

# Deep Neural Networks with Application to Transformer Failure Diagnosis

Hugo Riviere S. Moraes<sup>1</sup>, Adriana Rosa Garcez Castro<sup>2</sup>

<sup>12</sup>Instituto de Tecnologia – Universidade Federal do Pará (UFPA)

Caixa-postal: 8619 – 66.075.900 – Belém, PA - Brasil

hugo.adventist@gmail.com, adcastro@ufpa.br

**Abstract.** *This paper presents the outcomes of an application research of deep neural networks to diagnosis incipient faults in power transformer through dissolved gas-in-oil analysis (DGA). Two models was proposed, first, using a Stacked Autoencoder Network and later a Convolutional Neural Network. To the development of the system was used the database TC10 which contains data of faults usually found in electrical equipment in service. This database are described in norm IEC 60599. The outcomes achieved, considering the testing data was 100% and 96,5% of accuracy, showed that this models containing a wide applicability to the problem.*

**Resumo.** *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 (DGA). Dois modelos são propostos utilizando Redes Neurais Autocodificadoras Empilhadas e Redes Neurais Convolucionais. Para o desenvolvimento do sistema foi utilizada a base de dados TC10 de equipamentos faltosos inspecionados em serviço e usada para a publicação da norma IEC 60599. Os resultados alcançados, considerando a base de dados de testes (100% e 96,5% de acerto), se mostraram promissores, mostrando, desta forma, a grande aplicabilidade dos modelos para o problema.*

## 1. Introdução

O transformador de potência é um dos equipamentos elétricos mais importantes em um sistema de energia, portanto, é decisivo o seu bom funcionamento para a segurança do sistema. Quando em serviço, um transformador pode estar sujeito a uma variedade de estresses elétricos e térmicos que podem levar o equipamento a uma condição de falta, sendo que em condição de falta alguns gases são desenvolvidos e retidos no óleo isolante. A concentração destes gases e a relação entre estas concentrações permitem a determinação do tipo e severidade da falta ocorrida no transformador.

Alguns métodos tradicionais, baseados na análise dos gases dissolvidos em óleo vêm sendo utilizados pelas empresas de energia para auxílio na detecção de faltas incipientes. Os Métodos de Dörnenburg, de Rogers, Triângulo de Duval e IEC 60599, são alguns destes métodos, os quais têm sido implementados de forma improvisada ou modificada por várias empresas. Em geral, estes métodos permitem detectar dois tipos de faltas, elétrica e térmica, e se baseiam na comparação das concentrações de gases dissolvidos com um “benchmark” contendo valores típicos de concentrações de gases.

De acordo com [Miranda e Castro 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 obter um consenso quanto a valores limites dos gases e seus significados.

A natureza heurística e empírica dos métodos baseados em razões têm 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 pelas empresas de energia para se chegar ao diagnóstico. Neste caso, os especialistas da área devem utilizar, em conjunto, seus conhecimentos para chegar à conclusão. Existem casos em que o diagnóstico é impossível de ser obtido devido à inabilidade destes métodos para fornecer interpretação para algumas combinações de valores de razões dos gases (problema de “no decision”) [Macedo et al 2012].

Visando solucionar as dificuldades relacionadas aos métodos convencionais, alguns sistemas baseados em Inteligência Computacional vêm sendo propostos na literatura, com resultados bastante promissores [Cheng e Yu 2018]. Dentre os modelos propostos, destacam-se os baseados em Redes Neurais Artificiais, Lógica Difusa, sistemas Neuro-difusos, Teoria do Sistema Gray (GST), entre outros.

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 et al 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 e Castro 2017], tais como dados de séries temporais. Dentre as redes neurais profundas destacam-se as Unidades Long Short Term (LSTM), as Redes Neurais Autocodificadoras, as Redes Autocodificadoras Empilhadas (Stacked Autoencoders) e as Redes Neurais Convolucionais (Convolutional Neural Network - CNN).

