of scaling data are standardization and min-max scaling. For the present study, the input parameters are standardized to a Gaussian-type distribution with zero mean value and unit variance. The initialization of the ANNs is done with very small random weights as zero-initialized weights pose a problem of symmetry and can cause convergence problem. The Adam solver is implemented due to its demonstrated superior performance in comparison to standard gradient descent algorithm [40]. The first and second moments exponential decay rate is set to be 0.9 and 0.999, respectively. The loss function used by the Adam solver is MSE loss [40], whose tolerance is set to 0.0001 within 100 iterations.

Grid search and random search are two most widely used hyper-parameter optimization techniques. In grid search, set of trials is formed by assembling every possible combination of values a hyper-parameter can take. It must be noted that although random search is better than grid search and manual search [41], it suffers from the drawback of reproducibility of results. Implementation of grid search is simple, and parallelization is a trivial task. If the column is categorical (containing non-numeric values), then each level of that column was encoded as a new column taking only binary numeric values (0 or 1), also known as one-hot-encoding. This results in a new vector space where each category is orthogonal and equidistant to each other [40,39].

2.3. ANN default baseline model development

We first start with a general optimization to select the ANN model parameters including the hidden layer sizes, alpha and learning rate, etc. from the range of parameters to be tested in this study as summarized in Table 2. To perform the general optimization, an exhaustive list of numeric input parameters derived based on the geometric and operating conditions is selected that includes dimensional and non-dimensional parameters that can impact the heat transfer coefficient as shown in Table 2. To select the ANN model architecture, a comprehensive grid search is conducted

Table 2
Range of ANN hyperparameters and list of all parameters in the database.

ANN Model Hyperparameter	Values
Batch Size	16–128
Epochs	500–1000
Learning Rate	0.00025–0.01
Momentum	0.0–0.1
Number of Hidden Layers	2–12
Exhaustive Input Parameter Set	
Bd, Bo, Co, c_{pg}, c_{pf}, c_{vg}, c_{vf}, D, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, G, h_{fg}, k_f, k_g, P_c, P_F, P_H, Pe_f, Pe_g, P, Pr_f, Pr_g, q'', R, Re_f, Re_{fo}, Re_g, Re_{go}, Su_f, Su_g, T, We_f, We_{fo}, We_g, We_{go}, Xtt, x, β, μ_f, μ_g, ρ_f, ρ_g, σ	

Table 3
ANN default model parameters selected in this study.

Parameter	Value
Activation function	ReLu
L2 Regularization Parameter, α[*]	0.01
Solver	Adam
Epochs	1000
Batch Size	32
Learning rate, λ[*]	0.001
Exponential decay rate for estimates of first moment vector, $\beta 1$	0.9
Exponential decay rate for estimates of second moment vector, $\beta 2$	0.999
Tolerance	1.00E-07
Layer sizes[*]	400,200,100,80,50,10

[*] These ANN parameters are the subject of optimization later in the study.

based on Table 2. The final model architecture that the grid search provided for the exhaustive list of inputs is shown in Table 3. This model would serve the initial basis of our optimization to select the non-exhaustive list of input parameters going forward.

3. Results and discussions

3.1. Systematic approach to optimization

Because of large variability possible with modeling parameters and randomness associated with the parameter selection, a systematic approach needs to be followed to be able to reach a near-optimized model architecture. First, we select input parameters from previous physical based correlations. Different combinations of inputs are compared. Then according to data science analysis (Pearson's correlation coefficient and mutual information), additional input parameters will be selected and added to the input sets derived from the analysis of previous correlations. Finally, a comprehensives grid search is performed to optimize the ANN model hyperparameters to obtain the best performing model.

3.2. Correlation-based input parameters investigation

Past work by Qiu et al. [27], He et al. [42], Zhu et al. [43] and Domingos [44] have shown that adding more input parameters does not necessarily improve model performance. We can also see that using all the input parameters together as shown in Table 2 does not result in the best performance. Therefore, in this study, we want to investigate a better strategy to select the most relevant input parameters that should be used as inputs for the ANN model Table 1 summarizes most relevant flow boiling heat transfer coefficient correlations from literature. Utilizing each of the correlations, we select two sets of input parameters to be investigated for heat transfer coefficient predicting accuracy with the ANN model. The first set of parameters is selected based on the actual parameters in the individual correlations, both dimensional and non-dimensional. Consider for example the correlation by Lazarek and Black [10] as shown below:

$$h_{tp} = 30Re_{fo}^{0.857}Bo^{0.714}\frac{k_f}{D_h} \tag{13}$$

The parameters based on this correlation includes Re_{fo}, Bo, k_f, and D_h. The second set is based on only the physical properties, where non-dimensional parameters are broken down to its building compositions. Taking the composition of these equations into account, the parameter list for Lazarek and Black [10] will include k_f, D_h, G, μ_f, q''_H, h_{fg}. This is done for each individual correlation in Table 1. Table 4 shows the MAE obtained for the two sets of parameters for each of the individual correlations. Interestingly, Table 4 shows that across most correlations, dimensional parameters set yield favorable results, in terms of lower MAE, as compared to non-dimensional parameters set. This is very insightful as using dimensional over non-dimensional or vice-versa is an open discussion when it comes to machine learning modeling of physical phenomena like heat transfer coefficient or pressure drop in two-phase flows. Non-dimensional parameters are preferred in physics-based two-phase predictive models due to the generalization possible when developing a reliable design tool for thermal engineers. These results here show that those parameters on their own may not be able to capture the physics involved in boiling situations leading to the need for utilization of the derived dimensional parameters in the machine learning tools.

Furthermore, if we look closely in Table 4, it is also seen that correlations based on interpolation of the two boiling modes as discussed in Table 1, viz. Shah [5], Gungor and Winterton [7], and Liu and Winterton [8], Warrier et al. [12], Bertsch et al. [15], Li and Wu [15], Ducoulombier et al. [17] and Kim and Mudawar [3], have shown low MAE ($<$ 15%), hence, the parameters from these correlations are prime candidates for ANN model input parameters. In comparison, mostly the correlations based on single form terms or single boiling mode show higher MAEs. This points to the multi-mode definitions better representing the boiling phenomena which is also confirmed by the actual correlations MAEs which are lower for two boiling mode forms as shown in Table 4.

