Application of Deep Learning Convolutional Neural Network for Spray Characterization

Amin Heyrani Nobari[1], Farzaam Khorasani-Gerdehkouhi[1], Naib Gulam[1], Nasser Ashgriz[1]
[1]University Of Toronto Department of Mechanical and Industrial Engineering
5 King's College Rd, Toronto, Canada
amin.heyraninobari@mail.utoronto.ca; farzam.khorasani.gerdehkouhi@mail.utoronto.ca;
naib.gulam@mail.utoronto.ca; ashgriz@mie.utoronto.ca

Abstract – The application of a machine learning artificial intelligence (AI) for spray characterization is investigated. Images of sprays in regions where droplet formation has not taken place, therefore not allowing any insight into the spray droplets themselves. The objective is to bridge the gap from the process of droplet formation to the characteristics of the spray that is produced at the end of this process. To achieve this, convolutional neural networks (CNN) are trained to classify images of sprays that were captured at different operating fluid pressures. Even though this is not directly characterizing the spray, it provides evidence for the potential of machine learning methods in spray characterization as distinctions are made prior to spray formation meaning CNNs are able to distinguish patterns in sprays prior to the droplet formation process, hence proving the possibility of bridging the aforementioned gap. Our models were able to accurately identify images of sprays taken at different operating pressures. Moreover, the convolutional neural networks were further analysed to understand how they were able to make these distinctions, that are not easily visible to the human eye. For this gradient class activation maps were determined to understand the inner workings of the convolutional neural networks. These gradient class activation mappings could prove useful in determining new physical patterns that were previously unknown, which could contribute to a better understanding of sprays and the droplet formation process.

Keywords: Sprays, Atomization, Droplet Size Distribution, Artificial Intelligence, Convolutional Neural Networks, Machine Learning

1. Introduction

Sprays have a wide variety of applications including internal combustion engines, agricultural sprays, pharmaceutical powder production through spray drying process, aerosol generation for drug delivery, fire suppression, gas turbines, and many more [1]. The spray characterizations are needed in order to properly design systems that use sprays. Spray characterizations include spray droplet size, velocity, and number density distributions. Currently, spray characterizations are obtained outside of the actual industrial use of the spray and at a research stage of the system design. The same information is then used for the design of the actual system for the actual operating conditions. Since the operating conditions influences the spray characterizations, the design parameters have to be adjusted to optimize the system performance. This is a trial and error process with significant cost and time delays.

The objective of the present study is to develop an AI based method to characterize spray in the actual operating conditions. This can be achieved if the near nozzle information can be related to the spray characterization. Since it is much easier to determine the near nozzle information in most actual systems, the spray characterization in actual operating conditions can be determined in-situ and at low cost.

Our approach to this problem is through implementation of deep neural network artificial intelligence (AI) to produce statistically valid and accurate method of spray characterization with great predictive power. There are many examples of such scientific applications of deep neural networks, such as the work done by Ma, et al. [2], where deep neural networks were utilized for predicting the activity of potential drug molecules; the works of Wei, et al. [3] in predicting organic chemistry reactions using deep learning; and the most relevant to the work done in this study, the introduction of convolutional neural networks for image recognition pioneered by Krizhevsky, et al. [4], which classified the ImageNet database images with astonishing accuracy.

The present study examines the effectiveness of deep convolutional neural networks (CNN) as models for spray characterization. The most important characteristic of a spray is the droplet size distribution. The droplet size distribution is

obtained by a variety of methods. Most commonly used methods are either based on direct imaging of droplets and further image processing to determine the size of the droplets, or light scattering methods, such as the phase Doppler anemometry system that operates based on the Doppler effect in Mie scattering signals coming from the droplets. In a review done by Lee, et al. [5], it is mentioned that although the phase-Doppler method can be an accurate method of droplet sizing, because of difficulty in optical alignment and expensive equipment cost, it is mainly used by research laboratories and not in actual operating conditions.

