



PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA



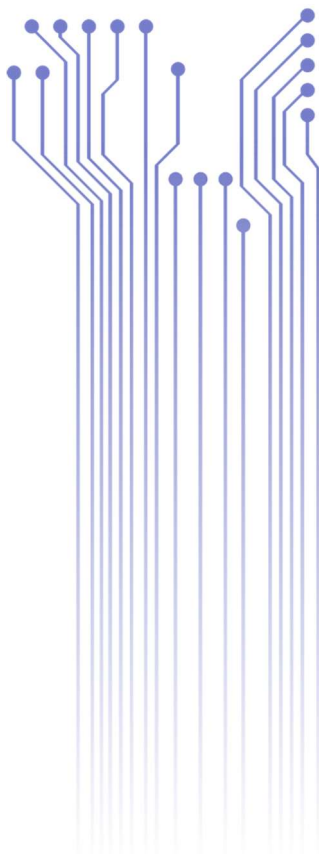
Universidade Federal de Sergipe



Johnny Herbert Paixão de Oliveira

DISSERTAÇÃO

**REPRESENTAÇÃO E CLASSIFICAÇÃO DE PERFIS  
TÉRMICOS DE ISOLADORES POLIMÉRICOS PARA  
DIAGNÓSTICO DE POLUIÇÃO SUPERFICIAL**



São Cristóvão  
2026

JOHNNY HERBERT PAIXÃO DE OLIVEIRA

REPRESENTAÇÃO E CLASSIFICAÇÃO DE PERFIS TÉRMICOS DE ISOLADORES POLIMÉRICOS  
PARA DIAGNÓSTICO DE POLUIÇÃO SUPERFICIAL

*Dissertação submetida à Coordenação do  
Programa de Pós-graduação em Engenharia  
Elétrica da Universidade Federal de Sergipe  
como parte dos requisitos necessários para a  
obtenção do grau de Mestre em Ciências no  
Domínio da Engenharia Elétrica.*

Orientador:

Professor Tarso Vilela Ferreira, D. Sc.

Coorientador:

Professor George Victor Rocha Xavier, D. Sc.

São Cristóvão  
2026

JOHNNY HERBERT PAIXÃO DE OLIVEIRA

REPRESENTAÇÃO E CLASSIFICAÇÃO DE PERFIS TÉRMICOS  
DE ISOLADORES POLIMÉRICOS PARA DIAGNÓSTICO DE  
POLUIÇÃO SUPERFICIAL

*Dissertação submetida à Coordenação do  
Programa de Pós-graduação em Engenharia  
Elétrica da Universidade Federal de Sergipe  
como parte dos requisitos necessários para a  
obtenção do grau de Mestre em Ciências no  
Domínio da Engenharia Elétrica.*

Aprovado em 24/07/2026

**Professor Ivan José da Silva Lopes, D. Sc.**  
Universidade Federal de Minas Gerais  
Avaliador

**Professor Jugurta Rosa Montalvão Filho, D. Sc.**  
Universidade Federal de Sergipe  
Avaliador

**Professor Wamberto José Lira de Queiroz, D. Sc.**  
Universidade Federal de Campina Grande  
Avaliador

**Professor Tarso Vilela Ferreira, D. Sc.**  
Universidade Federal de Sergipe  
Orientador

**Professor George Victor Rocha Xavier, D. Sc.**  
Universidade Federal de Sergipe  
Coorientador

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL  
UNIVERSIDADE FEDERAL DE SERGIPE**

O48r      Oliveira, Johnny Herbert Paixão de  
Representação e classificação de perfis térmicos de isoladores  
poliméricos para diagnóstico de poluição superficial / Johnny  
Herbert Paixão de Oliveira ; orientador Tarso Vilela Ferreira - São  
Cristóvão, 2026.  
157 f. : il.

Dissertação (mestrado em Engenharia Elétrica) – Universidade  
Federal de Sergipe, 2026.

1. Isoladores e isolamentos elétricos. 2. Termografia. 3.  
Modelos Ocultos de Markov. 4. Aprendizado do computador. I.  
Ferreira, Tarso Vilela orient. II. Título.

CDU 621.3





**UNIVERSIDADE FEDERAL DE SERGIPE  
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA  
COORDENAÇÃO DE PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA-PROEE**

**TERMO DE APROVAÇÃO**

**“Representação e Classificação de Perfis Térmicos de Isoladores  
Poliméricos para Diagnóstico de Poluição Superficial”**

Discente:

**Johnny Herbert Paixão de Oliveira**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Sergipe, como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica.

Aprovada pela banca examinadora composta por:

Documento assinado digitalmente  
 **JUGURTA ROSA MONTALVAO FILHO**  
Data: 24/07/2026 11:26:43-0300  
Verifique em <https://validar.iti.gov.br>

**Prof. Dr. Jugurta Rosa Montalvão Filho (PROEE/UFS)**  
**Presidente**

Documento assinado digitalmente  
 **IVAN JOSE DA SILVA LOPES**  
Data: 24/07/2026 11:37:00-0300  
Verifique em <https://validar.iti.gov.br>

**Prof. Dr. Ivan José da Silva Lopes (PPGEE/UFGM)**  
**Examinador Externo**

Documento assinado digitalmente  
 **WAMBERTO JOSE LIRA DE QUEIROZ**  
Data: 24/07/2026 17:49:31-0300  
Verifique em <https://validar.iti.gov.br>

**Prof. Dr. Wamberto Lira Queiroz (PPGEE/UFCG)**  
**Examinador Externo**

Documento assinado digitalmente  
 **JOHNNY HERBERT PAIXAO DE OLIVEIRA**  
Data: 30/07/2026 07:56:29-0300  
Verifique em <https://validar.iti.gov.br>

**Johnny Herbert Paixão de Oliveira**  
**Discente**

Cidade Universitária “Prof. José Aloísio de Campos”, 24 de julho de 2026.

Dedico este trabalho aos meus pais, exemplos de amor, dedicação e força, que sempre estiveram ao meu lado nos momentos mais desafiadores. Como se tudo isso não bastasse, foram eles que me ensinaram o verdadeiro significado de persistência e conquista.

## AGRADECIMENTOS

Primeiramente, agradeço a Deus, por me dar força, saúde e perseverança para enfrentar os desafios que surgiram ao longo deste percurso.

Aos meus pais, Joelma da Paixão e Sérgio Ricardo de Oliveira, pela base sólida que me proporcionaram, pelo apoio incondicional, pelo incentivo em todos os momentos e, acima de tudo, por acreditarem no meu potencial, mesmo quando eu duvidava de mim mesmo. Vocês são minha maior inspiração e este trabalho é também um reflexo do esforço e dedicação que sempre me ensinaram a ter, assim como minha irmã, Nicolay Rayanne Paixão de Oliveira, cuja presença, apoio e carinho foram fundamentais ao longo dessa caminhada.

A minha companheira, Rute Santos Porto Lima, pelo carinho, paciência e apoio inestimável ao longo desta jornada. Obrigado por estar ao meu lado nos momentos de estresse, pelas palavras de encorajamento e por compreender minhas ausências durante os períodos mais intensos de estudo. Seu suporte foi essencial para que eu chegasse até aqui com mais leveza e motivação.

Aos meus amigos e orientadores, o Prof. Dr. Tarso Vilela Ferreira e o Prof. Dr. George Victor Rocha Xavier, pela paciência, dedicação e ensinamentos fundamentais durante todo o desenvolvimento deste trabalho. Sua orientação não foi apenas acadêmica, mas também um incentivo para que eu buscasse sempre a excelência e o pensamento crítico.

Aos meus colegas e amigos do Laboratório de Automação Controle e Simulação (LACS), que estiveram ao meu lado nesta jornada universitária, compartilhando desafios, conquistas e muitas horas de estudo e dedicação. Em especial, agradeço a Bruno Vinicius Silveira Araujo, pelo apoio nos momentos difíceis, pelas trocas de conhecimento e pelas palavras de encorajamento. A parceria e amizade de vocês foram fundamentais para tornar essa caminhada mais leve e enriquecedora.

Por fim, externo meus agradecimentos a CAPES, pelas bolsas de estudos que recebi durante o mestrado.

*“Há uma força motriz mais poderosa que o vapor,  
a eletricidade e a energia atômica: a vontade”*

Albert Einstein

## RESUMO

O monitoramento de isoladores em sistemas elétricos de potência é essencial para reduzir o risco de falhas associadas à poluição superficial. Em ambientes litorâneos, a deposição de sais e a presença de umidade podem intensificar a corrente de fuga superficial e provocar aquecimentos localizados capazes de conduzir o equipamento à falha e causar interrupções no fornecimento de energia. Nesse contexto, a termografia infravermelha constitui uma técnica não invasiva promissora, pois permite avaliar a condição térmica do equipamento sem contato direto e sem interrupção da operação. Nesta dissertação, propõe-se uma metodologia para classificar perfis térmicos de isoladores poliméricos, visando distinguir condições limpas e poluídas. Para isso, perfis térmicos unidimensionais são extraídos de termografias, corrigidos em relação à temperatura ambiente, segmentados em janelas sucessivas e descritos por atributos relacionados ao nível térmico, à dispersão de temperatura: variância e desvio padrão, ao contraste, à forma da distribuição: assimetria e curtose, à variação espacial e à dependência espacial. São avaliadas três formas de representação da informação ao longo do perfil: agregada, sequencial e simbólica. A representação agregada resume os atributos obtidos nas janelas, enquanto as representações sequencial e simbólica preservam explicitamente a ordem espacial dessas janelas. Para a classificação, são avaliados máquina de vetores de suporte, floresta aleatória, rede neural recorrente de memória de longo e curto prazo, modelo oculto de Markov com misturas gaussianas e modelo oculto de Markov discreto. A base de dados é composta por 344 perfis térmicos, distribuídos em 69 grupos de captura, e contempla isoladores limpos e isoladores submetidos à poluição artificial e natural. A seleção dos atributos, o ajuste dos modelos e a estimativa inicial do desempenho são realizados por validação aninhada, com separação por grupos de captura, evitando que amostras provenientes de uma mesma captura sejam distribuídas entre treinamento e teste. Os resultados indicam que todos os modelos avaliados apresentam capacidade de distinguir perfis limpos e poluídos. Posteriormente, realiza-se uma análise de robustez, na qual os melhores modelos de cada forma de representação são avaliados em diferentes

repetições da divisão entre treinamento e teste, mantendo-se fixos os atributos selecionados, os hiperparâmetros e as estruturas definidas durante a validação aninhada. Nessa análise, a máquina de vetores de suporte aplicada à representação agregada apresenta o melhor desempenho, com acurácia balanceada de  $0,923 \pm 0,058$  e métrica multicritério de  $0,897 \pm 0,067$ . Conclui-se que a segmentação espacial do perfil térmico fornece informações relevantes para a identificação da poluição superficial e constitui uma abordagem promissora para o diagnóstico de isoladores poliméricos.

**Palavras-chave:** Isoladores poliméricos; Termografia; Perfis térmicos; Poluição superficial; Modelos Ocultos de Markov; Aprendizado de máquina.

# ABSTRACT

Monitoring insulators in electric power systems is essential for reducing the risk of failures associated with surface pollution. In coastal environments, salt deposition and the presence of moisture can intensify surface leakage current and cause localized heating capable of leading the equipment to failure and causing interruptions in the power supply. In this context, infrared thermography is a promising non-invasive technique because it enables the thermal condition of the equipment to be assessed without direct contact or interruption of operation. This dissertation proposes a methodology for classifying the thermal profiles of polymeric insulators, aiming to distinguish between clean and polluted conditions. To this end, one-dimensional thermal profiles are extracted from thermograms, corrected relative to the ambient temperature, segmented into successive windows, and described using features related to thermal level, temperature dispersion, including variance and standard deviation, contrast, distribution shape, including skewness and kurtosis, spatial variation, and spatial dependence. Three forms of representing the information along the profile are evaluated: aggregated, sequential, and symbolic. The aggregated representation summarizes the features obtained from the windows, whereas the sequential and symbolic representations explicitly preserve the spatial order of these windows. For classification, a support vector machine, a random forest, a long short-term memory recurrent neural network, a hidden Markov model with Gaussian mixtures, and a discrete hidden Markov model are evaluated. The dataset comprises 344 thermal profiles distributed across 69 acquisition groups and includes clean insulators as well as insulators subjected to artificial and natural pollution. Feature selection, model tuning, and the initial performance estimation are conducted through nested validation with separation by acquisition groups, thereby preventing samples originating from the same acquisition from being distributed between the training and test sets. The results indicate that all evaluated models are capable of distinguishing between clean and polluted profiles. Subsequently, a robustness analysis is conducted in which the best models for each representation are evaluated over different repetitions of the training–test split, while keeping fixed the selected features, hyperparameters, and structures defined during nested validation. In this analysis, the support vector machine applied to the aggregated representation achieves the best performance, with a balanced

accuracy of  $0.923 \pm 0.058$  and a multicriteria metric of  $0.897 \pm 0.067$ . It is concluded that the spatial segmentation of the thermal profile provides relevant information for identifying surface pollution and constitutes a promising approach for the diagnosis of polymeric insulators.

**Keywords:** Polymeric insulators; Thermography; Thermal profiles; Surface pollution; Hidden Markov Models; Machine learning.

# LISTA DE ILUSTRAÇÕES

|                                                                                                                                                                        |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figura 1. Comparação entre distância de arco e distância de escoamento.....                                                                                            | 28  |
| Figura 2. Exemplares de isolador. (a) Porcelana; (b) Vidro; (c) Polimérico. ....                                                                                       | 29  |
| Figura 3. Esquemático da matriz de temperatura e a extração de informações térmicas na região de interesse.....                                                        | 31  |
| Figura 4. Hiperplano ótimo para padrões linearmente separáveis. ....                                                                                                   | 48  |
| Figura 5. Estrutura da Random Forest com três árvores de decisão. ....                                                                                                 | 51  |
| Figura 6. Cadeia de Markov com 3 estados e transições selecionadas. ....                                                                                               | 54  |
| Figura 7. Etapa de indução do algoritmo <i>forward</i> . ....                                                                                                          | 57  |
| Figura 8. Probabilidade <i>forward</i> baseado na estrutura de treliça. ....                                                                                           | 58  |
| Figura 9. Etapa de recursão do algoritmo <i>Backward</i> . ....                                                                                                        | 59  |
| Figura 10. Arquitetura de uma rede neural recorrente genérica. ....                                                                                                    | 63  |
| Figura 11. Camadas da arquitetura de rede LSTM. ....                                                                                                                   | 64  |
| Figura 12. Estado de célula da arquitetura LSTM. ....                                                                                                                  | 64  |
| Figura 13. Porta de esquecimento na arquitetura de rede LSTM. ....                                                                                                     | 65  |
| Figura 14. Porta de entrada e vetores candidatos na arquitetura de rede LSTM. ....                                                                                     | 67  |
| Figura 15. Novo estado da célula na arquitetura de rede LSTM. ....                                                                                                     | 68  |
| Figura 16. Saída da rede LSTM. ....                                                                                                                                    | 69  |
| Figura 17. Fluxograma metodológico geral. ....                                                                                                                         | 87  |
| Figura 18. Termovisor. ....                                                                                                                                            | 88  |
| Figura 19. Arranjo experimental. ....                                                                                                                                  | 89  |
| Figura 20. Fluxograma de extração do perfil térmico. ....                                                                                                              | 94  |
| Figura 21. Perfil térmico antes e após o procedimento de ampliação. ....                                                                                               | 97  |
| Figura 22. Processo de segmentação do perfil térmico em janelas. ....                                                                                                  | 99  |
| Figura 23. Fluxograma da validação aninhada. ....                                                                                                                      | 104 |
| Figura 24. Termografia de isoladores. (a) Limpo; (b) Poluído naturalmente. ....                                                                                        | 115 |
| Figura 25. Exemplos de perfis térmicos. ....                                                                                                                           | 117 |
| Figura 26. Comparação dos perfis térmicos médios, em termos de $\Delta T$ , ao longo da direção longitudinal do isolador para diferentes condições de superfície. .... | 118 |
| Figura 27. Segmentação do perfil térmico em janelas. ....                                                                                                              | 118 |
| Figura 28. Construção da representação sequencial dos atributos. ....                                                                                                  | 119 |
| Figura 29. Construção da representação agregada dos atributos. ....                                                                                                    | 120 |
| Figura 30. Construção da representação simbólica do perfil térmico. ....                                                                                               | 121 |
| Figura 31. Frequência de seleção de atributos nas representações sequencial e agregada. ....                                                                           | 122 |
| Figura 32. Desempenho dos modelos na avaliação por perfil térmico. ....                                                                                                | 135 |
| Figura 33. Desempenho dos modelos na avaliação por grupo de captura. ....                                                                                              | 137 |
| Figura 34. Matriz de confusão normalizada do modelo SVM na análise de robustez. ....                                                                                   | 143 |
| Figura 35. Erro médio por subtipo experimental do modelo SVM. ....                                                                                                     | 144 |

# LISTA DE TABELAS

|                                                                                                       |     |
|-------------------------------------------------------------------------------------------------------|-----|
| Tabela 1. Matriz de confusão. ....                                                                    | 70  |
| Tabela 2. Quadro sinóptico.....                                                                       | 86  |
| Tabela 3. Níveis de ESDD do experimento 1.....                                                        | 90  |
| Tabela 4. Níveis de ESDD do experimento 2.....                                                        | 91  |
| Tabela 5. Níveis de ESDD do experimento 3.....                                                        | 92  |
| Tabela 6. Níveis de ESDD no experimento 4.....                                                        | 93  |
| Tabela 7. Composição da base de dados por classe e subtipo experimental.....                          | 116 |
| Tabela 8. Configurações simbólicas selecionadas em cada repetição externa da validação aninhada. .... | 125 |
| Tabela 9. Subconjuntos e configurações selecionados por repetição externa.....                        | 129 |
| Tabela 10. Métrica multicritério por grupo de captura em cada repetição externa.....                  | 132 |
| Tabela 11. Desempenho dos modelos campeões na análise de robustez.....                                | 141 |

# LISTA DE ABREVIATURAS E SIGLAS

## **Sigla/abreviatura Descrição**

|            |                                                                           |
|------------|---------------------------------------------------------------------------|
| CNN        | Convolutional Neural Network; Rede Neural Convolucional                   |
| DTW        | Dynamic Time Warping                                                      |
| ESDD       | Equivalent Salt Deposit Density; Densidade Equivalente de Depósito de Sal |
| F1         | Medida F1 ou F1-score                                                     |
| FLIR       | Forward Looking Infrared                                                  |
| FN         | Falso negativo                                                            |
| FP         | Falso positivo                                                            |
| GMM        | Gaussian Mixture Model(s)                                                 |
| HMM        | Hidden Markov Model                                                       |
| HMM-GMM    | HMM com emissões gaussianas                                               |
| IEC        | International Electrotechnical Commission                                 |
| IQR        | Interquartile Range; Intervalo Interquartil                               |
| KNN        | k-Nearest Neighbors; k-vizinhos mais próximos                             |
| LACS       | Laboratório de Automação, Controle e Simulação                            |
| LAT        | Laboratório de Alta Tensão                                                |
| LSTM       | Long Short-Term Memory                                                    |
| MAD        | Median Absolute Deviation                                                 |
| Mask R-CNN | Mask Region-based Convolutional Neural Network                            |
| mAP50      | mean Average Precision at 50%                                             |
| MCC        | Matthews Correlation Coefficient                                          |
| MLP        | Multilayer Perceptron                                                     |
| mRMR       | Minimum Redundancy Maximum Relevance;                                     |
| MTrip-YOLO | Arquitetura MTrip-YOLO baseada em YOLO11n                                 |
| NSDD       | Non-Soluble Deposit Density                                               |

**Sigla/abreviatura Descrição**

|         |                                           |
|---------|-------------------------------------------|
| NUPEG   | Núcleo de Pesquisa em Petróleo e Gás      |
| OOB     | Out-of-bag                                |
| PCA     | Principal Component Analysis              |
| RBF     | Radial Basis Function                     |
| R-CNN   | Region-based Convolutional Neural Network |
| RF      | Random Forest                             |
| RI      | Região de Interesse                       |
| RNN     | Recurrent Neural Network                  |
| RT-DETR | Real-Time Detection Transformer           |
| SDK     | Software Development Kit                  |
| SVM     | Support Vector Machine                    |
| UFS     | Universidade Federal de Sergipe           |
| UV      | Ultravioleta                              |
| VN      | Verdadeiro negativo                       |
| VP      | Verdadeiro positivo                       |
| YOLO    | You Only Look Once                        |
| YOLO11n | Versão nano da família YOLO11             |
| YOLOv5  | Versão 5 da família YOLO                  |

# LISTA DE SÍMBOLOS

|                   |                                                                                            |
|-------------------|--------------------------------------------------------------------------------------------|
| $R$               | Região de interesse térmica analisada.                                                     |
| $N$               | Número de pixels válidos da região de interesse ou número de estados em modelos de Markov. |
| $T_i$             | Temperatura, ou intensidade térmica calibrada, do $i$ -ésimo pixel.                        |
| $i$               | Índice de pixel, amostra, estado ou observação, conforme o contexto.                       |
| $x$               | Posição espacial, vetor de atributos ou vetor de entrada, conforme o contexto.             |
| $x_i$             | Vetor de atributos da $i$ -ésima amostra.                                                  |
| $x_{ij}$          | $j$ -ésimo atributo extraído da $i$ -ésima amostra.                                        |
| $p$               | Número total de atributos.                                                                 |
| $y_i$             | Rótulo verdadeiro associado à $i$ -ésima amostra.                                          |
| $\hat{y}$         | Classe prevista pelo modelo.                                                               |
| $f(\cdot)$        | Função de classificação ou função discriminante.                                           |
| $\theta$          | Conjunto de parâmetros ajustáveis do modelo.                                               |
| $\theta^*$        | Conjunto ótimo de parâmetros obtido no treinamento.                                        |
| $L(\cdot)$        | Função de perda utilizada para medir o erro de classificação.                              |
| $\Omega(\theta)$  | Termo de regularização associado à complexidade do modelo.                                 |
| $\lambda$         | Hiperparâmetro que controla a intensidade da regularização.                                |
| $\tau$            | Limiar de decisão probabilística.                                                          |
| $P(\cdot)$        | Probabilidade ou distribuição de probabilidade.                                            |
| $\bar{T}$         | Média térmica da região de interesse.                                                      |
| $T_{\text{med}}$  | Mediana dos valores de temperatura.                                                        |
| $E$               | Energia térmica calculada a partir dos valores de temperatura.                             |
| $E_n$             | Energia térmica normalizada.                                                               |
| $\sigma_T$        | Desvio padrão térmico.                                                                     |
| $A_T$             | Amplitude térmica.                                                                         |
| $T_{\text{max}}$  | Temperatura máxima observada.                                                              |
| $T_{\text{min}}$  | Temperatura mínima observada.                                                              |
| $Q_1$             | Primeiro quartil, correspondente ao percentil 25.                                          |
| $Q_3$             | Terceiro quartil, correspondente ao percentil 75.                                          |
| $P_{90} - P_{10}$ | Diferença entre os percentis 90 e 10.                                                      |
| $\gamma_1$        | Assimetria da distribuição térmica.                                                        |
| $\gamma_2$        | Curtose ou excesso de curtose, conforme a formulação adotada.                              |
| $\mu_r$           | $r$ -ésimo momento central da distribuição térmica.                                        |
| $H$               | Entropia térmica.                                                                          |
| $p_i$             | Probabilidade de ocorrência do $i$ -ésimo nível de temperatura.                            |
| $a$               | Inclinação do perfil térmico obtida por ajuste linear.                                     |
| $b$               | Termo independente do ajuste linear ou viés do hiperplano, conforme o contexto.            |

|                      |                                                                                             |
|----------------------|---------------------------------------------------------------------------------------------|
| $\Delta T_i$         | Diferença sucessiva entre temperaturas de posições vizinhas.                                |
| $\sigma_{\Delta T}$  | Desvio padrão das diferenças sucessivas.                                                    |
| $\Delta \bar{T}$     | Média das diferenças sucessivas.                                                            |
| TV                   | Variação total do perfil térmico.                                                           |
| $\kappa$             | Curvatura estimada a partir de diferenças de segunda ordem.                                 |
| $\bar{\kappa}$       | Curvatura média do perfil térmico.                                                          |
| $G_{\max}$           | Gradiente máximo observado no perfil térmico.                                               |
| $\rho(d)$            | Autocorrelação espacial para deslocamento $d$ .                                             |
| $d$                  | Deslocamento espacial usado no cálculo da autocorrelação.                                   |
| $F$                  | Estatística F usada para avaliar a relevância de atributos.                                 |
| $I(X;Y)$             | Informação mútua entre o atributo $X$ e a classe $Y$ .                                      |
| $p(x,y)$             | Distribuição conjunta entre atributo e classe.                                              |
| $p(x), p(y)$         | Distribuições marginais do atributo e da classe.                                            |
| $S$                  | Conjunto de atributos já selecionados.                                                      |
| $\rho(x_j, x_s)$     | Coefficiente de correlação de Pearson entre o atributo candidato e um atributo selecionado. |
| $\varepsilon$        | Constante pequena usada para evitar divisão por zero no critério mRMR por quociente.        |
| $\mathbf{w}$         | Vetor de pesos do hiperplano da SVM.                                                        |
| $\ \mathbf{w}\ $     | Norma euclidiana do vetor de pesos.                                                         |
| $C$                  | Parâmetro de penalização da SVM de margem suave.                                            |
| $\xi_i$              | Variável de folga associada à $i$ -ésima amostra na SVM.                                    |
| $\alpha_i$           | Multiplicador de Lagrange associado à $i$ -ésima amostra na formulação dual da SVM.         |
| $K(x_i, x_j)$        | Função kernel que mede a similaridade entre duas amostras.                                  |
| $c$                  | Constante do kernel polinomial.                                                             |
| $q$                  | Grau do kernel polinomial.                                                                  |
| $\gamma$             | Parâmetro do kernel RBF que controla a largura da função radial.                            |
| $B$                  | Número total de árvores no Random Forest.                                                   |
| $h_b(x)$             | Predição realizada pela $b$ -ésima árvore de decisão.                                       |
| $\text{mode}(\cdot)$ | Operador que retorna a classe mais votada.                                                  |
| $p_{mk}$             | Proporção de amostras da classe $k$ no nó $m$ de uma árvore de decisão.                     |
| $N_m$                | Número total de amostras no nó $m$ .                                                        |
| $R_m$                | Região do espaço de atributos associada ao nó $m$ .                                         |
| $I(\cdot)$           | Função indicadora, igual a 1 quando a condição é verdadeira e 0 caso contrário.             |
| $G_m$                | Índice de Gini no nó $m$ .                                                                  |
| $H_m$                | Entropia de classes no nó $m$ .                                                             |
| $IG$                 | Ganho de informação obtido por uma divisão em árvore de decisão.                            |
| $O$                  | Sequência de observações em um Modelo Oculto de Markov.                                     |
| $O_t$                | Observação no instante ou posição $t$ .                                                     |
| $Q$                  | Sequência de estados ocultos.                                                               |
| $q_t$                | Estado do sistema no instante ou posição $t$ .                                              |
| $S_i$                | $i$ -ésimo estado do modelo de Markov.                                                      |

|               |                                                                                                       |
|---------------|-------------------------------------------------------------------------------------------------------|
| $a_{ij}$      | Probabilidade de transição do estado $i$ para o estado $j$ .                                          |
| $A$           | Matriz de transição entre estados ocultos.                                                            |
| $B$           | Matriz ou conjunto de probabilidades/densidades de emissão.                                           |
| $b_j(O_t)$    | Probabilidade ou densidade de emissão da observação $O_t$ no estado $j$ .                             |
| $v_k$         | $k$ -ésimo símbolo observável em HMM discreto.                                                        |
| $\pi$         | Distribuição inicial dos estados.                                                                     |
| $\pi_i$       | Probabilidade inicial de o sistema estar no estado $i$ .                                              |
| $\lambda$     | Modelo Oculto de Markov completo, definido por $A$ , $B$ e $\pi$ .                                    |
| $\alpha_t(i)$ | Variável <i>forward</i> : probabilidade da sequência parcial até $t$ e término no estado $i$ .        |
| $\beta_t(i)$  | Variável <i>backward</i> : probabilidade do restante da sequência dado o estado $i$ no instante $t$ . |
| $\gamma_t(i)$ | Probabilidade a posteriori de estar no estado $i$ no instante $t$ .                                   |
| $\xi_t(i,j)$  | Probabilidade a posteriori de transitar do estado $i$ para o estado $j$ entre $t$ e $t+1$ .           |
| $T$           | Comprimento da sequência ou número de janelas do perfil térmico.                                      |
| $N$           | Número de estados ocultos do HMM, quando usado no contexto de modelos de Markov.                      |
| $M$           | Número de símbolos observáveis ou componentes de mistura, conforme o contexto.                        |

# SUMÁRIO

|          |                                                          |    |
|----------|----------------------------------------------------------|----|
| 1        | Introdução .....                                         | 20 |
| 1.1      | Motivação.....                                           | 21 |
| 1.2      | Objetivos .....                                          | 23 |
| 1.3      | Produção científica.....                                 | 23 |
| 1.4      | Estrutura do texto .....                                 | 25 |
| 2        | Fundamentação Teórica .....                              | 27 |
| 2.1      | Isoladores poliméricos e poluição superficial .....      | 27 |
| 2.2      | Termografia.....                                         | 30 |
| 2.3      | Extração dos atributos térmicos .....                    | 33 |
| 2.3.1    | Atributos de nível térmico .....                         | 34 |
| 2.3.2    | Atributos de dispersão e contraste térmico .....         | 35 |
| 2.3.3    | Atributos de forma da distribuição .....                 | 37 |
| 2.3.4    | Atributos de variação espacial.....                      | 39 |
| 2.3.5    | Atributos de dependência espacial .....                  | 42 |
| 2.4      | Seleção dos atributos.....                               | 42 |
| 2.4.1    | Relevância dos atributos.....                            | 43 |
| 2.4.2    | Redundância entre atributos .....                        | 44 |
| 2.4.3    | Método mrmr.....                                         | 45 |
| 2.5      | Classificadores .....                                    | 46 |
| 2.5.1    | Classificadores de entrada de dimensão fixa .....        | 46 |
| 2.5.1.1. | Máquina de vetor de suporte .....                        | 47 |
| 2.5.1.2. | Random forest .....                                      | 50 |
| 2.5.2.   | Classificadores de entrada sequencial .....              | 53 |
| 2.5.2.1. | Cadeia de Markov .....                                   | 53 |
| 2.5.2.2. | Modelo oculto de Markov .....                            | 54 |
| 2.5.2.3. | Modelo oculto de Markov com emissões gaussianas .....    | 60 |
| 2.5.2.4. | Rede LSTM.....                                           | 62 |
| 2.6      | Métricas de avaliação dos modelos.....                   | 69 |
| 2.6.1    | Matriz de confusão .....                                 | 69 |
| 2.6.2    | Métricas de desempenho .....                             | 70 |
| 2.6.3    | Desempenho médio e dispersão .....                       | 73 |
| 3        | Revisão Bibliográfica.....                               | 74 |
| 3.1      | Modelos ocultos de Markov no setor elétrico .....        | 74 |
| 3.2      | Monitoramento por termografia no setor elétrico .....    | 78 |
| 3.3      | Análise das contribuições dos trabalhos correlatos ..... | 83 |
| 4        | Metodologia .....                                        | 87 |
| 4.1      | Aquisição da base de dados.....                          | 88 |
| 4.1.1    | Experimento 1: isoladores em condição limpa.....         | 89 |

|       |                                                             |     |
|-------|-------------------------------------------------------------|-----|
| 4.1.2 | Experimento 2: poluição artificial uniforme .....           | 90  |
| 4.1.3 | Experimento 3: poluição artificial não uniforme.....        | 91  |
| 4.1.4 | Experimento 4: poluição natural .....                       | 92  |
| 4.2   | Extração do perfil térmico.....                             | 94  |
| 4.3   | Pré-processamento dos perfis térmicos .....                 | 95  |
| 4.3.1 | Identificação do grupo de captura .....                     | 95  |
| 4.3.2 | Correção pela temperatura ambiente .....                    | 96  |
| 4.3.3 | Ampliação da base de dados.....                             | 96  |
| 4.4   | Segmentação dos perfis térmicos em janelas .....            | 97  |
| 4.5   | Extração dos atributos térmicos .....                       | 99  |
| 4.6   | Representações do perfil térmico .....                      | 100 |
| 4.6.1 | Representação sequencial.....                               | 101 |
| 4.6.2 | Representação agregada .....                                | 101 |
| 4.6.3 | Representação simbólica .....                               | 102 |
| 4.7   | Validação aninhada .....                                    | 103 |
| 4.8   | Seleção de atributos e configuração .....                   | 104 |
| 4.8.1 | Seleção de atributos na representação contínua.....         | 104 |
| 4.8.2 | Construção dos subconjuntos .....                           | 106 |
| 4.8.3 | Seleção das configurações simbólicas.....                   | 107 |
| 4.9   | Treinamento dos modelos de classificação .....              | 108 |
| 4.9.1 | Ajuste de hiperparâmetros.....                              | 108 |
| 4.9.2 | Predição por grupo de captura.....                          | 109 |
| 4.9.3 | Avaliação por perfil térmico.....                           | 110 |
| 4.9.4 | Avaliação por grupo de captura.....                         | 110 |
| 4.9.5 | Ambiente computacional e implementação.....                 | 111 |
| 4.10  | Análise de robustez dos campeões .....                      | 112 |
| 5     | Resultados e discussões.....                                | 114 |
| 5.1   | Caracterização da base de dados e dos perfis térmicos ..... | 115 |
| 5.2   | Atributos térmicos e construção das representações .....    | 119 |
| 5.3   | Atributos selecionados na representação contínua.....       | 121 |
| 5.4   | Configurações simbólicas selecionadas .....                 | 124 |
| 5.5   | Desempenho dos modelos na validação aninhada.....           | 127 |
| 5.5.1 | Avaliação por perfil térmico.....                           | 134 |
| 5.5.2 | Avaliação por grupo de captura.....                         | 137 |
| 5.6   | Análise de robustez dos modelos campeões.....               | 140 |
| 6     | Conclusões .....                                            | 146 |
| 6.1   | Trabalhos futuros .....                                     | 148 |
|       | REFERÊNCIAS .....                                           | 151 |

# 1 INTRODUÇÃO

O desempenho de isoladores em sistemas elétricos de alta tensão constitui um elemento essencial para garantir a confiabilidade e a segurança operacional das subestações, particularmente em regiões litorâneas. A deposição de partículas higroscópicas provenientes da névoa salina, intensificada por ventos, variações de umidade e efeitos meteorológicos, forma camadas contaminantes capazes de modificar significativamente as propriedades elétricas da superfície externa dos isoladores. Em situação de alta umidade, essas camadas tornam-se condutivas, intensificando a distribuição da corrente de fuga e favorecendo o surgimento de descargas parciais, bandas secas e aquecimentos localizados ao longo do isolador. A evolução desses fenômenos pode culminar em arcos elétricos ou falhas por escoamento (*flashover*), ameaçando a integridade do sistema elétrico de potência (Santos e Sanz-Bobi, 2023).

Esses mecanismos de degradação, associados tanto às condições ambientais quanto ao estresse elétrico, representam riscos significativos para o sistema, podendo gerar interrupções, deterioração da qualidade do fornecimento e penalidades regulatórias às empresas que operam instalações de transmissão e distribuição. Nesse contexto, torna-se indispensável o emprego de estratégias de manutenção preditiva e monitoramento contínuo para assegurar o pleno funcionamento dos ativos e prevenir falhas catastróficas (Bindi *et al.*, 2023).

Tradicionalmente, a análise da condição de isoladores contaminados é realizada mediante ensaios laboratoriais padronizados, como a Densidade Equivalente de Depósito de Sal (do inglês *Equivalent Salt Deposit Density*, ESDD) e a Densidade de Depósito Não Solúvel (do inglês *Non-Soluble Deposit Density*, NSDD) e medições de corrente de fuga (Salem *et al.*, 2022), além da detecção de descargas corona por ultravioleta (UV) (Brito *et al.*, 2022). Embora amplamente empregadas, essas técnicas apresentam limitações importantes: os ensaios de deposição salina são longos e apresentam imprecisões na determinação dos valores medidos, além da possibilidade de destruição das amostras contaminadas durante o transporte para análise (Gouda *et al.*, 2022); a medição de

corrente de fuga, por sua vez, torna-se invasiva por exigir instrumentação dedicada, além de ser sensível a variações ambientais (Wang *et al.*, 2019; Salem *et al.*, 2022; Villalobos *et al.*, 2022); e as inspeções por UV normalmente apresentam metodologias ainda limitadas à aplicação em laboratórios ou ambientes controlados.

Diante dessas limitações operacionais e de aplicabilidade em campo, métodos não invasivos, aplicáveis diretamente *in situ* e capazes de capturar respostas físicas associadas ao estado do isolador passam a ser particularmente úteis. Nesse contexto, métodos capazes de registrar respostas térmicas associadas à condução superficial surgem como alternativas especialmente adequadas para antecipação de falhas (Shi, 2021).

## 1.1 MOTIVAÇÃO

Apesar dos avanços no monitoramento por termografia, observa-se que a maior parte das abordagens presentes na literatura trata os dados termográficos de forma predominantemente estática. Em muitos casos, as termografias são analisadas como imagens bidimensionais independentes, sendo processadas por técnicas de visão computacional ou modelos de aprendizado profundo, como redes neurais convolucionais (Tang e Jian, 2024). Embora essas abordagens sejam eficazes para a detecção, segmentação e classificação de padrões térmicos, não há como garantir que elas modelam de forma interpretável a continuidade espacial do campo térmico ao longo do eixo longitudinal do isolador. Assim, a relação ordenada entre regiões adjacentes, como lado fase, região central e lado terra, pode não ser explorada diretamente como uma sequência espacial estruturada. Essa limitação motiva a investigação de representações capazes de preservar a organização longitudinal dos perfis térmicos, permitindo avaliar se a dependência espacial entre janelas sucessivas contém informação discriminativa relevante para a classificação da condição do isolador.

No entanto, a distribuição de temperatura ao longo do isolador apresenta um comportamento intrinsecamente estruturado. O perfil térmico longitudinal reflete a interação entre fenômenos físicos como campo elétrico, gradientes de temperatura, corrente de fuga, presença de umidade e deposição não uniforme de contaminantes (Zhang e Meng, 2019). Dessa forma, cada ponto do perfil não deve ser interpretado de maneira isolada, mas sim como parte de uma sequência ordenada, na qual a relação entre posições consecutivas carrega informação relevante sobre o estado operacional do equipamento. Estudos na área de análise de séries temporais e reconhecimento de padrões

destacam que a modelagem adequada de dados sequenciais é fundamental quando há dependência entre observações adjacentes (Bishop, 2006).

Nesse contexto, a forma de representação do perfil térmico torna-se um elemento central para o desempenho dos sistemas de diagnóstico. Abordagens baseadas em agregação de atributos, como médias, desvios padrão ou estatísticas globais, embora úteis para redução de dimensionalidade, podem levar à perda de informações estruturais importantes, especialmente aquelas relacionadas à dinâmica espacial do aquecimento ao longo do isolador. Como consequência, padrões sutis associados a diferentes condições de poluição podem não ser adequadamente capturados por modelos que ignoram essa organização sequencial.

Por outro lado, modelos probabilísticos sequenciais oferecem uma alternativa promissora para esse tipo de problema. Entre eles, destacam-se os Modelos Ocultos de Markov (do inglês *Hidden Markov Model*, HMM), que caracterizar valores observáveis de saída de sistemas, que dependem de valores não observáveis. Esses modelos têm sido aplicados com sucesso em diversas áreas, incluindo reconhecimento de padrões, processamento de sinais e diagnóstico de sistemas industriais, devido à sua capacidade de capturar dependências estruturais e inferir comportamentos subjacentes não diretamente observáveis (Rabiner, 1989).

Apesar da maturidade desses modelos e de sua ampla aplicação em problemas de natureza sequencial, observa-se uma lacuna significativa na literatura no que se refere à sua utilização na modelagem de perfis térmicos de isoladores. Em particular, não são encontrados trabalhos que explorem explicitamente a natureza sequencial do perfil térmico longitudinal como uma fonte primária de informação para classificação de estados de poluição.

Diante desse cenário, a principal motivação deste trabalho reside na necessidade de investigar abordagens que não apenas utilizem dados termográficos, mas que também respeitem e explorem sua estrutura intrínseca. Parte-se da hipótese de que a modelagem sequencial do perfil térmico pode fornecer uma representação mais fiel do fenômeno físico observado, permitindo capturar padrões espaciais e transições características entre diferentes regiões do isolador. Consequentemente, espera-se que essa abordagem contribua para o desenvolvimento de métodos de diagnóstico mais robustos, sensíveis e interpretáveis.

## 1.2 OBJETIVOS

O objetivo geral deste trabalho é desenvolver e avaliar uma metodologia para representação e classificação de perfis térmicos de isoladores poliméricos, visando ao diagnóstico de condições associadas à poluição superficial por meio da comparação entre representações agregada, sequencial e simbólica e diferentes modelos de classificação.

Como objetivos específicos desta pesquisa, pode-se citar:

- Organizar, padronizar e ampliar uma base de dados composta por perfis térmicos extraídos de imagens termográficas de isoladores poliméricos em condições limpas e poluídas;
- Extrair perfis térmicos longitudinais dos isoladores, realizar a correção em relação à temperatura ambiente e segmentá-los em janelas sucessivas ao longo da direção espacial do isolador;
- Definir e documentar uma metodologia sistemática de extração, seleção e organização de atributos térmicos associados ao nível térmico, dispersão, contraste, forma da distribuição, variação espacial e dependência espacial;
- Construir e avaliar diferentes formas de representação dos perfis térmicos, contemplando as representações agregada, sequencial e simbólica;
- Desenvolver, treinar e avaliar modelos de classificação aplicados aos perfis térmicos, incluindo Máquina de Vetor de Suporte (do inglês *Support Vector Machine*, SVM), Floresta Aleatória (do inglês, *Random Forest*, RF); Memória de Curto Longo Prazo (do inglês, *Long Short-Term Memory*, LSTM) e os Modelos Ocultos de Markov (do inglês *Hidden Markov Model*, HMM);
- Comparar o desempenho dos modelos e das representações por meio de métricas de classificação e validação por grupos de captura, considerando a capacidade de distinguir isoladores limpos e poluídos.

## 1.3 PRODUÇÃO CIENTÍFICA

Durante o desenvolvimento do mestrado, o autor participou da elaboração de publicações científicas relacionadas ao monitoramento e ao diagnóstico de equipamentos elétricos, com ênfase no uso de técnicas de processamento de sinais, processamento de imagens, termografia, imagens ultravioleta e aprendizado de máquina. Essa produção

científica resultou em dois artigos publicados em periódico internacional e três trabalhos publicados em congresso, contribuindo para a consolidação das abordagens investigadas ao longo da pesquisa.

As publicações em periódico foram veiculadas na revista *Measurement* e abordam aplicações de técnicas computacionais no monitoramento de equipamentos elétricos. O primeiro trabalho trata do monitoramento de para-raios de ZnO por meio de redes neurais convolucionais, processamento de imagens e alinhamento de sinais, enquanto o segundo apresenta uma metodologia automatizada para o monitoramento de isolamento por imagens ultravioleta e aprendizado profundo. Esses trabalhos reforçam a inserção da pesquisa em temas associados à inspeção não invasiva e ao diagnóstico baseado em dados.

