



**SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO E CULTURA
UNIVERSIDADE FEDERAL DE SERGIPE
Pró-Reitoria de Pós-Graduação e Pesquisa
Programa de Pós-Graduação em Engenharia Elétrica**

**APLICAÇÃO DA TÉCNICA BOVW PARA IDENTIFICAR FALHAS
EM MÁQUINAS ROTATIVAS A PARTIR DE IMAGENS
TERMOGRÁFICAS DE BAIXA RESOLUÇÃO**

ELIEL TAVARES LINS

São Cristóvão - SE, Brasil
Dezembro de 2021



**SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO E CULTURA
UNIVERSIDADE FEDERAL DE SERGIPE
Pró-Reitoria de Pós-Graduação e Pesquisa
Programa de Pós-Graduação em Engenharia Elétrica**

**APLICAÇÃO DA TÉCNICA BOVW PARA IDENTIFICAR FALHAS
EM MÁQUINAS ROTATIVAS A PARTIR DE IMAGENS
TERMOGRÁFICAS DE BAIXA RESOLUÇÃO**

ELIEL TAVARES LINS

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica – PROEE, da Universidade Federal de Sergipe, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Eduardo Oliveira Freire

Coorientador: Prof. Dr. Lucas Molina

São Cristóvão - SE, Brasil
Dezembro de 2021



UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
COORDENAÇÃO DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA-PROEE

TERMO DE APROVAÇÃO

"Aplicação da técnica BOVW para identificar falhas em máquinas rotativas a partir de imagens termográficas de baixa resolução"

Discente:

Eliel Tavares Lins

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Sergipe, como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica.

Aprovada pela banca examinadora composta por:

Prof. Dr. Jugurta Rosa Montalvão Filho (PROEE/UFS)
Presidente

Prof. Dr. Douglas Bressan Riffel (PROEE/UFS)
Examinador Interno

Prof. Dr. Evandro Ottoni Teatini Salles (UFES)
Examinador Externo

Eliel Tavares Lins
Candidato

Cidade Universitária "Prof. José Aloísio de Campos", 09 de dezembro de 2021.

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

L759a Lins, Eliel Tavares
Aplicação da técnica BoVW para identificar falhas em máquinas rotativas a partir de imagens termográficas de baixa resolução / Eliel Tavares Lins ; orientador Eduardo Oliveira Freire. – São Cristóvão, 2021.
85 f. ; il.

Dissertação (mestrado em Engenharia elétrica) – Universidade Federal de Sergipe, 2021.

1 Engenharia elétrica. 2. Termografia. I. Freire, Eduardo Oliveira, orient. II. Título.

CDU: 621.3

Agradecimentos

Primeiramente agradeço a Deus pelo dom da vida, e por ter me dado forças para enfrentar mais esse desafio na minha vida.

Agradeço a todos os meus familiares que sempre me incentivaram para fazer esse Mestrado, sempre me motivando para que eu me dedicasse cada vez mais.

Agradeço a todos os professores do programa pelo conhecimento compartilhado e em especial ao meu orientador, professor Dr. Eduardo Freire, e meu coorientador professor Dr. Lucas Molina, pela paciência e disponibilidade em me ajudar ao longo de todo curso.

Agradeço a direção e a gerência industrial da empresa Sergipe Industrial Têxtil Ltda., da qual fui funcionário, por terem me liberado para fazer esse Mestrado e também por terem disponibilizado o Simulador de Falhas utilizado nessa pesquisa.

Agradeço ao professor Dr. Douglas Bressan, por ter disponibilizado a câmera FLIR utilizada nessa pesquisa, e faço um agradecimento especial ao colega Wellington Silva que me acompanhou durante toda fase de coleta de dados com essa câmera FLIR.

Também agradeço aos meus colegas de turma, com os quais sempre pude contar nos momentos de dificuldade e também nos momentos de descontração.

Enfim, agradeço a todos que fazem a UFS, desde seus professores, funcionários efetivos e terceirizados, sem os quais não seria possível a viabilidade desse curso.

Resumo da Dissertação apresentada ao PROEE/UFS como parte dos requisitos necessários para a obtenção do grau de Mestre (Me.)

APLICAÇÃO DA TÉCNICA BOVW PARA IDENTIFICAR FALHAS EM MÁQUINAS ROTATIVAS A PARTIR DE IMAGENS TERMOGRÁFICAS DE BAIXA RESOLUÇÃO

ELIEL TAVARES LINS

Dezembro de 2021

Orientadores: Prof. Dr. Eduardo Oliveira Freire e Prof. Dr. Lucas Molina

Esta dissertação de mestrado visa descrever a implementação de uma estratégia de detecção de falhas em máquinas rotativas utilizando termografia infravermelha. Ela foi baseada em imagens adquiridas com uma câmera de baixa resolução, e comparou os resultados com aqueles obtidos com o uso de uma câmera de média resolução para verificar se era possível usar câmeras termográficas de resolução mais baixa sem perda significativa de desempenho, reduzindo custos e tornando essa abordagem mais acessível. A técnica BoVW (*bag of visual words*) foi aplicada para o tratamento dos dados e categorização das falhas. A eficácia do método apresentado foi validada experimentalmente a partir do diagnóstico das falhas simuladas com base nas imagens termográficas dos dois tipos de câmeras, e, adicionalmente, ratificado por análise de vibração. Os resultados apresentaram uma taxa de acerto acima de 92% para ambas as câmeras, comprovando a viabilidade do uso de câmeras termográficas de baixa resolução para identificação de falhas em máquinas rotativas.

Palavras-chaves: Termografia infravermelha; *Bag of visual words*; Falha em rolamentos; Desbalanceamento; Desalinhamento de eixo; Extração de características.

Abstract of Dissertation presented to PROEE/UFS as a partial fulfillment of the requirements for the degree of Master.

APPLICATION OF THE BOVW TECHNIQUE TO IDENTIFY FAULTS IN ROTATING MACHINERY FROM LOW- RESOLUTION INFRARED THERMOGRAPHY IMAGES

ELIEL TAVARES LINS

Dezembro de 2021

Advisor: Prof. Dr. Eduardo Oliveira Freire and Prof. Dr. Lucas Molina

This master thesis aims to describe the implementation of a fault detection strategy in rotating machines using infrared thermography. It was based on images acquired with a low resolution camera and compared the results with those obtained with the use of a medium resolution camera to verify if it is possible to use lower resolution thermographic cameras without significant loss of performance, thus reducing costs and making this approach more accessible. The BoVW technique (bag of visual words) was applied for data treatment and fault categorization. The effectiveness of the presented method was validated experimentally from the diagnosis of the faults simulated based on the thermographic images of the two types of cameras, and, in addition, ratified by vibration analysis. The precision results were over 92% of the accuracy for both cameras, thus proving the viability of using low resolution thermographic cameras to identify faults in rotating machines.

Keywords: Infrared thermography; Bag of visual words; Bearing fault; Unbalanced; Misaligned shaft; Feature extraction.

Sumário

Lista de Figuras	i
Lista de Tabelas	v
1 Introdução	1
1.1 Objetivos	2
1.1.1 Objetivo Geral	2
1.1.2 Objetivos Específicos	2
2 Revisão Bibliográfica	4
3 Fundamentação Teórica	9
3.1 Radiação Infravermelha	9
3.1.1 Termografia	11
3.1.2 Emissividade	11
3.1.3 Equipamentos Termográficos	12
3.2 Processamento de Imagens	13
3.2.1 SIFT	14
3.2.2 SURF	17
3.2.3 SVM	21
3.2.4 <i>k-Means</i>	23
3.2.5 BoVW	24
3.3 Desalinhamento de eixos	27
3.4 Mancais de rolamentos	29
3.4.1 Falhas em mancais de rolamento	30
3.5 Acoplamentos	31
3.6 Desbalanceamento	32

3.7 Validação cruzada	33
3.7.1 <i>Hold-Out</i>	34
3.7.2 <i>K-fold</i>	34
4 Metodologia	35
5 Resultados e Discussão	44
6 Conclusões	52
Referências Bibliográficas	55
Apêndices	62

Lista de Figuras

Figura 1 – Configuração experimental (traduzido de JIA et al., 2019)	6
Figura 2 – <i>Hall</i> Industrial (LEEMANS et al., 2011)	7
Figura 3 – Configuração experimental (traduzido de SEO et al., 2012)	8
Figura 4 – Espectro eletromagnético com detalhe ao espectro infravermelho (adaptado de https://canal.cecierj.edu.br/anexos/recurso_interno/7620/download/6f656d6c84ad20247fc819d7f4cf434a).	10
Figura 5 - Imagem térmica (infravermelho) ilustrando (a) as falhas de isolamento térmico de uma casa; (b) a presença de câncer em uma região da mama; (c) a detecção da turbina de um jato militar. (Fonte: ZEIDAN, 2017)	10
Figura 6 – Representação esquemática de uma medição termográfica (Fonte: FLIR <i>Systems</i> , 1999)	12
Figura 7 - Processo da Diferença de Gaussianas em uma imagem. (a) Representa o conjunto de imagens geradas após o processo de convolução da Gaussiana em diferentes resoluções. (b) Resultado da subtração de imagens com resoluções próximas (adaptado de LOWE, 2004).	15
Figura 8 - Processo de descoberta do mínimo e máximo local da DOG. A descoberta é realizada com a comparação de um pixel (representado por X na imagem) com outros 26 vizinhos em regiões 3x3 da imagem atual com as imagens em níveis de suavização adjacentes (adaptado de LOWE, 2004)	16
Figura 9 - Um descritor de ponto de interesse é criado calculando primeiro a magnitude do gradiente e a orientação em cada ponto da amostra da imagem em uma região em torno da localização do ponto de	

interesse, conforme mostrado à esquerda. Estes são ponderados por uma janela gaussiana, indicada pelo círculo sobreposto. Essas amostras são então acumuladas em histogramas de orientação resumindo o conteúdo em sub-regiões 4x4, conforme mostrado à direita, com o comprimento de cada seta correspondendo à soma dos gradientes-magnitudes próximos àquela direção dentro da região (adaptado de LOWE, 2004). 17

