

# Utilização de CNNs e Sistemas Embarcados para Classificação de Defeitos em Rolamento de Motores Elétricos

Caio Marcos M. Alves, Luiz A. Pinto, Marco Antônio de S. L. Cuadros, Gustavo M. de Almeida

**Resumo** As redes de transdutores inteligentes são utilizadas em muitas áreas para monitoramento e automação do ambiente. O padrão IEEE 1451, através de diretrizes de hardware e software, proporciona aos fabricantes e desenvolvedores uma referência para o desenvolvimento de componentes para esse tipo de rede, além de facilitar o acesso e reconhecimento automático (*plug-and-play*) de transdutores em várias configurações físicas. Neste artigo, descreve a configuração de um dispositivo embarcado com um sistema para web de reconhecimento para uma rede de transdutores inteligentes baseada no padrão IEEE 1451. O sistema foi desenvolvido com foco no usuário possibilitando uma exibição amigável das informações dos nós presentes na rede. A realização dos testes foram baseadas em módulos TIM's no qual possui os TEDS descrita na literatura.

## 1 Introdução

Motores elétricos estão sujeitos a diversos tipos de falhas, sendo as mais comuns as associadas aos elementos de rolamento que, de acordo com **13' bonnett' 2008**, correspondem a 69% do total de falhas nesses equipamentos.

Uma vez que falhas em rolamentos alteram o regime de vibração, a aquisição e a análise desses sinais pode fornecer informações importantes sobre o estado de conservação desses componentes. Nesse contexto, diversos trabalhos foram propostos sobre a utilização de Machine Learning para a detecção de falhas em rolamentos de motores elétricos.

---

Caio Marcos M. Alves

Programa de Pós-Graduação em Engenharia de Controle e Automação, Instituto Federal do Espírito Santo (IFES), Mangueiras – Serra, Espírito Santo, Brasil.  
e-mail: caiomarcos@gmail.com

Luiz A. Pinto

Programa de Pós-Graduação em Engenharia de Controle e Automação, Instituto Federal do Espírito Santo (IFES), Mangueiras – Serra, Espírito Santo, Brasil.  
e-mail: pinto.luizalberto@ifes.edu.br

Marco Antônio de S. L. Cuadros

Programa de Pós-Graduação em Engenharia de Controle e Automação, Instituto Federal do Espírito Santo (IFES), Mangueiras – Serra, Espírito Santo, Brasil.  
e-mail: marcoantonio@ifes.edu.br

Gustavo M. de Almeida

Programa de Pós-Graduação em Engenharia de Controle e Automação, Instituto Federal do Espírito Santo (IFES), Mangueiras – Serra, Espírito Santo, Brasil.  
e-mail: gmaia@ifes.edu.br

---

*Anais do XV Encontro Anual de Ciência da Computação (EnAComp 2020)*. ISSN: 2178-6992.

Catalão, Goiás, Brasil. 25 a 27 de Novembro de 2020.

Copyright © autores. Publicado pela Universidade Federal de Catalão.

Este é um artigo de acesso aberto sob a licença CC BY-NC (<http://creativecommons.org/licenses/by-nc/4.0/>).

Uma comparação entre métodos baseados na análise dos sinais de vibração no domínio do tempo, da frequência e tempo-frequência está apresentada em **13'randall'2011**. Em **13'pandarakone'2019** as SVM (Support Vector Machine) foram utilizadas para classificar tipos de arranhões presentes nos elementos dos rolamentos através da análise das componentes de frequência da corrente elétrica. Em **13'tian'2016**, os autores utilizaram Análise de Componentes Principais (Principal Component Analysis - PCA) e curtose espectral para extração de descritores dos sinais de vibração. Para a classificação foi utilizado o vizinho mais próximo (*k-Nearest-Neighbor* - KNN). Em **13'janssens'2016** os autores utilizaram CNNs (*convolutional neural network*, ou redes neurais convolucionais) com apenas uma camada convolucional para diagnosticar falhas em rolamentos, com acurácia e precisão superiores a 93% utilizando. Outros trabalhos, como em **13'wen'2018** utilizam redes baseadas na LeNet-5 para classificar com sucesso o sinal original de vibração.

