



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

HUGO RIVIERE SILVA MORAES

**REDES NEURAIIS PROFUNDAS APLICADAS AO DIAGNÓSTICO DE FALTAS
INCIPIENTES EM TRANSFORMADORES IMERSOS EM ÓLEO ISOLANTE**

DM: 26 / 2019

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2019

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

HUGO RIVIERE SILVA MORAES

**REDES NEURAS PROFUNDAS APLICADAS AO DIAGNÓSTICO DE FALTAS
INCIPIENTES EM TRANSFORMADORES IMERSOS EM ÓLEO ISOLANTE**

DM: 26 / 2019

Dissertação submetida à Banca Examinadora do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Pará para obtenção do Grau de Mestre em Engenharia Elétrica na área de Computação Aplicada, elaborada sob a orientação da Profa. Dra. Adriana Rosa Garcez Castro.

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2019

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**REDES NEURAIIS PROFUNDAS APLICADAS AO DIAGNÓSTICO DE FALTAS
INCIPIENTES EM TRANSFORMADORES IMERSOS EM ÓLEO ISOLANTE**

AUTOR: HUGO RIVIERE SILVA MORAES

Dissertação de mestrado submetida à avaliação da banca examinadora aprovada pelo colegiado do programa de pós-graduação em Engenharia Elétrica da Universidade Federal do Pará e julgada adequada para obtenção do grau de Mestre em Engenharia Elétrica na área de Computação Aplicada com ênfase em Inteligência Computacional.

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2019

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

- S586r Silva Moraes, Hugo Riviere
REDES NEURAIIS PROFUNDAS APLICADAS AO
DIAGNÓSTICO DE FALTAS INCIPIENTES EM
TRANSFORMADORES IMERSOS EM ÓLEO ISOLANTE /
Hugo Riviere Silva Moraes. — 2019.
xv;53 f. : il. color.
Orientador(a): Prof^ª. Dra. Adriana Rosa Garcez Castro
Dissertação (Mestrado) - Programa de Pós-Graduação em
Engenharia Elétrica, Instituto de Tecnologia, Universidade Federal
do Pará, Belém, 2019.
1. Rede Neural Convolutacional. 2. Rede Autocodificadora.
3. análise gascromatográfica. 4. faltas incipientes. I. Título.

CDD 006.32

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**REDES NEURAIIS PROFUNDAS APLICADAS AO DIAGNÓSTICO DE FALTAS
INCIPIENTES EM TRANSFORMADORES IMERSOS EM ÓLEO ISOLANTE**

AUTOR: HUGO RIVIERE SILVA MORAES

DISSERTAÇÃO DE MESTRADO SUBMETIDA À AVALIAÇÃO DA BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA DA UNIVERSIDADE FEDERAL DO PARÁ E JULGADA ADEQUADA PARA OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE COMPUTAÇÃO APLICADA.

APROVADA EM ____/____/_____.

BANCA EXAMINADORA

Prof. Dr.

Prof. Dr.

Prof. Dr.

Prof. Dr.

VISTO:

Prof Dr.(COORDENADOR DO PPGGE/IETC/UFPA)

“Dedico este trabalho à minha esposa Gisele e a minha família”.

AGRADECIMENTOS

Primeiramente quero agradecer ao Criador, por me fazer quem eu sou e me dar a dádiva de conquistar estas vitórias e compartilhá-las com as pessoas que eu amo. Agradeço à minha família e esposa pelo apoio e incentivo de sempre. À minha orientadora, Prof.^a Adriana Rosa Garcez Castro, que depositou a confiança em mim desde à seleção e por sua paciência e apoio.

Aos meus professores e colegas de turma por todo o apoio durante as disciplinas. Ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pelo suporte e oportunidade, bem como a CAPES por prover a bolsa de mestrado, mostrando sempre cumprir a missão de fomentar a pesquisa e desenvolvimento desta instituição.

“Confie no Senhor de todo o seu coração e não se apoie em seu próprio entendimento; reconheça o Senhor em todos os seus caminhos, e ele endireitará as suas veredas. Não seja sábio aos seus próprios olhos; tema ao Senhor e evite o mal.” Provérbios 3:5-7

RESUMO

Realizar o diagnóstico de faltas incipientes em transformadores mostra-se um grande desafio pois, devido à variabilidade das condições em que ocorrem as faltas, torna-se muito difícil definir a origem e o tipo de falta. Métodos convencionais baseados na análise dos gases dissolvidos em óleo vêm sendo utilizados pelas empresas para realizar o diagnóstico de faltas, porém, estes métodos ainda necessitam ser aplicados em conjunto para se chegar a um resultado satisfatório, além de dependerem bastante do conhecimento de um especialista da área. Visando solucionar as dificuldades relacionadas aos métodos convencionais, alguns sistemas baseados em Inteligência Computacional vêm sendo propostos na literatura e têm apresentado resultados promissores. Considerando a importância do diagnóstico de faltas em transformadores, este trabalho apresenta os resultados do estudo desenvolvido da aplicação de Redes Neurais profundas ao diagnóstico de faltas. Dois modelos são propostos utilizando Redes Neurais Convolucionais e Redes Neurais Autocodificadoras Empilhadas. Para o desenvolvimento dos sistemas foi utilizada a base de dados TC 10 com situações de transformadores faltosos. Esta base foi utilizada para desenvolvimento do método IEC 60599, que é um dos principais métodos usados pelas concessionárias de energia para diagnóstico de transformadores através da análise dos gases dissolvidos em óleo. Os resultados promissores alcançados com os dois modelos propostos (acurácia de 100% na base de teste) se sobressaíram a outros encontrados na literatura, mostrando a grande aplicabilidade das redes neurais profundas ao problema de diagnóstico de faltas incipientes de transformadores. Portanto, mostra-se uma melhor alternativa que os métodos convencionais geralmente utilizados.

Palavras-chave: Rede Neural Convolucional. Rede Autocodificadora. Análise gascromatográfica. Faltas incipientes.

ABSTRACT

Diagnosing incipient faults in transformers is a major challenge because it is very difficult to define the source and type of fault, due to the variability in the conditions under which faults occur. Conventional methods based on the analysis of dissolved gases in oil have been used by companies to diagnose faults, however, these methods still need to be applied together to reach a satisfactory result, as well as relying heavily on the knowledge of a specialist. In order to solve the difficulties related to conventional methods, some systems based on Computational Intelligence have been proposed in the literature and have presented promising results. This paper presents the results of the study developed of the application of deep neural networks to fault diagnosis, considering then the importance of fault diagnosis in transformers. Two models are proposed using Convolutional Neural Networks and Stacked Autoencoding Neural Networks. For the development of the systems we used the TC 10 database with faulty transformer situations. This base was used to develop the IEC 60599 method, which is one of the main methods used by power utilities for transformer diagnostics through the analysis of dissolved gases in oil. The promising results achieved with the two proposed models (100% accuracy in the test base) show the great applicability of deep neural networks to the problem of incipient transformer fault diagnosis, however showing a great alternative to the conventional methods commonly used.

Keywords: Convolutional Neural Networks. Stacked Autoencoder. Gas chromatography. Incipient faults.

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1 - Rede Neural Autocodificadora. Destacando as partes que a compõe..... | 9 |
| Figura 2 - Rede Neural Stacked-Autocoder. | 11 |
| Figura 3 - Arquitetura de uma rede convolucional..... | 13 |
| Figura 4 - Operação de convolução entre uma camada de entrada e um filtro de convolução resultando em um neurônio da camada convolucional..... | 14 |
| Figura 5 - Max Pooling com um filtro 2x2 e stride igual a 2. | 15 |
| Figura 6 - Vista em corte lateral de um transformador de potência. | 21 |
| Figura 7- Esquema básico de um sistema de cromatografia. | 27 |
| Figura 8 - Coordenadas e zonas de faltas no triângulo de Duval. | 31 |
| Figura 9 - Representação gráfica do IEC 60599..... | 34 |
| Figura 10 - Arquitetura de Rede Neural Autocodificadora Empilhada..... | 37 |
| Figura 11 - Processo de pré-treinamento de uma rede autoencoder..... | 38 |
| Figura 12 - Treinamento de uma rede stacked sparse autoencoder. | 39 |
| Figura 13 - Taxas de erro de classificação variando o número de neurônios da primeira e segunda autocodificadora. | 40 |
| Figura 14 - Taxas de erro de classificação variando os neurônios na primeira autocodificadora. | 41 |
| Figura 15 - Taxas de erro de classificação encontradas variando os neurônios na segunda camada. | 41 |
| Figura 16 - Melhor topologia para o sistema de diagnóstico via Autocodificadora Empilhada. | 42 |
| Figura 17 - Matriz de confusão para base de dados de teste. | 42 |
| Figura 18 - Arquitetura da rede convolucional com maior desempenho. | 45 |
| Figura 19 - Matriz de confusão dos dados de teste da rede neural convolucional de melhor desempenho. | 45 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1- Características dos óleos parafínicos e naftênicos | 22 |
| Tabela 2- Tabela de definição do método Rogers. | 28 |
| Tabela 3 - Método de Rogers em função do código..... | 29 |
| Tabela 4 - Critérios do Método de Dörnenburg. | 30 |
| Tabela 5 - Concentração de gases característicos para validação do método. | 30 |
| Tabela 6 - Critério de Diagnóstico do Método Gás-Chave. | 32 |
| Tabela 7 - Critério do método IEC. | 33 |
| Tabela 8 - Divisão da base de dados aplicada nos sistemas de diagnóstico..... | 37 |
| Tabela 9 – Codificação do tipo de falta na saída da Auto codificadora Empilhada..... | 38 |
| Tabela 10–Arquitetura 1..... | 43 |
| Tabela 11–Arquitetura 2..... | 43 |
| Tabela 12–Arquitetura 3..... | 43 |
| Tabela 13 - Arquitetura4..... | 44 |
| Tabela 14–Arquitetura 5..... | 44 |
| Tabela 15 - Arquitetura6..... | 44 |
| Tabela 16 - Desempenho das arquiteturas de rede convolucional..... | 44 |
| Tabela 17 - Medidas de dispersão da acurácia obtidas em 100 divisões aleatórias. | 46 |
| Tabela 18 - Comparação de resultados dos modelos propostos e método IEC 60599..... | 46 |
| Tabela 19 - Exemplos de diagnóstico para alguns casos da base TC-10. | 47 |
| Tabela 20 - Resultados encontrados na literatura..... | 48 |

LISTA DE SIGLAS

SEP – Sistema Elétrico de Potência

DGA – *Dissolved Gas Analysis*

SAE – *Stacked Auto Encoder*

LSTM – *Long Short-Term Unit*

CNN – *Convolutional Neural Network*

OMI – Óleo Mineral Isolante

OVI – Óleo Vegetal Isolante

SCG – *Scaled Conjugate Gradient Backpropagation*

ANEEL – Agência Nacional de Energia Elétrica

LAQUANAM – Laboratório de Química Analítica e Ambiental

MLP – *Perceptron de Múltiplas Camadas*

ReLU – *Rectified Linear Unit*

SUMÁRIO

| | |
|--|-----------|
| 1. INTRODUÇÃO..... | 2 |
| 1.1. Descrição do problema | 2 |
| 1.2. Objetivo Geral | 4 |
| 1.3. Objetivos específicos..... | 4 |
| 1.4. Estrutura da dissertação | 5 |
| 2. REDES NEURAS PROFUNDAS | 6 |
| 2.1. Introdução | 6 |
| 3. DIAGNÓSTICO DE FALTAS EM TRANSFORMADORES | 18 |
| 3.1. Introdução | 18 |
| 3.2. Componentes de um Transformador de Potência | 20 |
| 3.3. Propriedades do Óleo Isolante..... | 21 |
| 3.3.1. Óleo Mineral Isolante | 21 |
| 3.3.2. Óleo Vegetal Isolante | 22 |
| 3.4. Impacto da Perda de Isolação na Vida Útil do Transformador | 23 |
| 3.5. Métodos de Análise dos Gases Dissolvidos em Óleo..... | 25 |
| 3.5.1. Método Clássico de Rogers | 28 |
| 3.5.2. Método de Dörnenburg..... | 29 |
| 3.5.3. Triângulo de Duval..... | 30 |
| 3.5.4. Método do Gás-Chave | 32 |
| 3.5.5. Método IEC 60599 | 33 |
| 4. APLICAÇÃO DE REDES NEURAS PROFUNDAS AO DIAGNÓSTICO DE TRANSFORMADORES..... | 36 |
| 4.1 Introdução | 36 |
| 4.2 Base de dados de transformadores com faltas | 36 |
| 4.3 Sistema de Diagnóstico de Faltas Via Autocodificadoras Empilhadas..... | 37 |
| 4.4 Sistema de diagnóstico de Faltas via Rede Neural Convolutacional..... | 43 |
| 4.5 Aplicação da amostragem aleatória | 45 |
| 5. CONCLUSÃO | 49 |

1. INTRODUÇÃO

1.1. Descrição do problema

Em um Sistema Elétrico de Potência (SEP) existem diversos tipos de transformadores que desempenham funções vitais para garantir o bom funcionamento do sistema. Dentre os diversos transformadores, destaca-se o transformador de potência, responsável por garantir a transmissão e distribuição de potência elétrica através da elevação e diminuição dos níveis de tensão, possibilitando assim a redução de perdas na transmissão de grandes quantidades de energia através de longas distâncias (FONSECA, 2014).

Os transformadores de potência precisam ser constantemente monitorados para garantia de seu bom estado e bom funcionamento, evitando assim problemas de interrupção no fornecimento de energia, o que poderia acarretar graves perdas sob os aspectos técnico, econômico e social.

Um transformador, quando em serviço, pode estar sujeito a uma variedade de estresses elétricos e térmicos que podem conduzir o equipamento para uma condição de falta (elétrica ou térmica). A detecção e eliminação destas faltas, em condição ainda incipiente, é primordial para assegurar o bom funcionamento do sistema de energia.

Entre os diversos métodos de diagnóstico de faltas incipientes de transformadores destacam-se os métodos baseados na análise dos gases dissolvidos em óleo isolante (DGA - *Dissolved Gas Analysis*). Este método de análise consiste na extração de amostras do óleo isolante para submeter à uma cromatografia gasosa, ou seja, a um processo de análise química instrumental a fim de se identificar e quantificar os níveis de gases dissolvidos no óleo. Os Métodos de Dörnenburg (DUVAL; DEPABLO, 2001), de Rogers (Rogers, 1978), Triângulo de Duval (Duval, 2002) e IEC 60599 (DUVAL; DEPABLO, 2001), são alguns destes métodos, sendo que estes estabelecem critérios para classificação do tipo de falta ocorrida em transformadores através da razão entre as concentrações de gases dissolvidos no óleo. O método IEC 60599 é um dos métodos mais utilizados pelas empresas de energia, sendo que utiliza a relação entre cinco gases comumente gerados em condição de falta: hidrogênio (H_2), acetileno (C_2H_2), etileno (C_2H_4), etano (C_2H_6) e metano (CH_4). Em geral, estes métodos permitem detectar dois tipos de faltas, a elétrica e a térmica, e se baseiam na comparação das concentrações de gases dissolvidos com um “*benchmark*” contendo valores típicos de concentrações de gases.

Garantir a confiabilidade do funcionamento dos transformadores mostra-se um desafio, devido à dificuldade de se obter diagnósticos precisos que apontem faltas incipientes. De acordo com (MIRANDA et al., 2005) a análise dos gases dissolvidos em óleo atualmente não pode ser considerada uma ciência, mas sim uma arte sujeita a variabilidade, visto que transformadores de diferentes tamanhos, estruturas e fabricantes apresentam características diferentes e, como consequência, é difícil se obter um consenso quanto aos valores limites dos gases e seus significados.

A natureza heurística e empírica dos métodos baseados em razões tem levado a diferentes interpretações, visto que a aplicação de métodos diferentes sobre um conjunto idêntico de dados pode produzir diferentes diagnósticos sobre a condição do transformador. Devido a imprecisão de um determinado método, diversos métodos devem ser utilizados ao mesmo tempo para se obter um diagnóstico satisfatório. Neste caso, especialistas devem utilizar seus conhecimentos em conjunto com os resultados para se chegar a uma conclusão. Existem casos em que o diagnóstico é impossível de ser obtido devido à inabilidade destes métodos para fornecer uma interpretação para algumas combinações de valores de razões dos gases (problema de “*no decision*”) (DA SILVA, 2013).

A escassez de dados confiáveis a respeito de transformadores faltosos destaca-se entre as dificuldades encontradas para diagnosticar faltas incipientes através de métodos convencionais (MIRANDA; CASTRO; LIMA, 2012). O próprio método apresentado na IEC 60599, norma de referência, se baseia em uma base de dados de transformadores faltosos considerada pequena.

