Thermography Inspection with Machine Learning for Malfunction Prediction in Power System Equipment

Ana María Garzón*, David Celeita‡**, Trung Dung Le***, Cesar Salazar*
Victor Sicachá*, Natalia Laiton*

* School of Engineering, Science and Technology, Universidad del Rosario, Bogotá, Colombia

** Faculty of Engineering, Universidad de la Sabana, Chía, Colombia

*** Université Paris-Saclay, CentraleSupélec, CNRS, Group of Electrical Engineering Paris (GeePs), Gif-sur-Yvette, France

(anamari.garzon@urosario.edu.co, david.celeita@unisabana.edu.co, trungdung.le@centralesupelec.fr, cesara.salazar@urosario.edu.co, victor.sicacha@urosario.edu.co, natalia.laiton@urosario.edu.co)

‡ Corresponding Author; David Celeita, Faculty of Engineering, Universidad de la Sabana, Chía, Colombia, Tel: +57 314 252 8234, david.celeita@unisabana.edu.co

Received: 13.04.2025 Accepted: 27.08.2025

Abstract- This paper presents a machine learning-based approach for early failure detection in power system equipment using thermographic image analysis. While existing studies often rely on complex deep learning models, our method introduces a lightweight yet effective solution combining image preprocessing, unsupervised segmentation, and classical classifiers. The key innovation lies in the integration of automated Region of Interest (RoI) detection and the evaluation of its impact on model performance. A dataset of 624 RGB thermographic images from motors and transformers is used to benchmark various machine learning algorithms, including KNN, SVM, Decision Tree, Random Forest, Naive Bayes, and Neural Networks. Performance metrics such as precision, recall, and F1-score are analyzed and show that RoI segmentation significantly improves classification accuracy (up to 32% for SVM and 15% for Neural Networks). The proposed method is resource-efficient and achieves strong results without requiring temperature vector data. These findings highlight the practical value of the approach for predictive maintenance and early fault diagnosis in industrial environments.

Keywords- Thermography inspection, image segmentation, predictive fault identification, machine learning, neural network.

1. Introduction

As power systems become increasingly complex, ensuring the reliable and efficient operation of smart equipment is critical to maintaining uninterrupted power supply [1]. Failures in key components can lead to severe consequences, including outages, safety risks, and significant financial losses. In response, machine learning techniques for equipment failure detection have gained considerable attention in recent years [2], [3]. Thermographic inspection data offers valuable insights into the operational state of components such as motors, transformers, and batteries [4], [5], and can be leveraged to build predictive models for early fault detection.

Predictive maintenance plays a vital role in enhancing the efficiency and reliability of power system equipment. Machine learning, when applied to thermographic imagery, has shown significant promise in identifying potential failures before they occur [6], [7]. Thermography is a non-destructive, non-intrusive diagnostic technique that captures surface temperature using infrared radiation. By analyzing thermal patterns, machine learning algorithms can detect anomalies that signal emerging faults [8]. Some studies, particularly in photovoltaic systems, employ deep learning for automated fault diagnosis. For instance, [9] proposes a convolutional neural network (CNN) followed by a residual network (ResNet) to identify and classify faults. In [10] a two-stage approach, where the first stage involves identifying potential

faults using a convolutional neural network (CNN), and the second stage involves diagnosing faults using a residual neural network (ResNet).

In [11] a new dataset for fault detection in electrical equipment is introduced, incorporating interpretable machine learning models like decision trees with feature selection and importance ranking. Other approaches rely on real-time inspection using temperature anomalies and SVM classifiers for fault identification in photovoltaic systems. Similarly, the review in [12] surveys the use of infrared thermography in energy audits, where various studies report high accuracy in fault detection and classification. Collectively, these efforts demonstrate the potential of combining thermal imaging with advanced learning techniques for equipment diagnostics across different energy systems. In overall, these studies demonstrate the potential of using thermal imaging and deep learning for fault diagnosis in photovoltaic systems [13] and energy audits in residential or commercial buildings.

To implement predictive maintenance using thermography, it is essential to collect and label a dataset of thermal images corresponding to the equipment's condition at capture time. These images must be preprocessed to remove noise and ensure consistency. A machine learning model can then be trained using this data with algorithms such as SVM, Random Forest, or Neural Networks, depending on the application's requirements. Once trained, the model can assess the condition of new images, allowing for timely maintenance and minimizing unexpected failures.

In this study, we propose an innovative framework for fault detection in power systems using machine learning applied to thermographic data. We validate the methodology through a real-world case study, demonstrating its potential to enhance reliability and reduce downtime in smart equipment operation. The paper is structured as follows: Section II introduces the types of thermographic inspections used in power systems. Section III details the methodology, comprising data acquisition, feature extraction, and machine learning-based classification. Section IV presents the case study and section V results and discussion. Finally, section VI summarizes certain challenges, limitations, and future work. Section VII concludes with final remarks, highlights of the contributions and future perspectives.

2. Types of Thermographic Inspection of Power Systems

When carrying out a literature review of the monitoring of power systems with IR images, it is possible to notice that there are two types of images: images obtained in closed spaces and images obtained in open spaces. Such classification is proposed on our approach (see Figure 1) to objectively categorize the methods depending on the physical environment of the power system equipment. Those located in open spaces typically monitor solar panels [14] and photovoltaic cells, or electrical transmission lines. The images are taken either by locating the camera in drones, or in a high or distant fixed point.

Two points must be considered in the preprocessing of these images: The first is the distinction of the background from the system or equipment to be monitored.

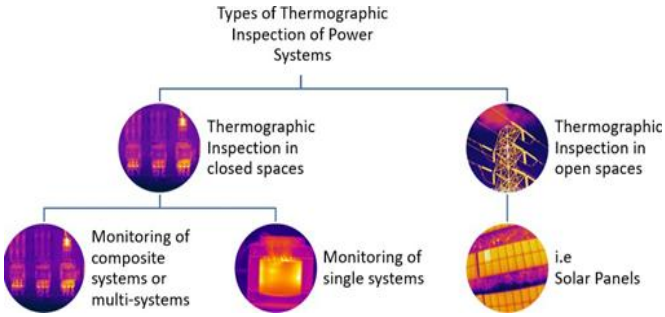

Fig. 1. Types of thermographic inspection of power systems.

Unlike the images that are taken in closed spaces, it must be considered that in this case the light and the temperature of the environment are variable and impossible to control. Thermal imaging cameras are sensitive to these factors which can make segmentation more difficult. On the other hand, the perspective in which the images have been taken is also an important point in their preprocessing.

