

Metodologia para a Seleção Ótima de Instâncias via Evolução Diferencial para a Classificação de Distúrbios de Qualidade da Energia Elétrica

***Nota: Informações sobre autores deverão ser informadas apenas na versão final, em caso de aprovação.**

Matheus do Val Oliveira
Departamento de Engenharia Elétrica e de
Computação
Universidade de São Paulo (EESC-USP)
São Carlos, Brasil
matheusdoval@usp.br

Moisés Júnior Batista Borges Davi
Departamento de Engenharia Elétrica e de
Computação
Universidade de São Paulo (EESC-USP)
São Carlos, Brasil
moisesdavi@usp.br

Mário Oleskovicz
Departamento de Engenharia Elétrica e de
Computação
Universidade de São Paulo (EESC-USP)
São Carlos, Brasil
olesk@sc.usp.br

Resumo—O crescimento das cargas chaveadas e não lineares no sistema elétrico, principalmente associado ao crescimento das fontes renováveis, tem aumentado a necessidade de metodologias eficientes para a classificação de Distúrbios da Qualidade da Energia Elétrica (DQEEs). Algoritmos baseados em Inteligência Artificial, como Redes Neurais Artificiais (RNAs), têm sido constantemente estudados para essa aplicação, mas exigem uma etapa de pré-processamento dos sinais para extrair instâncias representativas. Entretanto, nem todas as instâncias extraídas contribuem significativamente para a classificação. Para otimizar essa seleção, este trabalho propõe o uso da Evolução Diferencial Binária (EDB) aplicada às instâncias obtidas por meio da Transformada de Stockwell Discreta (TSD) em sinais de DQEE, utilizando RNAs como classificadores. Para avaliar a metodologia, sinais sintéticos representando nove tipos de DQEEs foram gerados no software Matlab. Após a aplicação da TSD, a EDB foi utilizada para avaliar diferentes combinações de instâncias em uma mesma topologia de RNA, visando selecionar a melhor configuração. O processo foi repetido considerando sinais com relação sinal-ruído (SNR) de 40 dB. Adicionalmente, a metodologia foi avaliada quando aplicada para o desenvolvimento de um classificador de faltas para um sistema de transmissão modelado no ATPDraw, ressaltando sua capacidade de generalização para outras tarefas além da classificação de distúrbios de qualidade de energia. Os resultados demonstraram acurácia superior a 98%, validando a eficácia do método proposto.

Palavras-Chave—Qualidade da Energia Elétrica, Evolução Diferencial Binária, Transformada de Stockwell, Redes Neurais Artificiais, Classificação de Faltas.

I. INTRODUÇÃO

A qualidade da energia elétrica tem se tornado um problema significativo tanto para consumidores quanto para as concessionárias, devido ao constante aumento do uso de dispositivos eletrônicos e de eletrônica de potência, de sistemas desequilibrados e de cargas não lineares [1]. A norma IEEE 1159-2019 [2] também aponta esses fatores como possíveis causas para a distorção das formas de onda das tensões e correntes medidas. Diversas metodologias têm sido empregadas para identificar esses fenômenos, geralmente aplicando técnicas de processamento de sinais para extrair informações representativas do sinal original, aqui denominadas *instâncias*. Essas instâncias são posteriormente

utilizadas em técnicas de reconhecimento de padrões, como as discutidas em [3]. Entre as ferramentas de pré-processamento de sinais destacadas em [3], ressalta-se a Transformada de Stockwell Discreta (TSD), que permite a extração de informações de sinais não estacionários simultaneamente nos domínios do tempo e da frequência.

Por exemplo, no estudo de [1], modificações na TSD e no uso das Máquinas de Vetores de Suporte (SVM) permitiram a classificação de nove Distúrbios da Qualidade da Energia Elétrica (DQEEs), alcançando taxas de acerto superiores a 97% em mais de 90% das classes em cenários com relação sinal-ruído (SNR) de 20 dB. Entretanto, estudos como o de [4] propõem melhorias na combinação entre a TSD e os classificadores, otimizando a seleção das instâncias utilizadas na classificação. Especificamente, os autores em [4] empregam a Otimização por Enxame de Partículas para selecionar as instâncias extraídas da TSD e pela Transformada Tempo-Tempo (TTT). Em seguida, avaliam diferentes algoritmos de aprendizagem para classificar 11 tipos de DQEEs, visando aumentar a robustez ao ruído. O método alcançou uma taxa de acerto de 95%, ao mesmo tempo em que reduziu a quantidade de instâncias utilizadas, sem recorrer a metodologias multiobjetivo. Por outro lado, os autores em [5] utilizam o Algoritmo Genético de Ordenação Não-Dominada II para minimizar simultaneamente o erro de classificação e a quantidade de características selecionadas para o classificador, que neste caso é baseado em Árvores de Decisão. Assim como no trabalho anterior, a TSD e a TTT são aplicadas na extração de características, permitindo a classificação de 15 tipos de DQEEs com acurácia máxima de 99,93%. Já em [6], a otimização multiobjetivo é realizada por meio da Otimização por Colônia de Formigas, também utilizando a TSD e a TTT. No entanto, neste estudo, três classificadores diferentes foram testados: as Árvores de Decisão, K-Vizinhos Mais Próximos e Máquinas de Vetores de Suporte. O objetivo foi avaliar a capacidade de generalização da metodologia em diferentes abordagens de aprendizado. Embora o melhor resultado tenha sido idêntico ao trabalho anterior (99,93%), este método conseguiu reduzir pela metade o número de instâncias necessárias, além de executar todo o processo em metade do tempo.