Visando solucionar as dificuldades relacionadas aos métodos convencionais, alguns sistemas baseados em Inteligência Computacional vêm sendo propostos na literatura, com resultados bastante promissores (CHENG; YU, 2018). Dentre os modelos propostos, destacam-se os baseados em Redes Neurais Artificiais (YI; WANG; WANG, 2016), (SOUAHLIA; BACHA; CHAARI, 2012), (COLORADO et al., 2010), Lógica Difusa (HUANG; YANG; HUANG, 1997), sistemas Neuro-difusos (NARESH; SHARMA; VASHISTH, 2008), (CASTRO; MIRANDA, 2005), (CASTRO; MIRANDA, 2004) e Teoria do Sistema Grey (GST) (CHENG; YU, 2018), (DONG; YAN; TANIGUCHI, 2003), (SONG BIN et al., 2002).

Atualmente, as redes neurais baseadas em aprendizado profundo (*Deep Learning*) vêm se destacando cada vez mais na literatura, com aplicações em diversas áreas. O aprendizado profundo permite que modelos computacionais aprendam representações de dados com múltiplos níveis de abstração (LECUN; BENGIO; HINTON, 2015). Esses métodos vêm sendo aplicados com sucesso em muitos campos como: reconhecimento de fala, reconhecimento de

objetos visuais, detecção de objetos, tradução de idiomas, entre outros. Além dos resultados promissores para problemas envolvendo a aplicação de dados de imagem 2-D, alguns autores vêm desenvolvendo pesquisas na área de aplicação das redes neurais profundas em problemas com dados 1-D (PENHA; CASTRO, 2018), tais como dados de séries temporais. Dentre as redes neurais profundas destacam-se as *Long Short Term Units*- LSTM, as Redes Autocodificadoras Empilhadas (*Stacked Auto-Encoders* - SAE) e as Redes Neurais Convolucionais (*Convolutional Neural Network* - CNN).

1.2. Objetivo Geral

Considerando os resultados promissores já apresentados na literatura da aplicação de redes neurais profundas para problemas de classificação através de dados 1-D, este trabalho apresenta os resultados obtidos da pesquisa de aplicação de redes neurais profundas para o problema de diagnóstico de faltas incipientes em transformadores baseado na análise dos gases dissolvidos em óleo. Para o desenvolvimento do sistema de diagnóstico foi utilizada a base de dados TC10 de equipamentos faltosos inspecionados em serviço usado para a publicação da norma IEC 60599 e publicado em (DUVAL; DEPABLO, 2001).

1.3. Objetivos específicos

Considerando o objetivo principal deste trabalho, foram estabelecidos os seguintes objetivos específicos:

- Desenvolver dois modelos utilizando redes neurais profundas. O primeiro modelo, baseado em uma Rede Autocodificadora Empilhada e o segundo modelo baseado em uma Rede Neural Convolucional;
- Avaliar o desempenho de ambos os modelos e realizar a comparação dos resultados a fim de estabelecer qual método foi mais preciso no diagnóstico das faltas incipientes;
- Comparar os resultados obtidos dos modelos de classificação de faltas com os resultados obtidos pelo método convencional IEC 60599;
- Comparar os resultados obtidos neste trabalho com resultados de sistemas de classificação baseados em Inteligência Computacionais já apresentados na literatura.

1.4. Estrutura da dissertação

Este trabalho está estruturado da seguinte forma:

Capítulo 2 – Apresenta a teoria em que as redes neurais profundas estão embasadas e um breve histórico, contendo os marcos do desenvolvimento das redes neurais profundas. São mostradas as principais aplicações destas redes sendo detalhado os dois modelos empregados neste trabalho.

Capítulo 3 – Apresenta uma visão geral a respeito de diagnóstico de faltas incipientes em transformadores isolados à óleo e descreve os principais métodos convencionais de diagnóstico de faltas em transformadores.

Capítulo 4 – Demonstra como os modelos de redes neurais profundas foram aplicados ao diagnóstico de transformadores, sendo exposto os resultados alcançados e realizada a comparação dos modelos propostos neste trabalho.

Capítulo 5 – Apresenta a conclusão do trabalho, avaliando o modelo proposto que obteve maior precisão no diagnóstico.

2. REDES NEURAIIS PROFUNDAS

2.1. Introdução

Métodos de aprendizagem profunda fazem parte do campo de estudo denominado *Representation Learning*, isto é, aprendizado de representações de dados. Segundo (BENGIO; COURVILLE; VINCENT, 2013), a aprendizagem profunda tornou-se um campo de pesquisa muito importante para a comunidade voltada para o aprendizado de máquina, com workshops regulares nas principais conferências mundiais, como NIPS¹ e ICML², e uma nova conferência dedicada a ele, a ICLR³.

Este campo de estudo parte do princípio que os humanos utilizam representações e abstrações para a resolução de problemas e o desenvolvimento do raciocínio. Ele propõe estabelecer técnicas de aprendizado de máquina que permitam a um sistema simular o comportamento do cérebro humano, detectando automaticamente representações de dados necessárias à extração de características provenientes de dados brutos (*raw data*). O objetivo principal é facilitar a extração de informação útil para a aplicação de classificadores ou preditores em tarefas de aprendizado de máquina.

O sucesso dos algoritmos de aprendizado de máquina geralmente depende de como o dado está sendo representando. Esta afirmação parte da hipótese de que diferentes representações podem embarçar e ocultar os diferentes fatores explicativos da variação por trás dos dados. Por exemplo, em uma tarefa de detecção de objetos em imagens tem-se a dificuldade de distinguir tais objetos de suas sombras, pois uma imagem é composta da interação de uma ou mais fontes de luz, das formas e das propriedades dos materiais presentes nela. No cenário em que sombras de um objeto estão projetadas em outros podem surgir formas complexas, criando a ilusão de contornos onde não existem e afetando drasticamente a tarefa de detecção. Portanto, para superar este tipo de desafio, tem-se empregado a abordagem de

¹*Neural Information Processing Systems* – NIPS é uma conferência organizada anualmente por uma organização sem fins lucrativos que tem a finalidade de promover o intercâmbio de pesquisas sobre sistemas de processamento de informações neurais sob os aspectos biológicos, tecnológicos, matemáticos e teóricos.

²*International Conference on Machine Learning* – ICML é uma conferência apoiada pela Sociedade Internacional de Aprendizagem Automática cujo principal objetivo é fomentar a pesquisa em aprendizado de máquina.

³*International Conference on Learning Representations* – ICLR é o principal encontro de profissionais dedicados ao avanço do ramo da inteligência artificial chamado aprendizagem de representação, mas, geralmente chamado de aprendizagem profunda.

utilizar uma vasta quantidade de exemplos não-rotulados para aprender representações que separam as várias fontes explicativas (BENGIO; COURVILLE; VINCENT, 2013).

Existem várias formas de aplicar *representation learning*, dentre elas estão as redes neurais que, devido ao aumento do poder de processamento dos computadores, tem sido objeto de estudo e implementação em diversos setores da sociedade. Basicamente, existem dois tipos de redes neurais: as tradicionais, ou superficiais (*shallow networks*) e as profundas (*deep networks*).

Este trabalho tem como objetivo a aplicação de redes neurais profundas ao problema de diagnóstico de faltas em transformadores e, portanto, para um bom entendimento sobre a proposta, este capítulo se dedica a fundamentação teórica sobre estas redes neurais.

2.2. Redes Neurais Profundas

Redes neurais de aprendizagem profunda se caracterizam como modelos computacionais compostos de múltiplas camadas de processamento que aprendem representações de dados através de um processo que envolve a extração de características em diferentes níveis de abstração. Estas redes são construídas pela composição de módulos simples, não-lineares, onde cada módulo transforma a representação de um nível mais tangível, dado por uma entrada sem pré-processamento (*raw input*), ou por uma entrada submetida a um pré-processamento, para um nível que seja mais abstrato. Com a composição de tais transformações, funções muito complexas podem ser aprendidas, possibilitando que o método de aprendizagem profunda extraia um conjunto de características úteis para detectar representações complexas, como, por exemplo, a de uma face humana (LECUN; BENGIO; HINTON, 2015).

A aprendizagem profunda foi inspirada por uma pesquisa desenvolvida nos anos 60, quando David H. Hubel e Torsten Wiesel, médicos neurofisiologistas, desenvolveram um projeto que tinha o objetivo de compreender o funcionamento do sistema visual (HUBEL; WIESEL, 1959). Foram realizados experimentos com gatos, que consistia em inserir um microelétrodo no córtex visual dos felinos e projetar padrões de luz e sombra em frente ao mesmo no intuito de observar a sua atividade cerebral. Células simples e células complexas foram encontradas no córtex visual felino. Essas células se ativam em resposta a certas propriedades das entradas sensoriais visuais, tal como na mudança de orientação dos padrões exibidos. Foi observado que células complexas exibem mais invariância espacial do que as simples. Essa observação inspirou os pesquisadores na área de Inteligência Artificial a desenvolver os modelos de redes neurais profundas (SCHMIDHUBER, 2015).

Os primeiros pesquisadores tentaram aplicar os métodos de aprendizagem profunda em reconhecimento de padrões com redes neurais multicamadas, porém, o problema não foi totalmente compreendido até meados de 1980, quando grupos de pesquisa diferentes descobriram que as arquiteturas multicamadas poderiam ser treinadas com um simples método chamado *Backpropagation* (SCHMIDHUBER, 2015). Apesar do sucesso com este algoritmo de aprendizado, ele era utilizado apenas para a resolução de problemas considerados superficiais. O seu potencial para aprendizado profundo só foi descoberto em 1991 com a tese do Dr. Hochreiter (HOCHREITER, 1991), considerado um marco histórico na pesquisa de métodos de aprendizagem profunda. Hochreiter descobriu a principal dificuldade de implementar o *Backpropagation* em aprendizagem profunda: as redes neurais tradicionais passavam pelo problema do desaparecimento ou explosão do vetor gradiente, isto é, com a aplicação de funções de ativação padrão, como a sigmoide, sinais de erro se acumulavam e eram retro propagados causando uma diminuição repentina ou crescimento absurdo de seus valores. Este problema ficou conhecido originalmente como *long time lag problem* (em português, problema de atraso de tempo).

Ao longo dos anos foram explorados vários meios de resolver este problema; dentre eles está o uso de pesos aleatórios no espaço de matrizes de peso das redes neurais sem depender de gradientes de erro (HOCHREITER; SCHMIDHUBER, 1996).

Um importante avanço na resolução deste problema foi alcançado por Geoff Hinton e seguidamente por outros pesquisadores (HINTON; SALAKHUTDINOV, 2006), (BENGIO; COURVILLE; VINCENT, 2013), (AURELIO RANZATO; BOUREAU; CUN, 2008). A ideia central refere-se ao uso de pré-treino não-supervisionado por camadas, ou seja, aplica-se aprendizado não-supervisionado para aprender uma nova transformação em cada camada a ser composta com as transformações aprendidas em camadas anteriores. Basicamente, cada iteração do aprendizado não-supervisionado adiciona uma camada de pesos na rede neural profunda. Por fim, o conjunto de camadas podem ser combinadas para inicializar uma tarefa de aprendizado profundo supervisionado.

Após o pré-treino, as características extraídas podem ser utilizadas como entrada de um algoritmo padrão de aprendizado de máquina, ou como pesos iniciais de um modelo de rede neural profunda. O pré-treino também pode ser aplicado no modo de aprendizado supervisionado. Apesar de não ser uma técnica melhor que a anterior, como é demonstrando em (BENGIO; LECUN, 2007), ainda assim é melhor que não utilizar o pré-treinamento.

Outra maneira de pré-treinamento é vista em (SEIDE et al., 2011). É realizado de forma supervisionada em cada camada da rede adicionada anteriormente a cada etapa da iteração, e

em seus experimentos essa variante discriminante produziu melhores resultados do que o pré-treinamento não-supervisionado.

As redes neurais profundas são responsáveis pelos avanços em reconhecimento de fala, processamento de imagem, texto, vídeo, fala e áudio. Além disso, estão sendo aplicadas em muitos outros domínios. Espera-se que em um futuro próximo o método de aprendizado não-supervisionado alcance proporções maiores de aplicação, afinal, busca-se cada vez mais automatizar o processo de reconhecimento de padrões (LECUN; BENGIO; HINTON, 2015).

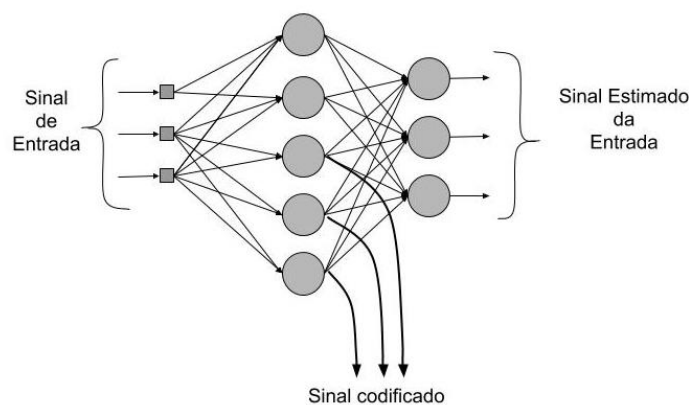
Neste trabalho, foram aplicados dois modelos de aprendizagem profunda, *Stacked Auto-Encoder* e Rede Neural Convolucional, que serão apresentados a seguir.

2.3. Rede Neural Autocodificadora Empilhada (*Stacked Auto-Encoder*– SAE)

A Rede Autocodificadora (em inglês, *Auto Encoder* - AE) é um caso especial da rede *Perceptron* de Múltiplas Camadas (em inglês, *Multi-Layer Perceptron* - MLP), onde o número de neurônios da camada de entrada é igual ao número de neurônios da camada de saída, sendo que a rede deve ser treinada para mapear a função identidade entre entrada e saída, ou seja, os padrões de entrada devem ser reconstruídos na saída da rede com a maior precisão possível. A Figura 1 apresenta a estrutura de uma rede AE.

A camada intermediária ou escondida da rede AE é responsável pela extração de características das informações fornecidas através dos neurônios de entrada. A camada intermediária permite que a rede armazene somente conhecimento relacionado às principais características do conjunto de dados de entrada, atuando assim como uma camada extratora de características.

Figura 1 - Rede Neural Autocodificadora. Destacando as partes que a compõe.



Fonte: Autor

Portanto, em uma AE a camada escondida, chamada de codificadora, é responsável por codificar os dados, enquanto a camada de saída é responsável por reconstruir as entradas originais através dos dados codificados. Desta forma, a rede realiza uma simples aproximação da função identidade, fazendo o papel de reconstruir na saída os dados de entrada.

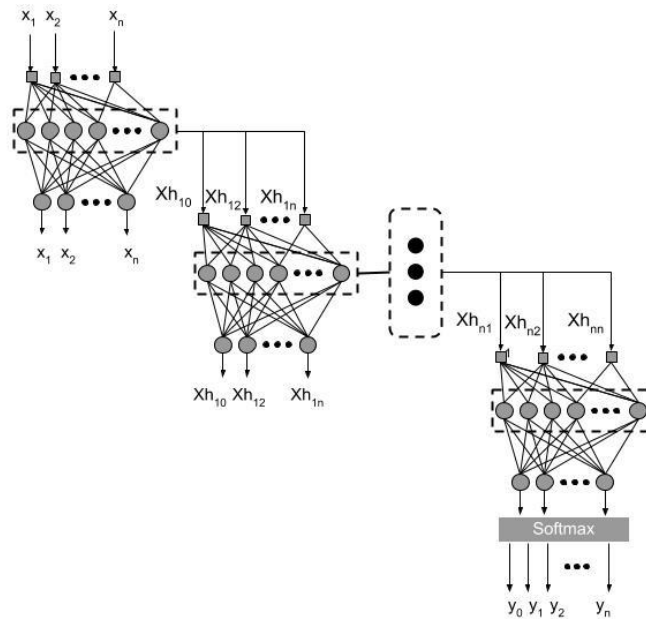
Este processo de codificação ocorre com a projeção da informação de um espaço de variáveis S para um espaço S' , emulando uma função f , e a decodificação ocorre com a projeção inversa de S' a S emulando a função inversa f' . Se o vetor de saída for igual ao de entrada, então é comprovado que a informação foi codificada corretamente, concluindo o processo de aprendizado da rede. Um vetor que não é codificado corretamente produzirá na saída um vetor distinto, indicando um erro eventualmente grande.

