

**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

Gustavo Maciel dos Santos



**TERMOGRAFIA E INTELIGÊNCIA ARTIFICIAL NA DETECÇÃO DE FALHAS
EM TRANSFORMADORES**

**Recife
2017**

Gustavo Maciel dos Santos

**TERMOGRAFIA E INTELIGÊNCIA ARTIFICIAL NA DETECÇÃO DE FALHAS
EM TRANSFORMADORES**

Tese submetida à Coordenação do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do grau de Doutor em Engenharia Elétrica.

Área de concentração: Processamento de Energia.

Orientador: Prof. Dr. Ronaldo Ribeiro Barbosa de Aquino.

Coorientadora: Prof. Dr^a Milde Maria da Silva Lira.

**Recife
2017**

Catálogo na fonte
Bibliotecária Valdicéa Alves, CRB-4 / 1260

S237t Santos, Gustavo Maciel dos.
Termografia e inteligência artificial na detecção de falhas em transformadores / Gustavo Maciel dos Santos - 2017.
97folhas, Il.e Tabs.

Orientador: Prof. Dr. Ronaldo Ribeiro Barbosa de Aquino.
Coorientadora: Prof^a. Dr^a. Milde Maria da Silva Lira.

Tese (Doutorado) – Universidade Federal de Pernambuco. CTG.
Programa de Pós-Graduação em Engenharia Elétrica, 2017.
Inclui Referências.

1. Engenharia Elétrica. 2. Redes neurais artificiais. 3. Lógica fuzzy.
4. Análise de gases dissolvidos. 5. Transformador. 6. Infravermelho. I. Aquino, Ronaldo Ribeiro Barbosa de. (Orientador). II. Lira, Milde Maria da Silva (Coorientadora). III. Título.

UFPE

621.3 CDD (22. ed.)

BCTG/2017 - 358



Universidade Federal de Pernambuco

Pós-Graduação em Engenharia Elétrica

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE
TESE DE DOUTORADO DE

GUSTAVO MACIEL DOS SANTOS

TÍTULO

**“TERMOGRAFIA E INTELIGÊNCIA ARTIFICIAL NA
DETECÇÃO DE FALHAS EM TRANSFORMADORES”**

A comissão examinadora composta pelos professores: RONALDO RIBEIRO BARBOSA DE AQUINO, DEE/UFPE; MILDE MARIA DA SILVA LIRA, DEE/UFPE, GERALDO LEITE TORRES, DEE/UFPE; JOSÉ MAURÍCIO DE BARROS BEZERRA, DEE/UFPE; BENEMAR ALENCAR DE SOUZA, DEE/UFCG e WASHINGTON LUIZ DE ARAÚJO NEVES, DEE/UGCG, sob a presidência do primeiro, consideram o candidato **GUSTAVO MACIEL DOS SANTOS APROVADO.**

Recife, 30 de agosto de 2017.

MARCELO CABRAL CAVALCANTI
Coordenador do PPGE

RONALDO RIBEIRO BARBOSA DE AQUINO
Orientador e Membro Titular Interno

BENEMAR ALENCAR DE SOUZA
Membro Titular Externo

MILDE MARIA DA SILVA LIRA
Coorientadora e Membro Titular Interno

WASHINGTON LUIZ DE ARAÚJO NEVES
Membro Titular Externo

GERALDO LEITE TORRES
Membro Titular Interno

JOSÉ MAURÍCIO DE BARROS BEZERRA
Membro Titular Interno

*Dedico este trabalho à minha esposa, às
minhas filhas, à minha mãe e meu pai (in
memorian)*

AGRADECIMENTOS

Com o tempo passando, percebi o peso da frase “não ter mais idade para isso” e com crianças em casa, elaborar um simples texto pode significar dias. Para continuar seguindo contei com muitos empurrões os quais sem eles não seria possível chegar até aqui. Com esses personagens, consegui manter firme no pensamento “ainda dá tempo”. A eles eu agradeço. Agradeço ao professor Dr. Ronaldo Aquino, pela oportunidade, seu apoio e sua atenção durante estes 9 anos de convivência e a Professora Doutora Milde Lira pela disponibilidade e empenho em me ajudar no decorrer desse trabalho. Sempre motivadora e acolhedora em todos os aspectos. Enquanto alguns jovens alunos falam que seus professores são difíceis, eu só faço me gabar dos professores que tenho, e isso eu agradeço sempre. Obrigado por não desistirem de mim.

Agradeço aos demais professores que colaboraram com este trabalho, avaliando e sugerindo. Acreditem, mesmo se tivemos uma pequena conversa, esse pouco, de alguma forma, está contido neste trabalho. Sintam-se abraçados.

Agradeço aos jovens alunos, estagiários que durante este longo período, tiveram a oportunidade de conhecer este trabalho e colaboraram. Dentre eles um especial agradecimento para Eduarda, Paula e Lucas que colaboraram fortemente na reta final.

Agradeço à minha esposa Paula Moraes, grande professora, que insistentemente corrigia minhas palavras. Quando eu dizia “...se eu conseguir terminar...” ela trocava para “...quando eu conseguir terminar...”. Obrigado por me motivar constantemente.

Agradeço às minhas filhas Maria Beatriz e Isabela pela compreensão e paciência. Poderemos brincar mais vezes agora.

Agradeço à minha mãe Conceição que me motivou desde o princípio com a frase “eu quero, eu posso, eu consigo...”

Agradeço ao meu pai Gildo(*in memoriam*). Seu sonho se tornou meu sonho. Nosso sonho se tornou realidade. Estás sempre em meu coração.

Deus, obrigado por mais este presente. Estou sempre a sua disposição.

“E, voltando para os seus discípulos, achou-os adormecidos; e disse a Pedro: Então nem uma hora pudeste velar comigo? Vigiai e orai, para que não entreis em tentação; na verdade, o espírito está pronto, mas a carne é fraca.”

Mateus 26:40,41

RESUMO

A análise dos gases dissolvidos no óleo isolante em transformadores de potência refrigerados a óleo é uma técnica bastante difundida para detecção de falhas incipientes. Contudo, esta técnica envolve procedimentos de segurança para a coleta das amostras de óleo, prazos de resposta dos laboratórios e em alguns casos há a necessidade de se retirar o transformador de operação. Retirar um transformador de operação em determinadas situações pode se tornar muito onerosa, pois há ambientes de produção que devem funcionar ininterruptamente e paradas como esta representam perda financeira. Termografia Infravermelha é uma técnica não destrutiva de medição de temperatura comumente utilizada para detectar anomalias e prever possíveis falhas sem interromper a operação do sistema. Este trabalho apresenta estudos com o uso da medição infravermelha de temperatura para detectar falhas incipientes no transformador baseada na Análise dos Gases Dissolvidos no Óleo Isolante. A metodologia desse estudo apresenta o uso sistemas inteligentes para analisar as temperaturas faciais do transformador e detectar falhas incipientes. Os resultados obtidos neste trabalho apresentam 86% e 83% de acertos de classificação utilizando Redes Neurais Artificiais e Lógica Fuzzy, respectivamente.

Palavras-chave: Redes neurais artificiais. Lógica fuzzy. Análise de gases dissolvidos. Transformador. Infravermelho

ABSTRACT

Dissolved Gas Analysis of insulating oil in refrigerated power transformer oil is a widespread technique for detecting incipient faults. However, this technique involves safety procedures for the collection of oil samples, laboratory response time and, in some cases, removing the transformer from operation. Removing a transformer from operation in certain situations can become very costly as there are production environments that must run uninterrupted so that stoppages such as this represent financial loss. Infrared Thermography is a non-destructive temperature measurement technique commonly used to detect anomalies and predict possible faults without disrupting system operation. This paper presents studies based on the use of infrared temperature measurement to detect incipient faults in transformers through Dissolved Gas Analysis of the insulating oil. This study's methodology uses intelligent systems to analyse transformer face temperatures and detect incipient faults. The results obtained in this work present 86 and 83% of classification correctness using Artificial Neural Networks and Fuzzy Logic, respectively.

Keywords: Artificial neural network. Fuzzy logic. Dissolved gas analysis. Transformer. Infrared

LISTA DE ILUSTRAÇÕES

Figura 1-1: (a) Imagem termográfica da face A do transformador, (b) imagem real do transformador e (c) indicações do posicionamento das faces.....	20
Figura 1-2: Arquitetura da rede neural	21
Figura 1-3: Sistema de Inferência Fuzzy	22
Figura 2-1: Sir Willian Herschel (1738 – 1832).....	25
Figura 2-2: Termovisor.....	26
Figura 2-3: Arranjo de plano focal	27
Figura 2-4: A radiação pode ser refletida, transmitida, absorvida ou emitida.....	30
Figura 2-5: FOV e IFOV	33
Figura 2-6: Método Refletor.....	34
Figura 4-1: União de dois conjuntos fuzzy.....	42
Figura 4-2: Intersecção de dois conjuntos fuzzy	43
Figura 4-3: Complemento do conjunto fuzzy.....	44
Figura 4-4: Gráfico da função de pertinência de uma variável linguística.....	44
Figura 4-5: Exemplo de função de pertinência triangular	45
Figura 4-6: Exemplo de função de pertinência trapezoidal	46
Figura 4-7: Exemplo de função de pertinência Gaussiana	46
Figura 4-8: Modelo geral em diagrama de blocos de um controlador fuzzy.....	47
Figura 4-9: Exemplo de um valor real com correspondência em dois valores linguísticos	48
Figura 5-1: Modelo do neurônio artificial	53
Figura 5-2: (a) Função de limiar. (b) Função linear por partes. (c) Função sigmóide para parâmetro de inclinação α variável.	55
Figura 5-3: Rede alimentada adiante ou acíclica com uma única camada de neurônios.....	56
Figura 5-4: Rede alimentada adiante ou acíclica totalmente conectada com uma camada oculta e uma camada de saída.....	57
Figura 5-5: Rede recorrente sem laços de auto-realimentação e sem neurônios ocultos	57
Figura 5-6: Rede recorrente com neurônios ocultos.....	58
Figura 5-7: Manifestação do supertreinamento	59
Figura 5-8: Grafo arquitetural de um perceptron de múltiplas camadas com duas camadas ocultas.....	64
Figura 5-9: Correção dos pesos por <i>backpropagation</i>	65
Figura 5-10: Erro em função do peso para uma única conexão	67

Figura 6-1: Medições de temperatura e umidade durante a captura de imagem pelo termovisor	73
Figura 6-2: Dados de placa de um transformador e medições das grandezas elétricas	74
Figura 6-3: Faces do transformador a serem capturadas pelo termovisor	74
Figura 6-4: Resultado da Análise do Óleo Isolante	76
Figura 6-5: Imagem termográfica da face C de um transformador convertida em matriz ..	78
Figura 6-6: Exemplo de um transformador objeto de medição	78
Figura 6-7: Imagem termográfica de uma das faces do equipamento	79
Figura 6-8: Programa FLIR QuickReport 1.2 S1	79
Figura 6-9: Percentual de variância explicada por cada componente principal	81
Figura 6-10: Arquitetura da RNA	84
Figura 6-11: Minimizando o erro nos dados de Treinamento e Validação	86

LISTA DE TABELAS

Tabela 2-1: Valores de emissividade de alguns materiais comuns	31
Tabela 2-2: Fator de Correção da temperatura em função da velocidade do vento	35
Tabela 3-1: Forma de dados para uma análise de componentes principais, com variáveis $X_1, X_2, \dots, X_p$ e observações de n casos	37
Tabela 6-1: Percentual de erro de classificação no conjunto de teste (RNA)	87
Tabela 6-2: Matriz Confusão (RNA).....	88
Tabela 6-3: Percentual de classificação correta por diagnóstico (RNA).....	88
Tabela 6-4: Percentual de erro de classificação no conjunto de teste (FIS)	89
Tabela 6-5: Matriz Confusão (FIS)	89
Tabela 6-6: Percentual de classificação correta por diagnóstico (FS).....	89

SUMÁRIO

1	INTRODUÇÃO	15
1.1	DESCRIÇÃO DO PROBLEMA.....	18
1.2	JUSTIFICATIVA.....	18
1.3	OBJETIVOS	19
1.3.1	Objetivos Gerais.....	19
1.3.2	Objetivos Específicos	19
1.4	METODOLOGIA	20
1.5	ESTRUTURA DO TRABALHO.....	22
2	TERMOGRAFIA INFRATERMELHA E TERMOVISORES	24
2.1	TERMOGRAFIA INFRATERMELHA.....	24
2.2	A HISTÓRIA DA TECNOLOGIA INFRATERMELHA	25
2.3	OPERAÇÃO DO TERMOVISOR	28
2.4	RESOLUÇÃO DE PROBLEMAS	28
2.5	CONCEITO DE CONSERVAÇÃO DE ENERGIA	29
2.6	EMISSIVIDADE	32
2.7	CAMPO DE VISÃO (FOV)	32
2.8	CAMPO DE VISÃO INSTANTÂNEO (IFOV).....	32
2.9	TONS DE COR.....	33
2.10	TÉCNICAS DE MEDIÇÃO TERMOGRÁFICA.....	33
3	ANÁLISE DE COMPONENTES PRINCIPAIS - ACP	36
3.1	PROCEDIMENTO PARA UMA ANÁLISE DE COMPONENTES PRINCIPAIS.	37
3.2	CONTRIBUIÇÃO DE CADA COMPONENTE:	39
3.3	ANÁLISE DE COMPONENTES PRINCIPAIS EM IMAGENS PARA RECONHECIMENTO DE PADRÕES	39
4	LÓGICA FUZZY	41
4.1	TEORIA DOS CONJUNTOS FUZZY	41
4.2	OPERAÇÕES COM CONJUNTOS FUZZY	42
4.3	VARIÁVEIS LINGÜÍSTICAS.....	44
4.4	FUNÇÕES DE PERTINÊNCIA	45
4.4.1	Triangular:	45
4.4.2	Trapezoidal.....	46

4.4.3 Gaussiana	46
4.5 CONTROLE FUZZY	47
4.6 FUZZYFICAÇÃO	47
4.7 BASE DE REGRAS	48
4.8 MÁQUINA DE INFERÊNCIA	49
4.9 DEFUZZYFICAÇÃO	49
4.10 ANFIS - ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM.....	50
5 REDES NEURAIS ARTIFICIAIS	52
5.1 INTRODUÇÃO	52
5.2 APRENDIZAGEM	58
5.2.1 Aprendizagem Por Correção De Erro.....	60
5.2.2 Aprendizagem baseada em memória	61
5.2.3 Aprendizagem Hebbiana.....	61
5.2.4 Aprendizagem competitiva.....	62
5.2.5 Aprendizagem de Boltzmann.....	63
5.3 REDE NEURAL BACKPROPAGATION.....	64
5.3.1 Introdução.....	64
5.3.2 Ajuste dos Pesos	65
5.3.3 Observações	69
5.4 PROJETO DE UMA REDE NEURAL	69
5.5 VANTAGENS	70
5.6 APLICAÇÕES	71
6 APLICAÇÃO DA METODOLOGIA PROPOSTA	72
6.1 INTRODUÇÃO	72
6.2 FERRAMENTAS E EQUIPAMENTOS UTILIZADOS	72
6.3 BANCO DE DADOS.....	73
6.4 VARIÁVEIS PARA O SISTEMA INTELIGENTE (ANFIS E RNA) REDE NEURAL	75
6.5 ANÁLISE DO ÓLEO ISOLANTE.....	75
6.6 PREPARAÇÃO DOS DADOS	76
6.6.1 Preparação dos Dados da Imagens Termográficas.....	77
6.6.2 Compressão dos dados utilizando Análise de Componentes Principais	80
6.7 REDE NEURAL	81
6.7.1 Introdução.....	81

6.7.2 Embaralhar os Dados	82
6.7.3 Normalização dos Dados	82
6.7.4 Arquitetura da Rede	83
6.8 LÓGICA FUZZY	84
6.8.1 Introdução.....	84
6.8.2 Sistema de Inferência Neuro-Fuzzy Adaptativo (ANFIS).....	84
6.8.3 Determinação do Sistema de Inferência Fuzzy (FIS).....	85
6.8.4 ANFIS.....	86
6.9 RESULTADOS.....	87
6.9.1 Abordagem Redes Neurais	87
6.9.2 Abordagem Lógica Fuzzy.....	88
7 CONCLUSÕES E CONSIDERAÇÕES FINAIS	90
7.1 CONTRIBUIÇÕES.....	90
7.2 SUGESTÕES DE TRABALHOS FUTUROS	92
7.3 CONSIDERAÇÕES FINAIS.....	92
REFERÊNCIAS	94

1 INTRODUÇÃO

Os transformadores de potência são extremamente importantes e as técnicas que monitoram o desempenho destes equipamentos sem retirá-los de operação devem ser largamente pesquisadas (SILVA, 2012). Considerando os transformadores de baixa tensão, para a coleta do óleo isolante a ser analisado, tem-se que retirá-lo de operação, pois não possuem em sua maioria a torneira para a coleta de material. A retirada de operação envolve paradas programadas podendo tornar-se muito onerosa, pois há ambientes de produção que devem funcionar ininterruptamente e paradas como essa representam perda financeira sem considerar a equipe técnica envolvida neste procedimento e o tempo disponibilizado para esta intervenção. Programas de manutenção preditiva que utilizam termografia podem ajudar. Isso tudo, somado a um bom custo/benefício que a termografia oferece, torna essa tecnologia bastante atraente do ponto de vista comercial (YADAV et al., 2008).

A inspeção termográfica é uma técnica não destrutiva realizada com a utilização de sistemas infravermelhos para a medição de temperaturas ou observações de padrões diferenciais de distribuição de calor, com o objetivo de propiciar informações relativas à condição operacional de um componente, equipamento ou processo. O infravermelho é uma frequência eletromagnética naturalmente emitida por qualquer corpo, com intensidade proporcional a sua temperatura.

Com o uso da termografia, várias técnicas e metodologias são adotadas por empresas. Uma dessas metodologias adotada por empresas como Petrobrás, Eletronuclear, Du Pont, Siemens e General Motors do Brasil, considera o uso de um critério de aquecimento para se definir um diagnóstico sobre o componente, o qual pode ser classificado como “Normal”, “Suspeita de Falha”, “Falha Potencial” ou “Falha Iminente” (VERATTI, 1999). Essa metodologia, denominada Critério Flexível de Classificação de Aquecimentos – CFCA, foi proposta em 1982 pelo Engº Atílio Bruno Veratti, após pesquisa em trabalhos semelhantes realizados por empresas norte-americanas, inglesas e francesas (ENGETERMO, 2009).

Para buscar a máxima operacionalidade, produtividade e confiabilidade do sistema, a aplicação da termografia na operação do sistema elétrico de Furnas considera as temperaturas e/ou aquecimentos internos dos equipamentos e classifica-os em “Baixa probabilidade de falha”, “Moderada probabilidade de falha”, “Alta probabilidade de falha” ou “Alta probabilidade de falha iminente” (ARAÚJO; BARBOSA; SINISCALCHI, 2008).

A CEMIG avalia para-raios de SiC e ZnO baseando-se na diferença de temperatura entre as fases dos equipamentos (ÁLVARES, 2008). Estudo desenvolvido pela LIGHT em parceria com o CEPREL diagnosticou defeitos em transformadores de potência utilizando como referência os valores máximos e mínimos de temperatura e a diferença entre estes valores, agrupando os equipamentos por famílias, aplicando-se um tratamento estatístico cujos resultados foram ratificados através de ensaios laboratoriais dos componentes inspecionados (BARRETO JR et al., 2001).

Todos os trabalhos com o uso de termografia na avaliação de transformadores e subestações, concentram suas técnicas em avaliações dos isoladores poliméricos. Uma aplicação de grande envergadura seria a avaliação do estado de transformadores que não incorporam a possibilidade de coleta de óleo isolante para análise, sem que seja necessário o desligamento da carga

Na concessionária de energia o número de transformadores de rede de distribuição pode ser de algumas centenas a milhares com alto custo de danos materiais decorrentes de suas falhas (SEFIDGARAN; MIRZAI; EBRAHIMZADEH, 2011). Os transformadores de potência estão entre os elementos mais caros e críticos em um sistema de potência. Falhas em transformadores podem ser catastróficas e quase sempre incluem danos internos irreversíveis. Portanto, todos os transformadores devem ser acompanhados de perto e continuamente a fim de garantir a sua disponibilidade máxima (NORTON, 2011). A manutenção adequada incluindo o condicionamento do óleo isolante podem prolongar por mais de 60 anos a vida-útil dos transformadores de potência (CHATTERJEE et al., 2013). De acordo com (JAFFERY; DUBEY, 2014) os investimentos em novos equipamentos para o sistema elétrico diminuíram drasticamente nos últimos 20 anos e muitos transformadores estão funcionando muito além de sua capacidade prevista. Métodos de manutenção periódica convencionais necessitam que o equipamento esteja fora de operação (off-line) para que se verifique as condições de operação. Isso tem levado a necessidade de desenvolvimento de ferramentas de diagnósticos mais precisas com monitoramento do sistema em operação (on-line) utilizando diversas técnicas analíticas além de inteligência artificial. Um dos métodos mais informativos para falhas incipientes nos transformadores de potência é a Análise dos Gases Dissolvidos (DGA). Tradicionalmente retira-se uma amostra do óleo isolante do transformador, envia-se para o laboratório e se aguarda o resultado da análise cromatográfica dos gases dissolvidos. Ao menos que exista qualquer suspeita de problema no transformador, as amostras podem ser colhidas em intervalos de até um ano. Isto significa que se uma falha

se desenvolver em um período mais curto do intervalo de amostragem esta pode levar ao comprometimento do transformador (CHATTERJEE et al., 2013).

A maioria das técnicas disponíveis no mercado baseiam-se na medição da concentração de gases em uma amostra extraída do transformador, principalmente utilizando a cromatografia. Estes dispositivos são caros e exigem operações off-line além da necessidade de preparação da amostra. Assim, novas técnicas, economicamente aceitáveis e confiáveis, devem ser exploradas (JAFFERY; DUBEY, 2014)

Recentemente a tecnologia de termografia infravermelha tem ganho mais reconhecimento e aceitação por ser uma técnica não invasiva e não destrutiva. Na técnica baseada na termografia infravermelha, o diagnóstico é realizado através da análise da imagem térmica capturada pela câmera infravermelha. Está implícito que a vida dos equipamentos elétricos é radicalmente reduzida com o aumento da temperatura. A técnica de medição de temperatura oferece muitas vantagens como tempo de resposta rápido, amplas faixas de temperatura, altamente confiável e inofensivo, alta resolução espacial e uma abordagem muito lucrativa para o acompanhamento dos sistemas de energia elétrica (JAFFERY; DUBEY, 2014).

Isto faz com que alguns pesquisadores busquem métodos de monitoramento on-line de transformadores de potência e de baixo custo. Hoje em dia os usos de técnicas avançadas de inteligência artificial são capazes de interpretar corretamente vários fenômenos e detectar com sucesso falhas incipientes (NORTON, 2011).

Vários trabalhos utilizando inteligência artificial são apresentados como propostas de interpretação do significado dos gases dissolvidos em comparação com os métodos IEC e IEEE como em (CASTRO; MIRANDA, 2005) utilizando Redes Neurais Artificiais e em (DA SILVA; GARCEZ CASTRO; MIRANDA, 2012) utilizando Lógica Fuzzy, além de (MORAIS; ROLIM, 2006) que propõe utilizar as duas técnicas

Este trabalho propõe utilizar a técnica de termografia infravermelha para diagnosticar a existência de uma falha incipiente no transformador, decodificando as imagens térmicas com técnicas de inteligência artificial: Redes Neurais Artificiais e Lógica Fuzzy, comparando os resultados com os métodos IEC e IEEE, utilizando Redes Neurais Artificiais e Lógica Fuzzy como técnica de classificação (BEZERRA et al., 2008).

1.1 DESCRIÇÃO DO PROBLEMA

De todas as técnicas apresentadas há dois elementos em comum: todas elas buscam critérios de classificação e todas necessitam de um especialista (profissional da área) para avaliação.

A aplicação da termografia infravermelha no sistema elétrico pode ser extremamente efetiva, prevenindo falhas e perdas de produção, mas para alcançar seus benefícios, as limitações das pessoas envolvidas devem ser conhecidas e respeitadas, sob risco de não se obter os resultados desejados (ARAÚJO; BARBOSA; SINISCALCHI, 2008). Porém, o diagnóstico do profissional em termografia poderá variar com o treinamento, motivação e capacidade visual.

Tratando-se de apenas algumas subestações se justifica o uso de tal profissional, mas quando se trata de empresas, setores ou redes comerciais que possuem vários transformadores, pode tornar a equipe de manutenção muito onerosa para que as ações preventivas ocorram em um período estrategicamente adequado.

Este trabalho propõe o uso de técnicas de inteligência artificial para indicar possíveis tendências de problemas de falhas em transformadores, utilizando as temperaturas registradas nas imagens termográficas, em substituição a análise do profissional de termografia.

