Monitoring of combustion regimes based on the visualization of the flame and machine learning

M P Tokarev¹, S S Abdurakipov¹, Gobyzov O A¹, A V Seredkin¹,², V M Dulin¹,²

¹ Kutateladze Institute of Thermophysics, Siberian Branch of RAS, 1 Lavrentyev Avenue, Novosibirsk, 630090, Russia
² Novosibirsk State University, 2 Pirogova Street, Novosibirsk, 630090, Russia

E-mail: mtokarev@itp.nsc.ru

Abstract. Development of modern intelligent monitoring and control systems in energy, allowing reducing the level of harmful emissions and energy intensity production is relevant. In the scientific literature usage of new efficient machine learning techniques for automatic extraction of features for the classification of combustion regimes is insufficiently covered. In this paper we describe a method for determining combustion regimes based on images of flames. To determine the combustion regimes, a convolutional neural network is trained using labeled data. It is shown that in the gas flame colour images the accuracy of the classification of regimes is up to 98%. Results of the convolutional neural network are compared to classification results of various linear models.

1. Introduction

Monitoring tools play an important role in maintaining or optimizing the combustion of fuels [1]. Analysis of the flame images can be used as an additional source of data, since the observed flame characteristics, such as its size, shape, front position, brightness, emission spectrum, and the nature of the variation of these parameters over time, carry a large amount of information on combustion regimes. A number of works in this direction refer to an attempt to formulate a method or an algorithm for estimating the parameters of the combustion process on the basis of an analysis of individual characteristics of the flame images. For example, in [2], [3] the authors showed the relationship between the fuel/air ratio and the temperature in the swirling flame during the combustion of coal and biofuels with the distribution of brightness and its pulsations in the images. An example of the use of information on the spectral composition of the radiation (flame color), directly related to the chemical reactions, can serve a method of planar pyrometry, described, for example, in [4], and many other works.

Another approach to monitoring, in addition to directly determining the dependence of individual characteristics of the flame image on the combustion parameters, can be a comparison of the flame image and the combustion regime, that is, the characteristic range of parameters in which this regime is implemented. To implement this approach, a problem of flame image classification based on a multidimensional feature vector should be solved. In the scientific literature to date the use of methods of automatic selection of features for the classification of combustion regimes has not been described.

For classification using the multidimensional feature vector as input data, there is a wide range of methods such as the support vector machine, a neural network (for example, a multilayer perceptron),...
gradient boosting over linear models or decision trees, etc. [5]. The effectiveness of such methods for solving computer vision problems depends on the expert knowledge of a researcher to create a characteristic description of images. It is important to note that the use of each pixel as an independent input parameter is difficult due to the fact that the number of input variables of the model is also extremely large. In this regard, dimension reduction methods, such as principal component analysis or clustering methods for combining groups of similar pixels are applied to images.

Another effective way to solve the problem of image classification is using convolutional neural networks (CNN) [6]. This type of neural network algorithms is successfully used for a wide class of tasks, including handwriting recognition, face recognition in images and many others. The basic operations for convolutional neural networks are convolution and subsampling (pooling) [7]. The use of convolutional layer reduces the number of parameters many times over the fully-connected layer, but requires more hyperparameters to be determined before the training procedure.

The standard approach to CNN learning is learning from labeled data (supervised learning). However, unsupervised learning can be also used to solve the problem. Generating a sample of labeled data for neural network training is considered laborious, but for the analysis and classification of flame images this approach is feasible, since the creation of the dataset is possible by specifying the values of the parameters of a burner device in a pre-known range corresponding to a certain combustion regime.

In this paper, we consider a method of automatic determination of the combustion regime by images of a swirling flame on the example of a laboratory gas burner device. To determine the combustion regimes, a convolutional neural network trained on a labeled dataset is used, and the result of its work is compared with the results of classification using several linear models.

2. Experimental setup

The flame images were collected on a laboratory gas burner setup, which is an open circuit connected to cylinders with combustible gas and a pipeline with compressed air. A calibrated gas flow control system was used to supply air and fuel. The fuel was mixed with an oxidizer in a long pipe, the output flow was formed by an axisymmetric tapered nozzle with a swirl generator. The swirling rate of the flow was varied using swirlers with different angle of inclination of blades. A more detailed description of the setup is given in [8].

The Reynolds number of the jet based on the average flow rate and air viscosity was 5,000. The regime was determined by a combination of two parameters: the swirling rate of the flow S and the ratio of fuel/air in the mixture Φ. In total, thirteen different combustion regimes of the propane-air mixture were registered. Registration was performed using a colour camera with a spatial resolution of 2 Mpixels for each regime, about $10^4$ images were recorded on a dark background from one angle. Examples of the visualization images and the parameters of the combustion regimes are given in figure 1.

