Denoising and fuel spray droplet detection from light-scattered images using deep learning

Veeraraghava Raju Hasti [a,*], Dongyun Shin [b]

[a] School of Mechanical Engineering, Purdue University, West Lafayette, IN,47907 United States of America
[b] School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN,47907 United States of America

HIGHLIGHTS

- A deep learning method proposed for denoising and droplets detection from light scatterd images.
- Performance assessment of standard CNN, modified ResNet, and modified U-Net architectures.
- Modified U-Net architecture out-performed for denoising light-scattered spray image.
- Potential for real-time processing of the experimental data using the deep neural networks.

GRAPHICAL ABSTRACT
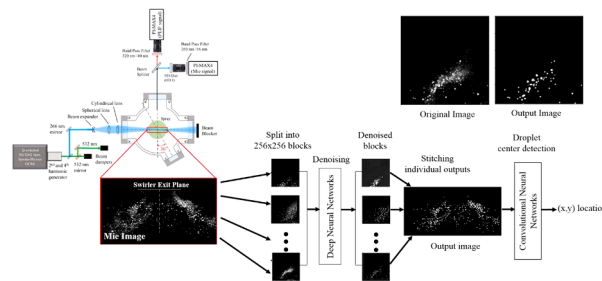
ARTICLE INFO

ABSTRACT

A deep learning-based method for denoising and detecting the gas turbine engine spray droplets in the light-scattered image (Mie scattering) is proposed for the first time. A modified U-Net architecture is employed in the proposed method to denoise and regenerate the droplets. We have compared and validated the performance of the modified U-Net architecture with standard conventional neural networks (CNN) and modified ResNet architectures for denoising spray images from the Mie scattering experiment. The modified U-Net architecture performed better than the other two networks with significantly lower Mean Squared Error (MSE) on the validation dataset. The modified U-Net architecture also produced images with the highest Power Signal to Noise Ratio (PSNR) compared to the other two networks. This superior performance of the modified U-Net architecture is attributed to the encoder-decoder structure. During downsampling, as part of the encoder, only the most prominent features of the image are selectively retained by excluding any noise. This reconstruction of the noise-free features has produced a more accurate and better denoised image. The denoised images are then passed through a center predictor CNN to determine the location of the droplets with an average error of 1.4 pixels. The trained deep learning method for denoising and droplet center detection takes about 2.13 s on a single graphics processing unit (GPU). This study shows the promise for real-time processing of the experimental data using the well-optimized network.

---

* Corresponding author at: 585, School of Mechanical Engineering, Purdue Mall, West Lafayette, IN 47907 United States of America
E-mail address: vhasti@purdue.edu (V.R. Hasti).

1. Introduction

Gas turbine combustion performance, pollutant formation, and operability limits strongly depend on the fuel spray atomization, droplet size distribution and velocity, droplet evaporation, and mixing characteristics [1]. High fidelity computational model development to accurately predict the lean blowout (LBO) equivalence ratio [2], fuel sensitivity to LBO limits [3], flame characteristics during the LBO [4] in a combustor depends on the accurate spray boundary conditions obtained from the experimental techniques. Thus, identifying the droplet size and distribution of the spray is a critical task in characterizing the sprays. Although the measurement technique can be selected depending on applications, the Phase Doppler Anemometry (PDA) and laser diffraction analyzer have been widely used in the spray community for measuring the drop size and drop distribution of the spray due to high accuracy and high spatial resolution [5]. However, the measurements with such techniques can be a trivial and time/resource-consuming process to construct a two-dimensional spray field since both methods are limited to single point-wise and line-of-sight measurements. The laser sheet drop sizing technique, such as laser-induced fluorescence (LIF)/Mie signal ratio using a single [6] or two-phase [7] Structured Laser Illuminated Planar Imaging (SLIPI), can provide a two-dimensional drop size map of the spray. However, the fundamental of LIF/Mie ratio drop sizing is not always valid [8], and the method may not be accurate [9] due to various factors such as multiple scattering [10], aromatic/fluorescent dye [11], and fuel concentration [12], and evaporation [13]. Furthermore, the experimental system for the SLIPI technique can be significantly complicated by adding multiple phases for a better resolution with mitigation of the multiple scattering effects [7].

The light scattered image often suffers from multiple scattering, when photons scattered from a particle are re-scattered from neighboring particles before reaching the detecting sensor. This results in a foggy and false particle image, especially for dense particle cases. Several conventional segmentation methods, such as the overlapping object recognition (OOR) algorithm [14], shape-criterion-based method (elliptical fitting) [15], concentric circular arrangement method [16], contour splitting algorithm [17], and touching cells splitting algorithm [18], require various segmentation parameters to be adjusted, such as thresholding for binarization, morphological operators, etc. However, these conventional methods suffer degraded performance with complex shape particles or particle clusters and noises in the image. Although the watershed method [19] can be employed more robustly in particle segmentation applications, it is still limited in segmenting overlapping objects.

Machine learning has been a widely used technique for image segmentation with the development of a convolutional neural network (CNN). An autoencoder neural network has been used to reconstruct the desired image from the original compressed image through dimensionality reduction for image denoising, image segmentation, and audio compression. One of the most widely used autoencoder architectures is the U-Net, primarily developed for medical image segmentation [20]. The U-Net architecture consists of an encoder (contracting path) and a decoder (expansive path), which gives it the U-shaped architecture. The encoder downsamples the image keeping only the most prominent features by removing the noise. Each convolution is followed by a rectified linear unit (ReLU) and a max-pooling operation. The decoder is the inverse operation of the encoder. The decoder upsamples the image and reconstructs the input image to get the desired output image. The skip connections between the corresponding layers in the encoder and decoder add the important features that may be removed during the downsampling in the image reconstruction process.

Several studies have employed deep learning algorithms using CNN in various particle segmentation applications. Li et al. [21] used a single convolution neural network to analyze shadowgraph images of bubbles. Their approach employs a two-channel-output U-Net model to create a binary particle image and a particle centroid image segmented through a marker-controlled watershed approach. However, their approach relies on high-quality synthesized particle images. Furthermore, the model accuracy is limited by the loss of particle information due to overlapped particles. Nobari *et al.* [22] adopted the VGG model developed by Simonyan et al. [23] to classify the spray images recorded at different fluid pressures. This method distinguished patterns in sprays before atomization, but it is limited to characterizing the primary breakup of the spray, such as bulk of liquid, liquid sheet, or ligament. Oktay et al. [24] employed the multiple output convolutional neural networks (MO—CNN) with Hough transform for simultaneous detection, localization, and segmentation of the nano-particles from transmission electron microscopy (TEM). Although this model provides the location, size, center, and radius of each particle simultaneously for both round and elliptical shapes, the model requires a masked dataset for training their model, which is time-consuming. Zhang et al. [25] proposed a Mask R-CNN segmentation model with an edge fitting method to determine the size and shape parameters of nano-particles from TEM images. However, the localization of each particle is missing. Various deep learning methods for particle measurement have been proposed for different applications. However, the deep learning model for the gas turbine engine spray application has not been proposed, especially for the spray droplets from laser-based Mie-scattered images.

