Machine Learning-Based Identification of Cellulose Particle Pre-Bridging and Bridging Stages in Transformer Oil

Nur Badariah Ahmad Mustafa[1], Marizuana Mat Daud[2]*,
Hidayat Zainuddin[3], Nik Hakimi Nik Ali[4], Fadilla Atyka Nor Rashid[5]
Institute of Power Engineering, Universiti Tenaga Nasional, Malaysia[1]
Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Malaysia[2]
Faculty of Electrical Engineering, Universiti Teknikal Malaysia, Malaysia[3]
School of Electrical Engineering, Universiti Teknologi MARA, Malaysia[4]
Faculty of Technology and Information Science, Universiti Kebangsaan Malaysia, Malaysia[5]

***Abstract*—The deterioration of transformer oil quality is influenced by factors including the presence of acids, water, and other contaminates such as cellulose particles and metal dust. The dielectric strength of the oil decreases over time and depending on the service conditions. This study introduces an efficient machine learning method to classify the pre-bridging and bridging stages by analyzing the formation of cellulose particle bridges in synthetic ester transformer oil. It is important to note that the pre-bridging and bridging stages indicate a pre-breakdown condition. The machine learning approach implements the combination of digital image processing (DIP) technique and support vector machine (SVM). The DIP technique, specifically the feature extraction method, captures the feature descriptors from the cellulose particles bridging images including area, MajorAxisLength, MinorAxisLength, orientation, contrast, correlation, homogeneity and energy. These descriptors are used in SVM to assess the pre-bridging and bridging stages in transformer oil without human intervention. Various SVM models were implemented, including linear, quadratic, cubic, fine Gaussian, medium Gaussian, and coarse Gaussian. The results achieved 96.5% accuracy using quadratic and cubic SVM models with the eight feature descriptors. This research has significant implications, allowing early detection of transformer breakdown, prolonging transformer lifespan, ensuring uninterrupted power plant operations, and potentially reducing replacement costs and electricity disruptions due to late breakdown detection.**

***Keywords*—Cellulose bridging; feature classification; feature extraction; oil deterioration; support vector machine; synthetic transformer oil**

I. Introduction

In recent decades, technology has grown particularly in the field of electrical power. Transformers showcase a significant example of the developed technology in the realm of electrical power. Recent research suggests that about one-third of transformer faults causes by the deterioration transformer insulation. Therefore, a variety of studies have been undertaken to investigate and comprehend the factors leading to insulation failures. Liquid dielectric oils or also known as insulating oil, is a specialized type of oil used in electrical equipment to provide insulation and cooling. It plays a crucial role in maintaining the reliability and safety of various high voltage electrical systems such as high voltage power transformers. The advantages of liquid dielectric oils are able to operate as insulation and as a heat exchanger. However, liquid dielectric oils are highly prone to contamination [1]. This is due to the fact that within a power transformer, the transformer oil is consistently exposed to metal components, the iron core, and pressboard insulation.

The digital image processing (DIP) technique has been extensively utilized across diverse fields, including manufacturing, medical imaging, meteorology, astronomy, remote sensing, and agriculture. It is also highly relevant in the electrical power sector. With the progress in artificial intelligence and computer vision technologies, new opportunities have emerged for pattern recognition in the study of cellulose bridging formation. For instance, Sinduja et al. [2] evaluated transformer oil quality using machine learning techniques. They compared various kernelized support vector machine (SVM) functions, including the sigmoid kernel function (SKF), radial basis kernel function (RBF), Gaussian kernel function (GKF), and Bayesian optimization (BO). Among these methods, BO demonstrated the highest recognition rate of 99.5%, utilizing features such as the transformer oil's resistivity, acidity, flash point, and dielectric dissipation factor (tan delta). Author in study [2] proposed decision tree method to predict and classify the incipient faults in transformer oil based on the five key gases: hydrogen, methane, ethane, ethylene and acetylene. However, the accuracy performance obtained using decision tree was 62.9%.

AI and ML technologies have revolutionized transformer health monitoring through real-time assessment systems that leverage edge computing and deep learning frameworks like TensorFlow. These systems enable immediate evaluation of critical indicators such as oil color, facilitating early anomaly detection [3]. Simultaneously, the extensive data generated by IoT devices in power transformers is effectively harnessed through AI techniques to optimize maintenance schedules, thereby reducing operational downtime and associated costs [4].

In healthcare applications, transformer-based deep learning models demonstrate significant capability in longitudinal health trajectory analysis, enabling prediction of disease onset and supporting continuous patient monitoring [5]. These advanced

AI platforms can integrate diverse data streams to generate personalized health recommendations, enhancing both patient management protocols and overall system efficiency [6]. Despite these considerable advantages in both industrial and healthcare domains, implementation challenges persist, particularly regarding data privacy protections and the development of robust algorithms capable of effectively processing complex health and operational data [7].

II. Related Works

Many works have been highlighted, implementing machine learning on various datasets and features in the field of transformer oil quality. Other works that related to transformer oil quality measurement using machine learning and artificial intelligence have been presented in Table I.

Previous research studies have demonstrated that the texture features published in study [12]-[14], led to improved image classification accuracy. Therefore, in this study, the DIP technique, specifically the feature extraction method is used to extract the morphological and texture feature descriptors of the cellulose particle bridging images. The cellulose particles bridging feature descriptors are then classified into pre-bridging and bridging stages using SVM. It is important to note that the pre-bridging and bridging stages are indicative of the pre-breakdown condition. A thicker bridging pattern and higher feature descriptor values signify a greater probability of the transformer oil approaching a breakdown condition. The proposed system that is the combination of DIP technique and SVM is operated without human intervention. Therefore, this finding enables the early detection of potential breakdowns in transformers, helps assess their lifespan, and protects them from failures, ensuring uninterrupted operation of power plants. Additionally, it can significantly reduce costs associated with transformer replacements caused by delayed breakdown detection and prevent power disruptions resulting from transformer failures.

