Development of algorithms for diagnostics as well as control of biomass gasification and combustion processes

Laura Yesmakhanova[1*], Waldemar Wójcik[2], Seitzhan Orynbayev[3], Indira Shedreyeva[4], Martebe Zhankuanyshev[5]

[1,3,4,5]*Taraz Regional University named after M.Kh.Dulaty, Taraz, Kazakhstan.*
[2]*Politechnika Lubelska Institute of Electronics and Information Technology, Lublin, Poland.*

(*Email: yesmakhanovalaura@gmail.com*)

Abstract

This study examines the development of algorithms for diagnosis as well as the control of biomass gasification and combustion processes. The scientific work is grounded in theory and research findings from both laboratory stations and industrial facilities that use biomass as fuel, typically co-fired with coal after pre-crushing. The practical significance of the work is that biomass was converted into gas or solid form to provide additional fuel during combustion at laboratory stations. The work specifically focuses on diagnosing complex processes, including biomass gasification and its co-combustion in solid form. The experimental studies have been carried out with a deep theoretical analysis, and the material presented in the study contains concise results concerning the diagnosis and control of gasification and combustion processes, both at laboratory and industrial scales. The research is carried out on laboratory and industrial scales using research equipment, including the flame control system, which allowed obtaining the measurement results necessary to analyze various diagnostic algorithms, which were an element necessary to develop the control of a complex object, using the process of co-combustion of coal dust and biomass as an example. The obtained results, both from laboratory bench tests and industrial conditions, enabled us to develop robust control using hybrid neural network structures. Scientific novelty–using robust control algorithms, the efficiency of the combustion process of both coal dust and co-combustion of coal dust with biomass is increased. To achieve this goal, flexible control of the complex process is used.

Keywords: Biomass, Borescope, Combustion process, Control, Diagnostics, Gasification process, Neural networks.

1. Introduction

In the context of the growing energy crisis, the effective utilization of renewable energy sources, which includes biomass, is becoming significant. It can be utilized after some transformation in solid, liquid, or gaseous form in the combustion process, which is a source of heat and electricity generation [1].

During the times of rapidly progressing climate change, i.e., global warming, the rational utilization of energy resources, including renewable energy sources, is becoming more and more crucial. Obtaining energy through combustion process imposes significant limitations on its operation due to strict environmental regulations. As a consequence, carrying out such a process is a very difficult task, and the process itself is complex.

Thermal processes require both well-developed diagnostics and appropriate control algorithms that are robust to frequently changing environmental conditions [1, 2].

The main step in utilizing biomass for energy production is its refining (transforming) into heat, gas, or oil. Until the 19th century, the most popular combustion was biomass, mainly wood. Over the years, further methods for its conversion have been developed (Figure 1).
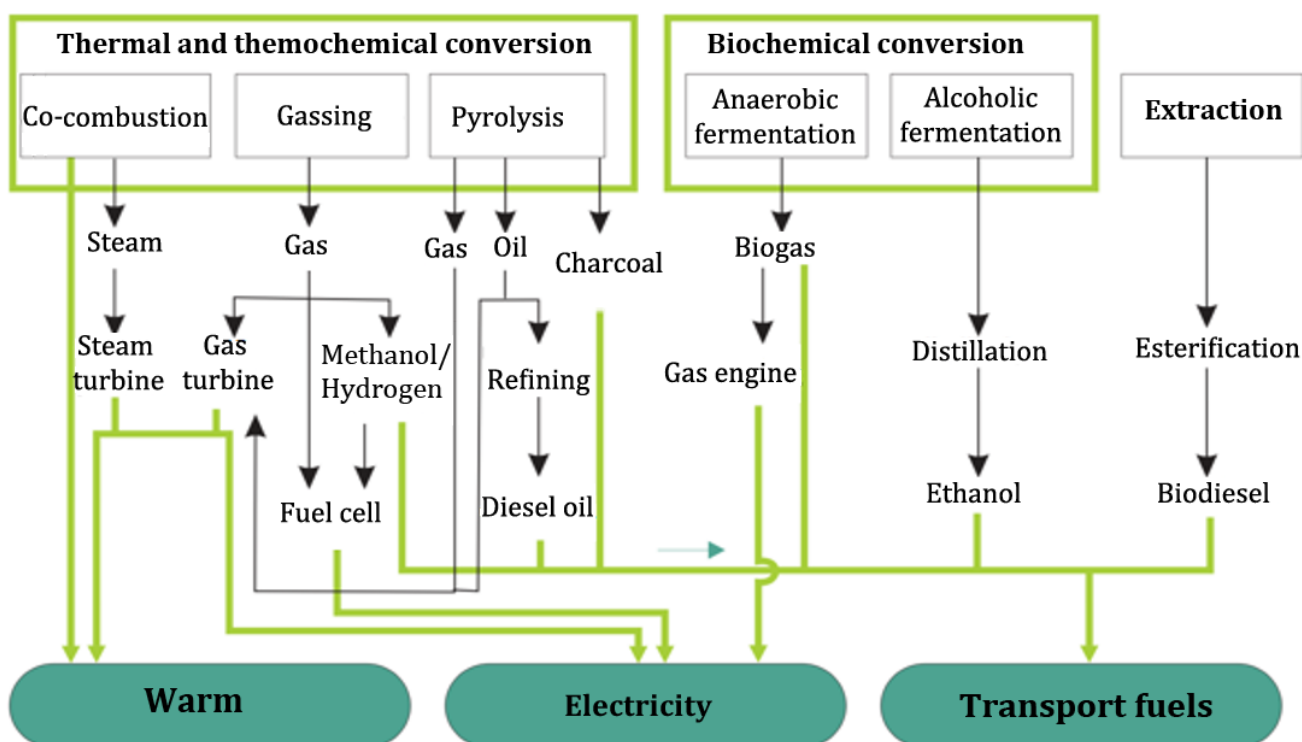

Figure 1.
Opportunities for biomass conversion into electricity, heat and transportation fuels.

For biomass combustion, the earliest furnaces were grate furnaces, followed later by fluidized bed furnaces. The thermal output of boilers with these furnaces usually does not exceed 50 MW. Table 1. summarizes the characteristics of biomass thermal processing for industrial purposes [3].

Table 1.
Characteristics of biomass thermal processing for industrial purposes.

Conversion technology	Base fuel	Fuel parameters		Achievable thermal efficiency
		Fragmentation	Humidity	
Biomass combustion in furnace power boilers	Sawdust, bark, chips, wood dust	0.5÷5 cm	10÷50%	20÷50 Megawatt (Exceptionally 4÷300 MW)
Biomass combustion in fluidized bed boilers	Various biodegradable wastes	< 5 cm	< 60%	~25 MW (Exceptionally up to 300 MW)
Co-combustion of biomass with coal in power boilers	Sawdust, bark, wood chips, wood dust	0.5 cm	< 25%	All medium and large efficiency boilers.
Gasification of biomass in a fixed bed	Wood chips, poultry manure, sewage sludge	0.5÷10 cm	< 20%	up to 50 MW

By the way, the co-combustion of biomass with coal is becoming more and more important in commercial energy. It can be concluded that direct co-combustion may be an acceptable way to replace coal with biomass. Biomass gasification on an industrial scale is primarily used in the wood or pulp and paper industries, i.e., where significant quantities of biomass waste, e.g., sulphite wastewater, need to be utilized. The construction of small biomass gasification plants is also a priority. Usually the construction of these facilities prioritizes local power generation.

Biomass, specified as one of the main sources of renewable energy, is increasingly used in electricity and heat production. The rational use of this specific fuel can bring significant environmental and economic benefits. Currently, the energy sector primarily uses biomass for its combustion in small boilers or for co-combustion with coal in existing power boilers [2]. The biomass gasification process is being intensively researched, and new pilot plants are being built, which make it possible to significantly increase the share of biomass gasification technology in electricity and heat production, especially in distributed systems. One of the main advantages of the gasification process is the possibility to use the process gas after its pre-treatment as fuel in gas-fired cogeneration systems, i.e., systems producing electricity, heat, and/or cold in combination (so-called cogeneration and trigeneration systems). These systems are characterized by, among other things, high efficiency, often exceeding 90%, low emissions of harmful substances, and the possibility of using poor-quality waste fuel [3].

The gasification process can be broadly defined as a series of chemical reactions combined with thermal processes to produce combustible gas from the organic matter of the fuel. This process takes place at elevated temperature and in the presence of a gasifying agent, which is most often air, oxygen, steam, or carbon dioxide [4]. The chemical composition of the gas produced in the gasification process depends on the following process factors:

- Temperature and pressure of gasification,
- The chemical composition of the feedstock, i.e., the gasified fraction and the gasifying agent,
- Determine the type of reactor and the duration of gas residence in each zone.

As the gasification temperature increases, the share of combustible components such as CO and H2, in the process gas increases, but at the same time the equilibrium share of the most caloric component of the process gas, CH4, decreases. Conducting the gasification process at the highest possible temperatures is associated with obtaining a gas with a lower organic content [5].

Internal combustion engines and gas turbines need to carry out the gas combustion process at the highest possible pressures. Such conditions favor the formation of methane and therefore the production of gas with a higher calorific value.

A crucial factor in producing process gas with the best energy value is the use of a gasifying agent containing the least amount of nitrogen, such as pure oxygen, a mixture of air and oxygen, or water vapor. Increasing the proportion of water vapor increases the amount of hydrogen and methane in the synthesis gas [6].
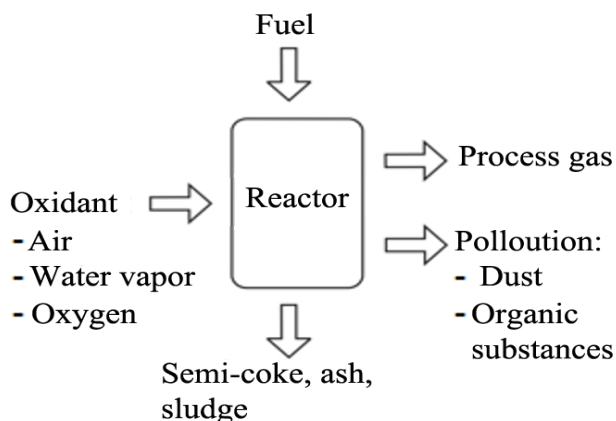

Figure 2.
General scheme of the gasification process.

The main product of the gasification process is a process gas, also called generator gas. The components of this gas are carbon monoxide, carbon dioxide, hydrogen, methane, and nitrogen. The gasification process also produces solids in the form of coke, ash, and slag, as well as liquids in the form of various types of tar and post-gasification water. Figure 2 shows a simplified general scheme of the gasification process.