The second point of the preprocessing is the distinction of the parts or regions of interest for monitoring once the background has been separated. In the case of solar panels, it is convenient to obtain information from each photovoltaic cell separately, and such location of the sensor must be carefully analyzed [15], some approaches have successfully detected snow layer with deep learning [16]. For example, this requires additional work if detailed information about the possible effect of faults in such equipment that can particularly change the production rate and efficiency [17], [18].

On the other hand, monitoring in closed spaces is typically carried out under constant ambient temperature and lighting conditions, which allows a better generalization to be made to distinguish the object from the background. Under this condition, simple systems, such as a single motor or transformer [19] where some ML-based approaches have demonstrated to be helpful for utilities to reduce the cost of preventative maintenance [20], or complex compound systems that include multiple machines can be monitored. In the second case, in addition to distinguishing the region of interest from the background, it is necessary to identify the different parts of the image to obtain a localized and accurate prediction.

2.1. Comparison with State-of-the-Art Approaches

Several recent studies have explored the use of machine learning for failure detection in power system equipment using thermographic imaging. For example, Manno et al. [10] proposed a deep learning strategy based on convolutional neural networks (CNN) and ResNet architectures for automatic fault diagnosis in photovoltaic systems. Similarly, Najafi et al. [11] introduced a dataset and applied interpretable ML models such as decision trees with feature selection for fault diagnosis in electrical equipment. Bommes et al. [9] adopted supervised contrastive learning for detecting anomalies in IR images of PV modules, while Cubukcu and Akanalci [13] explored real-time thermal inspection techniques using SVM. Compared to these approaches, our method offers several advantages:

- Low computational complexity: Unlike deep learning models that require large datasets and high computing power, our approach works effectively with limited image data (RGB bmp files), achieving competitive results with classical algorithms (KNN, SVM, Decision Tree, etc.).
- Segmentation-enhanced performance: Our method explicitly separates the region of interest using an automated segmentation pipeline, which significantly improves classification accuracy (up to 17% difference in Neural Networks for the motor case).
- Scalable vector representation: The technique of converting RGB images to vectors allows efficient storage and processing, enabling lightweight deployment in industrial systems.
- Hardware simplicity: The method is implemented with non-invasive thermographic inspections and standard IR cameras, making it viable for broad industrial use.

However, some limitations remain for further research:

- No use of temperature vector (T): While the dataset includes temperature data, only RGB channels were used, potentially limiting precision in fine-grained classification.
- Generalization across environments: The method performs well in controlled conditions but may require further adaptation to variations in lighting, background, and camera angle in real-world deployments.
- Absence of deep learning architectures: Though efficient, the method may not match the predictive power of large neural networks trained on extensive labeled data.
- These trade-offs position the proposed method as a lightweight yet effective alternative to more resource-intensive deep learning-based solutions, particularly suited for small-to-medium scale predictive maintenance deployments.

3. Methodology

The proposed methodology is designed to detect and classify equipment faults in both open and closed power system environments using thermographic image analysis and machine learning. The process comprises four core stages: (1) image acquisition, (2) image preprocessing, (3) region of interest (RoI) segmentation, and (4) classification using machine learning algorithms.

A graphical summary of the full pipeline is presented in Figure 2 to improve clarity and reproducibility. In the image acquisition stage, thermal images are captured using an infrared (IR) camera positioned at an appropriate distance and angle, depending on whether the system is indoors (controlled lighting) or outdoors (variable lighting). The dataset includes 624 images captured under a constant ambient temperature of 23°C, covering transformers and motors under both healthy and faulty conditions.

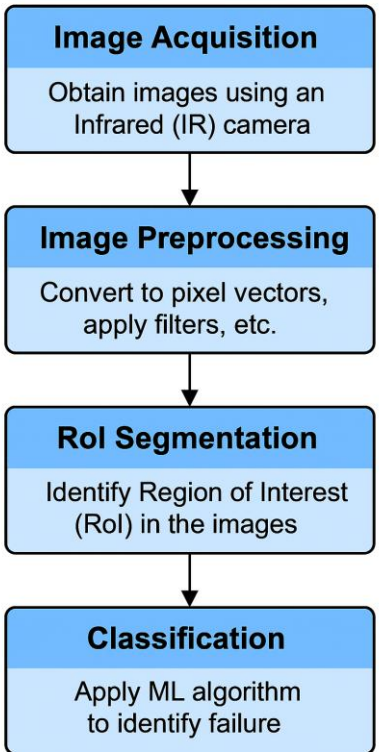

Fig. 2. Steps for automated detection of faults in power systems using IR images.

During the image preprocessing phase, the images are converted from RGB to grayscale or remain in RGB, depending on the equipment category. Vectorization of image data is then performed, converting each image into a structured pixel array suitable for machine learning inputs. Additional enhancements, such as contrast adjustment, morphological filtering, and color space conversion, are applied to highlight features relevant for fault detection.

The next step involves Region of Interest (RoI) segmentation, which is crucial to isolating the components most indicative of failure. Manual segmentation is feasible when the system and camera configuration are static. However, for broader scalability, automated segmentation is implemented using a pipeline that includes K-Means clustering, Otsu thresholding, histogram equalization, shadow removal, and Canny edge detection for identifying components such as cables and connectors.

In the final stage, classification, the segmented image data is fed into various machine learning algorithms including K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Decision Trees, Random Forests, Naive Bayes, and a simple Neural Network model. The goal is to determine whether the equipment is functioning correctly or experiencing a specific type of fault. This multi-model comparison ensures robustness and highlights the performance gains provided by segmentation. A complete overview of algorithm configurations, evaluation metrics, and results is available in our open-access GitHub repository [21].

To ensure reproducibility and transparency, the entire implementation (including data preprocessing routines, segmentation scripts, classification models, and evaluation

procedures) has been shared at 24. All models were implemented in Python 3.9 using OpenCV, Scikit-learn, and TensorFlow. The experiments were conducted on a workstation equipped with an Intel Core i7 processor, 32 GB RAM, and an NVIDIA RTX 3060 GPU. Hyperparameters such as k=5 for KNN, a maximum depth of 10 for decision trees, and a neural network trained for 100 epochs with a batch size of 32 are documented in the repository's notebooks.

4. Case Study Description: Closed Space, Single System

In this paper, the procedure carried out in [11] is analyzed, implemented, and modified, using the dataset created for that case study.

