Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Research on Infrared Thermal Fault Detection for Transformer Bushings Based on the LTF-YOLO Algorithm

Zeyuan Fu

2023212565@bupt.cn

Beijing University of Posts and Telecommunications

Article

Keywords: Transformer bushings, Infrared thermal fault diagnosis, YOLOv8, Intelligent power inspection

Posted Date: June 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9613303/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Research on Infrared Thermal Fault Detection for Transformer Bushings Based on the LTF-YOLO Algorithm

Zeyuan Fu[1]

[1] Beijing University of Posts and Telecommunications, Beijing City, China
Corresponding Author E-mail: 2023212565@bupt.cn
Data availability: The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

***Abstract*—Transformer bushings are critical insulation components in power systems and the early diagnosis of thermal faults plays an important role in ensuring safe and reliable operation. Conventional infrared inspection methods rely heavily on expert experience and suffer from low efficiency while existing deep learning based approaches often struggle with small scale hot spots weak temperature contrast and complex background interference leading to missed detections and false alarms. To address these challenges this study proposes a lightweight detection algorithm named LTF-YOLO that integrates three complementary improvement strategies.**

The proposed method is developed on the YOLOv8n baseline and enhances feature representation at multiple levels. A parameter free SimAM attention module is embedded at the end of the backbone network to strengthen spatial perception of early stage weak thermal anomalies based on neural energy modeling. A DySample dynamic upsampling mechanism is introduced to improve feature map reconstruction accuracy through content aware sampling point offsets which effectively enhances contour restoration for small scale bushing targets. In addition a C2f MSDA module is constructed to capture cross scale contextual information using parallel dilated convolutions enabling joint modeling of local details and global temperature rise patterns.

Experimental results on a self constructed infrared image dataset of transformer bushings demonstrate that the proposed LTF-YOLO algorithm achieves a mean average precision of 0.951 with only a slight increase in model parameters corresponding to an improvement of 6.25 percent compared with the baseline YOLOv8 model. In complex scene evaluation the performance gain further increases to 9.07 percent. These results indicate that the proposed approach satisfies the requirements for accurate identification of bushing overheating defects and its effectiveness and superiority are further validated through ablation and comparative experiments.

***Index Terms*—Transformer bushings, Infrared thermal fault diagnosis, YOLOv8, Intelligent power inspection**

I. INTRODUCTION

Power transformers serve as the core hub for energy conversion and transmission in power grids and their operating condition has a direct impact on supply reliability. As the key insulation component connecting internal windings with external transmission lines transformer bushings perform multiple functions including electrical insulation mechanical support and sealing protection [1,2]. During long term operation bushings are prone to local or overall abnormal temperature rise which is commonly referred to as thermal faults. The heating behavior of transformer bushings mainly originates from two fundamental mechanisms. One source is the Joule heat generated when load current flows through the current carrying conductor inside the bushing while the other source arises from dielectric losses including polarization loss and conduction loss under alternating electric fields. When contact defects insulation degradation or oil shortage occurs these heat generation mechanisms are significantly intensified leading to localized overheating [3,4]. In addition surface contamination under humid conditions can produce leakage current and internal partial discharge may introduce extra heat which further distorts the surface temperature distribution of the bushing. Statistical studies indicate that bushing related failures account for approximately twenty five to forty percent of all transformer faults and most of these failures experience a noticeable heating stage before developing into severe accidents such as explosion or fire. Therefore continuous and accurate monitoring of bushing operating temperature together with early fault diagnosis is of critical practical importance for preventing major equipment damage and ensuring the safe and economical operation of power grids [5-7].

Infrared thermography is a non contact technique that detects infrared radiation emitted from object surfaces and converts it into temperature distribution images which enables fast and intuitive identification of overheating defects in electrical equipment. It has become a standard and indispensable tool for condition monitoring in power systems. However conventional infrared diagnosis relies heavily on the experience of maintenance personnel. Large numbers of infrared images must be manually inspected to visually identify abnormal temperature regions and determine fault types based on operational guidelines. This practice suffers from several limitations including high labor cost and low efficiency which make it difficult to meet the growing demands of equipment scale and inspection frequency. The diagnosis process is also highly subjective and sensitive to individual experience and working conditions resulting in poor consistency. Early stage weak or concealed thermal defects are often overlooked and historical data are rarely explored for long term trend analysis[8-11].