Fig. 1 shows the performance of the four best input parameters sets that includes Bertsch et al. [15], and Liu and Winterton [8] based on the actual parameters set and Shah [5] and Gungor and Winterton [7] based on the derived dimensional parameters set. Based on these results, we downselect these four correlations' parameters for further optimization in the following sections using data science driven tools. These four correlations are referenced as the best four correlations in following discussions.

3.3. Bivariate analysis

Previous section selected input parameters (or features) based on physics based correlations. To augment the set of relevant input parameters, two data driven approaches – Pearson's correlation-based parameter selection and mutual information feature selection are used. Having a comprehensive set of relevant features is paramount to machine learning algorithm performance as they improve the performance with respect to model selection, robust generalization, and accurate prediction with unseen data [45].

To determine the bivariate linear correlation between variable (thermodynamic properties) including the dependent variable, heat transfer coefficient (h), Pearson's correlation coefficient is calculated. Pearson's correlation produces coefficient which measures the strength and direction of two continuous random variables and is parametric in nature – replying heavily on the mean values of the random variables [46]. Fig. 2 shows the Pearson's correlation coefficient (ρ) heat map between the relevant variables presents in the data, including 'h'. Two uncorrelated variables would have an absolute value close to zero, while positively correlated

Table 4
Input parameters list selected for optimization based on flow boiling correlations.

Serial No.	Correlation	Actual Correlation Parameters	MAE by ANN model	Dimensional Parameters	MAE by ANN model	MAE by original correlation
1	Lazarek and Black [10]	Bo, D_h, k_f, Re_{fo}	22.37%	D_h, G, h_{fg}, k_f, q''_H, μ_f	21.15%	33.27%
2	Shah [5]*	Bo, Co, D_h, Fr_f, Pr_f, Re_f, k_f	13.38%	$C_{p,f}$, D_h, G, h_{fg}, k_f, q''_H, u, x, μ_f, ρ_f, ρ_g	10.35%	35.41%
3	Cooper [6]	M, P_R, q''_H,	27.55%	(Same as actual)	N/A	38.37%
4	Gungor and Winterton [7]*	Bo, D_h, Fr_f, M, P_R, Pr_f, q''_H, Re_f, x_{tt},	14.09%	$C_{p,f}$, D_h, G, h_{fg}, k_f, M, P_R, q''_H, u, x, μ_f, μ_g, ρ_f, ρ_g	10.24%	59.18%
5	Liu and Winterton [8]*	D_h, Fr_f, k_f, M, P_R, Pr_f, q''_H, Re_{fo}, x, ρ_f, ρ_g	10.4%	$C_{p,f}$, D_h, G, k_f, M, P_R, q''_H, u, μ_f, ρ_f, ρ_g	11.37%	32.94%
6	Tran et al. [11]	Bo, We_{fo}, ρ_f, ρ_g	25.24%	D_h, G, h_{fg}, k_f, q''_H, ρ_f	21.25%	45.45%
7	Warrier et al. [12]*	Bo, D_h, k_f, Pr_f, Re_{fo}	21.73%	C_{pf}, D_h, G, h_{fg}, k_f, q''_H, x, μ_f	12.06%	39.79%
8	Yu et al. [13]	Bo, We_{fo}, ρ_f, ρ_g	25.92%	D_h, G, h_{fg}, q''_H, ρ_g, ρ_f	21.61%	84.77%
9	Agostini and Bontemps [14]	G, q''_H, x	30.84%	(same as actual)	30.84%	57.33%
10	Bertsch et al. [15]*	D_h, k_f, k_g, M, P_R, Pr_f, Pr_g, q''_H, Re_{fo}, Re_{go}, x, ρ_f, ρ_g, σ	10.11%	$C_{p,f}$, $C_{p,g}$, D_h, G, h_{fg}, L, M, k_f, k_g, P_R, R, q''_H, x, μ_f, μ_g, ρ_g, ρ_f, σ	11.10%	46.33%
11	Li and Wu [16]	Bo, Bd, D_h, k_f, Re_f	21.61%	D_h, G, h_{fg}, k_f, q''_H, x, μ_f, ρ_g, ρ_f	13.26%	46.93%
12	Ducoulombier et al. [17]*	P_R, Bo, D_h, M, k_f, x_{tt}, Pr_f, q''_H, Re_f, Re_{fo}	11.71%	$C_{p,f}$, D_h, G, h_{fg}, M, k_f, P_R, q''_H, x, μ_f, μ_g, ρ_f, ρ_g	12.96%	52.00%
13	Oh and Son [9]	D_h, k_f, Pr_f, Re_f, x_{tt}	26.36%	$C_{p,f}$, D_h, G, h_{fg}, k_f, q''_H, μ_f, μ_g, ρ_f, ρ_g	21.84%	56.71%
14	Kim and Mudawar [3]*	Bo, k_f, P_F, P_H, P_R, Pr_f, Re_f, We_{fo}, x_{tt}, x, ρ_f, ρ_g	14.14%	$c_{p,f}$, D_h, G, k_f, h_{fg}, P_F, P_H, P_R, q''_H, x, σ, μ_f, μ_g, ρ_g, ρ_f,	12.45%	27.37%

* Correlations are based on the interpolation of two boiling modes.

