

# Image driven deep learning based compact model to predict critical heat flux in direct immersion cooling via pool boiling

Pranay Nirapure, Ayushman Singh, Srikanth Rangarajan, Bahgat Sammakia  
Department of Mechanical Engineering (Binghamton University), Binghamton, U.S.A.  
srangar@binghamton.edu

**Abstract**—This paper introduces a concise model, based on convolutional neural networks (CNNs), to predict the performance of two-phase boiling heat transfer systems by leveraging data acquired through high-speed visualization techniques. The proposed approach utilizes high-speed digital imaging to meticulously capture and analyze bubble behavior on a heating component. Through this, a quantitative relationship is established between physical data and visual patterns, offering detailed insights into dynamic parameters such as bubble growth, departure, coalescence, and the static parameter of bubble diameter. The Convolutional Neural Network (CNN) is trained using visual patterns and physical data obtained from the visualization technique. Two distinct frameworks are employed for training: the first framework (A1) deploys Mask-R-CNN to extract physical data from images and utilizes a machine learning model trained on experimental data for interpolation. The second framework (A2) focuses on characteristic categorization, classifying test images into specific heat flux categories based solely on visual patterns derived from convolution operations. The study aims to quantify the relative weights assigned to the categorization and interpolation algorithms. Furthermore, this research endeavors to tackle a critical challenge encountered by AI models which is quantification of similarity in the training data to ensure optimal prediction accuracy.

**Index Terms**—Thermal management of power electronics, pool boiling, two-phase immersion cooling, Convolution Neural Network, Machine Learning

## NOMENCLATURE

$C_P$	Specific heat of liquid J/kg K
$h_{fg}$	Enthalpy of vaporization, J/kg
$C_{sf}$	Surface fluid factor
$\mu_f$	Dynamic viscosity of the liquid kg/m.s
$\alpha$	Surface tension-liquid-vapor interface (N/m)
$\rho_f$	Density of the liquid kg/m <sup>3</sup>
$\rho_g$	Density of the vapor kg/m <sup>3</sup>
$Pr_f$	Prandtl number of liquid
$\Delta T_{wall}$	Wall Superheat
$h$	Heat transfer coefficient(W/cm <sup>2</sup> K)
$q''$	Input heat flux (W/cm <sup>2</sup> )
$k_{cu}$	Thermal conductivity of copper

## I. INTRODUCTION

In recent decades, there has been an unprecedented surge in the demand for processing capabilities and data transfer, driven by the proliferation of applications such as artificial intelligence (AI) and machine learning (ML) which often

require significant computational power [1] - [2]. As more businesses and individuals rely on digital services, there is an increased demand for data storage, and computing density [3]. This growth is fueled by trends such as cloud computing, online streaming, e-commerce, and the increasing digitization of various industries. All these applications require the data centers to be extremely efficient both in terms of processing capabilities and in terms of using energy. But the operation of data centers or in general any electronic system is subjected to heat generation due to the inherent inefficiencies in electronic components. And their safe operation and performance depends heavily upon the operating temperature. All of the components dissipate substantial amount of thermodynamic low-grade energy in terms of heat. For safe and efficient operation of these components generated heat has to be removed from the system [1]. Hence, efficient heat dissipation

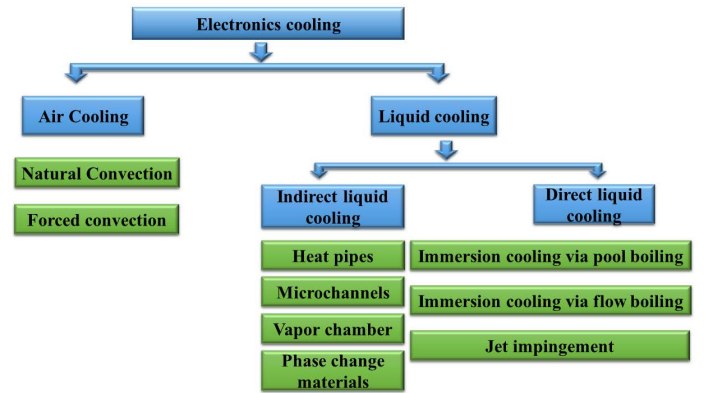


Fig. 1: Classification of cooling technologies based upon cooling medium

becomes crucial to prevent overheating, which can affect the performance and reliability of electronic components. Fig. 1 shows a classification of cooling techniques based upon cooling medium.

Figure 2 shows a comprehensive understanding of the capabilities of the traditional cooling methods in terms of range of heat flux they can handle. Natural convection of air is very low cost and highly reliable technique but is only limited to very low heat fluxes. Extension of this technique is forced convection air cooling which is very popular among modern data centers because of its reliability, it can achieve

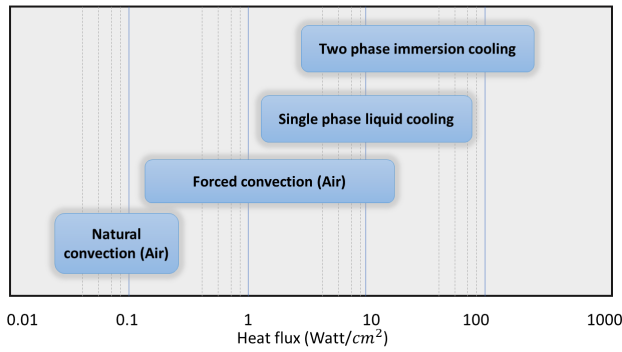


Fig. 2: Limits of cooling technologies

higher heat transfer coefficient than the natural convection and can reportedly meet the demands of upto  $37\text{W}/\text{cm}^2$  [4]. For the applications where higher heat transfer rate is required, single phase forced convection liquid cooling can be used. It facilitates with higher heat transfer coefficients due to sensible heating of fluid with higher heat capacity than air and can handle heat fluxes of about  $100\text{W}/\text{cm}^2$  [4]. It is not very common in the industry though because of the additional cost to cover the maintenance cost of pumps and other crucial infrastructure [5]. But with the increase in the demand of higher power density all the aforementioned cooling techniques fall short. Hence, it becomes crucial to look for more advanced cooling techniques. One such technique is two phase immersion cooling via pool boiling, as it involves the enthalpy of vaporization to convert the working fluid to the vapor. Enthalpy of vaporization is much higher than the sensible heat and this heat transfer also takes place at constant temperature, which makes it very efficient if seen from thermodynamic point of view. Hence, this technique provides highest heat transfer coefficient among all the mentioned techniques. This technique can reportedly handle very high heat fluxes of  $562\text{W}/\text{cm}^2$  for surface-fluid combination of water and vertically oriented GaN transistor coated with Parylene C and  $111\text{W}/\text{cm}^2$  for dielectric fluids [6]. S. Fan et al reported two phase immersion cooling systems with 72.2-79.3% higher COP values than the single-phase immersion cooling system in the server power range of 6.6-15.9 kW [4]

Recently, this technique has gained immense practical significance, especially in the electronics industry [12], driven by escalating demands for high-speed data processing and transfer. However, a prime challenge in harnessing it is, managing the critical heat flux (CHF), which represents its upper operational limit. Hence, accurately predicting CHF becomes paramount for real-time deployment and minimizing operational downtime [7]. Consequently, there has been a surge of interest in comprehending the underlying physics of boiling and a pressing need to develop models and techniques for predicting CHF [8] and optimizing the performance of two-phase boiling systems. Traditional measurement approaches often fall short, especially in the context of modern data centers and high-power density components with multiple hot-spots (Heterogeneous Integration). To delve deeper into the physics of boiling, a promising technique in use is high-speed

visualization of boiling phenomena, as it gives a direct insight into thermal as well as flow aspects of the problem. Certain researchers have been focusing on machine learning and deep learning to model two phase immersion cooling technique using various set of input data such as numerical data, boiling acoustics and visual data of boiling. Rokoni et al [16] uses high speed images as input and uses combination of supervised and unsupervised machine learning methods to produce future high speed images. Malakhov et al [17] uses high speed images as an input to the U-Net based CNN and Mask R-CNN to get nucleation site density, bubble departure diameters, frequency and bubble lifetime. Geraldo et al [18] uses non-boiling and boiling acoustic signals as input and predicts sodium boiling using Multiple linear regressions with classification model. A common research gap which has been identified from the studied literature was lack of test results for new unseen data to the model and analysis of the data going into the model as training data set. This work tries to address those problems to some extent.