Some of the limitations in the aforementioned methods can be mitigated and minimized by introducing deep learning AI. Specifically, by moving away from particle size measurements towards relating other aspects of the atomization process to the particle size distribution. This provides a pathway to move from early stages of atomization to characterizing the spray. Correspondingly, the need for the difficult optical setups and image processing will be eliminated. This stems from the fact that the particle measurement process will only be present once during the development of the database for training the deep neural network. In this study, images from the early stages of spray formation are used to distinguish between different operating conditions without any insight from the spray itself, the size distribution, or the physical parameters of the experiment. This is done using CNNs trained on the aforementioned images to distinguish between the different operating conditions. Here, we study the capability of CNNs in identifying and using physically relevant patterns in the spray formation process to provide useful information with respect to the spray.

2. Deep Learning and Convolutional Neural Networks

Deep learning refers to the machine learning methods that involve neural networks with more than two layers. This means that in deep neural networks, there exists many *hidden layers*, which are intermediate layers between the input and output. Each of these levels or layers is derived from imposing relatively simple non-linear functions on the previous level. This results in the conversion of raw data, such as image pixel values, into highly abstract formats from which incredibly complicated data relationships can be determined and used for sorting and categorization. In addition, while the number of layers and relative functions between them are specified during the creation of a deep learning model, the levels themselves are defined and altered through machine learning methods with no human intervention [6].

Deep learning models very often incorporate supervised learning. The work done in this study is an example of supervised learning. Essentially, this involves applying a training set consisting of many solved example problems to the model for learning purposes. Next, a validation set with solved problems that the model has not been introduced to yet is applied to gauge its performance. The results from this testing phase are used in conjunction with an error function to find the error in the model output in comparison to the known desired outputs for the testing set. This function is objective in nature and aims to minimize the error to the furthest extent possible. This is done by adjusting model parameters (weights) in a direction that reduces the error. These parameters are directly related to the layered representation discussed earlier and involve numerical coefficients that can be changed to alter the relationship between layer and model inputs and outputs. To make meaningful parameter changes and optimization, a gradient vector is created and used that can predict how a parameter adjustment would impact the errors [6]. Stochastic Gradient Descent (SGD) is one of the quickest and most common methods used to optimize parameters and involves using the input and the output along with its associated error to define an average gradient vector for a small set of training examples. Parameters are then adjusted according to this vector to minimise the error [6]. There have many methods derived from the SGD method, and in this study one of the methods that improves SGD significantly, called AdaMax [7], is used. It should be noted that the gradient vectors in a deep learning model with multiple layers are most easily determined using the backpropagation technique. This involves applying the derivative chain rule to find the gradient vector for the error function with respect to the input of a model layer by working backwards from the gradient of the next layer, from the computed output of the model to the raw input layer [6].

Convolutional neural networks (CNNs) are deep learning models that make very effective use of arrays of data, making them particularly suited for image recognition and classification. They make use of layered representations,

with the internal layers between the input and output layers being divided into convolutional, pooling, and fully connected layers [6].

Convolutional layers have elements, stylized as neurons, that are organised into feature maps. Each element of these maps is connected to locales in the feature maps of the previous layer in the model. This extends to the input layer which holds pixel values in most cases. The connections between layers are not random and are made through parameter sets called filter banks [6]. These are composed of filters or kernels stacked on top of each other. Each filter is relatively small (3x3 pixels, 6x6 pixels, etc.) but is convolved over the entirety of the input to produce a 2D activation (or feature) map from potentially 3D data matrices [8]. These filters hold parameters that are adjusted as a model is trained to identify specific features in the input. When features are identified, filters indicate them through activations which can be seen on the map. Filter banks sum the effects of multiple filters and pass this value into a non-linear function for abstraction. In most cases, this is the Rectified Linear Unit (ReLU) function which is a half-wave rectifier that responds to positive inputs [6].