Figura 10 - Atribuição de orientação: uma janela de orientação deslizante de tamanho $\frac{\pi}{3}$ detecta a orientação dominante das respostas da *Wavelet de Haar* ponderada gaussiana em cada ponto da amostra dentro de uma vizinhança circular em torno do ponto de interesse (BAY et al., 2008). . . . 18

Figura 11 - Para construir o descritor, uma grade quadrática orientada com 4 x 4 sub-regiões quadradas é colocada sobre o ponto de interesse (esquerda). Para cada quadrado, as respostas da *wavelet* são calculadas a partir de 5 x 5 amostras (para fins ilustrativos, foi mostrado apenas 2 x 2 subdivisões aqui). Para cada campo, foram coletadas as somas dx , $|dx|$, dy e $|dy|$, calculados relativamente à orientação da grade (direita) (BAY et al., 2008). 19

Figura 12 - Processo de análise multiescala usando redução do tamanho da imagem (esquerda) e usando integral de imagem com os filtros em diferentes escalas (direita) (adaptado de BAY et al., 2008). 20

Figura 13 - Filtros em escalas diferentes, 9x9 (esquerda) e 15x15 (direita) (BAY et al., 2008). 21

Figura 14 - Pontos de interesse detectados para um campo de girassol (BAY et al., 2008). 21

Figura 15 – Duas classes linearmente separáveis (CANAHUIRE CABELLO, 2016) . . .	22
Figura 16 – Banco de dados de imagens termográficas de um mancal de rolamento . . .	25
Figura 17 – Criação do vetor de características de um ponto de interesse da imagem (Fonte: Adaptado de SOARES, 2012)	26
Figura 18 - Exemplo de uma imagem em tons de cinza (a) e um histograma da frequência relativa de ocorrência de cada uma das 500 palavras visuais (b) (Fonte: autor).	27
Figura 19 – a) Alinhamento com régua; b) Alinhamento com relógio comparador; c) Alinhamento a laser – adaptado de http://www.flowcontrolnet work.com/the-importance-of-shaft-alignment/	28
Figura 20 - Rolamento fixo de uma carreira de esferas (adaptado de NSK – 2013) . . .	29
Figura 21 - Rolamento de rolos cilíndricos (adaptado de NSK – 2013)	29
Figura 22 - Oxidação na face do rebordo, na superfície da pista e no espaçamento das esferas (Fonte: adaptado de NSK, 2013).	31
Figura 23 - Acoplamentos rígidos: (a) Flange, (b) Luva, (c) Luva de compressão, (d) <i>Quill shaft</i> (Fonte: TUCKMANTEL, 2018)	32
Figura 24 - Acoplamentos com flexibilidade mecânica: (a) Engrenagens, (b) <i>Falk</i> , (c) Corrente de rolos, (d) <i>Hirth</i> , (e) <i>Oldham</i> e (f) <i>Schmidt</i> (Fonte: TUCKMANTEL, 2018)	32
Figura 25 - Sistema Rotor-Mancal (SENKO, 2013)	33
Figura 26 – Diagrama de fluxo - BoVW (LI et al., 2019)	35
Figura 27 – Simulador de falhas	36
Figura 28 – Layout de instalação do simulador de falhas	36
Figura 29– Câmeras termográficas	37
Figura 30 – Coleta de dados	38

Figura 31 – Simulação de falha: eixo desalinhado	39
Figura 32– Detalhes do contrapeso adicionado ao volante para simular a condição de desbalanceamento do conjunto	39
Figura 33 – Ilustração das opções de configuração do tamanho do blockwidth	41
Figura 34 – Ilustração do método de validação cruzada <i>k-fold</i> com $k = 10$	42
Figura 35 - Histograma das palavras visuais das imagens termográficas – FLIR T420	45
Figura 36 - Histograma das palavras visuais das imagens termográficas – FLUKE VT02	47

Lista de Tabelas

Tabela 1 – Especificações das câmeras (Fonte: Manual do usuário)	37
Tabela 2 – Especificações do Simulador de Falhas	38
Tabela 3 – Legenda das tabelas 4 e 5	47
Tabela 4 – Medição de vibração durante os testes – FLUKE VT02	47
Tabela 5 – Medição de vibração durante os testes – FLIR T420	48
Tabela 6 – Resultado da taxa de acerto	49
Tabela 7 – Tempo total de processamento dos resultados.	50
Tabela 8 – Resultados da validação cruzada <i>k-fold</i>	50
Tabela 9 – Resultados dos testes de similaridade por produto interno.	51

Capítulo 1

1 Introdução

A alta competitividade, em função da evolução, crescimento e globalização da economia, determinou que as empresas tenham um compromisso ainda maior com o contínuo aperfeiçoamento de seus produtos e processos. Segundo SIQUEIRA (2009), o desafio da otimização dos ativos de produção é um processo que envolve a avaliação de funções, tarefas e atividades com o objetivo de conseguir um equilíbrio entre atividades reativas, preventivas e preditivas para garantir a preservação das funções dos ativos. Este processo é conseguido através da identificação dos modos de falha dos equipamentos, assim como, das respectivas consequências das falhas.

Nessa linha, pode-se destacar a manutenção como importante aliada das empresas em busca de melhores resultados, seja aumentando a disponibilidade das máquinas para produção, ou reduzindo os custos gerais de produção. Numa classificação resumida, pode-se dividir a manutenção em corretiva, preventiva e preditiva. Com relação a manutenção preditiva, apresentam-se algumas técnicas para diagnóstico de falhas muito utilizadas nas indústrias, em especial, análise de vibrações, análise de lubrificantes, emissões acústicas, termografia, entre outras. Mas todas elas contribuindo para uma melhor previsão de possíveis falhas ou que um determinado componente venha a funcionar fora das especificações desejadas.

Para BEZERRA (2004), um dos componentes críticos das máquinas são os rolamentos. Os defeitos desses componentes podem ser causados durante o processo de fabricação ou uso. Caso essas falhas não sejam detectadas a tempo, podem levar a um mau funcionamento da máquina e até comprometer outros componentes. A pesquisa de técnicas adequadas para o monitoramento das falhas em rolamentos mostra-se, portanto, de grande importância.

Diante da relevância em se monitorar falhas de componentes de máquinas, procurou-se apresentar, através dessa dissertação, uma contribuição às novas alternativas de monitoramento, tornando assim, a previsibilidade de falhas cada vez mais precisa e de fácil aquisição dos dados.

Vale ressaltar que técnicas como a análise de vibração e emissão acústica têm grande importância na manutenção preditiva. Mas essas técnicas apresentam alguns pontos negativos, como o risco de distorções no sinal devido a interferências externas. Também vale destacar a necessidade de um técnico especializado para realizar medições de vibração com precisão, e assim poder apresentar um diagnóstico correto. Todos estes

fatores criaram as condições para que a termografia surgisse como alternativa a substituir ou complementar essas técnicas.

Neste trabalho, considera-se como sendo câmeras termográficas de alta resolução aquelas com resolução igual ou superior à VGA (640x480 pixels); câmeras termográficas com resolução a partir da QVGA (320x240 pixels) e abaixo de VGA são consideradas como sendo de média resolução e câmeras termográficas com resolução abaixo de QVGA são consideradas como de baixa resolução.

Nos trabalhos encontrados na literatura da área, verifica-se que a sua ampla maioria se baseia no processamento de imagens termográficas adquiridas por câmeras de média resolução ou de alta resolução. Tais câmeras, além de terem custo elevado, nem sempre estão disponíveis no ambiente industrial. Por outro lado, câmeras termográficas de baixa resolução podem ser encontradas com maior frequência nas indústrias em geral, pois já são utilizadas em inspeções termográficas de rotina. Além disso, são câmeras com custo muito mais acessível.

1.1 Objetivos

1.1.1 Objetivo Geral

O objetivo geral deste trabalho consiste em investigar a viabilidade da utilização de imagens termográficas de baixa resolução para identificação de falhas em máquinas rotativas utilizando, e se necessário adaptando, um método que já tenha se mostrado eficaz, porém, que se baseou em imagens termográficas de média e alta resolução, possibilitando assim, reduzir o custo e aumentar a disponibilidade e viabilidade deste método de inspeção sem contato.

1.1.2 Objetivos Específicos

Os objetivos específicos dessa dissertação, são apresentados a seguir:

- Realizar uma ampla revisão bibliográfica sobre detecção de falhas em máquinas rotativas aplicando a técnica da termografia;
- Identificar métodos de processamento de imagens termográficas aplicados a falhas em máquinas rotativas;

- Verificar a viabilidade desses métodos para imagens termográficas com baixa resolução;
- Adaptar e implementar um método para detecção de falhas em máquinas rotativas a partir de imagens com baixa resolução;
- Realizar uma comparação entre imagens termográficas de média resolução (FLIR T420) e de baixa resolução (Fluke VT02) para avaliar a taxa de acerto na detecção de falhas em máquinas rotativas;
- Utilizar o método clássico de medição de vibração para validar o método de análise de falhas em máquinas rotativas a partir das imagens termográficas;
- Disponibilizar acesso público ao banco de dados com todas as imagens utilizadas nessa pesquisa, de forma que possa ser utilizado em outros trabalhos.

Capítulo 2

Revisão Bibliográfica

Com o aumento da necessidade de serem mais competitivas, as indústrias se veem forçadas a incorporar a seus processos fabris tecnologias para otimização de custos e aumento de produtividade e eficiência (SOARES, 2014). Alta confiabilidade, baixo impacto ambiental e segurança das operações são questões importantes para todos os setores (PENG et al., 2010). Políticas ótimas de manutenção em termos de custo, tempo de inatividade e qualidade do equipamento devem ser identificadas, além do que, a manutenção contribui significativamente para o custo total da empresa (GARG E DESHMUKH, 2006). Esses fatos levaram pesquisadores e profissionais a verem a manutenção como uma operação chave dentro das empresas de manufatura e assim tentar desenvolver uma abordagem holística para o gerenciamento de manutenção (WAEYENBERGH E PINTELON, 2009; BOUSDEKIS et al., 2015).