In this study, authors propose a compact deep learning model (based upon convolutional neural networks) that leverages visual data obtained through high-speed visualization technique to forecast the performance of two-phase boiling heat transfer systems. This approach utilizes high-speed digital imaging to precisely study the behavior of bubbles on heating component, subsequently establishing a quantitative relationship between physical data extracted from images and visual patterns. The visualization technique offers a detailed insight into boiling phenomena, capturing dynamic parameters such as bubble growth, departure, and coalescence, as well as static parameter which is the bubble diameter. Visual patterns and physical data obtained from the visualization technique is employed to train a Convolutional Neural Network (CNN) through two distinct frameworks. The first framework employs Mask-R-CNN (A1) to extract physical data from the images and employs a machine learning model trained on experimental data for interpolation. The second framework deploys characteristic categorization (A2), classifying test images into specific heat flux categories solely based on visual features derived from convolution operations on the images. To predict the final heat flux, a weighted average of A1 and A2 is considered. Additionally, effect of relative weights assigned to the categorization and interpolation algorithms on prediction accuracy has also been studied. The prime reason of comparing the interpolation and the classification of the test images is based upon the proposed conjecture that range of heat flux is intertwined with the visual uncertainty which can be defined as lack of distinction in visual patterns within the high speed images of different heat fluxes. As it is discovered in later part of this work that degree of visual uncertainty associated with pair of heat fluxes can determine the validity of one of the two aforementioned frameworks in certain heat flux range. Hence, this work addresses a fundamental question that how visual uncertainty is related to heat flux, this has been studied in detail by the help of a parameter introduced in this paper called similarity index. If we can define the correlation of heat flux and visual uncertainty we can also deduce from there that which of the two algorithms can be used in a particular heat

flux regime.

## II. METHODOLOGY

### A. Experimental Setup

Fig.3 and Fig. 4 illustrates the schematic and photograph of experimental arrangement utilized for investigating boiling physics respectively. The setup comprises of a copper block equipped with a ceramic heater, facilitating heat supply through a compact die area measuring  $1\text{cm}^2$ . Heat is gradually applied to the copper block in increments of  $5\text{W}/\text{cm}^2$ , with a further increase of  $2\text{W}/\text{cm}^2$  during the initiation of bubble nucleation. The vaporized gas is subsequently condensed back into the pool by means of a copper condenser coil, with cold water or coolant circulating through a heat exchanger.

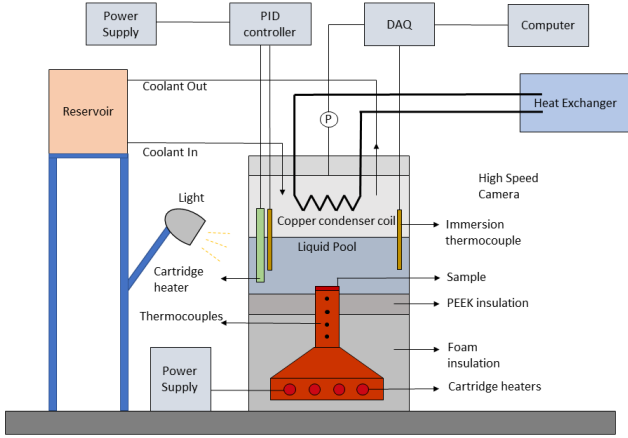


Fig. 3: Schematic of experimental setup

To capture the vapor bubble rise and departure, a Fastec IL-5 high-speed monochrome camera is used at 500 Frames per second.

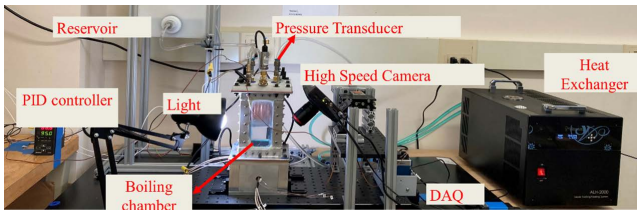


Fig. 4: Photograph of experimental setup

To study the saturated pool boiling, the liquid pool is maintained at the saturation temperature ( $100^\circ\text{C}$  at  $1\text{ atm}$  for water) using heaters immersed inside the liquid pool. The temperature of the pool is maintained by controlling the heaters with a PID controller. There are two thermocouples immersed in the pool as well to ensure the saturation condition inside the pool. The experiment is initiated by heating the liquid pool to reach a consistent saturation temperature for approximately 1 hour, thus ensuring uniform temperature distribution inside the liquid pool. Following this, the liquid

undergoes vigorous boiling for 30 minutes to degass non-condensable gases. Subsequently, heat is supplied from the copper block, and temperature data is systematically recorded at steady state for each heat flux value.

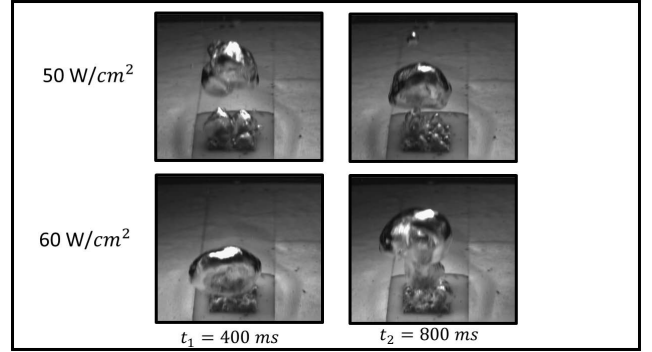


Fig. 5: High speed bubble departure images

Fig.5 shows the high speed images captured at different time instants for two different heat fluxes captured at steady state, it can be seen that even at the steady state, visual data for pool boiling at one time instant for a particular heat flux can be very distinct from data for the same heat flux for other time instants or it is also possible for visual data at particular time instant to be similar with data belonging to other heat flux at a different time instant. This brings us to the conclusion that the visual characteristics of boiling has to be studied carefully which has been done using a parameter called similarity index in sections further, which quantifies the visual uncertainty.

### B. Validation of Experimental data and data reduction

The saturated pool boiling with DI water experiment is performed for polished copper surface. The empirical correlations developed by Rohsenow [11] is used to validate the critical heat flux for the polished copper surface. The Rohsenow correlation is given as:

$$\frac{C_p \Delta T_{wall}}{h_{fg}} = C_{sf} \left[ \frac{q''}{\mu_f h_{fg}} \left[ \frac{\alpha}{g(\rho_f - \rho_g)} \right]^{\frac{1}{2}} \right]^{\frac{1}{3}} Pr_f$$

where,  $C_p$ : specific heat of liquid  $\text{J/kg K}$ ,  $h_{fg}$ : enthalpy of vaporization,  $\text{J/kg}$ ,  $C_{sf}$ : surface fluid factor (chosen as 0.0070 [21]),  $q''$ : nucleate pool boiling heat flux ( $\text{W}/\text{m}^2$ ),  $\mu_f$ : dynamic viscosity of the liquid  $\text{kg}/\text{m.s}$ ,  $\alpha$ : surface tension-liquid-vapor interface  $\text{N/m}$ ,  $\rho_f$ : density of the liquid  $\text{kg}/\text{m}^3$ ,  $\rho_g$ : density of the vapor  $\text{kg}/\text{m}^3$ ,  $Pr_f$ : Prandtl number of liquid. To be noted,  $C_{sf}$  is a parameter which takes into account the effect of surface characteristics and the interaction of that surface with the fluid and has to be chosen very carefully to accurately represent the characteristic interaction behavior which determines the critical heat flux.