Early studies on infrared thermal fault diagnosis of transformer bushings mainly relied on manual inspection and traditional image processing techniques. Tian Suhuimin [12] and co workers proposed a diagnostic framework based on infrared image preprocessing temperature anomaly detection and localization in which edge detection and threshold segmentation were used to identify fault regions. Although this type of method showed effectiveness in relatively simple environments it suffered from limited adaptability due to handcrafted feature extraction strong subjectivity in decision making and low processing efficiency which made it difficult to handle the large volume of infrared images generated during substation inspections and to detect weak defects under complex backgrounds. With the development of deep learning convolutional neural networks were introduced into bushing fault identification tasks. Liu Yunpeng [13] and colleagues constructed a dataset containing normal oil deficient and locally overheated conditions and employed convolutional neural networks to directly classify infrared images achieving significantly higher accuracy than traditional support vector machines and back propagation neural networks. However this approach operated at the image level and required prior cropping of bushing regions which prevented automatic fault localization and limited its ability to capture multi scale and small target features. To enable end to end detection and automatic localization subsequent studies shifted toward object detection frameworks represented by the YOLO series. Zhang Run [14] and co authors proposed an improved YOLO based method that incorporated attention mechanisms and optimized detection heads to achieve accurate bushing localization and combined image alignment with temperature feature extraction for fault diagnosis. To address the scarcity of abnormal samples Wang Ziyao [15] and colleagues introduced a cross domain data generation strategy based on generative adversarial networks which significantly improved recognition performance under limited data conditions. Focusing on deployment in ultra high voltage transformer applications Liu Chang [16] and co workers developed a lightweight detection model by redesigning multi scale feature extraction modules suppressing background interference and reducing model complexity which improved both accuracy and computational efficiency. Despite these advances existing methods still face challenges including the small proportion of bushings in infrared images insufficient multi scale feature fusion strong background interference caused by oil contamination and illumination changes and the inherent trade off between high detection accuracy and real time performance which indicates the need for further improvement of robustness and efficiency in infrared thermal fault diagnosis of transformer bushings.

To address these challenges this study takes the high performance YOLOv8 framework as the baseline and proposes an enhanced model that integrates three complementary improvement strategies to construct a lightweight and high accuracy intelligent diagnosis method for transformer bushing thermal faults referred to as the LTF-YOLO algorithm. First a parameter free SimAM [17] attention module is introduced into key layers of the backbone network to achieve adaptive feature enhancement based on neural energy modeling which improves sensitivity to weak thermal anomalies without increasing model parameters. Second a dynamic upsampling strategy based on content aware sampling is employed to replace conventional interpolation in the feature pyramid and segmentation head which significantly improves small target detail recovery and boundary accuracy [18]. Finally the original feature fusion module is redesigned using multi scale dilated attention to integrate parallel dilated convolutions and attention mechanisms which jointly optimize local detail preservation and global contextual perception [19].

II. LTF-YOLO Algorithm Principles

A. YOLOv8 Algorithm

YOLOv8 is a single stage object detection framework proposed by Ultralytics in 2023. Building upon the design philosophy of the YOLOv5 series it introduces a systematic redesign of both the backbone network and the detection head. A key architectural update is the replacement of the C3 module with the C2f module which adopts a gradient flow strategy inspired by ELAN structures. This design effectively alleviates gradient vanishing in deep networks while reducing the overall number of parameters. In terms of detection paradigm YOLOv8 abandons the traditional anchor based matching strategy and adopts an anchor free formulation with a decoupled head that separately handles classification and bounding box regression. In the regression branch distribution based focal loss is employed to model bounding box localization which improves positioning accuracy particularly for medium and small sized targets. In addition the framework unifies object detection instance segmentation and image classification within a single architecture allowing flexible switching between different task heads. From an engineering perspective YOLOv8 provides multiple model scales ranging from lightweight to large capacity variants and incorporates efficient design strategies such as depthwise separable convolutions. This enables the model to maintain competitive detection accuracy while significantly reducing computational complexity making it well suited for real time inference on edge computing devices.

B. LTF-YOLO Algorithm

This study takes YOLOv8n as the baseline model and focuses its improvements on the tail end of the backbone network the C2f modules in the neck and the upsampling strategy. Figure 1 clearly illustrates the overall architecture of the improved model.