Além das publicações em periódico, foram apresentados três trabalhos no 2025 *Workshop on Communication Networks and Power Systems* (WCNPS), contemplando temas relacionados à corrente de fuga, poluição em isoladores poliméricos e análise de perfis térmicos. Dentre essas publicações, destaca-se o artigo *Comparative Analysis of Thermal Profile Alignment Techniques for Surge Arresters*, que recebeu o prêmio de melhor artigo na área de sistemas elétricos de potência. Esse reconhecimento evidencia o potencial das técnicas de análise de perfis térmicos sequenciais com o Modelo Oculto de Markov, bem como sua relevância para o monitoramento de equipamentos em sistemas elétricos.

As publicações desenvolvidas durante o período do mestrado são listadas a seguir:

1. ARAUJO, B. V. S.; RODRIGUES, G. A.; OLIVEIRA, J. H. P. de; *et al.* *Monitoring ZnO surge arresters using convolutional neural networks and image processing techniques combined with signal alignment. Measurement*, v. 248, 116889, 2025. DOI: 10.1016/j.measurement.2025.116889.
2. RODRIGUES, G. A.; ARAUJO, B. V. S.; OLIVEIRA, J. H. P. de; *et al.* *Automated monitoring of insulation by ultraviolet imaging employing deep learning. Measurement*, v. 242, Part D, 116018, 2025. DOI: 10.1016/j.measurement.2024.116018.
3. OLIVEIRA, J. H. P. de; ARAUJO, B. V. S.; RODRIGUES, G. A.; *et al.* *Comparative Analysis of Thermal Profile Alignment Techniques for Surge Arresters. 2025 Workshop on Communication Networks and Power Systems*, 2025. DOI: 10.1109/WCNPS69127.2025.11295890.

4. OLIVEIRA, A.; ARAÚJO, E. S.; OLIVEIRA, J. H. P. de; *et al.* *Assessment of Leakage Current on Polymeric Insulators Under Non-Uniform Saline Pollution Conditions*. **2025 Workshop on Communication Networks and Power Systems**, 2025. DOI: 10.1109/WCNPS69127.2025.11295919.
5. LISBOA, J. L. de C.; OLIVEIRA, J. H. P. de; LEBRE, U. D. E. S.; *et al.* *Statistical Analysis of Insulation Leakage Currents Under Pollution Incidence in a 500 kV Substation*. **2025 Workshop on Communication Networks and Power Systems**, 2025. DOI: 10.1109/WCNPS69127.2025.11295915.

## 1.4 ESTRUTURA DO TEXTO

A presente dissertação está organizada em seis capítulos, incluindo este.

No Capítulo 2, desenvolve-se a fundamentação teórica necessária à pesquisa. Inicialmente, são abordados os conceitos relacionados aos isoladores elétricos, incluindo suas funções, tipos construtivos, aspectos associados aos isoladores poliméricos, poluição superficial e efeitos elétricos e térmicos decorrentes da contaminação. Em seguida, são discutidas as estratégias de monitoramento e diagnóstico, as técnicas invasivas e não invasivas de inspeção, os fundamentos da termografia infravermelha, os fatores que influenciam a medição termográfica e sua aplicação em isoladores.

Ainda no Capítulo 2, são apresentados os procedimentos de extração e interpretação de atributos térmicos, contemplando descritores de nível térmico, dispersão, forma da distribuição, variação espacial e dependência espacial. Também são discutidos os métodos de seleção de atributos, com ênfase na relevância, redundância e no método *Minimum Redundancy Maximum Relevance* (mRMR), bem como os classificadores empregados no trabalho, incluindo modelos de entrada fixa, como a Máquina de Vetor de Suporte e Random Forest, e modelos de entrada sequencial, com destaque para a rede *Long Short Term Memory* e Modelos Ocultos de Markov.

No Capítulo 3, realiza-se a revisão bibliográfica, reunindo trabalhos relacionados ao uso de técnicas de monitoramento, processamento de sinais, processamento de imagens, aprendizado de máquina e modelos probabilísticos aplicados ao diagnóstico de equipamentos elétricos. A partir dessa revisão, a pesquisa realizada é posicionada em

relação às abordagens existentes, com destaque para a lacuna associada à modelagem sequencial de perfis térmicos de isoladores.

No Capítulo 4, descreve-se a metodologia proposta, incluindo a organização da base de dados, o tratamento das imagens termográficas, a extração dos perfis e atributos térmicos, a seleção das variáveis mais relevantes e o treinamento dos modelos de classificação. Também são detalhados os critérios de avaliação adotados para a comparação do desempenho dos classificadores.

No Capítulo 5, são apresentados os resultados obtidos a partir da aplicação da metodologia proposta, incluindo a análise do desempenho dos atributos selecionados e dos modelos de classificação avaliados.

Por fim, no Capítulo 6, são reunidas as conclusões da pesquisa, com destaque para as principais contribuições do trabalho, suas limitações e as sugestões para estudos futuros voltados ao aprimoramento do diagnóstico de isoladores por meio de termografia e modelagem baseada em dados.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste Capítulo, são reunidos os fundamentos teóricos necessários ao desenvolvimento da metodologia proposta nesta dissertação. Inicialmente, são abordados os conceitos relacionados aos isoladores elétricos, com ênfase nos isoladores poliméricos, na poluição superficial e nos efeitos elétricos e térmicos decorrentes da contaminação. Em seguida, são apresentados os fundamentos da termografia infravermelha e sua aplicação no monitoramento de equipamentos elétricos, incluindo os fatores que influenciam a medição térmica. Também são descritos os procedimentos de extração e seleção de atributos térmicos, bem como os fundamentos dos modelos de classificação utilizados, abrangendo classificadores de entrada fixa e classificadores aplicados à análise de sequências. Por fim, são apresentadas as métricas empregadas para avaliar o desempenho dos modelos, permitindo a comparação das diferentes abordagens consideradas neste trabalho.

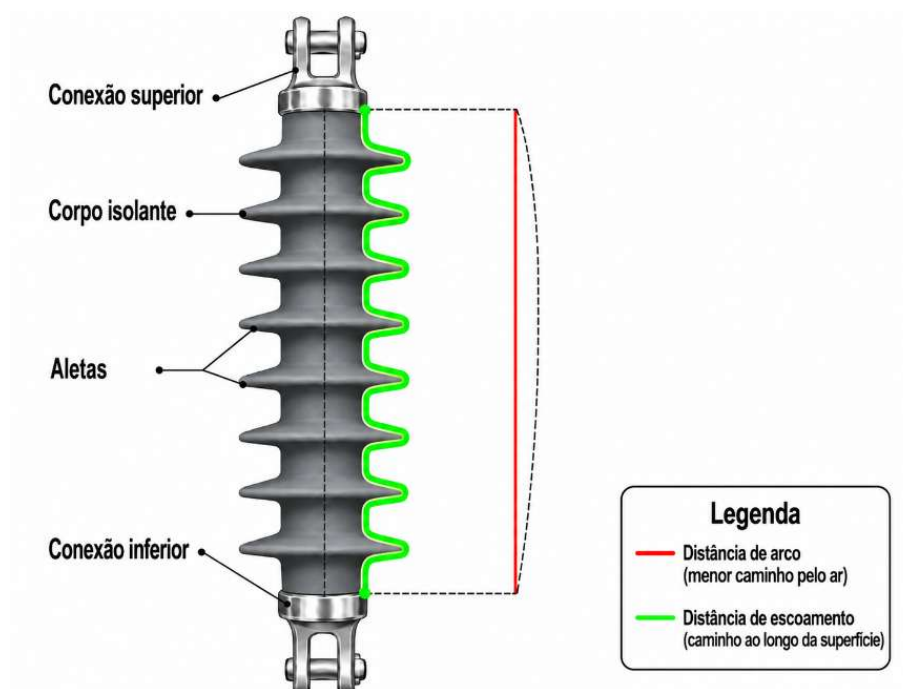
### 2.1 ISOLADORES POLIMÉRICOS E POLUIÇÃO SUPERFICIAL

Os isoladores elétricos são componentes essenciais em sistemas elétricos de potência, pois asseguram o isolamento entre partes energizadas e estruturas aterradas, ao mesmo tempo em que suportam esforços mecânicos associados aos condutores, barramentos e demais elementos do sistema. Dessa forma, seu estado operacional e confiabilidade está diretamente relacionado à continuidade do fornecimento de energia elétrica, uma vez que falhas no isolamento podem provocar descargas superficiais, desligamentos não programados e danos a equipamentos adjacentes (Frontin, 2013; Dias, 2021; Nóbrega, 2024).

Em redes de média e alta tensão, os isoladores estão sujeitos a solicitações elétricas, mecânicas, térmicas e ambientais. Do ponto de vista elétrico, devem suportar a tensão nominal do sistema e eventuais sobretensões sem permitir a formação de caminhos condutivos entre pontos submetidos a diferentes potenciais. Nesse contexto, a distância de escoamento é um parâmetro relevante, especialmente em ambientes contaminados, pois corresponde ao caminho superficial disponível para a circulação da corrente de fuga ao longo do isolador. Quando contaminantes e umidade se acumulam sobre essa

superfície, a condutividade superficial pode aumentar, favorecendo descargas parciais, formação de bandas secas e, em casos críticos, a ocorrência de *flashover* (Kuffel, Zaengl e Kuffel, 2000; Frontin, 2013; Salem *et al.*, 2022). Na Figura 1 ilustra-se a comparação entre a distância de arco e distância de escoamento em um isolador.

Figura 1. Comparação entre distância de arco e distância de escoamento.



Fonte: autoria própria.

Os isoladores podem ser classificados, de forma geral, em cerâmicos e poliméricos. Os isoladores cerâmicos apresentam elevada rigidez dielétrica e ampla tradição de uso no setor elétrico, porém possuem maior massa e maior fragilidade mecânica. Os isoladores poliméricos, também denominados compostos ou não cerâmicos, destacam-se pelo menor peso, facilidade de instalação, resistência mecânica adequada e desempenho favorável em ambientes contaminados, principalmente devido às propriedades do revestimento externo (Gorur, Cherney e Burnham, 1999; Dias, 2021; Nóbrega, 2024).

A estrutura típica de um isolador polimérico é composta por núcleo de fibra de vidro, revestimento polimérico e terminais metálicos. O núcleo é responsável pela resistência mecânica à tração, enquanto o revestimento atua como barreira dielétrica e ambiental, protegendo o isolador contra umidade, radiação ultravioleta, poluição e descargas superficiais. As aletas aumentam a distância de escoamento e dificultam a formação de filmes contínuos de água sobre a superfície. Os terminais, por sua vez,

realizam a conexão mecânica entre o isolador, o condutor e a estrutura de sustentação, transmitindo os esforços mecânicos ao núcleo (Zhao e Bernstorff, 1998; Gorur, Cherney e Burnham, 1999; Ferreira, 2007). Na Figura 2 ilustra-se, respectivamente, exemplares de isolador de porcelana, vidro e polimérico.

Figura 2. Exemplares de isolador. (a) Porcelana; (b) Vidro; (c) Polimérico.



Fonte: autoria própria.

Uma característica importante dos isoladores poliméricos é a hidrofobicidade, especialmente quando o revestimento é constituído por borracha de silicone. Essa propriedade reduz a tendência de formação de filmes contínuos de água sobre a superfície, favorecendo a formação de gotas isoladas e diminuindo a continuidade elétrica entre as regiões energizadas e aterradas. Com isso, há menor propensão à intensificação da corrente de fuga em condições úmidas. Entretanto, a exposição prolongada à radiação ultravioleta, temperatura, umidade, poluição, esforços mecânicos e estresse elétrico pode provocar envelhecimento do material, perda parcial da hidrofobicidade, fissuras, erosão, trilhamento, exposição do núcleo, corrosão das ferragens e comprometimento da interface entre núcleo, revestimento e terminais (Diniz, 2019; Qiao *et al.*, 2022; Salem *et al.*, 2023).

A poluição superficial em isoladores ocorre pela deposição de partículas contaminantes sobre a superfície isolante. Esses contaminantes podem ser de origem marinha, industrial, agrícola ou natural, incluindo sais, poeira, fuligem, partículas sólidas, resíduos orgânicos e compostos químicos. Em regiões litorâneas, a deposição de sais provenientes da névoa salina é particularmente relevante, pois esses compostos apresentam caráter higroscópico e podem formar camadas condutivas quando associados

à presença de umidade. Nessas condições, a corrente de fuga tende a aumentar, favorecendo a formação de aquecimentos localizados, bandas secas e descargas superficiais, resultando na degradação do equipamento (Gorur, Cherney e Burnham, 1999; Dias, 2021; Salem *et al.*, 2022; Ahmed *et al.*, 2022).

Do ponto de vista térmico, a passagem da corrente de fuga por regiões contaminadas produz dissipação de energia na superfície do isolador. Esse processo pode gerar elevações locais de temperatura, especialmente em regiões com maior concentração de contaminantes, maior condutividade superficial ou maior solicitação elétrica. Assim, a distribuição térmica ao longo do isolador passa a refletir, ainda que indiretamente, o estado de sua superfície. Em uma condição íntegra e limpa, espera-se baixa corrente de fuga e uma distribuição térmica relativamente mais uniforme. Em uma condição contaminada, envelhecida ou defeituosa, a superfície pode apresentar aumento localizado da condutividade, maior dissipação de potência e aquecimento localizado (Diniz, 2019; Dias, 2021; Salem *et al.*, 2023).

Essa relação entre poluição, corrente de fuga e aquecimento superficial fundamenta o uso da termografia infravermelha como técnica não invasiva para análise da condição de isoladores. Por meio da inspeção térmica, torna-se possível observar padrões de aquecimento associados à presença de contaminantes, perda de hidrofobicidade, intensificação da condução superficial e regiões de maior solicitação elétrica. No contexto desta dissertação, essa fundamentação é particularmente relevante porque justifica a extração de perfis térmicos longitudinais e sua análise como representação do comportamento espacial da temperatura ao longo do isolador.

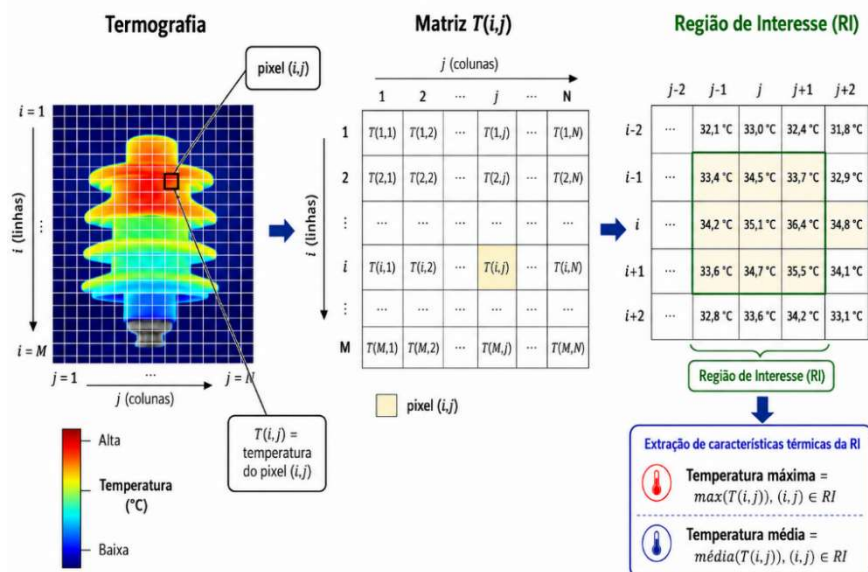
## 2.2 TERMOGRAFIA

A termografia é uma técnica de medição não invasiva que permite estimar a distribuição de temperatura aparente de uma superfície a partir da radiação infravermelha emitida, refletida e transmitida pelo objeto analisado. Essa técnica tem sido amplamente utilizada em inspeções de sistemas elétricos por possibilitar a avaliação térmica de equipamentos energizados, sem contato direto e sem necessidade de interrupção da operação. Sua aplicabilidade decorre do fato de que alterações físicas, elétricas ou superficiais em componentes elétricos podem se manifestar como variações térmicas detectáveis por sensores infravermelhos (Muniz e Mendes, 2019; Pan *et al.*, 2025).

A base física da termografia está associada ao princípio de que todo corpo com temperatura superior a 0 K emite radiação eletromagnética. A intensidade dessa emissão depende da temperatura absoluta do corpo e das propriedades radiativas da superfície. O corpo negro é utilizado como referência ideal, por absorver toda a radiação incidente e emitir a máxima radiação possível para uma dada temperatura. Entretanto, superfícies reais não se comportam como corpos negros perfeitos, pois parte da radiação incidente pode ser absorvida, refletida ou transmitida. Em materiais opacos, como grande parte dos componentes elétricos e isoladores cerâmicos ou poliméricos, a transmissividade é geralmente desprezível, de modo que a interpretação termográfica depende principalmente da emissão própria da superfície e da radiação refletida do ambiente (Santos, 2012; Muniz e Mendes, 2019).

O termovisor converte a radiação infravermelha recebida pelo detector em sinais elétricos que são processados para formar uma imagem térmica. Em câmeras modernas, esse processo ocorre por meio de uma matriz de sensores, geralmente denominada matriz de plano focal, na qual cada elemento sensível registra a radiação proveniente de uma pequena região da cena. Após etapas internas de calibração e compensação, a imagem resultante pode ser interpretada como uma matriz térmica, em que cada pixel está associado a um valor de temperatura aparente ou intensidade térmica calibrada (Muniz e Mendes, 2019; Pan *et al.*, 2025). Na Figura 3 ilustra-se um esquemático da matriz de temperatura e a extração de informações térmicas na região de interesse.

Figura 3. Esquemático da matriz de temperatura e a extração de informações térmicas na região de interesse.



Fonte: autoria própria.

A medição termográfica, entretanto, não corresponde apenas à temperatura real da superfície. O valor estimado pelo termovisor é influenciado por parâmetros relacionados ao objeto, ao ambiente e ao próprio sistema de medição. Entre os principais fatores estão emissividade, temperatura refletida, temperatura ambiente, umidade atmosférica, distância de aquisição, ângulo de observação, resolução espacial e ruído do sensor. Esses fatores podem alterar os valores absolutos de temperatura e, conseqüentemente, afetar a comparação entre inspeções realizadas em condições distintas (Muniz e Mendes, 2019; Barreira, Almeida e Simões, 2021; Pan *et al.*, 2025).

A emissividade é um dos parâmetros mais relevantes na medição termográfica, pois define a capacidade de uma superfície emitir radiação em relação a um corpo negro. Valores incorretos de emissividade inseridos no termovisor podem produzir erros expressivos na temperatura estimada, especialmente em superfícies metálicas, polidas, oxidadas de forma irregular ou contaminadas. Além disso, a emissividade pode variar com a umidade e com o estado superficial do material, o que é particularmente relevante para isoladores expostos ao ambiente e sujeitos à deposição de poluentes (Muniz e Mendes, 2019; Barreira, Almeida e Simões, 2021).

A temperatura refletida também deve ser considerada, pois parte da radiação registrada pelo termovisor pode não ser emitida pelo objeto, mas refletida por sua superfície. Esse efeito é mais pronunciado em materiais de baixa emissividade ou com comportamento especular, podendo produzir falsas indicações de aquecimento ou mascarar anomalias reais. Em inspeções de sistemas elétricos, esse fator é relevante em ferragens, conexões, carcaças metálicas e regiões cuja superfície tenha sido modificada por oxidação, umidade ou contaminação (Lafargue-Tallet *et al.*, 2022).

A temperatura ambiente, a umidade atmosférica e a distância de aquisição influenciam a transferência de calor e a transmissão da radiação infravermelha entre o objeto e o sensor. Variações na temperatura do ar modificam as condições de troca térmica por convecção e radiação, enquanto a radiação solar pode aquecer diretamente o componente ou gerar reflexos no termograma. A umidade e a distância, por sua vez, afetam a transmitância atmosférica, podendo atenuar a radiação recebida pelo termovisor e introduzir incertezas na medição, especialmente em inspeções de campo (Santos, 2012; Muniz e Mendes, 2019; Pan *et al.*, 2025).

O ângulo de aquisição e a resolução espacial também interferem na qualidade da informação térmica. Ângulos muito oblíquos podem modificar a emissividade aparente da superfície, reduzir a área efetivamente observada e aumentar a influência de reflexões.

Já a resolução espacial determina o menor detalhe térmico distinguível na imagem e depende da distância de aquisição, do conjunto óptico (lente, etc.) utilizado e do tamanho do detector. Quando o objeto ocupa poucos pixels na imagem, pequenas regiões aquecidas podem ser suavizadas ou não detectadas adequadamente, comprometendo a extração de atributos e a análise de perfis térmicos (Muniz e Mendes, 2019; Pan *et al.*, 2025).

Além desses fatores, o ruído de medição pode afetar a repetibilidade dos valores extraídos da imagem térmica. Esse ruído está associado a limitações do detector, discretização dos sinais, instabilidades eletrônicas, condições de calibração e variações ambientais durante a aquisição. Em aplicações baseadas em atributos numéricos, como a proposta nesta dissertação, essas variações podem influenciar medidas como temperatura máxima, média, desvio padrão, gradiente térmico e descritores espaciais. Por isso, procedimentos de padronização, correção ambiental e análise relativa são importantes para reduzir a influência de variações não associadas diretamente ao estado do isolador (Hou *et al.*, 2022).

### 2.3 EXTRAÇÃO DOS ATRIBUTOS TÉRMICOS

A extração de atributos térmicos consiste na transformação de uma imagem termográfica, de uma região de interesse ou de um perfil térmico em um conjunto de descritores numéricos capazes de representar propriedades relevantes do campo de temperatura. Em imagens térmicas, cada pixel pode ser interpretado como uma amostra espacial da temperatura aparente, de modo que a imagem deixa de ser apenas uma representação visual e passa a constituir uma matriz quantitativa de medições térmicas. Essa abordagem é coerente com estudos que utilizam termogramas para análise automática, nos quais atributos estatísticos, texturais e espaciais são empregados para caracterizar padrões de aquecimento, heterogeneidade térmica e anomalias locais (Han *et al.*, 2022; Periyasamy *et al.*, 2022; Santoro e Sesana, 2025).

Considere uma região térmica de interesse  $R$ , composta por  $L$  pixels válidos, conforme ilustrado na Figura 3, cujos valores de temperatura são representados por

$$T = \{T_1, T_2, \dots, T_L\}, \quad (1)$$

em que  $T_i$  representa a temperatura, ou intensidade térmica calibrada, do  $i$ -ésimo *pixel*. Quando a imagem é analisada como perfil unidimensional, pode-se representar o sinal térmico como

$$\mathbf{T} = \{T_1, T_2, \dots, T_l\}^T, \quad (2)$$

em que  $x$  indica a posição espacial. Assim, os atributos térmicos podem ser organizados em grupos complementares: atributos de nível térmico, dispersão e contraste, forma da distribuição: assimetria e curtose, variação espacial, dependência espacial e interpretação física. Essa organização é útil porque separa diferentes aspectos do fenômeno térmico: intensidade média do aquecimento, grau de heterogeneidade, presença de assimetrias, transições espaciais abruptas e persistência de padrões ao longo da superfície.

### 2.3.1 ATRIBUTOS DE NÍVEL TÉRMICO

Os atributos de nível térmico descrevem a magnitude global da temperatura na região analisada. Eles fornecem uma primeira estimativa do estado térmico predominante da superfície, indicando se a região apresenta aquecimento geral, resfriamento ou manutenção de um nível térmico aproximadamente homogêneo. Em termografia infravermelha, atributos como média, variância, assimetria, curtose e contraste têm sido empregados para estudar características estatísticas de termografias, especialmente porque imagens térmicas possuem propriedades diferentes das imagens digitais convencionais (Han *et al.*, 2022).

A média amostral da temperatura na região de interesse representa o valor médio da temperatura na região de interesse, cuja expressão é dada por

$$\mu_T = \frac{1}{L} \sum_{i=1}^L T_i. \quad (3)$$

Esse atributo expressa o nível térmico do conjunto de pixels. Em termos físicos, valores elevados de média podem indicar maior emissão térmica superficial, acúmulo de calor ou presença de fontes térmicas mais intensas. Em análises de falhas, regiões defeituosas podem apresentar maior temperatura média quando há retenção de calor, dissipação anômala ou concentração localizada de energia.

Outro atributo de nível térmico é a mediana, que corresponde ao valor central da distribuição ordenada de temperaturas, podendo ser expressa por

$$\tilde{T} = \text{mediana}(T_1, T_2, \dots, T_l). \quad (4)$$

A mediana é menos sensível a valores extremos do que a média. Por isso, torna-se útil quando a região analisada contém poucos pixels muito quentes ou muito frios, os quais poderiam deslocar artificialmente a média. Em superfícies com pontos de aquecimento isolados, a comparação entre média e mediana permite avaliar se o aquecimento é generalizado ou concentrado. Quando a média é substancialmente maior que a mediana, pode haver presença de pontos quentes localizados; quando ambas são próximas, o campo térmico tende a ser mais uniforme.

A energia térmica, no contexto de atributos de imagem ou sinal, pode ser definida como a soma dos quadrados dos valores térmicos

$$E_T = \sum_{i=1}^L T_i^2, \quad (5)$$

ou em sua forma normalizada

$$E_{T_{norm}} = \frac{1}{L} \sum_{i=1}^L T_i^2. \quad (6)$$

Esse atributo enfatiza temperaturas mais altas, pois o termo quadrático aumenta a contribuição dos pixels mais intensos. Dessa forma, a energia é especialmente sensível a regiões com aquecimento acentuado. Em descritores de textura, a energia também aparece associada à concentração ou uniformidade dos níveis de intensidade, sendo empregada em matrizes de coocorrência para caracterizar padrões texturais em termografias (Milosevic, Jankovic e Peulic, 2014; Periyasamy *et al.*, 2022).

### 2.3.2 ATRIBUTOS DE DISPERSÃO E CONTRASTE TÉRMICO

Os atributos de dispersão e contraste quantificam a variabilidade térmica dentro da região analisada. Enquanto os atributos de nível indicam o quanto uma região está aquecida em média, os atributos de dispersão indicam como esse aquecimento se distribui espacialmente. Essa distinção é importante porque duas regiões podem apresentar médias

semelhantes, mas padrões térmicos distintos: uma pode ser aproximadamente uniforme, enquanto outra pode conter alternância entre áreas frias e quentes.

O desvio padrão térmico mede o espalhamento dos valores em relação à média, sendo definido por

$$\sigma_T = \sqrt{\frac{1}{L-1} \sum_{i=1}^L (T_i - \mu_T)^2}. \quad (7)$$

Valores elevados de  $\sigma_T$  indicam maior heterogeneidade térmica. Em termografia, isso pode estar associado a aquecimento irregular, presença de materiais com propriedades térmicas distintas, deposição não uniforme de contaminantes, falhas internas, rugosidade superficial ou variações de emissividade. Em estudos de classificação de termografia, o desvio padrão e a variância são frequentemente utilizados para diferenciar regiões normais e anômalas, pois regiões alteradas tendem a apresentar maior variação térmica (Periyasamy *et al.*, 2022; Santoro e Sesana, 2025).

Ademais, a amplitude térmica é definida como a diferença entre a maior e a menor temperatura observada

$$A_T = T_{max} - T_{min}, \quad (8)$$

em que  $T_{max}$  e  $T_{min}$  indicam, respectivamente a temperatura máxima e mínima observada. Esse atributo fornece uma medida simples da faixa total de variação térmica. Em imagens térmicas, amplitudes elevadas sugerem forte contraste entre zonas quentes e frias. Entretanto, como depende apenas de dois valores, a amplitude é sensível a ruídos, pixels saturados ou reflexões térmicas pontuais.

Além da amplitude, o intervalo interquartil (do inglês *interquartile range*, IQR), pode ser utilizado como medida robusta de dispersão

$$IQR = Q_3 - Q_1, \quad (9)$$

em que  $Q_1$  é o primeiro quartil, correspondente ao percentil 25, e  $Q_3$  é o terceiro quartil, correspondente ao percentil 75. O IQR mede a dispersão da metade central dos dados e é menos sensível a valores extremos do que a amplitude. Montgomery e Runger (2018) destacam o intervalo interquartil como uma medida de variabilidade menos influenciada

por valores extremos, sendo útil quando se deseja avaliar a variabilidade predominante sem que poucos pontos muito quentes dominem a análise (Montgomery e Runger, 2018).

O desvio absoluto mediano (do inglês *Median Absolute Deviation*, MAD), também pode ser utilizado como medida robusta de dispersão

$$MAD = \text{mediana}(|T_i - \tilde{T}|). \quad (10)$$

Esse atributo representa a mediana dos desvios absolutos em relação à mediana térmica. Por ser robusto, o *MAD* é adequado para imagens em que há valores extremos, como pequenas regiões de reflexo, bordas com erro de segmentação ou pixels isolados com ruído. Em aplicações ambientais ou industriais, esse atributo pode auxiliar na diferenciação entre heterogeneidade térmica estrutural e alterações pontuais.

Por fim, a diferença entre percentis, como  $P_{90} - P_{10}$ , mede a faixa térmica ocupada pela maior parte dos dados, desconsiderando extremos inferiores e superiores. Esse atributo descreve o contraste térmico robusto entre regiões relativamente frias e relativamente quentes. Em comparação com a amplitude,  $P_{90} - P_{10}$  reduz a influência de ruídos extremos e fornece uma estimativa mais estável do contraste interno da região. Em imagens térmicas, esse atributo pode ser usado para caracterizar poluição não uniforme, acúmulo desigual de material, falhas localizadas ou diferenças de dissipação térmica entre partes da superfície.

### 2.3.3 ATRIBUTOS DE FORMA DA DISTRIBUIÇÃO

Os atributos de forma da distribuição descrevem o comportamento global do histograma térmico. Eles não se limitam ao nível médio nem à variabilidade, mas indicam como os valores estão distribuídos, se há predominância de cauda à direita ou à esquerda, concentração em torno da média ou maior aleatoriedade térmica. Esses atributos são importantes porque alterações físicas na superfície podem modificar a forma da distribuição sem, necessariamente, produzir grandes mudanças na média.

A assimetria mede o grau de falta de simetria da distribuição térmica em relação ao seu centro. Esse atributo é calculado a partir do terceiro momento central padronizado, sendo definido por

$$g_1 = \frac{m_3}{m_2^{3/2}}, \quad (11)$$

em que

$$m_i = \frac{1}{L} \sum_{i=1}^L (T_i - \mu_T)^i, \quad (12)$$

em que  $m_2$  e  $m_3$  representam, respectivamente, o segundo e terceiro momento central.

Em imagens térmicas, a assimetria permite avaliar se a distribuição dos valores de temperatura apresenta predominância de caudas em uma direção. Valores positivos indicam tendência à presença de uma cauda à direita, associada a temperaturas mais elevadas do que o padrão predominante da janela. Esse comportamento pode sugerir pontos quentes, aquecimento localizado ou pequenas regiões com maior emissão térmica. Valores negativos indicam tendência à presença de uma cauda à esquerda, associada a temperaturas mais baixas do que o padrão central, podendo indicar pontos frios, sombreamento térmico, perda local de calor ou regiões com menor emissão aparente.

Ademais, a curtose descreve o grau de concentração da distribuição e o comportamento de suas caudas. Zwillinger e Kokoska (2000) apresentam a curtose como uma medida relacionada à extensão do pico da distribuição, indicando que valores menores em magnitude estão associados a distribuições mais achatadas e uniformes (Zwillinger e Kokoska, 2000). Estatisticamente, ela é baseada no quarto momento central padronizado, sendo expressa por

$$g_2 = \frac{m_4}{m_2^2}, \quad (13)$$

em que  $m_4$  representa o quarto momento central e  $m_2$  o segundo momento central. Além da curtose  $g_2$ , pode-se utilizar o excesso de curtose, definido por

$$g_2 - 3 = \frac{m_4}{m_2^2} - 3. \quad (14)$$

Essa distinção é importante porque na formulação clássica de Pearson, uma distribuição normal apresenta curtose igual a 3, já na formulação de Fisher, subtrai-se 3, de modo que a distribuição normal passa a apresentar excesso de curtose igual a 0. Assim,

valores positivos indicam distribuição com maior presença de valores extremos ou caudas mais pronunciadas em comparação à normal; valores negativos indicam distribuição mais achatada, com menor concentração de extremos; e valores próximos de zero indicam comportamento semelhante ao de uma distribuição normal quanto à concentração e às caudas.

No contexto térmico, a curtose é relevante porque permite identificar se a janela analisada contém valores extremos de temperatura, mesmo quando a média e o desvio padrão não evidenciam completamente essa condição. Uma curtose elevada pode estar associada à presença de pontos de aquecimento intenso, transições térmicas concentradas ou pequenas regiões anômalas. Já valores reduzidos podem indicar uma distribuição térmica mais espalhada e uniforme, sem predominância de picos térmicos muito marcantes.

A entropia térmica mede o grau de desordem ou incerteza associado à distribuição dos valores de temperatura. Para seu cálculo, as temperaturas, originalmente contínuas, foram quantizadas linearmente em  $(Q)$  níveis, definidos por intervalos de mesma largura no domínio térmico considerado. Dessa forma, cada valor de temperatura foi associado a um dos  $(Q)$  intervalos, e a probabilidade  $(p_i)$  foi estimada pela razão entre o número de amostras pertencentes ao  $i$ -ésimo nível e o número total de amostras da janela analisada. A entropia térmica foi então calculada por

$$H = - \sum_{i=1}^Q p_i \log_2(p_i). \quad (15)$$

Uma entropia elevada indica maior diversidade de valores térmicos, sugerindo irregularidade, textura térmica complexa ou heterogeneidade espacial. Por outro lado, uma entropia baixa indica maior concentração de valores, associada a superfícies mais homogêneas.

#### 2.3.4 ATRIBUTOS DE VARIAÇÃO ESPACIAL

Os atributos de variação espacial descrevem como a temperatura se modifica ao longo da superfície. Diferentemente dos atributos anteriores, que analisam a distribuição dos valores sem necessariamente considerar a posição dos pixels, os atributos espaciais incorporam a ordem e a vizinhança dos dados. Essa característica é fundamental em

imagens e perfis térmicos, pois fenômenos como bordas térmicas, transições abruptas, falhas, rugosidade superficial e gradientes de aquecimento dependem diretamente da organização espacial das temperaturas.

A inclinação térmica pode ser calculada ao longo de um perfil unidimensional por transformação linear

$$T(x) \approx ax, \quad (16)$$

em que  $a$  representa a inclinação do perfil térmico. Valores positivos indicam aumento de temperatura ao longo da direção analisada, enquanto valores negativos indicam resfriamento. Em termos físicos, uma inclinação acentuada pode indicar fluxo térmico direcional, presença de fonte de calor em uma extremidade, deposição gradual de material ou transição entre zonas com diferentes propriedades térmicas.

Além disso, as diferenças sucessivas representam variações locais entre *pixels* vizinhos, cuja equação é dada por

$$\Delta T_i = T_{i+1} - T_i. \quad (17)$$

Esse atributo permite observar a direção e a intensidade das mudanças térmicas ao longo do perfil. Valores positivos indicam aumento de temperatura entre dois pontos consecutivos, enquanto valores negativos indicam redução. Quando o módulo de  $\Delta T_i$  é elevado, há uma transição térmica local acentuada, possivelmente associada a bordas térmicas, irregularidades superficiais ou mudanças abruptas nas propriedades da região analisada.

Além das diferenças sucessivas propriamente ditas, pode-se calcular o desvio padrão das diferenças sucessivas, que resume a variabilidade dessas mudanças locais, definida como

$$\sigma_{\Delta T} = \sqrt{\frac{1}{L-1} \sum_{i=1}^{L-1} (\Delta T_i - \overline{\Delta T})^2}, \quad (18)$$

em que  $\overline{\Delta T}$  representa a média das diferenças sucessivas. Enquanto  $\Delta T_i$  descreve cada variação local entre pontos vizinhos,  $\sigma_{\Delta T}$  indica o quanto essas variações oscilam ao longo da janela ou do perfil. Assim, valores baixos de  $\sigma_{\Delta T}$  sugerem mudanças térmicas mais regulares e aproximadamente constantes, enquanto valores elevados indicam alternância

intensa entre aumentos e reduções de temperatura, caracterizando maior irregularidade espacial do perfil térmico.

De maneira análoga, a variação total resume a quantidade acumulada de mudança térmica ao longo de um perfil, definida como

$$VT = \sum_{i=1}^{L-1} |T_{i+1} - T_i|. \quad (19)$$

Valores altos de variação total indicam perfis mais irregulares, com alternância frequente entre regiões quentes e frias. Valores baixos indicam transições suaves e maior uniformidade. Esse atributo é particularmente útil para avaliar irregularidade superficial, pois superfícies rugosas, contaminadas ou heterogêneas tendem a produzir maior oscilação térmica.

Ademais, a curvatura média pode ser estimada a partir de diferenças de segunda ordem, cuja equação é dada por

$$C_i = T_{i+1} - 2T_i + T_{i-1}, \quad (20)$$

e sua forma média pode ser definida como

$$\bar{C} = \frac{1}{L-2} \sum_{i=2}^{L-1} |T_{i+1} - 2T_i + T_{i-1}|. \quad (21)$$

A curvatura representa a mudança da inclinação do perfil térmico. Enquanto a primeira diferença indica variação local, a segunda diferença indica mudanças na tendência da variação. Assim, a curvatura média é útil para detectar inflexões térmicas, bordas suaves, regiões de transição e alterações não lineares no perfil.

Por fim, o gradiente máximo representa a maior variação térmica local observada, cuja equação é dada por

$$G_{max} = \max|\nabla T|, \quad (22)$$

em que

$$\nabla T = \frac{dT}{di} \approx T_{i+1} - T_i. \quad (23)$$

Gradientes elevados indicam bordas térmicas, interfaces, transições abruptas ou regiões com contraste espacial intenso.

### 2.3.5 ATRIBUTOS DE DEPENDÊNCIA ESPACIAL

Os atributos de dependência espacial avaliam a relação entre pixels vizinhos ou entre posições separadas por determinada distância. Eles são importantes, pois superfícies térmicas reais geralmente apresentam continuidade espacial: regiões próximas tendem a possuir temperaturas semelhantes quando o campo térmico é suave. Por outro lado, baixa dependência espacial pode indicar textura irregular, ruído, descontinuidades, contaminação não uniforme ou presença de múltiplos mecanismos térmicos atuando simultaneamente.

A autocorrelação espacial mede a semelhança entre o campo térmico e uma versão deslocada dele. Para um perfil unidimensional, pode ser definida como

$$\rho(k) = \frac{\sum_{i=1}^{N-k} (T_i - \mu_T)(T_{i+k} - \mu_T)}{\sum_{i=1}^N (T_i - \mu_T)^2}, \quad (24)$$

em que  $k$  representa o deslocamento espacial. Valores elevados de autocorrelação para pequenos deslocamentos indicam continuidade térmica e suavidade espacial. Valores baixos indicam variação rápida, textura irregular ou perda de organização espacial. Em processamento de imagens, a correlação espacial e a coocorrência entre pixels são fundamentos importantes para a análise de textura e dependência espacial (Milosevic, Jankovic e Peulic, 2014; Gonzalez e Woods, 2018).

## 2.4 SELEÇÃO DOS ATRIBUTOS

A extração de atributos a partir de perfis térmicos produz um conjunto amplo de variáveis candidatas, capazes de representar diferentes propriedades do comportamento térmico, como valores médios, máximos, mínimos, amplitudes, inclinações, variações espaciais, descritores estatísticos, medidas de dispersão e atributos derivados da organização do perfil. Embora esse conjunto inicial seja informativo, nem todos os atributos contribuem de maneira equivalente para a tarefa de classificação. Algumas variáveis podem apresentar baixa associação com a classe-alvo, enquanto outras podem expressar informações semelhantes às já contidas em atributos previamente considerados.

Nesse contexto, a seleção de atributos torna-se uma etapa essencial para reduzir a dimensionalidade, melhorar a interpretabilidade e diminuir o risco de sobreajuste dos modelos (Büyükkeçeci e Okur, 2023).

A seleção de atributos pode ser entendida como o processo de identificar, dentro de um conjunto original de variáveis, um subconjunto que retenha as informações mais importantes para a tarefa de classificação. Em problemas supervisionados, o objetivo não é simplesmente escolher os atributos individualmente mais fortes, mas encontrar um conjunto que, em combinação, represente adequadamente as diferenças entre as classes.

No caso dos atributos térmicos, esse problema é relevante porque muitas variáveis podem ser derivadas de uma mesma região do perfil ou de uma mesma propriedade física do campo térmico. Por exemplo, atributos como temperatura máxima, média, energia, amplitude térmica e determinadas estatísticas de dispersão podem apresentar relações entre si, especialmente quando calculados em janelas sucessivas de um mesmo perfil térmico.

#### 2.4.1 RELEVÂNCIA DOS ATRIBUTOS

A relevância de um atributo está associada à sua capacidade de diferenciar as classes analisadas. Em uma tarefa de classificação supervisionada, um atributo relevante é aquele cujos valores apresentam comportamento distinto entre os grupos, contribuindo para separar as amostras de classes diferentes. Assim, atributos que apresentam distribuições semelhantes entre as classes tendem a contribuir pouco para o processo de classificação, enquanto atributos com maior separação entre os grupos tendem a ser mais informativos.

Uma forma de avaliar a relevância de atributos contínuos é por meio da estatística  $F$  associada à análise de variância. Esse procedimento compara a variabilidade entre os grupos com a variabilidade dentro dos grupos, de modo que valores elevados de  $F$  indicam maior separação entre as classes em relação à dispersão interna de cada classe. Assim, quando utilizado para seleção de atributos, o teste  $F$  permite ranquear cada atributo segundo sua capacidade individual de discriminar as classes do problema. Montgomery e Runger (2018) apresentam a estatística  $F$  da ANOVA como a razão entre o quadrado médio dos tratamentos e o quadrado médio do erro, sendo essa razão utilizada para testar se há diferença significativa entre as médias dos grupos (Montgomery e Runger, 2018):

$$F = \frac{\frac{1}{K-1} \sum_{k=1}^K n_k (\bar{x}_k - \bar{x})^2}{\frac{1}{N-K} \sum_{k=1}^K \sum_{i=1}^{n_k} (x_{ki} - \bar{x}_k)^2}, \quad (25)$$

em que,  $K$  é o número de classes;  $n_k$  é o número de amostras da classe  $k$ ;  $N$  é o número total de amostras;  $x_{ki}$  é o valor do atributo para a  $i$ -ésima amostra da classe  $k$ ;  $\bar{x}_k$  é a média do atributo na classe  $k$ ; e  $\bar{x}$  é a média global do atributo.

#### 2.4.2 REDUNDÂNCIA ENTRE ATRIBUTOS

A relevância isolada de um atributo não garante que ele seja útil quando combinado com outros atributos. Dois atributos podem ser altamente relevantes para a classe e, ainda assim, redundantes entre si. A redundância ocorre quando diferentes variáveis carregam informações semelhantes, de modo que a inclusão de ambas adiciona pouca informação nova ao modelo. Em atributos térmicos, isso pode ocorrer quando descritores diferentes representam a mesma tendência física, como intensidade global de aquecimento, dispersão térmica ou variação espacial do perfil.