Fig. 6 shows the locations of the thermocouples  $T_1$ ,  $T_2$  and  $T_3$  for the temperature measurement.  $T_w$  represents Wall temperature and  $T_{sat}$  represents saturation temperature of liquid. Heat flux is calculated as shown below:

$$q'' = -k_{cu} \frac{\Delta T^*}{2\Delta x}$$

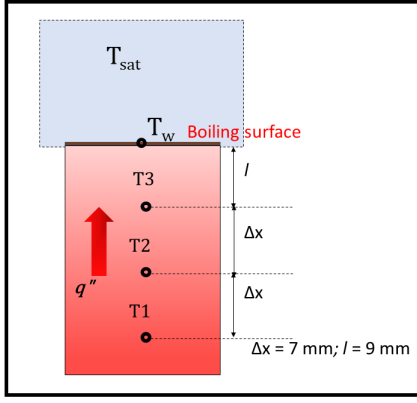


Fig. 6: Data reduction

$$\Delta T^* = 3T_3 - 4T_2 + T_1$$

$$T_w = T_3 - q'' \cdot \frac{l}{k_{cu}}$$

$$\Delta T_{wall} = T_w - T_{sat}$$

$$h = \frac{q''}{\Delta T_{wall}}$$

where,  $\Delta T_{wall}$ : Wall Superheat,  $h$ : Heat transfer coefficient ( $W/cm^2K$ ),  $q''$ : Input heat flux ( $W/cm^2$ ) and  $k_{cu}$ : thermal conductivity of copper.

### C. AI model architecture and workflow

In this study, Convolutional Neural Networks (CNNs) have been selected as the fundamental architectural framework to act as a surrogate model for the boiling physics investigated in the paper. CNNs stand out as an ideal choice for the analysis of visual data [9], primarily owing to their innate capacity to autonomously assimilate hierarchical features, especially for a complexity rich problem such as pool boiling. The ultimate deployment of our model is illustrated in the visual representation provided in Fig. 7. The evaluation procedure

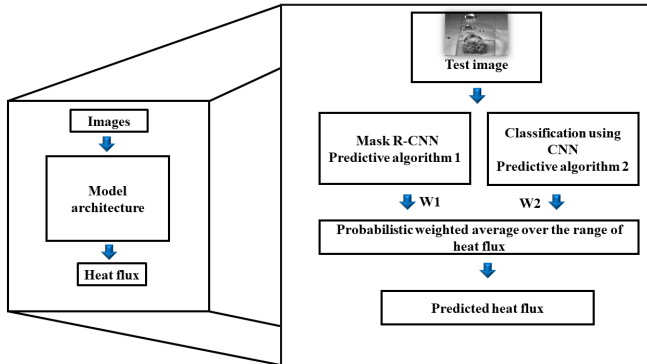


Fig. 7: AI architecture

involves presenting test images to the model, which subsequently undergoes through two distinct prediction algorithms. Singular weights are methodically assigned to each of these

algorithms and final prediction is calculated using a weighted average of A1 and A2 as:

$$[W_1 * H_1] + [(1 - W_1) * H_2]$$

where,  $W_i$  is weight given to corresponding heat flux  $H_i$  and  $W_1 + W_2 = 1$ .  $W_1$  signifies the weight assigned to algorithm A1 which by default assigns weight  $1 - W_1$  to algorithm A2.

Detailed work flow of individual algorithms are explained in subsections further. It is worth reiterating the fact that the time series of images corresponding to each steady state heat flux is used as a training data for the CNN models.

1) *Mask R-CNN with Machine Learning (A1)*: This algorithm can be comprehended in two parts:

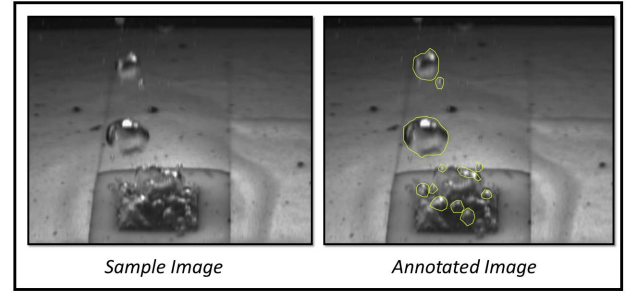


Fig. 8: Sample and Annotated Images

Training framework: The approach utilizes a combination of in-house image data capturing various bubble patterns and data from existing literature [14] to train a Mask region-based Convolutional Neural Network (MRCNN) [10]. This trained model is designed to identify and accurately measure the count and dimensions of bubbles occurring above the boiling surface, thereby containing information about operational heat flux of the heated surface. 30 distinct images have been manually annotated for the location and boundaries of the bubbles as shown in Fig.8, which acts as the guiding path for the supervised training of MRCNN model. These annotation

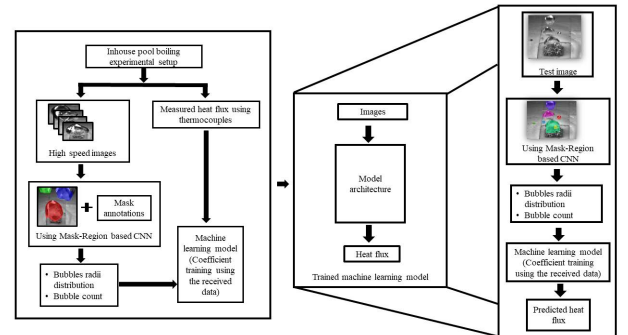


Fig. 9: Architecture of Algorithm 1

files are used along with weights pre-trained on MS COCO (Microsoft Common Objects in Context) [13] to extract the primitive features. Model have been trained for 25 epochs with aim of classifying the objects in the image under one of the two classes namely "Bubble" and "Background" to separate the bubbles from the background.



Once the Mask R-CNN model is ready (called as bubble detector from here on) it has been used to generate a data-set of bubble count and average bubble size corresponding to a heat flux. This data set with input as average bubble diameter and bubble count and output as measured heat flux is provided to linear regression-based machine learning model with two input variables of bubble size and bubble count. Coefficients are saved and are used statically in the deployment model.

Deployment framework: After preparing the Mask R-CNN model to function as a bubble detector and successfully training the machine learning model to correlate bubble radius and count with their respective heat flux values, the models are ready for implementation. Test image from the set up goes into the deployment model consisting of bubble detector which extracts the features like average bubble size and bubble count and passes it on to the trained machine learning model which then predicts the heat flux.

2) *CNN with interpolation (A2)* : The neural network architecture used comprises a total of nine layers, encompassing convolutional, batch normalization, max pooling, flatten, and dense layers. The initial layers involve two sets of Conv2D layers with 32 and 64 filters, each followed by batch normalization for stabilization and accelerated training. The choice of the number of filters in convolutional layers is a design decision which was decided based upon common observation from the similar object detection problems [19]. Early convolutional layers (with fewer filters, like 32) tend to learn low-level features such as edges, textures, and simple patterns. Deeper convolutional layers (with more filters, like 64) capture higher-level and more abstract features that result from combinations of low-level features. Additionally, Max pooling is strategically incorporated to reduce spatial dimensions after every two convolution layers. The subsequent flatten layer transforms the 2D feature maps into a one-dimensional vector. Fig. 10 shows architecture of algorithm A2.