Considerando os resultados promissores já apresentados na literatura da aplicação de redes neurais profundas para problemas de classificação, 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. Dois modelos são propostos e avaliados. O primeiro modelo, baseado em uma Rede Autocodificadora Empilhada e o segundo modelo baseado em uma Rede Neural Convolucional. Para o desenvolvimento do sistema foi utilizada a base de dados TC10 de equipamentos faltosos inspecionados em serviço usado para a publicação da norma IEC 60599 [IEC Publ. 60599 1999] e publicado em [Duval e dePabla 2001].

## **2. Rede Autocodificadora Empilhada**

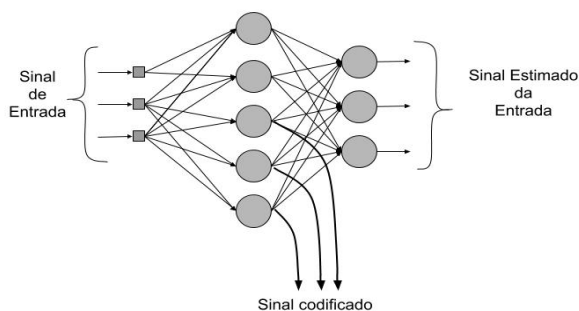
A Rede Neural Autocodificadora (Autoencoder) é um caso especial da rede Perceptron de Múltiplas Camadas (MLP), onde o número de neurônios da camada de entrada é igual ao número de neurônios da camada de saída, sendo que a rede deve ser treinada

para mapear a função de identidade entre entrada e saída, ou seja, os neurônios da camada de saída deverão ter valores aproximadamente iguais aos neurônios da camada de entrada.

A camada intermediária ou escondida da rede Autocodificadora é responsável pela extração de características contidas nos neurônios de entrada. As informações da entrada serão posteriormente reconstruídas nos neurônios de saída. A camada intermediária permite que a rede armazene somente conhecimento relacionado às principais características do conjunto de dados de entrada, atuando como uma camada extratora de características.

Quando o número de neurônios na camada intermediária é menor que o número de neurônios na camada de entrada/saída, tem-se uma rede neural Autocodificadora subcompleta, neste contexto, a camada intermediária deve aprender uma representação compacta do sinal de entrada. Por outro lado, quando o número de neurônios da camada intermediária é maior que o da camada de entrada/saída, tem-se a Autocodificadora completa. A decodificadora subcompleta vêm sendo geralmente utilizada para problemas de compressão de sinais, enquanto que a rede Autocodificadora completa vêm sendo utilizada para problemas de classificação, onde se utiliza para classificação do sinal as características extraídas na camada escondida da Autocodificadora.

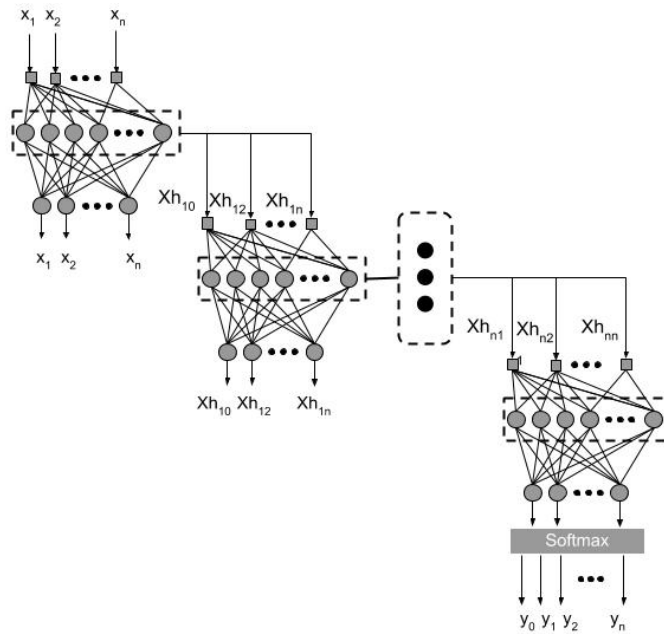
A Figura 1 apresenta a estrutura de uma rede neural Autocodificadora completa.