TABLE I. Related Works to Measure Transformer Oil Quality using Machine Learning and Artificial Intelligence

Authors (year)	Objective	Features	Findings
Firouzimagham et al. (2020) [8]	Conduct online transformer oil analysis utilizing spectroscopy techniques combined with a machine learning classifier.	Oil color	Accuracy of 80% using linear SVM.
Sun et al. (2021) [9]	• Identify the partial discharge pattern based on the phase resolved partial discharge (PRPD) spectrum using a new MobileNets CNN • Compare MobileNets CNN with other deep learning method	Feature from PRPD spectrum image	Proposed model algorithm - superior accuracy with 98.71% accuracy compared to others. Recognition accuracy rate of 4 classes: Tip discharge (100%), Surface discharge (96.31%), Air-gap discharge (98.53%) and Suspended discharge (100%).
Benmahamed et al. (2018) [10]	Evaluate the insulation condition of power transformer oil by applying K-Nearest Neighbors (KNN) and Naïve Bayes algorithms, utilizing dissolved gas analysis (DGA) data.	• DGA in ppm • DGA In percentage • Dörnenberg ratios • Rogers ratios • Duval triangle reports	KNN is superior to Naïve Bayes with accuracy of 82% (DGA in ppm), 86% (DGA in %), 92% (Duval triangle reports) and 84% (Dörnenberg ratios)
Bhatia et al. (2020) [11]	Assess power transformer using machine learning based regression and classification.	• interfacial tension values (IFT) • breakdown voltage (BDV) • acidity • colour • dissipation factor (DF) • water content.	SVM is superior to the other method with an accuracy of 92.3% for combination features of acidity, color, BDV and DF.

III. Methodology

Computer vision is a multidisciplinary field at the intersection of computer science, engineering and artificial intelligence (AI). The combination of this multidisciplinary fields enables the computers to interpret and understand the visual information similar to human eyes and brains. In this study, we applied the DIP and machine learning techniques, specifically SVM, to assess the condition of transformer oil before it reaches a breakdown state. Our evaluation focused on observing the formation of cellulose particle bridges within a controlled laboratory environment. The bridging experiment involved supplying HVDC to one electrode while grounding the other at room temperature. The voltage gradually increased in 1-minute intervals at each voltage level. Initially, the voltage was raised to 2kV, 7kV and finally 10kV before stepping back to 5kV until breakdown occurred. This stepwise allowed us to observe DEP phenomena, specifically particle mobility, from a static state until a complete, thicker bridge formed between the two spherical electrodes. The time intervals between voltage changes ranged from three to five seconds. All tests were conducted in three times to ensure consistent outcomes.

The experiment aimed to measure the breakdown voltage (BDV), both without and with the contaminants, from the start of bridging until full bridging. However, during the tests, images and videos were recorded at regular intervals to document the bridging process. Fig. 1 illustrates the full setup of the bridging experiment.

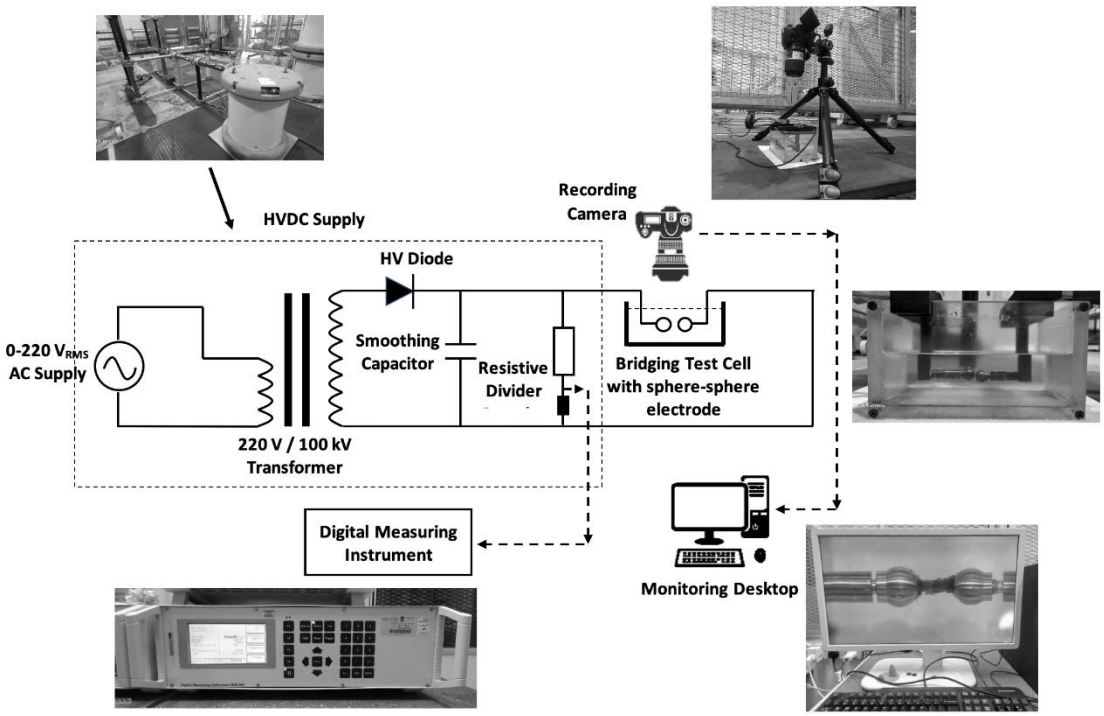


Fig. 1. The laboratory setup of bridging experiment.

The images of cellulose particle bridging were analyzed quantitatively using the DIP technique. Within this method, the images were segmented into distinct, meaningful regions to facilitate more detailed and precise analysis. The important features were then extracted from the segmented image region. The features obtained through the feature extraction process were used to classify the formation of cellulose particle bridges into pre-bridging and bridging stages by using SVM. Fig. 2 describes the overall methodology of cellulose particle bridging classification.

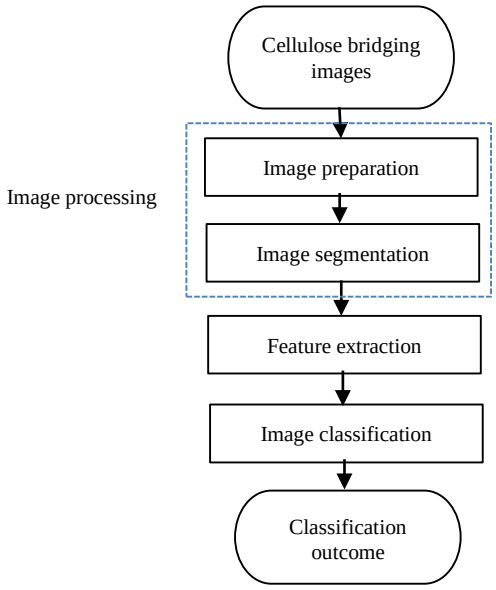


Fig. 2. Methodology of cellulose particles bridging classification.