The dataset consists of 624 RGB thermal images of 320x240 pixels, of which 369 correspond to images of induction motors and 255 to images of transformers. The images were taken by a Dalitech T4/T8 infrared thermal image camera at 23° Celsius of environment temperature. This instrument generates images with four channels (RGBT), where the first three channels are RGB, and the fourth channel is the temperature matrix T. However, in this study, we only work with the RGB channels, since images of the dataset were acquired in a controlled environment of 23° Celsius.

Regarding the transformer, 8 cases of short circuit failures in common core winding are considered to generate the faults artificially; while for the motor case, 8 different cases of artificial generated stuck rotor fault, cooling fan failure and stator windings failure are considered. The objective of both the investigation carried out by the researchers and by the present manuscript is to find the best way to distinguish and classify the failures described above.

The first step on this path is the pre-processing of the data. The images are converted to grayscale, a projection of RGB space in which each pixel takes a single value from 0 to 255, instead of taking 3 different component values. The advantage of working in grayscale is that this dimensionality reduction decreases the time and computational cost of the algorithms. Regarding the temperature, this is directly proportional to the pixel value. However, it should be considered that this reduction could lead to loss of information. [11] proposes to classify "cold" and "hot" images before segmenting the region of interest and the same will be done here. To carry out the classification, the data was divided into training and test sets, the pixels with the highest value from each of the images were extracted, and the behavior of these values in the training set was explored.

Figure 3 shows the range of maximum pixel intensities for cold and hot transformer images. A clear separation between the two groups highlights the effectiveness of using pixel intensity thresholds for binary temperature classification prior to segmentation. It is noticeable that there is a difference between the maximum value of the cold images and the hot images. Similar results are obtained for the motor. To perform a classification, use intermediate values (mean or median) to generate a classification threshold, establish a threshold through observation, or use a support vector machine classification that finds a hyperplane (or line) that conditions the classification.

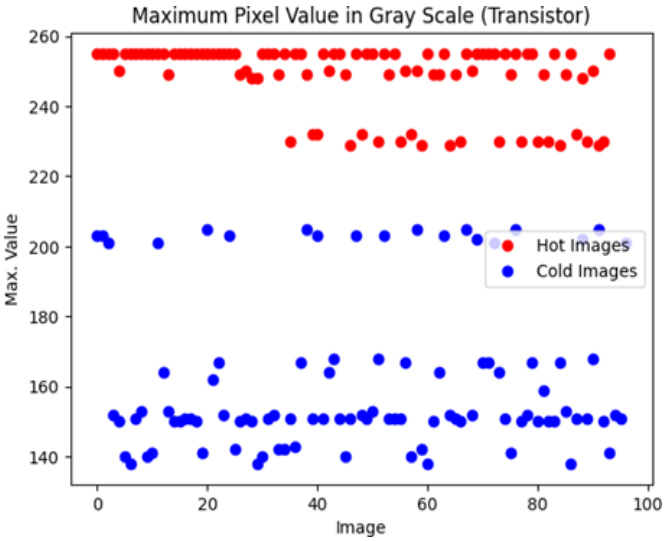

Fig. 3. Maximum pixel value (transformer dataset).

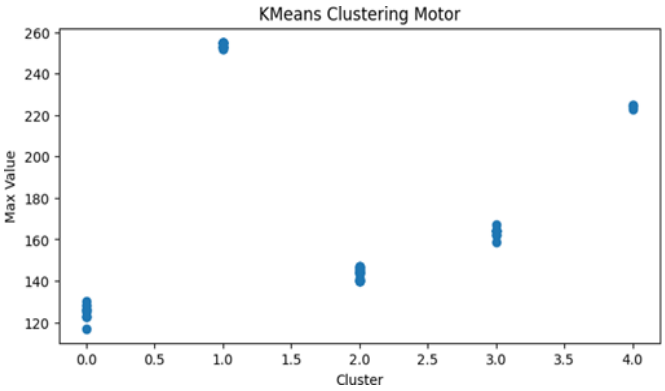

Fig. 4. K-Means clustering results for motor images (5 clusters).

The results obtained when using linear SVM are a perfect classification for both cases (which is not surprising since the classes are quite distant). However, from the plot of the train data, it is evident that the data is grouped in more than two subgroups focusing only on the maximum value pixel. In this sense, additional exploration of the dataset is done with KMeans, trying to distinguish those subgroups and the number of clusters is determined through the elbow method.

In the motor case study, the optimal number of clusters is 5; 3 clusters for cold images and 2 clusters for hot images as shown in Figure 4 that demonstrates how motor images naturally group into five clusters—three for cold conditions and two for hot—based on maximum pixel values. This clustering serves as a scalable preprocessing step for guided segmentation and classification. Later, the results for the transformer case study can show 6 clusters, therefore the methodology can be scalable.

5. Results and Discussion

By comparing the results obtained above with the failure that corresponds to each value, it is noticeable that the images that have the same fault label often are in the same cluster. It can be inferred that the maximum value pixel is a good guide

not only to create an effective preprocessing that facilitates the segmentation of the region of interest but also to approximate the final classification.

5.1. Identifying RoI (Region of Interest)

From the previously identified clusters, an image is chosen at random from each of the previous groups to create a manual mask of the region of interest, which will allow evaluation of the results obtained by the automation. The segmentation challenge, in this case, is that the images were taken at different angles, on a surface that reflects the light emitted by the equipment, and that includes cables external to the equipment that should be included in the RoI since they provide information about the state of the phases. In this sense, the algorithm to find the RoI follows the following steps:

5.2. Initial Segmentation Setting a Threshold

Segmentation of an image in two regions can be done by setting a threshold k such that if a pixel has a value that is greater than k it belongs to a class and if not, it belongs to the other. The threshold can be defined manually or by numerical methods by identifying inflection, local maximum, or minimum points in the histogram of the distribution of pixel intensity. However, it is a non-optimum procedure on a large scale. Then, a threshold function *f (x, y)* must be defined such that if a pixel with coordinates *(x, y)* evaluated in f has a value greater than *k* belongs to one region *R1*, and if the value is less or equal to *k* it belongs to the other region *R2*. This procedure can be summarized by the following function:

$$g(x,y) = \begin{cases} 1 \quad if \quad f(x,y) > k \\ \\ 0 \quad if \quad f(x,y) \leq k \end{cases} \tag{1}$$

The problem now consists of finding the most optimal and general function f and the threshold *k*. As explained in [22], it has been sought to use a PDF probability density function, but the limitation is that the calculation of functions requires many distribution assumptions that are not always fulfilled and cannot be easily calculated for atypical distributions.