A single convolutional layer will be composed of different feature maps with distinct filter banks. Each of these banks will be used to identify unique local features within input data and the usage of multiple distinct banks allows for the identification of common features throughout the entirety of the input. Data is then fed from one layer to the next, allowing for the detection of local relationships between features identified by previous convolutional layers. Pooling layers, on the other hand, are used to combine similar features detected by the convolutional layers into single larger features. They contain neuron elements that find the maximum value of a local set of convolutional elements in a feature map. Nearby pooling elements make use of input sets that do not neighbour each other as closely, resulting in the simplification of multiple input features into single larger combinations [6]. Fully connected layers are found near the output of a CNN and, as the name indicates, involve adjacent levels that are fully connected to each other, in contrast to convolutional and pooling layers. These layers make use of conventional ANN techniques and result in class values that dictate the model output [8].

As with other supervised deep learning models, CNNs make use of objective error functions, SGD, and backpropagation to adjust parameters and filters during model training. However, due to the large number of parameters and layers present, CNNs can be at risk of overfitting during training [8]. This involves training models that correspond entirely to the data sets with which they are trained rather than the general features present in these sets [9]. As such, they are unable to produce results when fed with unknown data. The most effective solution to this problem is to avoid making CNN models needlessly complex and making use of varied training sets with minimal in-set correlations.

3. Methodology

The present study involves (i) data gathering and labelling process, which is going to be the basis for training the CNN, and (ii) the CNN architecture and strategy for training. Both these aspects are discussed briefly to provide the insight into how the data was gathered and analysed.

3.1. Data Gathering and Experimental Setup

To be able to truly demonstrate the potential of CNNs, different types of datasets where gathered and labelled. However, the principles of how this process was done is common for all datasets. For any dataset, images of sprays were captured at different operating conditions (such as different pressures, flow rates, temperatures, viscosities, and more), and for each dataset the operating condition that was being altered for analysis was labelled. From here, a database of tens of thousands of images were produced, which was used to train multiple CNN models. The idea here is that by producing a dataset of sprays at different operating conditions and training a CNN to identify the operating conditions and predict them, it can be demonstrated that the use of CNNs for spray characterization is a valid method.

The experimental setup, as shown in Fig. 1, was intended to be simple and minimal, so as to mitigate the issues that are brought forth by traditional experimental methods for spray characterization. Therefore, a setup including only a camera and light source was used for capturing images of the spray. The camera was placed such that the images captured of the sprays would not include the actual spray, but the zones prior to particle formation, hence making it impossible for the CNN to retrieve any insight from the particles. This was done to ensure the primary objective of the research was not compromised, and the true ability of the CNN in identifying relevant features of the spray formation process were gauged and understood.

For this study the data was gathered using a flat-fan H1/4VV4001 nozzle at different operating pressures, and the CNN was trained to identify this operating pressure based on the images gathered from the nozzle.
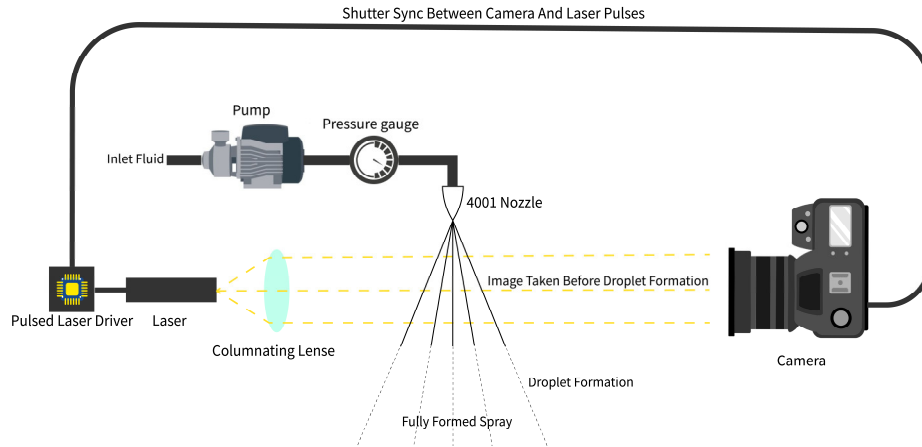