Inspirado nos trabalhos citados, este estudo explora as potencialidades das RNAs, com foco na arquitetura Perceptron de Múltiplas Camadas (PMC), para a classificação de DQEEs. A TSD é empregada como ferramenta de pré-processamento, enquanto que a Evolução Diferencial Binária (EDB) é utilizada para a seleção otimizada das instâncias mais relevantes. Por se tratar de uma metodologia baseada em técnicas computacionalmente mais simples quando comparadas às de [5] e [6], por exemplo, sua implementação é acessível e permite alcançar taxas de acerto comparáveis às dos trabalhos anteriores. Inicialmente, foram gerados nove conjuntos distintos de DQEEs, que passaram pelo pré-processamento com a TSD e pela seleção de instâncias via EDB. O mesmo procedimento foi repetido considerando a presença de ruído, com relação sinal-ruído (SNR) de 40 dB. Além disso, a metodologia foi aplicada à classificação de faltas em uma Linha de Transmissão (LT) utilizando os mesmos princípios de pré-processamento e otimização. O desempenho do modelo foi avaliado tanto em condições ideais quanto em cenários ruidosos (SNR = 40 dB).

O presente trabalho é dividido da seguinte forma. Tendo a Seção 1 realizado a introdução e contextualização do problema, a Seção 2 apresenta a fundamentação teórica necessária para a compreensão dos principais conceitos e ferramentas empregadas. A Seção 3 detalha a metodologia empregada neste trabalho, enfatizando a aplicação da TSD, a obtenção de instâncias, e a operação do algoritmo da EDB. A Seção 4 apresenta os resultados da implementação da metodologia para a classificação de DQEEs e para a classificação de faltas na LT. A Seção 5 apresenta as conclusões do trabalho, e é seguida pelas referências bibliográficas.

II. FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os principais conceitos e ferramentas utilizadas neste trabalho, destacando sua fundamentação teórica e justificando sua aplicação. A Seção II-A aborda a Transformada de Stockwell em suas formas contínua e discreta. Em seguida, a Seção II-B detalha as etapas do algoritmo da EDB. Por fim, a Seção II-C discute a utilização de RNAs, com foco na *toolbox* de Estatística e Aprendizagem de Máquina do Matlab.

A. Transformada de Stockwell

A Transformada de Stockwell Contínua (TSC) para um sinal temporal $h(t)$ é definida em [7] como:

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi f t} dt \quad (1)$$

Na equação (1), τ representa um parâmetro temporal, e f um parâmetro espectral. A TSC permite obter informações sobre as frequências presentes no sinal ao longo do tempo. A versão discreta da transformada, denominada Transformada de Stockwell Discreta (TSD), é aplicada a um sinal amostrado $h[kT]$, onde T representa o período de amostragem. Sua definição, conforme apresentada em [7] como:

$$S[jT, \frac{n}{NT}] = \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] e^{-\frac{2\pi^2 m^2}{n^2}} e^{\frac{i2\pi j}{N}}, n \neq 0 \quad (2)$$

sendo jT o índice correspondente ao tempo e $\frac{n}{NT}$ o índice associado à frequência, onde N sendo o tamanho do sinal. Os índices variam conforme $n, j = 0, 1, 2, \dots, N-1$. Como resultado, a TSD é representada por uma matriz, na qual cada

linha fornece informações sobre um determinada frequência, enquanto cada coluna contém informações referentes a um dado instante específico de tempo. O termo $H\left[\frac{n}{NT}\right]$ corresponde à Transformada de Fourier do sinal $h[kT]$.

B. Evolução Diferencial Binária

O algoritmo da Evolução Diferencial, desenvolvido por [8], é uma meta-heurística utilizada para minimizar funções não lineares em espaços contínuos e não diferenciáveis. Neste trabalho, emprega-se a versão binária deste algoritmo (EDB), conforme [9]. A implementação da EDB é dada pelas seguintes etapas:

1) Inicialização da População

A população inicial é representada pela matriz $X_{i,j}^{G=1}$, em que cada elemento da população tem D genes e que cada geração terá NP indivíduos (soluções candidatas), e $G=1$ indica a primeira geração. Cada indivíduo da população é gerado aleatoriamente, seguindo uma distribuição uniforme, conforme a equação:

$$X_{i,j}^{G=1} = \text{round}(\text{rand}(1)) \quad (3)$$

com $i = 1, 2, \dots, NP$ e $j = 1, 2, \dots, D$. No Matlab, $\text{rand}(1)$ gera valores aleatórios entre 0 e 1, e $\text{round}()$ arredonda para 0 ou 1, criando uma representação binária para cada indivíduo.

2) Mutação

Na etapa de mutação, cada gene da população sofre uma inversão probabilística, gerando uma nova matriz, M^G . A mutação segue a regra:

$$M_{i,j}^G = 1 - X_{i,j}^G, \text{ se } F < \text{rand}(1) \quad (4)$$

onde F é o coeficiente de mutação, que define a taxa de alteração dos genes.

3) Cruzamento

O cruzamento combina elementos das matrizes X^G (população original) e M^G (pós-mutação) formando uma nova matriz, U^G , conforme a equação:

$$U_{i,j}^G = \begin{cases} M_{i,j}^G, & \text{se } \text{rand}_j(1) \leq CR \text{ ou } j = \text{rnbr}(i) \\ X_{i,j}^G, & \text{caso contrário} \end{cases} \quad (5)$$

onde CR é o coeficiente de cruzamento. O cruzamento ocorre quando um número aleatório $\text{rand}_j(1)$ for menor que CR ou quando um índice aleatório $\text{rnbr}(i)$ coincidir com o índice j . Essa segunda condição garante que, para cada indivíduo, ao menos um gene será modificado [8].

4) Seleção

Após o cruzamento, cada indivíduo da população é avaliado com base na Função Objetivo (FO), comparando-se a matriz original, X^G , com a matriz pós-cruzamento, U^G . A nova população, X^{G+1} , é definida pela regra:

$$X_i^{G+1} = \begin{cases} U_i^G, & \text{se } FO^U < FO^X \\ X_i^G, & \text{caso contrário} \end{cases} \quad (6)$$

onde FO^U e FO^X representam os valores da FO para U^G e X^G , respectivamente. Para um problema de minimização, um indivíduo da nova geração será substituído somente se U^G apresentar um valor de FO menor do que X^G .