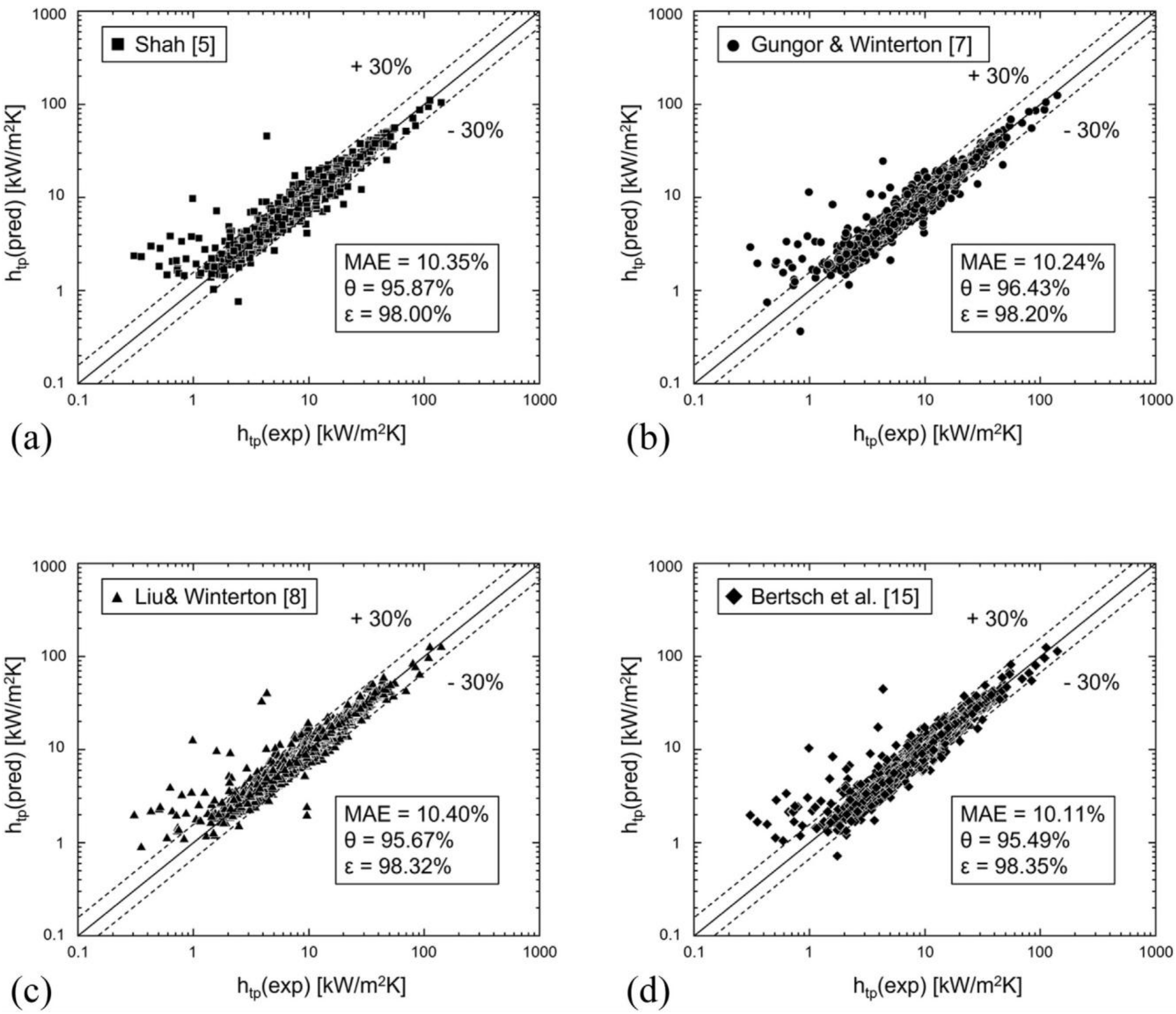

Fig. 1. Heat transfer coefficient as predicted by ANN with input set derived from previous correlations: (a) Shah [5] (b) Gungor and Winterton [7] (c) Liu and Winterton [8] (d) Bertsch et al. [15].

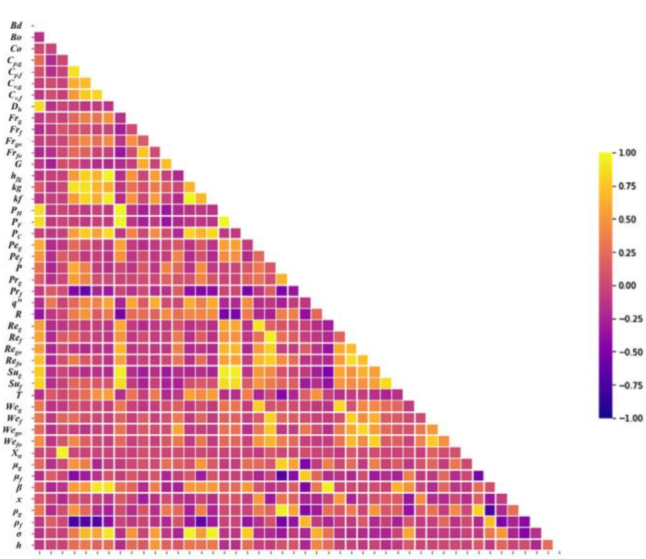

Fig. 2. Pearson correlation heatmap of the consolidated database.

variables have positive value and negatively correlated variables have negative value. Interestingly, the dimensionless parameters not included in most correlations viz. Fr_f, Fr_g, Fr_{fo} and Fr_{go} are correlated with the dependent variable, 'h', and only Fr_f is proposed by other correlations like Gungor and Winterton [7] and Liu and Winterton [8]. If we look at the Pearson's correlation coefficient between 'h' and Fr_{go}, it has a value of 0.516, higher than Fr_f (0.131). Thus, it becomes imperative to include all four of them into our model input parameter list. Fig. 3 shows the predictions when adding Fr_f, Fr_g, Fr_{fo} and Fr_{go} in the input parameters sets of the best four correlations including Shah [5] Gungor and Winterton [7], Liu and Winterton [8] and Bertsch et al. [15].

The second step of bivariate analysis is selecting parameters by Mutual information (MI). MI is a statistical measure between two random variables, x and y, quantifying the amount of information obtained about one random variable by knowing the other. MI and can capture non-linear relationships as well, unlike Pearson's coefficient and is invariant under transformations in the feature space that are invertible and differentiable [42]. The measure information is expressed as:

$$I(X;Y) = \int_X \int_Y p(x,y)\left(log\frac{p(x,y)}{p(x)p(y)}\right)dxdy \tag{14}$$