1.2 JUSTIFICATIVA

Atualmente, várias técnicas para o uso de imagens infravermelhas são desenvolvidas, mas são dependentes da avaliação de um profissional em termografia.

Não há estudos que utilizem técnicas de inteligência artificial para indicativo de defeitos em transformadores utilizando registros termográficos ou que sugiram a supressão da análise do profissional em termografia

Na avaliação de transformadores, várias técnicas são desenvolvidas para indicar possível defeitos nos equipamentos, porém se há a necessidade de avaliações em grande escala, vários profissionais capacitados são necessários.

O uso de técnicas de inteligência artificial pode auxiliar na varredura de várias amostras de registros termográficos de forma que o profissional em termografia poderia concentrar seus esforços apenas nas amostras que possuíssem indicativos de falhas, no caso

das avaliações do óleo isolante. Os profissionais em termografia seriam direcionados para análises complementares em equipamentos que já tivessem o diagnóstico de indicativo de falhas. Tal estratégia tornaria os recursos humanos disponíveis ainda mais eficientes.

Um aprofundamento subsequente neste tema pode também fazer surgir alternativas tecnológicas aos fabricantes destes equipamentos de medição (termovisores), adaptando-os ao uso por operadores sem conhecimento aprofundado das técnicas envolvidas na termografia. Tais adaptações poderiam ser alcançadas através da inserção das técnicas de inteligência artificial nos softwares de tratamento das imagens termográficas de seus instrumentos ou através da inserção no próprio hardware do termovisor.

1.3 OBJETIVOS

1.3.1 Objetivos Gerais

Desenvolver técnica de inteligência artificial que permita estabelecer o estado do óleo isolante de transformadores de baixa tensão a partir de suas imagens térmicas.

1.3.2 Objetivos Específicos

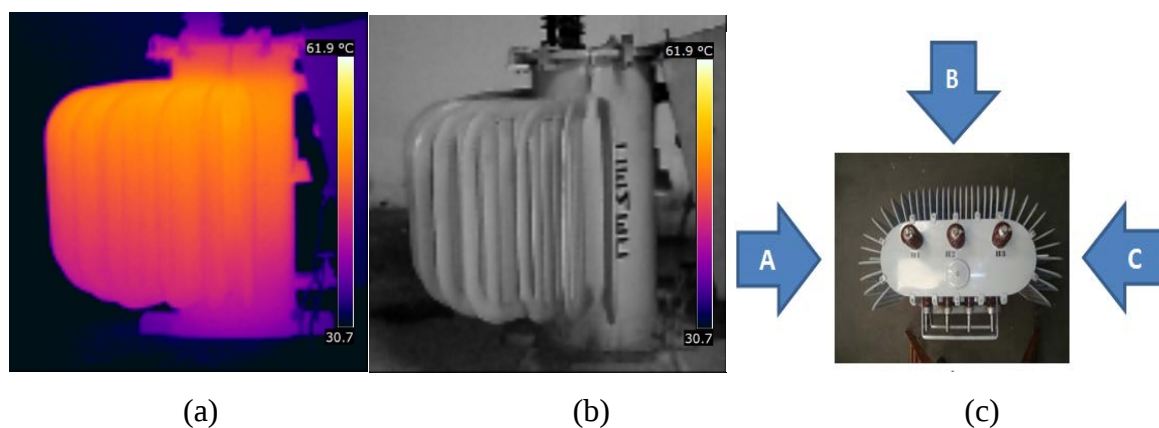
- Coletar e analisar o óleo isolante de transformadores de baixa tensão (análise físico-química e cromatográfica).
- Registrar as imagens termográficas de transformadores de baixa tensão e suas grandezas elétricas no instante deste registro.
- Utilizar a técnica de Análise de Componentes Principais nos registros de temperatura que representam as imagens termográficas e assim comprimir os dados registrados.
- Utilizar Redes Neurais Artificiais e Lógica Fuzzy para correlacionar as imagens termográficas com os resultados das análises físico-químicas dos óleos isolantes dos transformadores avaliados.

1.4 METODOLOGIA

Para este estudo são avaliados transformadores de baixa tensão instalados e operando em instalações comerciais distintas. São coletadas amostras do óleo isolante e realizadas análises físico-química e cromatográfica em laboratório.

Utilizando um termovisor, são registradas as imagens termográficas das três faces do transformador, como pode ser ilustrado na Figura 1-1. Estas imagens são convertidas em matrizes de dimensões 120x120, as quais constituem representações discretas das imagens originais de resolução de 120x120 pixels.

Figura 1-1: (a) Imagem termográfica da face A do transformador, (b) imagem real do transformador e (c) indicações do posicionamento das faces



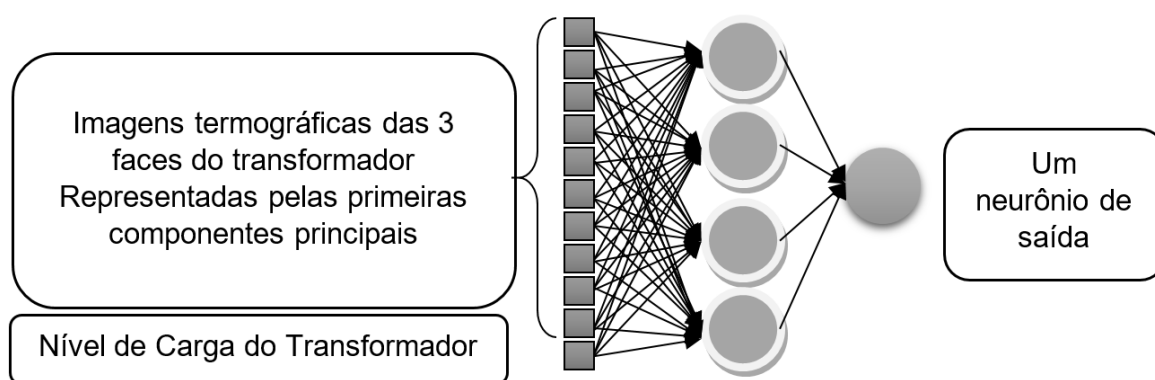
Fonte: próprio autor

Em uma imagem discreta de 120x120 o total de células desta matriz é equivalente a 14.400 registros e considerando que são 3 imagens a representar um transformador, são totalizados 43.200 registros. Logo, qualquer cálculo envolvendo dados dessa magnitude toma um esforço computacional elevado e fica clara a necessidade da redução da dimensionalidade, para que os dados possam ser manipulados (KITANI; THOMAZ, 2006). A Análise de Componentes Principais (ACP) reduz a dimensionalidade dos dados, concentrando as informações significativas nas primeiras componentes geradas pela transformação (PRICHOA; RIBEIRO, 2013). Tal transformação permite, ainda, descartar quaisquer componentes que explicam somente uma pequena proporção da variação nos dados (ARAÚJO; BARBOSA; SINISCALCHI, 2008).

Outra variável coletada no momento do registro das imagens termográficas é a potência da carga à qual estava submetido o transformador.

Ao se optar pelo uso de Rede Neural Artificial (RNA) deve ser lembrado que das decisões mais importantes em sua construção é a escolha das variáveis nas quais a rede irá se basear para poder aprender. Para o caso específico se deseja encontrar nas imagens termográficas características que as correlacionem com o resultado da análise físico-química do óleo isolante. Desta forma, foi definido que a RNA teria como variáveis de entrada a imagem termográfica, representada pelas primeiras componentes principais, e o nível de carga do transformador (relação entre a potência da carga a qual estava submetido no momento do registro da imagem termográfica e a potência máxima do transformador). Para o sinal de saída foi utilizado o resultado da análise do óleo isolante, conforme ilustração contida na Figura 1-2.

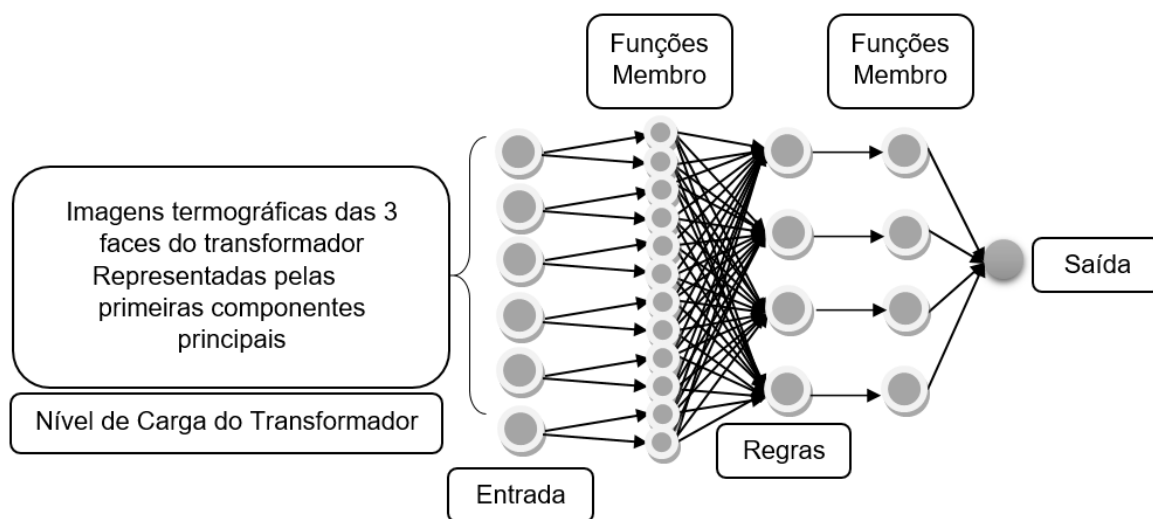
Figura 1-2: Arquitetura da rede neural



Fonte: próprio autor

A Lógica Fuzzy é uma técnica de inteligência artificial que avalia o grau de pertinência de determinado elemento. São estabelecidas funções membro para quantificar a probabilidade de o elemento fazer parte desse determinado grupo. No caso específico se deseja encontrar nas imagens termográficas características que as correlacionem com o resultado da análise físico-química do óleo isolante. Assim, da mesma forma que na formação de uma RNA, se define que o sistema de inferência fuzzy teria como variáveis de entrada a imagem termográfica representada pelas primeiras componentes principais e o nível de carga do transformador (relação entre a potência da carga a qual estava submetido no momento do registro da imagem termográfica e a potência máxima do transformador). Para o sinal de saída é utilizado o resultado da análise do óleo isolante, conforme ilustrações registradas na Figura 1-3.

Figura 1-3: Sistema de Inferência Fuzzy



Fonte: próprio autor

O resultado esperado é que sistemas em RNA para solucionar problemas complexos e LÓGICA FUZZY para solucionar problemas representados pelas suas relações de entrada e saída, demonstrem-se eficientes e deem margens para serem utilizadas como indicativo para decidir sobre uma intervenção em transformadores de baixa tensão. Se, por exemplo, uma empresa necessita efetuar uma intervenção, em vários equipamentos, com muito mais velocidade é possível efetuar a medição termográfica e, utilizando esta técnica, traçar um roteiro de intervenção priorizando os equipamentos cujas respostas da rede neural indicarem propensão a falha.

1.5 ESTRUTURA DO TRABALHO

Este trabalho está organizado em oito capítulos, que serão descritos a seguir. Neste primeiro capítulo é feita uma introdução ao tema, enfocando-se a justificativa, objetivos e importância.

No capítulo 2 apresenta-se uma revisão da literatura em Termografia e Termovisores com um enfoque nas técnicas para a coleta das imagens e as variáveis a serem coletadas para refinamento das temperaturas medidas.

No capítulo 3 tem-se uma revisão da literatura em Análise de Componentes Principais. O uso desta técnica é fundamental para a compressão dos dados das imagens

termográficas de forma a permitir o uso de sistemas computacionais para as emulações necessárias.

Nos capítulos 4 e 5 tem-se uma revisão de literatura das ferramentas de inteligência artificial a Lógica Fuzzy e as Redes Neurais Artificiais, destacando-se o uso desta técnica para correlacionar as imagens termográficas com as condições dos transformadores após a avaliação do óleo isolante.

No capítulo 6, apresenta-se o desenvolvimento da metodologia proposta.

Finalizando, no capítulo 7 são apresentadas as conclusões e propostas para trabalhos futuros.

2 TERMOGRAFIA INFRAVERMELHA E TERMOVISORES

A termografia permite a aferição das temperaturas superficiais dos objetos observados, mapeando imagens que identifiquem a distribuição dessas temperaturas ao longo do objeto. Todos os corpos a uma temperatura acima do zero absoluto ($-273,15\text{ }^{\circ}\text{C}$) emitem calor, que é constantemente absorvido e reemitido pelo próprio corpo e por tudo o que o rodeia. Termografia é o termo usado para descrever o processo de tornar essas radiações térmicas visíveis e susceptíveis de interpretação. De uma forma simplista, pode ser dito que a termografia é a técnica que estende a visão humana através do espectro infravermelho (CARAMALHO, 2012).

Os termovisores operam com base nos princípios da termografia infravermelha. Este tipo de aparelho é usado como ferramenta de teste de baixo custo, muitas vezes proporcionando lucratividade, para a resolução de problemas, manutenção e inspeção de sistemas elétricos e mecânicos.

2.1 TERMOGRAFIA INFRAVERMELHA

Termografia infravermelha é a ciência que envolve o uso de dispositivos ópticos eletrônicos para detectar e medir a radiação e correlacioná-la com a temperatura da superfície. Radiação é o movimento do calor que ocorre quando há propagação de energia radiante (ondas eletromagnéticas) sem a presença de um meio direto de transferência. A termografia infravermelha moderna é realizada através do uso de dispositivos ópticos eletrônicos para detectar e medir a radiação e correlacioná-la com a temperatura da estrutura ou do equipamento que está sendo inspecionado.

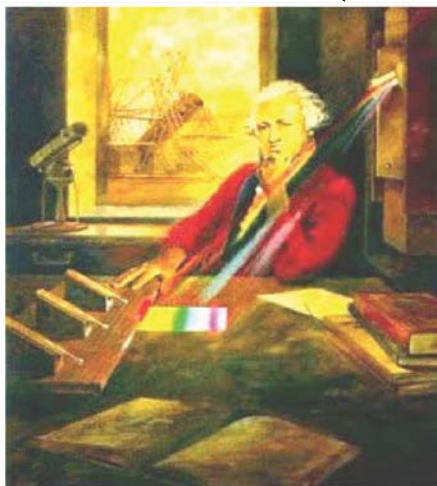
O fato dos seres humanos terem limitações físicas na detecção de calor motivou o desenvolvimento de dispositivos mecânicos e eletrônicos hipersensíveis ao calor. Esses dispositivos são a forma convencional utilizada para a inspeção térmica, abrangendo inúmeras aplicações.

2.2 A HISTÓRIA DA TECNOLOGIA INFRAVERMELHA

A palavra “infravermelho” é derivada de “além do vermelho”, referindo-se ao lugar ocupado por este comprimento de onda no espectro de radiação eletromagnética. O termo “termografia” deriva de radicais de palavras que significam “imagem da temperatura”. As raízes da termografia podem ser atribuídas ao astrônomo alemão Sir William Herschel que, em 1800 realizou experiências com a luz solar.

Herschel descobriu a radiação infravermelha atravessando a luz solar por um prisma e colocando um termômetro nas várias cores onde ele pode medir a temperatura de cada cor utilizando um termômetro sensível de mercúrio. Herschel, ilustrado na Figura 2-1, descobriu que a temperatura aumentava quando saía da luz vermelha e entrava em uma área à qual ele se referiu como “calor escuro”. “Calor escuro” era a região do espectro eletromagnético atualmente conhecida como calor infravermelho e reconhecida como radiação eletromagnética.

Figura 2-1: Sir Willian Herschel (1738 – 1832)



Fonte: (CARAMALHO, 2012)

Vinte anos mais tarde, o físico alemão Thomas Seebeck descobriu o efeito termoelétrico. Isso levou à invenção do termomultiplicador, uma versão inicial do termopar, por parte do italiano Leopoldo Nobili em 1829. Este simples dispositivo de contato é baseado na premissa de que a diferença de tensão entre dois metais diferentes muda com a temperatura. Macedonio Melloni, companheiro de Nobili, logo aprimorou o termomultiplicador criando a termopilha (uma combinação de termomultiplicadores em série) e enfocou a radiação térmica sobre ela de forma tal que fosse possível detectar o calor do corpo a uma distância de 9,1m.

Em 1880, o astrônomo americano Samuel Langley usou o bolômetro para detectar o calor do corpo de uma vaca a 304m de distância. Em vez de medir a diferença de tensão, o bolômetro mede a variação da temperatura. Sir John Herschel, filho de Sir William Herschel, usando um dispositivo chamado evaporígrafo, conseguiu produzir a primeira imagem infravermelha em 1840. A imagem térmica resultou da evaporação diferencial de uma fina película de óleo e foi vista através da reflexão da luz sobre a película de óleo.

Um termovisor é um dispositivo que detecta padrões de calor no espectro de comprimento de onda do infravermelho sem contato direto com o equipamento, conforme ilustrado na Figura 2-2.

Figura 2-2: Termovisor



Fonte: (FLUKE CORPORATION, 2009)

As versões iniciais dos termovisores ficaram conhecidas como detectores de fotocondução. De 1916 a 1918, o inventor americano, Theodore Case, fez experiências com detectores de fotocondução e conseguiu produzir um sinal através da interação direta com fótons em vez de através da geração de calor. O resultado foi um detector de fotocondução mais rápido e mais sensível. Durante as décadas de 1940 e 1950, a tecnologia de geração de imagens térmicas expandiu-se para atender um número cada vez maior de aplicações militares. Os cientistas alemães descobriram que, através do resfriamento do detector de fotocondução, era possível aumentar o desempenho geral.

Só depois da década de 1960 é que a geração de imagens térmicas foi usada para aplicações extramilitares. Embora os sistemas iniciais de geração de imagens térmicas fossem incômodos, lentos para a aquisição de dados e oferecessem baixa resolução, eles chegaram a ser usados para aplicações industriais tais como a inspeção de grandes sistemas de transmissão e distribuição elétricas. Avanços contínuos na década de 1970 voltados para

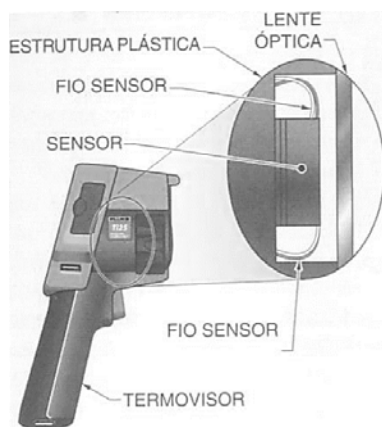
aplicações militares produziram os primeiros sistemas portáteis que podiam ser usados para aplicações tais como diagnósticos de construções e testes não-destrutivos de materiais.

Na década de 1970, os sistemas de geração de imagens térmicas eram duráveis e confiáveis, mas a qualidade das imagens era ruim quando comparada com os modernos termovisores. Já no início da década de 1980, a geração de imagens térmicas estava sendo amplamente usada para fins medicinais, na indústria de ponta e na inspeção de construções. Os sistemas de geração de imagens térmicas foram aprimorados para produzir imagens totalmente radiométricas de forma que as temperaturas radiométricas pudessem ser medidas em qualquer lugar da imagem. Imagem radiométrica é a imagem térmica que contém os cálculos das medidas de temperatura para vários pontos dentro de uma imagem.

Os resfriadores de termovisor foram aperfeiçoados para substituir o gás comprimido ou liquefeito anteriormente usado para resfriar esses aparelhos. Sistemas mais econômicos de geração de imagens térmicas de vidicon piroelétrico (VPE), baseados em tubos também foram desenvolvidos e amplamente usados. Embora não radiométricos, os sistemas VPE de geração de imagens térmicas eram leves, portáteis e funcionavam sem resfriamento.

No final da década de 1980, um novo dispositivo conhecido como arranjo de plano focal (APF), antes usado pelas forças armadas, foi lançado no mercado comercial. Arranjo de plano focal (APF) é um dispositivo sensor de imagens que consiste em um arranjo (normalmente retangular) de detectores sensíveis ao infravermelho no plano focal da lente, conforme ilustrado na Figura 2-3.

Figura 2-3: Arranjo de plano focal



Fonte: (FLUKE CORPORATION, 2009)

Este foi o avanço significativo com relação aos detectores originais e o resultado foi o aumento da qualidade das imagens e da resolução espacial. Os arranjos típicos encontrados

nos modernos termovisores têm pixels que variam entre 16 X 16 e 640 X 480. Um pixel, neste sentido, é o menor elemento independente de um APF que pode detectar energia infravermelha.

O desenvolvimento da tecnologia do APF que utiliza vários detectores tem aumentado desde o ano 2000.

Os custos desses sistemas diminuíram em mais de 10 vezes na última década e a qualidade tem aumentado drasticamente. Além disso, o uso de softwares de computador para o processamento de imagens também evoluíram substancialmente (FLUKE CORPORATION, 2009). Praticamente todos os modernos sistemas de infravermelho oferecidos no mercado utilizam softwares que facilitam a realização de análises e a geração de relatórios. É possível criar relatórios com rapidez e enviá-los por meio eletrônico através da Internet ou mantê-los em formato comum, tal como PDF, e gravá-los em vários dispositivos de armazenamento digital.

2.3 OPERAÇÃO DO TERMOVISOR

A finalidade de um termovisor é detectar a radiação infravermelha emitida pelo alvo. Alvo é um objeto que deve ser inspecionado com o uso de um termovisor.

A radiação infravermelha é focalizada através do sistema óptico do termovisor. O sinal produzido pelo termovisor é convertido em uma imagem eletrônica (termograma) sobre uma tela. Termograma é a imagem de um alvo eletronicamente processada sobre uma tela onde diversos tons de cor correspondem à distribuição da radiação infravermelha sobre a superfície do alvo.

2.4 RESOLUÇÃO DE PROBLEMAS

A termografia infravermelha tem uma função importante na resolução de problemas de operações comerciais e industriais. Geralmente, questões relativas à condição dos equipamentos surgem a partir da detecção de algum indício ou condição anormal. Isso pode ser tão obvio e simples quanto a observação de uma vibração, um som ou uma leitura de temperatura, ou ser tão sutil que a causa geradora do problema pode ser difícil ou impossível de discernir.

Um sinal térmico é uma imagem colorida representativa da energia infravermelha, ou do calor, emitida por um objeto. Os principais benefícios da termografia infravermelha

envolvem a possibilidade de rápida realização dos testes, sem destruir o equipamento. Além disso, como os termovisores não requerem contato, eles podem ser usados enquanto o equipamento ou componente está em operação.

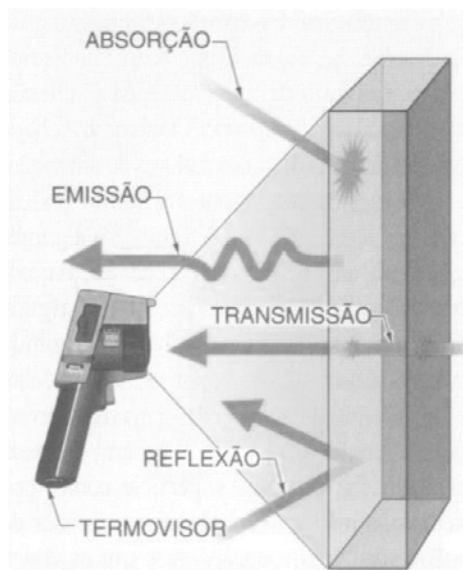
Nem sempre é simples saber exatamente quais condições são necessárias para se resolver o problema de um determinado equipamento. Além da experiência do profissional de termografia, uma profunda compreensão das variáveis, como transferência de calor, radiometria, utilização de câmera e funcionamento dos equipamentos a serem avaliados, são todos elementos necessários para uma solução de problemas bem-sucedida. Radiometria é a detecção e a medição de energia eletromagnética radiante, especificamente na parte infravermelha do espectro.

2.5 CONCEITO DE CONSERVAÇÃO DE ENERGIA

A luz e a radiação infravermelha se comportam de forma semelhante quando interagem com vários materiais. A radiação infravermelha é refletida por alguns tipos de superfícies, tais como a chapa de metal embaixo de queimadores de um fogão. As reflexões de corpos quentes e frios podem ser vistas através de termovisores em algumas superfícies, tais como metais brilhantes, às quais chamamos de “espelhos térmicos”. Em alguns casos, a radiação infravermelha é transmitida por uma superfície, como, por exemplo, a lente de um sistema de geração de imagens infravermelhas. A radiação infravermelha também pode ser absorvida por uma superfície como, por exemplo, a mão próxima de um queimador de fogão aceso. Neste caso, acontece uma mudança de temperatura fazendo com que a superfície emita mais energia.

Transmissão é a passagem de energia radiante através de um material ou estrutura. A radiação infravermelha também pode ser absorvida em uma superfície, causando a mudança de temperatura e a emissão de mais energia na superfície do corpo. Absorção é a captação da energia radiante. Emissão é a descarga de energia radiante. Embora um sistema de geração de imagens infravermelhas possa ler radiações refletidas, transmitidas e emitidas, somente a energia absorvida ou emitida é que afeta a temperatura da superfície, conforme podemos observar na ilustração da Figura 2-4.

