A Deep Learning-Based Thermal Analysis System for TNB Substation Monitoring

Azka Aftab[1], Asma Kausar[2] and Mohammad Ibrahim Shapiai[3]

***Abstract*— Thermal inspections in electrical substations are important for early fault detection and preventive maintenance. However, conventional inspection methods are still highly manual, subjective, and prone to error, especially when it comes to analyzing thermal signatures in overlapping or complex equipment compartments. This paper presents a deep learning-based framework for automating thermal fault detection in 11kV substations operated by Tenaga Nasional Berhad (TNB). The proposed system classifies substation compartments using MobileNetV2, segments regions of interest (ROIs) with the Segment Anything Model (SAM) and identifies anomalies based on ΔT thresholding (temperature difference analysis). Thermal and optical datasets were collected using FLIR and Fluke devices from actual substations, and performance was validated using accuracy, IoU, and other relevant metrics. The results achieved a classification accuracy exceeding 93% and a SAM-predicted segmentation IoU confidence score of up to 0.92, demonstrating the strong potential of the proposed approach for real-world deployment. This approach offers a scalable, interpretable and real-time fault analysis tool customized to Malaysia's energy infrastructure.**

I. INTRODUCTION

Malaysia's electricity infrastructure is highly dependent on the continuous operation of substations to provide statewide energy distribution. However, routine inspection and maintenance of substations continue to be a manual, labor-intensive process characterized by inconsistency and limited scalability. Thermal imaging has emerged as a crucial instrument for substation monitoring, facilitating non-contact, real-time identification of anomalous heat patterns that may indicate potential equipment malfunctions. Current methodologies use manual interpretation of thermograms, which is inherently subjective and open to human error and delays [1] [2].

The growing implementation of smart technology in power systems underscores the pressing necessity for automated and intelligent problem detection procedures. Conventional rule-based methods for evaluating temperature anomalies show restricted adaptability and suboptimal performance across varying environmental conditions [3] [4]. On the other hand, the emergence of deep learning has shown great potential in overcoming these constraints by facilitating intelligent feature extraction, classification, and segmentation directly from unprocessed thermal and optical data [5] [6].

[1]Azka Aftab is with the Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia `aftab@graduate.utm.my`

[2]Asma Kausar is with the Faculty of Business Administration, University of Tabuk, Tabuk, Saudi Arabia `aaftab@ut.edu.sa`

[3]Mohammad Ibrahim Shapiai is with the Faculty of Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia `md_ibrahim83@utm.my`

Notwithstanding global progress, most thermal fault detection systems are inadequately adapted to the Malaysian situation, especially concerning the 11kV switchgear compartments utilized in Tenaga Nasional Berhad (TNB) substations. These compartments include intricate designs and overlapping heat sources that pose difficulties for both manual and automated interpretation [7] [8]. Moreover, numerous current models need substantial computational resources and are incompatible with real-time applications in field settings [1] [9].

This work presents a lightweight AI-driven solution for thermal inspection and fault detection in TNB substations to address these shortcomings. The methodology utilizes MobileNetV2 for compartment categorization [1], the Segment Anything Model for precise region-of-interest segmentation [7], and ΔT-based scoring to evaluate thermal anomalies [10] [8]. The aim is to improve the efficiency, precision, and scalability of substation diagnostics in Malaysia utilizing a real-world dataset obtained from FLIR and Fluke thermal instruments.

II. LITERATURE REVIEW

Conventional rule-based fault scoring systems lack compartment-level accuracy and do not scale efficiently across various equipment manufacturers and configurations. Current AI methodologies generally focus just on classification or segmentation, without an integrated and explainable pipeline that encompasses both. Consequently, a distinct deficiency exists in providing a cohesive, pixel-accurate, and severity-sensitive diagnostic system for practical implementation in 11kV substations.