O monitoramento dos parâmetros operacionais do parque produtivo é uma ferramenta imprescindível para o planejamento das ações de intervenção da manutenção. Parâmetros como frequência e amplitude de vibração, temperatura de operação, granulometria de particulado em reservatórios de óleo são algumas das inúmeras fontes de informação utilizadas em modelos que determinam a existência de uma condição anormal, ou um processo de degradação em andamento, que, se não corrigido, evoluirá para uma falha. (CALADO, 2013).

Nas últimas décadas, o monitoramento de condições e diagnóstico de falhas mecânicas em máquinas rotativas tem recebido considerável atenção com o objetivo de aumentar a confiabilidade e segurança na operação de máquinas rotativas. No entanto, a análise de vibração ou emissão acústica pode ser intrusiva com seus sensores que precisam ser montados em contato direto com as máquinas (HUO, 2017).

Diante desse quadro, uma dessas linhas de pesquisa tem se direcionado para estudar a análise de vibração a partir de câmeras digitais convencionais, conforme pode-se citar a Dissertação de Mestrado de LEONARDI (2015). Esse trabalho apresentou o desenvolvimento de uma técnica para medições de deslocamento e vibrações sem contato a partir de quadros de imagens capturados. Foi utilizada uma câmera digital da marca CASIO® modelo EXILIM EX-ZR700, com sensor CMOS, configurada com resolução 640 x 480 pixels, e velocidade de filmagem de 30 quadros por segundo, que foram analisados e processados com o auxílio de ferramentas matemáticas de processamento de imagem. Entretanto, o autor afirmou que o método apresentou algumas dificuldades para câmeras de baixa resolução, exigindo que o ambiente possuísse boa iluminação.

Outra Dissertação de Mestrado nessa linha de pesquisa foi de RISSÁ (2017), que apresentou uma técnica de medida de deslocamento de baixa amplitude e frequência, a partir do processamento de imagens capturadas por meio de uma câmera digital, com uma resolução de 1920x1080 pixels a uma taxa de 30 fps (*frames per second*, ou seja, quadros por segundo). A técnica proposta foi submetida a testes realizados em uma máquina rotativa, composta por um motor elétrico acoplado a uma caixa de transmissão do tipo redutora de velocidades e um rotor. O autor afirmou que com o auxílio de um relógio comparador, juntamente com a câmera digital, foi possível realizar a detecção e análise de pequenos deslocamentos.

Por sua vez, o monitoramento de condições com base na medição de temperatura, utilizando a termografia por infravermelho (TI) como uma técnica alternativa à análise de vibração ou emissão acústica, é um teste não destrutivo, livre de interferência eletromagnética, além de poder fornecer grande cobertura de inspeção (HIDAYAT et al., 2018).

AZAD E RAMJI (2011) afirmam que ao utilizar câmeras termográficas, pode-se observar o calor gerado pelo atrito produzido em sistemas mecânicos. E que no caso de sistemas de acionamento mecânico, problemas em rolamentos podem ser identificados de maneira rápida e fácil, à medida que as falhas aparecem em componentes do rolamento (anel interno, anel externo, elemento girante ou gaiola), além da ausência ou excesso de lubrificação. As cargas radiais e/ou axiais impostas ao rolamento também produzem calor facilitando a análise dos mesmos sob a ótica da termografia, antevendo assim falhas e colapsos.

NUNEZ et al. (2016) destaca que ao utilizar a diferença da temperatura ambiente e da temperatura de um rolamento com falha até se atingir a temperatura de equilíbrio térmico, é possível utilizar a Termografia para avaliar falhas em rolamentos, indicando que gradientes de temperatura maiores que 5°C já sugerem anomalia que merece investigação pelo setor de manutenção.

SINGH (2014) utilizou pesos em seu experimento para causar desbalanceamento no eixo do sistema e buscar detectar pela termografia a presença de anomalias térmicas nos rolamentos. O mesmo afirmou que as forças radiais que causaram o desbalanceamento do eixo, em conjunto com as falhas por abrasão nas esferas de contato dos rolamentos, provocaram instabilidade e aquecimento nos rolamentos. Tal aquecimento foi detectado com o uso da termografia, e, em comparação com rolamento normal, percebeu-se que a termografia auxiliou na detecção de falhas nos rolamentos.

MOUSSA (2014) realizou um experimento para avaliação e monitoramento de falhas em rolamentos a partir de defeitos inseridos por meio de eletroerosão, vindo a concluir que as temperaturas nos rolamentos com falhas foram menores que as temperaturas do rolamento normal, justificando que há diminuição na área de contato das esferas com as pistas quando há presença de falha devido ao fato de que as esferas, ao invés de sempre rolarem pela pista, há momentos em que elas “saltitam” entre as pistas. Com a diminuição da área de contato, segundo o pesquisador, diminui-se o atrito gerado

e, por conseguinte diminui-se o calor gerado, refletindo na temperatura inferior do rolamento normal. Tais deduções reportadas por MOUSSA (2014), foram feitas partindo-se dos estudos de SAKAMOTO e TSUKIZOE (1978), SAKAMOTO (1987) e HESS e SOOM (1991) em que, sob carga harmônica, a área média de contato diminuiu, levando a força de fricção média a reduzir.

Outros autores têm se destacado no estudo de técnicas para identificar falhas em rolamentos de máquinas rotativas por meio da termografia. Diante da relevância de alguns desses trabalhos, será apresentado a seguir, de uma forma mais detalhada, para melhor contextualizar essa dissertação, os pontos mais relevantes de cada um desses artigos.

De acordo com JIA et al. (2019), seus resultados foram obtidos a partir de uma câmera termográfica Fluke Ti32, com resolução de 640 x 480 pixels, sendo o principal dispositivo para aquisição de dados, cujos parâmetros foram cuidadosamente definidos para se obter imagens de TI (Termografia Infravermelha) precisas. O objeto de diagnóstico utilizado pelo autor foi um rolamento de aço. Para seu estudo foi utilizado um experimento, conforme ilustrado na Fig. 1, onde a configuração de velocidade do motor foi gradualmente aumentada até 2000 rpm (rotações por minuto) para aquisição dos dados. Foram coletados dois conjuntos de dados (Grupo 1 e Grupo 2), onde cada um contém 100 imagens. O Grupo 1 é o conjunto de 100 imagens capturadas durante o estado transitório da máquina (começando 5 minutos após o início de operação da máquina). A velocidade de 2000 rpm foi mantida por 15 minutos até que a máquina atingisse o seu estado térmico estacionário. Então, outro conjunto de 100 imagens foi capturado como Grupo 2 (com 15 minutos para o final do experimento). Sabe-se que as imagens capturadas sob estado transitório podem variar significativamente, mesmo que apresentem a mesma característica de falha. Por outro lado, tal variação é significativamente reduzida para as imagens capturadas em estado estacionário.

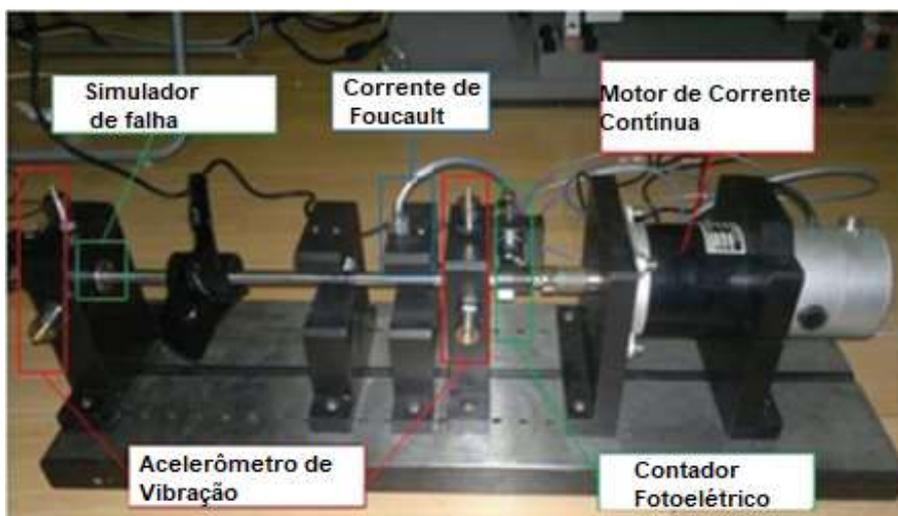


Figura 1 – Configuração experimental (traduzido de JIA et al., 2019).

Ainda de acordo com JIA et al. (2019), as ferramentas por eles utilizadas foram BoVW (*bag of visual word*) e CNN (*Convolutional Neural Network*) que foram comparadas nos experimentos. Para BoVW, a estrutura de diagnóstico foi realizada da

seguinte forma: i) captura das imagens de TI e pré-processamento das regiões de interesse dessas imagens, seguidas de transformação característica com algoritmo de detecção e descrição de pontos de interesse invariantes a escala e rotação (SIFT – *Scale Invariant Feature Transform*) para extrair características visuais das imagens de TI; ii) das características visuais extraídas, um vocabulário visual foi gerado, pelo qual cada imagem foi analisada para diagnóstico de falhas; e iii) os recursos extraídos do BoVW foram usados como entrada do algoritmo de classificação, o SVM (*Support Vector Machine*). Para a CNN, a estrutura foi organizada da seguinte forma: i) captura e pré-processamento das imagens TI da região de interesse, seguidas do redimensionamento da resolução da imagem para entrada da rede; ii) construção de uma estrutura de rede para aprendizagem e classificação das características de falhas de cada imagem aplicando a CNN; e iii) as características extraídas utilizando a CNN foram usadas como entrada no classificador SVM. E como resultado final, a saída do classificador SVM indicou o estado de falha das máquinas rotativas, incluindo qual componente está com defeito.