A solution to this problem exposed by the same source [22], is Otsu's method. It is an optimum procedure where f maximizes the variance between classes, basically carrying out a statistical method of discriminant analysis on the intensity of the pixels. An advantage of the Otsu's method highlighted by the authors is that its results are obtained from the histogram giving, as a result, an array of dimension 1. Access the complete process and the specifications of the Otsu formulas in the following repository: https://github.com/anamarigarzon/IR-Image-Analysis-in-Power-Systems

In this project, the cluster to which the image belongs is considered. Similar and satisfactory results are usually obtained for hot images by applying Otsu directly. For cold images, a histogram equalization procedure is performed, which standardizes the image values, some morphological erosion transformations and finally Otsu is applied.

There are multiple ways of finding this threshold function, which vary depending on the image distribution. A useful and simple but very manual way to establish it is to plot the histogram of the distribution of the pixel values in the image and determine possible segmentation points.

5.3. Shadow Elimination and Addition of The Missing Equipment Area

Next, a new Otsu segmentation is performed on the portion of the image that is outside the mask obtained in the previous step. In the images of the two clusters with a smaller maximum pixel value, the problem is that the equipment section has a temperature very close to that of the background, so it is advisable to perform morphological operations and extra contrast enhancement to increase differentiation. Figures 5 and 6 illustrate the sequential stages of RoI segmentation using Otsu's method and shadow correction. In Figure 5, the initial segmentation highlights the primary thermal region of interest using an automated threshold. However, the output still includes shadows and partially missed areas (especially in cold clusters where thermal contrast is low). To address this, Figure 6 shows the results after applying additional morphological operations and histogram equalization, followed by a second Otsu threshold. This refinement enables the algorithm to recover missed regions and exclude background reflections, resulting in a more accurate and complete RoI mask. This step is particularly critical when working with cold images where temperature gradients are less distinct and prone to misclassification without enhancement.

In hot images, the shadow reflected by the surface has an intensity that is close to the missing section of the image. So, in this case, the procedure to follow consists of performing an erosion of n iterations according to the image cluster, until these sections separate, and it is chosen to keep the section with the largest area. Finally, the section is dilated n times and added to the previous segmentation.

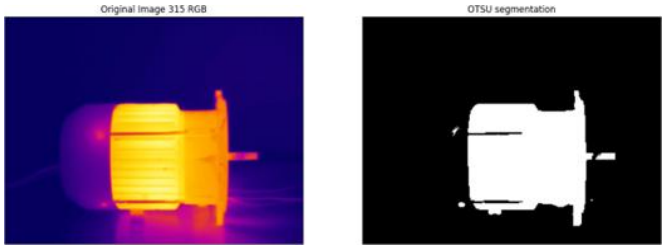

Fig. 5. Initial Otsu segmentation applied to a cold motor image.

Fig. 6. Enhanced segmentation with shadow removal and morphological operations.

5.4. Cable Segmentation

The segmentation after the shadow elimination ignores key elements due to their lower contrast with respect to the background, and those are the cables. It is important to consider the cables in the segmentation because they provide important information about the status of the phases.

Figure 7 compares the original grayscale thermographic image (left), the automatically generated final mask (center), and the manually annotated mask (right) used as ground truth for evaluation.

This example highlights the effectiveness of the proposed segmentation pipeline in replicating expert annotations. Despite challenges such as thermal reflection and cable inclusion, the final mask closely matches the manual mask—capturing not only the motor body but also finer details like cable extensions. This alignment validates the precision of the automated RoI segmentation strategy, which is critical for downstream classification accuracy. The segmentation quality is later quantified using the Jaccard Index, as shown in Figures 10 and 11. To identify them, the image is divided into regions in such a way that for the largest number of images the cables are in the same square region, and once this division is made, the Canny Edge Detector is used on those areas to find the cables [23]. Canny edge detection is a popular edge detection algorithm developed by John F. Canny that first removes noise using a 5x5 Gaussian filter. As it is described in the OpenCV documentation, it then finds the intensity gradient of the image using a sobel kernel on both the X and Y axes. Then the first derivative is obtained in the horizontal and vertical direction.

Then follows the Non-maximum Suppression stage, in which all the pixels of the image are scanned and those that do not constitute the edges are removed. For this, it is checked if each pixel is a local maximum in its surroundings in the direction of the gradient. If it is a maximum, it is considered for the next stage, if not, its value changes to zero.

Table 1 and Table 2 compare classification accuracy for various machine learning models applied to motor images, with and without RoI segmentation. The results show that most models benefit from segmentation, with the most notable improvements observed in Support Vector Machines (SVM), which gained +0.24 in accuracy, and Neural Networks, which improved by +0.17. These gains confirm that isolating the relevant thermal region enhances the model's ability to distinguish fault patterns. In contrast, Naive Bayes showed a slight performance drop (−0.04), likely due to its strong assumptions of feature independence, which may not hold in segmented thermal data. Overall, the results suggest that segmentation significantly boosts model robustness for motor fault detection.

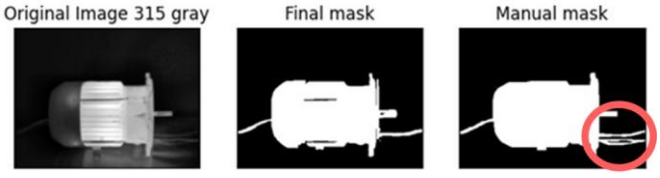

Fig. 7. Comparison between original image, generated mask, and manual ground truth (motor).

Table 1. Comparison of accuracy of methods with and without segmentation over the motor images

Method	Accuracy with Segmentation	Accuracy with no Segmentation	Difference of Accuracy
KNN	0.98	0.93	0.05
SVM	0.87	0.63	0.24
Decision tree	0.98	0.95	0.03
Random Forest	0.98	0.97	0.01
Naive Bayes	0.77	0.81	-0.04
Neural Network	0.81	0.64	0.17

Table 2. Comparison of accuracy of methods with and without segmentation over the transformer images

Method	Accuracy with Segmentation	Accuracy with no Segmentation	Difference of Accuracy
KNN	1.00	0.98	0.01
SVM	0.63	0.56	0.07
Decision tree	1.00	0.96	0.04
Random Forest	1.00	1.00	0.00
Naive Bayes	0.86	0.76	0.10
Neural Network	0.90	0.60	0.3

Figures 8 and 9 illustrate the complete Region of Interest (RoI) segmentation procedure for a motor (grayscale) and a transformer (RGB), respectively. These visualizations provide insight into how the proposed pipeline processes images with different thermal profiles and color spaces. In both cases, the segmentation algorithm successfully isolates the main equipment body and relevant peripheral elements, such as cables and connectors, which are essential for fault diagnosis.