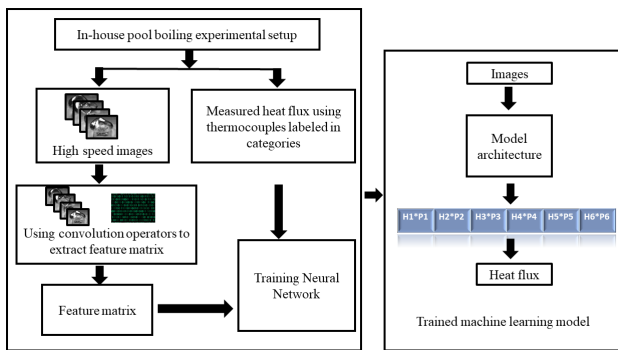


Fig. 10: Architecture of Algorithm 2

The deep learning prediction layers consist of a dense layer with 128 neurons (number of input neurons is a hyperparameter which can again be tuned, this particular value has been chosen based upon previous experience of authors and based upon observation from similar object detection problems) [?] and a final output layer with number of neurons based upon the number of classes we want to classify data into, signifying classification task which represent the number of heat fluxes

the model has been trained for. The choice of activation functions, including 'sigmoid' and 'softmax,' as well as the use of He uniform initialization for kernel weights, is aimed at enhancing the model's learning capabilities. Figure 11 shows the details of architecture of A2. Performance metrics used to

Layer type	Output shape	Kernel size	Activation
Input	(32, 32, 3)	-	-
Conv2D	(32, 32, 32)	(3, 3)	Sigmoid
128Batch Normalization	(32, 32, 32)	-	-
Conv2D	(32, 32, 32)	(3, 3)	Sigmoid
Batch Normalization	(32, 32, 32)	-	-
Max Pooling 2D	(16, 16, 32)	(2, 2)	-
Conv2D	(16, 16, 64)	(3, 3)	Sigmoid
Batch Normalization	(16, 16, 64)	-	-
Conv2D	(16, 16, 64)	(3, 3)	Sigmoid
Batch Normalization	(16, 16, 64)	-	-
Max Pooling 2D	(8, 8, 64)	(2, 2)	-
Flatten	4096	-	-
Dense	128	-	Sigmoid
Dense	3	-	Softmax

Fig. 11: Architecture details of A2

assesses the robustness of this model used is Accuracy which is plotted in Fig. 13. Training-validation loss is shown in Fig. 12, as it reaches close to 0 after 20 epochs, model can be regarded as consistent with the validation data.

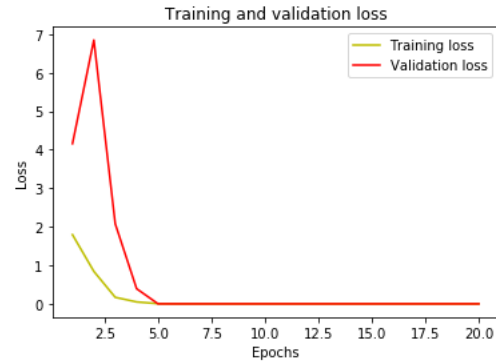


Fig. 12: Training and validation loss history

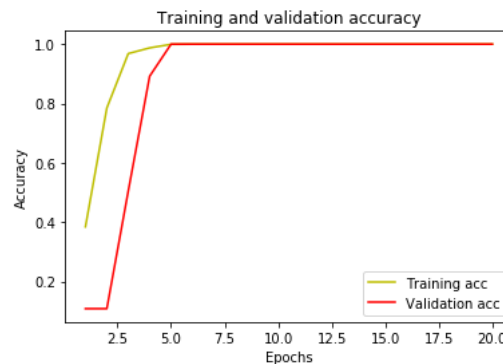


Fig. 13: Training and validation accuracy history

Fig. 14 shows the confusion matrix for data with eight classes of heat flux considered, this matrix is crucial in calculating the accuracy and other performance parameters for the CNN model, here columns represent the actual class and

rows represent the predicted class, each of the entries in the table represents the number of instances of actual class that were predicted as predicted class.

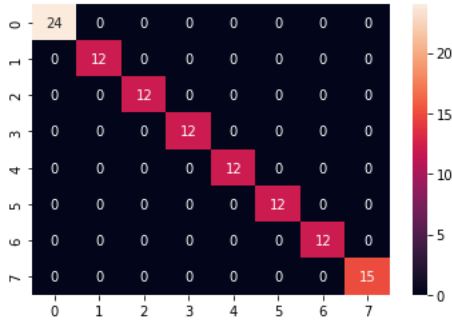


Fig. 14: Confusion Matrix

This model takes the probability of matching input image with a particular heat flux class and takes the average based upon the probabilities as shown in Fig.10 and gives an interpolated value of the heat flux rather than stating the class with maximum probability: Predicted heat flux using

$$A2 = \sum [P_i * H_i]$$

where,  $P_i$ : Probability of matching input image with a  $i_{th}$  heat flux class,  $H_i$ :  $i_{th}$  heat flux class

	A1	A2
Framework Used	Mask R-CNN	Features based CNN
Type of prediction	Interpolative	Classification
Type of feature used	Physical data extracted from images (bubble count and size)	Uses convolution operations to extract hierarchy of features
Computational Cost	Based upon training data	Based upon training data

Fig. 15: Characteristics of A1 and A2

Fig.15 shows the characteristics of two algorithms used in this work.

### III. RESULTS AND DISCUSSIONS

#### A. Experimental Results

Data is collected using the thermocouples and reduced as shown in Fig. 6. Apart from the numerical data high speed visualization has been used to acquire the visual data. Acquired numerical data has been used to calculate the heat flux. Furthermore, the calculation of heat flux was carried out by incorporating the measured surface superheat. The experimental data's validity was confirmed by Rohsenow Correlation and an additional experimental study from literature [15] (Wang et al) and illustrated in Fig. 16 through this verification process. The uncertainty in the thermocouples used to measure the temperature in the experiment was  $\pm 0.5$  °C.

The maximum uncertainty in the heat flux measurement was 7 %.

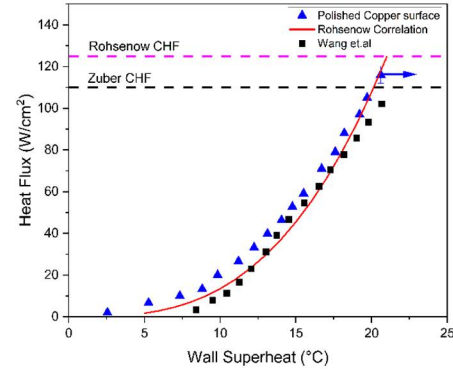


Fig. 16: Validation of Experimental data

This validation is crucial as this value of heat flux will go as a baseline truth in the AI model.

#### B. Similarity Analysis

1) *Similarity Index*: This section deals with two fundamental problems authors came across while defining the nature of the problem. First is to identify the problem of prediction as classification or interpolation which was briefly discussed at the end of section I. It becomes crucial to understand the correlation between visual uncertainty and range of heat flux as it will further decide the validity of a particular algorithm to be used in the specific heat flux range.

