JURNAL ELTEK
Vol. 24, No. 2, October 2026, pp. 80~92
p-ISSN: 1693 – 4024 | e-ISSN: 2355-0740
DOI: 10.33795/eltek.v24i2.11136

Reference-based degradation screening of 150-kV insulators using GWO-optimized one-class SVM under limited puncture-test data

Giovanni Dimas Prenata[1], Ahmad Ridhoi[2]
e-mail: gprenata@untag-sby.ac.id, ridhoi@untag-sby.ac.id
[1,2]Jurusan Teknik Elektro, Universitas 17 Agustus 1945 Surabaya

Article Information

Article History

Received August 15, 2026
Revised September 26, 2026
Published October 30, 2026

Keywords:

150-kV Insulator
One-class SVM
Grey Wolf Optimizer
Degradation Screening
Reference-based Learning
Anomaly Detection
Puncture-test Data

ABSTRACT

This study proposes a reference-based degradation screening framework for 150-kV ceramic, glass, and polymer insulators using a Grey Wolf Optimizer–One-Class Support Vector Machine (GWO-OCSVM) under limited puncture-test data. The dataset comprised 72 observations, including 36 replacement samples used exclusively as reference data and 36 existing samples reserved for screening and evaluation. Input features were constructed from normalized string position and voltage ratio relative to the material-specific reference mean. Model development employed nested leave-one-position-out validation, while the OCSVM parameters ν and γ were optimized by GWO using reference data only. REF-GWO-OCSVM achieved an AUC of 0.8299, improving over REF-OCSVM at 0.7535, while increasing the mean anomaly-score gap from 0.0560 to 0.1447. The final consensus parameters were ν=0.1338 and γ=1.8257. A conservative P_{95}=0.1970 threshold reduced reference rejection to 2/36 while flagging 9/36 existing samples. The results demonstrate that the proposed method enables reference-based anomaly screening without using existing-condition data during training or hyperparameter optimization.

Corresponding Author:
Giovanni Dimas Prenata,
Electrical Engineering Study Program,
Universitas 17 Agustus 1945 Surabaya,
Jalan Semolowaru Nomor 45, Sukolilo, Surabaya, Jawa Timur, Indonesia
Email: gprenata@untag-sby.ac.id

1. INTRODUCTION

High-voltage insulators constitute a fundamental part of overhead transmission systems because their electrical integrity determines whether energized conductors remain adequately isolated from grounded structures. The condition of an insulator cannot be assessed only from its nominal voltage rating, because dielectric deterioration, internal defects, contamination, and changes in electrical stress may modify its measured response before complete failure occurs. Puncture testing is specifically intended to investigate the ability of an insulator dielectric body to withstand steep electrical stress, and experimental work on compact puncture-test circuits has emphasized the importance of appropriate impulse generation and measurement arrangements for obtaining reproducible test information [1]. In addition, electrical stress along a string is not spatially uniform. Stray and disc capacitances alter the voltage distribution across individual units [2], while spatial electric-field investigations have demonstrated that the location of deteriorated or zero-value units affects the field distribution and can provide information concerning insulation condition [3]. Electric-field studies under adverse environmental conditions further confirm that electrical stress varies according to both position and the physical state of the insulation system [4]. These observations indicate that degradation screening from limited electrical measurements should preserve the relationship between the measured response and its physical context rather than relying on a single global threshold.

Environmental degradation provides additional evidence that an electrical measurement should be interpreted relative to an appropriate reference condition. Mathematical and experimental investigations of ceramic insulators have shown that pollution severity and its spatial distribution can substantially alter flashover voltage [5]. Experimental work involving non-uniform contamination has further used clean-condition flashover voltage as a reference for evaluating reductions associated with deteriorated operating conditions [6]. Leakage-current and

voltage characteristics have likewise been employed to evaluate the physical condition of porcelain, glass, and silicone-rubber insulators under different contamination and environmental conditions [7]. For porcelain insulation, time- and frequency-domain leakage-current indicators have been developed to distinguish variations in insulation health [8], whereas corresponding experimental studies on polymeric insulators have confirmed that leakage-current behavior changes systematically with pollution, wetting, and non-uniform contamination [9]. These studies collectively support the concept of reference-relative condition assessment, although the diagnostic variables, experimental conditions, and material domains considered in previous work differ from the limited puncture-test measurements addressed in the present study.

Machine-learning techniques have increasingly been employed to convert electrical measurements into quantitative insulation-condition information. Experimental validation using contaminated high-voltage insulators has demonstrated that machine-learning models can extract diagnostic patterns from leakage-current and environmental measurements [10]. Regression-oriented approaches have also been applied to predict continuous flashover-related parameters of polymeric insulators under contamination [11]. SVM-based condition identification has been combined with metaheuristic optimization to address the sensitivity of model performance to hyperparameter selection [12], while neural-network, SVM, and nearest-neighbor approaches have been evaluated for identifying artificial defects in high-voltage ceramic insulators [13]. Nevertheless, these approaches are predominantly supervised: their training process assumes that examples representing two or more known condition classes are available. Such an assumption becomes restrictive when only a reliable reference population is available and the condition or degradation severity of in-service specimens has not been independently labeled.

One-class learning offers a different formulation for this situation. Rather than learning a separating boundary from both reference and degraded classes, One-Class Support Vector Machine (OCSVM) estimates the support of a reference distribution using samples from that distribution alone and constructs a kernel-based decision function that is positive within the estimated reference region and negative outside it [14]. This formulation is particularly attractive for degradation screening because replacement or reference insulators can define the target condition without requiring degraded observations during model development. However, one-class models involve a fundamental trade-off between retaining legitimate reference observations and avoiding an excessively broad acceptance region. Research on one-class classifier optimization has therefore considered ways to control the volume of the accepted domain when actual outlier examples are unavailable [15]. OCSVM is also sensitive to hyperparameter selection, especially when only a limited number of reference samples are available, and calibration studies have shown that inappropriate parameter selection can strongly affect minimum-volume set estimation under small-sample conditions [16]. These characteristics motivate an optimization strategy that uses reference data only, so that subsequently evaluated in-service observations cannot leak into the construction of the normality boundary.