The motor example (Figure 8) shows the effectiveness of grayscale preprocessing combined with clustering and edge detection, while the transformer example (Figure 9) demonstrates how RGB color space processing handles more complex backgrounds and reflections. These figures confirm that the segmentation framework is flexible and adaptable to different image types and equipment configurations, key for generalizing across applications in industrial power systems.

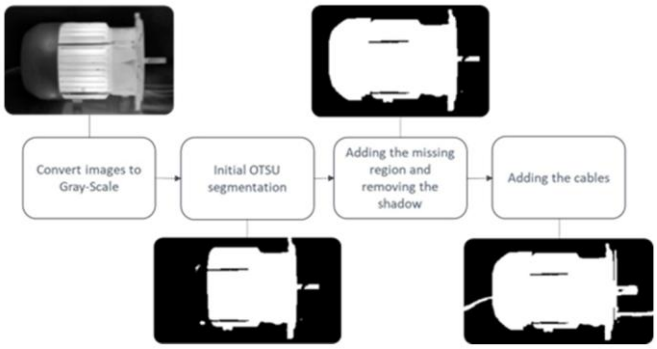

Fig. 8. RoI segmentation pipeline applied to motor image (grayscale).

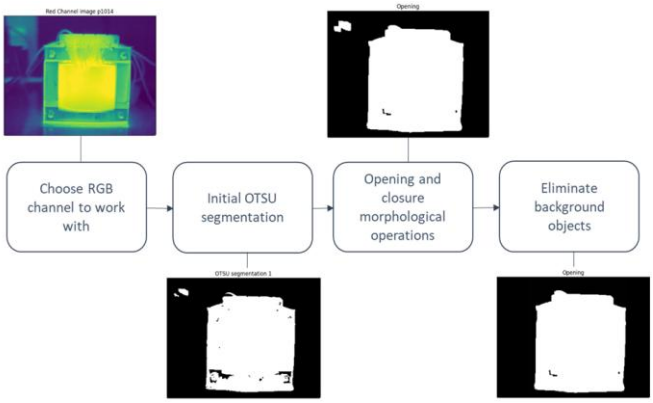

Fig. 9. RoI segmentation pipeline applied to transformer image (RGB).

The final stage is the Hysteresis Thresholding stage, which defines which edges are truly edges and which are not. Two values are needed for this stage, a minimum and all pixels below that minimum are discarded, and a maximum and all pixels above that maximum are considered safe edges. Those that are between both values are classified as edges or non-edges based on their connectivity with neighboring pixels. The precision of the masks generated by the previous process are measured with the Jaccard Score. This index compares the similarity between the manual masks and its respective automated masks. The Basic Jaccard Index according to [24] can simply be expressed as:

$$\mathcal{J}(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| + |A \cap B|} \qquad (2)$$

Where A, and B are two sets. In this context, |A∩B| is the number of pixels of the image that have the same intensity, and |A ∪ B| is the size of the image. The results obtained for the 10-compared masks (one for each of the clusters) are between 0.73 and 0.81 for cold images and between 0.87 and 0.94 for hot images.

Note the Jaccard scores for both equipment in Figures10 and 11, where independently from the quantity of identified clusters (five clusters for the motor with grey-scale, and six for the transformer with RGB), the results shown an accurate performance. In the case of the motor, the Jaccard score is located in between 79.133% to its maximum 89.765%. In the case of the transformer, the Jaccard score behaves better with a range from 84.076% to its maximum 90.607%.

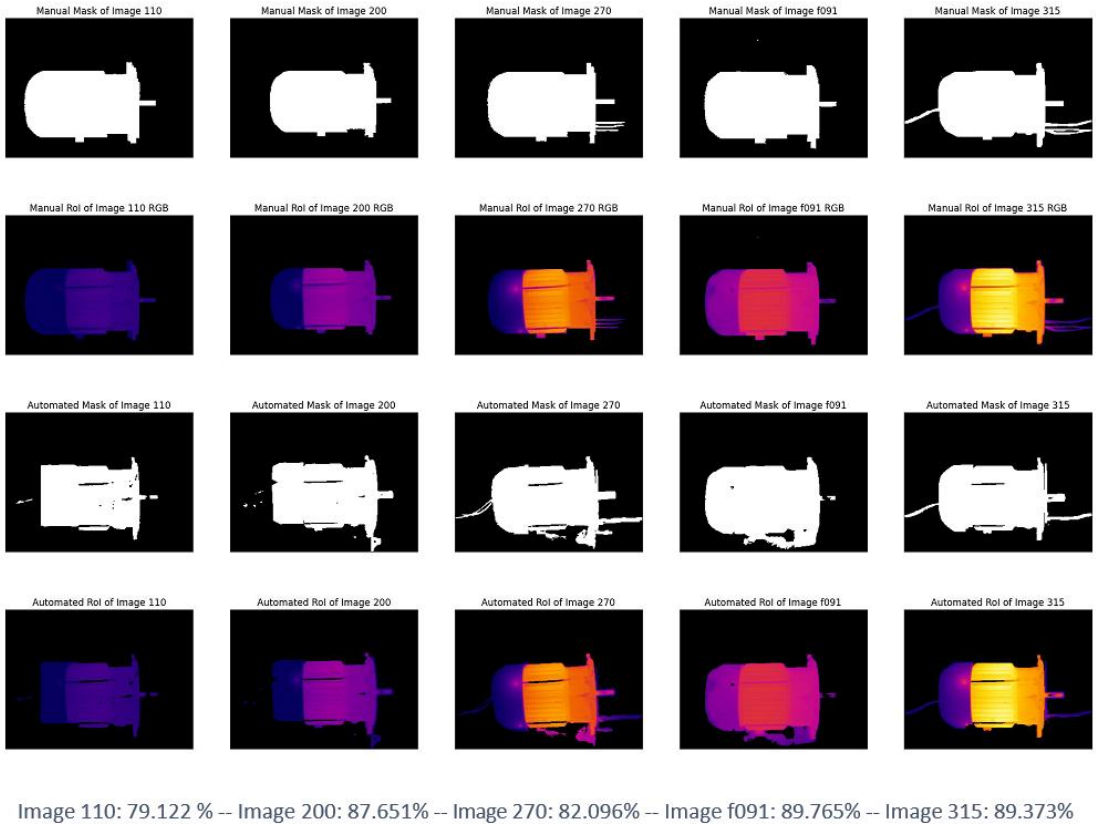

Fig. 10. Jaccard scores across motor image clusters: RoI identification, results of Jaccard scores and K-Means clustering - Jaccard scores per image - motor with grey-scale and five clusters.