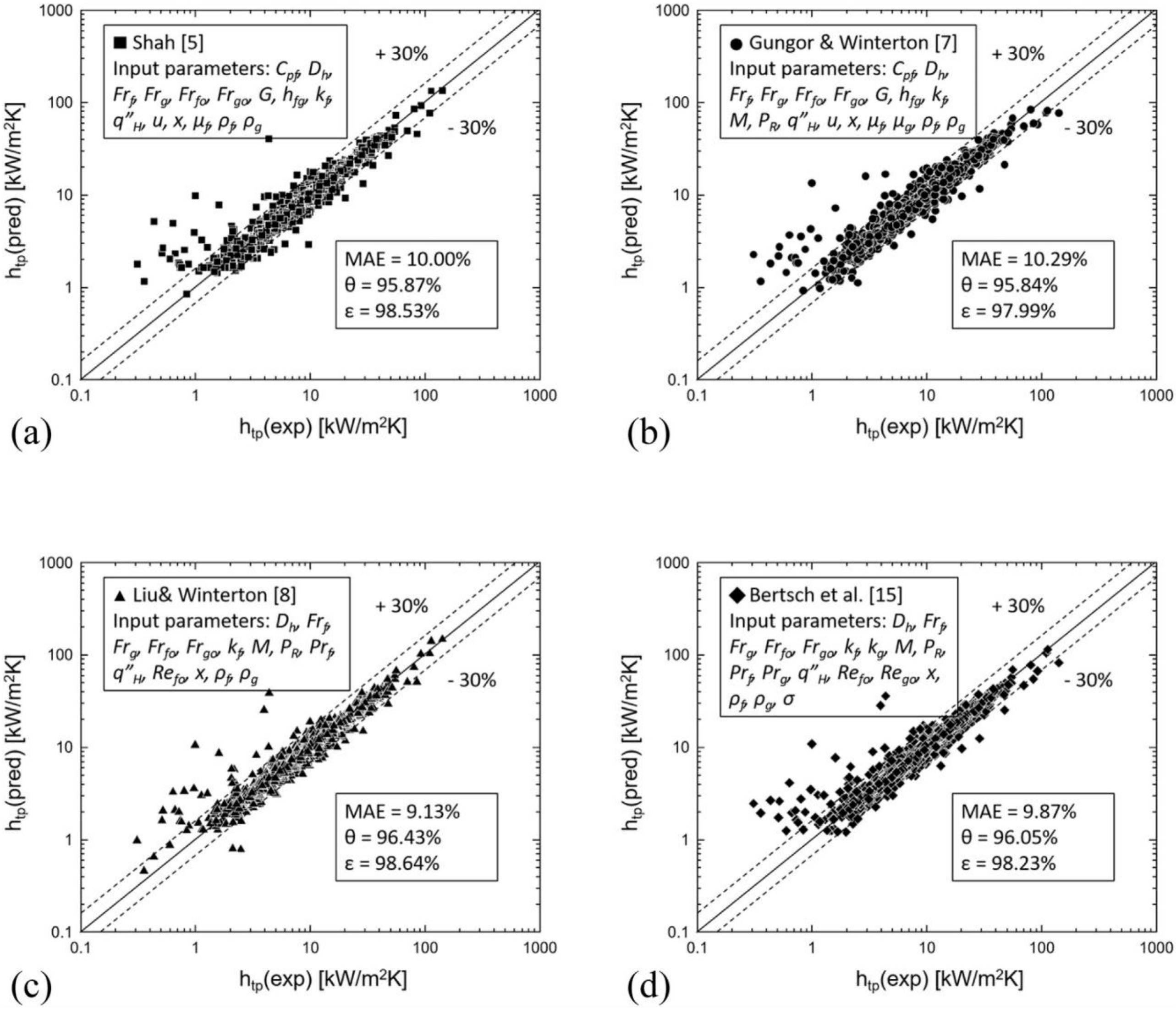

Fig. 3. Heat transfer coefficient as predicted by ANN with best four correlation input sets augmented by Pearson's correlation coefficient (a) Shah [5] (b) Gungor and Winterton [7] (c) Liu and Winterton [8] (d) Bertsch et al. [15].

Where *p(x)* and *p(y)* are the marginal probability density functions while *p(x,y)* is the joint probability density function of random variables x and y. In other words, mutual information determines the similarity between the joint distribution function to the product of the marginal distributions. In case the random variables, x and y, are completely independent (or unrelated) the integral of above equation will be zero. For interaction between more than two variables multi-information is an extension of MI, but not widely used due to difficulty in interpretation [42].

Fig. 4 shows the feature importance based on mutual information gain theory. The Su_g and Su_f are top predictors identified by MI for predicting heat transfer. Upon further augmenting the parameters based on this theory, following parameters are selected to be added to the input list –

Su_g, Su_f, Bd, Pr_f. The performance of the best four correlations adding this last subset of parameters is show in Fig. 5. Above discussion shows that ANN model performance can be further improved by adding parameters according to data science-based tools. A systematic approach which combines the correlation-based and data science based analysis can reduce the randomness associated with parameter selections techniques used for developing machine learning tools and improve the model performance.

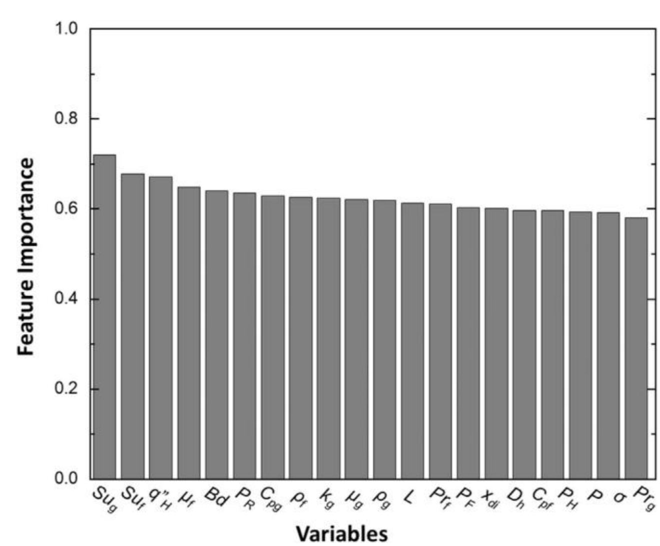

Fig. 4. Feature selection results obtained by mutual information (MI) feature selection method.