A. Comparison Analysis of Current Similar Systems

Thermal fault detection has advanced from empirical methods to AI-driven pipelines that incorporate CNNs, segmentation networks, and data fusion techniques. Lightweight CNNs like MobileNetV2 offer high accuracy with low computational demand [1], while models like SAM and DeepLabV3 enable fine-grained fault localization [11] [12]. However, three key limitations persist:

- Modular Integration: Few frameworks unify classification, segmentation, and thermal scoring in a cohesive pipeline [10] [13].
- Thermal Fidelity Loss: RGB conversions for CNN input often obscure critical gradients [8] [14].
- Real-time Deployment: Many high-performing models are too resource-intensive for edge devices [9].

Malaysian substations further face variability in compartment design and brand-specific behaviors. While transfer learning offers adaptation potential [6], pixel-aligned ΔT estimation remains a bottleneck.

While numerous research efforts have focused individually on thermal image classification, ROI segmentation, or temperature anomaly scoring, very few provide a unified diagnostic architecture that integrates all three components—compartment classification, segmentation of thermographic regions, and ΔT-based severity estimation—into a single end-to-end pipeline. This absence of modular integration has been recognized as a critical bottleneck in recent literature, particularly for systems designed for real-time or industrial use [10] [13].

For instance, many classification-based models such as those using EfficientNet or ResNet offer high accuracy in compartment recognition but fail to isolate fault-prone zones or quantify thermal deviations within those areas [6] [15]. Similarly, segmentation-focused architectures like DeepLabV3 or the Segment Anything Model (SAM) excel at spatial localization but typically lack embedded classification or severity scoring mechanisms [7] [12].

Even promising hybrid solutions, such as SAM integrated with DeepLabV3 or YOLO for region detection, are often evaluated in controlled lab conditions with standardized equipment and fixed imaging angles—settings that do not reflect the messy, reflective, and dynamically loaded environments typical of actual substations. This makes them fragile in the presence of field challenges like thermal reflection artifacts, sensor calibration drift, and equipment brand heterogeneity, especially in Malaysian substations where TNB utilizes a mix of VCB and RMU models from multiple manufacturers [2] [8] [16].

Taken together, these limitations highlight the current gap between academic models and deployable systems. In response, this paper presents a unified pipeline tailored to Malaysia's 11kV substations that combines MobileNetV2-based classification [1], SAM-driven segmentation [7], and pixel-aligned ΔT severity scoring—all optimized for lightweight deployment.

III. METHODOLOGY

This section outlines the structured research methodology employed to design, implement, and evaluate the proposed deep learning-based thermal fault detection system. Each phase is tailored to address specific technical challenges and ensure the system's robustness in real-world substation environments.

A. Research Design and Strategy

The research focuses on optimizing three interdependent tasks: compartment classification, ROI segmentation, and ΔT-based thermal severity scoring.

Instead of a system development lifecycle, the methodology is structured around four research phases: problem scoping, model design, dataset development and training, and performance evaluation. As shown in Figure 1 this approach ensures that each component of the proposed framework is tested, benchmarked, and validated against real-world substation data.

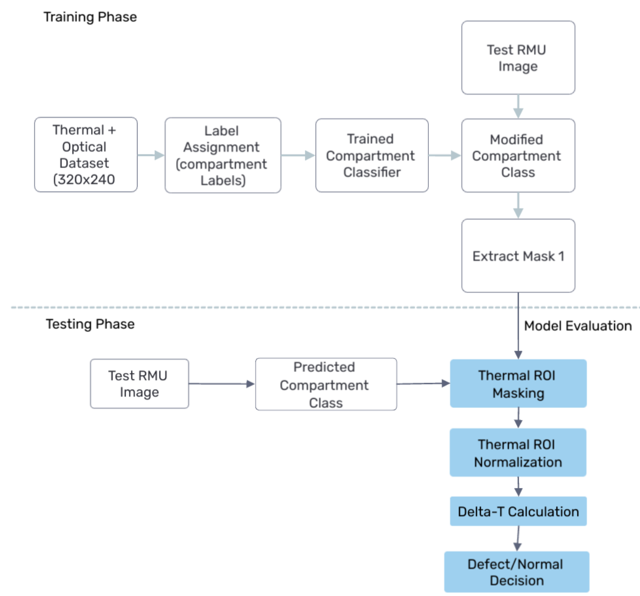

Fig. 1. Proposed System Workflow

B. Justification for Quantitative Method

As summarized in Table I quantitative methods were selected due to their ability to objectively evaluate model performance and ensure replicable results under defined conditions. By using standardized metrics such as accuracy, precision, SAM-predicted IoU confidence, and ΔT deviation, the framework can be rigorously compared to alternative solutions and optimized for deployment in live substations.