Fig. 1. Schematic of the experimental setup

The physical setup of the experiment included a camera and pulsed laser diode as a light source, while the pressure of the fluid was controller by adjusting a pump with respect to readings from a pressure gauge. For each experiment the pressure was adjusted to the desired value, and the spray was allowed to operate for a time to reach steady flow rate and pressure. At this point the camera was used to capture images of the spray at approximately 30 frames per second, while the pulsed laser was running in synch with the camera to provide short pulses of light. The short pulses(1μs) allow for limited exposure, which in turn eliminate motion blur in images. The image capturing was allowed to take place until 10,000 images (approximately 5 minutes) of each experiment were gathered at 18MP resolution.

3.2. Convolutional Neural Network Architecture and Approach

The architecture of the Neural Network used in this study is one that follows similar principles as that of VGG model developed by Simonyan, et al. [10]. This model uses grayscale images in place of RGB colour images, as there does not appear to be any useful insight in colour when it comes to the patterns in the fluid going through the atomization process. Also, by reducing the number of channels in the input of the CNN the number of parameters that must be trained are reduced three-fold, which allows for faster training of the model and a lighter model, which can be deployed on less powerful computers. The images gathered in the experiment are also reduced in size before being used in the CNN. This is because the architecture of the CNN has an input size of 150 px by 150 px, therefore the images are resized to 150 px by 150 px for the model. The approach in the model used in this study is to reduce the complexity of the model as much as possible, without lowering the accuracy of the model. To achieve this, many different architectures were tested with and models with different complexity levels were used to train for the spray images, and the results were compared to find the most effective yet light model that could identify the fluid pressure in the spray inlet with good accuracy. It was discovered that the effectiveness of a simple model with 3 layers of convolution with 3x3 receptive fields, were just as effective as larger and more complex models, and in fact increasing the effective receptive field size past 5x5 had negative effects on the model's ability to reach higher accuracy. Also, the number of convolution filters within each layer were minimized as well and the final result of the experimentation in the model led to a model shown in Fig. 2.
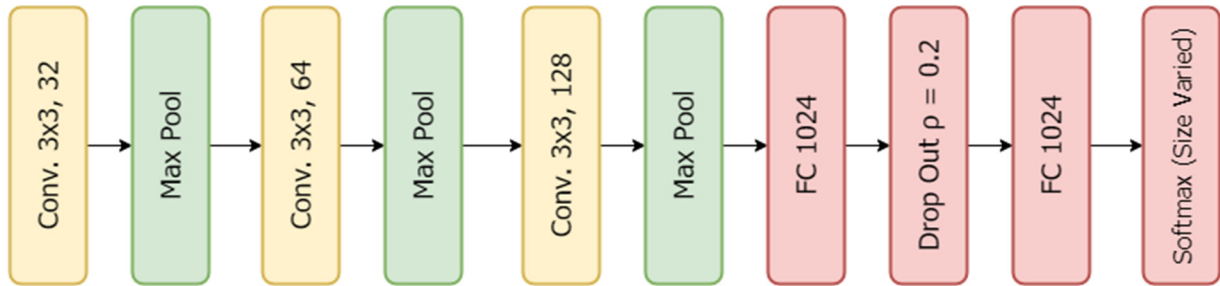
Fig. 2. Architecture of the CNN model used in this study

All layers in Fig. 2, except for the final layer have ReLu activation functions, while the last layer has a SoftMax activation function. As evident, there are three convolution layers followed by max pooling layers after each convolution layer, and after the third max-pooling layer a fully connected layer of size 1024 exists followed by another layer with similar size, with a drop out of density 0.2 placed between the two, and finally the model ends with a SoftMax layer with the size being set to the number of categories at the output, which is determined by the number of experimental operating conditions involved in each dataset. The strategies adopted in the training of the aforementioned model are also important to note. In the training process of the model a dropout layer was introduced before the last fully connected layer to limit the possibility of over-fitting during training, while also adopting the same augmentation methods employed by Simonyan, et al. [10] to further mitigate over-fitting. For the training of the model itself, the training algorithm used is the AdaMax method proposed by Kingma, et al. [7]. Finally, it is noteworthy to mention that for training 8000 images were used for each category, with 2000 images being set aside for validation of the model's accuracy, to validate the model's generalizability.