In the present work, we propose a deep learning-based method for the first time to denoise the Mie-scattering spray images resulting from the non-illuminated regions, possibly due to the multiple scattering, and to detect individual droplets in the denoised images. Multiple scattering in this study refers to a phenomenon in which a Mie photon scatters off several other droplets before reaching the camera sensor. These multiple scattered paths of photons can lead to undesired effects such as blurring, haziness, and detection of light intensities from non-illuminated areas [6]. Furthermore, the proposed model does not require a masked dataset, which significantly reduces human effort and error in preparing the training dataset. The spray images in the present study are obtained from the complex hybrid pressure-swirl airblast (HPSA) fuel injector under realistic gas turbine engine conditions. Our work also focuses on validating the performance and capability of three deep learning architectures, namely standard convolutional neural networks (CNN), modified ResNet, and modified U-Net for denoising the spray images. In addition, we have developed center predictor CNN to detect the location of the individual droplets from the denoised images. This study is aimed at reducing the post-processing time from the Mie images using the trained deep learning algorithm.

The paper is organized as follows: Details of the experimental setup and data collection are given in Section 2; deep learning methodology is explained in Section 3; hyperparameter tuning, training, validation, and key results are discussed in Section 4; and important conclusions are presented in Section 5.

2. Description of the experimental setup and data collection

2.1. Experimental system

The spray measurements are performed at the Purdue Variable Ambient Pressure Spray (VAPS) test rig using a hybrid pressure-swirl airblast (HPSA) atomizer. The detailed description of the VAPS test rig with flow line configuration can be found in the previous works (LBO [26] and cold-start [27]) by Shin *et al*. The HPSA atomizer was designed by the Parker-Hannifin Corporation, and the schematic of the internal layout of the atomizer is shown in Fig. 1. The HPSA atomizer is composed of two components: the prime injector and air swirlers. The prime injector has two fuel circuits for pilot and main lines, but only a pilot line is used in this study. The air swirler has three swirler passages: inner swirler, dome swirler, and outer swirler. These swirlers surround the prime injector. The airflow through these swirlers helps in atomizing the fuel spray and mixing. The mechanism of the atomization process is
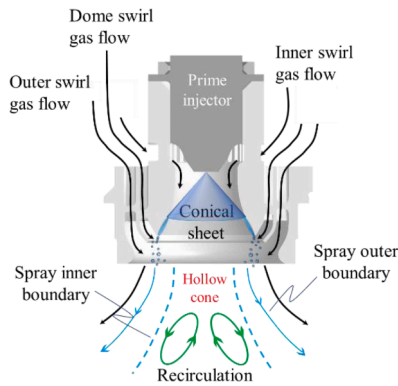
Fig. 1. Schematic diagram of hybrid pressure-swirl airblast (HPSA) atomizer. Black solid lines indicate the flow through three different swirlers: inner, dome, and outer. Blue solid lines indicate the outer boundary of the spray. Blue dashed lines indicate the inner boundary of the spray (hollow-cone region). Adapted and modified from previous studies at LBO [26] and cold-start conditions [27] by Shin et al.

as follows: (1) a hollow cone fuel spray exits the pilot nozzle of the prime injector and impinges on the pre-filming surface, (2) a fuel film forms on the surface and flows towards the prefilmer tip, (3) fuel ligaments are formed as the fuel film is torn and extends from the prefilmer tip, (4) the dome/outer swirling gas flows interact with the ligaments and disintegrate them into droplets. A detailed discussion of this atomizer can be found in Mansour et al. [29].

The pressure vessel of the VAPS test rig has four optical windows, as shown in Fig. 2. Three 127-mm diameter windows are oriented perpendicular to each other, and a 76.2-mm diameter window is oriented at $60°$ from one of the 127-mm windows. The heated nitrogen flow (394 K) is supplied directly into the pressure vessel to build a desired pressure and temperature in the vessel. The fuel line is directly connected to the pilot line of the prime injector. The operating conditions for the spray measurement are corresponding to vessel pressure $P_{amb} = 1.01$ bar, fuel temperature $T_{fuel} = 332$ K, atomizing gas temperature $T_{airbox} = 394$ K, fuel injection pressure differential $\Delta P_{pilot} = 1.72$ bar, and gas swirler pressure drop $\Delta P/P = 3\%$.
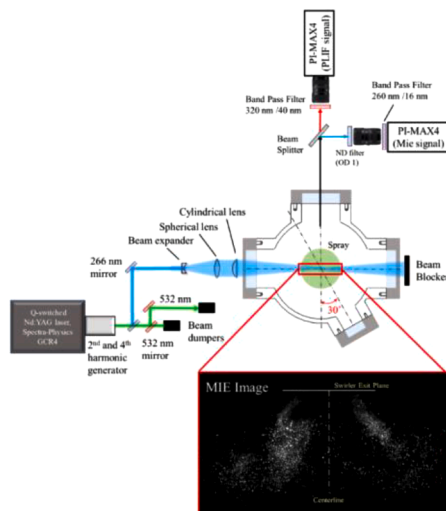
Fig. 2. Schematic diagram of simultaneous PLIF and Mie scattering image measurement system in the VAPS test rig. An instantaneous Mie image of the spray is also shown below. Distance from the swirler exit plane to the bottom of the spray is approximately 40 mm. Adapted and modified from work by Shin et al. [28].

2.2. Planar laser-induced fluorescence (PLIF) and Mie scattering measurement

Simultaneous fuel-PLIF and Mie scattering measurements are conducted for two-dimensional spray visualizations. Fig. 2 shows the schematic diagram of the measurement systems in the VAPS test rig. A Q-switched Nd: YAG laser is tuned to 266 nm through the second and fourth harmonic generation processes at 70 mJ/pulse with 10 Hz of repetition rate. This beam is expanded to approximately 40 mm in height and collimated into a cylindrical lens to create a sheet with approximately 0.5 mm thickness at the focused plane. This laser sheet enters the pressure vessel through the fused silica window. Jet-A, which has aromatic compounds such as alkyl-benzene and alkyl-naphthalenes, is used for this work. The fuel fluorescence light and Mie-scattered light are separated from a dichroic beam splitter and entered into two time-synchronized ICCD cameras (PI-MAX4) equipped with UV lenses (Nikkor 70~210 mm f/4.5 and Objectif UV 100 f/4.5). For the PLIF camera, a transmission filter centered at 320 nm with a bandpass of 40 nm is used to capture the fluorescence signals. For the Mie camera, a transmission filter centered at 260 nm with a bandpass of 16 nm is used to capture only 266 nm scattered light. Each of the 800 instantaneous PLIF and Mie-scattered images is recorded at 5 Hz. For the present work, only 800 Mie-scattered images are used and processed for analysis. One of the 800 instantaneous Mie-scattered images obtained at $P_{amb} = 1.01$ bar, e $T_{fuel} = 332$ K, $T_{airbox} = 394$ K $\Delta P_{pilot} = 1.72$ bar, and $\Delta P/P = 3\%$ is shown in Fig. 2.