Na Fig. 2, pode-se identificar como LEEMANS et al. (2011), a partir de uma câmera termográfica (FLIR A320), que foi colocada no teto do *hall* industrial (7 m de altura), adquiriu seus termogramas, os quais foram pré-processados em cinco Hertz (para um controle visual contínuo por um operador) e foram processados e registrados a cada 15 minutos. Duas formas foram utilizadas para detectar um desvio de temperatura. Pela primeira forma, o modelo era treinado durante um determinado período e, em seguida, o erro do modelo era monitorado. Qualquer desvio estaria induzindo um aumento do erro. A segunda maneira considerou um treinamento contínuo, sendo o modelo adaptado após cada medição. O limiar foi a temperatura média mais três desvios padrão dessa temperatura.

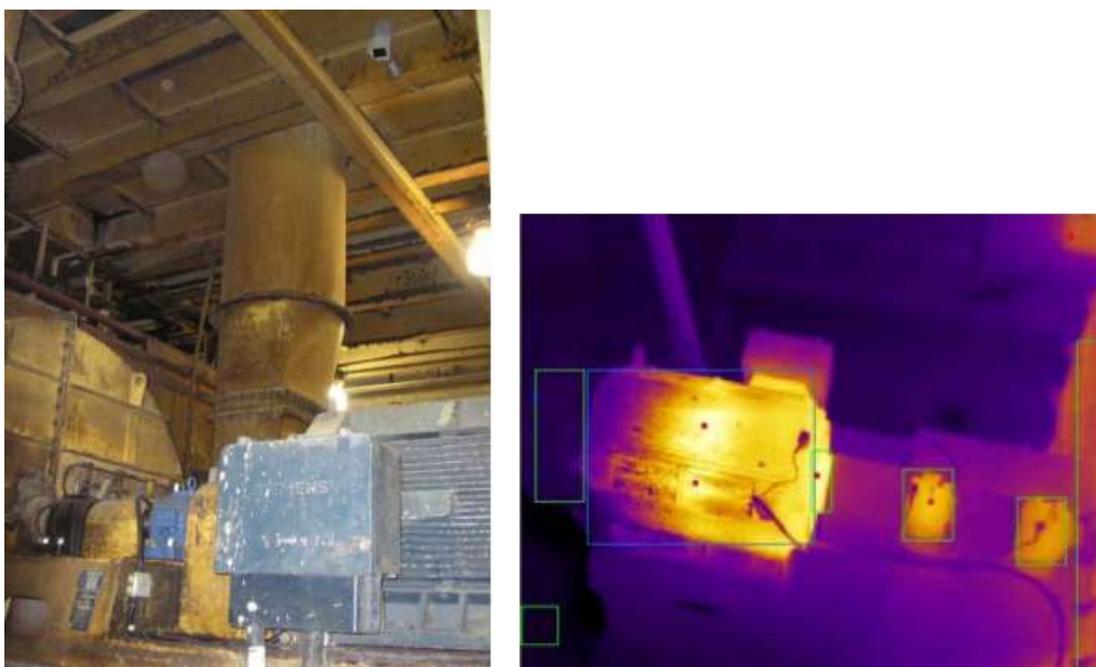


Figura 2 – *Hall* Industrial (LEEMANS et al., 2011).

LEEMANS et al. (2011) concluiu afirmando que o melhor modelo levou em consideração a temperatura ambiente, inércia térmica, a potência transmitida e as trocas de calor com um dissipador de calor fixo. Esse modelo precisava ser treinado continuamente para compensar as variações de temperatura devido ao processo. Além do que, ao medir-se com precisão a temperatura do rolamento e assim detectando-se um desvio de temperatura precoce, deve ser possível monitorar a evolução da temperatura e, assim, prever a vida útil restante do componente. Essas informações são cruciais para um programa de manutenção avançado.

De acordo com SEO et al. (2012), a partir do uso de uma câmera termográfica, FLIR Silver 450M, o experimento mediu 20 quadros por segundo durante 40 minutos, conforme é ilustrado a seguir na Fig. 3. A mudança das características de temperatura do rolamento foi observada em tempo real. Os dados foram calculados a partir da conexão entre o anel interno, externo e a esfera do rolamento no qual a geração de calor foi mais grave. Foram avaliadas as características de temperatura de acordo com as cinco condições do rolamento, sendo facilmente detectado, nos rolamentos danificados, o seguinte: 1) O gradiente de temperatura inicial, 2) O valor máximo de temperatura de acordo com cada condição, 3) A taxa de aumento de temperatura de acordo com a velocidade de rotação. Com base nesses resultados, o diagnóstico de anormalidade não-destrutivo sem contato, usando o método de termografia infravermelha, foi útil para o monitoramento de condições de elementos de máquinas rotativas.

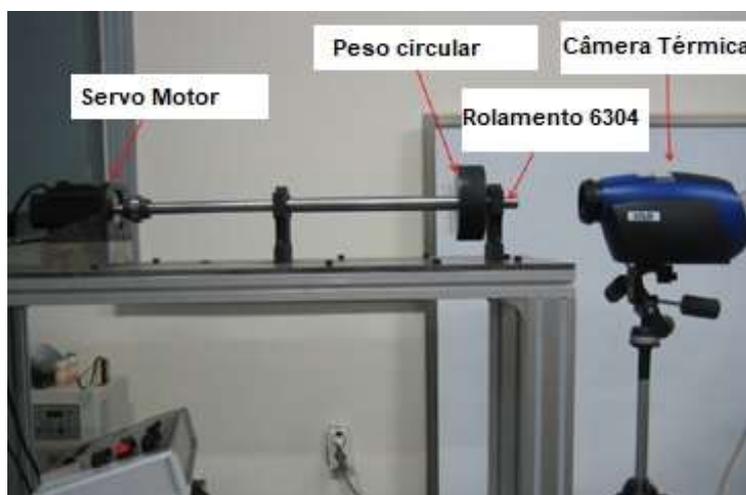


Figura 3 – Configuração experimental (traduzido de SEO et al., 2012).

Vale ressaltar a necessidade de mais estudos com câmeras termográficas de baixa resolução, tendo em vista que as publicações estudadas apresentam apenas imagens com média e alta resolução. Além de avaliar também se os métodos de processamento das imagens termográficas já desenvolvidos necessitam de algum aprimoramento, como forma de compensar a menor resolução das imagens usadas.

Capítulo 3

Fundamentação teórica

Neste capítulo são apresentados os conceitos básicos relacionados com a pesquisa, em que são abordados os seguintes temas: termografia, processamento de imagens, desalinhamento de eixos, mancais de rolamentos, acoplamentos, desbalanceamento e validação cruzada.

3.1 Radiação Infravermelha

A radiação infravermelha é conhecida desde o início do século XIX, mas foi somente a partir da segunda metade do século XX que os detectores infravermelhos foram desenvolvidos (principalmente para aplicações militares) e que o desenvolvimento deste tipo de radiação encontrou algumas aplicações (ZEIDAN, 2017). De acordo com MEOLA (2012), a radiação infravermelha é a energia irradiada pela superfície de um objeto cuja temperatura está acima do zero absoluto. A radiação emitida é uma função da temperatura do material, em que quanto mais alto a temperatura, maior será a intensidade da energia infravermelha emitida.

Existem três maneiras pelas quais a energia radiante que atinge um objeto pode ser dissipada: absorção, transmissão e reflexão (MOLLMANN et al., 2005). As frações da energia radiante total que estão associadas a cada desses modos de dissipação são chamados de absortividade, transmissividade e refletividade do corpo (HOWELL et al., 2010). Foi ilustrado através da Fig. 4 o espectro eletromagnético para facilitar o entendimento da natureza e origem da radiação infravermelha que é parte do espectro eletromagnético, em que seu comprimento de onda é maior do que o da luz visível.

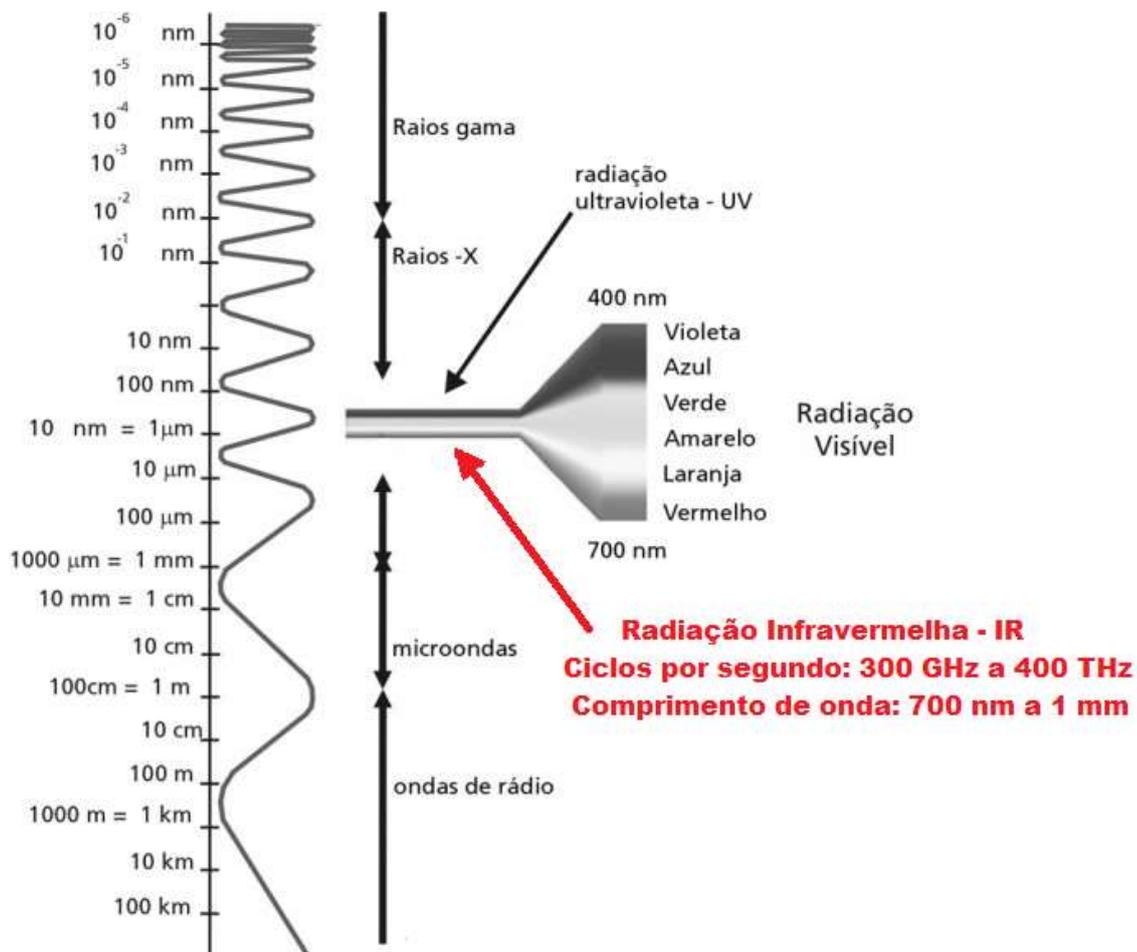


Figura 4 – Espectro eletromagnético com detalhe ao espectro infravermelho (adaptado de https://canal.cecierj.edu.br/anexos/recurso_interno/7620/download/6f656d6c84ad20247fc819d7f4cf434a)