**Figure 1. Rede Neural Autocodificadora completa.**

Em [Sanger 1989] foi demonstrado que as redes Autocodificadoras são inspiradas na análise de componentes principais (PCA), sendo que quando a camada intermediária trabalha com uma função linear o mapeamento de entrada para camada intermediária é similar ao do PCA, onde tem-se que a informação é projetada ao longo da direção dos eixos ortogonais (autovetores) com a variância minimizada, e com grande perda de informação. Entretanto, quando a função de ativação é não-linear, o mapeamento se torna semelhante ao da análise de componentes principais não-lineares (NLPCA), fornecendo desta forma melhores resultados de características extraídas [Japkowicz 2000]. Para 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 Neural Autocodificadora Empilhada (Stacked Autoencoder - SAE) é composta por várias Redes Neurais Autocodificadoras. A Figura 2 apresenta a estrutura de uma rede SAE.



**Figura 2. Rede Neural Autocodificadora completa.**

A primeira Autocodificadora utiliza os dados da aplicação os quais se deseja a extração de características para futura classificação através de dados rotulados. A segunda Autocodificadora é treinada a partir dos dados codificados pela primeira Autocodificadora, e assim por diante, para todas as Autocodificadoras da estrutura. Por último, temos a camada responsável pelo processo de classificação, onde o sinal codificado da última Autocodificadora treinada é fornecido como entrada para uma MLP com neurônio de saída com a função *softmax*. Com todas as autocodificadoras empilhadas, é realizado então 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 (neurônio com função *softmax*). Com a função *softmax* [Bishop 2006] utilizada na estrutura da SAE é possível obter a probabilidade de um determinado sinal de entrada oferecido a SAE pertencer a uma determinada classe, e esta função é definida como

$$S(y_j) = \frac{e^{y_j}}{\sum_j e^{y_j}} \quad (1)$$

$$\sum_i S_{y_i} \quad (2)$$

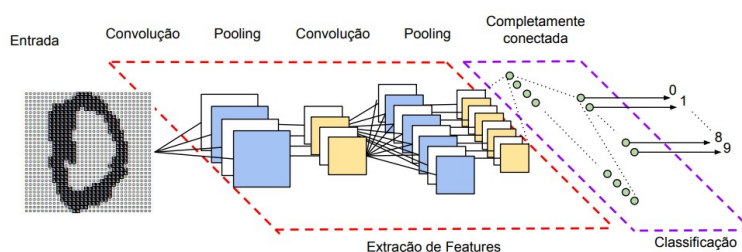
onde  $j$  é o índice dos neurônios de saída da rede,  $i$  corresponde ao índice de um neurônio de saída, ao qual se deseja calcular a probabilidade deste pertencer a uma determinada classe e  $S$  corresponde ao vetor de probabilidades de uma amostra pertencer a uma determinada classe.

### 3. Rede Neural Convolutacional

Uma rede neural convolutacional (Convolutional Neural Network – CNN) é um tipo de rede neural profunda inspirada no córtex visual humano. Este modelo é empregado, geralmente, para problemas de reconhecimento de imagem [LeCun et al 2015].

Uma rede convolutacional é constituída por uma estrutura de dados multidimensional. A arquitetura de uma rede convolutacional, conforme Figura 3, é

composta basicamente por três camadas: convolucional, *pooling* e totalmente conectada. Esta arquitetura está organizada em camadas, denominadas volumes de dados, pois são formados por 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 1999].



**Figure 3. Arquitetura de uma rede convolucional [Vargas 2016].**

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 as camadas ocorrem 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.