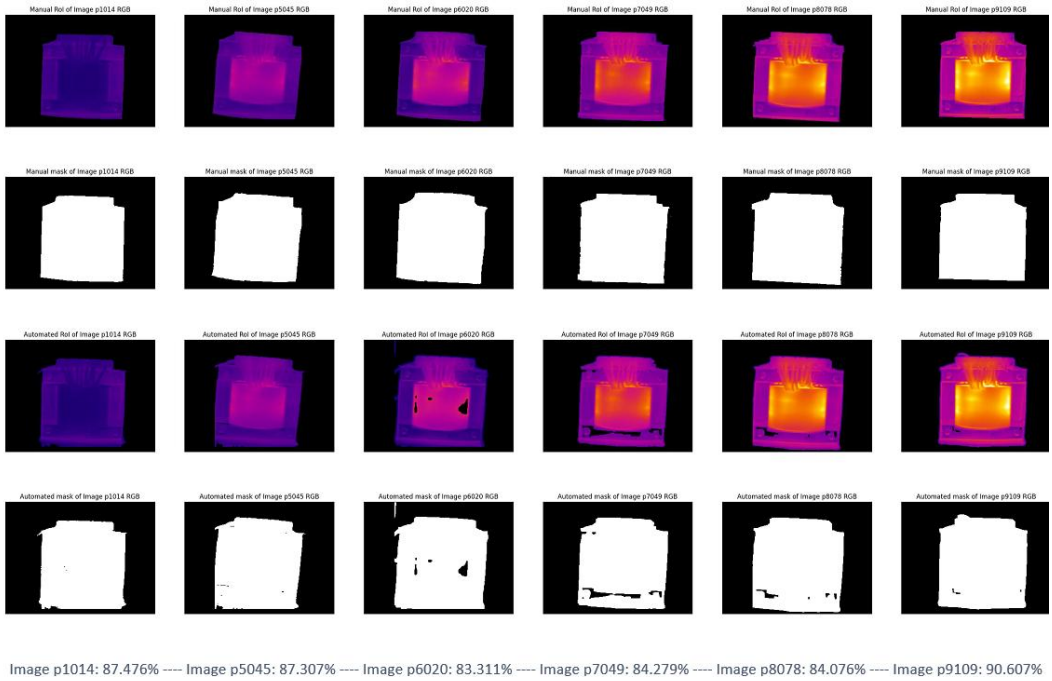

Fig. 11. Jaccard scores across transformer image clusters.: RoI identification, results of Jaccard scores and K-Means clustering - Jaccard scores per image - transformer with RGB and six clusters.

As can be seen in the comparison table for the motor I and the transformer II, the classification results have a favorable balance in most of the methods explored, with those based on trees having the highest accuracy. When comparing the classification results obtained with the images before being segmented (as they were in the original dataset) using the same methods and the same variables, we can notice a large improvement in the segmented images, with a difference of up to 0.17 percentage points. This helps us confirm the hypothesis that classification improves when the region of interest is separated from the background and is a great finding for future implementations.

5.5. Performance Metrics

To further evaluate the impact of Region of Interest (RoI) segmentation on classification accuracy, we computed precision, recall, and F1-score across six machine learning models for both segmented and non-segmented thermographic images. These weighted average metrics provide a comprehensive perspective on model performance beyond simple accuracy values.

5.5.1. Impact of segmentation

The results clearly demonstrate that RoI segmentation enhances model performance in most cases. For instance, the Support Vector Machine (SVM) classifier shows a substantial improvement:

- Precision: from 0.46 to 0.83
- Recall: from 0.64 to 0.88
- F1-score: from 0.52 to 0.84

Similarly, Neural Networks benefit from segmentation, improving F1-score from 0.63 to 0.78. Tree-based models (Decision Tree and Random Forest) already performed well without segmentation but still showed modest gains with RoI isolation (see Table 3). The only exception is the Naive Bayes model, which showed a slight performance drop, likely due to its simplistic assumptions and sensitivity to class overlap after segmentation.

Table 3. Comparison of performance metrics

Method	F1-Score without Segmentation	F1-Score with Segmentation	Δ F1-Score
KNN	0.94	0.99	+0.05
SVM	0.52	0.84	+0.32
Decision tree	0.96	0.99	+0.03
Random Forest	0.97	0.99	+0.02
Naive Bayes	0.77	0.71	-0.06
Neural Network	0.63	0.78	+0.15

5.5.2. Bar chart comparison of weighted average metrics

Figures 12, 13 and 14 display the weighted average precision, recall, and F1-score for all models, comparing the results with and without segmentation. These visualizations reinforce that segmentation not only boosts classification performance in most models but does so consistently across multiple metrics.

This evaluation confirms that accurate RoI segmentation plays a vital role in improving fault detection capabilities. It not only simplifies the classification task by removing background noise but also helps the algorithms focus on thermally relevant features, especially when using models sensitive to input dimensionality and distribution.

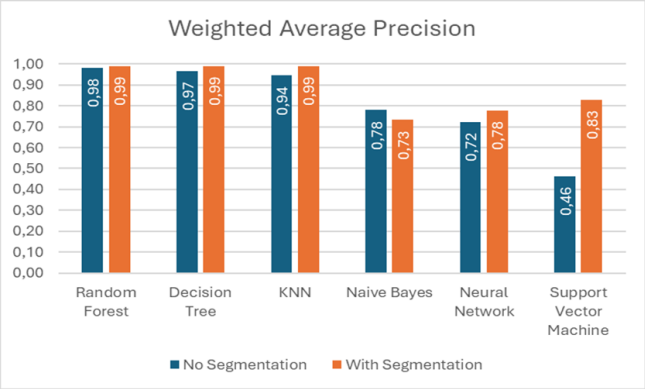

Fig. 12. Comparison of performance metrics: weighted average precision.

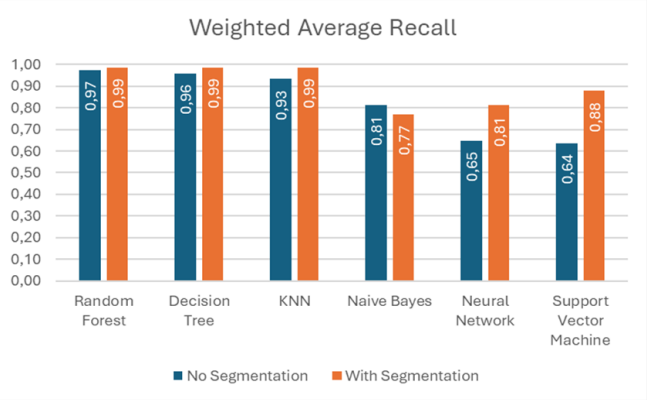

Fig. 13. Comparison of performance metrics: weighted average recall.