2.3. Noise from Mie scattering image

Fig. 3 shows the pixel intensity profile of a portion of an original instantaneous Mie image. In this work, the noise is defined as the signal that causes blurriness, haziness, undesired light intensities from non-illuminated areas. As shown in Fig. 3, the undesired light signals were generated in the dense particle region of the spray, possibly due to the multiple scattering. This study proposes a deep learning model to filter these undesired light signals and regenerate the spray droplet image. The pixel intensity profile shows that these undesired signals are relatively lower than the signals from the droplets. The signal intensity criteria for the droplets was determined based on the comparison of simultaneous PLIF and Mie images. Since both Mie and PLIF images were taken simultaneously, the cross-correlation was performed between the instantaneous PLIF and Mie images by identifying the pixel-to-pixel locations of the observable droplets on both images. Once the pixel-to-pixel locations were determined between the PLIF and Mie images, the difference of image can be obtained by subtracting the Mie image from the PLIF image for the signals from the vapor phase. Note that each PLIF and Mie image were normalized by the peak intensity value before the subtraction. By discriminating the vapor phase, the non-liquid region in the Mie image was able to be identified. Based on the signal intensity values in the non-liquid region in the Mie spray image, the signal intensity criteria for the droplets was determined.

2.4. Data preparation for training and validation of deep learning model

The original Mie images are first passed through a Gaussian filter with a filter size of 3 × 3. A Gaussian filter is a filter whose impulse response is a Gaussian. In image processing, a Gaussian filter is a weighted averaging filter with weights representing the value of Gaussian function. Since Gaussian filter is an averaging filter, it blurs or smoothens the image to remove any high frequency components (noise) from the image. It is a preferred choice of filter to be applied before carrying out edge detection as it smoothens out the image and reduces the possibility of categorizing noise as edge. A 3 × 3 Gaussian filter convolves with a 3 × 3 window around the center pixel to determine the new value of the center pixel. As the filter size increases, the effect of faraway pixels is also included. The Gaussian filter works fine in
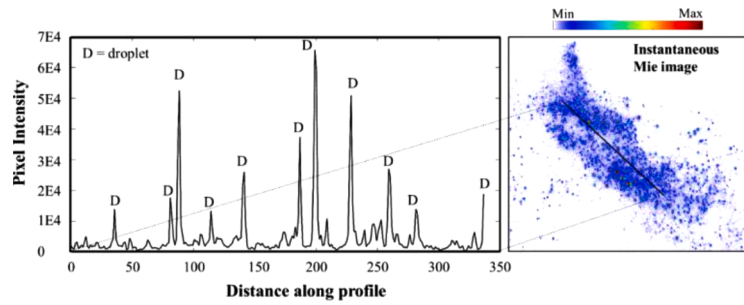
Fig. 3. Pixel intensity profile of the instantaneous Mie image.

removing various types of noises, however is not effective against salt and pepper noise. This noise appears when an image is transferred over a noisy digital channel. Also, since Gaussian filter is an averaging filter there is an inherent loss of image quality depending on the amount of smoothness. Hence, it cannot be used when noise level in the image is very high. The accuracy of the Gaussian filter is dependent on the amount of smoothening that is applied to the image. The Gaussian filter cannot reproduce the original image with 100% accuracy as there is loss of image quality. The filtered image is then passed through a Canny Edge detector which detects the outer boundary of the droplets. Once the edges are detected, a bounding box is drawn around each droplet to separate it from the other droplets. The center of the bounding box corresponds to the center of the droplet and the smaller dimension of the bounding box is considered as the diameter of the droplet. With this information about the center and radius of the droplets, a new clean image is generated with replacing each droplet as a circle with center and radius obtained from the previous step.

As an example, the original Mie scattered image collected from the experiment is shown in Fig. 4(a)) and the corresponding filtered image is shown in Fig. 4(b). The detected droplets are enhanced using a series of erosion-dilation steps dependent on each image. Each image is resized to 1024 × 512 and split into blocks of size 256 × 256 pixels for efficient processing with the neural networks. A larger block size greatly increases the computational requirements, whereas a smaller block size reduces the computational complexity significantly and improves accuracy with localized predictions. Figs. 4(c) and 4(d) show the magnified view of such one block (256 × 256 size) from the original and filtered image, respectively.

3. Deep learning methodology

The proposed deep learning methodology for denoising the fuel spray images from the Mie scattering and droplet center detection is shown in Fig. 5. We present details of this methodology, including the standard convolutional neural networks (CNN) [30], modified ResNet [31], and modified U-Net [20] architectures in this section. The original architectures are designed to be trained on a large dataset like ImageNet, making it infeasible for our case to be used directly. Hence, a lighter version of the original architectures is developed that can be efficiently trained on the available dataset without overfitting.

3.1. Standard CNN architecture

The standard CNN architecture is shown in Fig. 6, and it comprises the convolutional layers stacked one after another. No downsampling of the input image is performed to maintain the exact image resolution for output and input image. Each convolutional layer uses a filter of size 3 × 3 with a stride of 2 and additional padding to maintain the resolution. A total of 8 hidden layers are implemented, with the first 4 layers convolved with 16 filters and the last 4 with 32 filters. Each layer passes through a ReLU [32] activation and then batch normalization layer [33]. The final layer implements a 1 × 1 convolution to merge the 32 channels of the last hidden layer and create the final output image.

3.2. Modified ResNet architecture

The original ResNet architecture is designed to work as a feature extractor network connected to a classification head and trained on the ImageNet dataset. This architecture involves strided convolutions and pooling operations to reduce the dimension of the input. However, we
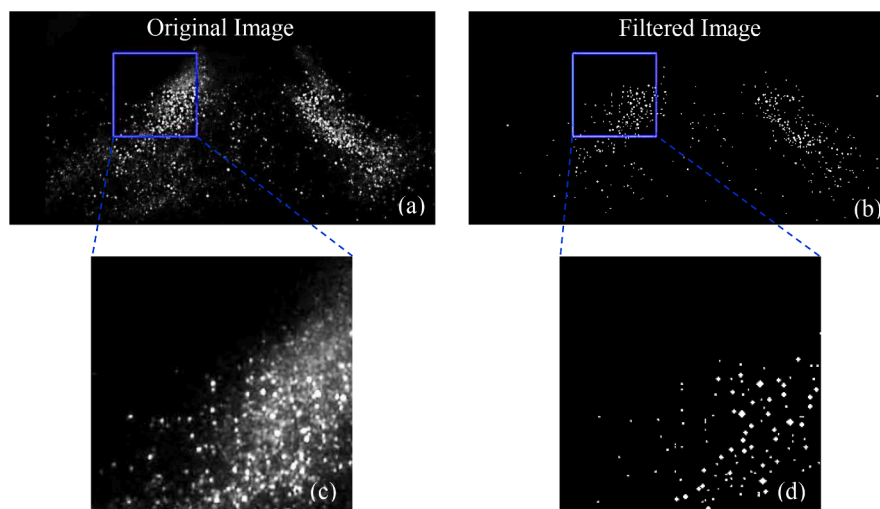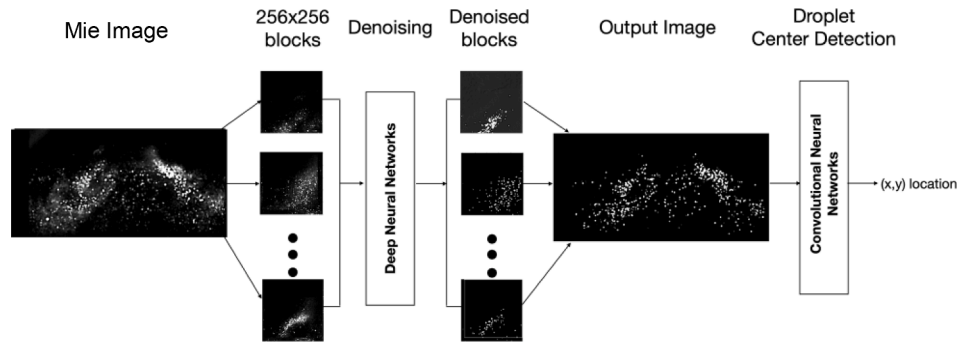

