Comparison of Traditional Image Segmentation Methods Applied to Thermograms of Power Substation Equipment

Renan de Oliveira Alves Takeuchi 1,*, Leandra Ulbricht 2, Fabiano Gustavo Silveira Magrin 1,3, Francisco Itamarati Secolo Ganacim 4, Leonardo Göbel Fernandes 3, Eduardo Félix Ribeiro Romaneli 1,3 and Jair Urbanetz Junior 1,3

Abstract: The variation in the thermal state of electrical energy substation equipment is normally associated with natural wear or equipment failure. This can be detected by infrared thermography, but technically it demands a long time to analyze these images. Computational analysis can allow an automated, more agile, and more efficient analysis to detect overheated regions in thermographic images. Therefore, it is necessary to segment the region of interest in the images; however, the results may diverge depending on the technique used. Thus, this article presents the improvement of four different techniques implemented in Python and applied in a substation under real operating conditions for a period of eleven months. The performance of the four methods was compared using eight statistical performance measures, and the efficiency was measured by the runtime. The segmentation results showed that the methods based on a threshold (Otsu and Histogram-Based Threshold) were fast, with processing times of 0.11 to 0.24 s, but caused excessive segmentation, presenting the lowest accuracy (0.160 and 0.444) and precision (0.004 and 0.049, respectively). The clustering-based methods (Cluster K-means and Fuzzy C-means) showed similar results to each other but were more accurate (0.936 to 1.000), more precise (0.965 to 1.000), and slower, with 2.55 and 38.8 s, respectively, compared to the threshold methods. The Fuzzy C-means method obtained the highest values of specificity, accuracy, and precision among the methods under analysis, followed by the Cluster K-means method.

Keywords: infrared thermography; substation equipment; abnormal operation; segmentation; image processing

1. Introduction

Substations play an essential role in the electrical power system, as they are responsible for the safe and reliable operation of the electrical network [1,2]. The failures observed in substations are predominantly caused by overloads, timeworn equipment, and unbalanced loads [3]. These failures can make corrective maintenance more costly, in addition to compromising the operational safety and reliability of the network [1,3,4]. Monitoring the equipment through infrared thermal images is one of the recommended methods to detect these failures, as an abnormal increase in temperature is linked to equipment problems and can lead to a severe failure of the power system [5–7]. Therefore,
the fast and accurate detection of such thermal phenomena is essential to ensure system integrity and reduce financial losses [1,5].

Infrared thermography (IRT) is a technique that consists of capturing images of the infrared radiation emitted by objects [6,8]. This technique has some advantages, as it is contact-free and free from electromagnetic interference from the substation equipment itself [8–10]. In addition, it is safe for those involved, and it allows us to monitor electrical equipment in real time, without interrupting the system’s operation [8,10]. On the other hand, IRT also has disadvantages, such as the measurement of only the surface temperatures, interference from environmental conditions, reading errors caused by the different types of emissivity of materials that make up the equipment, the high cost of high-quality cameras, and the need for specific technical knowledge [11].

Another aspect regarding the use of IRT in the traditional way is the need for a team to travel to the location to monitor the substation equipment, followed by the long time spent screening images to identify the indicators of failures [7]. Afterward, another team must return to carry out the repair. In addition, if the failure occurs between inspections, it can lead to an interruption in the system’s operation [5,7,12] because, according to the Brazilian standard ABNT-NBR-15763, a thermal inspection must be performed every six months. A consistent solution to the issues presented above is the application of the computational analysis of infrared images, increasing the efficiency of diagnostics based on infrared detection [2–4].

Several works have approached computational analysis in the field of thermography, using different image processing techniques and computer vision, as well as the inclusion of artificial intelligence algorithms for problem prediction and classification, which may require a high level of processing. In recent decades, significant progress has been made in image processing technology [13,14]. Processing must detect the overheated region of the electrical equipment, called the region of interest (ROI), then extract characteristic information from the ROI, and finally determine the indicators of failure according to predefined classification criteria. Among image processing techniques, the overheated area detection technique predominantly found in the literature in different applications is the infrared image segmentation method [3,5–8,12].

Different segmentation methods are found in the literature; these are classified as region-based [15], contour-based [5], thresholding-based [12], and clustering-based [16]. However, each method is not equally suitable for all specific types of images, because the methods found in the literature are usually applied to specific image types and applications. Thus, an algorithm may be suitable for one type of image but may not be suitable for another, as is the case with segmentation methods developed for application to optical images, which do not have good results when applied to infrared images, for example, the contour-based method, which had its use discontinued due to malfunction and was not tested in this work. This occurs because of the pixel density and low-intensity contrast, which reduces the definition between the edges of the image objects, making it difficult to divide the image into two planes, with the ROI being precisely in the foreground [8,17]. Another influencing factor is the location of the equipment; sometimes, it is installed in a very hostile environment, which can cause noise in the image acquisition.

The fields of application of the methods are varied, but authors usually do not clearly show the details of the algorithm steps, which makes the development of the methods difficult. Applying these methods in substations with hundreds of pieces of equipment can effectively reduce the inspection workload and the risk of making the wrong decision.

Therefore, this work seeks the best “absolute” method among the improvements of four traditional segmentation methods for detecting the overheating of electrical equipment. The contribution of this article is the proposal to detect hotspots and extract the characteristic information from the ROI to replace human screening, making the diagnosis more agile and efficient with a moderate level of image processing. For this, several standardized manual acquisitions were made weekly over a period of eleven
months, always during the morning to have similar load conditions, with approximately 7500 images at the end.