Na Fig. 5, pode se observar algumas importantes aplicações que surgiram a partir dos estudos realizados com a radiação infravermelha. Com a fabricação de detectores infravermelhos de alta sensibilidade, câmeras de infravermelho de alta qualidade tornaram-se disponíveis e têm aplicações em vários campos estratégicos, como medicina, engenharia, ciência, defesa, agricultura, meio ambiente e energia (ZEIDAN, 2017).

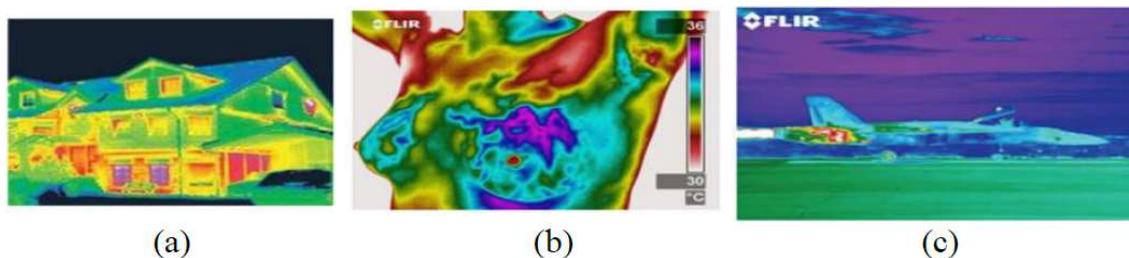


Figura 5 - Imagem térmica (infravermelho) ilustrando (a) as falhas de isolamento térmico de uma casa; (b) a presença de câncer em uma região da mama; (c) a detecção da turbina de um jato militar. (Fonte: ZEIDAN, 2017)

3.1.1 Termografia

De acordo com MAHIAGITE e MOORE (2001), a termografia é uma técnica não destrutiva que tem por objetivo analisar o perfil de temperaturas na superfície de um corpo, tornando possível detectar regiões de sua superfície que apresentem uma diferença de temperatura em relação ao restante do corpo, indicando assim a possibilidade da presença de defeitos nesta superfície ou abaixo dela. Posteriormente, pode-se correlacionar esta diferença de temperaturas a um mal funcionamento do equipamento inspecionado e também à presença de defeitos.

Para que ocorra essa transição, a energia térmica precisa ser transportada de um objeto para o outro de forma que as temperaturas, e conseqüentemente o fluxo de calor, permaneçam constantes (KAPLAN, 1999). Neste contexto, existem três formas de transferência de calor: condução, convecção e radiação.

A condução é a transmissão de calor em que a energia passa de um corpo para outro através das partículas do meio que os separa. Para esse processo ocorrer é necessário um meio material, não podendo ocorrer, portanto, no vácuo. A convecção é um processo de transmissão de calor através do movimento de massas de fluido, trocando de posição entre elas. A radiação é o processo de transmissão de calor por meio de ondas eletromagnéticas (ondas de calor), podendo ocorrer em meios materiais ou no vácuo (BÔAS, 1987).

3.1.2 Emissividade

Uma propriedade de extrema importância dos materiais para avaliar a sua aplicabilidade em termografia é a emissividade (ϵ). A emissividade é a propriedade que representa a razão entre a quantidade de energia total irradiada e/ou absorvida pelo corpo de prova a uma temperatura, e a de um corpo negro na mesma temperatura. O corpo negro é uma fonte hipotética de radiação que emite a energia máxima de radiação teoricamente possível, a uma certa temperatura, e tem seu valor de emissividade igual a 1 (um). Os demais são denominados de corpos cinza, e apresentam valores de emissividade entre 0 (zero) e 1 (um) (MALDAGUE, 2001).

A termografia apresenta um grande sucesso na inspeção de materiais com alta emissividade (entre 0,7 a 0,9). Este sucesso se deve à alta intensidade de radiação emitida por um material que possui uma alta emissividade, conforme é visto na lei de Stefan-Boltzmann e também pelo fato de materiais com baixa emissividade apresentarem uma alta taxa de reflexão, refletindo assim radiação oriunda de outras fontes, o que confere ruídos na imagem termográfica e reduz a intensidade dos dados de interesse do corpo inspecionado (MALDAGUE, 2001).

A emissividade é a variável mais importante para obtenção de dados termográficos quantitativos. Depende do tipo de material e da qualidade da superfície analisada. Quanto mais brilhante for a superfície do objeto maior será a sua refletância e

menor será a sua emissividade. Nestes casos, deve-se preparar a superfície com um material de emissividade conhecida a exemplo de pintá-lo com tinta a óleo fosca, cuja emissividade é 0,95. Não ter conhecimento desta variável impossibilita a leitura quantitativa do objeto em estudo (CARAMALHO, 2012; AHMADIPOUR e NEMATOLLAHI, 2015).

3.1.3 Equipamentos Termográficos

De acordo com SILVA (2016), os equipamentos termográficos ou termovisores, também conhecidos como câmeras térmicas ou de infravermelho, captam a radiação que resulta em uma dada temperatura e possibilita a câmera calcular e mostrar determinado coeficiente térmico do objeto. Quando um termovisor é direcionado para o objeto de estudo, a energia infravermelha é focada por um conjunto de lentes sobre um detector de infravermelho que envia a informação eletrônica para um sensor processar e gerar a imagem em aproximadamente $5\mu s$ para os equipamentos de sistemas de varredura e de 15ms para os sistemas FPA (*Focal Plane Array – Matriz de plano focal*).

A captação da imagem térmica ocorre da seguinte forma ilustrada na Fig. 6.

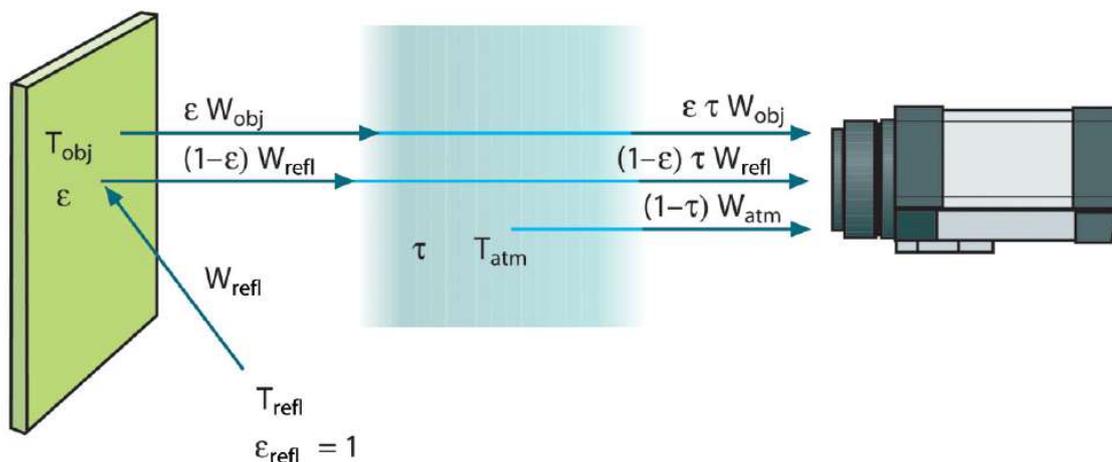


Figura 6 – Representação esquemática de uma medição termográfica (FLIR Systems, 1999).

USAMENTIAGA (2014) afirma que a potência total da radiação medida pela câmera (W_{tot}), apresentada na Fig. 6, é formulada a partir dos fatores influentes entre a câmera, a superfície do objeto e a atmosfera:

$$W_{tot} = \varepsilon \cdot \tau \cdot W_{obj} + (1 - \varepsilon) \cdot \tau \cdot W_{refl} + (1 - \tau) \cdot W_{atm} \quad (3.1)$$

Onde, ε é a Emissividade, τ é o Coeficiente de transmissibilidade da atmosfera, W_{obj} é a Energia radiada a partir do objeto, W_{refl} é a Energia refletida a partir dos corpos vizinhos e W_{atm} é a Energia relativa à absorção da atmosfera.

3.2 Processamento de Imagens

De acordo com PEDRINI E SCHWARTZ (2008), o processamento de imagens digitais consiste em um conjunto de técnicas para capturar, representar e transformar imagens com auxílio do computador. Ainda de acordo com os autores, tais técnicas permitem extrair e identificar informações das imagens, além de prover melhoria de qualidade visual e interpretação automática de seus conteúdos.