Os avanços em conectividade, eficiência energética e capacidade de processamento e memória dos hardwares para sistemas embarcados têm encorajado os pesquisadores a desenvolver sistemas embarcados para diagnóstico de falhas on-site. Em **13'gongora'2018**, uma RNA embarcada é desenvolvida para detectar falhas no rolamento através do sinal da corrente elétrica. Em **13'eren'2018** os autores utilizam redes neurais convolucionais compactas para classificar as falhas em rolamentos através de sinais de vibração. Em **13'markiewicz'2019** os autores propõem uma solução embarcada de baixo consumo para classificação de falhas utilizando redes *LSTM* (*Long Short-Term Memory*). Em **13'lu'2019** os autores utilizaram redes neurais convolucionais compactas, embarcadas em sistemas microcontrolados para classificar 7 tipos de falha em motores de indução. Os sinais de vibração foram convertido em imagens e classificados. Segundo os autores, as acurácias alcançadas são de 100%, 97,5% e 94,13% para cada teste, respectivamente.

Esse trabalho tenta demonstrar a viabilidade de se utilizar redes neurais convolucionais simples, em sistemas embarcados microcontrolados, para classificar falhas em rolamentos de motores elétricos através de sinais de vibração. Para os experimentos foi utilizado o conjunto de dados da Case Western Reserve University Bearing Data Center (CWRU), amplamente utilizado como benchmark em aplicações semelhantes (**13'smith'2015**).

O restante do trabalho contém as seguintes seções: a Seção 2 apresenta uma breve descrição das CNNs. Em seguida, na Seção 3, descreve-se o método “Espectro Logarítmico de Base 2”. Na Seção 4 o conjunto de dados CWRU é descrito. Na Seção 5 apresentam-se os dois modelos de CNN que serão avaliados. Na Seção 6 os processos de treinamento e classificação com os resultados obtidos são descritos. Finalmente, na seção 7 estão descritas as conclusões do trabalho.

## 2 Referencial Teórico

As Redes Neurais Convolucionais, constituídas por diversas camadas, são variações das redes perceptron, que utilizam filtros convolucionais para a extração de características das imagens em diferentes etapas de processamento.

Constituída por um conjunto de filtros, a camada convolucional extrai as características da imagem de entrada. Uma CNN possui várias camadas convolucionais que são capazes de extrair informações em níveis, sucessivamente, mais baixos de detalhamento. A camada de Pooling tem

como a principal função a redução da dimensão espacial da imagem de entrada para a camada convolucional seguinte.

A camada de normalização em lote é utilizada para reduzir a faixa de variação dos valores das unidades das camadas ocultas, o que resulta na aceleração do treinamento da rede. A camada totalmente conectada, tem todos os seus neurônios conectados a todos neurônios da camada anterior. Em problemas de classificação, essa camada combina as características para a classificação. A camada de classificação é a camada de saída da rede, que consiste em números positivos cuja soma é igual a 1, que correspondem às probabilidades das classes.

### 3 Espectro Logarítmico de Base 2 (B2LS)

Redes neurais convolucionais, em geral, têm como entrada imagens em duas dimensões, porém, sinais de vibração são unidimensionais, como é o caso dos sinais do conjunto de dados CWRU. Dessa forma, é necessário tratar estes sinais antes de utilizá-los no treinamento e classificação com a CNN.

O B2LS do sinal de vibração é utilizado para construir a imagem de entrada da CNN, pois no sinal original, a imagem é dominada pelo componente de frequência de maior potência, dificultando o reconhecimento das componentes de frequência de menor potência. O B2LS diminui essas diferenças, melhorando o contraste e acurácia da classificação (13' lu' 2019).

