Classificação de Defeitos em Rolamento de Motores Elétricos Utilizando CNN

Caio Marcos Malheiros Alves

*Programa de Pós-Graduação em Controle e Automação*  
*Instituto Federal do Espírito Santo*

## ***Resumo –* motores elétricos são onipresentes em processos industriais e no dia a dia. O diagnóstico de defeitos em tais máquinas é de suma importância. As técnicas mais recentes de diagnóstico utilizam CNN com grande sucesso, porém são implementadas para rodar em servidores off-line. Este trabalho analisa uma proposta de CNN desenvolvida para ser executada em um sistema embarcado in situ, online. A CNN foi testada com sucesso por seus autores utilizando um dataset proprietário deles. Os testes realizados aqui utilizam um dataset diferente, muito utilizado em benchmarks na área, o dataset da CWRU. Os testes realizados demonstram que a CNN proposta também obtém sucesso ao classificar defeitos do dataset CWRU, validando sua estrutura e demonstrando uma boa generalização para ser aplicada em montagens e sistemas diferentes do utilizado em seu desenvolvimento.**

## *I. Introdução*

Motores elétricos, em particular motores de indução, são peças fundamentais de processos industriais, pois são a principal forma de transformar energia elétrica em energia mecânica. O monitoramento de seu funcionamento é necessário para garantir confiabilidade dos processos, reduzir custos com manutenção e evitar acidentes. O monitoramento tem como objetivo identificar estados de funcionamento – correto, defeituoso, saudável, falha etc. - além de identificar qual tipo de falha ou defeito. Essa identificação pode ser feita utilizando de modelagens e simulações ou através da análise de sinais extraídos diretamente da máquina em questão, como temperatura, corrente e tensão elétrica, emissões acústicas e vibrações mecânicas.

Os rolamentos estão entre os mais cruciais elementos destas máquinas rotativas. Eles são responsáveis por dar suporte mecânico ao elemento rotativo, reduzir o coeficiente de atrito do movimento e garantir a precisão da rotação, portanto, uma vez que existe um defeito no rolamento, a estabilidade de todo o sistema é afetada, podendo causar graves acidentes. É estimado que defeitos nos rolamentos são responsáveis por entre 44% e 90% de todas as falhas em motores elétricos.

O diagnóstico e detecção de falhas em rolamentos são amplamente estudados há vários anos utilizando métodos de processamento de sinais e, mais recentemente, utilizando os mais variados métodos de machine learning. Essas análises podem ser utilizadas para extrair diferentes métricas e para diferentes estratégias. É possível, além de detectar defeitos, predizer a ocorrência de falhas antes que aconteçam e, a partir daí, identificar o estado de saúde do equipamento, finalmente estimando sua vida útil restante (RUL – rest of useful life).

As duas principais etapas para estes tipos de diagnóstico são 1) aquisição de dados e 2) processamento e análise dos dados. Tradicionalmente estas etapas são realizadas em separado (diagnóstico off line), onde primeiro é feita a aquisição dos dados in situ, diretamente no motor em questão, para depois, em um computador remoto, serem feitas todas as etapas de tratamento dos dados, processamento dos sinais e finalmente diagnóstico. Neste cenário, tem ocorrido um grande desenvolvimento de diferentes métodos inteligentes de diagnóstico. Se encontra na literatura hoje uma gama de algoritmos, redes neurais artificiais, redes convolucionais, redes profundas para detecção e predição de defeitos utilizando sinais de vibração de motores elétricos. Esses métodos têm atingido altos níveis de precisão na identificação de defeitos em rolamentos e na previsão de RUL. Estes avanços, porém, trazem também um custo computacional cada vez mais elevado, para ganhos cada vez mais marginais em precisão.

Porém, em um diferente contexto, se tratando de IIoT (industrial internet of things) e sistemas embarcados, microcontrolados e SBC (single board computers), também houve significantes avanços em conectividade, eficiência energética e capacidade de processamento e de memória. Equipamentos conectados já são uma realidade e cada vez mais se fala em “edge intelligence”. Mesmo assim, são raros os trabalhos e métodos de diagnóstico desenvolvidos para tais sistemas insitu e online, dada sua relativa limitação de recursos computacionais.