C. Redes Neurais Artificiais

Neste trabalho, as RNAs são utilizadas como classificador de padrões, adotando a arquitetura das redes PMC. O desenvolvimento e treinamento das redes foram realizados no software Matlab, por meio da *toolbox* de Estatística e

Aprendizagem de Máquina, que disponibiliza a função *fitnet* [10] para a criação de modelos específicos de classificação. A função *fitnet* emprega o algoritmo de treinamento Broyden-Fletcher-Goldfarb-Shanno (BFGS) de memória limitada, um método *quasi*-Newton de minimização da função de custo. Esse algoritmo realiza uma busca iterativa baseada em uma aproximação da matriz Hessiana. Tanto o algoritmo de treinamento quanto a função de custo, que é a entropia binária cruzada, adequada para problemas de classificação, não podem ser alteradas pelo usuário. Entretanto, é possível ajustar diversos hiper parâmetros do modelo, incluindo a quantidade de camadas intermediárias, a quantidade de neurônios por camada, as funções de ativação em cada camada, e a inicialização dos pesos sinápticos e *biases*. O detalhamento da implementação da função *fitnet* e sua adaptação à metodologia proposta é apresentada na Seção 3, uma vez que seu aprofundamento foge ao escopo desta seção.

III. METODOLOGIA

As principais etapas da metodologia podem ser identificadas no fluxograma da Figura 1 e são detalhadas nas Seções III-A, abordando a geração dos sinais em análise; III-B, descrevendo a aplicação da TSD nos sinais sintéticos e o processo da extração das instâncias; III-C, sobre a implementação da EDB e das RNAs; e III-D, tratando da análise dos resultados obtidos.

A. Geração e manipulação dos sinais

Os sinais representativos dos DQEEs foram gerados com base em [4], utilizando as equações apresentadas na Tabela I. Para isso, foram simulados, no Matlab, 9 distúrbios distintos, codificados como Operação Normal (C1), Afundamento (C2), Elevação (C3), Interrupção (C4), *Flicker* (C5), Transitório Oscilatório (C6), Harmônicos (C7), *Notch* (C8) e *Spike* (C9). Embora a norma IEEE 1159-2019 [2] não classifique *Spikes* como deformações típicas das formas de onda, esses eventos foram incluídos na análise para fins de comparação. Sua modelagem segue uma inversão de sinal em relação à equação dos *Notches*, conforme descrito na Tabela I. Assim, enquanto os *Notches* representam recortes em direção ao interior da onda, os *Spikes* correspondem a perturbações direcionadas ao exterior da forma de onda.

Para cada DQEE listado na Tabela I, foram gerados 250 casos distintos, totalizando 2250 amostras. Destas, 200 de cada tipo (1800 no total) foram utilizadas para o treinamento da RNA, enquanto as 50 amostras restantes de cada classe (450 no total) foram reservadas para a etapa de teste. A frequência fundamental adotada foi de 50 Hz, com base nos artigos de revisão, e a frequência de amostragem foi de 128 amostras por ciclo, resultando em 6,4 kHz. Os conjuntos de treinamento foram gerados explorando os valores extremos dos parâmetros que caracterizam cada distúrbio, garantindo que as RNAs fossem treinadas com as variações mais representativas dos fenômenos analisados. Já os dados de teste foram gerados com parâmetros aleatórios dentro dos intervalos especificados.

Como prática comum na área, a metodologia proposta foi avaliada em um sistema elétrico realista visando ressaltar sua capacidade de generalização para outras tarefas além da classificação de DQEEs. Neste trabalho, a metodologia foi aplicada para o desenvolvimento de um classificador de faltas em LTs. Para isso, uma LT de 400 kV e 300 km foi modelada em ATPDraw com base em [11], e é indicada na Figura 2. O

bloco *M* na figura indica o terminal da LT em que as tensões trifásicas foram medidas.

Para a elaboração dos conjuntos de treinamento e de teste, a escolha dos parâmetros foi baseada em [12]. Em relação à operação do sistema, sua frequência fundamental é de 50 Hz, e a frequência de amostragem utilizada é de 6,4 kHz. O conjunto de treinamento incluiu faltas monofásicas (AT, BT e CT), bifásicas (AB, BC, CA), bifásicas-terra (ABT, BCT, CAT) e trifásicas (ABC), considerando variações no Ângulo de Incidência de Falta (AIF) em 0, 45 e 90°, na resistência fase-terra (R_G) em 0, 25, 50 e 100 Ω , e na resistência entre fases (R_P) em 0, 1, 1,5 e 2 Ω . As faltas foram aplicadas ao longo de 5 a 95% do comprimento da LT, em incrementos de 5%. Para as faltas bifásicas-terra, R_P foi fixada em 1 Ω , enquanto R_G variou para garantir um conjunto de dados balanceado e evitar viés na RNA. No total, o conjunto de treinamento contou com 2280 casos. Já o conjunto de teste foi gerado com os mesmos tipos de falta, mas com diferentes parâmetros. Os AIFs considerados foram 15, 30, 60 e 75°, enquanto R_G assumiu valores de 15, 35 e 75 Ω e R_P variou entre 0,25, 1,25 e 1,75 Ω . As faltas foram aplicadas de 7 a 91% do comprimento da LT, em passos de 7%. Para as faltas bifásicas-terra, R_P foi fixada em 1,75 Ω . Dessa forma, o conjunto de teste totalizou 1580 casos.

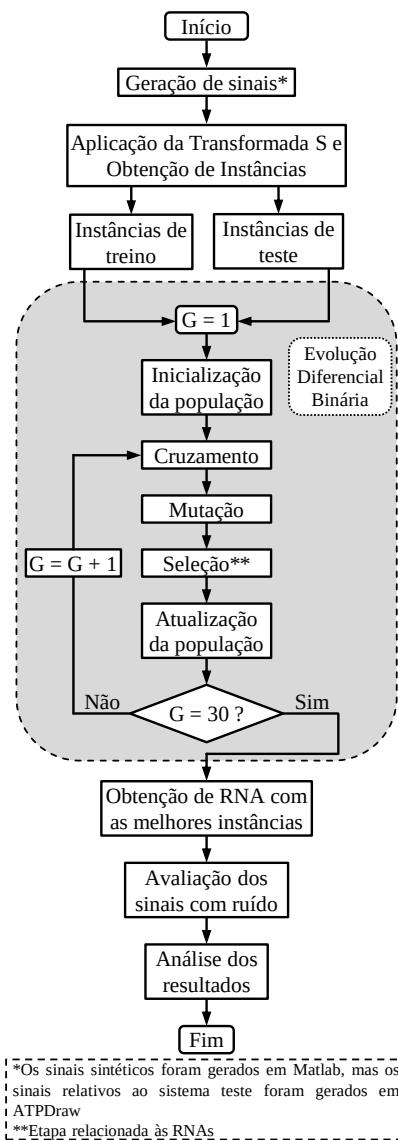


Fig. 1. Fluxograma geral do trabalho.