TABLE I
JUSTIFICATION FOR QUANTITATIVE METHODS

Reason	Justification
Objective Evaluation	Enables unbiased scoring of classification, segmentation, and thermal thresholds
Reproducibility	Experiments can be consistently replicated for academic and industrial benchmarking
Statistical Comparison	Supports head-to-head performance analysis between baseline and proposed models
Operational Relevance	Measures ΔT errors, runtime latency, and inference time to assess field readiness
Scalability	Aligns with future automation objectives in substations requiring minimal human oversight

C. Dataset

A real-world dataset was constructed using thermal imagery captured from Tenaga Nasional Berhad (TNB) substations across two primary switchgear categories: Vacuum Circuit Breakers (VCB) and Ring Main Units (RMU). All thermal images were obtained using FLIR E6 and Fluke

Ti401 Pro infrared cameras, representing both indoor (VCB) and outdoor (RMU) conditions. Each image corresponds to one of 11 switchgear compartments, with 40 images per compartment—yielding a total of 440 images.

As presented in Table II, the dataset was split in an 80:20 ratio for training and testing respectively, ensuring balanced representation of compartments across both subsets. Each image was further labeled based on equipment type and visual inspection of thermal anomalies, allowing for binary classification into "Normal" and "Fault" cases. Though no pixel-level ground truth masks were available, ΔT overlays and severity estimation were applied post-segmentation for evaluation purposes.

TABLE II
RESEARCH ACTIVITY PLANNING AND OUTCOMES

Data Split	**Training (80%)**	**Testing (20%)**	**Total**
Image Count	352 images	88 images	440 images
Compartments	11	11	11
Images per Compartment	32	8	40

D. Technologies Used

The system was developed using a carefully chosen set of tools and frameworks optimized for lightweight deployment, efficient processing, and thermal-vision compatibility. MobileNetV2 was used for compartment classification via TensorFlow and Keras, while segmentation tasks were handled using Meta AI's Segment Anything Model (SAM) loaded and deployed through PyTorch. OpenCV was used for thermal-to-RGB conversion, resizing, and preprocessing. Jupyter Lab was used as an experimentation platform.

IV. DESIGN AND IMPLEMENTATION

The proposed solution is a modular, AI-driven pipeline that automates thermal fault inspection for 11kV switchgear. It integrates lightweight classification, prompt-based segmentation, and ΔT-based severity analysis in a single end-to-end system optimized for reproducibility and real-world usability.

A. System Architecture

The system workflow consists of four core components:

- Thermal .csv matrices and optical .png images are resized to 320×240 resolution to ensure pixel-level spatial alignment shown in Figure 2. This standardization allows accurate overlay and consistent data formatting for downstream tasks [8] [14].
- Compartment Classification: A lightweight MobileNetV2 architecture classifies input images into one of 11 compartment types. Feature extraction is performed using pretrained ImageNet weights, followed by dense layers, dropout regularization, and softmax output for final prediction.
- ROI Segmentation: The Segment Anything Model (SAM), built on a Vision Transformer backbone (ViT-B), uses a center-point prompt to identify fault-prone regions [17]. The output mask is overlaid on the thermal matrix to isolate temperature zones of interest [1] [5].
- Scoring & Status Labelling: Once the region of interest is identified, the system extracts the max and min temperatures and computes ΔT using:

$$\Delta T = T_{max} - T_{min} \quad (1)$$

Any region with $\Delta T > 37°F$ is flagged as defective. This simple yet interpretable logic supports immediate binary labeling of each compartment.

B. Training and Evaluation Setup

Model training was conducted in JupyterLab using a stratified 80:20 train-test split. Optical-thermal pairs were fed to both classification and segmentation models. Evaluation metrics included:

- Classification: Accuracy, F1 Score
- Segmentation: SAM-predicted IoU confidence, Dice Coefficient
- ΔT Analysis: False Positives/Negatives, Visual Overlays.