4. Results

The results are presented in two parts. The first model that is presented is one that is trained on images taken further away from the spray with a larger field of view, and another is one that is taken close up with more of the intricacies of the flow captured (Fig. 3).
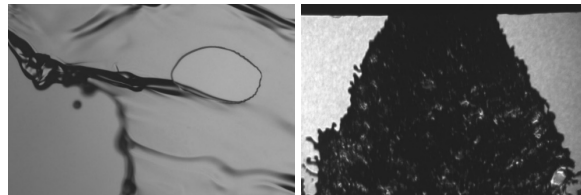

Fig. 3. Sample Images of close up (left) and far (right) images used for training.

4.1. Results from Close Up Image Training

The close-up images used for this model, were captured using a water glycerine mixture with 20% glycerine, at pressures of 50 psi, 70 psi, and 90 psi (Fig 4). Simple visual inspection of these images can hardly provide any discernible differences between them. The question was whether CNN can accurately differentiate these image from each other, and identify intricacies that slip the human eye.

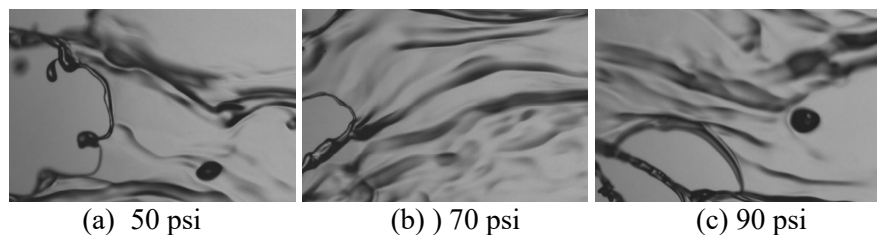(a) 50 psi   (b)) 70 psi   (c) 90 psi

Fig. 4. Sample closeup images of different sprays at different injection pressures.

A CNN with the previously mentioned structure was trained on the images to identify the pressures and classify any given images into the three pressure categories. The most accurate model after the training for this dataset was a 92% accuracy on the training set and 88% accurate on the validation set, which means that it can be expected that this model would be 88% accurate in classifying any image that falls under the three categories that the model was trained for.

4.2. Results from Far Image Training

For the images taken with a larger field of view, the data was gathered at 100 psi, 150 psi, 200 psi, 250 psi, 300 psi, 350 psi, and 400 psi (Fig 5) for water without any glycerine. In images from categories that are further apart the differences are visible to human eyes, however, in cases with pressures only 50 psi to 100 psi apart the differences are hardly visible. In this case as well, CNNs prove to be effective in identifying and classifying the images of the spray.

100 psi   150 psi   200 psi   250 psi
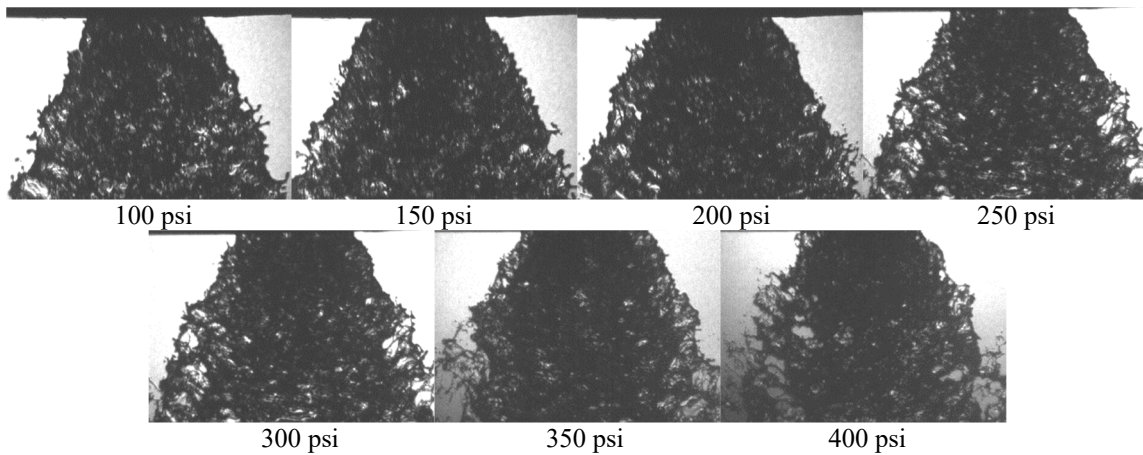
300 psi   350 psi   400 psi