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

$$V[n], n = 1, 2, \dots, 2MN \quad (1)$$

O B2LS de  $V[n]$  é obtido por

$$S[i] = \log_2 \left( \frac{FFT(V[n])}{MN} \right), i = 1, 2, \dots, 2MN \quad (2)$$

onde  $FFT(.)$  é a transformada rápida de Fourier. Considerando que  $S[i]$  é simétrico, a imagem  $I$  é construída da forma

$$I = \frac{\begin{bmatrix} S[1] & \dots & S[N] \\ \vdots & \ddots & \vdots \\ S[N(M-1)+1] & \dots & S[MN] \end{bmatrix}}{\max(S[i])}, i = 1, 2, \dots, 2MN \quad (3)$$

onde  $\max(S[i])$  é o maior valor entre todos componentes da B2LS.

### 4 O Conjunto de Dados CWRU

Este conjunto de dados elaborado pela Case Western Reserve University é utilizado como benchmark para modelos de classificação de falhas em rolamentos através de sinais de

vibração. A montagem utilizada para a aquisição dos dados é mostrada na Figura 1 - motor de 2 hp, acoplado a um dinamômetro.

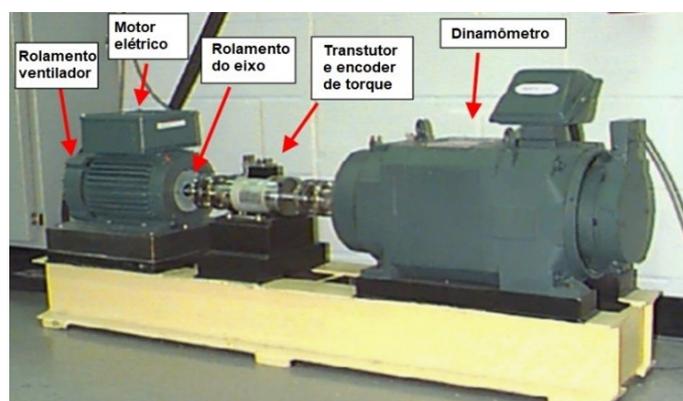


Figura 1: Equipamento utilizado na aquisição de dados do conjunto da CWRU.

Foram introduzidas cinco diferentes falhas nos elementos do rolamento que suportam o eixo, com diâmetros de 0,18 mm, 0,36 mm, 0,53 mm e 0,71 mm, localizados na esfera, na pista interna e em três posições diferentes da pista externa do rolamento. Assim foram coletados sinais de vibração para velocidades de 1797, 1772, 1750 (nominal) e 1730 rpm. As amostras foram coletadas a 48 kHz. Neste trabalho, foram utilizados os dados relativos ao defeito com 0,53 mm, nas 4 condições de carga, com dados coletados no lado do eixo do motor.

Foram coletados sinais para cada velocidades de 1797, 1772, 1750 e 1730 rpm, e para cada posição do defeito (normal, esfera, pista interna, pista externa oposto, pista externa ortogonal, e pista externa centralizado), totalizando 24 sinais.. Para este trabalho, os sinais foram agrupados de acordo com seus defeitos, gerando assim o total de seis classes, que são: RE oposto, RE centralizado, RE ortogonal, RI, Esfera, e Normal (C1, C2, C3, C4, C5 e C6 respectivamente). O gráfico da Figura 2 mostra um dos sinais de vibração do conjunto de dados CWRU.