TABLE I. DISTÚRBIOS DE QUALIDADE DA ENERGIA ELÉTRICA

DQEE	Código	Equação	Parâmetros
Operação Normal	C1	$A \sin(\omega t)$	$0.9 < A < 1.1$
Afundamento	C2	$\{1 - \alpha[u(t - t_1) - u(t - t_2)]\} \sin(\omega t)$	$0.1 < \alpha < 0.9, T < t_2 - t_1 < 9T$
Elevação	C3	$\{1 - \beta[u(t - t_1) - u(t - t_2)]\} \sin(\omega t)$	$0.1 < \beta < 0.8, T < t_2 - t_1 < 9T$
Interrupção	C4	$\{1 - \rho[u(t - t_1) - u(t - t_2)]\} \sin(\omega t)$	$0.9 < \rho < 1.0, T < t_2 - t_1 < 9T$
Flicker	C5	$[1 + \alpha \sin(2\pi\beta t)] \sin(\omega t)$	$0.1 < \alpha < 0.2, 5 < \beta < 20 \text{ Hz}$
Transitório Oscilatório	C6	$\sin(\omega t) + \beta e^{-\frac{(t-t_1)}{\tau}} \sin(\omega_n(t - t_1))[u(t - t_1) - u(t - t_2)]$	$0.1 < \beta < 0.8, 0.5T < t_2 - t_1 < 3T$ $8 < \tau < 30 \text{ ms}, 300 < f_n < 900 \text{ Hz}$
Harmônicos	C7	$\sum_{n=1}^4 a_{2n-1} \sin((2n-1)\omega t)$	$0.05 < a_3, a_5, a_7 < 0.15,$ $\sum_{n=1}^4 a_{2n-1}^2 = 1$
Notch	C8	$\sin(\omega t) - \text{sign}(\sin(\omega t)) \left\{ \kappa \sum_{n=0}^9 (u(t - (t_1 + 0,02n)) - u(t - (t_2 + 0,02n))) \right\}$	$0.1 < \kappa < 0.4,$ $0 < t_1, t_2 < 0.5, 0.01T < t_2 - t_1 < 0,05T$
Spike	C9	$\sin(\omega t) + \text{sign}(\sin(\omega t)) \left\{ \kappa \sum_{n=0}^9 (u(t - (t_1 + 0,02n)) - u(t - (t_2 + 0,02n))) \right\}$	$0.1 < \kappa < 0.4,$ $0 < t_1, t_2 < 0.5, 0.01T < t_2 - t_1 < 0,05T$

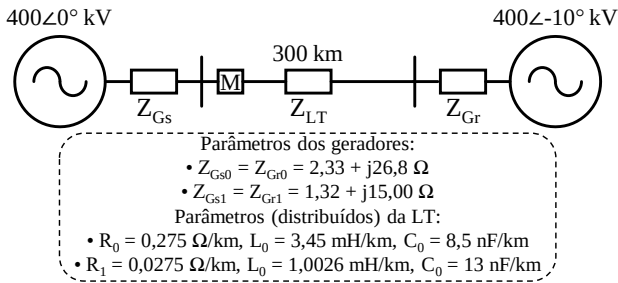
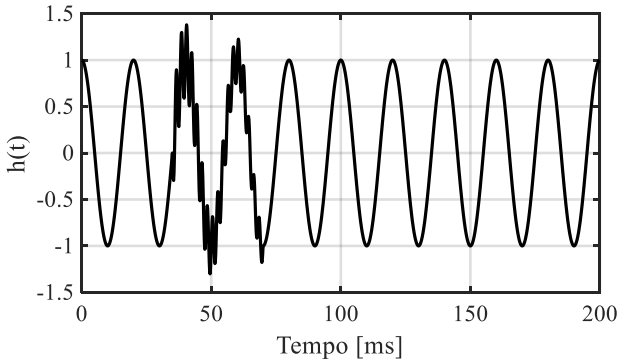
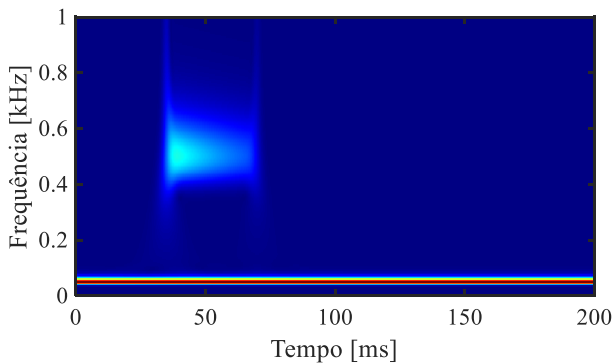


Fig. 2. Sistema elétrico simulado com base em [11].


 Fig. 3. Sinal $h(t)$, forma de onda representativa de transitório oscilatório.

 Fig. 4. Matriz resultante da aplicação da TSD sobre $h(t)$.

Além disso, para avaliar a robustez da metodologia, foi introduzido ruído gaussiano nos sinais sintéticos e simulados. O nível de ruído foi definido SNR de 40 dB a fim de verificar o impacto da degradação do sinal na capacidade de generalização das RNAs. Apesar de [13] sugerir níveis de ruído entre 50 e 48 dB, optou-se por utilizar 40 dB, como em [4] para avaliar o potencial da metodologia. Os conjuntos de treinamento e de teste originais (sem ruído) são utilizados na segunda etapa da metodologia, detalhada na próxima seção.