Cada filtro possui um tamanho que é determinado por um hiperparâmetro chamado campo receptivo (*receptive field*), portanto, este determina o tamanho da “janela” de convolução. O volume de saída é determinado por três hiperparâmetros: 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. 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 e Zisserman 2014].

Na camada de *pooling* geralmente é utilizada a função *max pooling*, 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.

A função comumente utilizada na ativação de neurônios é denominada Unidades Lineares Retificadas (ReLU). Uma função não-linear que exerce um grande impacto no desempenho do treinamento, reduzindo o tempo de treinamento [Krizhevsky et al 2012]. Esta função é definida como:

$$f(x) = \max(0, x) \quad (3)$$

A saída do último estágio é geralmente um vetor (não uma imagem) que alimenta uma MLP com a função *softmax* no neurônio de saída.

O treinamento na forma de aprendizado supervisionado ocorre com a aplicação do algoritmo padrão *Backpropagation*, onde uma função de otimização é aplicada na redução do erro. É comum aplicar-se a função SGDM - *Stochastic Gradient Descent With Momentum* [Murphy 2012], definida como:

$$\theta_{l+1} = \theta_l - \alpha \nabla E(\theta_l) \quad (4)$$

$$\theta_{l+1} = \theta_l - \alpha \nabla E(\theta_l) + \gamma(\theta_l - \theta_{l-1}) \quad (5)$$

onde  $l$  é o número de iterações,  $\alpha > 0$  é a taxa de aprendizado,  $\theta$  é vetor de otimização dos parâmetros para redução da taxa de erro, e  $E(\theta)$  é a função objetivo.

Em (5), temos a aplicação do *Momentum*, onde  $\gamma$  determina a contribuição do vetor gradiente de iterações anteriores para a iteração atual. A adição deste termo na atualização dos parâmetros reduz a oscilação da função objetivo [Murphy 2012].

#### 4. Sistemas de Diagnóstico de faltas incipientes em Transformadores

A produção de gases em óleo em transformadores pode ser um sinal de faltas incipientes causadas por muitos tipos de fenômenos. Alguns critérios para indicação das faltas no transformador via análise dos gases dissolvidos em óleo têm sido estabelecidos. Estes critérios se baseiam na análise das relações entre 5 gases comumente gerados em condição de falta: hidrogênio ( $H_2$ ), acetileno ( $C_2H_2$ ), etileno ( $C_2H_4$ ), etano ( $C_2H_6$ ) e metano ( $CH_4$ ). A Tabela 1 apresenta o critério IEC 60599, que é um dos métodos mais utilizados pelas concessionárias de energia para interpretação dos gases dissolvidos em óleo.

A Tabela 2 apresenta a composição da base de dados utilizada para desenvolvimento dos sistemas de diagnósticos propostos. Segundo [Miranda et al. 2012], é comum observar que os casos T2 e T3 estão agrupados em muitos estudos, e isto se deve ao fato do pequeno número de casos para estas faltas no banco de dados TC-10, sendo então inadequado para o desenvolvimento de sistemas de diagnóstico automático. Na Tabela 2 tem-se o grupo T2 que corresponde as faltas com  $T > 300^\circ C$ , que corresponde a união de T2 e T3 da tabela IEC.

A base de dados é composta por um total de 300 padrões. Para o desenvolvimento dos sistemas, esta base foi dividida em aproximadamente 70% dos dados para treino e 30% dos dados para teste, correspondendo então a 214 padrões para treino e 86 padrões para teste.

**Tabela 1. Critério IEC 60599**

Caso	Tipo de Falta	$C_2H_2/$ $C_2H_4$	$CH_4/$ $H_2$	$C_2H_4/$ $C_2H_6$
PD	Descarga Parcial	NS *	<0,1	<0.2
DL	Descarga Baixa Energia	> 1	0.1-0.5	> 1
DH	Descarga Alta Energia	0.6-2.5	0.1-1	> 2
T1	Falta Térmica- $T < 300^\circ C$	NS	> 1 mas NS	<1
T2	Falta Térmica- $300^\circ C < T < 700^\circ C$	<0.1	> 1	1-4
T3	Falta Térmica- $T > 700^\circ C$	<0.2	> 1	> 4

NS - Não significativo.