Second is very fundamental problem associated with the nature of classification models (A2 in this case). It is concerned with the data acquisition to train the model. In general for the supervised learning models performing classification task, the training data set has to be very carefully chosen so that it covers the maximum features of the target prediction class as well as it should not be excessive to introduce inaccuracy in prediction due to visually redundant data for multiple classes. To be noted, having redundant data may not necessarily harm the accuracy but it would increase the computational efforts at no further gain in accuracy improvement. In other words, the testing data should belong to the statistical distribution the training data follows. In this application, particularly there has to be an upper and lower limit of the gaps between the consecutive heat fluxes for which the model has to be trained. If the gap is too small the visual pattern will be similar for the two heat fluxes and it will be visually impossible to distinguish the two fluxes, hence there will be no further improvement in prediction accuracy with increase in training data. If the gap is too large then a substantial range of the visual pattern will be missing from the model, which can again impact accuracy of the classification model while deploying it, hence this study seeks an optimum gap for the training data of classification model A2 with the available data.

To quantify this visual uncertainty, a parameter called **similarity index** has been proposed in this work. It can

address both the aforementioned questions, as similarity index directly quantifies the visual uncertainty, value of similarity index in a heat flux zone can hint the validity of A1 or A2, and it also helps in deciding how far the heat flux classes should be for the training of the classification model. Similarity index is defined for a pair of heat fluxes and can be expressed as Percentage of image combination from two heat flux data sets having similarities more than 45% using a scale invariant feature transform (SIFT) detector. Similarity index is determined as shown in Fig.17. Similarity Index calculator

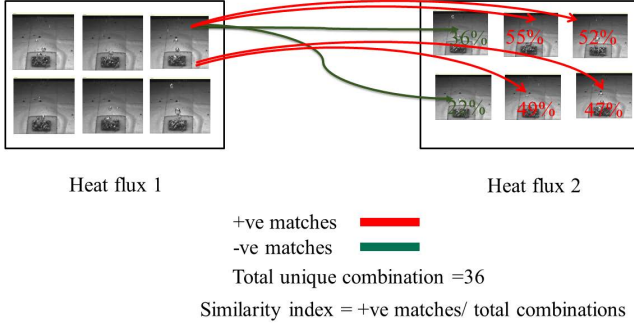


Fig. 17: Similarity Index

developed by the authors takes visual data acquired by the high speed camera and compares every possible combination of images in two heat flux data sets under consideration. A Visual similarity Score is calculated for combination of images using Scale Invariant Feature Transforms (SIFT) which is used to detect and match local features in an image, if the Visual Similarity Score is greater than 45% then it is considered as a positive match. Similarity Index can be expressed as percentage positive match out of total unique combination of images between two heat flux data sets under consideration.

$Q = 80 \text{ W/cm}^2$

Equally spaced time instants (ms)	0	200	400	600	800	1000	1200	1400
0	0.360	0.410	0.390	0.340	0.440	0.380	0.320	0.420
200	0.400	0.410	0.370	0.410	0.350	0.410	0.340	0.320
400	0.340	0.500	0.380	0.390	0.350	0.400	0.280	0.330
600	0.390	0.360	0.280	0.320	0.400	0.440	0.420	0.280
800	0.420	0.430	0.350	0.370	0.410	0.390	0.410	0.390
1000	0.410	0.410	0.340	0.410	0.340	0.400	0.360	0.440
1200	0.410	0.500	0.410	0.430	0.380	0.410	0.340	0.460
1400	0.350	0.420	0.390	0.360	0.360	0.380	0.390	0.380

$Q = 60 \text{ W/cm}^2$

Fig. 18: Similarity Score between individual images of two different heat fluxes

Fig.18 shows similarity scores for images captured for  $60 \text{ W/cm}^2$  and  $80 \text{ W/cm}^2$  at an interval of 200 ms for a particular test run. As threshold limit used for SIFT similarity score is 45%, similarity scores greater than 0.45 are indicated in red, low similarity scores are indicated in colors yellow,

$Q = 80 \text{ W/cm}^2$

Equally spaced time instants (ms)	0	200	400	600	800	1000	1200	1400
0	0.381	0.327	0.399	0.310	0.363	0.339	0.351	0.327
200	0.384	0.372	0.268	0.396	0.409	0.384	0.366	0.329
400	0.369	0.401	0.357	0.382	0.389	0.363	0.420	0.427
600	0.417	0.319	0.325	0.436	0.344	0.417	0.344	0.344
800	0.371	0.414	0.407	0.393	0.414	0.379	0.343	0.371
1000	0.395	0.437	0.462	0.454	0.336	0.395	0.412	0.445
1200	0.441	0.368	0.375	0.421	0.368	0.388	0.362	0.382
1400	0.368	0.368	0.364	0.346	0.307	0.342	0.364	0.382

$Q = 60 \text{ W/cm}^2$

Fig. 19: Similarity Score between individual images of two different heat fluxes for an independent test run

green and blue in decreasing order, this test run gives the similarity index of 6.25. Additionally, to confirm the repeatability and error in repeated experiments similarity index had been calculated for a different test run data, which gives the similarity index of 3.125 and similarity score matrix had been shown in figure 19. Hence, the similarity index is under 10 %. Moreover, this is just to stress upon the fact that similarity index is extremely low in higher heat flux regime than in lower heat flux regimes and similarity index had not been used in prediction models.

$Q_{av}$  is the mean of the two heat fluxes under consideration to find the similarity index. This similarity index has been calculated for consecutive heat fluxes from the collected data, effect of average heat flux of the considered heat fluxes ( $Q_{av}$ ) is plotted as shown in Fig.20 which shows that Similarity

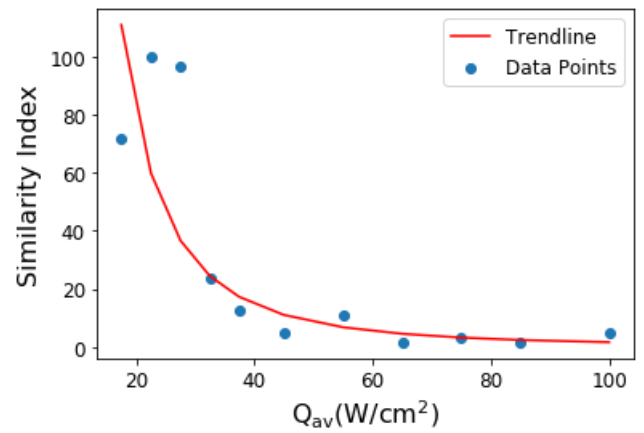


Fig. 20: Dependence of similarity on  $Q_{av}$

Index decreases with increase in heat flux, which shows that visual uncertainty decreases with increase in heat flux. This leads us to a very important conclusion that if visual uncertainty is lower at higher heat flux, then visual pattern will change quickly and drastically with small increase in the heat flux, hence if we use a classification algorithm (A2) which purely deals with visual features and not with physical data, then the data has to be acquired frequently and has to be closely spaced to achieve the desired accuracy which also

means the more effort to collect data and more computational effort to process that data to train the model.