Neste trabalho será avaliada uma proposta de rede neural convolucional (CNN) desenvolvida para ser utilizada em dispositivos restritos, como um SBC ou um sistema microcontrolado. A avaliação será feita utilizando o dataset da Case Western Reserve University (CWRU), uma vez que a CNN foi inicialmente desenvolvida e estudada utilizando um dataset elaborado pelos autores originais. Essa validação pretende demonstrar a viabilidade e capacidade de generalização da CNN em questão, ao testá-la utilizando um dataset que é tido como um dos datasets fundamentais para validar performance de diferentes algoritmos.

## *II. O dataset CWRU*

A montagem utilizada para a aquisição do dataset CWRU é mostrada na figura 1. Nela, se tem um motor de 2 hp a esquerda, um transdutor/encoder de torque ao centro e um dinamômetro acoplado a direita.

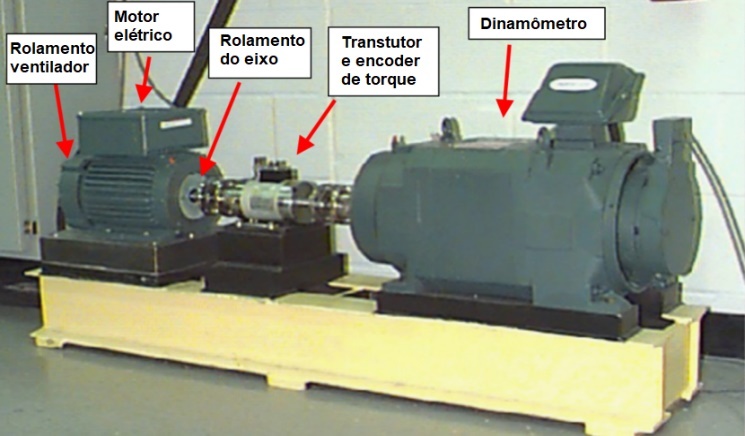


Figura 1 - Montagem para aquisição do CWRU dataset.

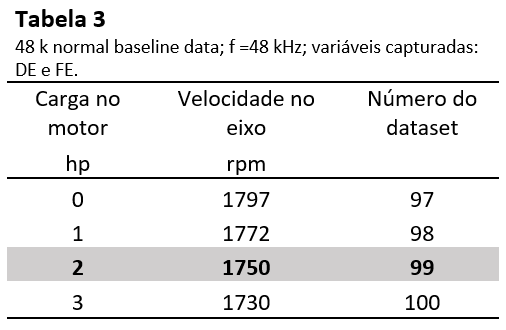
Foram introduzidos defeitos pontuais nos rolamentos com diâmetros de 7 mils (0,18 mm), 14 mils (0,36 mm), 21 mils (0,53 mm) e 28 mils (0,71 mm), localizadas na esfera, na parte interna do rolamento (RI) e na parte externa do rolamento (RE) em três posições diferentes. Dados de vibração foram coletados para cargas de 0 a 3 hp e velocidades de rotação de 1720 a 1797 rpm, utilizando dois acelerômetros tanto no lado do eixo do motor quanto no lado do ventilador do motor, e utilizando duas frequências de amostragem, 48 kHz e 12 kHz. O dataset é público e disponível no CWRU data center - <https://csegroups.case.edu/bearingdatacenter/home>.