**Tabela 2. Base de dados para Treino e Teste**

<b>Tipo de falta</b>	<i>Treino</i>	<i>Teste</i>
T1	62	15
T2	55	15
PD	24	6
DL	20	11
DH	53	39

#### **4.1. Sistema de Diagnóstico de Falhas via Autocodificadoras Empilhadas**

Para este modelo foi escolhida a arquitetura com duas Autocodificadoras empilhadas e cada Autocodificadora com 20 neurônios na camada escondida e função de ativação sigmoide. A entrada e saída para a primeira Autocodificadora são as 3 razões de gases adotadas pelo método IEC 60599. O sistema foi treinado para o diagnóstico de 5 tipos de falhas. Essa arquitetura foi escolhida após exaustivos treinos variando-se o número de neurônios das camadas escondidas das duas Autocodificadoras e analisando-se a porcentagem de acerto do sistema considerando o conjunto de teste.

#### **4.2. Sistema de diagnóstico de Falhas via Rede Neural Convolutacional**

Para este modelo foi escolhida uma arquitetura de rede neural convolutacional com 5 camadas, sendo uma camada de convolução, seguida de uma ReLU, uma totalmente conectada, e, por fim, uma camada que aplica a função *softmax* e uma camada de classificação. Essa arquitetura foi escolhida após exaustivos treinos variando-se o número de camadas da rede CNN e analisando-se a porcentagem de acerto do sistema considerando o conjunto de teste.

Para treinamento do sistema de diagnóstico via rede neural convolutacional foram utilizadas como entradas as 3 razões dos gases, assim como no modelo das redes Autocodificadoras empilhadas, porém, neste caso, foi necessário realizar um pré-processamento deste vetor de entrada, pois, a entrada para a rede neural convolutacional deveria ser uma matriz quadrada. Portanto, foi realizado o redimensionamento da matriz de entrada de treino para quatro dimensões, isto é, uma matriz 4-D, assumindo a dimensão 4x4x1x214. Desta forma, a CNN pode interpretar os dados 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, ou seja, indexam os valores das razões dos gases.

#### **4.3. Treinamento e Teste dos Modelos**

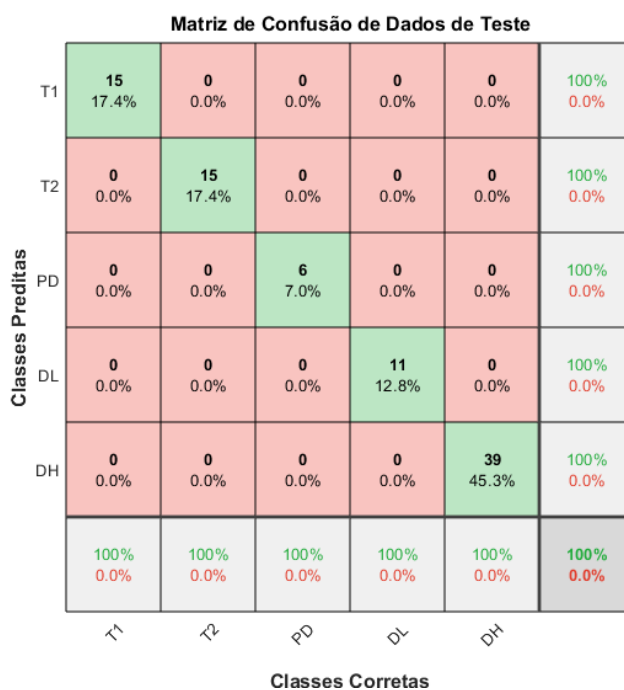
Os dois modelos de aprendizagem profunda foram treinados utilizando o *software* Matlab. A Tabela 3 apresenta a comparação dos resultados de classificação obtidos sobre a base de testes para ambos os modelos. A tabela também apresenta os resultados de classificação utilizando o critério IEC 60599 apresentado na Tabela 1.