A redundância pode ser definida como a similaridade entre um atributo candidato e os atributos já selecionados. Dentre as técnicas de correlação, destaca-se a correlação absoluta média de Pearson entre o atributo candidato e os atributos previamente escolhidos. Dessa forma, seja  $f_i$  um atributo candidato e  $S$  o conjunto de atributos já selecionados, a redundância média pode ser expressa por

$$Red(f_i, S) = \frac{1}{S} \sum_{f_j \in S} |\rho(f_i, f_j)|, \quad (26)$$

em que  $\rho(f_i, f_j)$  representa o coeficiente de correlação de Pearson entre o atributo candidato  $f_i$  e um atributo já selecionado  $f_j$ . O uso do valor absoluto é importante porque correlações negativas fortes também indicam dependência linear elevada. Assim, tanto correlações próximas de  $+1$  quanto de  $-1$  representam alta redundância entre atributos.

Essa formulação é coerente com a ideia central do método *Minimum Redundancy Maximum Relevance* (mRMR), que busca equilibrar atributos relevantes para a classe e pouco redundantes entre si. Embora a formulação original do mRMR utilize informação mútua tanto para relevância quanto para redundância, variações do método podem

empregar outras medidas, como a estatística  $F$  para relevância e a correlação para redundância, visando à maior eficiência computacional ou à adequação ao tipo de dado.

### 2.4.3 MÉTODO MRMR

O método *Minimum Redundancy Maximum Relevance*, ou mRMR, é uma estratégia de seleção de atributos baseada no princípio de que um bom subconjunto deve conter variáveis altamente associadas à classe e, ao mesmo tempo, pouco redundantes entre si. O método foi formalizado por Peng, Long e Ding (2005), que discutiram a relação entre máxima dependência, máxima relevância e mínima redundância para a seleção de atributos baseada em informação mútua (Peng, Long e Ding, 2005). Na formulação clássica, baseada em Informação Mútua, a relevância de um atributo  $f_i$  em relação à classe  $c$  pode ser representada conforme (26), cuja equação pode ser descrita como

$$Rel(f_i) = I(f_i; c). \quad (27)$$

Enquanto a redundância em relação ao conjunto já selecionado  $S$  é representada conforme (27).

O critério por diferença, pode ser escrito como

$$score(f_i) = Rel(f_i) - Red(f_i, S). \quad (28)$$

Já o critério por quociente, pode ser expresso como

$$score(f_i) = \frac{Rel(f_i)}{Red(f_i, S) + \epsilon}, \quad (29)$$

em que  $\epsilon$  é uma constante pequena adicionada para evitar divisão por zero e melhorar a estabilidade numérica.

Dessa forma, o critério por diferença privilegia atributos cuja relevância excede a redundância média em relação ao conjunto já selecionado. Já o critério por quociente favorece atributos que apresentam alta relevância proporcionalmente à redundância. Em ambos os casos, o objetivo é selecionar atributos que tragam informação discriminativa adicional, e não apenas atributos individualmente fortes.

## 2.5 CLASSIFICADORES

A classificação automática de perfis térmicos pode ser entendida como uma etapa de reconhecimento de padrões aplicada ao diagnóstico de isoladores. Nesse contexto, as informações térmicas extraídas de uma imagem ou de um perfil unidimensional de temperatura são convertidas em atributos capazes de representar a condição do componente. O objetivo é estimar, a partir desses atributos, se o isolador pertence à classe normal, associada ao isolador limpo, ou à classe alterada, associada ao isolador poluído ou defeituoso.

Esse problema se enquadra no aprendizado supervisionado, pois o classificador é treinado com exemplos previamente rotulados. Cada amostra do conjunto de treinamento possui uma entrada, formada por atributos térmicos, e uma saída conhecida, correspondente à classe do isolador. A partir desses exemplos, o modelo aprende uma função de decisão que deve ser capaz de generalizar para novas amostras ainda não utilizadas no treinamento (Murphy, 2012; James *et al.*, 2013).

### 2.5.1 CLASSIFICADORES DE ENTRADA DE DIMENSÃO FIXA

Os classificadores de entrada fixa operam sobre vetores de atributos de dimensão fixa. Nessa abordagem, cada isolador é representado por um vetor, cuja equação é dada por

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{ip}]^T, \quad (30)$$

em que,  $\mathbf{x}_i$  é o vetor de atributos da  $i$ -ésima amostra;  $p$  é o número total de atributos; e  $x_{ij}$  representa o  $j$ -ésimo atributo extraído dessa amostra.

A principal característica dessa abordagem é que a sequência original do perfil térmico é convertida em uma representação resumida. Isso permite o uso de modelos clássicos de classificação, com menor custo computacional.

Assim, os classificadores de entrada fixa são adequados para testar a hipótese de que estatísticas globais ou agregadas do perfil térmico são suficientes para distinguir isoladores limpos de isoladores poluídos. Entre os modelos, destacam-se a Máquina de Vetores de Suporte e o Random Forest.

### 2.5.1.1. MÁQUINA DE VETOR DE SUPORTE

A Máquina de Vetores de Suporte (do inglês *Support Vector Machine*, SVM) é um classificador supervisionado originalmente proposto para problemas de classificação em dois grupos. A ideia central é construir uma superfície de decisão que separe as classes com a maior margem possível. Cortes e Vapnik (1995) descrevem a SVM como uma máquina de aprendizado na qual os vetores de entrada são mapeados, possivelmente de forma não linear, para um espaço de características de alta dimensão, no qual uma superfície linear de decisão é construída. Essa formulação é especialmente relevante para problemas com poucos dados, pois o modelo busca uma fronteira de separação com boa capacidade de generalização (Cortes e Vapnik, 1995).

No caso linear, a superfície de decisão é um hiperplano cuja equação é dada por

$$g(x) = \mathbf{w}^T \mathbf{x} + b, \quad (31)$$

em que,  $\mathbf{w}$  é o vetor de pesos;  $\mathbf{x}$  é o vetor de atributos térmicos; e  $b$  é o termo de viés (Cortes e Vapnik, 1995; Haykin, 2009). A classificação é definida pelo sinal da função discriminante

$$\hat{y} = \text{sign}(\mathbf{w}^T \mathbf{x} + b), \quad (32)$$

em que,  $\hat{y}$  é a classe prevista. Se o valor de  $\mathbf{w}^T \mathbf{x} + b$  for positivo, a amostra é atribuída à classe positiva; se for negativo, é atribuída à classe negativa.

Para dados linearmente separáveis, a SVM busca o hiperplano que maximiza a margem entre as classes (Cortes e Vapnik, 1995; Haykin, 2009). Dessa forma, o problema de otimização consiste em encontrar um vetor de pesos  $\mathbf{w}$  e um viés  $b$  que minimize

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2, \quad (33)$$

sujeito a

$$y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1, i = 1, 2, \dots, N. \quad (34)$$

Nessas equações,  $\|\mathbf{w}\|^2$  é a norma euclidiana do vetor de pesos;  $y_i$  é o rótulo da  $i$ -ésima amostra;  $\mathbf{x}_i$  é o vetor de atributos da  $i$ -ésima amostra; e  $b$  é o viés do hiperplano. A

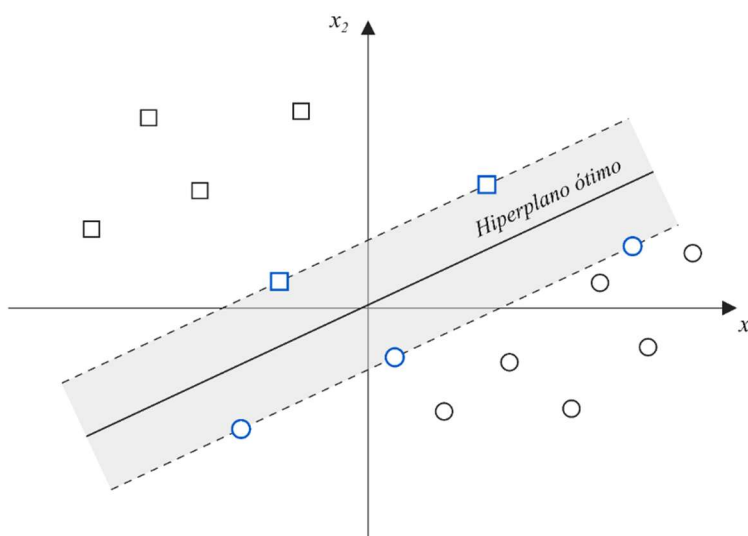
restrição garante que todas as amostras estejam corretamente classificadas e fora da região interna da margem.

Além disso, a largura total da margem é dada por:

$$Margem = \frac{2}{\|w\|}. \quad (35)$$

Dessa forma, quanto menor a norma de  $w$ , maior a margem entre as classes. Os pontos que se encontram sobre as fronteiras da margem são chamados vetores de suporte. Esses pontos são decisivos, pois determinam a posição do hiperplano ótimo. Amostras distantes da margem não alteram a fronteira de decisão, desde que permaneçam corretamente classificadas. Na Figura 4 ilustra-se a ideia de hiperplano ótimo para padrões linearmente separáveis. Os pontos destacados em azul e as linhas tracejadas são, respectivamente, os vetores de suporte e as margens.

Figura 4. Hiperplano ótimo para padrões linearmente separáveis.



Fonte: Adaptado de (Haykin, 2009).

No diagnóstico de perfis térmicos, os vetores de suporte podem ser interpretados como amostras críticas: isoladores cujos atributos térmicos estão próximos da fronteira entre a condição limpa e a condição poluída. Esses casos são importantes porque representam regiões de maior incerteza do diagnóstico.

Em aplicações reais, os dados dificilmente são separáveis linearmente. Perfis térmicos de isoladores limpos e poluídos podem apresentar sobreposição devido a ruídos de medição, estágios intermediários de poluição, variações ambientais ou limitações da extração de atributos. Para lidar com esse cenário, utiliza-se a SVM de margem suave, que introduz variáveis de folga cujo problema de otimização é dado por

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i, \quad (36)$$

sujeito a

$$\begin{aligned} y_i(\mathbf{w}^T x_i + b) &\geq 1 - \xi_i \\ \xi_i &\geq 0, i = 1, 2, \dots, N. \end{aligned} \quad (37)$$

Nessas equações,  $\xi_i$  é a variável de folga associada à  $i$ -ésima amostra;  $C$  é o parâmetro de penalização;  $\mathbf{w}$  e  $b$  definem o hiperplano; e  $y_i$  é a classe verdadeira. A variável  $\xi_i$  mede o quanto uma amostra viola a margem. Quando  $\xi_i = 0$ , a amostra está corretamente posicionada fora ou sobre a margem. Quando  $0 < \xi_i < 1$ , a amostra está dentro da margem, mas ainda do lado correto da fronteira. Quando  $\xi_i > 1$ , a amostra é classificada incorretamente (Cortes e Vapnik, 1995; Haykin, 2009).

O parâmetro  $C$  controla o equilíbrio entre margem larga e penalização de erros. Valores altos de  $C$  penalizam fortemente violações da margem, gerando modelos mais ajustados ao conjunto de treinamento. Valores baixos de  $C$  permitem maior tolerância a erros, favorecendo fronteiras de decisão mais suaves e potencialmente mais generalizáveis (Cortes e Vapnik, 1995; Haykin, 2009). A SVM também pode ser formulada em sua forma dual (Cortes e Vapnik, 1995; Haykin, 2009). Essa formulação é importante porque permite o uso de funções kernel. Dessa forma, o problema de otimização é definido como

$$\max_{\alpha} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j), \quad (38)$$

sujeito a

$$0 \leq \alpha_i \leq C \quad (39)$$

$$\sum_{i=1}^N \alpha_i y_i.$$

Nessas equações,  $\alpha_i$  são os multiplicadores de Lagrange;  $K(x_i, x_j)$  é a função kernel;  $C$  é o parâmetro de regularização; e  $y_i, y_j$  são os rótulos, respectivamente, da  $i$ -ésima e  $j$ -ésima amostra. A função kernel calcula a similaridade entre duas amostras em um espaço de características, sem exigir que esse espaço seja explicitamente construído (Cortes e Vapnik, 1995; Haykin, 2009).

Por fim, a função de decisão da SVM com kernel é dada por

$$g(x) = \sum_{i \in S} \alpha_i y_i K(x_i, x) + b, \quad (40)$$

em que,  $S$  é o conjunto de vetores de suporte. Apenas amostras com  $\alpha_i > 0$  contribuem diretamente para a decisão. Isso torna a SVM um classificador esparsos, pois a fronteira de decisão depende apenas de uma parte dos dados de treinamento.

Entre os kernels, destaca-se, o kernel de base radial, ou RBF:

$$K_{RBF}(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2). \quad (41)$$

Em síntese, no kernel RBF,  $\gamma$  controla a largura da função radial; valores elevados de  $\gamma$  tornam a fronteira mais sensível a variações locais, enquanto valores menores produzem fronteiras mais suaves.

No diagnóstico térmico, o kernel RBF pode ser útil quando a separação entre as classes depende de relações não lineares entre atributos, como temperatura máxima, variação térmica, gradiente espacial e dispersão entre janelas. Já o kernel linear pode ser adequado quando a separação entre isoladores limpos e alterados é aproximadamente linear no espaço de atributos.

A SVM é, portanto, adequada para avaliar se os atributos térmicos extraídos são capazes de formar uma fronteira de separação estável entre as classes. Entretanto, seu desempenho depende da normalização dos atributos e da seleção adequada de hiperparâmetros.

#### 2.5.1.2. RANDOM FOREST

Breiman (2001) define uma floresta aleatória como uma coleção de classificadores estruturados em árvore, na qual cada árvore depende de um vetor aleatório e emite um voto para a classe mais provável. A decisão final é obtida pela agregação dos votos das árvores. O método pertence à família dos *ensembles*, isto é, modelos que combinam vários classificadores individuais (árvores de decisão) para produzir uma escolha coletiva mais robusta (Breiman, 2001). O método também fornece estimativas internas de erro e medidas de importância de variáveis, o que é valioso para interpretar atributos térmicos relevantes.

Em uma árvore de decisão, cada nó divide o espaço de atributos de acordo com uma regra. Para avaliar a qualidade de um nó, calcula-se a distribuição das classes nele presentes, cuja equação é dada por:

$$\hat{p}_{mk} = \frac{1}{N_m} \sum_{z_i \in R_m} I(y_i = k), \quad (42)$$

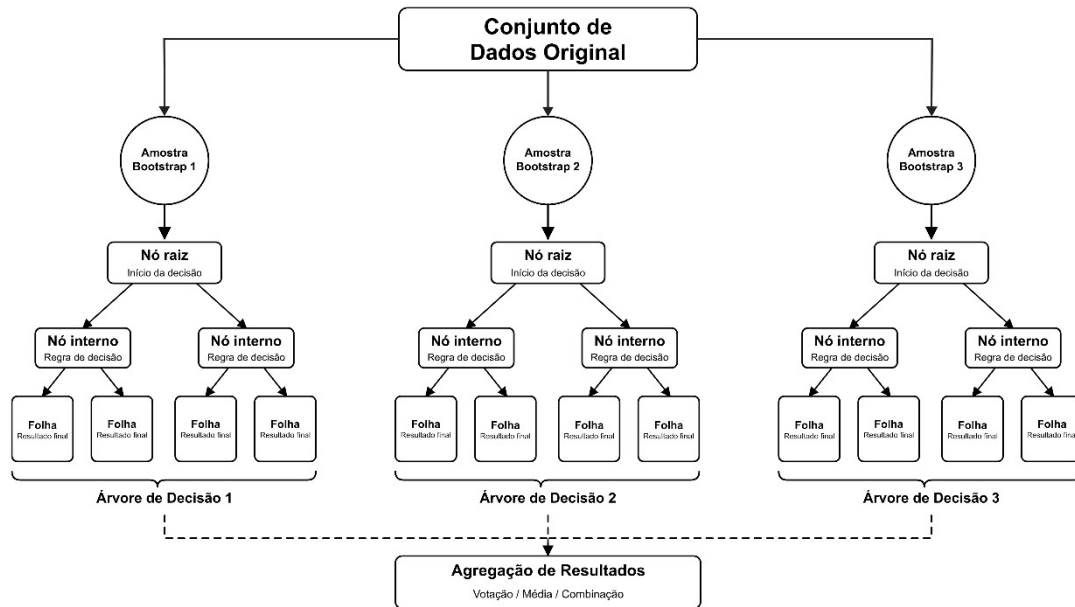
em que,  $\hat{p}_{mk}$  é a proporção de amostras da classe  $k$  no nó  $m$ ;  $N_m$  é o número total de amostras no nó;  $R_m$  é a região do espaço de atributos associada a esse nó; e  $I(\cdot)$  é a função indicadora, que vale 1 quando a condição é verdadeira e 0 caso contrário. A equação calcula, portanto, a composição de classes em um nó. Se quase todas as amostras do nó pertencem à mesma classe, o nó é puro; se as classes estão misturadas, o nó é impuro (Géron, 2019; Raschka e Mirjalili, 2019; Carvalho, 2024).

No Random Forest, a predição final é obtida por votação majoritária entre  $B$  árvores de decisão. Logo, a partir de um vetor aleatório  $z$  a predição final é dada por:

$$\hat{y}_{RF}(z) = \text{mode}\{h_1(z), h_2(z), \dots, h_B(z)\}, \quad (43)$$

Nessa equação,  $h_B(z)$  é a predição da  $b$ -ésima árvore;  $B$  é o número total de árvores; e  $\text{mode}$  retorna a classe mais votada. A interpretação é que cada árvore atua como um classificador individual, e a floresta decide pela maioria. Essa estratégia reduz o efeito de decisões instáveis de árvores isoladas (Géron, 2019; Raschka e Mirjalili, 2019; Carvalho, 2024). Na Figura 5 ilustra-se a estrutura da Random Forest.

Figura 5. Estrutura da Random Forest com três árvores de decisão.



Fonte: autoria própria.

Dessa forma, a votação também permite estimar uma probabilidade empírica de classe:

$$\hat{P}(y = k|z) = \frac{1}{B} \sum_{b=1}^B I(h_b(z) = k). \quad (44)$$

Logo, se 95 de 100 árvores classificam o perfil como defeituoso, a probabilidade estimada para a classe defeituosa é 0,95. Essa saída é útil para diagnóstico porque permite distinguir decisões muito confiantes de decisões próximas da fronteira entre classes.

Como cada árvore é treinada com uma amostra *bootstrap*, parte das amostras não é usada no treinamento daquela árvore. Essas amostras são chamadas *out-of-bag* (OOB). O erro OOB é dado por:

$$Erro_{OOB} = \frac{1}{n} \sum_{i=1}^n I(y_i \neq \hat{y}_{OOB,i}), \quad (45)$$

em que,  $\hat{y}_{OOB,i}$  é a predição da amostra  $i$  usando apenas as árvores que não utilizaram essa amostra no treinamento. O erro OOB funciona como uma estimativa interna de generalização. Em estudos com poucos dados, isso é valioso porque permite avaliar o modelo sem reservar uma grande parcela do banco de dados apenas para validação.

### 2.5.2. CLASSIFICADORES DE ENTRADA SEQUENCIAL

Classificadores sequenciais são utilizados quando a ordem das observações contém informação importantes. No caso de perfis térmicos, a sequência pode representar a posição da distribuição de temperatura ao longo do isolador. Assim, não se avalia apenas quanto o perfil aquece, mas onde esse aquecimento ocorre e como ele evolui ao longo do comprimento analisado.

A representação sequencial pode ser definida como:

$$X = (x_1, x_2, \dots, x_T), \quad (46)$$

em que,  $X$  representa o perfil térmico completo;  $x_T$  é o vetor de atributos da janela  $t$ ; e  $T$  é o número de janelas. A diferença em relação ao vetor de atributos fixos é que a posição  $t$  não é descartada. Isso permite que o classificador aprenda padrões como “aquecimento localizado no início do perfil”, “transição gradual para região quente” ou “pico térmico isolado seguido de retorno à normalidade”.

#### 2.5.2.1. CADEIA DE MARKOV

Segundo Rabiner (1989), um sistema pode ser descrito como estando em 1 de  $N$  estados distintos ( $S_1, S_2, \dots, S_N$ ) a cada instante de tempo. Em tempos discretos, o sistema pode mudar de estado, seguindo um conjunto de probabilidades associadas a essas transições. O estado do sistema no instante  $t$  é denotado por  $q_t$  (Rabiner, 1989).

Para uma descrição probabilística completa do sistema, seria necessário especificar o estado atual e todos os estados anteriores. No entanto, em uma cadeia de Markov discreta de primeira ordem, essa dependência ao passado é implicitamente capturada no estado atual. Assim, o estado futuro é condicionalmente independente dos estados passados, dado o estado presente, o que torna a modelagem mais eficiente e prática, isto é

$$P[q_t = S_j | q_{t-1} = S_i, q_{t-2} = S_k \dots] = P[q_t = S_j | q_{t-1} = S_i] \quad (47)$$

Além disso, são considerados apenas aqueles processos em que o lado direito de (47) é independente do tempo, ou seja, a transição entre estados é uma distribuição condicional entre o estado atual e o anterior, descrita da seguinte forma

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i], 1 \leq i, j \leq N, \quad (48)$$

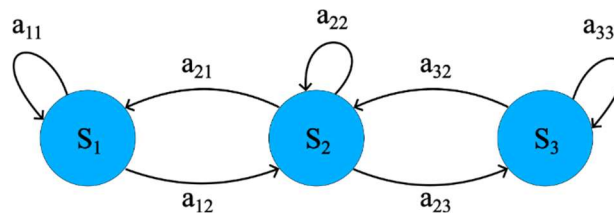
em que  $a_{ij}$  representa a probabilidade de transição do estado  $i$  para o  $j$ , com os coeficientes de transição de estados possuindo as seguintes propriedades

$$a_{ij} \geq 0;$$

$$\sum_{j=1}^N a_{ij} = 1. \quad (49)$$

O processo descrito é conhecido como um modelo de Markov observável, pois a saída do sistema consiste no conjunto de estados a cada instante de tempo. Nesse caso, cada estado está diretamente associado a um evento físico que pode ser observado, tornando o comportamento do sistema totalmente acessível por meio de suas transições de estado. Na Figura 6 ilustra-se uma cadeia de Markov com 3 estados e transições selecionadas.

Figura 6. Cadeia de Markov com 3 estados e transições selecionadas.



Fonte: adaptado de (Rabiner, 1989).

#### 2.5.2.2. MODELO OCULTO DE MARKOV

Os Modelos Ocultos de Markov, ou HMMs, estendem as Cadeias de Markov ao considerar que os estados que governam a sequência não são observados diretamente. Observa-se apenas uma sequência de sinais ou atributos produzidos por esses estados ocultos. Rabiner (1989) apresenta os HMMs como modelos estatísticos adequados para processos que geram sequências observáveis, que podem ser discretas ou contínuas, e organiza sua teoria em três problemas fundamentais: avaliação da probabilidade da sequência, estimação da sequência de estados e ajuste dos parâmetros do modelo (Rabiner, 1989).

Um HMM é definido por

$$\lambda = (A, B, \pi), \quad (50)$$

em que,  $\lambda$  representa o modelo completo;  $A$  é a matriz de probabilidade de transição entre estados ocultos;  $B$  representa as probabilidades ou densidades de emissão de estados; e  $\pi$  é a distribuição inicial dos estados. Essa equação é compacta, mas resume toda a estrutura probabilística do HMM. No perfil térmico,  $A$  descreve como regimes térmicos latentes se sucedem;  $B$  descreve como cada regime gera observações térmicas; e  $\pi$  descreve como a sequência tende a iniciar.

Ademais, a matriz de transição é definida como

$$A = \{a_{ij}\}, \quad (51)$$

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i).$$

em que,  $a_{ij}$  representa a probabilidade de o estado oculto passar de  $S_i$  para  $S_j$ . A diferença em relação à Cadeia de Markov observável é que esses estados não são diretamente medidos. Eles representam regimes latentes que explicam os atributos térmicos observados. Assim, uma sequência de janelas pode ser interpretada como resultado de uma trajetória oculta de estados térmicos (Rabiner, 1989).

Além disso, a distribuição inicial é dada por

$$\pi = \{\pi_i\}, \quad (52)$$

$$\pi_i = P(q_1 = S_i).$$

Essa expressão define a probabilidade de início em cada estado oculto. Em um HMM treinado para isoladores normais, pode ser esperado que  $\pi_i$  favoreça estados associados à normalidade térmica. Em um HMM treinado para defeitos, a distribuição inicial pode favorecer estados mais variáveis, dependendo da posição em que as alterações aparecem (Rabiner, 1989).

Por fim, para observações discretas, a matriz de emissão é definida por

$$B = \{b_j\}, \quad (53)$$

$$b_j(k) = P(o_t = v_k | q_t = S_j),$$

em que,  $o_t$  é a observação na posição  $t$ ;  $v_k$  é o  $k$ -ésimo símbolo observável; e  $b_j(k)$  é a probabilidade de o estado oculto  $S_j$  emitir esse símbolo. Se os atributos térmicos forem discretizados em categorias, como baixo, médio e alto aquecimento,  $B$  descreve quais categorias são mais prováveis em cada estado oculto.

Ademais, a probabilidade conjunta de uma sequência observada  $O$  e uma sequência oculta  $Q$  é dada por

$$P(O, Q | \lambda) = \pi_{q_1} b_{q_1}(o_1) \prod_{t=2}^T a_{q_{t-1}, q_t} b_{q_t}(o_t). \quad (54)$$

Essa equação combina três componentes: a probabilidade de iniciar no primeiro estado, a probabilidade de emitir a primeira observação e, para cada etapa seguinte, a probabilidade de transitar para o próximo estado e emitir a observação correspondente, em toda sequência de tamanho  $T$ . O termo  $b_{q_t}(o_t)$  conecta o estado oculto à observação (Rabiner, 1989). No diagnóstico, uma sequência terá alta probabilidade se suas transições e emissões forem compatíveis com o padrão aprendido para uma classe.

Por fim, a probabilidade da sequência observada é obtida somando todas as trajetórias ocultas possíveis, cuja equação é dada por

$$P(O | \lambda) = \sum_Q P(O, Q | \lambda). \quad (55)$$

Essa equação é conceitualmente importante porque os estados ocultos não são observados. Portanto, não se sabe qual trajetória  $Q$  produziu  $O$ . A solução probabilística é somar a contribuição de todas as trajetórias possíveis. O problema é que o número de trajetórias cresce exponencialmente com o comprimento  $T$ , tornando a soma direta inviável para sequências longas.

A partir do problema citado, Rabiner (1989) definiu o algoritmo *Forward* para a solução do primeiro problema (Rabiner, 1989). Esse algoritmo pode ser utilizado para quantificar a verossimilhança de uma sequência de observação, dado o Modelo Oculto de Markov, a partir da variável a seguir

$$\alpha_t(i) = P(O_1, O_2, \dots, O_t, q_t = S_i | \lambda). \quad (56)$$

Essa variável representa a probabilidade de observar a sequência parcial até  $t$  e terminar no estado  $S_i$ . Ela é acumulativa, resumindo todos os caminhos que poderiam ter

chegado ao estado  $S_i$  na posição  $t$ . Por isso, permite evitar a enumeração explícita de todas as trajetórias ocultas.

Para isso, o processo de inicialização é definido como

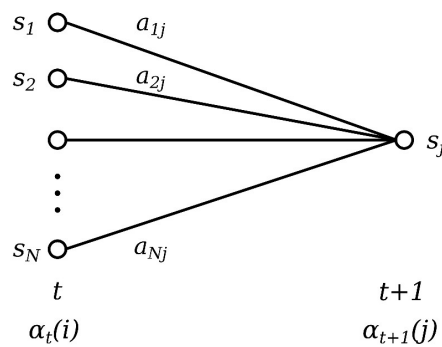
$$\alpha_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N. \quad (57)$$

Essa equação multiplica a probabilidade de iniciar no estado  $S_i$  pela probabilidade desse estado emitir a primeira observação. Ela estabelece o ponto de partida da recursão do algoritmo *Forward*, cuja equação é dada por

$$\alpha_{t+1}(j) = \left[ \sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(O_{t+1}), \quad 1 \leq t \leq T-1, 1 \leq j \leq N, \quad (58)$$

A soma dentro dos colchetes agrega todas as formas de chegar ao estado  $S_j$  na etapa  $t+1$ , vindo de qualquer estado anterior  $S_i$ . Depois, multiplica-se por  $b_j(O_{t+1})$ , que é a probabilidade de emitir a nova observação a partir do estado  $S_j$ . Essa equação mostra como o HMM atualiza a probabilidade da sequência à medida que novas janelas são processadas. Na Figura 7 ilustra-se o passo de indução, etapa principal do algoritmo *forward*.

Figura 7. Etapa de indução do algoritmo *forward*.



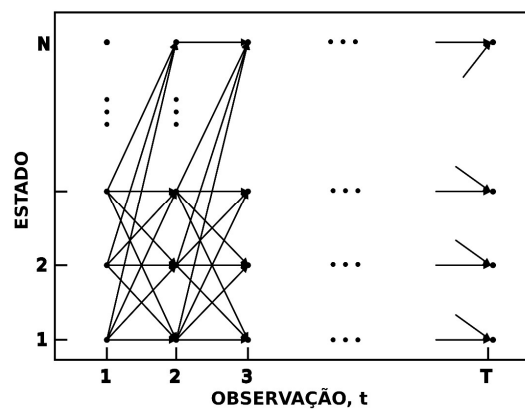
Fonte: adaptado de (Rabiner, 1989).

A probabilidade final da sequência é dada por

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i). \quad (59)$$

Esse método evita a necessidade de explorar todas as combinações possíveis de estados, tornando o processo computacionalmente mais eficiente. No entanto, devido à multiplicação sucessiva de probabilidades, os valores resultantes tendem a se aproximar de zero, especialmente em sequências longas, levando a problemas de subestimação numérica. Para contornar essa limitação, utiliza-se o logaritmo e suas propriedades, o que permite calcular o logaritmo da verossimilhança, mantendo a estabilidade numérica e facilitando a interpretação dos resultados (Rabiner, 1989). Na Figura 8 ilustra-se a estrutura de treliça para o cálculo da probabilidade *forward*.

Figura 8. Probabilidade *forward* baseado na estrutura de treliça.



Fonte: adaptado de (Rabiner, 1989).

De maneira análoga, uma variável *backward* é definida como

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, \dots, o_T | q_t = S_i, \lambda). \quad (60)$$

Essa variável mede a probabilidade de observar o restante da sequência, do tempo  $t + 1$  até  $T$ , dado que o estado atual é  $S_i$  (Rabiner, 1989). Enquanto  $\alpha$  resume o passado da sequência,  $\beta$  resume o futuro. As duas variáveis são complementares e permitem inferir probabilidades posteriores dos estados.

A inicialização é definida como

$$\beta_T(i) = 1. \quad (61)$$

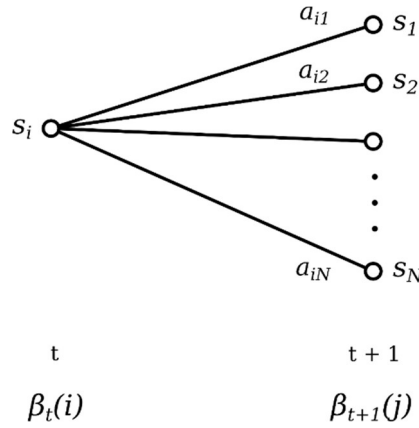
No último tempo, não há observações futuras a serem explicadas. Por isso, a probabilidade do “restante vazio” da sequência é definida como 1.

Ademais, a recursão é dada por

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(o_{t+1}) \beta_{t+1}(j) \quad (62)$$

Na Figura 9 ilustra-se a etapa de recursão do algoritmo *backward*.

Figura 9. Etapa de recursão do algoritmo *Backward*.



Fonte: adaptado de (Rabiner, 1989).

A partir da definição de  $\alpha$  e  $\beta$ , define-se uma variável cuja equação é dada por

$$\gamma_t(i) = P(q_t = S_i | O, \lambda). \quad (63)$$

$$\gamma_t(j, k) = \left[ \frac{\alpha_t(j) \beta_t(j)}{\sum_{j=1}^N \alpha_t(j) \beta_t(j)} \right]. \quad (64)$$

Nessa equação, o numerador combina a probabilidade de explicar o passado até  $t$  e o futuro após  $t$  passando pelo estado  $S_i$ . O denominador normaliza pela probabilidade total da sequência. Assim,  $\gamma_t(i)$  indica o quanto o estado  $S_i$  é provável na janela  $t$ , considerando todo o sinal. Essa quantidade pode ser usada para interpretar quais regiões do perfil térmico são associadas a regimes térmicos latentes.

Por fim, a partir das definições das variáveis  $\alpha$  e  $\beta$ , pode-se escrever  $\xi_t(i, j)$  na forma

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)},$$

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}.$$
(65)

Essa equação mede a probabilidade de estar no estado  $S_i$  em  $t$  e transitar para  $S_j$  em  $t + 1$ , dado todo o sinal observado. O numerador representa esse caminho específico. O denominador soma todos os caminhos de transição possíveis entre  $t$  e  $t + 1$ , garantindo normalização. Essa variável é essencial no treinamento do HMM, pois permite estimar quantas transições de cada tipo ocorreram em média.

Dessa forma, a partir do algoritmo de Baum-Welch (Rabiner, 1989) reestima-se a distribuição inicial ( $\bar{\pi}_i$ ) por

$$\bar{\pi}_i = \gamma_1(i).$$
(66)

Essa atualização diz que a nova probabilidade inicial do estado  $S_i$  é igual à probabilidade posterior de estar nesse estado no primeiro instante. Em outras palavras, o modelo ajusta sua crença inicial com base na inferência feita sobre os dados observados.

Além disso, a matriz de transição é reestimada conforme a equação

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i, j)}.$$
(67)

O numerador representa o número esperado de transições de  $S_i$  para  $S_j$ . Enquanto o denominador representa o número esperado de vezes em que o modelo esteve em  $S_i$  e poderia sair dele.

Para observações discretas, as emissões são reestimadas por

$$\bar{b}_j(k) = \frac{\sum_{t=1:O_t=v_k}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}.$$
(68)

#### 2.5.2.3. MODELO OCULTO DE MARKOV COM EMISSÕES GAUSSIANAS

Até este ponto, a modelagem por Modelos Ocultos de Markov tem sido frequentemente apresentada a partir de observações representadas por símbolos discretos

pertencentes a um alfabeto finito, o que permite associar a cada estado uma distribuição de probabilidade discreta. No entanto, em diversas aplicações reais, as observações assumem a forma de sinais ou vetores contínuos, tornando essa formulação limitada. Embora seja possível converter sinais contínuos em símbolos discretos por meio de técnicas de quantização, esse processo pode acarretar perda significativa de informação e comprometer o desempenho do modelo. Nesse contexto, torna-se relevante empregar HMM com densidades de observação contínuas, capazes de representar de forma mais adequada a variabilidade dos dados. Para isso, é necessário impor restrições à forma da função densidade de probabilidade associada às observações, de modo que seus parâmetros possam ser reestimados de maneira consistente. Uma das formulações mais utilizadas para esse fim é a representação da função densidade de probabilidade por meio de misturas finitas, que oferecem flexibilidade para modelar distribuições complexas em problemas envolvendo sinais contínuos (Rabiner, 1989), cuja equação é dada por

$$b_j(O) = \sum_{m=1}^M c_{jm} \mathfrak{R}[O, \mu_{jm}, U_{jm}], \quad 1 \leq j \leq N, \quad (69)$$

em que,  $O$  é o vetor de observações a ser modelado,  $c_{jm}$  é o coeficiente de mistura (ou ganho) para  $m$ -ésima componente de mistura no estado  $j$  e  $\mathfrak{R}$  é qualquer densidade log-côncava ou elipiticamente simétrica (por exemplo, Gaussiana), com vetor de média  $\mu_{jm}$  e matriz de covariância  $U_{jm}$  para  $m$ -ésima componente de mistura do estado  $j$ . Geralmente, uma densidade gaussiana é usada para  $\mathfrak{R}$ . Além disso, o ganho de mistura  $c_{jm}$  satisfaz as seguintes restrições estocásticas:

$$\begin{aligned} \sum_{m=1}^M c_{jm} &= 1, \quad 1 \leq j \leq N; \\ c_{jm} &\geq 0, \quad 1 \leq j \leq N, 1 \leq m \leq M. \end{aligned} \quad (70)$$

Por fim, as equações de reestimação dos parâmetros associados à densidade de mistura, isto é, os coeficientes de ponderação ( $\bar{c}_{jk}$ ), os vetores médios ( $\bar{\mu}_{jk}$ ) e as matrizes de covariância ( $\bar{U}_{jk}$ ), assumem a seguinte forma:

$$\bar{c}_{jk} = \frac{\sum_{t=1}^T \gamma_t(j, k)}{\sum_{t=1}^T \sum_{k=1}^M \gamma_t(j, k)}. \quad (71)$$

$$\bar{\mu}_{jk} = \frac{\sum_{t=1}^T \gamma_t(j, k) * O_t}{\sum_{t=1}^T \gamma_t(j, k)}. \quad (72)$$

$$\bar{U}_{jk} = \frac{\sum_{t=1}^T \gamma_t(j, k) * (O_t - \mu_{jk}) * (O_t - \mu_{jk})'}{\sum_{t=1}^T \gamma_t(j, k)}. \quad (73)$$

Em que:

$$\gamma_t(j, k) = \left[ \frac{\alpha_t(j)\beta_t(j)}{\sum_{j=1}^N \alpha_t(j)\beta_t(j)} \right] \left[ \frac{c_{jk} \Re[O, \mu_{jm}, U_{jm}]}{\sum_{m=1}^M c_{jk} \Re[O, \mu_{jm}, U_{jm}]} \right]. \quad (74)$$

O termo  $\gamma_t(j, k)$  representa uma generalização de  $\gamma_t(j)$ , utilizado no caso de misturas simples ou de densidades discretas, passando a considerar não apenas a probabilidade do sistema estar no estado  $j$  no instante  $t$ , mas também a contribuição do  $k$ -ésimo componente da mistura nesse estado. A fórmula de reestimação do parâmetro de transição  $a_{ij}$  permanece igual àquela empregada para densidades de observação discretas. A interpretação das equações de reestimação é direta: o coeficiente  $c_{jk}$  é atualizado pela razão entre o número esperado de vezes em que o sistema se encontra no estado  $j$  utilizando o  $k$ -ésimo componente da mistura e o número esperado total de vezes em que o sistema se encontra no estado  $j$ . De forma semelhante, a reestimação do vetor médio  $\mu_{jk}$  pondera cada termo do numerador pela observação correspondente, resultando no valor esperado da parte do vetor de observação explicada pelo  $k$ -ésimo componente da mistura. Por fim, a matriz de covariância  $U_{jk}$  é reestimada de modo análogo, descrevendo a dispersão esperada das observações em relação ao vetor médio correspondente ao componente  $k$  no estado  $j$  (Rabiner, 1989).

#### 2.5.2.4. REDE LSTM

A arquitetura *Long Short-Term Memory* (LSTM), é uma extensão das redes neurais recorrentes (RNNs) desenvolvida para superar o problema de gradientes explosivos ou dissipados que prejudica a aprendizagem de dependências de longo prazo em sequências. A LSTM incorpora um mecanismo explícito de controle de fluxo de informações por meio das chamadas portas (*gates*), permitindo regular, de forma adaptativa, o que deve ser lembrado, atualizado ou descartado ao longo do tempo

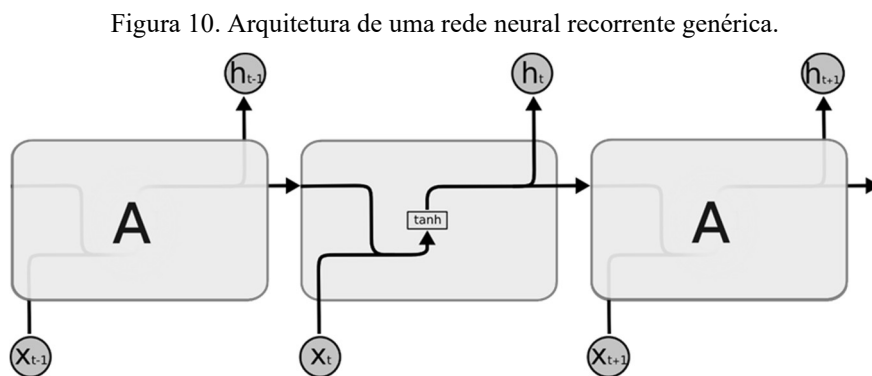
(Hochreiter, 1997). Revisões posteriores destacam a LSTM como uma arquitetura recorrente amplamente aplicada a séries temporais e dados sequenciais, justamente por controlar o fluxo de informação por meio de células de memória e portas (Houdt, Mosquera e Nápoles, 2020).

A atualização do estado oculto para uma rede recorrente simples é definida como:

$$h_t = \phi(W_{xh}x_t + W_{hh}h_{t-1} + b_h), \quad (75)$$

em que,  $x_t$  é o vetor de entrada na janela  $t$ ;  $h_t$  é o estado oculto atual;  $h_{t-1}$  é o estado oculto anterior;  $W_{xh}$  conecta a entrada ao estado oculto;  $W_{hh}$  conecta o estado anterior ao atual;  $b_h$  é o viés; e  $\phi$  é uma função de ativação. A equação mostra que a rede combina informação atual com memória anterior. Porém, ao propagar gradientes por muitos passos, produtos sucessivos podem fazer o sinal de erro desaparecer ou explodir.

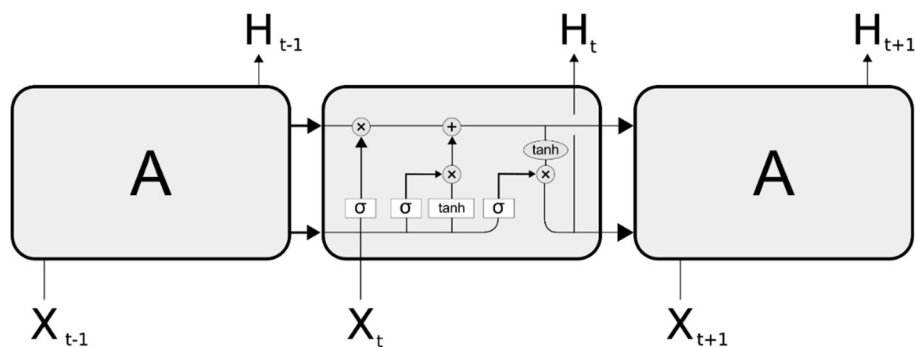
Todas as redes neurais recorrentes têm a forma de uma cadeia de módulos repetidos. Em RNNs padrão, esse módulo repetido apresenta uma estrutura muito simples, como uma única camada de ativação. Na Figura 10 ilustra-se a arquitetura de uma rede neural recorrente genérica.



Fonte: autoria própria.

As LSTMs também possuem essa estrutura em cadeia, mas o módulo repetido apresenta uma estrutura diferente. Em vez de uma única camada de rede neural, há quatro camadas, que interagem de uma maneira muito específica (Hochreiter, 1997; Houdt, Mosquera e Nápoles, 2020). Na Figura 11 ilustram-se as camadas da arquitetura de rede LSTM.

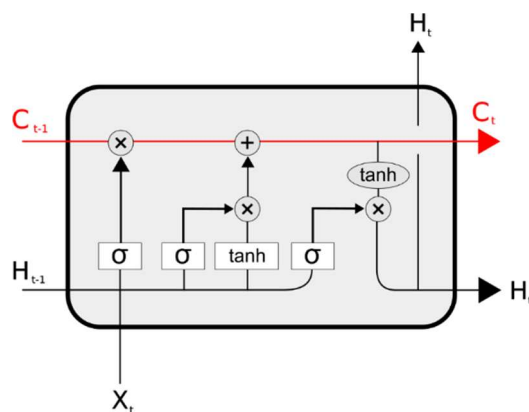
Figura 11. Camadas da arquitetura de rede LSTM.