As tabelas 1 a 3 mostram todos os experimentos e conjuntos de dados disponíveis no dataset CWRU. Nota-se uma grande variedade de dados, e em destaque estão os dados utilizados neste trabalho. Foi escolhido o defeito de 21 mils (0,53 mm) pois presume-se que apresente uma característica de mais fácil identificação do que defeitos menores, além disso todos os dados foram coletados (todos defeitos – RI, as três posições de RE e esfera) e estão disponíveis no dataset. A velocidade de 1750 rpm foi escolhida pois é a velocidade de rotação nominal do motor. Ao final, o total de classes a serem classificadas 6 – *Normal*, *RI*, *Esfera*, *RE ortogonal*, *RE centralizado* e *RE oposto*.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Tabela 1** |  |  |  |  |  |  |  |
| *"12 k drive end bearing fault data"*; f =12 kHz; variáveis capturadas: DE, FE e BA (falhas de 0,007" a 0,021"); DE (falha de 0,0028"). | | | | | | | |
| Dim. Falha | Carga no motor | Velocidade no eixo | Número do dataset | | | | |
| mm | hp | rpm | RI | Esfera | RE centralizado | RE ortogonal | RE oposto |
| 0,18 | 0 | 1797 | 105 | 118 | 130 | 144 | 156 |
|  | 1 | 1772 | 106 | 119 | 131 | 145 | 158 |
|  | 2 | 1750 | 107 | 120 | 132 | 146 | 159 |
|  | 3 | 1730 | 108 | 121 | 133 | 147 | 160 |
| 0,36 | 0 | 1797 | 169 | 185 | 197 | - | - |
|  | 1 | 1772 | 170 | 186 | 198 | - | - |
|  | 2 | 1750 | 171 | 187 | 199 | - | - |
|  | 3 | 1730 | 172 | 188 | 200 | - | - |
| **0,53** | 0 | 1797 | 209 | 222 | 234 | 246 | 258 |
|  | 1 | 1772 | 210 | 223 | 235 | 247 | 259 |
|  | **2** | **1750** | **211** | **224** | **236** | **248** | **260** |
|  | 3 | 1730 | 212 | 225 | 237 | 249 | 261 |
| 0,71 | 0 | 1797 | 3001 | 3005 | - | - | - |
|  | 1 | 1772 | 3002 | 3006 | - | - | - |
|  | 2 | 1750 | 3003 | 3007 | - | - | - |
|  | 3 | 1730 | 3004 | 3008 | - | - | - |

Cada conjunto de dados numerados do dataset CWRU corresponde a uma leitura de cerca de 10 segundos, portanto os conjuntos têm cerca de 120000 ou 480000 pontos (dependendo se foram coletados utilizando amostragem de 12 kHz ou 48 kHz) para cada sensor que foi utilizado. Por exemplo, o conjunto 236 possui 122281 pontos de aceleração para cada sensor utilizado neste teste em questão, ou seja, 122281 pontos para DE (lado do eixo), 122281 pontos para FE (lado do ventilador) e 122281 pontos para BA (base). Neste trabalho, foram utilizados apenas os dados relativos ao sensor posicionado no lado do eixo (DE) de cada conjunto.

Apenas como ilustração, apresento nas figuras 3 a 5 os gráficos dos valores de aceleração relativos aos conjuntos de dados 99, 211 e 248 respectivamente.



A close up of a device

Description automatically generated

Figura - 1s de amostras do conjunto 99

A picture containing object

Description automatically generated

Figura - 1s de amostras do conjunto 248

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Tabela 2** |  |  |  |  |  |  |  |
| *"48 k drive end bearing fault data"*; f =48 kHz; variáveis capturadas: DE e FE. | | | | | | | |
| Dim. Falha | Carga no motor | Velocidade no eixo | Número do dataset | | | | |
| mm | hp | rpm | RI | Esfera | RE centralizado | RE ortogonal | RE oposto |
| 0,18 | 0 | 1797 | 109 | 122 | 135 | 148 | 161 |
|  | 1 | 1772 | 110 | 123 | 136 | 149 | 162 |
|  | 2 | 1750 | 11 | 124 | 137 | 150 | 163 |
|  | 3 | 1730 | 112 | 125 | 138 | 151 | 164 |
| 0,36 | 0 | 1797 | 174 | 189 | 201 | - | - |
|  | 1 | 1772 | 175 | 190 | 202 | - | - |
|  | 2 | 1750 | 176 | 191 | 203 | - | - |
|  | 3 | 1730 | 177 | 192 | 204 | - | - |
| **0,53** | 0 | 1797 | 213 | 226 | 238 | 250 | 262 |
|  | 1 | 1772 | 214 | 227 | 239 | 251 | 263 |
|  | **2** | **1750** | **215** | **228** | **240** | **252** | **264** |
|  | 3 | 1730 | 217 | 229 | 241 | 253 | 265 |

A close up of a beach

Description automatically generated

Figura - 1s de amostras do conjunto 211

## *III. Preparação dos Dados de Vibração*

Redes neurais convolucionais normalmente têm como entrada imagens em duas dimensões (2D), porém, sinais de vibração são unidimensionais (1D). Por isso é necessário tratar estes sinais antes de utilizá-los no treinamento e classificação com a CNN.