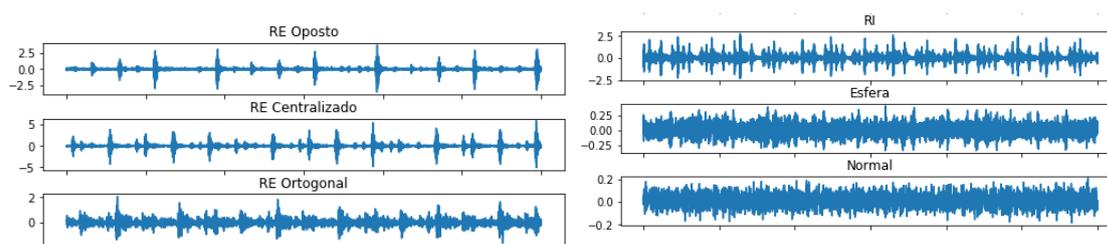


Figura 2: Exemplos de sinais de vibração do CWRU – 1750 rpm, 12 kHz. Fonte: autor.

## 5 CNN Classificadoras

Dois modelos de CNN foram utilizados para classificar os sinais apresentados na Seção 4. Ambos são propostos como modelos apropriados para serem executados em sistemas com

restrições de memória e processamento, portanto são relativamente simples, com poucas camadas e parâmetros. Estes dois modelos foram validados com sucesso usando um conjunto de dados proprietário dos autores que os desenvolveram (13<sup>lu</sup>2019). As tabelas 1 e 2 mostram respectivamente as arquiteturas do Modelo 1 e Modelo 2.

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
Total de parâmetros	71.283

Tabela 1: CNN Modelo 1

Camada	Parâmetros
1a convolução	Kernel: 2x2; Qtd. Kernel: 2
1a max-pooling	Kernel: 3x3; Stride: 2x2
Densa	8
Saída	6
Total de parâmetros	2.776

Tabela 2: CNN Modelo 2

## 6 Treinamento e Classificação

Apesar das CNNs aqui estudadas tenham como objetivo realizar classificação em um sistema microcontrolado, a etapa de treinamento utiliza algoritmos com demanda computacional acima do que um microcontrolador é capaz de entregar, além de não existirem algoritmos prontos e de fácil utilização para realizar treinamento embarcado. Portanto, todo o treinamento foi feito em um desktop. Para tal, o primeiro passo é separar os sinais originais do CWRU, que são compostos por cerca de 480000 pontos (10 segundos de captura a 48 kHz), em amostras menores. Assim, para cada sinal de 480000 pontos, foram criadas amostras de acordo com as equações (1) a (3), com  $M=N=28$ , sem sobreposição. Dessa forma, foi possível criar 234 amostras de 2048 pontos de cada um dos 24 sinais, ou 936 amostras para cada classe. Sobre esse conjunto de amostras (total de 5616), 70% foi separado para treinamento (3931 amostras), enquanto 30% (1685) foi utilizado para teste de classificação, no microcontrolador. A Figura 3 ilustra as etapas de pré-processamento, a partir do sinal original até a imagem utilizada para treinar e classificar.

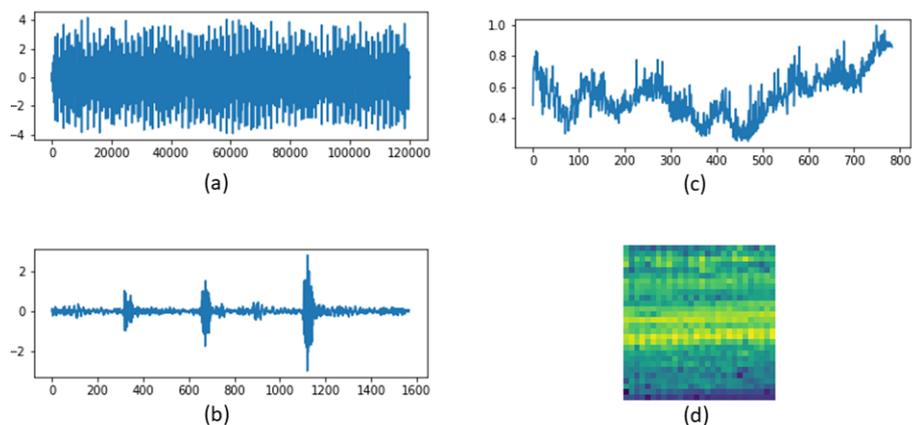


Figura 3: Transformação do sinal original em amostras para classificação. (a) sinal original do CWRU, (b) amostra de 2048 pontos, (c) B2LS da amostra e (d) imagem de 28 x 28 pixels para treinamento e classificação.