Fonte: autoria própria

A ideia principal para as LSTMs é o estado da célula, representado pela linha horizontal que atravessa a parte superior do diagrama, ilustrado na Figura 12.

Figura 12. Estado de célula da arquitetura LSTM.



Fonte: autoria própria.

O estado da célula funciona como uma esteira rolante. Ele percorre toda a cadeia em linha reta, com apenas algumas interações lineares pontuais. Assim, torna-se muito fácil para a informação fluir por essa estrutura sem sofrer alterações.

A LSTM possui a capacidade de remover ou adicionar informações ao estado da célula, processo cuidadosamente regulado por estruturas chamadas *gates* (Hochreiter, 1997; Houdt, Mosquera e Nápoles, 2020).

Os *gates* são mecanismos que permitem, de forma opcional, a passagem de informações. Eles são compostos por uma camada de rede neural sigmoide e por uma operação de multiplicação ponto a ponto.

A camada sigmoide gera valores entre zero e um, indicando a quantidade de cada componente que deve ser permitida a passar. Valores próximos de 0 bloqueiam a entrada de informação; valores próximos de 1 permitem sua passagem.

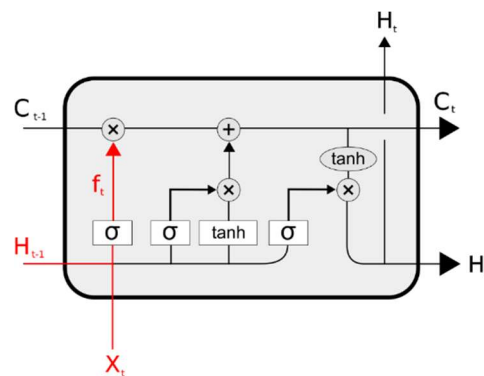
Uma LSTM possui três desses *gates*, responsáveis por proteger e controlar o estado da célula.

A primeira etapa de uma rede LSTM consiste em determinar quais informações serão descartadas do estado da célula (Hochreiter, 1997; Houdt, Mosquera e Nápoles, 2020). Essa decisão é realizada por uma camada sigmoide denominada porta de esquecimento (*forget gate*), cuja equação é dada por:

$$f_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f), \quad (76)$$

em que,  $\sigma$  é a função sigmoide, responsável por limitar os valores da porta ao intervalo entre 0 e 1. Assim, ela permite interpretar a saída como um grau de retenção da informação.  $\mathbf{W}_f$  é a matriz de pesos associada à porta de esquecimento. Esses pesos são parâmetros aprendidos durante o treinamento da rede e determinam a importância das informações de entrada e do estado oculto anterior.  $[\mathbf{h}_{t-1}, \mathbf{x}_t]$  representa a concatenação entre o estado oculto anterior e a entrada atual. O termo  $\mathbf{h}_{t-1}$  corresponde ao estado oculto produzido no instante anterior ( $t - 1$ ), contendo informações já processadas pela LSTM. Já  $\mathbf{x}_t$  representa a entrada recebida pela rede no instante atual  $t$ .  $\mathbf{b}_f$  é o vetor de viés da porta de esquecimento, também aprendido durante o treinamento. Ele permite ajustar a ativação da porta, contribuindo para o controle do quanto a informação anterior será mantida (Hochreiter, 1997; Houdt, Mosquera e Nápoles, 2020). Na Figura 13 ilustra-se a porta de esquecimento na arquitetura de rede LSTM

Figura 13. Porta de esquecimento na arquitetura de rede LSTM.



Fonte: autoria própria.

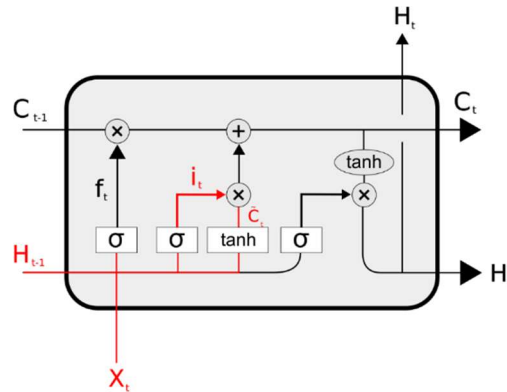
A etapa seguinte consiste em determinar quais novas informações serão armazenadas no estado da célula. Esse processo ocorre em duas partes. Inicialmente, uma camada sigmoide, denominada porta de entrada (*input gate*), define quais valores serão atualizados. Em seguida, uma camada com função de ativação, tangente hiperbólica, gera um vetor de novos valores candidatos ( $\tilde{C}_t$ ) que podem ser incorporados ao estado da célula (Hochreiter, 1997; Houdt, Mosquera e Nápoles, 2020). Posteriormente, essas duas informações são combinadas para produzir a atualização do estado, e suas equações são definidas como:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \quad (77)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c). \quad (78)$$

em que,  $i_t$  representa a porta de entrada no instante  $t$ , responsável por controlar quais novas informações serão incorporadas ao estado da célula da LSTM. A função  $\sigma$  corresponde à função sigmoide, que limita os valores da porta ao intervalo entre 0 e 1, indicando o grau de permissão para a entrada de novas informações. O termo  $\tilde{C}_t$  representa o estado candidato da célula, isto é, o conjunto de novas informações que podem ser adicionadas à memória da rede. A função  $\tanh$  é a tangente hiperbólica, que limita os valores do estado candidato ao intervalo entre -1 e 1. Os termos  $W_i$  e  $W_c$  são, respectivamente, as matrizes de pesos da porta de entrada e do estado candidato da célula, enquanto  $b_i$  e  $b_c$  são, respectivamente, os vetores de viés associados a essas operações. Já  $[h_{t-1}, x_t]$  representa a concatenação entre o estado oculto anterior  $h_{t-1}$ , que contém informações processadas até o instante  $t - 1$ , e a entrada atual  $x_t$ , correspondente aos dados recebidos pela rede no instante  $t$  (Hochreiter, 1997; Houdt, Mosquera e Nápoles, 2020). Na Figura 14 ilustra-se a porta de entrada e os vetores candidatos na arquitetura de LSTM.

Figura 14. Porta de entrada e vetores candidatos na arquitetura de rede LSTM.



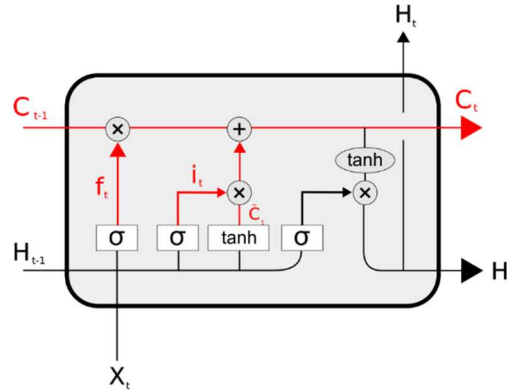
Fonte: autoria própria.

Em seguida, o estado da célula anterior ( $C_{t-1}$ ), é atualizado para formar o novo estado da célula, ( $C_t$ ). As etapas anteriores já determinaram quais informações devem ser mantidas, descartadas ou adicionadas; resta, portanto, efetivar essa atualização. Para isso, é definida a seguinte equação:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \quad (79)$$

em que,  $C_t$  representa o estado da célula no instante  $t$ , isto é, a memória atualizada da LSTM. O termo  $f_t$  corresponde à porta de esquecimento, responsável por controlar quais informações do estado anterior devem ser mantidas ou descartadas. Já  $C_{t-1}$  representa o estado da célula no instante anterior. O operador  $\odot$  indica a multiplicação elemento a elemento, aplicada entre vetores de mesma dimensão. O termo  $i_t$  representa a porta de entrada, que controla a quantidade de nova informação a ser incorporada à memória da célula. Por fim,  $\tilde{C}_t$  corresponde ao estado candidato da célula, contendo as novas informações que podem ser adicionadas ao estado atual (Hochreiter, 1997; Houdt, Mosquera e Nápoles, 2020). Na Figura 15 ilustra-se o novo estado da célula na arquitetura de rede LSTM.

Figura 15. Novo estado da célula na arquitetura de rede LSTM.



Fonte: autoria própria.

Por fim, é necessário determinar qual informação será produzida como saída. Essa saída é baseada no estado da célula, porém corresponde a uma versão filtrada desse estado. Dessa forma, define-se as seguintes equações:

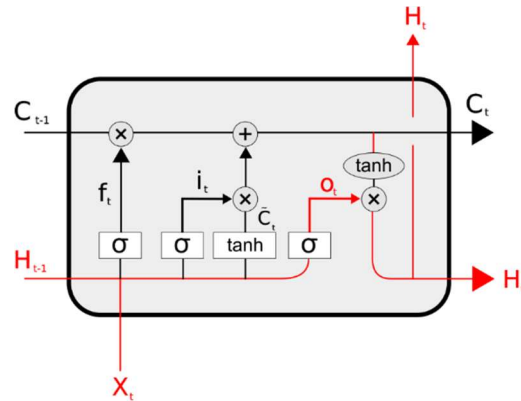
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \quad (80)$$

$$h_t = o_t \odot \tanh C_t. \quad (81)$$

em que,  $o_t$  representa a porta de saída no instante  $t$ , responsável por controlar quais informações do estado da célula serão utilizadas para formar a saída da LSTM. A função  $\sigma$  corresponde à função sigmoide, que limita os valores da porta ao intervalo entre 0 e 1, indicando o grau de liberação das informações armazenadas na célula. O termo  $W_o$  representa a matriz de pesos da porta de saída, enquanto  $b_o$  corresponde ao vetor de vies associado a essa porta, ambos aprendidos durante o treinamento da rede. Já  $[h_{t-1}, x_t]$  representa a concatenação entre o estado oculto anterior  $h_{t-1}$ , que contém informações processadas até o instante anterior, e a entrada atual  $x_t$ , correspondente aos dados recebidos pela rede no instante  $t$ .

Além disso,  $h_t$  representa o estado oculto no instante  $t$ , ou seja, a saída produzida pela LSTM após a atualização da memória. O operador  $\odot$  indica a multiplicação elemento a elemento. O termo  $\tanh(C_t)$  corresponde à aplicação da função tangente hiperbólica sobre o estado atual da célula  $C_t$ , limitando seus valores ao intervalo entre -1 e 1. Assim, o estado oculto  $h_t$  é obtido pela combinação entre as informações liberadas pela porta de saída  $o_t$  e o conteúdo transformado do estado da célula  $C_t$ . Na Figura 16 ilustra-se a saída da rede LSTM.

Figura 16. Saída da rede LSTM.



Fonte: autoria própria.

Após a definição das equações que descrevem o funcionamento interno da LSTM, o treinamento da rede consiste no ajuste dos parâmetros associados às portas de esquecimento, entrada e saída, bem como ao vetor candidato da célula. Esses parâmetros correspondem às matrizes de pesos  $W_f$ ,  $W_i$ ,  $W_c$  e  $W_o$ , além dos respectivos vieses  $b_f$ ,  $b_i$ ,  $b_c$  e  $b_o$ , apresentados nas equações (76) - (81).

## 2.6 MÉTRICAS DE AVALIAÇÃO DOS MODELOS

A avaliação de modelos de classificação exige métricas capazes de quantificar não apenas a proporção de acertos, mas também o tipo de erro cometido pelo classificador. Em problemas de diagnóstico, essa distinção é importante porque falsos positivos e falsos negativos podem ter consequências práticas diferentes.

### 2.6.1 MATRIZ DE CONFUSÃO

A matriz de confusão é uma representação tabular que sintetiza a relação entre as classes reais e as classes previstas por um classificador. Ela permite visualizar, de forma organizada, quantas amostras foram corretamente classificadas e quantas foram confundidas entre as classes. Diferentemente de uma métrica única, como a acurácia, a matriz de confusão explicita os tipos de erro cometidos pelo modelo, sendo, portanto, uma ferramenta central para a análise diagnóstica de classificadores.

Em um problema de classificação binária, a matriz de confusão organiza os resultados em quatro categorias fundamentais: verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos. Considerando, neste trabalho, a classe

positiva como “poluído” e a classe negativa como “limpo”, a matriz pode ser representada da seguinte forma:

Tabela 1. Matriz de confusão.

|                    | Predito Limpo | Predito Poluído |
|--------------------|---------------|-----------------|
| Verdadeiro Limpo   | VN            | FP              |
| Verdadeiro Poluído | FN            | VP              |

Fonte: autoria própria.

em que (VP) representa os verdadeiros positivos, isto é, amostras poluídas corretamente classificadas como poluídas; (VN) representa os verdadeiros negativos, isto é, amostras limpas corretamente classificadas como limpas; (FP) representa os falsos positivos, isto é, amostras limpas classificadas incorretamente como poluídas; e (FN) representa os falsos negativos, isto é, amostras poluídas classificadas incorretamente como limpas.

A matriz de confusão é a base para o cálculo das demais métricas de desempenho. Sua interpretação é especialmente relevante em aplicações diagnósticas, pois permite identificar qual tipo de erro predomina no modelo. Um número elevado de falsos negativos indica que o classificador deixa de detectar isoladores poluídos, o que pode representar risco operacional. Por outro lado, um número elevado de falsos positivos indica que o modelo gera muitos alarmes indevidos, podendo resultar em inspeções desnecessárias e perda de confiança no sistema de diagnóstico.

## 2.6.2 MÉTRICAS DE DESEMPENHO

A acurácia mede a proporção total de classificações corretas realizadas pelo modelo. Essa métrica considera simultaneamente os verdadeiros positivos e os verdadeiros negativos em relação ao total de amostras avaliadas, sendo definida por:

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN}, \quad (82)$$

A acurácia fornece uma medida geral do desempenho do classificador. Entretanto, sua interpretação deve ser feita com cautela quando há desbalanceamento entre as classes. Nesses casos, um modelo pode apresentar acurácia elevada simplesmente por favorecer a classe majoritária, mesmo que apresente baixo desempenho na classe minoritária. Por isso, a acurácia não deve ser utilizada isoladamente em problemas de diagnóstico.

A sensibilidade, também denominada taxa de verdadeiros positivos ou revocação, mede a capacidade do modelo de identificar corretamente as amostras pertencentes à classe positiva. No contexto desta dissertação, representa a capacidade de detectar corretamente isoladores poluídos. Essa métrica é definida por:

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \cdot \quad (83)$$

Valores elevados de sensibilidade indicam que o modelo apresenta baixa ocorrência de falsos negativos. Em termos diagnósticos, essa métrica está associada à segurança da detecção, pois modelos com baixa sensibilidade podem deixar de identificar condições anormais. Entretanto, a sensibilidade não informa a quantidade de falsos positivos gerados; por isso, deve ser analisada em conjunto com a especificidade e a precisão.

A especificidade, também conhecida como taxa de verdadeiros negativos, mede a capacidade do modelo de identificar corretamente as amostras pertencentes à classe negativa. Neste trabalho, representa a capacidade do classificador de reconhecer corretamente isoladores limpos, sendo definida por:

$$\text{Especificidade} = \frac{VN}{VN + FP} \cdot \quad (84)$$

Essa métrica é importante porque avalia a tendência do modelo a gerar falsos alarmes. Uma baixa especificidade indica que o classificador atribui indevidamente a classe poluída a muitos isoladores limpos. Em aplicações de manutenção, esse comportamento pode aumentar custos operacionais e reduzir a confiabilidade do sistema de apoio ao diagnóstico.

A partir da sensibilidade e da especificidade, define-se a acurácia balanceada, dada por:

$$\text{Acurácia balanceada} = \frac{\text{Sensibilidade} + \text{Especificidade}}{2} \cdot \quad (85)$$

A acurácia balanceada avalia se o modelo apresenta desempenho equilibrado entre as classes. Diferentemente da acurácia convencional, que pode ser influenciada pela classe majoritária, a acurácia balanceada penaliza classificadores que apresentam bom desempenho em uma classe e desempenho reduzido na outra.

Além disso, a precisão mede a proporção de predições positivas que estão corretas. Essa métrica é definida por:

$$Precisão = \frac{VP}{VP + FP}. \quad (86)$$

A precisão é importante quando se deseja avaliar a confiabilidade dos alarmes positivos gerados pelo modelo. Um modelo com baixa precisão produz muitos falsos positivos, isto é, classifica como poluídos muitos isoladores que, na realidade, estão limpos. Entretanto, a precisão não informa quantos isoladores poluídos foram efetivamente detectados, informação fornecida pela sensibilidade. Por esse motivo, precisão e sensibilidade devem ser interpretadas de forma conjunta.

O F1-score combina precisão e sensibilidade por meio da média harmônica entre essas duas métricas, sendo definido por:

$$F1 - score = \frac{2 \cdot Precisão \cdot Sensibilidade}{Precisão + Sensibilidade}. \quad (87)$$

O F1-score assume valores elevados apenas quando precisão e sensibilidade são simultaneamente altas. Dessa forma, ele é útil quando se deseja avaliar o equilíbrio entre detectar corretamente a classe positiva e evitar falsos alarmes positivos. Em problemas com duas classes, também pode ser calculado o F1 macro, definido como a média dos valores de F1 obtidos para cada classe. Essa forma de cálculo evita que o desempenho em uma classe majoritária domine a avaliação global do modelo, sendo adequada quando se deseja considerar o desempenho médio entre as classes.

Além dessas métricas, o coeficiente de correlação de Matthews, ou MCC, fornece uma medida global da qualidade da classificação considerando simultaneamente os quatro elementos da matriz de confusão (Matthews, 1975). O MCC é definido por:

$$MCC = \frac{VP \cdot VN - FP \cdot FN}{\sqrt{(VP + FP)(VP + FN)(VN + FP)(VN + FN)}}. \quad (88)$$

O MCC varia de -1 a 1. Valores próximos de 1 indicam concordância elevada entre rótulos reais e preditos; valores próximos de 0 indicam desempenho semelhante ao acaso; e valores próximos de (-1) indicam discordância sistemática entre as predições e os rótulos reais. A principal vantagem do MCC é que ele só assume valores elevados quando o classificador apresenta bom desempenho em todos os componentes da matriz

de confusão. Assim, o modelo precisa detectar corretamente a classe positiva, reconhecer corretamente a classe negativa e, ao mesmo tempo, manter baixos os falsos positivos e os falsos negativos.

### 2.6.3 DESEMPENHO MÉDIO E DISPERSÃO

Quando os modelos são avaliados em múltiplas repetições experimentais, uma única medida de desempenho pode não representar adequadamente o comportamento do classificador. Diferentes divisões entre treino e teste, diferentes sementes aleatórias ou diferentes rodadas de validação podem produzir variações nas métricas obtidas. Por isso, torna-se necessário apresentar não apenas o valor médio das métricas, mas também sua dispersão.

Essa análise é importante porque dois modelos podem apresentar médias semelhantes, mas comportamentos distintos ao longo das repetições. Um modelo pode alcançar desempenho médio elevado, porém com grande variabilidade, indicando sensibilidade à divisão dos dados ou à inicialização experimental. Outro modelo pode apresentar desempenho médio ligeiramente inferior, mas com menor dispersão, sugerindo maior estabilidade. Em aplicações de diagnóstico, essa estabilidade é relevante porque um modelo que apresenta bom desempenho apenas em algumas repetições pode ser menos confiável para aplicação prática.

### 3 REVISÃO BIBLIOGRÁFICA

Com base nos fundamentos apresentados nas seções anteriores e na consolidação do uso da termografia infravermelha no monitoramento de equipamentos elétricos, a presente revisão bibliográfica foi estruturada com foco em duas vertentes complementares e ainda pouco exploradas de forma integrada na literatura. A primeira concentra-se na aplicação de modelos probabilísticos sequenciais, em especial os Modelos Ocultos de Markov, no contexto do sistema elétrico de potência, abordando sua utilização em tarefas de modelagem dinâmica, detecção de padrões e análise temporal de sinais. A segunda vertente dedica-se às técnicas de classificação aplicadas ao diagnóstico de anomalias em equipamentos elétricos por meio de dados termográficos, englobando abordagens baseadas em aprendizado de máquina e processamento de sinais térmicos. Dessa forma, esta revisão busca não apenas contextualizar os avanços existentes em cada uma dessas áreas, mas também evidenciar a lacuna científica associada à exploração conjunta da dinâmica sequencial de perfis térmicos para fins de diagnóstico, o que fundamenta a proposta metodológica desenvolvida neste trabalho.

#### 3.1 MODELOS OCULTOS DE MARKOV NO SETOR ELÉTRICO

A aplicação dos Modelos Ocultos de Markov (HMM) em sistemas dinâmicos para a detecção de falhas foi sendo consolidada a partir de trabalhos como o de Smyth (1994), expandindo-se posteriormente para uma ampla gama de problemas em sistemas elétricos de potência, principalmente devido à sua capacidade de modelar processos estocásticos com dependência temporal e estados não observáveis (Smyth, 1994).

Na área de equipamentos elétricos, diversas abordagens foram desenvolvidas para o monitoramento e diagnóstico da degradação. Miya (2008) investigou o monitoramento da condição de buchas de papel impregnado a óleo utilizando HMM associado a redes neurais e modelos de mistura Gaussiana (do inglês, *Gaussian Mixtures Models*, GMM), empregando dados de análise de gases dissolvidos como variáveis de entrada. O método foi estruturado em etapas de detecção e classificação de falhas, demonstrando que o HMM, embora apresente maior custo computacional no treinamento, possui elevada capacidade de discriminação entre diferentes tipos de falhas devido à modelagem da

evolução temporal dos padrões (Miya, 2008). Mais recentemente, Zhan *et al.* (2024) utilizaram HMM para avaliação contínua de risco em transformadores, considerando concentrações de gases dissolvidos como variáveis dinâmicas para estimativa probabilística de falha, reforçando o papel dos HMMs como ferramenta robusta para análise de sistemas complexos e não estacionários (Zhan, 2024).

Posteriormente, Venkatesh *et al.* (2011) e Balasubramanian *et al.* (2016) aplicaram versões de densidade contínua do HMM no diagnóstico de isolamentos a óleo, com foco na identificação de padrões de descargas parciais. Nesses trabalhos, os sinais foram tratados como sequências estocásticas compostas por variáveis como magnitude da descarga, número de pulsos e fase de ocorrência. Em particular, Balasubramanian *et al.* (2016) realizaram experimentos sob diferentes condições de envelhecimento acelerado (térmico e elétrico), com tensões entre 13 kV e 21 kV, demonstrando que a matriz de transição de estados do HMM apresenta assinaturas distintas para cada condição de degradação, permitindo diferenciar estados saudáveis e degradados do sistema (Venkatesh, Gopal e Kannan, 2011; Balasubramanian e Venkatesh, 2016)

Mais recentemente, Hong e Lin (2019) utilizaram HMM aliado à análise dinâmica não linear de vibrações para classificar condições mecânicas de enrolamentos de transformadores. Nesse estudo, características não lineares foram extraídas por meio de técnicas como o gráfico de recorrência cruzada, gerando sequências de atributos que alimentam o HMM para classificação entre estados normal, degradado e anômalo. Posteriormente, Liu *et al.* (2021) propuseram um modelo semi markoviano de alta ordem para estimar a vida útil residual de disjuntores de alta tensão, incorporando explicitamente a duração dos estados e dependências de ordem superior. Esse modelo permite representar com maior fidelidade a dinâmica de degradação, superando limitações do HMM tradicional na modelagem de processos com tempos de permanência variáveis. De forma complementar, Sui *et al.* (2023) aplicaram HMM combinado com energia de pacotes de wavelet, explorando informações no domínio tempo-frequência para detectar afrouxamentos do núcleo e deformações estruturais, evidenciando a versatilidade do modelo em diferentes tipos de sinais físicos (Hong e Lin, 2019; Liu *et al.*, 2021; Sui *et al.*, 2023).

No estudo de máquinas elétricas, especialmente motores, observa-se uma evolução do uso do HMM para o prognóstico de falhas. Yusuf *et al.* (2013) combinaram HMM com modelagem Bayesiana Naïve para monitoramento de motores de indução, utilizando sinais elétricos como corrente e tensão, e explorando a capacidade do HMM

em representar padrões temporais enquanto o classificador Bayesiano reduz a incidência de falsos alarmes (Yusuf *et al.*, 2013).

Em seguida, Soualhi *et al.* (2015) empregaram HMM para prever o surgimento de falhas mecânicas, modelando o processo de degradação ao longo do tempo por meio de estados ocultos que representam diferentes níveis de dano. Nesse contexto, as transições entre estados permitem antecipar a evolução da falha, possibilitando aplicações em manutenção preditiva. Zhang *et al.* (2016) demonstraram a eficácia do HMM no diagnóstico online de curtos-circuitos em motores DC, utilizando sequências de sinais de corrente para identificar não apenas a ocorrência da falha, mas também o seu grau de severidade (Soualhi, Clerc e Razik, 2015; Zhang, Zhan e Ehsani, 2016).

Outra linha de aplicação relevante envolve a análise de linhas de transmissão, conversores e qualidade de energia. Huang *et al.* (2016) desenvolveram algoritmos baseados em HMM para detecção de interrupções em linhas de transmissão utilizando dados sincrofasoriais, modelando o estado das linhas como uma cadeia de Markov oculta e aplicando técnicas de inferência baseadas em filtros de partículas para lidar com ruído e incertezas nas medições (Huang, 2016).

Freire *et al.* (2019) aplicaram HMM diretamente em séries temporais multivariadas, utilizando sinais trifásicos de tensão e corrente sem necessidade de etapas complexas de pré-processamento. Os resultados indicaram uma taxa de erro de aproximadamente 0,03%, superando métodos tradicionais como Random Forest, SVM e KNN, além de reduzir significativamente o custo computacional. Xiao e Ai (2019) propuseram uma abordagem baseada em múltiplos HMMs, incorporando dados climáticos e espaciais para previsão de distúrbios na qualidade de energia, demonstrando melhoria na acurácia ao considerar fatores externos como condições meteorológicas (Freire *et al.*, 2019; Xiao e Ai, 2019).

Kouadri *et al.* (2020), por sua vez, combinaram HMM com Análise de Componentes Principais (do inglês *Principal Component Analysis*, PCA) para diagnóstico de falhas em conversores de sistemas eólicos, utilizando PCA para redução de dimensionalidade e extração de características relevantes antes da classificação probabilística pelo HMM. Adicionalmente, Jianquan *et al.* (2022) utilizaram HMM para modelagem preditiva de perdas em linhas de transmissão, transformando séries temporais de perdas em estados discretos por meio de técnicas de agrupamento *fuzzy* e capturando a relação entre variáveis operacionais e perdas ao longo do tempo (Kouadri *et al.*, 2020; Jianquan *et al.*, 2022).

Mais recentemente, observa-se a expansão dos HMMs para problemas associados à integração entre sistemas elétricos, dados operacionais e infraestruturas inteligentes. Oliveira *et al.* (2025) investigaram o uso de DTW estocástico baseado em HMM para o alinhamento de perfis térmicos de para-raios de alta tensão em subestações. Nesse estudo, o alinhamento térmico foi formulado a partir de uma estrutura de HMM com matriz de transição do tipo esquerda-direita, na qual os estados ocultos representam posições sucessivas do perfil de referência, enquanto as observações correspondem às temperaturas do perfil a ser alinhado. Os autores empregaram os algoritmos *Forward* e *Viterbi* para, respectivamente, avaliar a dissimilaridade entre os sinais e estimar o caminho ótimo de alinhamento. Embora o DTW com restrição adaptativa tenha apresentado desempenho superior em precisão e menor custo computacional, o DTW estocástico baseado em HMM demonstrou potencial para aplicações futuras em classificação automática de defeitos, sobretudo por incorporar uma interpretação probabilística da evolução dos perfis térmicos (Oliveira *et al.*, 2025).

No contexto de segurança cibernética em sistemas elétricos de potência, Zhang *et al.* (2025) propuseram um Modelo Oculto de Markov Dinâmico para predição de ataques cibernéticos em sistemas ciberfísicos de potência. Diferentemente de abordagens estáticas, o modelo considera a evolução dinâmica da intenção do atacante e ajusta a predição com base nas características do processo de ataque e defesa. O método combina redução de dados de alarme, seleção de atributos, agrupamento e otimização do HMM por meio do algoritmo de Baum-Welch, além de utilizar uma versão dinâmica do algoritmo de Viterbi para estimar os próximos estados de ataque. Os resultados foram obtidos a partir de uma plataforma física de simulação cibernética aplicada a um sistema real de 126 nós, demonstrando que os HMM também são adequados para representar estados não observáveis em ambientes de alta complexidade, nos quais apenas sequências de alarmes e eventos são diretamente medidas (Zhang *et al.*, 2025).

Além das aplicações em equipamentos e segurança cibernética, Xiao *et al.* (2025) aplicaram HMM à mitigação de desvios energéticos em sistemas acoplados de potência e transporte. Nesse trabalho, os estados ocultos representam condições operacionais não observáveis da camada de transporte, como padrões de mobilidade, níveis de carga das baterias e intenções de deslocamento de veículos elétricos, enquanto as observações são associadas às demandas agregadas de carregamento informadas por agregadores de veículos elétricos. A utilização do HMM permite estimar probabilisticamente a evolução da demanda de carregamento e reduzir a diferença entre os valores programados pelo

operador do sistema elétrico e a demanda efetivamente realizada. Essa aplicação evidencia a capacidade do HMM de atuar não apenas no diagnóstico de falhas, mas também em problemas de previsão, despacho e operação coordenada de infraestruturas elétricas sujeitas a incertezas temporais (Xiao *et al.*, 2025).

No contexto de monitoramento não intrusivo de cargas, Zoha *et al.* (2013) introduziram o uso do Modelo Oculto de Markov Fatorial para identificação de estados operacionais de dispositivos a partir de medições agregadas de energia, explorando a decomposição do consumo total em múltiplas cadeias de Markov independentes. Posteriormente, Luan *et al.* (2022) expandiram essa abordagem, incorporando explicitamente a duração dos estados e variáveis adicionais como potência reativa, melhorando o desempenho em cenários industriais. Ademais, Kumar (2023) propuseram uma versão modificada do HMM Fatorial com detecção de pontos de mudança, aumentando a eficiência em dados de baixa taxa de amostragem e permitindo melhor identificação de eventos de operação (Zoha, 2013; Luan *et al.*, 2022; Kumar, 2023)

### 3.2 MONITORAMENTO POR TERMOGRAFIA NO SETOR ELÉTRICO

A aplicação da termografia infravermelha em sistemas de inspeção para manutenção preditiva foi sendo consolidada a partir de avanços em sensores térmicos, processamento digital de imagens e inteligência computacional, expandindo-se posteriormente para uma ampla gama de problemas em sistemas elétricos de potência, principalmente devido à sua capacidade de avaliar condições térmicas anômalas sem interrupção da operação, de forma não destrutiva, segura e compatível com ambientes energizados.

Nos primeiros estudos, observa-se a utilização direta dos perfis térmicos como fonte de informação diagnóstica. Wanderley Neto *et al.* (2007) investigaram a avaliação da condição de para-raios de óxido de zinco a partir de perfis térmicos obtidos por termografia, nos quais defeitos foram induzidos experimentalmente, incluindo curto-circuito de varistores, presença de umidade, poluição superficial e deslocamentos estruturais, de modo a gerar variabilidade nas assinaturas térmicas. Esses perfis, após normalização, foram utilizados como entrada em uma rede neural artificial treinada com os algoritmos *Backpropagation* e *Resilient Propagation*, estruturada com três camadas intermediárias de 16 neurônios, resultando em uma precisão de 99,4% na detecção da

presença de falhas, embora ainda limitada na discriminação detalhada dos tipos específicos de defeito (Wanderley Neto, 2007).

Com a evolução das abordagens, passou-se a explorar de forma mais sistemática a extração de características estatísticas das imagens térmicas. Huda e Taib (2013) propôs um sistema baseado na extração de quinze atributos de primeira e segunda ordem, como contraste, entropia, homogeneidade e métricas diferenciais entre regiões de falha e regiões de referência, seguido de seleção por análise discriminante multivariada, resultando em um subconjunto de dez variáveis estatisticamente relevantes. Nesse cenário, o classificador por Análise Discriminante apresentou desempenho superior à rede *Multi Layer Perceptron* (82,40% contra 80,40%), evidenciando a importância da seleção de atributos dominantes na discriminação entre estados operacionais (Huda e Taib, 2013).

Na sequência, a literatura passou a enfatizar a segmentação automática da região de interesse como etapa fundamental do processo diagnóstico. Zou e Huang (2015) propuseram um método baseado em agrupamento *K-means* para segmentação dinâmica das regiões térmicas, eliminando a necessidade de delimitação manual, seguido de classificação a partir de uma Máquina de Vetor de Suporte (do inglês *Support Vector Machine*, SVM) com otimização de parâmetros combinada com validação cruzada e sequências caóticas. A abordagem utilizou atributos como temperatura máxima, média e área ativa dos pixels, atingindo níveis de acurácia próximos a 97,8%, superando métodos baseados em redes de retropropagação (Zou e Huang, 2015).

De forma complementar, Leksir *et al.* (2017) demonstraram que a utilização de SVM sobre imagens previamente segmentadas permite uma redução significativa de ruído e uma delimitação mais precisa das regiões de aquecimento em comparação com técnicas clássicas de limiarização, como *Otsu* e *Fuzzy C-Means*. Essa linha de investigação é reforçada por Singh (2021), que aplicou limiarização ótima e máxima entropia associadas a filtragem adaptativa para segmentação de isoladores de porcelana, extraíndo posteriormente atributos geométricos e térmicos, como área, eixos e intensidade para classificação via SVM com *kernel* Gaussiano, permitindo a discriminação de falhas em múltiplos níveis de severidade com elevada precisão, obtendo acurácia de 97%, e eficiência computacional (Leksir, Mansour e Moussaoui, 2017; Singh *et al.*, 2021).

Paralelamente, foram desenvolvidas abordagens baseadas em descritores de textura e relações espaciais. Ullah *et al.* (2017) utilizaram características extraídas da matriz de temperatura representando padrões de textura e energia térmica, associadas a

redes neurais *Multi Layer Perceptron*. Inicialmente, o modelo apresentou acurácia em torno de 79%, sendo posteriormente aprimorado com a introdução de segmentação via *graph cut*, baseada em minimização de energia, elevando o desempenho para mais de 84%, evidenciando o impacto da modelagem espacial na discriminação entre estados com e sem defeito (Ullah *et al.*, 2017).

Além disso, trabalhos como Li *et al.* (2019) avançaram na mitigação de ruídos térmicos, propondo uma reestruturação matricial baseada na dependência entre *pixels* vizinhos para eliminar interferências oriundas de excitações externas, como correntes parasitas, permitindo maior precisão na detecção de microfissuras. Em uma abordagem distinta, Novizon (2019) integrou dados termográficos com medições elétricas, como correntes de fuga, combinando essas informações em um modelo MLP para predição contínua da condição de para-raios, alcançando níveis de assertividade próximos a 98% e evidenciando o potencial da fusão de dados multimodais (Li *et al.*, 2019; Novizon *et al.*, 2019).

No contexto de aprendizado de máquina clássico, Najafi *et al.* (2020) propuseram uma abordagem baseada em *Random Forest* (RF) e *AdaBoost* associada a aprendizado interpretável, na qual os dados são previamente segmentados em estados “frios” e “quentes”, seguidos de extração de características via janela deslizante e classificação por árvores de decisão. O método alcançou acurácia superior a 93–95% em múltiplas categorias de falha, destacando-se pela interpretabilidade do modelo, que permite avaliar a importância relativa das variáveis no processo decisório (Najafi *et al.*, 2020).

Com o avanço do aprendizado profundo, observa-se uma transição para modelos capazes de realizar extração automática de características. Ullah *et al.* (2020) utilizaram uma Rede Neural Convolucional (do inglês *Convolutional Neural Network*, CNN) baseada em *AlexNet* para extrair representações profundas a partir de imagens térmicas, posteriormente classificadas por *Random Forest*, atingindo desempenho superior a 96%. Nesse contexto, a CNN atua como extratora de características complexas, enquanto o classificador final realiza a decisão, explorando a capacidade das redes profundas de capturar padrões espaciais não lineares (Ullah *et al.*, 2020).

Outras abordagens avançadas incluem o uso de técnicas multirresolução e novos descritores térmicos. Choudhary *et al.* (2021) aplicaram a Transformada Wavelet Discreta bidimensional para decomposição das imagens térmicas, seguida de redução de dimensionalidade via PCA e seleção baseada na distância de *Mahalanobis*, obtendo desempenho extremamente elevado na classificação de falhas em rolamentos. De forma

complementar, Li (2021) propôs descritores baseados em decomposição Wavelet para eliminar efeitos de aquecimento ambiental, como insolação, permitindo discriminar falhas reais a partir de gradientes térmicos e desvios locais (Choudhary, Goyal e Letha, 2021; Li *et al.*, 2021).

Nos trabalhos mais recentes, Wang *et al.* (2022) desenvolveram um sistema robótico controlado para inspeção termográfica automatizada, incorporando métricas como temperatura máxima, taxa de crescimento térmico e área radiativa, permitindo diagnósticos contínuos e automáticos em subestações remotas. Das *et al.* (2022) propuseram a incorporação de modelos híbridos capazes de integrar informações espaciais e temporais, a partir da arquitetura *Long-Short Term Memory* (LSTM) combinada a uma CNN, além disso, utilizou-se *Mask R-CNN* para segmentação e classificação de isoladores, explorando dependências sequenciais nas representações térmicas e alcançando F1-score superior a 98,8% na estimativa da gravidade da contaminação (Wang *et al.*, 2022; Das *et al.*, 2022).

Além disso, destacam-se arquiteturas profundas e estratégias híbridas. Sasikumar e Rajendran (2024) utilizaram *DenseNet* combinada com recursos robustos acelerados e redes neurais artificiais para diagnóstico de falhas em motores, atingindo acurácia de 98,7%. Shafique *et al.* (2024) exploraram uma abordagem semi-supervisionada baseada em *K-means* e SVM, reduzindo a dependência de dados rotulados e mantendo desempenho elevado na ordem de 94%. Por fim, Tang e Jian (2024) integraram modelos de detecção de objetos e segmentação semântica, como *RetinaNet* e *DeepLabV3+*, permitindo a identificação simultânea de múltiplas regiões críticas em ambientes complexos de subestações, com classificação baseada em métricas térmicas absolutas e relativas (Sasikumar e Rajendran, 2024; Shafique *et al.*, 2024; Tang e Jian, 2024).

Mais recentemente, Liang *et al.* (2025) abordaram o monitoramento termográfico de buchas de parede em sistemas de transmissão em corrente contínua de ultra-alta tensão, com foco na medição multiponto de temperatura sob interferência das saias isolantes. Os autores destacaram que a geometria periódica dessas saias pode gerar pontos de interferência nas imagens infravermelhas, dependentes do ângulo e da distância de aquisição, distorcendo a estimativa da distribuição térmica ao longo da bucha. Para superar essa limitação, foi proposto um método não supervisionado baseado em detecção de anomalias, comparação entre pixels e mecanismo de votação entre diferentes algoritmos, permitindo identificar e remover pontos de interferência sem necessidade de dados previamente rotulados. A abordagem foi testada em 161 imagens infravermelhas

de buchas e apresentou taxa de sucesso de 100% na identificação de sobreaquecimentos localizados, evidenciando a importância de técnicas adaptativas de pré-processamento para aumentar a confiabilidade da medição térmica em equipamentos com geometrias complexas (Liang *et al.*, 2025).

De forma diretamente relacionada à análise de perfis térmicos de para-raios, Araújo *et al.* (2025) propuseram uma metodologia voltada ao monitoramento de para-raios de óxido de zinco em subestações de alta tensão, combinando redes neurais convolucionais, processamento digital de imagens e alinhamento de sinais para análise de perfis térmicos. Diferentemente de abordagens centradas apenas na detecção de pontos quentes isolados, o método utiliza a termografia para identificar automaticamente a região de interesse, segmentar a coluna isolante, extrair a matriz de temperatura e construir o perfil térmico longitudinal do equipamento. A partir desse perfil, aplica-se o alinhamento por Dynamic Time Warping (DTW), permitindo comparar o equipamento analisado com um equipamento de referência ou com perfis históricos sob condições semelhantes. Essa estratégia possibilita não apenas indicar a presença de anomalias térmicas, mas também avaliar o grau de similaridade entre perfis, identificar desvios positivos ou negativos de temperatura e localizar espacialmente a região afetada no corpo do para-raios. Assim, a principal contribuição do estudo reside na transição de uma inspeção pontual baseada em temperaturas máximas para uma abordagem de monitoramento estrutural do comportamento térmico ao longo da coluna isolante, com potencial para diagnóstico mais abrangente e preditivo. Os resultados obtidos em imagens de uma subestação de 500 kV e em ensaios laboratoriais com equipamentos saudáveis e defeituosos indicaram desempenho satisfatório na detecção e segmentação da região de interesse, além de permitir a identificação correta de anomalias e de suas respectivas localizações, mesmo diante de variações de distância, ângulo de aquisição, erro de medição do termovisor e influências ambientais (Araujo *et al.*, 2025).

Ademais, Sun *et al.* (2025) avançaram no uso da termografia para alerta precoce de falhas em cubículos de alta tensão, propondo uma abordagem baseada em imagens infravermelhas associadas à otimização multiescala de características. O estudo parte do problema de elevação anormal de temperatura em compartimentos internos de painéis elétricos, nos quais a dissipação térmica limitada pode acelerar o envelhecimento dos componentes, reduzir a rigidez dielétrica e aumentar o risco de falhas severas. Para melhorar a qualidade das amostras e a capacidade discriminativa do modelo, os autores empregaram reforço multidirecional de bordas, mecanismo de atenção eficiente por canal

e otimização por pirâmide de características, integrando posteriormente essas informações a um modelo SVM otimizado por uma versão aprimorada do *Sparrow Search Algorithm*. A proposta alcançou acurácia de 98,8%, além de melhorar a capacidade preditiva em relação a métodos de referência, evidenciando a relevância de mecanismos de atenção e fusão multiescala na extração de padrões térmicos em equipamentos de subestação (Sun *et al.*, 2025).

Em uma linha mais recente associada a modelos profundos leves, Wang *et al.* (2026) desenvolveram o MTrip-YOLO, uma arquitetura baseada no YOLO11n aprimorada com Mamba-2 e mecanismo de *Triplet Attention* para detecção de falhas em buchas de alta tensão a partir de imagens infravermelhas. O método foi projetado para lidar com dois desafios recorrentes em inspeções termográficas de subestações: a presença de alvos pequenos, frequentemente observados a longas distâncias, e a interferência de objetos tubulares com morfologia semelhante, como barramentos, isoladores e outros componentes metálicos. O módulo Mamba-2 foi incorporado para capturar dependências espaciais de longo alcance com baixa complexidade computacional, enquanto o mecanismo de *Triplet Attention* foi utilizado para reforçar assinaturas térmicas fracas e suprimir ruídos de fundo. O conjunto de dados incluiu imagens de buchas normais, falhas por mau contato, falta de óleo e alta perda dielétrica. O modelo final alcançou mAP50 de 91,6%, com um número mínimo de 1,9 milhão de parâmetros e velocidade de inferência de 133 quadros por segundo, superando modelos como Faster R-CNN, RT-DETR e versões recentes do YOLO em diferentes métricas. Esses resultados indicam uma tendência de desenvolvimento de modelos termográficos voltados não apenas à acurácia, mas também à implantação em dispositivos embarcados, plataformas móveis, inspeções com drones e sistemas de monitoramento em tempo real (Wang *et al.*, 2026).