Fig. 4. (a), (b) Example of the Mie scattered spay image and the corresponding filtered image. (c), (d) Zoomed section of the original and filtered image.

Fig. 5. General outline for the deep learning methodology. The original Mie image is split into 256 × 256 blocks and denoised using deep learning. Denoised image blocks are then stitched together for the output image.
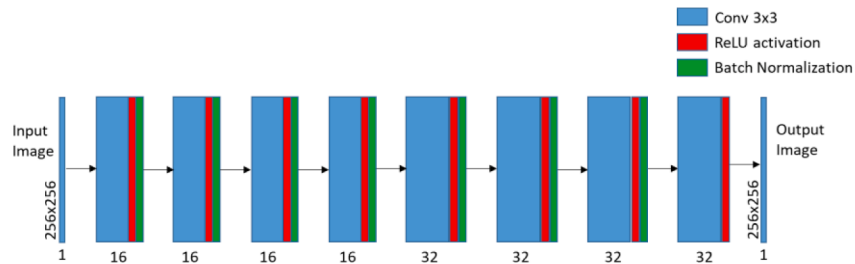

Fig. 6. Block diagram of the standard CNN architecture for denoising of fuel spray from Mie scattering images.

require the dimensions to remain the same for the image to image translation problem (the present study). Hence, a modified ResNet-inspired architecture is developed.

For the modified ResNet-inspired architecture shown in Fig. 7, the convolutional layers are replaced by convolutional blocks. Each convolutional block comprises two convolutional layers with a skip connection between the input image and the output image. The convolutional layers use a filter of size 5 × 5 followed by ReLU activation and batch normalization. Skip connection allows for a deeper network while preventing overfitting of the data. The first convolutional layer uses a 3 × 3 filter to get 32 feature maps of the input passed to the convolutional blocks. The output layer uses a 1 × 1 convolution with ReLU activation.

3.3. Modified U-Net architecture

The U-Net architecture was originally developed for semantic segmentation of medical images very efficiently. However, we propose to use this U-Net architecture with modifications for denoising the Mie images. The modifications are made to the number of feature maps and

the shape of the bottleneck layer to prevent overfitting of the network on the training data.

The U-Net architecture comprises an encoder and decoder network with the layers connected through the skip-connections. Every level of the encoder consists of a convolutional layer followed by a downsampling layer. A 3 × 3 filter is used in each convolutional layer. The convolutional layer is followed by a ReLU activation and batch normalization. Instead of the pooling layers, we use strided convolution with a stride length of 2 for downsampling the input. The bottleneck layer is set of size 8 × 8. The decoder network upsamples the encoder output from 8 × 8 back to 256 × 256 using transposed convolutions. Each level of the decoder comprises a convolutional layer followed by an upsampling layer. The convolution operation uses a 3 × 3 filter followed by the ReLU activation. Overfitting the network is prevented by using dropout layers with setting a dropout rate to a value of 0.2. The last layer uses Sigmoid activation to force the output between 0 and 1. The encoder and decoder levels are connected using skip connections that bypass the bottleneck layer to transfer information from the encoder section to the decoder section. The modified U-Net architecture is shown in Fig. 8.
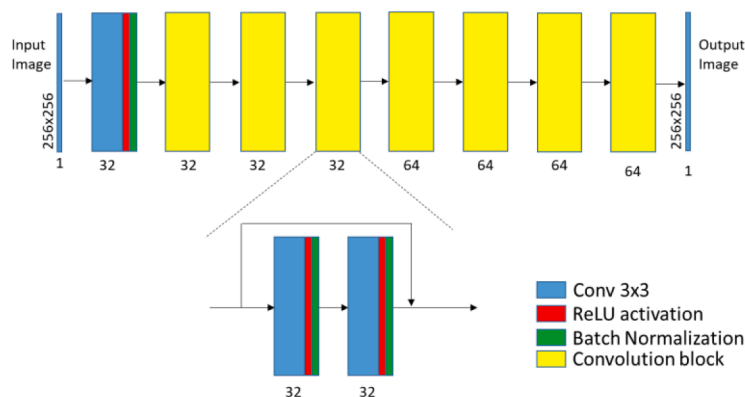

Fig. 7. Block diagram of the modified ResNet architecture for denoising of fuel spray from Mie scattering images.
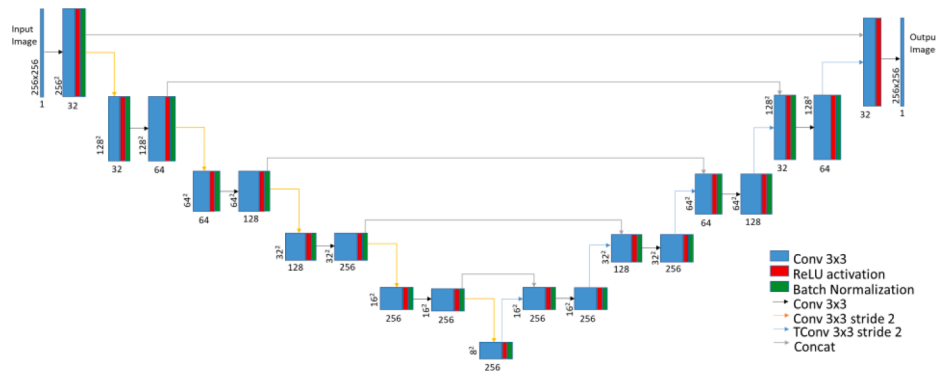
Fig. 8. Block diagram of the modified U-Net architecture for denoising of fuel spray from Mie scattering images.