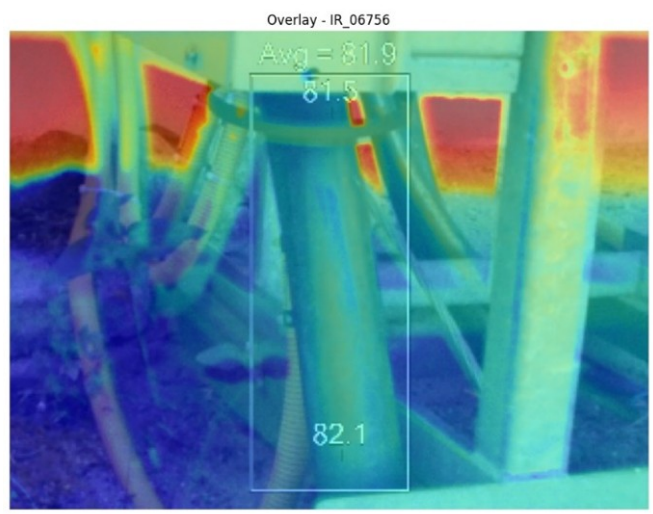

Fig. 2. Optical and Thermal Overlay Pixel-by-Pixel

C. Comparative Testing

To select the best model architecture, several deep learning models were compared:

- Classification Models: MobileNetV2, ResNet18, EfficientNet, VGG16
- Segmentation Models: SAM, U-Net, SegNet

SAM showed superior generalization with minimal tuning, while MobileNetV2 offered a reliable trade-off between accuracy and computational efficiency, achieving 88.89% validation accuracy.

D. Implementation Highlights

The final system integrates compartment classification, ROI segmentation, and thermal scoring into a streamlined, fully automated inspection pipeline. This modular architecture allows each component—classification, segmentation, and ΔT analysis—to function independently while still contributing to an end-to-end diagnostic framework. At its core, the pipeline begins by classifying the compartment using the MobileNetV2 model demonstrated in Figure 3, followed by point-based region segmentation shown in Figure 4 using the Segment Anything Model (SAM), and culminates in ΔT computation from pixel-aligned thermal matrices.

SAM's prompt-based segmentation was particularly advantageous, as it allowed for rapid and flexible region selection using a single coordinate input. This reduced the need for complex pre-annotated masks and enabled quick adaptation to multiple compartment shapes and layouts. Once the region of interest (ROI) was identified, OpenCV was used to align and resize thermal matrices for mask overlay. NumPy facilitated pixel-wise extraction of temperature values within each segmented ROI. The system then calculated the temperature difference (eq.1) and applied a predefined threshold of 37°F to determine the fault status. Compartments exceeding this threshold were flagged as defective, while others were labeled as normal.

One of the defining strengths of this implementation is its computational simplicity. All training, testing, and evaluation tasks were performed on a CPU-only runtime in Jupyter Lab, emphasizing the feasibility of deploying such a system without access to specialized GPU hardware. Despite limited resources, the models converged effectively and delivered reliable results across all compartments.