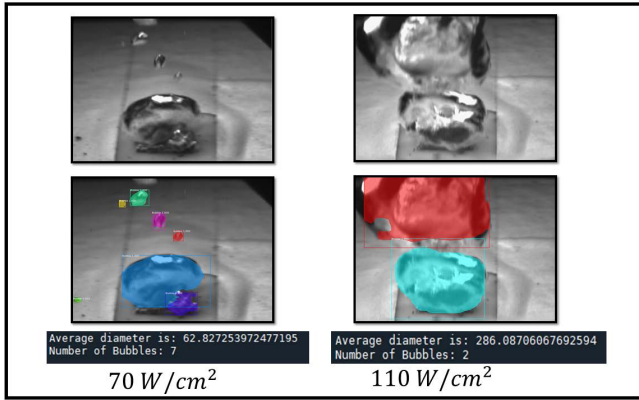


Fig. 21: Relation between bubble statics and heat flux at higher heat fluxes

Whereas, using an interpolation algorithm (A1) which uses physical data like bubble statics which is directly correlated to the physics of boiling may achieve better accuracy with relatively lesser computational effort, as shown in Fig.21 even with a gap of  $40W/cm^2$ , a trend (increase in average size of bubble and decrease in number of bubbles with increase in heat flux) can be traced between physical data and operating heat flux. Another question that arises is why interpolation

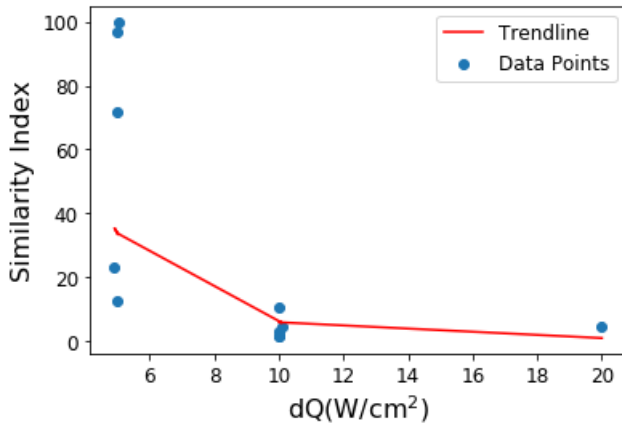


Fig. 22: Dependence of similarity on dQ

algorithm A1 cannot be utilized across the entire heat flux range. This can be attributed to the scarcity of the physical data that A1 is competent of extracting at lower heat fluxes. Observing Fig.23 it can be observed that most of the bubbles are sitting on the surface and did not depart, which makes it difficult for A1 to distinguish them and the physical data which is extracted is ambiguous in nature. Hence, A1 will be more accurate and beneficial computationally in higher heat flux zones while A2 will be more useful in lower heat flux zones, these results are explored in detail in subsection III-C

To address the second question posed at the start of this subsection, i.e. effect of similarity on (heat flux gap) dQ considered while training the classification model we can

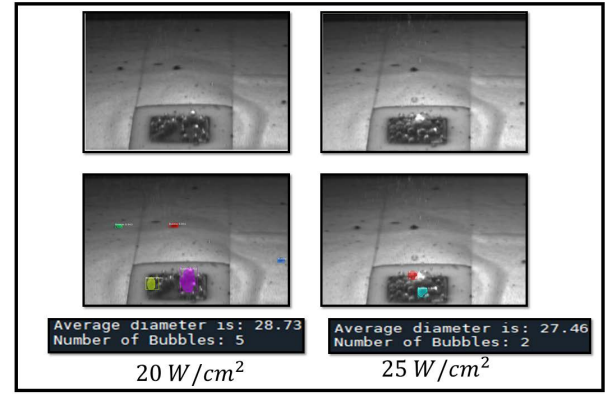


Fig. 23: Physical data at lower heat fluxes

observe Fig. 22, as the gap dQ increases from  $10W/cm^2$  to  $20W/cm^2$ , similarity index drops from 100% to 1.56% because of the variation in bubble pattern. With small dQ, heat flux patterns will exist in close vicinity and will be visually similar than for the higher dQ, hence we can observe higher similarity index in the lower range of dQ. Generally, considering a numerical model, increasing the grid division in computational domain increases the accuracy because physics is captured more meticulously in finer grid division than coarser grid in computational space. But here as we go for smaller dQ or closer heat flux data sets to train classification model within the visual uncertainty band, multiple heat fluxes may look similar and may not increase the prediction accuracy at all at a computational cost. This is illustrated in Fig.24, assuming a scenario where test image belongs to a heat flux between H1 and H2, there is a high chance that classification model may identify it as far as H4 due to high similarity within the visual uncertainty band. Visual uncertainty band can be comprehended as the range of heat flux for which the visual pattern of boiling is similar for all the heat fluxes within that range.

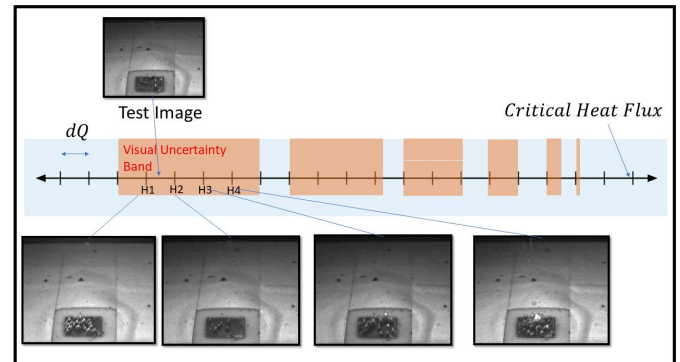


Fig. 24: Visual Uncertainty Band

Hence, this also leads us to a noteworthy result that to improve the accuracy in classification model(A2) dQ has to be carefully chosen which is studied in detailed in subsection III-B2.

2) *Effect of heat flux gap in training on the accuracy of prediction:* As this section deals with the gap in training data



of algorithm A2, all the prediction errors calculated are based on classification algorithm A2. To assess the initial influence of the heat flux gap in training classes on the accuracy of predictions, various gap scenarios were systematically tested. Exclusive sets of images, not included in the training and validation data sets, were employed to observe trends and validate our proposed methodology outlined in the preceding subsection III-B1. Observation from Fig. 25 and Fig. 26

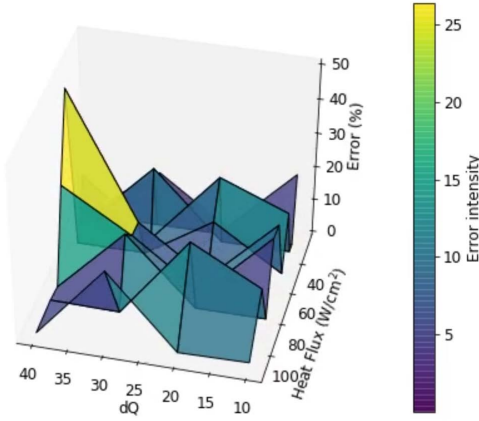


Fig. 25: Variation in prediction accuracy with Heat flux and dQ

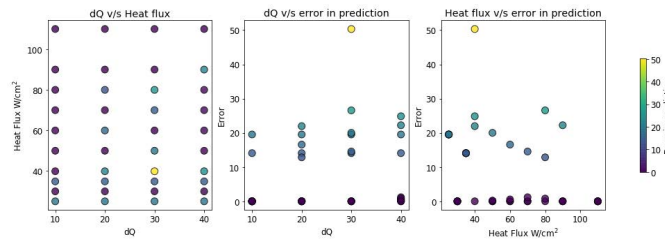


Fig. 26: Projections of Fig.25

reveals that Error drops with increase in heat flux which can be accounted by drop in similarity index, as model has been trained for constant dQ across the heat flux range in the single trial of testing dQ, error is expected to be declining.