B. Aplicação da Transformada de Stockwell Discreta e obtenção de instâncias

Para ilustrar a aplicação da TSD e sua utilização na extração de instâncias representativas, foi gerado um sinal de transitório oscilatório conforme a equação (C6) na Tabela I, em que $f = 50 \text{ Hz}$ é a frequência fundamental, $\omega = 2\pi f$ é a frequência angular correspondente, $\beta = 0.45$ é a amplitude máxima do transitório, $t_1 = 35 \text{ ms}$ e $t_2 = 70 \text{ ms}$ são, respectivamente, os instantes de início e término do transitório, $\tau = 40 \text{ ms}$ é a taxa de decaimento, $f_n = 500 \text{ Hz}$ é a frequência do transitório e $\omega_n = 2\pi f_n$ é a frequência angular associada. A Figura 3 apresenta o sinal $h(t)$, enquanto a Figura 4 exibe a matriz obtida pela aplicação da TSD no mesmo sinal.

Observando as Figuras 3 e 4, nota-se uma correspondência evidente entre o comportamento do sinal no domínio do tempo e sua representação no domínio tempo-frequência. Na Figura 3, a amplitude do transitório se reduz gradualmente ao longo do tempo, permitindo a identificação visual do início e do término da perturbação. A Figura 4, por sua vez, exibe a distribuição espectral do sinal ao longo do tempo. O início do transitório é evidenciado pelo espalhamento espectral, com predominância da frequência de 500 Hz. À medida que o tempo avança, observa-se uma atenuação progressiva da intensidade do espectro, acompanhada por um novo espalhamento espectral ao final da perturbação.

Para os sinais sintéticos, os 10 ciclos foram utilizados como entrada, resultando em inter-harmônicos respectivas TSDs. Para prosseguir nas análises, optou-se por excluir as componentes inter-harmônicas, ficando somente com as múltiplas inteiras de 50 Hz. Um estudo em paralelo foi realizado, por meio da aplicação da transformada de Fourier,

para verificar as amplitudes máximas e mínimas dos componentes harmônicos presentes em todos os sinais. Após essa verificação, foram selecionadas as frequências até a 19ª harmônica (950 Hz) para os DQEEs. Para o sistema de transmissão, como o interesse é sobre o instante do transitório, as tensões trifásicas foram medidas meio ciclo pré e meio ciclo pós falta, totalizando 128 pontos (um ciclo), não permitindo que houvessem componentes inter-harmônicos. Por meio da análise de Fourier realizada nos mesmos moldes que dos sinais sintéticos, foi verificada a presença de todo o espectro de frequências, de forma que este foi totalmente utilizado para a obtenção das instâncias no caso da LT.

Para extrair as instâncias que representam $h(t)$, utilizam-se as informações das colunas e linhas da matriz da TSD. Neste caso, as colunas contêm informações temporais, das quais se extraem os valores máximo (T_{max}), mínimo (T_{min}), médio (T_{mean}), desvio padrão (T_{std}) e valor quadrático médio (T_{rms}). As linhas, por sua vez, contêm informações espectrais, das quais se extraem os valores máximo (F_{max}), mínimo (F_{min}), médio (F_{mean}), desvio padrão (F_{std}) e valor quadrático médio (F_{rms}). Além dessas extrações, calcula-se, para cada conjunto, a soma ($M+m$) e a diferença ($M-m$) entre os valores máximos e mínimos, além dos valores médio (mean), desvio padrão (std) e quadrático médio (RMS) entre todos os valores. A Tabela II apresenta a organização das instâncias extraídas para cada DQEE, a qual lista um conjunto identificado por STA, que diz respeito à própria matriz da TSD.

TABLE II. CONJUNTO DE INSTÂNCIAS EXTRAÍDAS PELA TSD

	$M + m$	$M - m$	mean	std	RMS
STA	F1	F2	F3	F4	F5
T_{max}	F6	F7	F8	F9	F10
T_{min}	F11	F12	F13	F14	F15
T_{mean}	F16	F17		F18	F19
T_{std}	F20	F21	F22	F23	F24
T_{rms}	F25	F26	F27	F28	
F_{max}	F29	F30	F31	F32	F33
F_{min}	F34	F35	F36	F37	F38
F_{mean}	F39	F40		F41	F42
F_{std}	F43	F44	F45	F46	F47
F_{rms}	F48	F49	F50	F51	

C. Evolução Diferencial Binária

Conforme ilustrado no fluxograma da Figura 1, a execução do algoritmo de EDB inicia-se com a seleção das instâncias do conjunto de treinamento. O processo evolutivo começa com a geração inicial da população, representada pela iteração $G = 1$. Cada indivíduo da população é um vetor binário de dimensão 1×51 , onde cada gene corresponde a uma instância listada na Tabela I. A ativação ou desativação de uma instância é representada pelos valores 0 e 1, respectivamente. Dessa forma, se o i -ésimo gene do vetor for 1, então a instância F_i será utilizada como entrada para a RNA; caso contrário, será descartada. Após a inicialização da população, seguem-se as etapas de mutação e cruzamento, conforme o esquema evolutivo da EDB.

A etapa de seleção consiste na avaliação de cada indivíduo da população, determinando quais instâncias serão utilizadas para treinar a RNA. O fluxograma da Figura 5 ilustra esse processo. Para cada indivíduo, a RNA é treinada utilizando o conjunto de instâncias ativadas pelo vetor binário

correspondente. Em seguida, o conjunto de teste é avaliado imediatamente após o treinamento. O treinamento da RNA é encerrado segundo dois critérios, (1) a quantidade limite de épocas (LE), limitada em 200, e (2) um critério de parada antecipada (*early stopping*, ES), que caso a função de perda no conjunto de teste não diminua após 5 épocas consecutivas, o treinamento é interrompido para evitar *overfitting*. A FO utilizada é a acurácia balanceada média obtida a partir da matriz de confusão da RNA para o conjunto de teste. Para casos em que a quantidade de instâncias fosse maior do que 25, à FO original foi adicionado um fator de penalização, garantindo que uma quantidade menor de instâncias atinja máximas acurácias. Para evitar problemas de convergência, as instâncias passaram pela normalização z -score, consistindo em subtrair de cada instância a média das instâncias e dividir pelo desvio padrão das instâncias.