Grey Wolf Optimizer (GWO) provides a population-based mechanism for searching such OCSVM hyperparameters without requiring gradient information. The original GWO formulation models the leadership hierarchy of alpha, beta, delta, and omega wolves and updates candidate solutions through iterative encircling and hunting mechanisms [17]. Nevertheless, optimizing a model and assessing its performance using the same resampled observations can produce optimistic estimates. Nested cross-validation separates the parameter-selection process from the outer evaluation process and has been shown to substantially reduce model-selection bias [18]. More generally, the cross-validation structure should reflect the dependency and prediction objective of the data rather than being selected only for computational convenience [19]. After a continuous anomaly score has been obtained, receiver-operating-characteristic analysis and the area under the ROC curve provide threshold-independent measures of how consistently one group receives higher scores than another [20]. For one-class degradation screening, however, such labeled separation metrics should be used post hoc, rather than as optimization objectives, when the intention is to preserve genuinely reference-only model development.

Despite this progress, two gaps remain relevant to limited puncture-test condition assessment. First, recent high-voltage-insulator studies mainly use labeled contamination, defect, image, or leakage-current classes and therefore assume that representative fault labels are available [21] [22] [23]. More recent anomaly-detection work has shown the value of learning primarily from normal or reference examples when faults are rare [24] [25], while contemporary OCSVM research continues to examine how decision scores can be calibrated and interpreted probabilistically [26]. However, these developments have not directly addressed a small, position-indexed puncture-test dataset in which replacement units can define a credible reference population but Existing observations do not have independently verified damage labels. Second, the literature provides little guidance on how to combine reference-only OCSVM learning, position-grouped nested validation, and metaheuristic hyperparameter search without allowing the in-service observations to influence the learned reference boundary. These gaps motivate a reference-only screening formulation rather than another supervised condition classifier.

Four recent studies by the present authors provide directly relevant methodological and application context. A Support Vector Machine study on electrical-distribution reliability demonstrated the feasibility of margin-based learning under a small electrical dataset [27]. For dielectric-condition assessment, a transformer-oil study compared gradient-based, genetic-algorithm, and Grey Wolf Optimizer training for breakdown-voltage prediction, providing a

related example of metaheuristic optimization applied to limited high-voltage insulation data [28]. A subsequent transformer-oil study used Support Vector Regression and K-Nearest Neighbor regression with nested grouped validation to reduce information leakage and quantify split sensitivity in repeated experimental data [29]. In addition, unsupervised K-Means and Fuzzy C-Means clustering has been applied to thermal-condition patterns of high-voltage electrical equipment when predefined class labels are undesirable or incomplete [30]. The present study extends these lines of work in a different direction by combining replacement-only OCSVM learning, GWO hyperparameter optimization, and position-grouped nested validation for reference-based puncture-test screening.

Based on these gaps, this study proposes a Reference-Based GWO-Optimized One-Class SVM framework for degradation screening of 150-kV ceramic, glass, and polymer insulators under limited puncture-test data. The specific contribution is a leakage-controlled one-class pipeline in which only Replacement observations define and optimize the reference domain, physical position is preserved during nested validation, voltage is normalized to a material-specific Replacement reference, and the final decision threshold is calibrated from unbiased held-out reference scores. Four configurations are compared to separate the contributions of representation and GWO. The proposed REF-GWO-OCSVM is not claimed to be universally superior in discrimination when an optimized RAW model reaches the same AUC; rather, its contribution is competitive separation in a compact, physically interpretable reference-relative feature space with a slightly larger anomaly-score gap and lower consensus nu. The resulting method identifies strong deviations from a replacement-derived reference domain instead of assigning definitive ground-truth damage labels.

2. RESEARCH METHOD

Data Provenance and Experimental Context

The 72 values analyzed in this study were transcribed from an undergraduate field-study dataset conducted at PT PLN (Persero) UPT Surabaya. The source report describes an online diagnostic puncture-test procedure on 150-kV transmission-line insulators using an Isolometro/Transmission Tester, with measurements collected from installed insulator strings without intentionally de-energizing the line. The source cites IEC 1211:1994 in its description of puncture testing. Because the original instrument calibration certificate and full procedural log were not available to the present authors, this article reports the provenance as documented by the source dataset and does not independently claim a more specific standards-compliance level.

The source specification identifies ceramic NGK 120 kN/25000 LBS insulators (12 units; disc diameter 254 mm; creepage distance 455 mm), glass F12/146 DC (U120B) insulators (12 units; disc diameter 255 mm; creepage distance 320 mm), and a composite-polymer ISOL VIP-A004-120SB insulator (37 sheds; 16-mm core diameter; creepage distance 2830 mm). The puncture-test tables contain 12 indexed measurement points for each material-condition combination, producing 3 materials x 2 conditions x 12 points = 72 reported observations. For ceramic and glass, these indices correspond to the 12 string units; for polymer, the source table reports 12 indexed test points even though the specification lists 37 sheds, so the indices are treated only as source-reported measurement positions. The source does not report voltage-measurement uncertainty, repeat measurements at each position, or a separate repeatability study. Accordingly, each reported position value is treated as one observation, and this missing uncertainty information is explicitly considered when interpreting generalizability.

2.1. Dataset Organization and Reference-Only Learning Strategy

The experimental dataset contained three insulator materials ceramic, glass, and polymer under two conditions, Existing and Replacement. Each material-condition combination consisted of 12 physical positions. Consequently, the complete dataset contained:

$$N = 3 \times 2 \times 12 = 72 \tag{1}$$

The proposed OCSVM was developed exclusively from the Replacement observations:

$$N_{\text{ref}} = 3 \times 12 = 36 \tag{2}$$

whereas the number of Existing observations reserved for screening was:

$$N_{\text{exist}} = 36 \tag{3}$$

This separation follows the one-class learning principle in which only observations representing the target or reference distribution are required for model construction [14] [15]. Importantly, the 36 Existing observations were not supplied to either OCSVM training or GWO hyperparameter optimization. For each physical disc position p, with p ∈ {1, 2, …, 12}, a dimensionless spatial coordinate was defined as:

$$s_p = \frac{p-1}{P-1}, \quad P = 12 \tag{4}$$

which gives:

$$0 \le s_p \le 1 \tag{5}$$

The position variable was retained because electrical stress and voltage distribution along an insulator string can depend on spatial location [2] [3] [4]. The puncture-test dataset also provides a calculated leakage-current quantity according to the material-dependent relationship:

$$I = \frac{K}{V} \tag{6}$$

Because I is therefore a deterministic transformation of V in the available dataset, it was exported only for data auditing and was not treated as an additional independent OCSVM feature. Here, K is the material-specific scaling constant used by the source calculation (271 for ceramic, 221 for glass, and 199 for polymer). The available records do not contain a separate independently measured leakage-current channel; therefore, excluding I does not discard an independent electrical measurement, but only avoids duplicating voltage information through a deterministic transformation.