The gasification process should include only high-temperature endothermic chemical reactions, but in technical practice the term refers to the entire cycle of processes used to produce combustible gas. In a biomass gasification reactor, the following process stages can be distinguished [2, 7]:

- Drying-this stage involves the evaporation of humidity and the heating of fuel particles;
- Pyrolysis-this stage involves the thermal decomposition of the organic part of the fuel into gaseous, liquid, and solid products;
- Gasification is a process in which endothermic reactions are carried out in the absence of oxygen; the main product of this process is combustible gas;
- Combustion-at this stage, unreacted fuel particles and pyrolysis products are partially burned. This process is carried out mainly to provide heat for endothermic gasification reactions.

Many methods, under different pressure and temperature conditions, can carry out gasification, producing products of different chemical composition. Depending on its calorific value, we can divide process gas into three main groups:

- Gas with low calorific value-4÷6 MJ/m3n (gasification with air or vapor-air mixture);
- Gas with average calorific value-12÷18 MJ/m3n (gasification with oxygen and steam);
- Gas with high calorific value-32÷36 MJ/m3 n (hydrogenation).

Low-heat gas is used for direct combustion or as fuel for engines; medium-heat gas can be used as a feedstock to produce chemicals such as ammonia or methanol [8].

Since it is expensive to use oxygen for gasification, air is used on a scale of up to about 50 MW. The disadvantage of this system is that nitrogen dilutes the produced gas and reduces its calorific value to 4÷6 MJ/m3 (for comparison, the calorific value of natural gas is 36 MJ/m3). Gasification with oxygen leads to an increase of calorific value up to 10÷15 MJ/m3, and with water steam, up to 12÷20 MJ/m3. In contrast to the reaction with oxygen or air, the reaction of coal with water vapor is strongly endothermic and requires a significant amount of heat input at a temperature of at least 750°C and effectively above 850°C [9].

The total efficiency of biomass conversion into energy in pyrolysis and gasification processes is 75÷80% [9].

2. Research Methods

The Materials and Methods should be described with sufficient detail to allow One of the most popular methods of reducing greenhouse gas emissions, especially CO2, is the use of renewable fuels, including biomass. Biomass combustion is more beneficial to the environment than fossil fuel combustion because the content of harmful elements (mainly sulfur and nitrogen) is lower in biomass than in conventional coal. Moreover, plant biomass has a theoretically more favorable carbon dioxide balance than fossil fuels because it absorbs some of it during photosynthesis as it grows. Co-combustion of biomass with coal is a relatively cheap alternative to obtain thermal energy for electricity generation for existing power boilers, which can be utilized after appropriate adaptation of existing units [10].

However, co-firing of biomass and coal has some disadvantages, such as lower combustion efficiency, increased amount of slag, and corrosion of steel elements of the boiler. Physical and chemical parameters of biomass are not constant in time and, furthermore, differ from those of coal. Together with the fuel, a certain uncontrolled amount of air always enters the combustion chamber, which changes the normal (effective) value of the excess air ratio. Consequently, conditions change, and the fuel is no longer fully combusted. This can unpredictably change the state of the process and consequently increase the amount of pollutants emitted into the atmosphere. Therefore, it is necessary to keep the process in the desired state. This requires the use of a system that reads the process as non-invasively and without delay as possible [11].

There is no method that can directly quantify the parameters that are indicative of the combustion process quality taking place in a single burner. Therefore, these methods are indirect, primarily involving optical and acoustic techniques. While these are non-invasive methods, they provide spatially selective information about the ongoing combustion process with little or no delay. Many of these methods utilize various characteristics of the flame's optical emission as a source of information about the combustion process. For this purpose, both narrowband and broadband sensors operating in the ultraviolet, visible, and infrared ranges are used. The diagnostic methods developed to date relate the measured optical signals to important combustion parameters concerning flame stoichiometry, pollutant emissions, flame stability, fuel properties, operating conditions, and accident conditions. Optical data processing extracts certain temporal characteristics of the signal, including its mean intensity and standard deviation, frequency characteristics, and other frequency-dependent parameters [12].

In addition to the signal derived from the optical emission of the flame, flame images are also used as a source of information about the combustion process in combination with various computerized data processing techniques. In recent years, there has been a rapid development of machine learning algorithms, especially deep learning methods. Many of them work very well in the process of recognizing images, including flames. For example, convolutional neural networks provide high accuracy (94.4%) of fire detection based on videos recorded by a facility monitoring system. Combustion diagnostics also employ deep learning algorithms.

Researchers attempted to solve the problem of identifying undesirable states of the combustion process. Each of these states is characterized by a different value of excess air ratio. If this parameter is changed, the combustion conditions change, and the fuel may not burn completely. Detection of an undesirable condition may indicate improper operation of the furnace system. Information about the occurrence of a new undesirable condition would allow the operator or control system to restore the normal condition. The study obtained this information by analyzing and classifying the flame image using both traditional and deep learning methods [13]. The study also compared both learning methods, including in terms of combustion state detection performance. It is worth emphasizing the use of deep learning methods because it is a new approach in research on diagnostics of co-combustion of hard coal and biomass. No known research results have utilized deep learning methods for this purpose. The diagnosis method based on deep learning significantly simplifies the whole model-building cycle since the image analysis and data preprocessing steps are eliminated compared to traditional learning algorithms. Deep convolutional networks are able to reveal the intrinsic characteristics of individual combustion states based on raw data obtained from images while providing high classification efficiency [14].

The material used in this paper was collected by the staff of the Department of Electronics and Information Technology of the Lublin Technological University in the course of experiments conducted at the Institute of Power Engineering in Warsaw. This foundation has led to the writings of numerous scientific articles and monographs to date. The combustion tests of coal dust and chopped biomass were carried out on a test stand consisting of a cylindrical combustion chamber with a length of about 2.5 m and a diameter of about 0.7 m (Figure 3). Inside the chamber was a 1:10 scale model of a rotary dust-air burner compared to industrial burners and an oil burner with a gas igniter. A dust line supplied the fuel mixture and

primary air to the burner [15]. It was possible to independently regulate the fuel and air supply. The feeder speed adjustment varied the amount of coal and biomass fed. During the operation of the plant, the computer data acquisition system recorded a number of parameters, including primary and secondary air flow rate, average temperature in the chamber, gas composition at a distance of 1 m from the front of the burner, and at the outlet of the combustion chamber. On both sides of the combustion chamber, there were viewing holes to allow observation of the flame. For image recording, a digital camera MC1311 from Mikrotron GmbH with a full-frame transmission matrix equipped with a CameraLink interface was used. The camera was attached to a high-temperature boroscope mounted in the viewing port at an angle of approximately 450 to the burner axis (Figure 4). To protect the optical part from contaminants (soot, ash) and high temperatures, the optical fiber was cooled with water and blown with compressed air. Images with a resolution of 800x800 pixels were recorded in RGB mode with a color depth of 24 bits. Exposure time and refresh rate were kept constant at 500 frames per second. Video recordings of individual combustion variants were saved in the computer RAM as AVI files; the duration of each recording was 300 seconds [16].
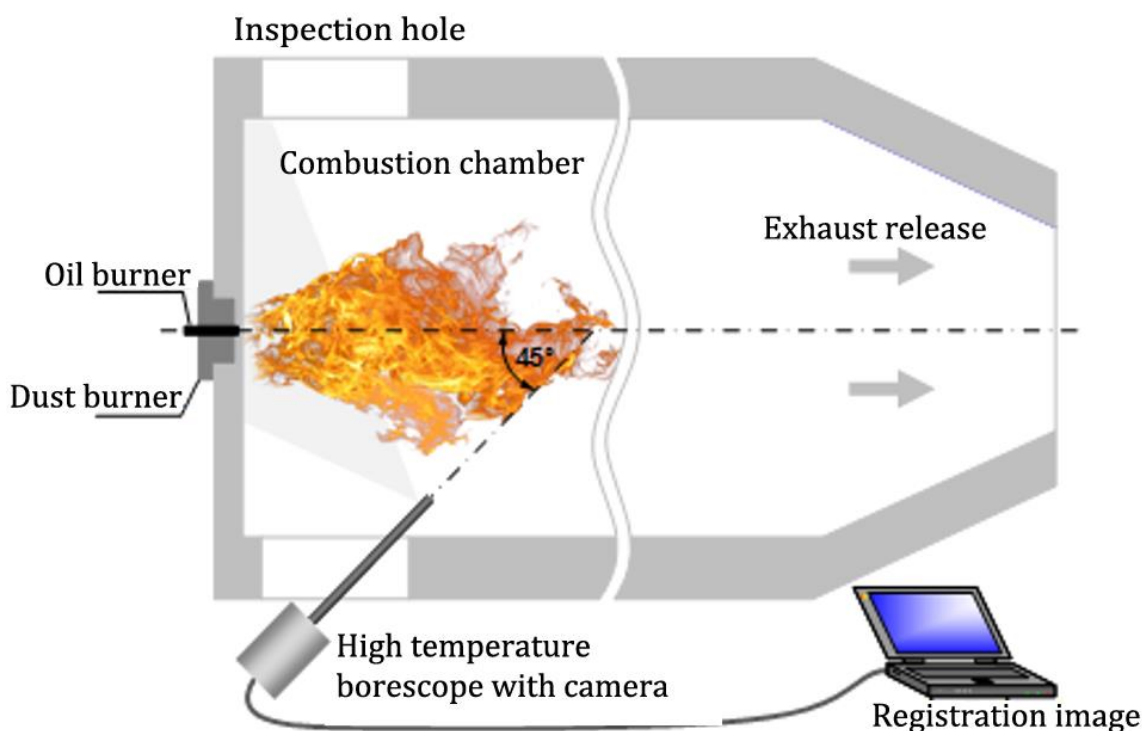

Figure 3.
Schematic diagram showing the chamber arrangement in the combustion chamber.

Mixture of pulverized coal and biomass in the form of chopped straw was burned together during the trials. Control measurements were made under constant fuel flow rate so that the estimated heat output of the burner (Pth) remained constant. The quantity of supplied air was controlled, which influenced the value of the excess air coefficient (λ). The coefficient λ is defined as the ratio of the actual air quantity where the fuel burns to the amount required for its complete combustion (stoichiometric quantity). The stoichiometric value, by contrast, can be computed based on analyzing the equations of chemical reactions occurring during combustion of a particular fuel.

Figure 4 illustrates a laboratory setup, which is located at the Department of Electronics and Information Technology of the Lublin University of Technology for testing the biomass gasification process. The chamber was attached to a high-temperature boroscope (shown by the red arrow), installed in the viewing window at an angle of approximately 450 to the burner axis.
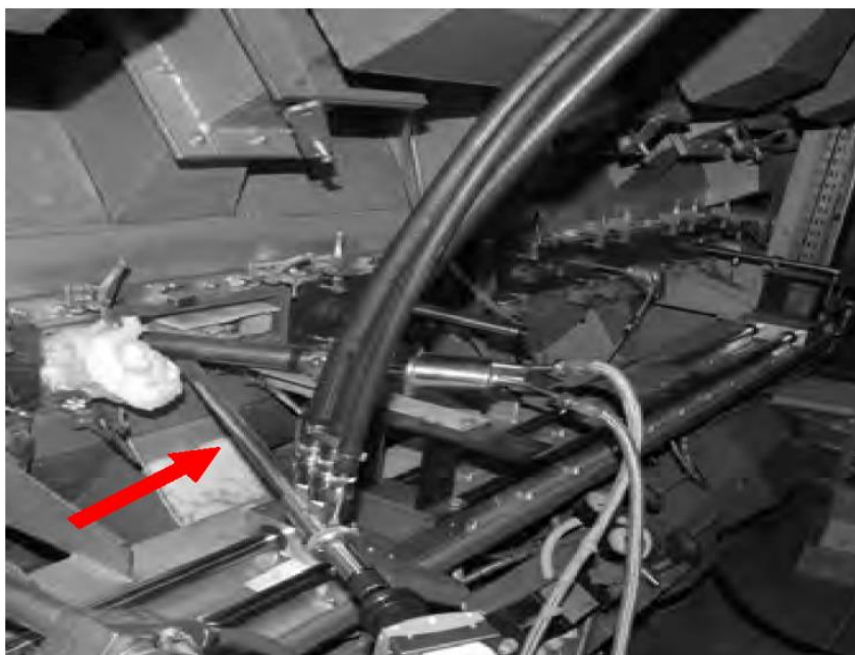
Figure 4.
Photograph of the combustion chamber with boroscope marking (Red arrow).

At co-combustion of pulverized coal and biomass mixture, complete combustion took place at the amount of air less than the stoichiometric amount, i.e., $\lambda \approx 0.75$, so the value of excess air ratio was taken as the value corresponding to "normal" combustion conditions. Proceeding from this, 3 values of the coefficient λ were taken during the tests. The first one corresponded to normal conditions ($\lambda = 0.75$). The second value corresponded to the lack of air ($\lambda = 0.65$), and the third to its excess ($\lambda = 0.85$). For heat output Pth = 250 kW, all three cases were considered; for Pth = 300 kW, the cases of normal combustion and excess air were considered; and for Pth = 400 kW, the cases of normal combustion and lack of air were considered. Tests were conducted for seven combustion variants with different Pth and λ settings and at a biomass content of 10%. Video recording started after the process in the combustion chamber had stabilized. This moment was determined on the basis of gas analyzer readings and the opinion of a specialist with many years of experience in the operation of industrial combustion chambers for pulverized coal and chopped biomass [2, 5].

Image analysis techniques diagnose the combustion process in two main ways. The first uses the flame shape and the second uses the flame texture. Determining the shape, of a flame is not an easy task. The location in space where an exothermic oxidation reaction occurs determines the flame occurrence. It is difficult to clearly delineate the boundaries where combustion occurs because the concentrations of reacting substances do not change dramatically. In the case of a dust flame, the main source of radiation is a cloud of high-temperature solid particles (coal dust, soot, ash, etc.). Their concentration also does not change dramatically, so the edge of the flame cannot be precisely and unambiguously defined, as, for example, in the case of a solid. In a flame, the phenomena of absorption and scattering of radiation also occur, which introduce additional ambiguities when defining its boundary. The problem of a clear definition of the flame edge complicates the whole analysis process. Keeping this in mind, the study used an approach based on the flame texture rather than the flame edge. This eliminates the task of accurately segmenting the flame to identify its geometric features. The second, more important advantage of using texture is that it provides much more information about the object of analysis than its shape.

It was assumed that the combustion conditions and emission parameters stabilized over the period of the tests. It was therefore concluded that a flame image recording fragment of about 3 seconds was sufficient for use in follow-up investigations. The video file fragments pertaining to the different combustion process variants were recorded as 800 x 800-pixel images. It was possible to convert each video file into 1676 PNG image files with a colour depth of 24 bits. VirtualDubMod 1.5.10.2 was used for this conversion. Looking at the contribution of the different colour components in creating the flame image showed that the red colour component (R) plays a key role. The histograms shown in Figure 5 illustrate this. The graphs generated for each combustion case show the average intensity values of the individual components for the entire dataset [17].
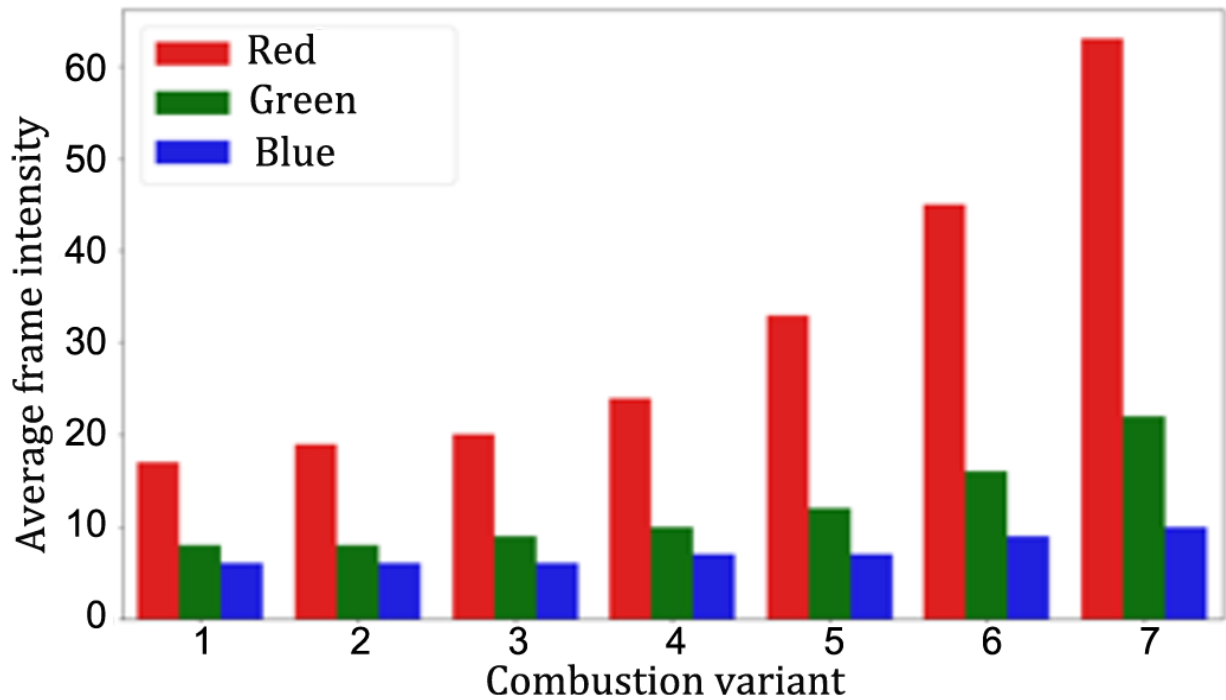
Figure 5.
Average value of the intensity of color components for individual variants of the combustion process. The presented results are based on the analysis of the whole set consisting of 1676 image frames for each variant.

The R-component played a dominant role, so for each image, its intensity, understood as the average value of all image pixels, was calculated. The obtained parameter became the basis for calculating the probability density function (Figure 6) and constructing box plots (Figure 7). The results presented in Figure 6 and Figure 7 show that there are significant differences in the intensity of the R component between individual Pth thermal powers. The higher the power, the higher the intensity of the R-component of the image frames. In turn, for the same thermal power, the intensity of the flames takes different values depending on the excess air ratio. An increase in excess air relative to the normal value decreases the intensity of the image, and, conversely, a decrease in excess air increases the intensity. Note that the brightness of the image background, which extracts the flame contour, closely correlates with the intensity level of the frames. Therefore, in the segmentation process it was necessary to take into account the changes in the intensity of image frames depending on the current thermal power and excess air ratio [18].
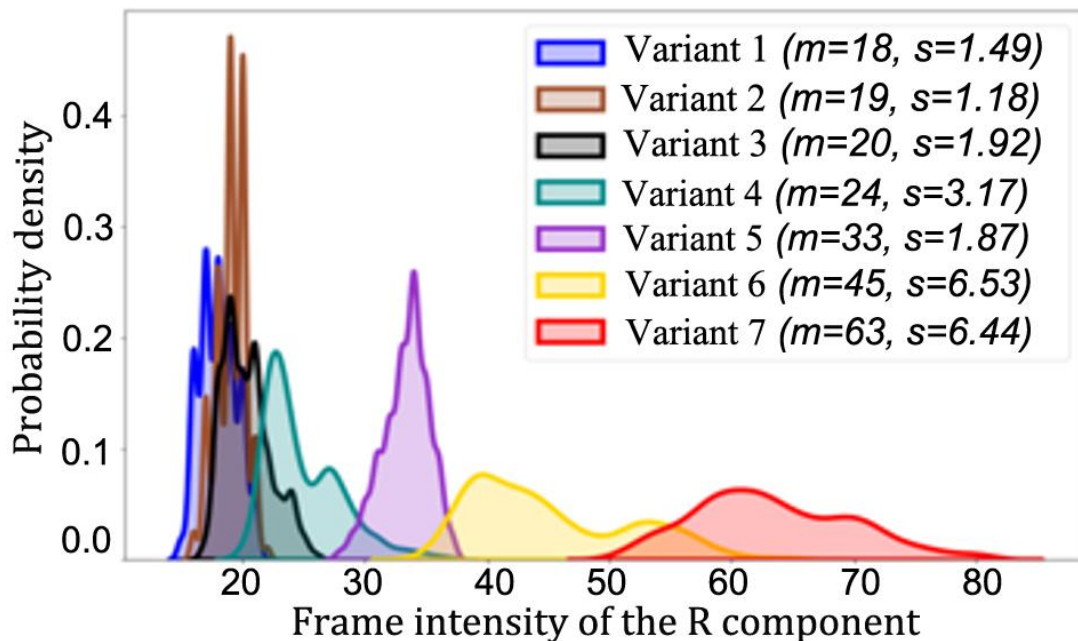

Figure 6.
Plots of the probability density function of the component R frame intensity for individual combustion variants (m - arithmetic mean, s - standard deviation).
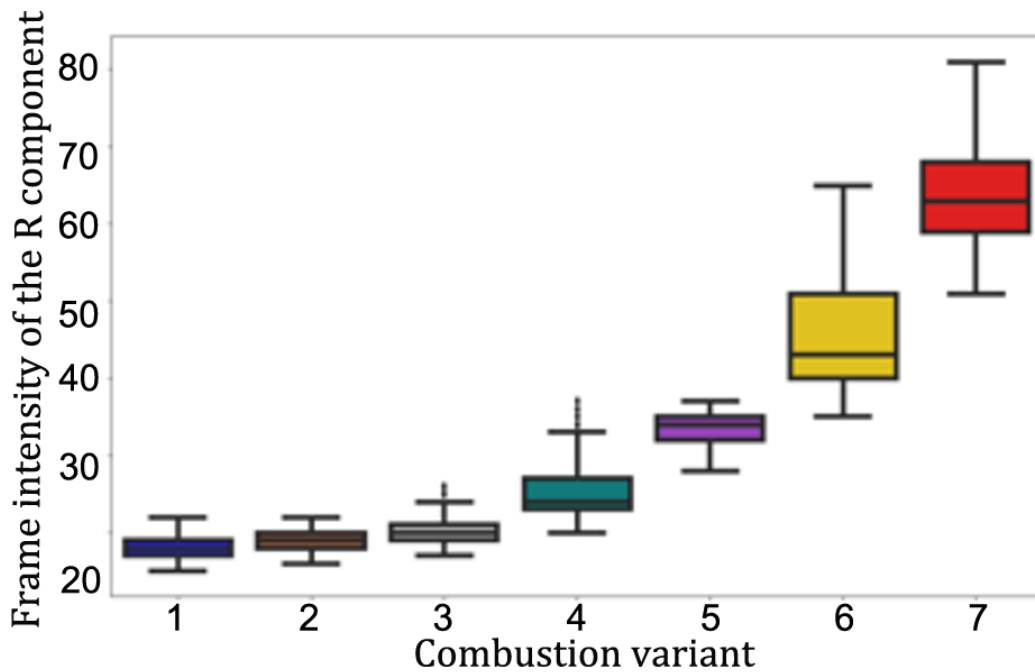
Figure 7.
Scatter plots of the R component frame intensities around the median.

In order to obtain the best conditions for flame contour segmentation, it was necessary to eliminate the background effect, the brightness of which varied depending on the parameter values. A non-linear gamma correction was used for this purpose. The flame's brightest pixels only slightly darkened, while the background's darkest pixels significantly darkened. This reduced the effect of the background on the quality of the flame contour segmentation. The amount of change in the intensity of individual pixels depended on the value of the gamma correction factor G. An adaptive selection of this value based on the intensity of the R-component of the current image frame was used. The type of function used to calculate the required value of G was determined empirically. Note that the gamma correction only acted as an auxiliary function to accurately determine the burning limits, without affecting the value of the feature descriptors. The resulting contour was then superimposed on the original image (without correction) to produce a directional diagram whose centre coincided with the centroid of the contour. Therefore, since the accuracy of flame boundary detection was obtained from the texture of the central part of the flame and not from its contour, the diagnostic performance was not affected.

The goal of the image processing step was to define a square region of interest whose center coincided with the centroid of the flame contour. The shape of the ROI was determined by the properties of the classifiers used. In the case of the ROI size, two opposite requirements had to be taken into account. On the one hand, the analyzed domain should be as large as possible to contain the maximum amount of information to effectively differentiate combustion variants. However, on the other hand, the smaller the region of interest, the shorter the processing time of the whole system and the less frame loss of the image. ROI is a framework for extracting texture features from an image [19].

The loss of a frame occurs when the extraction condition is not met, i.e., the ROI is not completely contained within the flame contour. Figure 8 illustrates the segmentation algorithm that was tested for different ROI sizes to estimate the percentage of image frames where the extraction condition was not met (Table 2). This frame loss issue is particularly critical because the frame area varies depending on the different combustion variants. In the case of variants 1÷3, the flame area is relatively large and homogeneous, and the image shows a low dynamic of texture changes. These variations occur over a large area, so the ROI should be large enough to account for the differences between the alternatives.
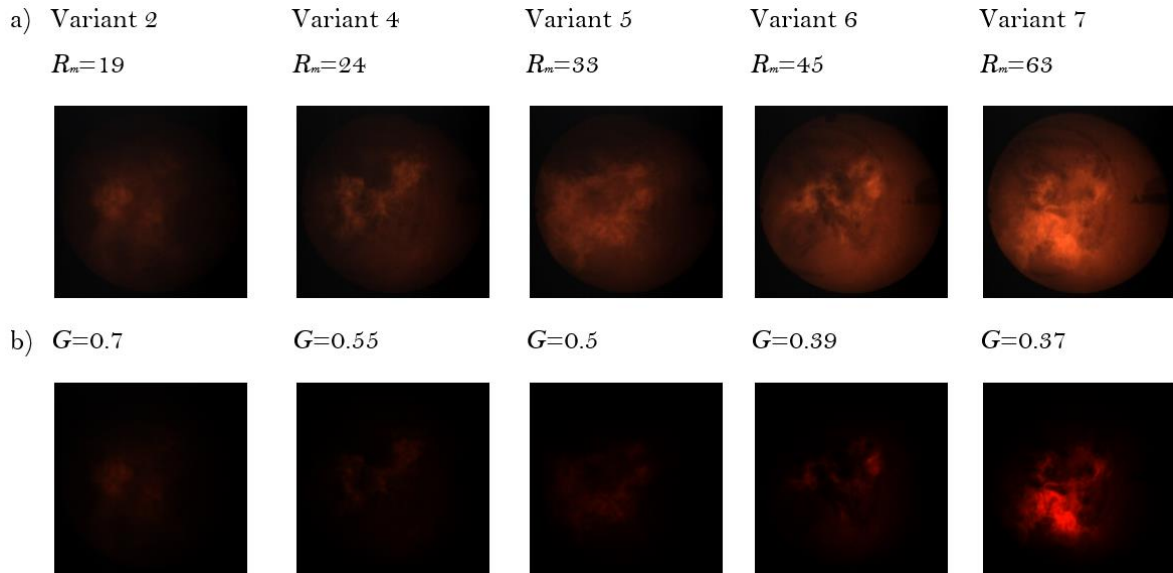
a) Variant 2   Variant 4   Variant 5   Variant 6   Variant 7

$R_m=19$ $R_m=24$ $R_m=33$ $R_m=45$ $R_m=63$

b) $G=0.7$ $G=0.55$ $G=0.5$ $G=0.39$ $G=0.37$

Figure 8.
Results of gamma correction of flame image samples belonging to different variants: a) Input images; b) Gamma-corrected images.

Adaptive selection of the gamma correction factor effectively eliminates the image background regardless of the average intensity of the R-component. This creates better conditions for flame region segmentation.
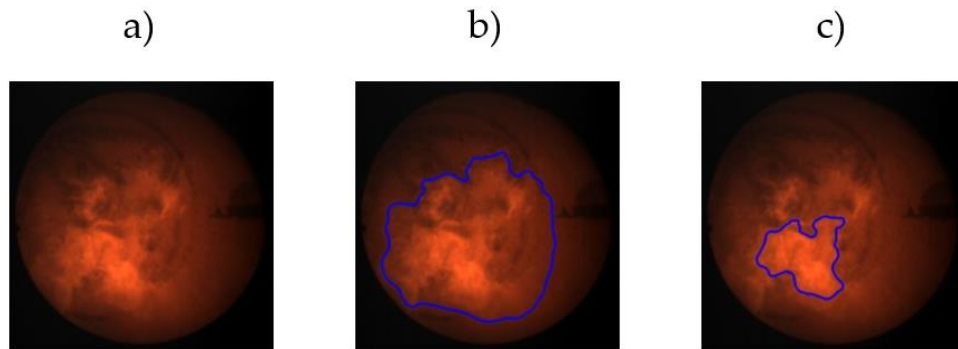
a)   b)   c)

Figure 9.
Effect of gamma correction on segmentation quality: a) Example of flame image related to variant #7; b) Segmentation result without gamma correction (The selected object also contains a significant background area); c) Segmentation result after gamma correction.

Figure 9 illustrates the effect of gamma correction on the segmentation quality.
The results show that the use of gamma correction significantly improves the quality of flame area segmentation.

Table 2.
Percentage of image frames rejected due to not fulfilling the ROI extraction condition for different ROI sizes (Numbers expressed in %). This condition means that the ROI should be completely inside the flame contour.

Variant	100×100	80×80	60×60	40×40
1	0.12	0.00	0.00	0.00
2	0.06	0.00	0.00	0,00
3	6.03	2.27	0.89	0.36
4	36.58	23.75	14.86	7.88
5	2.45	1.37	0.78	0.12
6	44.81	30.43	21.54	14.14
7	39.74	27,15	18.08	11.75

Flame area demonstrates greater irregularity and greater dynamics of texture changes for the variations 4÷7. The ROI size must be reduced compared to variants 1÷3 to reduce frame loss. In the case of options 4÷7, the texture changes between the following images occur in a relatively small area, so that the smallest ROI contains enough information to ensure effective classification. It is estimated that the loss of personnel cannot exceed 15%. On this basis, and in keeping with previous considerations, the size of the ROI has been assumed to be 100 x 100 pixels for variants 1÷3 and 40 x 40 pixels for variants 4÷7 [20].

Below is the list of the most crucial operations carried out during image preprocessing.
1. Loading the source image in PNG format, color depth 24 bits, and size 800x800 pixels (Figure a);

2. Calculating the intensity of the R-component of an image as the average intensity of all its pixels;

3. Perform adaptive gamma correction (Figure b);

4. Convert the image to 8-bit grayscale;

5. Normalize the histogram of the image;

6. Median filtering with a 33x33 pixel mask (Figure c);

7. Thresholding of the image by Otsu method (Figure d);

8. The system performs morphological operations such as closure and opening using a 3x3 pixel structural element. Detection of all contours in the image;

9. Finding the contour with the largest area (Figure e). The observation that statistically out of all the contours in the image, the one belonging to the flame has the largest area was used;

10. Calculating the coordinates of the centroid of the contour (Figure f). The following relations were used: x=m10/m00, y=m01/m00, where: m00 - normal moment of zero order and m10, m01 - normal moments of the first order;

11. Finding a square of size 100x100 or 40x40 pixels (depending on the burning variant), the center of which coincides with the centroid of the contour (Figure g);

12. Copying from the original image (without gamma correction), the region of interest is defined by the found square, provided that it is completely contained in the flame contour (Figure h). If this condition is not met, ROI extraction is not performed.
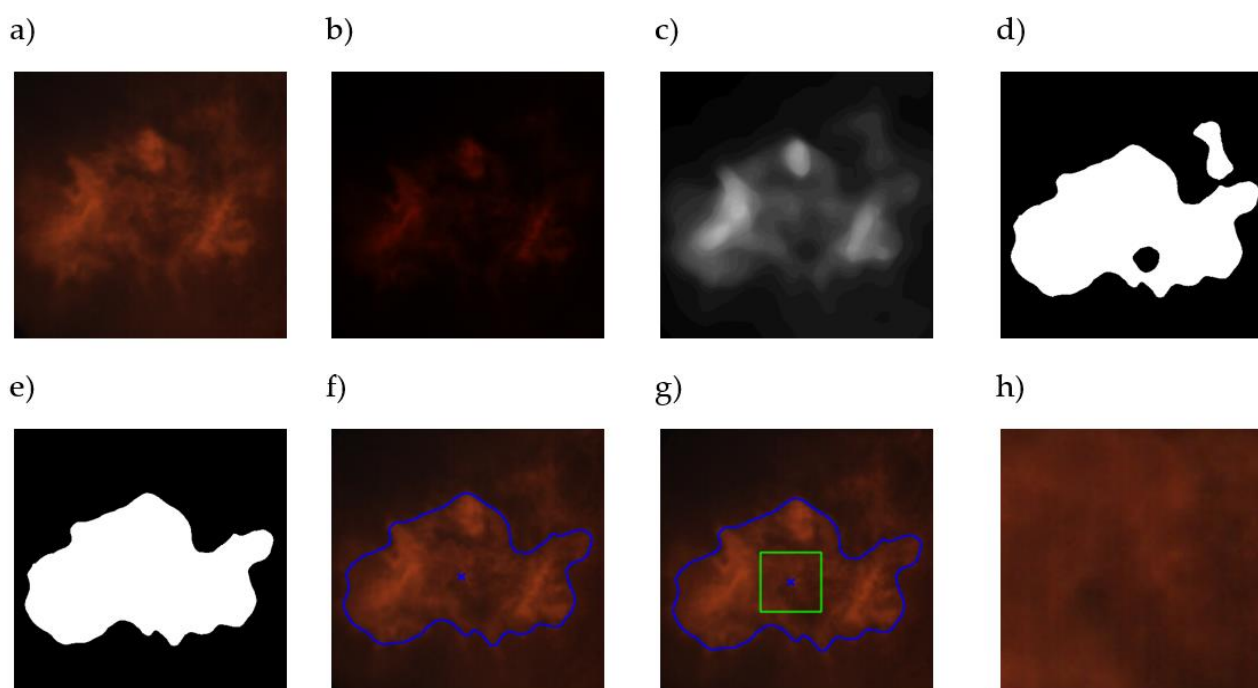

Figure 10.
Sample results of the most important operations performed during image preprocessing: (a) Original image; (b) Result of gamma correction by the function G=f(Rm); c) Grayscale image after histogram normalization and median filtering; d) Binary image after Otsu thresholding; e) Image after morphological operations (Closing, opening) and finding the contour of the largest area; (f) Image with the flame contour and its centroid marked (Blue cross); (g) Image with a marked region of interest (Green square of 100x100 or 40x40 pixels) Whose center coincides with the centroid of the contour; (h) ROI copied from the original image (Without gamma correction).

For better readability, the Figure 10 shows the center portion of the processed image containing the flame. The size of all images (except ROI) was 800x800 pixels.

The full dataset contained 3000 observations, half of which were from variant #5 and half from variant #4. This sample was randomly divided into training and test parts so that the first part was 70% and the second 30% of the full sample (Table. 3).

Table 3.
Number of subsets resulting from the division of a large data set.

Data set	Number of observations	Part of the complete set	Variant №5	Variant №4
Full	3000	1.0	1500	1500
Training	2100	0.7	1050	1050
Test	900	0.3	450	450

The resulting estimated texture parameters yielded 290 feature descriptors. The next step involved pre-processing the data. To the extent that the feature descriptors were measured at different scales (interval, ratio), the traits were scaled using

classical standardization $z_{ij} = (x_{ij} - \bar{x}_j)/s_j$: where: xij - value of feature j for observation i; zij - standardised value of attribute j for observation i; $\bar{x}_j$ - arithmetic mean of feature j; sj - standard deviation of feature j. Upon normalisation, all attributes had an interval scale and a normal distribution of N (0,1). Rescaling was performed once on the training data. Thereafter, the test data were transformed in the same way. Vector of mean values and variance obtained by standardizing the training data was used for each trait during this process [21, 22].

The next step involved a four-step data cleaning procedure.
1. Feature removal with constant values (variance equal to 0);
2. Deletion of features with almost constant values (variance less than 0.01);
3. Removal of features with repeated values;
4. Removal of mutually correlated features.

Pearson's correlation coefficient (r) was used, which reveals linear relationships between features and assumes normality of their distribution. The signs for which |r|>0.9 were removed, indicating a strong correlation.

Data cleaning reduced the number of features from 290 to 23 (Figure 11). Data preprocessing was carried out using the scikit-learn library and the Python programming language, which was used in further experiments.
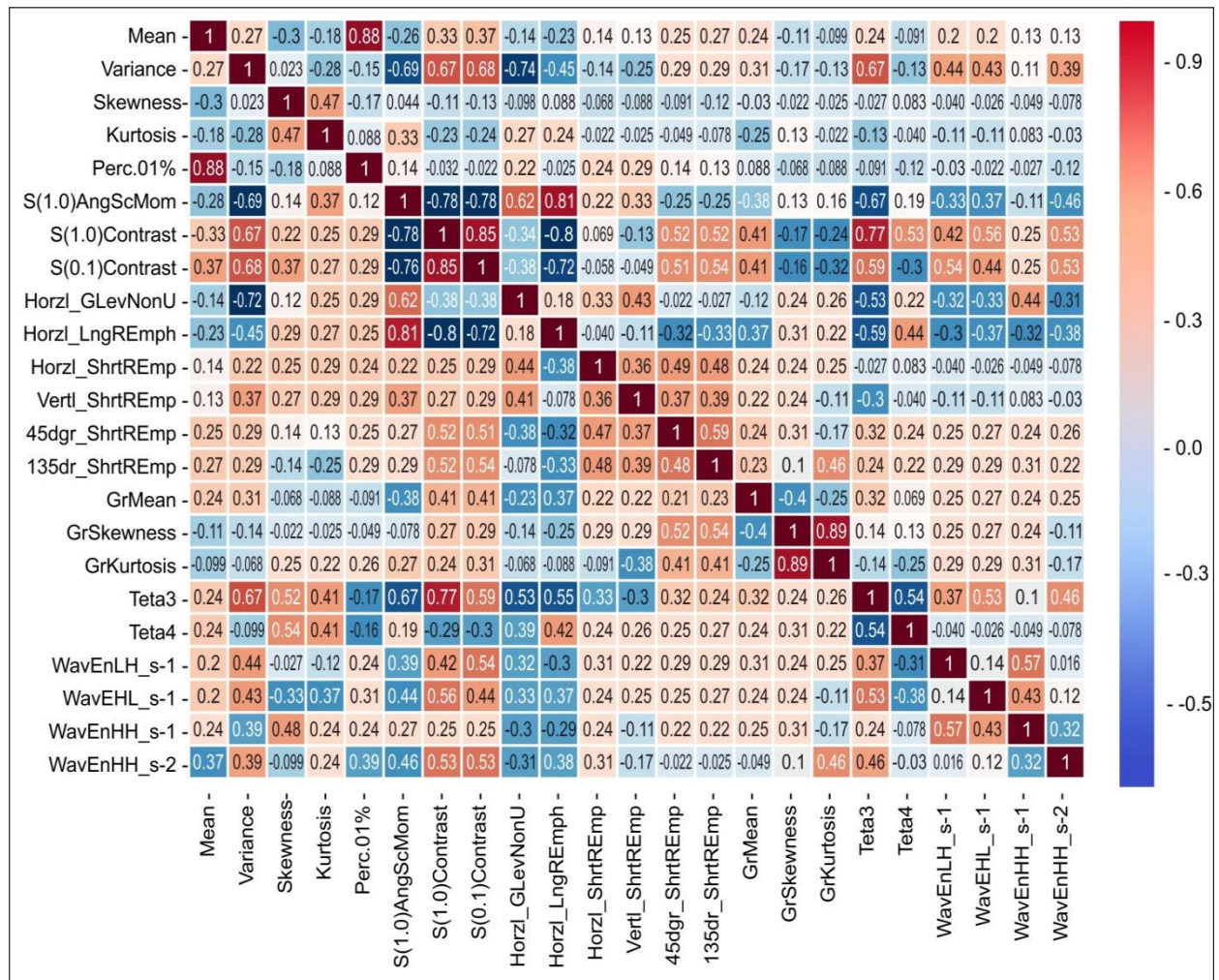

Figure 11.
Correlation matrix between the features obtained by cleaning a large dataset.

At the next stage of data preprocessing, we performed a feature selection. This is intended to reduce the original (complete) set of features to a certain subset comprising features that are relevant in terms of the applied criterion. Data cleaning resulted in a reduction from the full set of 290 features to just 23 features. This significant reduction was mainly due to the presence of a large number of features with high mutual correlation (|r|>0.9). Supplementary reductions were therefore cancelled at the selection stage, and the aim was to generate a rank list for each screening method used. With the ranking lists, a set of features was included to reflect their importance in the process of identifying observations related to particular combustion variants. The number of features appearing in the lists constructed by the individual methods was as follows: filtering and packing methods: 23; embedded methods: logistic regression: 11, AdaBoost: 8, LGBM: 10. The number of features appearing optimally for each selection method was selected in the next step, the learning classification models. So that excessive fitting of the models was avoided, feature selection was performed only on the training data. Its outputs were

then used to transform the test data. Feature selection was performed using scikit-learn, scikit-feature, ReliefF, MLxtend, and Light GBM libraries.

9 training sets (one for each selection method), arranged by feature categorisation, were used in the formation of the classification models. In this search scenario, the feature set was retained, being the one returned by a given selection method. Whilst training the models, the set of features in the training set was modified as per the ranking list. Such numbers ranged from 2 to 23 for filter and wrapper methods, from 2 to 11 for logistic regression, from 2 to 8 for AdaBoost method, and from 2 to 10 for LGBM approach. During the training process, the GridSearch method used to find the optimal values of the hyperparameters of a given model (setting) was improved. For this purpose, the GridSearchCV method available in the model_selection module of the scikit-learn library was used. The modeling was evaluated using 10-fold cross-validation. Motivating this, for each subset of features, the most effective classifier and the optimal values of its hyperparameters were derived. In the following, such a model will be called the optimal model [23].

The Figure 12 illustrates the sample results of the SFS verification method. The curves on the graph each correspond to a different classification method. On the other part, the points marked on the curves reflect the accurateness of the model verification, found to be optimal for a given number of features. With each classification method, 22 optimal models were built, as shown in the diagram. A global total of 242 such patterns were created using all methods and saved as files on disc.
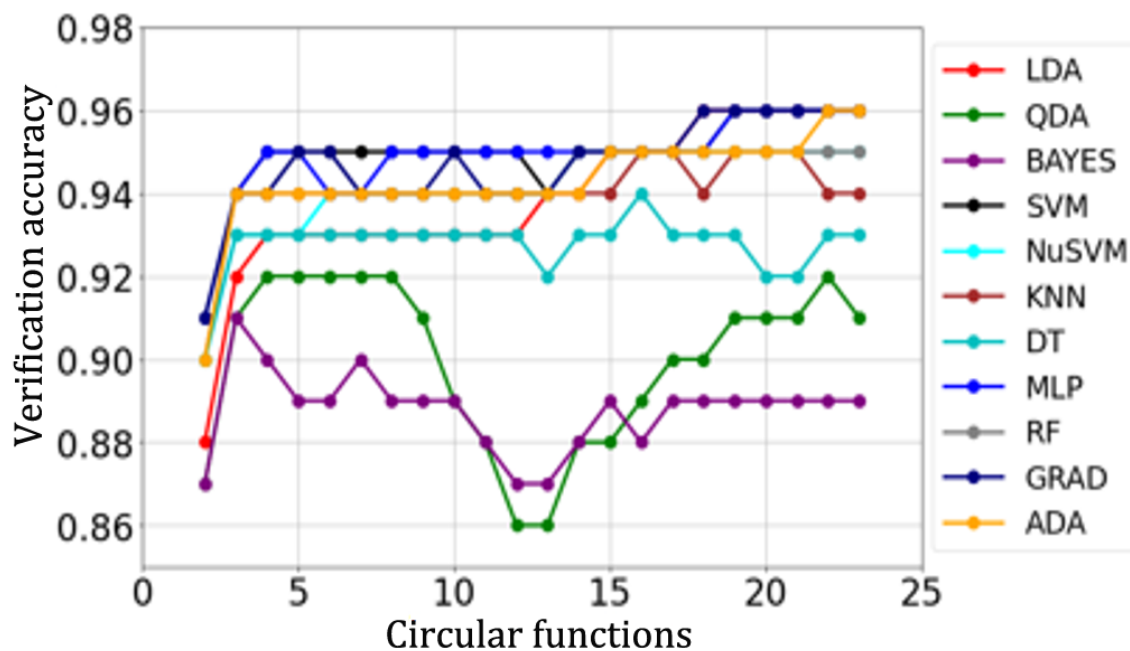

Figure 12.
Results of validation of optimal models for SFS method.

Based on the results presented in Figure 12 for each classification method, we identified the most efficient model as the one that provided the highest validation accuracy with the minimum number of features in the training set. Choosing the most efficient model clearly defined the optimal set of features used for its construction [24].

Table 4 summarizes the results of selecting the most effective models for specific classification methods. Based on the analysis of these data, the best models for the SFS method were found to be the LDA and GRAD models (rows are highlighted with gray background).

Table 4.
Results of selecting the most effective models for specific classification methods using the SFS method.

Classification method	Test accuracy	Optimal number of functions	Optimal model parameters
LDA - Linear discriminant analysis	0.96	18	Solver='svd'
QDA - Quadratic discriminant analysis	0.92	4	tol=1e-05
BAYES - Bayes' theorem	0.91	3	var_smoothing=1e-09
SVM - Support vector machine	0.96	19	C=1.0, gamma='scale', kernel='linear'
NuSVM - Symbolic model checker	0.95	14	Gamma='scale', kernel='rbf', nu=0.3
KNN - nearest neighbors	0.95	16	n_neighbors=5
DT - Decision tree	0.94	16	Criterion='gini', max_depth=7
MLP - Multilayer perceptron	0.96	19	Activation='relu', alpha=0.0001, hidden_layer_sizes=(3,), max_iter=1000, solver='lbfgs'
RF - Random forests	0.95	5	max_depth=9, n_estimators=100
GRAD - gradient-weighted	0.96	18	Loss='exponential', n_estimators=100
ADA - Active domain adaptation	0.96	22	n_estimators=75

The selection criterion for the identification of the optimal model was the analogous to the optimum model, i.e., the highest validation accuracy with the lowest function number in the training set. The verification accuracy for both models was 0.96, and 18 features were used to build them.

All filtering and casing methods utilised the best pattern selection procedure previously described for the SFS method. For the embedded methods, only the classifiers used in the selection process were configured and tuned (Table 5). Specifies for each model the selection method used to create the feature classification, the classification method, and the number of features used. These symbols present in the model symbol should be interpreted as follows: 1 - I position-selection method (A-analysis of variance, F-FISHER, R-RELIEF); 2-e position-ranking method; 3 position-number of features used. Table 5. Best classification models for large datasets and full feature sets. The letter "W" in the model designation denotes the embedded method [25].

Table 5.
Best classification models for large dataset and full feature set. The letter "W" in the model designation denotes the embedded method.

Model symbol	Feature selection method	Classification method	Number of functions
F-SVM-16	FISHER	SVM - Support vector machine	16
A-SVM-10	ANOVA - Analysis of variance	SVM - Support vector machine	10
A-MLP-10	ANOVA - Analysis of variance	MLP - Multilayer perceptron	10
R-SVM-17	RELIEF	SVM - Support vector machine	17
R-MLP-17	RELIEF	MLP - Multilayer perceptron	17
SFS-LDA-18	SFS - Sequential forward search	LDA - Linear discriminant analysis	18
SFS-GRAD-18	SFS - Sequential forward search	GRAD - Gradient-weighted	18
SBS-RF-13	SBS - Sequential backward search	RF - Random Forest	13
RFE-KNN-5	RFE - Recursive feature elimination	KNN - K-Nearest neighbor	5
W-ADA-8	W - wrapper	ADA - Active domain adaptation	8
W-LR-9	W	LR - Logistic regression	9
W-LGBM-7	W	LGBM - Light gradient boosting Machine	7

Figure 13 displays the results of the model validation and testing. The following models had the highest classification accuracy on the test set: R-MLP-17 (acq = 94%), SPS-GRAD-18 (acq = 95%), SBS-RF-13 (acq = 94%), and Sh. -LGBM-7 (ACC = 94%). The test accuracy of the remaining models was slightly worse and ranged from 92÷93%. The vast majority of the models achieved a test accuracy of 96%. Almost all models achieved a higher value of TNR index compared to TPR. This means that variant #5 (normal combustion conditions) was recognized with higher accuracy than variant #4 (excess air). The high quality of the four previously mentioned models is also confirmed by the ROC curves in Fig. On the graph (1-TNR, TPR), they are much closer to the point with coordinates (0, 1), meaning perfect classification, compared to the adequate

curves belonging to other models. The AUC parameters confirm this fact by expressing the size of the areas under the ROC curves. For the four best models mentioned above, the AUC value ranges between 0.987÷0.990, and for the other models, the mentioned range is 0.961÷0.975.

Figure 13 illustrates the results of validation and testing of the best classification models for a large dataset and a full set of features.

Figure 14 shows the ROC curves with AUC parameters for the best classification models built using the full set of features for a large dataset.
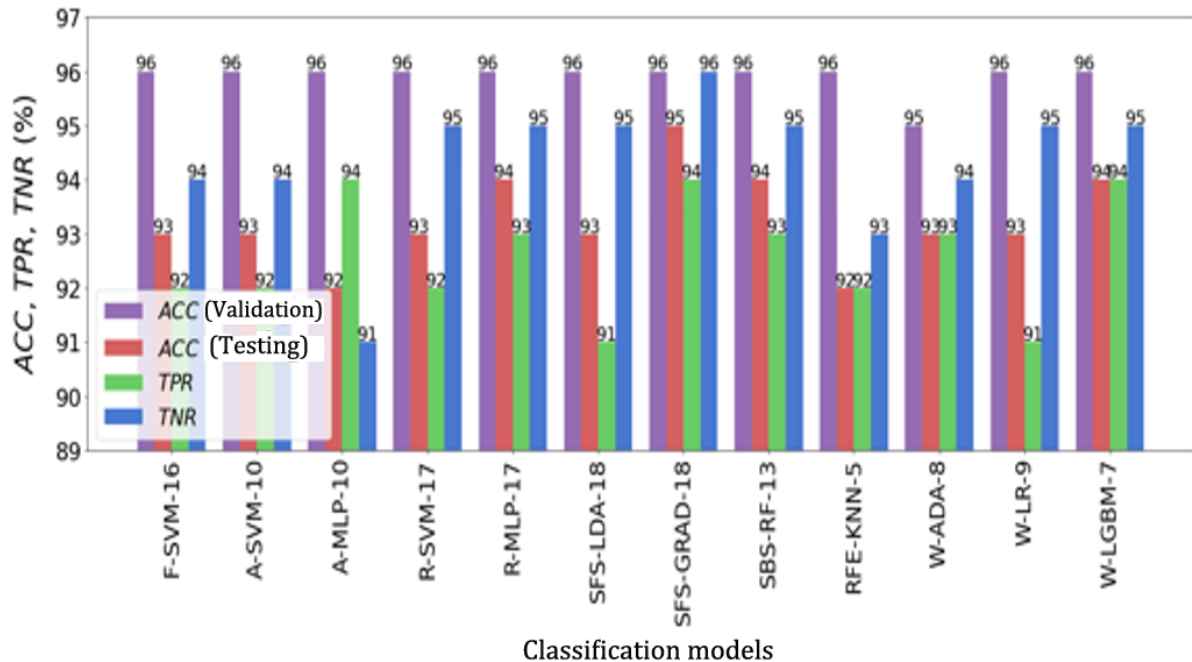

Figure 13.
Validation and testing results of the best classification models for a large dataset and full feature set.
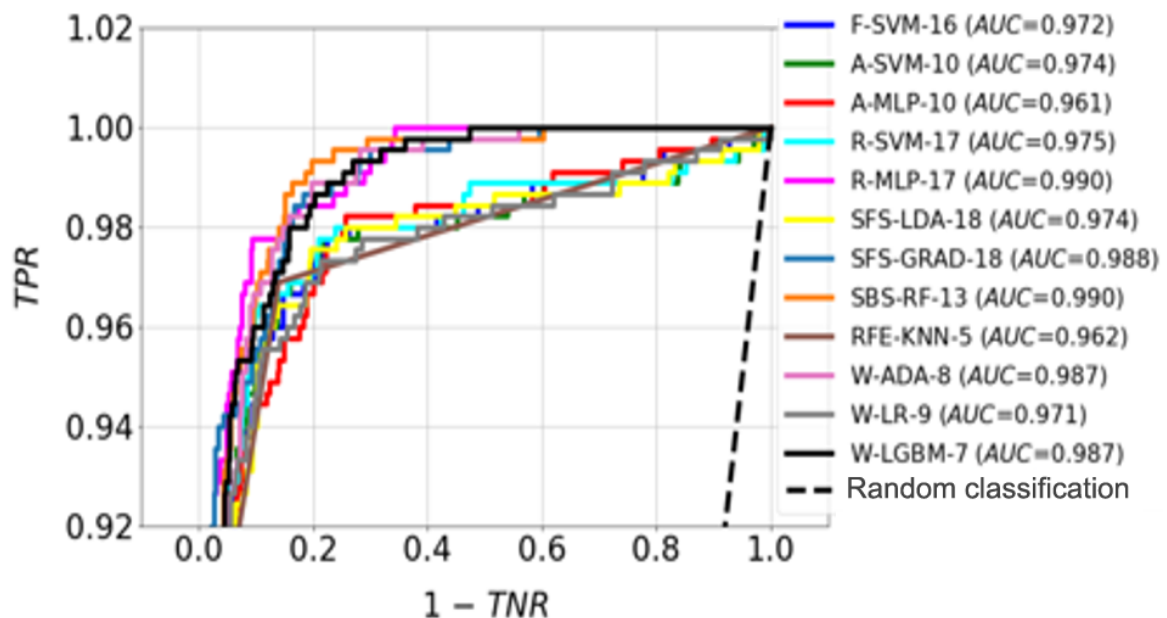

Figure 14.
ROC curves with AUC parameters for the best classification models built using the full feature set for a large dataset.

3. Results of the Study

The study's key finding, which this chapter presents, was that the ranking limited the number of features of a size where an increase no longer significantly increased validation accuracy. It was assumed that this significant increase could not be greater than 0.05. Figure 15 presents the idea of selecting the best model based on the full and limited set of features using the SFS method and GRAD classifier as examples.
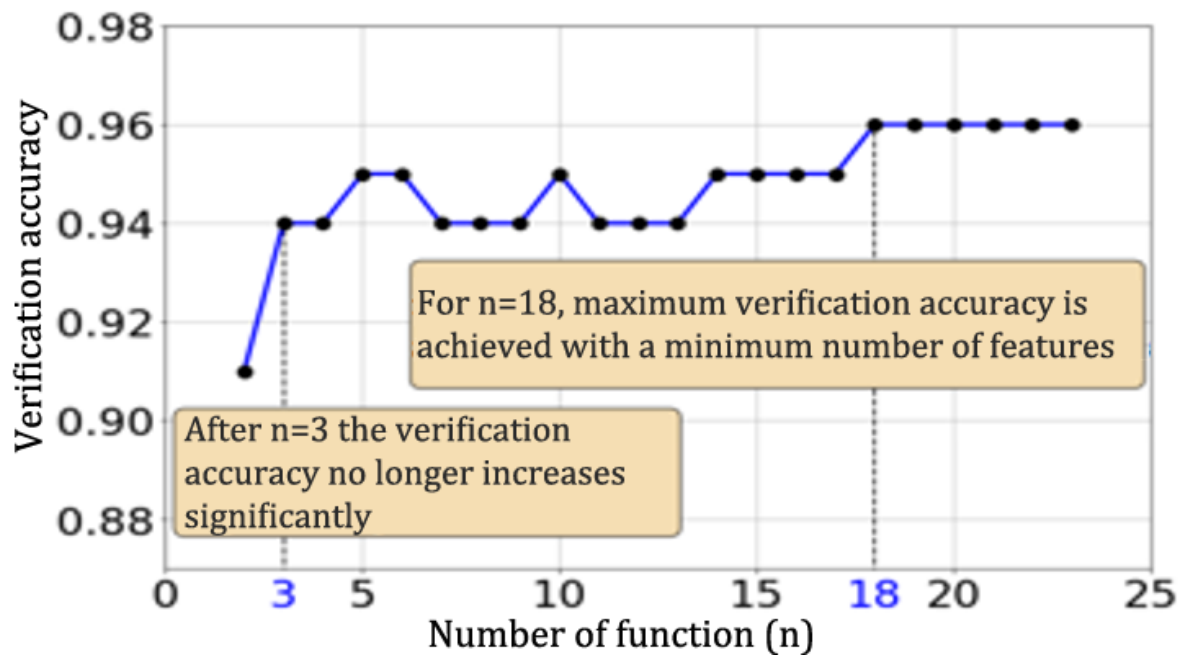
Figure 15.
Selection of the best model over the full and constrained feature set for the SFS method and GRAD classifier.

As can be seen from the progress of the accuracy check, it is justified to take n = 18 for the full set, since this is the minimum number at which the maximum achievable validation accuracy (0.96) is achieved. In the scenario with a limited set of features, we utilized the first 3 features from the ranking list, achieving a validation accuracy of 0.94. As can be seen in Figure 15, further increasing the size of the feature set is no longer associated with a significant increase in this accuracy. Therefore, a model was chosen whose validation accuracy is only 0.02 worse than the maximum accuracy of the GRAD classifier for the SFS method. However, by reducing the number of features from 18 to 3, we achieved a significant reduction in model complexity at the cost of a slight deterioration in accuracy. This situation has a positive impact on the model implementation and the performance of the potential prediction system [4, 5].

The principle of dividing the small dataset for deep learning methods was the same as for the large dataset. The training set previously used in traditional learning methods was also randomly divided into training and training parts, and the same test set was used without any modification. Table 6 contains information about the number of subsets. It should be recalled here that only color images were used in the study scenario using a small dataset.

Table 6.
Number of subsets obtained by dividing a small dataset.

Data set	Number of observations	Part of the full set	Variant #5	Variant #4
Complete	600	1.0	300	300
Training	240	0.4	120	120
Verification	180	0.3	90	90
Test	180	0.3	90	90

Figure 16 displays the plots of the training and validation loss curves. Their course proves that the VGG16, VGG19, MobileNetV2, and ResNet50 models fit the data well. The InceptionResNetV2 model fits slightly worse, and the worst performance was achieved by the Xception model, for which we observe the largest gap between the learning and validation curves. None of the models showed signs of overtraining.

Figure 17 presents the evaluation results. The best-performing models were VGG16 and VGG19, for which the ACC index reached 94%. These models also had a high accuracy in recognizing variant #4 (TPR), which was 96% for VGG16 and 98% for VGG19. The ROC curves shown in Figure 18 confirm the high efficiency of the VGG16 and VGG19 models. The AUC parameters for them are 0.984 and 0.986, respectively. For the other models, the area under the ROC curve is much smaller and varies in the range of 0.892÷0.968.

Using receiver operating characteristic (ROC) curves, we compare the performance of deep learning models.

An ROC curve shows the true positive rate (TPR), or sensitivity, compared to the false positive rate (FPR), or 1-specificity, for different classification score thresholds. The area under the ROC curve (AUC) corresponds to the integral of the curve (TPR values) over FPR values from zero to one. AUC provides an aggregate measure of performance over all possible thresholds.

A perfect classifier always correctly assigns positive class observations to the positive class and has a TPR of 1 for all thresholds.

A random classifier returns random score values and has the same values for FPR and TPR for all thresholds.
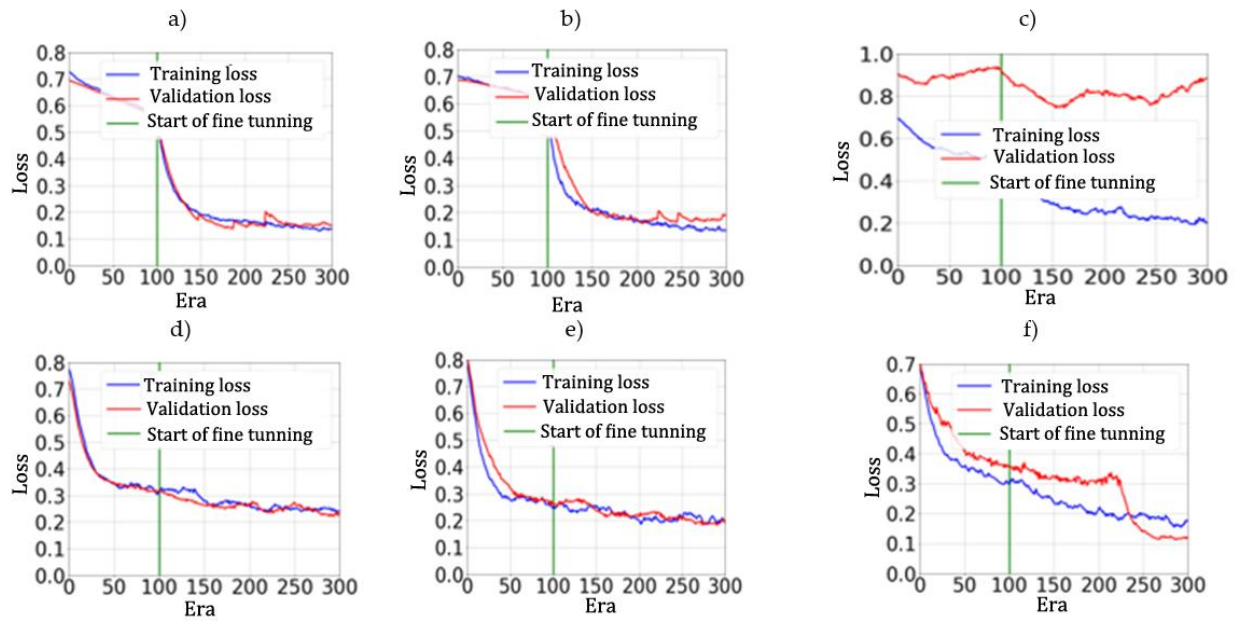
Figure 16.
Model training and validation loss plots for a small dataset and color images: a) VGG16; b) VGG19; c) Xception; d) MobileNetV2; e) ResNet50; f) InceptionResNetV2.
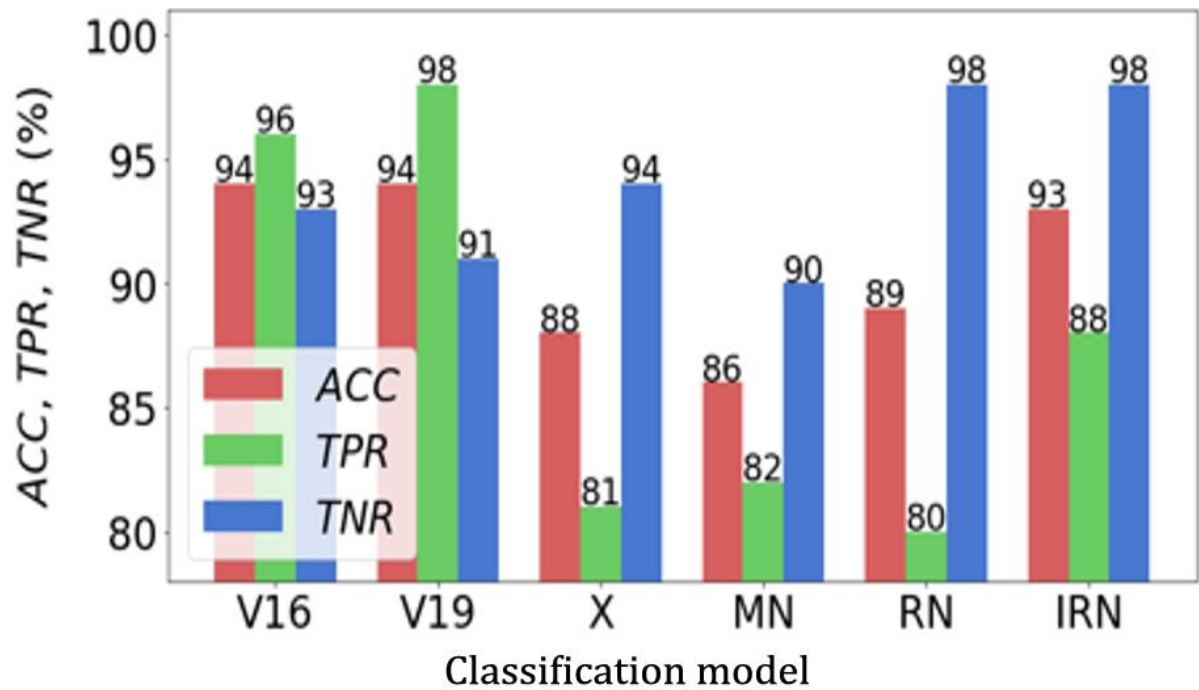

Figure 17.
Results of model testing for a small data set and color images. Symbols used: V16 - VGG16, V19 - VGG19, X - Xception, MN - MobileNetV2, RN - ResNet50, IRN - InceptionResNetV2.
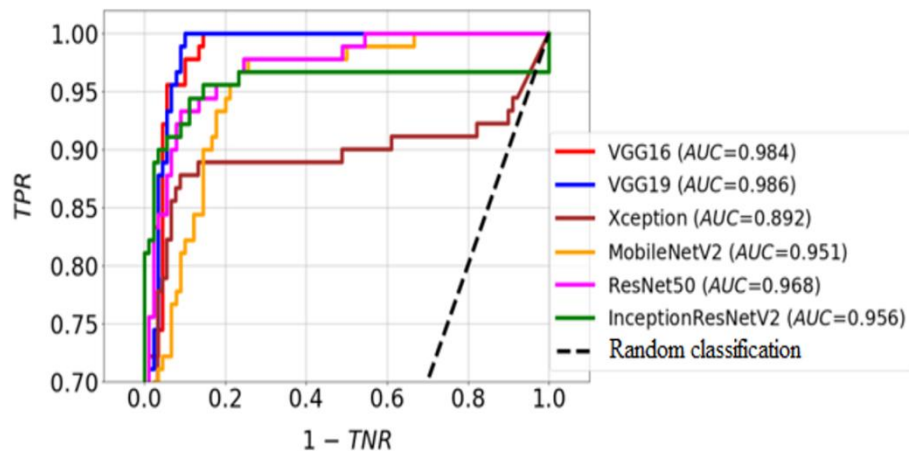
Figure 18.
ROC curves with AUC parameters for models built using a small dataset and color images.

4. Conclusions

The energetic utilization of biomass as a renewable energy source in the presented theoretical and empirical studies, including diagnostics and control of its conversion processes, has shown that in order to carry out these processes efficiently, it is necessary to conduct studies covering both technological processes in existing plants and those that will allow the introduction of both modifications, including control algorithms resistant to variability of technological parameters and environment, and newly developed ones.

It should be emphasized here that diagnostics of these processes is a very important issue. It is essential to ensure proper control over these processes. It should also be noted that each object requires an individual approach, i.e., appropriate research, the results of which will optimize the control processes.

The use of optical methods of control and diagnostics makes it possible to minimize delays occurring in other diagnostic systems as much as possible. At the same time, their resistance to electromagnetic interference and passivity should be emphasized, which is an important advantage.

References

[1]   L. N. Yesmakhanova, *Controlling the combustion of pulverized coal using advanced technologies [Text]: Monograph / L.N. Yesmakhanova.* Taraz: Taraz University, 2019.

[2]   L. Yesmakhanova *et al.,* "Simulating the coal dust combustion process with the use of the real process parameters," *ARPN Journal of Engineering and Applied Sciences,* vol. 16, no. 22, pp. 2395-2407, 2021.

[3]   M. Brudka, *Neural networks in robot control based on ultrasound images.* Politechnika Warszawska: Praca Doktorska, 2000.

[4]   W. Wojcik *et al.,* "Neural methods of interpretation of data obtained from optical sensor for flame monitoring in optical fibers: Applications," in *Proceedings of SPIE,* 2005, vol. 5952, pp. 409-414.

[5]   W. Wójcik and A. Smolarz, "Stabilization of NOx emissions from a single pulverized coal burner using NPC and neural combustion parameter estimation method," *Pomiary Automatyka Kontrola,* vol. 53, no. 11, pp. 20-23, 2007.

[6]   W. Wójcik, M. Kalita, A. Smolarz, and B. Pilek, "Controlling combustion process in power boiler by genetic algorithm and neural network photonics applications in astronomy, communications, industry, and high-energy physics experiments III," in *Proceedings of SPIE,* 2005, vol. 5775, pp. 348-353.

[7]   A. S. Askarova, "Three-dimensional modeling of the processes of formation of harmful substances during the combustion of low-grade coals in combustion chambers [Text] / A.S. Askarova, S.A. Bolegenova, V.Yu. Maximov," *Bulletin of the National Academy of Sciences of the Republic of Kazakhstan,* no. 6, pp. 15-18, 2010.

[8]   A. S. Askarova, "Effect of the computational grid size on the results of computer modeling of heat and mass transfer processes in combustion chambers," [Text] / A.S. Askarova, Bolegenova S.A., Maksimov V.Yu., Bekmukhamet A. // Materials," presented at the 18th All-Russian Scientific Conference of Young Scientists (VNKSF-18). - Krasnoyarsk: Publishing house of Russia, 2012, 2012.

[9]   O. N. Novikov, "Energy-ecological optimization of fuel combustion in boilers and furnaces by regulating the fuel-air ratio [Text] / ON. Novikov, D.G. Artamonov, A.L. Shkarovsky, M.A. Kochergin, A.N. Okatiev. - M.: Industrial Energy," p. 288, 2000.

[10]   A. E. Bryson Jr, W. F. Denham, and S. E. Dreyfus, "Optimal programming problems with inequality constraints," *AIAA Journal,* vol. 1, no. 11, pp. 2544-2550, 1963. https://doi.org/10.2514/3.2107

[11]   S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive network," *Complex Systems,* vol. 2, pp. 321-323, 1988.

[12]   K. Jaroszewski, "System diagnostyczny instalacjioc zyszczania spalin elektrowniwę glowejwy korzystują cysztucznesiecin euronowe, pracadoktorska," Ph.D. Thesis, Szczecin University of Technology, 2007.

[13]   I. Chyzh *et al.,* "Energy resolution of dual-channel opto-electronic surveillance system," in *Proceedings of the Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments, Wilga, Poland, 31 August–2 September 2020,* 2020, vol. 11581.

[14]   W. Wójcik and A. Smolarz, "Using the neural method of estimating combustion parameters to regulate the operation of a pulverized coal burner," *Pomiary Automatyka Kontrola*, vol. 51, pp. 30-33, 2005.

[15]   V. G. Kolobrodov, G. S. Tymchik, M. S. Kolobrodov, A. S. Vasyura, P. Komada, and Z. Azehsova, "The output signal of a digital optoelectronic processor," in *Proceedings of the SPIE Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments, Wilga, Poland, 3–10 June 2018; Volume 10808, pp. 256–263*, 2018.

[16]   K. Tainaka *et al.*, "Measurement techniques for soot in pulverized coal combustion fields," *Journal of the Society of Powder Technology, Japan*, vol. 55, no. 5, pp. 275-281, 2018. https://doi.org/10.4164/sptj.55.275

[17]   F. Tian, H. Zhang, and Y. Tian, "Design and implementation of furnace temperature measurement system for power plant coal-fired boiler in acoustic method," presented at the In MATEC Web of Conferences (Vol. 232, p. 04031). EDP Sciences, 2018.

[18]   W. Xu, Y. Yan, X. Huang, and Y. Hu, "Quantitative measurement of the stability of a pulverized coal fired flame through digital image processing and statistical analysis," *Measurement*, vol. 206, p. 112328, 2023. https://doi.org/10.1016/j.measurement.2022.112328

[19]   Y. Huang, X. Liu, and M. Xu, "Online measurement of soot formation distribution along time and axial in the volatile flames of coal and biomass using light scattering," *Journal of the Energy Institute*, vol. 105, pp. 33-41, 2022. https://doi.org/10.1016/j.joei.2022.07.013

[20]   J. Matthes *et al.*, "Camera based flame stability monitoring and control of multi-burner systems using deep learning based flame detection," *Thermal Science and Engineering Progress*, vol. 41, p. 101859, 2023. https://doi.org/10.1016/j.tsep.2023.101859

[21]   S. Li, T. Xu, P. Sun, Q. Zhou, H. Tan, and S. Hui, "NOx and SOx emissions of a high sulfur self-retention coal during air-staged combustion," *Fuel*, vol. 87, no. 6, pp. 723-731, 2008. https://doi.org/10.1016/j.fuel.2007.05.043

[22]   M. Mollo, A. Kolesnikov, and S. Makgato, "Simultaneous reduction of NOx emission and SOx emission aided by improved efficiency of a once-through benson type coal boiler," *Energy*, vol. 248, p. 123551, 2022. https://doi.org/10.1016/j.energy.2022.123551

[23]   C. Wang, Y. Liu, S. Zheng, and A. Jiang, "Optimizing combustion of coal fired boilers for reducing NOx emission using gaussian process," *Energy*, vol. 153, pp. 149-158, 2018. https://doi.org/10.1016/j.energy.2018.01.003

[24]   L.-G. Zheng, H. Zhou, K.-F. Cen, and C.-L. Wang, "A comparative study of optimization algorithms for low NOx combustion modification at a coal-fired utility boiler," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2780-2793, 2009. https://doi.org/10.1016/j.eswa.2008.01.088

[25]   L. Ma, Q. Fang, C. Yin, H. Wang, C. Zhang, and G. Chen, "A novel corner-fired boiler system of improved efficiency and coal flexibility and reduced NOx emissions," *Applied Energy*, vol. 238, pp. 453-465, 2019. https://doi.org/10.1016/j.apenergy.2019.01.084