As previously mentioned, the feature extraction technique produced morphological and texture values extracted from the cellulose particles bridging images by computation. This study began with the manual identification of pre-bridging and bridging stages of cellulose bridging formation images by an expert based on his knowledge and experience dealing with the bridging formation process. The bridging formation was observed by tracking the development of bridging thickness, starting from the initial stage to the formation of a thicker cellulose bridge. Subsequently, the identified pre-bridging and bridging images were processed through a feature extraction method to analyze both morphological (shape) and texture features within the segmented region of interest (ROI).

Both morphological and texture features were employed by the supervised machine learning algorithm, SVM, to perform the classification task using a training dataset. The training dataset consisted of sets of pre-bridging and bridging features samples. The algorithm identified repetitive features of the cellulose particles and used them to classify the test dataset, which comprised features that the SVM model had never seen before. A more detailed discussion of the SVM model is provided in the image classification section.

A. Image Preparation

As previously mentioned, the pre-bridging and bridging datasets were collected from conducted laboratory setup experiments. Fig. 3 illustrates the position of two spherical electrodes with the transformer oil filled inside the experiment box. The cellulose particle bridges formed between the two electrodes, and after a specific period, voltage was applied to one of the electrodes.

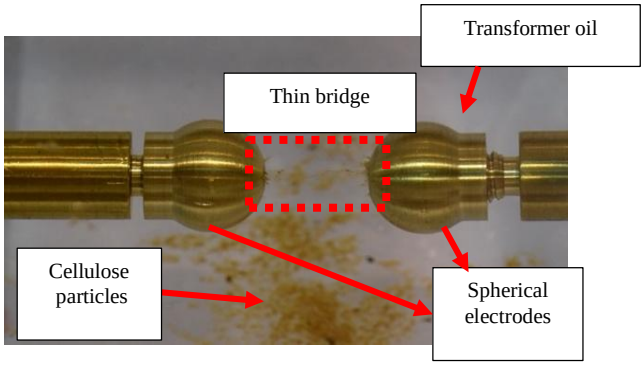


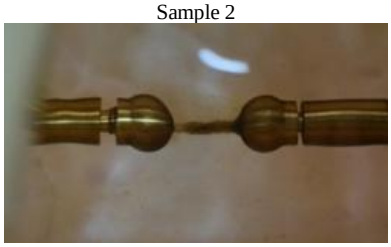
Fig. 3. Sample of captured image from bridging experiment.

The image datasets were collected, identified, and classified into the stages of initial bridge formation (pre-bridging) and active bridge formation (bridging), based on the thickness of the cellulose particle bridges. Images from the post-bridging stage were excluded from the analysis, as they were considered irrelevant, under the assumption that a breakdown and subsequent oil failure would have already occurred at that point. The categorization of the images into pre-bridging and bridging stages was carried out manually to ensure accurate image classification.

During the experiments, the formation of cellulose particle bridges progressed sequentially from the pre-bridging to the bridging stage. In the pre-bridging stage, cellulose particles were dispersed around the electrodes. Over time, these particles were gradually drawn toward both electrodes, forming a thin bridge. As the process continued, the cellulose particle bridge thickened significantly, signaling the beginning of a transformer breakdown condition. This was followed by an explosion of bright light between the two spherical electrodes. A few minutes later, the cellulose particles started to detach from the electrodes, marking the start of the post-breakdown phase in the cellulose particle bridging process.

For this study, 200 images were manually identified and categorized into pre-bridging and bridging stages, with 100 images acquired for each condition. The data samples obtained from feature extraction method were labelled as pre-bridging (denoted as 0) and bridging (denoted as 1). These samples were split into a 60:40 ratio, resulting in 120 training images (60 for each pre-bridging and bridging stages) and 80 testing images (40 for each pre-bridging and bridging stages). Although a large number of images were extracted, some were discarded due to quality issues, including blurriness, inadequate lighting, misalignment, and distortions such as waves. These issues could compromise the accuracy of image feature extraction. To address this, an image selection process was implemented to identify and use only clear, high-quality images for further analysis in the image processing stages. Table II displays example images corresponding to the pre-bridging and bridging phases.

TABLE II. SAMPLE IMAGES - PRE-BRIDGING AND BRIDGING FORMATION

Formation process	Sample images	
Pre-bridging: Initially, cellulose particles became polarized and moved in a scattered manner between the electrodes. Subsequently, the polarized cellulose attached to the electrodes. During this phase, charge transfer occurred between the electrodes. Some of the cellulose particles involved in this charge transfer began moving toward the opposite electrode.	Sample 1 	Sample 2 
Bridging: The bridging process begins with the formation of a thin bridge, which gradually thickens over time.	Sample 1 	Sample 2 

B. Image Segmentation

Image segmentation stage was conducted to extract the meaningful regions for more detail analysis. This phase is crucial for preparing images for the feature extraction process. Cellulose particle bridge images, initially in RGB format, were processed in the image segmentation stage to generate segmented regions of interest (ROI). The flowchart illustrating the image segmentation process is presented in Fig. 4.

The image segmentation stage involved two primary steps: vertex marking and background subtraction. During the vertex marking step, the tips of the electrodes were manually selected. The area from the tip of the left electrode to the tip of the right electrode was marked using a mouse cursor, as illustrated in Fig. 5. This area between the markers is identified as the cellulose bridging region of interest (ROI). Following this, the background subtraction step was performed to eliminate unwanted objects.

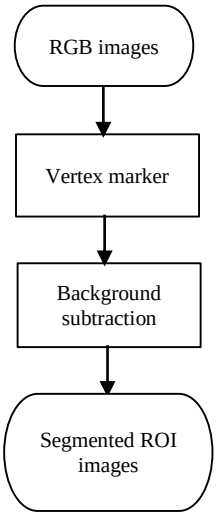


Fig. 4. Process of image segmentation.