### 3.3 ANÁLISE DAS CONTRIBUIÇÕES DOS TRABALHOS CORRELATOS

Os estudos que motivaram e fundamentaram esta dissertação foram apresentados ao longo deste capítulo. A revisão bibliográfica evidenciou que a termografia infravermelha tem sido amplamente empregada no monitoramento de equipamentos elétricos, contemplando desde a identificação de regiões aquecidas e a extração de informações térmicas até a aplicação de métodos automáticos de diagnóstico. Em

diferentes abordagens, observa-se o uso de imagens térmicas, regiões de interesse, mapas de temperatura, perfis térmicos e descritores extraídos da distribuição de temperatura para avaliar a condição de equipamentos como para-raios, isoladores, chaves seccionadoras, transformadores, motores e demais componentes de subestações e linhas de transmissão.

Os trabalhos analisados também demonstraram avanços importantes no processamento automatizado das informações térmicas. Alguns estudos utilizaram técnicas de segmentação e detecção de regiões de interesse para localizar equipamentos ou anomalias térmicas; outros exploraram descritores estatísticos, gradientes de temperatura, mapas térmicos, perfis longitudinais e métodos de aprendizado de máquina para classificar a condição operacional dos equipamentos. Nesse contexto, verificou-se que a literatura já apresenta contribuições relevantes no uso de redes neurais artificiais, máquinas de vetores de suporte, florestas aleatórias, redes convolucionais e modelos híbridos aplicados ao diagnóstico baseado em imagens térmicas.

Apesar desses avanços, observa-se que a maior parte das pesquisas representa a informação térmica de forma global, seja por meio da imagem completa, de regiões de interesse, de parâmetros estatísticos ou de descritores agregados. Mesmo nos estudos que exploram perfis térmicos ou curvas de temperatura ao longo do equipamento, a análise tende a se concentrar em valores globais, gradientes, diferenças térmicas ou procedimentos de alinhamento, sem tratar explicitamente o perfil térmico unidimensional como uma sequência espacialmente organizada de informações. Assim, a dependência entre regiões consecutivas do perfil, bem como a ordem em que essas regiões aparecem ao longo do isolador, ainda é pouco explorada como elemento discriminativo no diagnóstico.

Essa lacuna torna-se particularmente relevante quando se considera que o perfil térmico de um isolador não é apenas um conjunto de valores independentes de temperatura, mas uma sequência de observações distribuídas ao longo de sua geometria. Nesse sentido, modelos capazes de representar dependências sequenciais, como os Modelos Ocultos de Markov, apresentam potencial para o tratamento desse tipo de informação, uma vez que permitem modelar a evolução de estados não diretamente observáveis a partir de uma sequência de observações. A revisão bibliográfica mostrou aplicações bem-sucedidas desses modelos em problemas de diagnóstico, reconhecimento de padrões e monitoramento de sistemas, inclusive no setor elétrico. Entretanto, não foram identificadas aplicações voltadas especificamente à modelagem de perfis térmicos

obtidos por termografia infravermelha em isoladores, o que reforça uma lacuna metodológica e justifica a investigação dessa abordagem nesta dissertação.

Nesse cenário, esta dissertação diferencia-se por propor uma abordagem voltada à classificação de perfis térmicos de isoladores poliméricos, tratando o perfil unidimensional como uma estrutura espacialmente ordenada. Dessa forma, a contribuição central desta dissertação não está apenas no uso da termografia ou na extração de perfis térmicos, aspectos já presentes em parte da literatura, mas na avaliação sistemática de diferentes formas de representar a informação térmica ao longo do isolador. Ao comparar múltiplas representações, como, agregada, sequencial e simbólica, o trabalho busca investigar se a preservação da dependência espacial do perfil térmico acrescenta informação relevante para distinguir isoladores limpos e poluídos. Assim, a metodologia proposta amplia a discussão sobre o uso de perfis térmicos no diagnóstico de isoladores, deslocando o foco da simples análise global da temperatura para a modelagem da organização espacial do aquecimento ao longo do equipamento.

Tabela 2. Quadro sinóptico.

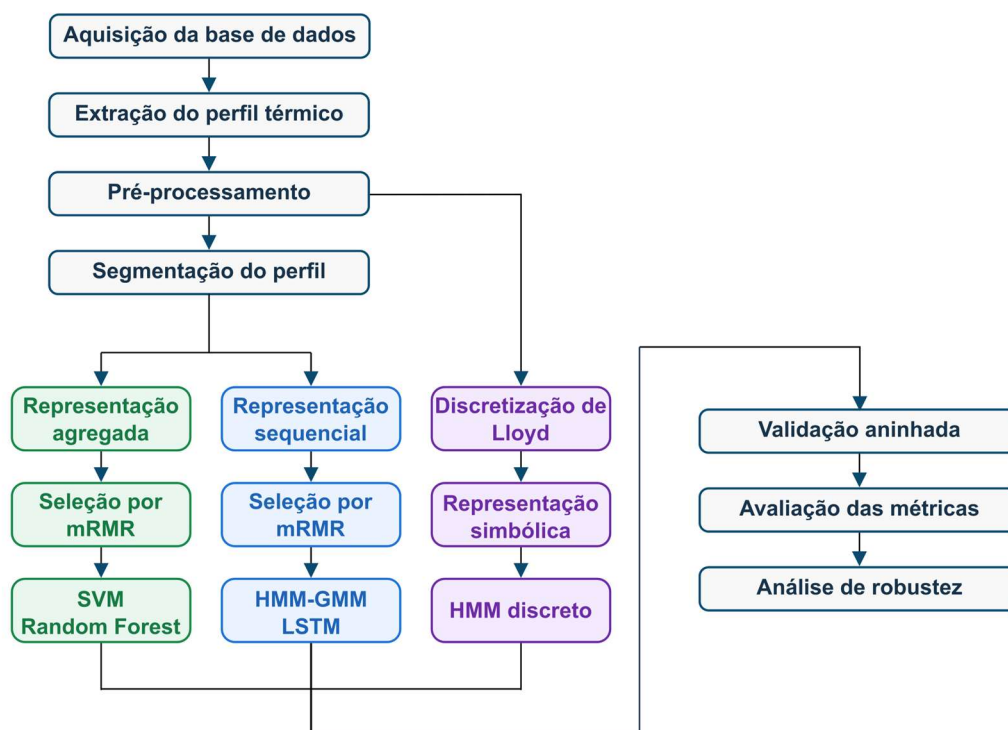
| Referência                       | Perfil térmico longitudinal | Avaliação de dependência espacial | Aplicação de HMM | Múltiplas representações | Correção relativa ao ambiente | Processamento automático | Classificação de integridade |
|----------------------------------|-----------------------------|-----------------------------------|------------------|--------------------------|-------------------------------|--------------------------|------------------------------|
| (Wanderley Neto, 2007)           | X                           |                                   |                  |                          |                               |                          | X                            |
| (Huda e Taib, 2013)              | X                           | X                                 |                  |                          | X                             |                          | X                            |
| (Zou e Huang, 2015))             |                             |                                   |                  |                          | X                             | X                        | X                            |
| (Ullah <i>et al.</i> , 2017))    |                             | X                                 |                  |                          | X                             |                          | X                            |
| (Ullah <i>et al.</i> , 2020)     |                             |                                   |                  |                          | X                             |                          | X                            |
| (Najafi <i>et al.</i> , 2020)    |                             | X                                 |                  |                          |                               |                          | X                            |
| (Choudhary, Goyal e Letha, 2021) |                             |                                   |                  |                          |                               |                          | X                            |
| (Li <i>et al.</i> , 2021)        | X                           |                                   |                  |                          |                               |                          | X                            |
| (Wang <i>et al.</i> , 2022)      | X                           |                                   |                  |                          | X                             |                          | X                            |
| (Wen e Jin, 2022))               | X                           |                                   |                  |                          | X                             |                          | X                            |
| (Sasikumar e Rajendran, 2024)    |                             |                                   |                  |                          |                               |                          | X                            |
| (Tang e Jian, 2024)              |                             |                                   |                  |                          | X                             |                          | X                            |
| (Shafique <i>et al.</i> , 2024)  |                             |                                   |                  |                          |                               | X                        | X                            |
| (Araujo <i>et al.</i> , 2025)    | X                           |                                   |                  |                          | X                             | X                        | X                            |
| (Liang <i>et al.</i> , 2025)     |                             |                                   |                  |                          | X                             | X                        | X                            |
| (Oliveira <i>et al.</i> , 2025)  | X                           |                                   | X                |                          | X                             | X                        |                              |
| (Wang <i>et al.</i> , 2026)      |                             |                                   |                  |                          |                               |                          | X                            |
| Presente Trabalho                | X                           | X                                 | X                | X                        | X                             | X                        | X                            |

Fonte: autoria própria.

## 4 METODOLOGIA

A metodologia deste trabalho foi estruturada com o objetivo de avaliar a capacidade de diferentes representações de perfis térmicos na classificação binária de isoladores poliméricos, distinguindo amostras em condição limpa de amostras submetidas à poluição superficial. Para isso, os perfis térmicos unidimensionais foram processados, transformados em diferentes representações e avaliados por modelos de classificação pertencentes a distintas famílias metodológicas. Na Figura 17 ilustra-se o fluxograma metodológico geral desta dissertação.

Figura 17. Fluxograma metodológico geral.



Fonte: autoria própria.

## 4.1 AQUISIÇÃO DA BASE DE DADOS

A base de dados utilizada nesta dissertação foi composta por termografias de isoladores poliméricos de média tensão, modelo Balestro, classe de tensão de 15 kV, avaliados sob diferentes condições de contaminação superficial. Os ensaios foram estruturados de modo a contemplar cenários com isoladores limpos, poluídos artificialmente e poluídos naturalmente, permitindo investigar a influência de diferentes naturezas de poluição no comportamento térmico dos equipamentos.

As imagens térmicas foram obtidas com o termovisor FLIR T840, equipamento com faixa de medição de  $-20\text{ }^{\circ}\text{C}$  a  $120\text{ }^{\circ}\text{C}$ , precisão de  $\pm 2\text{ }^{\circ}\text{C}$  ou  $\pm 2\%$  da leitura, lente de  $14^{\circ}$  e campo de visão de  $14^{\circ} \times 10^{\circ}$ . Antes das aquisições, a temperatura refletida foi determinada conforme as orientações do manual técnico da FLIR (FLIR, 2019), e a emissividade foi ajustada para 0,92, valor aproximado para o material constituinte dos isoladores poliméricos (Araujo *et al.*, 2025). Na Figura 18, apresenta-se o termovisor utilizado nos ensaios.

Figura 18. Termovisor.



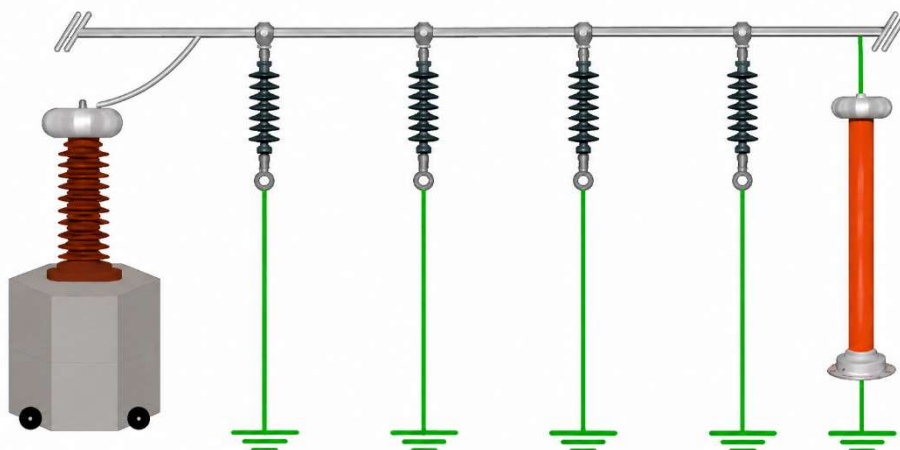
Fonte: (FLIR, 2019).

As termografias foram capturadas em ângulos frontais similares, mantendo-se distância aproximada de 3 m entre o termovisor e os isoladores. Além disso, durante os experimentos, cinco termografias foram obtidas para cada um dos quatro isoladores testados simultaneamente.

Os ensaios foram realizados no Laboratório de Alta Tensão da Universidade Federal de Sergipe (LAT/UFS), ambiente controlado e adequado para medições envolvendo equipamentos de alta tensão. Os isoladores foram energizados por meio de uma fonte de alta tensão alternada, com medição da tensão realizada a partir de divisor resistivo conectado ao barramento onde os isoladores foram alocados. Durante as aquisições, as condições ambientais, como temperatura ambiente e umidade relativa do

ar, foram monitoradas com termo-higrômetro calibrado, modelo EXBOM FEPRO-MUT50, a fim de aumentar a confiabilidade dos dados experimentais. Além disso, em cada experimento, os isoladores foram mantidos energizados durante 1 hora até a estabilização térmica antes da aquisição das imagens. Na Figura 19 é ilustrado o arranjo experimental.

Figura 19. Arranjo experimental.



Fonte: autoria própria.

Ao longo dos experimentos, os ensaios foram organizados em quatro grupos principais: ensaios somente com isoladores limpos, ensaios com poluição artificial uniforme, ensaios com poluição artificial não uniforme e ensaios com poluição natural. Essa organização permite agrupar experimentos de mesma natureza física, mesmo quando realizados em datas ou condições ambientais distintas, mantendo a coerência da análise e facilitando a formulação do problema de classificação. Em cada uma dessas condições, foram obtidas termografias dos quatro isoladores ensaiados simultaneamente, incluindo sempre o isolador limpo de referência. Essa organização permitiu comparar os perfis térmicos dos isoladores contaminados com uma condição limpa submetida ao mesmo ambiente experimental.

#### 4.1.1 EXPERIMENTO 1: ISOLADORES EM CONDIÇÃO LIMPA

O primeiro grupo experimental foi composto por ensaios realizados somente com isoladores em condição limpa, isto é, sem aplicação intencional de poluição superficial. Esse conjunto foi incorporado à base com o objetivo de ampliar a representatividade da classe Limpo, permitindo que os modelos de classificação observem variações térmicas

naturais associadas à aquisição, ao posicionamento, às condições ambientais e à própria variabilidade entre amostras limpas.

As condições ambientais registradas nesse experimento foram temperatura refletida de 30 °C, temperatura ambiente de 28,5 a 28,9 °C e umidade relativa de 74% a 76%. A tensão aplicada foi de 10 kV (fase-terra), valor adotado por representar uma condição de sobretensão em relação à operação nominal do isolador, favorecendo o estresse elétrico e térmico sobre a superfície contaminada. Essa condição teve como objetivo facilitar a estabilização do comportamento térmico durante o ensaio. Esse mesmo padrão de tensão aplicada foi adotado nos demais experimentos, de modo a manter a comparabilidade entre as amostras analisadas. A severidade da contaminação foi caracterizada por meio da densidade equivalente de depósito de sal, ESDD (*Equivalent Salt Deposit Density*), conforme a norma IEC 60815-1. Os níveis de ESDD considerados para os isoladores desse experimento são apresentados na Tabela 3.

Tabela 3. Níveis de ESDD do experimento 1.

| Isolador         | ESDD (mg/cm <sup>2</sup> ) | Grau de poluição |
|------------------|----------------------------|------------------|
| ISØ <sub>1</sub> | 0.0004                     | Atmosfera limpa  |
| ISØ <sub>2</sub> | 0.0004                     | Atmosfera limpa  |
| ISØ <sub>3</sub> | 0.0004                     | Atmosfera limpa  |
| ISØ <sub>4</sub> | 0.0004                     | Atmosfera limpa  |

Fonte: autoria própria.

#### 4.1.2 EXPERIMENTO 2: POLUIÇÃO ARTIFICIAL UNIFORME

O segundo grupo experimental corresponde aos ensaios com poluição artificial distribuída uniformemente sobre a superfície dos isoladores. O procedimento de aplicação da contaminação foi realizado com base na norma IEC 60815-1 (IEC, 2008), que orienta a caracterização de níveis de poluição em isoladores. Nesse cenário, foram avaliadas amostras com diferentes níveis de severidade de contaminação, além de uma amostra limpa de referência.

Esse experimento foi desenvolvido com o objetivo de gerar amostras que permitam analisar a resposta térmica dos isoladores diante de diferentes intensidades de poluição superficial uniformemente distribuída. Por apresentar distribuição mais homogênea da contaminação, esse cenário tende a produzir alterações térmicas mais

distribuídas ao longo do corpo do isolador, permitindo avaliar como o aumento do nível de poluição influencia os padrões térmicos observados nas termografias.

As condições ambientais registradas nesse experimento foram temperatura refletida de 31 °C, temperatura ambiente de 28,6 a 29,2 °C e umidade relativa de 78% a 82%. A severidade da contaminação foi caracterizada por meio da densidade equivalente de depósito de sal, ESDD (*Equivalent Salt Deposit Density*), conforme a norma IEC 60815-1 (IEC, 2008). Os níveis de ESDD considerados para os isoladores desse experimento são apresentados na Tabela 4.

Tabela 4. Níveis de ESDD do experimento 2.

| Isolador | ESDD (mg/cm <sup>2</sup> ) | Grau de poluição          |
|----------|----------------------------|---------------------------|
| ISØ      | 0.0004                     | Atmosfera limpa           |
| IS1      | 0.079                      | Contaminação pesada       |
| IS2      | 0.132                      | Contaminação muito pesada |
| IS3      | 0.593                      | Contaminação muito pesada |

Fonte: autoria própria.

#### 4.1.3 EXPERIMENTO 3: POLUIÇÃO ARTIFICIAL NÃO UNIFORME

O terceiro grupo experimental foi composto por ensaios com poluição artificial não uniforme. Diferentemente do experimento com poluição uniforme, neste caso a contaminação foi aplicada de forma concentrada em regiões específicas do isolador. Esse procedimento foi adotado para representar situações mais próximas de condições reais de operação, nas quais a deposição de contaminantes pode ocorrer de maneira irregular ao longo da superfície do equipamento.

Foram consideradas configurações em que a poluição foi concentrada em diferentes regiões do isolador, como o lado fase, o lado terra ou combinações de regiões ao longo do corpo isolante. Também foram consideradas amostras com poluição uniforme dentro desse conjunto experimental, utilizadas como referência para comparação com as configurações não uniformes. Essa organização permite avaliar se a distribuição espacial da contaminação produz alterações térmicas distintas nos perfis extraídos.

Nesse tipo de ensaio, a principal variável de interesse não é apenas a intensidade da poluição, mas também sua localização ao longo do isolador. A concentração de contaminantes em regiões específicas pode alterar localmente o campo elétrico, a

dissipação térmica e a distribuição de temperatura observada na termografia. Assim, os ensaios de poluição não uniforme permitem investigar se os descritores térmicos são sensíveis à posição espacial da contaminação.

As condições ambientais registradas foram temperatura refletida de 26 °C, temperatura ambiente de 24,1°C a 25°C e umidade relativa em torno de 88% a 90%. Na Tabela 5, apresentam-se as condições dos isoladores utilizados nesse experimento.

Tabela 5. Níveis de ESDD do experimento 3.

| Isolador | Distribuição | ESDD<br>(mg/cm <sup>2</sup> ) | Grau de<br>poluição          | Região |
|----------|--------------|-------------------------------|------------------------------|--------|
| ISØ      | -            | 0.001                         | Limpo                        | -      |
| IS4      | Não uniforme | 1.3                           | Contaminação<br>muito pesada | Terra  |
| IS5      | Não uniforme | 1.2                           | Contaminação<br>muito pesada | Fase   |
| IS6      | Uniforme     | 1.2                           | Contaminação<br>muito pesada | -      |

Fonte: autoria própria.

#### 4.1.4 EXPERIMENTO 4: POLUIÇÃO NATURAL

O quarto grupo experimental foi composto por isoladores submetidos à poluição natural, obtida por exposição em ambiente externo com incidência de salinidade. Esse tipo de ensaio busca representar uma condição mais próxima daquela observada em campo, especialmente em regiões litorâneas, onde a deposição de sais e partículas higroscópicas pode alterar o comportamento elétrico e térmico dos isoladores. Para isso, os isoladores foram posicionados em ambiente costeiro, situados a aproximadamente 1500 m do mar.

Ademais, os isoladores permaneceram expostos ao ambiente por aproximadamente a 117 dias, em local abrigado da chuva e sob influência de poluição salina. Essa condição permitiu a formação de uma camada de contaminação mais próxima daquela observada em campo, diferindo da poluição artificial tanto pela composição quanto pela distribuição espacial. Enquanto a poluição artificial apresenta composição controlada, geralmente associada ao uso de materiais como caulim e sais, a poluição

natural possui composição desconhecida e dependente das condições ambientais locais, incluindo salinidade, umidade, direção do vento e deposição de partículas.

As condições ambientais registradas durante o ensaio de poluição natural foram temperatura refletida de 29 °C, temperatura ambiente de 27,8°C a 29,3°C e umidade relativa de 74% a 79%. Na Tabela 6, apresentam-se as condições dos isoladores utilizados nesse experimento.

Tabela 6. Níveis de ESDD no experimento 4.

| Isolador | ESDD (mg/cm <sup>2</sup> ) | Grau de poluição          |
|----------|----------------------------|---------------------------|
| ISØ      | 0.0004                     | Limpo                     |
| IS7      | 0.45                       | Contaminação muito pesada |
| IS8      | 0.41                       | Contaminação muito pesada |
| IS9      | 0.43                       | Contaminação muito pesada |

Fonte: autoria própria.

Embora os experimentos tenham sido conduzidos sob diferentes tipos e níveis de poluição, o problema de classificação adotado neste trabalho foi formulado como uma tarefa binária. Dessa forma, os perfis térmicos extraídos dos isoladores foram agrupados em duas classes principais: limpo e poluído.

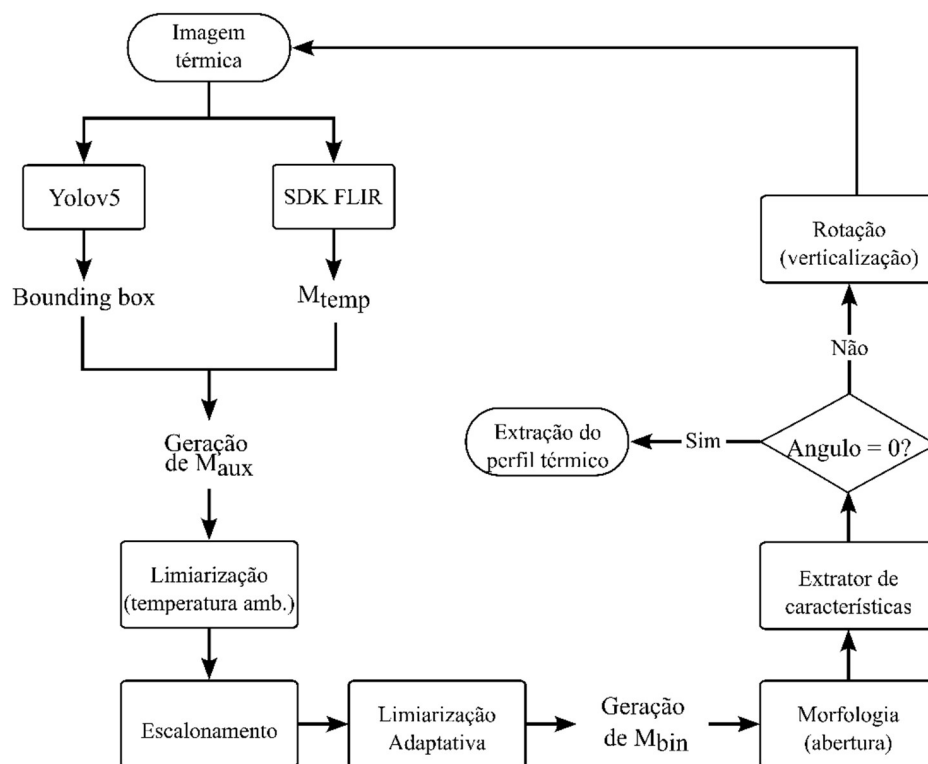
A classe “limpo” corresponde aos isoladores sem poluição presentes em cada grupo experimental, enquanto a classe poluído engloba os isoladores submetidos à poluição artificial uniforme, à poluição artificial não uniforme e à poluição natural. Essa formulação tem como objetivo avaliar se os atributos extraídos dos perfis térmicos são capazes de diferenciar a condição de referência limpa de diferentes cenários de contaminação superficial.

Dessa maneira, a base termográfica reúne 344 perfis térmicos, contemplando amostras obtidas em condições controladas de laboratório e em condição real de exposição ambiental. A classe Limpo é composta por 124 perfis, enquanto a classe Poluído reúne 220 perfis, distribuídos entre os subtipos Fase, Terra, Uniforme e Natural, com 20, 25, 55 e 120 perfis, respectivamente. Essa composição contribui para tornar a análise mais abrangente, pois permite que os modelos sejam avaliados frente a variações na intensidade, na distribuição espacial e na natureza da poluição presente nos isoladores.

## 4.2 EXTRAÇÃO DO PERFIL TÉRMICO

O método proposto não utiliza diretamente a termografia em sua forma original, mas sim um sinal unidimensional extraído a partir dela, correspondente ao perfil térmico central do equipamento. Para essa etapa, adotou-se o procedimento descrito por Araujo *et al.* (2025), no qual a imagem térmica é processada em conjunto com a matriz de temperaturas. Inicialmente, a região de interesse é identificada pela rede neural convolucional YOLOv5, que fornece a bounding box do equipamento. Em paralelo, o FLIR SDK é utilizado para obter a matriz de temperaturas  $M_{temp}$ . A partir dessas informações, gera-se a matriz auxiliar  $M_{aux}$ , composta pelas temperaturas delimitadas pela região de interesse. Em seguida, aplica-se um limiar baseado na temperatura ambiente, seguido de escalonamento, limiarização adaptativa e abertura morfológica, resultando na matriz binária  $M_{bin}$ . Essa matriz permite estimar o ângulo de inclinação do equipamento e, quando necessário, realizar sua rotação para a posição vertical. Ao final do processo, com a região de interesse devidamente segmentada e alinhada, extrai-se o perfil térmico ao longo do eixo central do equipamento. Na Figura 20, ilustra-se a metodologia proposta por Araujo *et al.* (2025) para a extração do perfil térmico.

Figura 20. Fluxograma de extração do perfil térmico.



Fonte: Adaptado de (Araujo *et al.*, 2025)

Dessa forma, cada perfil térmico unidimensional representa a variação espacial da temperatura ao longo do comprimento do equipamento, permitindo transformar a informação contida na imagem infravermelha em um sinal unidimensional adequado para análise estatística e modelagem sequencial. Essa representação simplifica a análise, preservando a principal distribuição térmica ao longo do isolador.

## 4.3 PRÉ-PROCESSAMENTO DOS PERFIS TÉRMICOS

Após a extração do perfil térmico, os dados foram submetidos a uma etapa de pré-processamento com o objetivo de transformar os registros térmicos em sinais unidimensionais adequados às etapas posteriores. Essa etapa é responsável por padronizar a leitura dos perfis, aplicar correções associadas às condições ambientais e, ampliação da base de dados.

### 4.3.1 IDENTIFICAÇÃO DO GRUPO DE CAPTURA

Além da classe do isolador, cada perfil térmico foi associado a um identificador de grupo de captura. Esse identificador representa o conjunto de amostras obtidas em uma mesmo instante de aquisição. No contexto deste trabalho, perfis pertencentes ao mesmo grupo de captura podem apresentar elevada semelhança entre si, uma vez que compartilham condições ambientais, configuração experimental e posicionamento relativo dos isoladores no mesmo instante de aquisição.

A identificação do grupo de captura é uma etapa essencial para o controle de vazamento de dados. Caso perfis pertencentes ao mesmo grupo fossem distribuídos simultaneamente entre os conjuntos de treino e teste, o modelo poderia ser favorecido por similaridades específicas da aquisição em vez de aprender padrões efetivamente associados à condição do isolador. Esse efeito poderia produzir estimativas excessivamente otimistas de desempenho.

Para evitar esse problema, o identificador de grupo de captura foi utilizado em todas as divisões da base de dados. Assim, amostras pertencentes ao mesmo grupo foram mantidas sempre no mesmo subconjunto experimental, seja ele de treino, validação ou teste. Essa estratégia garante que a avaliação seja realizada em grupos não vistos durante o treinamento, tornando a estimativa de desempenho mais rigorosa e mais próxima de uma situação real de aplicação.

#### 4.3.2 CORREÇÃO PELA TEMPERATURA AMBIENTE

Nesta dissertação, adotou-se como correção principal a subtração da temperatura ambiente do perfil térmico bruto. Dessa forma, para cada ponto do perfil, o valor de temperatura ambiente correspondente ao ensaio é subtraído do valor medido no isolador.

Essa correção transforma o perfil absoluto em um perfil térmico relativo ao ambiente. Com isso, valores positivos indicam regiões do isolador com temperatura superior à temperatura ambiente, enquanto valores próximos de zero indicam regiões com temperatura semelhante à condição ambiental. Essa representação reduz a influência de deslocamentos globais de temperatura entre ensaios realizados sob diferentes condições ambientais.

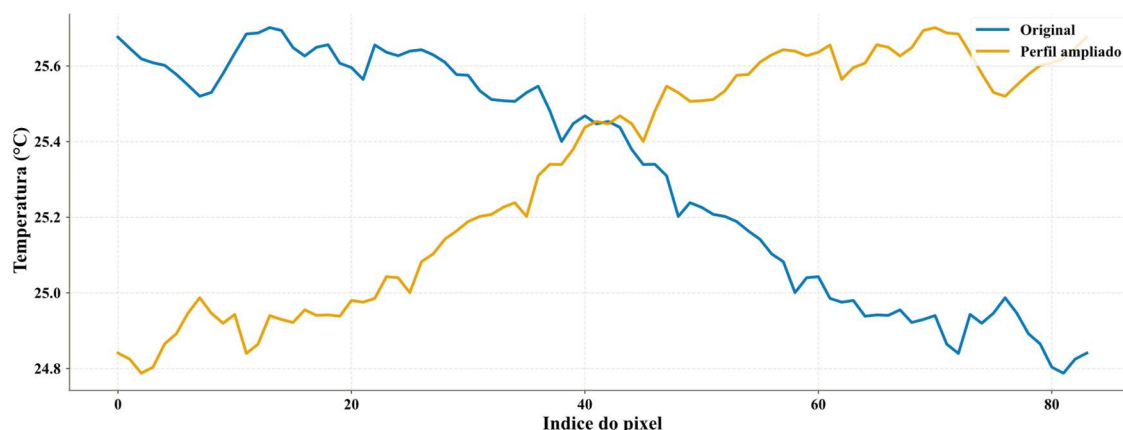
A correção pela temperatura ambiente é especialmente importante porque os experimentos foram realizados em momentos distintos e sob diferentes valores de temperatura e umidade relativa do ar. Sem essa correção, parte das diferenças observadas entre os perfis poderia estar associada às condições externas de aquisição, e não necessariamente ao estado superficial do isolador. Assim, a utilização do perfil relativo contribui para que os atributos extraídos estejam mais associados ao comportamento térmico do equipamento do que às variações ambientais entre os experimentos.

#### 4.3.3 AMPLIAÇÃO DA BASE DE DADOS

Essa etapa tem como objetivo aumentar a diversidade dos exemplos disponíveis durante o treinamento, contribuindo para reduzir a dependência dos modelos em relação à orientação específica do perfil e às variações particulares de cada aquisição.

A principal transformação considerada consiste na inclusão da amostra original e de uma versão reversa do perfil, obtida pela inversão da ordem dos pontos do sinal. Essa operação preserva os valores térmicos observados, mas altera sua orientação espacial. Dessa forma, o modelo passa a observar padrões térmicos equivalentes em sentidos opostos ao longo do perfil, o que pode favorecer a aprendizagem de características mais gerais do comportamento térmico. Na Figura 21 ilustra-se um perfil térmico antes e após o procedimento de ampliação.

Figura 21. Perfil térmico antes e após o procedimento de ampliação.



Fonte: autoria própria.

Apesar de ser descrita como ampliação da base de dados, essa etapa não foi aplicada indiscriminadamente a toda a base experimental. Para evitar vazamento de dados, a ampliação foi realizada somente sobre os dados de treino em cada divisão experimental. Os subconjuntos de validação e teste permanecem compostos apenas por perfis originais, sem versões ampliadas. Essa regra impede que uma transformação derivada de uma amostra de avaliação seja utilizada no treinamento, preservando a independência entre os conjuntos.

Mesmo após a ampliação dos dados, a identidade do grupo de captura original foi preservada. Assim, uma amostra original e sua versão reversa permaneceram associadas ao mesmo identificador de grupo. Essa decisão foi necessária para manter a coerência da separação por grupos e garantir que variações derivadas de uma mesma captura não fossem distribuídas entre subconjuntos distintos. Considerando a base total de 344 perfis térmicos, a divisão aproximada de 70% para treinamento e 30% para teste resultou em 241 perfis originais destinados ao treinamento e 103 perfis destinados ao teste. A ampliação foi aplicada exclusivamente à base de treinamento, dobrando o número de perfis disponíveis nessa etapa, de 241 para 482 perfis, sem alterar o conjunto de teste. Dessa forma, evitou-se que versões artificialmente derivadas de uma mesma captura influenciassem a avaliação final dos modelos.

#### 4.4 SEGMENTAÇÃO DOS PERFIS TÉRMICOS EM JANELAS

Após a extração dos perfis térmicos unidimensionais, cada perfil foi segmentado em janelas ao longo da direção longitudinal do isolador. Essa etapa teve como objetivo permitir uma análise local do comportamento térmico, uma vez que a presença de

poluição pode provocar alterações específicas em determinadas regiões do equipamento, como no lado fase, na região central ou no lado terra.

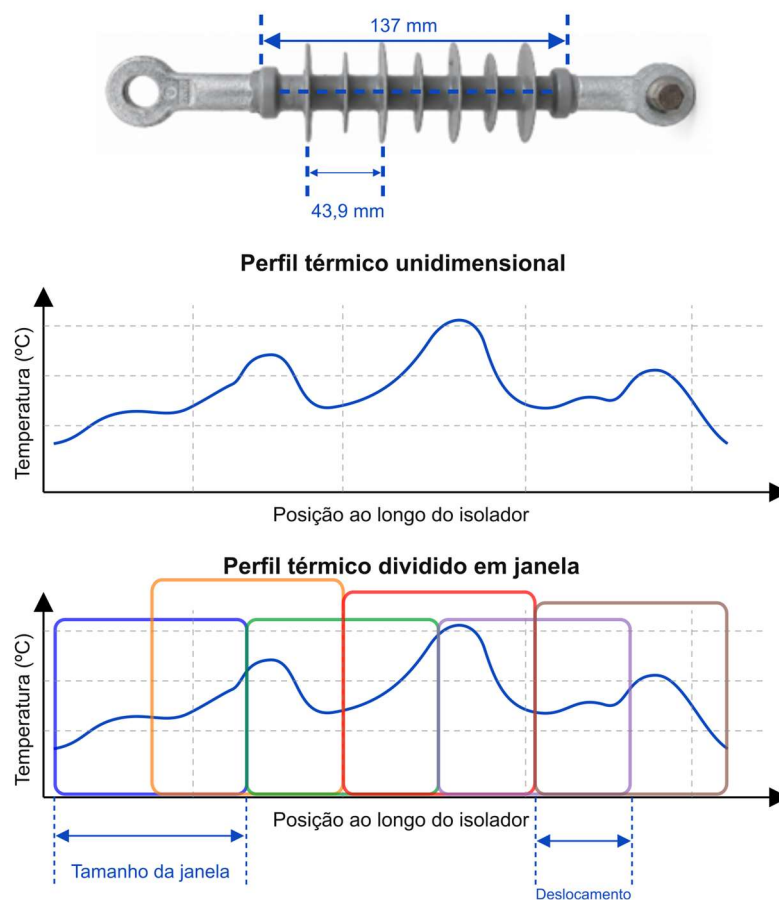
A segmentação em janelas possibilita que variações térmicas localizadas sejam preservadas durante a extração de atributos. Caso o perfil fosse analisado apenas de forma global, alterações pontuais ou regionais poderiam ser suavizadas ou até mesmo mascaradas por estatísticas calculadas sobre todo o comprimento do isolador. Assim, a divisão do perfil em trechos menores permite caracterizar melhor a heterogeneidade térmica associada à distribuição da poluição.

Em uma etapa exploratória inicial, o tamanho da janela e o deslocamento entre segmentos consecutivos foram tratados como parâmetros de configuração e avaliados por meio de uma busca exaustiva, realizada conjuntamente com a definição dos hiperparâmetros dos modelos ocultos de Markov. Nessa avaliação, o critério de informação bayesiano foi utilizado para comparar as diferentes configurações. A análise dos resultados obtidos indicou que os tamanhos de janela e os deslocamentos mais adequados apresentavam relação com características geométricas do isolador, especialmente com a distância entre aletas e o comprimento da região isolante.

A partir dessa constatação, a segmentação adotada neste trabalho passou a ser definida com base na geometria do equipamento, em vez de permanecer como um parâmetro determinado exclusivamente por otimização empírica. Essa escolha relaciona a divisão do perfil térmico à estrutura física do isolador, reduz a arbitrariedade da segmentação e permite que cada janela represente uma região fisicamente significativa do componente.

O deslocamento entre janelas consecutivas foi definido a partir de um passo associado ao tamanho da própria janela. No protocolo adotado, utilizou-se um passo correspondente à metade da janela, permitindo a sobreposição parcial entre segmentos consecutivos. Essa sobreposição aumenta a continuidade da análise espacial, reduz perdas de informação entre janelas adjacentes e favorece a detecção de transições térmicas graduais ao longo do perfil. Na Figura 22 ilustra-se o processo de segmentação do perfil térmico em janelas.

Figura 22. Processo de segmentação do perfil térmico em janelas.



Fonte: autoria própria.

Ao final da segmentação, cada perfil térmico encontra-se organizado como uma sequência de janelas, a qual serve como base para a etapa subsequente de extração de atributos. Nessa etapa, são calculadas medidas capazes de descrever o comportamento térmico local de cada segmento, preservando a organização espacial do perfil e possibilitando uma caracterização mais detalhada das variações térmicas ao longo do isolador.

## 4.5 EXTRAÇÃO DOS ATRIBUTOS TÉRMICOS

Após a segmentação dos perfis térmicos em janelas, realizou-se a extração dos atributos térmicos utilizados nas representações baseadas em características contínuas. Essa etapa teve como objetivo transformar cada trecho local do perfil em um vetor numérico capaz de descrever o comportamento térmico da região correspondente do isolador.

Os atributos considerados nesta etapa foram previamente definidos e equacionados na Seção 2.4. Assim, nesta seção, a extração é descrita sob o ponto de vista metodológico, enfatizando sua aplicação aos perfis térmicos segmentados e sua utilização nas etapas posteriores de representação e classificação.

Para cada janela do perfil térmico, foram calculados atributos pertencentes a diferentes grupos descritivos, incluindo atributos de nível térmico, dispersão e contraste, forma da distribuição, variação espacial e dependência espacial. Esses grupos foram adotados com o objetivo de representar propriedades complementares do sinal térmico, contemplando tanto características associadas à intensidade dos valores de temperatura quanto aspectos relacionados à heterogeneidade e à organização espacial do perfil.

Os atributos de nível térmico foram utilizados para caracterizar a magnitude dos valores de temperatura em cada janela. Os atributos de dispersão e contraste permitiram descrever a variabilidade térmica local. Os atributos de forma da distribuição foram empregados para sintetizar propriedades estatísticas da distribuição dos valores dentro da janela. Já os atributos de variação espacial foram utilizados para representar mudanças ao longo da posição no perfil, enquanto a dependência espacial foi considerada para descrever a relação entre pontos consecutivos do sinal.

## 4.6 REPRESENTAÇÕES DO PERFIL TÉRMICO

Após a extração dos atributos, os dados foram organizados em diferentes formas de representação. Essa etapa foi necessária porque os modelos avaliados neste trabalho operam sobre estruturas de entrada distintas. Enquanto alguns classificadores utilizam vetores de atributos de dimensão fixa, outros exploram a evolução sequencial dos atributos ao longo do perfil ou sequências simbólicas obtidas pela discretização do sinal térmico.

Nesta dissertação foram consideradas três representações principais: representação sequencial, representação agregada e representação simbólica. As duas primeiras são baseadas em atributos contínuos extraídos das janelas dos perfis térmicos. A terceira é baseada na discretização dos valores do perfil, produzindo sequências de símbolos utilizadas por modelos probabilísticos discretos.

#### 4.6.1 REPRESENTAÇÃO SEQUENCIAL

Na representação sequencial cada perfil térmico é descrito por uma sequência ordenada de vetores de atributos. Cada vetor corresponde aos atributos calculados em uma janela específica do perfil, preservando a posição relativa dessa janela ao longo do isolador.

Essa representação mantém a organização espacial do sinal térmico, permitindo que os modelos explorem a evolução dos atributos entre regiões consecutivas. Dessa forma, a amostra não é descrita apenas pelo conjunto de valores extraídos, mas também pela forma como esses valores se distribuem ao longo do perfil.

A representação sequencial é especialmente adequada para investigar alterações térmicas localizadas ou distribuídas de forma não homogênea. Em condições de poluição não uniforme, por exemplo, a contaminação pode estar concentrada em regiões específicas do isolador, produzindo mudanças térmicas cuja posição ao longo do perfil é relevante para a classificação. Assim, preservar a ordem das janelas permite que o modelo avalie padrões espaciais associados à condição do isolador.

Neste trabalho, a representação sequencial foi utilizada pelos modelos capazes de operar sobre sequências de atributos, como o HMM com emissões Gaussianas (HMM-GMM) e a LSTM.

#### 4.6.2 REPRESENTAÇÃO AGREGADA

Na representação agregada, os atributos calculados localmente em cada janela são resumidos ao longo de todo o perfil térmico por meio de estatísticas descritivas. Com isso, cada perfil deixa de ser representado como uma sequência de vetores e passa a ser descrito por um único vetor de dimensão fixa.

Essa representação sintetiza o comportamento global das propriedades térmicas extraídas nas janelas. Por exemplo, um atributo calculado localmente pode ser resumido por medidas como média, desvio-padrão, mínimo, máximo, percentis, mediana, assimetria ou curtose ao longo do perfil. Dessa forma, a representação agregada mantém informações sobre o nível global, a variabilidade, os extremos e a distribuição espacial dos descritores térmicos, ainda que não preserve explicitamente a ordem completa das janelas.

A principal vantagem dessa representação é permitir a utilização de classificadores tabulares, que exigem uma matriz de entrada em que cada linha corresponde a uma amostra e cada coluna corresponde a uma característica. Assim, a representação agregada fornece uma descrição compacta do perfil térmico, adequada para modelos como SVM e Random Forest.

Embora a agregação reduza a informação sequencial explícita, ela preserva indiretamente aspectos importantes da heterogeneidade térmica ao resumir como cada atributo varia ao longo das janelas. Por isso, essa representação permite avaliar se descritores globais derivados da representação sequencial são suficientes para diferenciar isoladores limpos e poluídos.

#### 4.6.3 REPRESENTAÇÃO SIMBÓLICA

Além das representações baseadas em atributos contínuos, este trabalho também considerou uma representação simbólica dos perfis térmicos. Nessa abordagem, o perfil térmico unidimensional é convertido em uma sequência de símbolos discretos por meio de um processo de quantização.