## 6.1 Treinamento

Para a execução do treinamento, foram utilizadas 982 amostras de cada classe, em batches de 10 amostras, durante 100 épocas. A Figura 4 mostra a acurácia durante o treinamento da CNN Modelo 1 e CNN Modelo 2. Os dois modelos alcançaram rapidamente altos valores de acurácia durante o treinamento em um computador desktop.

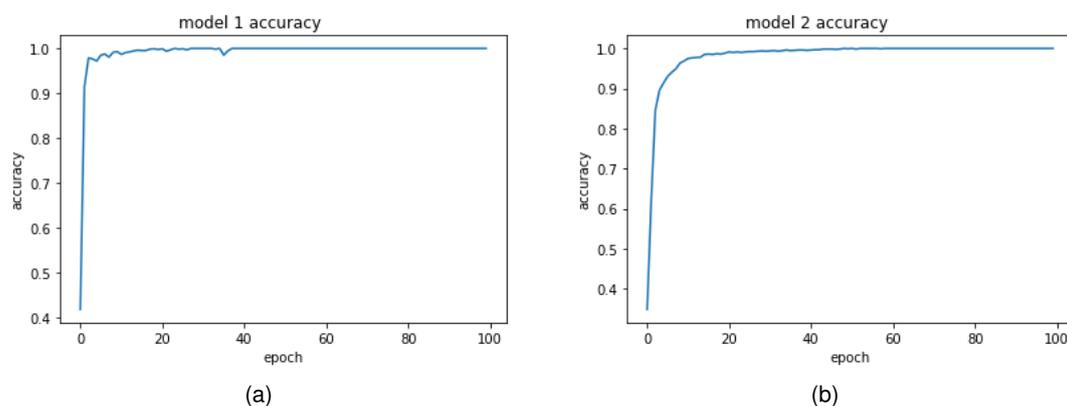


Figura 4: Evolução de acurácia durante treinamento da CNN Modelo 1 (a) e CNN Modelo 2 (b).

## 6.2 Classificação

Para realizar a classificação utilizando as redes apresentadas em um sistema microcontrolado, foi escolhida a placa de desenvolvimento B-L475E-IOT01A da STMicroelectronics, que utiliza o microcontrolador STM32L475VGT6. Este é um microcontrolador Arm Cortex-M4 de ultra-baixo

consumo, alto desempenho com 1 Mbyte de memória flash, 128 Kbyte de memória SRAM, clock principal de 80 MHz e FPU (floating point unit).

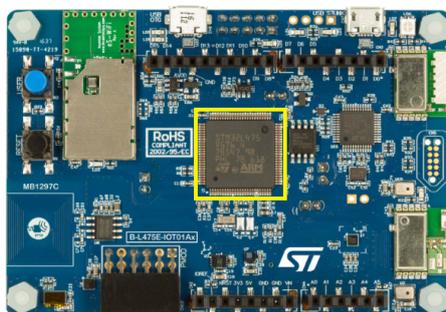


Figura 5: Placa de desenvolvimento B-L475E-IOT01A, com o microcontrolador STM32L475VGT6 em destaque, utilizado para classificação. Fonte: st.com.

Utilizando ferramentas disponibilizadas pela STMicroelectronics, as redes já treinadas foram convertidas para que pudessem ser embarcadas no microcontrolador. Uma vez realizada a conversão, é possível observar algumas métricas (Tabela 3 e Tabela 4) relativas a complexidade computacional (MAC - multiply-accumulate operations) e memória (ROM) demandada pelas redes classificadoras.