For the RAW representation, the voltage scaling parameters were calculated exclusively from the active training subset:

$$V_{\mathrm{norm},i} = \frac{V_i - V_{\min}^{(\mathrm{tr})}}{V_{\max}^{(\mathrm{tr})} - V_{\min}^{(\mathrm{tr})}} \tag{7}$$

The RAW feature vector was then:

$$\mathrm{x}_i^{\mathrm{RAW}} = \left[s_i,\, V_{\mathrm{norm},i},\, M_c,\, M_g,\, M_p\right] \tag{8}$$

where Mc, Mg, and Mp are one-hot indicators for ceramic, glass, and polymer.

For the proposed reference-based representation, the material-specific Replacement mean was estimated using training observations only:

$$\bar{V}_m^{(\mathrm{tr})} = \frac{1}{n_m^{(\mathrm{tr})}} \sum_{i \in \mathcal{T}_m} V_i \tag{9}$$

where Tm denotes the currently available training-reference observations belonging to material m. Each voltage was converted into a material-relative ratio,

$$q_i = \frac{V_i}{\bar{V}_{m(i)}^{(\mathrm{tr})}} \tag{10}$$

and the proposed feature vector became:

$$\mathrm{x}_i^{\mathrm{REF}} = [s_i,\, q_i] \tag{11}$$

Thus, material identity enters the representation through the reference scale in Eq. (9), rather than through three categorical dimensions. The resulting two-dimensional domain allows the OCSVM to learn the relationship between physical position and deviation from the material-specific Replacement reference.

2.2. One-Class Support Vector Machine and Anomaly Score

The proposed framework uses the OCSVM formulation introduced for estimating the support of an unlabeled reference distribution [14]. For a reference training set {xi}i=1n, the primal OCSVM problem can be expressed as:

$$\min_{\mathrm{w},\xi,\rho} \frac{1}{2} \| \mathrm{w} \|^2 + \frac{1}{\nu n} \sum_{i=1}^{n} \xi_i - \rho \tag{12}$$

subject to:

$$\mathrm{w}^T \Phi(\mathrm{x}_i) \geq \rho - \xi_i \tag{13}$$

and:

$$\xi_i \geq 0 \tag{14}$$

The parameter $0 < \nu \leq 1$ controls the one-class boundary by governing the trade-off between reference acceptance and boundary complexity [14]. The C++ implementation directly solves the corresponding dual problem,

$$\min_{\alpha} \frac{1}{2} \alpha^T \mathrm{K} \alpha \tag{15}$$

subject to:

$$0 \leq \alpha_i \leq \frac{1}{\nu n} \tag{16}$$

and:

$$\sum_{i=1}^{n} \alpha_i = 1 \tag{17}$$

An RBF kernel was employed:

$$K(\mathrm{x}_i, \mathrm{x}_j) = \exp\left[-\gamma \| \mathrm{x}_i - \mathrm{x}_j \|^2\right] \tag{18}$$

where $\gamma > 0$ controls the local sensitivity of the reference boundary. The dual optimization was implemented internally using pairwise SMO-like updates. If

$$g_i = (\mathrm{K}\alpha)_i \tag{19}$$

the algorithm selects an index i that can receive additional coefficient mass and an index j that can release coefficient mass. The local curvature is

$$\eta = K_{ii} + K_{jj} - 2K_{ij} \tag{20}$$

The unconstrained pairwise step is:

$$\Delta_0 = \frac{g_j - g_i}{\eta} \tag{21}$$

and the feasible update used by the program is

$$\Delta = \min\left[\Delta_0,\, U - \alpha_i,\, \alpha_j\right], \quad U = \frac{1}{\nu n} \tag{22}$$

The two coefficients are updated as

$$\alpha_i \leftarrow \alpha_i + \Delta \tag{23}$$

$$\alpha_j \leftarrow \alpha_j - \Delta \tag{24}$$

which preserves Eq. (17) throughout the optimization. Pairwise optimization of the one-class dual is consistent with the sequential optimization strategy used in the original OCSVM formulation [14].

The resulting decision function is

$$f(\mathrm{x}) = \sum_{i=1}^{n} \alpha_i\, K(\mathrm{x}_i, \mathrm{x}) - \rho \tag{25}$$

Accordingly, $f(\mathrm{x}) \geq 0$ indicates that the observation lies on the reference-like side of the canonical OCSVM boundary, whereas $f(\mathrm{x}) < 0$ indicates deviation outside that boundary. Because ρ varies among nested models, the raw decision function was transformed into a normalized anomaly score:

$$A(\mathrm{x}) = -\frac{f(\mathrm{x})}{|\rho|} \tag{26}$$

Consequently,

$$A(\mathrm{x}) > 0 \tag{27}$$

represents a canonical one-class rejection, while increasingly positive values indicate progressively stronger deviations from the learned reference domain.

2.3. Reference-Only GWO Hyperparameter Optimization

The OCSVM boundary can be highly sensitive to its parameterization, particularly under limited reference data [15] [16]. Therefore, GWO was employed to optimize ν and γ. The search intervals implemented in the program were

$$0.03 \leq \nu \leq 0.30 \tag{28}$$

and

$$-8 \leq \log_2 \gamma \leq 6 \tag{29}$$

equivalent to

$$2^{-8} \leq \gamma \leq 2^{6} \tag{30}$$

The important methodological restriction is that GWO fitness was calculated only from held-out Replacement/reference observations. Existing measurements were not used to make the optimized boundary more favorable to the final screening results.

Five reference-only components were calculated during inner validation. The held-out reference rejection rate was

$$R_{\mathrm{ref}} = \frac{N_{\mathrm{ref,rejected}}}{N_{\mathrm{ref,val}}} \tag{31}$$

The mean positive anomaly deficit was

$$D_{\mathrm{ref}} = \frac{1}{N_{\mathrm{ref,val}}} \sum_{i} \max\left[0, A_i\right] \tag{32}$$

The worst-material reference rejection rate was

$$R_{\mathrm{worst}} = \max_{m} R_{\mathrm{ref},m} \tag{33}$$

To prevent the optimizer from minimizing rejection merely by generating an excessively broad reference domain, a deterministic grid was evaluated over the reference feature space. The envelope compactness term was defined as

$$C_{\mathrm{env}} = \frac{N_{\mathrm{grid,accepted}}}{N_{\mathrm{grid,total}}} \tag{34}$$