Após a avaliação da FO, a população é atualizada. Caso o número máximo de gerações ($G_{max} = 30$) não tenha sido atingido, o contador de gerações é incrementado, e o processo é repetido a partir da etapa de mutação. Se o algoritmo de parada for atingido, o algoritmo é finalizado, e a melhor configuração de instâncias é utilizada para treinar a RNA final, que será avaliada na próxima etapa.

Para a aplicação da metodologia na classificação de faltas na LT, utilizou-se um vetor de 51 instâncias, mas no treinamento da RNA, foram consideradas três vezes esse número de instâncias, garantindo equilíbrio na escolha entre as fases das tensões. Os demais parâmetros da EDB e da RNA foram mantidos conforme a Seção IV-A.

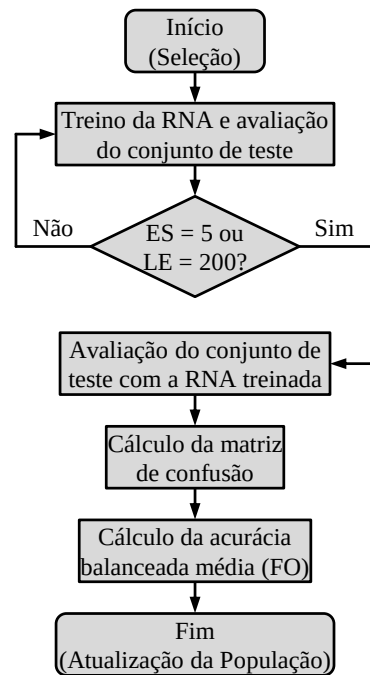


Fig. 5. Fluxograma representativo do processo de seleção da EDB.

D. Análise dos resultados

Após a conclusão do processo evolutivo da EDB, o indivíduo final da última geração define o subconjunto de instâncias selecionadas para a RNA definitiva. Com base nessa configuração, uma nova RNA é treinada e testada, utilizando exclusivamente as instâncias otimizadas. Para avaliar o desempenho do modelo, são geradas métricas de interesse, incluindo a matriz de confusão e as respectivas curvas de acurácia e de genes no melhor indivíduo em cada

geração. Além disso, a robustez da metodologia é testada ao aplicar os casos com ruído ao modelo treinado, obtendo-se uma segunda matriz de confusão para análise comparativa.

IV. RESULTADOS

Esta seção apresenta os resultados da metodologia proposta. A Seção IV-A aborda a classificação de DQEEs em sinais sintéticos, enquanto a Seção IV-II apresenta a sua aplicação na classificação de faltas em um LT, validando a sua eficácia em um sistema real.

A. Avaliação da metodologia proposta para a classificação de DQEEs

Após a geração dos conjuntos de treinamento e teste, a EDB foi aplicada com os seguintes parâmetros: $D = 51$; $NP = D$ (diferente de [9], entre 5D e 10D, devido ao esforço computacional para treinar/testar cada RNA); taxa de mutação F (definida aleatoriamente a cada iteração); $CR = 0.8$; e número máximo de gerações, $G_{max} = 30$. A RNA foi configurada com 25 neurônios em uma única camada oculta com função de ativação *ReLU*, e uma camada de saída com 9 neurônios e função de ativação *Softmax*. O tamanho da camada de entrada variou conforme o indivíduo da população gerado pela EDB. Para garantir diversidade na EDB, a semente de valores aleatórios do Matlab foi configurada como *shuffle*, evitando repetições. Já nos treinamentos das RNAs, a semente foi mantida em *default* para garantir a reprodutibilidade dos pesos sinápticos e vieses.

As Figuras 6 e 7 resumem o desempenho da EDB. A Figura 6 mostra a evolução da FO e o número de instâncias por geração. Até a 18ª geração, a RNA utilizava 23 instâncias com 98% de acerto. A partir da 22ª geração, apenas 18 instâncias foram selecionadas, mantendo uma acurácia de 98.22%, evidenciando que um maior número de instâncias não implica necessariamente em melhor desempenho. A Figura 7 apresenta a matriz de confusão da RNA treinada com as 18 instâncias selecionadas pela EDB. As classes mais desafiadoras foram afundamento, elevação e interrupção, possivelmente devido às semelhanças nas suas leis de formação, diferenciadas apenas pelos valores das constantes.

As instâncias selecionadas, conforme a Tabela II, foram: F1, F5, F6, F7, F17, F18, F25, F27, F28, F29, F30, F34, F35, F37, F41, F42, F44 e F50. Dessas, duas pertencem ao conjunto STA, sete são associadas à T e nove à F, sugerindo que as características espectrais têm maior relevância na classificação de DQEEs. A RNA treinada com essas instâncias foi testada também com sinais contendo ruído (SNR=40dB), conforme ilustrado na Figura 8. Apesar da melhora na identificação das classes C3 e C4, observou-se um aumento nos erros para a classe C5 e uma queda significativa na acurácia da classe C2, evidenciando a sensibilidade do modelo ao ruído para distúrbios específicos. Como já observado, C2 e C4 diferem apenas pela magnitude de seus limiares, de modo que variações nos sinais podem deslocar as respectivas instâncias da região relativa à classe original do DQEE.

B. Avaliação da metodologia proposta para a classificação de faltas em linhas de transmissão

Os resultados obtidos são apresentados nas Figuras 9, 10 e 11. A Figura 9 indica que com apenas 18 instâncias, a RNA atingiu 99.94% de acerto, superando configurações com instâncias anteriores. As matrizes de confusão para os sinais sem ruído e com SNR = 40 dB, respectivamente apresentadas

nas Figuras 10 e 11, mostram que o acréscimo de ruído resultou em maior confusão entre as faltas ABT e ABC, enquanto as acurácias de BCT e CAT permaneceram com os mesmos valores, indicando confusões entre a presença ou não do terra nas faltas envolvendo as fases BC e CA. Os demais tipos de faltas mantiveram um desempenho de 100% de acerto.