A discretização foi realizada a partir do método iterativo de Lloyd (Max, 1960; Lloyd, 1982). Para cada configuração experimental, o número de símbolos define a cardinalidade do alfabeto utilizado para representar o perfil. Assim, os valores contínuos de temperatura são substituídos por símbolos associados aos intervalos definidos pelo quantizador.

O discretizador é ajustado exclusivamente com os perfis pertencentes ao conjunto de treino de cada divisão experimental. Após esse ajuste, o mesmo discretizador é aplicado aos perfis de validação ou teste. Essa regra impede que a distribuição dos dados de avaliação influencie os limiares de discretização, preservando a independência entre treino e avaliação.

A representação simbólica permite que o perfil térmico seja analisado como uma sequência discreta, possibilitando o uso de modelos probabilísticos baseados em transições entre símbolos. Nesta dissertação, essa representação foi utilizada pelo HMM discreto, que modela a sequência observada por meio de estados ocultos e probabilidades de emissão simbólica.

Dessa forma, as três representações adotadas permitem avaliar diferentes formas de explorar a informação térmica dos isoladores. A representação sequencial preserva a

organização espacial dos atributos contínuos, a representação agregada resume o comportamento global dos descritores e a representação simbólica transforma o perfil em uma sequência discreta adequada à modelagem probabilística.

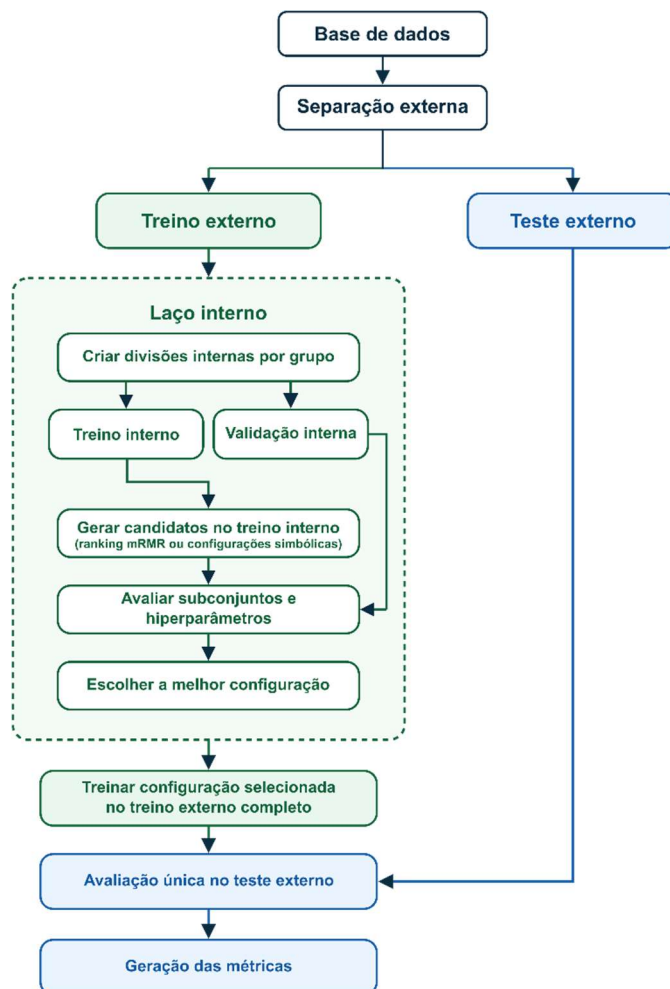
## 4.7 VALIDAÇÃO ANINHADA

A avaliação principal dos modelos foi conduzida por meio de validação aninhada com separação por grupos de captura. Esse protocolo foi adotado para estimar a capacidade de generalização dos classificadores de forma mais rigorosa, evitando que amostras obtidas em uma mesma condição de aquisição fossem distribuídas simultaneamente entre conjuntos de treino, validação e teste. Conforme discutido por Raschka e Mirjalili (2019), a validação cruzada e os procedimentos de ajuste de hiperparâmetros são etapas fundamentais para a avaliação adequada de modelos de aprendizado de máquina, especialmente quando diferentes configurações precisam ser comparadas antes da avaliação final (Raschka e Mirjalili, 2019).

A validação aninhada é composta por dois níveis: o laço externo e o laço interno. O laço externo é responsável pela estimativa final de desempenho, mantendo um conjunto de teste externo isolado durante todo o processo de seleção de atributos, escolha de hiperparâmetros e definição de configurações. O laço interno, por sua vez, é utilizado exclusivamente para tomada de decisões de modelagem, como seleção de subconjuntos de atributos, escolha do número de atributos, ajuste de hiperparâmetros e definição das configurações simbólicas. Essa organização é coerente com a recomendação de Lewis *et al.* (2023), que destacam a validação aninhada como uma estratégia adequada quando a seleção de atributos e o ajuste de modelos precisam ser incorporados ao processo de avaliação, evitando estimativas enviesadas de desempenho (Lewis *et al.*, 2023).

Neste trabalho, foram utilizadas cinco repetições externas. Em cada repetição externa, os grupos de captura foram separados em treino externo e teste externo. Dentro de cada treino externo, foram geradas dez divisões internas por grupo de captura, utilizadas para seleção de atributos, escolha de subconjuntos e definição de configurações. Dessa forma, a avaliação final foi baseada em múltiplas separações externas, enquanto as decisões de modelagem foram tomadas apenas a partir das divisões internas do respectivo treino externo. Na Figura 23 ilustra-se o fluxograma da validação aninhada.

Figura 23. Fluxograma da validação aninhada.



Fonte: autoria própria.

## 4.8 SELEÇÃO DE ATRIBUTOS E CONFIGURAÇÃO

Após a construção das representações dos perfis térmicos, foi realizada a etapa de seleção de atributos e configurações dos modelos. Essa etapa teve como objetivo reduzir a dimensionalidade das representações contínuas, evitar o uso de atributos redundantes ou pouco informativos e definir, de forma controlada, as configurações mais adequadas para cada família de classificador avaliada.

### 4.8.1 SELEÇÃO DE ATRIBUTOS NA REPRESENTAÇÃO CONTÍNUA

A seleção de atributos foi aplicada às representações baseadas em atributos contínuos, contemplando tanto a representação agregada quanto a representação sequencial. Embora essas representações tenham estruturas distintas, em ambos os casos, a seleção foi conduzida com o objetivo de identificar descritores térmicos relevantes para

a classificação binária entre as classes Limpo e Poluído, evitando a inclusão de variáveis redundantes.

Na representação agregada, cada perfil térmico foi descrito por um vetor de dimensão fixa. Nesse caso, cada coluna da matriz de atributos foi tratada como uma variável candidata. A relevância de cada atributo foi estimada pelo valor F ANOVA, e a redundância foi calculada pela correlação absoluta de Pearson entre o atributo candidato e os atributos previamente selecionados, conforme conceitos apresentados na Seção 2.5.

Na representação sequencial, cada perfil foi descrito por uma sequência de vetores de atributos, preservando a ordem espacial das janelas. Para essa representação, a seleção foi conduzida no nível de grupos de atributos. Cada descritor sequencial foi inicialmente representado por estatísticas-resumo, utilizadas apenas para ordenação. Assim, a relevância de um grupo foi calculada a partir das relevâncias de suas estatísticas, enquanto a redundância foi estimada pela média das correlações absolutas entre as estatísticas do grupo candidato e as estatísticas dos grupos já selecionados. Após a seleção, o modelo sequencial recebeu as sequências correspondentes aos atributos escolhidos.

Para a ordenação, inicialmente, selecionou-se o atributo com maior relevância individual em relação à classe. Em seguida, a cada nova iteração, os atributos candidatos ainda não selecionados foram avaliados considerando o equilíbrio entre relevância e redundância. No protocolo adotado, utilizou-se o critério por quociente, no qual a relevância do atributo candidato é ponderada pela redundância média em relação ao conjunto já selecionado (Zhao, Anand e Wang, 2019).

Para aumentar a estabilidade da seleção, a ordenação não foi obtida a partir de uma única divisão interna. Em cada repetição externa, o mRMR foi recalculado em múltiplas divisões internas por grupo de captura, sempre utilizando apenas o treino interno correspondente. As colocações produzidas nessas divisões foram então agregados para formar um ordenamento estável. Essa agregação considerou, principalmente, a frequência com que cada atributo apareceu entre as primeiras posições, sua posição média e seu valor mRMR médio.

Esse procedimento reduz a dependência da seleção em relação a uma única partição dos dados. Assim, atributos selecionados de forma recorrente em diferentes divisões internas recebem maior prioridade no ranking estável, enquanto atributos que aparecem apenas ocasionalmente tendem a ter menor influência na composição dos subconjuntos candidatos.

#### 4.8.2 CONSTRUÇÃO DOS SUBCONJUNTOS

A partir da ordenação estável produzida pelo processo de seleção, foram construídos subconjuntos cumulativos de atributos. Nessa abordagem, o subconjunto com um atributo contém apenas o atributo mais bem posicionado; o subconjunto com dois atributos contém os dois primeiros atributos; e assim sucessivamente, até o maior valor de  $k$  definido no protocolo experimental.

Essa estratégia permitiu avaliar a contribuição progressiva dos atributos à medida que novas variáveis são adicionadas ao modelo. Um atributo bem posicionado individualmente pode ser suficiente para produzir bom desempenho em determinados classificadores, enquanto outros modelos podem se beneficiar da combinação de múltiplos descritores complementares. Assim, a análise por subconjuntos permite investigar o equilíbrio entre desempenho, complexidade e interpretabilidade.

Os subconjuntos foram construídos separadamente para cada tipo de representação. Para os modelos tabulares, como SVM e Random Forest, os subconjuntos *top-k* foram formados a partir da ordenação estável dos atributos agregados. Para os modelos sequenciais, como HMM contínuo e LSTM, os subconjuntos foram formados a partir da ordenação estável dos grupos de atributos sequenciais.

A escolha do valor de  $k$  foi realizada no laço interno da validação aninhada, utilizando apenas os dados disponíveis no treino externo de cada repetição. Após a construção do ranking estável, os subconjuntos candidatos foram avaliados em divisões internas Monte Carlo por grupo de captura, utilizando o próprio classificador correspondente. Dessa forma, a seleção do número de atributos foi dependente do modelo, permitindo que cada família de classificador adotasse o subconjunto mais adequado à sua estrutura de entrada e ao seu viés de aprendizagem.

A métrica utilizada para selecionar o subconjunto vencedor foi definido a partir da média entre métricas calculadas no nível de grupo de captura, incluindo acurácia balanceada, F1 macro e coeficiente de correlação de Matthews. A priorização das métricas por grupo foi adotada porque ela reduz a influência de múltiplos perfis correlacionados pertencentes à mesma captura.

Após a escolha do subconjunto vencedor, os atributos selecionados foram congelados. Em seguida, o modelo foi treinado utilizando todo o conjunto de treino externo e avaliado uma única vez no teste externo. Dessa forma, o teste externo não

participou da definição da ordenação, da escolha do subconjunto ou de qualquer outra decisão do modelo.

#### 4.8.3 SELEÇÃO DAS CONFIGURAÇÕES SIMBÓLICAS

Para os modelos baseados em representação simbólica, a etapa de seleção não envolveu mRMR nem subconjuntos de atributos contínuos. Nessa abordagem, os modelos operam diretamente sobre sequências simbólicas obtidas pela discretização dos perfis térmicos. Portanto, a seleção concentrou-se nas configurações associadas à discretização e à modelagem probabilística das sequências.

As configurações avaliadas incluíram o número de símbolos utilizado na quantização Lloyd-Max e o número de estados ocultos do HMM discreto. Esses parâmetros influenciam diretamente a forma como o perfil térmico contínuo é convertido em uma sequência discreta e como os padrões de transição entre símbolos são modelados.

Neste trabalho a grade de configurações considera cinco valores para o número de símbolos: 4, 6, 8, 10 e 12. Para cada cardinalidade do alfabeto simbólico, avaliam-se diferentes números de estados ocultos para os modelos das classes Limpo e Poluído, variando de 2 a 5 estados em cada classe. Assim, a combinação entre número de símbolos, estados do HMM da classe Limpo e estados do HMM da classe Poluído resulta em 80 configurações candidatas.

A seleção dessas configurações foi realizada no laço interno da validação aninhada. Em cada divisão interna, o discretizador Lloyd-Max foi ajustado exclusivamente com os perfis do treino interno e aplicado posteriormente aos perfis de validação interna. Em seguida, as diferentes combinações de parâmetros foram avaliadas com base nas métricas definidas no protocolo experimental.

Após a escolha da configuração simbólica mais adequada no laço interno, o discretizador e o modelo correspondente foram ajustados novamente utilizando todo o treino externo. Somente depois dessa etapa o modelo foi avaliado no teste externo. Esse procedimento assegura que os limiares de discretização, os parâmetros de modelagem e a escolha da configuração sejam determinados sem acesso às amostras de teste.

Dessa forma, a seleção de atributos e configurações foi conduzida de maneira específica para cada família de modelo. Nos modelos contínuos, a seleção envolveu ranqueamento por relevância e redundância, agregação dos rankings internos, construção de subconjuntos e escolha do número de atributos. Nos modelos simbólicos, a seleção

envolveu a escolha dos parâmetros de discretização e dos modelos probabilísticos. Em ambos os casos, todas as decisões foram restritas aos dados de treinamento de cada etapa da validação, preservando a independência do teste externo.

## 4.9 TREINAMENTO DOS MODELOS DE CLASSIFICAÇÃO

Os modelos de classificação foram treinados para a tarefa binária de distinção entre as classes Limpo e Poluído.

Os classificadores foram incorporados de acordo com a representação de dados exigida por cada família de modelo. Os modelos SVM e Random Forest utilizaram vetores de atributos agregados; os modelos HMM contínuo e LSTM utilizaram sequências de vetores extraídos das janelas dos perfis térmicos; e o HMM discreto operou sobre sequências simbólicas obtidas pela discretização dos perfis.

O objetivo desta etapa não foi selecionar previamente um único classificador, mas submeter modelos com diferentes mecanismos de aprendizagem ao mesmo princípio de avaliação. Assim, todos os modelos foram treinados respeitando a separação por grupos de captura, a seleção de atributos ou configurações apenas com dados de treino e a avaliação final em grupos não vistos.

### 4.9.1 AJUSTE DE HIPERPARÂMETROS

Nesta dissertação, a seleção de hiperparâmetros foi tratada como parte do ajuste do modelo e, portanto, ocorreu sem acesso ao conjunto de teste externo. Essa decisão evita que o desempenho no teste influencie a escolha das configurações utilizadas no treinamento final.

Para os modelos SVM e Random Forest, o ajuste de hiperparâmetros realizou-se por busca em grade, com validação cruzada *k-fold* respeitando os grupos de captura (Raschka e Mirjalili, 2019). Nesse procedimento, combinações candidatas de hiperparâmetros são avaliadas nas divisões internas, e a configuração de melhor desempenho é selecionada para o treinamento posterior. No SVM, os principais parâmetros avaliados foram associados ao *kernel* RBF, especialmente os parâmetros de penalização e de alcance do kernel. Na *Random Forest*, foram avaliados parâmetros relacionados à quantidade de árvores, profundidade das árvores, critérios mínimos de divisão, amostragem, poda e ponderação das classes.

Para o HMM com emissões gaussianas, o ajuste de configuração envolveu principalmente o número de estados ocultos e o número de componentes da mistura. As combinações candidatas foram ajustadas separadamente para cada classe, e a seleção foi realizada considerando o critério de informação bayesiana (Murphy, 2012). Para a LSTM, devido ao maior custo computacional, o ajuste de hiperparâmetros foi realizado por busca aleatória, contemplando parâmetros como número de unidades recorrentes, *dropout*, taxa de aprendizagem e o tamanho do lote (*batchsize*).

No ramo simbólico, o ajuste combinou parâmetros da discretização com parâmetros dos classificadores. As configurações avaliadas incluíram o número de símbolos da quantização e o número de estados do HMM discreto definidos a partir do critério de informação bayesiana (Murphy, 2012).

Ademais, para os Modelos Ocultos de Markov foram ajustados modelos independentes para as classes Limpo e Poluído. Cada modelo foi estimado utilizando apenas as sequências pertencentes à sua respectiva classe no conjunto de treino. As configurações associadas ao número de estados ocultos e ao número de componentes da mistura foram definidas previamente ou avaliadas no processo interno de seleção, sem acesso ao teste externo.

Durante a classificação, cada sequência foi avaliada nos modelos das duas classes. O rótulo final foi atribuído à classe cujo modelo apresentou maior verossimilhança para a sequência observada.

Por fim, com uma configuração de hiperparâmetros já definida em uma etapa de referência, o algoritmo permitiu reutilizá-la como configuração fixa. Nessa situação, não ocorreu uma nova busca para cada repetição; os valores foram congelados, o modelo foi reajustado nos dados de treino correspondentes e apenas seu desempenho foi reavaliado. Essa separação evita confundir a variabilidade do treinamento com a variabilidade decorrente de novas escolhas de hiperparâmetros.

#### 4.9.2 PREDIÇÃO POR GRUPO DE CAPTURA

As predições foram inicialmente obtidas no nível de perfil térmico. Posteriormente, as predições dos perfis pertencentes ao mesmo grupo de captura foram agregadas por voto majoritário, produzindo uma predição no nível do grupo. As métricas foram registradas nos dois níveis, por perfil e por grupo, sendo o desempenho por grupo utilizado como referência principal na comparação dos modelos.

Dessa forma, o treinamento dos classificadores foi conduzido sob uma mesma estrutura experimental: seleção e ajuste apenas nos dados de treino, congelamento da configuração selecionada, reajuste final com o treino externo completo e avaliação única no teste externo. Essa organização permitiu comparar diferentes modelos e representações, mantendo o controle de vazamento de dados entre treino e avaliação.

#### 4.9.3 AVALIAÇÃO POR PERFIL TÉRMICO

Na avaliação por perfil, cada amostra recebeu uma predição individual do classificador. Essa análise permitiu verificar o desempenho direto do modelo sobre os perfis térmicos extraídos das termografias, considerando cada perfil como uma unidade independente de predição.

As métricas por perfil foram calculadas a partir da comparação entre os rótulos verdadeiros e os rótulos preditos para todos os perfis pertencentes ao conjunto de teste da repetição externa. Essa avaliação é útil para identificar a capacidade do modelo de classificar corretamente perfis individuais, bem como para analisar erros associados a variações locais ou específicas de determinados perfis.

Apesar de sua utilidade, a avaliação por perfil pode superestimar a quantidade efetiva de amostras independentes quando vários perfis pertencem ao mesmo grupo de captura. Por isso, os resultados por perfil foram complementados pela avaliação no nível dos grupos.

#### 4.9.4 AVALIAÇÃO POR GRUPO DE CAPTURA

Na avaliação por grupo de captura, as predições individuais dos perfis pertencentes ao mesmo grupo foram agregadas para produzir uma única decisão final. Essa agregação foi realizada por voto majoritário, no qual a classe atribuída ao grupo corresponde à classe mais frequente entre as predições dos perfis daquele grupo.

Essa estratégia foi adotada porque perfis provenientes de uma mesma captura tendem a compartilhar condições experimentais semelhantes. Assim, a avaliação por grupo reduz a influência de múltiplos perfis correlacionados e fornece uma estimativa mais conservadora do desempenho dos modelos.

As métricas por grupo foram calculadas a partir da comparação entre o rótulo verdadeiro de cada grupo de captura e o rótulo obtido após o voto majoritário. Esse nível

de avaliação foi utilizado como referência principal na seleção e comparação dos modelos, especialmente no laço interno da validação aninhada.

O critério de seleção dos subconjuntos de atributos, hiperparâmetros e configurações foi baseado em uma medida multicritério calculado no nível dos grupos de captura. Essa medida considerou a média entre acurácia balanceada, F1 macro e coeficiente de correlação de Matthews. Dessa forma, foram favorecidas configurações capazes de apresentar desempenho equilibrado entre as classes e consistente em grupos de captura não vistos.

A análise conjunta das métricas por perfil e por grupo permitiu verificar se o desempenho observado em amostras individuais se manteve quando as predições foram agregadas em decisões por captura.

Dessa forma, a avaliação adotada nesta dissertação buscou não apenas medir a taxa de acerto dos classificadores, mas também considerar o equilíbrio entre classes, a robustez das decisões e a independência experimental dos grupos de captura.

#### 4.9.5 AMBIENTE COMPUTACIONAL E IMPLEMENTAÇÃO

Todos os procedimentos computacionais utilizados nesta dissertação foram implementados em ambiente local, em um computador com sistema operacional Windows 11 Pro, processador AMD Ryzen 5 5600X, 16 GB de memória RAM e placa de vídeo NVIDIA GeForce RTX 3060 com 12 GB de memória dedicada. Esse ambiente foi utilizado para as etapas de organização da base de dados, pré-processamento dos perfis térmicos, extração e seleção de atributos, treinamento dos modelos, ajuste de hiperparâmetros, validação aninhada e geração dos resultados.

A implementação foi realizada em linguagem Python, com dependências mínimas definidas para garantir maior reprodutibilidade dos experimentos. As bibliotecas NumPy, Pandas e OpenPyXL foram utilizadas para manipulação de dados, organização das tabelas e leitura ou gravação de arquivos em formato de planilha. As bibliotecas SciPy e Statsmodels foram empregadas em operações estatísticas e análises auxiliares. A biblioteca Scikit-learn foi utilizada na implementação dos classificadores de entrada fixa, nas etapas de padronização, validação, cálculo de métricas e procedimentos de seleção e avaliação dos modelos. Para a otimização de hiperparâmetros, utilizou-se a biblioteca Optuna, enquanto a biblioteca TensorFlow foi empregada na implementação e treinamento dos modelos baseados em redes LSTM. A seleção de atributos pelo critério

mRMR foi realizada com apoio da biblioteca mrmr-selection. As bibliotecas Matplotlib e Seaborn foram utilizadas para a construção das figuras e visualização dos resultados. Por fim, Joblib e tqdm foram utilizadas, respectivamente, para armazenamento de objetos computacionais, paralelização ou persistência de resultados e acompanhamento do progresso das rotinas executadas.

As dependências mínimas utilizadas no ambiente computacional foram: NumPy  $\geq 1.21.0$ , Pandas  $\geq 1.3.0$ , Scikit-learn  $\geq 1.0.0$ , SciPy  $\geq 1.7.0$ , Statsmodels  $\geq 0.14.0$ , Matplotlib  $\geq 3.4.0$ , Seaborn  $\geq 0.11.0$ , Joblib  $\geq 1.0.0$ , tqdm  $\geq 4.62.0$ , Optuna  $\geq 3.0.0$ , TensorFlow  $\geq 2.8.0$ , mrmr-selection  $\geq 0.2.8$  e OpenPyXL  $\geq 3.1.5$ .

## 4.10 ANÁLISE DE ROBUSTEZ DOS CAMPEÕES

Após a avaliação dos modelos no protocolo de validação aninhada, foi realizada uma análise complementar de robustez com as configurações campeãs. O objetivo dessa etapa foi verificar a estabilidade do desempenho dos melhores modelos diante de novas repetições de treinamento e avaliação, mantendo congeladas as decisões já selecionadas anteriormente.

A análise de robustez não teve como finalidade realizar uma nova seleção de atributos, hiperparâmetros ou configurações. Em vez disso, os subconjuntos de atributos, os parâmetros dos classificadores e as configurações simbólicas escolhidas na etapa principal foram mantidos fixos. Dessa forma, a análise buscou estimar a variabilidade do desempenho associada às divisões dos dados, à inicialização dos modelos e ao próprio processo de treinamento, sem introduzir novas decisões de seleção.

Essa etapa é importante porque, em bases experimentais com número limitado de grupos de captura, pequenas mudanças na composição dos conjuntos de treino e teste podem influenciar o desempenho observado. Assim, a reavaliação dos modelos campeões em múltiplas repetições permite verificar se os resultados obtidos na validação aninhada são consistentes ou se dependem fortemente de uma divisão específica dos dados.

Para cada modelo campeão a configuração selecionada foi congelada antes da execução das repetições adicionais. Em seguida, o modelo foi reajustado nas novas divisões de treino correspondentes e avaliado nos respectivos conjuntos de teste, sempre respeitando a separação por grupos de captura. Esse procedimento manteve o mesmo princípio de controle de vazamento adotado nas etapas anteriores.

Os resultados da análise de robustez foram consolidados por meio de estatísticas descritivas das métricas de desempenho, incluindo medidas de tendência central e dispersão. Foram consideradas principalmente a acurácia balanceada, o F1 macro e o coeficiente de correlação de Matthews, calculados no nível dos grupos de captura. A variabilidade dessas métricas foi utilizada para avaliar a estabilidade dos modelos campeões

Além do desempenho médio, a análise também considerou a dispersão dos resultados. Modelos com desempenho elevado, mas alta variabilidade entre repetições, foram interpretados com cautela, pois podem indicar sensibilidade às divisões dos dados. Por outro lado, modelos com desempenho médio competitivo e menor variabilidade foram considerados mais estáveis para o problema analisado.

Dessa forma, a análise de robustez dos campeões complementou a validação aninhada, permitindo avaliar não apenas quais modelos apresentaram melhor desempenho, mas também quais configurações mantiveram comportamento mais consistente em diferentes repetições experimentais.

## 5 RESULTADOS E DISCUSSÕES

Neste Capítulo, apresentam-se e discutem-se os resultados obtidos a partir da metodologia proposta para a classificação binária de perfis térmicos de isoladores poliméricos. A análise foi conduzida com o objetivo de avaliar a capacidade de diferentes formas de representação dos perfis térmicos em distinguir isoladores em condição limpa de isoladores submetidos à poluição superficial. Para isso, foram consideradas representações baseadas em atributos agregados, atributos sequenciais e sequências simbólicas, as quais foram avaliadas por classificadores pertencentes a diferentes famílias metodológicas.

A apresentação dos resultados foi organizada de modo a acompanhar a sequência lógica do procedimento experimental. Inicialmente caracteriza-se a base de dados utilizada, descrevendo sua composição, a distribuição das amostras entre classes e subtipos experimentais, bem como aspectos relevantes dos perfis térmicos extraídos. Em seguida, são analisados os atributos térmicos obtidos a partir dos perfis segmentados, destacando sua relação com a separação entre as classes e sua contribuição para a construção das representações utilizadas pelos modelos.

Na sequência são apresentados os resultados da seleção de atributos e configurações, considerando tanto os modelos baseados em atributos contínuos quanto os modelos baseados em representação simbólica. Posteriormente, discute-se o desempenho dos classificadores no protocolo de validação aninhada, com ênfase nas métricas calculadas por perfil térmico e por grupo de captura.

Além da comparação direta entre modelos, este capítulo também discute o efeito das diferentes representações do perfil térmico sobre o desempenho da classificação. Dessa forma, busca-se verificar se a preservação da organização espacial do perfil, por meio de representações sequenciais ou simbólicas, fornece vantagens em relação à representação agregada. Por fim, são analisadas as matrizes de confusão, os padrões de erro e a robustez dos modelos campeões, permitindo avaliar não apenas o desempenho médio, mas também a estabilidade e a confiabilidade das configurações selecionadas.

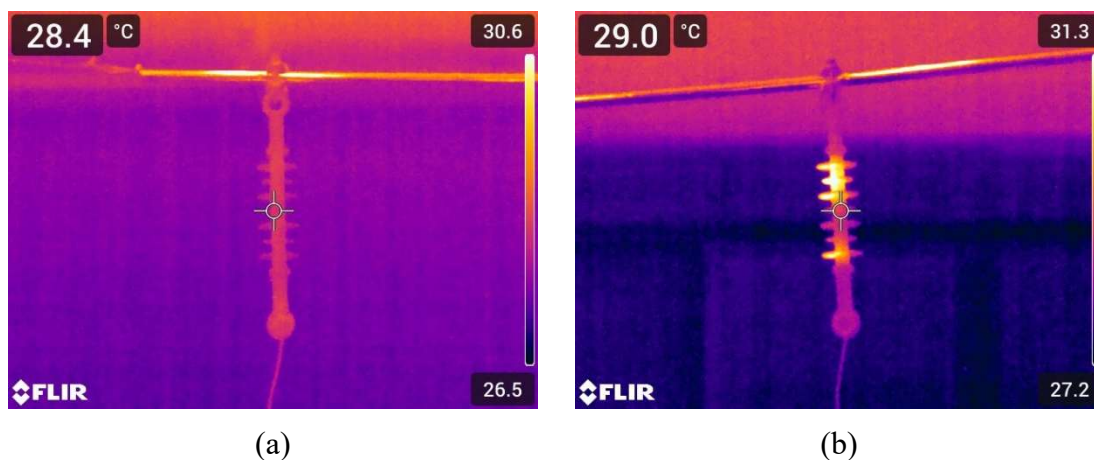
Assim, os resultados são interpretados considerando três aspectos principais: o desempenho quantitativo dos classificadores, a coerência física dos atributos e representações utilizados e a capacidade de generalização dos modelos diante de grupos

de captura não vistos durante o treinamento. Essa abordagem permite que a discussão ultrapasse a simples comparação de métricas, relacionando os resultados obtidos às características experimentais da base e às hipóteses metodológicas adotadas no trabalho.

## 5.1 CARACTERIZAÇÃO DA BASE DE DADOS E DOS PERFIS TÉRMICOS

A análise dos resultados inicia-se pela caracterização da base de dados utilizada nos experimentos, uma vez que a composição das amostras, a distribuição entre classes e a organização por grupos de captura influenciam diretamente a interpretação do desempenho dos modelos de classificação. A base considerada neste trabalho é composta por 344 perfis térmicos unidimensionais extraídos de termografias de isoladores poliméricos, organizados segundo uma formulação binária do problema: isoladores em condição limpa e isoladores submetidos à poluição superficial. Na Figura 24 ilustram-se termografias de um isolador limpo e poluído.

Figura 24. Termografia de isoladores.  
(a) Limpo; (b) Poluído naturalmente.



Fonte: autoria própria.

Na Tabela 7 é apresentada a composição da base de dados por classe e subtipo experimental.

Tabela 7. Composição da base de dados por classe e subtipo experimental.

| Classe  | Subtipo  | Número de perfis | Número de grupos |
|---------|----------|------------------|------------------|
| Limpo   | Limpo    | 124              | 25               |
| Poluído | Fase     | 20               | 4                |
| Poluído | Terra    | 25               | 5                |
| Poluído | Uniforme | 55               | 11               |
| Poluído | Natural  | 120              | 24               |

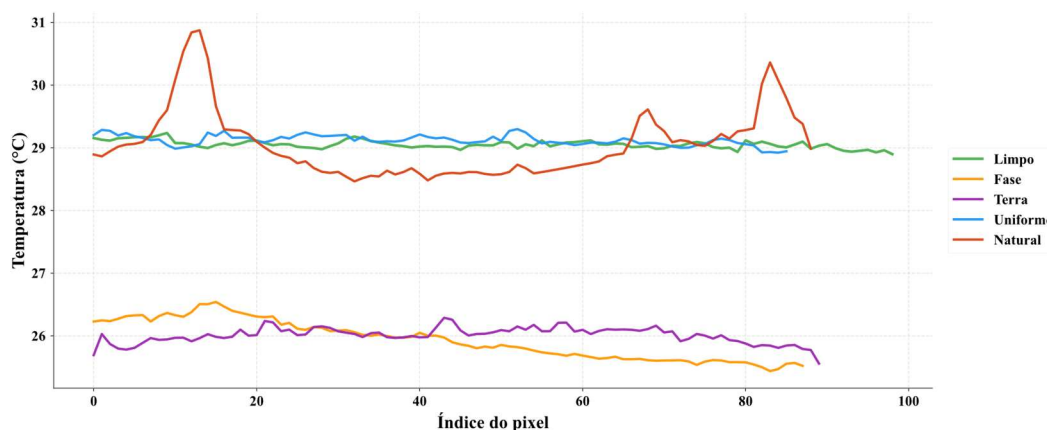
Fonte: autoria própria.

Observa-se que os 344 perfis térmicos estão distribuídos em 69 grupos de captura. Desse total, 124 perfis pertencem à classe Limpo, distribuídos em 25 grupos, enquanto 220 perfis pertencem à classe Poluído, distribuídos em 44 grupos. Entre os perfis poluídos, verifica-se a presença de subtipos associados a diferentes formas de contaminação, incluindo perfis com poluição concentrada no lado fase, no lado terra, com distribuição uniforme e perfis com poluição proveniente de névoa salina costeira, caracterizado como natural.

A distribuição das amostras evidencia que a base apresenta maior quantidade de perfis e grupos associados à classe Poluído. Esse aspecto é relevante para a avaliação dos classificadores, pois a simples proporção de acertos pode não representar adequadamente o desempenho quando há diferença no número de amostras entre as classes.

Ademais, a caracterização dos perfis térmicos também permite compreender a natureza dos sinais utilizados como entrada dos modelos. Os perfis possuem comprimento variável, entre 81 e 101 pontos, o que decorre do processo de extração do perfil térmico central a partir das imagens infravermelhas e das variações de enquadramento e posicionamento do isolador durante as aquisições. Ainda que exista essa variação no número de pontos, os perfis preservam a informação espacial fundamental do problema, isto é, a distribuição da temperatura ao longo do eixo longitudinal do isolador. Na Figura 25 são ilustrados exemplos de perfis térmicos.

Figura 25. Exemplos de perfis térmicos.

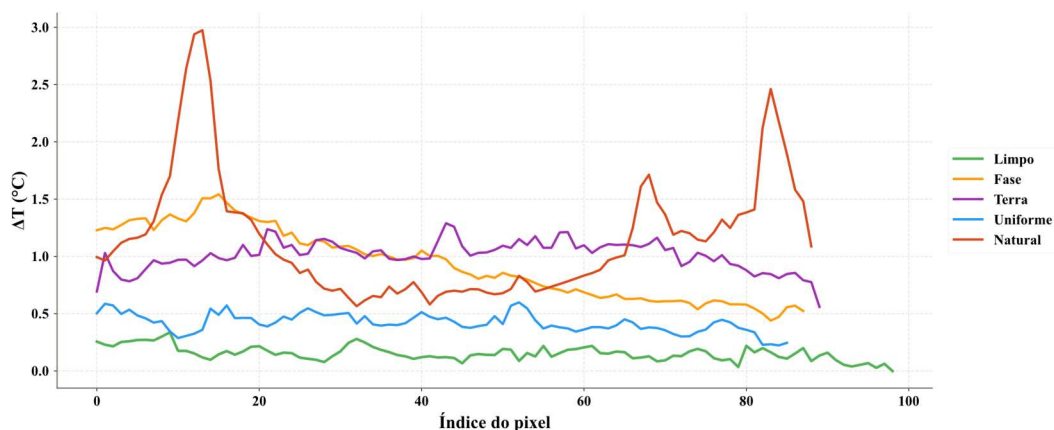


Fonte: autoria própria.

A comparação visual entre os sinais permite observar que a poluição superficial pode alterar tanto o nível térmico quanto a distribuição espacial da temperatura ao longo do isolador. Em determinadas condições, a diferença entre as classes pode manifestar-se por elevações térmicas mais distribuídas, enquanto em outras pode ocorrer de forma localizada, dependendo da região em que a poluição se concentra e da severidade da poluição. Essa característica justifica a adoção de diferentes representações do perfil térmico, uma vez que padrões globais e padrões espacialmente localizados podem exigir estratégias distintas de modelagem.

Outro fator relevante observado na base de dados é a variação das condições ambientais entre os ensaios. A temperatura ambiente registrada nos experimentos variou entre 24,1 °C e 29,3 °C. Essa variação pode introduzir deslocamentos globais nos valores absolutos dos perfis térmicos, dificultando a comparação direta entre amostras obtidas em condições distintas. Dessa forma, a correção pela temperatura ambiente, adotada no pré-processamento, tem como finalidade reduzir a influência desses deslocamentos e representar o perfil em termos de aquecimento relativo. Na Figura 26 são ilustrados os mesmos perfis representados na Figura 25 com a correção a partir da temperatura ambiente.

Figura 26. Comparação dos perfis térmicos médios, em termos de  $\Delta T$ , ao longo da direção longitudinal do isolador para diferentes condições de superfície.

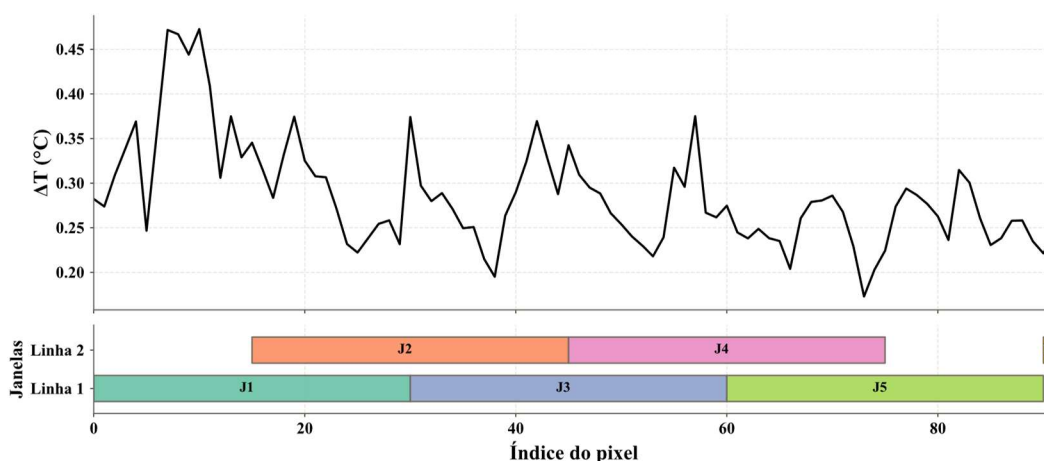


Fonte: autoria própria.

Conforme ilustrado na Figura 26, observa-se que a correção não altera a forma geral do sinal, mas desloca seus valores para uma escala relativa ao ambiente do ensaio. Com isso, os pontos do perfil passam a representar a diferença entre a temperatura medida no isolador e a temperatura ambiente correspondente.

Após a correção, os perfis foram segmentados em janelas ao longo do comprimento do isolador. Na Figura 27 ilustra-se a segmentação do perfil térmico em janelas.

Figura 27. Segmentação do perfil térmico em janelas.



Fonte: autoria própria.

Conforme apresentado na Figura 27, a divisão em janelas parcialmente sobrepostas permite caracterizar regiões sucessivas do isolador e, ao mesmo tempo, preservar a continuidade espacial entre segmentos adjacentes. Essa organização fornece

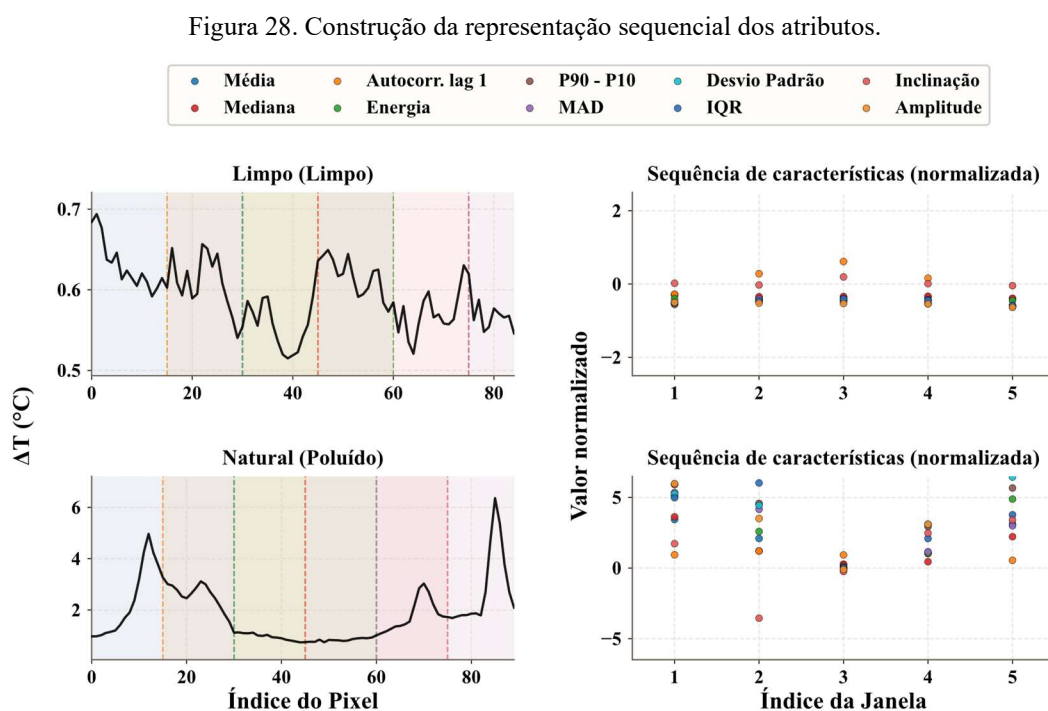
a base para a extração dos atributos térmicos locais e para a construção das diferentes representações avaliadas.

Portanto, a caracterização inicial da base evidencia três aspectos fundamentais para a análise dos resultados. O primeiro é a presença de desbalanceamento entre as classes, o que justifica a adoção de métricas balanceadas. O segundo é a existência de grupos de captura, que exige uma estratégia de validação capaz de evitar vazamento de dados e fornecer uma estimativa mais realista de generalização. O terceiro é a natureza espacial dos perfis térmicos, cujas alterações podem ocorrer de forma global ou localizada ao longo do isolador.

## 5.2 ATRIBUTOS TÉRMICOS E CONSTRUÇÃO DAS REPRESENTAÇÕES

Após a segmentação dos perfis térmicos em janelas sobrepostas, cada segmento passa a ser descrito por um conjunto de atributos térmicos locais. Essa etapa transforma o perfil unidimensional, originalmente representado por uma sequência de valores de temperatura, em uma sequência de vetores de atributos.

Na Figura 28 ilustra-se a construção da representação sequencial dos atributos.

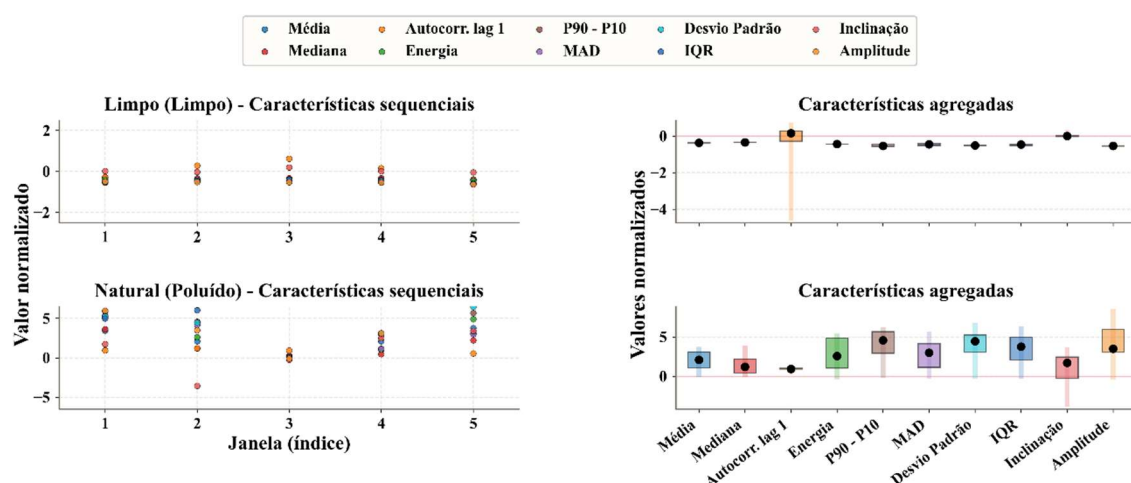


Fonte: autoria própria.