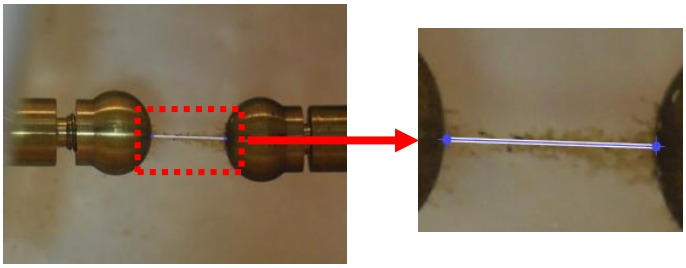


Fig. 5. Vertex markers selection for ROI determination.

During background subtraction process, the selected area between the electrodes was converted into a grayscale image. In this step, pixel corresponding to unwanted objects were identified. Any pixels with a value higher than the identified threshold value was considered an unwanted pixel. In general, the background subtraction process eliminated areas outside the marked regions, such as the electrodes, the white background, and any unwanted noise present in the image.

Mathematically, this process can be represented in terms of pixel intensity values. Let's denote the pixel intensity at position *(i,j)* in the current frame as *I(i,j)*, and the corresponding background model value as *B(i,j)*. The difference between them is expressed in Eq. (1).

$$D(i,j) = |I(i,j) - B(i,j)| \quad (1)$$

If $D(i,j)$ exceeds a predefined threshold, the pixel is considered part of the foreground. This process is repeated for all pixels in the frame. In general, background subtraction involves mathematically comparing pixel intensities between the current frame and a background model to detect changes caused by moving objects, producing a binary mask that differentiates between foreground and background elements.

C. Feature Extraction

This research utilized feature extraction methods to analyze the morphological and texture properties of the segmented region of interest (ROI). The analysis was based on pixel-level calculations of the extracted image features. The feature extraction algorithm was developed using MATLAB's built-in functions, particularly *regionprop* and *bwboundaries*, from the MATLAB Image Processing Toolbox. The *bwboundaries* function traces the contours of selected regions in binary images. The function *[B,L] = bwboundaries(BW, 'noholes')* was applied to compute the boundaries of the selected regions and superimpose them on the image. The 'noholes' parameter ensures that only the external boundaries of objects are detected, improving the algorithm's performance.

As illustrated in Fig. 6, the binary-form segmented ROI images were input into the feature extraction stage to identify and extract significant pixels from the cellulose particle bridging images. Morphological features, including area, MajorAxisLength, MinorAxisLength, and orientation, were calculated to quantify the shape of the ROI. Additionally, texture features such as contrast, correlation, energy, and homogeneity were computed to evaluate the texture of ROI. The extracted feature data were compiled and subsequently the segmented used in the classification stage.

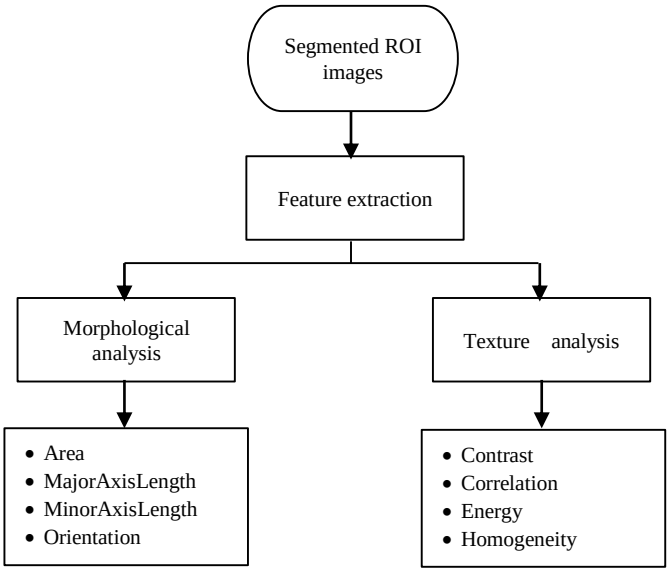


Fig. 6. Extracted features of segmented ROI images.

The properties of selected ROI were measured using *regionprop* function. In MATLAB, the shape measurement was determined based on the ROI properties such as 'Area', 'MajorAxisLength', 'MinorAxisLength' and 'Orientation'. Table III shows the description of shape measurement properties.

TABLE III. DESCRIPTION OF STATISTICAL TEXTURE MEASURES

Statistics	Description
Contrast	Measures the local variations present in the GLCM
Correlation	Assesses the likelihood of specific pixel pairs occurring together in the GLCM.
Energy	Quantifies the uniformity or textural consistency within the GLCM
Homogeneity	Evaluates how closely the elements in the GLCM are distributed along its diagonal.

The image texture was analyzed using the second-order texture analysis technique, Gray-Level Co-Occurrence Matrix (GLCM). The functions determine the frequency of pixel pairs

that share specific intensity values (gray levels) and particular spatial relationships within the image. For instance, GLCM computes the frequency of two neighboring pixels with identical intensities, either horizontally, vertically, or diagonally. In MATLAB, the *graycomatrix* function generates the gray-level co-occurrence matrix for a grayscale image. This matrix represents the frequency at which a specific pixel intensity pair appears at a given distance and angle.

The illustration on the ROI measurement using morphological properties is shown in Fig. 7. All measurement values were in pixels. The thickness and length of the bridging pattern formed between the electrodes were characterized using measurements of the minor and major axis lengths.

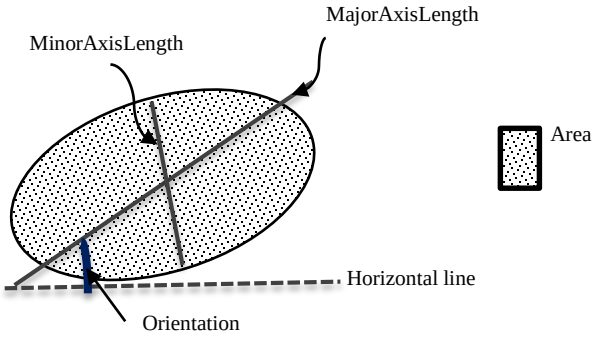


Fig. 7. Illustration of ROI morphological measurement.

Once the matrix is computed, the texture features can be extracted from it using *graycoprops* function. The measures help to quantify the image texture and provide relevant information of image patterns and structures. Common statistical texture measures are contrast, energy, homogeneity and correlation, expressed in Eq. (2) to Eq. (5), respectively.

$$\text{Contrast} = \sum(|i-j|^2 \times \text{p(i,j)}) \quad (2)$$

where the p(i,j) = pixel at location (i,j).

$$\text{Energy} = \sum(\text{p(i,j)})^2 \quad (3)$$

$$\text{Homogeneity} = \frac{\sum \text{p(i,j)}}{1+|i-j|} \quad (4)$$

$$\text{Correlation} = \sum(i-\mu i)(j-\mu j)\frac{p(i,j)}{\sigma_i \sigma_j} \quad (5)$$

Where the $\mu_i = \sum \text{p(i,j)i}$ and $\mu_j = \sum \text{p(i,j)j}$, while σ_i and σ_j are the standard deviation of values i and j references respectively. The description of statistical texture is shown in Table IV.

TABLE IV. DEFINITION OF MORPHOLOGICAL (SHAPE) PROPERTIES FORMATION

Shape Measurement Properties	Description
Area	The total pixel count within the chosen area.
MajorAxisLength	The major axis length (in pixels) of the selected region.
MinorAxisLength	The minor axis length (in pixels) of the selected region.
Orientation	Angle between horizontal line and the major axis line.

D. Image Classification: Pre-bridging or Bridging Condition

SVM-based image classification is a widely used and efficient method in machine learning and computer vision. As a supervised learning algorithm, SVM is suitable for handling multi-class classification tasks. In this study, SVM was employed to assign cellulose bridging stages (pre-bridging or bridging) to the input images based on the extracted features. Fig. 8 shows the SVM process in pre-bridging and bridging stages classification.

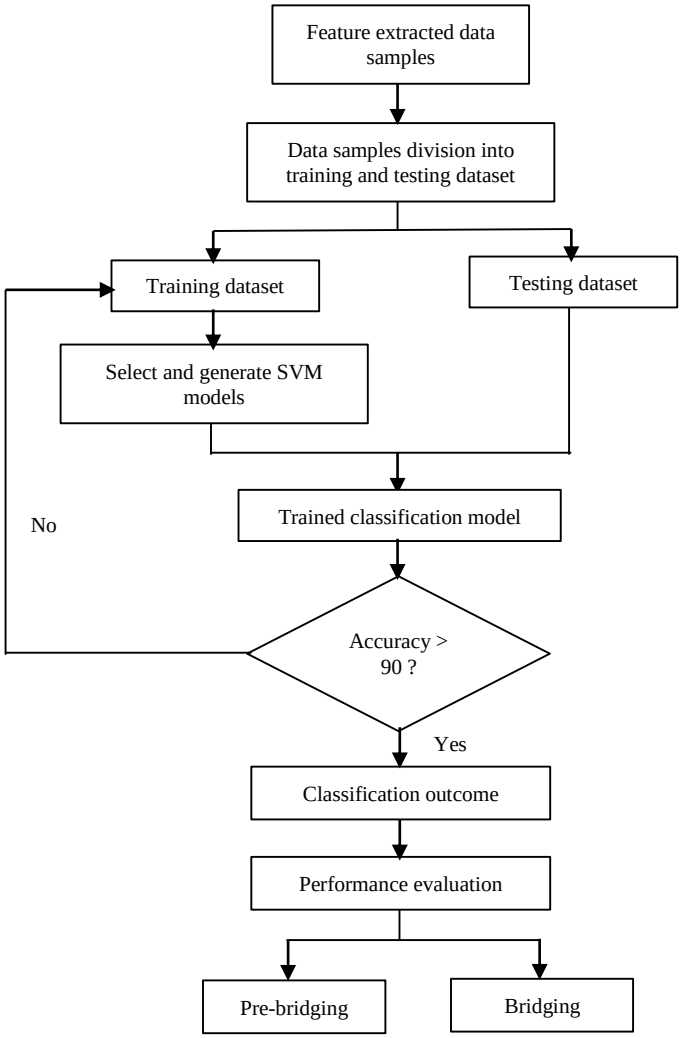


Fig. 8. Flowchart of classification module.

As previously mentioned, the feature extraction process identified eight essential descriptors: contrast, orientation, energy, correlation, homogeneity, MinorAxisLength, MajorAxisLength, and area. These descriptors were employed to define the pre-bridging and bridging patterns. The descriptors, along with the corresponding cellulose bridging stages, were organized and used to train the SVM models. To avoid overfitting, a 5-fold cross-validation method (k=5) was applied. The training utilized various SVM models, including linear, cubic, quadratic, coarse Gaussian, medium Gaussian, and fine Gaussian. Once the training was complete, the SVM model was prepared for data prediction using samples from the testing dataset.

To evaluate the SVM model's performance with different kernel type, the process described above was repeated with different set of feature descriptors: (a) Set 1: SVM models trained with eight feature descriptors, (b) Set 2: SVM models trained with five feature descriptors, and (c) Set 3: SVM models trained with four feature descriptors. The selection of the number of feature descriptors was based on identifying the best features which will be discussed in detail in results and discussion section. The list of selected feature descriptors is shown in Table V. In Set 1, all morphological and texture features were input into SVM models. In Set 2, energy descriptors were incorporated into the SVM models alongside morphological features, as consistent patterns were observed throughout the pre-bridging and bridging phases, shown in Fig. 10 and 11. In Set 3, only morphological features were used in the SVM models.

TABLE V. LIST OF SELECTED FEATURE DESCRIPTORS

	Feature descriptors
Set 1	• Area • MajorAxisLength • MinorAxisLength • Orientation • Contrast • Correlation • Homogeneity • Energy
Set 2	• Area • MajorAxisLength • MinorAxisLength • Orientation • Energy
Set 3	• Area • MajorAxisLength • MinorAxisLength • Orientation

As depicted in Fig. 8, the accuracy of SVM models was evaluated to measure their performance. The evaluation was conducted on the training dataset. The process involved feeding the SVM models with different kernel types with different sets of input features, as described in Sets 1, 2 and 3. Subsequently, the outcomes of the SVM models were compared with the ground data (image data) which manually determined by the expert.

E. Performance Evaluation

The performance of each SVM model was evaluated based on its accuracy value, where higher accuracy indicates better model performance. Eq. (6) describes the accuracy formula used in this study.

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP} \; OR \quad (6)$$

$$\frac{Correct\ predictions\ made}{Total\ number\ of\ predictions}$$

In Eq. (6), TP represents True Positive, TP represents True Negative, FP represents False Positive and FN represent False Negative. These terms are defined in Table VI.