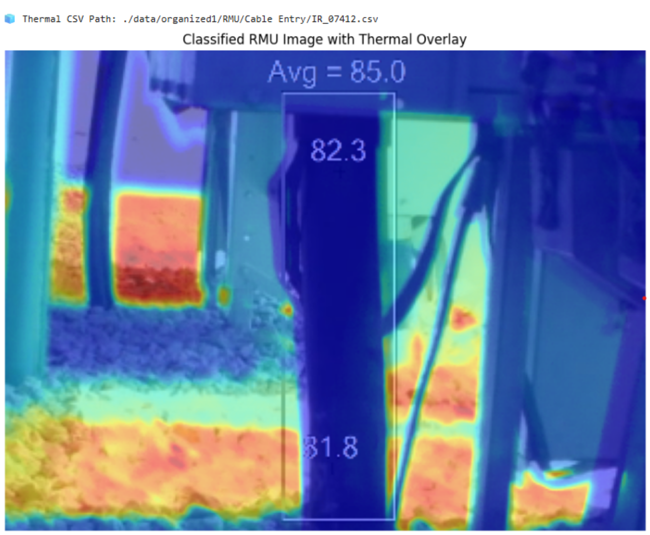

Fig. 3. Classified RMU Image with Thermal Overlay

The final output in Figure 5 includes comprehensive visual diagnostics, making the system intuitive and technician friendly. Using Matplotlib, the platform generates heatmap overlays, highlights ROI boundaries, and marks the hottest and coldest pixels. These visual cues allow inspection personnel to quickly verify anomalies and make informed decisions on-site. Additionally, the modular codebase is designed for future extension, including the integration of new models, real-time monitoring interfaces, or mobile deployment to edge devices like the NVIDIA Jetson Nano. The implementation details and source code of the proposed framework are publicly available for reproducibility.[1]

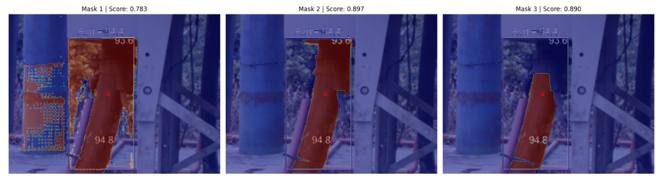

Fig. 4. Sample Segmentation Outputs

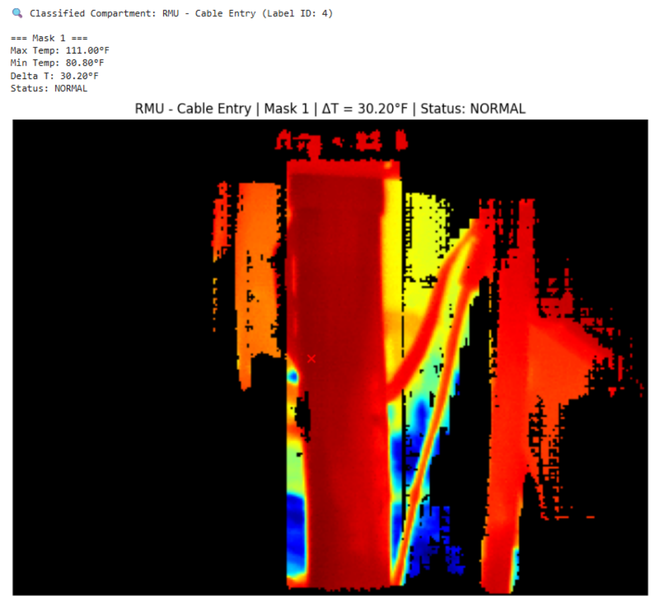

Fig. 5. ΔT Overlay with Status Labels

V. RESULTS AND ANALYSIS

Each module is analyzed using quantitative metrics and visual outputs to ensure generalizability, interpretability, and real-world feasibility.

The system was tested on thermal-optical image pairs from 11 distinct compartment types under both Ring Main Unit (RMU) and Vacuum Circuit Breaker (VCB) setups. The performance evaluation was based on an 80:20 stratified split of the full dataset (352 training, 88 testing). Each test image underwent full pipeline inference: classification, mask generation, ΔT computation, and visualization.

The MobileNetV2-based compartment classification model demonstrated rapid convergence during training, achieving approximately 93% training accuracy and 90% validation accuracy within nine epochs. The steady reduction in both training and validation loss indicates stable learning behavior and effective feature extraction from the thermal–optical image pairs. These results suggest that the pretrained MobileNetV2 architecture can reliably generalize

[1]Source code repository: `https://github.com/Azzkaa/FYP2/blob/main/1pipeline-Copy1.ipynb`

across different RMU and VCB compartments, making it suitable for automated thermal inspection in substation monitoring environments.

A. Classification Performance Analysis

To gain finer insight into class-wise reliability, Figure 6 depicts the model's per-class accuracy across all 11 compartments. The model achieved perfect (100%) accuracy for RMU - Cable Entry, VCB - Busbar Side, RMU - Fuse, and RMU - PT. These results are largely attributed to distinct thermal textures and compartment geometries that reduce inter-class confusion.