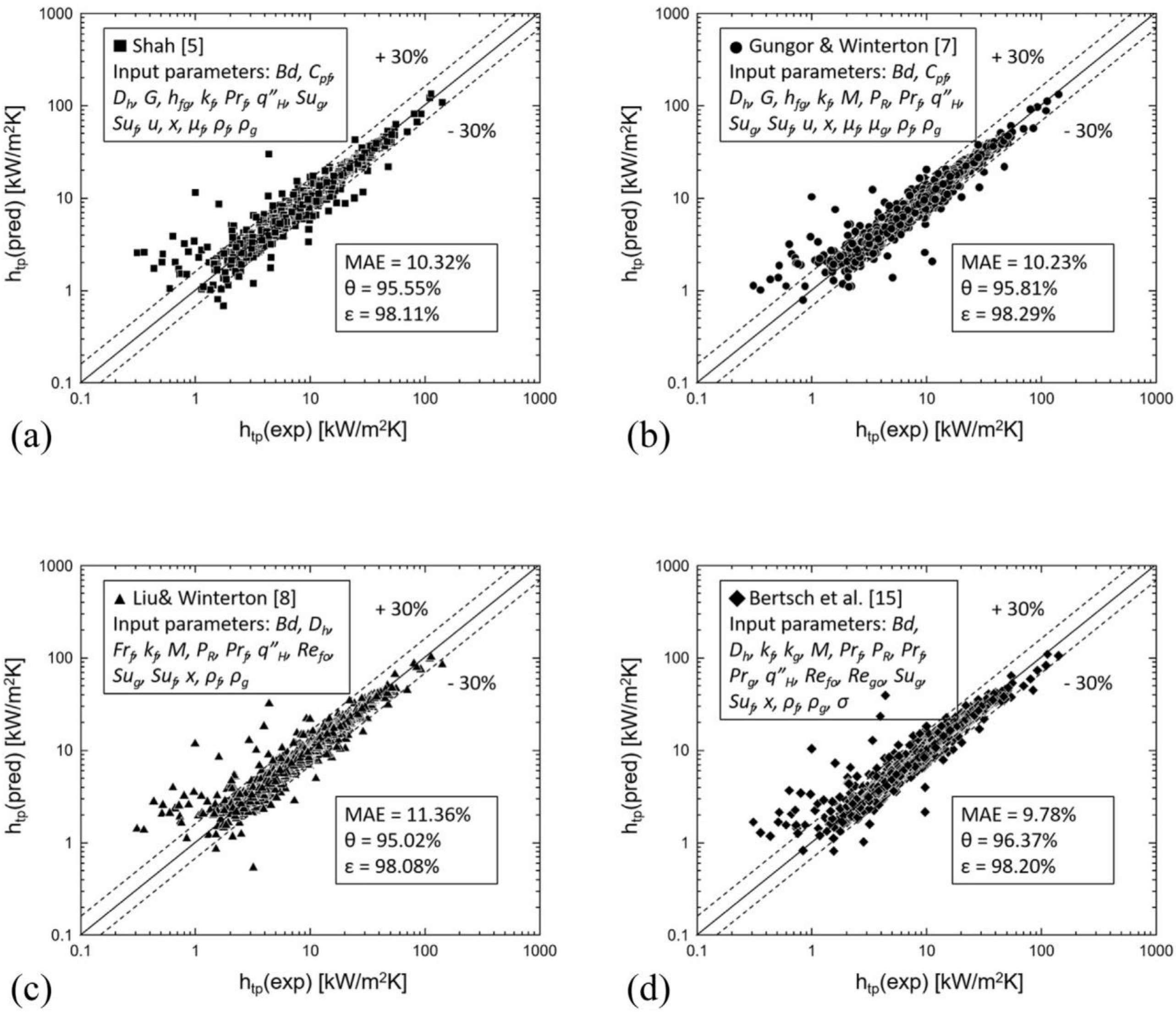

Fig. 5. Heat transfer coefficient as predicted by ANN with best four correlation input sets augmented by mutual information (a) Shah [5] (b) Gungor and Winterton [7] (c) Liu and Winterton [8] (d) Bertsch et al. [15].

3.4. Grid search

A final optimization to select the best performing list of input parameters is now necessary. Based on our previous reasoning of adding extra parameters to the correlations, we select four correlations, Shah [5], Liu and Winterton [8], and Gungor and Winterton [7], and Bertsch et al. [15] as the basis to perform a comprehensive grid search to add relevant data-driven parameters, and test all possible combinations of hidden layers and hyperparameters. To be able to select an optimal list of input parameters and the ANN hyper-parameters, the grid search investigated inputs over a range of hyper parameters summarized in Table 2. The search included running more than a million test cases based on the four correlation sets. It should be noted that if we did not downselect the initial sets using Section 3.2, we would have had to run number of tests that are a few orders higher than that. Table 5 shows the result of this analysis on the input sets based on the four correlations, with original input sets 'augmented' by the most impactful parameters identified by Pearson correlation coefficient analysis as well as Mutual information analysis. In all the four cases, the final MAE is close to 10%. All four correlations showed improvement in MAE when we added relevant parameters, however, the MAE did not reduce monotoncally with additionl parameters. Fig. 6(a) shows the performance of previous ANN model developed by Qiu et al. [27]. Fig. 6(b) shows the prediction of ANN model with input sets from Bertch's correlation. Here, the model was further optimized by adding data science informed paramters and the ANN hyperparameters are with a thorough grid search. Overall the best performed model obtained by grid search is shown in Fig. 6(c) with MAE of 8.48%. The final optimized input parameter set consists of C_{pf}, D_h, Fr_{go}, k_f, k_g, M, Pr_f, Pr_g, P_R, q''_H, Re_{fo}, Re_{go}, Su_g, x, ρ_f, ρ_g, σ, and is trained on the ANN model with network hidden layers (10, 20, 50, 100, 200, 400). The robustness of this ANN model architecture utilizing the selected input parameters and hyperparameters set is verified by investigating a large range of random state numbers, and the resulting uncertainty is shown to be low, only ±0.7%.

A clear framework for optimizing the neural network model parameters for predicting heat transfer coefficients was demonstrated in this study. This optimizing framework that combines correlations-based selection followed by data science-based selection can be applied not only to neural network models but other machine learning models. The results show success in predicting heat transfer coefficient, but it can be applied to other two-phase parameters including pressure drop and critical heat flux.

Table 5
'Augmented' input list optimization for selected flow boiling correlations.