Figura 2-4: A radiação pode ser refletida, transmitida, absorvida ou emitida



Fonte: (FLUKE CORPORATION, 2009)

Além disso, a quantidade de calor irradiado por uma superfície é determinada pela eficiência com que uma superfície emite energia. A maioria dos materiais não-metálicos, tais como superfícies pintadas ou a pele humana, emitem energia com eficiência. Isso significa que, à medida que sua temperatura aumenta, eles irradiam muito mais energia, assim como no caso do queimador do fogão.

Outros materiais, em grande parte metais sem pintura ou não muito oxidados, são menos eficientes em irradiar energia. Quando uma superfície metálica sem revestimento é aquecida, há comparativamente pouco aumento de transferência de calor irradiante, e é difícil ver a diferença entre uma superfície metálica fria e quente, seja com nossos olhos ou através de um sistema de geração de imagens térmicas. Metais não revestidos normalmente têm baixa emissividade (baixa eficiência de emissão). A emissividade é indicada por um valor entre 0,0 e 1,0. Uma superfície cujo valor seja 0,10, típica de cobre polido, emite pouca energia em comparação com a pele humana, que tem emissividade de 0,98.

Um dos desafios de se usar um termovisor é que esses instrumentos mostram uma energia que é invisível ao olho humano. Às vezes, isso pode ser confuso. Superfícies de baixa emissividade, tais como os metais, não só emitem energia de forma ineficiente, mas também refletem o calor ao seu redor. Quando é feita a leitura de uma superfície através de

um sistema de geração de imagens térmicas, a imagem mostra uma combinação de radiação infravermelha emitida e refletida.

Vários outros fatores podem afetar a emissividade do material. Além do tipo de material, a emissividade também pode variar com a condição da superfície. A temperatura e o comprimento de onda. A emissividade efetiva de um corpo também pode variar com o ângulo de visão.

Não é difícil caracterizar a emissividade da maioria dos materiais que não são metais polidos. Muitos materiais já foram caracterizados e os valores de suas emissividades podem ser encontrados em tabelas de emissividade. Os valores de emissividades só devem ser usados para orientação. Como a emissividade exata de um material pode variar com relação a esses valores, profissionais de termografia experientes também precisam entender como medir o valor real. A Tabela 2-1 apresenta os valores de emissividade de alguns materiais.

Tabela 2-1: Valores de emissividade de alguns materiais comuns

Material em teste	Emissividade	Material em teste	Emissividade
Asfalto	0,90 a 0,98	Tecido (preto)	0,98
Concreto	0,94	Pele (humana)	0,98
Cimento	0,96	Couro	0,75 a 0,80
Areia	0,90	Carvão (pó)	0,96
Terra	0,92 a 0,96	Verniz	0,80 a 0,95
Água	0,92 a 0,96	Verniz (fosco)	0,97
Gelo	0,96 a 0,98	Borracha (preta)	0,94
Neve	0,83	Plástico	0,85 a 0,95
Vidro	0,90 a 0,95	Madeira	0,90
Cerâmica	0,90 a 0,94	Papel	0,70 a 0,94
Mármore	0,94	Óxidos de Crômio	0,81
Gesso	0,80 a 0,90	Óxidos de Cobre	0,78
Argamassa	0,89 a 0,91	Óxidos de Ferro	0,78 a 0,82
Tijolo	0,93 a 0,96	Têxteis	0,90

Cavidades, frestas e buracos emitem energia térmica a uma taxa maior que as superfícies ao seu redor. O mesmo se aplica à luz visível. A pupila do olho humano é preta porque é uma cavidade, e a luz que entra nela é absorvida. Quando a luz é completamente

absorvida por uma superfície, dizemos que ela é “preta ou negra”. A emissividade de uma cavidade se aproxima de 0,98 quando sua profundidade é sete vezes maior que a sua largura (FLUKE CORPORATION, 2009).

2.6 EMISSIVIDADE

Metais não pintados ou não fortemente oxidados são difíceis de ler em um termovisor porque emitem muito pouco ou refletem muito. Independente de se estar apenas examinando os padrões térmicos ou realmente fazendo uma medida de temperatura radiométrica, é preciso levar em conta esses fatores.

Embora as tabelas de correção da emissividade possam ser úteis para entender como um material se comporta, na realidade uma correção feita na maioria das superfícies de baixa emissividade pode gerar erros inaceitavelmente grandes. Superfícies de baixa emissividade devem ser alteradas de alguma forma, como, por exemplo, cobrindo-as com fita isolante ou tinta, para aumentar a emissividade. Isso dá precisão e praticidade à interpretação e à medição.

2.7 CAMPO DE VISÃO (FOV)

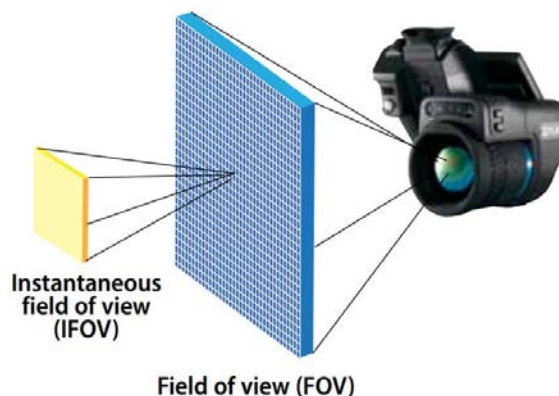
Campo de visão (FOV) é uma especificação que define o tamanho do que é visto no termovisor. A lente é a parte que tem maior influência sobre qual será o FOV, independentemente do tamanho do arranjo. Arranjos grandes, entretanto, fornecem mais detalhes, independente da lente usada, em comparação com arranjos estreitos.

2.8 CAMPO DE VISÃO INSTANTÂNEO (IFOV)

Campo de visão instantâneo (IFOV) é uma especificação usada para descrever a capacidade que um termovisor tem de decompor os detalhes espaciais (resolução espacial). Quando projetado a partir do detector através da lente, o IFOV dá o tamanho de um corpo que pode ser visto a uma determinada distância.

Uma medida do IFOV é a resolução de medição de um termovisor que descreve o corpo de menor tamanho que pode ser medido a uma determinada distância. As definições de FOV e IFOV podem ser observadas na ilustração da Figura 2-5.

Figura 2-5: FOV e IFOV



Fonte: (FLIR SYSTEMS, 2015)

2.9 TONS DE COR

Os tons de cor são um esquema usado para mostrar as variações e os padrões térmicos de uma imagem térmica. Seja em inspeções ou em análises, o objetivo é selecionar o tom que melhor identifique e comunique o problema. O ideal é que se escolha um termovisor que permita ao usuário selecionar ou mudar o tom desejado na câmera ou no software. Certas aplicações, por exemplo, podem ser melhor analisadas em tons monocromáticos, tais como uma escala em cinza ou em âmbar. Outras situações podem ser mais fáceis de analisar e explicar em tons de ferro, azul-vermelho ou outro de alto contraste. Uma ampla seleção de tons de cores disponíveis confere ao profissional maior flexibilidade nas inspeções, análise e relatórios térmicos.

2.10 TÉCNICAS DE MEDIÇÃO TERMOGRÁFICA

Uma câmara de infravermelhos mede e reproduz em imagens a radiação de infravermelhos emitida pelos objetos. O fato da radiação resultar da temperatura de superfície do objeto, possibilita que a câmara calcule e mostre essa temperatura. No entanto, a radiação medida pela câmara não depende apenas da temperatura do objeto.

Os seguintes parâmetros do objeto devem ser introduzidos na câmara:

- a emissividade do objeto,
- a temperatura aparente refletida,
- a distância entre o objeto e a câmara,

- a umidade relativa,
- a temperatura da atmosfera.

O principal parâmetro do objeto a definir corretamente é a emissividade que, sintetizando, consiste na medição da gama de radiação emitida pelo objeto, comparativamente à que é emitida por um corpo negro perfeito com a mesma temperatura.

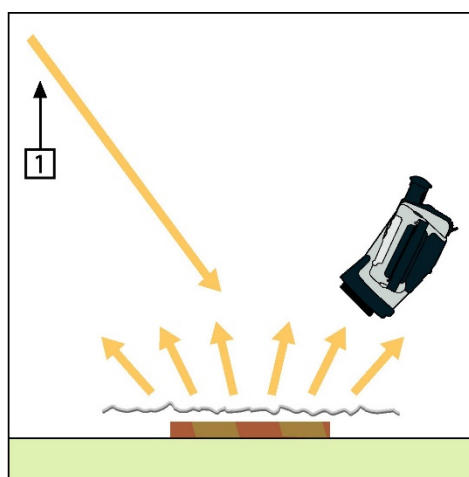
A distância consiste na medida linear entre o objeto e a lente frontal da câmara. Este parâmetro é utilizado para compensar dois fatos: que a radiação do alvo seja absorvida pela atmosfera entre o objeto e a câmara e que a radiação da própria atmosfera seja detectada pela câmara.

A câmara também pode compensar o fato da transmitância depender também da humidade relativa da atmosfera.

Quanto à temperatura aparente refletida, este parâmetro é utilizado para compensar a radiação refletida no objeto. Se a emissividade for baixa e a temperatura do objeto for relativamente diferente da refletida, será importante definir e compensar corretamente a temperatura aparente refletida.

Para a medição da temperatura refletida, podemos utilizar o Método Refletor que consiste em medir a temperatura refletida de uma folha de alumínio colocada à frente do objeto através da câmara com emissividade definida para 1,0 (Figura 2-6, onde 1 é a fonte de reflexão).

Figura 2-6: Método Refletor



Fonte: (FLUKE CORPORATION, 2009)

Além destes parâmetros a serem introduzidos na câmara, há fatores de perturbação que influenciam como o vento e os limites de distância de acordo com o tamanho do objeto que necessita de avaliação.

Durante a inspeção exterior, o efeito de refrigeração do vento deve ser tomado em consideração. Um sobreaquecimento medido a uma velocidade de vento de 5 m/s será aproximadamente duas vezes superior a 1 m/s. Uma temperatura excessiva medida a 8 m/s será 2,5 vezes mais alta a 1 m/s. Este fator de correção, que é baseado em medições empíricas, é normalmente aplicável até 8 m/s (FLUKE CORPORATION, 2009).

Existem, contudo, casos em que tem de efetuar a inspeção mesmo que o vento registre uma velocidade superior a 8 m/s. Existem muitos locais ventosos no mundo: ilhas, montanhas, etc. Porém é importante saber que os componentes sobreaquecidos detectados apresentariam uma temperatura consideravelmente superior a uma velocidade de vento inferior. O fator de correção empírico é indicado na Tabela 2-2.

Tabela 2-2: Fator de Correção da temperatura em função da velocidade do vento

Velocidade do vento (m/s)	Fator de correção
1	1
2	1,36
3	1,64
4	1,86
5	2,06
6	2,23
7	2,40
8	2,54

Quanto a ocorrência de uma medição de um objeto demasiadamente distante, pode-se considerar utilizar uma pequena regra geral para verificar se o instrumento é capaz de medir o objeto a uma determinada distância apenas multiplicando o IFOV por 3 (FLUKE CORPORATION, 2009). Exemplo: 25 graus correspondem a cerca de 437 mrad. Se a câmara tiver uma imagem de pixels 120×120 , o IFOV passa a ser $437/120 = 3,6$ mrad (3,6 mm/m) e a percentagem de dimensão do ponto é cerca de $1000/(3 \times 3,6) = 92:1$. Isto significa que a uma distância de 9,2 metros, o alvo tem de ter uma largura de pelo menos cerca de 0,1 metro ou 100 mm.

3 ANÁLISE DE COMPONENTES PRINCIPAIS - ACP

Um dos objetivos da análise de componentes principais é reduzir a dimensão do espaço original de entrada, projetando-o no subespaço de menor dimensão gerados pelas componentes mais relevantes. Além da extração de características, um outro interesse é extrair apenas aquelas mais relevantes dos dados, ou seja, projetar uma transformada que reduza o número efetivo de características, retendo ainda a maior parte do conteúdo de informação intrínseca dos dados. Este processo é conhecido como redução de dimensionalidade. Esta simplificação é extremamente atrativa do ponto de vista de implementação e de análise dos dados (LIRA et al., 2005). A solução baseada em componentes principais reduz o custo computacional do sistema classificador (CUENCA; SEIXAS; LEVY, 2009).

A reprodução da variabilidade total do sistema requer as p variáveis, porém, frequentemente, a maior parte dessa variabilidade pode ser explicada por um número pequeno $k < p$, de componentes principais. Nesse caso, existe praticamente a mesma quantidade de informações nas k componentes principais que nas p variáveis originais. As k componentes principais podem então substituir as p variáveis originais (MARQUES; MARQUES, 2005).

Análise de Componentes é um modelo fatorial no qual os fatores são baseados na variância total. Na análise de componentes, unidades são usadas na diagonal da matriz de correlação; esse procedimento implica computacionalmente que toda a variância é comum ou compartilhada (ARAUJO; COELHO, 2009). Análise de Componentes principais é um dos métodos multivariados mais simples. O objetivo da análise é tomar p variáveis $X_1, X_2, \dots, X_p$ e encontrar combinações destas para produzir índices $Z_1, Z_2, \dots, Z_p$ que sejam não correlacionados na ordem de sua importância, e que descreva a variação nos dados. A falta de correlação significa que os índices estão medindo diferentes "dimensões" dos dados, e a ordem é tal que $Var(Z_1) \geq Var(Z_2) \geq \dots \geq Var(Z_p)$, em que $Var(Z_1)$ denota a variância de (Z_1) . Os índices Z são então os componentes principais.

3.1 PROCEDIMENTO PARA UMA ANÁLISE DE COMPONENTES PRINCIPAIS

Uma análise de componentes principais começa com dados de p variáveis para n indivíduos, como indicado na Tabela 3-1.

Tabela 3-1: Forma de dados para uma análise de componentes principais, com variáveis $X_1, X_2, \dots, X_p$

e observações de n casos				
Caso	X_1	X_2	$\dots$	X_p
1	a_{11}	a_{21}	$\dots$	a_{1p}
2	a_{21}	a_{22}	$\dots$	a_{2p}
$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
n	a_{n1}	a_{n2}	$\dots$	a_{np}

O primeiro componente principal é então a combinação linear das variáveis $X_1, X_2, \dots, X_p$

$$Z_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p$$

que varia tanto quanto possível para os indivíduos, sujeitos à condição de que

$$Z_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p$$

Assim, $Var(Z_1)$, a variância de Z_1 , é tão grande quanto possível dada esta restrição sobre as constantes a_{1j} . A restrição é introduzida porque se isto não é feito, então $Var(Z_1)$ pode ser aumentada fazendo simplesmente crescer qualquer um dos valores a_{ij} .

O segundo componente principal

$$Z_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p$$

é escolhido de modo que $Var(Z_2)$ seja tão grande quanto possível sujeito à restrição de que

$$a_{21}^2 + a_{22}^2 + \dots + a_{2p}^2 = 1$$

e também à condição de que Z_1 e Z_2 tenham correlação zero para os dados. O terceiro componente principal,

$$Z_3 = a_{31}X_1 + a_{32}X_2 + \dots + a_{3p}X_p$$

é tal que a $Var(Z_3)$ seja tão grande quanto possível sujeita à restrição de que

$$a_{31}^2 + a_{32}^2 + \dots + a_{3p}^2 = 1$$

e também que Z_3 seja não correlacionada com ambas Z_1 e Z_2 . Posteriores componentes principais são definidos continuando da mesma maneira. Se existem p variáveis, então existirão no máximo p componentes principais.

Para se usar os resultados de uma análise de componentes principais, não é necessário saber como as equações, para os componentes principais, são obtidas.

Entretanto, é útil entender a natureza das equações. De fato, uma análise de componentes principais envolve encontrar os autovalores de uma matriz de covariâncias amostral.

A matriz de covariâncias é simétrica e tem a forma:

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{pmatrix}$$

em que o elemento c_{ii} na diagonal é a variância de X_i e o termo fora da diagonal c_{ij} é a covariância entre as variáveis X_i e X_j .

As variâncias dos componentes principais são os autovalores da matriz C . Existem p destes autovalores, alguns dos quais podem ser zero. Autovalores negativos não são possíveis para uma matriz de covariância. Assumindo que os autovalores estão ordenados como $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, então λ_i corresponde ao i -ésimo componente principal

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p$$

Em particular, $Var(Z_1) = \lambda_i$, e as constantes $a_{i1}, a_{i2}, \dots, a_{ip}$ são os elementos do correspondente autovetor, escalonado de modo que

$$a_{i1}^2 + a_{i2}^2 + \cdots + a_{ip}^2 = 1$$

Uma propriedade importante dos autovalores é que a soma deles é igual à soma dos elementos da diagonal (o traço) da matriz C . Isto é,

$$\lambda_1 + \lambda_2 + \cdots + \lambda_p = c_{11} + c_{22} + \cdots + c_{pp}$$

Porque c_{ii} é a variância de X_i e λ_i é a variância de Z_i , isto significa que a soma das variâncias dos componentes principais é igual à soma das variâncias das variáveis originais. Portanto, em certo sentido, os componentes principais contam com toda a variação nos dados originais.

A fim de evitar uma ou duas variáveis tendo indevida influência nos componentes principais, é usual codificar as variáveis $X_1, X_2, \dots, X_p$ para terem médias zero e variâncias um no início de uma análise. A matriz C então toma a forma

$$C = \begin{pmatrix} 1 & c_{12} & \cdots & c_{1p} \\ c_{21} & 1 & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & 1 \end{pmatrix}$$

em que $c_{ij} = c_{ji}$ é a correlação entre X_i e X_j . Em outras palavras, a análise de componentes principais é feita sobre a matriz de correlação. Neste caso, a soma dos termos da diagonal, e, portanto, a soma dos autovalores, é igual a p , o número de variáveis X .

3.2 CONTRIBUIÇÃO DE CADA COMPONENTE:

A contribuição de cada componente principal Z_i , em percentual, é calculada dividindo-se a variância de Z_i pela variância total. Representa a proporção de variância total explicada pelo componente principal Z_i .

$$\frac{Var(Z_i)}{\sum_{i=1}^p Var(Z_i)} \cdot 100 = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \cdot 100$$

A importância de um componente principal é avaliada por meio de sua contribuição, isto é, pela proporção de variância total explicada pelo componente. A soma dos primeiros k autovalores representa a proporção de informação retida na redução de p para k dimensões. Com essa informação podemos decidir quantos componentes vamos usar na análise, isto é, quantos componentes serão utilizados para diferenciar os indivíduos. Não existe um modelo estatístico que ajude nesta decisão. Para aplicações em diversas áreas do conhecimento o número de componentes utilizados tem sido aquele que acumula 70% ou mais de proporção da variância total (VARELLA, 2008).

$$\frac{Var(Z_i)}{\sum_{i=1}^p Var(Z_i)} \cdot 100 \geq 70\% \text{ onde } k < p$$

3.3 ANÁLISE DE COMPONENTES PRINCIPAIS EM IMAGENS PARA RECONHECIMENTO DE PADRÕES

Segundo a técnica de ACP, imagens podem ser tratadas como padrões em um espaço linear para efetuar reconhecimento estatístico. Sendo h o número de linhas de uma imagem e w o número de colunas, pode-se dizer que uma imagem é um padrão de $h \times w$ características ou um vetor no espaço $(h \times w)$ dimensional, o qual é chamado de “espaço de imagens”, representado por I .

Assim, dada uma imagem representada como uma matriz $h \times w$, pode-se construir sua representação como um vetor através de uma leitura coluna a coluna da imagem, colocando o valor de cada pixel da imagem em um vetor coluna Y (VASCONCELOS, 2007).

Neste trabalho, as imagens geradas pelo termovisor são de 120 X 120 pixels (matriz 120 X 120) e são consideradas 3 imagens (faces) para representar o objeto observado, portanto o vetor coluna Y totaliza 43.200 elementos. A ACP proporcionará a transformação do vetor coluna Y de forma que a sua dimensão seja menor e carregue o máximo de informação possível, para que os dados possam ser manipulados e armazenados de forma mais eficiente.

4 LÓGICA FUZZY

A lógica fuzzy permite soluções aproximadas. Lotfi Asker Zadeh foi o pioneiro na criação da lógica fuzzy que combinou os conceitos da lógica clássica e os conjuntos de Lukasiewicz (graus de pertinência sendo: 0, $\frac{1}{2}$ e 1, mais tarde expandiu um número infinito de valores entre 0 e 1) (SILVA et al., 2013).

Em 1965 foi publicado o trabalho de Conjuntos Fuzzy, por Lotf A. Zadeh, baseado na lógica multinível. Com este trabalho foi possível mostrar de forma matemática o tratamento dos aspectos imprecisos e ambíguos apresentados na lei da contradição. E a partir deste trabalho surge a expressão lógica fuzzy (MALUTTA, 2004). A lógica fuzzy, surgiu da necessidade do tratamento matemático dos aspectos vagos da informação sobre tudo dos problemas advindos da Engenharia Elétrica, pois esta área, estava repleta de ambiguidades, tais como a quantidade máxima de corrente elétrica deveria percorrer por um dado condutor, ou o grau de inclinação de mastro de navegação aéreo.

O termo em inglês “fuzzy” traduzido, tem o significado como algo vago, indefinido, incerto. Mas traduzido para o português os termos mais utilizados na área de inteligência artificial são nebuloso ou difuso. A lógica fuzzy trata de um raciocínio que busca classificar em números uma determinada realidade ou situação, que trabalha com muitas variáveis incertas e vagas, afim de facilitar o trabalho ou manipulação dos computadores. Comparando a lógica fuzzy com relação à lógica clássica, a lógica fuzzy apesar de ser imprecisa, contrário da lógica tradicional, ela reporta muito mais informações não estando restrita ao verdadeiro e falso. Isso permite que a lógica fuzzy descreva um determinado fato com muito mais detalhe e gradual, reduzindo assim a perda de informações, que consequentemente estará o mais coerente possível com a realidade em questão (MALUTTA, 2004).

4.1 TEORIA DOS CONJUNTOS FUZZY

O primeiro conceito está associado à definição de conjuntos Fuzzy. Um subconjunto A de um conjunto U é considerado subconjunto Fuzzy de U se for descrito como um conjunto de pares ordenados segundo o que se segue:

$$A = \{(x, \mu_A(x)); x \in U, \mu_A(x) \in [0, 1]\}$$

Onde:

$\mu_A(x)$ é uma função de pertinência que determina com que grau x está em A :

$\mu_A(x) = 1$	x pertence totalmente ao conjunto A ;
$0 < \mu_A(x) < 1$	x pertence parcialmente ao conjunto A ;
$\mu_A(x) = 0$	x não pertence ao conjunto A .