De acordo com o trabalho onde a CNN utilizada é apresentada, o espectro logarítmico de base dois (B2LS) do sinal de vibração é utilizado para construir a imagem de entrada da CNN. Se o sinal original fosse utilizado, a imagem seria dominada pelo componente de frequência de maior potência, dificultando o reconhecimento das componentes de frequência de menor potência. Assim, o B2LS diminui as diferenças de amplitude entre os diferentes componentes de frequência, melhorando o contraste e acurácia da classificação.

Sendo assim, supondo que a imagem de entrada da CNN tenha uma resolução de *M x N*, o sinal de vibração correspondente é dado por

O B2LS de é obtido por

Onde é a transformada rápida de Fourier. Considerando que é simétrico, a imagem em níveis de cinza é construída da forma

Onde é o maior valor entre todos componentes da B2LS.

## *IV. As Redes Neurais Convolucionais*

Dois modelos similares de CNN são testadas neste trabalho. Ambas foram propostas como modelos apropriados para serem utilizados em sistemas restritos, portanto são relativamente simples, com poucas camadas e parâmetros. A duas foram testadas com sucesso utilizando um dataset proprietário dos autores que propuseram tais redes convolucionais. As duas serão testadas utilizando o dataset CWRU, com o objetivo de demonstrar a capacidade de generalização e sua validade ao classificar dados similares, porém não exatamente os mesmos para os quais foram projetadas inicialmente.

As tabelas 4 e 5 mostram as estruturas dos dois modelos de CNN que serão testados.

|  |  |
| --- | --- |
| **Tabela 4** |  |
| Parâmetros CNN modelo 1 | |
| Camada | Parâmetros |
| 1a convolução | Kernel: 3x3; Qtd. Kernel: 16 |
| 1a max-pooling | Kernel: 3x3; Stride: 2x2 |
| 2a convolução | Kernel: 3x3; Qtd. Kernel: 32 |
| 2a max-pooling | Kernel: 3x3; Stride: 2x2 |
| Densa | 128 |
| Saída | 6 |

|  |  |
| --- | --- |
| **Tabela 5** |  |
| Parâmetros CNN modelo 2 | |
| Camada | Parâmetros |
| 1a convolução | Kernel: 2x2; Qtd. Kernel: 2 |
| 1a max-pooling | Kernel: 3x3; Stride: 2x2 |
| Densa | 8 |
| Saída | 6 |

As funções de ativação utilizadas são a “relu” nas camadas convolucionais e densas, e a “softmax” nas camadas de saída.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Tabela 6** |  |  |  |  |  |  |
| *CNN modelo 1* | | | | | |  |
| Teste | tam. conjunto de dados | tam. Imagem (px.) | pontos por imagem | imagens por classe | tempo por imagem (s) | acurácia |
| 1 | 120000 | 28x28 | 1568 | 76 | 0,13 | - |
| 2 | 2x120000 | 28x28 | 1568 | 153 | 0,13 | 100% |
| 3 | 120000 | 14x14 | 392 | 306 | 0,033 | - |
| 4 | 2x120000 | 14x14 | 392 | 612 | 0,033 | 98,9% |
| 5 | 120000 | 50x50 | 5000 | 24 | 0,42 | - |
| 6 | 2x120000 | 50x50 | 5000 | 48 | 0,42 | 100% |
| 7 | 480000 | 28x28 | 1568 | 306 | 0,033 | 63,1% |
| 8 | 2x480000 | 28x28 | 1568 | 612 | 0,033 | 99,3% |
| 9 | 480000 | 14x14 | 392 | 1224 | 0,00817 | - |
| 10 | 2x480000 | 14x14 | 392 | 2448 | 0,00817 | - |
| 11 | 480000 | 50x50 | 5000 | 96 | 0,104 | 68% |
| 12 | 2x480000 | 50x50 | 5000 | 192 | 0,104 | 100% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Tabela 7** |  |  |  |  |  |  |
| *CNN modelo 2* | | | | | | |
| Teste | tam. conjunto de dados | tam. Imagem (px.) | pontos por imagem | imagens por classe | tempo por imagem (s) | acurácia |
| 13 | 120000 | 28x28 | 1568 | 76 | 0,13 | - |
| 14 | 2x120000 | 28x28 | 1568 | 153 | 0,13 | 99,1% |
| 15 | 120000 | 14x14 | 392 | 306 | 0,033 | - |
| 16 | 2x120000 | 14x14 | 392 | 612 | 0,033 | 98,6% |
| 17 | 120000 | 50x50 | 5000 | 24 | 0,42 | - |
| 18 | 2x120000 | 50x50 | 5000 | 48 | 0,42 | - |
| 19 | 480000 | 28x28 | 1568 | 306 | 0,033 | - |
| 20 | 2x480000 | 28x28 | 1568 | 612 | 0,033 | - |
| 21 | 480000 | 14x14 | 392 | 1224 | 0,00817 | 36,5% |
| 22 | 2x480000 | 14x14 | 392 | 2448 | 0,00817 | 93,9% |
| 23 | 480000 | 50x50 | 5000 | 96 | 0,104 | - |
| 24 | 2x480000 | 50x50 | 5000 | 192 | 0,104 | - |