Thus, a larger accepted region receives a larger penalty. This follows the general one-class principle that optimization should consider not only target acceptance but also the volume of the accepted domain when actual outlier examples are unavailable [15]. Boundary complexity was represented by the support-vector fraction,

$$SV_{\text{frac}} = \frac{N_{\text{SV}}}{N_{\text{train}}} \tag{35}$$

The final reference-only GWO fitness implemented in the program was

$$F_{\text{ref}} = \begin{array}{l} 0.35R_{\text{ref}} + 0.15D_{\text{ref}} + 0.10R_{\text{worst}} \\ +0.30C_{\text{env}} + 0.10SV_{\text{frac}} \end{array} \tag{36}$$

and was minimized. GWO updates candidate hyperparameters using the alpha, beta, and delta solutions according to the original formulation [17]. The exploration coefficient decreased as:

$$a(t) = 2 - \frac{2t}{T-1} \tag{37}$$

while:

$$\mathrm{A} = 2a\mathrm{r}_1 - a \tag{38}$$

and:

$$\mathrm{C} = 2\mathrm{r}_2 \tag{39}$$

For the alpha wolf,

$$\mathrm{D}_\alpha = |\mathrm{C}_1\mathrm{X}_\alpha - \mathrm{X}| \tag{40}$$

$$\mathrm{X}_1 = \mathrm{X}_\alpha - \mathrm{A}_1\mathrm{D}_\alpha \tag{41}$$

Corresponding estimates are obtained from beta and delta,

$$\mathrm{X}_2 = \mathrm{X}_\beta - \mathrm{A}_2\mathrm{D}_\beta \tag{42}$$

$$\mathrm{X}_3 = \mathrm{X}_\delta - \mathrm{A}_3\mathrm{D}_\delta \tag{43}$$

and the updated wolf position becomes

$$\mathrm{X}(t+1) = \frac{\mathrm{X}_1 + \mathrm{X}_2 + \mathrm{X}_3}{3} \tag{44}$$

For every outer partition, the implementation used 10 wolves, 15 iterations, and five independent GWO repeats. The fixed baseline candidate was explicitly introduced into the initial population, and elitism retained the best solution obtained so far. Four configurations were evaluated:

$$\begin{array}{l} \mathrm{RAW-OCSVM}, \\ \mathrm{RAW-GWO-OCSVM}, \\ \mathrm{REF-OCSVM}, \\ \mathrm{REF-GWO-OCSVM} \end{array} \tag{45}$$

For the non-optimized baselines,

$$\nu = 0.10 \tag{46}$$

with

$$\gamma_{\text{RAW}} = 0.20 \tag{47}$$

and

$$\gamma_{\text{REF}} = 0.50 \tag{48}$$

2.4. Nested Position-Wise Validation and Final Degradation Screening

Because hyperparameter tuning forms part of model development, the same validation observations should not simultaneously be used for both parameter optimization and unbiased model evaluation. Nested cross-validation addresses this problem by placing model selection inside an inner loop and preserving an independent outer assessment loop [18], while the choice of fold structure should reflect the dependency structure of the observations [19]. The outer procedure therefore used 12-fold Leave-One-Position-Out (LOPO). For outer position po, all three Replacement observations from that physical position were excluded together:

$$\mathcal{V}_{p_o} = \{R_{c,p_o}, R_{g,p_o}, R_{p,p_o}\} \tag{49}$$

The resulting outer reference-training size was

$$N_{\text{outer,tr}} = 36 - 3 = 33 \tag{50}$$

GWO was then performed exclusively within these 33 observations. The remaining 11 physical positions were partitioned into three grouped inner folds, where ceramic, glass, and polymer observations belonging to the same position always remained in the same inner fold. Therefore, the outer position was excluded from both model fitting and hyperparameter optimization. After the optimal (ν, γ) pair had been determined for an outer fold, the OCSVM was retrained on its 33 reference observations and applied to both: (1) the three held-out Replacement observations; and (2) the three Existing observations at the same position. The second group was used only after model selection and therefore had no influence on the selected OCSVM boundary. Completion of the 12 outer folds produced

$$N_{\text{outer,ref}} = 36 \tag{51}$$

unbiased reference anomaly scores and

$$N_{\text{outer,exist}} = 36 \tag{52}$$

corresponding Existing-condition screening scores. The final hyperparameters were obtained from the 12 independently optimized outer models. For ν,

$$\nu_{\text{final}} = \text{median}(\nu_1, \dots, \nu_{12}) \tag{53}$$

whereas γ was aggregated in logarithmic space,

$$\gamma_{\text{final}} = 2^{\text{median}[\log_2\gamma_1,\dots,\log_2\gamma_{12}]} \tag{54}$$

The final REF-GWO-OCSVM was then trained using all 36 Replacement observations. A conservative screening threshold was derived exclusively from the unbiased outer-reference scores:

$$T_{95}^{*} = P_{95}\left(A_1^{(\text{ref})}, \dots, A_{36}^{(\text{ref})}\right) \tag{55}$$

The implementation additionally prevents the empirical threshold from becoming less restrictive than the canonical OCSVM boundary:

$$T_{95} = \max(0, T_{95}^{*}) \tag{56}$$

An Existing observation was finally flagged when

$$Flag_i = \begin{cases} 1, & A_i^{(\text{exist})} > T_{95}, \\ 0, & A_i^{(\text{exist})} \le T_{95}. \end{cases} \tag{57}$$

Thus, a calibrated flag denotes an Existing observation whose deviation from the Replacement-derived reference domain exceeds the 95th-percentile uncertainty observed for held-out references. It is not interpreted as a verified ground-truth failure label. For post-hoc evaluation only, the relative ranking of Existing and Replacement anomaly scores was summarized by the area under the ROC curve [20]. Because the implementation calculates AUC pairwise, it can be expressed as

The 95th percentile was selected as a conservative empirical operating point rather than as a physical failure threshold. In reference-only screening, an upper-tail percentile provides a transparent way to control the expected proportion of reference-like observations that are flagged when the future reference distribution resembles the held-out reference distribution. P95 therefore prioritizes specificity by targeting an approximately 5% upper-tail reference rejection, while accepting that some potentially abnormal Existing observations will remain unflagged. Because only 36 unbiased outer-reference scores are available, the percentile is sample-dependent; threshold sensitivity is consequently reported in the Results section.

$$AUC = \frac{W + \frac{1}{2}T}{N_E N_R} \tag{58}$$

where NE = 36 and N_R = 36, W is the number of Existing-reference score pairs satisfying AiE > AjR, and T is the number of tied pairs satisfying AiE = AjR. Crucially, Eq. (58) is not included in GWO fitness. AUC is computed only after the reference-only outer models have produced their scores. This preserves the central one-class principle of the proposed methodology: no Existing-condition information participates in reference-domain learning or hyperparameter optimization.