Correlation	Original input set	Number of additional inputs	Added Input list	Number of Hidden Layers	Hyperparameters (Batch size, Epochs, Learn rate)	MAE%
Shah [5]	$C_{p,f}, D_h, G, h_{fg}, k_f,$ $q''_H, u, x, \mu_f, \rho_f, \rho_g$	0	/	6	32, 1000, 0.001	10.35
		1	Fr_{go}	7	64, 1000, 0.00075	10.03
		2	Fr_{go}, k_g	6	64, 1000, 0.001	9.92
		3	Fr_{go}, k_g, Su_g	7	64, 1000, 0.0005	10.48
		4	Fr_{go}, k_g, Su_g, Su_f	7	64, 1000, 0.00075	9.77
Gungor and Winterton [7]	$C_{p,f}, D_h, G, h_{fg}, k_f, M,$ $P_R, q''_H, u, x, \mu_f, \mu_g, \rho_f, \rho_g$	0	/	6	32, 1000, 0.001	10.24
		1	Fr_{go}	7	128, 1000, 0.00075	9.89
		2	Fr_{go}, k_g	7	128, 1000, 0.00075	10.62
		3	Fr_{go}, k_g, Su_g	6	32, 1000, 0.001	9.42
		4	Fr_{go}, k_g, Su_g, Su_f	7	32, 1000, 0.001	9.58
Liu and Winterton [8]	$D_h, Fr_f, k_f, M, P_R, Pr_f,$ $q''_H, Re_{fo}, x, \rho_f, \rho_g$	0	/	6	32, 1000,0.001	10.40
		1	Fr_{go}	7	16, 1000, 0.00075	9.79
		2	$Fr_{go}, C_{p,f}$	7	32, 1000, 0.001	9.80
		3	$Fr_{go}, C_{p,f}, Su_g$	7	32, 1000, 0.00075	9.81
		4	$Fr_{go}, C_{p,f}, Su_g, Su_f$	7	64, 1000, 0.001	9.89
Bertsch et al. [15]	$D_h, k_f, k_g, M, P_R, Pr_f, Pr_g,$ $q''_H, Re_{fo}, Re_{go}, x, \rho_f, \rho_g, \sigma$	0	/	6	32, 1000, 0.001	10.11
		1	Fr_{go}	7	16, 1000, 0.001	9.59
		2	$Fr_{go}, C_{p,f}$	7	16, 1000, 0.00075	9.97
		3	$Fr_{go}, C_{p,f}, Su_g$	6	32, 1000, 0.001	8.48
		4	$Fr_{go}, C_{p,f}, Su_g, Su_f$	7	32, 1000, 0.001	9.82

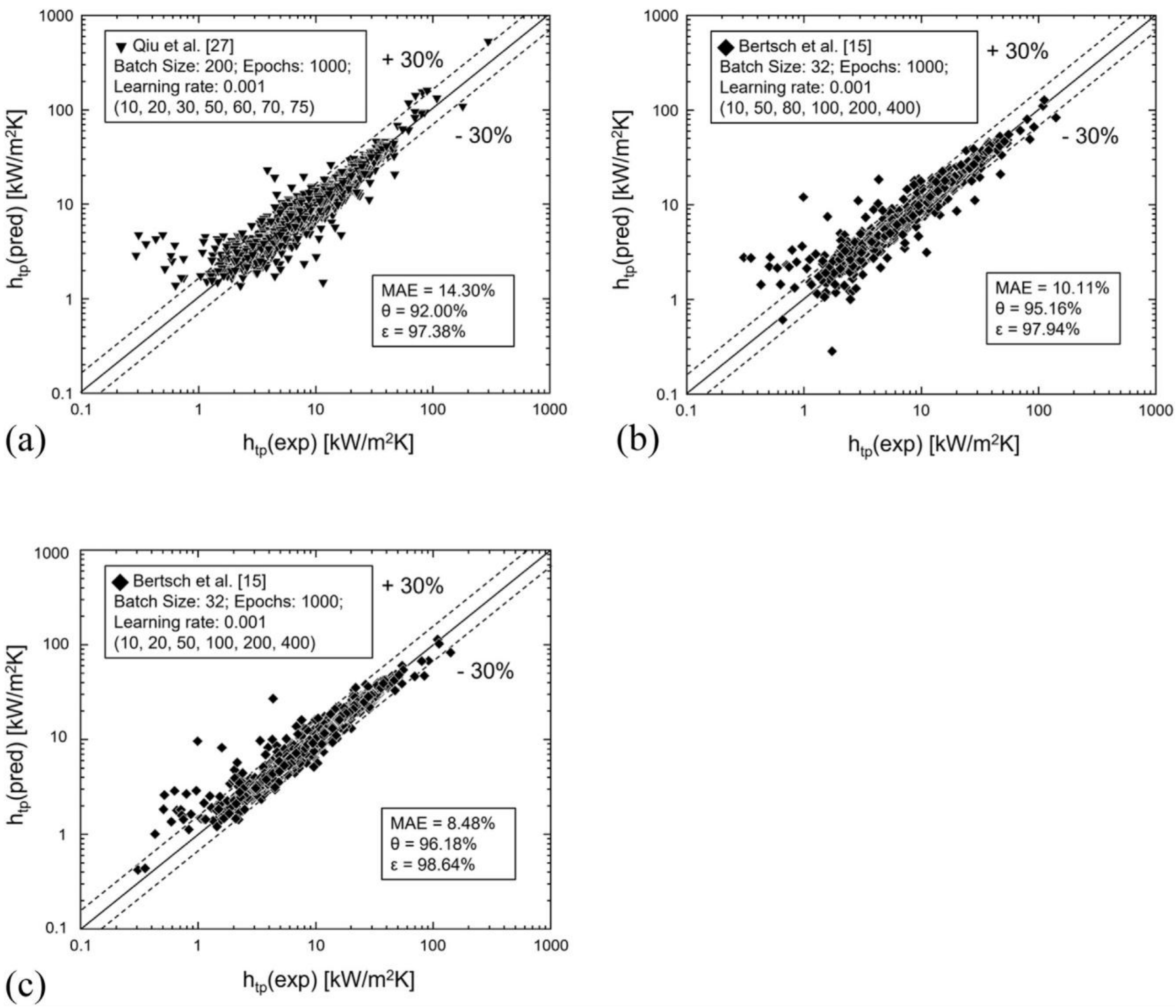

Fig. 6. Heat transfer coefficient predicted by (a) previous ANN model by Qiu et al. [27] (b) ANN model with input set derived from Bertsch et al. [15] (c) input parameter augmented by adding Fr_{go}, $C_{p,f}$, Su_g.

Table 6
Summary of database and optimized ANN model parameters.

Database	• 16,953 data points from 50 published sources • 16 Working fluids: CO_2, FC72, Propane, R11, R113, R123, R1234yf, R1234ze, R134a, R152a, R22, R236fa, R245fa, R32, R600a and water • Reduced pressures: 0.0046 < P_R < 0.77 • Hydraulic diameter: 0.15 mm < D_h < 6.5 mm • Mass velocity: 19 kg/m^{2}s < G < 1608 kg/m^{2}s • Liquid-only Reynolds number: 27 < $Re_{fo} = GD_h / \mu_f$ < 55,270 • Flow quality: 0 < x < 1
Optimized ANN model parameters	• Hidden layers (10, 20, 50, 100, 200, 400), • Regularization parameter (α): 0.01 • Batch size: 32 • Epochs: 1000 • Learning rate (λ): 0.001 • Inputs parameter set: C_{pf}, D_h, Fr_{go}, k_f, k_g, M, Pr_f, Pr_g, P_R, q''_H, Re_{fo}, Re_{go}, Su_g, x, ρ_f, ρ_g, σ.