4.2 OPERAÇÕES COM CONJUNTOS FUZZY

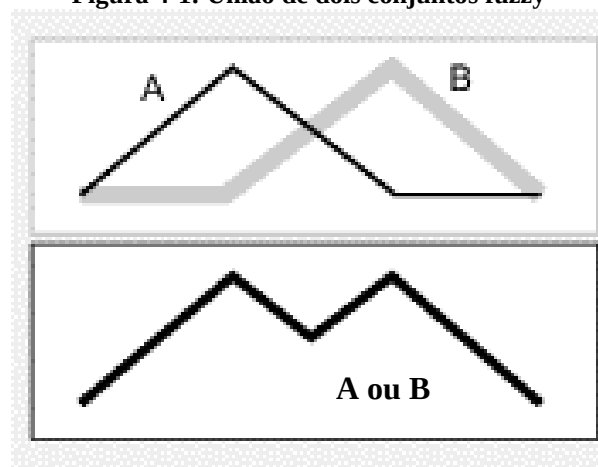
Na teoria clássica, quando um elemento x pertence a um conjunto A ou a um conjunto B , diz-se que $x \in A \cup B$. Quando um elemento x pertence simultaneamente aos conjuntos A e B , diz-se que $x \in A \cap B$. Também no caso de conjuntos nebulosos deseja-se fazer associações semânticas análogas e que resultariam em interpretações úteis para expressões do tipo $(\alpha \vee \beta)$, $(\alpha \wedge \beta)$, $(\neg \alpha)$. Portanto, necessita-se que sejam definidas, para conjuntos nebulosos, as operações com conjuntos (NASCIMENTO; YONEYAMA, 2000).

a) União ou disjunção dos dois conjuntos fuzzy

A união ou disjunção dos dois conjuntos fuzzy A e B resulta em um conjunto fuzzy C , o qual abrange os pontos máximos dos dois conjuntos unidos. Esse conjunto C pode ser descrito como $C = A \cup B$ ou $C = A \text{ ou } B$, cuja representação é dada por:

$$\mu_C(x) = \max(\mu_A(x), \mu_B(x)) = \mu_A(x) \vee \mu_B(x) \text{ (Figura 4-1)}$$

Figura 4-1: União de dois conjuntos fuzzy



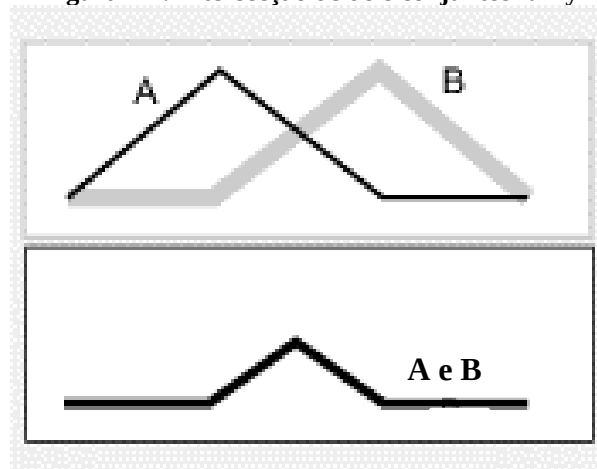
Fonte: (DEMUTH; BEALE; HAGAN, 2007)adaptado

b) Intersecção ou conjunção dos dois conjuntos fuzzy

A intersecção ou conjunção de dois conjuntos fuzzy A e B é utilizada para definir a região em comum entre esses dois conjuntos, resultando em um conjunto fuzzy D , este, é descrito como $D = A \cap B$ ou $D = A \text{ e } B$, cuja representação é dada por:

$$\mu_D(x) = \min(\mu_A(x), \mu_B(x)) = \mu_A(x) \wedge \mu_B(x) \text{ (Figura 4-2)}$$

Figura 4-2: Intersecção de dois conjuntos fuzzy



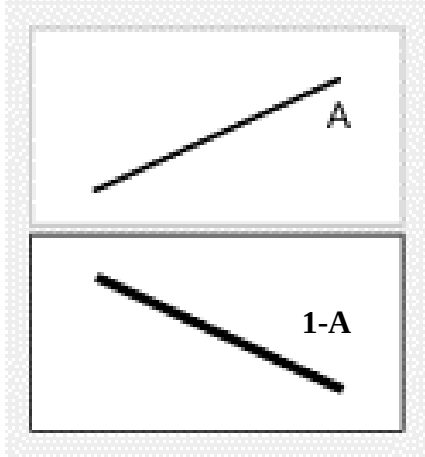
Fonte: (DEMUTH; BEALE; HAGAN, 2007)adaptado

c) Complemento de um conjunto fuzzy

A operação complemento é utilizada para definir a função de pertinência oposta de um subconjunto, ou seja, o complemento do subconjunto A , definido como $\neg A$, é formado pelos pontos opostos de A de dentro do intervalo $[0, 1]$. Essa operação, quando tratada nos extremos desse intervalo, é equivalente à operação “negação” da lógica booleana (SHHEIBIA, 2001). A representação dessa operação é descrita por:

$$\mu_{\neg A}(x) = 1 - \mu_A \text{ (Figura 4-3)}$$

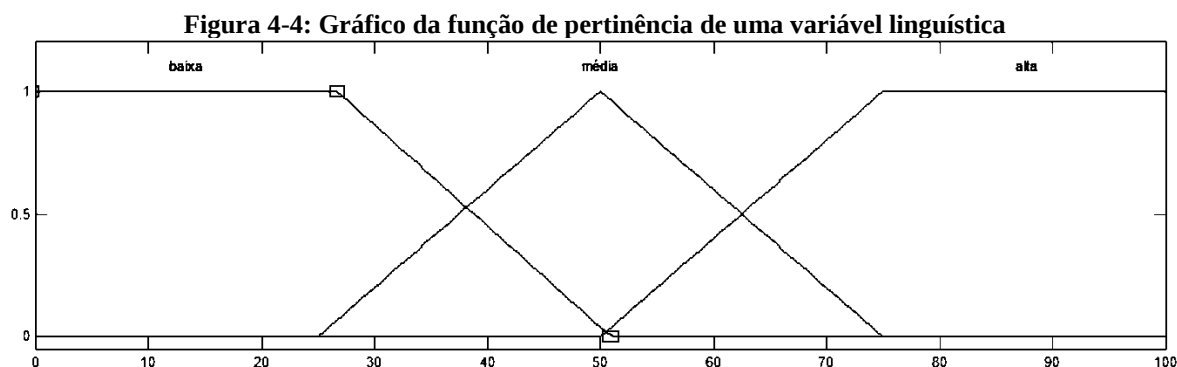
Figura 4-3: Complemento do conjunto fuzzy



Fonte: (DEMUTH; BEALE; HAGAN, 2007)adaptado

4.3 VARIÁVEIS LINGUÍSTICAS

Uma variável linguística é uma variável cujos valores são nomes de conjuntos Fuzzy. Por exemplo, a temperatura de um determinado processo pode ser uma variável linguística assumindo valores como “baixo”, “média” e “alta”. Estes valores são descritos por intermédio de conjuntos Fuzzy, representados por funções de pertinência, conforme Figura 4-4.



Fonte: próprio autor

Generalizando, os valores de uma variável linguística podem ser sentenças em uma linguagem específica construída a partir de termos primários (alto, baixo, pequeno, grande e gordo), de conectivos lógicos (negação *não*, conectivos *e* e *ou*), de modificadores (muito, pouco, levemente, extremamente) e de delimitadores (como parênteses). A principal função

de uma variável linguística é fornecer uma maneira sistemática para caracterização aproximada de fenômenos complexos ou mal definida. Em essência, a utilização do tipo de descrição linguística empregada pelos seres humanos, e não de variáveis quantificadas, permite o tratamento de sistemas analisados através de termos matemáticos convencionais (TANSCHKEIT, 2003).

4.4 FUNÇÕES DE PERTINÊNCIA

Cada conjunto Fuzzy é caracterizado pela função de pertinência, geralmente são representadas por $\mu_A(x)$. É através delas que será definido o quanto um determinado elemento pertence ao conjunto. De acordo com sua aplicação ou a maneira de representar em determinado contexto, existem diferentes tipos de funções de pertinência.

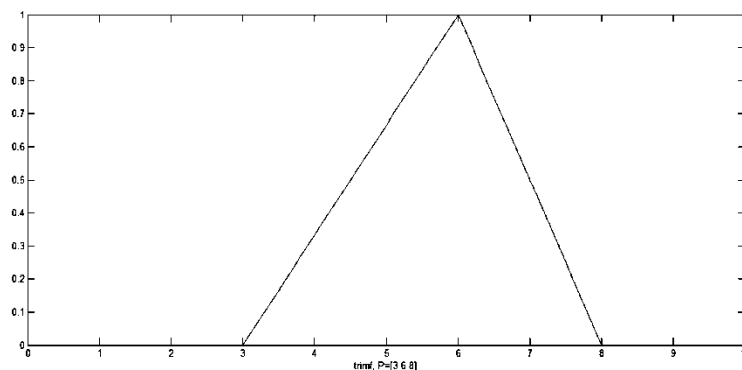
Funções de pertinência podem ser definidas a partir da experiência e da perspectiva do usuário, mas é comum fazer-se uso de funções de pertinência padrão, como, por exemplo, as de forma triangular, trapezoidal e gaussiana. Em aplicações práticas as formas escolhidas inicialmente podem sofrer ajustes em função dos resultados observados. A seguir mostraremos as funções de pertinência mais utilizadas:

4.4.1 Triangular:

É representada pela função μ_{tri} :

$$\mu_{tri}(x; a, b, c) = \begin{cases} 0, & \text{se } x < a \\ \frac{x-a}{b-a}, & \text{se } a \leq x \leq b \\ \frac{c-x}{c-b}, & \text{se } b \leq x \leq c \\ 0, & \text{se } x > c \end{cases}, \text{ com } a, b, c \in \mathbb{R} \text{ e } a \leq b \leq c \text{ (Figura 4-5).}$$

Figura 4-5: Exemplo de função de pertinência triangular



Fonte: próprio autor

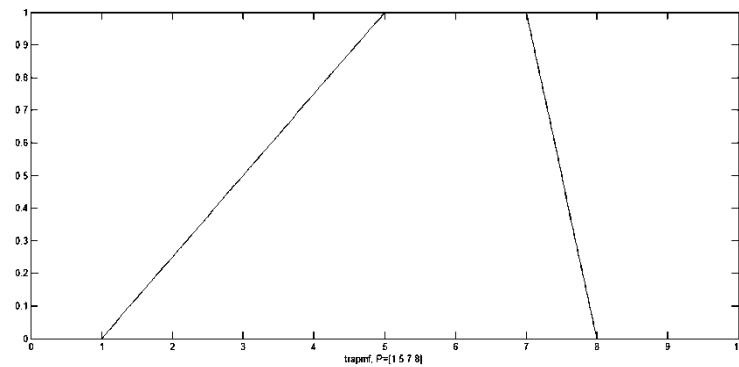
4.4.2 Trapezoidal

É representada pela função μ_{trap} :

$$\mu_{trap}(x; a, b, c, d) = \begin{cases} 0, & \text{se } x < a \\ \frac{x-a}{m-a}, & \text{se } a \leq x \leq m \\ 1, & \text{se } m \leq x \leq n, \text{ com } a, b, c, d \in R \text{ e } a \leq b \leq c \leq d. \\ \frac{b-x}{b-n}, & \text{se } n \leq x \leq b \\ 0, & x \geq b \end{cases}$$

(Figura 4-6)

Figura 4-6: Exemplo de função de pertinência trapezoidal



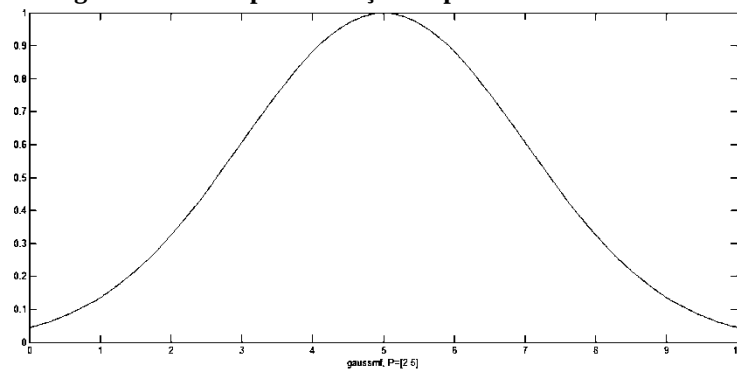
Fonte: próprio autor

4.4.3 Gaussiana

É representada pela função μ_{gauss} :

$$\mu_{gauss}(x; a, b, c) = ae^{-\frac{(x-b)^2}{2c^2}} \quad (\text{Figura 4-7})$$

Figura 4-7: Exemplo de função de pertinência Gaussiana



Fonte: próprio autor

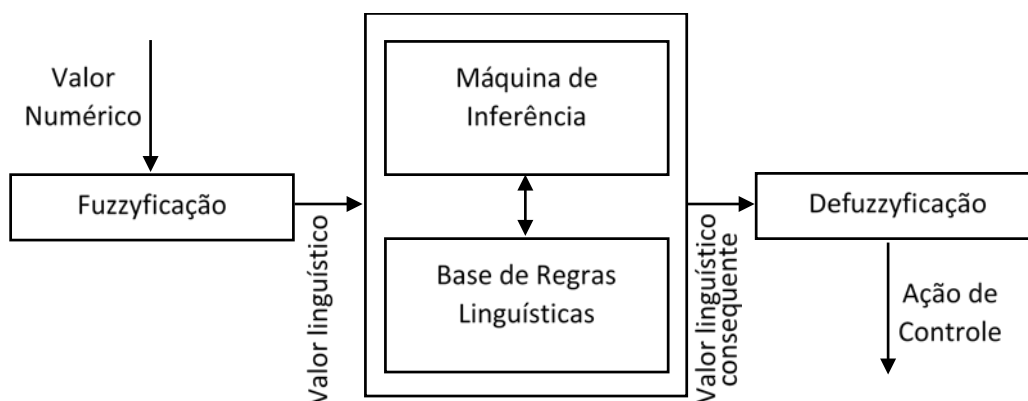
4.5 CONTROLE FUZZY

Sistemas de controle clássico apresentam muitas limitações para o controle de processos complexos, não lineares e multivariáveis. Essas limitações existem porque o sistema de controle clássico se baseia no modelo matemático, no qual, o sistema é descrito através de uma ou mais equações diferenciais, e devido à complexidade do processo esse modelo matemático se torna difícil de ser obtido e implementado. (BARROS e ALMEIDA, 2009)

A teoria de controle *fuzzy* é capaz de identificar e controlar uma planta não linear sem que aja um modelo matemático exato, através do uso de aprendizado e adaptação, ou mesmo pela extração do conhecimento de um especialista no controle do processo.

A Figura 4-8 é uma representação em diagramas de blocos da estrutura básica de um controlador *fuzzy* e um modelo geral para a identificação dos módulos que o compõem, fornecendo uma ideia de como funciona o fluxo de informação.

Figura 4-8: Modelo geral em diagrama de blocos de um controlador fuzzy



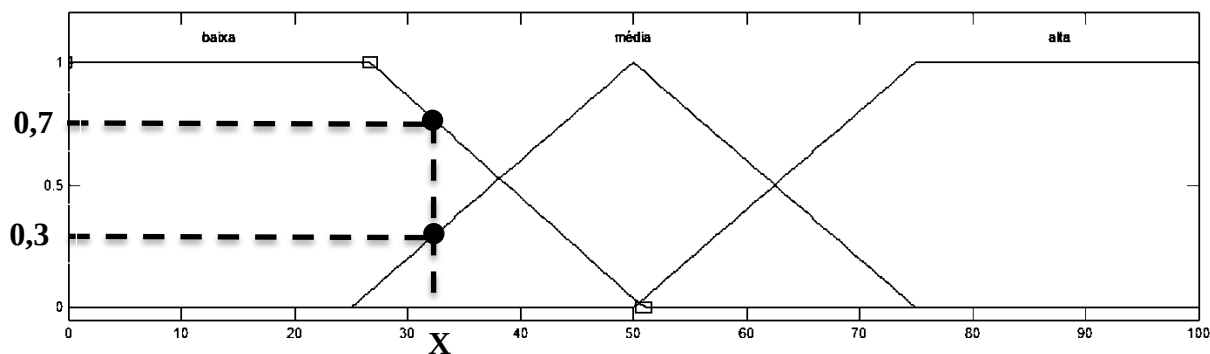
Fonte: próprio autor

4.6 FUZZYFICAÇÃO

Antes que se suceda qualquer operação com variáveis linguística é indispensável à tradução dos números reais do processo em membros de variáveis linguística com seus respectivos valores de pertinência. Um só valor real do processo pode ser membro de mais de um valor linguístico, sendo que, ele terá graus diferentes de relevância .

Na Figura 4-9, o valor X é ao mesmo tempo considerado como baixa e média, porém em graus diferentes de importância; simultaneamente 30% média e 70% baixa.

Figura 4-9: Exemplo de um valor real com correspondência em dois valores linguísticos



Fonte: próprio autor

A fuzzyficação é então, responsável por associar e transformar os valores reais do processo em valores sujeitos a manipulação por lógica *fuzzy* para dar suporte aos passos seguintes.

4.7 BASE DE REGRAS

A principal função da base de regras é representar de uma maneira estruturada a ação de controle de um operador do processo experiente (DRIANKOV; HELLENDORRN, HANS REINFRANK, 1996) e/ou de um engenheiro de controle na forma de um conjunto de regras de produção como:

se (estado do processo) *então* (saída de controle)

A parcela *se(if)* de uma regra, contém uma ou mais condições relacionadas ao estado do processo chamadas antecedentes, formando uma premissa; a parcela *então (then)* contém uma ou mais ações de controle chamadas consequente.

Para exemplificar a utilização da base de regras, toma-se o controle de uma caldeira, onde se estabelece uma relação entre a temperatura interna da unidade com a vazão necessária de água para o seu resfriamento, o que produz as seguintes regras:

se temperatura é alta *então* vazão de água é alta

se temperatura é baixa *então* vazão de água é baixa

se temperatura é estável *então* vazão de água é normal

Aparentemente simplista, o exemplo dado é capaz de manter estável a temperatura em uma caldeira, desde que os conjuntos difusos nele envolvidos tenham uma definição condizente com a realidade. O que vale observar é que para cada estado do processo é relacionada uma ação de controle.

4.8 MÁQUINA DE INFERÊNCIA

Os dados logo após serem fuzzyficados, são enviados à máquina de inferência, que avalia a ação de controle a ser tomada, a partir, do resultado das operações de conjunto *fuzzy* feito com as regras linguísticas acionadas. Existem muitos procedimentos inferenciais na lógica *Fuzzy*, porém os mais utilizados são o Mandani e o Takagi-Sugeno-Kang (MENDEL, 2001).

4.9 DEFUZZYFICAÇÃO

Os valores adquiridos na máquina de inferência são humanamente inteligíveis, mas da forma em que estão não são de grande utilidade ao processo. Então é utilizada a interface de defuzzyficação, onde os valores linguísticos são mapeados em valores numéricos de saída. Obtendo-se um valor discreto que melhor represente o valor linguístico consequente para que possa ser usado numa ação de controle no mundo real.

Os principais métodos de defuzzyficação são: Centro de gravidade, centro do máximo, média do máximo e a média ponderada (SHAW; SIMÕES, 1999).

Nesta etapa aplicam-se os valores obtidos na etapa de inferência em um dos métodos para se encontrar o valor de saída.

Centro de gravidade:

$$X_G = \frac{\int_{-\infty}^{+\infty} x \cdot \mu(x) \cdot dx}{\int_{-\infty}^{+\infty} \mu(x) \cdot dx}$$

Média dos máximos:

$$X_{MM} = \frac{\sum_{i=1}^M x_{m\acute{a}x}(i)}{M}$$

Centro de máximos:

$$X_{CM} = \frac{\sum_{i=1}^M X_{m\acute{a}x}(i) \cdot \mu[X_{m\acute{a}x}(i)]}{\sum_{i=1}^M \mu[X_{m\acute{a}x}(i)]}$$

4.10 ANFIS - ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

Durante a década de 1990, várias redes neuro-fuzzy apareceram na literatura. As mais importantes foram: Sistema de Inferência Neuro-Fuzzy Adaptativo (ANFIS) desenvolvido por (JANG, 1993); Rede de Controle com Aprendizagem Fuzzy Adaptativa (FALCON) desenvolvida por (LIN; LEE, 1991); Controle Neuro-Fuzzy (NEFCON) proposto por (NAUCK; KLAWONN; KRUSE, 1997); Controle Inteligente com Abordagem de Aproximação Generalizada (GARIC) desenvolvido por. (BERENJI; KHEDKAR, 1992). Estes desenvolvimentos deram origem a várias outras propostas. (SACCHI, 2009)

A ideia deste tipo de abordagem é definir os formatos das funções de pertinência dos Sistemas de Inferência Fuzzy (FIS) de forma eficiente, através da propriedade adaptativa das Redes Neurais. Os primeiros a utilizarem esta ideia foram (TAKAGI; SUGENO, 1985) (SUGENO; KANG, 1988). Posteriormente, (JANG, 1993) desenvolveu uma abordagem sistemática para a adaptação e ajuste dos parâmetros dos FIS, a qual se mostrou bastante eficiente na solução dos problemas por ele proposto.

Neste trabalho é utilizado o Sistema de Inferência Neuro-Fuzzy Adaptativo (ANFIS) desenvolvido por (JANG, 1993). Este modelo permite que os Sistemas de Inferência Fuzzy (FIS) ajuste seus parâmetros de forma adaptativa, ou seja, automaticamente durante o treinamento.

Utilizando um conjunto de dados entrada/saída, o ANFIS constrói um sistema de inferência fuzzy (FIS) cujos parâmetros das funções membros são ajustados usando o algoritmo backpropagation sozinho ou com o método dos mínimos quadrados. Isso permite os sistemas fuzzy aprenderem a partir dos dados que eles estão modelando.

Uma estrutura tipo rede, similar àquela das redes neurais, é usada para interpretar o conjunto entrada/saída. Os parâmetros associados com as funções membros se alteram durante o processo de aprendizagem. O cálculo desses parâmetros é simplificado através do vetor gradiente, que provê uma medida de eficiência que indica o desempenho do FIS na modelagem dos dados entrada/saída para um determinado conjunto de parâmetros. Qualquer rotina de otimização é aplicada para ajustar os parâmetros de forma a reduzir os erros medidos.

A abordagem da modelagem usada pelo ANFIS é similar a muitas técnicas de sistemas de identificação. Primeiro supõe-se uma estrutura de modelo parametrizado. Em seguida, coleta-se os dados entrada/saída em uma forma a ser usada pelo ANFIS para treinamento. Pode-se então usar ANFIS para treinar o modelo FIS que emula os dados de

treinamento, modificando os parâmetros das funções membros de acordo com o critério de erro escolhido.

O Modelo Validação é o processo pelo qual os vetores dos conjuntos de dados entrada/saída com o qual o FIS não foi treinado são apresentados ao modelo FIS treinado, para analisar o desempenho do modelo FIS na previsão de um conjunto de dados semelhantes aos valores de saída. O conjunto de validação é usado para controlar o potencial do modelo em super-ajustar os dados. Quando os conjuntos de validação e teste são apresentados ao ANFIS, o modelo FIS é selecionado para ter parâmetros associados ao menor erro no conjunto de validação.

A técnica de redes neurais será apresentada no capítulo seguinte.

5 REDES NEURAIS ARTIFICIAIS

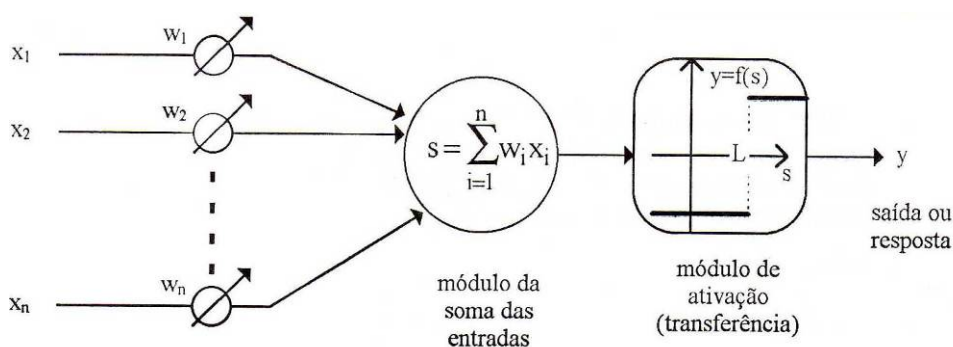
5.1 INTRODUÇÃO

A rede neural artificial é uma técnica de Inteligência Artificial que tenta simular em máquinas (computadores) o funcionamento do cérebro humano, de uma maneira simplificada. Ela é capaz de reconhecer padrões, extrair regularidades e detectar relações subjacentes em um conjunto de dados aparentemente desconexos. Além disso, ela apresenta habilidade de lidar com dados ruidosos, incompletos ou imprecisos, e de prever sistemas não lineares. Ela pode ser definida como a área das ciências da computação que visa o projeto de sistemas inteligentes, ou seja, sistemas que tentam emular algum tipo de inteligência, semelhante à de um ser humano, em termos de processos computacionais. Os sistemas inteligentes são assim denominados por exibirem características que associamos ao comportamento inteligente de um ser humano, como por exemplo: percepção, aprendizagem, raciocínio, comunicação e atuação em ambientes complexos (NILSSON, 1998).

Uma rede neural, segundo (HAYKIN, 2001), pode ser definida como, um processador maciçamente paralelamente distribuído, constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela assemelha-se ao cérebro em dois aspectos: (1) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem; (2) forças de conexão entre neurônios (os pesos sinápticos) são utilizadas para armazenar o conhecimento adquirido.

As redes neurais são formadas por neurônios e conexões entre eles. O neurônio (Figura 5-1) representa uma região onde informações são processadas. Seus três elementos básicos são: os pesos sinápticos, a função de soma e a função de transferência.

Figura 5-1: Modelo do neurônio artificial



Fonte: (LOESCH; SARI, 1996)

As conexões entre os neurônios, denominadas pesos sinápticos, são responsáveis pelo armazenamento das informações. Além disso, elas definem o efeito que a saída de um neurônio exerce sobre a entrada do neurônio seguinte. Os pesos sinápticos são de grande importância para uma rede neural, pois determinam toda a manipulação de valores da rede.

A função de soma processa os estímulos ponderados pelos respectivos pesos, ou seja:

$$x_j = \sum_i w_{ij} y_i \quad (5.1)$$

onde y_i é a saída gerada por cada neurônio da camada anterior.