TABLE VI. DEFINITION OF THE MODEL EVALUATION METHOD OF ACCURACY METRICS

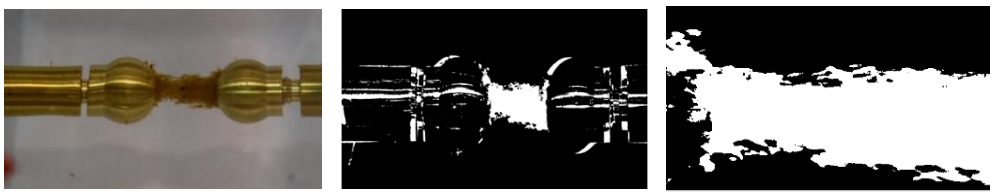
Evalution outcome		Definition
TP	True Positive	SVM model predicts the correct positive class
TN	True Negative	SVM model predicts the correct negative class
FP	False Positive	SVM model predicts the incorrect positive class
FN	False Negative	SVM model predicts the incorrect negative class

IV. RESULTS AND DISCUSSIONS

This work began with pre-processing cellulose particles image, focusing on segmenting the ROI (cellulose particle) and extracting the morphological and texture information. The output of the image segmentation stage is presented as a segmented ROI image (binary image), as shown in Fig. 9(b), and the original image displayed in Fig. 9 (a).

Texture is a significant characteristic of image data, as it helps identify objects or ROI within an image. In DIP, texture refers to the spatial variations in the brightness intensity of the pixels within an image. The texture of an image is characterized by a specific pattern of texture distribution that repeats sequentially throughout the image. In this study, texture provides additional contextual information about the image, complementing the morphological analysis of the ROI. The characteristic of the cellulose is observed in 180-degree horizontally placed view (see Fig. 9 (c)). This enhanced insight derived from texture complements the primary morphological-based analysis of the ROI and contributes to a more comprehensive understanding of the cellulose particle images.

In addition to texture's role, feature extraction was utilized to identify and capture key characteristics of cellulose particle bridging formation in the images. This section provides a detailed description of the results of feature extraction from the cellulose bridging images and the comparison of SVM models with different kernel types. The feature descriptors used to characterize the specific shape and pattern of the images were determined using feature extraction method. These eight feature descriptors played a crucial role in describing the cellulose bridging images.



(a) Original image (b) Segmented ROI image (c) Extracted cellulose image

Fig. 9. The process to extract Region Of Interest (ROI).

In this study, the images were classified into pre-bridging and bridging stages using SVM model which were designed and generated with various kernel types, including linear, quadratic, cubic, fine Gaussian, medium Gaussian and coarse Gaussian. As illustrated in Fig. 10 and Fig. 11, the morphological and texture features consistently displayed patterns during both the pre-bridging and bridging stages, effectively capturing the characteristics of the images. Thus, based on these observations, it was determined that the eight feature descriptors were important for distinguishing between image conditions.

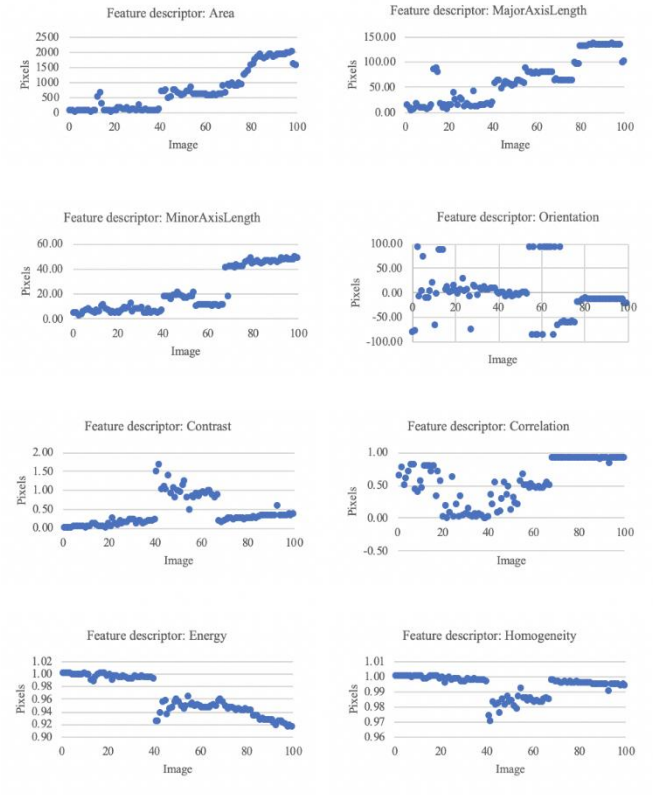


Fig. 10. Feature extraction outcomes of pre-bridging images.

In direct observation, area, MajorAxisLength, and MinorAxisLength demonstrated noticeable trends within the pre-bridging stage. These findings were attributed to the development process of the cellulose bridging structure, characterized by pixel accumulation leading to an expansion in both vertical and horizontal dimensions. However, in the case of the bridging stage, MajorAxisLength exhibited noticeable reduction from a certain point in time (image count). These findings were attributed to the detachment of pixels, which is indicative of a transformer breakdown.

The orientation feature descriptor has been shown to effectively differentiate between the pre-bridging and bridging stages. Upon observation, the orientation graph displayed a scattered pattern for the pre-bridging stage, while it remains relatively constant at an angle of 0 degree for the bridging stage. These findings were attributed to the initial development of the cellulose bridging structure, during which pixels scatter and coalesce to form a horizontally rectangular bridge shape. Furthermore, in terms of the energy feature descriptor, there was a high distribution of pixel intensity in pre-bridging stage but lower energy for the bridging stage. While homogeneity indicates the smoothness or regularity of texture based on the similarity of pixels' intensity. It exhibited a discernible incremental trend toward increased homogeneity as the bridging structure developed.