3.4. Metrics for performance evaluation

Each network is trained to optimize the Mean Square Error (MSE) between the predicted output and the true output. MSE is the preferred loss function as the aim is to minimize the absolute difference between pixels in the true denoised image and the predicted denoised image. The MSE as loss function (L) can be expressed as

$$L(y, f(x)) = \frac{1}{n} \sum^{|y - f(x)|^2} \qquad (1)$$

where, x is the input image, y is the true output, and $f(x)$ is the predicted output by the network. The sparsity of the droplets in the image signifies that most of the convolution outputs will be 0, resulting in very slow learning of the network. The Adam optimizer is used to handle the sparse gradients. MSE is an important parameter for evaluating the performance of a denoising algorithm. As in Eq. (1), it computes pixel-wise correctness of the predicted image compared to the original output. The smaller the MSE, the better the denoising results.

The Power Signal to Noise Ratio (PSNR) is the ratio between the maximum possible value of the pixel and the noise that affects the fidelity of the image. PSNR can be expressed as

$$PSNR = 10 log\left(I_{\mathrm{MAX}}^2 / MSE\right) \qquad (2)$$

where, I_{MAX} is the maximum pixel intensity in the image. As the images are normalized before processing, the maximum intensity value is 1. Hence, Eq. (2) can be simplified as

$$PSNR = -10 log(MSE) \qquad (3)$$

The higher PSNR value indicates a better denoising algorithm.

3.5. Droplet center detection

The Mie images denoised by the best performing deep neural networks are then passed to a center predictor convolutional neural network (CNN) to predict the center location of the droplets. The center predictor network is a 3-layer CNN with each layer using a 3×3 convolution filter and ReLU activation function. Each layer is followed by a batch normalization layer. The network is trained to optimize the MSE between the actual center pixels and the predicted center pixels. The deep learning methodology is implemented in Python 3.6 using the TensorFlow 2.5.0 library and trained on Google Colab with a single Nvidia T4 graphics processing unit (GPU).

4. Results and discussion

A thorough hyperparameter tuning was performed for each of the three networks to select the optimum values that result in the lowest MSE using the grid search method. For each hyperparameter, different

values were selected. Every combination of hyperparameters was run for a total of 10 epochs on a truncated dataset containing selectively chosen samples that represent the entire dataset for faster computation and adjustability. The range for the hidden layers was set as (5,10) for CNN, (10,15) for the modified ResNet, and (3,8) for the encoder section of the modified U-Net. The filter size was selected from 3×3, 5×5, and 7×7. Larger filter sizes are not suitable for images having very small objects. ReLU, tanh, and Leaky ReLU were the choices for activation function, and the optimizer was selected from Stochastic Gradient Descent and Adam. For the bottleneck layer, the choices were 4×4 to 32×32 with increments of 2. The learning rate was set from 1e-1 to 1e-4. The list of hyperparameters and the optimum values for each deep learning architecture are listed in Table 1. The standard CNN showed strong overfitting for hidden layers above 8, whereas skip-connection-based ResNet allowed for a deeper network with 13 layers. However, the rapid growth in trainable parameters prevented more depth. Modified U-Net style network with autoencoder network and skip connections yielded optimum results with 11 hidden layers and a bottleneck layer of size 8×8. The better performance is achieved with a filter size of 3×3 for standard CNN and Modified U-Net, and 5×5 for ResNet. All three networks showed better results with the ReLU activation and Adam optimizer.

In this study, the input data was split into training and validation pairs with an 80:20 ratio. Each of the 800 Mie images was resized and split into 8 blocks, which resulted in 6400 images. Out of 6400 images, 5120 images were used for the training set, and 1280 images were used for validation. The images were randomized before splitting to avoid clustering of similar images in the training set. Each training image went through an image augmentation pipeline comprising random crop, flip, random shift, and brightness effects to add more diversity and variability in the dataset, which dynamically increases the total number of training images fed into the network. With the optimum hyperparameters selected, each network ran on the entire dataset till the loss converges. The plot of loss vs. the number of epochs is shown in Fig. 9. As observed from the plots, the loss for all the three networks converges after 10 epochs. The convergence is determined when the change in the loss is less than 0.01.

The modified U-Net performed significantly better than the other two networks with substantially lower training loss and validation loss.

Table 1
Optimum Hyperparameters.

Parameters	Standard CNN	Modified ResNet	Modified U-Net
Hidden layers	8	13	11
Filter size	3×3	5×5	3×3
Activation Function	ReLU	ReLU	ReLU
Optimizer	Adam	Adam	Adam
Learning rate	1e-3	1e-3	1e-3
Bottleneck layer	–	–	8×8

(a) Standard CNN
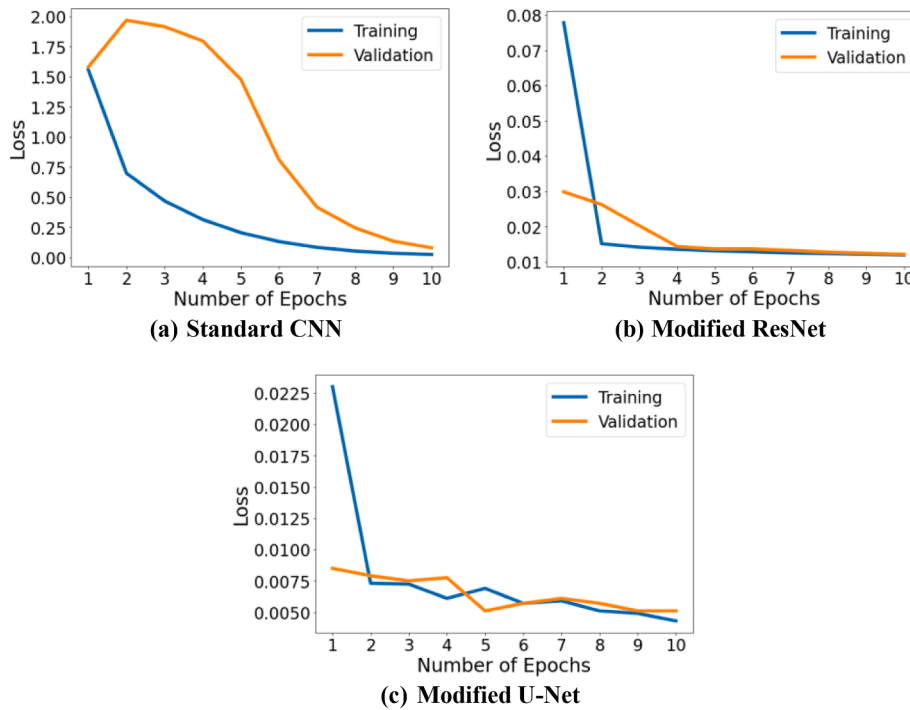

(b) Modified ResNet

(c) Modified U-Net

Fig. 9. Training and validation loss versus the number of epochs for three different deep learning architectures used for spray image denoising.

