Sub Topic: Fire

13th U.S. National Combustion Meeting
Organized by the Central States Section of the Combustion Institute
March 19–22, 2023
College Station, Texas

Classification of Airborne Firebrand Combustion State Using a Convolutional Neural Network

Savannah S. Wessies[1], *Nicolas Bouvet*[1,*], *and Eric D. Link*[1]

[1]*National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, USA*
[*]*Corresponding author:* nicolas.bouvet@nist.gov

Abstract: In wildland fire scenarios, firebrand transport allows for further propagation of the fire away from the main fire front. A better understanding of the number of firebrands and various characteristics of the firebrands would provide a more accurate assessment of the hazards associated with a given firebrand shower or flow. In recent years, NIST has designed and developed the Emberometer to characterize firebrand showers in the field. This device combines two imaging techniques, 3D Particle Velocimetry and 3D Particle Shape reconstruction, to characterize the firebrand showers in both time and space. By utilizing the NIST emberometer to monitor an airborne firebrand flow in an outdoor experiment, individual firebrands can be tracked and characterized over the duration of the test. In this work, we primarily focus on the methodology for determining the firebrands' combustion state: flaming or smoldering. Based on recorded video during testing, the differences between flaming and smoldering firebrands are readily apparent to the viewer. However, to code an algorithm for this task is nontrivial. The resulting program would be highly dependent on the features and decision thresholds the programmer deems important. Additionally, with these classification algorithms, it can be difficult to determine if the optimal solution has been found. However, convolution neural networks (CNNs), a type of machine learning tool, are widely used to classify images and do not have the same issues with arbitrary programmer-specific choices and optimization. Appropriately trained CNNs allow for the rapid classification of large image data sets, which greatly reduces the time required for classification compared to human sorting. In this work, two different approaches were utilized to determine the optimal CNN. First, a lightweight CNN (CNN A-D) was developed from the ground up. Second, transfer learning was applied to a set of pretrained, previously structured CNNs (CNN E-I). The CNNs were evaluated on an unseen subset of the data for accuracy, precision, recall, and f-measure. With the lightweight, newly developed CNNs, the optimal solution, based on a parametric study, had an accuracy of approximately 92 %. In general, the pretrained CNNs had higher accuracies around 95 %. Considering the other metrics and balancing concerns of overfitting and machine resource requirements, the optimal CNN was determined to be the case with transfer learning applied to CNN F.

Keywords: Firebrands, Machine learning, Image classification, Combustion state

1. Introduction

Firebrands, or combusting particles of vegetation or structural materials carried by the wind, help propagate wildland and Wildland-Urban-Interface (WUI) fires through the generation of spot fires. Over the past two decades, post-fire investigations have steadily identified firebrands as the primary cause of a significant number of structural losses at the WUI. In a case study of the Trails community (Rancho Bernardo, CA) affected by the Witch and Guejito fires in 2007, Maranghides and Mell identified that out of 274 residences, 74 were completely destroyed with an additional

16 sustaining various degrees of damage [1]. They found that two-thirds of the destroyed homes were ignited directly or indirectly by firebrands. Cohen and Stratton led post-burn investigations of residential destruction following the 2007 Grass Valley Fire near Lake Arrowhead in the San Bernardino Mountains [2]. Out of 199 destroyed or damaged homes examined, only 6 were identified as possibly ignited (directly) by the wildfire front. Most home ignitions were attributed to firebrands, either via direct assault onto structures or via generation of spot fires subsequently spreading to homes. Similar qualitative findings were outlined by Graham et al. in their investigation of the Fourmile Canyon fires that occurred in 2010 near Boulder, CO [3]. They concluded that 83 % of the home destroyed were due to direct firebrand ignition and/or surface fire spreading to homes.