Also careful observation of Fig. 25 reveals that error is increasing with increase in dQ which was expected because of inability of model to accommodate visual features in the training for the data not captured at higher dQ, for this work 4 different discrete values of dQ are considered (10, 20, 30 and 40) and captured data points were 25, 30, 35, 40, 50, 60, 70, 80 and 90, 110  $W/cm^2$ . Therefore, for certain data points covered in all of the dQ training, the error remains constant due to identical input data, even when there is a difference in the heat flux gap for that trial's training (trial is referred to as model analysis with a particular value of dQ). One critical data point which eludes all these exceptions is  $80W/cm^2$  as it was not included in the training of  $dQ=20, 30$  and  $40$ . For this case, prediction error first increases from 12% at  $dQ = 40W/cm^2$  to 26% at  $dQ = 30W/cm^2$  but then again drops to 12.8% at  $dQ = 20W/cm^2$ . It can be explained by the fact that training data occurred to consist the heat flux classes closer

to target class in both the trials with  $dQ = 20W/cm^2$  and  $dQ = 40W/cm^2$  (which lowered down the prediction error) which also generalizes the fact that availability of the data class during training in the vicinity of test images dictates the prediction accuracy, which is the gist of the argument that dQ dictates prediction accuracy. If dQ increases, the chances that the test image will lie in the vicinity of training classes will be low and prediction in the accuracy will be low, if the dQ is low and within the visual uncertainty band then prediction accuracy will be low.

To observe more general trend and correlation between dQ and prediction accuracy more data needs to be collected, model needs to be trained for smaller dQ and multiple intermediate points need to be tested which are at same vicinity for all the trails of dQ data set.

This observation indicates the existence of a correlation between the similarity index in data acquisition and the subsequent accuracy of prediction. Furthermore, the findings suggest that maintaining a constant dQ may not be justified, prompting the exploration of a multi-variable optimization approach to minimize prediction errors, by using a variable dQ based upon the visual pattern. The pursuit of a more nuanced optimization strategy serves as a crucial avenue for future research, emphasizing the need for a refined cost function. This prospective aims to further improve the precision of predictions.

3) *Effect of similarity index on prediction accuracy:* As previously mentioned, the selection of heat flux gaps during the training of the AI model is guided by the similarity index. This section explores how the similarity index directly influences prediction accuracy. Scenario with high similarity index has been examined, and the model has been trained with the corresponding gap for similarity index. Subsequently, similar heat fluxes are tested on the model to mitigate the impact of operating heat flux on the similarity index. In this specific case, a heat flux gap of  $dQ = 20W/cm^2$  was utilized for training the model, and two heat fluxes,  $25W/cm^2$  and  $30W/cm^2$ , were considered. A similarity index of 96.21% was determined using a similarity index calculator, which is the highest among all the observed folders. However, during prediction, A2 wrongly classified a test image belonging to  $25W/cm^2$  as  $30W/cm^2$ . This can be explained directly from the Fig. 26 which shows high similarity index at these heat fluxes, also it can be concluded from here that heat fluxes  $25W/cm^2$  and  $30W/cm^2$  lie in the visual uncertainty band.

### C. Prediction Accuracy and validity of A1 and A2

Building upon the earlier discussion of the advantages of algorithms A1 and A2 with respect to heat flux range, this section focuses on the performances of both the algorithms individually in predicting heat flux throughout that range.

Parity Plot in figure 27 illustrates the predictions of heat flux using A1. To be noted, this training is at heat flux gap of  $30W/cm^2$  in the machine learning model. At lower heat fluxes, as proposed in section III-B1 A1 gives top 20 % over-prediction points and is quite inaccurate as reason being physical data which A1 is capable of extracting is scarcely

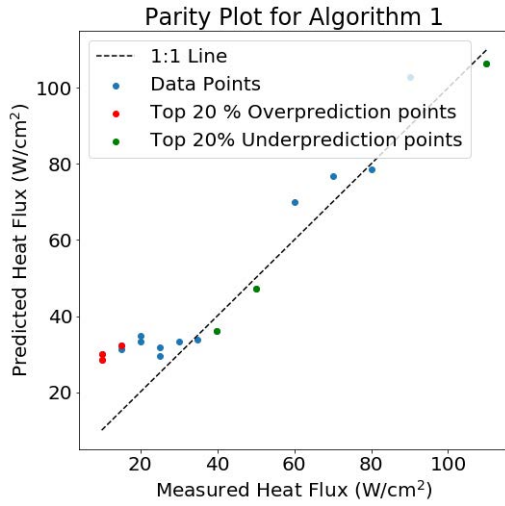


Fig. 27: Parity plot of algorithm 1 in entire heat flux range

available at lower heat fluxes which can be observed from Fig. 23. At higher heat fluxes even with the training gap of  $30W/cm^2$  this model predicts better than algorithm A2.

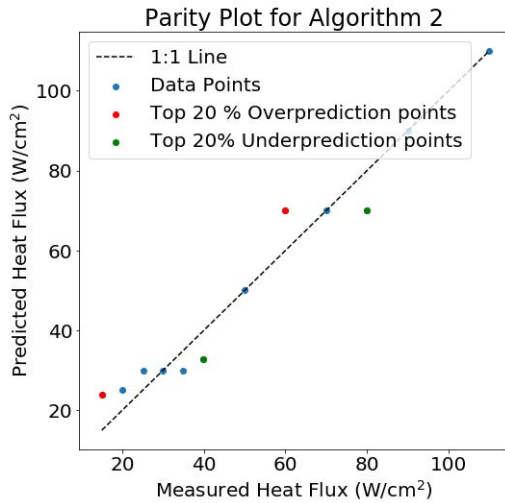


Fig. 28: Parity plot of algorithm 2 in entire heat flux range

Examining the Parity Plot in Fig. 28 for algorithm A2 reveals a noteworthy pattern in its predictive performance. This algorithm demonstrates accuracy better than A1, with average error of 22.6%, maximum error of 59.25 % and minimum error of 0 % in lower heat flux regimes ( $15W/cm^2$  to  $40W/cm^2$ ) as compared to average error of 39 %, maximum error of 108 % and minimum error of 3 % by A1. However, at higher flux regimes ( $40W/cm^2$  to  $80W/cm^2$ ) A2 gives an average error of 9.4% with maximum error of 17.71 % and minimum error of 0 % as compared to 8.6% of A1, with maximum error of 16.83 % and minimum error of 5.7 %. Which can be again explained by the proposed idea in subsection III-B1. Reason for limiting the higher heat flux range upto  $80W/cm^2$  is due to the absence of available data to test. This is because heat flux values beyond  $80W/cm^2$  were included in the training of the classification model, rendering it ineffective to use it as

test data and assess the robustness and reliability of A2 for those specific heat flux levels.

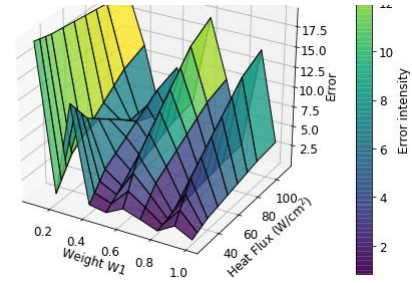


Fig. 29: Effect of individual algorithm on prediction accuracy with respect to heat flux

It can be observed from Fig. 29 that overall prediction error reduces over the heat flux which indicates prediction accuracy also depends upon heat flux range. Consequently, A2 appears to be particularly relevant in lower heat flux range if limited training data is used, whereas, A1 can be beneficial in mid to high flux ranges where physical data in terms of bubble statics is accurate representation of boiling physics.

Proposed methodology is the demonstration of a technique to utilize the visual data available to estimate the heat flux instead of using the methods requiring probes to derive the heat flux. Effect of changing the boiling surface area, surface and fluid characteristics may change the visual characteristics of boiling and has to be explored explicitly to draw conclusion about the generalization of the methodology. Fundamental motivation behind the work is that there is visual distinction between the different heat flux regimes which this work harnesses using the computer vision and deep learning.