Also, modified U-Net showed a faster convergence of loss for both training and validation than the other two networks. The standard CNN performed the worst with the highest loss among the three, with the validation loss showing a non-monotonous convergence. Further evaluation of the networks was performed using the metrics mentioned in Section 3.4 for different block sizes. The effect of block size on the performance of the three model architectures is quantitatively assessed, and results are shown in Table 2. The smaller block size takes less computational time but results in higher MSE and lower PSNR values. The block size of 256 × 256 pixels gives more accurate results with the lowest MSE values and highest PSNR values for all three architectures but takes more training time per epoch. The standard CNN with a simpler architecture takes lesser time for training per epoch, while the complex modified U-Net architecture takes more training time per epoch. Among three architectures, the modified U-Net architecture gives the lowest MSE and highest PSNR values with higher computational time. The modified U-Net architecture generated better-denoised images with the highest PSNR value of 22.757 dB.

The comparisons of denoised images resulted from each network are shown in Fig. 10. The original Mie image and the true denoised image are shown in Figs. 10(a) and 10(b). The output from the standard CNN, modified ResNet, and modified U-Net are shown in Fig. 10(c), Fig. 10 (d), and Fig. 10(e), respectively. The modified U-Net produced a more accurate output with clearly segmented droplets from the noise.
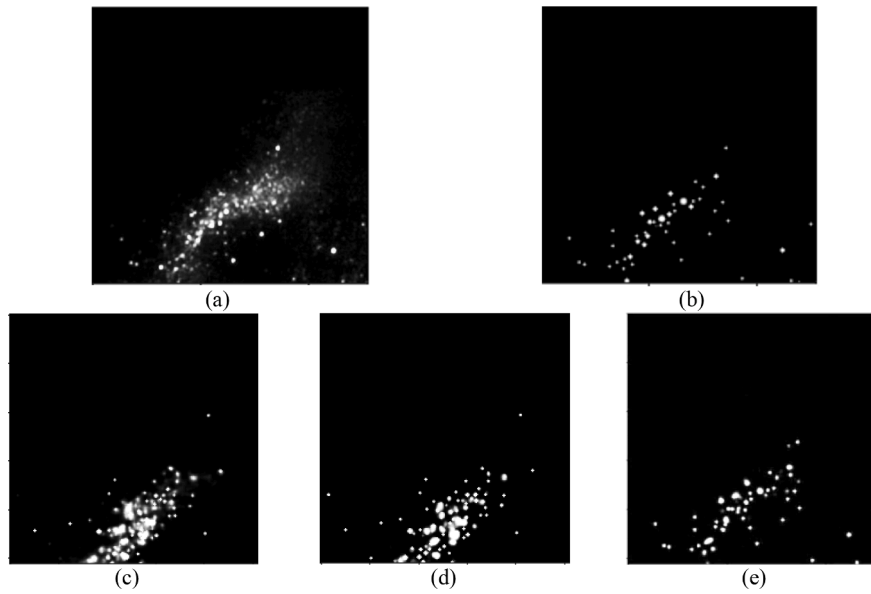
Table 2
Performance Evaluation.

Block Size	Model Architecture	Training Time/ Epoch	MSE	PSNR (dB)
64 × 64	Standard CNN	53s	0.0970	10.132
	Modified ResNet	89s	0.0132	18.79
	Modified U-Net	170s	0.0066	21.804
128 × 128	Standard CNN	72s	0.103	9.871
	Modified ResNet	101s	0.0141	18.507
	Modified U-Net	233s	0.0062	22.076
256 × 256	Standard CNN	169	0.0784	11.056
	Modified ResNet	191	0.0110	19.586
	Modified U-Net	447	0.0053	22.757

Although the standard CNN and modified ResNet performed comparably on denoising the image, multiple droplets formed single large clusters during the process. Furthermore, these two networks failed to detect the low-intensity droplets by removing them from the final output. This can be observed in the bottom right section of the image shown in Fig. 10(c) and Fig. 10(d) compared to Fig. 10(e).

The superior performance of modified U-Net over the other two networks can be attributed to its encoder-decoder structure. The downsampling part of the encoder selectively retains only the most prominent features of the image, which does not include any noise. Reconstruction of the noise-free features produces a more accurate and better-denoised image. However, it is an important task to decide how much to downsample the input image. Table 3 shows the results of the modified U-Net architecture with different bottleneck layer sizes. As we decreased the bottleneck layer dimensions from 32 × 32 to 8 × 8, MSE decreased. This is because only the most prominent features are retained, increasing the downsampling of the input image by disregarding unnecessary information. However, there is a certain threshold value for the MSE with the bottleneck size, as shown in Table 3. For the modified U-Net, the threshold value is 8 × 8. Any value below this threshold increased the MSE. In other words, on further downsampling the input image, we tend to lose even the important features of the input that causes inefficient reconstruction.

The additional skip connection between the encoder and decoder sections helped in retaining the original structure of the image, hence better and more accurate reconstruction of the droplets. The skip connection allowed some information from the encoder to bypass the bottleneck layer and join the decoder layer. This provided more information for the decoder to learn the original structure of the input image.

Our denoising algorithm was tested on the unseen original noisy Mie scattering spray image taken directly from the high-speed camera, as shown in Fig. 11(a). The trained model was used to denoise the unseen test image. The inference procedure involved splitting the input image into 8 blocks of size 256 × 256 pixels and passing them through the trained deep learning model. The denoised 8 blocks of size 256 × 256 pixels were then combined to generate the entire spray image, as shown in Fig. 11(b). The denoising algorithm took about 1.43 s on a single GPU.

Fig. 10. Image denoising result comparison. (a) Mie noisy image, (b) true denoised image, (c) standard CNN, (d) Modified ResNet, and (e) Modified U-Net.

Table 3
Comparison of loss for different bottleneck layer sizes.

Bottleneck size	MSE
4×4	0.0077
8×8	0.0042
16×16	0.0053
32×32	0.0059

The denoised images from the modified U-Net were then passed to a center predictor CNN to predict the x and y pixel location of every droplet in the image. The network was trained for 15 epochs, and the loss was optimized using the Adam optimizer. The comparison of training and validation losses with the number of epochs is shown in Fig. 12. The pixel locations were extracted by running a raster scan on the image. The performance results of the network are shown in Fig. 13 as a correlation between the true and predicted center locations.

The plots demonstrated that the predicted locations by the neural network agreed well with the true center locations. To evaluate the performance, the Average Error (AE) for the predicted x and y locations of every droplet was calculated. As shown in Eq. (4), the AE is defined as the mean of the absolute difference between the predicted location and the true location.

$$AE = \frac{1}{m} \sum |x_p - x| \tag{4}$$

where, x_p is the predicted location, x is the true location, and m is the total number of detected droplets. The test results for the AE in the predicted locations with the center detector CNN model for 3 representative images are summarized in Table 4.