2. Methodology

The analysis was developed in a real application, where all hotspots above 50 °C revealed in a substation under real operating conditions were evaluated between February and November 2021. In this period, five pieces of equipment with hotspots were identified and monitored.

It is a step-down distribution substation, with an area of approximately 6000 m², which has a dual 69 kV supply and lowers the voltage level to eight 13.8 kV bays. The power processed by the substation at peak hours is normally around 30 MW.

During this period, the substation was inspected weekly using a FLIR® T540 thermal camera, whose technical specifications are presented in Table A1 of Appendix A.

This thermal camera captures two images simultaneously, as shown in Figure 1: an optical image and an infrared image. Due to the size of the ROI being relatively small in relation to the whole image, the results of image processing are represented with a zoom in the following figures for better visualization for the reader, represented by the squares in red in Figure 1.

Figure 1. Images captured simultaneously by the thermal camera: (a) optical; (b) infrared. The red square represents the zoom used in the following figures to improve visualization.

The process of image fragmentation into regions facilitates its analysis and allows for obtaining more information about an ROI [15]. The ROI is defined by the segmented region of the object of interest area of the thermal image. The critical information acquired through it is the maximum temperature in that area and the average temperature of the pixels that form the ROI. In the present work, segmentation was performed using four different techniques implemented in Python, always on the same computer, with an i5-3230M 3rd-generation processor, ddr3 4 GB 1600 MHz RAM memory, and Intel HD Graphics 4000 graphics card, and their results are compared in Section 3. These techniques were applied to infrared images, with each pixel having a 16-bit temperature representation.

For this, initially, it was necessary to read the metadata of the radiometric files and extract the Raw Thermal Image file, which has pixel values between 0 and 65,535 (16 bits), referring to the A/D conversion of the radiation captured by the camera sensor. The equations that are applied use each pixel of this file as input (Appendix B), and the expected result is a matrix with values referring to the temperature in degrees Celsius using floating-point representation [18], which is subsequently normalized. This normalization is necessary because it is not possible to store an image with pixels using floating-point representation, so we chose to normalize the values so that they fit in a 16-bit integer variable through Equation (1).
where the maximum, minimum, and delta temperature values are stored in a supporting .txt file. \textit{temp}_c is the temperature in degrees Celsius, \textit{temp\_minimum}_c is the minimum temperature value in the image, and \(\Delta T\) is the difference between the maximum and minimum temperatures of the image. By applying the inverse of the equation, it is possible to retrieve the temperature values in degrees Celsius. Then, the four segmentation techniques were improved to work with 16-bit files and were performed as detailed in the algorithm steps.

First method: Image segmentation using the Otsu method consists of three steps: threshold determination, image thresholding, and ROI location. This method automatically finds the best threshold for the image. This threshold is determined by minimizing the variances within classes or, equivalently, by maximizing the variance between classes [17]. Otsu suggested that the optimal threshold can be obtained from Equation (2) [19].

\[
T = \text{Arg max}\{P_F P_B (\mu_F - \mu_B)^2\}
\]  

where \(P_F\) and \(P_B\) are the foreground and background probabilities, and \(\mu_F\) and \(\mu_B\) are the foreground and background variances, respectively. Then, the image is thresholded in both planes, with the foreground being the location of the ROI. To improve this method, a morphological erode filter is applied, which eliminates edges and slightly reduces the segmented area.

Second method: Image segmentation using the Histogram-Based Threshold method consists of three steps: threshold determination, image thresholding, and ROI location. This method also splits the image into the foreground and background but automatically determines the best threshold based on the shape properties of the histogram using different techniques, such as the analysis of the distance of the convex hull histogram, the smoothing of two peaks through autoregressive modeling, the coarser rectangular approximation for the histogram lobes, or the analysis of peaks and valleys [20]. The threshold \(T\) is in the visible and deep valley of the bimodal histogram [21].

The algorithm starts with the construction of the input image histogram by finding a threshold \(T\) that separates the histogram into two parts according to the probability criteria for the foreground and background areas. These are calculated using Equation (3) [20].

\[
P_B(T) = \sum_{g=0}^T p(g), \quad P_F(T) = \sum_{g=T+1}^G p(g)
\]  

The background is expressed as \(P_B\), \(0 \leq g \leq T\), and the foreground is expressed as \(P_F\), \(T + 1 \leq g \leq G\), where \(G\) is the maximum temperature value of a pixel, which, in this case, is 65,535 for a 16-bit image.

The next two methods are cluster methods, where the image is divided into regions that are represented by their respective centers, representing the pixel with the highest image temperature and the ROI in the cluster having the highest center temperature. In the first step in applying these algorithms, it is necessary to input the number of clusters, which are typically manually entered in other research, but this number may vary by image. To make this step automatic, an improvement was developed in this work through a preliminary algorithm for the automatic determination of the best number of clusters. Through the curve of the within-cluster sum-of-squares (Cluster K-means) and the final fuzzy partition coefficient (Fuzzy C-means), it was possible to apply the elbow method, which detected a dramatic reduction in the error, forming an elbow before stabilization [22]. In Figure 2, the curve of the within-cluster sum-of-squares is represented according to the variation in the number of clusters tested, ranging from 2 to 30, which presents the elbow in the case of eight clusters, that is, the greatest distance between the green line and the curve. In this example, the ideal number of clusters is 8.
Third method: Image segmentation using the Cluster K-means method consists of three steps: reporting the number of clusters, image clustering, and ROI location. This is the vector quantization method, which divides the dataset into k clusters in an interactive process and is represented by the average of all data in each cluster [23]. For each cluster, the midpoint, called the k-centroid, is calculated. Then, each pixel is randomly chosen, and the cluster value is assigned by comparing the shortest distance with the centroids. After associating all of the pixels, new k-centroids are calculated. A loop is generated, only being interrupted when the centers no longer change. Thus, the within-cluster sum-of-squares [24], that is, the variance given by Equation (4), is minimized.

\[ J = \sum_{k=1}^{K} \sum_{y=1}^{n} \|x_y - C_k\|^2 \]  

where \( \|x_y - C_k\|^2 \) is the distance between the pixel \( x_y \) and the centroid \( C_k \), and \( n \) is the number of pixels in the \( i^{th} \) cluster.

Fourth method: Image segmentation using the Fuzzy C-means (FCM) method consists of three steps: determining the number of clusters, image clustering, and ROI location. This method performs clustering based on fuzzy theory and provides more flexible clustering results, starting with initial cluster centers or arbitrary association values [25]. To improve this method, a standard seed was applied, avoiding the arbitrary assignment of the initial center in each execution of the algorithm. Each pixel is assigned a probability of belonging to this cluster, and the quadratic criterion is minimized, where clusters are represented by their respective centers [12]. FCM (\( J \)) is given by Equation (5).

\[ J = \sum_{j=1}^{c} \sum_{k=1}^{N} (u_{ik})^m \|y_k - v_i\|^2 \]  

FCM iterates a maximum number of \( t \) times while updating the degrees of association \( u_{ik} \) and cluster centers \( v_i \), found through Equations (6) and (7):

\[ u_{ik} = \left( \frac{\sum_{j=1}^{c} (\|y_i - v_{jk}\|^{2/(m-1)})}{\|y_i - v_{jk}\|} \right)^{-1} ; 1 \leq k \leq N; 1 \leq i \leq c \]  

\[ v_i = \frac{\sum_{k=1}^{N} (u_{ik})^m z_k}{\sum_{k=1}^{N}(u_{ik})^m} ; 1 \leq i \leq c \]  

Once the segmentation processes were completed, it was necessary to compare the performance of the different segmentation methods by conducting verification based on human evaluation. To standardize and perform a fair comparison between the methods, a reference standard was generated; that is, initially, it would be the region of overheating generated by a technician when detecting a hotspot during an inspection. This is the region that would be reported in the electrical-equipment-overheating technical report.
based on the analysis conducted by the technician in FLIR® Tools 6.4.18039.1003 software (Wilsonville, United States), adopted here as a human standard. Subsequently, it was identified that, in the case of the tool used, there was a limitation in the selection of the ROI due to its low accuracy, which, at best, would make the process costly over time until reaching the expected result. Therefore, an improved human standard was proposed and adopted as the reference or ground truth in this work. It was obtained by processing IRT through an algorithm in Python, where, initially, the threshold was manually entered after its definition by human criteria according to visual perception. Then, the image was separated into the background and foreground. The next section explains the selection of both patterns.

Then, several parameters corresponding to statistical performance measures were calculated. To compare the results of automatic segmentations with the improved human standard, a binary classification of pixels was performed, where TP is true-positive pixels, FP is false-positive pixels, TN is true-negative pixels, and FN is false-negative pixels [26].

From these data, statistical performance measures were also calculated, that is, criteria for selecting the best method, namely, sensitivity, specificity, accuracy, positive predictive value (PPV) or precision, negative predictive value (NPV), false-positive rate (FPR), false discovery rate (FDR), and false-negative rate (FNR), through the formulas shown in Table 1 [27].

| Measurements | Formula | Measurements | Formula |
|--------------|---------|--------------|---------|
| Sensitivity  | \( \frac{TP}{TP + FN} \) | NPV         | \( \frac{TN}{TN + FN} \) |
| Specificity  | \( \frac{TN + FP}{TP + TN} \) | FPR         | \( \frac{FP}{FP + TN} \) |
| Accuracy     | \( \frac{TP + FP + TN + FN}{TP} \) | FDR         | \( \frac{FP}{TP + FP} \) |
| PPV          | \( \frac{TP + FP}{TP + FP} \)     | FNR         | \( \frac{FN}{TP + FN} \) |

In addition to the statistical performance measures, the efficiencies of the methods in the form of the runtime were analyzed. For this metric, the runtime of a purely human analysis was compared with the automatic computational methods.

3. Results

In this section, the results of all hotspots detected in infrared images during the study period are presented, as well as analyzed and compared with each other regarding their performance and efficiency (runtime).

Case 1 detected a hotspot in a 13.8 kV Tandem disconnect switch (Figure 3a), and case 2 detected a hotspot in an insulator connection after passing through the circuit breaker at 69 kV (Figure 3b). Case 3 detected a hotspot on an expander bar of the 13.8 kV three-pole disconnect switch (Figure 3c), case 4 detected a hotspot in a 13.8 kV blade disconnect switch (Figure 3d), and case 5 detected a hotspot in a 13.9 kV-region three-pole disconnect switch connector (Figure 3e).

The pixel with the highest temperature in each infrared image was extracted, as can be seen in Figure 3a,e,i,m,q, with temperatures of 64.4 °C, 98.3 °C, 130.3 °C, 73.9 °C, and 76.1 °C, respectively.

After analyzing the images, a technician selected the ROI according to the tools available in the FLIR® Tools 6.4.18039.1003 software (Wilsonville, United States), resulting in the regions shown in Figure 3b,f,j,n,r. To maintain the same pattern of image processing and make it possible to compare them with the other methods, these ROIs were manually segmented by coloring the pixels using the GIMP® image editing software and separating
the IRT into two planes, as can be seen in Figure 3c,g,k,o,s, termed standard human segmentation for each case.