Fig. 5. Sample Images of far data

The most accurate model after the training for this dataset was a 99% accuracy on the training set and 97% accurate on the validation set, which means that it can be expected that this model would be 97% accurate in classifying any image that falls under the seven categories that the model was trained for. As evident the model is notably more accurate than the previous model, which is as expected since the model is distinguishing between images from experiments that are 50 psi apart, compared to the previous case where experiments were 20 psi apart.

4.3. Visualizing The CNN

To obtain some insight into the inner workings of the model, the feature maps [11] of the convolution layers are presented in Fig. 6 for one input example to demonstrate the image processing in each convolution layer. These feature maps alone are difficult to understand and as we dive deeper into the CNN, the feature maps become even more obscure and unrecognizable. However, these feature maps are useful in understanding the process the image is taken through before being flattened into a simple set of nodes in a neural network layer to be connected to the fully connected layers. Essentially the last 128 small images that are shown in each example are flattened into all the pixel values and these pixel values are treated as nodes of the layer that connects to the first fully connected layer. This image illustrates how the needed information in the input image are highlighted and passed forward to be used in the final layers of the model to determine the operating conditions. In essence, the convolutional layers are used to extract important features from the image while reducing the size of the image, hence eliminating irrelevant information, to prevent over-fitting, and reduce the number of parameters in the model. These images are commonly referred to as feature maps [11] as they reflect the process of feature extraction and highlighting in the CNN.
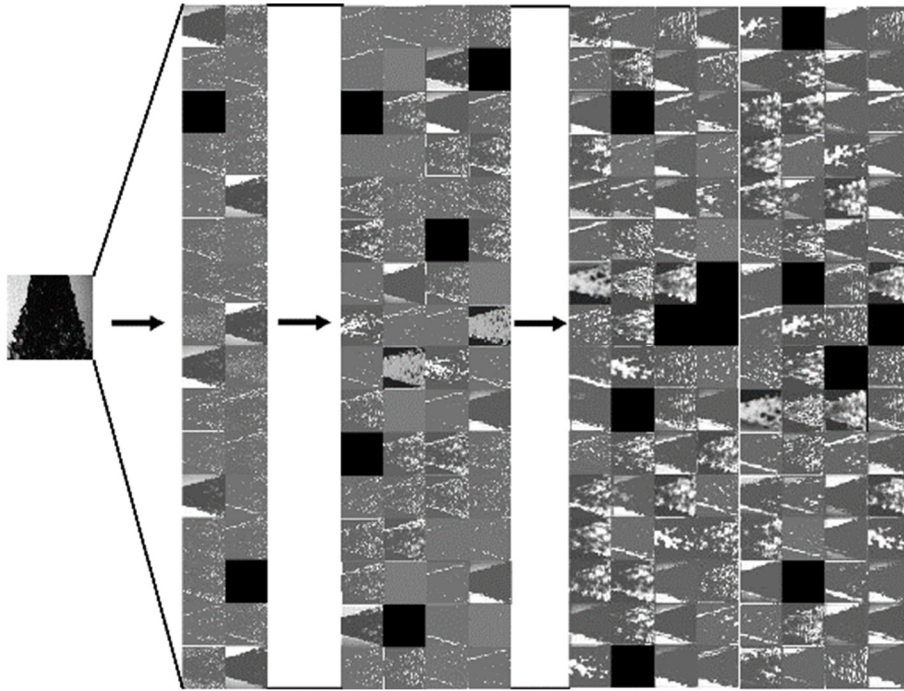
Fig. 6. Feature maps of the convolution layers for example inputs of each model.