#### D. Conclusions

- Visual Uncertainty has been quantified using the parameter called Similarity index, and it is established that uncertainty in the visual data of boiling is a strong function of heat flux range at which uncertainty is being explored, declining from 96.8% at  $25W/cm^2$  to 4% at  $100W/cm^2$ . This reiterates the fact that physical data based CNN model is an ideal tool to model boiling closer to the critical heat flux.
- Interpolation Algorithm A1, which utilizes physical data of boiling has been proved to be more accurate in higher heat flux range with an average error of 8.6 % with maximum error of 16.83 % and minimum error of 5.7 %, as compared to 9.4 % with maximum error of 17.71 % and minimum error of 0 % by A2. This conclusion can be accounted by the observation that A1 uses physical features extracted from the images which are more prominent and distinguishable at higher heat fluxes.
- Classification Algorithm A2, which utilizes visual features extracted using convolution operators on images has been proved to be better than A1 for lower heat flux ranges with an average error of 22% with maximum error of 59.25 % and minimum error of 0% as compared to

average error of 39% with maximum error of 108% and minimum error of 3% obtained while using A1.

- To assess the effect of similarity index on the prediction accuracy, a case study has been conducted for high similarity index of 96.21% for heat fluxes  $25W/cm^2$  and  $30W/cm^2$ , at training gap of  $dQ = 20W/cm^2$ . A2 predicted test image belonging to  $25W/cm^2$  as  $30W/cm^2$ .
- Effect of frequency of data collection (dQ) while training the classification model A2 has been studied and it is established that dQ dictates prediction accuracy. This analysis also address one of the research gaps and ensures that for all trial of dQ greater than  $10W/cm^2$  intermediate heat fluxes tested were unseen data for the model and the model is capable of predicting them.
- Study suggests that maintaining a constant dQ may not be justified and exploration of a multi-variable optimization approach to minimize prediction errors using a variable dQ based upon the visual pattern is required.
- In conclusion, this study successfully demonstrated the development of a compact model for pool boiling, leveraging experimental data-driven convolutional neural networks (CNN). This achievement establishes a promising framework that can be valuable for future investigations into pool boiling, particularly in the context of advancing next-generation electronics systems.

#### ACKNOWLEDGEMENT

Authors would like to thank Semiconductor Research Corporation (SRC Grant number 2878.027) and SUNY Research Foundation for their support.

#### REFERENCES

- [1] Ali Habibi Khalaj, Saman K. Halgamuge, A Review on efficient thermal management of air- and liquid-cooled data centers: From chip to the cooling system, *Applied Energy*, Volume 205, 2017, Pages 1165-1188, ISSN 0306-2619, <https://doi.org/10.1016/j.apenergy.2017.08.037>.
- [2] Yogendra Joshi, Pramod Kumar, *Energy Efficient Thermal Management of Data Centers*, Springer New York, NY, <https://doi.org/10.1007/978-1-4419-7124-1>
- [3] Rambo, J., Joshi, Y. Modeling of data center airflow and heat transfer: State of the art and future trends. *Distrib Parallel Databases* 21, 193–225 (2007). <https://doi.org/10.1007/s10619-006-7007-3>
- [4] Baris Burak Kanbur, Chenlong Wu, Simiao Fan, Fei Duan, System-level experimental investigations of the direct immersion cooling data center units with thermodynamic and thermoeconomic assessments, *Energy*, Volume 217, 2021, 119373, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2020.119373>.
- [5] Zhihao Zhang, Xuehui Wang, Yuying Yan, A review of the state-of-the-art in electronic cooling, *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, Volume 1, 2021, 100009, ISSN 2772-6711, <https://doi.org/10.1016/j.prime.2021.100009>.
- [6] Patrick Birbarah, Tarek Gebrael, Thomas Foulkes, Andrew Stillwell, Alexandra Moore, Robert Pilawa-Podgurski, Nenad Miljkovic, Water immersion cooling of high power density electronics, *International Journal of Heat and Mass Transfer*, Volume 147, 2020, 118918, ISSN 0017-9310, <https://doi.org/10.1016/j.ijheatmasstransfer.2019.118918>.
- [7] <https://resources.system-analysis.cadence.com/blog/msa2023-critical-heat-flux>
- [8] Zhang, L., Wang, C., Su, G. et al. A unifying criterion of the boiling crisis. *Nat Commun* 14, 2321 (2023). <https://doi.org/10.1038/s41467-023-37899-7>
- [9] Jee Hyun Seong, Madhumitha Ravichandran, Guanyu Su, Bren Phillips, Matteo Bucci, Automated bubble analysis of high-speed subcooled flow boiling images using U-net transfer learning and global optical flow, *International Journal of Multiphase Flow*, Volume 159, 2023, 104336, ISSN 0301-9322, <https://doi.org/10.1016/j.ijmultiphaseflow.2022.104336>.
- [10] He, Kaiming, et al. "Mask r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [11] Rohsenow, W. M. (July 29, 2022). "A Method of Correlating Heat-Transfer Data for Surface Boiling of Liquids." *ASME. Trans. ASME*. August 1952; 74(6): 969–975. <https://doi.org/10.1115/1.4015984>
- [12] <https://news.microsoft.com/source/features/innovation/datacenter-liquid-cooling/>
- [13] <https://cocodataset.org/home>
- [14] Seongchul Jun, Jinsub Kim, Donggun Son, Hwan Yeol Kim, Seung M. You, Enhancement of Pool Boiling Heat Transfer in Water Using Sintered Copper Microporous Coatings, *Nuclear Engineering and Technology*, Volume 48, Issue 4, 2016, Pages 932-940, ISSN 1738-5733, <https://doi.org/10.1016/j.net.2016.02.018>.
- [15] Xiaomeng Wang, Dani Fadda, Juan Godínez, Jungho Lee, Seung M. You, Effect of wettability on pool boiling heat transfer with copper microporous coated surface, *International Journal of Heat and Mass Transfer*, Volume 194, 2022, 123059, ISSN 0017-9310, <https://doi.org/10.1016/j.ijheatmasstransfer.2022.123059>.
- [16] A. Rokoni, L. Zhang, T. Soori, H. Hu, T. Wu, and Y. Sun, "Learning new physical descriptors from reduced-order analysis of bubble dynamics in boiling heat transfer," *Int. J. Heat Mass Transf.*, vol. 186, pp. 122501, 2022. DOI: 10.1016/j.ijheatmasstransfer.2021.122501.
- [17] . Malakhov, A. Seredkin, A. Chernyavskiy, V. Serdyukov, R. Mullyadzanov, and A. Surtaev, "Deep learning segmentation to analyze bubble dynamics and heat transfer during boiling at various pressures," *Int. J. Multiphase Flow*, vol. 162, no. January, pp. 104402, 2023. DOI: 10.1016/j.ijmultiphaseflow.2023.104402.
- [18] I. C. Geraldo, T. Bose, K. M. Pekpe, J. P. Cassar, A. R. Mohanty, and K. Paumel, "Acoustic monitoring of sodium boiling in a liquid metal fast breeder reactor from autoregressive models," *Nucl. Eng. Des.*, vol. 278, pp. 573–585, 2014. DOI: 10.1016/j.nucengdes.2014.07.026
- [19] Bhattiprolu, S. (2023). python for microscopists.158b - Transfer learning using CNN (VGG16) as feature extractor and Random Forest classifier
- [20] Bhattiprolu, S. (2021). python for microscopists. GitHub <https://github.com/bnsreenu/python-for-microscopists/blob/master/158b-transfer-learning-using-CNN-weights-VGG16-RF.py>
- [21] Jabardo, José Maria et al, Evaluation of the Rohsenow correlation through experimental pool boiling of halocarbon refrigerants on cylindrical surfaces. *Journal of The Brazilian Society of Mechanical Sciences and Engineering - J BRAZ SOC MECH SCI ENG*. 26. 10.1590/S1678-58782004000200015.