Conforme apresentado na Figura 28, a representação sequencial preserva a relação entre os atributos extraídos e suas respectivas posições ao longo do isolador. Essa estrutura permite que os modelos sequenciais avaliem não apenas os valores dos atributos térmicos, mas também sua evolução espacial entre janelas consecutivas. Desse modo, alterações regionais, transições térmicas e padrões de variação ao longo do perfil podem ser incorporados ao processo de classificação.

A partir da representação sequencial, também se constrói a representação agregada. Nessa abordagem, os atributos extraídos nas diferentes janelas são resumidos ao longo do perfil por meio de estatísticas descritivas, produzindo um único vetor de características para cada amostra. Assim, a sequência de vetores locais é convertida em uma representação tabular de dimensão fixa, adequada para classificadores que não operam diretamente sobre sequências. Na Figura 29, ilustra-se a construção da representação agregada a partir dos atributos locais extraídos nas janelas do perfil térmico.

Figura 29. Construção da representação agregada dos atributos.

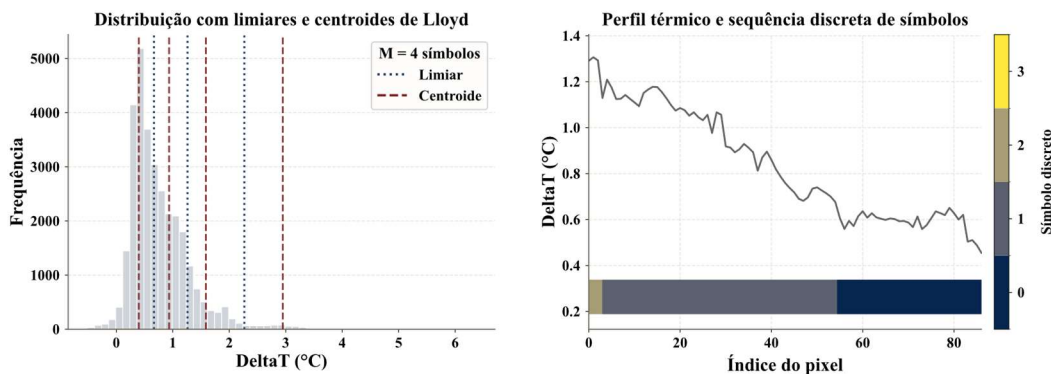


Fonte: autoria própria.

Conforme apresentado na Figura 29, a representação agregada sintetiza o comportamento dos atributos ao longo do perfil térmico. Para cada atributo calculado nas janelas, são obtidas medidas capazes de representar sua tendência central, dispersão, extremos e forma de distribuição ao longo do isolador. Com isso, a informação originalmente organizada como uma sequência passa a ser expressa em um vetor fixo de características, no qual cada coluna representa uma propriedade global derivada dos atributos locais.

Além das representações baseadas em atributos contínuos, nesta dissertação também é considerada uma representação simbólica dos perfis térmicos. Na Figura 30 apresenta-se o processo de quantização utilizado na construção da representação simbólica do perfil térmico.

Figura 30. Construção da representação simbólica do perfil térmico.



Fonte: autoria própria.

Conforme apresentado na Figura 30, a quantização utilizando o algoritmo de Lloyd divide a faixa de valores do perfil térmico em regiões associadas a símbolos discretos. Cada ponto do perfil contínuo é substituído pelo símbolo correspondente ao intervalo ao qual pertence. Dessa forma, o sinal térmico deixa de ser representado por valores reais de temperatura e passa a ser descrito por uma sequência de símbolos, preservando a ordem espacial dos pontos ao longo do perfil.

Dessa forma, a construção das representações estabelece a base para a comparação metodológica realizada neste trabalho. A representação sequencial preserva a organização espacial dos atributos térmicos extraídos nas janelas; a representação agregada resume essa sequência em um vetor fixo de características; e a representação simbólica converte o perfil térmico contínuo em uma sequência discreta de símbolos. A análise conjunta dessas abordagens permite investigar se a informação mais relevante para a classificação está associada ao comportamento global do perfil, à evolução espacial dos atributos ou aos padrões discretos de transição térmica.

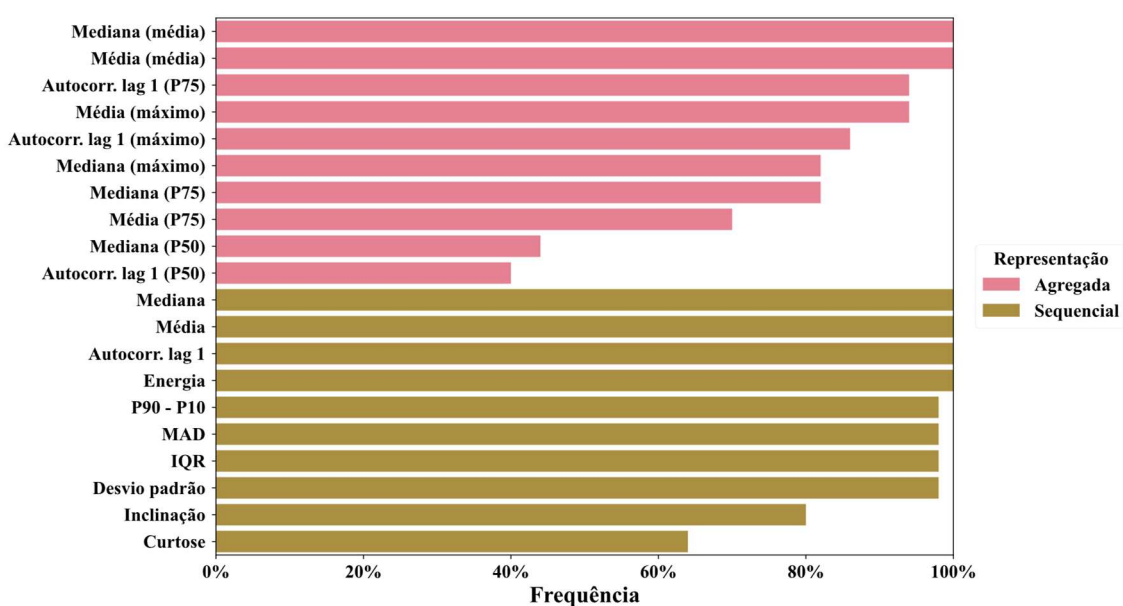
### 5.3 ATRIBUTOS SELECIONADOS NA REPRESENTAÇÃO CONTÍNUA

Após a construção das representações contínuas dos perfis térmicos, realiza-se a seleção de atributos com o objetivo de identificar os descritores mais relevantes para a

classificação binária entre as classes Limpo e Poluído. Essa etapa é aplicada às duas formas de representação contínua adotadas neste trabalho: a representação sequencial e a representação agregada. Embora ambas sejam derivadas dos atributos extraídos dos perfis segmentados em janelas, cada uma organiza a informação térmica de maneira distinta e, por isso, demanda uma interpretação específica dos resultados de seleção.

Na Figura 31, apresenta-se a frequência de seleção dos atributos nas representações sequencial e agregada, considerando os *rankings* produzidos pelo procedimento de seleção ao longo das divisões experimentais.

Figura 31. Frequência de seleção de atributos nas representações sequencial e agregada.



Fonte: autoria própria.

Conforme apresentado na Figura 31, na representação sequencial destacam-se os atributos Mediana, Média, Autocorr\_Lag1 e Energia, que aparecem com frequência máxima nos *rankings* analisados. Esse resultado indica que a distinção entre as classes está fortemente associada ao nível térmico local do perfil, representado pela média e pela mediana calculadas nas janelas. A recorrência desses atributos sugere que a poluição superficial produz alterações no patamar térmico do sinal, tornando as medidas de tendência central relevantes para a caracterização dos perfis.

Além dos atributos de nível térmico, a frequência elevada de Autocorr\_Lag1 indica que a dependência espacial do perfil também contribui para a classificação. Esse atributo representa a relação entre valores sucessivos do sinal, permitindo capturar o grau de persistência térmica ao longo do perfil. Sua recorrência sugere que os perfis limpos e

poluídos não diferem apenas pela magnitude dos valores de temperatura, mas também pela forma como esses valores se organizam espacialmente. Esse resultado é coerente com a hipótese de que a poluição superficial pode alterar a continuidade, a regularidade ou a persistência do padrão térmico ao longo do isolador.

O atributo Energia também aparece entre os mais frequentes na representação sequencial. Esse descritor está associado à intensidade global dos valores térmicos dentro das janelas e, portanto, complementa as medidas de nível. Sua presença indica que a magnitude acumulada do sinal térmico local contém informação relevante para a separação entre as classes. Em conjunto, Média, Mediana, Autocorr\_Lag1 e Energia mostram que a representação sequencial concentra informação discriminativa em atributos que descrevem tanto a intensidade térmica quanto a organização espacial do perfil.

Ainda na representação sequencial, observa-se elevada recorrência de atributos associados à dispersão e ao contraste térmico, como P90-P10, MAD, IQR e Desvio padrão. Esses atributos descrevem a variabilidade dos valores dentro das janelas e indicam o grau de heterogeneidade térmica local. Sua presença entre os atributos mais frequentes sugere que a poluição superficial também pode se manifestar por meio de alterações na variabilidade do perfil, especialmente quando a contaminação apresenta distribuição não uniforme. Nesses casos, o aquecimento pode ocorrer de maneira localizada, produzindo maior contraste entre diferentes regiões do isolador.

Na representação agregada, destacam-se principalmente os atributos Mediana (média) e Média (média), ambos associados ao resumo global do nível térmico ao longo das janelas. Esse resultado indica que, mesmo após a agregação dos atributos locais em um vetor fixo, as medidas de tendência central permanecem altamente relevantes para a classificação. A recorrência de Mediana (média) sugere que a média das medianas locais fornece uma descrição robusta do comportamento térmico do perfil, reduzindo a influência de variações pontuais e preservando a informação relacionada ao nível térmico geral do isolador.

Além das medidas de tendência central, a representação agregada também apresenta alta frequência de atributos como Autocorr\_Lag1 (p75), Média (máximo) e Autocorr\_Lag1 (máximo). Esses descritores indicam que a agregação preserva informações associadas às regiões de maior intensidade térmica e à dependência espacial do perfil. A presença de Média (máximo) sugere que valores elevados em determinadas janelas contribuem para a distinção entre as classes, o que pode estar relacionado a

aquecimentos localizados produzidos pela poluição. Já os atributos derivados da autocorrelação indicam que a regularidade espacial do sinal continua relevante mesmo quando a sequência original é resumida por estatísticas globais.

A comparação entre as representações mostra que os atributos de nível térmico aparecem como os mais recorrentes em ambas as abordagens. Esse comportamento indica que a diferença entre isoladores limpos e poluídos está fortemente relacionada ao patamar térmico dos perfis. Entretanto, a presença de atributos de autocorrelação, energia, dispersão e contraste demonstra que a classificação não depende exclusivamente da temperatura média ou mediana. A organização espacial do sinal e a variabilidade local também contribuem para a caracterização das classes, sobretudo em cenários nos quais a poluição não se distribui de forma homogênea ao longo do isolador.

Os resultados também indicam que a seleção favorece subconjuntos potencialmente parcimoniosos. A recorrência de poucos atributos em posições de destaque sugere que parte significativa da informação discriminativa pode estar concentrada em um conjunto reduzido de descritores. Esse aspecto é relevante para a interpretação dos modelos, pois reduz a complexidade da representação e facilita a análise física dos resultados. Em um problema experimental com número limitado de grupos de captura, a redução da dimensionalidade também contribui para diminuir o risco de ajuste excessivo aos dados de treinamento.

De modo geral, a seleção de atributos nas representações contínuas mostra que os descritores mais relevantes estão associados principalmente ao nível térmico, à energia, à dependência espacial e à variabilidade local do perfil. Essa combinação de propriedades é coerente com o fenômeno físico investigado, uma vez que a poluição superficial pode modificar tanto a intensidade do aquecimento quanto à forma como a temperatura se distribui ao longo do isolador. Dessa maneira, os resultados da seleção fornecem uma base interpretativa para a avaliação dos classificadores, pois indicam quais informações térmicas são priorizadas antes da etapa de treinamento e comparação dos modelos.

## 5.4 CONFIGURAÇÕES SIMBÓLICAS SELECIONADAS

Na Tabela 8 apresentam-se as configurações simbólicas selecionadas em cada repetição externa da validação aninhada. São indicados o número de símbolos adotado na quantização Lloyd-Max e o número de estados ocultos definidos para os HMMs das classes Limpo e Poluído a partir do critério de informação bayesiana.

Tabela 8. Configurações simbólicas selecionadas em cada repetição externa da validação aninhada.

| Repetição<br>externa | Número de símbolos | Estados<br>(Limpo) | Estados<br>(Poluído) |
|----------------------|--------------------|--------------------|----------------------|
| 1                    | 4                  | 2                  | 5                    |
| 2                    | 12                 | 5                  | 4                    |
| 3                    | 4                  | 4                  | 5                    |
| 4                    | 6                  | 5                  | 5                    |
| 5                    | 10                 | 2                  | 2                    |

Fonte: autoria própria.

Conforme apresentado na Tabela 8, não há uma configuração simbólica única selecionada em todas as repetições externas. O número de símbolos varia entre 4, 6, 10 e 12, indicando que a discretização mais adequada depende da composição dos grupos disponíveis em cada divisão da validação aninhada. Esse comportamento sugere que a representação simbólica apresenta sensibilidade à distribuição dos perfis entre treino e teste, especialmente em uma base com número limitado de grupos de captura.

A seleção de configurações com 4 símbolos em duas repetições externas indica que, em parte das divisões, uma representação simbólica mais compacta é suficiente para descrever padrões relevantes dos perfis térmicos. Uma discretização com menor número de símbolos reduz a complexidade da sequência observada e pode tornar a estimação das probabilidades de emissão mais estável. Essa simplificação também pode reduzir a influência de pequenas flutuações numéricas do perfil, fazendo com que o HMM discreto modele padrões mais gerais do comportamento térmico.

Por outro lado, a seleção de configurações com 10 e 12 símbolos em outras repetições indica que, em determinadas divisões, uma discretização mais detalhada é favorecida pelo processo de seleção. O aumento do número de símbolos permite preservar variações mais finas do perfil térmico, mantendo maior resolução na transformação do sinal contínuo em sequência discreta. Entretanto, essa maior granularidade também aumenta a complexidade da representação, pois amplia o número de possíveis observações simbólicas e exige maior quantidade de dados para estimar as probabilidades do modelo de forma robusta.

Em relação ao número de estados ocultos, observa-se que as configurações selecionadas também variam entre as repetições externas. Para a classe Limpo, são selecionados modelos com 2, 4 ou 5 estados, enquanto para a classe Poluído são

selecionados modelos com 2, 4 ou 5 estados. Essa variação indica que a estrutura interna necessária para representar cada classe não é fixa em todas as divisões experimentais. Em algumas partições, modelos mais simples são selecionados; em outras, o processo de escolha favorece estruturas com maior número de estados ocultos.

A diferença entre o número de estados selecionado para as classes Limpo e Poluído também merece atenção. Nas repetições 1 e 3, por exemplo, a classe Poluído é representada por HMMs com maior número de estados do que a classe Limpo. Esse comportamento é coerente com a composição da base, uma vez que a classe Poluído reúne diferentes formas de contaminação superficial, incluindo poluição uniforme, não uniforme e natural. Essa maior diversidade interna pode exigir uma estrutura probabilística mais flexível para representar os padrões simbólicos associados aos perfis contaminados.

Nas repetições 2 e 4, por outro lado, a classe Limpo é representada por modelos com maior número de estados do que a classe Poluído. Esse resultado indica que a complexidade estrutural selecionada não depende apenas da classe em si, mas também dos grupos presentes em cada divisão externa. Como as partições são realizadas por grupos de captura, a variabilidade dos perfis disponíveis para treinamento pode favorecer diferentes configurações do modelo, mesmo dentro da mesma formulação metodológica.

A repetição 5 seleciona uma configuração com 10 símbolos e 2 estados para ambas as classes. Essa escolha representa uma combinação em que a discretização do perfil é relativamente detalhada, mas a estrutura oculta dos modelos permanece simples. Desprende-se, daí, que o número de símbolos e o número de estados não aumentam necessariamente de forma conjunta. Em outras palavras, a seleção pode favorecer maior resolução simbólica sem ampliar a quantidade de estados ocultos, ou o inverso, dependendo da organização dos dados em cada repetição.

De modo geral, as configurações selecionadas mostram que a representação simbólica envolve um equilíbrio entre simplificação do sinal e preservação da informação térmica. Um número reduzido de símbolos produz sequências mais compactas e modelos potencialmente mais estáveis, mas pode eliminar diferenças sutis entre regiões do perfil. Um número maior de símbolos preserva mais detalhes do sinal contínuo, porém aumenta a complexidade da estimação. Da mesma forma, HMMs com poucos estados tendem a representar padrões mais gerais, enquanto modelos com mais estados podem capturar maior diversidade estrutural nas sequências simbólicas.

Assim, a seleção das configurações simbólicas evidencia que o HMM discreto não apresenta uma configuração única dominante entre as repetições externas. A variação no número de símbolos e no número de estados ocultos indica que a representação simbólica é sensível à composição dos grupos disponíveis em cada divisão da validação aninhada. Por essa razão, a definição dessas configurações é incorporada ao laço interno de seleção, evitando escolhas fixas ou arbitrárias. A avaliação quantitativa do desempenho obtido por essa representação é apresentada na seção seguinte, em conjunto com os demais classificadores, permitindo comparar todos os modelos sob o mesmo protocolo experimental.

## 5.5 DESEMPENHO DOS MODELOS NA VALIDAÇÃO ANINHADA

Essa etapa representa a estimativa principal de generalização dos classificadores, uma vez que o conjunto de teste externo permanece isolado durante todo o processo de seleção de atributos, ajuste de hiperparâmetros e escolha de configurações. Dessa forma, os resultados apresentados nesta seção correspondem ao desempenho dos modelos em grupos de captura não utilizados nas etapas internas de decisão.

A comparação contempla classificadores associados às três formas de representação consideradas neste trabalho. Na representação agregada, são avaliados os modelos SVM e Random Forest. Na representação sequencial, são avaliados o HMM com emissões gaussianas e a LSTM. Na representação simbólica, avalia-se o HMM discreto. Essa organização permite comparar não apenas diferentes classificadores, mas também diferentes formas de estruturar a informação térmica extraída dos perfis.

Os resultados são analisados em dois níveis, o primeiro considera cada perfil térmico como uma unidade individual de predição, já o segundo considera o grupo de captura como unidade de decisão, por meio da agregação das predições individuais por voto majoritário.

Antes da análise dos resultados médios, apresenta-se a distribuição dos subconjuntos de atributos e das configurações selecionadas em cada repetição externa da validação aninhada. Essa análise é relevante porque permite verificar se os modelos convergem para escolhas semelhantes entre as partições ou se as decisões do laço interno variam em função da composição dos grupos de treino e validação.

Na Tabela 9 apresentam-se os subconjuntos de atributos ou configurações selecionadas para cada modelo em cada repetição externa. Para os modelos SVM e

Random Forest, são apresentados os atributos da representação agregada. Para os modelos HMM-GMM e LSTM, são apresentados os atributos da representação sequencial. Para o HMM discreto, são apresentados o número de símbolos da quantização e o número de estados ocultos selecionado para cada classe.

Tabela 9. Subconjuntos e configurações selecionados por repetição externa.

| Repetição externa | SVM                                                                                                                                       | Random Forest                                                                                                     | HMM-GMM        | LSTM                                                                                         | HMM discreto                  |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|----------------|----------------------------------------------------------------------------------------------|-------------------------------|
| 1                 | Mediana (média), Autocorr_Lag1 (p75), Média (máximo), Média (média), Autocorr_Lag1 (máximo)                                               | Mediana (média), Autocorr_Lag1 (p75), Média (máximo), Média (máximo)                                              | Mediana        | Mediana                                                                                      | M = 4; Limpo = 2; Poluído = 5 |
| 2                 | Autocorr_Lag1 (p75), Mediana (máximo), Média (p75), Média (máximo), Mediana (média), Mediana (p75), Média (média), Autocorr_Lag1 (máximo) | Autocorr_Lag1 (p75), Mediana (máximo), Média (p75), Média (máximo), Mediana (média), Mediana (p75), Média (média) | Mediana, média | Mediana, média                                                                               | M=12; Limpo=5; Poluído=4      |
| 3                 | Mediana (média)                                                                                                                           | Mediana (média)                                                                                                   | Mediana, média | Mediana, Média, Autocorr_Lag1, Energia, Curtose, MAD, IQR, P90_P10, Desvio_Padiao, Amplitude | M=4; Limpo=4; Poluído=5       |

|   |                 |                                                                                                                                   |                                                                  |                                           |                             |
|---|-----------------|-----------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------|-------------------------------------------|-----------------------------|
| 4 | Mediana (média) | Mediana (média),<br>Média (média),<br>Autocorr_Lag1<br>(p75) Média<br>(máximo),<br>Autocorr_Lag1<br>(máximo),<br>Mediana (máximo) | Média, Mediana,<br>Autocorr_Lag1, Energia, P90-<br>P10, IQR, MAD | Média                                     | M=6; Limpo=5;<br>Poluído=4  |
| 5 | Mediana (média) | Mediana (média),<br>Média (média),<br>Autocorr_Lag1<br>(p75) Média<br>(máximo),<br>Autocorr_Lag1<br>(máximo),<br>Mediana (máximo) | Média                                                            | Média, Mediana,<br>Autocorr_Lag1, Energia | M=10; Limpo=2;<br>Poluído=2 |

---

Fonte: autoria própria.

Conforme apresentado na Tabela 9, o SVM tende a selecionar subconjuntos mais compactos na representação agregada. Em três das cinco repetições externas, o modelo utiliza apenas o atributo Mediana (média). Esse comportamento indica que, para parte das divisões, uma medida robusta do nível térmico global é suficiente para produzir separação consistente entre as classes. Nas repetições em que o SVM seleciona subconjuntos maiores, os atributos adicionais também se relacionam principalmente ao nível térmico, aos valores máximos e à autocorrelação do perfil, sugerindo contribuição complementar da intensidade térmica e da organização espacial agregada.

A Random Forest apresenta maior variação no tamanho dos subconjuntos selecionados. Enquanto uma das repetições utiliza apenas Mediana (média), outras selecionam combinações com quatro, seis ou sete atributos. Nos subconjuntos da Random Forest, repetem-se atributos associados ao nível térmico, como Mediana (média), Média (média) e Média (máximo), e atributos relacionados à dependência espacial, como Autocorr\_Lag1 (p75) e Autocorr\_Lag1 (máximo). Isso indica que o modelo se beneficia da combinação entre intensidade térmica global e organização espacial resumida.

Nos modelos sequenciais, observa-se predominância dos atributos Mediana e Média. No HMM-GMM, esses atributos aparecem isoladamente ou combinados em diferentes repetições, reforçando a importância do nível térmico local preservado ao longo das janelas. Em uma das repetições, o HMM-GMM seleciona um subconjunto mais amplo, incluindo Autocorr\_Lag1, Energia, P90\_P10, IQR e MAD, o que indica que, em determinadas partições, atributos de dependência espacial, energia e dispersão também contribuem para a modelagem probabilística das sequências.

A LSTM apresenta comportamento semelhante quanto à recorrência de Média e Mediana, mas também seleciona subconjuntos mais amplos em algumas repetições. Na terceira repetição externa, por exemplo, são selecionados dez atributos, incluindo descritores de nível, energia, forma da distribuição, dispersão e contraste. Esse resultado sugere que, em determinadas divisões, a LSTM pode explorar uma descrição mais rica da sequência de janelas. Entretanto, em outras repetições, o modelo seleciona apenas um ou dois atributos, indicando que a contribuição de descritores adicionais depende da composição dos grupos disponíveis em cada laço externo.

Conforme discutido anteriormente a partir da análise da Tabela 8, para o HMM discreto, observa-se uma variação tanto no número de símbolos quanto no número de estados ocultos das classes Limpo e Poluído. Essa variação reforça que a representação simbólica não apresenta uma configuração única dominante entre as repetições externas,

mostrando que a melhor estrutura simbólica depende da distribuição dos grupos de captura em cada divisão da validação aninhada.

Ademais, na Tabela 10 apresenta-se o desempenho individual dos modelos em cada repetição externa, considerando uma métrica multicritério calculada no nível dos grupos de captura. Esse critério corresponde à média entre acurácia balanceada, F1 macro e coeficiente de correlação de Matthews, permitindo avaliar simultaneamente o equilíbrio entre classes, a qualidade média das predições e a correlação global entre rótulos verdadeiros e preditos.

Tabela 10. Métrica multicritério por grupo de captura em cada repetição externa.

| Repetição externa | SVM   | Random Forest | HMM-GMM | LSTM  | HMM discreto |
|-------------------|-------|---------------|---------|-------|--------------|
| 1                 | 0,865 | 0,865         | 0,865   | 0,865 | 0,865        |
| 2                 | 1,000 | 0,939         | 0,939   | 0,939 | 0,939        |
| 3                 | 0,929 | 0,858         | 0,858   | 0,829 | 0,829        |
| 4                 | 0,865 | 0,865         | 0,865   | 0,865 | 0,865        |
| 5                 | 0,865 | 0,807         | 0,807   | 0,807 | 0,791        |

Fonte: autoria própria.

Conforme apresentado na Tabela 10, o SVM apresenta o maior valor da métrica multicritério em todas as repetições externas. Esse resultado indica que a representação agregada associada ao classificador de margem mantém desempenho consistente quando a decisão é analisada no nível dos grupos de captura. A segunda repetição externa apresenta os maiores valores para todos os classificadores, enquanto a quinta repetição externa apresenta os menores valores para Random Forest, HMM-GMM, LSTM e HMM discreto. Esse comportamento mostra que a dificuldade do problema varia entre as partições externas, possivelmente em função da composição dos grupos de captura presentes no teste.

Observa-se também que Random Forest e HMM-GMM apresentam valores idênticos do multicritério em todas as repetições externas. Embora esses modelos utilizem representações distintas, agregada e sequencial, respectivamente, esse resultado sugere que ambas as formas de representação capturam informações discriminativas semelhantes no nível dos grupos de captura. A LSTM e o HMM discreto apresentam desempenho próximo, mas ligeiramente inferior em algumas repetições, especialmente na terceira e na quinta divisões externas.

A análise por repetição externa mostra que a dificuldade do problema não é uniforme entre as partições. Como a separação é realizada por grupos de captura, cada teste externo reúne condições experimentais específicas, podendo conter grupos mais ou menos difíceis de classificar. Dessa forma, diferenças entre as repetições podem refletir variações na composição dos grupos, nos subtipos de poluição presentes, nas condições ambientais ou na similaridade térmica entre amostras limpas e poluídas. Essa observação reforça a importância da validação aninhada com múltiplas repetições externas, pois uma única divisão poderia fornecer uma estimativa pouco representativa do desempenho real dos modelos.

Além disso, a comparação entre as Tabela 9 e Tabela 10 evidencia que a repetição de um mesmo subconjunto completo não ocorre de maneira uniforme entre os modelos. No caso do SVM, observa-se a maior estabilidade entre as configurações selecionadas, pois o subconjunto unitário formado apenas por Mediana (média) aparece em três das cinco repetições externas, correspondendo a uma frequência de 60%. Esse resultado indica que, para esse classificador, uma representação agregada bastante compacta é recorrente e suficiente para produzir desempenho elevado no nível dos grupos de captura.

Na Random Forest, por outro lado, não se observa repetição de um mesmo subconjunto completo entre as cinco repetições externas. O modelo seleciona combinações distintas de atributos em cada repetição externa, com tamanhos variando entre um e sete atributos. Esse comportamento indica maior variabilidade na composição dos subconjuntos escolhidos, ainda que o desempenho permaneça competitivo. Assim, a Random Forest parece depender menos de uma configuração fixa e mais da combinação de diferentes descritores selecionados em cada partição.

Nos modelos sequenciais, o HMM-GMM apresenta repetição parcial dos subconjuntos selecionados. O subconjunto formado por Mediana e Média aparece em duas das cinco repetições externas, correspondendo a uma frequência de 40%. Os demais subconjuntos selecionados aparecem apenas uma vez cada, incluindo os subconjuntos unitários formados por Mediana ou Média e o subconjunto mais amplo com sete atributos. Esse resultado mostra que o HMM-GMM apresenta certa recorrência em torno de uma representação sequencial compacta, embora também varie conforme a composição dos grupos de captura.

Na LSTM, não há repetição de um mesmo subconjunto completo entre as cinco repetições externas. O modelo seleciona subconjuntos de tamanhos distintos, variando de um a dez atributos. Essa variação sugere maior sensibilidade da LSTM à composição das

divisões externas, uma vez que, em algumas partições, o modelo opera com apenas um atributo sequencial, enquanto em outras utiliza uma representação mais ampla do perfil térmico.

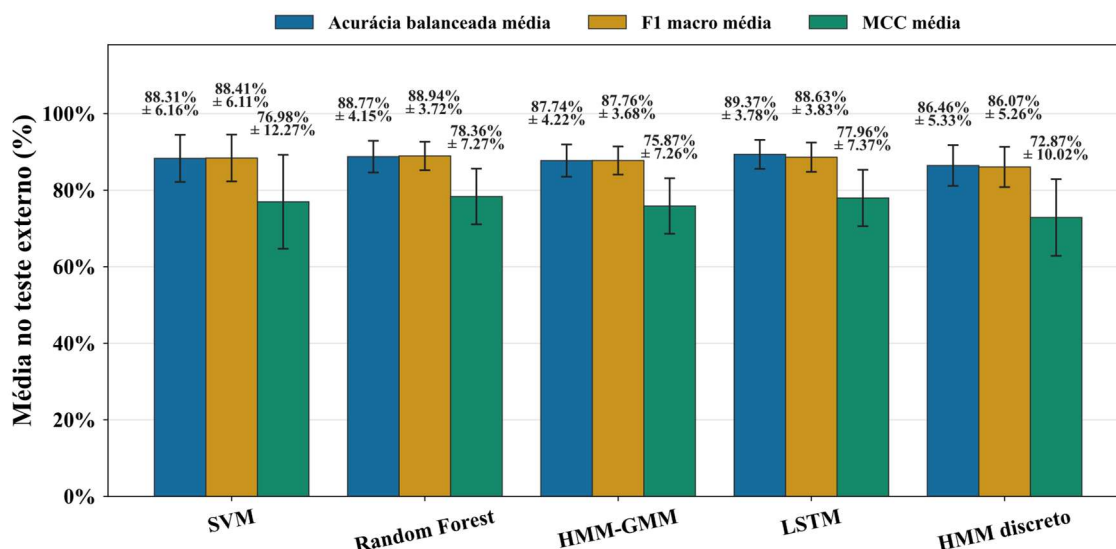
Para o HMM discreto, a análise de frequência assume natureza distinta, pois não há subconjuntos de atributos, mas configurações simbólicas. Nesse caso, nenhuma configuração completa se repete entre as cinco repetições externas. Embora o número de símbolos igual a 4 apareça em duas repetições, os números de estados ocultos associados às classes Limpo e Poluído são diferentes. Assim, considera-se que as configurações completas do HMM discreto apresentam frequência unitária, reforçando a sensibilidade da representação simbólica à composição dos grupos em cada divisão externa.

Dessa forma, a análise da frequência dos subconjuntos completos indica que o SVM é o modelo com maior estabilidade na seleção, pois repete o mesmo subconjunto em 60% das repetições externas. O HMM-GMM apresenta estabilidade intermediária, com repetição de um subconjunto em 40% das repetições. Random Forest, LSTM e HMM discreto não apresentam repetição de configurações completas, o que sugere maior dependência das escolhas realizadas no laço interno em cada partição da validação aninhada.

#### 5.5.1 AVALIAÇÃO POR PERFIL TÉRMICO

Na Figura 32 apresenta-se o desempenho médio dos modelos considerando a avaliação por perfil térmico individual. Nessa análise, cada perfil extraído da termografia recebe uma predição própria, e as métricas são calculadas a partir da comparação entre os rótulos verdadeiros e os rótulos estimados para todos os perfis pertencentes aos conjuntos de teste externos.

Figura 32. Desempenho dos modelos na avaliação por perfil térmico.



Fonte: autoria própria.

Conforme apresentado na Figura 32, os modelos apresentam desempenhos relativamente próximos no nível dos perfis individuais, com acurácia balanceada média variando aproximadamente entre 86,46% e 89,37%. Esse resultado indica que todas as representações avaliadas contêm informação discriminativa relevante para a distinção entre isoladores limpos e poluídos. Ainda assim, observam-se diferenças entre os modelos, tanto em termos de desempenho médio quanto de estabilidade entre as repetições externas.

A LSTM apresenta a maior acurácia balanceada média na avaliação por perfil, com valor de aproximadamente 89,37% ± 3,78%. Além disso, a LSTM apresenta F1 macro médio de 88,63% ± 3,83% e MCC médio de 77,96% ± 7,37%, valores próximos aos melhores resultados observados entre os modelos. A menor dispersão na acurácia balanceada, em comparação com a maioria dos demais métodos, também indica maior estabilidade da LSTM entre as repetições externas, sugerindo menor sensibilidade à composição dos grupos de treinamento e teste. Esse resultado sugere que a preservação da organização espacial dos atributos ao longo das janelas contribui para a classificação individual dos perfis térmicos. Como a LSTM opera sobre sequências, o modelo é capaz de explorar a evolução dos atributos ao longo do isolador, o que pode ser vantajoso quando as alterações térmicas associadas à poluição aparecem de forma localizada ou gradual.

A Random Forest também apresenta desempenho elevado na avaliação por perfil, com acurácia balanceada média próxima de 88,77% ± 4,15%, F1 macro médio de 88,94%

$\pm 3,72\%$  e MCC médio de  $78,36\% \pm 7,27\%$ . Esses valores indicam que a representação agregada preserva informação relevante para a classificação, mesmo sem manter explicitamente a ordem espacial completa das janelas. Nesse caso, as estatísticas globais dos atributos térmicos parecem ser suficientes para capturar parte importante das diferenças entre as classes. Além disso, os desvios-padrão relativamente baixos indicam boa estabilidade, especialmente no F1 macro e no MCC, reforçando a robustez do modelo agregado.

O SVM apresenta acurácia balanceada média de aproximadamente  $88,31\% \pm 6,16\%$ , F1 macro médio de  $88,41\% \pm 6,11\%$  e MCC médio de  $76,98\% \pm 12,70\%$ . Esse desempenho é próximo ao da Random Forest e mostra que uma representação agregada compacta, especialmente baseada em atributos de nível térmico e dependência espacial, produz resultados competitivos. No entanto, o SVM apresenta maior variabilidade entre as repetições externas, principalmente no MCC. Essa maior dispersão sugere que o modelo é mais sensível à composição dos grupos utilizados em cada divisão externa, podendo apresentar oscilações mais acentuadas quando determinados padrões térmicos aparecem com menor representatividade no conjunto de treinamento.

O HMM com emissões gaussianas apresenta acurácia balanceada média de aproximadamente  $87,74\% \pm 4,22\%$ , F1 macro médio de  $87,76\% \pm 3,68\%$  e MCC médio de  $75,87\% \pm 7,26\%$  na avaliação por perfil. Embora esse desempenho seja ligeiramente inferior ao da LSTM e dos modelos agregados, o resultado mostra que a modelagem probabilística sequencial também é capaz de explorar padrões relevantes nos atributos extraídos das janelas. A diferença em relação à LSTM pode estar associada à maior flexibilidade do modelo recorrente para capturar relações não lineares entre as observações sequenciais. Além disso, os desvios-padrão são moderados e próximos aos observados para a Random Forest e para a LSTM, indicando estabilidade razoável entre as repetições externas.

O HMM discreto apresenta acurácia balanceada média de aproximadamente  $86,46\% \pm 5,33\%$ , F1 macro médio de  $86,07\% \pm 5,26\%$  e MCC médio de  $72,87\% \pm 10,02\%$  no nível dos perfis. Embora seja o menor desempenho médio entre os modelos avaliados nessa unidade de análise, o resultado ainda indica capacidade discriminativa da representação simbólica. Isso sugere que a discretização do perfil térmico preserva parte dos padrões relevantes para a classificação, ainda que a transformação de valores contínuos em símbolos implique perda de resolução térmica. A maior variabilidade do MCC, quando comparada à Random Forest, à LSTM e ao HMM-GMM, indica que o

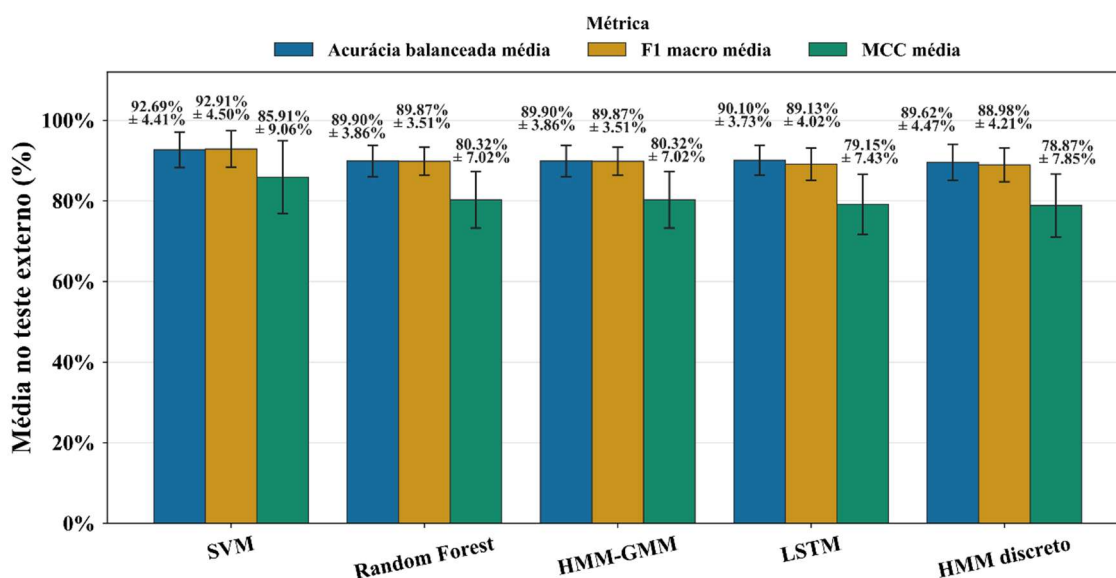
desempenho do HMM discreto é menos estável entre as repetições externas. Esse comportamento pode estar relacionado à sensibilidade do processo de quantização, uma vez que pequenas diferenças nos atributos contínuos podem ser agrupadas em símbolos semelhantes, reduzindo a capacidade do modelo de distinguir padrões térmicos mais sutis.

De modo geral, a avaliação por perfil mostra que os modelos baseados em representações contínuas apresentam desempenho ligeiramente superior ao HMM discreto. Esse comportamento é esperado, pois as representações contínuas preservam maior quantidade de informação numérica do perfil térmico. Ainda assim, a diferença entre os modelos não é extrema, indicando que a informação térmica discriminativa está presente tanto nos atributos contínuos quanto na sequência simbólica obtida por quantização.

### 5.5.2 AVALIAÇÃO POR GRUPO DE CAPTURA

Na Figura 33 ilustra-se o desempenho médio dos modelos considerando a avaliação por grupo de captura. Nessa análise, as predições dos perfis pertencentes a um mesmo grupo são agregadas por voto majoritário, produzindo uma única decisão para cada grupo experimental.

Figura 33. Desempenho dos modelos na avaliação por grupo de captura.



Fonte: autoria própria.

Conforme apresentado na Figura 33, a avaliação por grupo de captura produz desempenho médio superior ao observado na avaliação por perfil para a maioria dos

modelos. Esse comportamento indica que erros pontuais em perfis individuais tendem a ser compensados quando as decisões são agregadas no nível do grupo. Assim, o voto majoritário contribui para uma classificação mais estável da condição experimental, reduzindo a influência de variações locais ou ruídos presentes em perfis isolados.

Na avaliação por grupo, o SVM apresenta o melhor desempenho médio entre os modelos avaliados, com acurácia balanceada de aproximadamente  $92,69\% \pm 4,41\%$ , F1 macro de  $92,91\% \pm 4,40\%$  e MCC de  $85,91\% \pm 9,06\%$ . Esse resultado indica que a representação agregada, quando associada ao SVM, é altamente competitiva para a classificação dos grupos de captura. A superioridade do SVM nesse nível de avaliação sugere que a síntese global dos atributos térmicos fornece uma descrição suficientemente discriminativa para separar grupos limpos e poluídos. Entretanto, observa-se que o MCC apresenta a maior variabilidade entre os modelos, o que indica que, embora o SVM alcance o melhor desempenho médio, sua estabilidade ainda depende da composição dos grupos presentes em cada repetição externa.

A LSTM apresenta o segundo maior valor de acurácia balanceada por grupo, aproximadamente  $90,10\% \pm 3,73\%$ , F1 macro médio de  $89,13\% \pm 4,02\%$  e MCC médio de  $79,15\% \pm 7,43\%$ . Esse resultado confirma que a representação sequencial também é relevante para a classificação, especialmente por preservar a organização espacial dos atributos ao longo do perfil. A dispersão da acurácia balanceada é uma das menores entre os modelos avaliados, indicando boa estabilidade nessa métrica. No entanto, o MCC médio fica abaixo dos valores obtidos pelo SVM, pela Random Forest e pelo HMM-GMM, o que sugere que, embora a LSTM apresente boa capacidade de generalização por grupo, seu equilíbrio global entre acertos e erros das classes não supera o das melhores abordagens agregadas e probabilísticas nesse cenário.

O HMM com emissões gaussianas apresenta acurácia balanceada por grupo de aproximadamente  $89,90\% \pm 3,86\%$ , F1 macro de  $89,87\% \pm 3,51\%$  e MCC de  $80,32\% \pm 7,02\%$ . Esse desempenho é muito próximo ao da LSTM em termos de acurácia balanceada e superior em MCC. Além disso, o modelo apresenta uma das menores variabilidades entre as repetições externas, principalmente no F1 macro e no MCC. Isso indica que o HMM-GMM, apesar de ser um modelo probabilístico mais estruturado e menos flexível que a LSTM, apresenta boa capacidade de representar os padrões sequenciais dos perfis térmicos no nível dos grupos de captura.

A Random Forest também apresenta acurácia balanceada por grupo de aproximadamente  $89,90\% \pm 3,86\%$ , F1 macro de  $89,87\% \pm 3,51\%$  e MCC de  $80,32\% \pm$

7,02%. Esse resultado é semelhante ao obtido pelo HMM-GMM, embora os dois modelos utilizem representações diferentes. Enquanto a Random Forest explora atributos agregados em forma tabular, o HMM-GMM utiliza sequências de atributos contínuos. A proximidade entre os desempenhos sugere que tanto a síntese global dos atributos quanto a modelagem sequencial probabilística capturam informações relevantes para a classificação. Além disso, a baixa variabilidade observada indica que a Random Forest apresenta comportamento estável entre as repetições externas, reforçando a robustez da representação agregada no nível de grupo.

O HMM discreto apresenta acurácia balanceada por grupo de aproximadamente  $89,62\% \pm 4,47\%$ , F1 macro de  $88,80\% \pm 4,21\%$  e MCC de  $78,87\% \pm 7,85\%$ . Esses valores são superiores aos observados na avaliação por perfil, mostrando que a agregação por grupo beneficia também a representação simbólica. O desempenho do HMM discreto é próximo ao dos demais modelos, o que indica que a sequência simbólica obtida por quantização Lloyd-Max preserva informação discriminativa suficiente para produzir decisões consistentes no nível experimental. Entretanto, a variabilidade ligeiramente maior em comparação à Random Forest e ao HMM-GMM sugere que a discretização pode tornar o modelo mais sensível à divisão dos dados, especialmente quando pequenas diferenças térmicas entre perfis são convertidas em símbolos semelhantes.

A comparação entre as Figura 32 e Figura 33 evidencia que a avaliação por grupo altera a interpretação do desempenho dos modelos. No nível dos perfis, a LSTM apresenta a maior acurácia balanceada média, enquanto, no nível dos grupos, o SVM assume a melhor posição. Esse resultado evidencia que o melhor modelo pode depender da unidade de análise adotada. Como o grupo de captura representa uma unidade experimental mais conservadora e reduz a influência de perfis correlacionados, os resultados por grupo são adotados como referência principal na comparação dos classificadores.

A superioridade do SVM na avaliação por grupo sugere que a informação global sintetizada pela representação agregada é fortemente discriminativa para o problema analisado. Esse resultado não invalida a importância das representações sequenciais, mas indica que, para a base considerada, a preservação explícita da ordem espacial dos atributos não produz ganho suficiente para superar a representação agregada no critério principal de comparação. Ainda assim, os desempenhos da LSTM e do HMM-GMM mostram que a organização sequencial dos perfis contém informação relevante e pode contribuir para a classificação em determinadas partições.

Outro aspecto importante é a proximidade entre os desempenhos dos modelos na avaliação por grupo. Com exceção do SVM, que apresenta desempenho mais elevado, os demais classificadores apresentam acurácia balanceada média em torno de 0,896 a 0,901. Essa proximidade indica que diferentes representações do perfil térmico conseguem capturar aspectos relevantes do fenômeno, seja por meio de estatísticas globais, de sequências de atributos contínuos ou de padrões simbólicos discretizados.

Do ponto de vista metodológico, os resultados reforçam a importância da avaliação por grupos de captura. A melhora observada em relação à avaliação por perfil indica que a decisão por voto majoritário reduz a influência de erros individuais e fornece uma estimativa mais estável da condição do isolador. Além disso, como as divisões externas são realizadas por grupos, o desempenho apresentado reflete a capacidade dos modelos em generalizar para capturas não vistas, tornando a comparação mais rigorosa do que uma avaliação baseada apenas em perfis individuais.

De modo geral, a validação aninhada mostra que todos os modelos avaliados apresentam capacidade de distinguir perfis térmicos limpos e poluídos. O SVM com representação agregada apresenta o melhor desempenho no nível dos grupos de captura, enquanto a LSTM se destaca na avaliação por perfil. O HMM-GMM, a Random Forest e o HMM discreto apresentam resultados próximos entre si, indicando que diferentes mecanismos de modelagem são capazes de explorar a informação térmica disponível. Esses resultados fornecem a base para a comparação entre representações e para a análise posterior dos padrões de erro e da robustez dos modelos campeões.

## 5.6 ANÁLISE DE ROBUSTEZ DOS MODELOS CAMPEÕES

Após a avaliação dos modelos no protocolo de validação aninhada, realiza-se a análise de robustez das configurações campeãs. Essa etapa tem como objetivo verificar se os desempenhos observados na etapa principal se mantêm estáveis quando os modelos são reavaliados em novas divisões experimentais, mantendo congeladas as decisões previamente selecionadas. Dessa forma, a análise de robustez não realiza uma nova seleção de atributos, hiperparâmetros ou configurações simbólicas, mas avalia a consistência dos modelos já escolhidos.

A robustez é especialmente relevante neste trabalho devido ao número limitado de grupos de captura disponíveis e à variabilidade experimental associada aos diferentes tipos de poluição superficial. Como as divisões são realizadas por grupos, alterações na

composição dos conjuntos de treino e teste podem influenciar o desempenho observado. Assim, um modelo com bom desempenho médio na validação aninhada pode não ser necessariamente o mais confiável se apresentar elevada variação entre repetições adicionais.

Nesta etapa, os modelos campeões são avaliados considerando a mesma métrica multicritério adotado na seleção principal, calculado no nível dos grupos de captura. Esse critério corresponde à média entre acurácia balanceada, F1 macro e coeficiente de correlação de Matthews. Dessa forma, a análise considera simultaneamente o equilíbrio entre classes, a qualidade média das predições e a consistência global entre rótulos verdadeiros e preditos.

Na Tabela 11 apresenta-se o desempenho médio e a dispersão dos modelos campeões na análise de robustez. Os resultados são apresentados no nível dos grupos de captura, uma vez que essa unidade de análise representa a decisão experimental mais conservadora adotada neste trabalho.

Tabela 11. Desempenho dos modelos campeões na análise de robustez.

| Modelo        | Representação | Acurácia balanceada | F1-macro | MCC     | Métrica multicritério |
|---------------|---------------|---------------------|----------|---------|-----------------------|
| SVM           | Agregada      | 0,923 ±             | 0,923 ±  | 0,845 ± | 0,897 ±               |
|               |               | 0,058               | 0,062    | 0,083   | 0,067                 |
| HMM-GMM       | Sequencial    | 0,911 ±             | 0,906 ±  | 0,814 ± | 0,877 ±               |
|               |               | 0,066               | 0,074    | 0,088   | 0,076                 |
| LSTM          | Sequencial    | 0,906 ±             | 0,900 ±  | 0,809 ± | 0,872 ±               |
|               |               | 0,068               | 0,077    | 0,093   | 0,079                 |
| HMM discreto  | Sequencial    | 0,903 ±             | 0,896 ±  | 0,807 ± | 0,869 ±               |
|               | simbólica     | 0,063               | 0,070    | 0,086   | 0,073                 |
| Random Forest | Agregada      | 0,903 ±             | 0,896 ±  | 0,807 ± | 0,869 ±               |
|               |               | 0,063               | 0,070    | 0,086   | 0,072                 |

Fonte: autoria própria.

Conforme apresentado na Tabela 11, o SVM com representação agregada apresenta o maior valor médio do critério multicritério, atingindo  $0,897 \pm 0,067$ . Esse resultado confirma a tendência observada na validação aninhada, na qual o SVM também se destaca no nível dos grupos de captura. A manutenção do melhor desempenho médio na análise de robustez indica que a configuração selecionada para esse modelo não

depende apenas de uma partição específica dos dados, mas apresenta comportamento consistente quando reavaliada em novas divisões.

O desempenho do SVM também deve ser interpretado a partir da simplicidade da configuração selecionada, conforme discutido nas seções anteriores, esse modelo apresenta forte recorrência do atributo Mediana (média), frequentemente selecionado como subconjunto unitário. Assim, o resultado da robustez sugere que uma representação agregada compacta, baseada em uma medida robusta do nível térmico global, é capaz de produzir desempenho elevado e consistente na distinção entre grupos limpos e poluídos.

O HMM-GMM apresenta o segundo maior multicritério médio, com valor de  $0,877 \pm 0,076$ . Esse resultado mostra que a representação sequencial com atributos contínuos também preserva informação discriminativa relevante. Embora o desempenho médio seja inferior ao do SVM, o HMM-GMM mantém resultado competitivo, indicando que a modelagem probabilística da evolução dos atributos ao longo das janelas é capaz de representar padrões térmicos associados à condição superficial dos isoladores.

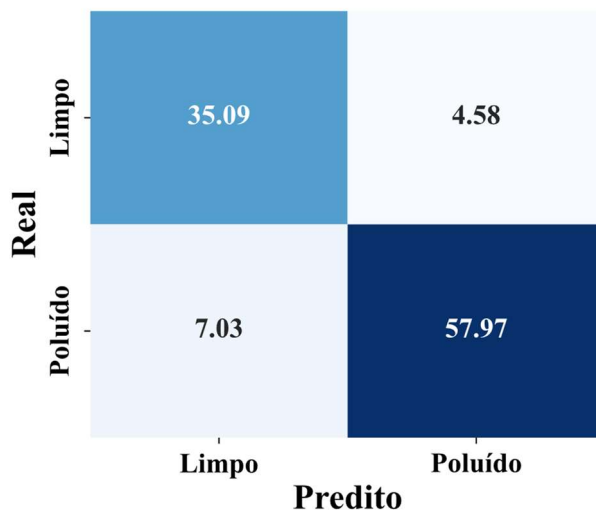
A LSTM apresenta multicritério médio de  $0,872 \pm 0,079$ , valor próximo ao obtido pelo HMM-GMM. Esse resultado confirma que a representação sequencial é relevante para o problema, mas também sugere que o aumento da flexibilidade do modelo recorrente não resulta, neste conjunto de dados, em ganho expressivo de robustez em relação ao HMM-GMM. Uma possível interpretação é que a quantidade limitada de grupos de captura restringe o benefício de modelos mais flexíveis, especialmente quando comparados a classificadores mais simples ou a modelos probabilísticos com estrutura mais controlada.

O HMM discreto e a Random Forest apresentam multicritério médios semelhantes, ambos próximos de 0,869. No caso do HMM discreto, esse resultado indica que a representação simbólica preserva parte relevante da informação térmica dos perfis, mesmo após a quantização dos valores contínuos. Entretanto, a discretização implica perda de resolução térmica, o que pode limitar o desempenho em comparação com representações contínuas. Já a Random Forest, embora utilize a mesma representação agregada do SVM, apresenta desempenho médio inferior, sugerindo que o classificador de margem explora de forma mais eficiente os atributos selecionados para este problema.

Após a análise quantitativa da robustez, avalia-se a matriz de confusão do modelo final escolhido. Considerando o maior multicritério médio e a simplicidade da configuração, o SVM com representação agregada é adotado como modelo final. Na

Figura 34 apresenta-se a matriz de confusão desse modelo, permitindo observar a distribuição dos acertos e erros entre as classes Limpo e Poluído.

Figura 34. Matriz de confusão normalizada do modelo SVM na análise de robustez.

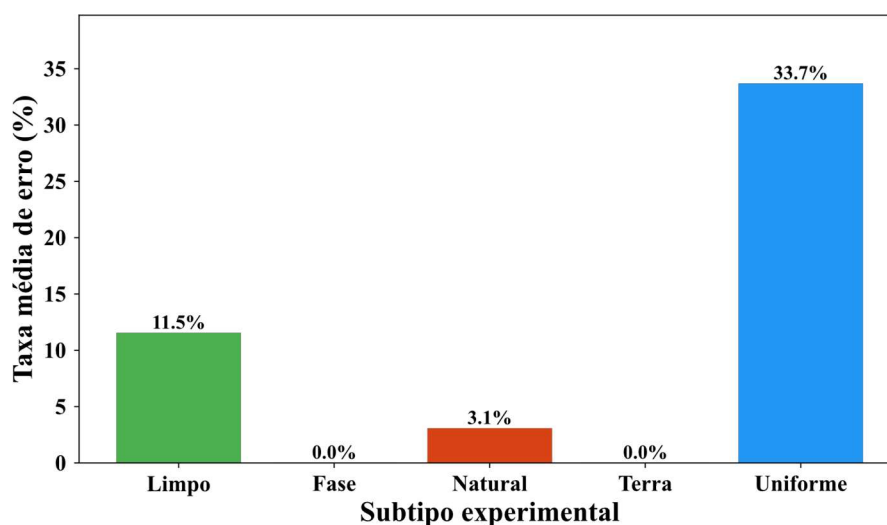


Fonte: autoria própria.

A análise da matriz de confusão é importante porque permite diferenciar os tipos de erro cometidos pelo classificador. No contexto do diagnóstico de isoladores, falsos negativos são particularmente relevantes, pois correspondem a grupos poluídos classificados como limpos. Esse tipo de erro pode comprometer a identificação preventiva de uma condição superficial indesejada. Por outro lado, falsos positivos representam grupos limpos classificados como poluídos, o que tende a produzir uma decisão mais conservadora, embora possa levar a inspeções adicionais.

Além da análise por classe, também se investiga o erro médio do modelo final em função do subtipo experimental. Essa análise complementa a matriz de confusão, pois permite identificar quais condições de poluição são mais difíceis para o classificador. Na Figura 35 ilustra-se o erro médio por subtipo experimental para o modelo final.

Figura 35. Erro médio por subtipo experimental do modelo SVM.



Fonte: autoria própria.

Conforme apresentado na Figura 35, o maior erro médio ocorre no subtipo Uniforme, com valor de 33,69%. Esse resultado indica que a poluição uniforme representa a condição mais desafiadora para o modelo final entre os subtipos avaliados. Uma possível explicação é que, ao produzir alterações mais distribuídas ao longo do perfil, a poluição uniforme pode gerar padrões térmicos menos localizados e, em alguns casos, mais próximos da variação global observada em perfis limpos, conforme observado na Figura 32. Assim, a distinção entre limpo e poluído pode se tornar mais difícil quando a alteração térmica não se concentra em regiões específicas do isolador.

A classe Limpo apresenta erro médio de 11,55%, indicando que parte dos grupos limpos é classificada como poluída. Esse tipo de erro sugere uma tendência conservadora do modelo em algumas situações, possivelmente associada à presença de variações térmicas naturais nos isoladores limpos ou a diferenças ambientais entre os ensaios. Embora falsos positivos não representem falha direta na detecção de poluição, eles são relevantes do ponto de vista operacional, pois podem indicar a necessidade de inspeções desnecessárias.

O subtipo Natural apresenta erro médio de 3,08%, valor inferior ao observado para Limpo e Uniforme. Esse resultado sugere que os perfis associados à poluição natural apresentam características térmicas suficientemente distintas da condição limpa para serem identificados pelo modelo final na maioria das repetições. Já os subtipos Fase e Terra não apresentam erro médio foi nulo, indicando que as condições de poluição localizadas nessas regiões são corretamente reconhecidas pelo SVM. Esse

comportamento pode estar associado ao fato de que contaminações concentradas produzem alterações térmicas mais marcadas em regiões específicas do perfil, favorecendo a separação por atributos agregados que capturam alterações de nível e intensidade.

De modo geral, a análise de robustez confirma que o SVM com representação agregada é o modelo mais adequado entre os campeões avaliados, considerando o critério multicritério por grupo de captura. O modelo apresenta o maior desempenho médio, mantém coerência com os resultados da validação aninhada e utiliza uma representação simples e interpretável. Ainda que as representações sequencial e simbólica apresentem desempenho competitivo, elas não superam o SVM quando se considera o equilíbrio entre desempenho, estabilidade e simplicidade. Assim, os resultados indicam que, para a base analisada, a informação discriminativa principal dos perfis térmicos pode ser adequadamente representada por atributos agregados, especialmente aqueles associados ao nível térmico global do sinal. Entretanto, o desempenho alcançado pelos modelos sequenciais mostra que a organização espacial das janelas ao longo do isolador também contém informação relevante para a classificação. Dessa forma, os resultados não descartam a abordagem sequencial, mas indicam que, nas condições experimentais e com a quantidade de dados disponível neste estudo, a preservação explícita da ordem espacial não proporciona ganho suficiente para superar a representação agregada. A representação sequencial permanece, portanto, como uma alternativa promissora para bases mais amplas e diversificadas, nas quais padrões térmicos localizados, transições entre regiões e dependências espaciais possam ser mais bem explorados.

## 6 CONCLUSÕES

Neste estudo, foi apresentada uma metodologia para a classificação de perfis térmicos de isoladores poliméricos, visando ao diagnóstico de condições associadas à poluição superficial. A abordagem proposta foi fundamentada na extração do perfil térmico longitudinal do isolador, na correção das temperaturas em relação ao ambiente, na segmentação do perfil em janelas, na extração de atributos térmicos e na avaliação de diferentes modelos de classificação. Dessa forma, buscou-se verificar se as informações térmicas distribuídas ao longo do isolador poderiam ser utilizadas como indicadores do seu estado operacional.

A metodologia adotada permitiu estruturar os perfis térmicos em diferentes formas de representação, contemplando a representação agregada, a representação sequencial e a representação simbólica. Essa organização possibilitou comparar modelos que utilizam vetores fixos de atributos, como SVM e Random Forest, com modelos que preservam a estrutura sequencial do perfil, como HMM, HMM-GMM e LSTM. Assim, foi possível avaliar não apenas o desempenho dos classificadores, mas também o efeito da forma de representação dos dados térmicos sobre a capacidade de distinguir isoladores limpos e isoladores poluídos.

A base de dados utilizada foi composta por perfis térmicos extraídos de termografias de isoladores em diferentes condições experimentais, incluindo isoladores limpos, isoladores com poluição artificial concentrada no lado fase, poluição artificial concentrada no lado terra, poluição artificial uniforme e poluição natural. Essa composição permitiu avaliar o comportamento dos modelos diante de diferentes manifestações da contaminação superficial, considerando variações na distribuição espacial da poluição e na origem do processo de contaminação.

É importante destacar que a avaliação dos modelos foi conduzida considerando a separação por grupos de captura. Esse procedimento tornou a análise mais rigorosa, pois impediu que perfis térmicos provenientes de uma mesma condição experimental fossem utilizados simultaneamente no treinamento e no teste. Dessa forma, os resultados obtidos refletem de maneira mais conservadora a capacidade dos modelos em generalizar para novas capturas, reduzindo o risco de superestimação do desempenho.

Na etapa de validação aninhada, verificou-se que todos os modelos avaliados apresentaram capacidade de distinguir perfis térmicos limpos e poluídos. Na avaliação por perfil térmico, a LSTM apresentou desempenho destacado, indicando que modelos recorrentes podem explorar padrões locais existentes na sequência de janelas do perfil térmico. Entretanto, na avaliação por grupo de captura, que representa a decisão experimental mais conservadora adotada neste trabalho, o SVM com representação agregada apresentou o melhor desempenho geral.

Além disso, a análise de robustez confirmou a superioridade do SVM com representação agregada. O modelo apresentou acurácia balanceada de  $0,923 \pm 0,058$ , F1-macro de  $0,923 \pm 0,062$ , coeficiente de correlação de Matthews de  $0,845 \pm 0,083$  e métrica multicritério de  $0,897 \pm 0,067$ . Esses resultados indicam que a configuração selecionada apresentou desempenho elevado e comportamento estável diante de novas divisões experimentais, demonstrando que sua escolha não dependeu apenas de uma partição específica dos dados.

Embora o SVM com representação agregada tenha apresentado o melhor desempenho final, os modelos sequenciais também forneceram resultados relevantes. O HMM-GMM, em especial, apresentou desempenho próximo ao modelo final, demonstrando que a modelagem sequencial do perfil térmico é uma alternativa promissora para o diagnóstico de isoladores. Esse resultado reforça a hipótese de que a organização espacial do perfil térmico contém informação discriminativa. No entanto, para a base analisada, a preservação explícita da sequência não resultou em desempenho superior ao da representação agregada.

A representação simbólica, utilizada no HMM discreto, também apresentou desempenho competitivo. Contudo, observou-se que a discretização pode reduzir parte da informação térmica disponível, principalmente quando os padrões de separação entre as classes dependem de variações sutis de temperatura. Ainda assim, essa abordagem permanece relevante por sua interpretabilidade e pela possibilidade de representar o perfil térmico como uma sequência de estados ou símbolos associados a diferentes regiões térmicas do isolador.

Na análise dos erros do modelo final, verificou-se que a poluição uniforme foi o subtipo experimental mais desafiador. Esse comportamento pode estar associado ao fato de que esse tipo de contaminação tende a produzir alterações térmicas menos localizadas ao longo do perfil, dificultando a separação em relação à condição limpa. Por outro lado,

poluições mais concentradas ou naturalmente distribuídas podem gerar padrões térmicos mais característicos, facilitando sua identificação pelos classificadores.

Dessa forma, conclui-se que a análise de perfis térmicos constitui uma abordagem promissora para o diagnóstico de isoladores poliméricos submetidos à poluição superficial. A metodologia proposta mostrou-se capaz de transformar informações termográficas em atributos representativos da condição do equipamento, permitindo a classificação entre isoladores limpos e poluídos por meio de modelos de aprendizado de máquina.

Outrossim, os resultados obtidos indicam que a representação agregada dos atributos térmicos, quando associada a um processo adequado de seleção de atributos e validação, pode fornecer desempenho robusto mesmo sem preservar explicitamente toda a sequência espacial do perfil. Esse resultado é particularmente relevante em cenários com quantidade limitada de dados, nos quais modelos mais simples e com menor flexibilidade podem apresentar maior estabilidade e melhor capacidade de generalização.

Portanto, o método desenvolvido contribui para o aprimoramento de técnicas não invasivas de monitoramento de isoladores, oferecendo uma alternativa baseada em termografia, extração de atributos e classificação automática. A proposta pode auxiliar no desenvolvimento de ferramentas de apoio à manutenção preditiva, permitindo a identificação de padrões térmicos associados à poluição superficial e contribuindo para a segurança operacional de sistemas elétricos de potência.

## 6.1 TRABALHOS FUTUROS

Como trabalhos futuros, recomenda-se a ampliação da base de dados utilizada neste estudo, contemplando um número maior de isoladores, grupos de captura e condições experimentais. Essa ampliação permitiria avaliar a capacidade de generalização dos modelos em cenários mais diversos, incluindo diferentes níveis de poluição, diferentes materiais isolantes, distintas geometrias de isoladores e variações ambientais mais amplas.

Outra possibilidade de continuidade consiste em expandir o problema de classificação para além da distinção binária entre isoladores limpos e poluídos. Em estudos futuros, pode-se investigar a classificação multiclases, considerando separadamente os diferentes subtipos de poluição, como poluição concentrada no lado fase, poluição concentrada no lado terra, poluição uniforme e poluição natural.

Ademais, recomenda-se aprofundar a investigação dos modelos sequenciais. Embora o SVM com representação agregada tenha apresentado o melhor desempenho final neste estudo, os resultados obtidos pelos modelos HMM-GMM e LSTM indicam que a estrutura espacial do perfil térmico possui informação relevante. Dessa forma, trabalhos futuros podem avaliar diferentes tamanhos de janela, diferentes taxas de sobreposição, novas estratégias de segmentação e arquiteturas sequenciais mais ajustadas à natureza dos perfis térmicos.

No caso dos Modelos Ocultos de Markov, estudos futuros podem explorar diferentes números de estados ocultos, diferentes estruturas de transição e formas alternativas de modelagem das emissões. Também seria interessante investigar a relação entre os estados ocultos estimados e regiões físicas específicas do isolador, buscando associar os estados do modelo a fenômenos como aquecimento localizado, formação de bandas secas, variações de condutividade superficial ou regiões de maior solicitação elétrica.

Além da aplicação dos Modelos Ocultos de Markov para classificação, trabalhos futuros podem investigar seu uso como modelo generativo para a criação de perfis térmicos sintéticos. Uma vez ajustado aos perfis reais, o HMM pode ser utilizado para gerar novas sequências térmicas que preservem, em alguma medida, as transições espaciais e os padrões estatísticos observados nos dados experimentais. Essa abordagem pode contribuir para a ampliação da base de dados, especialmente em situações nas quais a aquisição de novas amostras laboratoriais ou de campo seja limitada por custo, tempo, disponibilidade de equipamentos ou dificuldade de reprodução controlada das condições de poluição.

Nesse contexto, os perfis sintéticos gerados poderiam ser avaliados por meio de comparações estatísticas e estruturais com os perfis reais, considerando atributos térmicos, variações locais, dependência espacial entre janelas e distribuição das temperaturas ao longo do isolador. Além disso, esses perfis poderiam ser empregados em estratégias de aumento de dados, treinamento complementar de classificadores e análise de cenários intermediários de poluição. Dessa forma, o uso generativo do HMM permitiria não apenas ampliar artificialmente a base disponível, mas também investigar a capacidade do modelo de representar padrões térmicos associados a diferentes condições de superfície.

Quanto à representação simbólica, uma possibilidade de aprimoramento consiste em desenvolver métodos de discretização mais adaptativos. A definição dos símbolos

pode ser baseada na distribuição estatística dos dados, em limiares fisicamente interpretáveis ou em regiões térmicas associadas à geometria do isolador. Essa melhoria pode reduzir a perda de informação provocada pela quantização e tornar o HMM discreto mais competitivo e interpretável.

Também é possível explorar abordagens baseadas na matriz térmica bidimensional, em vez de utilizar apenas o perfil térmico unidimensional. Essa alternativa permitiria analisar a distribuição espacial completa da temperatura sobre o isolador, podendo capturar padrões que não são totalmente representados pelo perfil longitudinal. Para isso, poderiam ser avaliadas técnicas de processamento de imagem, extração de regiões de interesse e modelos de classificação baseados em imagens térmicas.

Além disso, recomenda-se avaliar a aplicação da metodologia em condições reais de campo. A validação em ambiente operacional permitiria verificar o desempenho do método diante de variações naturais de iluminação, distância, ângulo de aquisição, vento, umidade e interferências presentes em subestações e linhas de transmissão. Essa etapa é fundamental para transformar a metodologia proposta em uma ferramenta prática de apoio à inspeção e à manutenção preditiva.

Por fim, trabalhos futuros podem investigar a aplicação da abordagem proposta a outros equipamentos de alta tensão cujas falhas ou processos de degradação se manifestem por alterações térmicas, como buchas de transformadores, polos de disjuntores, transformadores de potencial, transformadores de corrente e conexões elétricas. Essa expansão permitiria verificar a generalidade da metodologia e contribuir para o desenvolvimento de sistemas mais abrangentes de diagnóstico termográfico em equipamentos elétricos.

## REFERÊNCIAS

AHMED, R. *et al.* Field-Dependent Pollution Model under Polluted Environments for Outdoor Polymeric Insulators. **Polymers**, 2022.

ARAUJO, Bruno Vinicius S. *et al.* Monitoring ZnO surge arresters using convolutional neural networks and image processing techniques combined with signal alignment. **Measurement**, 248, 2025.

BALASUBRAMANIAN, Madanmohan; VENKATESH, S. Assessment of Partial Discharge Signature s in Transformer Oil Insulation using Hidden Markov Model. **International Journal on Electrical Engineering and Informatics**, 2016.

BARREIRA, E.; ALMEIDA, R.M.S.F.; SIMÕES, M.L. Emissivity of Building Materials for Infrared Measurements. **Sensors**, 2021.

BINDI, M. *et al.* A Comprehensive Review of Fault Diagnosis and Prognosis Techniques in High Voltage and Medium Voltage Electrical Power Lines. **Energies**, 2023.

BISHOP, Christopher M. **Pattern Recognition and Machine Learning**. [S.l.]: Springer, 2006.

BREIMAN, L. Random Forests. **Machine Learning**, 2001.

BRITO, K.B. *et al.* Development of DIP-based algorithm for extraction of UV video attributes from corona discharges on polymeric insulators. **International Journal of Electrical Power & Energy Systems**, 2022.

BÜYÜKKEÇECİ, M.; OKUR, M. C. A Comprehensive Review of Feature Selection and Feature Selection Stability in Machine Learning. **Gazi University Journal of Science**, 2023.

CARVALHO, Itaiara Felix. **Identificação de fontes de descargas parciais por atributos extraídos da envoltória do sinal**. [S.l.]. 2024.

CHOUDHARY, Anurag; GOYAL, Deepam; LETHA, Shimi Sudha. Infrared Thermography-Based Fault Diagnosis of Induction Motor Bearings Using Machine Learning. **IEEE Sensors Journal**, 2021.

CORTES, C.; VAPNIK, V. Support-vector networks. **Mach Learn**, 1995.

DAS, Arup Kumar *et al.* Convolutional neural network and Bi-directional long short memory hybrid deep network aided infrared image classification framework for non-contact monitoring of overhead insulators. **IET Signal Processing**, 2022.

DIAS, Bruno Albuquerque. **GESTÃO TÉCNICA DE ISOLADORES POLIMÉRICOS UTILIZANDO APRENDIZADO DE MÁQUINA**. [S.l.]. 2021.

DINIZ, Larissa. **APRIMORAMENTO DA INSPEÇÃO DE ISOLADORES POLIMÉRICOS POLIMÉRICOS POR TERMOGRAFIA E CAMPO ELÉTRICO**. [S.l.]. 2019.

FERREIRA, Tarso Vilela. **Estudo do Trilhamento de Núcleo em Isoladores Poliméricos**. [S.l.]. 2007.

FLIR. **User's manual FLIRT8xxseries**. FLIR Systems. [S.l.]. 2019.

FREIRE, J. C. Arouche *et al.* Transmission Line Fault Classification Using Hidden Markov Models. **IEEE Access**, 2019.

FRONTIN, Sergio O. **Equipamentos de alta tensão – prospecção e hierarquização de inovações tecnológicas**. Brasília: Goya Editora LTDA., 2013.

GÉRON, Aurélien. **Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow**. [S.l.]: O'Reilly Media, Inc., 2019.

GONZALEZ, Rafael C.; WOODS, Richard E. **Digital Image Processing**. [S.l.]: Pearson, 2018.

GORUR, Ravi S.; CHERNEY, Edward A.; BURNHAM, Jeffrey T. **Outdoor Insulators**. Phoenix, Arizona: [s.n.], 1999.

GOUDA, Osama E. *et al.* Pollution Severity Monitoring of High Voltage Transmission Line Insulators Using Wireless Device Based on Leakage Current Bursts. **IEEE Access**, 2022.

HAN, S. *et al.* Statistical analysis of infrared thermogram for CNN-based electrical equipment identification methods. **Applied Artificial Intelligence**, 2022.

HAYKIN, Simon. **Neural Networks and Learning Machines**. [S.l.]: Pearson, 2009.

HOCHREITER, Sepp and Schmidhuber, Jürgen. Long Short-Term Memory. **Neural Computation**, 1997.

HONG, Kaixing; LIN, Guanxi. State classification of transformers using nonlinear dynamic analysis and Hidden Markov models. **Measurement**, 2019.

HOU, F. *et al.* Review on Infrared Imaging Technology. **Sustainability**, 2022.

HOUDT, Greg Van; MOSQUERA, Carlos; NÁPOLES, Gonzalo. A Review on the Long Short-Term Memory Model. **Artificial Intelligence Review**, 2020.

HUANG, Qingqing and Shao, Leilai and Li, Na. Dynamic Detection of Transmission Line Outages Using Hidden Markov Models. **IEEE Transactions on Power Systems**, 2016.

HUDA, A.S. Nazmul; TAIB, Soib. Application of infrared thermography for predictive/preventive maintenance of thermal defect in electrical equipment. **Applied Thermal Engineering**, 2013.

IEC. **IEC 60815-1: Selection and dimensioning of high-voltage insulators intended for use in polluted conditions – Part 1: Definitions, information and general principles**. Genebra. 2008.

JAMES, Gareth *et al.* **An Introduction to Statistical Learning**. Springer: [s.n.], 2013.

JIANQUAN, Fang *et al.* **Prediction Method of Line Loss Data Based on Hidden Markov Model**. 2022 IEEE 6th Conference on Energy Internet and Energy System Integration (EI2). [S.l.]: [s.n.], 2022.

KOUADRI, Abdelmalek *et al.* Hidden Markov model based principal component analysis for intelligent fault diagnosis of wind energy converter systems. **Renewable Energy**, 2020.

KUFFEL, E.; ZAENGL, W. S.; KUFFEL, J. **High Voltage Engineering - Fundamentals**. [S.l.]: [s.n.], 2000.

KUMAR, Partik and Abhyankar, Abhijit R. **A Time Efficient Factorial Hidden Markov Model Based Approach for Non-Intrusive Load Monitoring**. 2023 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE). [S.l.]: [s.n.]. 2023.

LAFARGUE-TALLET, Thomas *et al.* Active thermo-reflectometry for absolute temperature measurement by infrared thermography on specular materials. **Sci Rep**, 2022.

LEKSIR, Y. Laib dit; MANSOUR, M.; MOUSSAOUI, A. Localization of thermal anomalies in electrical equipment using Infrared. **Infrared Physics & Technology**, 2017.

LEWIS, MJ *et al.* nestedcv: an R package for fast implementation of nested cross-validation with embedded feature selection designed for transcriptomics and high-dimensional data. **Bioinformatics Advances**, 2023.

LI, H. *et al.* A Weighted Estimation Algorithm for Enhancing Pulsed Eddy Current Infrared Image in Ecp Non-Destructive Testing. **Appl. Sci.**, 2019.

LI, Te *et al.* Temperature distribution characteristics and heat defect judgment method based on temperature gradient of suspended composite insulator in operation. **IET Generation, Transmission & Distribution**, 2021.

LIANG, Huajie *et al.* Infrared image-based multi-point temperature monitoring for ultra-high-voltage direct current wall bushings subject to insulation shed interference. **High Voltage**, 2025.

LIU, Q. *et al.* A Novel Health Prognosis Method for a Power System Based on a High-Order Hidden Semi-Markov Model. **Energies**, 2021.

LLOYD, S. Least squares quantization in PCM. **IEEE Transactions on Information Theory**, 1982.

LUAN, Wenpeng *et al.* Industrial load disaggregation based on Hidden Markov Models. **Electric Power Systems Research**, 2022.

MATTHEWS, B.W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. **Biochimica et Biophysica Acta (BBA) - Protein Structure**, 1975.

MAX, J. Quantizing for minimum distortion. **IRE Transactions on Information Theory**, 1960.

MILOSEVIC, M; JANKOVIC, D; PEULIC, A. Thermography based breast cancer detection using texture features and minimum variance quantization. **EXCLI journal**, 2014.

MIYA, W. S. and Mpanza, L. J. and Marwala, T. and Nelwamondo, Fulufhelo V. **Condition monitoring of oil-impregnated paper bushings using Extension Neural Network, Gaussian Mixture and Hidden Markov Models**. [S.l.]: [s.n.]. 2008.

MONTGOMERY, Douglas C.; RUNGER, George C. **Applied statistics and probability for engineers**. [S.l.]: Wiley, 2018.

MUNIZ, Pablo Rodrigues; MENDES, Mariana Altoé. **Termografia infravermelha aplicada à manutenção elétrica - Dos fundamentos ao diagnóstico**. Vitória, ES: Edifes, 2019.

MURPHY, Kevin P. **Machine Learning: A Probabilistic Perspective**. [S.l.]: The MIT Press, 2012.

NAJAFI, Mohamad *et al.* Fault Diagnosis of Electrical Equipment through Thermal Imaging and Interpretable Machine Learning Applied on a Newly-introduced Dataset. **2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)**, 2020.

NÓBREGA, Samuel Cezarino da. **APLICAÇÃO DE TÉCNICAS DE VISÃO COMPUTACIONAL NO RECONHECIMENTO E NA DETECÇÃO DE DEFEITO EM ISOLADORES POLIMÉRICOS DE SUSPENSÃO**. [S.l.]. 2024.

NOVIZON *et al.* ondition based monitoring of gapless surge arrester using electrical and thermal parameters. **IOP Conference Series: Materials Science and Engineering**, 2019.

OLIVEIRA, J. H. P. de *et al.* **Comparative Analysis of Thermal Profile Alignment Techniques for Surge Arresters**. 2025 Workshop on Communication Networks and Power Systems (WCNPS). [S.l.]: [s.n.]. 2025.

PAN, Dong *et al.* Interference Factors and Compensation Methods When Using Infrared Thermography for Temperature Measurement: A Review. **IEEE Transactions on Instrumentation and Measurement**, 2025.

PAPAIOANNOU, N. *et al.* The Role of Mutual Information Estimator Choice in Feature Selection: An Empirical Study on mRMR. **Information**, 2025.

PENG, Hanchuan; LONG, Fuhui; DING, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2005.

PERIYASAMY, Shobhana *et al.* Support vector machine based methodology for classification of thermal images pertaining to breast cancer. **Journal of Thermal Biology**, 2022.

QIAO, X. *et al.* Aging of Polymeric Insulators under Various Conditions and Environments: Another Look. **Energies**, 2022.

RABINER, L. R. A tutorial on hidden Markov models and selected applications in speech recognition. **Proceedings of the IEEE**, 1989.

RABINER, L. R. A tutorial on hidden Markov models and selected applications in speech recognition. **Proceedings of the IEEE**, v. 77, p. 257-286, 1989.

RASCHKA, Sebastian; MIRJALILI, Vahid. **Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2**. [S.l.]: Packt, 2019.

SALEM, Ali Ahmed *et al.* Polymeric Insulator Conditions Estimation by Using Leakage Current Characteristics Based on Simulation and Experimental Investigation. **Polymers**, 2022.

SALEM, Ali Ahmed *et al.* Investigating and Modeling Ageing Effects on Polymeric Insulator Electrical Properties. **IEEE Access**, 2023.

SANTORO, Luca; SESANA, Raffaella. A physics-informed framework for feature extraction and defect segmentation in pulsed infrared thermography. **Engineering Failure Analysis**, 2025.

SANTOS, Héctor de; SANZ-BOBI, Miguel Á. A machine learning approach for condition monitoring of high voltage insulators in polluted environments. **Electric Power Systems Research**, 220, 2023.

SANTOS, Laerte dos. **Classificação e Modelagem de Fatores de Influência sobre Inspeções Termográficas em Ambientes Desabrigados**. [S.l.]. 2012.

SASIKUMAR, B; RAJENDRAN, P. FAULT DETECTION AND CLASSIFICATION USING DENSENET, SURF-BASED ANN, AND INFRARED THERMOGRAPHY. **ICTACT Journal on Image and Video Processing**, 2024.

SHAFIQUE, Usman *et al.* Infrared Thermography-Based Insulator Fault Classification via Unsupervised Clustering and Semi-Supervised Learning. **IEEE Access**, 2024.

SHI, Zhigang, Zhao, Yunlong, Liu, Zhanshuang, Zhang, Yanan, Ma, Le. Diagnosis and Classification Decision Analysis of Overheating Defects of Substation Equipment Based on Infrared Detection Technology. **Scientific Programming**, 2021.

SINGH, Laxman *et al.* Design of thermal imaging-based health condition monitoring and early fault detection technique for porcelain insulators using Machine learning. **Environmental Technology & Innovation**, 2021.

SMYTH, Padhraic. Hidden Markov models for fault detection in dynamic systems. **Pattern Recognition**, 1994.

SOUALHI, A.; CLERC, G.; RAZIK, H. **Fault prognosis based on Hidden Markov Models**. 2015 IEEE Workshop on Electrical Machines Design, Control and Diagnosis (WEMDCD). [S.l.]: [s.n.]. 2015.

SUI, Xuewei *et al.* **Research on Power Transformer Fault Diagnosis Based on Improved Wavelet Packet Energy and Hidden Markov Model**. 2023 IEEE 6th International Electrical and Energy Conference (CIEEC). [S.l.]: [s.n.]. 2023.

SUN, Qiwen *et al.* **Infrared Image-Based Fault Warning for Substation High-Voltage Switchgear via Multi-Scale Feature Optimization**. 2025 International Conference of Clean Energy and Electrical Engineering (ICCEEE). [S.l.]: [s.n.]. 2025.

TANG, Z.; JIAN, X. Thermal fault diagnosis of complex electrical equipment based on infrared image recognition. **Scientific Reports**, 2024.

TANG, Z.; JIAN, X. Thermal fault diagnosis of complex electrical equipment based on infrared image recognition. **Sci Rep**, 2024.

ULLAH, I. *et al.* Predictive Maintenance of Power Substation Equipment by Infrared Thermography Using a Machine-Learning Approach. **Energies**, 2017.

ULLAH, I. *et al.* Deep Learning Image-Based Defect Detection in High Voltage Electrical Equipment. **Energies**, 2020.

- VENKATESH, S.; GOPAL, S.; KANNAN, K. Bharathi. **A NOVEL HYBRID CONTINUOUS DENSITY HIDDEN MARKOV MODEL-PROBABILISTIC NEURAL NETWORK FOR MULTIPLE SOURCE PARTIAL DISCHARGE PATTERN RECOGNITION**. [S.l.]: [s.n.]. 2011.
- VILLALOBOS, Rodrigo J. *et al.* A New Current Transducer for On-Line Monitoring of Leakage Current on HV Insulator Strings. **IEEE Access**, 2022.
- WANDERLEY NETO, Estácio Tavares. Monitoramento e Diagnóstico de Pára-Raios a ZnO através de Redes Neurais Artificiais. **Tese de doutorado**, Agosto 2007.
- WANG, Jianguo *et al.* Leakage Current Response Mechanism of Insulator String With Ambient Humidity on Days Without Rain. **IEEE Access**, 2019.
- WANG, Jianguo *et al.* Online Monitoring of Electrical Equipment Condition Based on Infrared Image Temperature Data Visualization. **IEEJ Trans Elec Electron Eng**, 2022.
- WANG, Z. *et al.* Improved YOLO11 with Mamba-2 (SSD) and Triplet Attention for High-Voltage Bushing Fault Detection from Infrared Images. **Energies**, 2026.
- WEN, F.; JIN, J. Detection of Superficial Defects in the Insulation Layers of Aviation Cables by Infrared Thermographic Technique. **Coatings**, 2022.
- XIAO, Fei; AI, Qian. Data-Driven Multi-Hidden Markov Model-Based Power Quality Disturbance Prediction That Incorporates Weather Conditions. **IEEE Transactions on Power Systems**, 2019.
- XIAO, Siyi *et al.* **Hidden Markov Model-Based Energy Deviation Mitigation for Coupled Power and Transportation**. 2025 IEEE International Conference on Energy Technologies for Future Grids (ETFG). [S.l.]: [s.n.]. 2025.
- YUSUF, Syed *et al.* **Fault classification improvement in industrial condition monitoring via Hidden Markov Models and Naïve Bayesian modeling**. 2013 IEEE Symposium on Industrial Electronics & Applications. [S.l.]: [s.n.]. 2013.
- ZHAN, Zhongqiang and Yang, Dingqian and Hu, Gongduo and Chen, Gang and Xu, Guanghu and Chen, Jie. **Transformer Risk Assessment Method Based on Hidden Markov Model**. 2024 International Conference on HVDC (HVDC). [S.l.]: [s.n.]. 2024.
- ZHANG, Bo *et al.* Hidden Markov Model-Based Cyberattack Prediction in Power Systems. **IEEE Transactions on Smart Grid**, 2025.
- ZHANG, D.; MENG, F. Research on the Interrelation between Temperature Distribution and Dry Band on Wet Contaminated Insulators. **Energies**, 2019.
- ZHANG, Jiayuan; ZHAN, Wei; EHSANI, M. **On-line fault diagnosis of electric machine based on the Hidden Markov Model**. 2016 IEEE Transportation Electrification Conference and Expo (ITEC). [S.l.]: [s.n.]. 2016.
- ZHAO, Tiebin; BERNSTORF, R. A. Ageing tests of polymeric housing materials for non-ceramic insulators. **IEEE Electrical Insulation Magazine**, v. 14, p. 26-33, 1998.

ZHAO, Zhenyu; ANAND, Radhika; WANG, Mallory. **Maximum Relevance and Minimum Redundancy Feature Selection Methods for a Marketing Machine Learning Platform**. 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA). [S.l.]: [s.n.]. 2019.

ZOHA, Ahmed and Gluhak, Alexander and Nati, Michele and Imran, Muhammad Ali. **Low-power appliance monitoring using Factorial Hidden Markov Models**. 2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing. [S.l.]: [s.n.]. 2013.

ZOU, Hui; HUANG, Fuzhen. A novel intelligent fault diagnosis method for electrical equipment using. **Infrared Physics & Technology**, 2015.

ZWILLINGER, Daniel; KOKOSKA, Stephen. **Standard probability and statistics tables and formulae**. [S.l.]: Chapman & Hall/CRC, 2000.