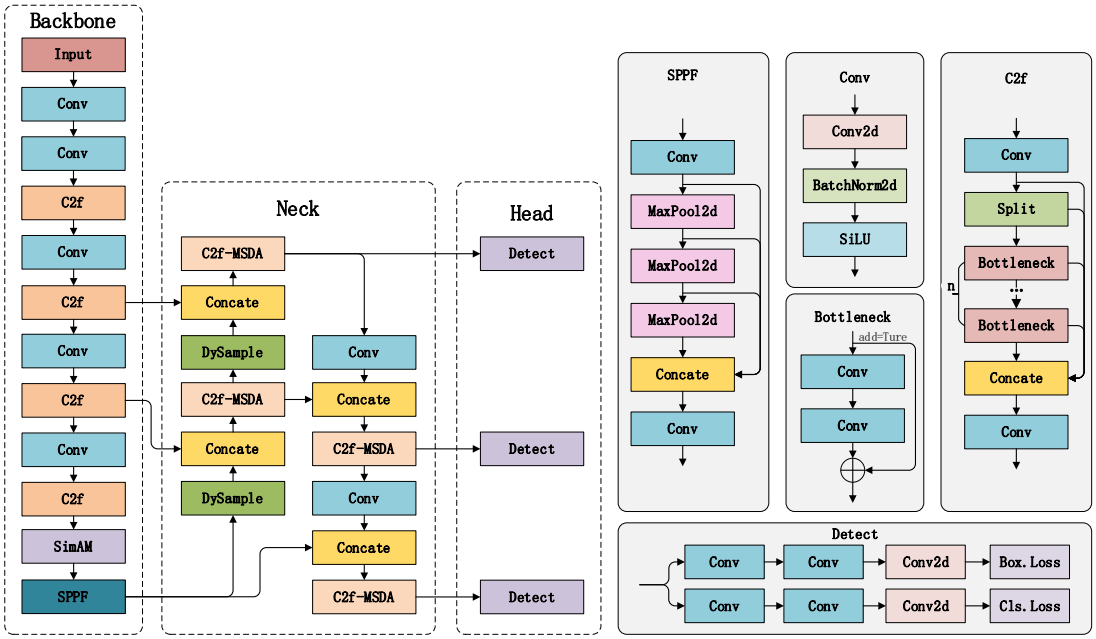


Fig. 1. LTF-YOLO Algorithm Structure.

For the proposed LTF-YOLO method this study takes the YOLOv8 model as the baseline and introduces a set of targeted enhancements that form a coordinated improvement framework. First a parameter free SimAM attention module is integrated. Driven by a neuroscience inspired energy function this module can automatically identify and strengthen neuron responses that differ markedly from their surrounding context. As a result early and subtle local thermal anomalies in infrared images are more clearly highlighted while background thermal noise is effectively suppressed. Next a DySample dynamic upsampling strategy is adopted. By using a lightweight network to predict content dependent sampling point offsets it replaces fixed interpolation kernels and enables more precise feature map reconstruction. This significantly improves the contour recovery of small scale bushing targets as well as the boundary quality of segmentation masks. Finally an MSDA multi scale dilated attention module is employed. Parallel multi branch dilated convolutions are used to capture contextual information under different receptive fields and the features are adaptively fused through attention weights. This allows the model to simultaneously understand fine local details and global structural patterns and effectively addresses the detection difficulty caused by large differences in fault physical scales. Together these three improvements systematically tackle key challenges in transformer bushing infrared diagnosis including weak early fault signatures loss of small target details strong interference from complex backgrounds and the coexistence of multi scale faults. The result is a clear gain in diagnostic accuracy while maintaining a lightweight model suitable for practical deployment.

C. Simple Attention Module

SimAM is an attention mechanism based on local self similarity within feature maps and operates as a purely feed forward module that introduces almost no additional computational cost or model parameters. In this study the SimAM attention mechanism assigns adaptive weights to each point in the feature map across both channel and spatial dimensions which directly enhances the contrast between transformer bushing thermal fault regions and their surrounding background in infrared images. By amplifying weak local thermal anomalies while suppressing complex background interference commonly found in substations the method effectively improves feature discrimination and at the same time meets the strict requirements for lightweight design and real time performance in practical engineering deployment.

The implementation of SimAM relies entirely on forward computation without the need for parameter learning through backpropagation. The process begins with feature extraction by a convolutional neural network where the input feature map is denoted as X with dimensions C by H by W representing channel height and width respectively. Each value at a specific spatial location and channel is treated as an individual neuron. For each pixel position the mechanism evaluates its similarity to neighboring pixels by computing the normalized mean of the squared differences between the pixel value and those within its local neighborhood. This operation provides an indirect measure of local contrast relative to the surrounding region and enables the network to emphasize neurons that carry more distinctive information for fault characterization.

$$S_{i,j} = \frac{1}{N} \sum_{k \in \Omega_{i,j}} \left\| x_{i,j} - x_k \right\|_2^2 \tag{1}$$

In this context $\Omega_{i,j}$ denotes the local neighborhood of pixel $x_{i,j}$ excluding the pixel itself and N represents the number of pixels within that neighborhood. During implementation SimAM performs centering by using the mean value of the entire feature map and then computes the squared difference after centering to simplify the calculation. Based on the resulting $s_{i,j}$ obtained from this process the attention weight $w_{i,j}$ is subsequently generated through the corresponding formulation.

$$w_{i,j} = \frac{1}{1+\exp(-\frac{1}{4}(\frac{s_{i,j}}{\sigma_{i,j}^2+\varepsilon}-1))} \tag{2}$$

In this formulation $\sigma^2_{i,j}$ represents the normalized form of $s_{i,j}$ which is approximated using the mean and standard deviation of $s_{i,j}$ computed over the entire feature map or a local region. A small constant ε is introduced to prevent numerical instability caused by division by zero. Through this normalization s_{ij} is mapped into the range between zero and one and serves as the attention weight. The resulting attention map is then multiplied elementwise with the original feature map X to produce a reweighted feature representation. This process enables adaptive feature adjustment and enhances the expressive capability of the network.