In spite of the known hazards firebrands pose to structures in the WUI, there are relatively few studies on airborne firebrand flows [4–7]. There is even less known about the thermal/combustion characteristics of the firebrands in these flows. Understanding the proportion of flaming versus smoldering firebrands is essential in order to determine how severe a firebrand exposure may be. Although some information has been reported for the firebrand temperature [6–8], there is currently no approach performing systematic distinction of airborne firebrand combustion state. The combustion state is fundamental for two reasons: i) a flame anchored to a firebrand complicates contour detection, thus compromising accurate sizing operations, and ii) flaming firebrands have been shown to elicit ignition propensities in target fuels as compared to their smoldering counterparts [9].

This paper intends to help alleviate the current limitations in the literature related to firebrand combustion state. To do so, it leverages data collected using a measurement system called the Emberometer [10]. This system was designed to characterize airborne firebrand flows and allows for motion tracking of burning particles in full 3D space and time. The device was used to study an artificially generated firebrand shower that produced mixed amounts of smoldering and flaming particles. A sub-set of 3D-tracked firebrand images was used to develop a method to classify firebrand combustion state in an airborne setting.

While the differences between flaming and smoldering firebrand images are often apparent to the human eye, the task of coding an algorithm to classify them using traditional image processing tools is non-trivial. Such an approach typically requires hand-crafted feature extraction, known to be an arduous task [11], on which the classification step would entirely depend. As both feature extraction and classification decisions bear some arbitrariness due to programmer-specific choices (e.g., important feature to be retained, decision threshold to be applied), it is usually difficult to assess if the algorithm developed is well suited to/optimized for the problem to be tackled. To circumvent those difficulties, we turn towards machine learning tools, and more specifically Convolutional Neural Networks (CNNs), which are widely utilized in the field of visual imagery that routinely uses image and video data as inputs [12]. CNNs have been shown to be particularly well suited for complex image classification problems, including in recent years, fire-related issues such as fire detection [13–15] and firebrand cluster detection in wildfires [16]. Additionally, adequately trained CNNs may perform classification tasks extremely rapidly which represents significant time savings as compared to human screening when very large datasets need to be processed (over 71,000 individual firebrand images in the current dataset). Here, we present the approach used to develop/choose the CNNs of interest, including some details on network architectures, training procedures, and selected performance metrics. To the authors' knowledge, this is the first time such an approach is leveraged to determine if individual airborne firebrands are in the flaming or

(a) (b)

Figure 1: (a) Labeled picture of Emberometer, and (b) An example of an acquired frame, camera #4 view

smoldering combustion state.

2. The Emberometer

The Emberometer system[1] is devoted to the characterization of airborne firebrand flows typically encountered in wildland/WUI fires. The system merges two optical techniques, 3D Particle Tracking Velocimetry (3D-PTV) as well as 3D Particle Shape Reconstruction (3D-PSR), in a single field-deployable device. A brief summary is provided here. Further details can be found in [10, 17, 18]. The system shown in Figure 1 is composed of four compact cameras (Sony DSC-RX10M3, $\approx$ 20 Megapixels) with large diameter built-in lenses. To enhance firebrand detection, the cameras have undergone full spectrum conversion and are equipped with infrared filters (Hoya R72) ensuring light transmission above 760 nm. A sample image is provided in Figure 1a. The cameras (and controlling electronics) are embedded in hard shells equipped with dedicated optical windows made of quartz. Shells are attached to the extremities of a collapsible X-shaped aluminum stand using custom multiaxial mounts. Using microcontrollers embedded in the camera shells and custom-developed control hubs, the mechanical initialization operations and long range system control can be conducted remotely in field settings. Additionally, video feedback to the user is provided via a 5 GHz long range, low latency HD video transmitter.

The video streams acquired by the system are rendered into high bit depth color image sequences by extracting individual video frames via a commercial software. Frame synchronization is performed using an in-house MATLAB code that also transposes the sequences into low bit

[1]Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

depth grayscale color space. Grayscale image sequences are then processed by the PTV software, and raw particle tracking outputs are subsequently handled by a series of in-house FORTRAN routines, used to gather the complete time history of a particle's 3D motion. The routines include an algorithm that identifies and patches particle broken trajectories. The firebrand combustion state analysis developed below uses firebrand high bit depth color images (background subtracted), cropped and numbered according to the final tracking outputs.