## V. Testes e Resultados

Para treinar e finalmente utilizar a CNN para classificar os sinais, é preciso construir um banco de imagens representando cada sinal. Essas imagens, construídas a partis do B2LS do sinal como mostrado anteriormente, devem ser grandes o bastante para carregarem em si informações suficientes para inferência, mas também devem ter um tamanho tal que as CNN relativamente simples sejam capazes de processar. Outro fator relevante neste trabalho é a quantidade de amostras que se tem, quantidade essa limitada pelo dataset utilizado, que foge do controle do autor, portanto, imagens maiores que carregam mais informação acarreta naturalmente numa menor quantidade de imagens que se pode construir utilizando os dados disponíveis.

Cada imagem, isto é, cada amostra utilizada para a CNN treinar/inferir, utiliza MxNx2 pontos de amostras de cada conjunto de dados. Por exemplo, uma imagem 28x28 utiliza 1568 pontos (ver seção III). Em um conjunto de dados de duração de 10 segundos, capturados a 12 kHz, se tem cerca de 120000 pontos, nos dando 76 imagens, o que é um valor pequeno para treinar e validar uma CNN, como será mostrado nos testes. A alternativa para se obter uma maior quantidade de imagens é diminuir a resolução, por exemplo utilizando imagens 14x14, o que permite construir 306 imagens para treinamento/validação da CNN. Porém como comentado, uma imagem menor carrega menos informações sobre o sinal original, também dificultando o treinamento da rede. Para contornar tal problema, foram feitos testes onde se utiliza os conjuntos de dados de 48 kHz e 12 kHz juntos, igualando a taxa de amostragem e juntando os sinais, dobrando a quantidade de imagens que podem ser construídas. A figura 6 mostra alguns exemplos de imagens construídas a partir do sinal de vibração e que são utilizadas como entrada para a CNN.

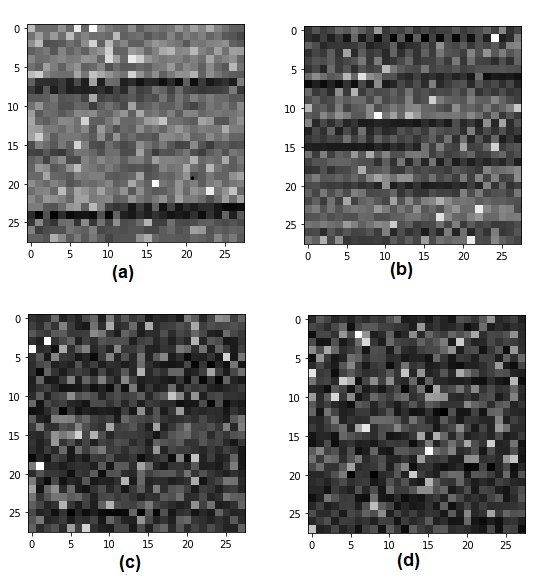


Figura - Exemplos de imagens geradas a partir dos conjuntos de dados. (a) conjunto 260, (b) conjunto 211, (c) conjunto 224 e (d) conjunto 99