By computing local similarity transforming it into attention weights and recalibrating features accordingly the SimAM attention mechanism significantly improves the network sensitivity to informative patterns without introducing any additional trainable parameters. Considering the complex background of transformer bushing infrared images and the weak thermal signatures present at the early stage of overheating faults SimAM is embedded into the YOLOv8 backbone at the layer preceding the SPPF module. This design enhances the extraction of discriminative features while maintaining the original computational cost and leads to a measurable improvement in detection accuracy.

D. DySample Dynamic Oversampling Mechanism

DySample is an extremely lightweight and efficient dynamic upsampling method. It adopts a point based sampling strategy which differs from earlier kernel based dynamic upsampling approaches and therefore requires far fewer parameters floating point operations GPU memory and inference latency. In this work the adaptive nature of DySample allows the upsampling process to adjust automatically to different infrared thermal patterns of transformer bushings such as uniform heating under normal conditions or localized intense overheating under fault conditions. By enabling more precise reconstruction of feature maps during upsampling it effectively enhances the model's ability to detect small scale thermal anomalies in infrared images while maintaining high efficiency and practicality for real world applications.

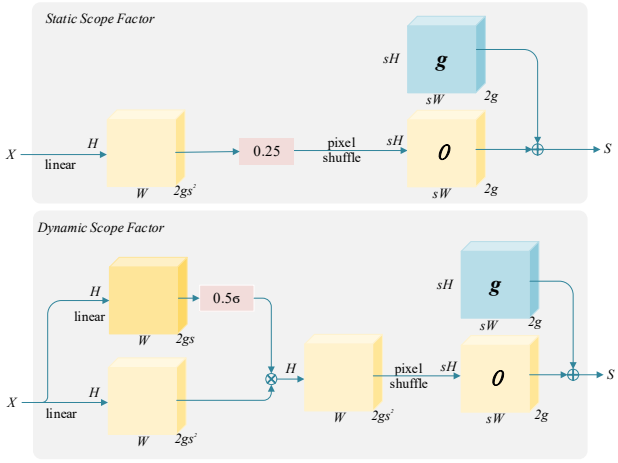


Fig. 2. DySample Dynamic Oversampler and Sample Point Generator.

The figure illustrates the sampling point generator used in DySample. Given an upsampling scale factor s and an input feature map X with dimensions C by H by W, O is defined as the reference point for subsequent offset prediction. A linear layer with C input channels and $2s$ squared output channels is then applied to generate an offset map with dimensions $2s$ squared by H by W, which means that for each input location the model predicts s by s groups of offsets. These offsets are rearranged using pixel shuffle to obtain a tensor of size 2 by sH by sW. Here σ denotes the sigmoid function, which is used to produce a scaling factor that constrains the offset range. The final sampling set is obtained by dynamically adding the predicted offsets to the original sampling grid G, and the upsampled feature map X prime with dimensions C by sH by sW is produced accordingly. In addition, a static range factor is introduced by combining a linear layer with pixel shuffle and a fixed scaling coefficient, and in this work the static range factor is set to 0.25 to stabilize the offset magnitude.

$$O = 0.25\text{Linear}(X) \tag{3}$$

In addition to the linear projection and pixel shuffle operations, a dynamic range factor is introduced to further improve flexibility. This factor is generated by applying a sigmoid function together with a fixed coefficient of 0.5 and is obtained through a linear projection of the input features. As a result, a pointwise range scaling factor is produced for each spatial location, allowing the sampling offsets to adapt more precisely to local feature variations while maintaining numerical stability and low computational cost.

$$O = 0.25\text{sigmoid}(\text{Linear}_1(X)) \times \text{Linear}_2(X) \tag{4}$$

D. Multi-Scale Expansion Attention Module

For infrared thermal fault detection of transformer bushings the scale of faults varies widely ranging from millimeter level micro cracks that are highly localized to overall temperature rise across the entire porcelain bushing at a global level which makes it difficult for conventional models to represent these patterns effectively at the same time. In addition strong thermal interference is often present near the bushing region such as heat emitted from adjacent conductors or reflections caused by sunlight which can easily draw model attention toward the

background and lead to missed detections or false alarms. To address these issues this study enhances the original YOLOv8 framework by introducing a multiscale dilated attention mechanism. The proposed module expands the receptive field without information loss by applying parallel dilated convolutions with dilation rates of one three and five allowing local fine grained features and global structural patterns to be captured simultaneously within a single network layer. As a result the model is able to detect subtle temperature anomalies associated with small cracks while also recognizing large scale heating patterns caused by moisture ingress or insulation degradation. Moreover the module adaptively learns the relative importance of features at different scales so that small dilation branches are emphasized for tiny fault targets whereas larger dilation branches dominate when extensive abnormal heating occurs. This adaptive scale selection enables the model to respond flexibly to different operating scenarios rather than relying on manually defined fixed scales and simplified assumptions. The overall structure of the MSDA module is illustrated in Figure 3.