The Embrometer was deployed in an outdoor setting at the Frederick County Public Safety Training Facility in Frederick, MD. The Emberometer was set downstream of a firebrand generator composed of a centrifugal blower connected in line with cylindrical stainless-steel duct elements (see additional details in [17]). The fuel was 350 g ± 1 g of dry (moisture content < 6 %) birch/maple dowels (diameter 6.4 mm ± 0.1, length 51 mm ± 0.4 mm) with chamfered edges. The background of the test was a mix of asphalt, forested areas (mixed vegetation), and clear sky.

3. Methods

The Emberometer utilizes four different camera views to track individual airborne firebrands in 3D. Accordingly, there may be up to four separate images available for a single firebrand at a given time. The combination of the four images into a single image is referred to as a firebrand quadview. Samples of firebrand quadview images are provided in Fig. 2. A firebrand may be visually identified as: flaming on the four available quadrants (column 1), smoldering on the four available quadrants (column 2), or a combination of flaming and smoldering depending on the view considered (column 3). In all cases, firebrand images span a wide range of visual appearances, with various particle sizes, shapes, and colors. As each quadrant may be independently rated "flaming" or "smoldering," each individual view is fed independently to the CNN in the form of a RGB image with dimensions (71 x 71) pixels, for both model training and routine classification purposes. Note that not all firebrand quadviews provide useful information. If, for some reason, the firebrand has not been detected for a specific view and time step (i.e., firebrand is not visible, or is detected but the particle correspondence step has failed), the corresponding quadrant will be entirely black (see lower left quadrant in upper right quadview in Fig. 2). These quadrants are automatically identified, labelled accordingly, and subtracted from the CNN analysis.

Two approaches, each with trade offs, were considered. First, design a lightweight CNN from the ground up. The classification task only entails two categories ("smoldering" or "flaming"), so the network architecture can presumably be kept very simple, but training of the network must be performed from scratch. Second, implement transfer learning on a set of pretrained, previously structured CNNs[19]. In this case, training is facilitated by prior learning at the expense of network footprint/complexity. In the present work, both approaches are compared to identify the CNN with the best overall performance. The newly designed and previously structured CNNs were implemented using the Deep Learning Toolbox in MATLAB (2022).

The architecture for the new CNNs (CNN A-D) is summarized in Figure 3. To optimize this network, parametric testing was conducted, and variables including filter size and number of filters per convolutional layer were considered. It was found that adding additional convolutional layers had little effect on the CNN's performance. Additionally, transfer learning was implemented on five previously structured CNNs, pretrained on the ImageNet dataset [20]. The five selected networks were chosen due to availability, prevalence in the literature, and variation in architecture. The first, CNN E, consists of 5 convolution layers and 3 fully connected layers with 61 million pa-