3. Used methods of clustering and classification

To obtain a basic assessment of the quality of the problem solution, a group of linear methods of machine learning with and without labeling was implemented and tested. We have used several linear classification methods: k-nearest neighbors (kNN), support vector machine (SVM) with linear kernel function, and logistic regression algorithm. Besides, several clustering-based methods of unsupervised learning, including a method of K-means and agglomerative hierarchical clustering were tested [5]. Multi-class classification by SVM and logistic regression methods were implemented using the "one-vs-all" strategy, which is to train a separate binary classifier for each class. For the application of these algorithms, the dimension of the data was previously reduced by decomposing the ensemble of vector representations of images by the principal component analysis (PCA) - the first 100 components of the decomposition were used as the input feature vector for the tested algorithms and contained more than 95% of the dispersion of the raw input data. The number of components was chosen empirically, so that further increase in their number did not lead to a significant change in the results of clustering.
Hyperparameters of linear classification models were optimized using the procedure of stratified three-fold cross-validation [5].

Figure 1. Examples of the flame images in different combustion regimes.

The obtained results demonstrated a number of problems inherent in the methods of unsupervised learning for the classification problem. As an example, in figure 2 the results of classification of 100 main components of the regimes by the K-means method are shown. The visualization in figure 2 was performed using the method of visualization of multidimensional variables t-SNE (t-distributed stochastic neighbor embedding) [9]. In the figure 2, the point cloud number is the corresponding combustion regime number, and the color of the points is the class assigned by the algorithm. It is seen that the algorithm combines several different, but visually similar combustion regimes into one cluster, at the same time splitting other more variable regimes into several clusters, which leads to low classification accuracy. Machine learning algorithms on labeled data have shown higher classification accuracy (at 88-89%) with a small standard deviation in cross-validation on different subsets (table 1).

Figure 2. Classification of images using the K-means method (unsupervised). t-SNE visualization in reduced-dimensional space on 100 PCA components. The t-SNE components are placed along the coordinate axes.
Table 1. Accuracy of linear models.

| Classification model               | Cross-validation accuracy | Cross-validation standard deviation |
|------------------------------------|---------------------------|-------------------------------------|
| Agglomerative clustering           | 27.91%                    | 1.55%                               |
| K-means                            | 27.69%                    | 1.67%                               |
| SVM                                | 89.13%                    | 0.51%                               |
| Logistic regression                | 89.16%                    | 0.52%                               |
| k-nearest neighbours (kNN)         | 88.57%                    | 0.55%                               |

A small standard deviation indicates the stability of the algorithm when classifying new data sets. Note that the selection of a different, more relevant method of forming the feature vector can significantly improve the result. For example, to generate the feature vector from images, it is possible to use a pre-trained convolutional neural network (Inception, ResNet, etc.).

4. CNN model description

The neural network was designed to solve the problem of classification of flame images obtained from a video stream of the digital camera. The proposed method of classification is based on the classical convolutional neural network that combines convolutional layers and subsampling layers. The CNN architecture, that is, the layers used and their sequence, was chosen by analogy with the AlexNet network architecture (table 2), which demonstrated a high result in the international ImageNet ILSVRC image classification competition in 2012 [10]. The number and dimension of layers were reduced to achieve the higher speed required to classify images from the on-line video stream. The data in the neural network is transmitted sequentially from the input layer to the output layer through hidden layers. First, the RGB image goes through two convolution operations in a row with a 3x3 kernel. A similar operation is applied after the fourth layer. The subsampling layers reduce the feature maps resolution, while retaining their structure, by selecting the maximum value with a 3x3 transformation kernel and step of 2. The dropout layers regularize and prevent neural network from overfitting [11]. The output of the last fully connected layer is fed to the final Softmax layer, which calculates the probability of the image belonging to the class. For convolutional and fully connected layers of the network, the activation function "linear rectifier" ReLu [6] was used, the calculation of which is much less resource-intensive than the calculation of the sigmoidal function or hyperbolic tangent.

Table 2. The configuration of the convolutional neural network model.

| Layer (type)                     | Output Shape | Param #   |
|----------------------------------|--------------|-----------|
| input_1 (InputLayer)             | 64, 64, 3    | 0         |
| conv2d_1 (Conv2D)                | 64, 64, 16   | 448       |
| conv2d_2 (Conv2D)                | 64, 64, 16   | 2320      |
| max_pooling2d_1 (MaxPooling2D)   | 32, 32, 16   | 0         |
| dropout_1 (Dropout)              | 32, 32, 16   | 0         |
| conv2d_3 (Conv2D)                | 32, 32, 8    | 1160      |
| conv2d_4 (Conv2D)                | 32, 32, 8    | 584       |
| max_pooling2d_2 (MaxPooling2D)   | 16, 16, 8    | 0         |
| dropout_2 (Dropout)              | 16, 16, 8    | 0         |
| flatten_1 (Flatten)              | 2048         | 0         |
| dense_1 (Dense)                  | 512          | 1.049.088 |
| dropout_3 (Dropout)              | 512          | 0         |
| dense_2 (Dense)                  | 13           | 6.669     |
| Total params: 1.060.269           |              |           |
The cross-entropy function was used as a loss function, and the share of correct answers (accuracy) was used as a quality metric. Additional metrics were precision and recall for each class, characterizing the share of correct operations in the total number of operations and false misses of the algorithm, respectively. The adaptive inertia method (Adam) was used to optimize the neural network weight vector. Adam is much less likely to "stuck" in the local minima of the loss function when searching for the optimum and shows the best results with minimal tuning of parameters. Like other hyperparameters, Adam parameters (the learning rate and learning rate decay) were selected for the cross validation by grid search parameters.