As can be seen in these figures, a significant number of pixels were not selected due to the limitation mentioned above, which would compromise the analysis of automatic methods, thus requiring an improved human standard. The results of the improved human pattern segmentation are shown in Figure 3d,h,l,p,t.

Figure 3. Process for obtaining ground-truth segmentation: (a,e,i,m,q) hotspot identified by the thermography technician during inspection; (b,f,j,n,r) ROI selected by the technician in FLIR Tools 6.4®
software; (c,g,k,o,s) ROI segmentation selected by the technician performed manually in GIMP® image editing software—human standard segmentation; (d,h,l,p,t) ROI segmentation resulting from the algorithm with the threshold defined by human criteria according to visual perception—improved human standard segmentation.

Then, the thermographic images were submitted to the automatic segmentation algorithms shown in this work. The results of these techniques are shown in Figure 4, where the ROI of the improved human pattern is superimposed over the automatic segmentations in order to make the process more visually accessible to the reader. During the analyzed period, varying environmental conditions were observed. Such conditions may influence the segmentation process when applied directly, as is the case of thermographic images on days with the presence of clouds.

As shown in Figure 4a,b,e,i,j,m,n,q,r, many pixels were marked as hotspots, i.e., FP (shown in white), due to the characteristics of infrared images with a high pixel density and the soft edges of the image objects, reinforcing that the use of segmentation methods developed for application in optical images cause excessive segmentation when applied to infrared images.
For the case of Figure 4k,l,o,p,s,t too many pixels were detected as FN (shown in red, highlighting in the enlarged region) due to excessive segmentation. In the case of Figure 4f, only two pixels were detected as FN; this is linked to the high variance in temperature, which impacted the correct choice of the threshold. In Figure 4c,d,g,l,s,t, a small number of pixels were detected as FP, whereas in Figure 4h, no pixels were incorrectly detected: i.e., all pixels were correctly determined as TP.

The five examples below show the processing time and the maximum and average temperatures of the ROI resulting from each method. The data are presented in Table 2.

The camera’s temperature measurement accuracy is ±2 °C or ±2% of the reading (as presented in Thermal Camera Technical Specifications—Appendix A).

Table 2. Maximum and average temperatures of the segmented ROI of the analyzed cases and the runtime of each method.

| IRT Analyzed | Segmentation Technique | ROI's Maximum Temperature [°C] | ROI's Average Temperature [°C] | Runtime [s] |
|--------------|------------------------|---------------------------------|----------------------------------|------------|
| Case 1       | Human Standard         | 64 ± 2                          | 59 ± 2                           | 105.00     |
|              | Optimized Human Standard | 64 ± 2                          | 60 ± 2                           | 113.00     |
|              | Otsu                   | 65 ± 2                          | 49 ± 2                           | 0.12       |
|              | Histogram-Based Threshold | 65 ± 2                          | 35 ± 2                           | 0.15       |
|              | Cluster K-means        | 65 ± 2                          | 59 ± 2                           | 2.65       |
|              | Fuzzy C-means          | 65 ± 2                          | 59 ± 2                           | 21.00      |
| Case 2       | Human Standard         | 98 ± 2                          | 85 ± 2                           | 73.00      |
|              | Optimized Human Standard | 98 ± 2                          | 80 ± 2                           | 92.10      |
|              | Otsu                   | 99 ± 2                          | 33 ± 2                           | 0.15       |
|              | Histogram-Based Threshold | 99 ± 2                          | 83 ± 2                           | 0.24       |
|              | Cluster K-means        | 99 ± 2                          | 72 ± 2                           | 2.90       |
|              | Fuzzy C-means          | 99 ± 2                          | 80 ± 2                           | 38.80      |
| Case 3       | Human Standard         | 130 ± 2                         | 71 ± 2                           | 50.00      |
|              | Optimized Human Standard | 130 ± 2                         | 70 ± 2                           | 90.00      |
|              | Otsu                   | 140 ± 2                         | 42 ± 2                           | 0.14       |
|              | Histogram-Based Threshold | 140 ± 2                         | 40 ± 2                           | 0.11       |
|              | Cluster K-means        | 140 ± 2                         | 122 ± 2                          | 2.75       |
|              | Fuzzy C-means          | 140 ± 2                         | 72 ± 2                           | 18.60      |
| Case 4       | Human Standard         | 73 ± 2                          | 61 ± 2                           | 58.00      |
|              | Optimized Human Standard | 73 ± 2                          | 61 ± 2                           | 110.00     |
|              | Otsu                   | 74 ± 2                          | 31 ± 2                           | 0.15       |
|              | Histogram-Based Threshold | 74 ± 2                          | 31 ± 2                           | 0.18       |
|              | Cluster K-means        | 74 ± 2                          | 67 ± 2                           | 2.85       |
|              | Fuzzy C-means          | 74 ± 2                          | 67 ± 2                           | 33.76      |
When comparing the maximum temperatures between the improved human standard and the automatic computational methods, the error identified was equal to 0.98%, 0.88%, 7.09%, 0.24%, and 0.17% for cases 1, 2, 3, 4, and 5, respectively. These temperature differences are related to the imperfect mathematical conversion of bits into temperature, which is not sufficient to compromise the response of any method. The error in case 3, despite being high, does not compromise the method because, as the camera was in the range of 20–120 °C, the software ran into a temperature limit (130.3 °C). However, when we extracted the metadata, this range did not matter, as we accessed the data directly from the sensor.

Therefore, there was an agreement between the methods regarding the maximum extracted temperature. However, the average temperatures showed high variation because of the different segmented regions. With the excessive segmentation of pixels, there was a significant drop in the average temperature, as the average temperature ended up being calculated with pixels from a non-hot or even cold region (Otsu and Histogram-Based Threshold, with the exception of case 2).

This same factor also affected the average temperature. The Otsu and Histogram-Based Threshold methods were the most affected ones, but the Cluster K-means method also had problems in case 2. In cases 3 and 4, the Cluster K-means and Fuzzy C-means methods presented a higher average temperature due to sub-segmentation.

All of the automatic computational methods proved to be agile concerning the improved human standard. The threshold methods segmented the areas of interest with a runtime between 0.11 and 0.24 s, while the Cluster K-means method segmented with a runtime between 2.55 and 2.90 s, and the FCM method segmented with a runtime between 8.50 and 38.80 s. This longer FCM runtime was caused by the automatic determination process of the number of clusters (used in pre-processing), which represented about 95% of the total time.

Considering that a technician from the research partner energy company takes around five minutes to analyze the IRT results for each image, even if they use the proposed FCM technique, the longest runtime will take approximately thirty seconds. This represents a saving of approximately four minutes and thirty seconds per image analyzed; that is, this automatic processing method could be ten times faster. Therefore, using an automatic segmentation method provides speed and reduces the effort and processing time, even for the slowest case, which is acceptable for the application proposed in this article. As expected, automatic segmentation methods are much faster than a human; however, there is a large variation in the computational time between the methods, and, depending on the application, this can be a determining factor.

The validation of numerical results to diagnose electrical equipment failures was performed after segmentation. Table 3 refers to the TP, FP, TN, and FN data obtained through the binary classification of pixels in the comparison between automatic segmentation and improved human standard segmentation.
Table 3. Binary pixel classification of the segmented ROIs of the analyzed cases, where TP is true-positive pixels, FP is false-positive pixels, TN is true-negative pixels, and FN is false-negative pixels.

| IRT Analyzed | Segmentation Technique          | Number of Pixels in Segmented Image |        | TN     | FP     | FN     |
|--------------|---------------------------------|-------------------------------------|--------|--------|--------|--------|
|              |                                 |                                     | TP     |        |        |        |
| Case 1       | Otsu                            | 81                                  | 4310   | 507    | 2      |        |
|              | Histogram-Based Threshold        | 83                                  | 702    | 4115   | 0      |        |
|              | Cluster K-means                  | 83                                  | 4814   | 3      | 0      |        |
|              | Fuzzy C-means                    | 83                                  | 4814   | 3      | 0      |        |
| Case 2       | Otsu                            | 8                                   | 2671   | 2221   | 0      |        |
|              | Histogram-Based Threshold        | 6                                   | 4892   | 0      | 2      |        |
|              | Cluster K-means                  | 8                                   | 4885   | 7      | 0      |        |
|              | Fuzzy C-means                    | 8                                   | 4892   | 0      | 0      |        |
| Case 3       | Otsu                            | 429                                 | 3826   | 2045   | 0      |        |
|              | Histogram-Based Threshold        | 429                                 | 3531   | 2340   | 0      |        |
|              | Cluster K-means                  | 23                                  | 5871   | 0      | 406    |        |
|              | Fuzzy C-means                    | 332                                 | 5863   | 8      | 97     |        |
| Case 4       | Otsu                            | 338                                 | 4854   | 6508   | 0      |        |
|              | Histogram-Based Threshold        | 338                                 | 4250   | 7112   | 0      |        |
|              | Cluster K-means                  | 161                                 | 11362  | 0      | 177    |        |
|              | Fuzzy C-means                    | 158                                 | 11362  | 0      | 180    |        |
| Case 5       | Otsu                            | 618                                 | 3114   | 1058   | 10     |        |
|              | Histogram-Based Threshold        | 628                                 | 2842   | 1330   | 0      |        |
|              | Cluster K-means                  | 583                                 | 4166   | 6      | 45     |        |
|              | Fuzzy C-means                    | 583                                 | 4166   | 6      | 45     |        |

Statistical performance measures, namely, sensitivity, specificity, accuracy, PPV, NPV, FPR, FDR, and FNR, are shown in Table 4. Here, for a perfect result, the first five indices must be equal to 1, and the last three must be equal to 0.

Table 4. Results of the eight statistical measures of performance of the four IRT image segmentation methods of the analyzed cases.

| IRT Analyzed | Segmentation Technique | Sensitivity | Specificity | Accuracy | PPV     | NPV     | FPR     | FDR     | FNR   |
|--------------|------------------------|-------------|-------------|----------|---------|---------|---------|---------|-------|
| Case 1       | Otsu                   | 0.976       | 0.895       | 0.896    | 0.138   | 1.000   | 0.105   | 0.862   | 0.024 |
|              | Histogram-Based Threshold | 1.000      | 0.146       | 0.160    | 0.020   | 1.000   | 0.854   | 0.980   | 0.000 |
|              | Cluster K-means        | 1.000       | 0.999       | 0.999    | 0.965   | 1.000   | 0.001   | 0.035   | 0.000 |
|              | Fuzzy C-means          | 1.000       | 0.999       | 0.999    | 0.965   | 1.000   | 0.001   | 0.035   | 0.000 |
| Case 2       | Otsu                   | 1.000       | 0.546       | 0.547    | 0.004   | 1.000   | 0.454   | 0.996   | 0.000 |
|              | Histogram-Based Threshold | 0.750     | 1.000       | 1.000    | 1.000   | 1.000   | 0.000   | 0.000   | 0.250 |
|              | Cluster K-means        | 1.000       | 0.999       | 0.999    | 0.533   | 1.000   | 0.001   | 0.467   | 0.000 |
|              | Fuzzy C-means          | 1.000       | 1.000       | 1.000    | 1.000   | 1.000   | 0.000   | 0.000   | 0.000 |
| Case 3       | Otsu                   | 1.000       | 0.652       | 0.675    | 0.173   | 1.000   | 0.348   | 0.827   | 0.000 |
|              | Histogram-Based Threshold | 1.000     | 0.601       | 0.629    | 0.155   | 1.000   | 0.399   | 0.845   | 0.000 |
|              | Cluster K-means        | 0.054       | 1.000       | 0.936    | 1.000   | 0.935   | 0.000   | 0.000   | 0.946 |
|              | Fuzzy C-means          | 0.774       | 0.999       | 0.983    | 0.976   | 0.984   | 0.001   | 0.024   | 0.226 |
| Case 4       | Otsu                   | 1.000       | 0.427       | 0.444    | 0.049   | 1.000   | 0.573   | 0.951   | 0.000 |
|              | Histogram-Based Threshold | 1.000     | 0.374       | 0.392    | 0.045   | 1.000   | 0.626   | 0.955   | 0.000 |
|              | Cluster K-means        | 0.476       | 1.000       | 0.985    | 1.000   | 0.985   | 0.000   | 0.000   | 0.524 |
|              | Fuzzy C-means          | 0.467       | 1.000       | 0.985    | 1.000   | 0.984   | 0.000   | 0.000   | 0.533 |
| Case 5       | Otsu                   | 0.984       | 0.746       | 0.777    | 0.369   | 0.997   | 0.254   | 0.631   | 0.016 |
|              | Histogram-Based Threshold | 1.000     | 0.681       | 0.723    | 0.321   | 1.000   | 0.319   | 0.679   | 0.000 |
|              | Cluster K-means        | 0.928       | 0.999       | 0.989    | 0.990   | 0.989   | 0.001   | 0.010   | 0.072 |
|              | Fuzzy C-means          | 0.928       | 0.999       | 0.989    | 0.990   | 0.989   | 0.001   | 0.010   | 0.072 |
PPV or precision is very important for this analysis, as it provides the ratio of selected pixels to total pixels that should be selected, so when PPV is equal to 1, the true selection is the maximum. However, it is desirable that the segmentation method combines good accuracy and PPV results.

Some statistical measures effectively reflected cases where segmentation was excessive, such as specificity (when values are less than 1) and FDR and FPR (when values are greater than zero).

As the analyzed methods showed low FN, some statistical measures did not reveal a significant direct relationship in the analysis of the results, such as sensitivity, NPV, and FNR. When the sensitivity has a high value, it is understood that the TP pixels are correct, but in the case of excessive area segmentation, the sensitivity will also be high, as the TP pixels will be inside the excessive area segmentation. When the sensitivity has a low value and the FNR has a high value, it is understood that there is sub-segmentation, as in the results of case 3 in the K-means method. Jaffery [26] reported that excessive sensitivity leads to excessive area segmentation, but this work disagrees with this statement, as excessive area segmentation is related to the presence of FP but not FN pixels.

The segmentation obtained through the Otsu method was not able to segment only the superheated area, causing excessive area segmentation, represented by high FDR and FPR indices, which should be equal to zero but showed values between 0.254 and 0.996 in cases 1 to 5, with the exception of FPR in case 1, with 0.105, and low specificities of 0.427 to 0.895 in cases 1 to 5, which should be equal to 1. The excessive area segmentation caused by this method agrees with the results shown in [5,12]. Low accuracy indices and PPV also disfavor this method. However, it has good results for the segmentation of electrical equipment, similar to the results presented by Fan et al. [3]. Thus, this method can be useful for pre-segmentation, aimed at removing interference from the background and then associating it with another method.

Despite being the second-fastest method, the segmentations obtained through the Histogram-Based Threshold method did not show a good pattern in the results because of the characteristics of the infrared images. In general, when the definition between the edges of the objects in the image is poor, the separation between the two planes is poor. In cases 1, 3, 4, and 5, FDR and FPR were close to 1 and the specificity was between 0.146 and 0.681, which characterize excessive area segmentation, which is also reflected in the low accuracy and PPV indices. In case 2, the results were the opposite: FDR and FPR were equal to zero, specificity equaled 1, and the accuracy and PPV indices were high, characterizing correct segmentation. These opposing results make the method fragile. There are recommendations for the use of this method for images that present a bimodal histogram, and it is not considered good for all types of images, as shown in [21,28,29].

The Cluster K-means method segmented the ROI, showing high indices of specificity, accuracy, and PPV between 0.936 and 1.000 and low indices of FPR and FDR that varied between 0.001 and 0.035, except for case 2, where the FDR of 0.467 is considered high, and the PPV of 0.533 is considered low, both caused by the proportionality between the eight TP and seven FP pixels. However, it did not accurately detect the overheated area in all cases, as shown in Figure 4c,g,k,o,s. Due to the size of the ROI and the proportionality with the FP, the error was high in case 2, but considering the results obtained in case 1 and sub-segmentation in cases 3, 4, and 5, this is the second method recommended in this work. The results presented by Fan et al. [3] showed that the K-means method was not able to accurately discover the real heating area. Muhammad et al. [16] recommended this method because it is more agile than FCM but did not analyze the quantitative performance of the methods.

The FCM clustering method was the only method that segmented the superheated area and reached the maximum values of the statistical measures in case 2. In case 1, only seven FP pixels were detected, which led to a slight drop in the statistical measures of specificity and accuracy (0.999) and PPV (0.965) and low FPR and FDR indices (0.001 and 0.035, respectively). In case 3, only 8 FP pixels, in addition to the 97 FN pixels, were
detected, which led to a greater drop in the statistical measures of specificity and accuracy, with 0.999 and 0.983, respectively. It also had a PPV of 0.976 and great FPR and FDR indices. Cases 4 and 5 produced results that were identical or very close to those of the Cluster K-means method. Our work found results similar to some studies, such as those that showed the longest runtime for FCM, having shown greater precision in [3,16] and being different from studies that showed excessive area segmentation in [8,12]. The possibility of its combination with other segmentation methods (Otsu + FCM) was presented by Fan et al. [3], with promising results. Therefore, this is the first method recommended in this work.

The limitation of this work was due to the selection of segmentation methods based on the number of citations. Thus, some more recently developed and more efficient methods may not have been selected due to the lower number of citations. According to the nature of the work, methods that require a lot of computational time and investment in hardware were not applied, for example, methods that use neural networks, which require more investment, time for training, and trained technicians to create references for the network to follow and a longer application time after the completion of the network.

4. Conclusions

The detection of equipment overheating can guarantee the integrity of the electrical power system, as this is one of the indications of a failure in electrical equipment inside substations. A computational analysis instead of a human technical analysis provides agility in the automatic detection of overheated regions of electrical equipment. Still, a variety of segmentation methods in the literature may or may not be suitable for infrared.

As our final result, we recommend the Fuzzy C-means method. This was the method that obtained higher scores in most cases for sensitivity, specificity, accuracy, positive predictive value, and negative predictive value and lower scores for the false-positive rate, false discovery rate, and false-negative rate in the analyzed infrared images.

Author Contributions: Conceptualization, R.d.O.A.T., L.U., and E.F.R.R.; data curation, R.d.O.A.T. and L.G.F.; formal analysis, R.d.O.A.T., L.U., F.G.S.M., and F.I.S.G.; funding acquisition, E.F.R.R.; investigation, R.d.O.A.T.; methodology, R.d.O.A.T., L.U., and F.I.S.G.; project administration, E.F.R.R.; software, R.d.O.A.T., F.I.S.G., and L.G.F.; supervision, L.U. and J.U.; validation, L.U., F.G.S.M., F.I.S.G., and L.G.F.; writing—original draft, R.d.O.A.T.; writing—review and editing, R.d.O.A.T., L.U., and L.G.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC was funded with R&D resources provided by COPEL-DIS and regulated by ANEEL, grant number PD-2866-0528/2020.

Acknowledgments: The authors would like to thank UTFPR for the support and infrastructure made available for the development of this research and COPEL-DIS for the support through project PD-2866-0528/2020—Development of a Methodology for Automatic Analysis of Thermal Images.

Conflicts of Interest: The authors declare no conflict of interest.
Appendix A

Table A1. FLIR® T540 Thermal Camera: Technical Specifications.

| Technical Specifications       | Value                                      |
|-------------------------------|--------------------------------------------|
| IR Resolution                 | 464 × 348 pixels                           |
| Thermal Sensitivity/NETD      | <40 mK, 24°C @ 30°C (86°F)                 |
| Accuracy                      | ±2 °C (±3.6 °F) or ±2% of reading           |
| Field of View                 | 24° × 18°                                  |
| Focal Length                  | 17 mm (0.67 in.), 24°                      |
| Spectral Range                | 7.5–14 µm                                  |
| Lens                          | 24°                                        |
| Detector Type and Pitch       | Uncooled microbolometer, 17 µm             |
| Emissivity Correction         | Yes, variable from 0.01 to 1.0 or selected from the materials list |
| Object Temperature Range      | 20 °C to 120 °C (−4 °F to 248 °F)          |
|                              | 0 °C to 650 °C (32 °F to 1202 °F)          |
|                              | 300 °C to 1500 °C (572 °F to 2732 °F)      |

Appendix B

The following pertains to the implemented equations used to return the temperature value in degrees Celsius from the raw values of the metadata of the images acquired with FLIR cameras.

Table A2. Input arguments and intermediary variables.

| Variables   | Explanations Regarding the Variables |
|-------------|---------------------------------------|
| E           | Emissivity                            |
| OD          | Object distance from thermal camera in meters |
| RTemp       | Apparent reflected temperature (°C) of the environment impinging on the object of interest |
| ATemp       | Atmospheric temperature (°C) for infrared transmission loss |
| IRWTTemp    | Infrared window temperature (°C)      |
| IRT         | Infrared window transmission          |
| RH          | Relative humidity expressed as percent |
| PR1         | PlanckR1 — calibration constant for FLIR cameras |
| PB          | PlanckB — calibration constant for FLIR cameras |
| PF          | PlanckF — calibration constant for FLIR cameras |
| PO          | PlanckO — calibration constant for FLIR cameras |
| PR2         | PlanckR2 — calibration constant for FLIR cameras |
| ATA1        | ATA1 — atmospheric attenuation constant to calculate atmospheric τ |
| ATA2        | ATA2 — atmospheric attenuation constant to calculate atmospheric τ |
| ATB1        | ATB1 — atmospheric attenuation constant to calculate atmospheric τ |
| ATB2        | ATB2 — atmospheric attenuation constant to calculate atmospheric τ |
| ATX         | ATX — atmospheric attenuation constant to calculate atmospheric τ |
| emiss.wind  | Window emissivity                      |
| h2o         | Variable with the result of converting relative humidity to water vapor pressure |
| τ           | Transmission through atmosphere        |
| raw_ref1    | Radiance reflecting off the object before the window |
| raw_ref1.attn| Attenuated radiance reflecting off the object before the window (in raw units) |
| raw_atm1    | Radiance from the atmosphere before the window |
| raw_atm1.attn| Attenuated radiance from the atmosphere before the window (in raw units) |
| temp_celsius| Estimated temperature (°C) from an infrared thermal imaging file |
\[ \text{emiss.\ wind} = 1 - IRT \]  