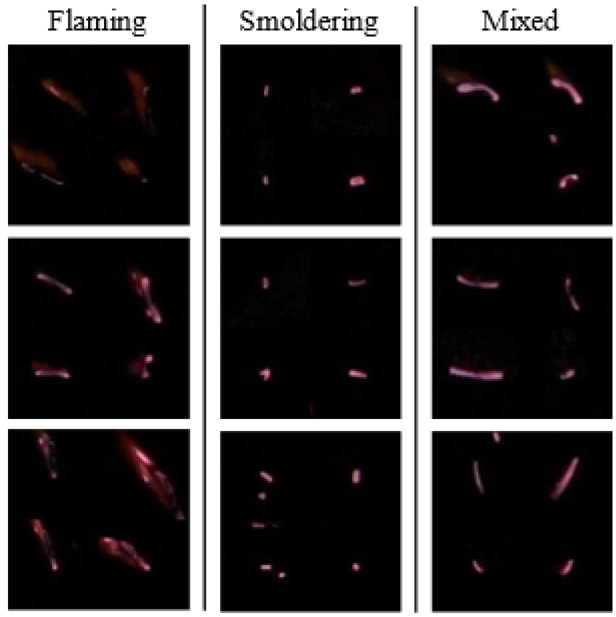

Figure 2: Quadview image samples (background subtracted) for flaming, smoldering, and mixed conditions firebrands. Each quadview represents a single firebrand at a single time step. A completely black square (see lower left quadrant of upper right quadview) indicates that the firebrand was not detected for that specific view/timestep combination.

rameters. The network utilizes overlapping pooling and local response normalization to optimize image classification accuracy. With 22 layers and 7 million parameters, CNN F had an optimal sparse structure with dense units. This network had a less traditional structure with 9 inception modules to improve computational resource allocation. CNN G followed a bottleneck design with 18 layers and 11.7 million parameters. This network uses residual mapping, skip connections, and heavy batch normalization in its architecture. Also with 18 layers, CNN H had only 1.24 million parameters. By down sampling later in the network architecture, this network had large activation maps in convolutional layers. Additionally, this network utilized 8 fire modules which consisted of squeeze layers followed by expand layers. The final CNN (CNN I) had the highest number of parameters at 138 million and 16 layers. This CNN followed the traditional, uniform architecture, and all convolutional layers had 3x3 filters.

For the purposes of training and transfer learning of the CNNs, a subset of the airborne firebrand population tracked with the Emberometer (Sect. 2) was considered. A total of 293 different particles were carefully reviewed and hand labeled as smoldering or flaming. The review process included all time steps and quadrant views available for a firebrand (excluding black quadrants as described above and quadrants for which large flames attributed to the firebrand generation device were visible). The labeling was performed on color images without background subtraction allowing for detailed dynamic examination of the firebrand silhouettes across time steps. This process resulted in 23,873 firebrand images classified as either flaming or smoldering. The majority of the hand labeled images (88.8 %) fell in the smoldering category, an observation consistent with the firebrand generation process. To ensure the trained CNNs did not become biased towards the smoldering classification, a subset of the smoldering images was selected at random. This subset

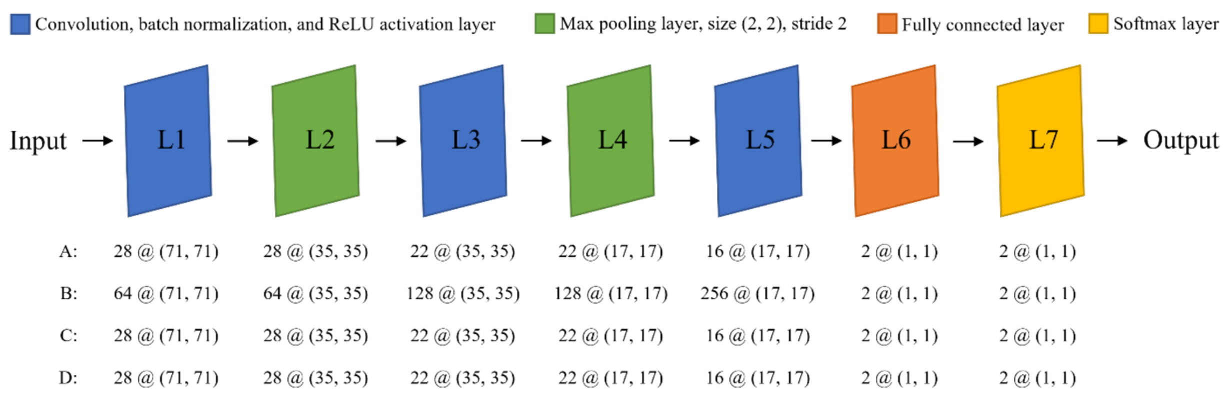

Figure 3: Architecture of the lightweight CNN developed in this work. A, B, C and D correspond to the different filter configurations specified in Table 1

had the same number of images as the flaming image subset (approx. 2,700). These hand classified images were separated randomly into training (80 %), validation (10 %), and test (10 %) datasets, which were held constant to allow for comparison between different CNNs.