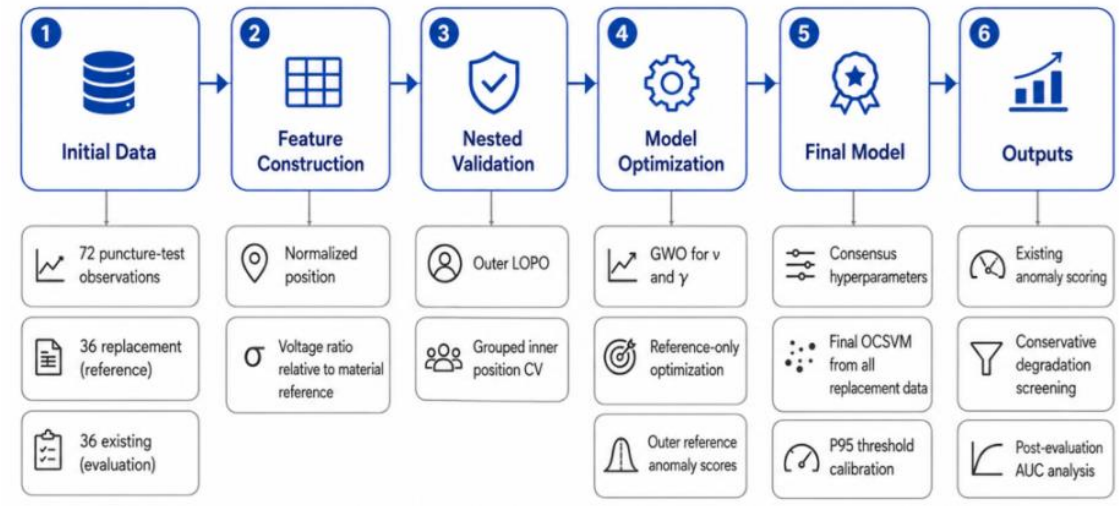


Figure 1. Model workflow

3. RESULTS AND DISCUSSION

3.1. Reference-Only One-Class Modeling and GWO Optimization

The proposed framework was evaluated using 72 puncture-test observations obtained from ceramic, glass, and polymer insulators. Only the 36 replacement-condition observations were used to construct the one-class reference domain, whereas the 36 existing-condition observations were reserved exclusively for subsequent screening and post-hoc condition-separation analysis. Model development employed nested leave-one-position-out validation. At each outer fold, the replacement observations from the same physical position across all three materials were withheld simultaneously, while GWO optimization was conducted only within the remaining reference observations using grouped three-fold position-based validation. Therefore, the existing-condition samples did not influence either OCSVM boundary construction or hyperparameter selection. Table 1 compares the four investigated configurations.

Table 1. Nested-LOPO one-class screening performance

Model	Reference rejection	Existing canonical flag	Mean reference score	Mean existing score	Score gap	AUC
RAW-OCSVM	0.2222	0.4167	-0.0222	-0.0059	0.0163	0.6956
RAW-GWO-OCSVM	0.4167	0.9167	-0.0053	0.1385	0.1438	0.8299
REF-OCSVM	0.2222	0.5000	-0.0520	0.0039	0.0560	0.7535
REF-GWO-OCSVM	0.4167	0.9167	-0.0012	0.1435	0.1447	0.8299

The baseline comparison demonstrates that reference normalization improved one-class condition separation even without metaheuristic tuning. The AUC increased from 0.6956 for RAW-OCSVM to 0.7535 for REF-OCSVM, while the mean anomaly-score gap between Existing and Replacement increased from only 0.0163 to 0.0560. This result indicates that expressing voltage relative to the material-specific replacement mean reduces part of the absolute scale difference among ceramic, glass, and polymer insulators and allows the OCSVM to represent their shared reference behavior in a substantially more compact two-dimensional domain consisting of normalized position and reference-relative voltage.

The effect of GWO was considerably stronger. REF-GWO-OCSVM increased the AUC from 0.7535 to 0.8299, corresponding to an absolute gain of approximately 0.0764. More importantly, the mean anomaly-score gap increased from 0.0560 to 0.1447, meaning that the separation between held-out Replacement and Existing observations became approximately 2.6 times larger after optimization. The mean Existing score became positive at 0.1435, whereas the mean held-out reference score remained close to the canonical boundary at -0.0012. This behavior indicates that the optimized reference envelope assigned systematically higher anomaly scores to the Existing measurements even though those measurements had not participated in GWO fitness calculation. The overall numerical results reported by the program confirm AUC values of 0.6956, 0.8299, 0.7535, and 0.8299 for RAW-OCSVM, RAW-GWO-OCSVM, REF-OCSVM, and REF-GWO-OCSVM, respectively.

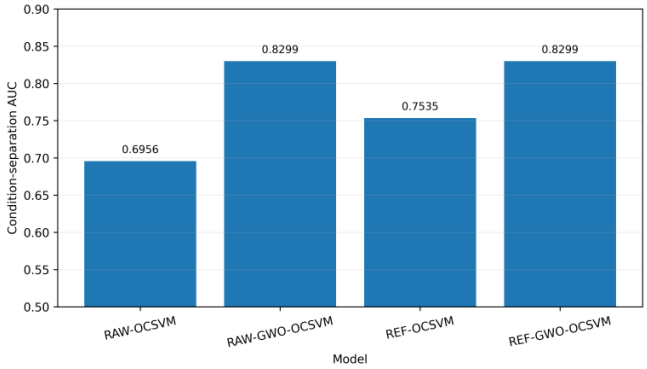


Figure 2. Condition-separation AUC obtained by the four nested-LOPO one-class configurations

A relevant finding is that REF-GWO-OCSVM and RAW-GWO-OCSVM reached the same AUC of 0.8299. Therefore, the results do not support a claim that reference normalization is universally superior in overall ranking discrimination after GWO optimization. Its value is instead seen in the physically interpretable two-dimensional feature space and the slightly larger mean anomaly-score gap, 0.1447 compared with 0.1438 for RAW-GWO-OCSVM. The reference-based model reached the same AUC with a substantially lower consensus nu, 0.1338 versus 0.2949 for the raw optimized configuration. Thus, the proposed representation should be interpreted as a compact reference-relative formulation that preserves physical meaning while remaining competitive with the optimized raw feature space.

The final proposed consensus parameters were ν=0.133754 and γ=1.825726. Neither the minimum nor maximum γ boundary was reached across the 12 outer folds, while ν reached its maximum boundary in 3 of the 12 folds. This indicates that the kernel-width search was well contained within the defined range, although some outer partitions favored a relatively aggressive upper-bound ν. Consequently, the selected hyperparameters should be regarded as fold-dependent regularization parameters rather than a universal parameterization of all insulator strings.

The GWO convergence behavior was also stable in aggregate. Across the outer folds and repeated runs, the mean best reference-only fitness of REF-GWO-OCSVM decreased from approximately 0.2812 at iteration 1 to 0.2589 at iteration 15, a reduction of about 7.95%. Considering all 60 repeat-fold endpoints (12 outer folds x 5 repeats), final reference-only fitness was 0.2589 +/- 0.0129 with a range of 0.2310-0.2832. The selected outer-fold nu values ranged from 0.0339 to 0.3000 and gamma from 1.3530 to 2.5592, showing that the limited dataset does not imply a single sharply determined hyperparameter pair. Most fitness improvement occurred during early and middle iterations, followed by progressively smaller changes. These results support the search budget while also showing that hyperparameter estimates remain fold-dependent under small-sample conditions.

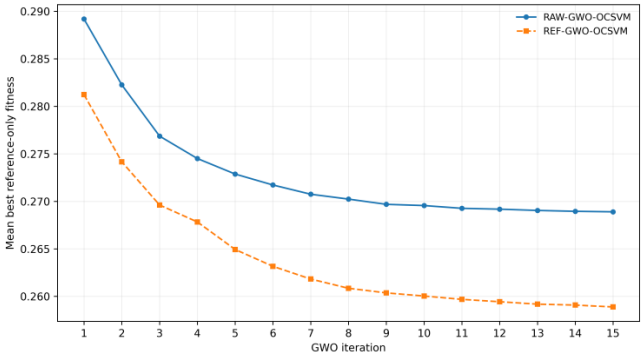


Figure 3. Mean GWO convergence for raw and reference-based OCSVM models.

An important distinction emerges between the canonical OCSVM boundary and the final calibrated screening threshold. At the canonical decision boundary A=0, REF-GWO-OCSVM rejected 15 of 36 held-out Replacement observations, corresponding to 41.67%, while flagging 33 of 36 Existing observations, or 91.67%. Thus, the optimized canonical boundary offered high sensitivity to Existing-condition deviation but was too restrictive to serve directly as a conservative screening threshold. This result justifies the subsequent P_{95}-based calibration rather than hiding the relatively high canonical reference rejection.

3.2. Calibrated Degradation Screening across Materials and Positions

The operational threshold was constructed from the distribution of unbiased outer Replacement anomaly scores. The 95th-percentile threshold was T95=0.196954. After calibration, 2 of 36 Replacement observations (5.56%) exceeded the threshold, whereas 9 of 36 Existing observations (25.00%) were flagged. This operating point intentionally trades flag coverage for lower reference false-alarm burden: compared with the canonical boundary, reference rejection decreased from 41.67% to 5.56%, while the Existing flag rate decreased from 91.67% to 25.00%. Thus, P95 is appropriate for conservative prioritization when unnecessary maintenance triggered by false alarms is undesirable. The nine Existing observations are screening candidates whose electrical response is unusually non-reference-like; they are not confirmed insulation failures.

The distinction between the reference and Existing distributions is visible in Fig. 3. Replacement scores are centered close to or below zero, whereas the Existing-score distribution is shifted toward positive values. Their overlap explains why the calibrated flag rate is substantially lower than the canonical 91.67% flag rate. Nevertheless, an AUC of 0.8299 demonstrates meaningful ranking ability across the full score distributions.

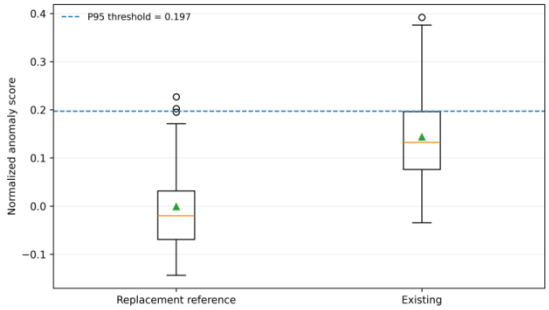


Figure 4. Distribution of nested-LOPO anomaly scores for Replacement and Existing observations. The dashed line denotes the calibrated P_{95} threshold.

Table 2. Material-wise performance of REF-GWO-OCSVM

Material	Mean reference score	Mean Existing score	Reference rejection at P_{95}	Existing flag rate at P_{95}	AUC
Ceramic	-0.0125	0.1502	0/12	3/12	0.8646
Glass	0.0100	0.1199	1/12	2/12	0.7639
Polymer	-0.0010	0.1605	1/12	4/12	0.8542

Material-specific results are summarized in Table 2. Ceramic produced the highest condition-separation AUC of 0.8646, closely followed by polymer at 0.8542, while glass produced 0.7639. Polymer had the largest mean Existing anomaly score (0.1605) and 4/12 calibrated flags; ceramic contributed 3/12 and glass 2/12. These differences should not be interpreted as proof that one material is intrinsically more degradation-prone, because the dataset contains different constructions and no independently verified damage severity. Rather, the model indicates that the Existing-to-Replacement voltage-pattern departure is more readily separated for ceramic and polymer than for glass in this specific dataset. Recent condition-monitoring studies likewise show that contamination severity, defect type, and sensing modality can substantially alter the diagnostic signature of high-voltage insulators [21]-[23]. The material-specific AUCs therefore describe separability of the present electrical patterns, not universal material reliability.

The positional results provide an even clearer interpretation of the screening behavior. At positions 1 and 2, all three Existing materials exceeded the calibrated threshold, accounting for six of the nine flags; ceramic and polymer were also flagged at position 3, and polymer contributed the remaining flag at position 10. Accordingly, eight of nine screening candidates occurred at positions 1-3. This clustering is physically plausible because voltage and electric-field distributions along insulator strings are non-uniform and are influenced by stray capacitances, hardware geometry, and the location of electrically weak units [2]-[4]. However, the source dataset does not explicitly document whether index 1 corresponds to the line end or tower end for every material, and the polymer indices are source-reported test points rather than individual sheds. Consequently, the present results support position dependence but do not justify assigning the position-1-3 cluster to a specific terminal-side stress mechanism without further field metadata.

Table 3. Position-wise calibrated screening characteristics

Position	Mean reference score	Mean Existing score	Reference flags	Existing flags
1	0.1648	0.3329	1	3
2	0.0329	0.3329	0	3
3	-0.0279	0.2267	0	2
4	-0.0305	0.1640	0	0
5	0.0076	0.1250	0	0
6	-0.1143	0.0028	0	0
7	-0.1386	-0.0130	0	0
8	-0.0384	0.0997	0	0
9	0.0782	0.1047	0	0
10	-0.0387	0.1830	0	1
11	-0.0866	0.0621	0	0
12	0.1776	0.1016	1	0

The position-dependent concentration of high scores is important because OCSVM was not trained with Existing-condition labels. In particular, the mean Existing anomaly score reached approximately 0.333 at both positions 1 and 2, substantially exceeding the P_{95}=0.197 threshold. Conversely, positions 6 and 7 showed mean Existing scores close to zero or negative and produced no calibrated flags. The screening therefore identifies localized departures from the replacement-derived reference domain rather than merely separating all Existing observations from all Replacement observations.

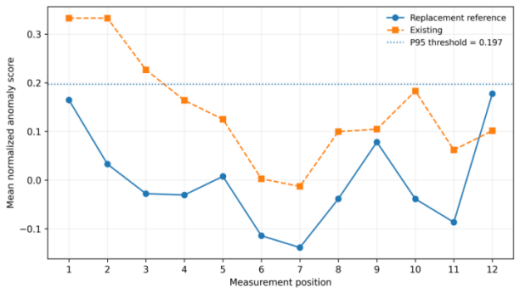


Figure 5. Position-wise mean anomaly scores for held-out Replacement and Existing observations. The dashed line denotes the calibrated P_{95} threshold.

The final OCSVM trained on all 36 Replacement observations used 8 support vectors, corresponding to 22.22% of the reference dataset, with ρ=0.455164. The support vectors were distributed across ceramic, glass, and polymer observations and included several measurements near the beginning, minimum-profile region, and terminal part of the string. This indicates that the final boundary was determined by reference observations distributed across multiple spatial regions rather than by one material alone.

Taken together, the results show that GWO considerably improves the ability of an OCSVM trained solely on replacement data to rank Existing-condition observations as increasingly non-reference-like. At the same time, the difference between the canonical flag rate of 91.67% and the calibrated P_{95} flag rate of 25.00% demonstrates why threshold calibration is essential. The former is sensitive but rejects too many valid held-out references; the latter sacrifices flag coverage to maintain a reference rejection rate of only 5.56%. Consequently, the proposed method should be interpreted as a high-specificity reference-assisted degradation screening approach, not as a supervised degradation classifier or a definitive failure-detection system.

The small sample size remains the principal uncertainty in this study. A position-block bootstrap of the 12 outer LOPO blocks (10,000 resamples, resampling all three materials together within a selected position) gave an AUC bootstrap mean of 0.8342, SD 0.0676, and a 95% interval of 0.6890-0.9537 around the observed AUC of 0.8299. For the P95 threshold, leave-one-position sensitivity analysis using the remaining 33 held-out reference scores produced a mean threshold of 0.1929 +/- 0.0120 and a range of 0.1640-0.1980 around the full-sample value 0.1970. These intervals are not evidence of population-level precision; instead, they quantify how strongly the reported operating characteristics can vary when a small number of position blocks changes.

Table 4. Small-sample stability and sensitivity analysis

Quantity	Point estimate	Variation / sensitivity
Condition-separation AUC	0.8299	Position-block bootstrap: mean 0.8342, SD 0.0676, 95% CI 0.6890-0.9537
P95 threshold	0.1970	Leave-one-position: 0.1929 +/- 0.0120, range 0.1640-0.1980
Final GWO reference fitness	0.2589	60 repeat-fold endpoints: 0.2589 +/- 0.0129, range 0.2310-0.2832

4. CONCLUSION

This study developed a reference-based GWO-optimized One-Class SVM for degradation screening of 150-kV ceramic, glass, and polymer insulators using limited puncture-test data. The proposed strategy trained and optimized the one-class reference domain exclusively from 36 Replacement observations, while 36 Existing observations were excluded from training and hyperparameter tuning. Nested leave-one-position-out evaluation further ensured that the held-out physical position was excluded from both reference-model construction and GWO optimization.

The proposed REF-GWO-OCSVM achieved a condition-separation AUC of 0.8299, compared with 0.7535 for REF-OCSVM and 0.6956 for RAW-OCSVM. GWO increased the mean reference-to-Existing anomaly-score gap from 0.0560 to 0.1447. RAW-GWO-OCSVM achieved the same AUC of 0.8299, indicating that reference normalization should not be claimed to universally outperform optimized raw features; rather, it provides comparable separation through a more compact and physically interpretable reference-relative representation.

The canonical optimized OCSVM boundary was relatively restrictive, rejecting 41.67% of held-out references while flagging 91.67% of Existing observations. Calibration using the unbiased outer-reference P_{95} score produced a substantially more conservative threshold of 0.196954, reducing reference rejection to 2/36 (5.56%) while identifying 9/36 (25.00%) Existing observations as strong screening candidates. Material-specific AUCs were 0.8646 for ceramic, 0.7639 for glass, and 0.8542 for polymer, and most calibrated deviations were concentrated at physical positions 1–3.

These results support reference-only one-class learning as a conservative prioritization mechanism when explicit degradation labels are unavailable, but the findings should be regarded as preliminary. The 72 observations represent only 36 reference and 36 Existing measurements from one source dataset, no position-level repeated measurements or instrument uncertainty were available, and the bootstrap analysis shows non-negligible uncertainty in AUC. The flagged observations therefore represent departures from a replacement-derived electrical reference, not independently confirmed insulation failures. Validation on independent and substantially larger datasets should include multiple strings per material, repeated measurements, documented measurement uncertainty, explicit position orientation, and independently verified physical-condition labels before the framework is interpreted as a field diagnostic classifier.

Funding and Competing Interests
This research received no external funding. The authors declare that there are no competing financial or personal interests that could have influenced the work reported in this paper.

5. REFERENCES

[1] D. R. de Mello, J. A. P. Rodrigues, J. A. D'Affonseca Cardoso, and F. de A. A. F. Vieira, "The influence of a compact test circuit in the puncture test," in Proc. 2010 *Int. Conf. High Voltage Engineering and Application (ICHVE)*, 2010, doi: 10.1109/ICHVE.2010.5640707.

[2] N. Tonmitr, K. Tonmitr, and E. Kaneko, "The effect of controlling stray and disc capacitance of ceramic string insulator in the case of clean and contaminated conditions," *Procedia Computer Science*, vol. 86, pp. 333–336, 2016, doi: 10.1016/j.procs.2016.05.090.

[3] L. Zheng, P. Yin, J. Li, H. Liu, T. Li, and H. Luo, "Research on diagnostic methods for zero-value insulators in 110 kV transmission lines based on spatial distribution characteristics of electric fields," *Energies*, vol. 18, no. 6, Art. no. 1534, 2025, doi: 10.3390/en18061534.

[4] J. Lu, J. Hu, Z. Fang, X. Qiao, and Z. Zhang, "Electric field distribution and AC breakdown characteristics of polluted novel lightning protection insulator under icing conditions," *Energies*, vol. 14, no. 22, Art. no. 7493, 2021, doi: 10.3390/en14227493.

[5] C. A. Badachi and P. Dixit, "Prediction of pollution flashover voltages of ceramic string insulators under uniform and non-uniform pollution conditions," *Journal of Electrical Systems and Information Technology*, vol. 3, no. 2, pp. 270–281, 2016, doi: 10.1016/j.jesit.2016.01.003.

[6] A. A. Salem et al., "Influence of contamination distribution in characterizing the flashover phenomenon on outdoor insulator," *Ain Shams Engineering Journal*, vol. 14, no. 12, Art. no. 102249, 2023, doi: 10.1016/j.asej.2023.102249.

[7] A. A. Salem et al., "Leakage current characteristics in estimating insulator reliability: Experimental investigation and analysis," *Scientific Reports*, vol. 12, Art. no. 14974, 2022, doi: 10.1038/s41598-022-17792-x.

[8] A. A. Salem et al., "Monitoring porcelain insulator condition based on leakage current characteristics," *Materials*, vol. 15, no. 18, Art. no. 6370, 2022, doi: 10.3390/ma15186370.

[9] A. A. Salem, K. Y. Lau, Z. Abdul-Malek, N. Mohammed, A. M. Al-Shaalan, A. A. Al-Shamma'a, and H. M. H. Farh, "Polymeric insulator conditions estimation by using leakage current characteristics based on simulation and experimental investigation," *Polymers*, vol. 14, no. 4, Art. no. 737, 2022, doi: 10.3390/polym14040737.

[10] U. A. Khan et al., "Experimental validation of machine learning for contamination classification of polluted high voltage insulators using leakage current," *Scientific Reports*, vol. 15, Art. no. 13246, 2025, doi: 10.1038/s41598-025-97646-4.

[11] Arshad, J. Ahmad, A. Tahir, B. G. Stewart, and A. Nekahi, "Forecasting flashover parameters of polymeric insulators under contaminated conditions using the machine learning technique," *Energies*, vol. 13, no. 15, Art. no. 3889, 2020, doi: 10.3390/en13153889.

[12] L. Song, Q. Liang, H. Chen, H. Hu, Y. Luo, and Y. Luo, "A new approach to optimize SVM for insulator state identification based on improved PSO algorithm," *Sensors*, vol. 23, no. 1, Art. no. 272, 2023, doi: 10.3390/s23010272.

[13] A. S. Haiba and A. E. Gad, "Artificial neural network analysis for classification of defected high voltage ceramic insulators," *Scientific Reports*, vol. 14, Art. no. 1513, 2024, doi: 10.1038/s41598-024-51860-8.

[14] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001, doi: 10.1162/089976601750264965.

[15] D. M. J. Tax and R. P. W. Duin, "Uniform object generation for optimizing one-class classifiers," *Journal of Machine Learning Research*, vol. 2, pp. 155–173, 2001.

[16] A. Thomas, V. Feuillard, and A. Gramfort, "Calibration of One-Class SVM for MV set estimation," in *Proc. IEEE Int. Conf. Data Science and Advanced Analytics (DSAA)*, 2015, doi: 10.1109/DSAA.2015.7344789.

[17] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[18] S. Varma and R. Simon, "Bias in error estimation when using cross-validation for model selection," *BMC Bioinformatics*, vol. 7, Art. no. 91, 2006, doi: 10.1186/1471-2105-7-91.

[19] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Surveys*, vol. 4, pp. 40–79, 2010, doi: 10.1214/09-SS054..

[20] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.

[21] A. A. Salem et al., "Classification of RTV-coated porcelain insulator condition under different profiles and

levels of pollution," *Scientific Reports*, vol. 14, Art. no. 22759, 2024, doi: 10.1038/s41598-024-73520-7.
[22] M. Monemi, S. M. Shahrtash, and M. Kalantar, "Failure type and failure level detection of insulators according to monitored leakage current," *Heliyon*, vol. 10, no. 14, Art. no. e34143, 2024, doi: 10.1016/j.heliyon.2024.e34143.
[23] Z.-Y. Xu and X. Tang, "Transmission line insulator defect detection algorithm based on MAP-YOLOv8," *Scientific Reports*, vol. 15, Art. no. 10288, 2025, doi: 10.1038/s41598-025-92445-3.
[24] L. Das, B. Gjorgiev, and G. Sansavini, "An improved anomaly detection model for automated inspection of power line insulators," *Engineering Applications of Artificial Intelligence*, vol. 158, Part B, Art. no. 111431, 2025, doi: 10.1016/j.engappai.2025.111431.
[25] Z. Xie, S. Wu, T. Yang, N. Zhan, et al., "Unsupervised visual anomaly detection for high voltage insulators using context aware multi scale memory networks," *Discover Applied Sciences*, vol. 8, Art. no. 463, 2026, doi: 10.1007/s42452-026-08435-1.
[26] Z. Que and C.-J. Lin, "One-Class SVM Probabilistic Outputs," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 36, no. 4, pp. 6244-6256, 2025, doi: 10.1109/TNNLS.2024.3395148.
[27] G. D. Prenata, "Klasifikasi Keandalan Sistim Distribusi Tenaga Listrik di PT. PLN (Persero) UP3 Surabaya Selatan Menggunakan Support Vector Machine (SVM)," *Jurnal Teknik Elektro*, vol. 16, no. 2, pp. 62-70, 2023, doi: 10.9744/jte.16.2.62-70.
[28] G. D. Prenata, "Predicting Breakdown Voltage of Transformer Oil under Copper/Iron Contamination: A Comparative Study of Gradient vs Metaheuristic Training," *ELKHA: Jurnal Teknik Elektro*, vol. 18, no. 1, pp. 65-73, 2026, doi: 10.26418/elkha.v18i1.101026.
[29] G. D. Prenata and A. Ridho'i, "Breakdown Voltage Prediction of Transformer Oil under Copper/Iron Contamination Using SVR and KNN Regression: A Baseline Study on Small Experimental Data," *El Sains: Jurnal Elektro*, vol. 8, no. 1, pp. 81-87, 2026, doi: 10.30996/elsains.v8i1.133785.
[30] G. D. Prenata and A. Ridho'i, "Thermal Behavior Clustering of High-Voltage Electrical Equipment Using K-Means and Fuzzy C-Means," *Elkom: Jurnal Elektronika dan Komputer*, vol. 19, no. 1, pp. 229-238, 2026, doi: 10.51903/elkom.v19i1.3700.