5. Conclusion

In the present study, a deep learning-based denoiser and droplet location predictor were developed for the first time to denoise and predict the location of spray droplets in the light scattered Mie images. The data for the deep learning algorithm was obtained by capturing the Mie scattering images of gas turbine engine sprays at engine relevant condition. A thorough comparison of different deep learning algorithms such as standard CNN, modified ResNet, and modified U-Net was performed for the denoising tasks. The modified U-Net architecture performed the best with an MSE value of 0.0053 and a PSNR value of
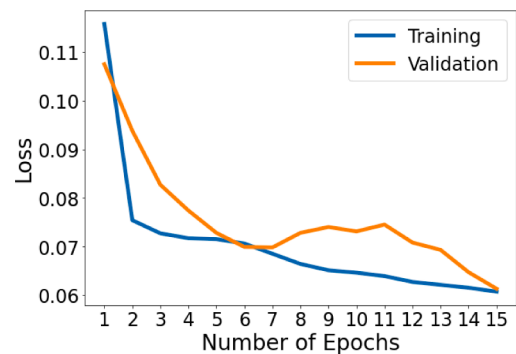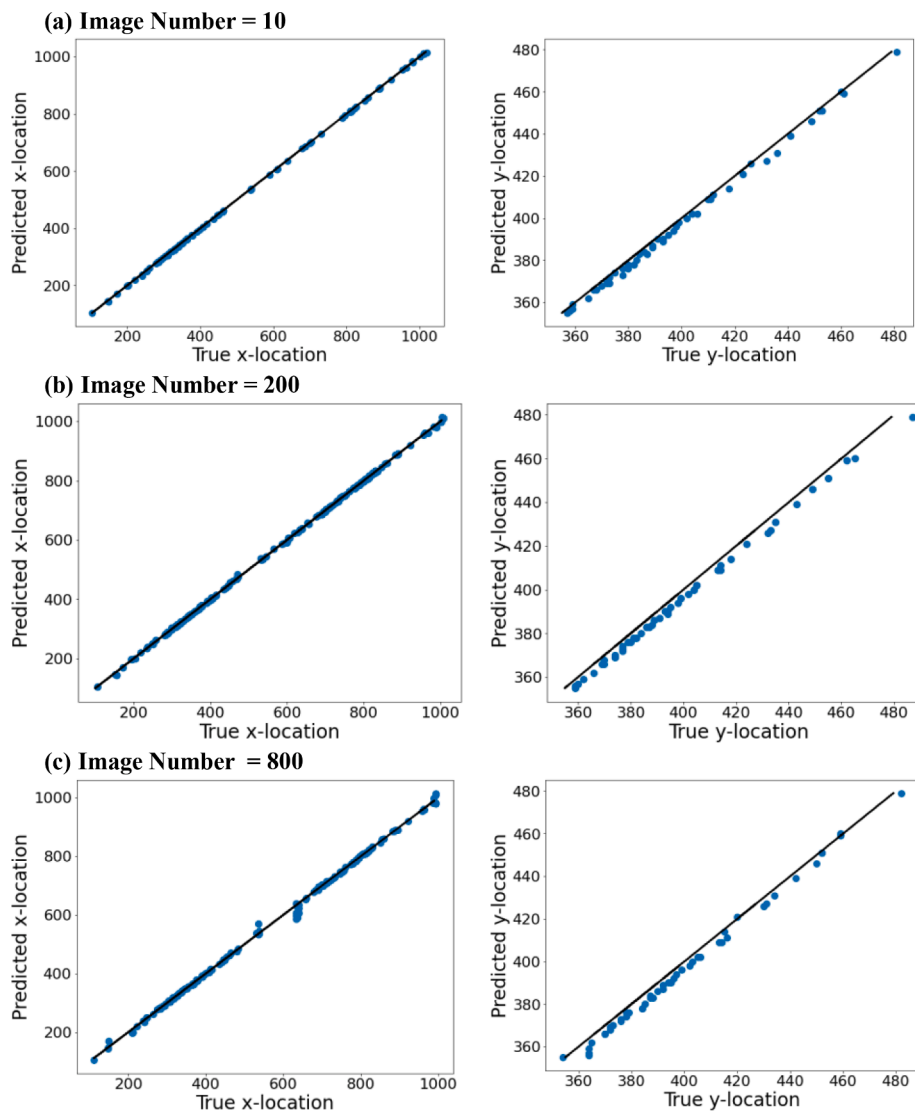

Fig. 12. Training and Validation loss vs. Epochs for Center Prediction CNN.
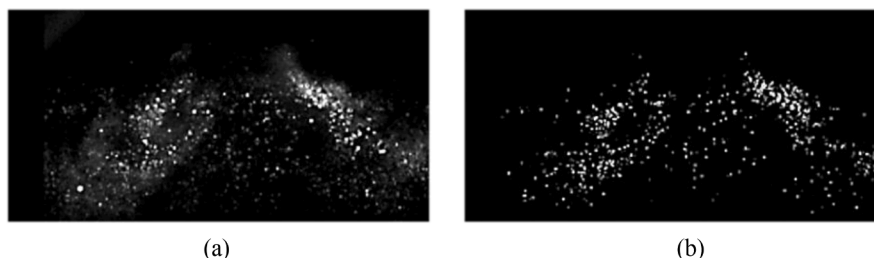
Fig. 11. Evaluation of trained model on a test image. (a) original image, (b) denoised image.

(a) Image Number = 10

(b) Image Number = 200

(c) Image Number = 800

Fig. 13. The correlation between the true and predicted locations of the droplets using the center predictor convolutional neural networks (CNN). The black solid line represents the 45° slope line.

Table 4
The average error in the predicted x and y location for 3 different Mie spray images.

Image Number	AE x-location	AE y-location
10	1.35	0.927
200	0.822	1.396
800	1.104	1.667

22.757 dB on the validation dataset. For the droplet location, a center prediction network based on CNN architecture was performed on the denoised images. The proposed network predicted the droplet center with an AE value of 1.35 pixels for x location and 0.927 pixels for y location. The uncertainty of the deep learning model is $\pm 3.4\%$ and $\pm 3.9\%$ in the predictions for the x and y pixel location of the droplets respectively. The validated deep learning model significantly reduces the time taken per image for denoising from 168 s to 2.13 s on a single GPU. This makes it a powerful tool for real-time usage on the fly.

The deep learning-based denoising algorithm automatically differentiates the droplets from noises that cause blurriness and undesired signals from non-illuminated regions. This ensures that the background noise is not misclassified as a droplet and does not interfere with the

droplet location prediction. In addition, the proposed model provides an efficient and robust method for determining the droplet center location. A well-optimized network is expected to have the capability of producing results at near real-time speeds. While the proposed model is beneficial, it has some limitations. Manual preparation of labeled datasets for training is highly time-consuming, and an image processing-based approach that uses the Canny edge detector is unable to detect and separate overlapping droplets. In addition, the proposed model is limited to round-shaped droplets. To overcome these limitations, an unsupervised or semi-supervised learning approach will be explored in future studies for generating ground truth datasets. The model can be extended to determine the quantitative information about the droplets, such as diameter, sphericity, and droplet size distribution using deep learning methods. These deep learning models will be further tested and fine-tuned utilizing the more accurate "true images" from SLIPI technique with less multiple scattering effects.