Camada	MAC	ROM (bytes)
1a convolução + max-pool	128.830	498
2a convolução + max-pool	468.836	16.182
Densa	65743	229.534
Saída	664	2.738
Total	664.074	284.952

Tabela 3: Complexidade - CNN Modelo 1.

Camada	MAC	ROM (bytes)
1a convolução + max-pool	10.336	44
Densa	2.716	10.849
Saída	145	211
Total	13.184	11.104

Tabela 4: Complexidade - CNN Modelo 2.

Para cada um dos dois modelos de CNN, a rede é embarcada e as amostras separadas para o teste de classificação são enviadas, uma a uma, ao microcontrolador através de comunicação serial. Para cada amostra, são executados no microcontrolador o pré-processamento (B2LS) e classificação, com o resultado desta classificação sendo enviado de volta ao computador desktop. Os tempos necessários para executar o pré-processamento e classificação de cada amostra foram capturados usando um osciloscópio e estão apresentados na Figura 6 e Figura 7.

Como esperado, o tempo necessário para a CNN Modelo 1 executar a classificação de uma amostra é significativamente maior - cerca de vinte vezes - do que o tempo necessário para a

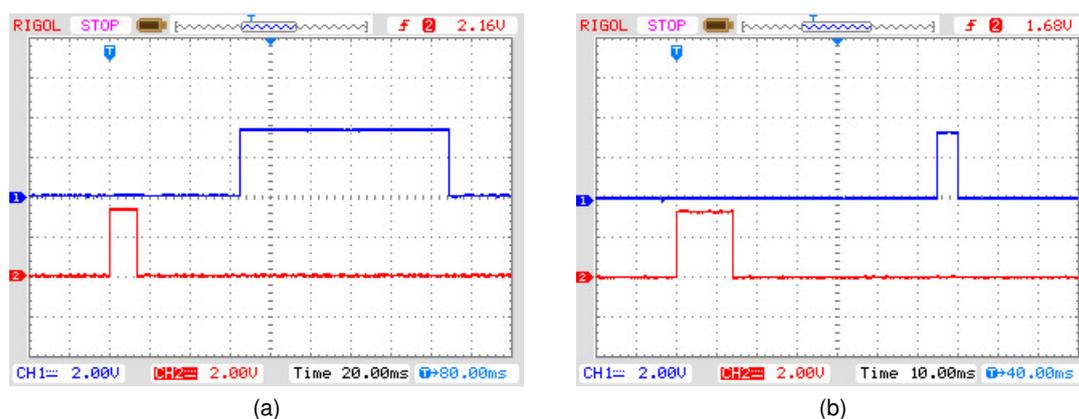


Figura 6: (a) Exemplo de ciclo consistindo de pré-processamento (traço inferior) e classificação (traço superior) para CNN Modelo 1. (b) Tempo para executar o pré-processamento (B2LS) de cada amostra: 14,4 ms.