(A1)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 65 (accessed on 20 September 2022)

\[ h_2o = RH \cdot e^{(1.5587+0.06959 \cdot \text{ATemp} - 0.0027816 \cdot \text{ATemp}^2 + 0.00000068455 \cdot \text{ATemp}^3)} \]  

(A2)

Reference: MINKINA, Waldemar; DUDZIK, Sebastian. Infrared thermography: errors and uncertainties. John Wiley & Sons, 2009, p. 171; and: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 67 (accessed on 20 September 2022).

\[ \tau = ATX \cdot e^{-\left[ \frac{DB}{2 \cdot \sqrt{ATX+\sqrt{ATX}}} \right]} + (1 - ATX) \cdot e^{-\left[ \frac{DB}{2 \cdot \sqrt{ATX+\sqrt{ATX}}} \right]} \]  

(A3)

Reference: MINKINA, Waldemar; DUDZIK, Sebastian. Infrared thermography: errors and uncertainties. John Wiley & Sons, 2009, p. 171; and: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 69 (accessed on 20 September 2022).

\[ \text{raw} \_\text{ref}1 = \frac{PR1}{PR2 \cdot e^{(\text{ATemp}+273.15) \cdot PF}} - PO \]  

(A4)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 75 (accessed on 20 September 2022).

\[ \text{raw} \_\text{ref}1 \_\text{attn} = \frac{1 - E}{E} \cdot \text{raw} \_\text{ref}1 \]  

(A5)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 76 (accessed on 20 September 2022).

\[ \text{raw} \_\text{atm}1 = \frac{PR1}{PR2 \cdot e^{(\text{ATemp}+273.15) \cdot PF}} - PO \]  

(A6)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 78 (accessed on 20 September 2022).

\[ \text{raw} \_\text{atm}1 \_\text{attn} = \frac{(1 - \tau) \cdot \text{raw} \_\text{atm}1}{\tau \cdot E} \]  

(A7)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 79 (accessed on 20 September 2022).

\[ \text{raw} \_\text{wind} = \frac{PR1}{PR2 \cdot e^{(h_2o \cdot \text{ATemp}+273.15) \cdot PF}} - PO \]  

(A8)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 81 (accessed on 20 September 2022).

\[ \text{raw} \_\text{wind} \_\text{attn} = \frac{\text{raw} \_\text{wind} \cdot \text{emiss} \_\text{wind}}{E \cdot \tau \cdot IRT} \]  

(A9)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 82 (accessed on 20 September 2022).

\[ \text{raw} \_\text{atm}2 = \text{raw} \_\text{atm}1 \]  

(A10)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 87 (accessed on 20 September 2022).

\[ \text{raw} \_\text{atm}2 \_\text{attn} = \frac{(1 - \tau) \cdot \text{raw} \_\text{atm}2}{E \cdot IRT \cdot \tau^2} \]  

(A11)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 88 (accessed on 20 September 2022).

\[ \text{raw} \_\text{obj} = \frac{\text{raw} \_\text{obj}}{E \cdot IRT \cdot \tau^2} \cdot \text{raw} \_\text{atm}1 \_\text{attn} \cdot \text{raw} \_\text{atm}2 \_\text{attn} \cdot \text{raw} \_\text{wind} \_\text{attn} \cdot \text{raw} \_\text{ref}1 \_\text{attn} \]  

(A12)

Reference: https://github.com/gtatters/Thermimage/blob/master/R/raw2temp.R line 90 (accessed on 20 September 2022).

\[ \text{temp} \_\text{celsius} = \frac{PB}{\ln\left[ \frac{PR1}{PR2 \cdot (\text{raw} \_\text{obj} + PO) + PF} \right] + 273.15} \]  

(A13)
24. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–137. https://doi.org/10.1109/TIT.1982.1056489.

25. Bezdek, J.C.; Ehrlich, R.; Full, W. FCM: The fuzzy c-meansclustering algorithm. *Comput. Geosci.* **1984**, *10*, 191–203. https://doi.org/10.1016/0098-3004(84)90020-7.

26. Jaffery, Z.A. Performance comparison of image segmentation techniques for Infrared images. In Proceedings of the 2015 Annual IEEE India Conference (INDICON), Online, 17–20 December 2015; IEEE: New Delhi, India; pp. 1–5. https://doi.org/10.1109/INDICON.2015.7443391.

27. Shanmugam, C.; Sekaran, E.C. IRT image segmentation and enhancement using FCM-MALO approach. Infrared Physics & Technology **2019**, *97*, 187–196. https://doi.org/10.1016/j.infrared.2018.12.032.

28. Aqeel, E.F. *The Use of Threshold Technique in Image Segmentation*; Al-Mustansryah University, Collage of Education Computer Department: Iraq, Baghdad 2010; pp. 3–5.

29. Abdullah, E. A Review on Global Thresholding Methods for Image Binarization. *J. Eng. Sci. Res.* **2020**, *2*, 38–49. https://doi.org/10.46387/bjesr.789534.