Já a função de transferência, também chamada de função de ativação, limita a amplitude do intervalo do sinal de saída do neurônio para algum valor finito, geralmente no intervalo normalizado $[0, 1]$ ou $[-1, 1]$.

$$y_j = f(x_j) \quad (5.2)$$

Aqui são identificados três tipos básicos de funções de ativação (HAYKIN, 2001):

a) Função de Limiar. Para este tipo de função de ativação, descrito na Figura 5-2 (a), temos:

$$f(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases} \quad (5.3)$$

b) Função Linear por Partes. Para a função linear por partes descrita na Figura 5-2 (b) temos:

$$f(x) = \begin{cases} 1, & x \geq \frac{1}{2} \\ x, & -\frac{1}{2} < x < \frac{1}{2} \\ 0, & x \leq -\frac{1}{2} \end{cases} \quad (5.4)$$

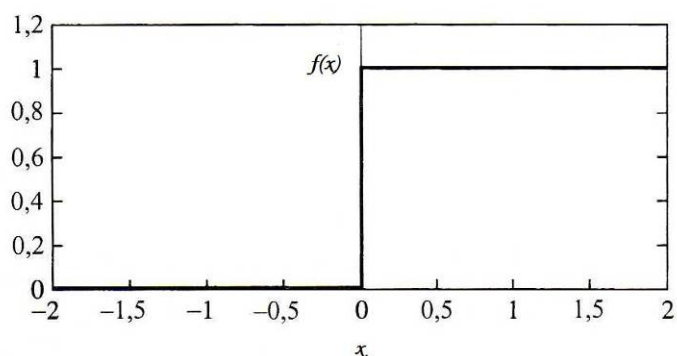
onde $(-\frac{1}{2}$ e $\frac{1}{2})$ é o intervalo que define a saída linear e 0 e 1 são os limites mínimo e máximo da função.

c) Função Sigmóide. A função sigmóide, cujo gráfico tem a forma de s, é de longe a forma mais comum de ativação utilizada na construção de redes neurais artificiais. Ela é definida como uma função estritamente crescente que exhibe um balanceamento adequado entre comportamento linear e não-linear. Um exemplo de função sigmóide é a função logística, definida por:

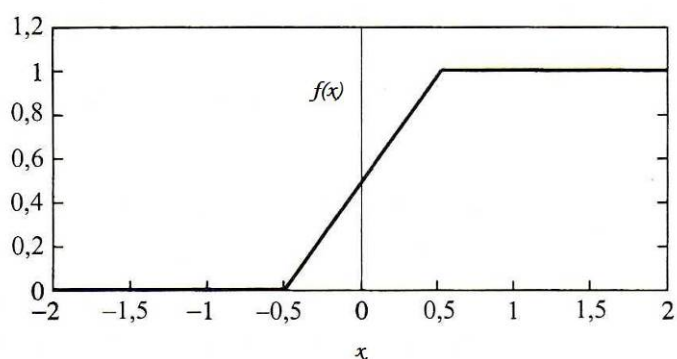
$$f(x) = \frac{1}{1 + \exp(-\alpha x)} \quad (5.5)$$

Onde α é o parâmetro de inclinação da função sigmóide. Variando-se o parâmetro α , obtemos funções sigmóides com diferentes inclinações, como ilustrado na Figura 5-2 (c).

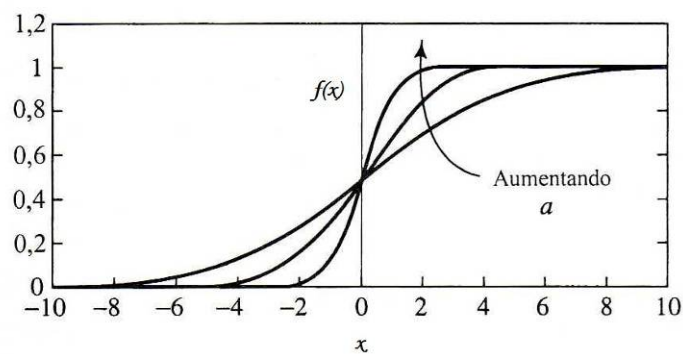
Figura 5-2: (a) Função de limiar. (b) Função linear por partes. (c) Função sigmóide para parâmetro de inclinação α variável.



(a)



(b)



(c)

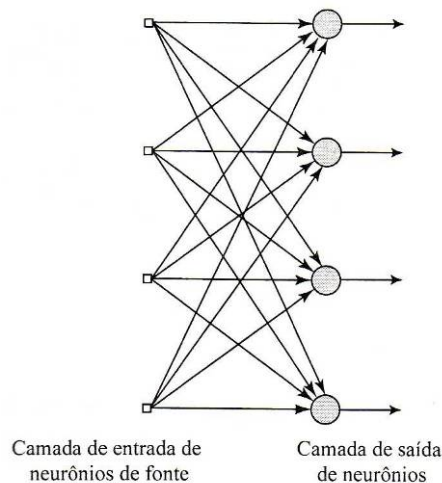
Fonte: (HAYKIN, 2001) adaptado

Além dos três elementos básicos já citados, o neurônio pode ainda apresentar um bias que tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação (HAYKIN, 2001). O termo bias age como um peso extra nas conexões das unidades cuja entrada é sempre um (FAUSETT, 1994).

Em geral, podemos identificar três classes de arquiteturas de rede fundamentalmente diferentes (HAYKIN, 2001):

a) **Redes Alimentadas Adiante com Camada Única:** Em uma rede neural em camadas, os neurônios estão organizados na forma de camadas. Na forma mais simples de uma rede em camadas, temos uma camada de entrada de nós de fonte que se projeta sobre uma camada de saída de neurônios, mas não vice-versa. Ela é ilustrada na Figura 5-3.

Figura 5-3: Rede alimentada adiante ou acíclica com uma única camada de neurônios

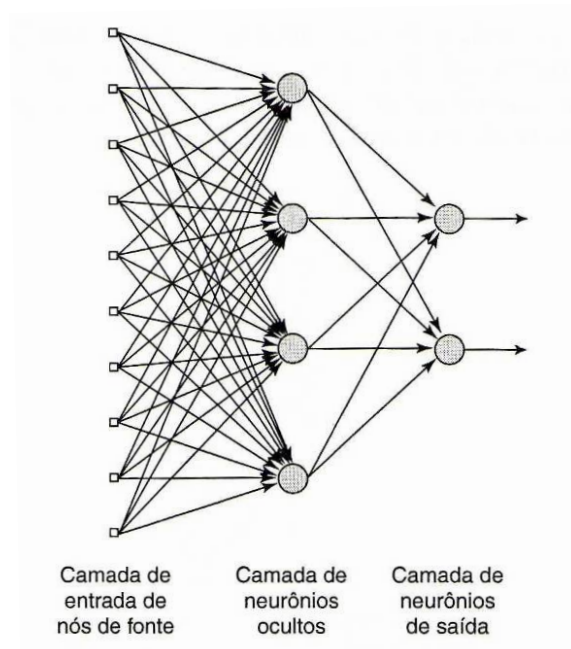


Fonte: (HAYKIN, 2001)

b) **Redes Alimentadas Diretamente com Múltiplas Camadas:** A segunda classe de uma rede neural alimentada adiante se distingue pela presença de uma ou mais camadas ocultas, cujos nós computacionais são chamados correspondentes de neurônios ocultos ou unidades ocultas. A função dos neurônios ocultos é intervir entre a entrada externa e a saída da rede de maneira útil. Adicionando-se uma ou mais camadas ocultas, tornamos a rede capaz de extrair estatísticas de ordem elevada. A Figura 5-4 ilustra uma rede neural de múltiplas camadas alimentada adiante para o caso de uma única camada oculta.

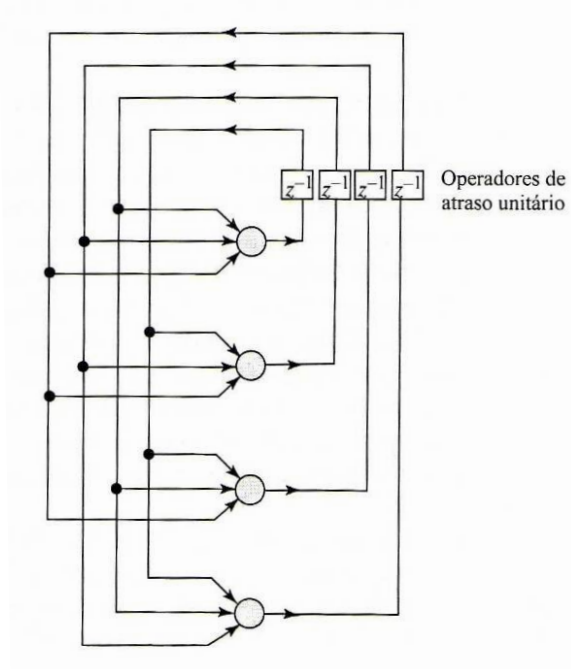
c) **Redes recorrentes:** Uma rede neural recorrente se distingue de uma rede neural alimentada adiante por ter pelo menos um laço de realimentação. Uma rede recorrente pode consistir, por exemplo, de uma única camada de neurônios com cada neurônio alimentando seu sinal de saída de volta para as entradas de todos os outros neurônios como ilustrado nas Figuras 5-5 e 5-6. Os laços de realimentação envolvem o uso de ramos particulares compostos de elementos de atraso unitário (representados por z^{-1}), o que resulta em um comportamento dinâmico não-linear, admitindo-se que a rede neural contenha unidades não-lineares.

Figura 5-4: Rede alimentada adiante ou acíclica totalmente conectada com uma camada oculta e uma camada de saída



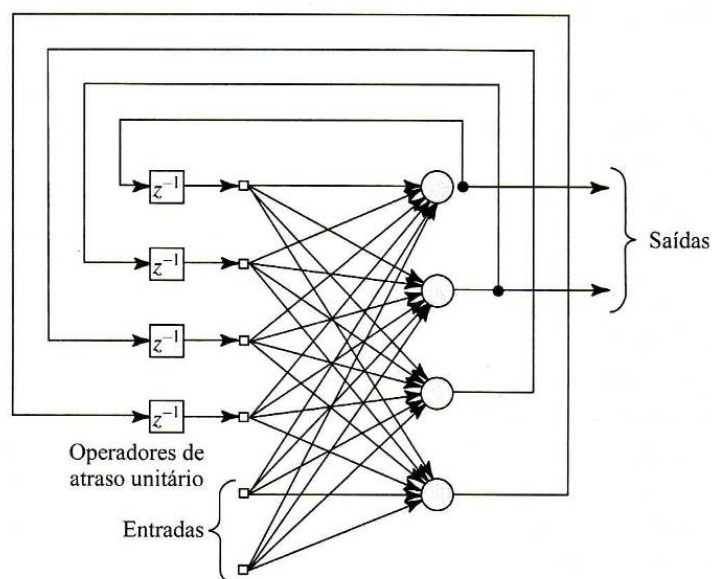
Fonte: (HAYKIN, 2001)

Figura 5-5: Rede recorrente sem laços de auto-realimentação e sem neurônios ocultos



Fonte: (HAYKIN, 2001)

Figura 5-6: Rede recorrente com neurônios ocultos



Fonte: (HAYKIN, 2001)

A rede neural deve ter a capacidade de generalização, ou seja, ela deve ser capaz não apenas de classificar as entradas para as quais ela recebe treinamento, mas também de generalizar e classificar entradas que não tenham sido apresentadas. Isto é possível graças a um processo de aprendizagem ao qual a rede é submetida. Esta propriedade permite que a rede encontre respostas corretas mesmo quando os dados disponíveis para as entradas estão incompletos ou danificados.

O desenvolvimento de uma rede neural ainda é um processo de tentativa e erro. A seleção da rede envolve a escolha da topologia da rede (ou arquitetura), da função de transferência e do algoritmo de aprendizagem.

5.2 APRENDIZAGEM

Segundo (LOESCH; SARI, 1996), a Psicologia possui uma dicotomia: de um lado as aproximações cognitivas e de outro a tecnologia behaviorista (estudo do comportamento). O campo de redes neurais recebe contribuições de ambas as visões. O sistema computacional aprende a generalizar e com o mesmo progresso dos animais através do treinamento behaviorista.

A propriedade mais importante das redes neurais é a habilidade de aprender de seu ambiente e com isso melhorar seu desempenho. Isto é feito através de um processo iterativo

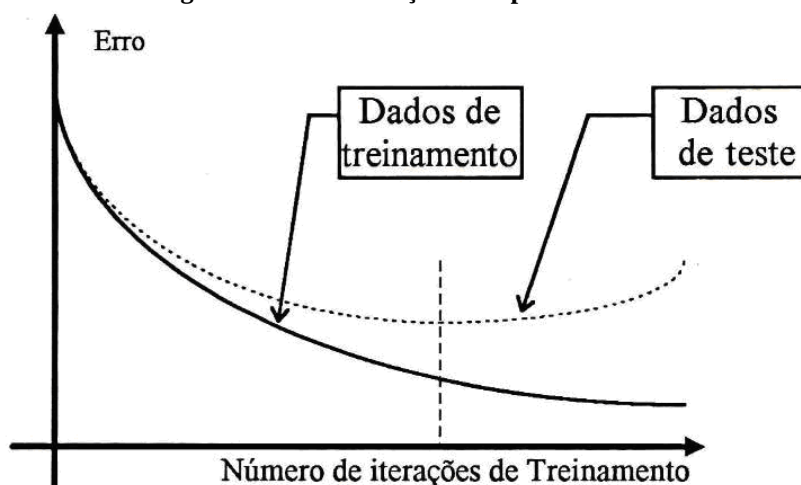
de ajustes aplicado a seus pesos, o treinamento. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

O processo de aprendizagem nas redes neurais acontece internamente por meio do ajuste dos pesos sinápticos das conexões durante a exposição dos exemplos, em resposta à quantidade de erros gerados pela rede.

As redes neurais são treinadas para aprender a partir dos dados de entrada. Assim como o cérebro humano, elas aprendem a partir de experiências e não através de programação. Por este motivo, deve-se tomar bastante cuidado com a formação do conjunto de treinamento. Este conjunto deve ser gerado a partir de dados históricos, ou seja, a partir de experiências e fatos ocorridos no passado.

Existe uma concepção errônea perigosa relativa ao treinamento iterativo. Ela se apóia no supertreinamento da rede, baseado na idéia de que se deve tentar treinar a rede para reduzir mais e mais o erro do conjunto de treinamento para aumentar seu desempenho. Embora parcialmente verdadeira, ela pode conduzir a um superajustamento ao conjunto de treinamento, perdendo, além de certo ponto, desempenho com relação ao conjunto de testes. A Figura 5-7 plota o erro de uma rede neural para o conjunto de treinamento e para o conjunto de teste. Após passar de N iterações o erro do conjunto de testes cresce. Deve-se parar o treinamento quando o ponto N foi atingido, para obter o máximo desempenho operacional da rede (LOESCH; SARI, 1996).

Figura 5-7: Manifestação do supertreinamento



Fonte: (LOESCH; SARI, 1996)

Dois paradigmas de aprendizagem são apresentados a seguir:

a) Aprendizagem supervisionada (ou aprendizagem com professor), quando é utilizado um agente externo que indica à rede a resposta desejada para o padrão de entrada. O ajuste dos pesos ocorre quando o sistema compara a saída da rede com a resposta desejada previamente conhecida.

b) Aprendizagem não-supervisionada (ou aprendizagem sem professor), quando não existe um agente externo indicando a resposta desejada para os padrões de entrada. A rede neural utiliza os neurônios como classificadores, e os dados de entrada como elementos de classificação. Esse tipo de rede trabalha essas entradas e se organiza de modo a classificá-las mediante algum critério de semelhança.

Em um processo de aprendizagem, os pesos dos neurônios são ajustados através de um algoritmo de aprendizagem. O algoritmo de aprendizagem é um conjunto pré-estabelecido de regras bem-definidas para resolução de um problema de aprendizagem (HAYKIN, 2001). Ele tem como objetivo encontrar pesos para a rede que permitam que esta gere saídas compatíveis com as desejadas. Os algoritmos existentes diferem entre si pela forma como ocorre o ajuste dos pesos sinápticos dos neurônios, ou seja, pela regra de aprendizagem adotada.

De acordo com (HAYKIN, 2001), existem cinco regras básicas de aprendizagem através das quais os pesos sinápticos de uma rede podem ser ajustados: aprendizagem por correção de erro, baseada em memória, hebbiana, competitiva e aprendizagem de Boltzmann.

5.2.1 Aprendizagem Por Correção De Erro

A aprendizagem por correção de erro é baseada no paradigma de aprendizagem supervisionada no qual a saída desejada para cada padrão de entrada é fornecida para a rede. O sinal de saída gerado pela rede, representado por $y_k(n)$, é comparado com a resposta desejada, representada por $d_k(n)$, produzindo um sinal de erro $e_k(n)$. Este erro é dado por:

$$e_k(n) = d_k(n) - y_k(n) \quad (5.6)$$

O sinal de erro é então utilizado para ajustar os pesos das conexões com o objetivo de aproximar o sinal de saída $y_k(n)$ da resposta desejada $d_k(n)$, reduzindo o erro $e_k(n)$.

5.2.2 Aprendizagem baseada em memória

Nesta regra de aprendizagem, todas as experiências passadas são armazenadas em uma grande memória de exemplos de entrada-saída classificados corretamente: (x_i, d_i) , onde $i = 1 \dots n$, x_i representa um padrão de entrada e d_i representa a resposta desejada correspondente (HAYKIN, 2001). Quando se deseja classificar um vetor de teste X (não visto antes), o algoritmo responde buscando e analisando os dados de treinamento em uma "vizinhança local" de X .

Esta regra envolve dois ingredientes essenciais, e a forma como eles são definidos é que vai diferenciar os algoritmos de aprendizagem baseada em memória entre si. Os ingredientes são: o critério utilizado para definir a vizinhança local do vetor de teste X e a regra de aprendizagem aplicada aos exemplos de treinamento em uma "vizinhança local" de X (HAYKIN, 2001).

5.2.3 Aprendizagem Hebbiana

O postulado de aprendizado de Hebb é a mais antiga regra de aprendizagem existente [HEBB, 1949 *apud* HAYKIN, 2001]. O princípio básico desta regra é: se dois neurônios em ambos os lados de uma conexão são ativados simultaneamente, então a força desta conexão é seletivamente aumentada. A forma mais simples de aprendizagem hebbiana é descrita por (HAYKIN, 2001):

$$w_{jk}(n+1) = w_{jk}(n) + \eta y_k(n) x_j(n) \quad (5.7)$$

onde w_{jk} é o peso sináptico do neurônio k , x_j e y_k são os sinais pré-sináptico e pós-sináptico, respectivamente, deste peso, e η é uma constante positiva que define a taxa de aprendizagem.

5.2.4 Aprendizagem competitiva

Na regra de aprendizagem competitiva, os neurônios da camada de saída competem entre si para se tornarem ativos, considerando que somente um neurônio pode estar ativo em um determinado instante. Este fenômeno é conhecido como '*winner-take-all*', isto é, 'o vencedor leva tudo'.

Para um neurônio ser o vencedor, isto é, estar ativo, seu campo local induzido v_k para um padrão de entrada x deve ser o maior dentre todos os neurônios da rede (HAYKIN, 2001). Quando isto acontece, o sinal de saída y_k deste neurônio é igual a um. Caso contrário o sinal de saída é colocado em zero, como demonstrado a seguir:

$$y_k = \begin{cases} 1 & \text{se } v_k > v_j \text{ para todos } j, j \neq k \\ 0 & \text{caso contrário} \end{cases} \quad (5.8)$$

Nesta regra, a variação Δw_{jk} que é aplicada ao peso w_{jk} é definida por:

$$\Delta w_{jk} = \begin{cases} \eta(x_j - w_{jk}) & \text{se o neurônio } k \text{ vencer a competição} \\ 0 & \text{se o neurônio } k \text{ perder a competição} \end{cases} \quad (5.9)$$

onde η é a taxa de aprendizagem. Como efeito desta regra de aprendizagem, o vetor de peso w_k é movido na direção do padrão de entrada x , a partir do neurônio vencedor k .

Este tipo de aprendizagem é adequado para descobrir características nos dados de entrada que podem ser utilizadas para agrupar padrões similares.

5.2.5 Aprendizagem de Boltzmann

A regra de aprendizagem de Boltzmann é um algoritmo de aprendizagem estocástico que realiza o ajuste dos pesos baseando-se na probabilidade e na mecânica estatística. A rede neural que utiliza esta regra é denominada máquina de Boltzmann (HAYKIN, 2001).

Os neurônios nesta máquina formam uma estrutura recorrente e podem assumir dois estados: ligado (+1) ou desligado (-1). Os estados de cada neurônio na máquina determinam o valor de uma função de energia que caracteriza esta rede:

$$E = -\frac{1}{2} \sum_j \sum_k w_{jk} x_j x_k \quad j \neq k \quad (5.10)$$

onde x_j é o estado do neurônio j e w_{jk} é o peso sináptico entre os neurônios j e k . A restrição $j \neq k$ é para garantir que nenhum dos neurônios tenha auto-realimentação. O equilíbrio é alcançado quando esta função de energia alcança um mínimo.

A máquina escolhe um neurônio ao acaso e, em um determinado passo do processo de aprendizagem, troca seu estado de $+x_k$ para $-x_k$, a uma pseudotemperatura T , com probabilidade:

$$P(x_k \rightarrow -x_k) = \frac{1}{1 + \exp(-\Delta E_k/T)} \quad (5.11)$$

onde ΔE_k é a variação de energia resultante da troca.

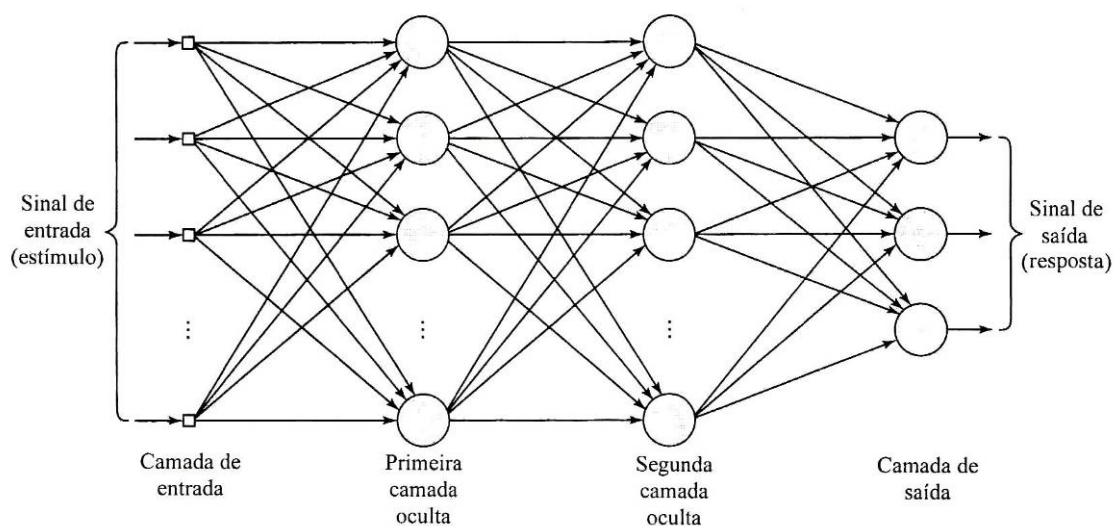
Os neurônios desta rede podem ser visíveis ou ocultos e o sistema é operado em dois modos: condição presa, no qual os neurônios visíveis estão presos em estados específicos determinados pelo meio; e condição de operação livre, onde os neurônios visíveis e ocultos podem operar livremente (HAYKIN, 2001).

5.3 REDE NEURAL *BACKPROPAGATION*

5.3.1 Introdução

A rede neural comumente denominada *backpropagation* é na verdade uma Rede Alimentada Diretamente com Múltipla Camada (Figura 5-8), treinada pelo algoritmo *backpropagation* (algoritmo de retropropagação do erro). O desenvolvimento do algoritmo de retropropagação representa um marco nas redes neurais (HAYKIN, 2001). Este trabalho foi o primeiro que possibilitou o ajuste dos pesos em redes multicamadas *feedforward*, abrindo caminho para a elaboração de redes neurais mais genéricas. Devido à grande popularidade do método *backpropagation*, seu nome é utilizado para denominar as redes que o utilizam no seu treinamento.

Figura 5-8: Grafo arquitetural de um perceptron de múltiplas camadas com duas camadas ocultas



Fonte:(HAYKIN, 2001)

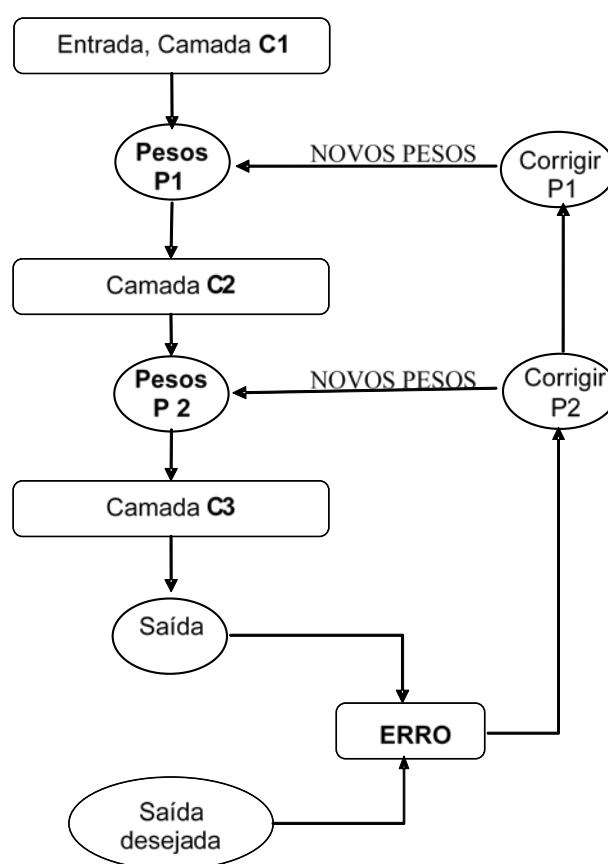
A atratividade do método *backpropagation* vem do conjunto de equações bem definidas e explícitas para correção dos pesos da rede. Este algoritmo consiste em realizar a retropropagação do erro gerado na comparação entre a saída da rede e a saída desejada com o objetivo de minimizar o erro total da saída gerada pela rede. O treinamento de uma rede através deste algoritmo envolve três etapas: a propagação dos dados da camada de entrada para a camada de saída da rede, o cálculo e a retropropagação do erro gerado pela rede, e o ajuste dos pesos (FAUSETT, 1994).

Na primeira etapa, estímulos de entrada são apresentados à rede e as ativações fluem até chegarem à camada de saída, gerando um resultado. Já na segunda e terceira etapas, o

resultado obtido pela rede é comparado com a saída desejada e o erro gerado é computado para as unidades de saída. Os pesos conectados às unidades de saída são então ajustados para reduzir este erro. Em seguida, o erro da camada de saída é utilizado para derivar estimativas de erro para as unidades da(s) camada(s) oculta(s), para que o erro seja então propagado para trás até a conexão da camada e entrada. O fluxo de informação deste processo é ilustrado de uma forma resumida na Figura 5-9.

O método *backpropagation* atualiza os pesos incrementalmente, depois de analisar cada par entrada-saída. Depois da apresentação de todos os pares entrada-saída diz-se que uma época foi concluída. Este treinamento, em geral, requer muitas épocas.

Figura 5-9: Correção dos pesos por *backpropagation*



Fonte: próprio autor

5.3.2 Ajuste dos Pesos

Uma regra de aprendizagem por correção de erro, a regra delta, é onde está baseado o ajuste dos pesos realizado pelo algoritmo *backpropagation*. Este ajuste é feito baseado na retropropagação do erro através da qual o erro gerado pelos neurônios na camada de saída é distribuído para os demais neurônios da rede.

Mesmo conhecendo o erro global da rede, não é possível determinar os pesos exatos para poder corrigi-lo. Entretanto, com base nesta informação, pode-se estabelecer a direção na qual os pesos devem ser ajustados para minimizar o erro quadrado total da saída da rede. Conhecida esta direção, é possível ajustar os pesos até que o menor erro global seja atingido. O ajuste de um peso w_{ij} que define seu valor para a próxima iteração é definido por:

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n) \quad (5.12)$$

A variação Δw_{ij} que é aplicada ao peso w_{ij} deve ser proporcional ao sinal de entrada x_j , que é definido pelos sinais de saída da camada anterior ponderados pelos pesos, e ao erro gerado na saída. Ela é dada por:

$$\Delta w_{ij}(n) = \eta \delta_j(n) x_j(n) \quad (5.13)$$

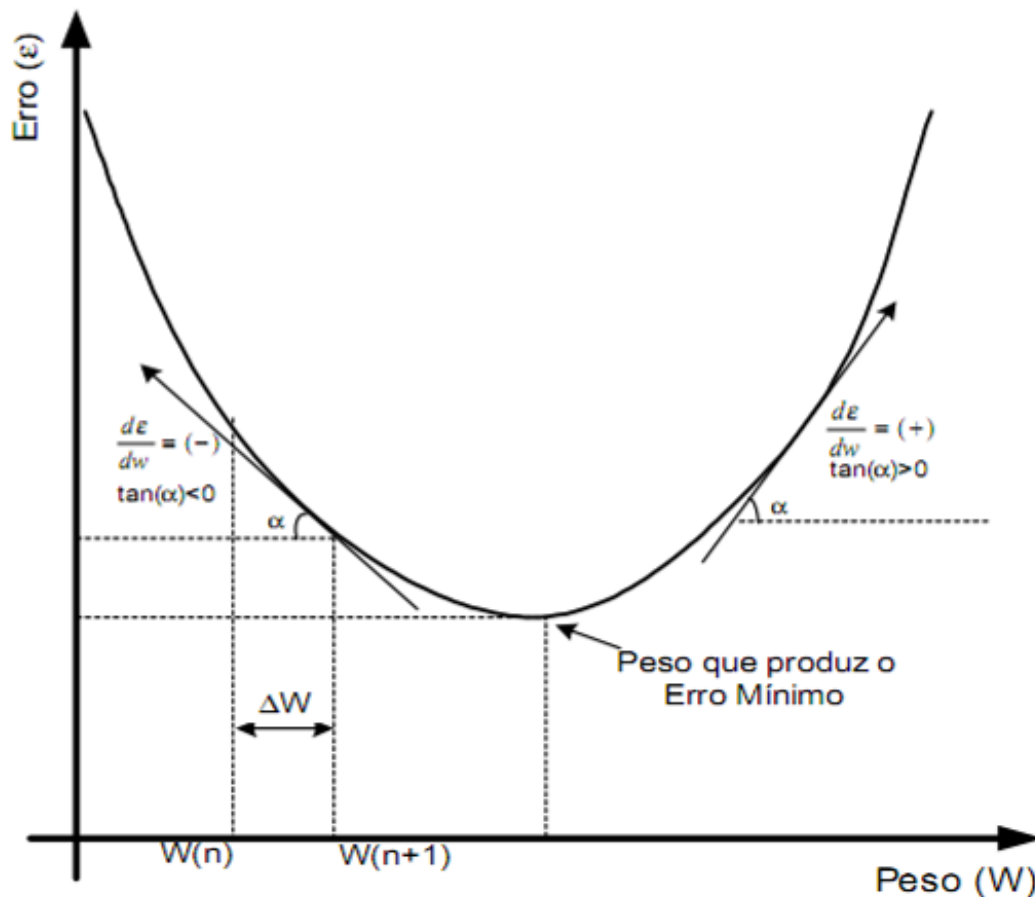
onde η é a taxa de aprendizagem e δ_j é o gradiente local do erro para o neurônio j .

A taxa de aprendizagem é um valor positivo, geralmente menor do que 1, que regula a intensidade com que as atualizações dos parâmetros (pesos) serão efetuadas. Taxas muito baixas, próximas de zero, tendem a fazer com que o aprendizado seja bastante lento, porém taxas muito altas, próximas de 1, podem fazer com que a rede oscile, como se estivesse aprendendo e desaprendendo, e às vezes nem consiga chegar a um patamar aceitável de aprendizado. O valor da taxa de aprendizado não precisa permanecer fixo durante todo o treinamento. Em algumas implementações ela pode ser adaptativa e controlada pela própria rede.

O gradiente local do erro é determinado através do método gradiente descendente. Ele é o termo responsável pela distribuição do erro da camada de saída para as camadas anteriores. O ajuste dos pesos (Δw) deve ser realizado na direção contrária ao gradiente, conforme mostra a Figura 5-10. Se o peso $w(n)$ (valor do peso na iteração n) está à esquerda do erro mínimo, o ajuste Δw deve ser positivo para que $w(n+1)$ (valor do peso da próxima

iteração) esteja mais próximo do valor de w que minimiza o erro. Por outro lado, se o peso $w(n)$ está à direita do erro mínimo, o ajuste Δw deve ser negativo.

Figura 5-10: Erro em função do peso para uma única conexão



Fonte: (LOESCH; SARI, 1996)

O ajuste dos pesos para os neurônios da camada de saída em um algoritmo *backpropagation* é diferenciado do ajuste dos pesos para os neurônios da camada oculta. A seguir será demonstrado o processo de ajuste dos pesos para cada caso separadamente.

- Ajuste dos pesos na camada de saída

Devido ao fato da aprendizagem ser do tipo supervisionada, o resultado desejado para a camada de saída l é fornecido para a rede. Com isto, pode-se fazer uma comparação deste com o resultado obtido pela rede nesta camada, gerando um sinal de erro que é utilizado para realizar o ajuste dos pesos dos neurônios desta camada:

$$\varepsilon_{l,j}(n) = d_{l,j}(n) - y_{l,j}(n) \quad (5.14)$$

Tendo o valor do erro, seu gradiente local é definido como:

$$\delta_{l,j}(n) = \varepsilon_{l,j}(n) \frac{\partial y_{l,j}(n)}{\partial x_{l,j}(n)} \quad (5.15)$$

A expressão para o ajuste dos pesos entre a camada de saída l e a camada $l-1$ é então determinada como:

$$\Delta w_{l-1,i,j}(n) = \eta \delta_{l,j}(n) y_{l-1,i}(n) \quad (5.16)$$

- Ajuste dos pesos na camada oculta

Da mesma forma que os neurônios na camada de saída devem apresentar saídas próximas aos alvos (d), os neurônios na(s) camada(s) oculta(s) também devem exibir saídas determinadas, contudo desconhecidas. A saída desejada para este tipo de camada não é informada para a rede. Neste ponto o algoritmo *backpropagation* justifica seu nome, retropropagando o erro gerado pelos neurônios na camada de saída para as camadas internas, distribuindo o erro para cada um dos neurônios nas camadas ocultas.

O gradiente local do erro para a(s) camada(s) oculta(s) é então definido como:

$$\delta_{l,j}(n) = \sum_k [\delta_{l+1,k}(n) w_{l,j,k}(n)] \frac{\partial y_{l,j}(n)}{\partial x_{l,j}(n)} \quad (5.17)$$

E o ajuste dos pesos entre os neurônios da camada l e da camada $l-1$ é determinado pela Equação (5.16) a qual também determina este ajuste para a camada de saída.

5.3.3 Observações

a) A escolha da função de transferência em uma rede neural *backpropagation* deve obedecer aos requisitos de continuidade, diferenciabilidade e monotonicidade. Estes requisitos são exigidos pelo algoritmo *backpropagation* para permitir que uma expressão analítica, para o ajuste dos pesos da rede, seja obtida.

b) Um termo extra que pode ser adicionado ao ajuste dos pesos na tentativa de melhorar performance da rede é a taxa de momentum. Esta taxa é um parâmetro de uso opcional, de valor positivo menor do que 1, cuja utilização visa imprimir uma dinâmica no treinamento tal que, eventualmente, possibilite que o algoritmo livre-se de mínimos locais durante o processo de busca pelo mínimo global. Para utilizá-la, os pesos de um ou mais padrões de treinamento anteriores devem ser salvos. Sua utilização é recomendada quando alguns dados de entrada são muito diferentes da maioria (FAUSETT, 1994). Normalmente resulta em uma aprendizagem mais rápida, mas pode causar instabilidade em alguns casos se for muito grande. Quando a taxa de momentum μ é utilizada, a expressão para o ajuste dos pesos entre a camada de saída l e da camada $l-1$ passa a ser definida como:

$$\Delta w_{l-1,i,j}(n) = \eta \delta_{l,j}(n) y_{l-1,i}(n) + \mu [w_{l-1,i,j}(n) - w_{l-1,i,j}(n-1)] \quad (5.18)$$

5.4 PROJETO DE UMA REDE NEURAL

Diante de um projeto de uma rede neural, não pensamos mais em procedimentos, regras ou fórmulas algorítmicas de processamento de dados, mas sim em tipos de dados de entrada, dados de saída e tratamento de dados. "O projeto de uma rede neural é baseado diretamente nos dados do mundo real, permitindo-se que o conjunto de dados fale por si mesmo" (HAYKIN, 2001).

O projeto de um sistema neural consiste de diversas etapas que devem ser executadas em sequência, de forma interativa e até mesmo com diversos ciclos de repetição. A construção do sistema começa pela identificação e coleta dos dados históricos relevantes para o problema. O passo seguinte é a preparação e adequação dos dados ao formato requerido pela rede neural, ou seja, a formatação dos dados. Neste processo cria-se uma

escala, estabelecendo um novo intervalo válido dentro do qual todos os dados são colocados. Os intervalos mais utilizados são $[-1,1]$ ou $[0,1]$. Este processo também é conhecido como normalização dos dados.

Após a escolha de uma representação para os dados do problema, deve-se separar os dados em dois conjuntos: o conjunto de treinamento o qual é gerado a partir de dados históricos, ou seja, a partir de experiências e fatos ocorridos no passado; e o conjunto de teste com o qual o funcionamento da rede é testado. Cada dado do conjunto de teste é apresentado uma única vez ao sistema.

O próximo passo é realizar a escolha do modelo neural a ser adotado e definir a topologia da rede. Em seguida realiza-se o desenvolvimento, treinamento e otimização do modelo, seguidos pela validação do mesmo. Na etapa de validação, faz-se uma comparação do resultado obtido pela rede com o resultado desejado.

Por fim, após a validação do modelo, realiza-se a aplicação do mesmo.

5.5 VANTAGENS

Uma das vantagens da rede neural artificial que mais se destaca é a sua capacidade de modelar e prever sistemas não-lineares. Este é o grande diferencial das redes neurais quando comparadas com outros métodos como, por exemplo, modelos estatísticos. Além desta, existem inúmeras outras vantagens atribuídas às redes neurais, dentre as quais se destacam:

- capacidade de encontrar soluções eficientes para problemas do mundo real;
- habilidade de lidar com dados ruidosos, incompletos ou imprecisos;
- capacidade de análise e reconhecimento de padrões;
- capacidade de resolver problemas práticos sem a necessidade da definição de regras ou de modelos precisos;
- capacidade de buscar a solução através de um método próprio de treinamento e auto-aprendizado;
- alta capacidade de adaptação e generalização;

A principal limitação das redes neurais é não ter capacidade de explicação. Os dados entram na rede e uma previsão sai, mas o tipo de relacionamento entre as variáveis não é revelado, assim como detalhes de como as redes raciocinam com os dados para chegar às conclusões também não são fornecidos. Por este motivo, as redes neurais são recomendadas

para serem aplicadas em áreas de conhecimento cujas teorias ainda não conseguem explicar adequadamente o comportamento dos fenômenos observados ou em áreas de grande complexidade que não necessitam de modelos precisos da realidade física do problema.

5.6 APLICAÇÕES

Mesmo com algumas restrições, a área de redes neurais tem demonstrado sua potencialidade em diversas aplicações, superando expectativas e gerando resultados até então não alcançados com qualquer outra técnica, seja computacional ou convencional.

As redes neurais artificiais podem ser treinadas para encontrar soluções, interpretar e classificar dados, reconhecer padrões, aproximar funções e prever eventos futuros. Atualmente, são inúmeras as áreas nas quais as redes neurais têm sido aplicadas, e os bons desempenhos alcançados têm incentivado pesquisadores a fazer das redes neurais uma alternativa de solução para problemas nas mais diversas áreas de atuação.

6 APLICAÇÃO DA METODOLOGIA PROPOSTA

6.1 INTRODUÇÃO

Neste capítulo apresenta-se o desenvolvimento de uma ferramenta inteligente para interpretar imagens termográficas de transformador de baixa tensão objetivando um pré-diagnóstico de falha (indicativo de falha incipiente).

Os sistemas inteligentes utilizados serão Sistema de Inferência Neuro-*Fuzzy* Adaptativo (ANFIS) e rede neural *feedforward* com um algoritmo de retropropagação.

Como entrada de dados serão utilizadas as imagens termográficas em valores discretos de medição de temperatura após pré-processadas por Análise de Componentes Principais

Como alvo (resposta da rede/sistema inteligente) serão utilizadas as conclusões estabelecidas por avaliações físico-químicas e cromatográficas do óleo isolante dos transformadores.

6.2 FERRAMENTAS E EQUIPAMENTOS UTILIZADOS

Para o desenvolvimento das Análises de Componentes Principais, Sistema de Inferência Neuro-*Fuzzy* Adaptativo (ANFIS) e rede neural *feedforward* com um algoritmo de retropropagação, foi utilizado o programa computacional MATLAB, versão 8.1.0.604 (R2013a).

Todas as simulações foram realizadas em um computador com Processador Intel Core i5-4210U; 2,40GHz; 8,00 GB de memória RAM e sistema operacional Windows 10 Pro. O equipamento teve sua utilização restringida a esse fim enquanto se processava a execução das redes.

As imagens termográficas foram obtidas utilizando um termovisor FLIR série-i40, com resolução infravermelha de 14.400 pixels e campo de visão (FOV) 25°x25°. Com o uso do programa FLIR QuickReport 1.2 S1 essas imagens foram convertidas em matrizes de

dimensões 120x120 caracterizando as temperaturas medidas, em uma representação discreta da imagem de resolução de 120x120 pixels.

6.3 BANCO DE DADOS

Os dados para desenvolver este trabalho são as medições de temperatura através da termografia de transformadores de baixa tensão, além da coleta do óleo isolante para obtenção de laudo sobre as condições do equipamento através de análise físico-química e cromatográfica.

De forma complementar, para ajustar as medições termográficas, também são medidas a temperatura atmosférica, velocidade do vento, umidade relativa do ar (Figura 6-1) e distância do ponto de medição ao instrumento, além da imagem de uma folha de alumínio no local direcionado ao transformador, de forma a compensar a temperatura refletida.

Figura 6-1: Medições de temperatura e umidade durante a captura de imagem pelo termovisor



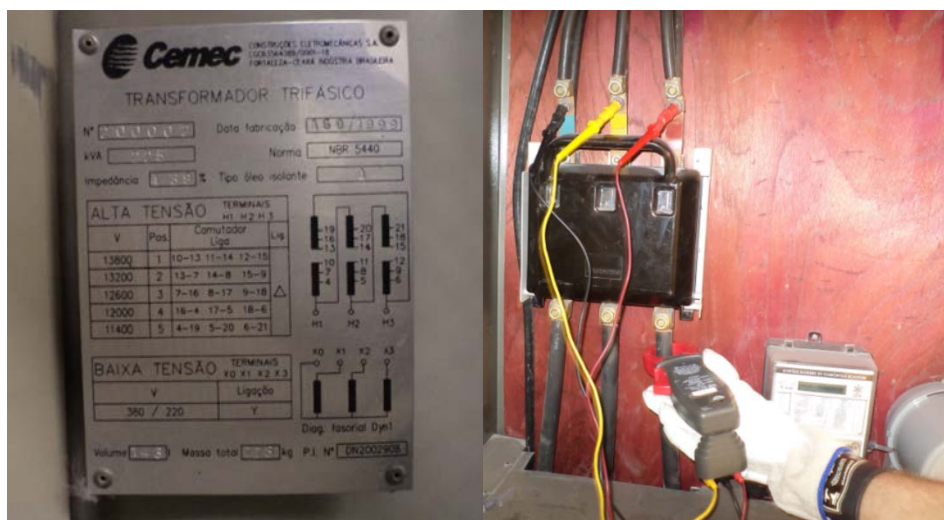
Fonte: próprio autor

As medidas elétricas de tensão, corrente e fator de potência da carga, além da potência de placa do transformador, ilustradas na Figura 6-2, também são registradas de forma a se obter o nível de carregamento do transformador.

A formação da base de dados se iniciou em 2012 sendo acompanhados 26 transformadores de potência (112,5kVA, 13,8kV/380V, com aspectos construtivos similares). Nos períodos de manutenção preventiva que poderiam variar entre seis meses e um ano, foram coletadas amostras do óleo isolante, capturadas imagens termográficas

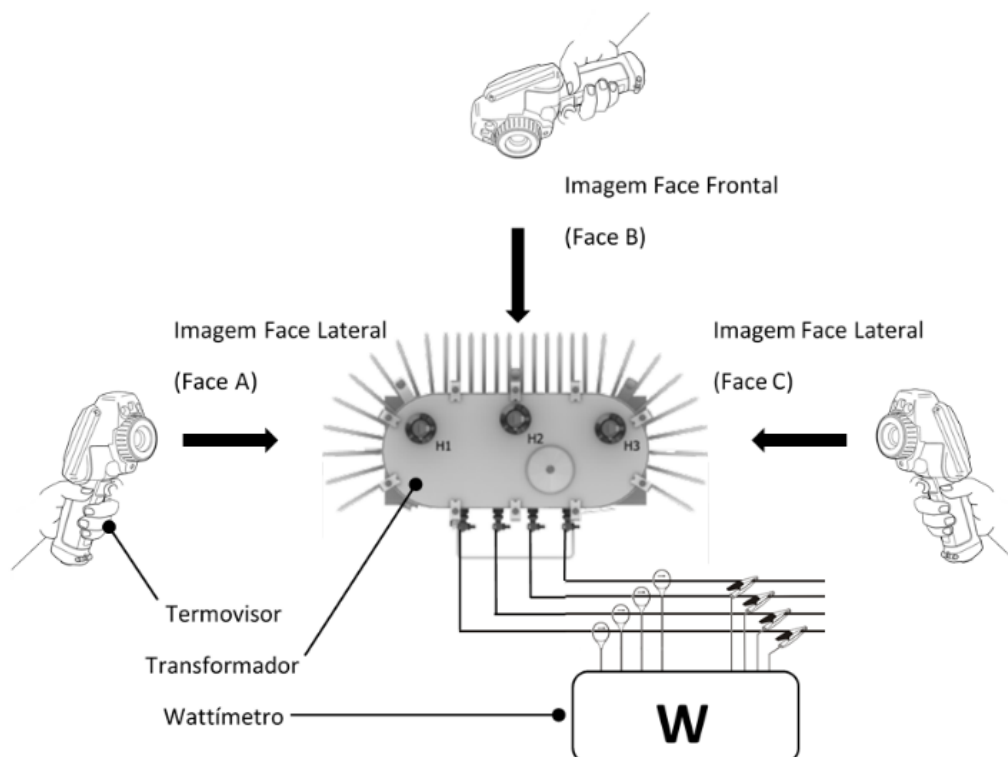
infravermelhas de três lados de suas faces: a sua face frontal (Face B) e as faces laterais (Faces A e C); e seu nível de carregamento elétrico foi medido (Potência Ativa), conforme ilustrações contidas na Figura 6-3. Formando uma base de dados que atualmente conta com 127 registros.

Figura 6-2: Dados de placa de um transformador e medições das grandezas elétricas



Fonte: próprio autor

Figura 6-3: Faces do transformador a serem capturadas pelo termovisor



Fonte: próprio autor

Para o registro das imagens com o termovisor além da medição das variáveis ambientais e elétricas, apenas um profissional foi necessário, com o tempo médio de uma hora para este fim. Considerando-se que para a coleta do óleo isolante os serviços foram executados exclusivamente nos fins de semana pela necessidade do desligamento de energia e a mesma equipe não conseguiu executar mais do que uma subestação por dia, em um mês se consegue executar no máximo oito subestações ou, extrapolando-se para o prazo de um ano, se consegue no máximo 96 subestações. Para os registros das imagens com o termovisor, considerando-se que não há restrições de dia, pois não há a necessidade de desligamento de energia, um mesmo profissional é capaz de avaliar ao menos quatro subestações por dia, ou extrapolando-se para o prazo de 30 dias, é capaz de avaliar 120 subestações.

6.4 VARIÁVEIS PARA O SISTEMA INTELIGENTE (ANFIS E RNA) REDE NEURAL

Uma das decisões mais importantes na construção de um sistema inteligente é a escolha das variáveis nas quais o sistema irá se basear para poder aprender. No caso estudado, é desejado que o sistema entenda a relação entre as imagens termográficas e o resultado das análises do óleo isolante.

É evidente que a imagem termográfica é um dado de entrada, porém qual imagem utilizar? A opção neste caso foi o uso das imagens de três faces do transformador (Figura 6-1) semelhante a sugestão de (ÁLVARES, 2008) utilizada para isoladores e (KITANI e THOMAZ, 2008) para reconhecer faces. Outro dado de entrada considerado é o carregamento do transformador no momento desta medição, justificado pelo fato de que a temperatura do equipamento varia com a carga.

O sinal de saída seria o resultado do laudo do óleo isolante coletado.

6.5 ANÁLISE DO ÓLEO ISOLANTE

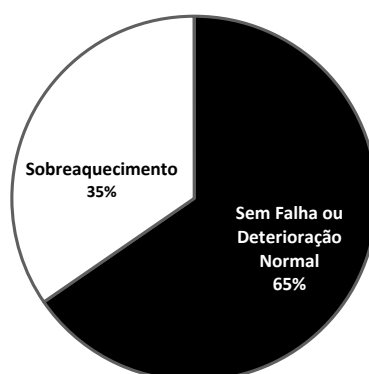
Há mais de quarenta anos o monitoramento de diagnóstico de defeitos em transformadores foi introduzido pela primeira vez no sistema elétrico (NING et al., 2011). Dentre as técnicas relevantes tem-se a Análise dos Gases Dissolvidos (DGA) que ganhou aceitação em todo o mundo como meio de detecção de falhas incipientes e é amplamente usado. Análise de gases dissolvidos (DGA) é uma parte da avaliação não destrutiva de transformadores de potência, através da monitorização de vários parâmetros do óleo

(CHATTERJEE; BHATTACHARJEE; ROY, 2012). Ambas metodologias da IEC (IEC 60599, 2015) e IEEE (C57.104-2008) dispõem guias de referência para a interpretação dos dados DGA. Diversos gases são gerados durante a operação de um transformador. Os principais gases responsáveis pelas falhas são o metano (CH_4), acetileno (C_2H_2), etano (C_2H_6), etileno (C_2H_4) (AL-JANABI et al., 2015), hidrogênio (H_2), o monóxido de carbono (CO) e o dióxido de carbono (CO_2) (GHONEIM; TAHA, 2016) e os gases propano (C_3H_8) e propileno (C_3H_6) (CHRISTIAN; GLÄSER, 2017)

Pela técnica da Análise de Gases Dissolvidos, utilizando a metodologia do IEC e IEEE, resultou que 35% dos transformadores avaliados apresentaram propensão a falha por sobreaquecimento, da mesma forma que em (SEFIDGARAN; MIRZAEI; EBRAHIMZADEH, 2011), conforme Figura 6-4.

Figura 6-4: Resultado da Análise do Óleo Isolante

**TÉCNICA DA ANÁLISE DE GASES DISSOLVIDOS
(IEC E IEEE)**



Fonte: próprio autor

6.6 PREPARAÇÃO DOS DADOS

Tendo o entendimento de como seriam organizadas as variáveis para a preparação do banco de treinamento dos sistemas inteligentes propostos, partiu-se para a adequação dos dados de forma a poder utilizar o Matlab na construção do sistema.

6.6.1 Preparação dos Dados da Imagens Termográficas

Parametrização das medições:

De posse das imagens termográficas, foram iniciados os ajustes nas variáveis medidas em conformidade com as grandezas ambientais aferidas, conforme descrito no Capítulo 2 (item 2.10)

Detalhe apenas para emissividade escolhida: considerando que os transformadores possuem normalmente as mesmas cores e materiais de pintura, optou-se pelo uso da emissividade no valor de 0,96, por ser um intervalo entre a emissividade do Verniz e do Verniz Fosco (Tabela 2-1).

De acordo com (JAFFERY; DUBEY, 2014), a imagem termográfica é exclusivamente baseada na distribuição de calor do sistema e a segmentação de imagens dos termogramas se torna complexo devido ao seu baixo contraste de intensidade e excessivo contraste de distribuição, sendo a extração da região quente dentro do termograma uma tarefa árdua.

Para contornar esta dificuldade, utilizou-se a matriz de temperatura gerada pelo próprio termovisor. Esta matriz é o conjunto das temperaturas da menor área que o equipamento termovisor é capaz de medir. Ou seja, o menor pixel da imagem termográfica. O conjunto desta matriz forma a imagem termográfica. Matematicamente, uma imagem de resolução $m \times n$ pixels é representada pela matriz T de dimensão $m \times n$, onde m é o número de linhas e n é o número de colunas da matriz e cada elemento da matriz é uma temperatura de pixels, como mostra a Equação (6.1)

$$IMAGEM = T = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{m1} & \cdots & t_{mn} \end{bmatrix} \quad (6.1)$$

No entanto, é possível representar também essa matriz como um vetor T d -dimensional, onde $d = m \cdot n$, então o vetor T pode ser representado como em (6.2):

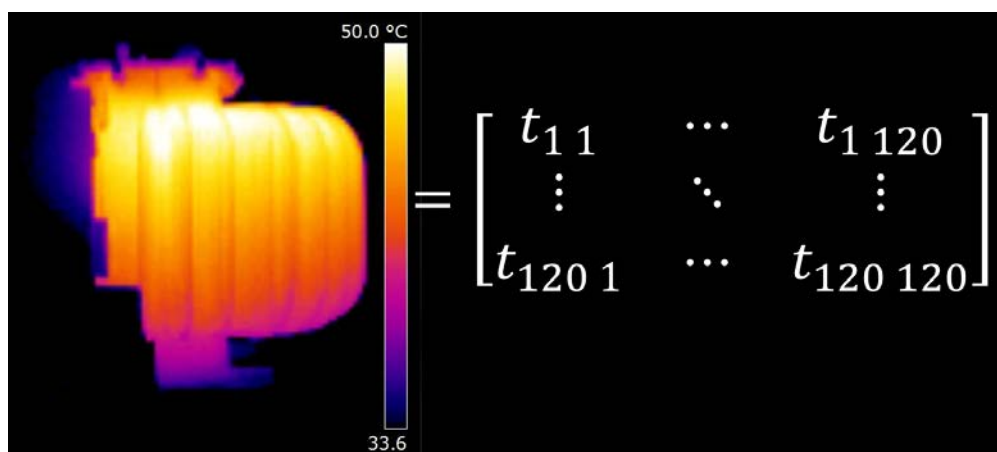
$$T = [t_{11} \quad \cdots \quad t_{1n} \quad \cdots \quad t_{m1} \quad \cdots \quad t_{mn}] \quad (6.2)$$

Sendo o transformador representado pelo conjunto das imagens das três faces (face A, face B e face C), então a imagem do transformador será um vetor T com as três faces (6.3):

$$T = [T_{FaceA} \quad T_{FaceB} \quad T_{FaceC}] \quad (6.3)$$

As imagens termográficas foram obtidas utilizando um termovisor FLIR série-i40, com resolução infravermelha de 14.400 píxels e campo de visão (FOV) 25°x25°. Com o uso do programa FLIR QuickReport 1.2 S1 estas imagens foram convertidas em matrizes de dimensões 120×120 com as medições das temperaturas, uma representação discreta da imagem de resolução de 120×120 píxels, conforme ilustrações contidas na Figura 6-5.

Figura 6-5: Imagem termográfica da face C de um transformador convertida em matriz



Fonte: próprio autor

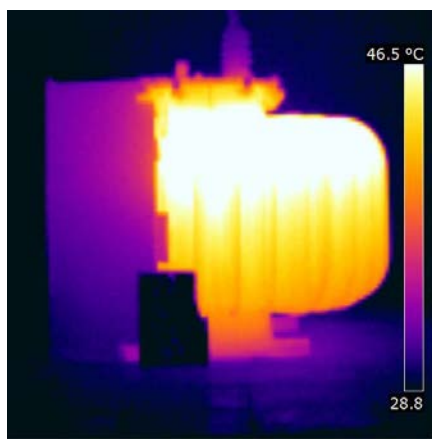
A Figura 6-6 apresenta um exemplo de um transformador objeto de medição com uma folha de alumínio, a Figura 6-7 apresenta a imagem termográfica do equipamento e a Figura 6-8 o uso do programa FLIR QuickReport 1.2 S1 para a edição da imagem com os campos para ajustes dos parâmetros.

Figura 6-6: Exemplo de um transformador objeto de medição



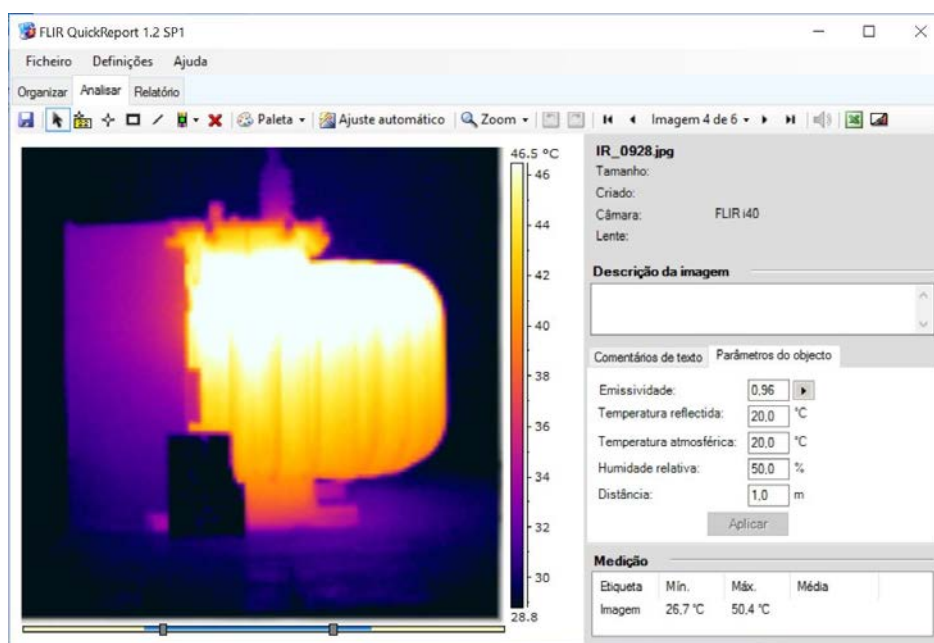
Fonte: próprio autor

Figura 6-7: Imagem termográfica de uma das faces do equipamento



Fonte: próprio autor

Figura 6-8: Programa FLIR QuickReport 1.2 S1



Fonte: próprio autor

Em uma imagem discreta de 120×120 temos uma matriz com 14.400 elementos. Sabendo que cada transformador teve três registros de imagens (três faces), são totalizados 43.200 elementos. Logo, qualquer cálculo envolvendo dados dessa magnitude toma um esforço computacional elevado e fica evidente a necessidade da redução dimensional, para que os dados possam ser manipulados (KITANI; THOMAZ, 2006). A Análise de Componentes Principais (PCA) reduz a dimensionalidade dos dados, concentrando as informações significativas nas primeiras componentes geradas pela transformação (PRICHOA; RIBEIRO, 2013). Através deste processo foram descartados

quaisquer componentes que explicam somente uma pequena proporção da variação nos dados (ARAUJO; COELHO, 2009).

6.6.2 Compressão dos dados utilizando Análise de Componentes Principais

A PCA é uma das mais populares técnicas estatísticas que reduz as dimensões dos conjuntos de dados, mas preserva a estrutura de correlação dos dados (NAGPAL; BRAR, 2014).

A PCA é realizada utilizando a função *pca* do software MATLAB:

```
[wcoeff, score, latent, tsquared, explained] =  
pca (BD, 'VariableWeights', 'variance');
```

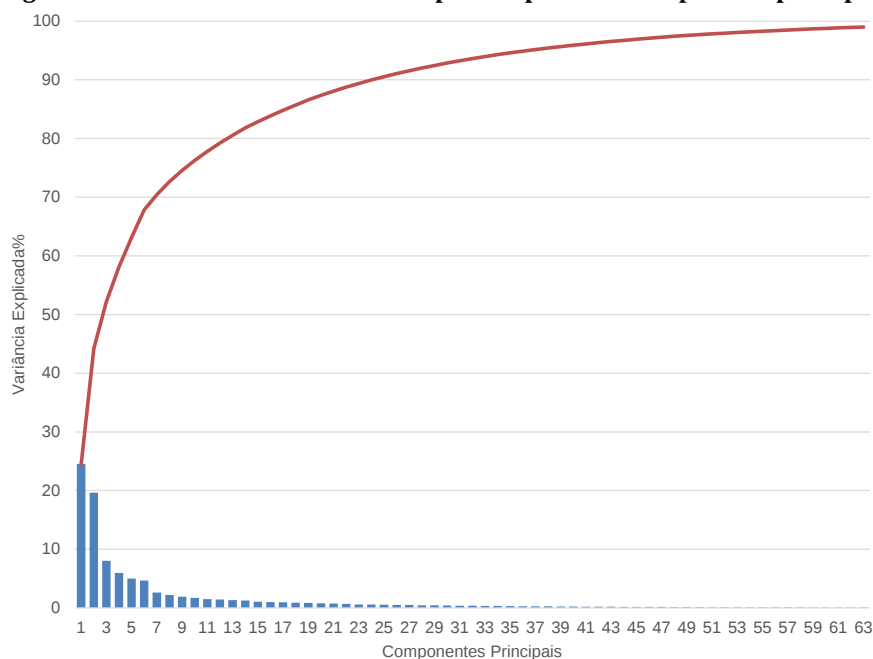
Onde *BD* é a Base de Dados das imagens termográficas, uma matriz 127×43200 , onde cada linha representa cada conjunto de três imagens termográficas registradas de cada transformador.

A primeira saída (*wcoeff*) contém os coeficientes da Componentes Principais, a segunda saída (*score*) contém as coordenadas dos dados originais no novo sistema de coordenadas definido por componentes principais em ordem decrescente. A terceira saída (*latent*) é um vetor que contém a variância explicada pelo componente principal correspondente. Cada coluna de *score* tem uma variância da amostra igual à linha correspondente em *latent*. A quarta saída (*tsquared*) returns the Hotelling's T-squared statistic for each observation.

A quinta saída, (*explained*) é um vetor que contém a variação percentual explicada pelo componente principal correspondente.

Executando esta função na Base de Dados das imagens termográficas (matriz 127×43200) é obtida uma nova Base de Dados de Componentes Principais, cujas 37 e 63 primeiras componentes explicam 95% e 99% do universo, respectivamente, conforme ilustração registrada na Figura 6-9.

Com isso, tem-se definida a base de dados que será utilizada para avaliação em sistemas inteligentes (RNA e Lógica Fuzzy). Será, portanto, uma matriz que conterà 127 registros com as informações que representam as imagens termográficas, adicionada dos registros do nível de carregamento do transformador e o resultado da análise do óleo isolante.

Figura 6-9: Percentual de variância explicada por cada componente principal

Fonte: próprio autor

6.7 REDE NEURAL

6.7.1 Introdução

As redes neurais são poderosas ferramentas para resolver tarefas complexas de aprendizagem (FERREIRA; LUDERMIR; DE AQUINO, 2013). As redes neurais artificiais (RNAs) são aplicadas extensivamente para problemas de reconhecimento de padrões. Elas são adaptativas, capazes de lidar com relacionamentos não-lineares, e são capazes de generalizar soluções para um novo grupo de dados (MORAIS; ROLIM, 2006). Seguindo esta característica, o processo de detecção de falhas incipientes em transformadores usando uma RNA pode ser visto como o processo de associar entradas (no caso as imagens de termografia infravermelha) para saídas (estado do transformador)

Uma Rede Neural Artificial é caracterizada por ter em sua arquitetura muitas unidades de processamento de baixo nível com alto grau de interconectividade através de conexões ponderadas (pesos). Entre muitos conceitos a RNA Multilayer Feedforward Neural Network é a mais utilizada comumente (CASTRO; MIRANDA, 2005)

A utilização de uma rede neural artificial exige que uma série de escolhas, não triviais, sejam feitas na busca de um modelo que apresente um resultado considerado satisfatório. Dentre as principais escolhas estão: a topologia da rede, o algoritmo de

aprendizagem, a função de ativação, a taxa de aprendizagem, a taxa de momentum e o número ideal de épocas.

Serão apresentados a seguir os passos utilizados para o desenvolvimento da rede neural.

Frequentemente, é dito que o projeto de uma rede neural utilizando o algoritmo de retropropagação é mais uma arte do que uma ciência, significando que muitos dos numerosos fatores envolvidos no projeto são o resultado da experiência particular de cada um. (HAYKIN, 2001, p. 205).

6.7.2 Embaralhar os Dados

De acordo com (HAYKIN, 2001), nas tarefas de classificação de padrões usando a aprendizagem por retropropagação sequencial, uma técnica simples bastante utilizada é tornar aleatória a ordem em que os exemplos são apresentados ao perceptron de múltiplas camadas de uma época para a seguinte. Idealmente, a aleatoriedade garante que os exemplos sucessivos apresentados à rede em uma época raramente pertençam à mesma classe. Assim, o desempenho do algoritmo de retropropagação é melhorado.

6.7.3 Normalização dos Dados

Cada variável deve ser pré-processada (FERREIRA; LUDERMIR; DE AQUINO, 2013) de modo que o seu valor médio, calculado sobre todo o conjunto de treinamento, ou seja próximo de zero, ou seja pequeno comparado com o desvio padrão (HAYKIN, 2001). A normalização é realizada para assegurar que todas as variáveis utilizadas tenham a mesma importância durante o treinamento, por isso as funções de ativação dos neurônios artificiais são limitadas. Então, a normalização deve colocar os dados entre os limites inferior e superior da função de ativação (DE AQUINO et al., 2009; LIRA et al., 2006).

Todos os dados foram normalizados para o intervalo $[-1,1]$ utilizando (6.4)

$$X_{norm} = \frac{(2 \cdot X - X_{max} - X_{min})}{X_{max} - X_{min}} \quad (6.4)$$

Onde X é o valor da variável a ser normalizada e X_{norm} é o valor normalizado. $X_{\text{máx}}$ e $X_{\text{mín}}$ são o valor máximo e mínimo em toda a série de dados.

6.7.4 Arquitetura da Rede

Utilizou-se o modelo de rede neural feedforward, com um algoritmo de retropropagação, no qual os pesos são ajustados baseados na regra de aprendizagem por correção de erro. A função de ativação sigmóide na forma de uma tangente hiperbólica foi escolhida para os neurônios da camada oculta e para o neurônio da saída, apesar de (VENKATASAMI; LATHA; KASIRAJAN, 2012) optar pelo uso da função de ativação linear para o neurônio de saída, o uso da função tangente hiperbólica apresentou melhores resultados. Para o treinamento, utilizou-se o algoritmo Levenberg-Marquardt que é uma variação do algoritmo Back-Propagation. Em geral, em problemas de aproximação de função, para as redes que contêm até algumas centenas de pesos, o algoritmo Levenberg-Marquardt terá a convergência mais rápida. Esta vantagem é especialmente perceptível se o treinamento muito exato é necessitado. Em muitos casos, este algoritmo é capaz de obter erros médios quadráticos mais baixos do que outros algoritmos (DEMUTH; BEALE; HAGAN, 2007).

No MATLAB a função utilizada foi a “*patternnet*”, o treinamento foi limitado em 10.000 épocas e o desempenho da rede foi medido pelo erro médio quadrático.

Os dados foram divididos em três conjuntos: Conjunto de treinamento, conjunto de validação e conjunto de teste (FERREIRA; LUDERMIR; DE AQUINO, 2013) na proporção de 60%, 20% e 20% respectivamente.

Segundo (DE AQUINO et al., 2006, 2007), a arquitetura deve ser definida modificando o número de neurônios da camada oculta e selecionada aquela que apresenta a melhor performance no processo de treinamento.

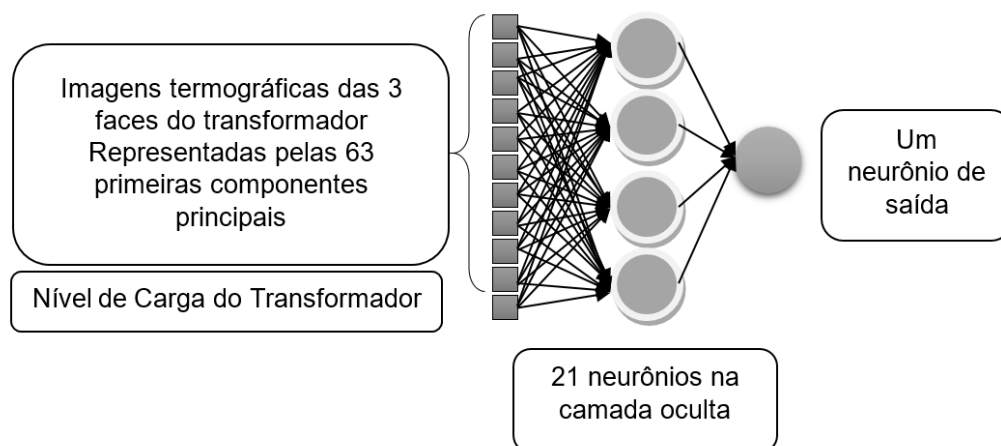
As redes neurais foram testadas com o número de neurônios da camada oculta variando de 3 até 30 neurônios. A rede foi inicializada 20 vezes para cada conjunto de neurônios na camada oculta.

Com este teste, foi verificado que o número de neurônios na camada oculta que apresentou a menor média dos erros foi a arquitetura com 21 neurônios, sendo, portanto, a arquitetura final definida como: (64-21-1), representada na Figura 6-10. Após determinar a melhor arquitetura para o problema em questão, procurou-se uma melhor inicialização de

pesos para esta arquitetura a fim de realizar as simulações em uma rede que tenha boa escolha de pesos iniciais.

Foram testadas 20 redes com inicializações distintas para cada arquitetura e se identificou as que apresentaram o menor erro médio

Figura 6-10: Arquitetura da RNA



Fonte: próprio autor

6.8 LÓGICA FUZZY

6.8.1 Introdução

A Lógica Fuzzy tem sido estudada neste tipo de problema por outros pesquisadores e mostrou-se muito poderosa. Uma vantagem considerável da Lógica Fuzzy é que a sua entrada de dados não precisa ser enorme e pré-processada (DE AQUINO et al., 2010).

6.8.2 Sistema de Inferência Neuro-Fuzzy Adaptativo (ANFIS)

O sistema inteligente ANFIS foi avaliado para esta questão. Tentou-se neste caso encontrar um Sistema Fuzzy com regras que pudessem indicar um resultado semelhante ao da análise do óleo isolante, tendo como dados de entrada as imagens termográficas (devidamente trabalhadas como já apresentado) e o nível de carregamento do transformador.

Por se tratar de uma correlação complexa, optou-se pelo Sistema Neuro-Fuzzy, que utiliza Redes Neurais para encontrar os melhores ajustes dos parâmetros das Funções de Pertinência.

6.8.3 Determinação do Sistema de Inferência Fuzzy (FIS)

Data Clustering

Agrupamento de dados numéricos é a base de muitos algoritmos de classificação e modelagem do sistema. O propósito do agrupamento (*Clustering*) é identificar conjuntos naturais de dados de um grande conjunto de dados para produzir uma representação concisa do comportamento de um sistema.

O MATLAB possui ferramentas para Lógica Fuzzy que permitem descobrir clusters em dados de treinamento de entrada e saída. É possível usar as informações de cluster para gerar um sistema de inferência fuzzy tipo Sugeno. Baseado nos dados de entrada e saída, é possível gerar modelos otimizados usando um número mínimo de regras. As regras se particionam de acordo com as qualidades fuzzy associadas com cada um dos agrupamentos de dados. Foi utilizada a função de linha de comando, *genfis2* para realizar automaticamente este tipo de geração de FIS.

Se não há uma ideia clara quantos clusters deve haver para um determinado conjunto de dados, *Subtractive clustering*, é um algoritmo para estimar o número de clusters e os centros de cluster em um conjunto de dados. As estimativas de clusters, que são obtidas a partir da função *subclust* do Matlab, podem ser usadas para inicializar o método de identificação de modelo (como ANFIS). A função *subclust* localiza os clusters, utilizando o método *Subtractive clustering*.

A função *genfis2* se baseia na função *subclust* para modelar um sistema de inferência fuzzy tipo Sugeno considerando os dados de treinamento de entrada e saída.

MODELO FIS

Utilizando a função *genfis2*, no banco de dados, obteve-se um modelo FIS com os parâmetros customizados.

Uma vantagem importante do uso de um método de *clustering* para encontrar regras é que as regras resultantes são mais adaptadas aos dados de entrada do que eles estão em uma FIS gerada sem *clustering*. Isto reduz o problema de uma propagação excessiva de regras quando os dados de entrada têm uma grande dimensão.

6.8.4 ANFIS

De posse do modelo FIS, utilizou-se recurso de otimização do ANFIS para melhorar esta modelagem.

A função do MATLAB aplicada é a *anfis* utilizando uma parte dos dados para treinamento e outra para validação e com isso evitar um supertreinamento (*overfitting*) que ocasiona a perda da generalidade.

Overfitting pode ser detectada quando os erros nos dados de validação começam a aumentar ao mesmo tempo em que o erro nos dados de treinamento continua a diminuir.

Para verificar o modelo para *overfitting*, foi utilizada a função *anfis* com a inclusão de dados de validação para treinar o modelo para 200 épocas.

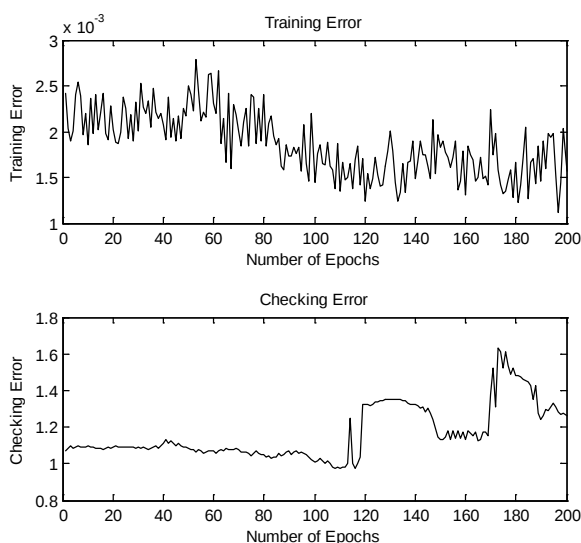
O comando utilizado segue:

```
[fismat3, trnErr, stepSize, fismat4, chkErr] = ...
  anfis([datin datout], fismat, [200 0 0.1], [], ...
  [chkdatin chkdatout]);
```

Onde: *fismat3* é a estrutura FIS para o menor erro com o uso dos dados de treinamento e *fismat4* é a estrutura instantânea FIS tomada quando o erro com o uso dos dados de validação atinge um mínimo.

Os gráficos de erro nos dados de Treinamento (*training*) e Validação (*checking*), visualizados na Figura. 6-11, após o menor valor do erro nos dados de validação, o qual ocorre na época 112, apresentam o erro nos dados de validação aumentando mesmo com o ANFIS continuando a minimizar o erro dos dados de treinamento até a época 200.

Figura 6-11: Minimizando o erro nos dados de Treinamento e Validação



Fonte: próprio autor

6.9 RESULTADOS

6.9.1 Abordagem Redes Neurais

A partir dos dados pré-processados, foram realizadas algumas simulações com diferentes números de neurônios na camada oculta. Avaliou-se 20 RNA's alterando os pesos iniciais. A Tabela 6-1 mostra o erro percentual de classificação no conjunto de teste variando o número de neurônios na camada oculta. O conjunto de teste possui 26 amostras compostos por 13 amostras de cada classe. A RNA_9 com 21 neurônios na camada oculta foi selecionada, pois apresentou o menor erro médio quadrático no conjunto de validação, além de ter mostrado o melhor desempenho no conjunto de teste.

Após determinar a melhor inicialização dos pesos, foi obtida a Rede Neural Artificial que reconhece com o menor erro calculado a correlação entre os dados de entrada (Imagens Termográficas) e de saída (Resultado da Análise do Óleo Isolante).

A Matriz de Classificação ou também chamada de Matriz Confusão, apresentada na Tabela 6-2, indica em sua diagonal a quantidade de casos em que o modelo acertou a classificação, e fora da diagonal ficam os casos em que o modelo indicou uma classe diferente.

Os resultados do erro total por classe da RNA_9 são apresentados na Tabela 6-3 e mostram que a RNA conseguiu atribuir corretamente 92% dos registros dos transformadores diagnosticados sem falhas e 79% dos registros dos transformadores diagnosticados com falha. O percentual total de classificação correta é de 86%, cujo valor está na faixa de bom desempenho mostrada por certos artigos que abordam problemas similares usando RNA's (MORAIS; ROLIM, 2006) e (BEZERRA et al., 2008).

Tabela 6-1: Percentual de erro de classificação no conjunto de teste (RNA)

	20 Neurônios na Camada Oculta	21 Neurônios na Camada Oculta	22 Neurônios na Camada Oculta
RNA_1	24	26	26
RNA_2	26	26	28
RNA_3	17	30	15
RNA_4	26	28	24
RNA_5	23	24	25
RNA_6	26	25	16
RNA_7	17	25	15
RNA_8	26	17	23
RNA_9	26	14	27

	20 Neurônios na Camada Oculta	21 Neurônios na Camada Oculta	22 Neurônios na Camada Oculta
RNA_10	23	22	22
RNA_11	23	27	28
RNA_12	25	27	29
RNA_13	26	23	27
RNA_14	21	24	20
RNA_15	22	25	18
RNA_16	24	26	17
RNA_17	21	16	18
RNA_18	26	23	20
RNA_19	27	23	22
RNA_20	24	19	25

Tabela 6-2: Matriz Confusão (RNA)

Classe Real	Classe atribuída pela RNA		TOTAL
	<i>Sobreaquecimento</i>	<i>Sem Falha ou Deterioração Normal</i>	
<i>Sobreaquecimento</i>	49	13	62
<i>Sem Falha ou Deterioração Normal</i>	5	60	65
TOTAL	54	73	127

Tabela 6-3: Percentual de classificação correta por diagnóstico (RNA)

DIAGNÓSTICO	Amostras	Percentual de Acerto
Sem Falha ou Deterioração Normal	65	92%
Sobreaquecimento	62	79%

6.9.2 Abordagem Lógica Fuzzy

A função *genfis2* possui um argumento chamado de *cluster radius* (ou raio de agrupamento em uma tradução livre). Do modelo FIS, realizou-se algumas simulações com diferentes números de *cluster radius*. Foi utilizada a função *genfis2* para gerar um modelo a partir de dados usando *clustering*. O *genfis2* exige que seja especificada a variável *cluster radius*. O *cluster radius* indica o alcance da influência de um *cluster* quando você considera o espaço de dados como um hipercubo de unidade. Especificar um pequeno *cluster radius* geralmente produz muitos *clusters* pequenos nos dados e resulta em muitas regras. Especificar um *cluster radius* grande geralmente produz alguns *clusters* grandes nos dados e resulta em menos regras. O *cluster radius* é especificado como o terceiro argumento do *genfis2*. Foram avaliados oito modelos FIS alterando o *cluster radius*. A Tabela 6-4 mostra o erro de porcentagem de classificação no conjunto de teste variando o valor do *cluster*

radius. O conjunto de teste possui 26 amostras compostas por 13 amostras de cada classe. O FIS_7 usando um *cluster radius* de valor 0,8 foi selecionado, pois apresentou o menor erro médio quadrático no conjunto de validação, além de ter mostrado o melhor desempenho no conjunto de teste.

Tabela 6-4: Percentual de erro de classificação no conjunto de teste (FIS)

	Cluster Radius	Erro percentual
FIS_1	0,2	27
FIS_2	0,3	24
FIS_3	0,4	23
FIS_4	0,5	29
FIS_5	0,6	26
FIS_6	0,7	27
FIS_7	0,8	17
FIS_8	0,9	20

Após determinar a melhor modelagem do Sistema de Inferência Fuzzy, foi obtido o sistema que reconhece com o menor erro a correlação entre os dados de entrada (Imagens Termográficas) e de saída (Resultado da Análise do Óleo Isolante).

A Matriz Confusão é apresentada na Tabela 6-5. Os resultados do erro total por classe da FIS_7 são apresentados na Tabela 6-6 e mostram que a Lógica Fuzzy conseguiu atribuir corretamente 82% dos registros dos transformadores diagnosticados sem indicativo de propensão a falha e 85% dos registros dos transformadores diagnosticados com indicativo de propensão falha. O percentual total de classificação correta é de 83%.

Tabela 6-5: Matriz Confusão (FIS)

Classe Real	Classe atribuída pela FIS		TOTAL
	Sobreaquecimento	Sem Falha ou Deterioração Normal	
Sobreaquecimento	53	9	62
Sem Falha ou Deterioração Normal	12	53	65
TOTAL	65	62	127

Tabela 6-6: Percentual de classificação correta por diagnóstico (FS)

DIAGNÓSTICO	Amostras	Percentual de Acerto
Sem Falha ou Deterioração Normal	65	82%
Sobreaquecimento	62	85%

7 CONCLUSÕES E CONSIDERAÇÕES FINAIS

As duas simulações propostas (RNA e FIS) detém a mesma finalidade de buscar um sistema inteligente de reconhecimento de imagens termográficas de forma a correlacionar com as condições operativas de um transformador de potência.

A RNA apresentou baixo percentual de erro de classificação 14% no conjunto de teste, enquanto o FIS mostrou 17%.

Todos os sistemas, após treinamento, conseguem convergir para um resultado global satisfatório, com diferença apenas na capacidade de generalização. A RNA apresentou um resultado ligeiramente melhor quando analisado de uma forma global, e a FIS apresentou um maior equilíbrio de resultados entre as classes.

7.1 CONTRIBUIÇÕES

Desenvolvimento de técnica de inteligência artificial para estabelecer o estado do óleo isolante de transformadores de baixa tensão a partir de suas imagens térmicas.

Os resultados mostraram que a abordagem por Redes Neurais Artificiais ou pelo Sistema de Inferência Fuzzy para diagnosticar a falha incipiente de superaquecimento em transformadores de potência de semelhantes classes de tensão e aspectos construtivos eram bem adequados.

Os transformadores de distribuição não possuem acesso fácil para coletar amostra do óleo isolante então há a necessidade de desligá-lo. Destaca-se aí a importância do procedimento não invasivo.

Portanto, foi mostrado que sistemas inteligentes são capazes de correlacionar imagens termográficas com as falhas incipiente de sobreaquecimento do transformador, servindo como um primeiro diagnóstico e um indicativo para uma avaliação mais aprofundada com técnicas invasivas ou avaliações de especialistas.

Eficientização no planejamento da equipe de manutenção com redução nos custos em recursos humanos e redução no intervalo de tempo de avaliação dos equipamentos

As técnicas de manutenção devem ser desenvolvidas sob a estratégia da redução dos tempos de intervenção nos equipamentos, buscando obter o menor tempo de indisponibilidade para o serviço.

A manutenção preditiva se baseia na evolução supervisionada de parâmetros significantes da deterioração do componente, permitindo alongar e planejar intervenções.

Os transformadores de potência são extremamente importantes e todas as pesquisas voltadas para monitorar o desempenho destes equipamentos são válidas. Como técnica mais utilizada para determinar as condições de funcionamento, a análise do óleo isolante serve como referência de meta a ser atingida por qualquer técnica de inteligência artificial.

Considerando que a coleta do óleo isolante de transformadores de baixa tensão necessariamente não pode ser realizada sem acarretar em indisponibilidade do sistema, a termografia infravermelha surge como técnica alternativa, não invasiva.

Propondo automatizar o resultado das inspeções termográficas, os sistemas inteligentes RNA e ANFIS se apresentaram como técnicas a serem utilizadas de forma a aumentar o desempenho dos recursos humanos envolvidos.

Entende-se que o maior fator de dificuldade é obter uma base de dados com características adequadas para o estudo apresentado, considerando que as intervenções em transformadores não ocorrem em intervalos curtos de tempo.

A partir da formação da base de dados, constatou-se que para a intervenção na subestação para coleta do óleo isolante, no mínimo, dois profissionais foram necessários (um técnico em eletrotécnica e um auxiliar), com um tempo médio de 4 horas para execução (considerando o desligamento e religação da concessionária). Nota-se, então, a contribuição deste trabalho nos procedimentos operacionais convencionais na avaliação de transformadores de baixa tensão. Aplicando-se esta técnica, há uma redução de 50% no recurso humano necessário e a redução do prazo da avaliação de 12 meses para 30 dias representa cerca de 90% de redução neste prazo. Significa que o intervalo máximo entre as avaliações de um mesmo equipamento (considerando um parque com 96 transformadores e apenas um profissional), será de apenas 30 dias com esta técnica, o que levaria um ano nos procedimentos convencionais.

Há muitos trabalhos acadêmicos na avaliação de transformadores, mas foi percebido que são focados em transformadores de grande potência que possuem muitos elementos

internos que proporcionam a facilidade na obtenção de dados, além de naturalmente serem geridos por empresas especializadas.

Portanto, este trabalho apresenta uma técnica inovadora, baseada em aspectos de “campo” e não laboratoriais, focando o que pode ser chamado de “varejo” do mercado de energia elétrica, uma forma de proporcionar uma ferramenta inteligente para um processo de manutenção de subestações de pequeno porte, sem se limitar a esses tipos de equipamentos.

7.2 SUGESTÕES DE TRABALHOS FUTUROS

Esta proposta possui uma vasta gama de opções de melhoria, considerando que estes estudos foram realizados sem incentivos financeiros.

Algumas sugestões seriam: avaliar aplicação desta técnica em outros transformadores de diferentes classes de tensão, aspectos construtivos, falhas incipientes, entre outros parâmetros; utilizar outros tipos de topologia de redes neurais artificiais e lógica fuzzy considerando que a realização de testes com diferentes topologias de rede pode vir a melhorar o desempenho do modelo, aproximando ainda mais a saída obtida do resultado desejado; e utilizar outros algoritmos de treinamento.

7.3 CONSIDERAÇÕES FINAIS

Em 2014 os primeiros resultados deste trabalho foram submetidos em forma de artigo ao CONTECC (congresso técnico científico da engenharia e agronomia) sendo selecionado como um dos melhores trabalhos, conferido certificado de reconhecimento ao mérito, e apresentado oralmente no evento(CONFEA, 2014). Em 2017 foi aceito e publicado na revista internacional Electrical Engineering.

A expectativa é que este projeto desperte o interesse de mais pesquisadores, empresas e até fabricantes de instrumentos de termografia, de forma a aprofundar a pesquisa na obtenção de base de dados. Essas pesquisas proporcionariam a utilização de outros tipos de redes neurais ou sistemas Fuzzy. Ou até a tentativa de implementar um certo tipo de programação no próprio equipamento.

Enfim, as alternativas são várias e este assunto acarreta na aproximação do acadêmico as condições reais e simples do “campo” que estão carentes de soluções inovadoras.

REFERÊNCIAS

- AL-JANABI, S. et al. Design and evaluation of a hybrid system for detection and prediction of faults in electrical transformers. **International Journal of Electrical Power & Energy Systems**, v. 67, p. 324–335, maio 2015.
- ÁLVARES, R. C. **Diagnóstico de Falhas em Pára-raios Utilizando Termografia** Universidade Federal de Minas Gerais. [s.l.] UFMG, 2008.
- ARAÚJO, R. A.; BARBOSA, L. C.; SINISCALCHI, R. T. **Os Impactos da Aplicação da Termografia na Operação do Sistema Elétrico de Furnas e as Ações Adotadas para Buscar a Máxima Operacionalidade, Produtividade e Confiabilidade do Sistema**. São Paulo: 2008
- ARAUJO, W. O. DE; COELHO, C. J. **Análise de Componentes Principais (PCA)**. Anápolis: University Center of Anápolis, 2009.
- BARRETO JR, J. T. et al. Diagnóstico de Campo da Condição Operativa de Subestações 138kV através de Termovisão. **Congresso de Inovação Tecnológica em Energia Elétrica**, 2001.
- BERENJI, H. R.; KHEDKAR, P. Learning and Tuning Fuzzy Logic Controllers through Reinforcements. **IEEE Transactions on Neural Networks**, p. 724–740, 1992.
- BEZERRA, J. M. B. et al. **Application of Pattern Recognition Techniques to non Invasive Insulation Monitoring**. Conference Record of the 2008 IEEE International Symposium on Electrical Insulation. **Anais...IEEE**, jun. 2008Disponível em: <<http://ieeexplore.ieee.org/document/4570287/>>
- C57.104-2008: IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers**. . [s.l: s.n.]. Disponível em: <<http://dx.doi.org/10.1109/IEEESTD.2009.4776518>>.
- CARAMALHO, A. **25 anos em termografia**. 1ª ed. [s.l: s.n.].
- CASTRO, A. R. G.; MIRANDA, V. An interpretation of neural networks as inference engines with application to transformer failure diagnosis. **International Journal of Electrical Power & Energy Systems**, v. 27, n. 9–10, p. 620–626, nov. 2005.
- CHATTERJEE, A. et al. Usage of nanotechnology based gas sensor for health assessment and maintenance of transformers by DGA method. **International Journal of Electrical Power & Energy Systems**, v. 45, n. 1, p. 137–141, fev. 2013.
- CHATTERJEE, A.; BHATTACHARJEE, P.; ROY, N. K. Mathematical model for predicting the state of health of transformers and service methodology for enhancing their life. **International Journal of Electrical Power & Energy Systems**, v. 43, n. 1, p. 1487–1494, dez. 2012.
- CHRISTIAN, B.; GLÄSER, A. The behavior of different transformer oils relating to the generation of fault gases after electrical flashovers. **International Journal of Electrical Power & Energy Systems**, v. 84, p. 261–266, jan. 2017.
- CONFEEA. **Conteccc promove divulgação de trabalhos técnicos e científicos em Teresina**. Disponível em: <<http://www.confeca.org.br/cgi/cgilua.exe/sys/start.htm?infoid=19255&sid=10>>. Acesso em: 28 jul. 2017.
- CUENCA, W. M.; SEIXAS, J. M.; LEVY, A. F. **Análise de Componentes Principais para**

Identificar Descargas Parciais em Transformadores de Potência. p. 6, 2009.

DA SILVA, A. C. M.; GARCEZ CASTRO, A. R.; MIRANDA, V. Transformer failure diagnosis by means of fuzzy rules extracted from Kohonen Self-Organizing Map.

International Journal of Electrical Power & Energy Systems, v. 43, n. 1, p. 1034–1042, dez. 2012.

DE AQUINO, R. R. B. et al. **A Hybrid Intelligent System for Short and Mid-term Forecasting for the CELPE Distribution Utility**. The 2006 IEEE International Joint Conference on Neural Network Proceedings. **Anais...IEEE**, 2006Disponível em: <<http://ieeexplore.ieee.org/document/1716455/>>

DE AQUINO, R. R. B. et al. Using Genetic Algorithm to Develop a Neural-Network-Based Load Forecasting. In: [s.l: s.n.]. p. 738–747.

DE AQUINO, R. R. B. et al. **Application of wavelet and neural network models for wind speed and power generation forecasting in a Brazilian experimental wind park**. 2009 International Joint Conference on Neural Networks. **Anais...IEEE**, jun. 2009Disponível em: <<http://ieeexplore.ieee.org/document/5178791/>>

DE AQUINO, R. R. B. et al. **A fuzzy system for detection of incipient fault in power transformers based on gas-in-oil analysis**. International Conference on Fuzzy Systems. **Anais...IEEE**, jul. 2010Disponível em: <<http://ieeexplore.ieee.org/document/5584511/>>

DEMUTH, H.; BEALE, M.; HAGAN, M. **Matlab: Neural Network Toolbox User's Guide**. Natick: Math Works, 2007.

DRIANKOV, D.; HELLENDORF, HANS REINFRANK, M. **An Introduction to Fuzzy Control**. 2ª ed. Berlin: Springer Verlag, 1996.

ENGETERMO. **Soluções em Termografia**, 2009.

FAUSETT, L. **Fundamentals of Neural Networks - Architecture, Algorithms, and Applications**. [s.l.] Prentice Hall International, Inc, 1994.

FERREIRA, A. A.; LUDERMIR, T. B.; DE AQUINO, R. R. B. An approach to reservoir computing design and training. **Expert Systems with Applications**, v. 40, n. 10, p. 4172–4182, ago. 2013.

FLIR SYSTEMS. **Temperature Guns Versus Thermal Imaging Technology**. Disponível em: <<http://www.flir.com/science/blog/details/?ID=71984>>. Acesso em: 26 jul. 2017.

FLUKE CORPORATION. **Introducao aos principios da termografia**. Everett: The Snell Group, 2009.

GHONEIM, S. S. M.; TAHA, I. B. M. A new approach of DGA interpretation technique for transformer fault diagnosis. **International Journal of Electrical Power & Energy Systems**, v. 81, p. 265–274, out. 2016.

HAYKIN, S. **Neural Networks - A Comprehensive Foundation**. 2. ed. Singapore: Pearson Education, 2001.

IEC 60599: Mineral oil-filled electrical equipment in service - Guidance on the interpretation of dissolved and free gases analysis. . [s.l: s.n.].

JAFFERY, Z. A.; DUBEY, A. K. Design of early fault detection technique for electrical assets using infrared thermograms. **International Journal of Electrical Power & Energy Systems**, v. 63, p. 753–759, dez. 2014.

- JANG, J. S. R. ANFIS: Adaptive-Network-based Fuzzy Inference Systems. **IEEE Transactions on Systems, Man and Cybernetics**, p. 665–685, 1993.
- KITANI, E. C.; THOMAZ, C. E. **Um tutorial sobre análise de componentes principais para o reconhecimento automático de faces**. São Paulo: University Center of FEI, 2006.
- LIN, C. T.; LEE, C. S. G. Neural-Network-based Fuzzy Logic Control and Decision System. **IEEE Transactions on Computers**, p. 1320–1336, 1991.
- LIRA, M. M. S. et al. CLASSIFICAÇÃO DE DISTÚRBIOS ELÉTRICOS UTILIZANDO WAVELETS, COMPONENTES PRINCIPAIS E REDES NEURAIS ARTIFICIAIS. n. January 2005, 2005.
- LIRA, M. M. S. et al. **Improving Disturbance Classification by Combining Multiple Artificial Neural Networks**. The 2006 IEEE International Joint Conference on Neural Network Proceedings. **Anais...IEEE**, 2006Disponível em: <<http://ieeexplore.ieee.org/document/1716569/>>
- LOESCH, C.; SARI, S. T. **Redes neurais artificiais: fundamentos e modelos**. Blumenau: Da FURB, 1996.
- MALUTTA, C. **MÉTODO DE APOIO À TOMADA DE DECISÃO SOBRE A ADEQUAÇÃO DE ATERROS SANITÁRIOS UTILIZANDO A LÓGICA FUZZY**. [s.l.] UFSC, 2004.
- MARQUES, J. M.; MARQUES, M. A. M. As componentes principais no descarte de variáveis em um modelo de regressão múltipla. **Revista da FAE**, v. 8, n. 1, p. 93–101, 2005.
- MENDEL, J. M. **Uncertain Rule-Based Fuzzy Logic Systems**. Los Angeles, CA: Prentice-Hall, 2001.
- MORAIS, D. R.; ROLIM, J. G. A Hybrid Tool for Detection of Incipient Faults in Transformers Based on the Dissolved Gas Analysis of Insulating Oil. **IEEE Transactions on Power Delivery**, v. 21, n. 2, p. 673–680, abr. 2006.
- NAGPAL, T.; BRAR, Y. S. Artificial neural network approaches for fault classification: comparison and performance. **Neural Computing and Applications**, v. 25, n. 7–8, p. 1863–1870, 1 dez. 2014.
- NASCIMENTO, J.; YONEYAMA, T. **Inteligência Artificial Controle e Automação**. 1^a ed. São Paulo: Edgard Bluche, 2000.
- NAUCK, D.; KLAWONN, F.; KRUSE, R. **Foudations of Neuro-Fuzzy Systems**. New York: Wiley, 1997.
- NILSSON, N. J. **Artificial Intelligence: A New Synthesis**. San Francisco: Morgan Kaufmann Publishers, Inc, 1998.
- NING, L. et al. A time-varying transformer outage model for on-line operational risk assessment. **International Journal of Electrical Power & Energy Systems**, v. 33, n. 3, p. 600–607, mar. 2011.
- NORTON, E. C. Condition monitoring and assessment of power transformers using computational intelligence. **International Review of Economics & Finance**, v. 20, n. 4, p. 850–851, out. 2011.
- PRICHOA, C. E.; RIBEIRO, S. R. A. **Aplicação da análise de componentes principais em dados extraídos automaticamente de imagens de satélite landsat 5 TM**. Anais XVI

Simpósio Brasileiro de Sensoriamento Remoto - SBSR. **Anais...**2013

SACCHI, R. **POLÍTICA DE OPERAÇÃO PREDITIVA ESTABILIZADA VIA TERMO INERCIAL UTILIZANDO “ANALYTIC SIGNAL”, “DYNAMIC MODELLING” E SISTEMAS INTELIGENTES NA PREVISÃO DE VAZÕES AFLUENTES EM SISTEMAS HIDROTÉRMICOS DE POTÊNCIA.** [s.l.] USP, 2009.

SEFIDGARAN, M.; MIRZAIE, M.; EBRAHIMZADEH, A. Reliability model of the power transformer with ONAF cooling. **International Journal of Electrical Power & Energy Systems**, v. 35, n. 1, p. 97–104, nov. 2011.

SHAW, I. S.; SIMÕES, M. G. **Controle e Modelagem Fuzzy.** [s.l.] Edgard Blucher, 1999.

SILVA, A. M. DE S. **Ferramenta inteligente de classificação do diagnóstico de falhas incipientes em transformadores de potência imersos em óleo.** [s.l.] Federal University of Pernambuco, 2012.

SILVA, L. M. DA et al. Modelagem fuzzy aplicada na detecção da vulnerabilidade à erosão costeira. **Boletim de Ciências Geodésicas**, v. 19, n. 4, p. 746–764, dez. 2013.

SUGENO, M.; KANG, G. T. Structure Identification on Fussy Model. **Fuzzy Sets and Systems**, p. 15–33, 1988.

TAKAGI, T.; SUGENO, M. Fuzzy Identification of Systems and Its Applications to Modeling and Control. **IEEE Transactions on Systems, Man and Cybernetics**, p. 116–132, 1985.

TANSCHKEIT, R. **Lógica Fuzzy, Raciocínio Aproximado e Mecanismos de Inferência.** Rio de Janeiro: [s.n.].

VARELLA, C. A. A. **Análise de componentes principais.** [s.l.] Universidade Federal Rural do Rio de Janeiro, 2008.

VASCONCELOS, S. **Análise de Componentes Principais.** [s.l.: s.n.].

VENKATASAMI, A.; LATHA, P.; KASIRAJAN, K. **Application of neural networks for transformer fault diagnosis.** 2012 IEEE 10th International Conference on the Properties and Applications of Dielectric Materials. **Anais...IEEE**, jul. 2012Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6318975>

VERATTI, A. B. Sistema Básico de Inspeção Termográfica. p. 6, 1999.

YADAV, R. et al. Condition Based Maintenance of Power Transformer : A Case Study. p. 6–8, 2008.