However, VCB - CB Open Door and RMU - Cable showed slightly reduced accuracy (∼75–78%). These compartments may share overlapping visual or thermal characteristics, leading to prediction ambiguity. This insight directs attention toward possible dataset augmentation or additional domain-specific regularization.

1) Sample-Based Classification Output: To visualize real-world classification performance, Figure 6 presents ten thermal-optical test images annotated with predicted labels, true labels, and ΔT statistics. In 8 of the 10 samples, the predictions were accurate. The two misclassifications occurred in compartments with subtle thermal contrast or ambiguous door configurations—consistent with known prediction weaknesses.

This qualitative output highlights the model's robustness to varying lighting, surface reflections, and enclosure angles, while also identifying edge cases that could benefit from focused re-training.

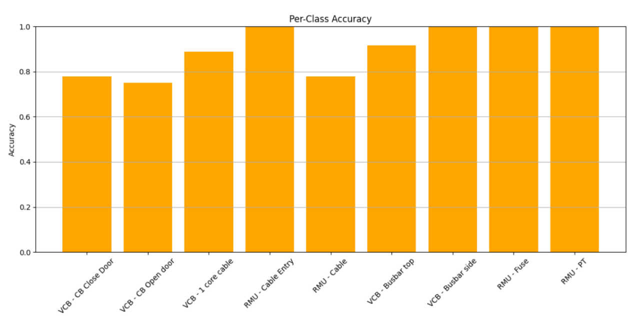

Fig. 6. Per-Class Accuracy Bar Chart

2) Confusion Matrix Analysis: The confusion matrix shown in Figure 7 further affirms the classification model's reliability. Diagonal dominance across all classes reflects strong class-wise prediction consistency. Minimal off-diagonal errors suggest very limited confusion between structurally or thermally similar compartments. For example, VCB - CB Close Door occasionally confused with VCB - CB Open Door—a reflection of shared enclosure geometry and overlapping surface heatmaps. Interestingly, VCB - 3 Core Cable lacks evaluation samples due to dataset imbalance. While class weights were applied during training to counteract this, it remains a future priority to collect more balanced data.

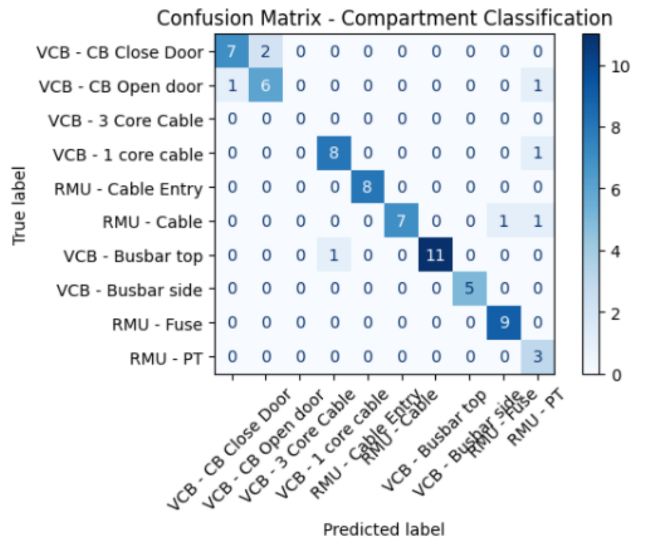

Fig. 7. Confusion Matrix – Compartment Classification

3) F1 Score Analysis: F1 scores provide a balanced view of precision and recall per class. Figure 8 reveals that 9 out of 11 classes scored above 0.85, with multiple classes (e.g., RMU - Cable Entry, VCB - Busbar Side) achieving $>$ 0.95. VCB - CB Open Door and RMU - PT yielded the lowest F1 scores (0.68–0.75), suggesting the need for more diverse examples or improved segmentation support.

The absence of an F1 score for VCB - 3 Core Cable once again reflects its absence in the test subset and indicates the need for synthetic augmentation or focused field capture [10] [18].