The learning rate was the first parameter considered during training. Based on the validation loss trends, 0.0001 was selected as the optimal value. The number of iterations was determined by a validation patience of 5. Additionally, for the transfer learning cases, the learning rate for the weights and biases was increased by a factor of ten to decrease required training time, and the mini-batch size was varied between 5 and 15. The majority of the training of the CNNs occurred on a hexa-core platform (Intel(R) Core (TM) i7-8700 CPU @ 3.2 GHz) equipped with a NVIDIA Quadro P1000 GPU (4.0 GB). One of the pretrained networks, CNN I, seemed to somewhat exceed available resources and was trained using an octa-core platform (Intel(R) Xeon (R) W-1370P @ 3.6 GHz) equipped with a NVIDIA Quadro RTX 4000 GPU (8 GB).

3.1 Performance Metrics

Several metrics are used to compare the performances of the different CNNs implemented in this work. The most common metric is the network Accuracy, defined as:

$$Accuracy = \frac{true\ predictions}{total\ number\ of\ predictions} \tag{1}$$

The Accuracy reflects the overall success of the classification task, independently of class awareness. The Precision and Recall metrics were used for deliver class-dependent information:

$$Precision = \frac{TP}{TP + FP} \quad \text{and} \quad Recall = \frac{TP}{TP + FN} \tag{2}$$

where TP (true positive) stands for the number of cases with an identical hand label and CNN prediction within a specified class, FP (false positive) is the number of cases identified by the CNN as part of a class but actually belonging to the other (e.g., for the flaming class, the firebrand is predicted as flaming by the CNN but is actually smoldering per the hand label), and FN (false negative) is the number of cases identified by the CNN as part of the other class but actually belonging to the class considered (e.g., for the flaming class, the firebrand is predicted as smoldering by the

CNN but is actually flaming per the hand label). Both Precision and Recall can be combined in a single metric, the F-measure or F-score, hereby noted F, which represents their harmonic mean:

$$F = 2\left(\frac{Precision \cdot Recall}{Precision + Recall}\right) \tag{3}$$

Accuracy, derived for both the validation and test sets, was closely monitored to ensure overfitting of the training data did not occur. F-measures for both smoldering and flaming classes were monitored to identify any possible classification bias towards one of the categories.

4. Results and Discussion

The CNNs were trained and validated on a subset of the available flaming and smoldering firebrand images. The newly designed CNN varied in filter size and number of filters per convolutional layer. A representative subset of the Accuracy and F-measure results for these CNNs is shown in Table 1. Additional combinations of filter size and numbers were trained, but the performance metrics were lower than those presented. For these simple 3-layer CNN variations, the best Accuracy for the validation and test sets was just under 92 %. These results were obtained for the CNN with filter size 3x3 and the number of filters, from the first convolutional layer to the last, equal to 64, 128, and 256. For this case, the Precision and Recall for the flaming class were 0.93 and 0.91 whereas the Precision and Recall for the smoldering class were 0.91 and 0.93, respectively. Adding additional layers or larger numbers of filters did not elicit a marked improvement to the performance metrics.

Table 1: Comparison of the different performance metrics for the CNNs trained to classify the images as flaming or smoldering. For the newly designed CNNs in the first 4 rows, FS stands for Filter Size, and NF for Number of Filters. The first number corresponds to the number of filters in the first convolutional layer, and so on. The CNNs trained via transfer learning presented in the last five rows of the table had a mini-batch size of 5.