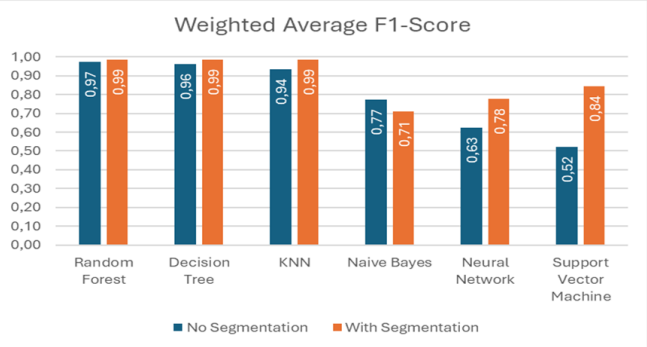

Fig. 14. Comparison of performance metrics: weighted average f1-score.

6. Challenges, Limitations, and Future Work

While the proposed method demonstrates strong performance in failure detection using thermographic image analysis and classical machine learning, certain challenges and limitations remain. First, the approach currently uses only the RGB components of the thermographic images, without leveraging the temperature matrix (T channel) available in the raw dataset. Although this simplification allows for more lightweight processing and storage, it may limit the precision of fault localization and classification, especially in edge cases with subtle thermal variations. Integrating the temperature channel in future iterations could improve granularity and diagnostic power. Second, the method was validated using a dataset captured under controlled ambient conditions (23 °C, fixed camera setup). As a result, its robustness under varying environmental conditions (such as outdoor lighting changes, occlusions, different camera angles, and background clutter) has not yet been tested. Adapting the segmentation pipeline for more dynamic settings, including real-world substations or field-deployed equipment, remains a key step for operational deployment.

Additionally, while the current segmentation and classification processes show high accuracy with classical ML models, they rely on handcrafted thresholds and feature selection. Future work may explore the use of convolutional neural networks (CNNs) and transfer learning to automate feature extraction and improve generalization across fault types and equipment classes. Finally, although manual masks were used to evaluate segmentation quality, automating this validation process using labeled ground truth in larger datasets would allow for broader and more objective performance assessment. These directions offer a roadmap for expanding the applicability, scalability, and precision of the proposed solution.

Another promising direction for future work involves integrating the temperature matrix (T-channel) captured by the thermal imaging camera. Although the current approach uses only the RGB components of thermographic data to ensure lightweight processing and general applicability, the T-channel offers access to absolute pixel-level temperature values. By combining these thermal gradients with spatial and textural features from the image, the classification model could better detect subtle fault conditions, especially in early-stage anomalies where visual cues are minimal. This multimodal fusion (incorporating both visual and thermal dimensions) has the potential to increase diagnostic accuracy and robustness, particularly in applications involving complex equipment or varying ambient conditions.

7. Conclusions

The way of encoding and decoding the images (in vectors with their RGB values) is a very efficient way that takes up a minimum amount of space. It is convenient to save images in the form of strings in databases, to use a great quantity of information without having to make large use of memory.

The initial binary classification by temperatures (cold-hot) recommended by the authors of the dataset is very effective not only in carrying out a pre-classification of the failure but also in determining what methods to use to segment the images and achieve better Jaccard scores. The segmentation was satisfactory and for future work, it can be adapted to different environments so that similar results are obtained in places with variable lighting or camera angles. The classification methods obtained the results documented in the literature both for the case of segmented images and for non-segmented images, which means that the best results are obtained with tree-based methods. It is possible to notice the advantages in classification with segmented images by observing how in all cases except one, the accuracy improves significantly in comparison with non-segmented images. This

tells us that the procedure carried out previously is a huge advantage when it comes to identifying failures, regardless of the method.

The proposed method has advantages over newer proposals, and that is that although we know that more precise and generalized segmentations can be obtained using large Neural Network models, the standard presented in this paper is efficient, economical and does not require sophisticated tools to execute. It should be noted that the results obtained were obtained only from the image data (bmp files) but that thermographic images usually include a temperature vector. The results that were obtained are very favorable despite this fact, and it should be considered that with this additional information, the given methods must work more precisely and efficiently.

Acknowledgements

This work was funded by the IEEE Foundation Grant 94012 - IAS Zucker Faculty Grant 2022 “Smart Industrial Power Systems with Advanced Data Analytics for Proactive and Predictive Maintenance” and it was partially funded by the starting grant IV-TFA056 entitled “Machine learning for Smart Energy Systems” Research Direction at Universidad del Rosario. The authors express their sincere gratitude to Universidad de La Sabana for their invaluable support in the development of this paper. This work was also supported by the project "Synergizing Data and Simulation: Digital Twins and Cloud Environments for Acceleration and Optimization of Energy Transition (SDS-DTC4ET)" at the Faculty of Engineering, Universidad de La Sabana.

The authors would like to thank to the Center of Resources for Learning and Research (CRAI) at Universidad del Rosario for their help with the heuristic state-of-the-art for this manuscript.

References

[1] T. Ahmad, H. Zhu, D. Zhang, R. Tariq, A. Bassam, F. Ullah, S. AlGhamdi, and S. S. Alshamrani, “Energetics systems and artificial intelligence: Applications of industry 4.0,” Energy Reports, vol. 8, pp. 334–361, 2022.

[2] A. S. Relkar, “Risk analysis of equipment failure through failure mode and effect analysis and fault tree analysis,” Journal of Failure Analysis and Prevention, vol. 21, no. 3, pp. 793–805, Jun 2021.

[3] V. Galaz, M. A. Centeno, P. W. Callahan, A. Causevic, T. Patterson, I. Brass, S. Baum, D. Farber, J. Fischer, D. Garcia, T. McPhearson, D. Jimenez, B. King, P. Larcey, and K. Levy, “Artificial intelligence, systemic risks, and sustainability,” Technology in Society, vol. 67, p. 101741, 2021.

[4] D. López-Pérez and J. Antonino-Daviu, “Application of infrared thermography to failure detection in industrial induction motors: Case stories,” IEEE Transactions on Industry Applications, vol. 53, no. 3, pp. 1901–1908, 2017.