As instâncias escolhidas para a classificação das faltas foram: F3, F4, F5, F7, F8, F9, F19, F21, F22, F23, F26, F29, F31, F34, F35, F37, F47 e F49. Somente a critério de visualização, as instâncias F8, T_{max}^{mean} , dos casos de teste são plotadas na Figura 12, indicando a distribuição tridimensional destas informações, e indicando regiões de sobreposição entre as faltas bifásicas e bifásicas-terra. Vale destacar que apesar das faltas trifásicas darem a impressão de serem somente um ponto, as instâncias ficaram concentradas em uma região muito próxima após a normalização *z-score*. Apesar disso, a distribuição tridimensional destas instâncias reflete a capacidade da EDB em procurar os conjuntos de instâncias mais adequados para a classificação das faltas.

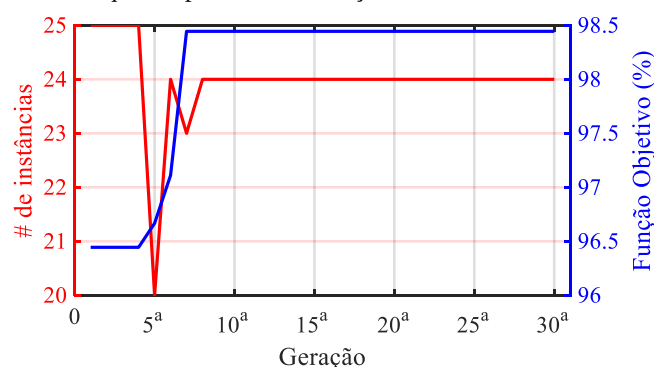


Fig. 6. Número de instâncias e respectivos valores da FO ao longo das gerações (DQEEs).

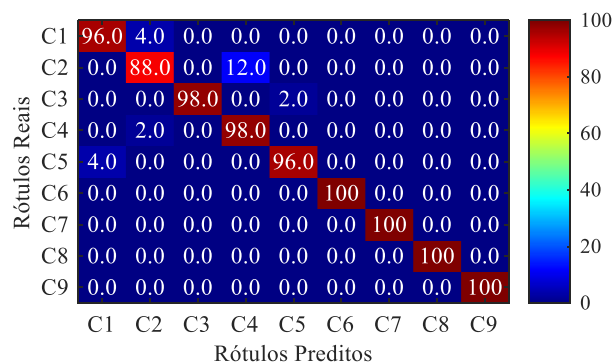


Fig. 7. Matriz de confusão - conjunto de teste (DQEEs) sem ruído (%).

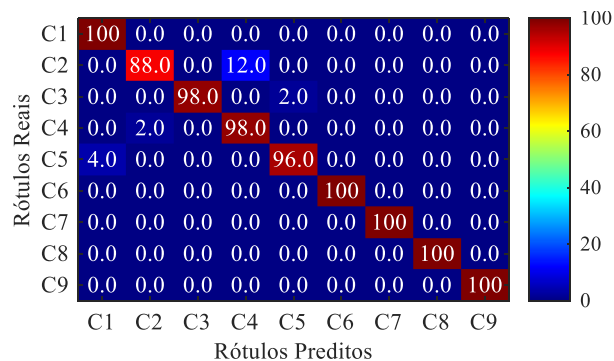


Fig. 8. Matriz de confusão - conjunto de teste (DQEEs) com SNR = 40 dB (%).

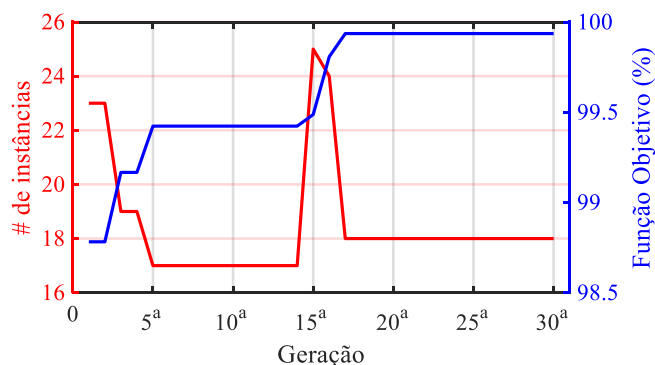


Fig. 9. Número de instâncias e respectivos valores da FO ao longo das gerações (LT).

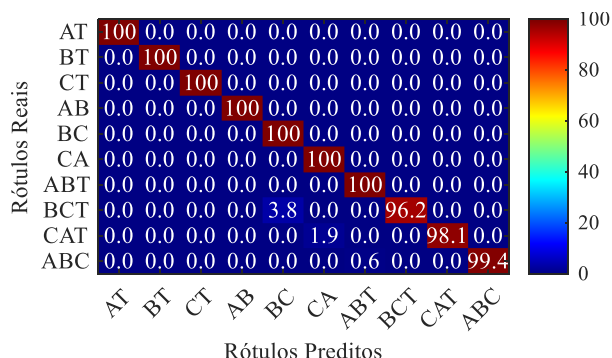


Fig. 10. Matriz de confusão - conjunto de teste (LT) sem ruído (%).

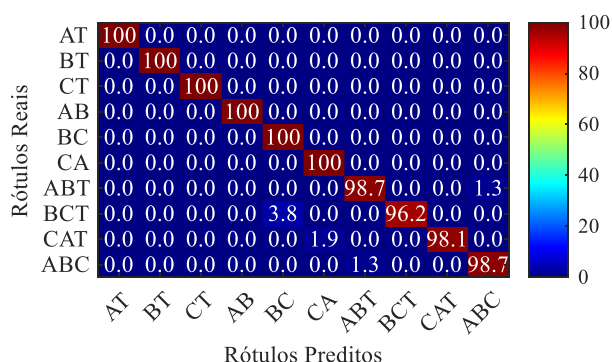


Fig. 11. Matriz de confusão - conjunto de teste (LT) com SNR = 40 dB (%).