Os dois sistemas apresentaram resultados superiores ao método tradicional IEC 60599. O sistema classificador baseado na rede neural Autocodificadora empilhada apresentou 100% de acerto no conjunto de teste. Entretanto, a rede neural convolucional, apesar de não ter alcançado 100% de acerto, mostrou resultados promissores para o problema, tendo errado apenas 3 casos de classificação no conjunto de teste.

**Tabela 3. Comparação de resultados dos Modelos proposto.**

Modelo	% de acerto do conjunto de teste	Número total de faltas não identificadas ou casos com diagnóstico incorreto
SAE	100	0 erros
CNN	96.51	3 erros
IEC 60599	89.53	7 não identificadas e 12 erros

As Figuras 4 e 5 apresentam a matriz de confusão para o caso das Redes Autocodificadoras Empilhadas e para o caso da Rede Neural Convolucional, respectivamente.



**Figura 5. Matriz de confusão do conjunto de dados de teste para rede Autocodificadora empilhada.**



**Matriz de Confusão de Dados de Teste**

Classes Preditas	T1	14 16.3%	1 1.2%	0 0.0%	0 0.0%	0 0.0%	93.3% 6.7%
	T2	0 0.0%	14 16.3%	1 1.2%	0 0.0%	0 0.0%	93.3% 6.7%
	PD	1 1.2%	0 0.0%	5 5.8%	0 0.0%	0 0.0%	83.3% 16.7%
	DL	0 0.0%	0 0.0%	0 0.0%	11 12.8%	0 0.0%	100% 0.0%
	DH	0 0.0%	0 0.0%	0 0.0%	0 0.0%	39 45.3%	100% 0.0%
		93.3% 6.7%	93.3% 6.7%	83.3% 16.7%	100% 0.0%	100% 0.0%	96.5% 3.5%
	↖	↖	PD	DL	DH		Classes Corretas

**Figure 6. Matriz de confusão do conjunto de dados de teste para rede convolucional.**

Os resultados comparativos com o IEC demonstram que o conjunto de testes utilizado não é especificamente um conjunto fácil para diagnóstico, visto que o método IEC não conseguiu identificar 7 casos e identificou 12 casos erroneamente. Os dois modelos propostos conseguiram identificar corretamente todas os casos que o IEC não consegue identificar (*no-decision*).

A Tabela 4 apresenta a classificação para alguns casos do banco de dados TC-10, considerando os dois métodos propostos e o método IEC.

**Tabela 4. Exemplos de diagnóstico para alguns casos da base TC-10**

$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$	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	DH

A Tabela 5 apresenta alguns resultados de sistemas de diagnóstico de faltas incipientes em transformadores já apresentados na literatura.

**Tabela 5. Resultados de trabalhos já apresentados na Literatura**

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

N - Estado normal ou saudável

Uma comparação direta de resultados deve ser realizada com cuidado visto que, em cada sistema apresentado, foram utilizadas bases de treinamento e testes diferentes. Porém, no geral, pode-se observar que os sistemas propostos apresentaram resultados competitivos com os já apresentados na literatura.

## 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 do sistema foi utilizada a base de dados TC10 de equipamentos faltosos inspecionados em serviço usado para a publicação da norma IEC 60599.

O primeiro modelo, baseado em uma Rede Autocodificadora Empilhada, apresentou 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 considerado promissor para a aplicação. Em um total de 86 padrões, errou apenas 3 casos, mostrando também ser um modelo com boa generalização dos dados.

Os dois modelos conseguiram eliminar o problema de *no-decision* característico do IEC 60599. Em todos os casos que o IEC não consegue chegar a um diagnóstico, os dois modelos apresentaram o tipo de falta corretamente. Isto se deve ao fato dos dois modelos terem características de aproximadores não-lineares.