5. CNN Analysis

So far it is established that the CNN has been successful in identifying features in the input images, that allow it to accurately distinguish between the spray images that are taken at different operating pressures. This demonstrates the ability of CNNs in bridging the gap between the atomization process before the spray is fully formed to the spray itself and its characteristics. This means that machine learning AI and specifically CNNs are able to provide an alternative and powerful model for spray characterization. However, this would not provide any further insight into the physics of the problem, and mostly serves as a black box with very little new information produced. CNNs are not necessarily too complex to understand and analyse, despite common belief. There have been many methods developed to understand the inner workings of CNNs, and one of the most recent methods developed for this purpose is the gradient class activation mapping (Grad-CAM) introduced by Selvaraju, et al. [12]. This method uses the gradients of any specific output (in this study this would be any output pressure) flowing into the final convolutional layer to produce a localization map high lighting the parts of an input image that are most important in making predictions. Figure 7 shows several Grad-CAM images. In these images the effects of the features extracted in the CNN (Fig. 6) on the decision the CNN makes at the end (the pressure classification) is investigated and the parts of the images that features were extracted from are highlighted. These are not all the features, but those features that influenced the classification in the CNN. The images presented below are localization maps, meaning the colours simply present the relative importance of each pixel on the decision the CNN makes; therefore no value is assigned to the contour maps. Areas that are red have the most influence while moving towards blue has less or negative influence on the decision. It is also important to note that since the localization map is normalized, in some cases the image is mostly red with some blue, which reflects that the decision was likely made by exclusion of features rather than inclusion.
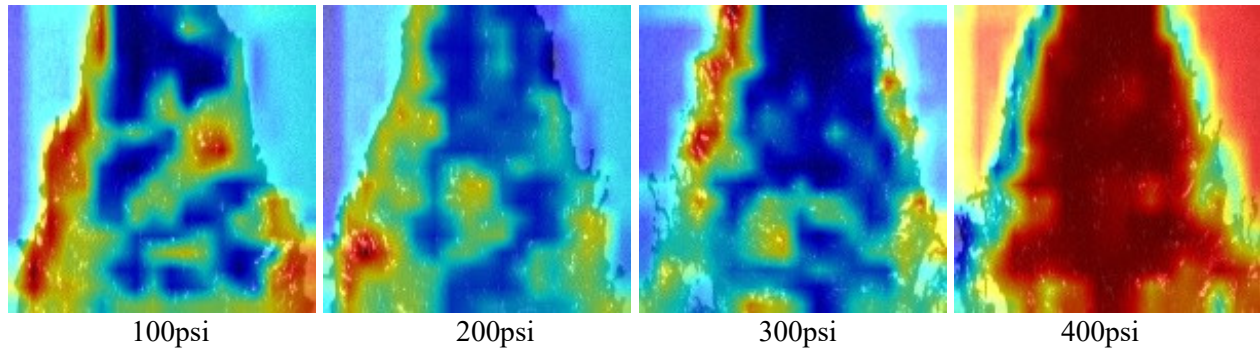
| 100psi | 200psi | 300psi | 400psi |

Fig. 7. Grad-CAM images of sprays at different pressures.