	CNN Architecture	**Accuracy**		**F-measure**		**Size on Disk** (MB)
		Validation	Test	Flaming	Smoldering	
Newly Designed	A. FS: 3x3| NF: 28, 22, 16	88.9	89.1	0.89	0.89	0.07
	B. FS: 3x3| NF: 64, 128, 256	91.8	92.0	0.92	0.92	1.84
	C. FS: 5x5| NF: 28, 22, 16	90.2	90.4	0.90	0.91	0.16
	D. FS: 7x7| NF: 28, 22, 16	91.2	90.8	0.90	0.91	0.23
Transfer Learning	CNN E	94.76	95.30	0.95	0.95	202
	CNN F	95.1	95.1	0.95	0.95	21.2
	CNN G	95.3	93.8	0.94	0.94	39.7
	CNN H	93.8	92.5	0.92	0.93	2.59
	CNN I	95.9	95.7	0.96	0.96	477

A distinct improvement to the performance criteria was noticed when transfer learning was applied to the pretrained CNNs. The last 5 rows of Table 1 represent the transfer learning cases, which generally performed a few percentage points better for each metric as compared to the newly designed CNN. The results presented in the table were found using a mini-batch size of 5. Training was also conducted with mini-batch sizes equal to 10 and 15, however, the results were observed to be marginally worse for every pretrained CNN. Using transfer learning on CNN H produced an improvement of the performance metrics as compared to the newly designed CNN. However, the performance was not as high as the other pre-trained networks. In general, the other transfer learning cases had validation and test accuracies around 95 %. The transfer learning on CNN G had high accuracy with the validation data, but the test accuracy was 1.5 % lower than the validation set, possibly indicating some overfitting. For the remaining networks, F-measures for both smoldering and flaming classes were found to be identical. Based on the various performance metrics and the memory requirements (those of CNN E and CNN I were approximately 10 and 22 times higher than that of CNN F, CNN I requiring additional computing resources), the CNN trained using transfer learning on CNN F was adopted as the firebrand classification model and was further used to process the leftover smoldering firebrand images (approx. 18,500) previously set apart during the training dataset generation step. An accuracy of 95.4 % was obtained, a value consistent with those reported in Table 1 for the validation and test datasets.

5. Conclusions

Leveraging machine learning tools and Emberometer-acquired firebrand images, a methodology for firebrand combustion state (smoldering/flaming) identification was presented. A set of Convolutional Neural Networks (CNN), including a custom lightweight and several pretrained networks, were trained on labeled images and individual network performances were compared. The highest accuracy resulting from the simple, newly designed CNN was approximately 92 %. In general, implementing transfer learning improved the classification accuracy compared to the lightweight CNNs. It was found that the pretrained CNN F (22 layers, 7 million parameters) offered the best compromise (based on accuracy, F-measure, and size metrics), with predictions systematically exceeding 95 % accuracy for the validation and test sets considered. The use of this CNN greatly reduces the total time required for combustion state classification and drastically reduces the necessary man hours after the initial training of the CNN. Moving forward, the combustion state classification could be used to improve size classification of tracked firebrands. Additionally, the flaming/smoldering data could help quantify the exposure hazard of a given firebrand flow. Finally, as more experiments are conducted with the Emberometer, additional data can be added to the training set leading to an increase in the accuracy of the classification and expanded applicability of the CNN.

6. Acknowledgements

The authors would like to thank Mr. Stephen Fink, Mr. Rik Johnsson, Mr. Marco Fernandez, and the Frederick County Public Safety Training Facility in Frederick, MD for facilitating the data collection. Additionally, we thank Dr. Andy Tam for providing insight regarding the machine learning approach.

References

[1] A. Maranghides and W. E. Mell, A case study of a community affected by the Witch and Guejito Fires, NIST Technical Note 1635, National Institute of Standards and Technology, Gaithersburg, MD (2009).

[2] J. D. Cohen and R. D. Stratton, Home destruction examination: Grass Valley Fire, Lake Arrowhead, California, Tech. Paper R5-TP-026b. Vallejo, CA: US Department of Agriculture, Forest Service, Pacific Southwest Region (Region 5). 26 p. (2008).