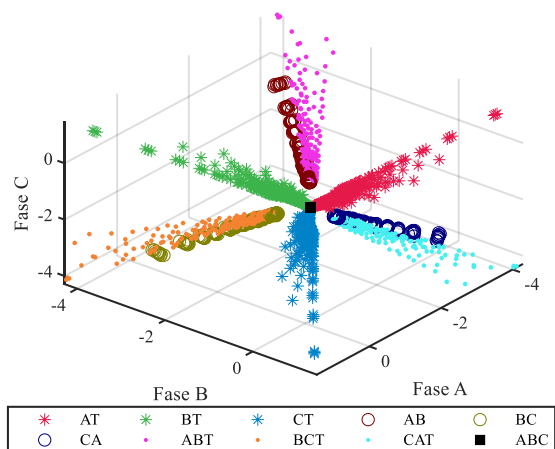


Fig. 12. Representação tridimensional das instâncias F8 (trifásicas) dos casos de teste.

V. CONCLUSÕES

O presente trabalho propôs uma metodologia baseada na EDB para a otimizar a seleção de instâncias na classificação de DQEEs. A TSD foi empregada na extração de instâncias, enquanto as RNAs foram utilizadas como classificadores. Os resultados demonstram a eficácia da abordagem proposta, alcançando uma taxa de acerto de 98.22% para a classificação de DQEEs e 99.94% para a classificação de faltas em LTs. Apesar de a metodologia ter se mostrado eficiente, algumas melhorias podem ser exploradas em trabalhos futuros. Como a abordagem prioriza ferramentas computacionalmente simples, a otimização multiobjetivo foi descartada. No entanto, pode-se investigar a seleção não apenas das instâncias de entrada, mas também da quantidade de camadas na RNA, do número de neurônios por camada e das funções de ativação utilizadas. Uma limitação relevante das RNAs aplicadas neste estudo é a fixação dos pesos sinápticos para um determinado conjunto de instâncias, o que pode impactar a variabilidade dos resultados. Para mitigar este efeito, uma alternativa seria a inicialização aleatória dos pesos, repetindo múltiplas vezes o treinamento e teste da rede. Dessa forma, seria possível analisar a variação das acurácias e seus parâmetros estatísticos, identificando a melhor configuração de topologia e conjunto de instâncias para cada problema de classificação.

AGRADECIMENTOS

Os autores gostariam de agradecer ao Laboratório de Sistemas de Energia Elétrica (LSEE), da Escola de Engenharia de São Carlos (EESC - USP) pela infraestrutura proporcionada, assim como a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo suporte financeiro disponibilizado.

REFERENCES

- [1] J. Li, Z. Teng, Q. Tang and J. Song, "Detection and Classification of Power Quality Disturbances Using Double Resolution S-Transform and DAG-SVMs," in IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 10, pp. 2302-2312, Oct. 2016, doi: 10.1109/TIM.2016.2578518.
- [2] "IEEE Recommended Practice for Monitoring Electric Power Quality," in IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009), vol., no., pp.1-98, 13 Aug. 2019, doi: 10.1109/IEEESTD.2019.8796486.
- [3] S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.
- [4] M. Mishra, "Power quality disturbance detection and classification using signal processing and soft computing techniques: A comprehensive review," *International Transactions on Electrical Energy Systems*, vol. 29, no. 8, p. e12008, 2019. DOI: 10.1002/2050-7038.12008.
- [5] C. -Y. Lee and Y. -X. Shen, "Optimal Feature Selection for Power-Quality Disturbances Classification," in IEEE Transactions on Power Delivery, vol. 26, no. 4, pp. 2342-2351, Oct. 2011, doi: 10.1109/TPWRD.2011.2149547.
- [6] U. Singh and S. N. Singh, "Optimal Feature Selection via NSGA-II for Power Quality Disturbances Classification," in IEEE Transactions on Industrial Informatics, vol. 14, no. 7, pp. 2994-3002, July 2018, doi: 10.1109/TII.2017.2773475.
- [7] U. Singh and S. N. Singh, "A new optimal feature selection scheme for classification of power quality disturbances based on ant colony framework," in Applied Soft Computing, vol. 74, pp. 216-225, 2019, doi: https://doi.org/10.1016/j.asoc.2018.10.017.
- [8] R. G. Stockwell, L. Mansinha and R. P. Lowe, "Localization of the complex spectrum: the S transform," in IEEE Transactions on Signal Processing, vol. 44, no. 4, pp. 998-1001, April 1996, doi: 10.1109/78.492555.
- [9] R. Storn and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," in Journal of Global Optimization, vol. 11, pp. 341-359, 1997.

- [9] C. Deng, B. Zhao, Y. Yang and A. Deng, "Novel binary Differential Evolution without scale factor F," Third International Workshop on Advanced Computational Intelligence, Suzhou, China, 2010, pp. 250-253, doi: 10.1109/IWACI.2010.5585113.
- [10] MathWorks, "fitnet: Fit a neural network classifier," MathWorks Documentation, 2025. [Online]. Disponível em: <https://www.mathworks.com/help/stats/fitnet.html>.
- [11] M. J. B. B. Davi, R. de B. Iscuissati, M. Oleskovicz, and F. V. Lopes, "Investigating the potential of Machine Learning for fault location on Inverter-Based Resource interconnection lines: Insights and recommendations," **Electric Power Systems Research**, vol. 231, p. 110366, 2024. doi: 10.1016/j.epsr.2024.110366.
- [12] P. Ray, B. K. Panigrahi, and N. Senroy, "Hybrid methodology for fault distance estimation in series compensated transmission line," **IET Generation, Transmission & Distribution**, vol. 7, no. 5, pp. 431–439, 2013. doi: 10.1049/iet-gtd.2012.0243.
- [13] R. H. G. Tan and V. K. Ramachandaramurthy, "Numerical model framework of power quality events," **European Journal of Scientific Research**, vol. 43, no. 1, pp. 30–47, 2010. ISSN: 1450-202X.