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. The authors declare that they have no

known financial interests/personal relationships which may be considered as potential competing interests.

Acknowledgments

The authors sincerely acknowledge Professor Robert P. Lucht for providing the spray images from his experimental laboratory. This research was funded by the U.S. Federal Aviation Administration Office of Environment and Energy through ASCENT, the FAA Center of Excellence for Alternative Jet Fuels, and the Environment, Project 29A through FAA Award Number 13-C-AJFE-PU-011 under the supervision of Dr. Cecilia Shaw and Dr. Anna Oldani. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the FAA.

References

[1] Lefebvre AH, Dilip RB. Gas turbine combustion: alternative fuels and emissions. CRC press; 2010.

[2] Hasti, V.R., Navarkar, A., and Gore, J.P., A data-driven approach using machine learning for early detection of the lean blowout, Energy and AI, 2021, 100099.

[3] Hasti V.R., et al., Lean blowout (LBO) computations in a gas turbine combustor, 2018 Joint propulsion conference, aiaa propulsion and energy forum, (AIAA 2018-4958).

[4] Hasti, V.R., Kundu, P., Kumar, G., Drennan, S.A., Som, S., and Gore. J.P., A numerical study of flame characteristics during lean blowout in a gas turbine combustor, 2018 Joint propulsion conference, aiaa propulsion and energy forum, (AIAA 2018-4955).

[5] Albrecht HE, Damaschke N, Borys M, Tropea C. Laser doppler and phase doppler measurement techniques. Springer science & business media; 2013.

[6] Mishra YN, Kristensson E, Berrocal E. Reliable LIF/Mie droplet sizing in sprays using structured laser illumination planar imaging. Opt Express 2014;22(4): 4480–92.

[7] Mishra YN, Kristensson E, Koegl M, Jönsson J, Zigan L, Berrocal E. Comparison between two-phase and one-phase SLIPI for instantaneous imaging of transient sprays. Exp Fluids 2017;58(9):1–17.

[8] Charalampous G, Hardalupas Y. Method to reduce errors of droplet sizing based on the ratio of fluorescent and scattered light intensities (laser-induced fluorescence/ Mie technique). Appl Opt 2011;50(20):3622–37.

[9] Domann R, Hardalupas Y. A study of parameters that influence the accuracy of the planar droplet sizing (PDS) technique. Particle & Particle Syst Characterization: Measurement and Description of Particle Properties and Behav Powders and Other Disperse Syst 2001;18(1):3–11.

[10] Powell, J.W., and Chia-fon, F.L., An investigation of multiple scattering in a hollow-cone spray. sae paper 2007-01-0648, 2007.

[11] Pastor JV, Lopez JJ, Juliá JE, Benajes JV. Planar laser-induced fluorescence fuel concentration measurements in isothermal diesel sprays. Opt Express 2002;10(7): 309–23.

[12] Domann R, Hardalupas Y. Quantitative measurement of planar droplet Sauter mean diameter in sprays using planar droplet sizing. Particle & Particle Syst Characterization: Measurement and Description of Particle Properties and Behav in Powders and Other Disperse Syst 2003;20(3):209–18.

[13] Viskup, R., Diesel and gasoline engines: intechOpen, 2020..

[14] Fantini E, Tognotti L, Tonazzini A. Drop size distribution in sprays by image processing. Comput Chem Eng 1990;14(11):1201–11.

[15] Honkanen M, Saarenrinne P, Stoor T, Niinimäki J. Recognition of highly overlapping ellipse-like bubble images. Measurement Sci Technol 2005;16(9): 1760.

[16] Strokina N, Matas J, Eerola T, Lensu L, Kälviäinen H. Detection of bubbles as concentric circular arrangements. Mach Vis Appl 2016;27(3):387–96.

[17] Zhong S, Zou X, Zhang Z, Tian H. A flexible image analysis method for measuring bubble parameters. Chem Eng Sci 2016;141:143–53.

[18] Bai X, Sun C, Zhou F. Splitting touching cells based on concave points and ellipse fitting. Pattern Recognit 2009;42(11):2434–46.

[19] Beucher S. Use of watersheds in contour detection. Proceedings of the international workshop on image processing. CCETT; 1979.

[20] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. International conference on medical image computing and computer-assisted intervention. Springer; 2015. p. 234–41.

[21] Li J, Shao S, Hong J. Machine learning shadowgraph for particle size and shape characterization. Measurement Sci Technol 2020;32(1):015406.

[22] Nobari, A.H., Khorasani-Gerdehkouhi, F., Gulam, N., and Ashgriz, N., Application of deep learning convolutional neural network for spray characterization, Proceedings of the 7th international conference on fluid flow, heat and mass transfer (FFHMT'20). November 2020.

[23] Simonyan, K., and Zisserman, A., Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556, 2014.

[24] Oktay AB, Gurses A. Automatic detection, localization and segmentation of nano-particles with deep learning in microscopy images. Micron 2019;120:113–9.

[25] Zhang, F., Zhang, Q., Xiao, Z., Wu, J., and Liu, Y., Spherical nano-particle parameter measurement method based on Mask R-CNN segmentation and edge fitting, Proceedings of the 2019 8th international conference on computing and pattern recognition. 2019, pp. 205–212.

[26] Shin D, Bokhart AJ, Rodrigues NS, Sojka PE, Gore JP, Lucht RP. Nonreacting spray characteristics for alternative aviation fuels at near-lean blowout conditions. J Propulsion and Power 2020;36(3):323–34.

[27] Shin, D., Bokhart, A.J., Rodrigues, N.S., Sojka, P.E., Gore, J.P., and Lucht, R.P., Spray characteristics of standard and alternative aviation fuels at cold-start conditions. AIAA J, accessed July 30, 2021.

[28] Shin D, Satija A, Lucht RP. Spray characteristics of standard and alternative aviation fuels at high ambient pressure conditions. Experimental Thermal and Fluid Sci 2021;130:110511.

[29] Mansour, A., Benjamin, M., Burke, T., Odar, A., and Savel, B., Hybrid atomizing fuel nozzle. U.S. patent 6547163 B1. April 15, 2003.

[30] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 2012;25:1097–105.

[31] He, K., Zhang, X., Ren, S., and Sun, J., Deep residual learning for image recognition, Proceedings of the ieee conference on computer vision and pattern recognition. 2016, pp. 770–8.

[32] Agarap, A.F., Deep learning using rectified linear units (relu). arXiv preprint arXiv: 180308375, 2018.

[33] Ioffe, S., and Szegedy, C., Batch normalization: accelerating deep network training by reducing internal covariate shift, International conference on machine learning. PMLR, 2015, pp. 448–456.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle		5. Report Date
		6. Performing Organization Code
7. Author(s)		8. Performing Organization Report No.
9. Performing Organization Name and Address		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address		13. Type of Report and Period Covered
		14. Sponsoring Agency Code
15. Supplementary Notes		
16. Abstract		
17. Key Words	18. Distribution Statement	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized