CENTRO UNIVERSITÁRIO FEI
YURI MARINHO OLIVATTI

DETECÇÃO AUTOMÁTICA DE PROBLEMAS ELÉTRICOS POR MEIO DA ANÁLISE COMPUTACIONAL DE IMAGENS TÉRMICAS

São Bernardo do Campo
2019

YURI MARINHO OLIVATTI

DETECÇÃO AUTOMÁTICA DE PROBLEMAS ELÉTRICOS POR MEIO DA ANÁLISE COMPUTACIONAL DE IMAGENS TÉRMICAS

Dissertação de Mestrado apresentada ao Centro Universitário FEI, como parte dos requisitos necessários para obtenção do título de Mestre em Engenharia Elétrica, orientado pelo Prof. Dr. Plinio Thomaz Aquino Junior.

São Bernardo do Campo

2019

Marinho Olivatti, Yuri.
DETECÇÃO AUTOMÁTICA DE PROBLEMAS ELÉTRICOS POR MEIO DA ANÁLISE COMPUTACIONAL DE IMAGENS TÉRMICAS / Yuri Marinho Olivatti. São Bernardo do Campo, 2019.
106 p. : il.

Dissertação - Centro Universitário FEI.
Orientador: Prof. Dr. Plinio Thomaz Aquino Junior.
Coorientador: Prof. Yuri Marinho Olivatti.

1. Termografia. 2. Energia elétrica. 3. Inteligência Artificial. 4. Machine Learning. I. Thomaz Aquino Junior, Plinio, orient. II. Título.

Elaborada pelo sistema de geração automática de ficha catalográfica da FEI com os dados fornecidos pelo(a) autor(a).

centro universitário FEI

APRESENTAÇÃO DE DISSERTAÇÃO
ATA DA BANCA EXAMINADORA

Programa de Pós-Graduação Stricto Sensu em Engenharia Elétrica

Mestrado

PGE-10

Aluno: Yuri Marinho Olivatti **Matrícula:** 116315-3

Título do Trabalho: Detecção automática de problemas elétricos por meio da análise computacional de imagens térmicas.

Área de Concentração: Inteligência Artificial Aplicada à Automação e Robótica

Orientador: Prof. Dr. Plinio Thomaz Aquino Junior

Data da realização da defesa: 26/03/2019

ORIGINAL ASSINADA

Avaliação da Banca Examinadora:

_____

_____

_____

_____

_____

São Bernardo do Campo, / / .

MEMBROS DA BANCA EXAMINADORA	
Prof. Dr. Plinio Thomaz Aquino Junior	Ass.: _____
Prof. Dr. Renato Camargo Giacomini	Ass.: _____
Prof. Dr. Carlos Alberto Kamienski	Ass.: _____

A Banca Julgadora acima-assinada atribuiu ao aluno o seguinte resultado:

APROVADO ☒ REPROVADO ☐

VERSÃO FINAL DA DISSERTAÇÃO

APROVO A VERSÃO FINAL DA DISSERTAÇÃO EM QUE FORAM INCLUÍDAS AS RECOMENDAÇÕES DA BANCA EXAMINADORA

_____

Aprovação do Coordenador do Programa de Pós-graduação

_____

Prof. Dr. Carlos Eduardo Thomaz

AGRADECIMENTOS

Agradeço a Deus por todas as oportunidades que tive e pela realização do mestrado. Agradeço a minha família e amigos por terem me apoiado a todo momento, e permitido que eu chegasse até aqui.

Ao Prof. Dr. Rodrigo Filev Maia, pelo convite e ingresso ao mestrado, por todas as conversas, sugestões, apoio e suporte e apoio durante todo o mestrado.

Ao Prof. Dr. Plínio Thomaz Aquino Junior pela orientação que me deu, pela paciência durante todo o período do mestrado e produção da dissertação e todos os ensinamentos que compartilhou.

Ao Prof. Dr. Renato Giacomini, pelo apoio dado nos experimentos, e todas as ideias e sugestões, viabilizando esse trabalho.

Ao Acácio e técnicos do CLE (Centro de Laboratórios Elétricos) por todo o suporte, infraestrutura e recursos fornecidos para os experimentos.

Ao Centro Universitário da FEI que forneceu apoio financeiro e toda a infraestrutura necessária.

A Vivo/Telefônica na parceria com a FEI e laboratório de IoT pelo financiamento do projeto de mestrado, recursos e infraestrutura.

Agradeço aos membros do Laboratório de IoT que estiveram sempre ao meu lado, fornecendo ideias e auxílio, e em especial a Cris que sempre esteve comigo nessa parceria de mestrado, compartilhando ideias, pesquisando e refletindo junto sobre descobertas e me dando apoio durante o mestrado.

RESUMO

A energia elétrica está entre os mais importantes recursos necessários na vida contemporânea. Apesar disso, parcela significativa do que é despendido na geração de energia é desperdiçado, já que parte da energia gerada não chega no consumidor final. Com perdas de dinheiro e recurso, além dos impactos ambientais envolvidos na produção energética, se faz necessário uma melhor gestão da energia na totalidade do processo. Considerando as elevadas taxas de perda de energia nas usinas e linhas de transmissão devido a problemas elétricos e irregularidades não previstas, um dos desafios nesta gestão é justamente como reduzir as falhas e otimizar a distribuição de energia. O gerenciamento de recursos em cidades é justamente um dos tópicos de estudo em cidades inteligentes e internet das coisas, conceitos apoiados em tecnologias de informação e comunicação, que visam promover eficiência nos processos já existentes e solucionar problemas. A tecnologia de sensoriamento de múltiplos componentes como ambiente de IoT é um caminho para o monitoramento remoto e gestão da rede elétrica, permitindo a análise apurada, em tempo real, do que realmente está acontecendo e realizar a identificação das regiões de anomalias na rede elétrica. Sempre considerando a utilização da tecnologia e a viabilidade de soluções mais eficientes, uma técnica para inspeção elétrica cada vez mais usada é a de termografia. As câmeras térmicas, que podem ser consideradas como sensores não invasivos, permitem medir o calor emitido por equipamentos e componentes elétricos, obter perfis térmicos e perceber pontos de variação de temperatura, informações que podem indicar potenciais falhas e anomalias na rede e em equipamentos elétricos. Este trabalho propõe o estudo de um método para realizar essa análise de forma automática por intermédio do uso de algoritmos de IA e *machine learning* para a identificação e classificação de problemas elétrico. Um experimento foi proposto e executado, gerando um *dataset* de imagens térmicas em um cenário de problemas elétricos. Por meio da análise realizada, foi possível validar a possibilidade da utilização de classificadores para identificar problemas elétricos em determinado contexto, bem como entender as melhores combinações de algoritmos de processamento de imagens e classificação para cada tipo de problema estudado, permitindo a obtenção de um método geral para coleta de dados, processamento de imagens térmicas e validação de classificadores.

Palavras-chave: Termografia. Energia elétrica. Inteligência Artificial. Machine Learning.

ABSTRACT

Electrical energy is among the most important resources needed in the modern life. Still, a significant part of the power generation is wasted, since some of the energy generated does not reach the final consumer. With waste of money and resources, besides the environmental impacts involved in energy generation, it is necessary a better management of electrical energy throughout the whole process. Considering the high rates of energy loss in power plants and transmission lines due to electrical problems and unforeseen conditions, one of the challenges in this management is how to reduce failures and optimize energy distribution. The management of resources in cities is one of the topics of study in smart cities and internet of things, concepts based on information and communication technologies, to promote efficiency in existing processes and to solve problems. The multicomponent sensing technology as the IoT environment is a solution for remote monitoring and power management, allowing real-time analysis of what is happening and identifying fault regions in the power grid. thermography is a technique for electrical inspection being more and more used, considering the use of technology and the feasibility of more efficient solutions. Thermal cameras, which can be considered as non-invasive sensors, can measure the heat emitted by electrical equipment and components. They allow to obtain thermal profiles and detect points of temperature variations, indicative of potential failures and anomalies in the network and in electrical equipment. This work proposes to perform detection of potential failures automatically, through the use of artificial intelligence algorithms for the identification and classification of electrical problems. An experiment was performed, generating a dataset of thermal images in a scenario of electrical problems. Through the proposed analysis, it was possible to validate the use of classifiers to identify electrical problems in a given scenario, as well as to understand the best combinations of algorithms of image processing and classification for each type of problem studied, and allow to obtain a method for data collection, thermal image processing and classifiers validation.

Keywords: Thermography. Electrical energy. Artificial Intelligence. Machine Learning.

LISTA DE ILUSTRAÇÕES

Figura 1: Faixas do espetro de luz 22
Figura 2: Procedimento geral das soluções em termografia e energia 37
Figura 3: Metodologia de detecção de problemas elétricos 56
Figura 4: Metodologia da etapa de processamento de imagem 57
Figura 5: Metodologia da etapa de extração de características 60
Figura 6: Metodologia da etapa de análise/classificação 63
Figura 7: Placa de circuito impresso desenvolvida 65
Figura 8: Tentativa de simulação com lâmpadas 66
Figura 9: Tentativa de simulação com fios elétricos. 68
Figura 10: Mesa de testes 69
Figura 11: Mesa de testes 70
Figura 12: Validação da imagem térmica a ser classificada (parte 1) 74
Figura 13: Validação da imagem térmica a ser classificada (parte 2) 75
Figura 14: Ruído provocado por outras fontes de calor Fone: Autor 78
Figura 15: Ruído provocado pelo mal-contato entre o borne e o fio 79
Figura 16: Exemplo de imagem térmica utilizada no processamento, pré-processada em grayscale 80
Figura 17: Imagem térmica do caso normal, com otsu (superior) e k-means (inferior) 81
Figura 18: Imagem térmica do caso com mal-contato, com otsu (superior) e k-means (inferior) 81
Figura 19: Teste dos modelos para os diferentes métodos de segmentação 87
Figura 20: Teste dos modelos para as diferentes características extraídas 88
Figura 21: Teste dos modelos para os diferentes métodos de segmentação 91
Figura 22: Teste dos modelos para as diferentes características extraídas 92
Figura 23: Teste dos modelos para as diferentes características extraídas 95
Figura 24: Teste dos modelos para os diferentes métodos de segmentação 96

LISTA DE TABELAS

Tabela 1: Trabalhos em termografia e energia elétrica (Continua) 38
Tabela 2: Trabalhos em termografia e algoritmos de IA (Continua)..... 47
Tabela 3: Trilhas da placa de circuito impresso 67
Tabela 4: Materiais e equipamentos a serem utilizados nos experimentos 68
Tabela 5: Problemas elétricos a serem simulados 71
Tabela 6: Simulações executadas em cada trilha..... 77
Tabela 7: Parâmetros estatísticos e térmicos extraídos da imagem processada 82
Tabela 8: Validação dos algoritmos de classificação 85
Tabela 9: Resultados dos algoritmos de classificação..... 86
Tabela 10: Validação dos algoritmos de classificação 89
Tabela 11: Resultados dos algoritmos de classificação..... 90
Tabela 12: Acurácia utilizando k-means e otsu, para cada simulação 91
Tabela 13: Validação dos algoritmos de classificação 93
Tabela 14: Resultados dos algoritmos de classificação..... 94
Tabela 15: Resultados dos algoritmos de classificação para cada simulação 97

LISTA DE ABREVIATURAS E SIGLAS

ANN	Artificial Neural Network
DTT	Dynamic Thermal Tomography
EM	expectation maximization
FCM	Fuzzy C-Means
FIR	Far-infrared
GLCM	Matriz de co-ocorrência em grayscale
LDA	Linear Discriminant Analysis
LWR	Long-wavelength Infrared
NIR	Near-infrared
NN	Nearest Neighbours
PCA	Principal Component Analysis
PCT	Principal Component Thermography
RBF	Radial Basis Function
RF	*Random Forest*
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
TIC	Tecnologia de Informação e Comunicação
UNCSTD	United Nations Commission on Science and Technology for Development
VANTs	Veículos aéreos não tripulados

SUMÁRIO

1 INTRODUÇÃO 11
1.1 OBJETIVO 15
1.2 JUSTIFICATIVA 15
1.3 ORGANIZAÇÃO DO TRABALHO 16
2 CONCEITOS FUNDAMENTAIS 18
2.1 PROBLEMAS ELÉTRICOS 18
2.1.1 ***Problemas de contato*** 19
2.1.2 ***Sobrecarga e subtensão*** 19
2.1.3 ***Curto-circuito*** 20
2.1.4 ***Instalações indevidas*** 20
2.2 TERMOGRAFIA 21
2.3 INTELIGÊNCIA ARTIFICIAL E MACHINE LEARNING 24
2.3.2.1 k-NN *27*
2.3.2.2 Redes Neurais *28*
2.3.2.3 SVM *29*
2.3.2.4 Árvores de decisão *31*
2.3.2.5 Naive Bayes *32*
3 REVISÃO BIBLIOGRÁFICA 34
3.1 CIDADES INTELIGENTES E ENERGIA 34
3.2 TERMOGRAFIA 36
3.2.1 ***Trabalhos correlatos em outras áreas*** 46
4 METODOLOGIA 55
4.1 METODOLOGIA DE PESQUISA 55
4.1.1 Processamento de imagem 57
4.1.2 Extração de características 59
4.1.3 Classificação e detecção de problemas 62
4.2 PLANEJAMENTO EXPERIMENTAL 64
4.2.1 Protocolo Experimental 66

4.2.1.1 Placa de circuito impresso 66

4.2.1.2 Materiais e equipamentos 68

4.2.1.3 Bancada **de testes** 69

4.2.2 Execução do experimento piloto 70

4.2.2.1 Simulação do problema de sobrecarga/sobrecorrente 72

*4.2.2.2 Simulação do problema de contato*s **elétricos** 73

4.2.3 Validação das imagens coletadas 74

5 RESULTADOS 76

5.1 CAPTURA DE IMAGENS 76

5.2 PROCESSAMENTO E EXTRAÇÃO DE CARACTERÍSTICAS 80

5.3 CLASSIFICAÇÃO E RESULTADOS 83

6 CONSIDERAÇÕES FINAIS 98

REFERÊNCIAS 101

1 INTRODUÇÃO

A energia elétrica é um importante recurso básico para a vida. Considerando sua fundamental relevância no mundo moderno, há a busca contínua pela otimização da gestão do recurso nas cidades e no campo, na demanda por maior eficiência no processo de geração, distribuição e armazenamento. Apesar disso, os números são preocupantes. No Brasil, em 2015, de toda a energia que foi gerada e transmitida pelo sistema de distribuição, 13,6% não chegou ao consumidor final, o que significa um desperdício de 500 milhões de MWh, equivalente a aproximadamente 11 bilhões de reais (ABRADEE, 2016; ANEEL, 2018). Esse problema também é de preocupação mundial, onde em 2014, pelo planeta inteiro, estima-se uma perda de 8% da energia gerada (WORLD BANK, 2014). Define-se perda elétrica como a diferença entre a energia gerada para ser transmitida no sistema e a energia que é efetivamente faturada, e pode ocorrer por fatores técnicos ou comerciais (ANTMANN, 2009; MARTINS, 2016). As perdas comerciais são caracterizadas por fatores externos ao sistema elétrico, como erros na medição, falta de pagamento por parte do consumidor e principalmente por roubos de equipamentos e ligações clandestinas. As perdas técnicas em geral acontecem naturalmente e são dadas pela dissipação de energia nas linhas de transmissão e distribuição, o denominado efeito Joule (ANTMANN, 2009). Apesar disso, há diversas condições que podem agravar essas perdas técnicas, como transformadores sobrecarregados, contatos deteriorados, corrosões e componentes com idade avançada, dentre outros. É possível mitigar o problema com a renovação das redes de alimentação, correta manutenção e substituição dos equipamentos e componentes, inspeção constante da rede elétrica e melhorias no processo de distribuição de energia (ABRADEE, 2016; KIESSLING et al., 2003; MARTINS, 2016).

Outro problema recorrente, além das perdas técnicas e comerciais, é o de interrupção do fornecimento de energia, também proveniente das mesmas causas. Problemas técnicos na rede, componentes e equipamentos antigos ou defeituosos, ligações clandestinas e adulteração da rede geram riscos de acidentes, incêndio e de queda de energia (FÁTIMA; RAIA, 2007; KIESSLING et al., 2003).

Em busca de minimizar as perdas elétricas e os problemas associados, é necessária a inspeção de toda a rede elétrica em períodos regulares. No entanto, inspecionar longas linhas de transmissão não é uma tarefa simples, além de custosa. A inspeção manual, a pé, apesar de ser a mais comum e barata, demanda tempo, esforço e muitas vezes é inviável, devido a regiões de difícil acesso, solo irregular e longas extensões, frequentemente o caso da rede elétrica no Brasil (ABRADEE, 2016; ANTMANN, 2009; BÜHRINGER et al., 2010; FÁTIMA; RAIA,

2007). Nestes casos, a inspeção se dá por helicópteros, que são missões bastante custosas e com algumas dificuldades envolvidas, já que é necessário extremo cuidado e voos bastante próximos da linha de transmissão. Com as novas tecnologias que vêm surgindo e se tornando acessíveis, passa-se a propor a ideia de utilizar VANTs (Veículos aéreos não tripulados), como drones, e técnicas de navegação para percorrer certas extensões de linhas de transmissão e até mesmo monitorar estações e subestações. Essa proposta elimina muitos dos problemas de inspeção, sendo uma solução menos custosa e que permite verificar longas dimensões territoriais. Apesar disso, ainda há a necessidade de inspeção das imagens capturadas por um especialista para a averiguação de problemas (NGUYEN; JENSSEN; ROVERSO, 2018).

Considerando o uso de tecnologias para inspeção, detecção de problemas e busca por maior eficiência no sistema, bem como algoritmos que tornem o processo mais inteligente e automatizado, os conceitos de Internet das Coisas (IoT) e cidades inteligentes surgem na perspectiva de lidar com esses aspectos.

Com a rápida urbanização ao longo do planeta nos últimos tempos, se aumenta a demanda por recursos como energia, água, gás e transporte, além da necessidade de um controle e gerenciamento de lixo e poluição. Uma das principais preocupações nos projetos de cidades inteligentes é justamente em como prover tais recursos de forma eficiente, reduzindo custos e evitando desperdícios, além de minimizar seus impactos, como apontado pelo relatório da UNCSTD (*United Nations Commission on Science and Technology for Development*) de 2016 (CUNHA et al., 2016; MOHANTY; CHOPPALI; KOUGIANOS, 2016; UNCTAD, 2016a). O relatório apresenta o conceito de *smart energy*, destacando a atenção a tópicos como medição, monitoramento, distribuição e otimização dos recursos de energia elétrica em cidades inteligentes. O sistema de *smart energy* consiste na integração dessas fontes de energia verde e sustentáveis, de forma descentralizada, onde toda a cidade pode contribuir com a geração de energia; na distribuição eficiente de energia, com a utilização de ICTs (tecnologia de informação e comunicação), como *smart grids* e *smart meters*; e na otimização do consumo de energia (MOHANTY; CHOPPALI; KOUGIANOS, 2016). A UNCSTD aponta em seu relatório que sistemas de gerenciamento de *smart energy* utilizam sensores, medidores avançados, controles digitais, fontes de energia renovável e ferramentas de análise para otimizar a distribuição e consumo; para tal, existem diversas inovações em infraestrutura para cidades inteligentes, como "*microgrids*, tecnologias para *smart grids*, geração de energia renovável distribuída, armazenamento de energia, e usinas virtuais" (UNCTAD, 2016a). *Smart grids* são apontadas como um elemento chave de infraestrutura para *smart energy,* definidas como um "sistema de distribuição de energia do ponto de geração ao ponto de consumo integrado com

uma ICT para aprimorar operações na rede elétrica, serviços ao cliente e benefícios ambientais" (UNCTAD, 2016b). Além de algumas cidades inteligentes que começam a implementar infraestrutura e serviços de *smart energy*, há países em desenvolvimento buscando soluções em *smart grids*. No Brasil começa-se a implantar *smart meters* para monitorar fraudes na rede elétrica. Na China, 27% dos medidores convencionais já foram substituídos pelos *smart meters* (MOHANTY; CHOPPALI; KOUGIANOS, 2016; UNCTAD, 2016a).

No cenário de cidades inteligentes e gestão de energia, se busca a utilização eficiente das VANTs e de sensores inteligentes, porém há algumas dificuldades envolvidas. A implantação de sensores na rede elétrica requer um grande investimento, um problema em grandes cidades ou longas linhas de transmissão. Se inspecionar redes elétricas já é uma missão complicada, instalar equipamentos e sensores muitas vezes é inviável, por questões de esforço e dinheiro. A solução de monitoramento por VANTs, embora atrativa, esbarra em um problema técnico; ainda é difícil realizar a detecção de falhas elétricas, já que através de câmeras convencionais é possível apenas verificar problemas decorrentes de características visíveis a olho nu, como corrosão de contatos e fios, e deterioração de equipamentos. A termografia é uma tecnologia não invasiva, que permite monitorar à distância e realizar diagnósticos, sem instalações complexas ou obstruções na rede elétrica.

A termografia por infravermelho (IR) é uma técnica bastante relevante no contexto de energia elétrica. Conforme Bagavathiappan et al. (2013), a temperatura é um dos indicadores mais comuns das condições de equipamentos e componentes; muitos dos defeitos e problemas causam variações anormais na distribuição de temperatura, devido ao calor gerado e as propriedades de transferência de calor (MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012). As câmeras térmicas, portanto, podem monitorar o comportamento térmico de equipamentos elétricos por meio da energia infravermelha emitida por eles. Apesar de não ser um indicador de todos os possíveis problemas elétricos, normalmente muitas falhas elétricas acontecem depois de sintomas de calor anormal produzido por um aumento da resistência elétrica esperada para o sistema (MALDAGUE, 2002; MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012). Conforme Maldague (2002), foi constatado que a maior parte dos superaquecimentos ocorrem devido a problemas de conexão do condutor e conexões com folga ou corroídas, enquanto outra parte se dá devido a defeitos encontrados em equipamentos elétricos. Em alguns casos, a sobrecarga também é uma das causas.

A termografia tem diversas aplicações relevantes, não só no campo da eletricidade, como na produção, manutenção preditiva, medicina, detecção de incêndio em florestas, programas de eficiência térmica em edifícios, monitoramento do trânsito em estradas,

agricultura e biologia, detecção de vazamentos de gás, vazamentos de água e testes não-destrutivos. Em comum a todos esses casos, constata-se que perfis de temperatura anormais podem indicar algum possível problema. Em geral, a maioria das aplicações utilizando termografia trabalham com a detecção de anomalias na análise de problemas ou falhas em potencial (MALDAGUE, 2002; MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012). A termografia é uma técnica bastante relevante no cenário das manutenções preventivas, permitindo realizar uma análise em grandes áreas. A sua vantagem vem da possibilidade de realizar análises sem implicar no desligamento do sistema ou de equipamentos em específico, além de ser um método relativamente seguro, já que não necessita do contato direto com os elementos inspecionados. Assim, a inspeção por infravermelho vem se consolidando como uma ferramenta de diagnóstico da condição de equipamentos elétricos. Porém, há diversas variáveis que impedem que as imagens termográficas de infravermelho indiquem com precisão se realmente há um problema. Essas imagens devem ser interpretadas por especialistas na área, que também deve ter conhecimento dos parâmetros envolvidos no sistema (BAGAVATHIAPPAN et al., 2013; MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012). Recentemente, começam a ser propostos sistemas automáticos de diagnóstico, que ajudam a predizer ou detectar os problemas elétricos com base nas características térmicas capturadas. Normalmente, as regiões de maior aquecimento nas imagens são fortes indicadores, que aliadas a algoritmos de processamento de imagens e inteligência artificial (IA), normalmente algoritmos de *machine learning*, podem dar informações a respeito das reais condições. A ideia geral no diagnóstico automático de imagens térmicas para aplicações elétricas consiste no processamento da imagem capturada para a extração das características relevantes e da região de interesse (ROI), que junto a outros parâmetros pertinentes a respeito do sistema inspecionado, servem de entrada para algoritmos de IA, normalmente voltados a predição e classificação (BAGAVATHIAPPAN et al., 2013; MALDAGUE, 2002; MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012; ŽARKOVIĆ; STOJKOVIĆ, 2017).

A utilização das imagens térmicas em conjunto com algoritmos para a inspeção de equipamentos e instalações elétricas, desde as usinas e linhas de transmissão, até o consumidor final em residências, edifícios e indústrias, mostra ser uma maneira promissora de identificar problemas e reduzir os elevados números de perda de energia e problemas ocasionados, identificando e classificando o possível problema elétrico. Assim, nesse cenário de cidades inteligentes, inspeções elétricas e monitoramento remoto para detecção de problemas, constata-se a importância de melhor entendimento dos métodos que vêm surgindo em termografia e o

conjunto de algoritmos de classificação pertinentes para tais aplicações, motivando o desenvolvimento deste projeto de pesquisa.

1.1 OBJETIVO

Este trabalho tem como objetivo propor um método para a classificação de problemas provenientes de equipamentos e componentes elétricos que resultam em variações de calor, por meio do estudo de algoritmos para a análise computacional de imagens térmicas.

1.2 JUSTIFICATIVA

É conhecido que a gestão e distribuição de energia elétrica é um grande problema da atualidade. Dado a necessidade de otimização dos recursos de energia elétrica, este trabalho busca estudar os cenários de distribuição de eletricidade, considerando o sensoriamento e algoritmos de inteligência artificial com foco na detecção de anomalias na rede. Há algumas soluções propostas para a gestão de energia elétrica e tecnologias que podem ser utilizadas, porém a maioria consiste em um sensoriamento invasivo, implicando também em uma infraestrutura para implantação e operação dos sensores, algo difícil, custoso e muitas vezes inviável em grandes cidades ou redes elétricas de longas extensões. As soluções de monitoramento remoto através de VANT's e outros veículos não tripulados é interessante, pois não envolve modificações na rede elétrica atual e resolve o problema de inspeções em longas extensões e em regiões de difícil acesso. É uma solução, entretanto pode ser limitada, já que a única fonte de dados são imagens obtidas por câmeras. Há pouco o que pode se concluir a partir delas, a não ser problemas físicos que sejam visíveis a olho nu, portanto a tomada de decisão se torna ineficaz.

Com câmeras térmicas, é possível obter o comportamento térmico dos elementos inspecionados, inclusive equipamentos elétricos. Ainda é necessário, porém, a análise de um especialista para realizar a avaliação do problema. Portanto, verifica-se a necessidade da utilização de técnicas de análise de dados para obtenção de conhecimento e tomada de decisão. A base de conhecimento produzida pode ser utilizada por algoritmos de inteligência artificial para a tomada de decisão. Isso também se aplica para a termografia; a partir das imagens coletadas, pode-se utilizar algoritmos para tratamento e análise de imagens, para identificação, classificação e predição de possíveis problemas gerados pelo estado diversificado da rede e anomalias.

Com a identificação e detecção de anomalias de forma autônoma por algoritmos de *machine learning*, através de um sistema não invasivo e que não requer modificações na infraestrutura elétrica atual das cidades e industrias, como aconteceria com outros tipos de sensoriamento, é possível identificar potenciais problemas elétricos, como estudado pelas pesquisas na área, e assim evitar maiores perdas, que apresentam números elevados ao redor do mundo e é um problema recorrente no Brasil.

1.3 ORGANIZAÇÃO DO TRABALHO

A metodologia utilizada na elaboração deste trabalho consiste em, a princípio, realizar o estudo do cenário de distribuição elétrica e as tendências para o futuro, entendendo as contribuições provenientes das novas tecnologias e projetos em cidades inteligentes. A partir disso, se deu a listagem dos trabalhos existentes e dos problemas encontrados até então, para sugerir uma solução por meio de termografia e algoritmos computacionais. A partir do estudo de trabalhos correlatos com base nessa mesma solução, se propõe um método e experimentos para comprovar a proposta.

Os conceitos fundamentais para entendimento da proposta do trabalho são apresentados na seção 2, descrevendo e trazendo a explicação teórica de energia e problemas elétricos, de termografia e dos principais algoritmos utilizados.

Na seção 3, a revisão bibliográfica é apresentada com o estudo do que há na literatura quanto o problema de energia elétrica, cidades inteligentes, termografia e algoritmos para detecção automática de falhas. Os trabalhos correlacionados, cada um com seus métodos, são apresentados com a solução proposta, algoritmos utilizados e resultados obtidos.

Na seção 4, considerando o estudo realizado, apresenta-se a proposta desse trabalho. Em 4.1 e 4.2 são detalhados a metodologia da pesquisa e a metodologia de experimento, com o método que será proposto e testes para obtenção de dados e validação. Na metodologia da pesquisa, foi apresentado o método para a detecção de problemas elétricos por meio da termografia e algoritmos, baseado nas referências bibliográficas. Esse método depende de diversas condições, onde diferentes técnicas e algoritmos podem ser aplicados. Em seguida, a metodologia de experimento é apresentada, com simulações de problemas elétricos e captura de imagens térmicas, que servem para compor a base de dados utilizada pelo método de detecção.

Na seção 5, os resultados da execução do experimento e aplicação da metodologia são apresentados, trazendo a simulação dos problemas elétricos apresentados no experimento

proposto na seção 4. Com esse teste, os resultados do método estudado são obtidos, indicando a viabilidade da aplicação da metodologia proposta e análise comparativa dos diferentes algoritmos selecionados.

Por fim, a seção 6 apresenta as considerações finais. Discute-se a contribuição conseguida pelo mapeamento das pesquisas e da proposta de um método geral para a detecção de problemas elétricos, bem como do experimento e validação da metodologia.

2 CONCEITOS FUNDAMENTAIS

Nas subseções a seguir os conceitos mais relevantes para entendimento do trabalho proposto são apresentados, considerando as áreas de energia elétrica e termografia, e os principais problemas elétricos, os conceitos para detecção de falhas elétricas por algoritmos de inteligência artificial e *machine learning*.

2.1 PROBLEMAS ELÉTRICOS

Considerando o cenário de perdas elétricas nos sistemas de energia por condições diversas, se faz necessário caracterizar os tipos de perdas e quais os problemas causadores, a fim de reduzir o desperdiço gerado. Esses problemas em sistemas elétricos podem ser estendidos por todo o sistema de energia nas cidades, desde a geração elétrica e distribuição, até o armazenamento e consumo, em veículos, residências e indústrias. Dentro desses problemas, os principais encontrados na literatura se dão por problemas de conexão (contatos com folga ou deteriorados), capacitores defeituosos ou sobrecarga, todos eles ocorrendo nos diferentes tipos de componentes elétricos do sistema (disjuntores, transformadores, linhas de transmissão, para-raios). Nos transformadores, problemas comuns se dão por conexões corroídas, sobrecarga elétrica ou baixo nível de fluído. Por outro lado, no caso de motores e geradores, além do problema de sobrecarga ou tensão abaixo do esperado, também há problemas de superaquecimento, mal isolamento ou bloqueio na passagem de resfriamento (MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012; SURIYAMONGKOL, 2002). Conforme apontado na pesquisa feita por Izykowski (2014), a taxa de falhas que ocorrem nos sistemas de geração e distribuição de energia são:

- Sobrecarga nas linhas de transmissão: 50%;
- Cabos: 10%;
- Disjuntores, fusíveis e aparelhagem de proteção: 15%;
- Transformadores: 14%;
- Equipamento de controle: 3%.

Nas subseções a seguir, os principais problemas elétricos encontrados na literatura são detalhados.

2.1.1 *Problemas de contato*

Os problemas de contato ou de conexão são caracterizados por mal encaixe e folgas, ou deteriorações nas ligações entre os condutores, por deformações, corrosão ou ferrugem. Eles representam falhas em sistemas elétricos e grandes perdas de energia nas linhas de transmissão (FOURNIER, 2001; MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012).

Emendas de cabo são considerados como um ponto crítico nesse contexto. Falhas muitas vezes ocorrerem devido ao superaquecimento do conector. O calor produzido nessas junções é baseado na maior resistência encontrada em relação ao restante do material condutor, que consiste na resistência do contato entre o fio e o material em que é encapado, do contato que ocorre entre eles e do contato entre os filamentos do fio (MOUSTAFA; ABD ELMAKSOUD, 2015). Conexões com folga causam uma mudança de resistência. Com isso, a eletricidade passa a ocupar uma pequena área da conexão defeituosa, que seria requerido para o fluxo de energia propriamente dito. Com isso, ocorrerá um aumento da resistência da conexão, e consequentemente, da temperatura. Qualquer mudança na resistência fará com que os equipamentos consumam mais energia do que a carga planejada.

2.1.2 *Sobrecarga e subtensão*

A sobrecarga é um problema que normalmente ocorre quando o sistema não é capaz de suprir a demanda necessária, em geral inesperada, criada por alguma influência externa. Com isso, se aumenta o consumo de corrente. Quando muita corrente passa pelos fios, eles se aquecem e podem derreter. Além disso, os equipamentos envolvidos também podem ser comprometidos. Disjuntores normalmente são utilizados para perceber correntes anormais e cortar o fluxo de energia, para evitar tal superaquecimento (The Electricity at Work Regulations 1989, 2015).

Outro problema que causa aumento na corrente é o de subtensão, onde é fornecido uma tensão abaixo do esperado. Transformadores e geradores devem operar a uma tensão abaixo da regulamentada na máxima carga possível, o que eleva a corrente elétrica e causa aumento de calor, além de submeter a um estresse incomum. O mesmo pode acontecer na indústria e em linhas de transmissão. O desbalanceamento de tensão pode causar também problemas em motores, devido ao superaquecimento causado. Máquinas e motores podem apresentar perda de desempenho e consumir mais corrente, superaquecendo e podendo sofrer danos (SEYMOUR; HORSLEY, 2008; SUDHA, 2009). O caso inverso, a sobretensão, também

caracteriza um problema elétrico, onde há um desbalanceamento de tensão, acima dos valores nominais estipulados, que normalmente resultam no superaquecimento de sistemas elétricos. Além disso, podem causar problemas e danificar no isolamento de máquinas e equipamentos elétricos. A sobretensão em equipamentos elétricos é caracterizada por um problema de desequilíbrio de carga, onde a sequência positiva da tensão é maior do que o valor nominal. Equipamentos que normalmente produzem certa quantidade de calor podem apresentar uma mudança anormal, produzindo mais calor, devido ao estresse causado pela sobretensão (SEYMOUR; HORSLEY, 2008; SUDHA, 2009).

2.1.3 *Curto-circuito*

O curto-circuito é um conhecido problema elétrico que pode ocorrer em situações em que há impedância zero ou baixa impedância. Quando a energia fornecida a partir de uma fonte de energia é desviada antes de chegar ao ponto em que a carga é destinada, há um curto-circuito (ESEN; ORAL; AKINCI, 2015).

Causam queda do fornecimento elétrico no sistema, normalmente representando grandes gastos no sistema de distribuição. Como resultado, podem também danificar outros equipamentos e componentes associados ao sistema, devido ao superaquecimento e forças mecânicas envolvidas. Em linhas de transmissão, o curto-circuito pode causar distúrbios no sistema de distribuição e até mesmo queda total no fornecimento. Outro efeito decorrente seria uma considerável redução de tensão em alimentadores, mesmo aqueles saudáveis, que estejam conectados a esse sistema com falha, causando correntes anormais puxadas por motores ou até mesmo resultando em sua parada, algo que representa um grande problema na área industrial. Por exemplo, tais problemas podem acontecer quando o isolamento entre dois condutores é comprometido ou uma quebra de um componente ocorre, onde sua resistência cai para zero, e assim a condutividade se torna alta (The Electricity at Work Regulations 1989, 2015).

2.1.4 *Instalações indevidas*

Muitos dos problemas elétricos anteriormente apresentados decorrem de instalações indevidas de equipamentos e componentes, onde há resistência elevada ou especificações de carga incorretas dadas condições de demanda não previstas. Alguns exemplos comuns são o uso não adequado de graxa inibitória, defeitos nas superfícies ou mal isolamento entre os condutores, e utilização inadequada de disjuntores, interruptores, fusíveis ou outros aparatos para proteção da rede elétrica, possivelmente com especificações incorretas. Fios de fusíveis

em tamanhos inadequados também são frequentes. Muitas perdas em cabos ocorrem quando condutores de espessuras abaixo do recomendado são utilizados. Isso causa envergamento dos condutores e aumento de temperatura. Desbalanceamento onde fios de neutro de tamanho abaixo do recomendado são utilizados, sendo menores que fios de fase também causam problemas de superaquecimento. Outras falhas comuns acontecem quando transformadores, após sua manutenção, possuem seu núcleo não propriamente apertado. Transformadores operando em carga subótima, devido a previsão incorreta de demanda, é outro problema de instalação imprópria (MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012; SURIYAMONGKOL, 2002).

É interessante destacar que dentre os problemas elétricos apontados, muitos deles apresentam como efeito variação da produção de calor. As propriedades elétricas dos componentes e equipamentos envolvidos tem direta relação com a temperatura. Considerando esse cenário, sugere-se o estudo dos principais problemas neste meio, dado as possíveis variações térmicas e perfis de temperatura, que podem ser percebidos por meio da termografia.

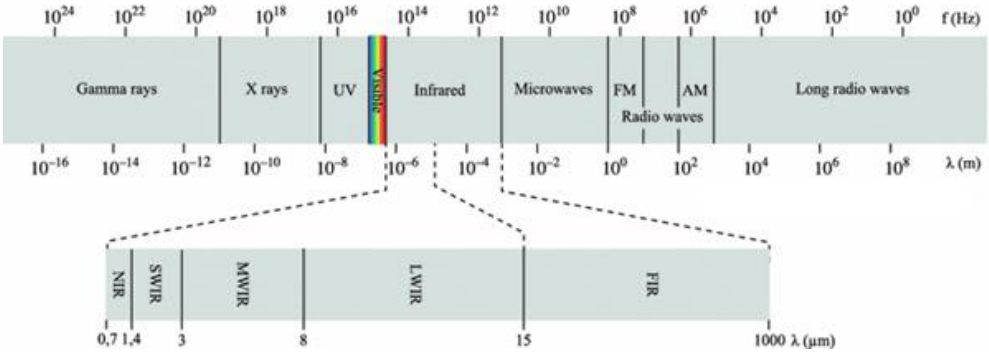
2.2 TERMOGRAFIA

A termografia é uma técnica que consiste na captura de imagens por meio da medição da emissão de radiação infravermelho pelos objetos, muitas vezes referida como radiação térmica. Essa radiação tem um espectro de onda bastante amplo entre 0,7 e 1000 μm. Dispositivos de visão noturna trabalham apenas na faixa de 0.7–1.4 μm, conhecida como *near-infrared*, uma faixa de espectro mais baixa onde a cena é iluminada com radiação infravermelha que será capturada. Essa é uma abordagem de sensor ativo. Por outro lado, em muitos casos se prefere uma abordagem passiva, onde a própria radiação emitida pelos objetos é capturada pelo dispositivo (ROGALSKI, 2003).

O problema no caso da abordagem de sensor passivo é que a radiação infravermelha que os objetos emitem estão em espectros de onda mais altos que requerem tecnologias mais avançadas de captura. Uma propriedade relevante nesse caso é que o comprimento de onda e intensidade da radiação emitida por qualquer objeto depende da sua temperatura. Isso foi percebido anteriormente, em 1800, quando a radiação infravermelha foi descoberta por William Hershel, usando prismas refratando a luz solar e medindo com termômetros. (GADE; MOESLUND, 2014; MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012). Foi possível perceber que posicionando os termômetros na parte vermelha da luz visível refratada, foram obtidas as maiores medições de temperatura. A partir disso, se percebeu que há uma relação

direta entre o calor e a radiação infravermelha emitida. É possível observar que conforme a temperatura da superfície de um objeto arbitrário aumenta, a intensidade da radiação também aumenta e o espectro de onda diminui. Isso quer dizer que para valores muito altos de temperatura, a radiação atinge o espectro visível, situação na qual é possível enxergar perceber; por exemplo, o ferro submetido a altas temperaturas, apresentando cor avermelhada. Em geral, para temperaturas mais baixas, é necessário algum dispositivo específico para capturar em um espectro de onda maior, como o *mid-infrared* (MWR) e *long-infrared* (LWR). Na imagem a seguir (Figura 1), as faixas dos espectros de onda são apresentadas. As ondas no espectro visível vão apenas até 0,7 µm o que explica como a radiação na cor vermelha desse espectro apresenta maior calor, sendo o mais próximo possível do espectro infravermelho (GADE; MOESLUND, 2014; MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012).

Figura 1: Faixas do espetro de luz



Fonte: Gade, 2014.

Ondas no espectro infravermelho começam a partir de 0,7 µm, sendo atribuídas a diferentes classificações. A classificação *near-infrared* (NIR) e *shortwave-infrared* vão de 0,7 µm a 3 µm, faixa de interesse em aplicações como na vigilância noturna e na agricultura. Câmeras nessas duas faixas, porém, ainda sofrem influência da luz visível refletida, não sendo tão úteis para a termografia, caso em que se espera capturar apenas a radiação de calor. Câmeras térmicas trabalham normalmente com as faixas de infravermelho mais altas. É importante destacar que essas faixas não cobrem o infravermelho inteiro, pois alguns trechos desse espectro não são para a termografia. As faixas, LWR e FIR, como destacado na figura acima, estão entre 2–5 µm e 8–14 µm, respectivamente. Elas escapam da influência da atmosfera, que bloqueia a transmissão das ondas em algumas faixas. Além disso, câmeras térmicas, por trabalharem com essas duas faixas de comprimentos de onda mais altos, também capturam apenas radiação

infravermelho, diretamente relacionado ao calor, evitando influência da luz visível (MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012),

O princípio físico por trás dessas faixas e a radiação infravermelha emitida, e que é capturada pelas câmeras térmicas, consiste na propriedade de que todo o objeto emite radiação, energia transportada em forma de ondas eletromagnéticas. A quantidade de energia emitida a partir da superfície é diretamente proporcional à sua emissividade. Todo material possui um coeficiente de emissividade que diz o quanto de radiação térmica é emitida de acordo com sua temperatura. Todos os objetos acima do zero absoluto (0 Kelvin ou -273.15ºC) emitem radiação, portanto podem ter sua temperatura medida por meio da termografia. Dispositivos de captura de imagens térmicas tem como princípio a composição de uma imagem de padrões térmicos, calibrada para medir a força de emissividade de diferentes superfícies em várias temperaturas, por meio de lentes específicas para focar a radiação emitida em um detector infravermelho, que converte a resposta elétrica obtida em uma informação visual, uma imagem digital onde as diferentes cores na escala correspondem aos níveis de temperatura. Dessa forma, quando se conhece a emissividade da superfície, dado a faixa de espectro do detector infravermelho, é possível estimar sua temperatura (BALARAS; ARGIRIOU, 2002; ROGALSKI, 2003).

A termografia é conhecida por ser um método de avaliação e teste não-destrutivo, que permite a inspeção e análise por meio de um detector infravermelho sem contato direto com a superfície de interesse (MALDAGUE, 2002). A princípio utilizado na área militar para monitoração, a sua capacidade de medição de temperatura à distância, de forma não-invasiva, permite que a tecnologia seja utilizada em diversos outros domínios. Na medicina se há uma ampla área de aplicação, já que a temperatura muitas vezes é associada a saúde e condições físicas; tumores podem ser detectados, é possível realizar o diagnóstico em tempo real durante certas cirurgias, além de informações a respeito de lesões nervosas. Se tratando de agricultura e meio ambiente, há aplicações relacionadas ao estudo e desenvolvimento de soluções para proteger plantas e árvores sensíveis ao frio e gelo, bem como monitoramento das condições de poluição do mar e de rios (MEOLA; CARLOMAGNO, 2004). Ela também pode ter bastante uso na manutenção de equipamentos mecânicos e elétricos, permitindo o monitoramento das suas condições, que muitas vezes pode ser percebida pela variação de temperatura e padrões de distribuição de calor incomuns. A análise dessas imagens pode permitir que muitas companhias prevejam problemas futuros e se antecipem a eles, tomando medidas corretivas antes que uma falha mais grave ocorra.

Existe um grande potencial na utilização de câmeras térmicas na área de energia, dado as vantagens da inspeção não-invasiva propiciada pelas câmeras térmicas e as propriedades dos sistemas elétricos, diretamente relacionados com a produção de calor. Conhecendo a emissividade ε dos equipamentos e componentes elétricos, uma razão normalmente em torno de 0.85 e 0.95, é possível medir a sua temperatura e obter perfis térmicos a fim de detectar falhas ou perceber possíveis problemas (TAVARES; COSTA; MAIA, 2006).

2.3 INTELIGÊNCIA ARTIFICIAL E MACHINE LEARNING

A Inteligência Artificial (IA) vem sendo cada vez mais utilizada na automatização de funções que envolvem o raciocínio humano e na tomada de decisão. Ela está relacionada com o uso da computação em tarefas associadas a conceitos como "*conhecimento, percepção, raciocínio, aprendizagem, compreensão e habilidades cognitivas similares*" (SALOKY; ŠEMINSKÝ, 1999). Conforme indicado por Russel, a IA sempre visou reproduzir as faculdades humanas como "*criatividade, autoaperfeiçoamento e uso da linguagem*"(RUSSELL; NORVIG, 2013). Considerando isso, a IA consiste basicamente no estudo das habilidades mentais, oferecendo ferramentas para resolver problemas complexos que normalmente requerem inteligência. Dentre essas ferramentas, há diversas técnicas poderosas como sistemas especialistas, sistemas de lógica *fuzzy*, redes neurais artificias (ANN), *neuro-fuzzy*, algoritmos genéticos e SVMs (*Support Vector Machine*) (SIDDIQUE; YADAVA; SINGH, 2003).

A IA carrega consigo uma diversidade de técnicas e algoritmos, como sistemas especialistas, visão computacional, robótica, programação automática, suporte a planejamento e decisão (SALOKY; ŠEMINSKÝ, 1999). O foco neste trabalho é no aspecto de aprendizado, ou seja, as ferramentas e algoritmos disponíveis em *machine learning*. Este aprendizado se refere à capacidade de um sistema de melhorar seu desempenho ao longo do tempo, adquirindo novos conhecimentos, reorganizando o que previamente já conhece, entendendo seus significados e suas relações. As técnicas de *machine learning* são bastante úteis quando queremos obter informações quanto a problemas elétricos e imagens térmicas, devida sua capacidade de aprendizado para identificar padrões e correlacioná-los com uma possível falha, dado que se há exemplos de treinamento, de casos conhecidos.

As subseções a seguir trazem os algoritmos e técnicas de inteligência artificial e *machine learning* mais utilizados nos trabalhos em termografia para detecção de anomalias, por meio de predição e classificação.

2.3.1 Lógica Fuzzy e Neuro-fuzzy

A lógica *fuzzy* traz o que se conhece da teoria de conjuntos, porém incluindo aspectos da incerteza humana, tornando processo de decisão mais natural. Se na teoria de conjuntos clássica um elemento pode pertencer a um conjunto ou não, em conjuntos *fuzzy* pode-se dizer que um elemento pertence a ele com algum grau de pertinência. A ideia geral de um sistema de lógica *fuzzy* é a utilização de variáveis linguísticas na definição de regras, compondo uma base de conhecimento, que será utilizada para inferir algo (DERNONCOURT, 2013).

Assim, sistemas de inferência *fuzzy* são baseados em uma série de regras, que podem ser consideradas como afirmações qualitativas. A inferência se dá por um processo de raciocínio, por meio composição das variáveis e regras, para chegar a alguma conclusão das informações qualitativas na base de conhecimento (SIDDIQUE; YADAVA; SINGH, 2003).

Essas regras são normalmente descritas em linguagem natural. Por exemplo: *"Se {temperatura do equipamento é alta} e {carga regulamentada é baixa}, então {prioridade de manutenção é alta}"*. A variável "prioridade de manutenção" pertence ao conjunto *fuzzy* "alta" com grau de pertinência que depende da sua premissa, "temperatura do equipamento" e "carga regulamentada", e os graus de pertinências envolvidos, bem como o operador *fuzzy* (AND, OR ou NOT), neste caso AND, dado que temos "e" em linguagem natural. Para determinar o grau de uma preposição *fuzzy* como essa, é necessário obter o grau das preposições por meio da função de pertinência, aplicação da implicação/inferência *fuzzy* em cada uma das regras, e por fim agregação dos resultados. Esse resultado agregado gera um conjunto *fuzzy* resultante, que através de operações de *defuzzificação,* poderá ser utilizado para tomada de decisão (DERNONCOURT, 2013).

Sistemas de inferência *fuzzy* tem como grande vantagem permitir a transparência e interpretação dos dados. Os dados são associados com variáveis linguísticas, por exemplo, {baixo, normal, alto}, e pode-se entender exatamente quais variáveis e quais regras contribuíram decisivamente para o resultado final. Sistemas *fuzzy* tipicamente utilizam regras *se-então* e mecanismos de inferência, que idealmente correspondem ao conhecimento de um especialista e o processo de tomada de decisão humano para um dado problema (PÉREZ-ORTIZ et al., 2016).

Em resumo, o sistema de inferência baseado em lógica *fuzzy* recebe valores de entrada, *fuzzy* ou não, e faz sua *fuzzyficação*, processo onde eles são associados às suas variáveis linguísticas. As variáveis pertencem a conjuntos *fuzzy* com algum grau de pertinência. Em seguida, há um mecanismo de inferência, que por meio das variáveis linguísticas, operadores

fuzzy e regras na base de conhecimento, faz a inferência gerando um conjunto agregado. Por fim, o *defuzificador* permite transformar o conjunto *fuzzy* resultante em uma única decisão de saída (DERNONCOURT, 2013).

Esse tipo de algoritmo pode ser útil na classificação de problemas de energia e inspeção elétrica, representando o processo de raciocínio e tomada de decisão de um especialista, trabalhando com a incerteza e informações vagas. Dado as medições de temperatura a partir da câmera térmica e as especificações técnicas, além de outras medidas elétricas, regras podem ser formuladas, indicando a gravidade de um problema ou prioridade de manutenção. A grande limitação desse caso se dá pelas regras propriamente ditas. É necessário possuir um bom conhecimento do domínio, para modelar a base de conhecimento utilizada na inferência. Em muitos casos, não se há a informação ou conhecimento necessário para relacionar grande parte dos parâmetros envolvidos na aplicação com algum resultado ou possíveis decisões. Por exemplo, é possível perceber variações de temperatura em um equipamento específico, mas não se conhece quais variações são aceitáveis ou não, bem como qual a distribuição de temperatura adequada ou pontos que podem aquecer mais. A ideia, nesse caso, é treinar um modelo que seja capaz de generalizar os casos conhecidos e aprender padrões, permitindo classificar novos casos. Para isso, se faz necessário ter casos de falha já conhecidos e inspecionados por termografia, indicar todas as medidas do equipamento obtidas e a qual classe ela pertence (o tipo de falha encontrado). A partir disso, realiza-se o treinamento e obtenção de um modelo, sem a necessidade de estabelecer regras específicas, como é o caso de sistemas *fuzzy*.

A seguir são detalhadas técnicas de classificação com aprendizagem, o que reduz o conhecimento especialista necessário, correlacionando as variáveis de entrada e reconhecendo padrões.

2.3.2 Classificadores de machine learning

Os problemas de classificação normalmente envolvem os conceitos de aprendizado supervisionado e não-supervisionado. O aprendizado não-supervisionado busca analisar as informações disponíveis e descobrir padrões em dados não etiquetados, analisar se há grupos que podem ser diferenciados ou *clusters* presentes nos dados. No caso do aprendizado supervisionado, os dados de entrada já possuem uma classe estabelecida. O objetivo neste caso é aprender um modelo com base nesses dados, de forma com que possa reproduzir a mesma classificação para os dados providos previamente, e além disso generalize essa representação para produzir saídas corretas para novas entradas ainda não apresentadas, tarefa conhecida

como predição (LORENA; DE CARVALHO, 2007; PÉREZ-ORTIZ et al., 2016). O procedimento geral de treinamento envolve variáveis de entrada e uma variável de saída esperada, par que representa exemplos de treinamento. A partir de uma base com um conjunto desses exemplos, busca-se aprender uma função que mapeie tais variáveis com a classe de saída, caracterizando sua capacidade preditiva (NG, 2000).

Os algoritmos de classificação supervisionados buscam separar classes do problema, na maior margem de acerto possível, usando apenas o conjunto de dados de treinamento. Uma característica importante que eles possuem é a de lidarem com ruídos e exemplos incorretos, graças sua capacidade de generalização. Esse trabalho utilizará algoritmos de aprendizado supervisionado, conforme descrito nas subsecções a seguir, dado que a classe das imagens é conhecida. Se espera a utilização de um conjunto de treinamento com imagens térmicas de situações conhecidas de problemas elétricos (PÉREZ-ORTIZ et al., 2016). A seguir, serão apresentados os principais algoritmos de classificação encontrados na literatura e que são utilizados para a classificação de imagens térmicas.

2.3.2.1 k-NN

Os algoritmos de vizinhos mais próximos (*nearest-neighbours*, ou NN) são considerados os métodos mais simples de aprendizado e classificação. A ideia básica deles é memorizar o conjunto de treinamento e então predizer a classe de qualquer nova instancia com base nas classes dos seus vizinhos mais próximos do conjunto de treinamento (BEN-DAVID; SHALEV-SHWARTZ, 2014). O princípio geral está em se assumir que os parâmetros que descrevem os dados/pontos do domínio são relevantes à sua classe, e, portanto, a classificação deve acontecer de forma com que pontos próximos aos outros provavelmente possuem a mesma classe. É um método baseado em similaridade de parâmetros/características. O algoritmo NN atribui a classe do vizinho mais próximo para cada novo padrão apresentado durante a fase de testes. Pode-se avaliar utilizando mais de um vizinho próximo, onde a classe é decidida pela classe que aparece em maioria nos vizinhos mais próximos, ou baseado no balanceamento de pesos por distância; essa técnica é conhecida como *k-NN*, onde são utilizados na classificação os *k* vizinhos mais próximos (PÉREZ-ORTIZ et al., 2016). O *k-NN* não gera um modelo a partir de dados de treinamento, como ocorre em muitos dos classificadores por aprendizado supervisionado. A sua função de classificação é obtida apenas localmente. Ele requer, portanto, que o conjunto de dados seja armazenado em sua totalidade ou pelo menos a maioria dele, e que os dados estejam disponíveis na etapa de teste e classificação de novos exemplos, algo

computacionalmente custoso. Assim, o processo de predição pode ser lento, principalmente para grandes conjuntos de dados. Um outro problema do *k-NN* é sua sensibilidade a ruídos, já que é um método baseado na similaridade de parâmetros. A sua vantagem, no entanto, está na simplicidade de implementação e facilidade em testar um conjunto de dados, sem muita parametrização e treinamento (BISHOP, 2013). Muitos trabalhos em termografia utilizam k-NN obtendo altas taxas de acerto, apresentando-se como uma possibilidade para esse trabalho, em energia elétrica.

2.3.2.2 Redes Neurais

As redes neurais artificiais (ANNs), dentre diversos propósitos, são utilizadas para classificação por meio do aprendizado supervisionado. As redes neurais podem capturar o conhecimento do domínio a partir de exemplos de treinamento. A técnica surgiu com a ideia de criar um modelo computacional do cérebro. Nela, o conhecimento implícito é obtido por meio do treinamento utilizando dados de entrada e a classificação esperada. O erro entre a saída esperada e a efetivamente obtida pode ser utilizado para alterar o peso das conexões entre os neurônios e criar um modelo que generalize os diversos casos ensinados (SIDDIQUE; YADAVA; SINGH, 2003).

Em resumo, conforme descrito por Russel e Norvig (2013), as redes neurais consistem em unidades de processamento, conectadas por ligações direcionadas, que propagam a ativação de uma unidade (neurônio) a outra. Essas ligações possuem também um peso numérico associada a elas, que determina a força e o sinal da conexão. Cada uma das unidades calcula a soma ponderada de suas entradas, aplicando uma função de ativação para obter uma saída. Essa função pode ser baseada em um limiar de ativação, caracterizando um típico neurônio *perceptron*, ou pode ser uma função logística, como a sigmoide. As funções de ativação não lineares permitem trabalhar com problemas que não são separáveis linearmente, propriedade importante nas redes neurais. Essa rede tem neurônios conectados de alguma forma, rede com alimentação para a frente, onde os neurônios de uma camada têm conexão somente com neurônios da próxima camada, formando um gráfico acíclico dirigido. Uma rede desse tipo representa uma função de sua entrada atual. Essas redes normalmente estão dispostas em camadas, podendo conter múltiplas camadas ocultas. É comum que em problemas que envolvam mais de duas classes haja uma unidade de saída para cada classe. Com uma camada oculta e uma camada de saída, é possível representar funções contínuas não lineares (RUSSELL; NORVIG, 2013).

As redes neurais *multilayer perceptrons* são redes *feedforward*, com alimentação para frente, onde cada neurônio recebe como entrada uma soma ponderada das saídas dos neurônios conectados a ele pelas arestas de entrada. É uma classe de rede neural bastante utilizada para a classificação e reconhecimento de padrões, que geralmente consiste na utilização de múltiplas camadas de neurônios, representando, matematicamente, de acordo com Bishop (2013), modelos de regressão com não-linearidade continua. Ou seja, do ponto de vista estatístico, as ANNs podem ser vistas como modelos de regressão linear generalizados, com uma combinação linear, porém de projeções não-lineares dos dados de entrada (BEN-DAVID; SHALEV-SHWARTZ, 2014; BISHOP, 2013).

O processo de aprendizagem supervisionada da rede neural se dá normalmente por ajustar os pesos das conexões, de forma interativa. Uma das principais técnicas utilizadas para tal ajuste é o *back propagation*, que faz essa atualização por meio da otimização por gradiente descendente (BEN-DAVID; SHALEV-SHWARTZ, 2014). O tamanho da rede é uma escolha importante. Uma rede muito grande pode ser capaz de memorizar todos os exemplos, mas não generalizar bem novos parâmetros de entrada. As redes neurais também estão sujeitas a superadaptações quando existem muitos parâmetros no modelo. Assim como muitos parâmetros podem ajustar todos os dados, eles podem ser incapazes de generalizar, como acontece em modelos de poucos parâmetros. Modelar uma rede exige tomar uma série de decisões, como o número de camadas ocultas e seus tamanhos, bem como a taxa de aprendizado dos neurônios. Sugere-se a utilização de técnicas de validação cruzada para poder escolher a melhor arquitetura que oferece a boa previsão para os conjuntos de validação (RUSSELL; NORVIG, 2013).

2.3.2.3 SVM

As SVMs também podem ser utilizadas para problemas de classificação. Elas permitem criar funções a partir de um conjunto de dados de treinamento devidamente classificados. Para o caso de funções de classificação, as SVMs obtêm um hiperplano no espaço de possíveis entradas, no domínio do problema, buscando separar os exemplos positivos dos negativos, o que caracteriza um classificador binário (SIDDIQUE; YADAVA; SINGH, 2003). Em resumo, a ideia básica da SVM consiste na separação de duas classes por meio de um hiperplano especificado por um vetor normal w e um termo de viés b. O hiperplano de separação ótimo é aquele que maximiza a distância entre este plano e os pontos mais próximos de cada classe (a margem entre eles). A ideia de se obter o hiperplano que melhor separa os dois conjuntos se baseia em evitar a suscetibilidade de cometer erros quando novos dados são classificados, já

que essas margens largas evitam o superajustamento do modelo aos dados de treinamento (LORENA; DE CARVALHO, 2007).

Junto com as SVMs, são utilizadas funções de kernel, em conjunto com uma função não-linear de mapeamento, para permitir limites de decisão não-lineares. Esta função kernel mapeia dados de entrada do espaço original para um espaço de características mais elevado, onde os dados podem ser separados linearmente. Uma das principais funções utilizadas é a gaussiana. Neste caso, as SVMs podem ser consideradas como classificadores tão bons quanto as ANNs, dado que a maior parte dos dados no mundo real são não-linearmente separáveis (LORENA; DE CARVALHO, 2007; PÉREZ-ORTIZ et al., 2016).

As SVMs definem funções centradas nos pontos dos dados de treinamento e selecionando um subconjunto deles durante o treinamento. Apesar de envolver otimização não-linear no treinamento, a função objetivo é convexa, e assim a solução do problema de otimização é relativamente direta. O número de funções no modelo resultante é geralmente muito menor que no número de pontos de dados, representando um modelo computacionalmente mais eficiente, principalmente para muitos dados, diferentemente do *k-NN* (BISHOP, 2013).

Além da aplicação de técnicas de kernel, outra generalização proposta é a de utilização de margens rígidas e margens suaves. As margens rígidas definem fronteiras lineares a partir de dados linearmente separáveis por um hiperplano. Para lidar com conjuntos de treinamento mais gerais, permite-se que alguns dados possam violar as fronteiras definidas pelo hiperplano, por meio das margens suaves. Elas permitem a presença de ruídos e *outliers* (dados fora da curva), por meio da utilização de variáveis de folga, que relaxam as restrições impostas na otimização (LORENA; DE CARVALHO, 2007; PÉREZ-ORTIZ et al., 2016).

As SVMs são robustas lidando com dados de grande dimensionalidade, característica que comumente resulta em classificadores super ou subajustados. Uma das principais vantagens das SVMs em relação as redes é que há um único mínimo global devido a convexidade do problema de otimização, enquanto as ANNs possuem mínimos locais na função objetivo minimizada (LORENA; DE CARVALHO, 2007). Além disso, a alta dimensionalidade, que pode gerar uma distribuição esparsa no domínio de entrada, pode ser uma dificuldade para as redes neurais generalizarem, produzindo modelos superajustados em consequência dos algoritmos de otimização usados para seleção de parâmetros e medições estatísticas para selecionar o melhor modelo (GUNN, 1998). Outra propriedade observada é que as SVMs originalmente lidam com a classificação binária, mas podem ser reformuladas para tratar

problemas de múltiplas classes, dividindo os dados em cada uma delas (métodos um-contra-um ou um-contra-todos) (PÉREZ-ORTIZ et al., 2016).

Não é uma tarefa simples decidir qual dos métodos, melhor se adequa para realizar a classificação. Ambos os classificadores possuem boa capacidade de generalização e trabalham com dados linearmente e não-linearmente separáveis (sendo a SVM com kernel gaussiano). SVM funciona melhor que *k-NN* para alta dimensionalidade de dados e quando há ruídos e *outliers*. ANN também possui bons resultados, apresentando em geral uma alta acurácia preditiva, mas apresentando a complexidade de refinamento de parâmetros de decisão como números de neurônios, camadas e taxa de aprendizado (BEN-DAVID; SHALEV-SHWARTZ, 2014; BISHOP, 2013).

Quando tratamos de problemas de energia elétrica e termografia, por meio de câmeras térmicas podemos obter medidas de equipamentos e componentes, indicativos de seu estado. Porém, o trabalho de analisar esses dados e entender o que eles representam é um trabalho que envolve a análise de especialistas na área. Conhecendo casos em condições normais e em falha de sistemas elétricos e suas medições termográficas, é possível treinar algoritmos de classificação e realizar a análise de problemas de forma automática, como é o objetivo desse trabalho.

2.3.2.4 Árvores de decisão

As árvores de decisão representam uma categoria de classificadores bastante utilizada devido a sua "interpretabilidade", de forma com que as árvores possam ser expressadas de forma equivalente a um conjunto de regras, sendo bastante intuitivo (PÉREZ-ORTIZ et al., 2016). Além disso, ainda possuem boa acurácia em diversos casos reais. O treinamento ocorre normalmente por algoritmos de aprendizado indutivo, construindo a árvore de forma recursiva. Alguns dos algoritmos mais comuns são o CART, ID3 e C4.5. No caso de trabalhos de energia, elas são interessantes para gerar um modelo compreensível e entender quais são os parâmetros mais decisivos na classificação, permitindo também extrair regras para detecção de falhas.

As árvores de decisão predizem a classe associada a uma instância percorrendo do nó raiz de uma árvore até alguma das suas folhas. A cada nó, o filho sucessor é escolhido com base na separação do espaço de entrada. Essa separação muitas vezes é baseada em um parâmetro em específico do conjunto de entrada, mas também pode ser dado por um conjunto de regras predefinidas de divisão. Uma folha dessa árvore representa uma classe em específico (BEN-DAVID; SHALEV-SHWARTZ, 2014).

No processo de aprendizado, a construção dessas árvores de decisão consiste em abordagens gulosas, muitas vezes recursivas, onde a árvore é construída gradualmente e decisões ótimas são tomadas durante a construção de cada nó. É um método de construção que funciona bem na prática, apesar de não garantir uma árvore de decisão ótima globalmente (BEN-DAVID; SHALEV-SHWARTZ, 2014).

A combinação de múltiplas árvores é uma solução também adotada, podendo aumentar o desempenho obtido em relação a uma única árvore de decisão. Um método comum são as *Random Forests* (RFs), onde várias árvores de decisão são construídas usando amostras aleatórias do conjunto de entrada. As diversas árvores compõem uma floresta, onde a decisão é dada pela agregação das decisões de todos as árvores, por exemplo por votação (PÉREZ-ORTIZ et al., 2016).

Trabalhos em imagens térmicas utilizam árvores de decisão com o propósito de melhor interpretar os resultados e entender quais parâmetros são mais relevantes para a classificação.

2.3.2.5 Naive Bayes

Em classificadores do bayesianos, o treinamento é baseado em construir um modelo probabilístico das variáveis de entrada, que são usadas durante a fase de testes para predizer a classificação de um novo exemplo. Um modelo simples de classificador bayesiano é o *Naive Bayes*, que carrega o "ingênuo" no nome pois assume que todos os atributos de entrada são independentes, dado a sua classe (BEN-DAVID; SHALEV-SHWARTZ, 2014). Com isso, apesar de possuir grande simplicidade, permitem modelar facilmente o problema e são computacionalmente muito eficientes. É um método que elimina grande parte da complexidade de classificadores bayesianos e apresenta, de acordo com (PÉREZ-ORTIZ et al., 2016), boa acurácia de classificação. Esse método, porém, não considera as relações entre as variáveis de entrada encontradas, muitas vezes complexas, e por isso é utilizada em casos específicos. Em problemas onde essa correlação é determinante para a classificação, o método pode ser pouco eficiente.

Naive Bayes é um classificador bastante útil quando a dimensionalidade dos dados é alta, onde a estimação de densidade do espaço é mais complicada. Além disso, é um método insensitivo ao tamanho da base de dados, sendo bastante utilizados em pequenos conjuntos. Apesar de violar os compromissos entre as variáveis, o modelo ainda possui bom desempenho pois nem sempre as correlações são satisfeitas no mundo real (BISHOP, 2013).

As imagens térmicas podem trazer informações relevantes quando a diferentes tipos de condições elétricas. Conhecendo as características térmicas mais descritivas dessas imagens quanto ao comportamento elétrico de equipamentos e componentes, bem como a que classe essas imagens pertencem, pode-se obter modelos que identifiquem padrões e classifiquem essas imagens, de acordo com o que já se conhece. O objetivo é obter um modelo que classifique novos exemplos ainda não vistos, visando a detecção automática de problemas elétricos. A metodologia é apresentada na próxima seção, a proposta deste trabalho.

3 REVISÃO BIBLIOGRÁFICA

Visando entender melhor o cenário de energia elétrica, para propor uma metodologia geral de detecção automática de problemas elétricos, trabalhos relacionados à energia elétrica ou termografia relevantes para tal objetivo são investigados na literatura. Algoritmos e métodos em cidades inteligentes podem ser interessantes, já que também consistem em técnicas inteligentes, a partir de parâmetros medidos da rede elétrica, para detecção de problemas. Além disso, também é feito o estudo dos métodos e técnicas utilizados em termografia e seus principais algoritmos para a detecção de falhas elétricas.

3.1 CIDADES INTELIGENTES E ENERGIA

Ainda não há um conceito único do que é uma cidade inteligente. Existem diversas iniciativas que são intituladas como cidades inteligentes, que em geral tratam de cidades que fazem o uso de informação e tecnologias digitais para melhorar seus serviços, torna-los mais flexíveis, eficientes e sustentáveis, provendo também melhor qualidade de vida aos habitantes (MOHANTY; CHOPPALI; KOUGIANOS, 2016). Uma definição cabível dada por Mohanty et al. (206), diz que *"...uma cidade inteligente sustentável é uma cidade inovadora que usa as tecnologias de informação e comunicação (TICs) e outros meios para melhorar a qualidade de vida, eficiência das operações e serviços urbanos, e competitividade, garantindo também que cumpra com as necessidades da geração presente e futura, em respeito aos aspectos econômicos, sociais e ambientas*".

As definições e pesquisas em cidades inteligentes em geral apontam a importância da otimização dos recursos da cidade, como transporte, água, educação, saúde, energia, subsistência e segurança, demonstrando que a utilização da infraestrutura da cidade de forma "inteligente" para uso eficiente da energia é um fundamental tópico de estudo em *smart cities*. (ALBINO; BERARDI; DANGELICO, 2015; AVALOS; SILVA; LARIOS, 2015; MORA; LARIOS, 2015).

Conforme (MOHANTY; CHOPPALI; KOUGIANOS, 2016), nos últimos anos, além das formas tradicionais de energia, como energia potencial, térmica, química, entre outras, temos novas formas de energia e geração, associadas à termos como energia limpa/*green energy*, energia renovável, e *smart energy*. A energia limpa está relacionada com um consumo de energia que cause o mínimo de impacto possível no meio ambiente; a energia autossustentável/renovável diz respeito às fontes de energia que durarão pelas próximas

gerações e podem ser regeneradas mais rápidas que consumidas. Nesse cenário, surge um conceito mais amplo, o de *smart energy*, onde essas novas formas de energia são incorporadas com TICs, nos diferentes aspectos de energia nas cidades, envolvendo geração, distribuição e *smart grids, smart storage e smart consumption* (MOHANTY; CHOPPALI; KOUGIANOS, 2016).

A infraestrutura utilizada em sistemas de *smart energy* é um fator essencial. Ela é a responsável pela coleta de dados quanto ao consumo de energia e a disponibilização dos dados para monitoramento remoto. Com ela, é possível otimizar o consumo de energia elétrica em um sistema de *smart energy*, através do uso eficiente de armazenamento de energia, *smart meters* e gerenciamento da distribuição de energia (UNCTAD, 2016a). A base dessa infraestrutura é provida pelas *smart grids*, redes de distribuição inteligentes responsáveis por integrar, de forma eficiente, confiável e sustentável, a integração entre consumidores e provedores de energia elétrica, de forma a reduzir as perdas e fornecer energia com qualidade. Elas permitem a integração de diversos tipos de fontes de energia, como por exemplo a energia eólica e a fotovoltaica (MOHANTY; CHOPPALI; KOUGIANOS, 2016). Na arquitetura de distribuição de energia inteligente, o fluxo de energia elétrica e informação é bidirecional, tornando possível que os próprios consumidores também forneçam energia e contribuam com a rede, aspecto que representa um fornecimento descentralizado. Também possibilitado pela *smart grid* é o fluxo bidirecional de informações, onde os sensores podem enviar informações de medições e ser operados. As *smart grids* são também interessantes para as concessionárias de energia, com a possibilidade de identificar de forma precisa quedas no fornecimento da rede ou como viabilizar um reestabelecimento (CALVILLO; SÁNCHEZ-MIRALLES; VILLAR, 2016).

De acordo com (MOHANTY; CHOPPALI; KOUGIANOS, 2016), o uso de tecnologia de informação tem um papel fundamental nas *smart grids* e permitem o gerenciamento de demanda do uso de energia, o despache de energia gerada por turbinas eólicas e painéis solares, serviços independentes de localização para PEVs (*Plug-in Electric Vehicles*), permitindo a recarga de veículos elétricos, e melhorar o relacionamento com os consumidores.

Cidades inteligentes e *smart grids* trazem consigo novas ideias na gestão de energia, processos, tecnologia e métodos computacionais que permitem otimizar e reduzir perdas. Esta pesquisa se caracteriza como um trabalho em cidades inteligentes, visando a gestão do recurso com foco em detecção de anomalias. A ideia de sensores invasivos com medidores instalados diretamente na rede elétrica não é interessante por questões de viabilidade, mas, ainda neste âmbito, há técnicas não invasivas de sensoriamento, como a termografia, que podem ser utilizadas para monitoramento remoto e obtenção de dados relevantes a respeito da condição

da rede e/ou equipamentos elétricos, desde a geração nas usinas e transmissão, até a sua utilização pelo consumidor final.

3.2 TERMOGRAFIA

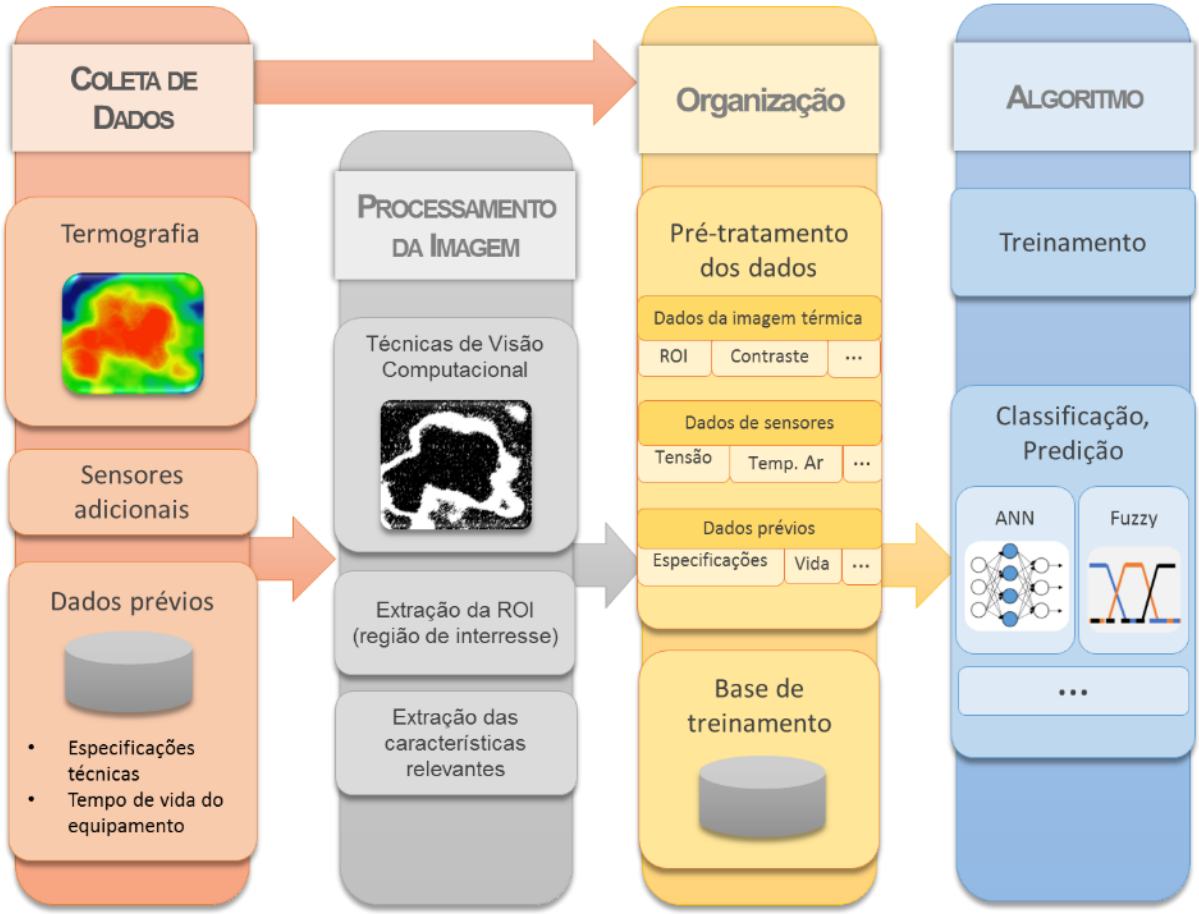
Com o objetivo de analisar de forma automática a condição de equipamentos e componentes elétricos, foram estudadas as principais soluções de trabalhos em energia elétrica, termografia e as técnicas computacionais utilizadas. Conforme (MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012), as falhas em sistemas elétricos podem ser classificadas em algumas categorias, como conexões com folga ou defeito, curto circuito ou circuito aberto, sobrecarga, desbalanceamento de carga e instalação imprópria de componentes. Há alguns fatores que devem ser considerados no equipamento a ser inspecionado e na câmera a ser utilizada para inspeção. Um dos fatores é o ambiental; deve-se levar em consideração vento, umidade e radiação solar, porém o fator temperatura do ar ambiente é o mais relevante dentre todos. Além disso, há os fatores técnicos, como as especificações da câmera, a distância e ângulo em relação ao equipamento, a emissividade do que está sendo inspecionado, gradiente termal entre o ponto de calor máximo interno e a superfície, e a carga elétrica sendo submetida (MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012).

Conforme os trabalhos estudados e apresentados na Tabela 1, o grande desafio em inspeções elétricas por meio da termografia se dá em relacionar todos esses fatores e parâmetros relevantes, e realizar, de forma automática, através de métodos computacionais, uma análise que dê indícios da real condição dos equipamentos, evitando o trabalho extensivo de analisar enormes conjuntos de imagem. Para realizar essa análise, grande parte dos trabalhos na área passaram a estudar e propor técnicas de processamento de imagens e visão computacional, bem como a inclusão de algoritmos de inteligência artificial para a predição e classificação de problemas.

A *Figura 2* ilustra o processo em comum proposto na maioria dos trabalhos. É necessário, a princípio, obter as informações relevantes do sistema inspecionado, além da imagem térmica propriamente dita; informações prévias já conhecidas, como as especificações técnicas e idade média dos equipamentos, ou informações obtidas por sensores adicionais, como medidores de corrente e tensão. Essas informações podem compor uma base de dados, assim como servir de parâmetro para os algoritmos que realizarão o processamento das imagens térmicas capturadas. O objetivo nesse processamento na maioria das vezes são dois, extrair a região de interesse (ROI), normalmente aquela que apresenta as maiores temperaturas, e extrair

as características (*features*) relevantes, que devem ser úteis na posterior detecção do problema. Como proposto por (ULLAH et al., 2017) dados como pico máximo, contraste, histograma, entropia, entre outros. Todos esses dados podem passar por uma etapa de tratamento prévio para servir de entrada para algoritmos computacionais responsáveis pela análise, assim como, em muitos casos, eles também compõem uma base de dados, utilizada para treinamento de algoritmos de IA.

Figura 2: Procedimento geral das soluções em termografia e energia



Fonte: Autor

Em busca de um método geral para classificação de imagens térmicas, dentre todas as possibilidades apresentadas na literatura, o processo indicado na *Figura 2* generaliza a metodologia utilizada nos trabalhos correlatos. Ele em geral consiste na coleta de dados termográficos do sistema a ser inspecionado, além de informações previamente conhecidas, como especificações técnicas. Esses dados são provenientes de equipamentos termográficos e sensores, bem como dados conhecidos do sistema. Em alguns desses casos, técnicas de visão computacional são utilizadas para extração da região de interesse, e em outros, são selecionadas

manualmente. Essa seleção é importante para se obter a região de interesse de estudo, eliminando ruídos e obtendo a área com informações relevantes para a análise do problema. A análise final, após a extração das regiões de potencial problema, pode ficar a cargo de um especialista, acontecer por comparação térmica ou através de algoritmos de IA. Em ambos os casos, há a etapa de organização dos dados, onde todos os dados coletados a partir do processamento da imagens térmica, dos sensores e especificações conhecidas são tratados para utilização em algoritmos de classificação ou predição, compondo uma base de dados. Na análise, normalmente são utilizados algoritmos de inteligência artificial, normalmente de predição ou classificação. Os principais algoritmos utilizados em trabalhos correlatos são demonstrados a seguir, na *Tabela 1*. Evidencia-se nestes trabalhos a importância dos algoritmos de visão computacional e de inteligência artificial na realização da detecção automática de problemas elétricos.

Tabela 1: Trabalhos em termografia e energia elétrica (Continua)

Trabalho	Técnica(s) de proc. de imagem	Método(s) de classificação	Aplicação	Resultados
(KORENDO; FLORKOWSKI, 2001)	Seleção manual da região	Método de Coeficientes Invariantes	Equipamentos elétricos e disjuntores	A hipótese de detecção de problemas elétricos foi verificada. Para o caso de disjuntores, é possível utilizar parâmetros de regiões específicas, e de acordo com a taxa de diferença de temperatura encontrada, diagnosticar um possível problema. Com isso, verifica-se a possibilidade de utilização em outrors equipamentos elétricos específicos, dada suas geometrias, com os parâmetros necessários para tal.
(SMEDBERG, 2006; WRETMAN, 2006)	Segmentação "top-down", extração da ROI por busca de estruturas repetidas (SIFT)	ANN	Equipamentos elétricos	Utilizando de conhecimento prévio, o trabalho faz a segmentação da imagem por meio da identificação de regiões/objetos similares na imagem. A partir disso, foi possível extrair parâmetros relevantes, como a temperatura máxima, a temperatura relativa em relação a outras regiões similares, a distância média de temperatura e de histograma. Testes com a segmentação manual foram feitos, com uma taxa de erro de 9%; para a segmentação automática proposta, o erro encontrado foi de 14,9%.

Tabela 1: Trabalhos em termografia e energia elétrica (Continuação)

(LI et al., 2006)	Limiarização e obtenção de "momentos" da imagem	SVM e Subtração	Instalações de Energia	A técnica utilizada consegue classificar o objeto em 4 tipos de equipamentos de instalações elétricas: transformador, disjuntor, transformador de tensão e transformador de corrente, em tempo real, com uma taxa de 90% de acerto. Com isso, o diagnóstico se deu comparando a imagem original com a atual e realizando uma subtração, obtendo possíveis anomalias.
(WONG et al., 2009)	Intensidade RGB e limiarização	-	Monitoramento da condição de máquinas em fábricas	Um algoritmo de processamento de imagem feito especialmente para detectar sobreaquecimento de máquinas em fábrica obteve uma alta taxa de acerto, 96% das imagens, de 10 mil, foram classificadas corretamente.
(CHOU; YAO, 2009)	Limiarização (*otsu*) e morfologia	-	Instalações de Energia	Com uma base de 100 imagens térmicas, de capacitores, transformadores e outros equipamentos de linhas de transmissão, atingiu 91% de acurácia. Não foi utilizado algoritmo de IA, apenas comparação simples da temperatura de componentes superaquecidos em relação a uma temperatura de referência.
(ALMEIDA et al., 2009)	Segmentação por Watershed Transform	Neuro-fuzzy	Diagnóstico de para-raios em instalações elétricas	Foi utilizado uma base de 764 medições, utilizando imagens térmicas e outros parâmetros relevantes, como material, índice de poluição e tensão suportada, sendo 40% destinado para treinamento do sistema neuro-fuzzy, e 60% para testes. A média de acerto foi de 88,25% na classificação em 4 categorias.
(RAHMANI; HADDADNIA; SERYASAT, 2010)	Limiarização e extração de momentos	SVM	Subestações de energia	Extraem-se as características das imagens utilizando os momentos de Zernike. Com isso, são formados vetores de características que foram utilizados por um classificador SVM, para classificar imagens entre dois tipos de problemas que ocorrem nos terminais e bases de fusíveis. A taxa de acerto foi de 83%, considerando um conjunto de 40 imagens térmicas, com variações entre elas, em posicionamento, rotação e escala.
(DE OLIVEIRA; LAGES, 2010)	Limiarização (*otsu*) e morfologia	-	Linhas de transmissão	Não foi utilizado algoritmo de IA, apenas comparação simples da temperatura da região de maior calor em relação a uma temperatura de referência e diagnóstico humano, com base em uma tabela padrão de temperaturas aceitáveis.

Tabela 1: Trabalhos em termografia e energia elétrica (Continuação)

(SHAFI'I; HAMZAH, 2010)	Extração de dados RGB (pixel a pixel)	ANN	Falhas internas de equipamentos elétricos	O trabalho usa diretamente os dados RGB de todos os pixels da imagem térmica, classificando o problema interno de equipamentos elétricos como nível baixo, intermediário, médio e alto, com uma elevada taxa de acerto de 99,38%. Uma base 336 imagens foi utilizada, com 84 imagens de cada classe de problema, com 50% delas destinadas ao treinamento de uma rede neural de três camadas.
(ZARKOVIC; STOJKOVIC, 2015)	-	Fuzzy e ANN	Substações de energia	Com a utilização de parâmetros previamente conhecidos, como a temperatura anterior de superaquecimento, além da tensão e a temperatura da região superaquecida atual (por meio da imagem térmica), o diagnóstico é feito por dois sistemas em paralelo, um fuzzy e um por rede ANN. Ambos tiveram resultados compatíveis, obtendo um diagnóstico razoável da urgência em que o equipamento deve ser substituido ou realizada manutenção.
(JAFFERY et al., 2017)	Comparação pixel a pixel com imagem de referência	Fuzzy	Painel Fotovoltaico	A partir de imagens térmicas de referência, as imagens a serem analisadas são comparadas pixel a pixel, gerando um *score* de coincidência, por semelhança. Um sistema fuzzy é utilizado com tal *score* para diagnóstico do painel fotovoltaico em classes de problema.
(DOS SANTOS; DE AQUINO; LIRA, 2017)	PCA (Principal Component Analysis)	ANN e Neuro-fuzzy	Falha em transformadores	A partir de fotografias térmicas de três faces (frontais e laterais) de transformadores, uma base com 127 registros foi montada. Se utiliza a matriz de temperatura de cada uma das imagens, e por se tratar de 3 faces, a matriz tem uma alta dimensionalidade. Com o PCA, foi possível reduzir a dimensionalidade do dataset utilizado, ainda que extraindo as principais componentes que explicam o universo, e com isso, viabilizar o uso de uma rede neural utilizando tais componentes como entrada, para classificar o transformador como com falha ou normal, com 86% de acerto. Também foram realizados testes com um sistema neuro-fuzzy, com 83% de acerto.
(ULLAH et al., 2017)	Seleção manual	ANN + Graph cut	Instalações de Energia	Utilizando uma base de 150 imagens térmicas e uma rede ANN, foi possível classificar componentes elétricos em substações como "com defeito" ou "sem defeito", com uma taxa de 79,78% de acerto, melhorada ao adotar o *graph cut*, alcançando 84% de acerto.

Tabela 1: Trabalhos em termografia e energia elétrica (Conclusão)

(LAIB DIT LEKSIR; MANSOUR; MOUSSAOUI, 2018)	FCM, *otsu* e SVM (Limiarização)	-	Instalações de Energia	O autor propõe diversos métodos para extração da região de interesse e limiarização, obtendo a região de possível problema, para imagens térmicas de equipamentos elétricos. O método extraiu com maior precisão a região com problema foi utilizando o SVM, que após o treinamento com os pixels das regiões de defeito, pôde classificar os pixels das imagens em pixel de "defeito" ou de "background" com bastante precisão, e assim binarizar as imagens evidenciando as regiões problemáticas. Apesar disso, não faz a detecção automática.
(MLAKIĆ; NIKOLOVSKI; MAJDANDŽIĆ, 2018)	CNN (Convolution Neural Network)		Transformadores em Instalações de energia	Uma rede CNN, técnica de Deep Learning, é utilizada para localizar a região de calor em transformadores. As diversas camadas da CNN dão informações relevantes quanto a imagem térmica e podem ser utilizadas para diagnóstico do problema. De acordo com as diferenças encontradas em relação a imagens anteriores, é possível diagnosticar se há um problema ou não.

Fonte: Autor

Conforme a *Figura 2* e a *Tabela 1*, a metodologia consiste, em princípio, no processamento da imagem térmica, para extração de informações relevantes para detecção do problema. O procedimento geral tomado consiste em primeiramente realizar o pré-processamento da imagem, para remoção de possíveis ruídos e aperfeiçoamento, adequando-a para os algoritmos que a utilizará. Após isso, a segmentação é realizada para decompor a imagem em partes relevantes para a análise, removendo informações irrelevantes e ruídos, e obtendo as regiões de interesse que serão utilizadas na detecção. Essa etapa é bastante importante, pois consiste em dividir a imagem em plano de fundo e nos objetos/partes de interesse. O quão mais preciso a segmentação for, mais fácil será a tarefa de reconhecimento, classificação e tomada de decisão, normalmente as etapas subsequentes do processo. Em geral, se utiliza toda a informação disponível do problema para a extração do objeto de interesse com precisão, que varia dependendo da situação (ALMEIDA et al., 2009; LAIB DIT LEKSIR; MANSOUR; MOUSSAOUI, 2018; LI et al., 2006; PALIMKAR, 2016; WRETMAN, 2006).

Há trabalhos que propõem a detecção diretamente da análise da região de interesse extraída. Em geral, consistem na interpretação direta das medidas de temperatura, onde as regiões de maior calor podem indicar um possível problema, como é feito no trabalho de Wong

et al. (2009). Nesse caso, se extrai as regiões onde o valor absoluto de temperatura ultrapassa certo limite definido, e assim a falha pode ser diagnosticada. A temperatura pode ser obtida pelo valor máximo de pixel ou pela média da região. Esse método caracteriza uma análise quantitativa, que nem sempre possui precisão, já que fatores externos como temperatura ambiente, umidade, emissividade, entre diversos, influenciam no valor de temperatura. Outros trabalhos propõem uma análise qualitativa, onde após a extração das regiões de maior calor, se analisa sua temperatura relativa ΔT, em relação a outras partes da imagem ou do equipamento, que estão sob condições similares. Uma forma de comparar a similaridade entre dois objetos é através do cálculo da distância de histograma entre duas regiões. Outra maneira, bastante utilizada, consiste em extrair as regiões de maior calor por *limiarização*. A temperatura média da região de calor extraída é comparada com a média da região externa. A diferença de temperatura pode indicar um possível problema, que deve ser julgado por um especialista na área ou por regras pré-definidas de variações de temperatura aceitáveis (CHOU; YAO, 2009; DE OLIVEIRA; LAGES, 2010; WRETMAN, 2006). Alguns dos trabalhos apresentados, como o de Laib Dit Leksir et al. (2018), apesar da utilização de algoritmos mais complexos, que em geral requerem ajuste de mais parâmetros e maior processamento, consistem puramente na extração da região de interesse, que seria a região onde há a possível falha, e não a detecção se há uma falha ou não.

Quando se há o interesse em detectar um possível problema, além da comparação da temperatura absoluta ou relativa, pode-se utilizar outros parâmetros conhecidos e/ou medidos, e com a análise de um especialista, classificar o problema em diferentes tipos de falhas, ou em diferentes graus. Grande parte dos trabalhos trazem estudos e metodologias para a realização da detecção automático de falhas elétricas, partindo da obtenção da região de interesse e realizando a extração de características relevantes dela, para a classificação do problema e assim a tomada de decisão. São propostos diversos algoritmos de inteligência artificial, que com diferentes parâmetros de entrada, característicos do problema estudado, podem realizar a classificação ou predição de possíveis falhas elétricas, cada um com sua taxa de acerto, como indicado na *Tabela 1.*

Algoritmos de IA se mostraram a solução mais utilizada para a detecção automática, conforme indicado pelos trabalhos na *Tabela 1*, devido sua capacidade de generalização, detecção de padrões e remoção de ruído. Há trabalhos que lidam com o problema apenas matematicamente, como o caso de Korendo e Florkowski (2001). O trabalho consiste em comparar a temperatura da área de problema com outras partes sob condições semelhantes. Assim, se utiliza do cálculo de taxas térmicas e fatores invariantes a carga e ambiente, como

neste caso, onde se mede as taxas de temperaturas das regiões entre os terminais de cabo e o conector principal, de um disjuntor. O método, porém, é específico para cada aplicação e cada equipamento dos quais se quer inspecionar. Para esse tipo de solução, as características geométricas da região de interesse e pontos específicos de estudo devem ser definidos com antecedência. Por outro lado, com a utilização de algoritmos de inteligência artificial, cabe apenas a definição de parâmetros relevantes a serem extraídos da imagem térmica, além de outras informações conhecidas, com a técnica mais apropriada para classificação/predição. A técnica mais utilizada nos trabalhos correlatos é a de redes neurais artificiais (ANNs). Elas permitem reconhecer padrões em imagens térmicas e construir superfícies de separação bastante complexas, permitindo sua classificação. É importante destacar a importância de uma base de dados diversificada e com conteúdo suficiente para o treinamento das redes neurais.

O trabalho proposto por Shafi'i e Hamzah (2010) utiliza o espaço de cores RGB e as temperaturas como dados de entrada para a rede neural, classificando a severidade de um possível problema elétrico em 4 classes, e obtendo uma elevadíssima taxa de sucesso, de 99,38%. A base de dados é grande, com 252 exemplos, de diversas imagens térmicas de uma subestação, porém não foi informado se há variedade de equipamentos e diferentes falhas elétricas. É possível que o sistema tenha se superespecializado para algumas situações em específico. Um dos problemas dessa abordagem é a capacidade de computacional necessária; trabalhar com todos os pixels da imagem é uma tarefa que exige alto processamento tanto no treinamento como na tarefa de classificação. Abordagens mais realistas visam a redução dos dados, eliminando possíveis ruídos e mantendo o que é mais relevante para o processo de análise, apesar da menor precisão.

Dos Santos et al. (2017) dedicaram-se em transformadores e consideram a utilização de imagens térmicas de três faces diferentes do equipamento, laterais e frontal. Para a análise de um registro em específico, três imagens, representadas numa matriz tridimensional, devem ser analisadas. Com o algoritmo PCA (*Principal Component Analysis*) foi reduzida a dimensionalidade dos dados, extraindo as principais componentes que explicam o universo e tornar viável a utilização de uma ANN, para identificação de padrões e assim classificação do equipamento em duas classes, "normal ou envelhecimento natural" ou "superaquecendo", onde uma taxa de 86% de acerto foi obtida. É interessante observar um parâmetro adicional de entrada além das componentes das imagens térmicas, que é o nível de carga do transformador no momento da fotografia. Em outros trabalhos envolvendo subestações e linhas de transmissão, é possível perceber que há a extração de diferentes características das imagens térmicas e utilização de outros parâmetros externos, no treinamento e testes com as redes

neurais. Parâmetros bastante utilizados são a temperatura absoluta máxima da região superaquecida, a temperatura relativa, a distância média de temperatura em relação a regiões similares, o histograma e diferenças de histograma da região de interesse, e o seu gradiente (SMEDBERG, 2006; WRETMAN, 2006; ZARKOVIC; STOJKOVIC, 2015). O trabalho de Ullah et al. (2017) traz a extração de ainda mais características visando o aumento da precisão de classificação, com parâmetros estatísticos de primeira e segunda ordem, entre eles, média, variância, desvio padrão, energia, contraste, entropia, assimetria e curtose. A taxa de acerto obtida pela rede neural proposta de três camadas foi de 84%, classificando os componentes em "normal" ou "com defeito".

Além das características extraídas das imagens, também pode-se obter dados provenientes de conhecimento prévio e sensores adicionais, como ocorre no trabalho de Zarkovic e Stojkovic (2015), que utiliza a temperatura anterior do último superaquecimento registrado e a medição de tensão/corrente. Dados como esse nem sempre estão disponíveis, mas quando sim, podem ser úteis no processamento da imagem, identificação de padrões e classificação.

Outra técnica bastante utilizada na classificação dessas imagens térmicas é a *Support Vector Machines* (SVM). Não diferente da abordagem com redes neurais, nesses trabalhos não se usa a informação completa da imagem. Há a extração de características relevantes das imagens para a utilização. SVMs possuem uma alta capacidade de generalização, bastante utilizadas para aproximação de função e classificação, devido sua capacidade de mapear os dados de entrada em um espaço dimensional de ordem mais elevada, onde podem se tornar linearmente separáveis (WRETMAN, 2006). O trabalho de Li et al. (2006) utiliza do método para reconhecer de qual o tipo de equipamento elétrico se trata a imagem; os principais parâmetros da imagem térmica foram considerados como cor, textura e características morfológicas. Técnicas de extração de características foram utilizadas para obtenção de características de imagens independente de movimento e outras transformações. Um vetor de características é extraído e utilizado no classificador SVM proposto, consistindo em um conjunto de classificadores binários para categorizar a imagem entre 4 tipos diferentes de imagens.

Em outro trabalho similar, de Rahmani et al. (2010), as características obtidas se dão pela extração de características das imagens pelos momentos de Zernike. Eles são momentos invariantes, que denotam características das imagens que não variam dependendo de escala, posição, orientação ou reflexão. Assim, se obtém vetores de características por meio desses momentos, estes diretamente usados como entrada no classificador SVM, com um kernel

gaussiano, realizando a classificação entre dois tipos de problemas, falha no terminal do cabo de fusíveis e falha nas bases superior e inferior de fusíveis, com 83% de acerto nos testes realizados. Um diferencial nesse trabalho é que a classificação não é feita para definir se há defeito ou não, ou graus de defeito, mas sim em qual tipo de defeito elétrico se caracteriza (RAHMANI; HADDADNIA; SERYASAT, 2010).

Uma técnica utilizada na área de IA é a lógica *Fuzzy*, por meio de sistemas/controladores *fuzzy*. Um dos seus benefícios é a capacidade de trabalhar com informações imprecisas, fazendo uso de variáveis e regras linguísticas, de forma similar ao modo de pensar humano, tendo bastante uso em sistemas de detecção. Para este caso, é necessário utilizar regras pré-estabelecidas com base no conhecimento existente em problemas elétricos e nas consequentes condições térmicas. Os dados de entrada, assim, são "*fuzzyficados"*, interpretados e comparados com as regras do sistema, onde é possível chegar em alguma conclusão (JAFFERY et al., 2017; ZARKOVIC; STOJKOVIC, 2015). No trabalho apresentado por Zarkovic e Stojkovic (2015), os sinais de entrada utilizados foram a idade do equipamento, nível de tensão, temperatura da região superaquecida (extraída da imagem térmica), temperatura do ambiente e temperatura anterior de superaquecimento (parâmetro previamente conhecido). O sistema infere, a partir da temperatura relativa da região de calor, em relação a temperatura ambiente, e da comparação com a temperatura do último superaquecimento conhecido, se o equipamento possui um alto grau de urgência para manutenção ou troca.

De forma similar, o trabalho de Jaffery (2017) compara a imagem térmica de painéis fotovoltaicos defeituosos com a de painéis em condições normais, extraindo pixel a pixel o grau de correlação entre as duas. E, a partir de um sistema *fuzzy,* por regras que se utilizam desse grau de correlação com a temperatura da região superaquecida, se classifica a falha do painel em níveis de severidade. O grande problema nos trabalhos apresentados que utilizam *sistemas fuzzy* é a necessidade de conhecimento e experiência de especialistas na área de energia e termografia para definição das regras; além disso, são regras exclusivas para aquela aplicação em específico. Outra técnica também interessante é a de *neuro-fuzzy*. Com ela, é possível reduzir de certa forma o conhecimento prévio necessário, já que se utilizam de redes neurais e sua capacidade de aprendizado para encontrar o melhor ajuste dos parâmetros das funções de pertinência, permitindo o uso de regras ainda mais complexas. Alguns trabalhos a utilizaram em conjunto de parâmetros medidos como tensão, índice de poluição, material e fabricante, além das medidas térmicas extraídas da imagem. Apesar de expandir as possibilidades, nesses sistemas ainda é necessário conhecer quais os dados de entrada necessários para a análise e formular a regras dado o conhecimento de como tais parâmetros se correlacionam na detecção

da falha elétrica, não eliminando completamente a necessidade de um especialista (ALMEIDA et al., 2009; DOS SANTOS; DE AQUINO; LIRA, 2017).

No trabalho de Dos Santos et al. (2017) foi feita uma comparação entre uma rede ANN e um sistema *neuro-fuzzy*, para a mesma aplicação e mesmos dados de entrada. Enquanto o primeiro obteve melhor taxa geral de acerto, classificando em duas categorias, normal e superaquecido, com 92% e 79% de acerto, respectivamente, o sistema *neuro-fuzzy* teve melhor balanço de acerto, classificando em normal com 82% de acerto e superaquecendo com 85%. Os dois sistemas se mostraram interessantes para detecção, sendo o primeiro mais efetivo para casos onde se deseja encontrar casos que estão em condição normal, e o segundo quando não se há distinção ou preferência pelas classes que se quer classificar corretamente.

A técnica de termografia se mostrou pertinente no campo de inspeção e detecção de problemas de equipamentos elétricos em geral. Como descrito, há diversas propostas e aplicações de algoritmos computacionais na área. Esses algoritmos são estudados ao longo desse trabalho, visando propor um método geral para análise e detecção de problemas de equipamentos e instalações elétricas e comparar a acurácia de classificação obtida pela comparação de diferentes algoritmos.

3.2.1 *Trabalhos correlatos em outras áreas*

Assim como as pesquisas realizadas no domínio de energia elétrica, os trabalhos de termografia em outras áreas também possuem uma característica em comum, que é a identificação de anomalias por meio da assimilação de variações térmicas anormais. Portanto, também é de interesse estudar os algoritmos e técnicas de processamento de imagens utilizadas, suas aplicações, métodos e parâmetros relevantes, e analisar suas apropriações para a metodologia de análise termográfica em energia elétrica.

A *Tabela 2* apresenta alguns desses trabalhos, com as técnicas de processamento de imagens e/ou inteligência artificial.

Tabela 2: Trabalhos em termografia e algoritmos de IA (Continua)

Trabalho	Técnica(s) de proc. de imagem	Método de classificação	Aplicação	Relação com a pesquisa
(HUO et al., 2017)	2D-DWT (Transformada discreta wavelet)	NN (Nearest Neighbour (NN) + AG (Algoritmo Genético)	Diagnóstico de falhas em rolamentos de máquinas rotatórias	Seleção de características que melhor indiquem variações térmicas para classificação de falhas
(PAGE, 2009)	PPT, TSR, PCA, limiarização e morfologia	-	Detecção de falhas em painéis de fibra de carbono	Detecção de anomalia por observação de pontos de calor anormais
(LIU et al., 2015)	Clusterização por *K-means* e expectation-maximization	-	Diagnóstico de problemas de diabetes de pé	Diagnóstico de problemas por comparação térmica de regiões similares
(DUDZIK, 2015)	PCA	ANN	Detecção de defeito em materiais	Classificaçõ em "defeito" ou "não-defeito" com base na variação térmica
(ACHARYA et al., 2012)	Extração de textura por matriz de co-ocorrencia	SVM	Diagnóstico de câncer de mama	Classificação de "problema" ou "não-problema" por meio da análise de distribuição de temperaturas fora do normal
(KARVELIS et al., 2014)	SIFT	*Naive Bayes* e C4.5	Diagnóstico de falhas em motores de indução	É feita a segmentação de equipamentos em regiões de interesse e sua classificação pela diferença térmica obtida
(ESTEVE AGELET et al., 2012)	PCA	PLSDA, SIMCA, *k-NN*, SVM	Utiliza apenasTrabalha com a faixa *near-infrared*, para detecção de danos em grãos de soja e milho	Apesar de ser um trabalho bem distante de energia, utiliza-se de grãos danificados que emitem radiação/calor diferente do normal e podem ser medidos por termografia e classificados
(MADRUGA et al., 2010)	Extração de parâmetros estatístico de ordens elevadas (HOS)	-	Plastico reforçado de fibras de carbono	Análise de parâmetros estatísticos de uma sequência termográfica. A distribuição de temperatura pode indicar se há um defeito.

Tabela 2: Trabalhos em termografia e algoritmos de IA (Continuação)

(SENNI et al., 2014)	Thresholding	-	Detecção de contaminação e "corpos estranhos" em biscoitos	Pontos de diferença de temperatura em relação ao caso normal são utilizados para identificar problemas
(USAMENTIAGA et al., 2014)	Várias técnicas de processamento para imagens térmicas: PCT, DTT, entre outros	-	Estudo de defeitos em espécimes de fibra de carbono e técnicas de termografia em geral	Identificação de defeitos, aquecendo uma superfície e percebendo as diferenças de distribuição de temperatura
(SVENSSON, 2014)	Thresholding	AdaBoost, *Random Forest*, SVM-RBF	Detecção de vazamentos por meio de termografia e sistema de informação geográfica (GIS0)	Análise de imagens aéreas em infravermelho para a identificação de pontos com temperatura maior que o normal, indicativo de possível problema
(LIU et al., 2017)	Extração de features + PCA	SVM	Diagnóstico de defeitos em servidores	Obtenção da temperatura e tamanho de *hospots* (pontos de maior calor) para análise e diagnóstico de problemas
(SINGH; NAIKAN, 2017)	Extração da matriz de perfil de temperatura da imagem térmica	-	Diagnóstico de falhas em motores de indução	Análise de perfis térmicos apresentados em situações normais e de defeito para detecção de problemas
(ROZO; ALVAREZ; DOMINGUEZ, 2012)	Segmentação por Watershed + Clusterização por EM (Expectation-maximization)	-	Segmentação de imagens térmicas	Não tem uma aplicação em específica, porém, as regiões mais quentes que podem apontar um possível defeito

Tabela 2: Trabalhos em termografia e algoritmos de IA (Conclusão))

(ARAKI; NOMURA; WAKAMI, 1993)	FCM (Fuzzy C-Means)	-	Segmentação de imagens térmicas para obtenção do número de pessoas ocupantes	Identificar a posição dos pontos de calor em uma imagem e segmentá-los. Neste caso é feito para detecção de pessoas, útil para o presente trabalho na detecção de pontos de possível defeito
(WANG; ZHANG; SHEN, 2010)	Extração de características por SCD (Shape Context Descriptor)	LDA e Adaboost	Detecção de pedestres em sistemas de monitoramento	A relação com o presente trabalho se dá em obter e segmentar os pontos de calor da imagem, que podem ser utilizados na devida detecção do que a aplicação específica desejada
(GHANTY et al., 2008)	Clusterização por K-Means	ANN	Estimar largura e profundidade de penetração de escamas de solda	A segmentação de imagens térmicas é feita para obter informações e analisar a região de calor. No caso de energia, seria útil para detectar defeitos em equipamentos
(WONG et al., 2009)	Limiarização	-	Detecção de invasores em condições de baixa luminosidade	É possível detectar anormalidades extraindo as regiões de maior calor e comparando variações térmicas em uma sequência de imagens
(PAPEŽ; PALFY; TURK, 2008)	Algoritmo de detecção de bordas e segmentação manual	ANN	Diagnóstico de síndrome do túnel do carpo	Diagnóstico com base em imagens térmicas de situações onde há problema e não há, onde há uma variação térmica, possibilitando treinar um classificador
(FANG; WU, 2007)	Filtros de passa-baixas e passa-alta + limiarização	ANN	Reconhecimento de objetos distantes, de baixa resolução	A detecção de formas com certa geometria por termografia pode ser útil no cenário de energia quando se conhece o formato da distribuição de calor do equipamento

Fonte: Autor

Na maioria dos trabalhos que se utilizam de termografia para detecção de anomalias, não se conhece exatamente quais valores de temperatura são aceitáveis e quando caracterizam uma falha, assim como nos casos apresentados de energia elétrica. Por isso, há a utilização de diversos algoritmos de processamento de imagens, para a extração das regiões de estudo e obtenção de dados relevantes das imagens térmicas, que podem dizer algo a respeito de um possível problema, além de dar suporte à utilização de algoritmos de IA para classificação.

Se tratando do processamento de imagens, a metodologia geral adotada não é diferente dos trabalhos em equipamentos elétricos, consistindo no tratamento da imagem para melhoramento de contraste e extração de ruídos, na segmentação para obtenção da região de interesse, e na extração de características relevantes, de acordo com as necessidades específicas da aplicação. No trabalho de Singh e Naikan (2017), por exemplo, apenas se extrai uma matriz do perfil de temperatura, da imagem térmica de motores. A partir dela, parâmetros como média e desvio padrão de temperatura são obtidos, onde após comparados com uma imagem térmica padrão e com as variações aceitáveis pré-definidas, é possível identificar anomalias de forma direta, por comparação. É um caso específico, onde não é necessário algoritmos de IA e o processamento da imagem é trivial.

Alguns trabalhos em termografia tratam exclusivamente de processamento de imagens térmicas, sem envolver algoritmos para classificação do problema ou tomada de decisão. Existem diversas propostas de algoritmos para segmentação de imagens térmicas, como é o caso da clusterização por EM (*Expectation-maximization*), como proposto por Rozo et al. (2012), que obteve resultados similares a outros algoritmos mais utilizados, como a segmentação por *watershed*, porém com a vantagem de ser menos sensível a falta de contraste, e a variações de posição e escala, características relevantes se tratando de inspeções térmicas. Outra técnica, proposta por Araki et al. (1993), foi a segmentação por *Fuzzy C-Means*, que teve altas taxas de acerto na segmentação de pessoas, já que a temperatura humana normalmente apresenta diferenças consideráveis em relação ao ambiente em lugares fechados, porém não foi feita a comparação com outros algoritmos. Ghanty et al. (2008) utiliza clusterização por *k-means* para classificar a imagem em plano de fundo e primeiro plano, onde a diferença de temperatura entre os dois normalmente é elevada. A clusterização por EM e *k-means* também é utilizada por Liu et al. (2015) para segmentação. Ainda nos casos em que não se faz necessário algoritmos de IA para a identificação de problemas, há o trabalho de Wong et al. (2009), utilizando a técnica de limiarização para remoção do plano de fundo e obtenção da região de interesse. Elaborou-se um algoritmo que compara a imagem atual com imagens anteriores, dado o monitoramento em tempo real, que é capaz de detectar pessoas que surgem na cena, possíveis

invasores. Existe uma quantidade razoável de trabalhos que utilizam a limiarização para obtenção da região de interesse, uma técnica simples que em muitos casos extrai corretamente a região de estudo, principalmente quando se sabe que há razoáveis diferenças de temperatura entre o plano de fundo e o objeto de interesse (FANG; WU, 2007; PAGE, 2009; SENNI et al., 2014; SVENSSON, 2014; WONG et al., 2009).

Há aplicações em que o diagnóstico ou detecção de falhas totalmente automático não é uma necessidade, embora tenha-se interesse em utilizar algoritmos de IA, como acontece no trabalho de Papez et al. (2008). O diagnóstico de síndrome do túnel do carpo pode ser feito através da utilização de redes neurais com as medidas de temperatura de diversos pontos específicos da mão. Nesse caso, a segmentação da imagem térmica deve ser exatamente nas regiões pré-definidas, portanto opta-se pela seleção manual do especialista.

Em complemento, há aqueles trabalhos que buscam reduzir o máximo possível da subjetividade humana, ou até mesmo tornar o processo de análise totalmente automático, utilizando de algoritmos mais complexos, normalmente para classificação. Uma das etapas mais importantes dentro das metodologias apresentadas é a de extração de características (*features*) das imagens, processo que ocorre na transição entre o processamento da imagem e a análise, onde informações relevantes são obtidas da imagem térmica e servem de entrada para algoritmos de inteligência artificial, sendo fundamental para o sucesso da análise. (FANG; WU, 2007; HUO et al., 2017; KARVELIS et al., 2014; LIU et al., 2015; SVENSSON, 2014; USAMENTIAGA et al., 2014). Conforme Usamentiaga et al. (2014), como as imagens térmicas obtidas podem ser interpretadas como sequência de números (valores de temperatura), esses números podem ser tratados por funções estatísticas, para descrever comportamentos e detectar variações. Medidas conhecidas como momentos estatísticos dão detalhes quanto a distribuição dos valores, como a média e variância. Dois dos principais momentos de estatística de alta ordem (acima da terceira potência) utilizados em termografia são a assimetria, muito utilizada para obter características sem ser afetada por um aquecimento não-uniforme ou superfícies de formatos diferentes, e a curtose, utilizada para perceber regiões com defeito, dado a difusividade (USAMENTIAGA et al., 2014). Existem outros momentos, como os momentos de Hu, que são invariantes a posição, escala e rotação, bastante úteis em casos nos quais não é possível capturar imagens de uma posição fixa. Outras características comumente utilizadas nesses trabalhos compreendem a homogeneidade, energia e entropia, por exemplo (FANG; WU, 2007; LIU et al., 2017). Quando se há maior conhecimento sobre a aplicação, pode-se extrair características que trazem indícios de alguma propriedade em particular. No trabalho de Karvelis et al. (2014), uma das características extraídas consiste na razão entre as temperaturas de regiões

segmentadas da imagem; em, Huo et al. (2017), se utiliza a soma de todos os pixels da imagem; em Svensson (2014), são extraídas propriedades bastante específicas da aplicação, como componentes conectadas, “retangularidade” e “circularidade”. Ainda considerando a extração de características, o algoritmo PCA foi utilizado em diversos trabalhos, como visto na *Tabela 2*.

O PCA é um método estatístico que permite transformar as características de um objeto, resultando em um novo conjunto de características, ordenadas em variância. Selecionando um número de componentes principais desse conjunto, é possível reduzir a dimensionalidade dos dados. Assim, apesar de gerar novas características que não possuem interpretação física, esse método permite explicar as principais fontes de variância, utilizando menos variáveis. O PCA elimina informações redundantes e reduz a correlação entre os vetores de características., permitindo destacar diferenças e similaridades entre os dados. Com a redução da dimensionalidade dos dados, também se reduz a complexidade do classificador, fazendo deste um método muito utilizado em conjunto com SVM e ANN. A maioria dos trabalhos que fazem seu uso tratam de uma sequência temporal de imagens, envolvendo muitos dados para se processar. Outros métodos utilizados para sequências termográficas são o PCT (*Principal Component Thermography*) e o DTT (*Dynamic Thermal Tomography*), técnicas que permitem obter informações a respeito de variações entre as imagens. (DUDZIK, 2015; ESTEVE AGELET et al., 2012; LIU et al., 2017; PAGE, 2009; USAMENTIAGA et al., 2014).

As técnicas mais utilizadas nos trabalhos em termografia e energia elétrica para a detecção de falhas foram as redes neurais (ANN) e classificadores SVM. Nas pesquisas em outras áreas de aplicação, como apresentado na *Tabela 2*, isso não é diferente. São amplamente utilizados em problemas de classificação, como acontece em diversos dos trabalhos com termografia, normalmente com o objetivo de categorizar imagens térmicas de acordo com a gravidade ou tipo de defeitos encontrados, em máquinas, motores, diferentes superfícies e materiais, além dos casos em que se deseja diagnosticar problemas médicos. A característica em comum na maioria das aplicações se dá pelas variações térmicas encontradas, e estas imagens devem ser classificadas de acordo com as características percebidas e o possível problema que elas representam. O processo de análise humana passa a ser substituído por algoritmos de inteligência artificial, onde em alguns trabalhos, foi obtido uma taxa de acerto ainda maior. Embora prefira-se respostas exatas, muitas vezes é difícil modelar um problema usando métodos matemáticos determinísticos. Esses algoritmos, por outro lado, são capazes de encontrar características em comum em um conjunto de dados que podem parecer não ter relação, embora ela exista. Devido sua natureza de aprendizado supervisionado, com base em

exemplos de treinamento conhecidos, esses algoritmos visam determinar um hiperplano que separa os exemplos positivos dos exemplos negativos, permitindo realizar uma classificação. Com uma modelagem adequada e uma base suficientemente grande, muitos trabalhos em termografia conseguiram utilizar redes neurais e classificadores SVM para reconhecimento de padrões e detecção automática de problemas/falhas (ACHARYA et al., 2012; ESTEVE AGELET et al., 2012; GHANTY et al., 2008; LIU et al., 2017; SIM, 2002; SVENSSON, 2014; USAMENTIAGA et al., 2014; WANG; ZHANG; SHEN, 2010; WONG et al., 2009).

Por outro lado, a utilização dos classificadores apresentados, dependendo do número de variáveis de entrada, pode ser computacionalmente custoso. Quando se sabe das particularidades do problema, é possível utilizar outros algoritmos. Por exemplo, o trabalho de Wang et al. (2010) utiliza SCD para extrair descritores, que obtém informações pertinentes para a identificação de formas e objetos específicos, com o objetivo de identificar a forma humana para monitoramento de pedestres. Em vez de utilizar um classificador como o SVM, como muitos trabalhos propõem, preferiu-se o LDA *(Linear Discriminant Analysis*), que é mais eficiente e é aplicável considerando a propriedade de que o conjunto de dados é normalmente distribuído. Existem diversos classificadores para casos em particular, como esse. Nesse mesmo trabalho, também pode-se observar que foi utilizado o AdaBoost, método que permite se obter um classificador forte a partir da combinação de classificadores considerados como fracos, resultando em uma classificação mais precisa. O método se baseia em aplicar diferentes pesos para as amostras e classificadores fracos, atualizando os pesos durante o treinamento, onde a cada iteração os pesos das amostras mal classificadas são incrementados. Esses classificadores fracos devem ser simples, não importando sua taxa de acerto, apenas necessário que realizem uma classificação pelo menos melhor que aleatório (SVENSSON, 2014).

Um dos problemas apontados anteriormente é o de decidir quais das características obtidas na etapa de extração são mais relevantes na classificação. Com essa informação, seria possível tornar o processo mais eficiente, utilizando apenas aquelas que mais influenciam na decisão final. Nesse aspecto, o trabalho de Huo et al. (2017), que faz a extração de diversas características de domínios de tempo e frequência, propõe um método para seleção dos parâmetros mais importantes. Por não haver exatamente uma direção de quais são as melhores características, é sugerido que, em vez de utilizar do conhecimento prévio de acordo com a experiência do autor, um método híbrido para classificação e seleção de variáveis seja utilizado. Por meio do uso de um classificador NN (*Nearest Neighbors*), método bastante comum de classificação, o trabalho visa analisar imagens térmicas de rolamentos defeituosos, em conjunto com a aplicação de um algoritmo genético (GA). A proposta consiste na seleção de duas das

características extraídas das imagens, formando um espaço de características 2D, que é utilizado com o classificador NN para analisar a efetividade dos parâmetros escolhidos. A taxa de sucesso é utilizada como função de aptidão do algoritmo genético, que após diversas gerações, seleciona os melhores indivíduos, que representam uma solução sub-ótima. Essa solução indica, portanto, quais são as duas características selecionadas que maximizam a taxa de sucesso do classificador NN. Sabendo quais os parâmetros mais relevantes para esta aplicação em específico, é possível então utilizá-los para a classificação de outras imagens de forma mais eficiente, além de servirem de entrada para outros classificadores.

Outros classificadores, menos utilizados, mas que podem ser úteis na classificação de imagens térmicas, são aqueles baseados em árvores de decisão e em *Naive Bayes*, como proposto por Karvelis et al. (2014), onde se viu a possibilidade de utilizá-lo, com resultados promissores. Métodos como *Naive Bayes* são úteis quando a base de treinamento não é muito grande, ou quando o grau de interação entre os parâmetros de entrada é pequeno. Algoritmos *Random Forest* e C4.5 também são interessantes para se verificar a importância de cada característica e para a formulação de regras de classificação, visto que o "conhecimento" adquirido pode ser interpretado facilmente, diferentemente de outros métodos caixa-preta (KARVELIS et al., 2014; SVENSSON, 2014).

A partir do estudo dos trabalhos em termografia e dos algoritmos utilizados, é possível concluir que há uma grande similaridade com as pesquisas apresentadas na área de energia, bem como há também a utilização e proposta de outros métodos de processamento de imagens e IA, conforme a diversidade de aplicações. Com o mapeamento realizado, é possível quais são os principais métodos que podem ser utilizados na proposta deste trabalho.

4 METODOLOGIA

Este trabalho visa propor um método para a análise de falhas elétricas por meio das técnicas de termografia, processamento de imagens e inteligência artificial. A partir da revisão bibliografia realizada, pode-se perceber a viabilidade da solução, onde diversas técnicas na área foram utilizadas com o fim de classificar imagens em diferentes tipos de problema.

Há uma grande diversidade de métodos e algoritmos de IA e *machine learning* para análise termográfica, porém a análise comparativa dos resultados entre todos esses algoritmos para imagens térmicas é escassa, principalmente no que se diz respeito a acurácia de classificação e desempenho das soluções apresentadas. Na metodologia de pesquisa, se propõe um método que considere as principais técnicas e algoritmos que podem ser utilizados em condições específicas nos cenários em energia elétrica e imagens térmicas. A partir disso, pretende-se verificar o método sugerido, na metodologia de experimento, onde devido à escassez de *datasets* de imagens térmicas, utilizando uma simulação elétrica simples, em ambiente controlado, é possível realizar a coleta de dados e obtenção das imagens térmicas que irão compor nosso *dataset*, para comparação desses algoritmos, sua eficácia e desempenho.

É importante destacar que se propõe um método que generalize a metodologia nos trabalhos correlatos encontrados na literatura. Cada trabalho apresenta suas particularidades, mas é necessário um método geral que englobe todas essas técnicas para que seja possível realizar uma análise comparativa coerente.

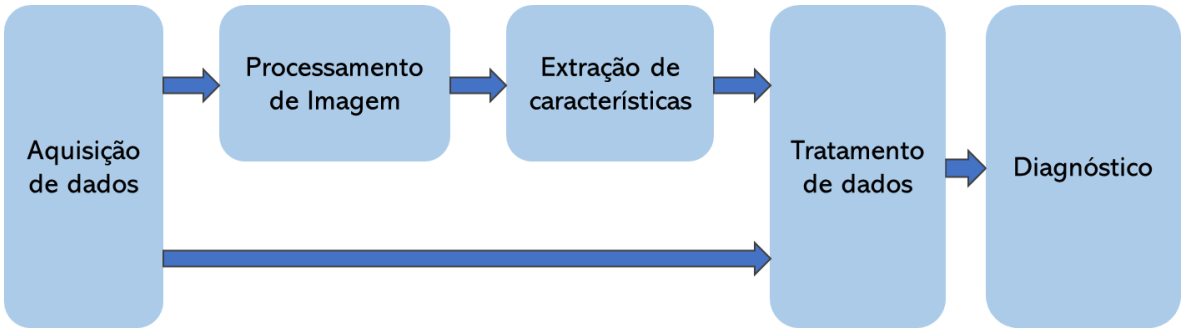
4.1 METODOLOGIA DE PESQUISA

Por meio da revisão dos trabalhos correlatos, é possível observar que o método geral utilizado consiste no processamento da imagem térmica, onde são ressaltadas as regiões de interesse para estudo, se extraem as informações e características relevantes para análise, e por fim algoritmos de inteligência artificial são utilizados para detecção de possíveis problemas, permitindo a tomada de decisão. O processo pode ser aplicado em diversos casos envolvendo equipamentos e componentes elétricos, como transformadores, disjuntores, linhas de transmissão e equipamentos em geral encontrados em estações e usinas de energia, bem como em painéis fotovoltaicos e baterias. Além disso, há muita semelhança encontrada em outras áreas de aplicação, como ocorre na detecção de vazamentos de água e gás, inspeção de defeitos em máquinas rotatórias e rolamentos, monitoramento autônomo de detecção de animais em plantações e florestas, monitoramento pedestres e invasores, diagnóstico de complicações

médicas, entre outros. Existe uma grande semelhança nesses trabalhos, onde através da observação de manchas de calor incomuns, pode-se pressupor um possível problema.

Diversas técnicas e métodos podem ser utilizados nos domínios onde variações de calor representam possíveis anormalidades ou diferenças de comportamento, já que possuem propriedades semelhantes, conforme visto na revisão bibliográfica apresentada. Com base nos trabalhos que foram apresentados na seção 3.2, o método proposto nesse trabalho é ilustrado na Figura 3, que será detalhado nas subseções a seguir.

Figura 3: Metodologia de detecção de problemas elétricos



Fonte: Autor

Embora exista uma diversidade de métodos, em geral elas consistem em filtros e técnicas para melhoria da qualidade e contraste, onde as regiões de interesse podem ser extraídas por métodos de segmentação, até mesmo por técnicas diretas como a limiarização. Esses possuem como característica em comum manchas de calor onde a diferença térmica pode ser facilmente notada, existindo um limiar de separação. Trabalhos mais sofisticados em processamento de imagens vão mais adiante, como Chou (2009), Karvelis et al. (2014) e Li et al. (2006), visando encontrar o problema de maneira totalmente visual, ao realizar a detecção por meio da extração da região de interesse e observação de variações em sua temperatura, por comparação com imagens anteriores ou outros métodos. Nos trabalhos que se utilizam de algoritmos de inteligência artificial busca-se, além de uma precisão maior de classificação, reduzir a subjetividade humana, por meio da detecção de anomalias através da análise computacional das imagens térmicas. Nesse caso, normalmente se extraem uma série de parâmetros, características relevantes das imagens dado a aplicação desejada, utilizados por algoritmos de IA, de classificação supervisionada. Além disso, esses algoritmos podem ser alimentados por outros dados de entrada, proveniente de medições de sensores ou especificações já conhecidas. De acordo com as referências bibliográficas, na maioria dos

métodos se utilizam classificadores na detecção de falhas, categorizando cada caso em "problema" ou "não-problema", ou em níveis de severidade.

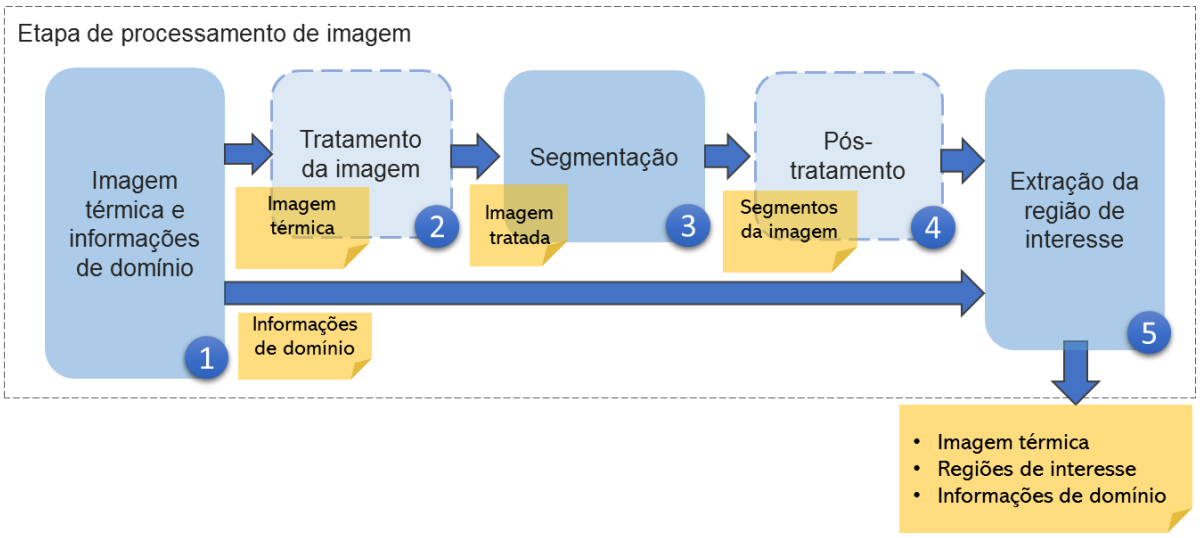
O método proposto na *Figura 3* é detalhado nas seções a seguir, de acordo com as particularidades da aplicação em energia elétrica que se deseja analisar, as diferentes variáveis envolvidas e algoritmos, como visto na revisão bibliográfica e trabalhos correlatos.

4.1.1 Processamento de imagem

A etapa de processamento da imagem consiste em realizar as operações necessárias sobre as imagens com o objetivo de realizar sua análise, prepará-las para a extração de características ou para utilização direta com algoritmos de IA. Essas características podem ser diversas, dependendo da aplicação desejada. Muitos dos métodos de análise e extração de características em termografia dependem da identificação da região de calor, onde há um possível problema. A partir disso, pode ser feito uma análise de forma visual, por inspeção humana, ou por análise das características térmicas, como apontado por alguns trabalhos na área, como o de Wong et al. (2009), Chou e Yao (2009), e De Oliveira e Lages (2010).

A etapa de processamento de imagem proposta nesse trabalho, como ilustrada *na Figura 4*, visa o tratamento e a extração da região de interesse de acordo com os métodos geralmente adotados na área de termografia, bem como na de processamento de imagens, permitindo realizar a análise da imagem e detecção de problemas de maneira automática.

Figura 4: Metodologia da etapa de processamento de imagem



Fonte: Autor

O procedimento geral consiste em realização a segmentação da imagem nas regiões que se caracterizam como manchas térmicas. Essa metodologia é similar aos trabalhados na área, com diferentes algoritmos de segmentação. As informações de domínio e outras propriedades da imagem podem ser utilizadas (*Figura 4.1*), junto com as regiões segmentadas, para a extração da região de interesse. As etapas de tratamento de imagem e pós-tratamento são consideradas como opcionais (*Figura 4.2* e *Figura 4.4*), visto que não são executadas na maioria dos trabalhos, apenas em algumas situações onde a qualidade da imagem não é suficiente para que os métodos de processamento de imagem retornem resultados precisos.

Nesse trabalho, se propõe a utilização da limiarização como principal forma de segmentação, que ocorre na etapa ilustrada na *Figura 4.3*. Uma característica na grande maioria das aplicações elétricas é que os equipamentos e componentes envolvidos possuem temperatura superior ao restante do ambiente, considerado como plano de fundo. Conhecendo o limiar de separação entre os dois, é possível extrair a região de interesse de estudo, sendo uma técnica simples e amplamente utilizada. Porém, há diversos casos onde não se conhece esse valor. Uma direção é obter a medida de intensidade média de toda a imagem, onde possivelmente a região de interesse estará em um limiar superior; de forma similar, método de *otsu*, por meio do uso do histograma de intensidades da imagem, permite identificar de forma automática o limiar que melhor separa dois conjuntos, um de menor e outro de maior intensidade, assim separando a região de interesse (de maior calor) do plano de fundo (ALMEIDA et al., 2009; CHOU; YAO, 2009; DE OLIVEIRA; LAGES, 2010; LAIB DIT LEKSIR; MANSOUR; MOUSSAOUI, 2018; PAGE, 2009; WONG et al., 2009).

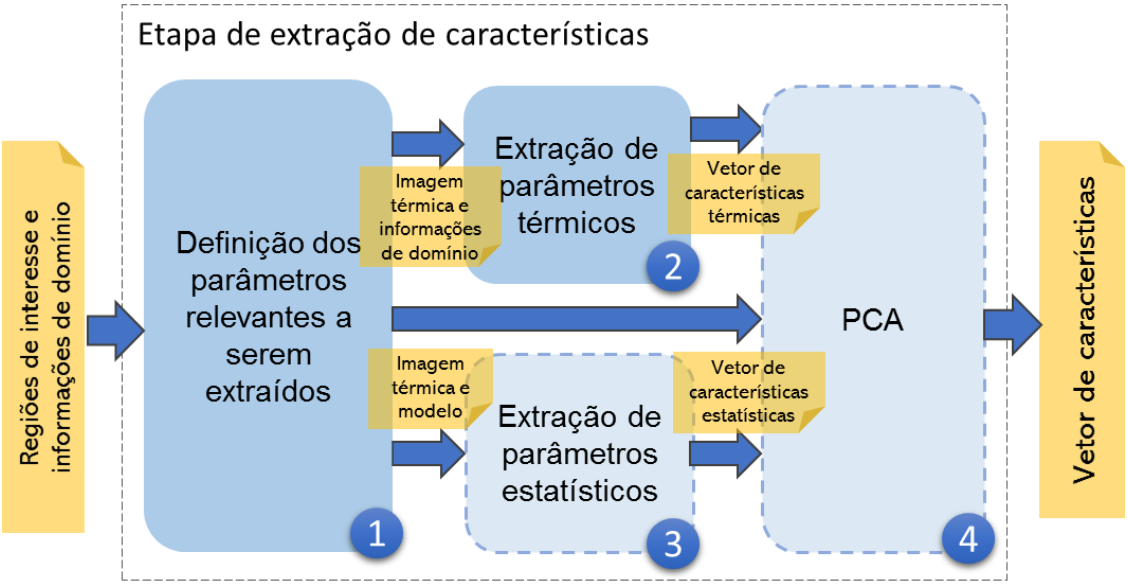
Nesse trabalho, se propõe a utilização da limiarização por *otsu* e da segmentação por *k-means* que obtiveram bons resultados nos trabalhos correlatos (ALMEIDA et al., 2009; CHOU; YAO, 2009; DE OLIVEIRA; LAGES, 2010; LAIB DIT LEKSIR; MANSOUR; MOUSSAOUI, 2018; PAGE, 2009; WONG et al., 2009). Comparou-se os resultados que foram obtidos por ambos e documentou-se as situações em que se saem melhor. A imagem, a princípio deve ser convertida para *grayscale*, para uso com tais algoritmos. Após a segmentação, a morfologia (Figura 4.4) pode ser aplicada para remover ruídos e conectar componentes da imagem. Com isso, se extrai a região de interesse (*Figura 4.5*), removendo o plano de fundo e obtendo a região de calor. Após o processamento da imagem e extração da região de interesse, é necessário utilizar as informações adquiridas para realizar a análise e detecção de falhas. A próxima etapa consiste na extração de características (*Figura 5*), onde diversos parâmetros da imagem e da região extraída são obtidos, sendo relevantes para a classificação e análise das condições elétricas do equipamento ou componente inspecionado.

4.1.2 Extração de características

A etapa de extração de características é fundamental na análise da imagem térmica. O objetivo nesta etapa é obter, a partir da imagem já processada, as informações que a represente com o máximo de detalhes possível e sejam relevantes para a detecção de problemas. Essas informações então servem de entrada para algoritmos de análise e classificação. É possível utilizar a informação pixel-a-pixel, fornecendo os valores de intensidade *grayscale* ou RGB, de todos pixels da imagem para o algoritmo de IA. Esse tipo de abordagem envolve uma série de questões. A primeira se dá pelo alto custo computacional envolvido e tempo para treinamento do classificador e testes. Além disso, pode haver uma grande quantidade de ruído, já que diversos pixels irrelevantes são considerados pelos classificadores, resultando em um modelo que descreve o ruído, o que compromete sua capacidade de generalização. Mesmo após a extração da região de interesse, obter as informações de todos os pixels da região de interesse ainda requer elevada capacidade computacional.

Nessa etapa, busca-se extrair características da imagem, traduzindo-a em uma quantidade menor de informações, mas que continuem sendo relevantes para os algoritmos de IA. A decisão das características que são extraídas é de grande importância, dado que descritores que representam mal uma imagem ou tragam informações irrelevantes para aquela aplicação podem comprometer todo o processo de análise. Nesse caso, não importará quais algoritmo e métodos são utilizados, pois não haverá informação suficiente para que haja uma classificação precisa. Essa fase, portanto, é essencial para o sucesso da classificação e as características devem ser selecionadas levando em consideração o objetivo da análise, e se essas informações são pertinentes ao tipo de classificação que se deseja. A *Figura 5* propõe um procedimento para a extração de características no domínio de termografia e energia elétrica.

Figura 5: Metodologia da etapa de extração de características



Fonte: Autor

Os trabalhos na área propõem a extração de características referente as condições térmicas do equipamento e do ambiente (*Figura 5.2)*. São as diferenças térmicas, justamente, um dos principais indícios de que um equipamento está com problema ou não e podem ser utilizadas para analisar o seu comportamento elétrico, como foi feito nos diversos trabalhos apresentados na literatura (ALMEIDA et al., 2009; CHOU; YAO, 2009; KARVELIS et al., 2014; SHAFI'I; HAMZAH, 2010). Portanto, propõe-se a extração de parâmetros térmicos para a classificação que será realizada. Nos trabalhos estudados, a obtenção dessas características se dá a partir da imagem original e da região de interesse extraída. A partir da temperatura mínima e máxima na imagem, que são valores conhecidos obtidos da termografia, é possível obter os valores de temperatura das regiões de calor, dado a intensidade de *grayscale*. As duas maneiras de encontrar a temperatura se baseiam em encontrar o pixel de maior intensidade na região e converter o valor para temperatura, ou tirar a média de intensidade de todos os pixels da região de calor.

O método escolhido depende da precisão da segmentação. Embora a temperatura média seja mais interessante, se não há certeza que a região segmentada é exatamente a região de calor, ou seja, há ruídos trazendo informações além do que interessa, a média de temperatura pode ser mais baixa do que realmente é. Nesses casos, é mais indicado obter apenas a temperatura máxima nessa região. Esse valor absoluto de temperatura normalmente não é relevante, considerando que outras condições, como a temperatura ambiente pode influenciar nessa medida. A temperatura da região em relação a outras regiões é um parâmetro que pode

indicar uma possível anomalia. Isso ocorre quando há componentes similares na mesma imagem, onde a diferença de temperatura entre eles é um parâmetro interessante. Em outros casos, se obtém a temperatura da região de interesse em relação a temperatura do plano de fundo (ambiente).

A temperatura absoluta ou relativa dificilmente é suficiente para a análise. Desconsiderando o caso em que existam dados tabelados de quais são os limites temperatura aceitáveis para dada aplicação, situação em que é fácil realizar a classificação, correlacionar valores de temperatura e concluir que há um possível problema nem sempre é possível. Mais informações a respeito da distribuição térmica devem ser obtidas. É através do uso delas, que podem a princípio parecem dizer pouco, que com algoritmos de IA pode-se criar correlações junto a exemplos de classificação e perceber padrões. Parâmetros estatísticos são bastante utilizados para isso, como indicado na *Figura 5.3*.

A partir do histograma de intensidades da imagem, normalmente em *grayscale*, é possível realizar a análise matemática das diferenças e semelhanças entre as diversas amostras, como por meio da média e variância. Elas, além de outras medidas, dizem algo a respeito de uma distribuição. Diferenças térmicas no objeto de estudo e na imagem térmica podem ser percebidas através da correlação dessas medidas. A forma tradicional mais encontrada na literatura faz uso da GLCM (matriz de co-ocorrência em *grayscale*) e da sua análise estatística. Além da média e variância, os parâmetros mais utilizados com sucesso nos trabalhos em termografia e energia foram o desvio padrão, assimetria, curtose, energia e entropia. Em especial, as medidas de assimetria e curtose descrevem bem problemas elétricos; a primeira representando a quão assimétrica é a forma que representa a distribuição de temperaturas, e a última, os picos nessa distribuição. Normalmente há diferenças notáveis nessas medidas conforme pontos da imagem apresentam temperaturas diferentes do normal, por isso mostraram-se bons descritores para serem utilizados na classificação. Momentos estatísticos invariantes também são bastante utilizados, devido a sua capacidade de representar informações não dependendo de aspectos como posição, escala ou rotação. São uteis em aplicações onde a câmera térmica ou objeto de estudo nem sempre estão fixos no mesmo local.

Por fim, métodos para a redução de dimensionalidade podem ser utilizados. O PCA (*Figura 5.4)*, amplamente utilizado nesse domínio, pode ser aplicado tanto nos parâmetros térmicos ou estatísticos extraídos, bem como sobre a imagem original. Normalmente, aqui se considera o número de características extraídas e a dificuldade, tanto em termos de desempenho como acurácia, de utilizá-los em um classificador. Essa técnica gera um novo vetor de características, ordenado por variância, onde é possível selecionar as características com maior

variabilidade, portanto as mais significantes, para serem utilizadas na classificação. Um método sem a extração de características seria aplicar o PCA sobre todos os pixels da imagem, para obter o vetor de características que será utilizado na classificação. Essa também pode ser considerada uma análise estatística da imagem, que permite obter as características mais dinâmicas, que causam variabilidade. Deve-se evitar essa abordagem, porém, quando se há bastante ruído, já que estes podem ser os responsáveis pela imagem e outras informações relevantes serão perdidas.

A descrição das propriedades de uma região é enfim representada por um vetor de característica multidimensional, como visto até aqui. Esse vetor, contendo diversas informações que descrevem a imagem térmica é utilizado como entrada para algoritmos de IA para realizar a analise desejada. Neste trabalho, serão analisadas ambas as combinações de no resultado final, as características térmicas e estatísticas isoladamente, e a classificação utilizando um conjunto com as duas combinadas.

4.1.3 Classificação e detecção de problemas

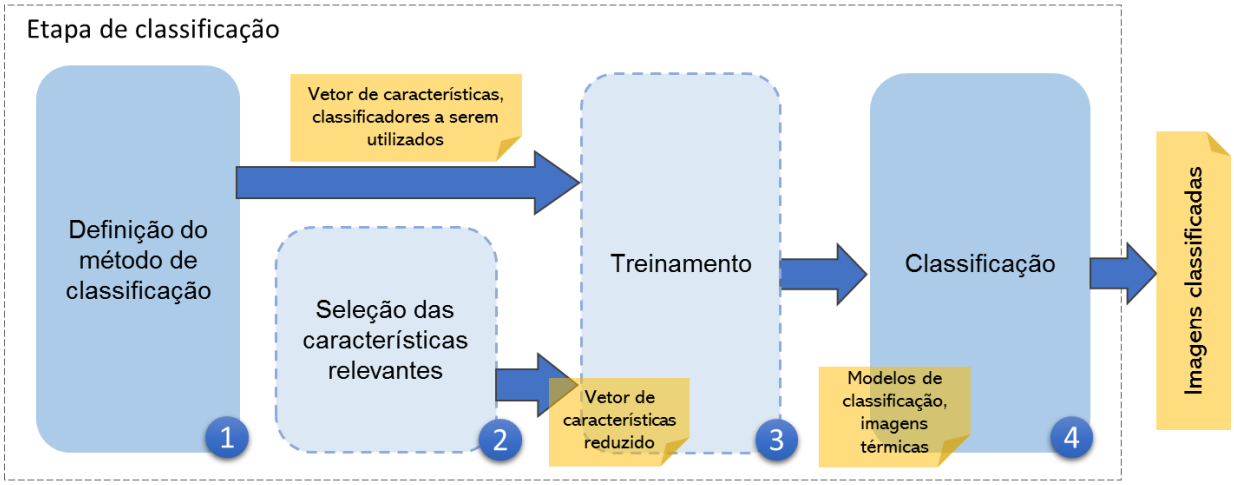
A fase de classificação busca, finalmente, realizar a análise da imagem e identificar possíveis problemas. No método proposto, considera-se a classificação da imagem em diferentes categorias, como o nível de severidade ou tipos de problema. Existem diversas abordagens para a detecção do problema. Geralmente a escolha se dá pelo treinamento e utilização de classificadores (*Figura 6.3* e *Figura 6.4*). Os métodos mais utilizados na literatura são SVM e redes neurais, devido suas capacidades de reconhecimento de padrões e generalização, mesmo para conjuntos de entrada não linearmente separáveis, problemas bastante complexos de resolver. Enquanto SVM são classificadores originalmente lineares, mas que podem ser estendidos para transformar o espaço de características e assim realizar a classificação, as ANNs (redes neurais) podem ser treinadas para construir superfícies complexas de separação. Os dois métodos carregam consigo complexidade de parametrização e exigem, para grandes conjuntos de dados, capacidade computacional considerável.

Há casos em específico que classificadores mais simples podem ser utilizados. Se o conjunto de dados for simples e houverem bons descritores, ou seja, características *k-NN* resolve. Se o conjunto de dados for normalmente distribuído, o classificador mais eficiente possível é o LDA, que nesse caso se sairia muito melhor que uma SVM. A flexibilidade das ANNs e classificadores SVM, porém, é o que faz deles os algoritmos mais utilizados. Normalmente, não se há conhecimento específico ou uma direção a respeito da distribuição e

significância dos dados. Também há técnicas também como as árvores de decisão, que por exemplo são mais rápidas de serem treinadas em relação a redes neurais. Por outro lado, o treinamento *on-line* das mesmas é mais complicado, já que muitas vezes implica na reconstrução da árvore. A grande vantagem se dá pela sua interpretabilidade, permitindo até mesmo extrair regras de classificação a partir do modelo gerado.

Ao se utilizar classificadores, se deseja correlacionar os parâmetros obtidos com o problema, por isso, é necessário possuir uma grande base de informações com exemplos de todos os casos em que se quer classificar. Normalmente, eles são sensíveis ao conjunto de treinamento e requerem uma grande quantidade de dados, que sejam suficientes para identificar padrões e generalizar cada caso. *Naive Baiyes* é uma outra técnica que permite realizar a classificação, conforme aparece nos trabalhos correlatos, que devido a sua capacidade de abstrair as relações existentes entre as variáveis de entradas, é um método útil ao menos para testes rápidos, com um conjunto de dados ainda pequeno. A *Figura 6* ilustra o procedimento proposto nesta etapa.

Figura 6: Metodologia da etapa de análise/classificação



Fonte: Autor

É importante destacar que mais de um classificador pode ser utilizado. Há diversas maneiras de se realizar tal abordagem; neste trabalho propõe-se comparar o resultado de diversos classificadores e realizar a tomada de decisão por votação, onde vários classificadores dão o resultado indicado e por maioria se define a qual classe aquele caso pertence.

Nesse trabalho, propõe-se utilizar redes neurais, SVM, *k-NN*, árvores de decisão, *Random Forests* e *Naive Bayes* para obter modelos de classificação (*Figura 6.3)* e comparar seus resultados. A seleção de características relevantes (*Figura 6.2)*, uma etapa opcional, não

foi utilizada, dado que o vetor de características proveniente da extração de características não é grande o suficiente. Com isso, as imagens podem ser classificadas a partir dos modelos obtidos, cada um com suas taxas de acurácia.

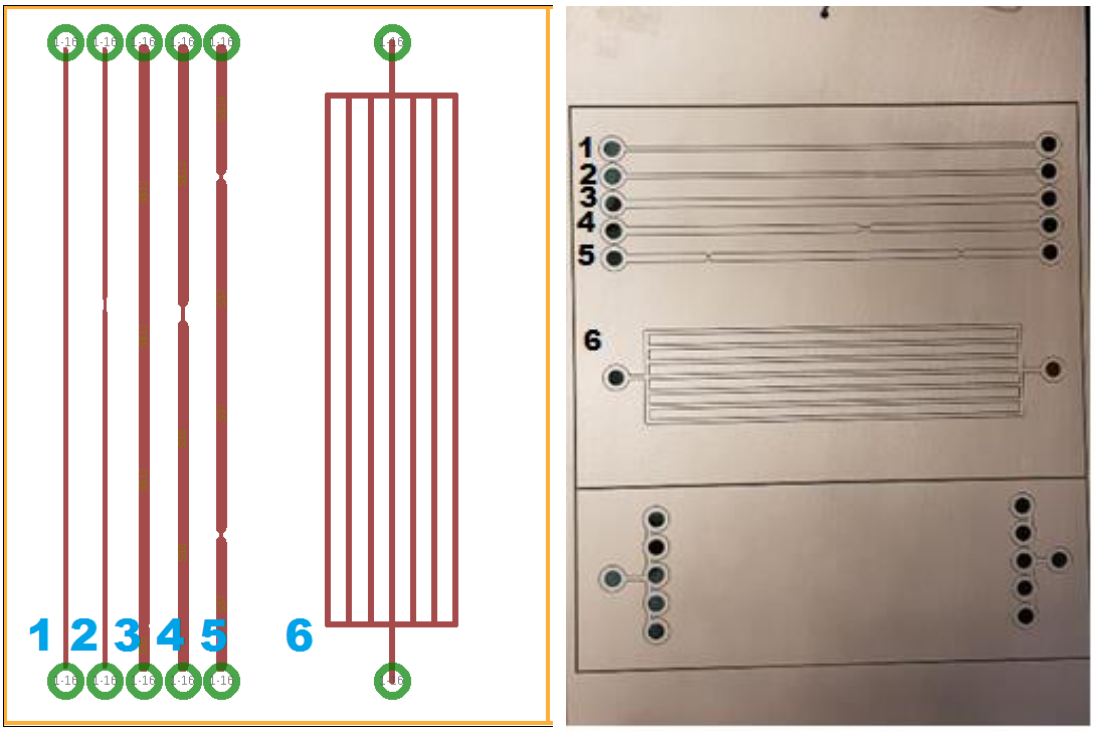
4.2 PLANEJAMENTO EXPERIMENTAL

Com o experimento proposto, pretende-se validar a metodologia apresentada para inspeção de problemas elétricos. O desafio na inspeção e detecção automática desses sistemas se dá nas particularidades que o domínio apresenta. Apesar de problemas elétricos em geral contribuírem com produção maior ou menor de calor, as diversas variáveis envolvidas devem ser compreendidas para assimilar quais variações de temperatura realmente indicam um problema ou não. Outro desafio é a obtenção de tais dados, já que não existem *datasets* públicos em tal aplicação de imagens térmicas, e realizar a coleta de dados em ambiente real é um processo extensivo.

A aplicação deste trabalho propõe a utilização de uma placa de circuito impresso, em um ambiente controlado, com a simulação de condições adversas no circuito de fornecimento de energia. A placa contém trilhas em diferentes larguras, algumas com pequenas descontinuidades, onde há menor contato. Essa é uma aplicação onde as variáveis mais relevantes são conhecidas, tanto medidas de dimensões, como as medidas elétricas. É possível controlar as variáveis elétricas envolvidas nas simulações, bem como a temperatura do ambiente, possibilitando verificar o comportamento do sistema submetido a diferentes condições, situação típica de problemas neste domínio. Através do experimento proposto, pretende-se capturar imagens por meio de uma câmera térmica, posteriormente utilizadas na metodologia proposta de processamento de imagens, classificação e detecção, seguido da análise dos resultados. Apesar de se tratar de um caso distante da situação real, como em linhas de transmissão e subestações, permite a obtenção de imagens térmicas para validação da metodologia de classificação.

A imagem apresentada na *Figura 7* mostra a placa construída para realização dos experimentos. Seu funcionamento é detalhado na seção 4.2.1.1.

Figura 7: Placa de circuito impresso desenvolvida

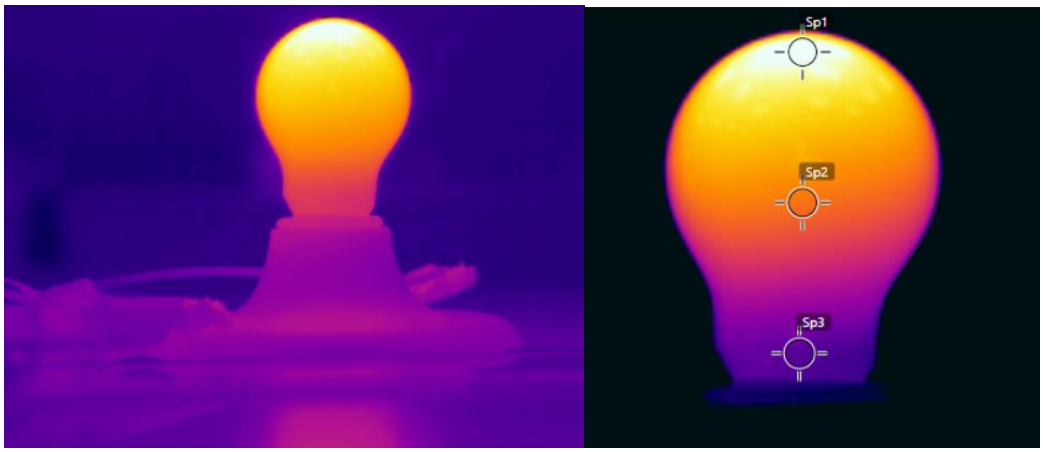


Fonte: Autor.

Esse é um experimento que tem como objetivo obter imagens térmicas de uma aplicação em energia com suas diversas particularidades, para verificar a metodologia proposta, perceber a sua capacidade de generalização e entender como as múltiplas variáveis influenciam na detecção realizada. A simulação com uma placa e suas trilhas propicia tal cenário, dado que os problemas elétricos diversos, como em outras aplicações reais, podem causar variações de calor. Apesar de não ser um caso que pode ser extrapolado para o cenário real, como situações em subestações e linhas de transmissão, ainda é relevante e suficiente pois fornece imagens que permitem a validação da metodologia de classificação.

Esse experimento foi baseado em uma tentativa mal-sucedida, de uma simulação com lâmpadas. Por ser um experimento com poucas variáveis e em ambiente controlado, imaginou-se obter imagens conforme a variação de temperatura da mesma, onde o calor seria proporcional à tensão/corrente, e, portanto, indicativo de um possível problema.

Figura 8: Tentativa de simulação com lâmpadas



Fonte: Autor.

O que não se considerou, porém, foi que a lâmpada produz radiação, que é capturada no espectro da imagem térmica, e, portanto, seria uma variável não-elétrica influenciando no experimento, que consequentemente interferem indesejavelmente nos resultados de classificação, impossibilitando a validação do mesmo. Com a experiência obtida deste experimento, foi possível propor a placa de circuito impresso, que apesar de não representar um cenário real, não inclui variáveis e ruídos não relacionados à energia (diferentemente do caso das lâmpadas).

4.2.1 Protocolo Experimental

Assim como em outras áreas no domínio de energia de energia, o sistema elétrico pode ter suas propriedades em particular e estar submetido a diferentes condições. Neste experimento, as diferentes espessuras de trilhas e as regiões de mal-contato são um aspecto determinante que pode influenciar na classificação e detecção, tornando esse um caso passível de estudo e relevante para análise da metodologia proposta. As subseções a seguir detalham todos os aspectos considerados e materiais necessários para a realização dos experimentos.

4.2.1.1 Placa de circuito impresso

A placa de circuito impresso apresentada na *Figura 7* foi confeccionada para realização do experimento. Ela possui seis trilhas com diferentes características, e, portanto, diferentes perfis térmicos. Essas variações são importantes para realizar simulações e coletar dados em uma aplicação específica, para observação do comportamento dos algoritmos e técnicas propostas na metodologia do projeto, e análise da precisão da classificação para as diferentes

condições dadas. A *Tabela 3* apresenta as diferentes trilhas presentes na placa e suas características.

Tabela 3: Trilhas da placa de circuito impresso

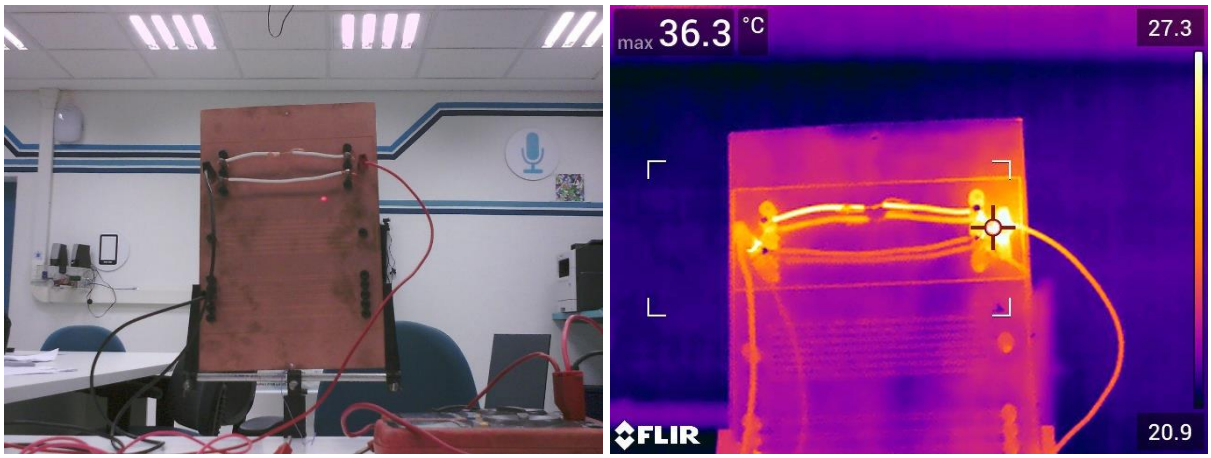
Trilhas	Características
Trilha 1	2mm de espessura, totalmente contínua
Trilha 2	2mm de espessura, com uma descontinuidade mais estreita, com 1cm
Trilha 3	3mm de espessura, totalmente contínua
Trilha 4	3mm de espessura, com uma descontinuidade mais estreita, com 1cm
Trilha 5	3mm, com duas descontinuidades, de um 1cm de espessura
Trilha 6	Conjunto com 7 trilhas unidas e continuas, de 2mm de espessura

Fonte: Autor.

As trilhas mais espessas esquentam mais rapidamente que as trilhas quando submetidas a alguma corrente elétrica. As trilhas 2, 4 e 5 possuem pequenas descontinuidades, regiões de menor contato em que se espera uma maior resistência, portanto maior produção de calor. Esse é um problema elétrico comum em que se deseja validar a metodologia. As diferentes espessuras (2 e 3mm) foram incluídas para introduzir uma variação na produção de calor pra diferentes correntes, que deve ser generalizada pelos algoritmos de classificação. A trilha 6 apresenta outra variabilidade, onde há um conjunto maior de trilhas e requerem mais corrente para aquecerem, apresentando uma distribuição térmica diferente, também a ser incluída nos casos de treinamento dos classificadores.

Vale considerar que a placa também tem bornes para encaixe de fios elétricos, que também faria parte do experimento proposto, onde um fio com uma região descontínua (corroída) seria utilizado, como mostrado na *Figura 9*. A partir de testes para verificar a viabilidade de tal simulação, foi decidido retirá-lo do experimento pelas novas variáveis envolvidas e problemas encontrados. O primeiro é que o calor do fio reflete na placa elétrica; é um problema que pode ser tratado por meio dos algoritmos de processamento de imagens, porém aumentaria a complexidade. Além disso, a descontinuidade do fio deveria provocar um maior aquecimento, pela maior resistência encontrada, porém, devido a espessura dos fios, não foi possível provocar tal fenômeno. O fio é fino o suficiente para que a câmera térmica não consiga capturar o calor em tal região de descontinuidade.

Figura 9: Tentativa de simulação com fios elétricos.



Fonte: Autor.

4.2.1.2 Materiais e equipamentos

Para a montagem da bancada de testes, realização dos experimentos e medições elétricas, além da placa, são utilizados componentes elétricos e equipamentos que permitiram simulação e observação das diferentes condições, como fios elétricos, multímetro e fonte de tensão, além da câmera térmica. A tabela a seguir detalha o material/equipamento, e a maneira como serão utilizados para a execução dos experimentos.

Tabela 4: Materiais e equipamentos a serem utilizados nos experimentos (continua)

Item	Utilização
Fios elétricos	São utilizados para a transmissão de energia de uma fonte elétrica para a placa de circuito impresso.
Placa de circuito impresso	Contém as trilhas de cobre onde é realizada a análise térmica, bem como bornes para a conexão elétrica com uma fonte de tensão.
Multímetro	Permite medir a corrente e tensão elétrica no elemento. Essas informações são importantes na simulação de problemas elétricos e correlação com as temperaturas atingidas pelas trilhas.

Tabela 5: Materiais e equipamentos a serem utilizados nos experimentos (continuação)

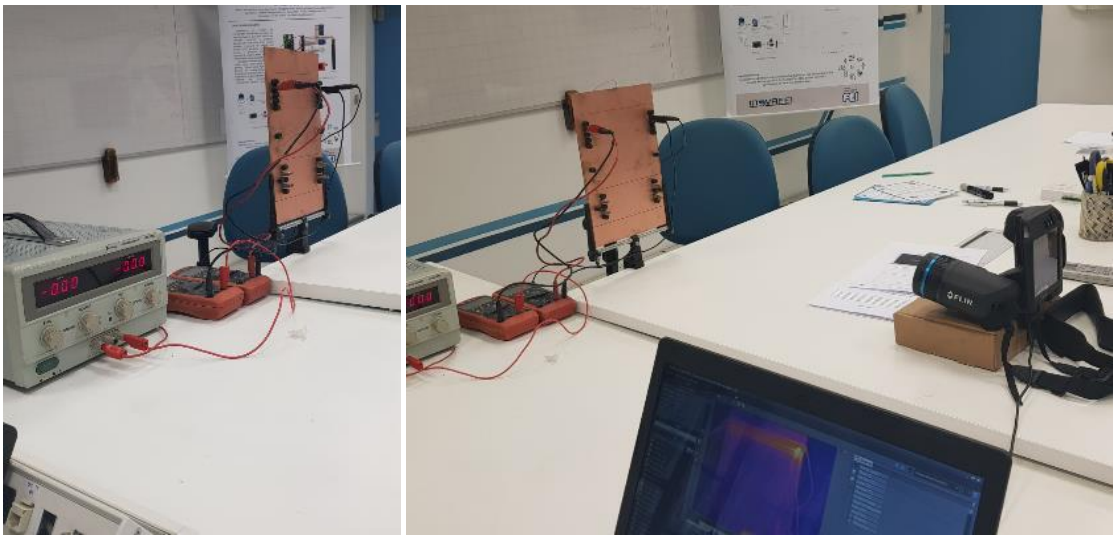
Item	Utilização
Fonte de tensão 30V, 5A	Bastante utilizado nas bancadas de testes permite o fornecimento de energia AC/DC, controlando tensão e corrente. Para o experimento, o dispositivo a ser utilizado permite converter uma tensão de entrada de 220V para uma de saída entre 0-30V, e corrente de saída de até 5A. São utilizados dois para obter uma corrente de até 10A.
Câmera térmica FLIR T530 SC	A câmera permite coletar o perfil térmico das trilhas da placa, uma malha de temperatura com os pontos mais quentes e mais frios das trilhas e todo o sistema elétrico. As imagens térmicas capturadas são utilizadas para realizar a detecção de problemas elétricos. O modelo utilizado possui resolução de 320x240 pixels e sensibilidade térmica de 30 mK.

Fonte: Autor

4.2.1.3 Bancada de testes

Uma típica bancada de testes consiste na utilização de um circuito elétrico, alimentado por uma fonte de tensão de 0-227v. Todas as trilhas possuem bornes em suas extremidades e permitem o fornecimento de energia e medição elétrica por meio de um multímetro. Os dois terminais ligados em paralelo com a placa permitem a medição de tensão, enquanto dois terminais ligados em série com o circuito serão utilizados para medição da corrente.

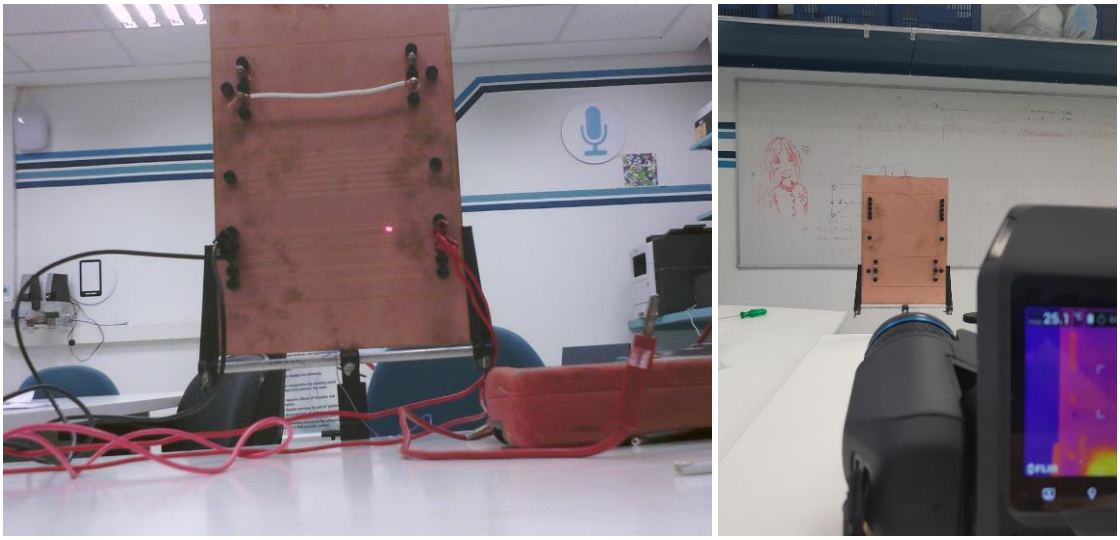
Figura 10: Mesa de testes



Fonte: Autor.

Para o experimento proposto, os bornes das seis trilhas, um de cada vez, serão alimentados por uma fonte de alimentação, variando a corrente fornecida. Como indicado na *Figura 11*, a câmera fica posicionada frontalmente com a placa, com algumas pequenas variações.

Figura 11: Mesa de testes



Fonte: Autor.

4.2.2 Execução do experimento piloto

Para a execução do experimento é necessário conhecer as variáveis envolvidas e entender como elas podem afetar no resultado dos testes. Os parâmetros elétricos e de temperatura devem ser obtidos para estudar os resultados da simulação, bem como coletar dados relevantes para os algoritmos do método proposto:

- Tensão e corrente: medidos e calculados através de um multímetro;
- Temperatura das trilhas e do ambiente: obtidos através da câmera térmica.

Na execução dos experimentos, propõe-se a simulação dos problemas elétricos mais encontrados e pertinentes. Conforme visto nos trabalhos em termografia e energia elétrica, e apontado por Mostafavi et al. (2012), a maior parte dos problemas elétricos encontrados por termografia se dão pelo fornecimento impróprio de energia ou problemas de contato. Na tabela a seguir, esses problemas são apresentados, com o procedimento para serem simulados nos experimentos e os prováveis efeitos no sistema elétrico proposto.

Tabela 6: Problemas elétricos a serem simulados

Problema	Procedimento	Efeito esperado
Sobretensão / sobrecorrente	• Variação gradual da corrente de entrada • Prover corrente acima do limite recomendado	Aquecimento dos fios, aquecimento das trilhas e contatos acima do normal
Problemas de contato ou corrosão	• Provocar alta resistividade entre os contatos dos condutores • Simula-se condutores com pouco contato, por meio de desníveis nas trilhas da placa	Sobreaquecimento em pequenas regiões, onde ocorre o contato dos condutores

Fonte: Mostafavi et al. (2012), Navani (2013), Seymour e Horsley (2008), Suriyamongkol (2002).

Para a simulação do problema de sobrecorrente, é necessário saber qual a corrente máxima que a placa deve suportar, e qual corrente representa uma situação de sobrecorrente. Uma corrente de falha é um fenômeno caracterizado por ser uma corrente acima do normal, além do que o sistema deve suportar, por algum intervalo de tempo. Um processo utilizado para analisar esse tipo de corrente é a integral de Joule, muito utilizada para a análise dos efeitos térmicos gerados pelas correntes de falta.

De acordo com o livro de Filho (2007) e regulamentado pela norma NBR 5410/2004 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2004), é definido que para o sistema estar em condições normais (seguro), a integral de Joule não pode ser superior à integral necessária para aquecer o condutor da sua temperatura máxima até a temperatura limite de curto-circuito. A norma também possui uma simplificação, válida somente para correntes de falha que durem de 0,1 a 5 segundos, como é caso deste trabalho, onde é valido a seguinte fórmula:

$$I^2_{cs} \, . \, T \leq K^2.S^2$$

Sendo $K^2.S^2$ a integral de Joule para aquecimento do condutor desde sua temperatura máxima, onde K é uma constante definida na norma, de acordo com o material do condutor e S é seção do condutor em mm²; I_{cs} é a corrente de falta; e T é o tempo de duração da corrente de falta (de 0,1 a 5 segundos).Para esse trabalho, os valores a serem atribuídos são:

- I = 10A
- K = 115
- S = 2mm de espessura, 1,25 mm de largura = 2,5mm²
- T = (entre 1 e 5 segundos)

Conforme apresentado com Filho (2007), já foram calculados e tabelados os valores para a integral de Joule $K^2.S^2$ para condutores de cobre, como o utilizado nesse trabalho, para diversos tamanhos de seção (mm²). Para 2,5mm², a integral de Joule é de 82,6. Portanto, $I^2.T$ deve ser menor que 82,6, ou seja, $I^2 \leq 82,6 / T$. Escolhendo arbitrariamente o tempo como 1, e como 5 segundos, temos que:

- $I^2 \leq 82,6 / 1$, portanto $I \leq 9,08A$
- $I^2 \leq 82,6 / 5$, portanto $I \leq 4,06A$

Portanto, conforme os resultados, se o tempo de duração de sobrecorrente for de 1 a 5 segundos, devemos aplicar uma corrente superior a 9,08A para que esta seja considerada uma sobrecorrente (corrente de falta), ou seja, exceder o que é regulamentado pela norma NBR 5410/2004. Assim, será utilizado o limite de 10A para caracterizar tal fenômeno.

4.2.2.1 Simulação do problema de sobrecarga/sobrecorrente

A simulação de sobrecarga/sobrecorrente busca representar condições impróprias de fornecimento de energia, que podem causar problemas elétricos e danificar equipamentos. A seguir, são apresentadas as condições para execução da simulação, permitindo sua reprodutibilidade:

- Temperatura ambiente: entre 15 e 25°C
- Tensão: 5v;
- Corrente inicial: 0A; corrente final: 10A;

O procedimento para a execução do experimento é descrito nos passos a seguir, simulando o problema de sobrecorrente:

i. Ligar lado negativo e positivo da fonte de tensão em uma das trilhas da placa;
ii. Pontas de prova do multímetro posicionada nos bornes e realizando a medição;
iii. Temperatura do ambiente ajustada;
iv. Câmera térmica posicionada de frente com a placa;
v. Ligar circuito com fonte fornecendo tensão de 5v, corrente em 0;
vi. Capturar imagens continuamente até que a temperatura esteja estável;
vii. Aumentar corrente em 1A e repetir dos passos ***v*** a ***viii***, até que se alcance 2A.
viii. Aumentar a corrente subitamente para 10A, e reduzir para 0A, capturando as imagens em todo o processo.

Os passos devem se repetir ligando a fonte nas diferentes trilhas na placa. Para o problema de sobrecorrente, foram coletadas imagens nas trilhas 1, 3 e 6, além da combinação de duas trilhas ligadas ao mesmo tempo, a trilha 1 e 3. Os passos de *v* a *vii* permitem a coleta de imagens das trilhas em condições normais, enquanto a o passo *viii*, de aumento súbito de corrente, representa um fenômeno similar ao de sobrecorrente nas linhas de transmissão.

Espera-se, com essa simulação, perceber condições elétricas indevidas através do calor acima do usual apresentado pela placa.

*4.2.2.2 Simulação do problema de contato*s elétricos

A simulação de problemas de contatos busca representar condições em que há alta resistividade entre os condutores, por algum defeito ou falha de instalação. A seguir, são apresentadas as condições para execução da simulação, permitindo sua reprodutibilidade:

- Temperatura do ambiente: entre 15 e 25°C
- Tensão: 5v;
- Corrente inicial: 0A; corrente final: 4A;

O procedimento para a execução do experimento é descrito nos passos a seguir, simulando o problema de contato:

i. Ligar lado negativo e positivo da fonte de tensão em uma das trilhas da placa;
ii. Pontas de prova do multímetro posicionada nos bornes e realizando a medição;
iii. Temperatura do ambiente ajustada;
iv. Câmera térmica posicionada de frente com a placa;
v. Ligar circuito com fonte fornecendo tensão de 5v, corrente em 0;
vi. Capturar imagens continuamente até que a temperatura esteja estável;
vii. Aumentar corrente em 1A e repetir dos passos ***v*** a ***viii***, até que se alcance 4A.

Os passos devem se repetir ligando a fonte nas diferentes trilhas na placa. Para o problema de mal-contato, são coletadas imagens nas trilhas 2, 4 e 5, além da combinação de duas trilhas ligadas ao mesmo tempo, a trilha 3 (que não apresenta falha) com a trilha 4. Além disso, os passos devem se repetir para as trilhas 1 e 3 individualmente, para obtenção de imagens onde não há problema, que posteriormente também serão utilizadas na validação da metodologia.

Espera-se, com essa simulação, perceber a produção de calor fora do usual em alguns pontos do sistema. Embora não afete diretamente a placa e provoque danos, a alta resistividade

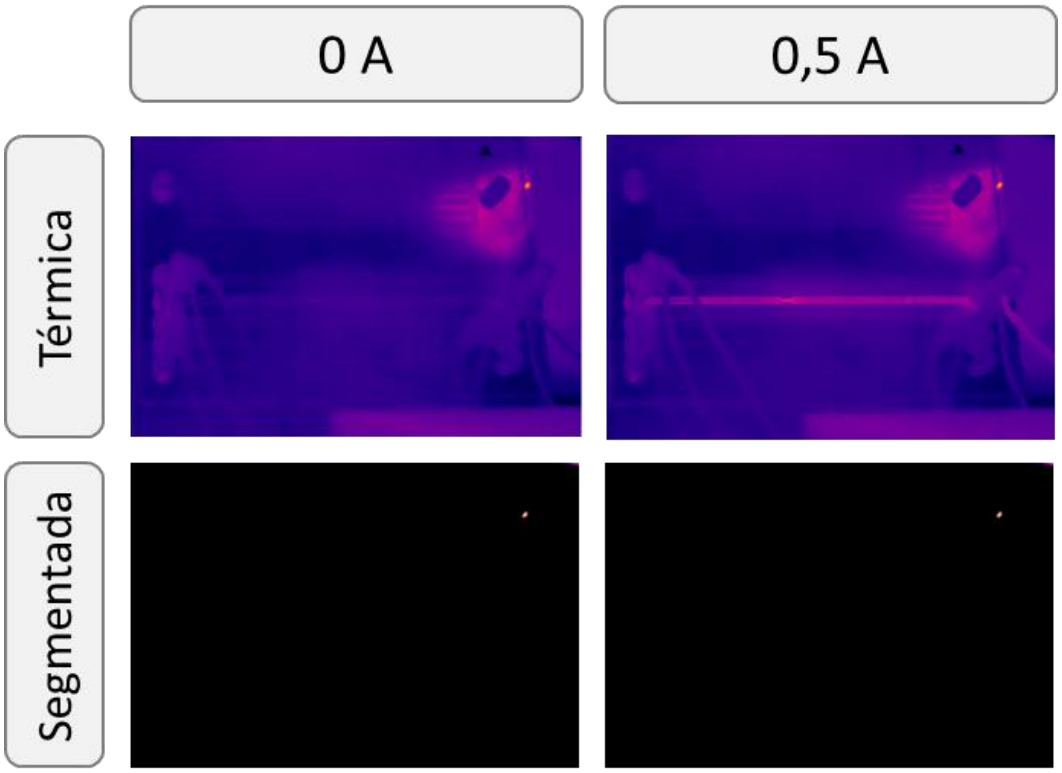
nesses pontos é um problema muito comum responsável por grandes perdas de energia em sistemas elétricos reais.

4.2.3 Validação das imagens coletadas

A simulação de problemas de mal contato busca representar condições em que há um maior calor produzido na região descontínua, justamente um problema de contato. Para isso, é necessário validar se a simulação realmente possibilita que a câmera térmica obtenha imagens que indiquem a possibilidade de tal falha.

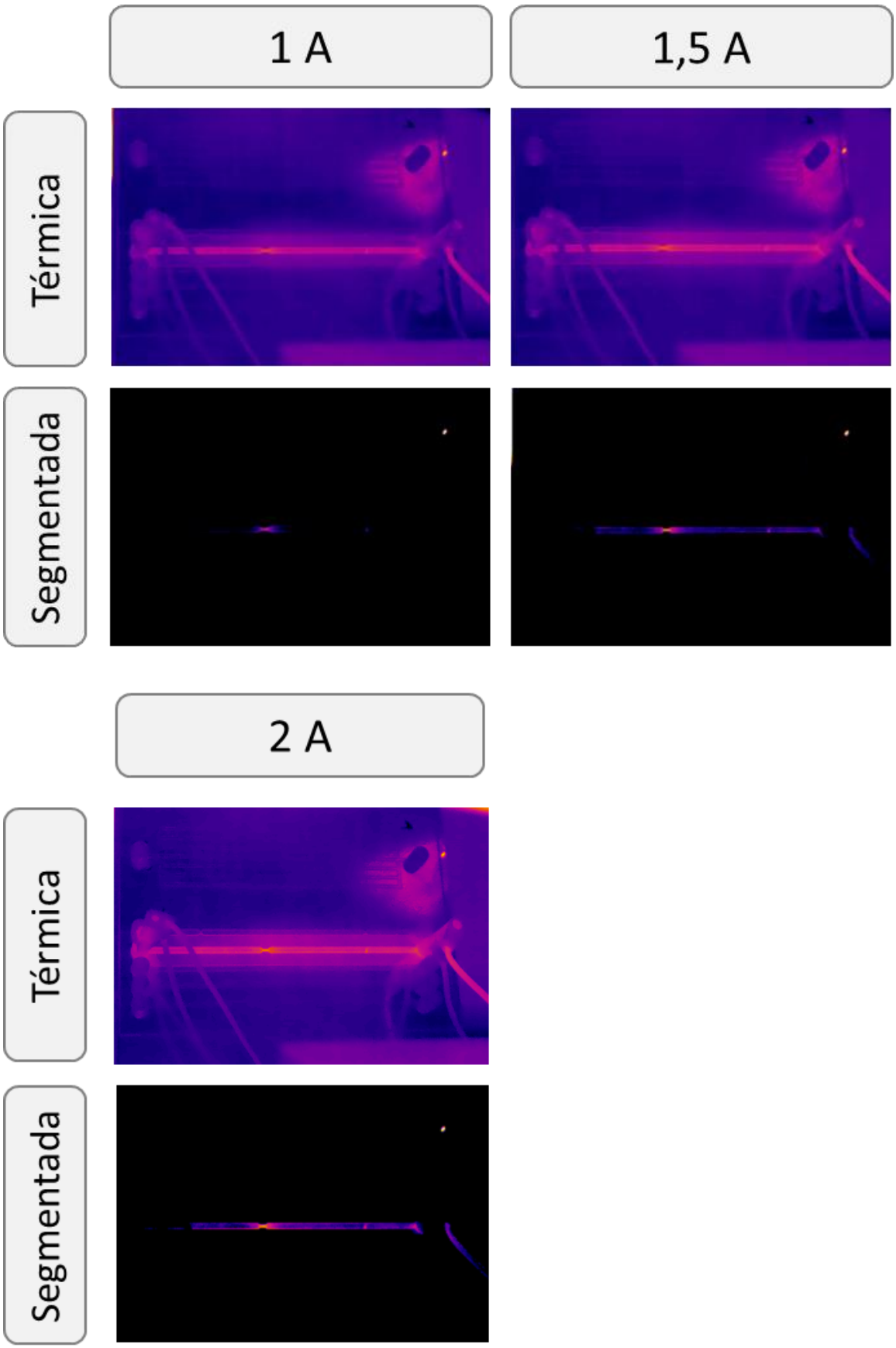
Como apresentado na *Figura 12* e *Figura 13*, é possível observar na imagem térmica, que após o aumento de corrente (começando de 0A, e sendo acrescido de 0,5A), a trilha inteira esquenta, porém com um maior aquecimento na região onde há a falha. Na imagem segmentada (a partir de um limiar definido de forma manual), é possível perceber que a partir de 1A começa-se a evidenciar a região de mal-contato, pois ela possui mais calor em relação ao restante da imagem. Com 1,5A e 2A, ela se torna ainda mais evidente. Com isso, é valido que um algoritmo de classificação também seja capaz de identificar tal falha, dado as variáveis corretas e exemplos de classificação suficientes ao classificador.

Figura 12: Validação da imagem térmica a ser classificada (parte 1)



Fonte: Autor.

Figura 13: Validação da imagem térmica a ser classificada (parte 2)



Fonte: Autor.

5 RESULTADOS

A partir da metodologia de pesquisa e experimento propostos, se deseja validar os métodos, técnicas e algoritmos propostos para os casos em energia elétrica, entender suas condições de uso e realizar uma análise comparativa da acurácia dos algoritmos de inteligência artificial e sua capacidade de classificação. Por meio do experimento proposto, foi possível compor um banco de dados para as seguintes condições elétricas:

- Sobrecarga (corrente acima do suportado);
- Problemas de mal-contato.

A base de dados foi composta por imagens coletadas por meio da câmera térmica. Essas imagens são filtradas, onde aplica-se a segmentação (recorte) e tratamento de imagem manual, por meio do software FLIR tools, que posteriormente são comparadas com o processamento de imagens automático.

A validação da metodologia se dá pela análise dos resultados obtidos pela combinação de métodos de processamento de imagens e classificadores na detecção desses problemas. Os métodos de processamento de imagens utilizados foram a segmentação manual (humana), limiarização por *otsu* e clusterização por *k-means*. Classificadores foram treinados com o conjunto de entrada para obtenção de modelos que generalizem os casos de estudo e possam indicar potenciais problemas elétricos. Os classificadores utilizados foram *k-NN*, SVM (linear e gaussiano), redes neurais, árvores de decisão (CART), *Random Forest* e *Naive Bayes*, bastante presentes na literatura e trabalhos correlatos.

5.1 CAPTURA DE IMAGENS

As informações a serem utilizadas na composição da base de dados foram obtidas por meio da captura de imagens térmicas pela câmera FLIR T530. O processo consiste em obter uma sequência termográfica durante todo o experimento, onde se faz uma gravação de imagens térmicas, capturando 30 frames por segundo, conforme a capacidade do dispositivo. As imagens foram etiquetadas de acordo com a classe a que pertencem, caso normal, de sobrecorrente ou mal-contato. A base de dados, além das imagens térmicas, também inclui a largura da trilha, as medidas elétricas obtidas através do multímetro, e as medidas termográficas associadas à câmera térmica (como emissividade). O conjunto formado por uma imagem térmica e tais parâmetros consiste em uma amostra, com etiquetas quanto ao problema simulado.

O experimento foi executado como detalhado na Seção 4.2, na placa projetada para as simulações de problemas elétricos. Foram realizadas duas simulações, para problemas de mal-contato (Simulação 1) e problemas de sobrecorrente (Simulação 2), como indicado na *Tabela 6*. Em ambas as simulações, foram coletadas imagens da placa em condições elétricas normais, bem como do problema simulado. As trilhas foram selecionadas de acordo com o descrito na Seção 4.2, sendo as trilhas 1, 3 e 6 sem descontinuidades, e as trilhas 2, 4 e 5 com descontinuidades, propícias para as simulações de mal-contato.

A quantidade de simulações realizadas, conforme apresentado na *Tabela 6,* não foi planejado de início. A execução das simulações continuou a ser executada até que fosse possível se obter uma quantidade de imagens suficientes para compor um *dataset*, permitindo gerar classificadores válidos. Números inferiores aos apresentados impediram que os classificadores generalizassem bem, trazendo resultados enviesados e de baixa acurácia.

Tabela 7: Simulações executadas em cada trilha

Simulação	Quantidade	Condição	Trilhas	Total de imagens
Simulação 1	34 experimentos	Normal	Trilha 1 Trilha 3 Trilha 6	2550
		Mal-contato	Trilha 2 Trilha 4 Trilha 5 Trilhas 4 e 5	3400
Simulação 2	42 experimentos	Normal	Trilha 1 Trilha 3 Trilha 6 Trilha 1 e 3	3360
		Sobrecorrente		1680

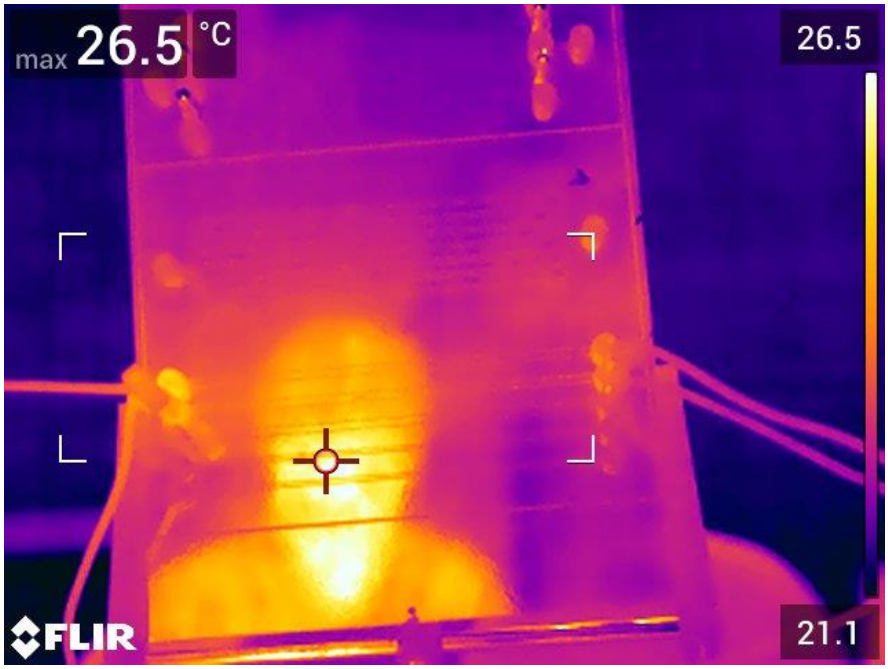
Fonte: Autor

A princípio, antes das simulações serem executadas efetivamente para a coleta das imagens do *dataset,* testes preliminares foram realizados com a placa confeccionada para o experimento. A placa se portou bem às duas simulações e não foi danificada, sendo suficiente para a realização de todas as coletas de imagem nas simulações deste trabalho.

Durante os testes, foi possível perceber que o calor na trilha se estabilizava em média após 5 segundos depois da aplicação de uma faixa de corrente. Assim, no processo de captura,

a cada cinco segundos após a mudança de correte, foi possível capturar imagens válidas para classificação. Um outro detalhe importante é que a placa de cobre reflete o calor radiado por outras fontes de calor, como indicado na *Figura 14*. Normalmente a temperatura é bastante inferior às trilhas e pode ser desprezada. Porém, quando a corrente aplicada é baixa (0 ou 0,5A), a temperatura das trilhas é baixa, e o reflexo do calor humano por exemplo fica bastante evidente, se tornando um ruído que pode atrapalhar na classificação. Essa é uma dificuldade que foi encontrada e que não foi possível tratar pelo método proposto, e, portanto, para a coleta de imagens foi necessário isolar o experimento de qualquer outra fonte de calor.

Figura 14: Ruído provocado por outras fontes de calor



Fonte: Autor

Outro ruído que pode influenciar na classificação e não foi possível tratar foram os bornes e os fios elétricos que alimentam a placa, como indicado na *Figura 15*. Eles aparecem nas imagens coletadas e esquentam conforme há o aumento da corrente. Uma forma de amenizar tal efeito foi garantir que os bornes e os fios estavam muito bem encaixados, de forma a reduzir qualquer resistência causada nas regiões de contato.

Figura 15: Ruído provocado pelo mal-contato entre o borne e o fio

Fonte: Autor

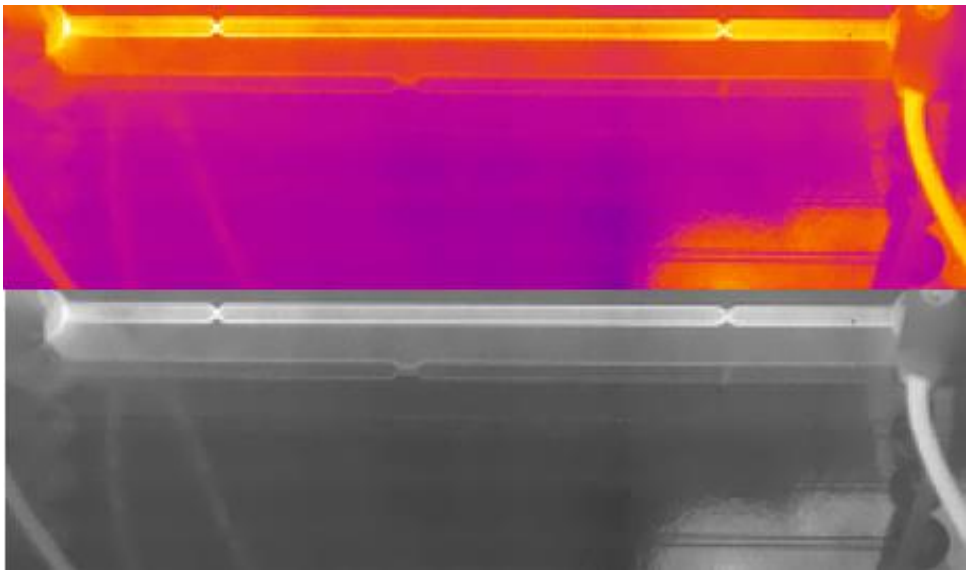
Após estes testes preliminares e adequação dos experimentos, as simulações para coleta de imagens se deram início, conforme detalhado na *Tabela 6* acima. Na Simulação 1, são capturadas imagens em 5 faixas de corrente, de 0A a 4A. Para cada captura, foram selecionadas 5 imagens, considerando que a maior parte delas é redundante. Assim, nos 34 experimentos realizados, se obteve 5x5x34 = 850 imagens. Para o caso de mal-contato, essas 850 são obtidas para as 4 trilhas onde há o problema, como indicado na *Tabela 6*. Assim, 850 x 5 = 3400 imagens obtidas. Também são obtidas imagens onde não há problema, sendo as 850 imagens coletadas de 3 trilhas, 2550 imagens. Cada simulação levou em torno de 20 minutos para ser executada. O processo é análogo para as outras simulações. No caso da Simulação 2, são apenas 3 faixas de corrente, sendo duas deles para a condição normal, e uma delas para o caso de sobrecorrente (10A). Para cada captura, selecionou-se 10 imagens, resultando em 1680 imagens de sobrecorrente e 3360 imagens para o caso normal. Nas duas simulações se obteve imagem da condição normal, que foram utilizadas a seguir para treinar o classificador para o caso onde não há problema. Por fim, a base de dados é filtrada para que esteja balanceada entre as condições, para se evitar um treinamento enviesado. Para isso, são utilizadas na totalidade as 5080 amostras dos casos de mal-contato e sobrecorrente, e selecionadas 5080 amostras do caso normal.

Com a composição da base de dados, o próximo passo é o de processamento de imagem, onde operações são aplicadas sobre as imagens para extração de informações que contribuirão na posterior classificação.

5.2 PROCESSAMENTO E EXTRAÇÃO DE CARACTERÍSTICAS

Após a captura, o primeiro passo do tratamento da imagem térmica consiste na conversão da imagem para *grayscale*, permitindo sua segmentação. Devido a resolução da câmera, não foi necessário nenhum melhoramento de contraste.

Figura 16: Exemplo de imagem térmica utilizada no processamento, pré-processada em *grayscale*

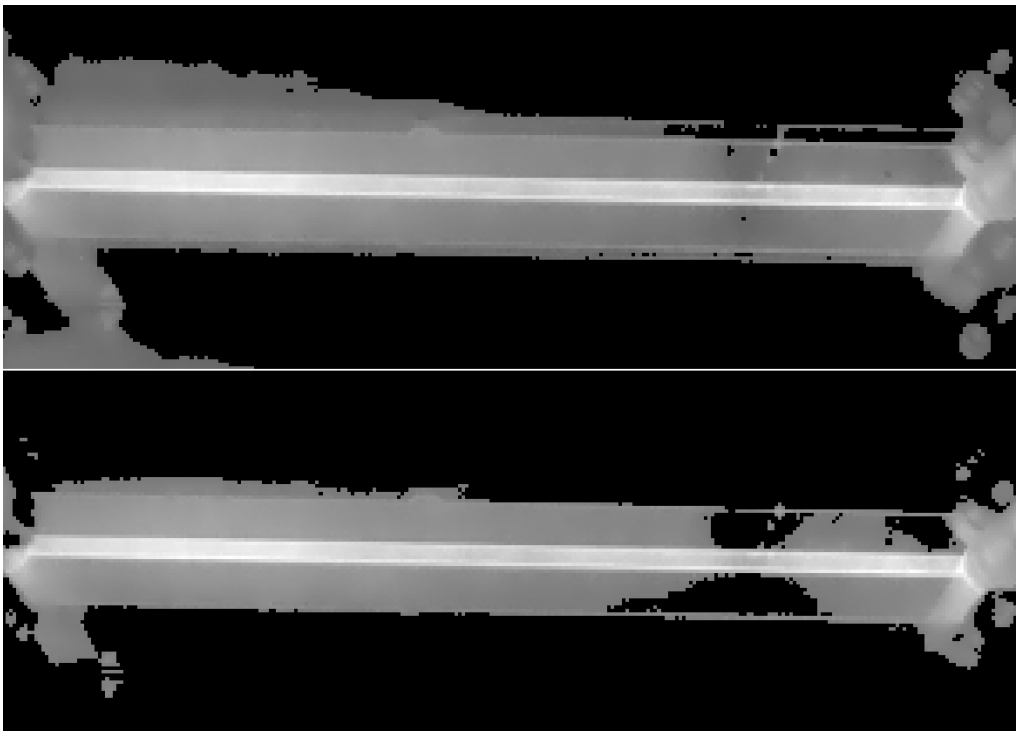


Fonte: Autor

O método a ser verificado para a segmentação foi a limiarização por *otsu*, que é uma técnica simples, de fácil implementação, e eficaz para o proposto de falhas elétricas e imagens térmicas, já que apresentam diferenças razoáveis de temperatura. *K-means* é um método mais custoso computacionalmente, porém também bastante utilizado em aplicações de segmentação de imagem térmica, portanto também foi incluso na validação da metodologia. O objetivo dessa etapa se dá em selecionar as regiões mais quentes da imagem, possivelmente as trilhas da placa e outros componentes elétricos envolvidos, extraindo o plano de fundo das imagens das regiões de calor. Além das duas imagens, segmentadas por *otsu* e *k-means*, a imagem original sem segmentação também foi utilizada. Essa é uma imagem que contém ruídos e interferências externas, caso importante para verificar o comportamento dos classificadores. Exemplos de

imagens segmentadas por tais algoritmos são apresentados na *Figura 17* para o caso normal, e na *Figura 18* para o caso de problema elétrico.

Figura 17: Imagem térmica do caso normal, com *otsu* (superior) e *k-means* (inferior)



Fonte: Autor

Figura 18: Imagem térmica do caso com mal-contato, com *otsu* (superior) e *k-means* (inferior)

Fonte: Autor

Antes de realizar a análise dos classificadores, é necessário completar a base de dados e exemplos de treinamentos utilizados com a extração de características das imagens térmicas. Como proposto em 4.1.2, pode-se extrair parâmetros estatísticos e relativos à termografia. As principais características que foram extraídas são apresentadas na *Tabela 7*. Enquanto os parâmetros térmicos foram obtidos por meio da análise de temperatura das regiões de calor em comparação com outras regiões ou com o plano de fundo, os parâmetros estatísticos foram extraídos por meio da análise do histograma da imagem ou da sua matriz de correlação em escala de cinza (GLCM).

Tabela 8: Parâmetros estatísticos e térmicos extraídos da imagem processada

Parâmetros estatísticos baseados no histograma da imagem	Parâmetros térmicos
• Intensidade media • Mediana da intensidade • Desvio padrão • Assimetria • Curtose • Energia • Entropia	• Temperatura máxima • Temperatura mínima • Temperatura máxima da região de interesse relativa ao ambiente • Temperatura média da região de interesse relativa ao ambiente • Tamanho da região mais quente (em pixels)
Parâmetros estatísticos baseados na matriz de co-ocorrência em escala de cinza (GLCM)	
• Contraste • Correlação • Entropia	

Fonte: Autor

O procedimento até o momento, antes da classificação, gerou um banco de dados com imagens térmicas, sua segmentação e características relevantes. Esses dados foram etiquetados e assim prontos para utilização. Os dados presentes na base de dados são:

- Imagem térmica;
- Delimitação da região de interesse;
- Vetor de características: parâmetros estatísticos e térmicos da imagem:
- Medições elétricas;
- Etiquetas/classes: condição elétrica, trilha.

Para o teste dos classificadores, como executado a seguir, essa etapa é fundamental. Caso a segmentação seja falha, obtendo regiões maiores do que a de interesse, não obtendo a

região de interesse correta ou incluindo ruídos, a precisão da classificação pode cair, já que é um caso onde não haverá informações suficientes ou corretas, de acordo com o modelo formulado. Para os primeiros testes do método, propõe-se a verificação da extração da região de interesse de forma manual. Verifica-se todas as imagens, de forma visual, observando se as delimitações de regiões de calor são realmente as mais quentes e de pontos críticos. Caso os limites estejam fora ou incluindo muito ruído, eles devem ser ajustados. Após a validação dos algoritmos de classificação, decisão dos melhores parâmetros e características, o processamento pode ser feito de forma totalmente automática, conforme inicialmente proposto, onde se validará os classificadores para as imagens segmentadas por *otsu* e *k-means*, sendo possível assim realizar a análise dos resultados e da acurácia obtida, algo que possibilitará investigar e discutir o quanto prováveis ruídos e *outliers* (casos fora da curva) podem influenciar nesses algoritmos.

5.3 CLASSIFICAÇÃO E RESULTADOS

Vale considerar que a grande maioria dos trabalhos correlatos trabalham somente com a métrica acurácia. De 16 trabalhos apresentados na Tabela 1, de trabalhos em energia elétrica e termografia, apenas 2 utilizam métricas adicionais, além da acurácia (*recall* e precisão). Com o objetivo de realizar a análise comparativa dos principais algoritmos e técnicas propostos, utilizando um método geral baseado neles, este trabalho também realizará a análise por meio da acurácia. A maioria dos trabalhos na área consiste em solucionar um problema, onde apenas fazem uso de algoritmos de IA e *machine learning*, sem precisamente fazer uma análise mais profunda dos algoritmos e seus resultados. Os resultados, de acordo com os autores, foram satisfatórios, indicando que a métrica de acurácia em tais casos foi suficiente (ALMEIDA et al., 2009; CHOU; YAO, 2009; DE OLIVEIRA; LAGES, 2010; LI et al., 2006; MOSTAFAVI; MIRMAJLESSI; FATHOLLAHI, 2012; PAPEŽ; PALFY; TURK, 2008; RAHMANI; HADDADNIA; SERYASAT, 2010; WONG et al., 2009; ZARKOVIC; STOJKOVIC, 2015).

Na escolha das métricas para análise da classificação, é importante destacar também o problema de desbalanceamento de classes. Por exemplo, um classificador que sempre classifique como negativo, e uma base de dados extremamente balanceada, onde a maioria dos exemplos são negativos, resultará num classificador com alta acurácia, que errará apenas os poucos exemplos positivos. Nesse caso, métricas como precisão e recall são importantes, indicando a relação da acurácia com o conjunto de dados, sua distribuição, os falsos positivos e falsos negativos (ABD ELRAHMAN; ABRAHAM, 2013; ALI; SHAMSUDDIN;

RALESCU, 2015; GUO et al., 2009; JUBA; LE, 2017).. Enquanto precisão indica, dos exemplos classificados como corretos, quais realmente eram corretos, *recall* indica entre os exemplos de uma determinada classe, quais realmente eram daquela classe. Essas métricas tendem a variar bastante da acurácia quando há desbalanceamento de classes. Quando a base de dados é balanceada, porém, a acurácia é uma métrica relevante pois traz a taxa de acerto entre os exemplos positivos e negativos, indicando o quão bem classifica exemplos de diferentes classes, uma medida significativa quanto à taxa de acerto do modelo (ABD ELRAHMAN; ABRAHAM, 2013; ALI; SHAMSUDDIN; RALESCU, 2015; GUO et al., 2009; JUBA; LE, 2017).

Validando a metodologia, os algoritmos de classificação utilizados foram redes neurais *(multilayer perceptron),* SVM (em duas versões, linear e com kernel de função de base radial), *k-NN*, *Naive Bayes* e árvores de decisão (CART) e *Random Forest*, por meio de uma ferramenta que permite esse processamento e análise de dados, no caso MATLAB. O ambiente escolhido provê um *framework* completo para treinamento, teste e validação, com tais métodos já implementados, permitindo a aplicação da metodologia. Há algumas decisões a serem tomadas, como os parâmetros envolvidos em cada um desses algoritmos. Na rede neural, deve-se escolher o número de neurônios na camada oculta e a taxa de aprendizado; no classificador SVM, deve-se definir o parâmetro de custo C, que permite uma flexibilidade entre permitir erros de treinamento e forçar margens rígidas, bem como a largura do kernel gaussiano; no classificador *k-NN*, deve-se escolher o número k de vizinhos e a métrica de distância. Os parâmetros foram selecionados empiricamente, onde a partir de valores iniciais, foram sofrendo ajustes até se encontrasse a melhor acurácia de classificação. Eles são descritos a seguir:

- Rede Neural: 10 neurônios na camada escondida
- SVM com kenel linear: C=4,5; Escala = 1
- SVM com kernel gaussiano: C=5,2; Escala = 1
- *k-NN*: k (vizinhos) = 5
- Árvore de Decisão: 100 divisões
- *Random Forest*: 100 divisões

A partir dessa parametrização, os algoritmos foram utilizados nas etapas de treinamento e validação do modelo. Um modelo é obtido por meio do treinamento com os exemplos obtidos na base de dados, para cada um dos classificadores utilizados. Com a validação cruzada, se valida o modelo obtido e seu desempenho para dados ainda desconhecidos, pelo método *k-fold*. Como a base de dados está balanceada, é possível que os *k* subconjuntos também esteja balanceado, com a proporção de exemplos de cada classe mantida. Os *k* classificadores

aprendidos tem cálculo de erro e acurácia agregados, obtendo uma medida confiável a respeito da capacidade de representação e generalização do modelo. O conjunto obtido com os experimentos de mal-contato foram de 5100 imagens, metade dos casos de mal-contato, e outra metade com casos normais, ambos com as características térmicas e estatísticas extraídas. Os classificadores foram treinados com 2550 desses exemplos, com validação cruzada por k-fold, sendo k=5. Na fase de testes, os classificadores foram testados com os 2550 exemplos restantes, com os resultados apresentados nas tabelas *Tabela 8* e *Tabela 9*. Vale considerar que modelos foram obtidos separadamente para os diferentes conjuntos de características extraídas (térmicas e estatísticas), bem como para os diferentes métodos de segmentação propostos (manual, limiarização por *otsu* e *k-means*), com o propósito de se avaliar a metodologia para as diferentes técnicas e algoritmos propostos.

Tabela 9: Validação dos algoritmos de classificação

Algoritmo	Conjunto de características	Segmentação manual	Segmentação por *otsu*	Segmentação por *K-means*
k-NN	Térmicas	0,995433	0,992289	0,989894
	Estatísticas	1	1	0,999202
	Ambas	0,998925	0,999202	0,999468
Árvore de decisão	Térmicas	0,974475	0,964371	0,959053
	Estatísticas	1	1	1
	Ambas	0,987769	0,999734	1
SVM Linear	Térmicas	0,99597	0,988035	0,983246
	Estatísticas	1	0,999202	0,996012
	Ambas	1	0,999202	0,995744
SVM Gaussiano	Térmicas	0,997044	0,989099	0,985639
	Estatísticas	1	1	0,998936
	Ambas	0,998936	0,998671	0,998936
Naive Bayes	Térmicas	0,799787	0,77825	0,598777
	Estatísticas	0,633874	0,742888	0,741824
	Ambas	0,863334	0,828237	0,689178
Rede Neural	Térmicas	0,997	0,857	0,986
	Estatísticas	0,999	1	1
	Ambas	1	1	1
Random Forest	Térmicas	0,992555	0,985908	0,9827
	Estatísticas	1	1	1
	Ambas	1	1	1

Fonte: Autor

Em geral, os resultados da classificação dos modelos obtidos após o treinamento apareceram com em torno de 99% de acurácia, com apenas algumas exceções. Tirando o *Naive Bayes*, com valores a baixo da média, os algoritmos obteram taxas de acerto elevadas,

principalmente as *Random Forest* e redes neurais, com 100% de acerto em diversos casos, de características utilizadas e de segmentação, representando uma ótima capacidade de assimilar os dados a partir do modelo treinado. É interessante observar que Naive Bayes é um algoritmo que exige menos recursos computacionais, é bastante eficiente, porém desconsidera a relação entre as variáveis de entrada, assumindo-as como independentes uma das outras. Esses resultados abaixo da média indicam que para o caso de imagens térmicas em problemas elétricos, há relações a serem consideradas e por isso, algoritmos do tipo podem não ter bons resultados. Na tabela a seguir, se verifica a capacidade desses modelos lidarem com dados que não foram utilizados no treinamento.

Tabela 10: Resultados dos algoritmos de classificação

Algoritmo	Conjunto de características	Segmentação manual	Segmentação por *otsu*	Segmentação por *K-means*
k-NN	Térmicas	0,811455	0,738578	0,676854
	Estatísticas	0,897248	0,899441	0,800414
	Ambas	0,741988	0,885611	0,885216
Árvore de decisão	Térmicas	0,821446	0,960463	0,922583
	Estatísticas	0,762062	0,823613	0,70497
	Ambas	0,797564	0,822256	0,818885
SVM Linear	Térmicas	0,774065	0,734322	0,776793
	Estatísticas	0,923372	0,813688	0,982171
	Ambas	0,866292	0,69828	0,887028
SVM Gaussiano	Térmicas	0,833885	0,77119	0,708157
	Estatísticas	0,705163	0,744655	0,985017
	Ambas	0,738445	0,799436	0,965636
Naive Bayes	Térmicas	0,542257	0,589238	0,340804
	Estatísticas	0,591419	0,660428	0,677157
	Ambas	0,66997	0,621474	0,613146
Rede Neural	Térmicas	0,748	0,71	0,72
	Estatísticas	0,812	0,641	0,85
	Ambas	0,863	0,834	0,86
Random Forest	Térmicas	0,95761	0,848522	0,939783
	Estatísticas	0,759383	0,889515	0,846108
	Ambas	0,753171	0,942377	0,904099

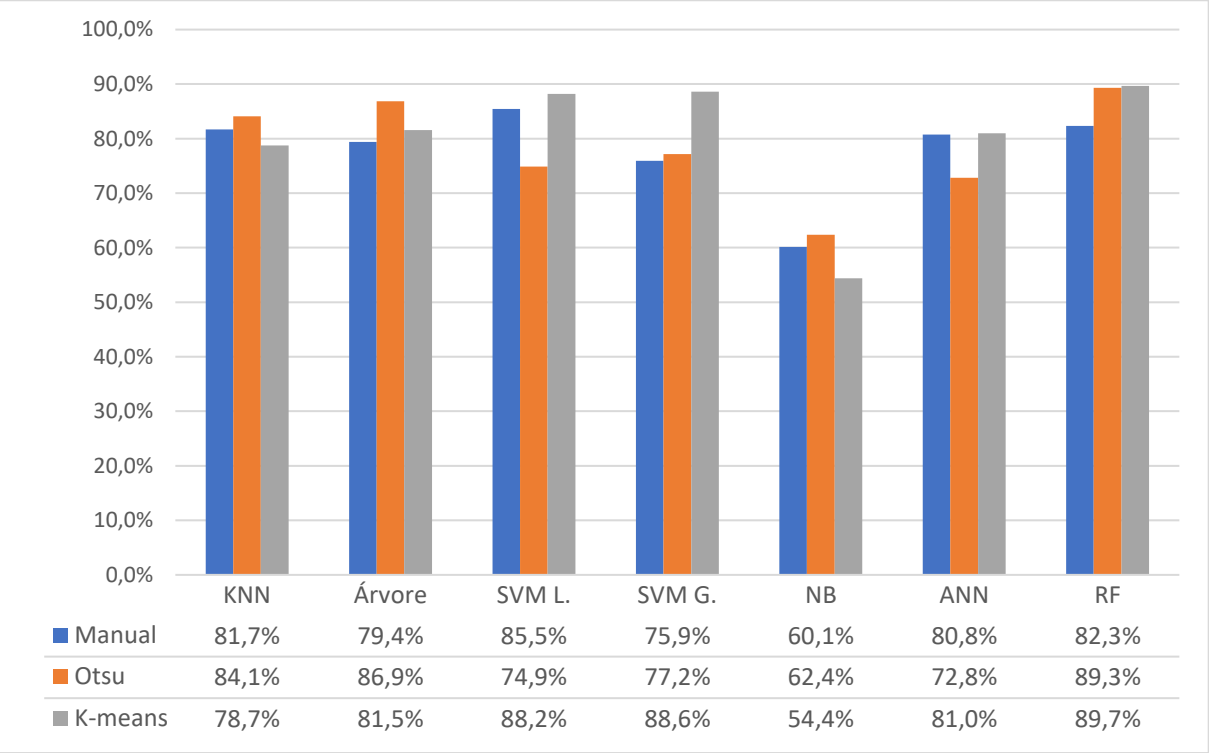
Fonte: Autor

A *Tabela 9* apresenta os resultados dos testes. Lidando com novos dados, percebe-se que houve uma queda. Contudo, tirando o *Naive Bayes*, os resultados foram acima de 70%, indicando que na maioria dos casos, os classificadores escolhidos identificam corretamente se há problema de mal-contato ou não. Algumas combinações de características e técnicas de processamento se mostraram mais eficientes para essa aplicação elétrica em específico, como

as características térmicas e segmentação por *otsu*, nas árvores de decisão, com 96% de acerto, assim como o SVM com kernel Gaussiano, para todas as características térmicas e estatísticas extraídas e segmentação por *k-means*, que também alcançou 96% de acurácia. As *Random Forest* com *otsu* e *k-means* também tiveram bons resultados, acima de 90%.

A seguir, são apresentadas as médias de acerto de cada classificador, sem considerar as características extraídas, apenas o tipo de segmentação que foi utilizado.

Figura 19: Teste dos modelos para os diferentes métodos de segmentação



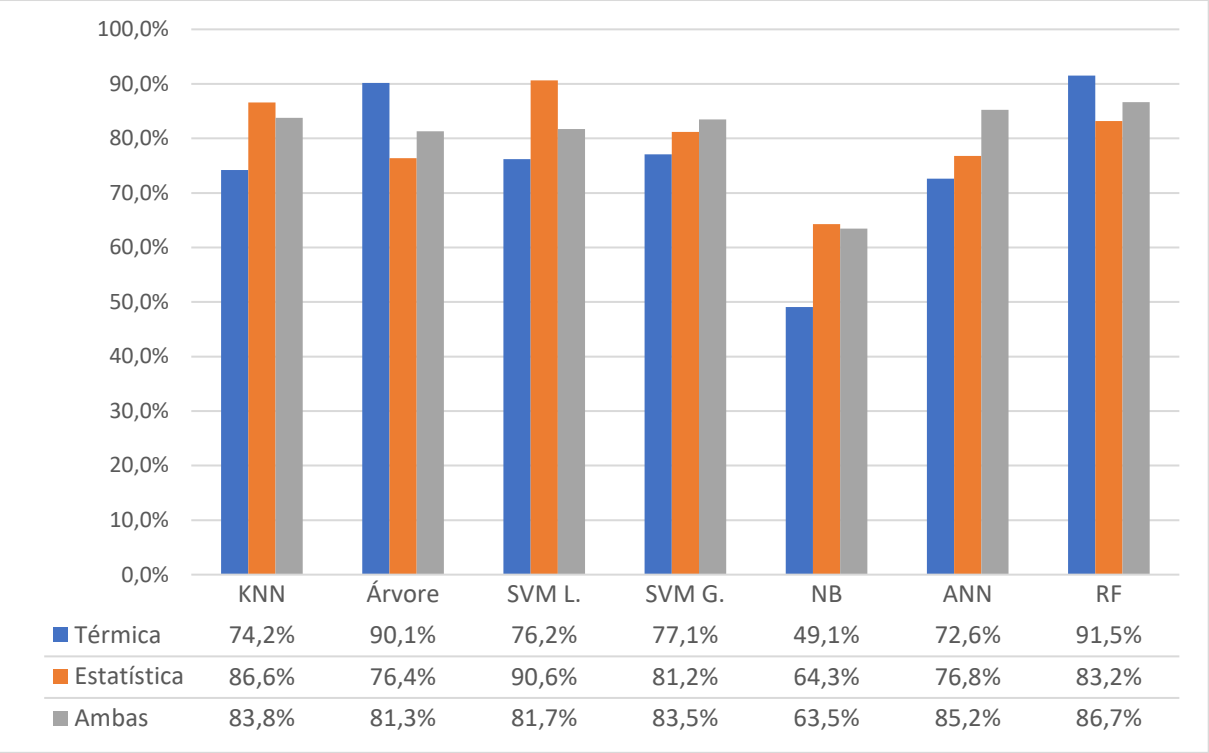
	KNN	Árvore	SVM L.	SVM G.	NB	ANN	RF
Manual	81,7%	79,4%	85,5%	75,9%	60,1%	80,8%	82,3%
Otsu	84,1%	86,9%	74,9%	77,2%	62,4%	72,8%	89,3%
K-means	78,7%	81,5%	88,2%	88,6%	54,4%	81,0%	89,7%

Fonte: Autor

Esse gráfico permite analisar o comportamento de cada classificador em relação à técnica de segmentação escolhida. A segmentação por *k-means* é a que aparece com as melhores taxas de acerto pelos classificadores, sendo a técnica que permite maior acurácia na maioria dos casos. Por outro lado, a segmentação por *otsu* se saiu a melhor, quando utilizada com *k-NN* e Árvore de Decisão. A segmentação manual teve resultado aproximadamente igual a por *k-means*, se tratando de redes neurais. Em média, *k-means* se apresenta como a técnica que permite maior acurácia. Outra análise que pode ser feita é quanto as características extraídas. A seguir, são apresentados os valores médios de acuraria de cada classificador, para cada conjunto de características, independentemente do método de processamento de imagens escolhido.

As melhores combinações foram as árvores de decisão com as características térmicas, e as *Random Forests*, também com características térmicas somente. Apesar dessas características se destacarem com tais algoritmos, é interessante observar que para os outros cinco algoritmos utilizados, elas são as que resultam em pior acurácia. Em média, as características estatísticas ou a combinação de ambas trazem melhores resultados.

Figura 20: Teste dos modelos para as diferentes características extraídas



Fonte: Autor

Após os testes de classificação com o problema de mal-contato, realizou-se os mesmos testes para a obtenção de modelos que classifiquem problemas de sobrecorrente, com as imagens térmicas onde correntes acima do normal são aplicadas sobre a placa. O conjunto de imagens com sobrecorrente é de 1680 no total. Utiliza-se 1680 do caso normal, para um conjunto de treinamento balanceado, resultando em 3360 imagens. Metade das imagens de cada caso são utilizadas no treinamento, e a outra metade de cada caso nos testes. O treinamento dos classificadores foi feito com validação cruzada com k = 5. A *Tabela 10* a seguir apresenta os resultados da validação do treinamento.

Tabela 11: Validação dos algoritmos de classificação

Algoritmo	Conjunto de características	Segmentação manual	Segmentação por *otsu*	Segmentação por *K-means*
k-NN	Térmicas	1	1	0,9997
	Estatísticas	1	1	1
	Ambas	1	1	0,9997
Árvore de decisão	Térmicas	0,9918	0,9944	0,9939
	Estatísticas	1	1	1
	Ambas	0,9899	1	1
SVM Linear	Térmicas	0,9984	0,9915	0,9888
	Estatísticas	1	0,9992	0,9979
	Ambas	1	1	0,9997
SVM Gaussiano	Térmicas	0,9992	0,9918	0,9894
	Estatísticas	1	1	1
	Ambas	1	1	0,9997
Naive Bayes	Térmicas	0,8006	0,7806	0,603
	Estatísticas	0,6317	0,7429	0,7429
	Ambas	0,8631	0,8394	0,6932
Rede Neural	Térmicas	0,9995	0,9997	0,9997
	Estatísticas	1	1	1
	Ambas	1	1	1
Random Forest	Térmicas	0,9999	1	0,9999
	Estatísticas	0,9995	1	1
	Ambas	1	1	1

Fonte: Autor

Para todos os algoritmos, a taxa de acerto foi bastante elevada, muito próxima de 100%. *Naive Bayes* foi o único que apresentou resultados abaixo disso. Em comparação com a validação dos modelos para o problema de mal-contato, esses se demonstram ser mais fáceis de se classificar para tais classificadores, com acurácia superior em todos os casos. A *Tabela 11* a seguir apresenta os testes realizado com os modelos obtidos para o problema de sobrecorrente, utilizando os 1680 exemplos restantes que não fizeram parte do treinamento.

Os resultados de acurácia dos testes são variados, mas com exceção de *Naive Bayes*, verifica-se que todos os classificadores possuem mais de 90% de acerto, para qualquer combinação de segmentação ou características extraídas. Imagens térmicas do problema de sobrecorrente são distintas o suficiente da situação normal para que os classificadores obtenham resultados elevados na maioria dos casos.

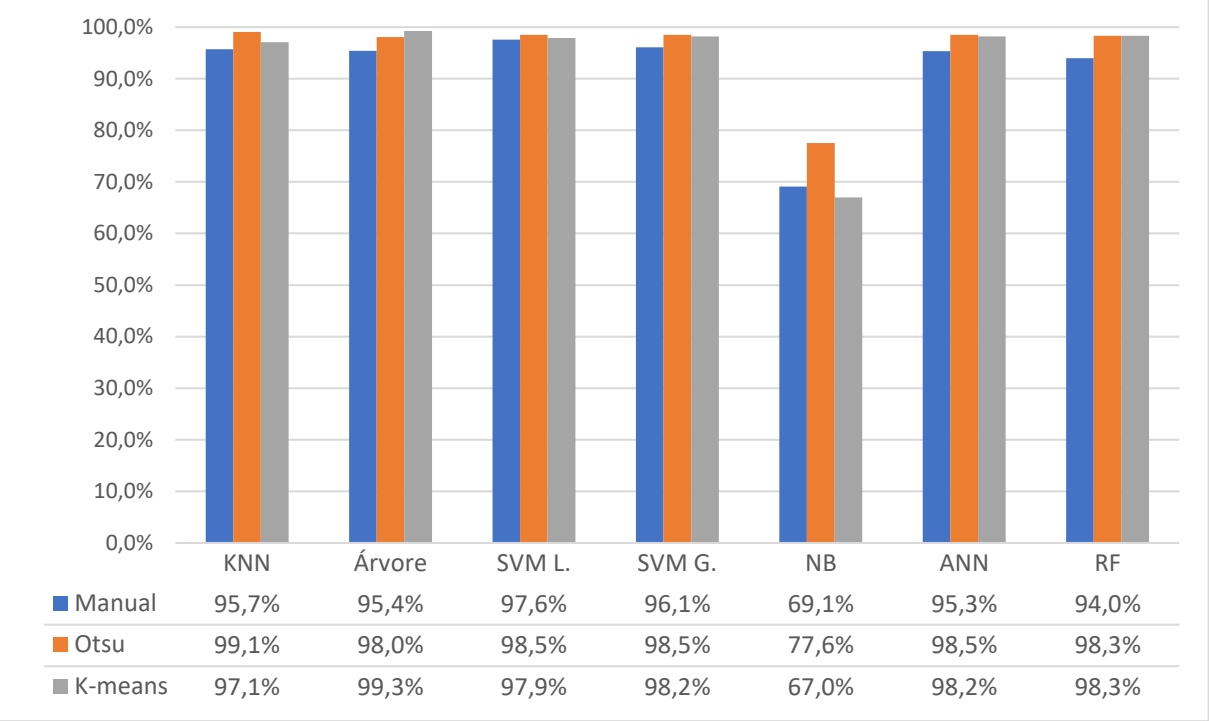
Tabela 12: Resultados dos algoritmos de classificação

Algoritmo	Conjunto de características	Segmentação manual	Segmentação por *otsu*	Segmentação por *K-means*
k-NN	Térmicas	0,907344	0,978018	0,985595
	Estatísticas	0,92917	0,992402	0,989593
	Ambas	0,975232	0,984735	0,968406
Árvore de decisão	Térmicas	0,934913	0,981268	0,960961
	Estatísticas	0,936815	0,986706	0,976097
	Ambas	0,96382	0,996808	0,979096
SVM Linear	Térmicas	0,924127	0,970961	0,957881
	Estatísticas	0,930997	0,993513	0,966338
	Ambas	0,986176	0,996385	0,967588
SVM Gaussiano	Térmicas	0,976856	0,978402	0,97979
	Estatísticas	0,929793	0,982878	0,997259
	Ambas	0,982277	0,980911	0,974676
Naive Bayes	Térmicas	0,729036	0,762496	0,586925
	Estatísticas	0,556333	0,735604	0,712444
	Ambas	0,76515	0,832486	0,675906
Rede Neural	Térmicas	0,90453	0,976335	0,973032
	Estatísticas	0,952332	0,991136	0,974928
	Ambas	0,962191	0,986371	0,984075
Random Forest	Térmicas	0,936956	0,986428	0,975774
	Estatísticas	0,947467	0,987652	0,967937
	Ambas	0,958709	0,978592	0,980687

Fonte: Autor

O gráfico a seguir apresentado pela *Figura 21* mostra a análise da taxa de acerto nos testes realizados, em relação ao algoritmo de segmentação utilizado com cada classificador. A segmentação por *otsu* apresentou na maioria dos casos os melhores resultados, apesar de bastante próximos da segmentação por *k-means*. É interessante observar que esse resultado se opõe ao que mal-contato, onde *k-means* é o algoritmo de segmentação que permite melhor taxa de acerto. Assim, para cada caso isolado, é importante observar que diferentes técnicas permitem melhor resultados. Esses casos são apresentados na *Tabela 12*, onde pode-se observar que para problemas de mal-contato, em maioria *k-means* se saiu melhor, enquanto para sobrecorrente, em maioria *otsu* teve melhores resultados.

Figura 21: Teste dos modelos para os diferentes métodos de segmentação



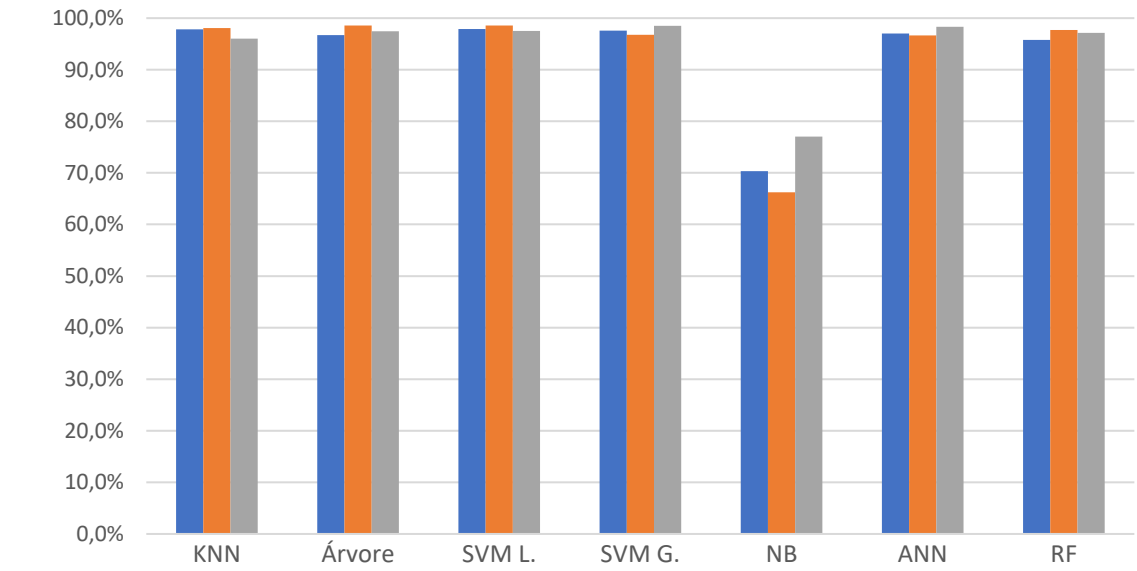
Fonte: Autor

Tabela 13: Acurácia utilizando *k-means* e *otsu*, para cada simulação

Algoritmo	Conjunto de características	Segmentação por *otsu*	Segmentação por *K-means*
k-NN	Sobrecorrente	0,985051667	0,981198
	Mal-contato	0,84121	0,787495
Árvore de decisão	Sobrecorrente	0,988260667	0,972051
	Mal-contato	0,868777333	0,815479
SVM Linear	Sobrecorrente	0,986953	0,963936
	Mal-contato	0,748763333	0,881997
SVM Gaussiano	Sobrecorrente	0,980730333	0,983908
	Mal-contato	0,771760333	0,88627
Naive Bayes	Sobrecorrente	0,776862	0,658425
	Mal-contato	0,623713333	0,543702
Rede Neural	Sobrecorrente	0,984614	0,977345
	Mal-contato	0,728333333	0,81
Random Forest	Sobrecorrente	0,984224	0,974799
	Mal-contato	0,893471333	0,896663

Fonte: Autor

Figura 22: Teste dos modelos para as diferentes características extraídas



	KNN	Árvore	SVM L.	SVM G.	NB	ANN	RF
Térmica	97,8%	96,7%	97,9%	97,6%	70,4%	97,0%	95,8%
Estatística	98,1%	98,6%	98,6%	96,7%	66,2%	96,7%	97,7%
Ambas	96,0%	97,4%	97,5%	98,5%	77,1%	98,3%	97,1%

Fonte: Autor

Como apresentado pela *Figura 22*, há um equilíbrio entre os resultados obtidos por características estatísticas ou a combinação de ambas as características, e somente térmica fica um pouco a baixo, apesar de serem diferenças abaixo de 1%. Conclui-se que para o caso de sobrecorrente, a escolha do algoritmo de segmentação é mais decisiva do que as características extraídas, possuindo maior influência no resultado final. Todos os classificadores obtiveram acurácia acima de 95% para qualquer combinação, com exceção de *Naive Bayes*, que como discutido, não considera as relações entre os parâmetros de entrada e obteve resultados baixos.

Por fim, os últimos testes foram realizados para verificar a capacidade geral dos modelos de classificar as condições elétricas da placa, como com problema ou sem, considerando os casos de mal-contato e sobretensão. Nesse caso, se avalia a acurácia dos classificadores em acertar se há realmente um problema, independente de qual seja. Para isso, será utilizado o total de imagens coletadas para mal-contato e sobretensão, com 5080 imagens de problema, e balanceado com mais 5080 imagens do caso normal. 60% dessas imagens foram utilizadas no treinamento, e 40% nos testes. Após o treinamento por validação cruzada, os resultados de acurácia são apresentados na *Tabela 13*.

Tabela 14: Validação dos algoritmos de classificação

Algoritmo	Conjunto de características	Segmentação manual	Segmentação por *otsu*	Segmentação por *K-means*
k-NN	Térmicas	0,5282	0,5397	0,7092
	Estatísticas	0,6718	0,6718	0,6747
	Ambas	0,6805	0,6776	0,6948
Árvore de decisão	Térmicas	0,6287	0,531	0,6374
	Estatísticas	0,692	0,692	0,692
	Ambas	0,4506	0,7006	0,692
SVM Linear	Térmicas	0,5741	0,7667	0,7149
	Estatísticas	0,4276	0,3902	0,3874
	Ambas	0,6891	0,6661	0,6632
SVM Gaussiano	Térmicas	0,5569	0,7638	0,7437
	Estatísticas	0,4276	0,3902	0,3931
	Ambas	0,6718	0,669	0,6632
Naive Bayes	Térmicas	0,6057	0,508	0,6948
	Estatísticas	0,669	0,669	0,669
	Ambas	0,7034	0,6977	0,6661
Rede Neural	Térmicas	0,5569	0,7983	0,7552
	Estatísticas	0,6546	0,6546	0,6374
	Ambas	0,5885	0,6632	0,6747
Random Forest	Térmicas	0,6173	0,7544	0,7332
	Estatísticas	0,6678	0,7213	0,5991
	Ambas	0,6948	0,7018	0,6712

Fonte: Autor

As SVMs e redes neurais foram as que obteram melhores resultados de classificação. Algumas combinações trouxeram acurácias bastante abaixo, como as características estatísticas com SVM. Por outro lado, as características térmicas para esse classificador se saíram bem em relação ao restante, trazendo acertos acima de 70% para os casos de segmentação automática; as redes neurais chegaram a aproximadamente 80% de acerto. Por outro lado, as árvores de decisão e *Random Forest* trouxeram resultados mais balanceados. As árvores acertaram com em média 69% para características estatísticas, e as *Random Forest*s e *k-NN* acertaram com em média 68% todos os casos, para ambas características. Diferentemente dos classificadores obtidos para os casos isolados anteriormente, o *Naive Bayes* obteve resultados próximos dos outros, sendo uma alternativa viável quando se deseja um classificador mais rápido com boa acurácia. Na *Tabela 14,* são apresentados os resultados de acurácia ao testar tais modelos com novos exemplos de teste, as 40% amostras não utilizadas no treinamento.

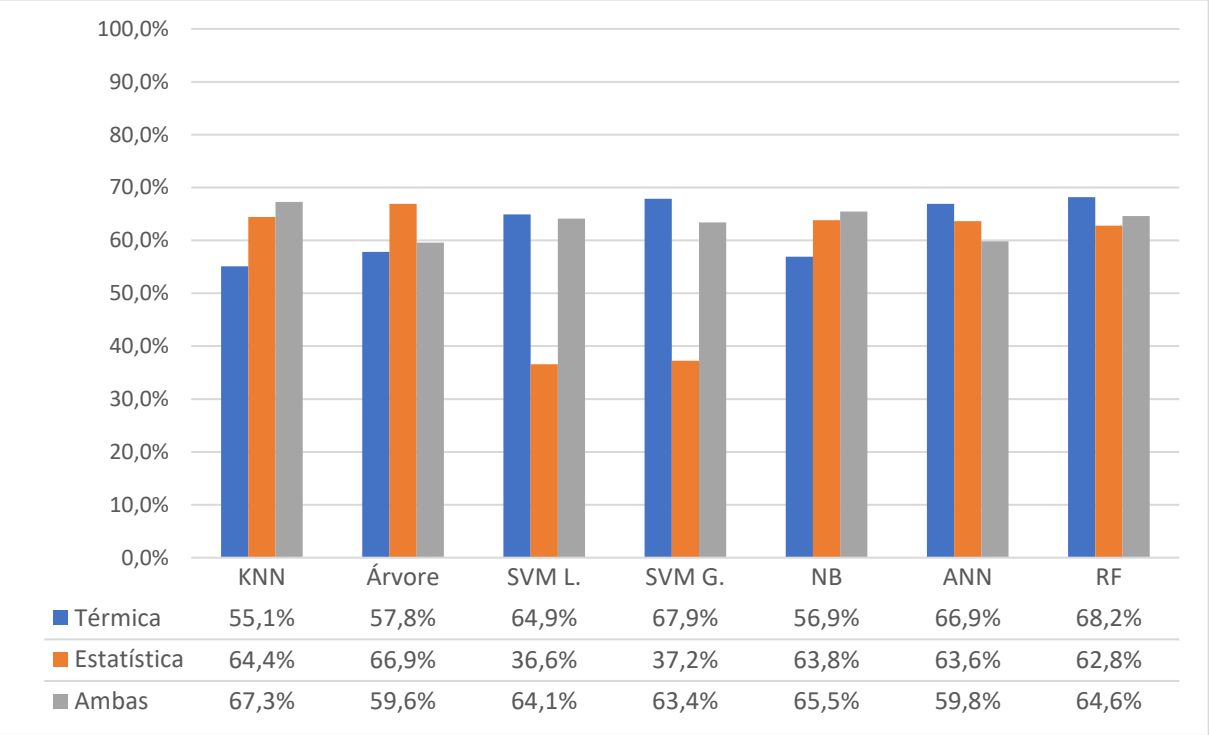
Tabela 15: Resultados dos algoritmos de classificação

Algoritmo	Conjunto de características	Segmentação manual	Segmentação por *otsu*	Segmentação por *K-means*
k-NN	Térmicas	0,443757	0,514131	0,694578
	Estatísticas	0,634022	0,65191	0,646063
	Ambas	0,674728	0,672431	0,67113
Árvore de decisão	Térmicas	0,59622	0,529632	0,608695
	Estatísticas	0,638161	0,677803	0,691119
	Ambas	0,434197	0,677056	0,677032
SVM Linear	Térmicas	0,501148	0,75588	0,691067
	Estatísticas	0,355791	0,377494	0,364354
	Ambas	0,616202	0,649256	0,658563
SVM Gaussiano	Térmicas	0,539549	0,759296	0,737064
	Estatísticas	0,378929	0,375597	0,362346
	Ambas	0,59989	0,662868	0,640039
Naive Bayes	Térmicas	0,548117	0,48554	0,673537
	Estatísticas	0,602015	0,652361	0,661045
	Ambas	0,645672	0,675474	0,643275
Rede Neural	Térmicas	0,473042	0,798096	0,736683
	Estatísticas	0,625552	0,647859	0,636076
	Ambas	0,494097	0,655282	0,645084
Random Forest	Térmicas	0,605	0,736487	0,704658
	Estatísticas	0,598609	0,705541	0,579717
	Ambas	0,632297	0,689043	0,61778

Fonte: Autor

A taxa de acerto caiu em relação aos testes com apenas o problema de mal-contato, ou apenas o problema de sobrecorrente. São mais dados e variáveis envolvidas nesse caso, e os classificadores não foram tão capazes para os dados de entrada, como aconteceu nos casos anteriores. Algumas taxas foram realmente baixas, outras melhores. *Naive Bayes* se saiu muito bem, considerando que teve taxas baixas para os casos anteriores. As SVMs para ambas as características também tiveram bons resultados, assim como *k-NN*. As redes neurais obteram resultados próximos disso, desconsiderando o caso de segmentação manual. As duas maiores acurácias obtidas foram para o caso de SVM com kernel Gaussiano, com 76% de acerto, e as redes neurais, com 80% de acerto; em ambos os casos, foram utilizadas características térmicas e segmentação por *otsu*.

Figura 23: Teste dos modelos para as diferentes características extraídas

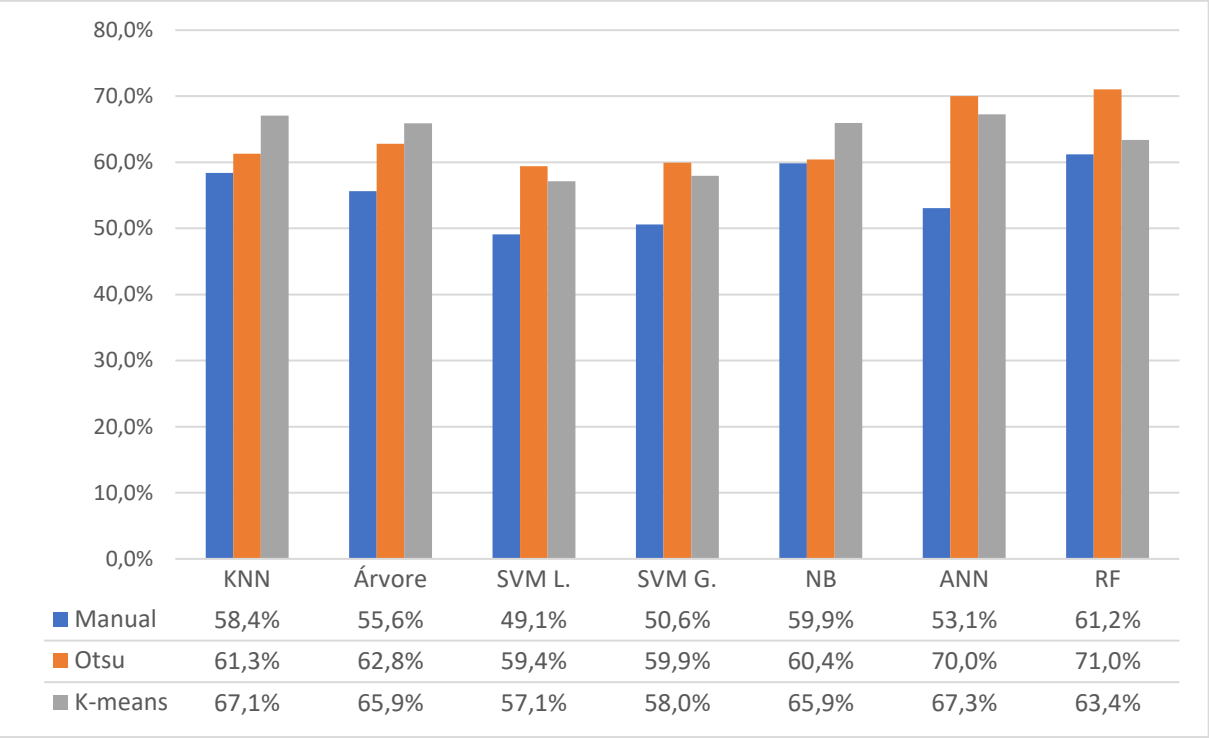


Fonte: Autor

A *Figura 23* indica que as características estatísticas foram as com resultados mais baixos, em média, principalmente se tratando das SVMs, apesar de terem se saído bem com as árvores de decisão, apenas. O conjunto com ambas for melhor para *Naive Bayes*, Árvore e *k-NN*, e as térmicas para SVM, ANN e *Random Forest*, representando um equilíbrio entre esses dois conjuntos. Porém os resultados mais elevados foram obtidos pelas características térmicas, tanto com redes neurais como SVMs.

Na *Figura 24*, temos os resultados de testes dos classificadores para as diferentes formas de segmentação. Nela, verifica-se que a segmentação manual representa as piores taxas de acerto dentre todos os modelos obtidos. Em três casos a segmentação por *k-means* teve maior acurácia, e em quatro casos a segmentação por *otsu*. Além disso, é possível notar que as maiores taxas de acerto foram pela segmentação por *otsu*, que obteram taxas de 70% e 71% de acerto para redes neurais e *Random Forest*s, respectivamente. *K-means* se saiu bem com redes neurais e *k-NN*, com uma média de 67% de acerto. Uma conclusão que se pode tirar é que para este caso de problemas elétricos, o processamento de imagens e segmentação melhorou a taxa de acerto obtida.

Figura 24: Teste dos modelos para os diferentes métodos de segmentação



	KNN	Árvore	SVM L.	SVM G.	NB	ANN	RF
Manual	58,4%	55,6%	49,1%	50,6%	59,9%	53,1%	61,2%
Otsu	61,3%	62,8%	59,4%	59,9%	60,4%	70,0%	71,0%
K-means	67,1%	65,9%	57,1%	58,0%	65,9%	67,3%	63,4%

Fonte: Autor

Como conclusão dos resultados, temos que para o caso de mal-contato, *k-means* se saiu como o melhor método de segmentação, com resultados superiores a segmentação por *otsu*. Com relação às características extraídas, as características térmicas foram as que trouxeram melhores resultados. As melhores taxas de acertos foram obtidas com árvores de decisão e SVM com kernel Gaussiano. Para o caso de sobrecorrente, temos que as combinações de características e algoritmos de segmentação pouco influenciam nos resultados de classificação, sempre trazendo taxas de acerto superiores a 90%, com exceção de *Naive Bayes*. Por fim, para o caso de problemas elétricos, considerando essas duas condições, a taxa de acerto cai, em relação aos outros casos, onde se há maior incerteza. As *Random Forest*s, SVMs e redes neurais foram as que obtiveram melhores resultados, com 74%, 76% e 80% de acurácia, respectivamente, sendo as características térmicas e segmentação por *otsu* as técnicas que trouxeram as maiores taxas. *Naive Bayes* teve um resultado médio elevado para qualquer combinação, em relação à sua taxa de acerto nos casos anteriores, chegando a 67% de acerto para segmentação por *k-means* e características térmicas, sendo também um algoritmo interessante, pois é computacionalmente mais eficiente. Esses resultados são sumarizados na *Tabela 15*.

Esses resultados indicam que diferentes combinações de características e algoritmos de processamento podem ser utilizadas para casos em específico, para obter bons resultados de classificação. Para diferentes casos, como visto, variando as técnicas escolhidas a taxa de acerto melhora; não é possível determinar que sempre *otsu* ou *k-means*, é superior, nem mesmo qual o conjunto de características mais adequado. Para esse caso específico de placa de circuito impresso e problemas elétricos, as características térmicas e segmentação por *otsu* se saíram um pouco melhores que as outras combinações, assim como SVMs, redes neurais e *Random Forest*es. Também pode-se concluir que em todos os casos a segmentação manual foi superada pelos algoritmos de segmentação.

Tabela 16: Resultados dos algoritmos de classificação para cada simulação

Classificador	Simulação	Acurácia (%)
k-NN	Sobrecorrente	99,24%
	Mal-contato	89,94%
	Geral	69,46%
Árvore de decisão	Sobrecorrente	99,68%
	Mal-contato	96,05%
	Geral	69,11%
SVM Linear	Sobrecorrente	99,64%
	Mal-contato	98,22%
	Geral	75,59%
SVM Gaussiano	Sobrecorrente	99,73%
	Mal-contato	98,50%
	Geral	75,93%
Naive Bayes	Sobrecorrente	83,25%
	Mal-contato	67,72%
	Geral	67,55%
Rede Neural	Sobrecorrente	99,11%
	Mal-contato	86,30%
	Geral	79,81%
Random Forest	Sobrecorrente	98,77%
	Mal-contato	95,76%
	Geral	73,65%

Fonte: Autor

A base de dados foi composta por imagens coletadas por meio da câmera térmica. Essas imagens são filtradas, onde aplica-se a segmentação (recorte) e tratamento de imagem manual, por meio do software FLIR tools, que posteriormente são comparadas com o processamento de imagens automático. O repositório de imagens e material complementar esta pesquisa é disponibilizada em https://github.com/yolivatti/electrical-ir-analysis.

6 CONSIDERAÇÕES FINAIS

Este trabalho busca entender os cenários para propor um método de detecção automática de problemas e falhas em sistemas elétricos por meio de algoritmos de *machine learning*. A partir da revisão na literatura, conclui-se que há diversos trabalhos que visam resolver alguns dos problemas existentes em perdas elétricas por meio da inspeção não invasiva propiciada pelas câmeras térmicas, já que esse é um domínio onde a temperatura é uma informação significativa.

Apesar das diversas propostas em termografia e energia, não se há um consenso de quando utilizar cada método, quais as técnicas mais adequadas dado as condições específicas da aplicação e das imagens térmicas coletadas, bem como uma análise comparativa dos algoritmos de classificação. Como primeiro resultado, foi possível propor uma metodologia geral, com base nas diversas soluções em termografia, energia elétrica e inteligência artificial, desde a etapa da obtenção da imagem térmica e dos dados relevantes, até a etapa da detecção. O processo consiste ainda no processamento para tratar a imagem e obter a região de calor, na extração de parâmetros relevantes e por fim em utilizar os dados obtidos para classificação.
Foi possível propor um experimento para simular alguns dos principais problemas elétricos apresentados na literatura, por meio de uma placa de circuito impresso e simulação de diferentes problemas elétricos nela. Através de equipamentos para medições e uma câmera térmica para coleta da temperatura dos elementos envolvidos, pode-se obter uma base de dados para a verificação da metodologia proposta. Com isso, esse trabalho entrega um *dataset* de imagens térmicas, rotulado de acordo com a condição elétrica apresentada.

Com o experimento executado, contata-se que é possível classificar as imagens térmicas coletadas em diferentes categorias, por meio do uso de algoritmos de processamento de imagens, extração de características relevantes e classificadores. Foi possível aplicar diferentes algoritmos de classificação a fim de detectar qual tipo de condição elétrica a placa e suas trilhas se encontravam. Pode-se concluir que os classificadores mais utilizados na literatura, para o esse caso, obtiveram taxas de acertos elevadas, próximas a 100%, para o caso de sobrecorrente simulada nas trilhas elétricas. Por outro lado, para problemas de mal-contato, essa taxa cai um pouco, ficando entre 70% e 96% de acerto. Apenas *Naive Bayes* se saiu abaixo do esperado, com taxas sempre inferiores aos outros classificadores para qualquer combinação de técnicas utilizadas. Por fim, para o caso considerando quaisquer dos problemas elétricos na placa, os classificadores obtiveram resultados um pouco a baixo, com a maior complexidade e número de variáveis envolvidas. As taxas de acerto neste caso estiveram em torno de 60%, com algumas

combinações obtendo resultados superiores, como as *Random Forest*s, SVMs e redes neurais para segmentação por *otsu* e características térmicas, com 74%, 76% e 80% de acerto, respectivamente. *Naive Bayes* também obteve bons resultados, com 67% em média para *k-means*, em oposição aos resultados anteriores nos outros casos.

Alguns pontos ainda ficaram abertos. O primeiro deles se dá quanto a dimensionalidade dos dados. As características extraídas das imagens térmicas foram consideradas a partir da literatura e trabalhos correlatos, para problemas elétricos. Porém, apenas um subconjunto deles pode ser útil para a classificação, em casos específicos. Reduzir a dimensionalidade do conjunto de dados e utilizar algoritmos de otimização para selecionar as características mais relevantes pode ser interessante, porém, aumentaria o escopo desse trabalho, com mais permutações e combinações diferentes a se estudar. Outro caso não considerado aqui é o de utilizar mais classes do que "problema" e "não problema". Na maioria dos trabalhos revisados, só se trabalha com duas classes, mas fica claro a possibilidade de, por exemplo, classificar casos conforme o tipo de condição elétrica. Nossa base de dados contém etiquetas conforme o problema elétrico simulado (mal-contato, sobrecorrente ou normal), permitindo que o método seja expandido para classificação em mais categorias.

Com base nos experimentos e resultados obtidos, é possível validar a utilização da metodologia proposta utilizando algoritmos de classificação e processamento de imagens, após a captura termográfica da cena, para casos onde a variação de calor em diferentes pontos da imagem representa algum tipo de situação. Não é possível generalizar qual o melhor algoritmo ou técnica. Para o caso de placas elétricas, as características térmicas e segmentação por *otsu* se saiu melhor, levando em consideração a maioria dos casos. Porém, ao analisar casos específicos, temos que outras combinações funcionam melhores com certos classificadores, como *k-means* com SVM para problemas de mal-contato. Essa constatação leva a conclusão que é possível variar as combinações e incluir até mesmo outros conjuntos de características para a classificação de diferentes tipos de problemas, elétricos ou não, com imagens térmicas Sempre que variações térmicas na cena em estudo representar algum tipo de situação, temos características que podem ser extraídas da imagem e utilizadas por classificadores para a obtenção de um modelo de classificação.

Um grande desafio para trabalhos futuros será a coleta e composição de base de dados para cenários reais. Obter imagens térmicas é desafiador, pois é necessário a obtenção de imagens situação normal e anômala, em diversas posições no sistema inspecionado e sob diferentes condições climáticas, e o grande número de variáveis envolvidas obriga a obtenção

de uma base de dados bastante numerosa para a obtenção de bons resultados com algoritmos de classificação.

Como contribuição, esse trabalho deixa um método para coleta de imagens térmicas e a análise computacional por meio de classificadores de *machine learning*. O trabalho deixa uma análise comparativa de algoritmos de classificação e processamento de imagens térmicas, para um problema específico de problema elétrico, que pode ser expandido para outros tipos de problemas e outras situações em cenários de energia elétrica. O roteiro experimental e a base de pesquisa *(dataset)* é disponibilizada, permitindo a reprodutibilidade desse trabalho, a inclusão de novas variáveis e diferentes algoritmos, e a solução de outros tipos de problemas. Esse trabalho pode contribuir para futuras pesquisas que visem avanço computacional na área termográfica, permitindo a análise e identificação de características e desenvolvimento de novos algoritmos, bem como para pesquisas que busquem solucionar problemas e otimizar a gestão de recursos em cidades inteligentes, por meio da termografia.

REFERÊNCIAS

ABD ELRAHMAN, S. M.; ABRAHAM, A. A Review of Class Imbalance Problem. **Journal of Network and Innovative Computing**, v. 1, p. 332–340, 2013.

ABRADEE. **Base de dados de perdas elétricas da ABRADEE**. Disponível em: <http://www.abradee.com.br/setor-de-distribuicao/banco-de-dados/comercial>. Acesso em: 23 abr. 2018.

ACHARYA, U. R. et al. Thermography Based Breast Cancer Detection Using Texture Features and Support Vector Machine. **Journal of Medical Systems**, v. 36, n. 3, p. 1503–1510, 2012.

ALBINO, V.; BERARDI, U.; DANGELICO, R. M. Smart cities: Definitions, dimensions, performance, and initiatives. **Journal of Urban Technology**, v. 22, n. 1, p. 1–19, 2015.

ALI, A.; SHAMSUDDIN, S. M.; RALESCU, A. L. Classification with class imbalance problem: A review. **International Journal of Advances in Soft Computing and its Applications**, v. 7, n. 3, p. 176–204, 2015.

ALMEIDA, C. A. L. et al. Intelligent thermographic diagnostic applied to surge arresters: A new approach. **IEEE Transactions on Power Delivery**, v. 24, n. 2, p. 751–757, 2009.

ANEEL. **Base de dados de perdas elétricas da ANEEL**. Disponível em: <http://www.aneel.gov.br/metodologia-distribuicao/-/asset_publisher/e2INtBH4EC4e/content/perdas/654800>. Acesso em: 23 abr. 2018.

ANTMANN, P. Reducing technical and non-technical losses in the power sector. **Background Paper for the World Bank Group Energy Sector Strategy**, n. July 2009, p. 1–35, 2009.

ARAKI, S.; NOMURA, H.; WAKAMI, N. Segmentation of thermal images using the fuzzy C-means algorithm. **[Proceedings 1993] Second IEEE International Conference on Fuzzy Systems**, p. 719–724, 1993.

AVALOS, M.; SILVA, I.; LARIOS, V. M. Application of Business Analytics on Big Data Scenarios to obtain Actionable Insights in Systems or Services i n Mega - Cities : Multiple use cases from the Living Lab at the Smart Cities Innovation Center at Guadalajara University. p. 1–7, 2015.

BAGAVATHIAPPAN, S. et al. Infrared thermography for condition monitoring - A review. **Infrared Physics and Technology**, v. 60, n. December 2017, p. 35–55, 2013.

BALARAS, C. A.; ARGIRIOU, A. A. Infrared thermography for building diagnostics. **Energy and Buildings**, v. 34, n. 2, p. 171–183, 2002.

BEN-DAVID, S.; SHALEV-SHWARTZ, S. **Understanding Machine Learning: From Theory to Algorithms**. [s.l: s.n.].

BISHOP, C. M. **Pattern Recognition and Machine Learning**. [s.l: s.n.]. v. 53

BÜHRINGER, M. et al. Inspection of High Voltage Power Lines -a New Approach. **Mobile**

Robotics - Solutions and Challenges - Proceedings of the Twelfth International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines, p. 455–462, 2010.

CALVILLO, C. F.; SÁNCHEZ-MIRALLES, A.; VILLAR, J. Energy management and planning in smart cities. **Renewable and Sustainable Energy Reviews**, v. 55, p. 273–287, 2016.

CHOU, Y.-C.; YAO, L. Automatic Diagnostic System of Electrical Equipment Using Infrared Thermography. **2009 International Conference of Soft Computing and Pattern Recognition**, p. 155–160, 2009.

CUNHA, M. A. et al. Smart Cities: transformação digital das cidades. p. 161, 2016.

DE OLIVEIRA, J. H. E.; LAGES, W. F. Robotized inspection of power lines with infrared vision. **2010 1st International Conference on Applied Robotics for the Power Industry, CARPI 2010**, p. 1–6, 2010.

DERNONCOURT, F. Introduction to Fuzzy Logic Control. In: **Massachusetts Institute of Technology 21**. [s.l: s.n.].

DOS SANTOS, G. M.; DE AQUINO, R. R. B.; LIRA, M. M. S. Thermography and artificial intelligence in transformer fault detection. **Electrical Engineering**, 2017.

DUDZIK, S. Two-stage neural algorithm for defect detection and characterization uses an active thermography. **Infrared Physics and Technology**, v. 71, p. 187–197, 2015.

ESEN, V.; ORAL, B.; AKINCI, T. C. The determination of short circuits and grounding faults in electric power systems using time-frequency analysis. **Journal of Energy in Southern Africa**, v. 26, n. 2, p. 123–132, 2015.

ESTEVE AGELET, L. et al. Feasibility of near infrared spectroscopy for analyzing corn kernel damage and viability of soybean and corn kernels. **Journal of Cereal Science**, v. 55, n. 2, p. 160–165, 2012.

FANG, Y.-C.; WU, B.-W. Neural network application for thermal image recognition of low-resolution objects. **Journal of Optics A: Pure and Applied Optics**, v. 9, n. 2, p. 134–144, 2007.

FÁTIMA, M. DE; RAIA, R. Graziella Costa Gonçalves. 2007.

FILHO, J. M. **Instalações Elétricas Industriais**. 8a. ed. Rio de Janeiro: Editora LTC, 2007.

FOURNIER, D. **Diagnostic of overheating underground distribution cable joints**. 16th International Conference and Exhibition on Electricity Distribution (CIRED 2001). **Anais**...IEEE, 2001Disponível em: <http://digital-library.theiet.org/content/conferences/10.1049/cp_20010726>

GADE, R.; MOESLUND, T. B. Thermal cameras and applications: A survey. **Machine Vision and Applications**, v. 25, n. 1, p. 245–262, 2014.

GHANTY, P. et al. Artificial neural network approach for estimating weld bead width and

depth of penetration from infrared thermal image of weld pool. **Science and Technology of Welding and Joining**, v. 13, n. 4, p. 395–401, 4 maio 2008.

GUNN, S. R. **Support vector machines for classification and regression.** [s.l: s.n.].

GUO, X. et al. On the Class Imbalance Problem. n. October, p. 192–201, 2009.

HUO, Z. et al. Self-adaptive Fault Diagnosis of Roller Bearings using Infrared Thermal Images. p. 6113–6118, 2017.

IZYKOWSKI, J. **POWER SYSTEM FAULTS**. [s.l: s.n.].

JAFFERY, Z. A. et al. Scheme for predictive fault diagnosis in photo-voltaic modules using thermal imaging. **Infrared Physics and Technology**, v. 83, p. 182–187, 2017.

JUBA, B.; LE, H. S. Precision-Recall versus Accuracy and the Role of Large Data Sets. p. 1–14, 2017.

KARVELIS, P. et al. An automated thermographic image segmentation method for induction motor fault diagnosis. **IECON Proceedings (Industrial Electronics Conference)**, p. 3396–3402, 2014.

KIESSLING, F. et al. **Overhead Power Lines**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003.

KORENDO, Z.; FLORKOWSKI, M. Thermography-based diagnostics of power equipment. p. 33–42, 2001.

LAIB DIT LEKSIR, Y.; MANSOUR, M.; MOUSSAOUI, A. Localization of thermal anomalies in electrical equipment using Infrared Thermography and support vector machine. **Infrared Physics & Technology**, v. 89, n. December, p. 120–128, 2018.

LI, B. et al. HV Power Equipment Diagnosis Based on Infrared Imaging Analyzing. **2006 International Conference on Power System Technology**, p. 1–4, 2006.

LIU, C. et al. Automatic detection of diabetic foot complications with infrared thermography by asymmetric analysis. **Journal of Biomedical Optics**, v. 20, n. 2, p. 26003, 2015.

LIU, H. et al. An Efficient Algorithm for Server Thermal Fault Diagnosis Based on Infrared Image. **Journal of Physics: Conference Series**, v. 910, n. 1, 2017.

LORENA, A. C.; DE CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 2007.

MADRUGA, F. J. et al. Infrared thermography processing based on higher-order statistics. **NDT and E International**, v. 43, n. 8, p. 661–666, 2010.

MALDAGUE, X. P. Introduction to NDT by active infrared thermography. **Materials Evaluation**, v. 60, n. 9, p. 1060–1073, 2002.

MARTINS, E. C. Análise Das Perdas Elétricas Em Baixa Tensão Do Sistema De Distribuição Da Cooperativa De Distribuição De Energia- Cersul. p. 34, 2016.

MEOLA, C.; CARLOMAGNO, G. M. Recent advances in the use of infrared thermography. **Measurement Science and Technology**, v. 15, n. 9, 2004.

MLAKIĆ, D.; NIKOLOVSKI, S.; MAJDANDŽIĆ, L. Deep Learning Method and Infrared Imaging as a Tool for Transformer Faults Detection. **Journal of Electrical Engineering**, v. 6, n. 6, p. 98–106, 2018.

MOHANTY, S. P.; CHOPPALI, U.; KOUGIANOS, E. Everything you wanted to know about smart cities: The Internet of things is the backbone. **IEEE Consumer Electronics Magazine**, v. 5, n. 3, p. 60–70, jul. 2016.

MORA, O. B.; LARIOS, V. M. Urban Operating Systems for Sensor Network Management in Smart Cities. **IEEE Guadalajara GDL CCD White Papers**, p. 1–4, 2015.

MOSTAFAVI, H. A.; MIRMAJLESSI, S. M.; FATHOLLAHI, H. **Thermal Imaging for Enhancing Inspection Reliability: Detection and Characterization**. [s.l.] InTech, 2012.

MOUSTAFA, G.; ABD ELMAKSOUD, T. Failure analysis and experimental investigations to the joint resistance of compressed connectors in aluminum cables. **Journal of Electrical Engineering**, v. 15, n. 2, p. 94–100, 2015.

NG, A. **Supervised learningCS229 Lecture notes**, 2000.

NGUYEN, V. N.; JENSSEN, R.; ROVERSO, D. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. **International Journal of Electrical Power and Energy Systems**, v. 99, n. September 2017, p. 107–120, 2018.

PAGE, M. G. Image Processing for Automated Flaw Detection in Pulsed Thermography. **Proceeding of 6th International Workshop-NDT Signal Processing**, n. 1, p. 1–9, 2009.

PALIMKAR, N. H. Fault Prediction in Electrical Equipments using Thermographic Inspection ". v. 5, n. 4, p. 685–688, 2016.

PAPEŽ, B. J.; PALFY, M.; TURK, Z. Infrared Thermography Based on Artificial Intelligence for Carpal Tunnel Syndrome Diagnosis. **Journal of International Medical Research**, v. 36, n. 6, p. 1363–1370, 2008.

PÉREZ-ORTIZ, M. et al. A review of classification problems and algorithms in renewable energy applications. **Energies**, v. 9, n. 8, p. 1–27, 2016.

RAHMANI, A.; HADDADNIA, J.; SERYASAT, O. Intelligent fault detection of electrical equipment in ground substations using thermo vision technique. **ICMEE 2010 - 2010 2nd International Conference on Mechanical and Electronics Engineering, Proceedings**, v. 2, n. Icmee, p. 150–154, 2010.

ROGALSKI, A. Infrared detectors: Status and trends. **Progress in Quantum Electronics**, v. 27, n. 2–3, p. 59–210, 2003.

ROZO, T. J. R.; ALVAREZ, J. C. G.; DOMINGUEZ, C. G. C. Infrared thermal image segmentation using expectation-maximization-based clustering. **STSIVA 2012 - 17th Symposium of Image, Signal Processing, and Artificial Vision**, p. 223–226, 2012.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence: a modern approach**. 3rd. ed. [s.l.] Elsevier Editora Ltda, 2013.

SALOKY, T.; ŠEMINSKÝ, J. Artificial Intelligence and Machine Learning. **Department of Automation and Control, Faculty of Mechanical Engineering, Technical University of Košice, Slovak Republic**, p. 1–6, 1999.

SENNI, L. et al. On-line automatic detection of foreign bodies in biscuits by infrared thermography and image processing. **Journal of Food Engineering**, v. 128, p. 146–156, 2014.

SEYMOUR, J.; HORSLEY, T. The seven types of power problems. **EngineerIT**, n. JUNE, p. 60–65, 2008.

SHAFI'I, M. A.; HAMZAH, N. Internal Fault Classification using Artificial Neural Network. **the 4th International Power Engineering and Optimization Conf. (PEOCO2010)**, n. June, p. 352–357, 2010.

SIDDIQUE, A.; YADAVA, G. S.; SINGH, B. Applications of artificial intelligence techniques for induction machine stator fault diagnostics: review. **4th IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives, 2003. SDEMPED 2003.**, n. August, 2003.

SIM, L. S. J. Computerized detection of breast cancer with arti cial intelligence and thermograms. v. 26, n. 4, p. 152–157, 2002.

SINGH, G.; NAIKAN, V. N. A. Infrared thermography based diagnosis of inter-turn fault and cooling system failure in three phase induction motor. **Infrared Physics and Technology**, v. 87, p. 134–138, 2017.

SMEDBERG, M. Thermographic Decision Support--Detecting and Classifying Faults in Infrared Images. 2006.

SUDHA, M. **Development and implementation of embedded protection scheme for three phase squirrel cage induction motor fault analysis**. [s.l.] Anna University, 2009.

SURIYAMONGKOL, D. Non-Technical Losses in Electrical Power Systems. **Ohio University**, n. November, p. 97, 2002.

SVENSSON, M. **Classification of leakage detections acquired by airborne thermography of district heating networks**. [s.l.] Institutionen för systemteknik, 2014.

TAVARES, E.; COSTA, E. G.; MAIA, M. J. A. Influence of emissivity and distance in high voltage equipments thermal Imaging. **2006 IEEE PES Transmission and Distribution Conference and Exposition: Latin America, TDC'06**, n. September, 2006.

TÉCNICAS, A. B. DE N. **NBR 5410: Intalações elétricas de baixa tensão**, 2004.

The Electricity at Work Regulations 1989. 3rd. ed. London: HSE Books, 2015. v. 1

ULLAH, I. et al. Predictive Maintenance of Power Substation Equipment by Infrared Thermography Using a Machine-Learning Approach. **Energies**, v. 10, n. 12, p. 1987, 2017.

UNCTAD. Smart cities and infrastructure Report of the Secretary-General Economic and Social Council. v. 3074, n. February, p. 18, 2016a.

UNCTAD. Issues Paper On Smart Cities and Infrastructure Advance Unedited Draft. n. January, p. 1–61, 2016b.

USAMENTIAGA, R. et al. Infrared Thermography for Temperature Measurement and Non-Destructive Testing. **Sensors**, v. 14, n. 7, p. 12305–12348, 2014.

WANG, W.; ZHANG, J.; SHEN, C. Improved human detection and classification in thermal images. **Proceedings - International Conference on Image Processing, ICIP**, n. February 2014, p. 2313–2316, 2010.

WONG, W. K. et al. **An Effective Surveillance System Using Thermal Camera**. 2009 International Conference on Signal Acquisition and Processing. **Anais**...IEEE, abr. 2009Disponível em: <http://ieeexplore.ieee.org/document/5163816/>. Acesso em: 16 mar. 2018

WORLD BANK. **Base de dados do World Bank**. Disponível em: <https://data.worldbank.org/indicator/EG.ELC.LOSS.ZS>. Acesso em: 23 abr. 2018.

WRETMAN, D. Finding regions of interest in a decision support system for analysis of infrared images. 2006.

ZARKOVIC, M. D.; STOJKOVIC, Z. Fuzzy logic and artificial neural network-based thermography approach for monitoring of high-voltage equipment. **International Journal of Electrical Engineering Education**, v. 52, n. 1, p. 81–96, 2015.

ŽARKOVIĆ, M.; STOJKOVIĆ, Z. Analysis of artificial intelligence expert systems for power transformer condition monitoring and diagnostics. **Electric Power Systems Research**, v. 149, p. 125–136, 2017.