A summary of the database employed to train the flow boiling mini/micro-channel heat transfer coefficient ANN model, and the final optimized input parameters and hyper-parameters are summarized in Table 6.

4. Conclusions

In this study, we verified the viability of systematically optimizing an Artificial Neural Network method for predicting flow boiling heat transfer in mini/micro-channel heatsinks. Utilizing machine learning, heat transfer coefficient is predicted on a test set by training an ANN model with a training set. The focus of this study is to find appropriate input parameters based on physics based and data-driven reasoning. Hyperparameters optimization was also performed on the ANN model using grid search technique to restrict MAE within 10%. Details of the key findings are as follows:

1. The consolidated database consisting of 16,953 data points for flow boiling heat transfer in mini/micro-channels was utilized for training the model [27]. A thorough optimization based on correlations from prior literature on flow boiling heat transfer, Pearson coefficient correlations mapping, and mutual information feature selection was conducted in this study.
2. The best-performing case with a reasonable number of input parameters is found to be one trained by the ANN model with (10, 20, 50, 100, 200, 400) hidden layers, 0.01 regularization parameter (α) and 0.001 learning rate (λ), with inputs parameter set consists of C_{pf}, D_h, Fr_{go}, k_f, k_g, M, Pr_f, Pr_g, P_R, q''_H, Re_{fo}, Re_{go}, Su_g, x, ρ_f, ρ_g, σ . This test case achieved an MAE of 8.48% with an uncertainty of ±0.7%. The ANN method outperforms universal correlations in predicting heat transfer coefficient for across multiple fluids and operating conditions.
3. A systematic approach which combines the correlation-based and data science based analysis is developed. It can reduce the randomness associated with parameter selections techniques used for developing machine learning tools and improve the model performance.
4. These results demonstrate great improvements in accuracy, and a potentially useful framework for optimizing machine learning models for predicting heat transfer coefficients that can be implemented on other two-phase flow configurations and parameters.

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

The data that has been used is confidential.

Acknowledgment

This project was partly supported by the Office of Naval Research (ONR), under Grant Number N00014-22-1-2618. The views and conclusions contained herein are those of the authors only and should not be interpreted as representing those of ONR, the U.S. Navy or the U.S. Government.

References

[1] I. Mudawar, Recent advances in high-flux, two-phase thermal management, J. Therm. Sci. Eng. Appl. 5 (2013), doi:10.1115/1.4023599.
[2] I. Mudawar, Two-phase microchannel heat sinks: theory, applications, and limitations, J. Electron. Packag. 133 (2011) 041002-1–041002-31.
[3] S.M. Kim, I. Mudawar, Universal approach to predicting saturated flow boiling heat transfer in mini/micro-channels – Part II. Two-phase heat transfer coefficient, Int. J. Heat Mass Transf. 64 (2013) 1239–1256, doi:10.1016/j.ijheatmasstransfer.2013.04.014.
[4] D. Garg, V.K. Dhir, A unified three-dimensional numerical model for boiling curve in a temperature controlled mode1, J. Heat Transf. 141 (2018), doi:10.1115/1.4041798.
[5] M.M. Shah, Chart correlation for saturated boiling heat transfer: equations and further study, ASHRAE Trans. 88 (1982) 185–196 ASHRAE.
[6] M.G. Cooper, Saturation nucleate pool boiling - a simple correlation, Inst. Chem. Eng. Symp. Ser. (1984) 785–793, doi:10.1016/B978-0-85295-175-0.50013-8.
[7] K.E.E. Gungor, R.H.S.H.S. Winterton, A general correlation for flow boiling in tubes and annuli, Int. J. Heat Mass Transf. 29 (1986) 351–358, doi:10.1016/0017-9310(86)90205-X.
[8] Z. Liu, R.H.S. Winterton, A general correlation for saturated and subcooled flow boiling in tubes and annuli, based on a nucleate pool boiling equation, Int. J. Heat Mass Transf. 34 (1991) 2759–2766, doi:10.1016/0017-9310(91)90234-6.
[9] H.K. Oh, C.H. Son, Evaporation flow pattern and heat transfer of R-22 and R-134a in small diameter tubes, Heat Mass Transf. 47 (2011) 703–717 Waerme-Und Stoffuebertragung, doi:10.1007/s00231-011-0761-4.
[10] G.M. Lazarek, S.H. Black, Evaporative heat transfer, pressure drop and critical heat flux in a small vertical tube with R-113, Int. J. Heat Mass Transf. 25 (1982) 945–960, doi:10.1016/0017-9310(82)90070-9.
[11] T.N. Tran, M.W. Wambsganss, D.M. France, Small circular- and rectangular-channel boiling with two refrigerants, Int. J. Multiphase Flow 22 (1996) 485–498, doi:10.1016/0301-9322(96)00002-X.
[12] G.R. Warrier, V.K. Dhir, L.A. Momoda, Heat transfer and pressure drop in narrow rectangular channels, Exp. Therm. Fluid Sci. 26 (2002) 53–64, doi:10.1016/S0894-1777(02)00107-3.
[13] W. Yu, D.M. France, M.W. Wambsganss, J.R. Hull, Two-phase pressure drop, boiling heat transfer, and critical heat flux to water in a small-diameter horizontal tube, Int. J. Multiphase Flow 28 (2002) 927–941, doi:10.1016/S0301-9322(02)00019-8.
[14] B. Agostini, A. Bontemps, Vertical flow boiling of refrigerant R134a in small channels, Int. J. Heat Fluid Flow 26 (2005) 296–306, doi:10.1016/J.IJHEATFLUIDFLOW.2004.08.003.
[15] S.S. Bertsch, E.A. Groll, S.V. Garimella, A composite heat transfer correlation for saturated flow boiling in small channels, Int. J. Heat Mass Transf. 52 (2009) 2110–2118, doi:10.1016/J.IJHEATMASSTRANSFER.2008.10.022.
[16] W. Li, Z. Wu, A general correlation for evaporative heat transfer in micro/mini-channels, Int. J. Heat Mass Transf. 53 (2010) 1778–1787, doi:10.1016/J.IJHEATMASSTRANSFER.2010.01.012.
[17] M. Ducoulombier, S. Colasson, J. Bonjour, P. Haberschill, Carbon dioxide flow boiling in a single microchannel – Part II: heat transfer, Exp. Therm. Fluid Sci. 35 (2011) 597–611, doi:10.1016/J.EXPTHERMFLUSCI.2010.11.014.
[18] J.C. Chen, Correlation for boiling heat transfer to saturated fluids in convective flow, Ind. Eng. Chem. Process Des. Dev. 5 (1966) 322–329, doi:10.1021/i260019a023.
[19] F.W. Dittus, L.M.K. Boelter, Heat transfer in automobile radiators of the tubular type, Int. Commun. Heat Mass Transf. 12 (1985) 3–22, doi:10.1016/0735-1933(85)90003-X.
[20] H.K. Forster, N. Zuber, Dynamics of vapor bubbles and boiling heat transfer, AIChE J. 1 (1955) 531–535, doi:10.1002/aic.690010425.