[3] R. Graham, M. Finney, C. McHugh, J. Cohen, D. Calkin, R. Stratton, L. Bradshaw, and N. Nikolov, Fourmile Canyon fire findings, Gen. Tech. Rep. RMRS-GTR-289. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station (2012).

[4] A. Tohidi, M. Gollner, C. Alfano, and Q. Quarles, Computer vision techniques for firebrand detection and characterization, 16th international conference on automatic fire detection and the suppression, detection and signaling research and applications (SUPDET 2017) (2017), pp. 12–14.

[5] M. Almeida, D. Viegas, and J. Raposo, Analysis of firebrand release on the spot fire mechanism, Advances in Forest Fire Research 2014, ed Viegas D.X. (2014) 61–72.

[6] A. Filkov and S. Prohanov, Particle tracking and detection software for firebrands characterization in wildland fires, Fire Technol. 55 (2019) 817–836.

[7] S. Zen, J. C. Thomas, E. V. Mueller, B. Dhurandher, M. Gallagher, N. Skowronski, and R. M. Hadden, Development of a field deployable firebrand flux and condition measurement system, Fire Technol. 57 (2021) 1401–1424.

[8] A. Filkov, S. Prohanov, E. Mueller, D. Kasymov, P. Martynov, M. El Houssami, J. Thomas, N. Skowronski, B. Butler, M. Gallagher, et al., Investigation of firebrand production during prescribed fires conducted in a pine forest, Proc. Combust. Inst 36 (2017) 3263–3270.

[9] A. Fernandez-Pello, C. Lautenberger, D. Rich, C. Zak, J. Urban, R. Hadden, S. Scott, and S. Fereres, Spot fire ignition of natural fuel beds by hot metal particles, embers, and sparks, Combust. Sci. Technol. 187 (2015) 269–295.

[10] N. Bouvet, E. D. Link, and S. A. Fink, Development of a new approach to characterize firebrand showers during Wildland-Urban Interface (WUI) fires: a step towards high-fidelity measurements in three dimensions, NIST Technical Note 2093, National Institute of Standards and Technology, Gaithersburg, MD (2020).

[11] W. Rawat and Z. Wang, Deep convolutional neural networks for image classification: A comprehensive review, Neural Comput. 29 (2017) 1–98.

[12] S. Geetha, C. Abhishek, and C. Akshayanat, Machine vision based fire detection techniques: a survey, Fire Technol. 57 (2021) 591–623.

[13] C. Chaoxia, W. Shang, and F. Zhang, Information-guided flame detection based on faster r-cnn, IEEE Access 8 (2020) 58923–58932.

[14] Y. Cheng, K. Chen, H. Bai, C. Mou, Y. Zhang, K. Yang, Y. Gao, and Y. Liu, An efficient fire detection algorithm based on multi-scale convolutional neural network, Fire Mater. 46 (2022) 981–992.

[15] Y. Wang, L. Dang, and J. Ren, Forest fire image recognition based on convolutional neural network, J. Algorithms Comput. 13 (2019) 1–11.

[16] H. Pan, D. Badawi, C. Chen, A. Watts, E. Koyuncu, and A. E. Cetin, Deep neural network with Walsh-Hadamard transform layer for ember detection during a wildfire, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022), pp. 257–266.

[17] N. Bouvet, E. D. Link, and S. A. Fink, A new approach to characterize firebrand showers using advanced 3D imaging techniques, Exp. Fluids 62 (2021) 1–14.

[18] N. Bouvet, E. D. Link, S. A. Fink, and E. D. Kuligowski, On the use of time-resolved three-dimensional diagnostics to characterize firebrand showers in the WUI, Advances in Forest Fire Research 2018, ed Viegas D.X. (2018) 826–836.

[19] *MATLAB version 9.12.0.1884302 (R2022a)*, The Mathworks, Inc., Natick, Massachusetts, 2022.

[20] *ImageNet*, Last Modified March 11, 2021, URL: `https://www.image-net.org/`.