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. 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

- Bacha, K., Souahlia, S., & Gossa, M. (2012). Power transformer fault diagnosis based on dissolved gas analysis by support vector machine. *Electric power systems research*, 83(1), 73-79.
- Cheng, L., & Yu, T. (2018). Dissolved Gas Analysis Principle-Based Intelligent Approaches to Fault Diagnosis and Decision Making for Large Oil-Immersed Power Transformers: A Survey. *Energies*, 11(4), 913.
- Cristopher, M., & Bishop, S. (2006). *Pattern recognition and machine learning*. Springer, New York, NY.
- Castro, A. R. G., Miranda, V., & Lima, S. (2011, September). Transformer fault diagnosis based on autoassociative neural networks. *16th International Conference on Intelligent System Application to Power Systems (ISAP)*, (pp. 1-5). IEEE.
- Castro, A., & Miranda, V. (2010, May). Sistema inteligente para diagnóstico de faltas incipientes em transformadores baseado em redes neurais auto-associativas. In *Proceedings of SBSE 2010-Simpósio Brasileiro de Sistemas Elétricos*.
- da Silva, A. C. M., Castro, A. R. G., & Miranda, V. (2012). Transformer failure diagnosis by means of fuzzy rules extracted from Kohonen Self-Organizing Map. *International Journal of Electrical Power & Energy Systems*, 43(1), 1034-1042.
- de Paiva Penha, D., & Castro, A. R. G. (2017). Convolutional neural network applied to the identification of residential equipment in non-intrusive load monitoring systems.

- In 3rd International Conference on Artificial Intelligence and Applications (pp. 11-21).
- Duval, M., & dePabla, A. (2001). "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," in IEEE Electrical Insulation Magazine, vol. 17, no. 2, pp. 31-41.
- Fei, S. W., & Zhang, X. B. (2009). Fault diagnosis of power transformer based on support vector machine with genetic algorithm. *Expert Systems with Applications*, 36(8), 11352-11357.
- IEC Publ. 60599. (1999, March). Interpretation of the analysis of gases in transformers and other oil-filled electrical equipment in service.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
- Japkowicz, N. (2000, July). Learning from imbalanced data sets: a comparison of various strategies. In *AAAI workshop on learning from imbalanced data sets* (Vol. 68, pp. 10-15).
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436.
- LeCun, Y., Haffner, P., Bottou, L., & Bengio, Y. (1999). Object recognition with gradient-based learning. In *Shape, contour and grouping in computer vision* (pp. 319-345). Springer, Berlin, Heidelberg.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. Adaptive Computation and Machine Learning. The MIT Press, Cambridge, Massachusetts.
- Rao, S., de Medeiros Martins, A., & Príncipe, J. C. (2009). Mean shift: An information theoretic perspective. *Pattern Recognition Letters*, 30(3), 222-230.
- Sanger, T. D. (1989). Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural networks*, 2(6), 459-473.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- V. Miranda, A. R. G. Castro & S. Lima. (2012, April). Diagnosing Faults in Power Transformers With Autoassociative Neural Networks and Mean Shift. *IEEE Transactions on Power Delivery*, vol. 27, no. 3, pp. 1350-1357.
- Vargas, A. C. G., Paes, A., & Vasconcelos, C. N. (2016, July). Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In *Proceedings of the XXIX Conference on Graphics, Patterns and Images* (pp. 1-4).
- V. Miranda and A. R. G. Castro. (2005, Oct.). "Improving the IEC table for transformer failure diagnosis with knowledge extraction from neural networks," in *IEEE Transactions on Power Delivery*, vol. 20, no. 4, pp. 2509-2516.
- Yang, H. T., & Liao, C. C. (1999). Adaptive fuzzy diagnosis system for dissolved gas analysis of power transformers. *IEEE Transactions on Power Delivery*, 14(4), 1342-1350.