YADAV E YADAV (2009) definem uma imagem como uma função bidimensional, $f(x, y)$, em que x e y são coordenadas espaciais em um determinado plano. Uma imagem é considerada digital quando (x, y) e os valores de amplitude de f são finitos em quantidades discretas. A amplitude de f no ponto (x, y) , denominado pixel, é chamada de intensidade ou escala de cinza da imagem neste ponto, sendo normalmente representada como potência de 2. Uma imagem digital é então representada por uma matriz de dimensões $M \times N$, em que M representa as linhas e N as colunas. As coordenadas de origem são $(0, 0)$ e as coordenadas da primeira linha são: $(0, 0)$, $(0, 1)$, $(0, 2)$, ..., $(0, N-1)$. Cada pixel está localizado em uma coordenada (x, y) .

CANAHUIRE CABELLO (2016) afirma que um algoritmo de extração de características de uma imagem deve, em um primeiro passo, extrair da forma mais eficaz e robusta possível, determinadas características que proporcionem a máxima informação possível. Essas características devem cumprir as seguintes condições:

- O custo computacional para a extração de características não deve ser excessivo. O tempo total da extração deve ser o menor possível;
- A localização deve ser precisa. O erro cometido na estimação das características também deve ser o menor possível;
- Devem ser robustas e estáveis. Deveriam permanecer ao longo de uma sequência de imagens similares;
- Devem conter a maior informação possível da cena, isso significa que as características devem permitir a extração de informação de tipo geométrico.

Características locais, também chamadas de pontos de interesse, são caracterizadas por mudanças bruscas de intensidade na região. Esses recursos locais são geralmente classificados em cantos, bordas e manchas (KASPERS, 2011). Daí destaca-se a segmentação como importante aliada na identificação desses pontos de interesse.

A segmentação é considerada um pré-processamento essencial para muitas aplicações. Ela consiste em dividir a imagem em várias partes, cada uma contendo informações próprias de cor, intensidade, formato e textura. Por exemplo, na identificação de caracteres, primeiro se faz a segmentação das regiões da imagem que possuem caracteres, e depois é que se faz a classificação dos mesmos (CARVALHO JÚNIOR, 2010). Para auxiliar no processo da segmentação, destacam-se o SIFT e o SURF (*Speeded Up Robust Features*) para realizar a detecção e descrição dos pontos de interesse.

3.2.1 SIFT

Para BORTH et al. (2016), o SIFT e o SURF são os dois métodos mais populares na extração de uma coleção de vetores de características locais da imagem, chamados de pontos de interesse. Cada ponto de interesse é composto de uma parte pequena da imagem e a quantidade de pontos detectados pode variar de uma imagem para outra. Esses métodos são invariantes à rotação, escala da imagem e mudança de iluminação.

Lowe (2004), afirma que um aspecto importante na abordagem do SIFT é que ele gera um grande número de pontos de interesse que cobrem densamente a imagem em toda a gama de escalas e locais. Lowe (2004) exemplifica ainda que uma imagem típica de tamanho 500x500 pixels dará origem a cerca de 2.000 pontos de interesse estáveis (embora esse número dependa do conteúdo da imagem e das opções de vários parâmetros). A quantidade de pontos de interesse é particularmente importante para o reconhecimento de objetos, onde a capacidade de detectar pequenos objetos em fundos desordenados requer que pelo menos 3 descritores sejam correspondidos corretamente de cada objeto para uma identificação confiável (Lowe, 2004).

A função $L(x, y, \sigma)$ define o espaço de escala de uma imagem, sendo esta produzida por uma operação de convolução de uma gaussiana, $G(x, y, \sigma)$, em uma imagem, $I(x, y)$, como apresentado na Equação 3.2:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3.2)$$

onde $*$ é o operador de convolução nas coordenadas e $G(x, y, \sigma)$ é definido de acordo com a Equação 3.3:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (3.3)$$

onde σ representa o nível de suavização da gaussiana G .

Para uma maior eficiência na detecção de pontos de interesse estáveis, o SIFT utiliza uma função de diferença entre Gaussianas (DOG - *difference of Gaussians*), $D(x, y, \sigma)$, em que pode ser calculada a partir da diferença entre duas gaussianas com escalas diferentes, sendo uma delas influenciada por um multiplicador k como apresentado na Equação 3.4.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (3.4)$$

O fator k é responsável pelo nível de suavização da imagem. Substituindo a Equação 3.2 na Equação 3.4 obtém-se a diferença entre gaussianas definida pela Equação 3.5

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3.5)$$

Na Fig. 7 foi ilustrada uma visão geral da diferença entre Gaussianas (DOG). De início, a imagem $I(x, y)$, foi incrementalmente convoluída com o filtro da gaussiana $G(x, y, \sigma)$, produzindo um conjunto de novas imagens em diferentes níveis de suavização, chamado de oitavas ou iterações. Um conjunto de quatro imagens gaussianas são

agrupadas em cada oitava, a Fig. 7(a) ilustra esse processo. Em seguida imagens com níveis de suavização próximos, ou seja, que possuem diferentes σ contidas na oitava são subtraídas, produzindo uma imagem resultante denominada diferença entre Gaussiana (DOG), como apresentado na Fig. 7(b). Ao final desse processo, um novo valor foi atribuído ao fator k e multiplicado à variável σ , produzindo uma nova oitava em um nível de suavização diferente, a DOG foi novamente calculada e o processo repetido formando uma pirâmide Gaussiana. A quantidade de oitavas a cada iteração foi empiricamente determinada como 3 (LOWE, 2004).

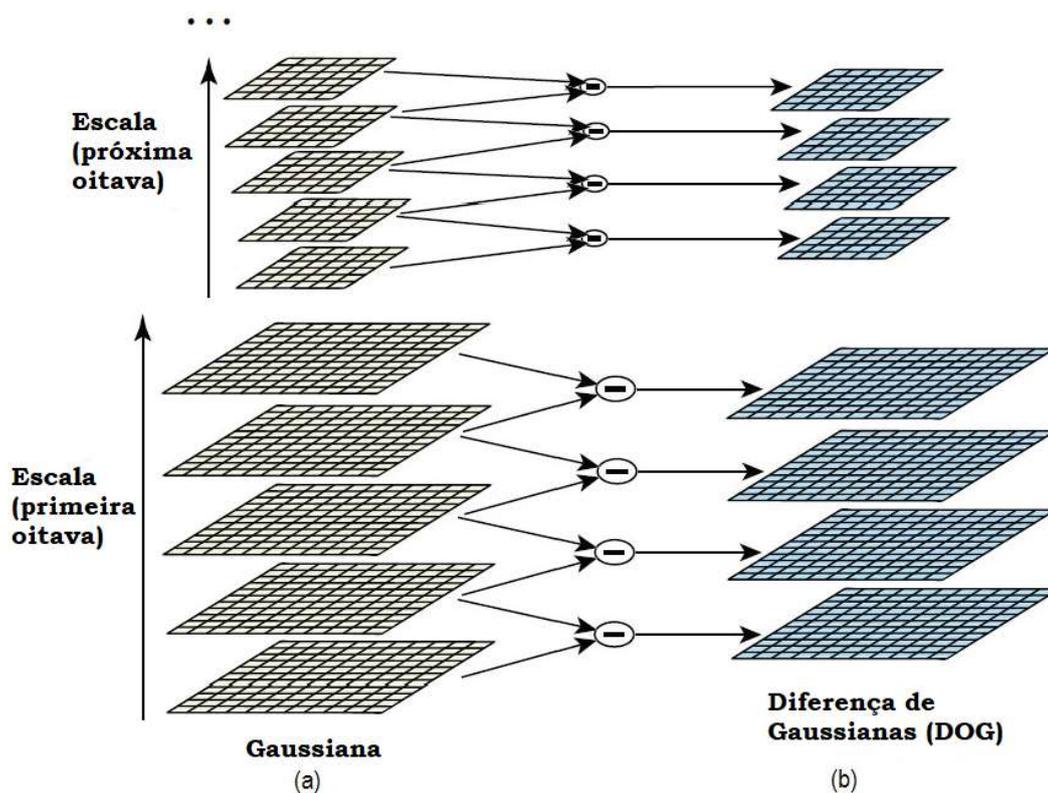


Figura 7 — Processo da Diferença de Gaussiana em uma imagem. (a) Representa o conjunto de imagens geradas após o processo de convolução da Gaussiana em diferentes resoluções. (b) Resultado da subtração de imagens com resoluções próximas (adaptado de LOWE, 2004).

Ao se verificar que as DOG foram encontradas no primeiro estágio, a descoberta do mínimo e máximo local das DOG foi necessária, pois, objetiva encontrar e localizar candidatos a pontos representativos nas imagens em cada nível de suavização, etapa esta presente no segundo estágio. O processo de descoberta ocorre através da comparação de intensidade de pixel utilizando uma janela de tamanho 3×3 . O elemento central nesta janela tem sua intensidade comparada com os 26 vizinhos próximos, sendo oito contidos na mesma DOG e nove localizados na mesma posição tanto na DOG de nível superior quanto inferior, conforme ilustrado na Fig. 8 (LOWE, 2004). Com a descoberta dos

pontos candidatos, em seguida é necessário o ajuste das informações de localização, escala e posicionamento. Essas informações permitem filtrar pontos em baixos contrastes, que possuem alta sensibilidade a ruídos, ou mal localizado em uma borda. Utiliza-se a expansão de Taylor (BROWN; LOWE, 2002) para, no espaço escala da função $D(x, y, \sigma)$, determinar a interpolação da localização do máximo, ou seja, diminuir a quantidade de pontos candidatos que representam a mesma característica.