[5] M. Salameh, S. Wilke, B. Schweitzer, P. Sveum, S. Al-Hallaj, and M. Krishnamurthy, “Thermal state of charge estimation in phase change composites for passively cooled lithiumion battery packs,” IEEE Transactions on Industry Applications, vol. 54, no. 1, pp. 426–436, 2018.

[6] S. Han, F. Yang, H. Jiang, G. Yang, D. Wang, and N. Zhang, “Statistical analysis of infrared thermogram for cnn-based electrical equipment identification methods,” Applied Artificial Intelligence, vol. 36, no. 1, p. 2004348, 2022.

[7] A. M. Garzón, N. Laiton, V. Sicachá, D. F. Celeita, and T. D. Le, “Smart equipment failure detection with machine learning applied to thermography inspection data in modern power systems,” in 2023 11th International Conference on Smart Grid (icSmartGrid), 2023, pp. 01–05.

[8] T. Mian, A. Choudhary, and S. Fatima, “Multi-sensor fault diagnosis for misalignment and unbalance detection using machine learning,” IEEE Transactions on Industry Applications, vol. 59, no. 5, pp. 5749–5759, 2023.

[9] L. Bommes, M. Hoffmann, C. Buerhop-Lutz, T. Pickel, J. Hauch, C. Brabec, A. Maier, and I. Marius Peters, “Anomaly detection in IR images of PV modules using supervised contrastive learning,” Progress in Photovoltaics: Research and Applications, vol. 30, no. 6, pp. 597–614, 2022.

[10] D. Manno, G. Cipriani, G. Ciulla, V. Di Dio, S. Guarino, and V. Lo Brano, “Deep learning strategies for automatic fault diagnosis in photovoltaic systems by thermographic images,” Energy Conversion and Management, vol. 241, p. 114315, 2021.

[11] M. Najafi, Y. Baleghi, S. A. Gholamian, and S. Mehdi Mirimani, “Fault diagnosis of electrical equipment through thermal imaging and interpretable machine learning applied on a newly-introduced dataset,” in 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), 2020, pp. 1–7.

[12] E. Lucchi, “Applications of the infrared thermography in the energy audit of buildings: A review,” Renewable and Sustainable Energy Reviews, vol. 82, pp. 3077–3090, 2018.

[13] M. Cubukcu and A. Akanalci, “Real-time inspection and determination methods of faults on photovoltaic power systems by thermal imaging in turkey,” Renewable Energy, vol. 147, pp. 1231–1238, 2020.

[14] L. Cardinale-Villalobos, R. Rimolo-Donadio, and C. Meza, “Solar panel failure detection by infrared uas digital photogrammetry: a case study,” International Journal of Renewable Energy Research, vol. 10, p. 114315, 2020.

[15] H. Oufettoul, S. Motahhir, G. Aniba, and I. Ait Abdelmoula, “Sensor placement strategy for locating

photovoltaic array failures," in 2022 11th International Conference on Renewable Energy Research and Application (ICRERA), 2022, pp. 360–368.

[16] O. Ozturk, B. Hangun, and O. Eyecioglu, "Detecting snow layer on solar panels using deep learning," in 2021 10th International Conference on Renewable Energy Research and Application (ICRERA), 2021, pp. 434–438.

[17] M. R. Tur, I. Colak, and R. Bayindir, "Effect of faults in solar panels on production rate and efficiency," in 2018 International Conference on Smart Grid (icSmartGrid), 2018, pp. 287–293.

[18] K. Natarajan, P. K. Bala, and V. Sampath, "Fault detection of solar pv system using svm and thermal image processing," International Journal of Renewable Energy Research, vol. 10, p. 114315, 2020.

[19] S. A. Saleh, M. E. Valdes, P. E. Sutherland, M. Haj-Ahmed, and E. W. Zundel, "Phaselet-based arc flash relay against low voltage side arcing current faults in mv-lv power transformers," IEEE Transactions on Industry Applications, vol. 59, no. 6, pp. 7462–7474, 2023.

[20] J. Penfield and M. Holland, "Machine learning based online monitoring of step-up transformer assets in electrical generating stations," in 2021 9th International Conference on Smart Grid (icSmartGrid), 2021, pp. 32–36.

[21] A. M. Garzón, "Predictive and Proactive Maintenance in Power Systems – IR Image Analysis," GitHub repository, 2024. https://github.com/anamarigarzon/Predictive_and_Proactive_Maintenance_in_Power_Systems

[22] R. C. Gonzalez and R. E. Woods, Image Segmentation: Thresholding. Pearson, 2018.

[23] J. Lee, H. Tang, and J. Park, "Energy efficient canny edge detector for advanced mobile vision applications," IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 4, pp. 1037–1046, 2018.

[24] L. da F. Costa, "Further generalizations of the jaccard index," CoRR, vol. abs/2110.09619, 2021.

Appendix

Pseudocode: RoI Segmentation Using Otsu's Method with Preprocessing

```
Input: RGB image I(x, y)

1. Convert I(x, y) to grayscale if needed:
      I_gray(x, y) ← RGB2Gray(I)

2. If image cluster == "cold":
      I_eq(x, y) ← HistogramEqualization(I_gray)
      I_proc(x, y) ← MorphologicalErosion(I_eq)
   Else:
      I_proc(x, y) ← I_gray

3. Apply Otsu's thresholding:
      k ← OtsuThreshold(I_proc)
      Mask_initial(x, y) ← 1 if I_proc(x, y) > k else 0

4. Shadow correction (for cold images):
      I_shadow ← I_gray − Mask_initial
      k2 ← OtsuThreshold(I_shadow)
      Mask_shadow(x, y) ← 1 if I_shadow(x, y) > k2 else 0
      Mask_combined ← Mask_initial ∪ MorphologicalDilation(Mask_shadow)

5. Cable detection (optional enhancement):
      ROI_region ← DefineCableRegion(I)
      Edges ← CannyEdgeDetector(ROI_region)
      Mask_cables ← MorphologicalPostProcess(Edges)

6. Final RoI mask:
      Mask_final ← Mask_combined ∪ Mask_cables

Output: Binary RoI mask Mask_final(x, y)
```

Notes for readers: Step 2 is applied only to "cold" clusters to improve contrast before thresholding. Steps 4–5 are optional enhancements that increase segmentation quality for low-contrast or detailed structures (e.g., cables). The function OtsuThreshold() selects the threshold k that maximizes inter-class variance, per the classic Otsu method.

Flow diagram: RoI Segmentation Using Otsu's Method with Preprocessing