[21] L. Cheng, G. Ribatski, L. Wojtan, J.R. Thome, New flow boiling heat transfer model and flow pattern map for carbon dioxide evaporating inside horizontal tubes, Int. J. Heat Mass Transf. 49 (2006) 4082–4094, doi:10.1016/j.ijheatmasstransfer.2006.04.003.
[22] V.E. Schrock, L.M. Grossman, Forced convection boiling in tubes, Nucl. Sci. Eng. 12 (1962) 474–481, doi:10.13182/nse62-a26094.
[23] H. Hausen, Wärmeübertragung Im Gegenstrom, Gleichstrom und Kreuzstrom, Springer, Berlin Heidelberg, 1950, doi:10.1007/978-3-642-53135-4.
[24] S.S. Kutateladze, Boiling heat transfer, Int. J. Heat Mass Transf. 4 (1961) 31–45, doi:10.1016/0017-9310(61)90059-X.
[25] E. Hihara, C. Tanaka, Boiling heat transfer of carbon dioxide in horizontal tubes, Thermal Engineering Heat Transfer Summer Conference 3 (2007) 843–849, doi:10.1115/HT2007-32885.
[26] S.W. Churchill, R. Usagi, A general expression for the correlation of rates of transfer and other phenomena, AIChE J. 18 (1972) 1121–1128, doi:10.1002/aic.690180606.
[27] Y. Qiu, D. Garg, L. Zhou, C.R. Kharangate, S.M. Kim, I. Mudawar, An artificial neural network model to predict mini/micro-channels saturated flow boiling heat transfer coefficient based on universal consolidated data, Int. J. Heat Mass Transf. 149 (2020) 119211, doi:10.1016/J.IJHEATMASSTRANSFER.2019.119211.
[28] S. Walczak, N. Cerpa, Heuristic principles for the design of artificial neural networks, Inf. Softw. Technol. 41 (1999) 107–117, doi:10.1016/S0950-5849(98)00116-5.
[29] K. Goudarzi, A. Moosaei, M. Gharaati, Applying artificial neural networks (ANN) to the estimation of thermal contact conductance in the exhaust valve of internal combustion engine, Appl. Therm. Eng. 87 (2015) 688–697, doi:10.1016/j.applthermaleng.2015.05.060.
[30] L. Huang, C. Ma, Y. Li, J. Gao, M. Qi, Applying neural networks (NN) to the improvement of gasoline turbocharger heat transfer modeling, Appl. Therm. Eng. 141 (2018) 1080–1091, doi:10.1016/j.applthermaleng.2018.06.062.
[31] H. Peng, X. Ling, Neural networks analysis of thermal characteristics on plate-fin heat exchangers with limited experimental data, Appl. Therm. Eng. 29 (2009) 2251–2256, doi:10.1016/j.applthermaleng.2008.11.011.
[32] W. Chang, X. Chu, A.F. Binte Shaik Fareed, S. Pandey, J. Luo, B. Weigand, E. Laurien, Heat transfer prediction of supercritical water with artificial neural networks, Appl. Therm. Eng. 131 (2018) 815–824, doi:10.1016/j.applthermaleng.2017.12.063.
[33] C.K. Tan, J. Ward, S.J. Wilcox, R. Payne, Artificial neural network modelling of the thermal performance of a compact heat exchanger, Appl. Therm. Eng. 29 (2009) 3609–3617, doi:10.1016/j.applthermaleng.2009.06.017.
[34] F. Nie, H. Wang, Y. Zhao, Q. Song, S. Yan, M. Gong, A universal correlation for flow condensation heat transfer in horizontal tubes based on machine learning, Int. J. Therm. Sci. 184 (2023) 107994, doi:10.1016/J.IJTHERMALSCI.2022.107994.
[35] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 20 (2001) 1189–1232.
[36] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (1943) 115–133, doi:10.1007/BF02478259.
[37] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (1986) 533–536, doi:10.1038/323533a0.
[38] D. Svozil, V. Kvasnička, J. Pospíchal, Introduction to multi-layer feed-forward neural networks, in: Chemometrics and Intelligent Laboratory Systems, Elsevier, 1997, pp. 43–62, doi:10.1016/S0169-7439(97)00061-0.
[39] W. Gao, S. Kannan, S. Oh, P. Viswanath, Estimating mutual information for discrete-continuous mixtures, Adv. Neural Inf. Process. Syst. 30 (2017) 5988–5999.
[40] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, International Conference on Learning Representations (ICLR). (2015).
[41] P. Cerda, G. Varoquaux, B. Kégl, Similarity encoding for learning with dirty categorical variables, Mach. Learn. 107 (2018) 1477–1494.
[42] Y. He, C. Hu, B. Jiang, Z. Sun, J. Ma, H. Li, D. Tang, Data-driven approach to predict the flow boiling heat transfer coefficient of liquid hydrogen aviation fuel, Fuel 324 (2022) 124778, doi:10.1016/J.FUEL.2022.124778.
[43] G. Zhu, T. Wen, D. Zhang, Machine learning based approach for the prediction of flow boiling/condensation heat transfer performance in mini channels with serrated fins, Int. J. Heat Mass Transf. 166 (2021) 120783, doi:10.1016/J.IJHEATMASSTRANSFER.2020.120783.
[44] P. Domingos, A few useful things to know about machine learning, Commun. ACM 55 (2012) 78–87, doi:10.1145/2347736.2347755.
[45] J.R. Vergara, P.A. Estévez, A review of feature selection methods based on mutual information, Neural Comput Appl 24 (2014) 175–186.
[46] J.J. Berman, in: Understanding your data, Morgan Kaufmann, Boston, MA, USA, 2016, pp. 135–187.