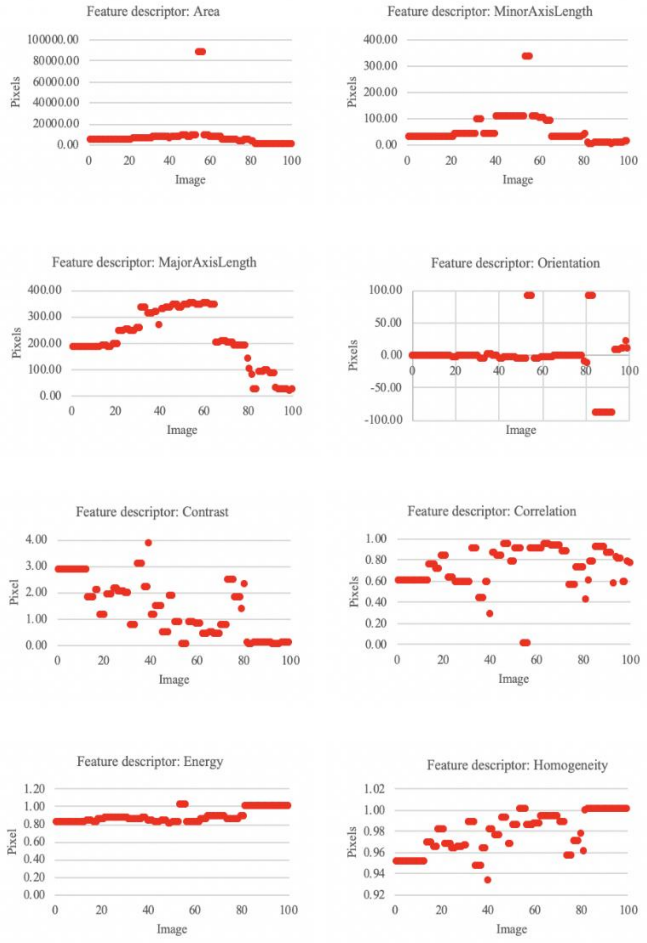


Fig. 11. Feature extraction outcomes of bridging images.

Accuracy evaluations of the SVM models were conducted using the pre-bridging and bridging patterns obtained from the feature extraction process, with assessments made on both the training and testing datasets. It is proven that the SVM models achieved a good classifier accuracy which is more than 80% when tested using training dataset. To assess the robustness of the SVM models, the models were tested using testing dataset. The findings of the accuracy performance for the training and testing dataset is shown in Fig. 12. In summary, the quadratic and cubic SVM models for the testing dataset exhibited better accuracy values compared to the training dataset, while fine Gaussian SVM model showed comparable accuracy values between the two datasets. Other models, such as linear, medium Gaussian and coarse Gaussian resulted in lower accuracy values for the testing dataset compared to the training dataset. However, the differences were considered acceptable as they were not substantial.

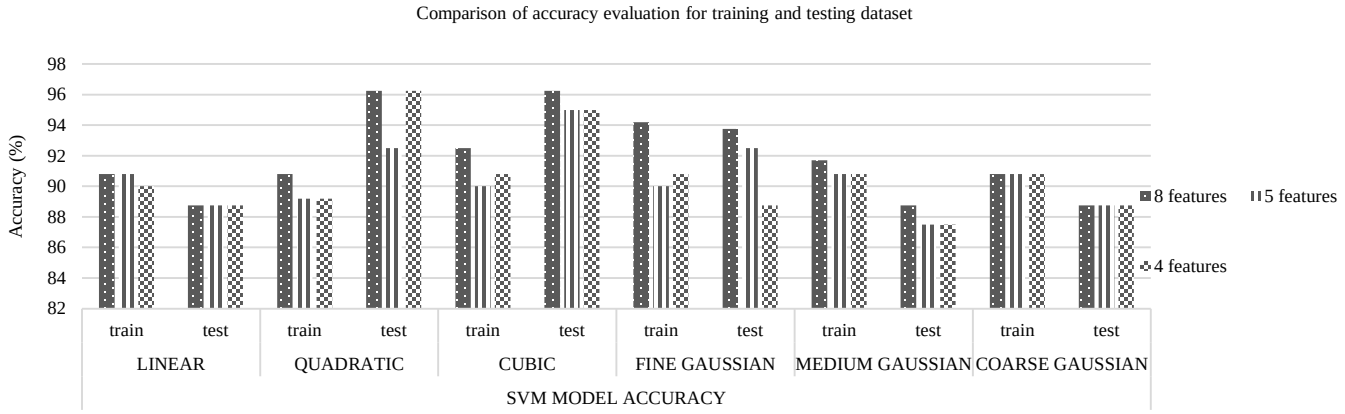


Fig. 12. Comparison of SVM kernel for 4, 5, and 8-features.

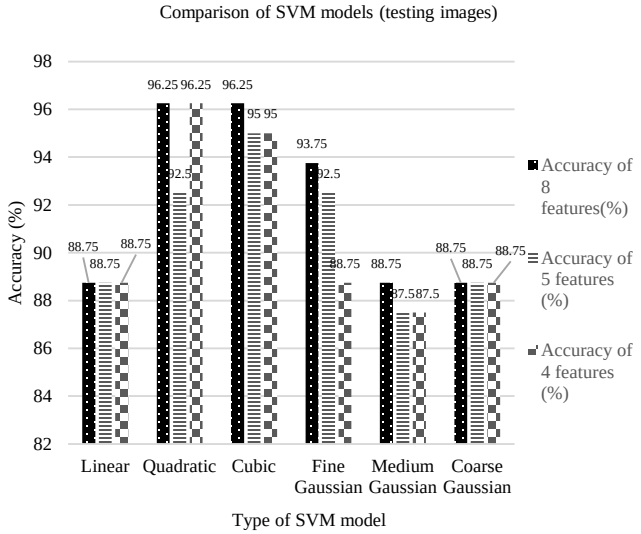


Fig. 13. Comparison of SVM models in terms of accuracy percentage.

Fig. 13 presents the accuracy performance comparison of SVM models using testing dataset. Notably, each of the SVM models exhibits exceptionally high accuracy when employing eight distinct feature descriptors (area, MajorAxisLength and MinorAxisLength, orientation, contrast, correlation, energy, and homogeneity). However, it was evident that the SVM model utilizing quadratic and cubic kernel types stands out with the most accurate predictions, achieving a remarkable accuracy rate of 96.25%. This suggests that all the features exhibit distinct characteristics that effectively distinguish between the pre-bridging and bridging stages. The high accuracy of the quadratic and cubic SVM models could be contributed by their non-linearity, which transforms the data into higher-dimensional space. This means that the quadratic and cubic SVMs are capable of identifying curved or non-linear decision boundaries. Linear SVM model had the lowest accuracy (88.75%) due to its simple algorithm and low kernel compared to the other applied SVM models. Based on these findings, it is evident that the linear SVM is not suitable for handling non-linearly separable data.