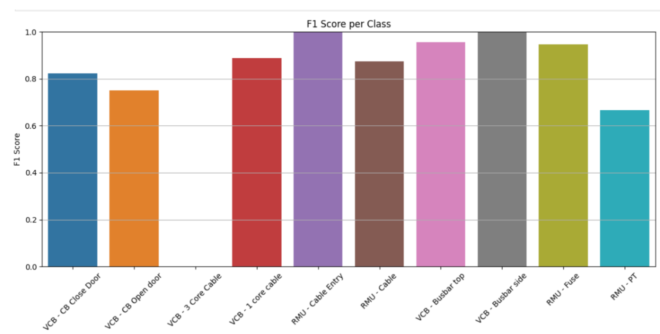

Fig. 8. F1 Score per Class

B. Segmentation Performance with SAM

1) Visual and Quantitative Evaluation: The SAM-based segmentation module was tested for its ability to localize thermal ROIs using a single center-point prompt. Figure 9 displays segmentation masks sorted by predicted SAM-predicted IoU confidence. The top-performing mask achieved a SAM-predicted IoU confidence score of 0.92, closely followed by 0.91. These values indicate strong model confidence and visually consistent ROI localization.

The lowest SAM-predicted IoU confidence score (0.73) still demonstrated reasonable thermal integrity, proving that

SAM can function with minimal supervision [7] [12] [17].

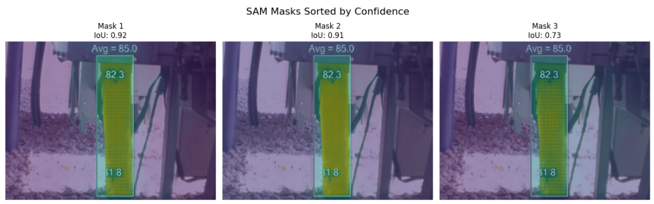

Fig. 9. SAM Masks Sorted by Confidence

2) Segmentation Consistency Metrics: Table III summarizes the SAM-predicted confidence and thermal consistency metrics for the generated masks. The three masks achieved predicted IoU confidence scores of 0.92, 0.91 and 0.73, with a mean of 0.85 and a low standard deviation of 0.09, indicating a stable ROI location. In addition to the predicted confidence of the IoU, thermal consistency was assessed using the average ROI temperature and thermal contrast ΔT. All masks retained the same average ROI temperature of 85.0°C, while the extracted thermal region maintained a ΔT of 50.5°C, confirming that the segmented ROI remained thermally consistent for anomaly analysis. Since ground-truth masks were not available, the reported IoU values represent SAM-predicted confidence scores rather than direct segmentation accuracy [5] [7].

TABLE III

COMPACT SAM SEGMENTATION CONSISTENCY AND THERMAL ROI METRICS

Metric	**Mask 1**	**Mask 2**	**Mask 3**	**Overall**
Predicted IoU confidence	0.92	0.91	0.73	Mean: 0.85
Confidence drop	0.00%	1.09%	20.65%	Range: 0.19
Avg. ROI temperature	85.0°C	85.0°C	85.0°C	Stable
Thermal contrast ΔT	50.5°C	50.5°C	50.5°C	Stable

C. Thermal Overlay and ΔT Scoring Interpretation

Following successful classification and segmentation, each image underwent ΔT scoring using the eq. 1.

A fixed threshold of 37°F was used as a cut-off for labeling regions as either Normal or Faulty. This threshold was derived from empirical fault diagnosis guidelines used in local substation inspection SOPs [10] [8].

VI. CONCLUSIONS AND DISCUSSION

This study presents a unified deep learning framework for automated thermal fault detection in TNB substations by integrating MobileNetV2 for compartment classification, SAM for ROI segmentation, and ΔT-based severity assessment. Experimental results show high classification accuracy, reliable segmentation, and consistent thermal scoring, confirming the feasibility of the proposed pipeline for real-world substation monitoring.

The modular architecture supports independent optimization of classification, segmentation, and thermal analysis while maintaining an efficient end-to-end diagnostic pipeline. Its lightweight design enables CPU-based and edge deployment, such as on Jetson Nano, while the classification outputs, segmentation masks, and ΔT overlays provide interpretable visual cues for rapid anomaly identification.