The full dataset consisted of 39,000 labeled RGB-images of size 1920x1080 pixels – 3,000 images for each of the 13 regimes. The dataset was divided into 3 sub-samples: training - 56% of images, validation - 14% and test subset - 30%. When splitting the dataset, just as for linear methods, mixing was used, while the balance of classes in each sub-sample was retained. The resolution of the images was reduced to 64x64 RGB-pixels to reduce the learning time. This size of the input images was chosen empirically, based on the trade-off between the classification accuracy and training and prediction time. The convergence of the CNN training with the selected hyperparameters was achieved at 25 epochs (figure 3). Learning of the final model took 24.2 min, average classification time per image was 5.1 ms. The hyperparameters optimization procedure took 40.8 hours. Algorithms for data preprocessing, classification, and construction of the CNN model were implemented in the Python programming language using NumPy, SciPy, Pandas, Scikit-Learn, OpenCV, Theano, Tensorflow, and Keras libraries. For data processing multiprocessor calculations on one computer (part of a computing cluster), on the basis of two processors Intel Xeon E5 v2, 8 x DDR3 16GB were used.

![Figure 3](image)

**Figure 3.** The dependence of the cross-entropy loss function of the algorithm and the model quality (accuracy) metric on the number of epochs.

### 5. CNN classification results

In figure 4 a classification error matrix is shown for 13 classes. The average accuracy of classification in all regimes on the deferred sample was 97.9%. The presented confusion matrix shows that an increased error occurs when classifying modes 0, 7 and modes 3, 4, 8. These modes have also been classified with an error when using linear methods, and as it is seen from the above figures (see figure 1), are hardly distinguishable visually.

However, even for these regimes, the quality of classification using CNN is much higher than the average quality of classification using linear methods, and for modes with high variability (modes 1, 11, 12) the proposed model demonstrates the confident classification with an accuracy of 99%.

### 6. Conclusion

The obtained results show that the use of the convolutional neural network pre-trained on the labeled data allows classifying the combustion regimes according to the images of flame with a high average accuracy of up to 98%. The accuracy of the neural network greatly exceeds the accuracy of the linear machine learning algorithms. It is found that algorithms based on object clustering (without labeling) tend to combine visually similar regimes into one cluster, splitting other more variable regimes into
several clusters, which significantly reduces the classification accuracy. However, the results of image clustering can be used as predictors in higher-level models. The resulting classification accuracy of the neural network of images with high variability allows expecting the effectiveness of the method in applied problems, such as monitoring of coal-fired boilers operation.

Figure 4. Confusion matrix and view of the flame regimes with less accurate prediction.

Acknowledgments
The work was supported by the Ministry of Education and Science of the Russian Federation (grant Agreement No. 14.604.21.0162).

References
[1] Sheishenov Z O, Borzov S M and Kozik V I 2009 Thermal engineering 56 № 3 254-8
[2] González-Cencerrado A, Peña B and Gil A 2012 Applied Energy 94 375–84
[3] González-Cencerrado A, Peña B and Gil A 2015 Fuel Processing Technology 130 299–310
[4] Draper T S, Zeltner D, Tree D R, Xue Y and Tsiava R 2012 Applied Energy 95 38–44
[5] Witten I, Frank E, Hall M and Pal C 2017 Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques (San Francisco: Morgan Kaufmann) p 654
[6] Hagan M T, Demuth H B, Beale M H and De Jesus O 2014 Neural Network Design 2nd Edition (USA: Martin Hagan) p 800
[7] Aghdam H and Heravi E 2017 Guide to Convolutional Neural Networks (Springer) p 282
[8] Alekseenko S V, Dulin V M, Kozorezov Y S, Markovich D M, Shtork S I and Tokarev M P 2011 Flow, turbulence and combustion 87 № 4 569-95
[9] Van der Maaten L J P and Hinton G E 2008 Journal of Machine Learning Research 9 2579-605
[10] Krizhevsky A, Sutskever I, and Hinton G 2012 Advances in neural information processing systems 25 1097-105
[11] Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R 2014 J. Machine Learning Res. 15 1929–58