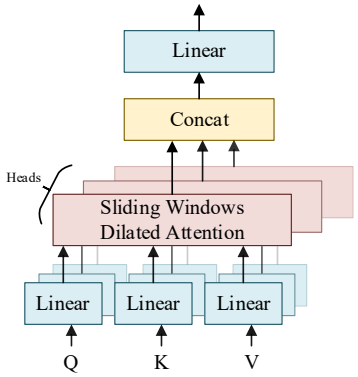


Fig. 3. MSDA Structure.

Below is a single-paragraph academic English translation written in a natural American journal style. It avoids lists punctuation like colons and parentheses keeps the language professional but readable and is phrased to reduce an obvious AI tone.

In the MSDA module the feature map channels are first divided into multiple heads and self attention is then performed within each head using different dilation rates. This parallel design improves both learning capacity and computational efficiency. By assigning distinct dilation rates to different heads the module focuses on features at multiple spatial scales and effectively enlarges the receptive field without increasing parameters or reducing resolution. The three branches are responsible for capturing short range medium range and long range contextual information respectively. The resulting features are then concatenated and passed through a linear layer for unified aggregation which allows the network to interpret image content across scales and form a more comprehensive representation. Through this mechanism MSDA is able to preserve fine local details while simultaneously modeling broader contextual patterns which leads to stronger feature expressiveness. In this study MSDA is integrated into the C2f module to form the proposed C2f MSDA structure which enhances robustness to varying object scales and input resolutions while maintaining high computational efficiency. The overall architecture of the C2f MSDA module is illustrated in Fig. 4.

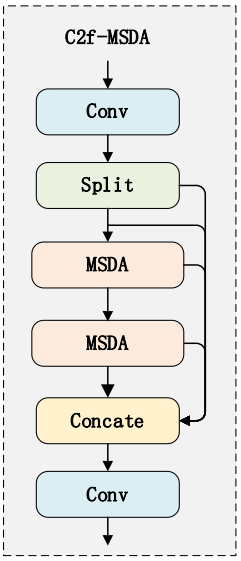


Fig. 4. C2f-MSDA Module Structure.

III. Experiments and Results Analysis

A. Test Dataset Generation

To comprehensively evaluate the effectiveness of the proposed LTF-YOLO algorithm for infrared thermal fault detection of transformer bushings and to examine its performance in complex real-world scenarios, a diverse and representative infrared image dataset of transformer bushings was constructed. The dataset integrates samples from multiple sources including publicly available datasets and images collected in practical engineering applications in order to ensure sufficient diversity and coverage.

The data were obtained from three main sources. First, through cooperation with ultra high voltage power companies and operation and maintenance units in several provinces and cities, 1327 infrared images were collected during routine inspections and special tests at substations. These images cover voltage levels ranging from 220 kV to 1000 kV and include bushings from different manufacturers and models under various environmental conditions such as sunny rainy foggy daytime and nighttime scenarios. Second, 165 infrared images of transformer bushings with overheating faults were gathered from open online resources. Third, to address the scarcity of typical and especially severe fault samples in real operating conditions, multi-physics simulations were carried out. By constructing accurate three dimensional bushing models and setting boundary conditions such as material properties load current and ambient temperature, steady state and transient temperature fields were simulated for normal conditions as well as typical fault cases including cracks oil loss moisture ingress through modified dielectric loss factors and joint overheating caused by increased contact resistance. The simulated temperature fields were rendered as infrared thermal images resulting in 231 additional samples.

All collected images were carefully screened to remove those that were blurred overexposed severely occluded or otherwise invalid. The remaining images were manually annotated by professionals with power engineering backgrounds using the LabelImg tool and random sampling checks confirmed an annotation accuracy exceeding 98 percent. To enhance model generalization and reduce overfitting various data augmentation strategies were applied to the training set to simulate complex inspection conditions. These included geometric and photometric transformations Gaussian noise injection and Mosaic augmentation which expanded the training data to twice the size of the original set. The final dataset contains 3716 high quality annotated infrared images and was randomly divided into training validation and test sets at a ratio of 8 to 1 to 1 while ensuring balanced category distributions across all subsets.

B. Training Environment and Hyperparameter Settings

All experiments in this study were conducted on a unified hardware-accelerated computing platform. The system was equipped with an NVIDIA GeForce RTX 5060 graphics processing unit and an Intel Core i5-12600KF central processing unit. The implementation was based on a Python 3.12 environment using the PyTorch 2.8.0 deep learning framework with CUDA 12.8 enabled to support parallel acceleration. During training the stochastic gradient descent optimizer was employed with a momentum value of 0.937 and a weight decay coefficient of 0.0005. The batch size was configured using an adaptive strategy that automatically adjusted to available GPU memory in order to maximize training throughput. The maximum number of training epochs was set to 3000 and an early stopping mechanism was applied such that training was terminated when no improvement was observed on the validation set for 200 consecutive epochs thereby reducing the risk of overfitting.

C. Evaluation Indicators

To systematically quantify the overall performance of the proposed method this study establishes an evaluation framework that jointly considers accuracy efficiency and deployment feasibility. Model accuracy is primarily assessed using the mean average precision at an IoU threshold of 0.5 which is obtained by integrating the area under the precision recall curve for each defect category and then computing their arithmetic mean. This metric provides a comprehensive measure of detection robustness across different confidence levels and reflects the model's ability to maintain stable recognition performance under varying decision thresholds.

$$\mathrm{m}AP = \frac{1}{N}\sum_{i=1}^{N}\int_{0}^{1} P_i(R)dR \tag{5}$$

In this formulation N denotes the total number of discharge defect categories while $P_i(R)$ evaluated at R represents the interpolated precision of the i th category at a given recall level R. In addition to accuracy computational complexity is evaluated using the number of billions of floating point operations required for a single forward inference which directly reflects the model's potential for real time processing on edge devices. Model compactness is further assessed through the number of trainable parameters and the disk storage size of the model which together constrain practical deployability under resource limited conditions. The parameter count indicates peak memory usage during inference while the model size affects storage requirements and loading overhead.

D. Ablation Experiment

To verify the contribution of each proposed enhancement an ablation study was conducted and the results are reported in Table 1. The experiments start from the original YOLOv8 model and progressively incorporate the improved modules while monitoring changes in key performance metrics. SimAM DySample and MSDA are added in sequence to evaluate the individual and cumulative effects of each component on mean average precision model parameter count model size and computational cost measured in GFLOPs. The detailed results of this step by step analysis are summarized in Table 1.

TABLE I
LTF-YOLO Ablation Study.

Serial	Baseline	SimAM	DySample	MSDA	mAP	GFLOPs	Parameters/M	Size/M
1	√				0.895	8.1	3.01	6.04
2	√	√			0.908	8.1	3.01	6.04
3	√	√	√		0.925	8.5	3.13	6.17
4	√	√	√	√	0.951	8.9	3.21	6.26

The ablation study shows that each proposed module contributes positively to overall model performance with a consistent increase in mAP and an improvement of 0.071 compared with the original YOLOv8 baseline. These results demonstrate that the proposed MDW YOLOv8 model effectively enhances the accuracy of infrared thermal fault detection for transformer bushings. Representative detection examples produced by the final model are presented in Figure 5.

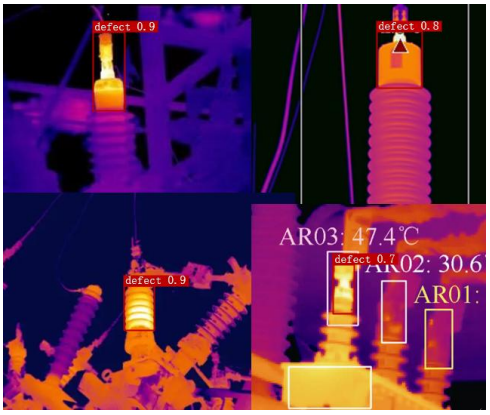


Fig. 5. Algorithm Detection Example.

During the ablation study the SimAM attention mechanism was first introduced. Compared with the baseline YOLOv8 model the mAP increased by 0.013 while the number of parameters and the model size remained unchanged. This result indicates that the parameter free SimAM mechanism effectively enhances the representation of weak features in transformer bushing regions while preserving the lightweight nature of the model. On this basis the original upsampling operation was then replaced with the DySample module which further improved mAP by 0.017 with only a slight increase in parameter count and model size. This demonstrates that DySample enhances feature reconstruction in overheated fault regions through adaptive upsampling. Finally the C2f module was upgraded with MSDA leading to an additional mAP gain of 0.026 while GFLOPs increased to 8.9 and the parameter count rose to 3.21 which highlights the strength of multi scale modeling enabled by MSDA. Although the final model introduces a moderate increase in computational cost it achieves an accuracy above 0.95 and maintains a compact model size which is suitable for subsequent edge deployment. Overall these ablation results confirm the effectiveness of the proposed LTF-YOLO improvements for infrared thermal fault detection of transformer bushings.

E. Comparative Experiment

To evaluate the advantages of LTF-YOLO and to comprehensively verify the superiority of the proposed improved model over mainstream object detection methods this study conducts comparative experiments on the same test set. The proposed approach is compared with several representative object detection algorithms including other variants from the YOLO family as well as Faster R-CNN[22] RepViT[23] and RT-DETR[24]. All experiments are carried out under identical experimental settings and the corresponding results are reported in Table 2.

TABLE II
The LTF-YOLO comparison test.

Algorithm	mAP	GFLOPs	Parameters/M	Size/M
YOLOv5	0.884	7.1	2.51	5.1
YOLOv6	0.877	11.9	4.23	8.4
YOLOv8	0.895	8.1	3.01	6.0
YOLOv9	0.876	7.7	2.04	4.1
Faster R-CNN	0.902	143.5	41.2	67.8
RepViT	0.883	11.7	4.12	7.3
RT-DETR	0.921	105.2	29.26	55.7
MDS-YOLO	0.951	8.9	3.21	6.26

Based on the comparative results discussed above the proposed LTF-YOLO algorithm demonstrates a clear overall advantage in detection accuracy while achieving an effective balance between precision and efficiency. Among the competing methods RT DETR delivers the second highest accuracy at 0.921 but requires substantially more parameters and computational cost than the proposed approach. In comparison with other lightweight models from the YOLO family the proposed model consistently attains higher accuracy and although its model size and computational load are slightly greater than those of some extremely lightweight variants models with comparable accuracy typically exhibit much larger model sizes and significantly higher GFLOPs than LTF-YOLO.

F. Performance Comparison in Complex Scenarios

To further examine the capability of the LTF YOLO model under specific challenges the test set was divided into representative scenarios to reflect its performance across different levels of difficulty. In this study four typical scenarios were defined including small scale targets with pixel areas smaller than 32 by 32 low contrast defects where the temperature difference between the fault region and the background is below five degrees Celsius complex background interference with multiple nearby high temperature sources and densely arranged multi target scenes captured at close range for three phase bushings. The experimental results for these scenarios are presented in Table 3.

TABLE III
Performance of the LTF-YOLO Model in Various Challenging Scenarios.

Scene Type	Sample size	YOLOv8 mAP	LTF-YOLO mAP	Increase
Small goals	361	0.852	0.933	+9.5%
Low contrast	452	0.836	0.921	+10.1%
Strong background interference	256	0.863	0.947	+9.7%
Multi-object Overlap	327	0.871	0.932	+7.0%

The above experimental results show that across all four predefined scenarios the LTF YOLO algorithm consistently achieves much higher accuracy than the original YOLOv8 model which provides strong evidence for the effectiveness of the proposed improvements. In particular the enhanced model successfully addresses the inherent challenges of transformer bushing infrared thermal fault diagnosis such as small target size weak feature responses complex background interference and large variations in fault scale. As a result LTF YOLO delivers a substantial improvement in diagnostic accuracy while maintaining a high level of inference efficiency.

IV. CONCLUSIONS

To address the key challenges in transformer bushing infrared thermal fault diagnosis including weak early stage features large variations in physical scale and strong interference from complex backgrounds this study develops a

lightweight detection model referred to as LTF YOLO. By systematically integrating parameter free feature enhancement through the SimAM attention mechanism content adaptive upsampling via DySample and multi scale dilated attention using MSDA the proposed algorithm significantly improves the sensitivity and localization accuracy of abnormal thermal regions in infrared images while maintaining a low computational cost of 8.9 GFLOPs. Ablation studies and comparisons with mainstream detection methods confirm the synergistic benefits of the three improvements. In particular the C2f MSDA module employs parallel convolution branches with dilation rates of one three and five which effectively overcomes the limitation of single scale receptive fields in jointly modeling point like crack defects and region level temperature rise caused by moisture degradation. Meanwhile the dynamic sampling strategy of DySample substantially enhances the resolution recovery of feature maps produced by the decoupled head. Experiments conducted under four challenging scenarios that closely resemble real engineering conditions demonstrate strong robustness to low signal to noise thermal anomalies and small scale heating targets with the mean average precision consistently exceeding 0.92. These results indicate that the proposed approach is well suited for practical deployment in complex substation environments. The main conclusions are as follows:

1) A transformer bushing infrared image dataset was established that contains 3716 valid samples and provides a valuable foundation for subsequent research.

2) An LTF-YOLO model tailored for transformer bushing overheating fault identification was developed and achieves higher detection accuracy than the original YOLOv8 while maintaining a lightweight design which demonstrates the effectiveness of the proposed approach.

3) Extensive experiments were conducted to verify the effectiveness of each improved module and comparative evaluations show that the overall performance surpasses that of seven other algorithms enabling reliable and accurate identification of transformer bushing overheating faults.