Future work will expand the dataset, improve pixel-level fault localization, compare SAM with other pretrained segmentation models, and adopt adaptive severity thresholds to enhance diagnostic precision and scalability across diverse substation environments.

ACKNOWLEDGMENT

The authors extend their sincere thanks to Tenaga Nasional Berhad (TNB) and the contributing field engineers for their support, insights, and thermal image contributions throughout the project.

REFERENCES

[1] H. Zhang, Z. Luo, and L. Feng, "Lightweight thermal fault classification using mobilenetv2 for real-time power substation inspection," *IEEE Access*, vol. 13, pp. 10 796–10 809, 2025.

[2] A. Rahmoune, T. Benslimane, and D. Zegour, "A rule-based thermal scoring approach for substation diagnostics," *Engineering Failure Analysis*, vol. 62, pp. 1–10, 2016.

[3] H. Mu, D. Li, and W. Ren, "Rule-based thermal fault analysis for high-voltage accessories," *Energy Reports*, vol. 7, pp. 1049–1056, 2021.

[4] K. Leksir and B. Medjahed, "Infrared thermal image segmentation using svm and clustering algorithms," *Journal of Electrical Systems*, vol. 15, no. 2, pp. 245–257, 2019.

[5] Q. Lin, Y. Zhao, and F. Zhou, "Lmdfusion: Lightweight multi-scale dense fusion for thermal-optical image segmentation," *Pattern Recognition Letters*, vol. 173, pp. 21–29, 2024.

[6] Y. Mahmoud, A. Elbaz, and F. Salem, "Transfer learning for thermal anomaly detection in power systems using efficientnet and resnet variants," *Applied Soft Computing*, vol. 145, p. 110970, 2024.

[7] Y. Chen, B. Xu, and Z. Yuan, "Roi segmentation in substation thermograms using sam and deeplabv3 fusion," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 2, pp. 1984–1993, 2024.

[8] D. Hoffmann, B. Müller, and F. Seidel, "Component-level deltat estimation in substation thermograms using region-based cnns," *Energies*, vol. 13, no. 17, p. 4506, 2020.

[9] Y. Lu, R. Cheng, and Y. Fang, "Lightweight mobilenetv2 for thermal fault classification in substation inspection," *Sensors*, vol. 23, no. 2, p. 798, 2023.

[10] F. Bommes, P. Leclerq, and J. Hay, "Cnn-based roi detection with deltat-based severity scoring for rmu diagnostics," *Electric Power Components and Systems*, vol. 49, no. 14-15, pp. 1215–1225, 2021.

[11] Y. Chen, Z. Wei, and D. Huang, "Infrared image-based substation fault detection using attention-enhanced cnn," *Sensors*, vol. 22, no. 11, p. 4178, 2022.

[12] Y. Li, L. Deng, and F. Yang, "Infrared image segmentation via deep self-attention and multi-factor similarity," *Infrared Physics & Technology*, vol. 132, p. 104608, 2024.

[13] R. Balakrishnan, A. Arumugam, and R. Kandasamy, "Thermal fault detection using ml-based preprocessing and normalization for electrical switchgear," *International Journal of Thermal Sciences*, vol. 184, p. 107971, 2022.

[14] A. Venegas, J. Santos, and L. Ruiz, "Ai-enhanced automated thermal inspection in industrial substations," *IEEE Access*, vol. 10, pp. 110 234–110 245, 2022.

[15] D. Alwar, U. Farooq, and M. Khan, "Adaptive ai models for fault localization in substation switchgear," *Electric Power Systems Research*, vol. 213, p. 108919, 2022.

[16] Z. Hrgović and I. Pavić, "Substation topology reconfiguration using reinforcement learning and ptdf analysis," *Energies*, vol. 17, no. 1, p. 345, 2024.

[17] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, *et al.*, "Segment anything," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 4015–4026.

[18] A. Aparna, R. Gupta, and P. Mishra, "Transfer learning for road surface defect detection using thermal images," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2191–2201, 2019.