Através de um bom treinamento uma autocodificadora pode ser utilizada para reconstruir padrões de entrada que possuem correlação em suas estruturas, por exemplo, é possível submeter à rede exemplos de imagens de uma mesma categoria e obter na saída a reconstrução dela.

Para o treinamento de uma autocodificadora pode ser utilizado o algoritmo *Backpropagation* padrão, com minimização do erro médio quadrático entre os valores dos neurônios de entrada e saída.

Uma Rede Autocodificadora Empilhada (em inglês, *Stacked Auto-Encoder - SAE*) é composta por diversas autocodificadoras empilhadas e por uma camada onde a função *Softmax* é aplicada. A estrutura de uma SAE pode ser observada na Figura 2. Cada rede autocodificadora compartilha os dados codificados na camada escondida, que servem de entrada para uma autocodificadora posterior sendo as saídas de cada uma descartada, exceto da última. Por fim, temos a camada responsável pelo processo de classificação, onde o sinal codificado na camada escondida da última rede AE treinada é fornecido como entrada para uma rede MLP que possui como saída a função *Softmax*. Logo, é realizado um treinamento a fim de que a MLP possa ajustar, a partir dos dados rotulados, seus pesos sinápticos para minimizar o erro entre a saída desejada para o sistema classificador e a saída da SAE.

Figura 2 - Rede Neural Stacked-Autocoder.



Fonte: Autor

A camada com a MPL tem no neurônio de saída a função *Softmax* que é comumente empregada em problemas de classificação onde se tem múltiplas classes.

$$S(x^T w_j) = \frac{e^{x^T w_j}}{\sum_{k=1}^K e^{x^T w_k}} \quad (1)$$

onde j corresponde ao índice da classe em que se deseja avaliar, w representa a matriz de pesos sinápticos de determinada amostra, x é o vetor de entrada e k indica os índices pertencentes as outras classes avaliadas no modelo.

Para o treinamento da MLP com saída com função *softmax* é utilizado o algoritmo de aprendizado *Backpropagation*.

2.4. Redes Neurais Convolucionais (*Convolutional Neural Network – CNN*)

Nos últimos anos pode-se presenciar o avanço exponencial no uso de *Deep Learning*, ou Aprendizado profundo, em tarefas como processamento de imagem, de áudio, entendimento de linguagem natural, entre outras. Entretanto, as primeiras aplicações bem-sucedidas desta técnica remontam da década de 80. A rede neural convolucional desenvolvida por LeCun e outros, nos laboratórios da AT&T Bell Labs, foi um caso pioneiro de sucesso. A sua arquitetura foi batizada com o nome de LeNet 1. À época, a rede foi utilizada para reconhecimento de escrita, utilizando uma base de dados de códigos postais escritos à mão (LECUN et al., 1989).

Uma rede convolucional é constituída por uma estrutura de dados multidimensional. A arquitetura de uma rede convolucional, conforme Figura 4, é composta basicamente por cinco camadas: de entrada, convolucional, *pooling*, função de ativação e totalmente conectada. Esta arquitetura está organizada em camadas, formada por volumes de dados de três dimensões: largura, altura e profundidade. As duas primeiras camadas possuem níveis, onde são dispostos os neurônios ativados nas operações de convolução e *pooling*, respectivamente. Cada nível, chamado de mapa de características (*feature maps*), é responsável por extrair características de uma entrada. A quantidade de mapas é determinada pelo número de filtros (ou *kernels*) definidos para a camada convolucional (LECUN et al., 1999).

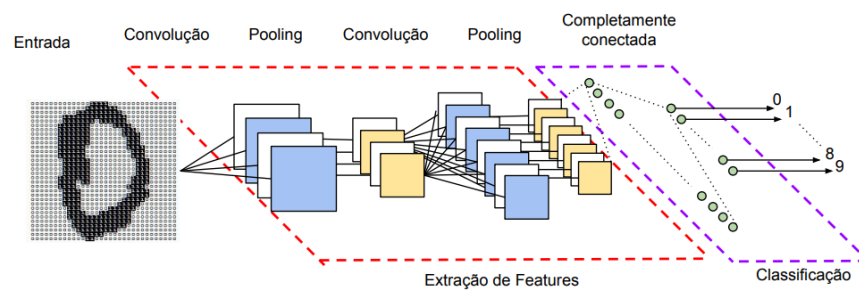
De forma mais detalhada, tem-se as seguintes camadas:

- De entrada: como o próprio nome sugere, nesta camada são inseridos os dados de entrada, sendo que em algumas aplicações podem ser dados brutos (*raw input*), como imagens em seu formato binário, ou os dados podem ser submetidos a um pré-processamento simples, como por exemplo, transformação de sinais em imagens com valores de intensidade de cor de cada pixel correspondendo a intensidade do sinal (WEN et al., 2017);
- De convolução: nesta camada é realizada a operação que denomina este tipo de rede, ou seja, a convolução, que ocorre entre o campo receptivo (*receptive field*) de uma imagem na camada de entrada e a primeira camada oculta da rede. O mesmo procedimento repete-se entre camadas mais profundas, onde são empregados os conceitos de compartilhamento de pesos (*weight sharing*), que será descrito posteriormente;
- ReLU: nesta camada a função linear retificada (*Rectified Linear Unit - ReLU*) é aplicada para realizar a ativação dos neurônios da camada anterior. Em (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), é mostrado que o treinamento de uma rede convolucional de 4 camadas com esta função foi 6 vezes mais rápido do que com o uso da função tangente hiperbólica;
- *Pooling*: nesta camada as principais características extraídas de uma camada anterior são agrupadas reduzindo a quantidade de parâmetros livres e desta forma prevenindo o *overfitting*;
- Totalmente conectada: esta camada constitui-se basicamente de uma MLP totalmente conectada que faz o papel de calcular os *scores* das classes, onde pode-se analisar qual classe se sobressai, ou seja, aquela que tem maior aderência à amostra submetida na entrada.

O uso de redes convolucionais com aprendizado supervisionado ocorre através da aplicação do algoritmo *Backpropagation*, onde o cálculo do erro é realizado a partir de um método chamado de gradiente descendente estocástico conjugado (*Scaled Conjugate Gradient*

- SCG), que consiste em apresentar para a rede o vetor de entradas, calcular as saídas e os erros, calcular o gradiente médio e ajustar os pesos. O processo é repetido em mini lotes de amostras do conjunto de treinamento até que a média da função objetivo se estabilize (erro mínimo). É chamado de estocástico porque a cada mini lote de exemplos é obtida uma estimativa do gradiente médio considerando todos os exemplos. Este procedimento simples geralmente encontra um bom conjunto de pesos de forma surpreendentemente mais rápida do que quando comparado a técnicas de otimização muito mais elaboradas (BOTTOU; BOUSQUET, 2008).

Figura 3 - Arquitetura de uma rede convolucional.

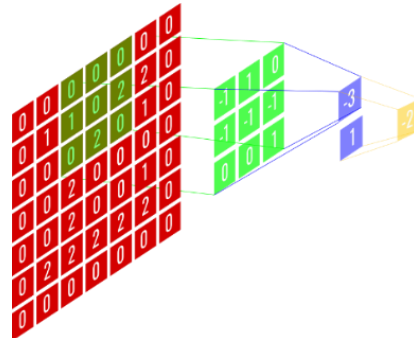


Fonte: Vargas et al., 2016.

Após o treinamento, o desempenho do sistema é medido em um conjunto diferente de exemplos chamado de conjunto de teste. Isso serve para testar a capacidade de generalização da rede - sua capacidade de produzir respostas corretas em amostras que não foram apresentadas durante o treinamento.

As Redes Neurais Convolucionais realizam a conexão entre os neurônios de suas camadas de uma forma diferente de redes neurais tais como a MLP. A conexão entre camadas ocorre de forma local, por exemplo, um neurônio da primeira camada convolucional se conecta apenas a uma sub-região de neurônios da camada de entrada. Portanto, é realizada a operação de convolução discreta entre saídas de uma camada anterior e um conjunto de filtros da camada de convolução. A Figura 5 apresenta de maneira detalhada a operação de convolução (LECUN et al., 1999).

Figura 4 - Operação de convolução entre uma camada de entrada e um filtro de convolução resultando em um neurônio da camada convolucional.



Fonte: Autor

Cada filtro possui um tamanho que é determinado por um hiperparâmetro chamado campo receptivo, portanto, este determina o tamanho da “janela” de convolução. Já o volume de saída é determinado por três hiperparâmetro: profundidade, *stride* e *zero-padding*. O primeiro controla a profundidade do volume de saída e corresponde ao número de filtros que se deseja utilizar na camada. O segundo determina a quantidade de “saltos” que se realiza no volume de entrada para aplicar a convolução discreta. E o terceiro determina a quantidade de zeros que se deseja adicionar nas bordas do volume de entrada para que seja possível realizar a operação de convolução discreta (SIMONYAN; ZISSERMAN, 2014).

É comum inserir periodicamente uma camada de agrupamento (*pooling*) entre sucessivas camadas convolucionais. Sua função é reduzir progressivamente o tamanho espacial da representação para reduzir a quantidade de parâmetros e computação na rede neural e, portanto, também controlar a superposição. A camada de *pooling* opera independentemente em cada fatia de profundidade da entrada e redimensiona-a espacialmente.

Na camada de *pooling* geralmente é utilizada a função *maxpooling*, uma função que extrai o valor máximo entre neurônios de uma determinada região de um mapa de características, portanto, reduz o volume de amostras da camada anterior.

Como pode ser observado na Figura 5 a operação *maxpooling* consiste em subamostrar espacialmente o volume de uma camada convolucional. Portanto, é definido um tamanho de filtro, comumente 2x2, onde a operação é realizada. Note que a operação consiste em identificar o valor máximo contido em um filtro, desta forma a operação é realizada em toda a profundidade do volume, sendo então esta profundidade preservada.

Figura 5 - Max Pooling com um filtro 2x2 e stride igual a 2.



Fonte: Autor

A função comumente utilizada na ativação de neurônios é denominada Unidade Linear Retificada (ReLU). Uma função não-linear que exerce um grande impacto no desempenho, reduzindo o tempo de treinamento de um modelo. Esta função é definida como,

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

Por fim, tem-se a camada totalmente conectada que representa a saída da rede neural. Nesta camada encontram-se todas as características extraídas pela rede para a classificação dos padrões. Por fim, é aplicada a função *Softmax* para o cálculo da probabilidade de uma determinada amostra pertencer à determinada classe.

$$y_r(x) = \frac{e^{(a_j(x,w))}}{\sum_{k=1}^K e^{(a_k(x,w))}} \quad (3)$$

$$a_j(x, w) = w^T x \quad (4)$$

onde $0 \leq y_r \leq 1$, $\sum_{j=1} y_j = 1$.

A última camada da rede calcula a perda de entropia cruzada para problemas de classificação de múltiplas classes com classes mutuamente exclusivas. Os valores de probabilidade obtidos da função *Softmax* são utilizados como entrada para esta camada, onde cada entrada é atribuída para uma das K classes mutuamente exclusivas usando a função de entropia cruzada. Esta função serve como métrica de comparação entre a saída da rede e a saída desejada. A entropia cruzada penaliza aquelas saídas que divergem das saídas desejadas (*targets*) e recompensa aquelas que possuem valor aproximado.

$$E(\theta) = -\sum_{i=1}^n \sum_{j=1}^k t_{ij} \ln(y_j(x_j, \theta)) \quad (5)$$

onde n e k representam a quantidade de amostras e classes, respectivamente; t_{ij} é o indicador de que a amostra i_{th} pertence à classe j_{th} , θ é o vetor de parâmetros. A saída para a amostra i é definida como $y_j(x_i, \theta)$, que neste caso é o valor da função *Softmax*.

2.4.1. Treinamento da Rede Neural Convolutacional

O treinamento de uma rede convolutacional consiste em duas etapas, chamadas de *forward pass* e *backward pass*. A primeira etapa inicia com a inserção de amostras na rede para serem submetidas a operações como convolução, *pooling*, *padding* e ativação dos neurônios através de funções como a ReLU. Por fim, obtém-se a saída da rede após a conversão de camadas de convolução para as camadas totalmente conectadas presentes em redes neurais tradicionais. Desta forma, inicia-se a segunda etapa, onde é utilizada uma métrica para o cálculo do erro entre a saída desejada e a saída fornecida pela rede. O valor dado pela métrica é utilizado como parâmetro para a aplicação do algoritmo de otimização *Backpropagation*, que utiliza uma técnica denominada (*Scaled Conjugate Gradient Backpropagation*– SCG) (MØLLER, 1993) para ajustar os parâmetros livres da rede e obter um desempenho com melhor generalização.

A técnica SCG possui algumas propriedades que a tornam vantajosa em relação à aprendizagem por lote (*Batch Learning*). Com o uso desta técnica a atualização dos pesos se torna on-line, ou seja, o algoritmo calcula o gradiente a cada conjunto de amostras aplicado na entrada, e desta forma, o ruído presente na atualização dos pesos pode ser uma vantagem, pois existe uma probabilidade maior de se encontrar o mínimo global (LECUN et al., 1998). Portanto, a cada iteração, uma única amostra do conjunto de treinamento é escolhida e o gradiente é calculado,

$$W(t + 1) = W(t) - \eta \frac{\partial E^t}{\partial W} \quad (6)$$

onde $W(t+1)$ representa a atualização dos pesos, $W(t)$ o conjunto de pesos atual, η é a taxa de aprendizado e E^t o erro associado à saída da rede.

Por conta da natureza estocástica do gradiente, o treinamento não assegura que a atualização dos pesos sempre reduzirá o valor do gradiente a cada iteração.

Para melhorar a convergência do gradiente, isto é, a oscilação do seu valor, pode ser aplicado um termo denominado *momentum*:

$$\Delta w(t + 1) = \frac{\eta \partial E_{t+1}}{\partial w} + \mu \Delta w(t). \quad (7)$$

O termo de momentum é responsável por determinar a velocidade de aprendizagem da rede. O Gradiente Descendente Estocástico com *Momentum*, η , determina a contribuição do gradiente encontrado na iteração anterior para a atual.

3. DIAGNÓSTICO DE FALTAS EM TRANSFORMADORES

3.1. Introdução

Os transformadores possuem um papel crítico no funcionamento de um Sistema Elétrico de Potência, sendo a manutenção deste tipo de equipamento essencial, visto que a ocorrência de um sinistro ou até mesmo o mau funcionamento do transformador pode ocasionar consequências drásticas, como a interrupção no fornecimento de energia, causando graves perdas econômicas. Quando um transformador falha, os custos para uma concessionária são muito altos visto que este problema envolve a substituição do equipamento, limpeza e tentativa de reparação, e até a compra de energia de outras empresas. Isso causa grande impacto no consumidor, que pode passar a pagar mais caro pela energia consumida (FILHO, 2012).

Geralmente, emprega-se a manutenção preventiva como medida de preservação da vida útil de um transformador, porém, sabe-se que este tipo de manutenção não é a ideal porque não prevê a condição do equipamento com exatidão. A manutenção preventiva obedece a um plano previamente elaborado e possui um cronograma pré-estabelecido, porém, só é possível conhecer a condição real do transformador no momento dos ensaios e verificações programadas.

O estado-da-arte apresenta a manutenção preditiva como um ramo promissor na área de engenharia de manutenção, pois a condição do equipamento é monitorada constantemente neste tipo de manutenção, dando maior confiabilidade para a detecção de faltas incipientes, sendo neste caso, necessário empregar técnicas e ferramentas eficientes para realizar a manutenção.

Segundo a ANEEL o relatório de análise de desempenho da operação do sistema de transmissão brasileiro apontou que em 2017 ocorreram 225 desligamentos com devolução maior que 1 minuto, sendo 120 em instalações operadas de forma remota e 105 de forma local, ocasionados por consequência de falha/defeito e explosão ocorrida em transformadores de três enrolamentos. O relatório aponta que o tempo de recomposição em instalações operadas remotamente é 211 vezes maior que nas operadas localmente e a causa mais crítica é de indeterminação no tipo de defeito ocorrido nas instalações operadas de forma remota. O relatório conclui sugerindo que as concessionárias avaliem os desligamentos a fim de determinar a causa das falhas e repasse as informações detalhadas para que seja elaborado um relatório mais preciso. Isto indica que se faz necessário o incentivo do desenvolvimento de

ferramentas inteligentes de monitoramento, como a desenvolvida por (MORAIS, 2008) realizada em parceria com a Eletronorte, o Núcleo de Energia, Sistemas e Comunicação (NESC-UFPA) e o Laboratório de Química Analítica e Ambiental (LAQUANAM-UFPA). O projeto consistiu em monitorar o óleo isolante na chave comutadora de um transformador de 10 kVA com o objetivo de obter resultados mais abrangentes e confiáveis em relação às suas condições de operação. No sistema proposto foram instalados sensores no transformador para a realização da coleta de dados das seguintes propriedades físico-químicas do óleo isolante: acidez, oxigênio dissolvido e resistividade. O protótipo do sistema de monitoramento foi instalado no LAQUANAM em face da impossibilidade de se realizar esses ensaios em campo. O sistema consiste no uso de uma bomba centrífuga para retirada e circulação do óleo da chave comutadora do transformador para escoar em uma tubulação onde estão instalados três sensores: de acidez (pH), de condutividade ($\Omega.m$) e de oxigênio (O_2). Foi desenvolvido um *software* em Matlab para tratar e analisar os dados extraídos dos sensores, portanto, foram utilizadas três variáveis de entrada: concentração de oxigênio, nível de acidez e de resistividade; e uma variável de saída, a condição do óleo, que possui dois estados, “satisfatório” e “não-satisfatório”. O sistema de diagnóstico foi desenvolvido com base em lógica *fuzzy*. Foram realizadas três simulações com este sistema difuso, em todas o sistema indicou a chance de 67% do óleo estar em condição satisfatória e 33% em condição não-satisfatória. Estes resultados podem contribuir para o setor elétrico em vários aspectos, como para a redução do número de intervenções emergenciais de manutenção, estender a vida útil do transformador, reduzir custos de manutenção, entre outros.

Diversos tipos de análises podem servir de base para o desenvolvimento de ferramentas de manutenção preditiva em transformadores, dentre elas: análise físico-química do óleo isolante, análise cromatográfica de gases dissolvidos em óleo, análises termográficas e análise de vibração. Além disso, podem ser utilizados os dados obtidos em ensaios de isolamento, relação de transformação, resistência ôhmica dos enrolamentos, checagem do nível de óleo, limpeza, estado da sílica gel, estado do aterramento, pontos de oxidação, etc.

Dentre as análises mencionadas, a análise cromatográfica de gases dissolvidos em óleo é utilizada por diversos métodos, como os mencionados em(CHENG; YU, 2018): método IEC 60599, método de Rogers, método de Dörnenburg, método de diagrama triangular de Duval, método de diagrama com dominância de gás, hidrogênio-acetileno-etileno (HAE)–método baseado em diagrama triangular, método de descarga térmica - TD.

Neste trabalho será dado enfoque nos transformadores de potência que possuem como meio isolante o óleo, portanto, para um bom entendimento da proposta, a seguir será descrita a

composição dos transformadores e a propriedade do óleo isolante, assim como a relação entre a deterioração deste isolante e a vida útil do equipamento. Nas seções seguintes serão também descritos os principais métodos de análise cromatográfica. Por fim, será abordado o método IEC 60599, escolhido como método para comparação de resultados com os métodos de diagnóstico propostos neste trabalho.

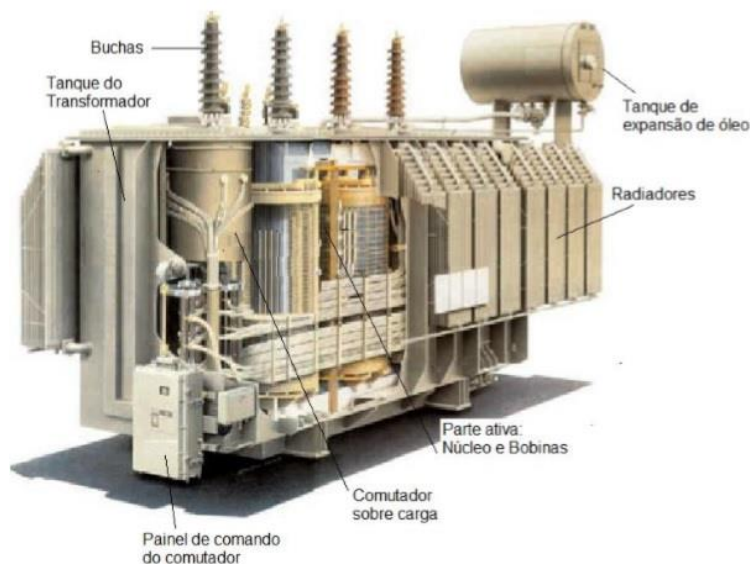
3.2. Componentes de um Transformador de Potência

Um transformador de potência é constituído por diversas partes como apresentado na Figura 6, sendo elas:

- Parte Ativa: consiste nos enrolamentos das bobinas de alta e baixa tensão, núcleo (que pode ser envolvido ou envolvente), sistemas de isolamento e conexões;
- Sistema de Preservação: é construído para aumentar a vida útil do equipamento evitando a contaminação do óleo isolante. Consiste em um tanque de expansão, tubulação para acesso do óleo ao tanque principal e bolsa ou membrana de borracha;
- Sistema de comutação: consiste em um conjunto eletromecânico, com a finalidade de alterar as relações de tensão no transformador de potência, em carga: executa-se a comutação com o equipamento em serviço; ou sem tensão, quando é necessário realizar o desligamento para alterar o *tap* de comutação;
- Enrolamento de Alta e Baixa Tensão: é o material condutor disposto na forma de espiras por onde a corrente elétrica flui;
- Núcleo: feito de material ferromagnético, é responsável por transmitir o fluxo magnético entre os lados primário e secundário;
- Buchas de Alta e Baixa Tensão: É o ponto de conexão das bobinas de alta e baixa tensão ao meio externo, são construídas com um material isolante, geralmente cerâmica ou polímero sintético;
- Tanque: é utilizado para abrigar a parte ativa e o óleo isolante;
- Radiadores: fazem parte do sistema de refrigeração e são responsáveis por realizar a dispersão de calor;
- Relé de pressão súbita: conhecido também como relé de Buchholz, é um dispositivo de proteção própria contra falta de óleo, acumulação de gases e falhas dielétricas catastróficas dentro do equipamento;

- Comutador de Taps: realiza a troca da relação de transformação do equipamento. Este equipamento opera de forma eletromecânica onde a operação é realizada através do acionamento remoto ou local;
- Tanque de expansão do óleo: permite a expansão do volume de óleo do transformador por conta das variações de temperatura a que o equipamento é submetido.
- Óleo isolante: É utilizado para realizar o isolamento e refrigeração do transformador. É encontrado na forma de líquido mineral ou vegetal em quantidade suficiente para propiciar o isolamento das bobinas e a transferência de calor para a superfície externa do tanque.

Figura 6 - Vista em corte lateral de um transformador de potência.



Fonte: Manual Siemens, 2008.

3.3. Propriedades do Óleo Isolante

3.3.1. Óleo Mineral Isolante

O óleo mineral isolante (OMI) é constituído por uma mistura de hidrocarbonetos (saturados e aromáticos) em maior proporção e de não-hidrocarbonetos (compostos de nitrogênio, de enxofre e de oxigênio). Possui cor amarela e é transparente. Pode ser de origem naftênica, mais usual, ou parafínica, tipo mais comum existente nas jazidas do Brasil.

A classificação do óleo mineral isolante é realizada segundo a norma internacional ASTM D 2140, dependendo da viscosidade, densidade e índice de refração (STOCCO, 2009). A tabela 1 apresenta as características que diferenciam as duas principais classes.

Tabela 1-Características dos óleos parafínicos e naftênicos.

| CARACTERÍSTICAS | PARAFÍNICOS | NAFTÊNICOS |
|-------------------------------|--------------------|-------------------|
| Ponto de fluidez | Alto | Baixo |
| Índice de Viscosidade | Alto | Baixo |
| Resistência à Oxidação | Grande | Pequena |
| Oleosidade | Pequena | Grande |
| Resíduo de Carbono | Grande | Pequena |
| Emulsibilidade | Pequena | Grande |

Fonte: MOURA; CARRETEIRO, 1987

Largamente utilizado, o óleo mineral é um produto derivado do petróleo e extraído através de um processo de fracionamento por destilação. Este processo resulta em vários subprodutos que são separados em frações de diversas faixas de ponto de ebulição, isto é, sob o efeito da temperatura as frações mais leves evaporam. A separação por destilação se baseia na diferença de volatilidade dos componentes e é promovida através de sucessivas vaporizações e condensações.

Após o processo de fracionamento e separação, realiza-se o refino e purificação do óleo isolante, a fim de minimizar a presença de hidrocarbonetos não-saturados.

Sua deterioração ocorre principalmente pela ocorrência de arcos e descargas elétricas que podem provocar a precipitação de partículas de carvão, a formação de ácidos e a liberação de gases. A água e o calor também comprometem a isolação dos transformadores. A deterioração da isolação sólida (papel isolante) e líquida dá origem à água e outras substâncias que podem servir de catalisadores, que juntamente com o ferro do tanque, do núcleo, dos parafusos e suportes, o cobre do enrolamento e a umidade, aceleram o processo, o que causa perda da propriedade dielétrica do óleo.

3.3.2. Óleo Vegetal Isolante

O óleo vegetal isolante (OVI) é extraído de grãos como: canola, soja, milho, entre outros, que passam por processos de degomagem, neutralização, branqueamento e desodorização (WILHELM; TULIO; UHREN, 2009). Além de atender às características dielétricas e refrigerantes, apresenta algumas vantagens em relação ao óleo mineral isolante, principalmente por ser considerado fluido de segurança (apresentam pontos de combustão acima de 300°C), biodegradável e proveniente de fonte renovável.

Os produtos de oxidação de um óleo vegetal não são nocivos ao papel isolante, ao contrário, evitam a sua degradação precoce. E, caso os produtos de oxidação do óleo vegetal polimerizarem, este processo não gera borra ou depósitos sobre os materiais internos do transformador. Em contrapartida, o óleo mineral isolante está sujeito a passar por um processo de oxidação que culmina no surgimento de ácidos que, ao atingir uma determinada concentração, podem agredir o papel isolante presente na parte ativa dos transformadores. Tais compostos podem, ainda, polimerizar e formar uma borra, que dificulta a transferência de calor para o meio ambiente se estiverem presentes na parte ativa ou nos radiadores.

A rigidez dielétrica do OMI e do OVI diminui à medida que a saturação relativa de água, causada pela presença de umidade no interior do tanque, atinge aproximadamente 50%. Mas, o OVI por ser um fluido polar possui maior afinidade com água do que o OMI, já que este é um fluido apolar. Portanto, o ponto de saturação do OVI é 20 vezes maior do que do OMI, isto é, a quantidade de água necessária para saturar o OVI é 20 vezes maior.

A umidade no OMI atua como agente catalisador na decomposição da celulose presente no isolamento das bobinas, diminuindo a vida útil do transformador. No OVI ocorre o efeito contrário, a vida útil do papel isolante é estendida na presença de óleos vegetais por causa da polaridade do fluido (WILHELM; TULIO; UHREN, 2009).

3.4. Impacto da Perda de Isolação na Vida Útil do Transformador

Na isolação do transformador estão presentes, basicamente, materiais de tipo sólido e líquido. O primeiro é encontrado na forma de material celulósico, sendo que os principais são: papel e papelão kraft, papel manilha e *pressboard* (MORAIS, 2008). E o segundo está na forma de óleo isolante, e como já mencionado, existem dois tipos: mineral e vegetal.

A vida útil de um transformador depende diretamente da condição destes materiais. Para que ela seja estendida é necessário evitar que tanto o papel isolante seja degradado por água e composto ácidos, quanto o óleo isolante seja deteriorado pela ação do calor, descargas elétricas e, também, da água presente no interior do equipamento. Na ocorrência do envelhecimento ou da contaminação dos mesmos, a segurança do equipamento é comprometida, podendo resultar em faltas incipientes que reduzem sua vida útil. Logo, medidas de prevenção devem ser adotadas para evitar que o isolamento seja comprometido, porém, mesmo com esta prevenção a deterioração ocorrida pela presença de umidade, de calor e de campos elétricos no interior do equipamento é praticamente inevitável. Assim, surge a necessidade de realizar inspeções de forma contínua através de métodos de diagnóstico de faltas incipientes.

Os principais tipos de falta em transformadores são de origem elétrica e térmica (PANTOJA, 2014). As principais falhas elétricas são:

- Descargas parciais de baixa intensidade: aparecem devido à baixa eficiência de materiais isolantes ou à existência de gases na isolação. Não é uma falha que gera sinais visíveis de deterioração;
- Descargas parciais de alta intensidade: as causas são as mesmas que na de baixa intensidade, porém, apresentam sinais visíveis de deterioração;
- Descargas de baixa energia: ocorrem com o mau contato de conexões, sendo caracterizadas pela presença de faíscas em ruptura de condutores ou descargas descontínuas em contatos de comutadores devido às rugosidades nas superfícies dos contatos gerados pelo continuado processo de comutação sob carga;
- Descargas de alta energia: ocorrem devido aos curtos-circuitos internos dos enrolamentos em conexões ou descargas entre condutores não-isolados.

As faltas térmicas e suas respectivas causas, são:

- Sobreaquecimento de 150°C até no máximo 300°C: surge devido ao sobreaquecimento dos enrolamentos ou pontos quentes no núcleo, ocasionado por perdas excessivas na magnetização, refrigeração insuficiente ou sobrecargas;
- Sobreaquecimento local de 300°C até no máximo 1000°C: ocorre devido às correntes de Foucault no núcleo, no tanque, na armação ou ainda devido ao mal contato em seletores de alta tensão, gerado pelo processo contínuo de comutação sob carga, culminando na carbonização do óleo isolante;
- Sobreaquecimento local acima de 1000°C: ocasionado por curto-circuito entre as lâminas do núcleo, com pontos de fusão ou condutores fundidos e pontos metálicos com início de fusão.

Em faltas ocasionadas por estresse elétrico, quando ocorre um arco elétrico, os principais gases produzidos são C_2H_2 e H_2 , e uma quantidade considerável de C_2H_4 e CH_4 . Como o desenvolvimento deste tipo de falta ocorre rapidamente, os gases são dissolvidos no óleo do transformador e então se reúnem no relé de gás. Quando ocorrem descargas de baixa energia, tais como ocorrência de arcos intermitentes de baixa energia, há, principalmente, uma produção de H_2 , em quantidade menor, sendo também produzido CH_4 . Desta forma, o componente e o conteúdo dos gases dissolvidos em óleo geralmente estão altamente relacionados com a localização da falta, a velocidade do fluxo de óleo e a duração da falta.

Quanto às falhas de descarga parcial, elas são um fenômeno de colapso local e repetitivo que ocorre no gás e os pontos marcados no papel isolante, sendo estruturados pela fraqueza do isolamento e pela concentração do campo elétrico. Quando ocorre uma descarga parcial, o conteúdo característico do componente de gás é determinado pela densidade de energia da descarga elétrica (CHENG; YU, 2018).

Em faltas onde ocorrem sobreaquecimento, causados tanto por efeitos elétricos quanto térmicos, o óleo do transformador e o material orgânico de isolamento vão envelhecendo gradualmente e se decompondo, produzindo uma pequena quantidade de hidrocarbonetos de baixo peso molecular e outros gases, tal como CO_2 e CO . Quando o estresse térmico não afeta a parte sólida da isolação, os gases produzidos são compostos sobretudo de hidrocarbonetos com baixo peso molecular, e entre eles, estão presentes os gases característicos CH_4 e C_2H_4 correspondendo a mais de 80% do total dos hidrocarbonetos. Geralmente, a quantidade de C_2H_2 não excede a 2% do total de hidrocarbonetos quando o aquecimento está abaixo de 500°C . Aquecimentos acima de 800°C produzem um pequeno montante de C_2H_2 , mas tal quantidade não ultrapassa o equivalente a 6% desses produtos. Quando o aquecimento afeta a parte sólida do isolamento, além do C_2H_2 , há a produção de mais gases CO_2 e CO , sendo que esta produção tem relação direta com o aumento da temperatura.

3.5. Métodos de Análise dos Gases Dissolvidos em Óleo

A partir dos indícios de que existe uma relação direta entre as concentrações de gases dissolvidos no óleo isolante e as faltas incipientes, foram desenvolvidos métodos para diagnosticar o tipo de falta que transformadores podem apresentar. Desta forma, podem-se tomar medidas para prolongar a vida útil do equipamento. Porém, ainda é muito difícil obter um consenso dos critérios que determinam a classificação destas faltas, ou seja, ainda é muito difícil aplicar regras que estabeleçam a relação entre os limites das quantidades dos gases dissolvidos e os tipos de faltas ocorridas. O principal obstáculo no desenvolvimento destas regras é a falta de correlação positiva entre os gases determinantes na identificação das falhas encontradas nos transformadores.

A análise dos gases dissolvidos em óleo não é considerada uma ciência, mas sim uma arte sujeita a variabilidade, visto que transformadores de diferentes tamanhos, estruturas e fabricantes apresentam características diferentes e, como consequência, é difícil obter um consenso quanto a valores limites dos gases e seus significados (MIRANDA et al., 2005). A presença e a concentração dos gases dependem de algumas características dos equipamentos

analisados, por exemplo, tipo, localização, grau de saturação dos gases dissolvidos em óleo, tipos dos materiais em contato com o óleo, entre outros (MORAIS, 2008).

Esta análise ocorre através do uso de técnicas como, a cromatografia gasosa (BAKAR; ABU-SIADA; ISLAM, 2014), o monitoramento online do hidrogênio (DE PABLO et al., 2011) e espectroscopia fotoacústica (STADLER, 2012). Portanto, a partir destas técnicas analíticas, pode-se aplicar métodos de interpretação dos dados obtidos para o diagnóstico do tipo de falta ocorrida no transformador.

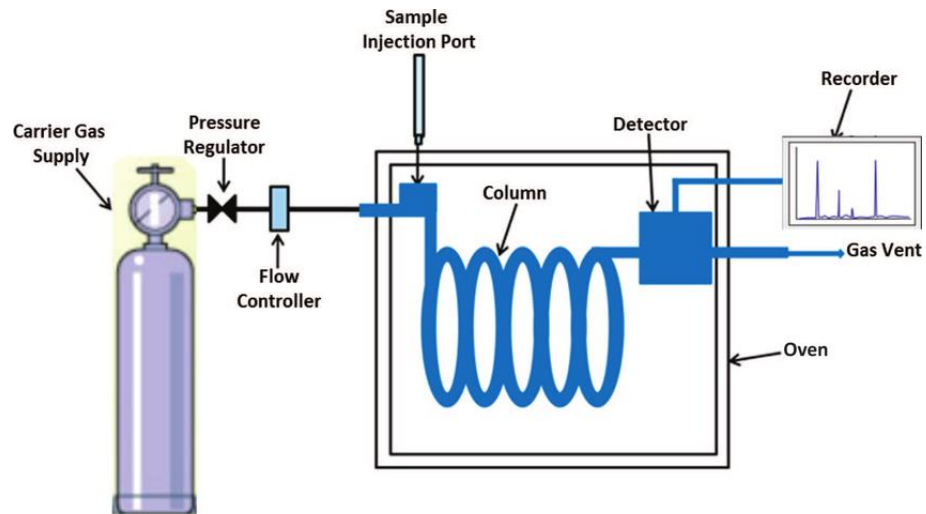
Dentre as três técnicas mencionadas, a análise por cromatografia gasosa é considerada a melhor, por ser a mais robusta e confiável. Já é utilizada a mais de 60 anos por diversos laboratórios ao redor do mundo e se tornou mais popular após a publicação de manuais relevantes por parte do IEEE, IEC e ASTM (BAKAR; ABU-SIADA; ISLAM, 2014). É utilizada para analisar tanto os gases combustíveis (hidrogênio, monóxido de carbono, metano, etano, etileno e acetileno), quanto os não-combustíveis (dióxido de carbono, nitrogênio e oxigênio). Pode, também, ser utilizada para analisar os gases presentes na região dentro do tanque do transformador, acima do óleo isolante.

O processo para a análise dos gases inicia-se através da coleta do óleo isolante para a extração do gás dissolvido. Podem ser aplicadas técnicas de extração como a extração a vácuo, coluna de remoção, ou amostragem de espaço livre. A extração a vácuo é adequada para a extração de uma fração de cada um dos gases, enquanto a coluna de remoção extrairá quase todo gás. Já a técnica de amostragem de espaço livre irá extrair uma fração de cada gás no espaço livre presente no transformador. A configuração básica de uma cromatografia gasosa está representada na Figura 7.

Após a coleta de amostra do óleo isolante, esta é injetada (*Sample Injection Port*) para que os componentes dos gases sejam separados na coluna cromatográfica. Como pode ser observado na Figura 7, a coluna possui uma estrutura em forma de bobina e no interior dela está presente uma fase estacionária, que dependendo do tipo de substância, podem estar no estado sólido ou líquido. Após a injeção da amostra, inicia-se o processo de separação dos componentes presentes na mesma. A amostra é vaporizada e transferida para a coluna. Argônio, hélio, nitrogênio e hidrogênio são normalmente usados como gases transportadores, denominados gases de arraste, para transferir a amostra do injetor, através da coluna, para o detector. Portanto, a separação ocorre através da interação entre as substâncias e a fase estacionária. Existem três possibilidades de mudança de estado da substância presente na amostra vaporizada. Primeiro, ela pode permanecer no estado gasoso, isto é, continuar fluindo dentro da coluna. Segundo, pode mudar para o estado líquido, dissolvendo na fase estacionária

ou, por último, condensar nesta fase. Simultaneamente, os componentes gasosos retidos na fase estacionária viajam com os componentes transportados pelo gás de arraste, porém, seu progresso através da coluna será muito mais lento do que se continuasse a viajar com o gás de arraste. Este processo ocorre centenas de vezes no interior da coluna. O tempo de passagem pela coluna varia entre os componentes e depende da temperatura da coluna e da composição química da fase estacionária.

Figura 7- Esquema básico de um sistema de cromatografia.



Fonte: Bakaret al., 2014.

Os componentes gasosos são transportados para fora da coluna um por um, dependendo do tempo de retenção, e interagem com os detectores, gerando sinais elétricos. Esses sinais são registrados por um sistema de coleta de dados e plotados, produzindo um cromatograma. Os gases podem ser identificados usando o tempo de retenção, e a concentração de um determinado gás é determinada pela magnitude do sinal elétrico associado. Vários tipos de detectores estão disponíveis, por exemplo, por condutividade térmica, ionização em chama, nitrogênio-fósforo, fotometria em chama, captura de elétrons, emissão atômica e condutividade eletrolítica (BAKAR; ABU-SIADA; ISLAM, 2014).

Como mencionado, existem diversos métodos de análise gascromatográfica que foram desenvolvidos ao longo dos anos. Este trabalho discorrerá sobre os principais métodos, que estão relatados nas seções abaixo.

3.5.1. Método Clássico de Rogers

O método de Rogers (ROGERS, 1978) foi desenvolvido com base na avaliação termodinâmica realizada por Halstead (HALSTEAD, 1973) que concluiu que a taxa de evolução de qualquer hidrocarboneto gasoso em particular variava com a temperatura, e que a uma determinada temperatura haveria uma taxa máxima de evolução desse gás e que cada gás atingiria sua taxa máxima a uma temperatura diferente. Esta consideração levou à escolha de quatro relações de gases para o diagnóstico de faltas incipientes, são elas: CH_4/H_2 , $\text{C}_2\text{H}_6/\text{CH}_4$, $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$, $\text{C}_2\text{H}_2/\text{C}_2\text{H}_4$.

Em suma, o método de Rogers utiliza as relações de concentração de gases e limites de variação destas relações para emitir diagnóstico de faltas em transformadores, baseado no uso de códigos que são gerados a partir das regras estabelecidas por estes limites de variação. Passou por adaptações ao longo dos anos para se adequar a correlação encontrada entre a concentração de gases e os tipos de falta ocorridas. Na Tabela 1 abaixo, são apresentadas as faixas de variação das relações e seus respectivos códigos. Na tabela 2 é apresentada a combinação dos códigos para determinar o diagnóstico final.

Tabela 2- Tabela de definição do método Rogers.

| Relação de Gases | Faixa de Variação | Código |
|---|--------------------------|---------------|
| CH_4/H_2 | ≤ 0.1 | 5 |
| $\text{C}_2\text{H}_6/\text{CH}_4$ | $> 0.1, < 1$ | 0 |
| $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$ | $\geq 1, < 3$ | 1 |
| $\text{C}_2\text{H}_2/\text{C}_2\text{H}_4$ | ≥ 3 | 2 |

Tabela 3 - Método de Rogers em função do código.

| CH ₄ /H ₂ | C ₂ H ₆ /CH ₄ | C ₂ H ₄ /C ₂ H ₆ | C ₂ H ₂ /C ₂ H ₄ | Diagnóstico |
|---------------------------------|--|--|--|--|
| 0 | 0 | 0 | 0 | Deterioração Normal |
| 5 | 0 | 0 | 0 | Descargas Parciais |
| ½ | 0 | 0 | 0 | Sobreaquecimento <150°C |
| ½ | 1 | 0 | 0 | Sobreaquecimento 200°C –200°C |
| 0 | 0 | 1 | 0 | Sobreaquecimento de condutores |
| 1 | 0 | 1 | 0 | Correntes de circulação nos enrolamentos |
| 1 | 0 | 2 | 0 | Correntes de circulação no núcleo e tanque, sobreaquecimento em conexões. |
| 0 | 0 | 0 | 1 | Descarga contínua |
| 0 | 0 | ½ | ½ | Arco em alta energia |
| 0 | 0 | 2 | 2 | Descarga contínua com baixa potência |
| 5 | 0 | 0 | ½ | Descarga parcial envolvendo o papel |

3.5.2. Método de Dörnenburg

Em 1977, Dörnenburg classificou as faltas em dois tipos: elétricas e térmicas. O método utiliza seis concentrações de gases organizadas em quatro razões, das quais duas são principais e outras duas auxiliares. Uma característica peculiar deste método é que ele não estabelece

critérios para caracterizar se o isolamento do transformador está em condições normais (DORNENBURG; GERBER, 1967).

A Tabela 4 apresenta os dados para análise em função das relações entre os gases e os tipos de faltas incipientes, já na Tabela 5 são mostrados os valores de concentração individual dos gases para validação de aplicação do método.

Tabela 4 - Critérios do Método de Dörnenburg.

| Relações entre concentrações de gases | Relações principais | | Relações auxiliares | |
|---------------------------------------|---------------------|----------------------------------|---|--|
| | Tipo de Falha | CH ₄ / H ₂ | C ₂ H ₂ / C ₂ H ₄ | C ₂ H ₆ /C ₂ H ₂ |
| Sobreaquecimento | > 1 | < 0.75 | > 0.4 | < 0.3 |
| Descarga de baixa Intensidade | < 0.1 | Não significativo | > 0.4 | < 0.3 |
| Outros tipos de descargas | >0.1 e < 1 | > 0.75 | < 0.4 | >0.3 |

Fonte: IEEE Std. C57.104 (2008, p. 16) (Adaptado).

Tabela 5 - Concentração de gases característicos para validação do método.

| Tipo de Gás | H ₂ | CH ₄ | C ₂ H ₆ | C ₂ H ₄ | C ₂ H ₂ |
|------------------------|----------------|-----------------|-------------------------------|-------------------------------|-------------------------------|
| Concentração ppm (v/v) | 200 | 50 | 15 | 60 | 15 |

O método é válido se ao menos um dos gases em cada uma das relações da Tabela 4 excederem o valor limite da Tabela 5, caso contrário, as relações não são significantes e o transformador deve ser investigado por métodos alternativos, por exemplo, o uso do método do gás-chave, que será mencionado posteriormente.

3.5.3. Triângulo de Duval

Este método foi desenvolvido a partir do padrão IEEE 60599 e a base de dados de transformadores faltosos IEC TC 10. O método consiste em utilizar uma ferramenta gráfica para diagnosticar os tipos de faltas levando em conta apenas a concentração relativa dos gases acetileno, etano e metano. Pode-se observar na Figura 8 que a ferramenta tem o formato de triângulo e possui sete zonas de faltas, cobrindo três tipos de faltas de origem elétrica (descargas parciais, descargas de baixa e alta energia) e outras três de origem térmica (faltas térmicas

abaixo de 300°C, faltas entre 300°C e 700°C, e acima de 700°C), além de identificar um tipo de falta que mescla descargas elétricas com sobreaquecimento.

Neste método, são utilizados apenas três tipos de gases (CH_4 , C_2H_4 , e C_2H_2), pois, segundo (Duval e Dukarm, 2005), são suficientes para obter um diagnóstico confiável. As coordenadas são definidas pela relação percentual da concentração de cada gás em relação ao montante gerado e dissolvido em óleo:

$$\text{CH}_4(\%) = \frac{m}{m+e+a} \quad 8a$$

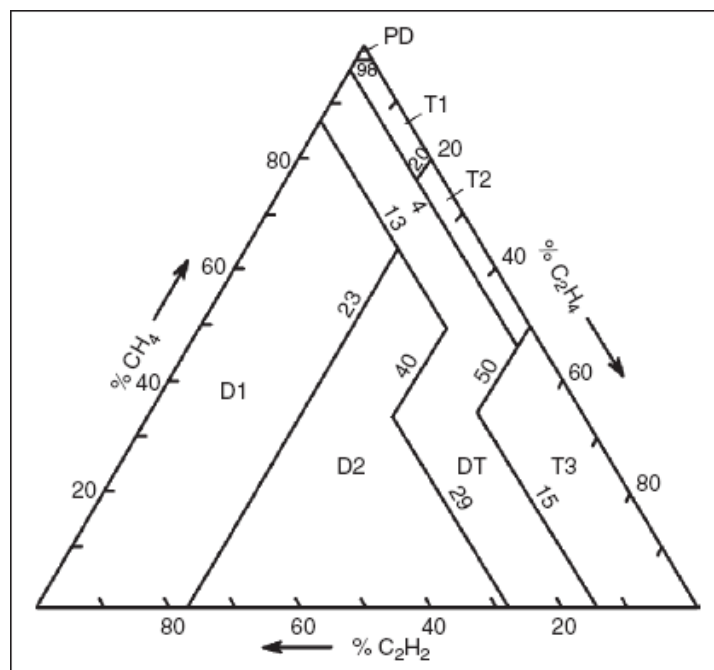
$$\text{C}_2\text{H}_4(\%) = \frac{e}{m+e+a} \quad 8b$$

$$\text{C}_2\text{H}_2(\%) = \frac{a}{m+e+a} \quad 8c$$

onde $m = \text{CH}_4$, $e = \text{C}_2\text{H}_4$, e $a = \text{C}_2\text{H}_2$ em ppm.

As faltas estão divididas em sete zonas, sendo elas: D1 - faltas elétricas de baixa energia, D2 - falta elétrica de alta energia, DT - falta térmica com a presença de descarga elétrica, T1 - falta térmica abaixo de 300°C, T2 - falta térmica entre 300°C e 700°C e T3 - falta térmica acima de 700°C. Pode-se observar que a relação de porcentagem entra as concentrações dos gases determina as coordenadas para a identificação do tipo de falta.

Figura 8 - Coordenadas e zonas de faltas no triângulo de Duval.



Fonte: Duval, 2002.

3.5.4. Método do Gás-Chave

O método se baseia na predominância de determinado gás em relação ao Total de Gases Dissolvidos – TGCD – no óleo isolante. O diagnóstico das faltas incipientes é atribuído pela concentração do gás predominante no momento que determinada falta ocorre sob diferentes condições de operação. O cálculo do TGCD é realizado considerando as concentrações dos seguintes gases: hidrogênio (H_2), metano (CH_4), etano (C_2H_6), etileno (C_2H_4), acetileno (C_2H_2) e monóxido de carbono (CO). Neste método as concentrações absolutas, medidas em ppm, e a taxa de geração (ppm/dia) dos gases são utilizadas para determinar o tipo e a intensidade de determinadas falhas.

A Tabela 6 apresenta os critérios de classificação empregados pelo método. Vale ressaltar que as porcentagens apresentadas foram obtidas a partir de diversos experimentos práticos, porém, este método não é aceito como uma ferramenta de diagnóstico confiável pois leva a muitos resultados incorretos (DUVAL; DUKARM, 2005).

Tabela 6 - Critério de Diagnóstico do Método Gás-Chave.

| Falta | Gás-Chave | Critério |
|------------------------------|------------------------------|--|
| Arco | Acetileno (C_2H_2) | Grande quantidade de H_2 e C_2H_2 são produzidos, com uma quantidade menor de CH_4 e C_2H_4 . CO_2 e CO podem ser formados se a falta atingir a celulose. |
| Corona | Hidrogênio (H_2) | É produzido H_2 e CH_4 , com pequena quantidade C_2H_6 e C_2H_4 . CO e CO_2 podem ser comparados se as cargas atingirem a celulose. |
| Sobreaquecimento do óleo | Etileno (C_2H_4) | É produzido C_2H_4 e CH_4 , junto com uma proporção menor de H_2 e C_2H_6 . Vestígios de acetileno podem ser formados se a falha for grave ou envolver contatos elétricos. |
| Sobreaquecimento da celulose | Monóxido de Carbono (CO) | É produzido uma grande quantidade de CO_2 e CO . Gases a base de hidrocarbonetos, tal como CH_4 e C_2H_4 serão formados se a falta envolver estrutura impregnada de óleo. |

3.5.5. Método IEC 60599

O método publicado na norma 60599 da *Internacional Electrotechnical Comission* – IEC consiste na identificação de faltas incipientes em transformadores utilizando os dados provenientes da cromatografia de 5 tipos de gases. Este método utiliza as mesmas razões propostas por Rogers, entretanto os valores que delimitam os tipos de faltas são diferentes. Na tabela são apresentados os valores que estabelecem o critério de classificação.

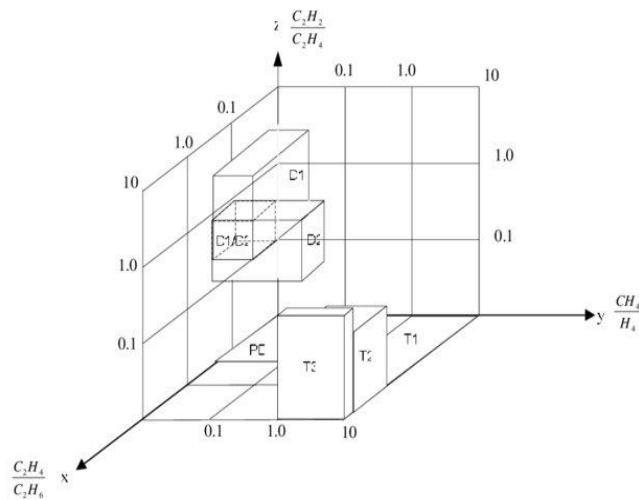
Como todos os outros métodos, o IEC 60599 também sofre do problema de “*no decision*” onde tem-se casos onde o método não é capaz de chegar a um diagnóstico do tipo de falta. Do código apresentado na tabela 7 podemos observar que o método apresenta problema de sobreposição entre as faltas do tipo D1 e D2, resultando em uma atribuição de faltas ambígua para determinados casos. Na Figura 9 o problema de “*no decision*” e sobreposição de falta D1 e D2 pode ser visualizado.

Tabela 7 - Critério do método IEC.

| Caso | Falta Característica | <u>C₂H₂</u> | <u>CH₄</u> | <u>C₂H₄</u> |
|------|-----------------------------|-----------------------------------|-----------------------|-----------------------------------|
| | | C ₂ H ₄ | H ₂ | C ₂ H ₆ |
| PD | Descarga Parcial | NS | <0.1 | <0.2 |
| D1 | Descarga de Baixa Energia | >1 | 0.1-0.5 | >1 |
| D2 | Descarga de Alta Energia | 0.6-2.5 | 0.1-1 | >2 |
| T1 | Falta Térmica < 300°C | NS | >1 mas NS | <1 |
| T2 | Falta Térmica 300°C – 700°C | <0.1 | >1 | 1-4 |
| T3 | Falta Térmica > 700°C | <0.2 | >1 | >4 |

NS – Valor não-significativo

Figura 9 - Representação gráfica do IEC 60599.



Fonte: BRAGA; CASTRO, 2008.

A tabela IEC foi construída a partir de informações de um banco de dados faltosos, o IEC TC 10 *database*, publicado em (DUVAL; DEPABLO, 2001).

3.6 Desvantagens do Uso de Métodos Tradicionais

O uso de métodos tradicionais apresenta certas desvantagens, tais como:

- a) É difícil obter um consenso quanto aos limites dos gases e seus significados, já que existem inúmeros fabricantes e modelos de transformadores;
- b) A natureza heurística e empírica dos métodos baseados em razões tem levado à diferentes interpretações dos dados;
- c) Às vezes é impossível de se obter o diagnóstico devido a ambiguidade ou até mesmo a falta de conclusão por parte destes métodos (problema de *no decision*). Devido à imprecisão de um método, é necessário utilizar vários simultaneamente para tentar chegar à uma conclusão, onde, muitas vezes é exigido o conhecimento de especialistas;
- d) O método IEC apresenta sobreposição entre os limites que definem os critérios de identificação das falhas, portanto, geram ambiguidade na conclusão do diagnóstico;
- e) Os métodos não possuem meios para mostrar ao profissional a evolução de uma falha incipiente após ser detectada.

Apesar de estes métodos serem utilizados por diversas empresas e laboratórios ao redor do mundo, ainda busca-se aprimorá-los e até mesmo utilizar novas abordagens para solucionar o problema da classificação de faltas incipientes em transformadores, como por exemplo, com o auxílio de Inteligência Computacional, onde tem-se obtido resultados promissores (CHENG;

YU, 2018). Isto se deve porque estes métodos tradicionais possuem certas limitações (DA SILVA, 2013)

4. APLICAÇÃO DE REDES NEURAIIS PROFUNDAS AO DIAGNÓSTICO DE TRANSFORMADORES

4.1 Introdução

Este capítulo tem como objetivo apresentar os sistemas propostos para o diagnóstico de faltas incipientes em transformadores baseado na análise dos gases dissolvidos em óleo.

Foram desenvolvidos dois modelos de sistema para classificação de faltas, sendo um modelo baseado na rede Autocodificadora Empilhada (*Stacked Autoencoder* – SAE) e o segundo modelo baseado na Rede Neural Convolutacional (*Convolutional Neural Network* – CNN). Para o desenvolvimento dos sistemas foram utilizados dados da base TC 10 de equipamentos faltosos, publicada em (DUVAL; DEPABLO, 2001).

Os resultados obtidos para os dois sistemas propostos foram comparados com o resultado de classificação obtido através do Método IEC 60599 que é um dos métodos mais utilizados pelas concessionárias de energia para interpretação dos gases dissolvidos em óleo. Uma comparação com resultados de sistemas de classificação de faltas já apresentados na literatura também será apresentada.

4.2 Base de dados de transformadores com faltas

A base de dados de equipamentos com faltas inspecionados em serviço usada na publicação IEC 60599 (DUVAL, 2002) e apresentada em (DUVAL; DEPABLO, 2001) foi utilizada para desenvolvimento dos sistemas propostos neste trabalho. Adicionalmente, uma base de dados derivada da literatura e uma base de dados obtida da CELPA (Centrais Elétricas do Pará- Brasil) também foi utilizada. A base é composta de 300 amostras onde 210 amostras foram usadas para treino dos sistemas e 90 foram utilizadas para teste dos sistemas. A Tabela 8 apresenta a composição da base de dados utilizada para desenvolvimento dos sistemas de diagnósticos propostos.

Segundo (MIRANDA; CASTRO; LIMA, 2012), é comum observar que os casos T2 e T3 (considerando a Tabela IEC 60599) estão agrupados em muitos estudos apresentados da literatura que utilizam a base de dados TC 10, e isto se deve ao fato do pequeno número de casos para estas faltas, sendo então inadequado para o desenvolvimento de sistema de diagnóstico automático. Na Tabela 8 tem-se o grupo T2 que corresponde as faltas com

temperatura $T > 300^{\circ}\text{C}$, que corresponde a união de T2 e T3 da Tabela 7. Portanto, ambos os sistemas propostos empregaram como saída as 5 classes representadas na Tabela 8 e como entradas as 3 concentrações de gases usadas pelo método IEC 60599 (acetileno/etileno, metano/hidrogênio, etileno/etano).

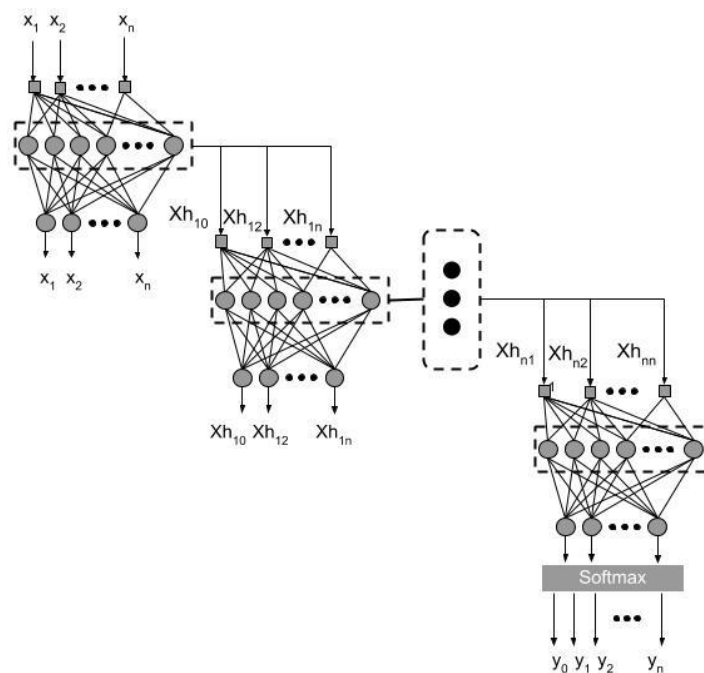
Tabela 8 - Divisão da base de dados aplicada nos sistemas de diagnóstico.

| Caso | Quantidade de amostras | |
|--------------|------------------------|-----------|
| | Treino | Teste |
| T1 | 54 | 23 |
| T2 | 48 | 22 |
| PD | 21 | 9 |
| D1 | 22 | 9 |
| D2 | 65 | 27 |
| Total | 210 | 90 |

4.3 Sistema de Diagnóstico de Falhas Via Autocodificadoras Empilhadas

A Figura 10 apresenta a estrutura proposta para classificação de falhas baseada nas redes Autocodificadoras Empilhadas.

Figura 10 - Arquitetura de Rede Neural Autocodificadora Empilhada.



Fonte: Autor

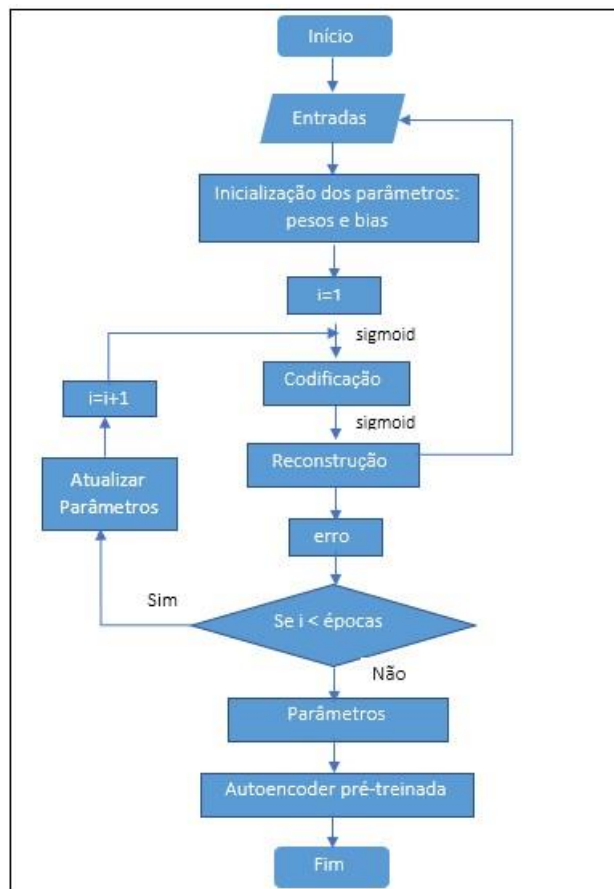
O sistema utiliza como entradas as 3 razões dos gases de acordo com o estabelecido pelo critério IEC 60599 e fornece na sua saída o tipo de falta, codificada de acordo como apresentado na Tabela 9.

Tabela 9 – Codificação do tipo de falta na saída da Auto codificadora Empilhada

| Tipo de falta | Valor codificado | Codificação da saída |
|---------------|------------------|----------------------|
| T1 | 1 | [1 0 0 0 0] |
| T2 | 2 | [0 1 0 0 0] |
| PD | 3 | [0 0 1 0 0] |
| D1 | 4 | [0 0 0 1 0] |
| D2 | 5 | [0 0 0 0 1] |

Este modelo possui duas fases de desenvolvimento. Na primeira fase as redes autocodificadoras devem ser treinadas para que características sejam extraídas a partir da camada intermediária de cada autocodificadora. A Figura 11 apresenta um fluxograma da fase de treinamento de uma única rede autocodificadora.

Figura 11 - Processo de pré-treinamento de uma rede autoencoder.

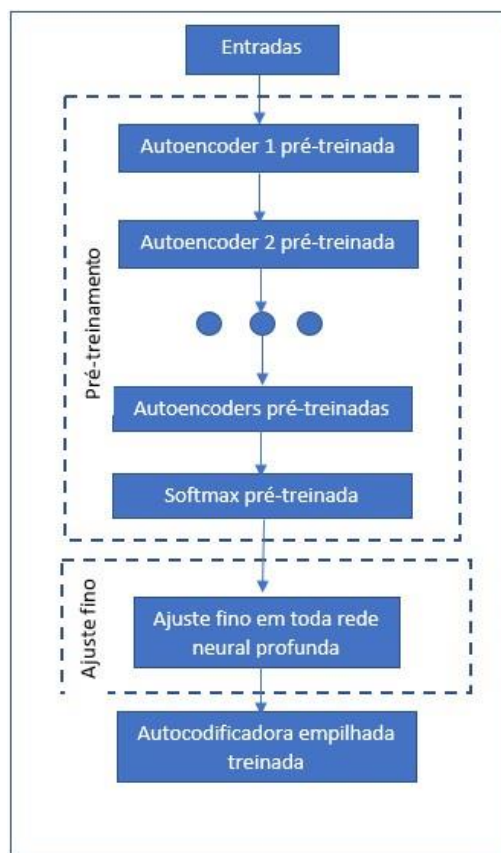


Fonte: adaptado de (QI et al., 2017).

O critério de parada utilizado para o treinamento de uma autocodificadora é o número máximo de épocas e os parâmetros atualizados em cada iteração correspondem aos pesos e *bias* presentes nas camadas da rede. As redes autocodificadoras são treinadas para reconstruir nas suas saídas os padrões apresentados nas suas entradas, logo o vetor de entrada e de saída devem ser os mesmos.

A Figura 12 apresenta o fluxograma de treinamento considerando as redes autocodificadoras empilhadas.

Figura 12 - Treinamento de uma rede stacked sparse autoencoder.



Fonte: Adaptado (QI et al., 2017) .

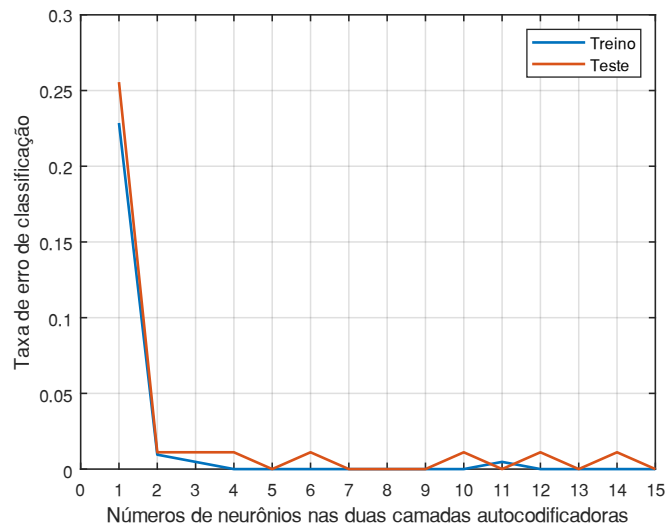
Para o desenvolvimento do sistema de diagnóstico proposto foram utilizadas 2 autocodificadoras empilhadas. Para o treinamento da primeira autocodificadora foi utilizado como padrões de entrada/saída as 3 razões dos gases. Após o treinamento desta rede, as características extraídas através da camada escondida desta autocodificadora foram apresentadas como padrões de entrada/saída para o treinamento de uma segunda autocodificadora.

Ao final do treino da segunda autocodificadora, as características extraídas da sua camada escondida são aplicadas como entrada na camada *Softmax*, onde o algoritmo de

aprendizado *Scaled Conjugate Gradient* - SCG, é utilizado para realização do treinamento supervisionado, tendo agora como saída da camada *Softmax* o tipo de falta do transformador codificado conforme Tabela 9.

Durante o desenvolvimento das duas redes autocodificadoras, para determinação da melhor topologia de cada rede, foram realizados para cada rede diversos treinamentos (15) variando-se o número de neurônios na camada escondida. A métrica de comparação do desempenho para definição das melhores topologias foi a taxa de erro de classificação considerando a estrutura geral classificadora. Nos primeiros treinamentos foram utilizadas topologias iguais em cada autocodificadora, ou seja, para ambas as redes foi modificado o número de neurônios em cada treinamento, iniciando em 1 neurônio até atingir o valor de 15 neurônios. A Figura 13 apresenta a curva de variação da taxa de erro de classificação em função da variação do número de neurônios na camada escondida das autocodificadoras.

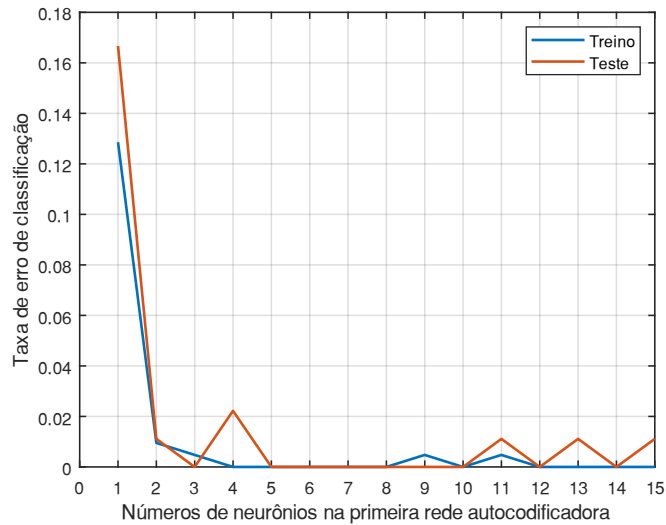
Figura 13 - Taxas de erro de classificação variando o número de neurônios da primeira e segunda autocodificadora.



A menor taxa de erro alcançada foi de 0%, ou seja, não houve erro de classificação para os dados de teste em alguns casos, como pode ser observado na Figura 13.

Na tentativa de se obter um melhor desempenho do sistema, considerando custo computacional, foram realizados mais 15 treinamentos, variando-se agora apenas o número de neurônios da camada escondida da primeira rede autocodificadora e mantendo-se fixo o número de neurônios da segunda autocodificadora em 5 neurônios. Os resultados obtidos da taxa de erro de classificação nesta fase são apresentados na Figura 14.

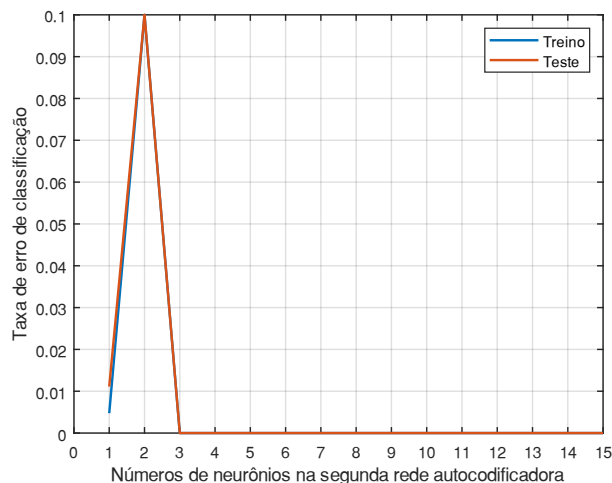
Figura 14 - Taxas de erro de classificação variando os neurônios na primeira autocodificadora.



A melhor topologia encontrada para a primeira autocodificadora foi com 3 neurônios. Tendo sido alcançado uma taxa de 0% de erro, ou seja, do total de padrões de teste não houve erro de classificação. Por fim, foram realizadas mais 15 simulações, mantendo-se fixo o valor de neurônios da primeira autocodificadora em 3, e variando-se o número de neurônios da segunda autocodificadora, na tentativa de se obter a melhor topologia da segunda autocodificadora com menor custo computacional. Os resultados encontrados podem ser observados na Figura 15.

A melhor topologia encontrada para a segunda autocodificadora foi com 3 neurônios, com uma taxa de erro de classificação de 0%.

Figura 15 - Taxas de erro de classificação encontradas variando os neurônios na segunda camada.



Portanto, após todos os treinamentos realizados, a melhor topologia encontrada para o sistema de classificação foi de 3 neurônios na camada escondida da primeira autocodificadora e 3 neurônios na camada escondida da segunda autocodificadora (Figura 16).

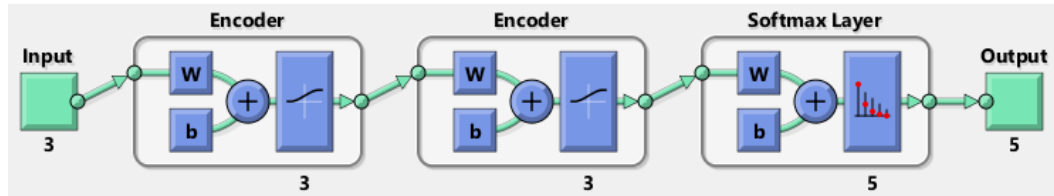


Figura 16 - Melhor topologia para o sistema de diagnóstico via Autocodificadora Empilhada.

A Figura 17 apresenta a matriz de confusão do sistema de diagnóstico de faltas desenvolvido com a autocodificadora representada na Figura 16, considerando apenas a base de dados de testes. É importante salientar que para a escolha da melhor topologia do sistema de classificação considerou-se apenas os melhores resultados para a base de teste, que é a base de dados que não é utilizada em nenhum momento para o treino das redes neurais, ou seja, a base de dados que geralmente se utiliza para verificação da capacidade de generalização do sistema.

Figura 17 - Matriz de confusão para base de dados de teste.

Matriz de Confusão

| | | | | | | | |
|------------------|---|---------------|--------------|--------------|--------------|--------------|--------------|
| Classes Preditas | 1 | 23 25.6% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 100% 0.0% |
| | 2 | 0 0.0% | 22 24.4% | 0 0.0% | 0 0.0% | 0 0.0% | 100% 0.0% |
| | 3 | 0 0.0% | 0 0.0% | 9 10.0% | 0 0.0% | 0 0.0% | 100% 0.0% |
| | 4 | 0 0.0% | 0 0.0% | 0 0.0% | 9 10.0% | 0 0.0% | 100% 0.0% |
| | 5 | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 27 30.0% | 100% 0.0% |
| | | | 100% 0.0% | 100% 0.0% | 100% 0.0% | 100% 0.0% | 100% 0.0% |
| | | 1 | 2 | 3 | 4 | 5 | |
| | | Classes Reais | | | | | |

A partir da matriz de confusão podemos verificar que o sistema de diagnóstico proposto a partir das autocodificadoras empilhadas apresentou um excelente desempenho, tendo alcançado 100 % de acurácia, considerando todos os padrões da base de teste.

4.4 Sistema de diagnóstico de Faltas via Rede Neural Convolutacional

Para o desenvolvimento do sistema de diagnóstico de faltas baseado em uma rede neural convolutacional foram utilizadas como entradas as 3 razões dos gases, assim como no modelo das redes autocodificadoras empilhadas. Foi necessário realizar um redimensionamento da matriz com dados de treino para quatro dimensões, isto é, uma matriz 4-D, assumindo a dimensão $1 \times 3 \times 1 \times 214$. Desta forma, a CNN pode interpretar os dados de entrada como uma matriz numérica 4-D, onde as primeiras três dimensões referem-se à altura, largura e canais e a última dimensão deve indexar as amostras individuais.

Foram realizados diversos treinamentos alterando-se a arquitetura da rede CNN. O valor da taxa de aprendizado permaneceu inalterado para todas as arquiteturas testadas e com o valor igual a 0.01. As arquiteturas testadas são apresentadas da Tabela 10 até a Tabela 15.

Tabela 10–Arquitetura 1.

| Nome da camada | Número de filtros | Tamanho do filtro | Stride | Saída da Camada |
|-------------------|-------------------|-------------------|--------|-----------------|
| Convolução | 4 | 1x1x1 | 1 | 1x3x4 |
| Pooling | 4 | 1x2 | 1 | 1x2x4 |
| Convolução | 8 | 1x1x4 | 1 | 1x2x8 |
| Pooling | 8 | 1x2 | 1 | 1x1x8 |

Tabela 11–Arquitetura 2.

| Nome da camada | Número de filtros | Tamanho do filtro | Stride | Saída da Camada |
|-------------------|-------------------|-------------------|--------|-----------------|
| Convolução | 8 | 1x1 | 1 | 1x3x8 |
| Pooling | 8 | 1x2 | 1 | 1x2x8 |
| Convolução | 16 | 1x1 | 1 | 1x2x16 |
| Pooling | 16 | 1x2 | 1 | 1x1x16 |

Tabela 12–Arquitetura 3.

| Nome da camada | Número de filtros | Tamanho do filtro | Stride | Saída da Camada |
|-------------------|-------------------|-------------------|--------|-----------------|
| Convolução | 4 | 1x1x1 | 1 | 1x3x4 |
| Pooling | 4 | 1x2 | 1 | 1x2x4 |
| Convolução | 8 | 1x1x4 | 1 | 1x2x8 |
| Pooling | 8 | 1x2 | 1 | 1x1x8 |

Tabela 13 - Arquitetura4.

| Nome da camada | Número de filtros | Tamanho do filtro | Stride | Saída da Camada |
|----------------|-------------------|-------------------|--------|-----------------|
| Convolução | 14 | 1x1x1 | 1 | 1x3x14 |
| Pooling | 14 | 1x2 | 1 | 1x2x14 |
| Convolução | 8 | 1x1x14 | 1 | 1x2x8 |
| Pooling | 8 | 1x2 | 1 | 1x1x8 |

Tabela 14–Arquitetura 5.

| Nome da camada | Número de filtros | Tamanho do filtro | Stride | Saída da Camada |
|----------------|-------------------|-------------------|--------|-----------------|
| Convolução | 8 | 2x2x1 | 1 | 1x3x8 |
| Pooling | 8 | 1x2 | 1 | 1x2x8 |
| Convolução | 16 | 1x1x8 | 1 | 1x2x16 |
| Pooling | 16 | 1x2 | 1 | 1x1x16 |

Tabela 15 - Arquitetura6.

| Nome da camada | Número de filtros | Tamanho do filtro | Stride | Saída da Camada |
|----------------|-------------------|-------------------|--------|-----------------|
| Convolução | 8 | 2x2 | 1 | 1x3x8 |
| Pooling | 8 | 1x1 | 1 | 1x2x8 |

A tabela 16 apresenta os resultados de acurácia obtidos para todas as arquiteturas treinadas considerando a base de dados de teste.

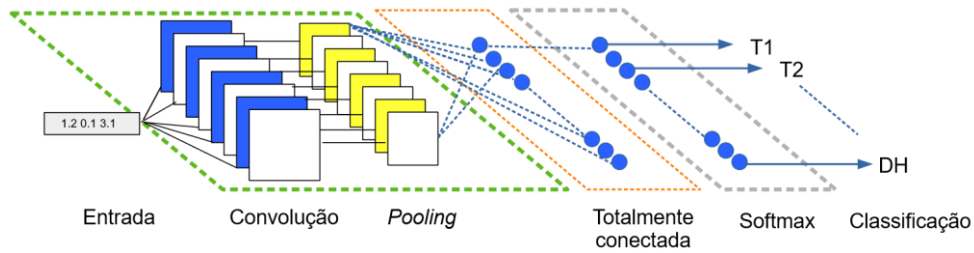
Tabela 16 - Desempenho das arquiteturas de rede convolucional.

| Arquitetura | Acurácia Teste |
|---------------|----------------|
| Arquitetura 1 | 0.8488 |
| Arquitetura 2 | 0.8953 |
| Arquitetura 3 | 0.8837 |
| Arquitetura 4 | 0.8721 |
| Arquitetura 5 | 0.9651 |
| Arquitetura 6 | 1 |

A arquitetura de rede neural convolucional com melhor desempenho corresponde à arquitetura 6 que possui 5 camadas: uma camada de convolução, seguida de uma camada de *pooling*, uma camada totalmente conectada, uma camada que aplica a função *softmax* e por fim, uma camada que realiza a classificação através da aplicação de entropia cruzada.

A Figura 18 apresenta a arquitetura da rede CNN com melhor desempenho e a Figura 19 apresenta a matriz de confusão obtida para esta rede considerando a base de dados de teste.

Figura 18 - Arquitetura da rede convolucional com maior desempenho.



O melhor resultado obtido foi alcançado através da aplicação da arquitetura mais simples, com apenas uma camada de convolução e uma de *pooling*. Portanto, pode-se inferir que o aumento das camadas da rede a torna muito complexa para a tarefa de classificação das faltas incipientes, tendo como consequência a diminuição de seu desempenho. A acurácia alcançada por este modelo nestas características foi de 100%.

Figura 19 - Matriz de confusão dos dados de teste da rede neural convolucional de melhor desempenho.

Matriz de Confusão

| | | | | | | | |
|------------------|----|---------------|--------------|--------------|--------------|--------------|--------------|
| Classes Preditas | T1 | 23 25.6% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 100% 0.0% |
| | T2 | 0 0.0% | 22 24.4% | 0 0.0% | 0 0.0% | 0 0.0% | 100% 0.0% |
| | PD | 0 0.0% | 0 0.0% | 9 10.0% | 0 0.0% | 0 0.0% | 100% 0.0% |
| | DL | 0 0.0% | 0 0.0% | 0 0.0% | 9 10.0% | 0 0.0% | 100% 0.0% |
| | DH | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 27 30.0% | 100% 0.0% |
| | | | 100% 0.0% | 100% 0.0% | 100% 0.0% | 100% 0.0% | 100% 0.0% |
| | | T1 | T2 | PD | DL | DH | |
| | | Classes Reais | | | | | |

4.5 Aplicação da amostragem aleatória

Para uma análise mais completa do desempenho das redes propostas para diagnóstico de faltas, foi aplicado o método de amostragem aleatória nos dois modelos propostos, onde em

cada modelo foram realizadas com partições aleatórias de treino e teste, a fim de avaliar seus desempenhos quando diferentes combinações de dados são apresentadas em treinamento. Isto é necessário, pois evita a possibilidade de, em uma divisão, o conjunto de teste seja favorecido com amostras de dados “mais fáceis” de serem aprendidas por cada modelo. Então, para tornar os resultados menos dependentes de uma única divisão foi realizada a amostragem aleatória com reposição. A Tabela 17 apresenta os resultados obtidos nesta análise.

Tabela 17 - Medidas de dispersão da acurácia obtidas em 100 divisões aleatórias.

| Modelo | Medidas de dispersão da acurácia de cada modelo | | | |
|--------|---|--------|---------------|--------|
| | Média | | Desvio Padrão | |
| | Treino | Teste | Treino | Teste |
| CNN | 0.9901 | 0.9794 | 0.0048 | 0.0153 |
| SAE | 0.9923 | 0.9833 | 0.0375 | 0.0139 |

Portanto, considerando os resultados obtidos, o modelo que obteve o melhor desempenho foi a rede neural Autocodificadora Empilhada, que além de obter a maior média de valor de acurácia, também obteve o menor desvio padrão. Com estes resultados, pode-se inferir que ambos modelos possuem uma baixa sensibilidade à variação das amostras nas partições de treino e teste.

4.6 Comparação com resultados da Tabela IEC e modelos propostos na literatura

Para realizar uma comparação de resultados dos dois modelos propostos com os resultados obtidos a partir do método IEC 60599, a mesma base de dados usada para treino e teste dos modelos foi aplicada ao critério IEC. A Tabela 18 apresenta os resultados para os 3 métodos.

Tabela 18 - Comparação de resultados dos modelos propostos e método IEC 60599

| Modelo | % de acerto dos conjuntos | |
|-----------|---------------------------|--------|
| | Treino | Teste |
| SAE | 100% | 100% |
| CNN | 98,1% | 100% |
| IEC 60599 | 94.76% | 91.11% |

Os dois sistemas apresentaram resultados superiores ao método tradicional IEC 60599. Considerando as porcentagens de desempenho apresentadas na Tabela 19, pode-se observar que o sistema classificador baseado na rede neural Autocodificadora Empilhada apresentou 100% de acerto no conjunto de teste e treino. A rede neural convolucional, apesar de ter alcançado um desempenho menor no conjunto de treino, também alcançou um desempenho de 100% no conjunto de teste. Nenhum dos modelos apresentou problemas de *no-decision*, isto é, não houve a indeterminação no momento de classificar as amostras, ao contrário do método IEC que não classificou 7 amostras e errou a classificação de 12.

A Tabela 19 apresenta a classificação de alguns casos do banco de dados TC-10, considerando os dois métodos propostos e o método IEC.

Tabela 19 - Exemplos de diagnóstico para alguns casos da base TC-10.

| C_2H_2/C_2H_4 | CH_4/H_2 | C_2H_4/C_2H_6 | Faltas | SAE | CNN | IEC 60599 |
|-----------------|------------|-----------------|--------|-----|-----|-----------|
| 0.0417 | 1.1628 | 0.4444 | T1 | T1 | T1 | T1 |
| 0.0198 | 1.8438 | 4.0 | T2 | T2 | T2 | T2 |
| 0.0001 | 0.1102 | 0.0001 | PD | PD | PD | NI |
| 1.1667 | 0.1065 | 0.1000 | PD | PD | PD | NI |
| 0.0001 | 0.0476 | 0.0001 | PD | PD | PD | PD |
| 1.0 | 0.1667 | 1.0 | DL | DL | DL | NI |
| 4.0 | 0.1607 | 4.0 | DL | DL | DL | DL |
| 0.6667 | 0.2250 | 4.0 | DH | DH | DH | D |

NI- Não identificado

Tabela 20 - Resultados encontrados na literatura.

| Referência | Número de amostras | | %de diagnóstico corretos | | N. de Saídas |
|------------------------|--------------------|-------|--------------------------|-------------|--------------|
| | Total | Teste | Treino | Teste | |
| Y Zhang et al. 1996 | 40 | (?) | (?) | 95 | 3+N |
| Wang 1998 | 188+22 | 60 | 99.3 a 100 | 93.3 a 96.7 | 5+N |
| YC Huang et al. 2003 | 220 + 600 | 0 | 95.12 | --- | 4+N |
| HT Yang, CC Liao, 1999 | 561 | 280 | 93.88 | 94.9 | 4+N |
| Guardado et al. 2001 | 69 | 33 | 100 | 100 | 5+N |
| Castro e Miranda 2005 | 431 | 139 | 100 | 97.8 | 3 |
| Miranda e Castro 2005 | 318 | 88 | 100 | 99.4 | 5 |
| G. Lv. et al 2005 | 75 | 25 | 100 | 100 | 3+N |
| WH Tang et al 2008 | 168 | (?) | (?) | 80 | 3+N |
| LX Donget al 2008 | 168 | (?) | (?) | 88.3 | 3+N |
| MH Wang et al. 2009 | 21 | 0 | 100 | --- | 8+N |
| SW Fei, XB Zhang 2009 | 142 | (?) | (?) | 94.2 | 3+N |
| NAM Isa et al. 2011 | 160 | 40 | 100 | 100 | 3+N |
| Castro e Miranda 2011 | 318 | 88 | 100 | 100 | 5 |
| K Bacha et al. 2012 | 94 | 30 | (?) | 90 | 6+N |
| Miranda e Castro 2012 | 1400 +318 + 34 | 100 | 100 | 100 | 5+N |

N- Estado normal ou saudável

A Tabela 20 apresenta alguns resultados de classificação para sistemas de diagnóstico já propostos na literatura. Uma comparação direta de resultados deve ser realizada com cuidado visto que, em cada sistema apresentado foram utilizadas bases de treinamento e testes diferentes, tendo caso em que amostras virtuais foram criadas para treinamento do sistema, como em (Miranda e Castro, 2012). Porém, no geral, pode-se observar que os sistemas propostos apresentaram resultados competitivos com os já apresentados na literatura e inovam ao aplicarem aprendizado profundo.

Portanto, como pode ser observado na tabela acima, trabalhos como o de (Miranda e Castro, 2012), obtiveram a mesma precisão nos resultados de diagnóstico, assim como os modelos propostos neste trabalho.

5 CONCLUSÃO

Este trabalho apresentou os resultados obtidos da pesquisa de aplicação de redes neurais profundas para o problema de diagnóstico de faltas incipientes em transformadores baseado na análise dos gases dissolvidos em óleo. Para o desenvolvimento dos sistemas foi utilizada a base de dados TC10 de equipamentos faltosos inspecionados em serviço e usada para a publicação da norma IEC 60599. O primeiro modelo, baseado em uma Rede Autocodificadora Empilhada, apresentou um resultado bastante promissor, tendo obtido 100% de acerto na base de teste para o sistema. Este resultado mostra que para este caso temos um modelo que apresenta uma boa generalização, ou seja, responde adequadamente para dados não utilizados na fase de treino do sistema. O segundo modelo, baseado em uma Rede Neural Convolutiva, apresentou também um resultado de 100% de acerto, considerado promissor para a aplicação. Os dois modelos conseguiram eliminar o problema de *no-decision* característico do IEC 60599. Em todos os casos que o IEC não conseguiu chegar a um diagnóstico, os dois modelos apresentaram o tipo de falta corretamente. Isto se deve ao fato de os dois modelos terem características de aproximadores não-lineares.

Foi aplicado o método de amostragem aleatória para verificar a sensibilidade de ambos os modelos em relação ao desempenho da rede quando aplicada diferentes partições de dados de teste. Portanto, em cem partições feitas aleatoriamente pôde-se observar que os dois modelos alcançaram resultados satisfatórios, demonstrando que, apesar de dispor de um pequeno conjunto de dados para o treinamento, são capazes de extrair as características determinantes para uma classificação correta dos tipos de falta. Considerando os resultados alcançados, mostrou-se que as Redes Neurais Autocodificadoras empilhadas e as Redes Neurais Convolutivas (trabalhando com dados 1-D) são adequadas para o problema de diagnóstico de faltas em transformadores.

Por fim, foi realizada a comparação entre os resultados obtidos e outros resultados encontrados na literatura, onde ficou constatado que este trabalho se destaca entre os de melhor desempenho.

Neste trabalho foram desenvolvidos dois modelos de aprendizagem profunda, com intuito de aplicá-los na tarefa de classificação de faltas incipientes em transformadores imersos em óleo. Sabe-se que este assunto é de grande relevância para o setor elétrico, visto que os defeitos que surgem em transformadores em sua grande maioria não tem suas causas mapeadas, e em parte isso se deve à dificuldade de extrair informação útil por meio de ensaios para realizar o diagnóstico nestes equipamentos, portanto, o uso de redes neurais profundas surge como uma

alternativa à solução deste problema, já que a partir de dados de concentrações dos gases encontrados em seus tanques pode-se realizar o treinamento de um modelo e buscar os limiares das classes de faltas.

Como trabalho futuro pretende-se realizar o estudo da aplicabilidade de outras redes neurais profundas para o problema. Pretende-se também o desenvolvimento de uma metodologia para criação de dados virtuais considerando a base de dados TC-10, na tentativa de se obter uma base de dados maior para o desenvolvimento de sistemas de diagnóstico mais precisos.

REFERÊNCIAS BIBLIOGRÁFICAS

- AURELIO RANZATO, M.; BOUREAU, Y.; CUN, Y. L. Sparse Feature Learning for Deep Belief Networks. In: PLATT, J. C. et al. (Eds.). **Advances in Neural Information Processing Systems 20**. [s.l.] Curran Associates, Inc., 2008. p. 1185–1192.
- BAKAR, N.; ABU-SIADA, A.; ISLAM, S. A review of dissolved gas analysis measurement and interpretation techniques. **IEEE Electrical Insulation Magazine**, v. 30, n. 3, p. 39–49, 2014.
- BENGIO, Y.; COURVILLE, A.; VINCENT, P. Representation learning: a review and new perspectives. **IEEE transactions on pattern analysis and machine intelligence**, v. 35, n. 8, p. 1798–828, 2013.
- BENGIO, Y.; LECUN, Y. Scaling Learning Algorithms towards AI. n. 1, p. 1–41, 2007.
- BOTTOU, L.; BOUSQUET, O. **The tradeoffs of large scale learning**. Advances in neural information processing systems. **Anais**. 2008
- BRAGA, M.; CASTRO, A. **EXTRAÇÃO DE CONHECIMENTO DE UMA REDE NEURAL APLICADA À PREVISÃO DE CARGA A CURTO PRAZO**. (C. B. de Automática, Ed.)Juiz de Fora: 2008
- CASTRO, A. G. R.; MIRANDA, V. **An Interpretation of Neural Networks as Transformer Failure Diagnosis**. International Conference on Probabilistic Methods Applied to Power Systems. **Anais**. Ames: IEEE, 2004. Disponível em: <<https://ieeexplore.ieee.org/document/1378823/>>
- CASTRO, A. R. G.; MIRANDA, V. Knowledge discovery in neural networks with application to transformer failure diagnosis. **IEEE Transactions on Power Systems**, v. 20, n. 2, p. 717–724, 2005.
- CHENG, L.; YU, T. Dissolved gas analysis principle-based intelligent approaches to fault diagnosis and decision making for large oil-immersed power transformers: A survey. **Energies**, v. 11, n. 4, p. 1–69, 2018.
- COLORADO, D. et al. Optimal operation conditions for a single-stage heat transformer by means of an artificial neural network inverse.pdf. **Science**, p. 1281–1290, 2010.
- DA SILVA, A. C. M. **Extração De Conhecimento Em Forma De Regras Difusas a Partir De Mapas Auto-Organizáveis De Kohonen – Aplicação Em Diagnóstico De Falhas**

Incipientes Em Transformadores. [s.l.] Universidade Federal do Pará, 2013.

DE PABLO, A. et al. **On-line condition monitoring of power transformers: A case history.** 2011 Electrical Insulation Conference (EIC). **Anais.** 2011

DONG, M.; YAN, Z.; TANIGUCHI, Y. **Fault diagnosis of power transformer based on model-diagnosis with grey relation.** Proceedings of the 7th International Conference on Properties and Applications of Dielectric Materials (Cat. No.03CH37417). **Anais.** 2003

DORNENBURG, E.; GERBER, O. E. Analysis of dissolved and free gases for monitoring performance of oil-filled transformers. **Brown Boveri Rev**, v. 54, n. 2/3, p. 104–111, 1967.

DUVAL, M. A review of faults detectable by gas-in-oil analysis in transformers. **IEEE Electrical Insulation Magazine**, v. 18, n. 3, p. 8–17, 2002.

DUVAL, M.; DEPABLO, A. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases. **IEEE Electrical Insulation Magazine**, v. 17, n. 2, p. 31–41, 2001.

DUVAL, M.; DUKARM, J. J. Improving the reliability of transformer gas-in-oil diagnosis. **IEEE Electrical Insulation Magazine**, v. 21, n. 4, p. 21–27, 2005.

FILHO, G. L. **Comparação entre os critérios de diagnósticos por análise cromatográfica de gases dissolvidos em óleo isolante de transformador de potência.** [s.l.] Universidade Federal do Pará, 2012.

FONSECA, J. F. **JACKSON FARIAS FONSECA ELEMENTOS DE MANUTENÇÃO EM TRANSFORMADORES DE POTÊNCIA Guaratinguetá - SP 2014.** [s.l.] Universidade Estadual Paulista, 2014.

HALSTEAD, W. D. "A thermodynamic assessment of the formation of gaseous hydrocarbons in faulty transformers. **J. Inst. Petroleum** 59, p. 239–241, 1973.

HINTON, G. E.; SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. **science**, v. 313, n. 5786, p. 504–507, 2006.

HOCHREITER, S. Untersuchungen zu dynamischen neuronalen Netzen. **Diploma, Technische Universität München**, v. 91, n. 1, 1991.

HOCHREITER, S.; SCHMIDHUBER, J. Bridging long time lags by weight guessing and Long Short-Term Memory. **Spatio temporal models in biological and artificial systems**, v. 37, p. 65–72, 1996.

HUANG, Y.-C.; YANG, H.-T.; HUANG, C.-L. Developing a new transformer fault diagnosis system through evolutionary fuzzy logic. **IEEE Transactions on Power Delivery**, v. 12, n. 2, p. 761–767, 1997.

HUBEL, D.; WIESEL, T. RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT ' S STRIATE CORTEX By D . H . HUBEL * AND T . N . WIESEL * From the Wilmer Institute , The Johns Hopkins Hospital and In the central nervous system the visual pathway from retina to striate cortex provides an. p. 574–591, 1959.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. **Imagenet classification with deep convolutional neural networks**. Advances in neural information processing systems. **Anais.2012**

LECUN, Y. et al. Backpropagation applied to handwritten zip code recognition. **Neural computation**, v. 1, n. 4, p. 541–551, 1989.

LECUN, Y. et al. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2324, 1998.

LECUN, Y. et al. Object recognition with gradient-based learning. In: **Shape, contour and grouping in computer vision**. [s.l.] Springer, 1999. p. 319–345.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. 2015.

MIRANDA, V. et al. Improving the IEC Table for Transformer Failure From Neural Networks. **October**, v. 20, n. 4, p. 2509–2516, 2005.

MIRANDA, V.; CASTRO, A. R. G.; LIMA, S. Diagnosing faults in power transformers with autoassociative neural networks and mean shift. **IEEE Transactions on Power Delivery**, v. 27, n. 3, p. 1350–1357, 2012.

MØLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. **Neural networks**, v. 6, n. 4, p. 525–533, 1993.

MORAIS, L. DOS R. **SISTEMA INTELIGENTE PARA MONITORAMENTO E DIAGNÓSTICO DE TRANSFORMADORES BASEADO NA ANÁLISE FÍSICO-QUÍMICA DO ÓLEO E ANÁLISE DOS GASES DISSOLVIDOS EM ÓLEO**. [s.l.] Universidade Federal do Pará, 2008.

MOURA, C. R. S.; CARRETEIRO, R. P. **Lubrificantes e Lubrificação**. Rio de Janeiro, Ed. **Técnica Ltda**, 1987.

NARESH, R.; SHARMA, V.; VASHISTH, M. An Integrated Neural Fuzzy Approach for Fault Diagnosis of Transformers. **IEEE Transactions on Power Delivery**, v. 23, n. 4, p. 2017–2024, 2008.

PANTOJA, M. P. F. **UM MÉTODO FUZZY PARA PROGRAMAÇÃO DE INSPEÇÕES EM TRANSFORMADORES BASEADO NA ANÁLISE DE GASES COMBUSTÍVEIS E PARÂMETROS FÍSICO-QUÍMICOS DO ÓLEO ISOLANTE**. [s.l.] Universidade Federal do Pará, 2014.

PENHA, D. DE P.; CASTRO, A. R. G. **Convolutional Neural Network Applied to the Identification of Residential Equipment in Nonintrusive Load Monitoring Systems**. ResearchGate, 2018

QI, Y. et al. Stacked Sparse Autoencoder-Based Deep Network for Fault Diagnosis of Rotating Machinery. **IEEE Access**, v. 5, p. 15066–15079, 2017.

ROGERS, R. R. Ieee And IEc Codes To Interpret Incipient Faults In Transformers/ Using Gas In Oil Analysis. **IEEE Transactions on Electrical Insulation**, v. EI-13, n. 5, p. 349–354, 1978.

SCHMIDHUBER, J. Deep learning in neural networks : An overview. **Neural Networks**, v. 61, p. 85–117, 2015.

SEIDE, F. et al. **Feature engineering in context-dependent deep neural networks for conversational speech transcription**. 2011 IEEE Workshop on Automatic Speech Recognition & Understanding. **Anais**. 2011.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.

SONG BIN et al. **Study on the fault diagnosis of transformer based on the grey relational analysis**. Proceedings. International Conference on Power System Technology. **Anais**.2002

SOUAHLIA, S.; BACHA, K.; CHAARI, A. Electrical Power and Energy Systems MLP neural network-based decision for power transformers fault diagnosis using an improved combination of Rogers and Doernenburg ratios DGA. **International Journal of Electrical Power and Energy Systems**, v. 43, n. 1, p. 1346–1353, 2012.

STADLER, A. Analyzing UV/Vis/NIR and Photo-Acoustic Spectra: A Note to the Band Gap of $Ti_xSi_{1-x}O_2$. **IEEE Transactions on Semiconductor Manufacturing**, v. 26, n. 1, p. 156–161, 2012.

STOCCO, M. B. C. Avaliação do potencial de aplicação de óleos vegetais como fluidos isolantes em transformadores de distribuição da rede elétrica. 2009.

WEN, L. et al. **A new data-driven intelligent fault diagnosis by using convolutional neural network**. 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). **Anais**. 2017.

WILHELM, H. M.; TULIO, L.; UHREN, W. Produção e uso de óleos vegetais isolantes no setor elétrico. **Engenharia**, n. 592, p. 120–124, 2009.

YI, J.; WANG, J.; WANG, G. Improved probabilistic neural networks with self-adaptive strategies for transformer fault diagnosis problem. v. 8, n. 1, p. 1–13, 2016.