Data availability: The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

EFERENCES

[1] Yuan Z, Sun G, Tang H, et al. Types and mechanisms of condenser transformer bushing failures[J]. IEEE Electrical Insulation Magazine, 2023, 39(5): 26-36.

[2] Jiang J, Zhao J, Liu J, et al. Temperature Field Distribution of Transformer Bushings Considering Radiation Effects[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2025.

[3] Dai Q, Liu Y, Yu X, et al. Temperature segment compensation method of dissipation factor for insulation diagnosis in converter transformer bushing[J]. IEEE Access, 2022, 10: 10310-10316.

[4] Liu H, Wang Y, Sun J, et al. Research on temperature prediction and fault early warning of transformer bushings[C]//2024 IEEE 7th International Electrical and Energy Conference (CIEEC). IEEE, 2024: 1161-1164.

[5] Lin M, Wu G, Liu K, et al. Review of detection methods for typical faults in transformer bushings[J]. IEEE Electrical Insulation Magazine, 2024, 40(2): 33-44.

[6] Luo X, Liu R, Long Z, et al. Diagnosis and Elimination of an Overheating Fault for Valve-Side Structure in Converter Transformers[C]//2022 IEEE International Conference on High Voltage Engineering and Applications (ICHVE). IEEE, 2022: 1-4.

[7] Zhu R, Liu K, Lin M, et al. Study of Thermal Field Distribution under Internal Crack Defects in the Valve-side Bushing of the Converter Transformer[C]//2023 Panda Forum on Power and Energy (PandaFPE). IEEE, 2023: 1278-1286.

[8] Carrijo D, Pinto M M, Gomes G, et al. Bushing anomaly detection: An initial assessment[J]. Journal of Control, Automation and Electrical Systems, 2022, 33(5): 1477-1488.

[9] Wilson A N, Gupta K A, Koduru B H, et al. Recent advances in thermal imaging and its applications using machine learning: A review[J]. IEEE Sensors Journal, 2023, 23(4): 3395-3407.

[10] Hou F, Zhang Y, Zhou Y, et al. Review on infrared imaging technology[J]. Sustainability, 2022, 14(18): 11161.

[11] Wilson A N, Gupta K A, Koduru B H, et al. Recent advances in thermal imaging and its applications using machine learning: A review[J]. IEEE Sensors Journal, 2023, 23(4): 3395-3407.

[12] Tian Suhuimin, Zhang Yuning, Hu Yanru. Diagnostic Analysis of Transformer Bushing Heating Faults Based on Infrared Images [J]. China Plant Engineering, 2024, (16): 204-206.

[13] LIU Yunpeng, DONG Wangying, XU Ziqiang, et al. Infrared Image Recognition Method on Fault of Transformer Bushing Based on Convolutional Neutral Networks [J]. High Voltage Apparatus, 2021, 57 (10): 134-140.

[14] ZHANG Run, QIU Zhibin, TONG Zhipeng, et al. Detection Method of Transformer Bushing Thermal Defect Based on Infrared Image Target Detection and Skew Correction [J]. Southern Power System Technology, 2024, 18 (09): 59-68.

[15] WANG Ziyao, LIANG Huishi, WANG Wei, et al. Online Identification Method for Transformer Bushing Anomaly Based on Data Domain Transformation of Thermal Images and Inception-CNN Network [J]. High Voltage Engineering, 2023, 49 (08): 3425-3436.

[16] LIU Chang, WANG Le, HU Jiaxiang, et al. VisualDetectionAlgorithmforOilLevelofUltra-highVoltageTransformer BushingBasedonImprovedYOLOv8n [J]. Hebei Electric Power, 2025, 44 (05): 25-31.

[17] Yang L, Zhang R Y, Li L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks[C]//International conference on machine learning. PMLR, 2021: 11863-11874.

[18] Liu W, Lu H, Fu H, et al. Learning to upsample by learning to sample[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2023: 6027-6037.

[19] Jiao J, Tang Y M, Lin K Y, et al. Dilateformer: Multi-scale dilated transformer for visual recognition[J]. IEEE transactions on multimedia, 2023, 25: 8906-8919.

[20] Pei S, Wang W, Hu C, et al. MDS-YOLO Model-Based Defect Detection Method for Porcelain Insulators Using Infrared Images[J]. IET Generation, Transmission & Distribution, 2025, 19(1): e70032.

[21] Wang G, Chen Y, An P, et al. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios[J]. Sensors, 2023, 23(16): 7190.

[22] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016:779-788.

[23] Wang A, Chen H, Lin Z, et al. Repvit: Revisiting mobile cnn from vit perspective[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 15909-15920.

[24] Zhao Y, Lv W, Xu S, et al. Detrs beat yolos on real-time object detection[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 16965-16974.