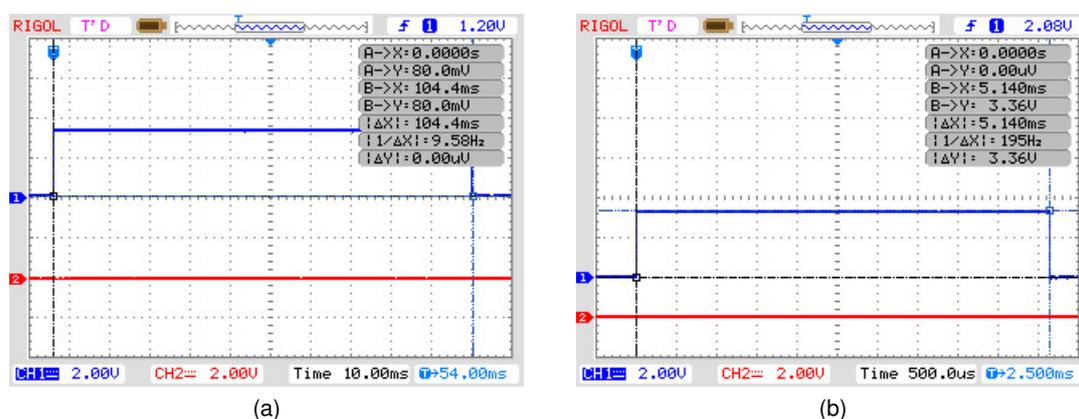


Figura 7: Tempo para classificar cada amostra de 28 x 28 pixel no microcontrolador. (a) CNN Modelo 1: 104,4 ms e (b) CNN Modelo 2: 5,14 ms. Observar escala de 10 ms por divisão em (a) e 0,5 ms em (b).

CNN Modelo 2 realizar a classificação de uma mesma amostra, uma vez que possui 26 vezes mais parâmetros, requisitando do microcontrolador 51 vezes mais operações MAC.

Após classificar todas as 1685 amostras utilizando cada CNN, os resultados foram consolidados nas matrizes de confusão apresentadas na Tabela 6 e na Tabela 7. A partir delas tira-se a acurácia, com resultados maiores que 99%. A Tabela 5 traz a acurácia obtida por outros quatro modelos de CNN que também classificaram o conjunto de dados CWRU, juntamente com os resultados obtidos utilizando a CNN Modelo 1 e CNN Modelo 2, apresentadas neste trabalho. Porém, as redes aqui apresentadas realizaram a classificação em um sistema embarcado microcontrolado, o que é um grande diferencial frente a maioria das soluções propostas até hoje na literatura.

Modelo	Acurácia (%)
CNN Modelo 1	99,94
CNN Modelo 2	99,94
(13'lu'2017)	92,60
(13'shao'2018)	99,15
(13'jia'2018)	99,92
(13'wen'2018)	99,79

Tabela 5: Comparação de acurácia na classificação do conjunto de dados CWRU.

		Condição prevista					
		C1	C2	C3	C4	C5	C6
Condição real	C1	272	0	0	0	0	0
	C2	0	285	0	0	0	0
	C3	0	0	287	0	0	0
	C4	0	0	0	282	0	0
	C5	0	0	0	0	289	0
	C6	0	1	0	0	0	269

Tabela 6: Matriz Confusão para CNN Modelo 1.

		Condição prevista					
		C1	C2	C3	C4	C5	C6
Condição real	C1	272	0	0	0	0	0
	C2	0	283	2	0	0	0
	C3	0	0	287	0	0	0
	C4	0	0	0	275	1	6
	C5	0	0	0	0	289	0
	C6	0	0	0	0	0	270

Tabela 7: Matriz Confusão para CNN Modelo 2.

## 7 Conclusão

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 simples, sendo executadas em sistemas embarcados extremamente restritos. As matrizes confusão mostram que é possível classificar com alta acurácia tais sinais, mesmo com as variações em velocidade, o que é um dificultador. Mais ainda, as CNN utilizadas para os testes neste trabalho foram inicialmente desenvolvidas pelos autores **13'lu'2019** utilizando um conjunto de dados 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 que é de grande relevância para a viabilidade deste tipo de estratégia de monitoramento, uma vez que diminui a necessidade de desenvolvimentos de arquiteturas específicas para cada sistema a ser monitorado, o que é um problema já levantado em **13'randall'2011**, que de acordo com o autor inviabilizaria modelos de aprendizado como solução de detecção de falhas em rolamento de motores.

Uma vez que classificar o CWRU é considerado relativamente simples e um problema já resolvido (**13'zhang'2020**), para o futuro deve-se aprofundar em outras características além da eficácia, como custo computacional do classificador, possibilidade de transferência de aprendizado, otimização da arquitetura da CNN, e aplicação em conjuntos de dados mais complexos, em ambientes reais.