As evident the CNN has been focusing on the boundary of the spray to make distinctions between the spray images, which is a physically relevant part of the images as they reflect the turbulence patterns and intensity, which are changing with pressure. Applying the same approach for the close-up images, the Grad-CAMs are presented in Fig. 8. As evident, here the visible patterns of turbulent in the flow have been focused on for decision making as well.
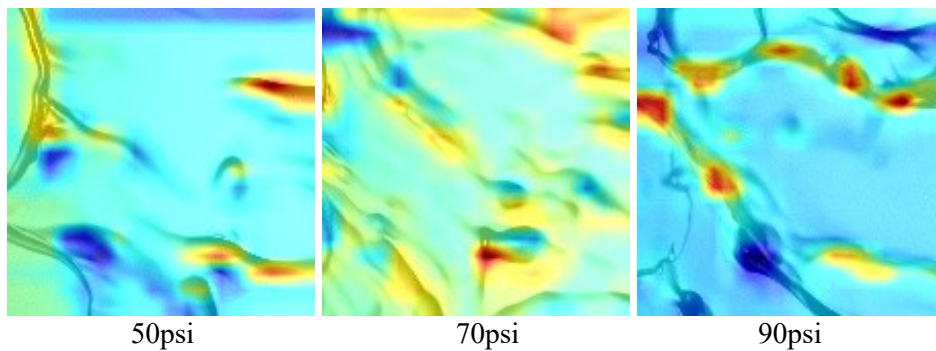

| 50psi | 70psi | 90psi |

Fig. 8. Grad-CAM images of closeup images of sprays at different pressures.

6. Conclusion

The results of the present study demonstrate the capabilities of CNNs and in general machine learning artificial intelligence in spray characterization. This was demonstrated by achieving a 94% accuracy in categorizing images of sprays with operating pressures 50 psi apart. Our study showed that CNNs were able to identify sprays that were operating at different inlet fluid pressures, without gaining any insight into the spray droplets themselves. This means that the CNN was able to distinguish between different sprays using the pre-breakup patterns in the spray, which shows the potential for CNNs to bridge the gap between the droplet formation and breakup process and the spray droplets at the end of this process, which would primarily be the droplet size distribution and other spray characteristics. The work presented in this study, shows machine learning AI as a viable tool for spray characterization with a statistical approach.

Furthermore, we were able to visualize the inner workings of the model using the Grad-CAM method to further our understanding of sprays. This introduces a new methodology for extracting information from experimental data. Image base data from sprays that is labelled and categorizable can be used in training CNNs and if successful, by visualizing the CNN's decision-making process through Grad-CAM or any other methods details can be extracted from the results. These details can provide previously unknown information about the droplet formation process. Moreover, the same approach can be potentially implemented in the future to develop models that directly relate spray images prior to droplet formation to the droplet size distribution, and provide a generalizable model for many types of nozzles and spray systems that can be deployed in industry and research to accelerate the development and assessment of spray systems and mitigate the need for specialized and expensive equipment.

References

[1] N. Ashgriz, Handbook of atomization and sprays, New York: Springer, 2011.

[2] J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl and V. Svetnik, "Deep Neural Nets as a Method for Quantitative Structure−Activity," *Journal of Chemical Information and Modeling,* vol. 55, pp. 263-274, 2015.

[3] J. N. Wei, D. Duvenaud and A. Aspuru-Guzik, "Neural Networks for the Prediction of Organic Chemistry Reactions," *ACS Central Science,* vol. 2, pp. 725-732, 2016.

[4] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in neural information processing systems,* pp. 1097-1105, 2012.

[5] S. Y. Lee and Y. D. Kim, "Sizing of Spray Particles Using Image Processing Technique," *KSME International Journal,* vol. 18, pp. 879-894, 2004.

[6] Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning," *Nature,* vol. 521, pp. 436-439, 2015.

[7] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv,* 2014.

[8] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *ArXiv e-prints,* pp. 3-8, 2015.

[9] "overfitting," Lexico, [Online]. Available: https://www.lexico.com/definition/overfitting. [Accessed 16 03 2020].

[10] K. Simonyan and A. Zisserman, "very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv,* 2014.

[11] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks," *European conference on computer vision,* pp. 818-833, 2014.

[12] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," *International Journal of Computer Vision,* vol. 128, p. 336–359, 2019.