Com isso em mente, as tabelas 6 e 7 resumem os testes elaborados, os parâmetros relacionados a cada teste e a acurácia de classificação de cada um. Testes sem valor de acurácia exibiram valores menores que 20% e foram desconsiderados. Nota-se que foram realizados testes variando parâmetros que influenciam na quantidade de imagens disponíveis para treinamento e validação e na quantidade de informação que elas carregam, que depende do seu tamanho e do tempo em segundos que elas correspondem no sinal de vibração. Este balanço – informação x quantidade de imagens – aparenta ser crucial no sucesso da CNN.

Além desses parâmetros exibidos nas tabelas, existem outros que poderiam ser testados, principalmente em relação ao treinamento da CNN, como por exemplo tamanho do batch e número de épocas, porém, para efeito de comparação e para limitar o escopo deste trabalho, todos os treinamentos foram realizados com batch de 10 e 100 épocas. Estes valores se apresentaram adequados em várias oportunidades, como se vê nas tabelas 6 e 7, portanto pode-se admitir que são valores razoáveis para testes.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Tabela 8** | | |  |  |  |  |  |
| Matriz confusão teste 11 | | | | | | | |
|  |  | Condição prevista | | | | | |
|  |  | 0 | 1 | 2 | 3 | 4 | 5 |
| Condição real | 0 | 29 | 0 | 0 | 0 | 0 | 0 |
| 1 | 2 | 13 | 0 | 0 | 7 | 0 |
| 2 | 16 | 0 | 6 | 5 | 0 | 0 |
| 3 | 6 | 0 | 4 | 8 | 0 | 0 |
| 4 | 3 | 10 | 0 | 0 | 11 | 1 |
| 5 | 0 | 0 | 0 | 0 | 4 | 19 |

Além das acurácias exibidas nas tabelas, temos como exemplo na tabela 8 a matriz de confusão para o teste 11. Outras matrizes não são relevantes, pois ou a rede teve sucesso total em classificar, ou não obteve sucesso algum, ou obteve sucesso parcial, porém não prático, como no teste 11.

## VI. Conclusões

O resultado mais importante e relevante que se pode tirar dos testes realizados neste trabalho é que é possível classificar sinais de vibração relativos a falhas no rolamento de motores elétricos utilizando redes neurais convolucionais bastante simples. Resultados como os do teste 22, 16 e 2 mostram que dado o correto equilíbrio entre número de amostras e tamanho das amostras, é possível classificar com alta acurácia tais sinais. Mais ainda, as CNN utilizadas para os testes neste trabalho foram inicialmente desenvolvidas utilizando um dataset completamente diferente do utilizado aqui, o que demonstra também que além de simples, estas redes não estão limitadas a serem utilizadas em montagens específicas e podem ser generalizadas para diferentes equipamentos.

O desenvolvimento de redes neurais e técnicas de machine learning para implementação em sistemas embarcados deve continuar, dada a tendência de rápido crescimento do setor de inteligência das coisas. Este trabalho mostra algumas direções interessantes, como por exemplo, o mínimo de informação necessária em uma amostra para que se possa inferir uma classe a partir dela para que ao mesmo tempo se trabalhe com informações relevantes e com redes pequenas. Seria interessante também ver mais estudos aplicando sistemas de classificação já desenvolvidos em datasets diferentes dos que foram utilizados para desenvolvê-los.

## Referências

S. LU, G. Qian, Q. He et al, Insitu Motor Fault Diagnosis Using Enhanced Convolutional Neural Network in an Embedded System, IEEE Sensors Journal, 2019.

S. Zhang, S. Zhang, B. Wang, et al, Machine Learning and Deep Learning Algorithms for Bearing Fault Diagnostics – A Comprehensive Review, arXiv, 2019.

W.A. Smith, R.B. Randall, Rolling Element Bearing Diagnostics Using the Case Western Reserve University Data: A Benchmark Study, Mech. Syst. Signal Process., 2015.