The Gaussian SVM model is also known for its higher kernal and its complex algorithms. However, for this application, gaussian SVM unable to produce higher accuracy values for the three sets compared to quadratic and cubic SVM. Thus, it can be concluded that the choice of SVM kernel depends on the characteristics of the image dataset. The combination of morphological and texture features (eight feature descriptors) demonstrated higher accuracy value compared to the SVM models with five and four feature descriptors. The higher accuracy indicates that the model is capable of accurately distinguishing between the pre-bridging and bridging stages.

V. Conclusion

This study utilizes digital image processing (DIP) and support vector machine (SVM) tools to assess transformer oil condition and predict potential breakdowns. The pre-breakdown conditions were categorized into pre-bridging and bridging stages for evaluation. The DIP technique specifically feature extraction method, was employed to determine the important cellulose bridging characteristics based on cellulose bridging formation images. The feature descriptors were fed into SVM models to classify the cellulose bridging structure patterns as either pre-bridging or bridging stages. This study involved the implementation and evaluation of various SVM models, such as linear, quadratic, cubic, fine Gaussian, medium Gaussian, and coarse Gaussian. Notably, the quadratic and cubic SVM models yielded an impressive accuracy rate of 96.5%, showcasing their effectiveness in identifying pre-bridging and bridging stages in pre-breakdown conditions. The evaluation of accuracy performance was also conducted using three sets of feature descriptors, and the findings showed that SVM models (quadratic and cubic) with eight feature descriptors resulted in higher accuracy values. This demonstrates that morphological (shape) and texture features played significant roles in analyzing the cellulose bridging structure. The significance of this work lies in its potential to revolutionize transformer maintenance practices. By enabling early detection of bridging faults, it contributes to the extension of transformer lifespans and the enhancement of the reliability of power plant operations. Furthermore, the research has the potential to reduce the financial burden associated with transformer replacement due to late breakdown detection. Additionally, it offers a valuable

solution to prevent electricity disruptions caused by transformer failures, further underscoring its importance in ensuring the stability and continuity of power supply. In summary, the successful application of DIP and SVM in cellulose particle bridging pattern recognition offers a promising approach to enhancing transformer health monitoring and maintenance in power systems. Future research may explore further refinements and real-world implementations of this approach, ultimately advancing the reliability and efficiency of power generation and distribution.

ACKNOWLEDGMENT

The authors express their gratitude to the Ministry of Higher Education for supporting this project through the Fundamental Research Grant Scheme (FRGS), grant number FRGS/1/2020/TK0/UNITEN/02/17.

REFERENCES

[1] S. Mahmud, G. Chen, I. O. Golosnoy, G. Wilson, and P. Jarman, "Bridging in contaminated transformer oil under AC, DC and DC biased AC electric field," in 2013 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, 2013, pp. 943–946.

[2] M. Sinduja, R. V. Maheswari, and B. Vigneshwaran, "Transformer oil quality assessment using machine learning techniques," in 2022 International Conference on Computer Communication and Informatics (ICCCI), 2022, pp. 1–5.

[3] N. M. Lindsay and A. N. K, "Design of Transformer Health Monitoring System Using Tensor Flow Architecture," pp. 974–979, Dec. 2024.

[4] R. Zemouri, "Power Transformer Prognostics and Health Management Using Machine Learning: A Review and Future Directions," Jan. 2025.

[5] H. Moen *et al.*, "Towards modeling evolving longitudinal health trajectories with a transformer-based deep learning model," Dec. 2024, doi: 10.48550/arxiv.2412.08873.

[6] Á. Periáñez *et al.*, "The Digital Transformation in Health: How AI Can Improve the Performance of Health Systems," *Health Systems and Reform*, vol. 10, no. 2, Oct. 2024, doi: 10.1080/23288604.2024.2387138.

[7] A. Arbi and M. Israr, "Empowering Cyber-Physical Systems through AI-driven Fusion for Enhanced Health Assessment," *International Journal of Data Informatics and Intelligent Computing*, vol. 3, no. 3, pp. 16–23, Aug. 2024, doi: 10.59461/ijdiic.v3i3.127.

[8] Y. Benmahamed, Y. Kemari, M. Teguar, and A. Boubakeur, "Diagnosis of power transformer oil using KNN and Naive Bayes classifiers," in 2018 IEEE 2nd International Conference on Dielectrics (ICD), Budapest, Hungary, 2018, pp. 1–4.

[9] N. K. Bhatia, A. H. El-Hag, and K. B. Shaban, "Machine learning-based regression and classification models for oil assessment of power transformers," in 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), 2020, pp. 400–403.

[10] N. B. A. Mustafa, I. D. Ramasamy, F. H. Nordin, N. H. N. Ali, H. Zainuddin, and M. M. Daud, "Characterization of cellulose bridging pattern in transformer oil using feature extraction technique," in 2022 IEEE International Conference on Power and Energy (PECon), 2022, pp. 219–224.

[11] S. Liao, M. Law, and A. Chung, "Dominant local binary patterns for texture classification," IEEE Transactions on Image Processing, vol. 18, pp. 1107–1118, 2009.

[12] R. A. Raj, D. Sarathkumar, S. K. Venkatachary, and L. J. B. Andrews, "Classification and prediction of incipient faults in transformer oil by supervised machine learning using decision tree," in 2023 3rd International Conference on Artificial Intelligence and Signal Processing (AISP), 2023, pp. 1–6.

[13] D. Firouzimagham, P. Aminaie, Z. Shayan, M. Sabouri, and M. H. Asemani, "Online transformer oil analysis based on spectroscopy technique and machine learning classifier: Experimental setup," in 2020 15th International Conference on Protection and Automation of Power Systems (IPAPS), 2020, pp. 30–36.

[14] Y. Sun, S. Ma, S. Sun, P. Liu, L. Zhang, J. Ouyang, and X. Ni, "Partial discharge pattern recognition of transformers based on MobileNets convolutional neural network," Applied Sciences, vol. 11, no. 15, p. 6984, 2021.