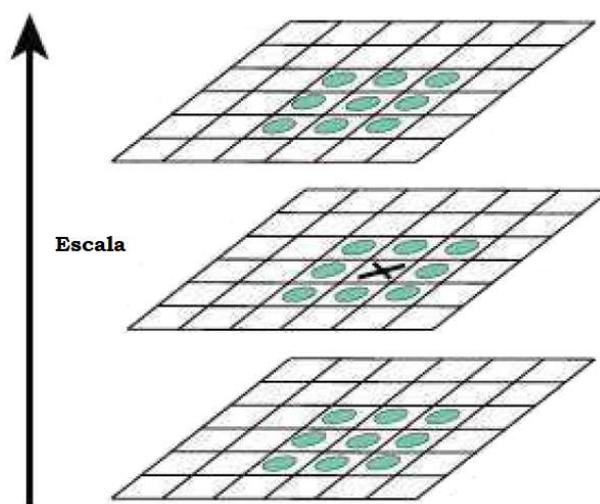


Figura 8 - Processo de descoberta do mínimo e máximo local da DOG. A descoberta é realizada com a comparação de um pixel (representado por X na imagem) com outros 26 vizinhos em regiões 3x3 da imagem atual com as imagens em níveis de suavização adjacentes (adaptado de LOWE, 2004).

A Fig. 9 ilustra o cálculo do descritor de ponto de interesse apresentando uma matriz de descritores 2x2 calculada a partir de um conjunto de amostras 8x8. Primeiro, as magnitudes e orientações do gradiente da imagem são amostradas em torno da localização do ponto de interesse, usando a escala do ponto de interesse para selecionar o nível de desfoque gaussiano da imagem (LOWE, 2004). Para obter a invariância da orientação, as coordenadas do descritor e as orientações do gradiente são giradas em relação à orientação do ponto de interesse. Para melhor eficiência, os gradientes são pré-calculados e ilustrados com pequenas setas em cada local de amostra no lado esquerdo da Fig. 9.

Ainda de acordo com LOWE (2004), uma função de ponderação gaussiana com metade da largura da janela do descritor é usada para atribuir um peso à magnitude de cada ponto da amostra. Isso é ilustrado com uma janela circular no lado esquerdo da Fig. 9. O objetivo desta janela gaussiana é evitar mudanças repentinas no descritor com pequenas mudanças na posição da janela, e dar menos ênfase aos gradientes que estão longe do centro do descritor, pois, são os mais afetados por erros de registro. O descritor de ponto de interesse é mostrado no lado direito da Fig. 9, o qual permite uma mudança significativa nas posições do gradiente, criando histogramas de orientação em regiões da amostra 4x4. A figura mostra oito direções para cada histograma de orientação, com o comprimento de cada seta correspondendo à magnitude dessa entrada do histograma.

Uma amostra de gradiente à esquerda pode deslocar até 4 posições de amostra enquanto ainda contribui para o mesmo histograma à direita, alcançando assim o objetivo de permitir maiores mudanças de posição locais.

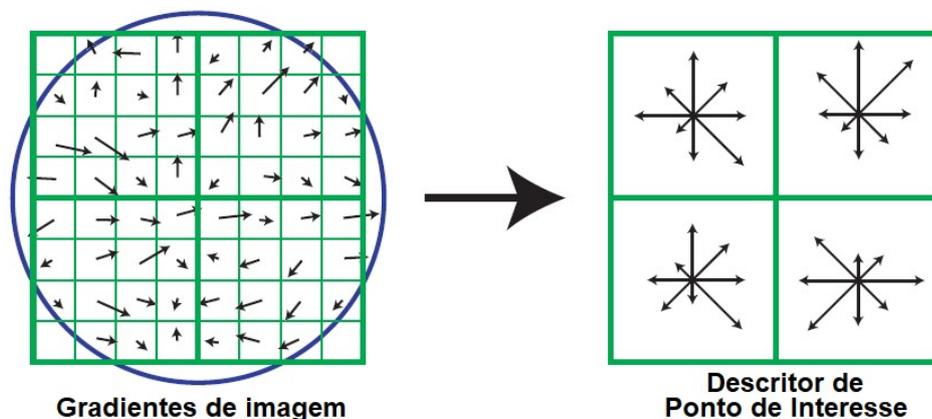


Figura 9 - Um descritor de ponto de interesse é criado calculando primeiro a magnitude do gradiente e a orientação em cada ponto da amostra da imagem em uma região em torno da localização do ponto de interesse, conforme mostrado à esquerda. Estes são ponderados por uma janela gaussiana, indicada pelo círculo sobreposto. Essas amostras são então acumuladas em histogramas de orientação resumindo o conteúdo em sub-regiões 4x4, conforme mostrado à direita, com o comprimento de cada seta correspondendo à soma dos gradientes-magnitudes próximos àquela direção dentro da região (adaptado de LOWE, 2004).

3.2.2 SURF

O SURF foi proposto por Herbert Bay et al. (BAY et al., 2006) sendo inspirado no SIFT, mas com o objetivo de ser rápido na extração e comparação de características. O SURF utiliza apenas características de imagens de tons de cinza, ou seja, não utiliza qualquer informação relacionada às cores da imagem (JURASZEK, 2014). Este método tem sido aplicado por várias razões, dentre as quais destacam-se: a redução do custo computacional, devido a redução da dimensionalidade a ser examinada; e a diminuição da redundância incorporada, uma vez que os pixels não se movem de forma independente e possuem um elevado grau de correlação (MONTANARI, 2016).

De acordo com BAY (2006), para a extração do detector, a primeira etapa consiste em construir uma região quadrada centrada em torno do ponto de interesse e direcionada ao longo da orientação selecionada. Para extrair as informações que melhor descrevem esta área em volta do ponto de interesse, o SURF utiliza convoluções a partir do método *Haar Wavelet*. A extração das informações de orientação resulta em um vetor de 64 posições que descrevem o ponto de interesse (JURASZEK, 2014).

Os pontos de interesse são localizados através da aplicação do cálculo do determinante máximo das matrizes Hessianas, sendo que estas são interpoladas em escala

e espaço de imagem com o método proposto por LOWE (1999). A interpolação do espaço de escala é tarefa importante devido às primeiras camadas de cada oitava possuírem uma grande diferença em relação a escala.

O descritor SURF aplica uma distribuição de intensidade na vizinhança do ponto de interesse detectado de forma semelhante ao descritor SIFT (LOWE, 2004). O descritor possui três etapas principais: atribuição de orientação, extração e indexação rápida para correspondência. A primeira etapa consiste na fixação de uma região circular no entorno do ponto de interesse para posterior análise das informações nela contidas. O principal objetivo desta etapa é tornar os pontos de interesse invariantes a rotação, de forma que as orientações calculadas sejam reproduzíveis. Diante desse objetivo, primeiramente são calculadas as respostas de *Wavelet de Haar* das direções x e y dentro da região circular formada no entorno do ponto de interesse em uma determinada escala s . Tanto a fase de amostragem quanto o tamanho das *wavelets* são dependentes da escala (BAY et al., 2008).

BAY et al. (2008) ainda afirma que as respostas da *wavelet* são calculadas e ponderadas com uma Gaussiana ($\sigma = 2s$) centrada no ponto de interesse. As respostas são representadas como pontos em um espaço com a força da resposta horizontal ao longo da abscissa e a força da resposta vertical ao longo da ordenada. De acordo com a Fig. 10 a orientação dominante é estimada calculando-se a soma de todas as respostas dentro de uma janela de orientação deslizante de tamanho $\frac{\pi}{3}$. As respostas horizontais e verticais dentro da janela são somadas. As duas respostas somadas produzem então um vetor de orientação local. O maior desses vetores em todas as janelas define a orientação do ponto de interesse.

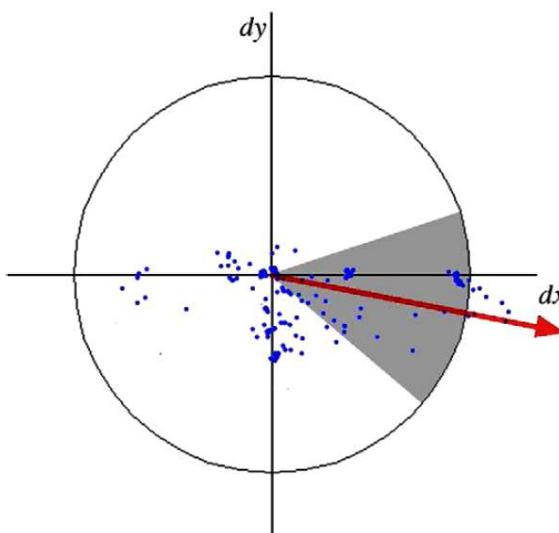


Figura 10 - Atribuição de orientação: uma janela de orientação deslizante de tamanho $\frac{\pi}{3}$ detecta a orientação dominante das respostas da *Wavelet de Haar* ponderada gaussiana em cada ponto da amostra dentro de uma vizinhança circular em torno do ponto de interesse (BAY et al., 2008).

A Fig. 11 ilustra o processo da etapa de extração de características para a formação do descritor. Inicialmente foi criada uma região quadrada centralizada no entorno do ponto de interesse com orientação já atribuída. Esta região ou janela possui tamanho igual a vinte vezes a escala s nesse ponto e são divididas em sub-regiões de tamanho 4×4 , preservando assim importantes informação espaciais. Para cada sub-região são aplicadas as *Wavelet de Haar* nas direções horizontais, dx , e verticais, dy , sendo estas calculadas em relação ao ponto de interesse. O vetor de características gerado pelo descritor consiste no somatório entre as direções e seus valores absolutos, onde $v = \sum dx, \sum dy, \sum |dx|, \sum |dy|$. Concatenando os somatórios, como resultado, o vetor descritor final possui tamanho 64.

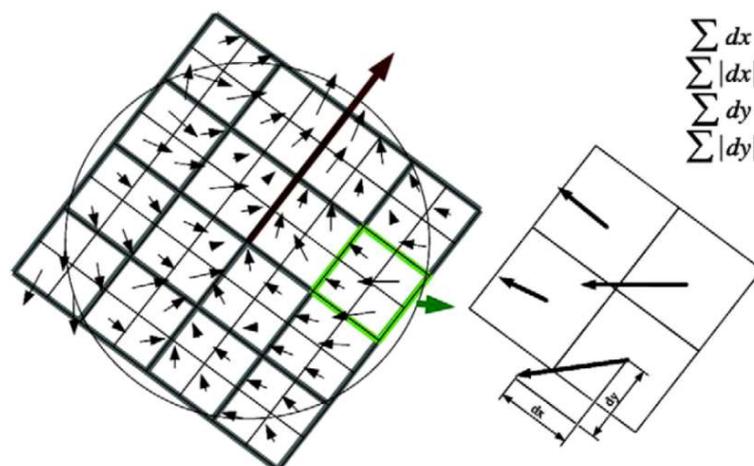


Figura 11 - Para construir o descritor, uma grade quadrática orientada com 4×4 sub-regiões quadradas é colocada sobre o ponto de interesse (esquerda). Para cada quadrado, as respostas da *wavelet* são calculadas a partir de 5×5 amostras (para fins ilustrativos, foi mostrado apenas 2×2 subdivisões aqui). Para cada campo, foram coletadas as somas dx , $|dx|$, dy e $|dy|$, calculados relativamente à orientação da grade (direita) (BAY et al., 2008).

Para um melhor entendimento do que está ilustrado na Fig. 12, será apresentada a técnica de imagem integral, que é muito útil em computação visual por permitir a realização de diversos cálculos em subregiões de uma imagem. Cálculos como média, soma e *wavelets de Haar* podem ser calculados de maneira muito eficiente, reduzindo significativamente o tempo de processamento das imagens. Uma forma de representação, apresentada inicialmente por CROW (1984), está ilustrada a seguir:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

Ainda de acordo com CROW (1984), cada ponto x, y de uma imagem contém o somatório da intensidade dos pixels da origem até a sua localização. Desta forma, com uma única passada na imagem, é possível computar a imagem integral. E com estas

informações, é possível calcular qualquer soma retangular utilizando apenas quatro referências.

De acordo com BAY et al. (2008), o processo de análise multiescala é realizado através do aumento gradativo do tamanho do filtro em uma pirâmide. A Fig. 12 demonstra esse processo, de forma que uma pirâmide de imagens consiste em apresentar uma sequência na qual cada imagem subsequente tem um nível de resolução menor que a anterior. Como pode ser observado na Fig. 12, esse processo difere de métodos usuais de análise multiescala em pirâmide que usam redução do tamanho da imagem.

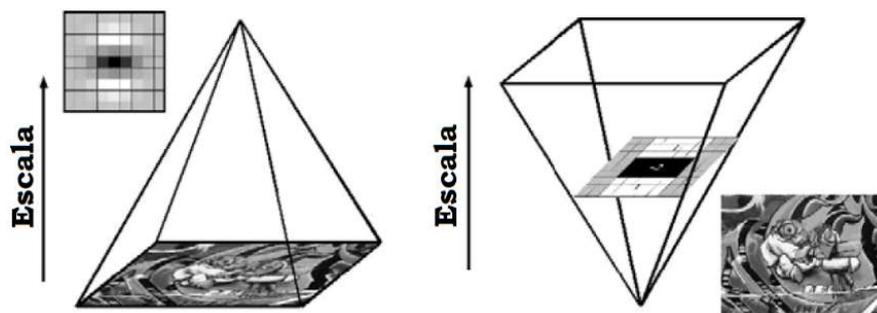


Figura 12 - Processo de análise multiescala usando redução do tamanho da imagem (esquerda) e usando imagem integral com os filtros em diferentes escalas (direita) (adaptado de BAY et al., 2008).

Ainda de acordo com BAY et al. (2008), este método possibilita otimizar o desempenho do algoritmo em termos de processamento, uma vez que as operações envolvem apenas alguns cálculos com a integração. As respostas às diferenças de escala e rotação são analisadas unicamente com a transformação dos filtros, tanto em relação a direção como ao tamanho.

A Fig. 13 representa dois filtros em escalas diferentes, ilustrando a aplicação do filtro em diferentes escalas. O mesmo está relacionado ao conceito de oitavas, que consiste em agregar valores intermediários entre as escalas de análise. Por fim, para localizar os pontos de interesse sobre escalas, é aplicada uma supressão não máxima em uma vizinhança 3x3x3. A supressão não máxima consiste em encontrar pontos que se sobressaem em relação à sua vizinhança.

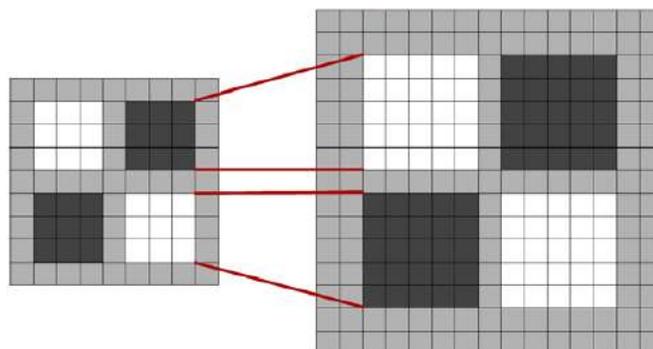


Figura 13 - Filtros em escalas diferentes, 9x9 (esquerda) e 15x15 (direita) (BAY et al., 2008).

A Fig. 14 apresenta o resultado obtido após a execução da primeira parte do método, em que é possível observar que dentre os vários círculos, os seus centros representam os pontos de interesse que foram detectados na imagem. O tamanho do círculo está diretamente ligado à escala em que o ponto foi encontrado.

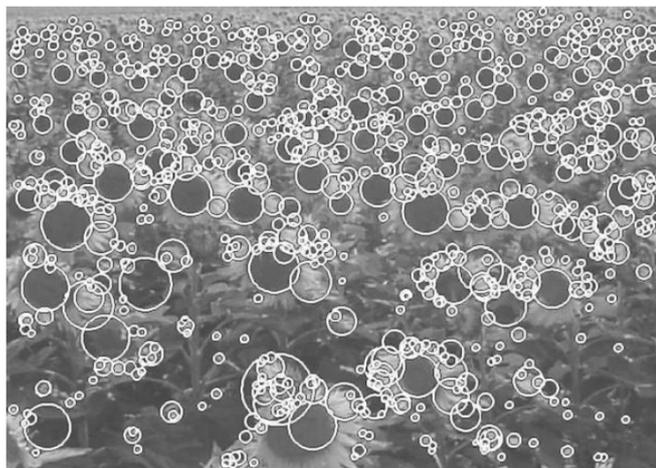


Figura 14 - Pontos de interesse detectados para um campo de girassol (BAY et al., 2008).

3.2.2 SVM

Uma técnica de classificação que tem recebido considerável atenção além de ter suas raízes na teoria de aprendizagem estatística é a SVM (*Support Vector Machines*), tendo sido proposto inicialmente por CORTES e VAPNIK (1995). Para CANAHUIRE CABELLO (2016), o classificador SVM serve unicamente para classificar conjuntos que sejam linearmente separáveis e que sejam completamente separáveis sem superposição entre as classes. Já para JURASZEK (2014), é possível a separação dos dados não lineares de forma linear desde que a SVM utilize os *kernels* para efetuar manipulações nos dados, tornando-os linearmente separáveis. A utilização dos *kernels* de forma correta exige que a distribuição dos dados seja conhecida. JURASZEK (2014) ainda afirma que existem diversos tipos de *kernels* disponíveis, entre eles o linear, polinomial, gaussiano ou radial e *sigmoid*.

Na Fig. 15 é mostrado um gráfico de um conjunto de dados que contém exemplos que pertencem a duas classes diferentes, representadas como quadrados e círculos (CANAHUIRE CABELLO, 2016). Se um conjunto de dados é linearmente separável, isto significa que se pode encontrar um hiperplano de tal modo que todos os quadrados residam em um lado da linha e todos os círculos residam no outro lado. No entanto, como mostrado na Fig. 11, há um número infinito de possíveis hiperplanos. Para saber qual é o melhor hiperplano, pode-se intuitivamente definir um critério para estimar o valor das linhas.

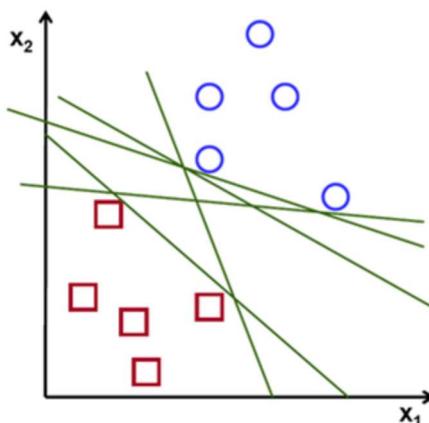


Figura 15 – Duas classes linearmente separáveis (CANAHUIRE CABELLO, 2016).

Na maior parte dos problemas reais, os dados não podem ser separados por uma superfície de decisão linear. Uma possibilidade para esses casos, são a aplicação de uma transformação $\phi: \mathbb{R}^p \rightarrow \mathbb{R}^q$ no espaço de características, de forma que os dados, neste novo espaço de dimensão q ($q > p$) sejam linearmente separáveis. No entanto, a aplicação direta da transformação pode tornar o treinamento das SVMs computacionalmente intratável, caso a dimensão do espaço transformado \mathbb{R}^q cresça muito rapidamente em relação ao espaço \mathbb{R}^p (por exemplo, na transformação polinomial, usualmente adotada com SVMs, $q = O(2^p)$). Um importante resultado na teoria subjacente às SVMs é que o problema de otimização utiliza as instâncias de treinamento somente para computar os produtos internos entre os vetores de características, $x_i \cdot x_j$, $i, j \in \{1 \dots N\}$. Portanto, se for definida uma função kernel $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$, não será necessário aplicar explicitamente a transformação sobre cada instância individualmente, mas apenas computar a matriz de valores $K(x_i, x_j)$ para todos os pares de vetores (x_i, x_j) do conjunto de treinamento antes de resolver o problema de otimização (BURGES, 1998).

De acordo com BISHOP (2006), em função dos ruídos e *outliers*, os dados de entrada raramente são linearmente separáveis em aplicações do mundo real. Em razão disso, o problema de otimização pode ser alterado de modo a permitir alguns erros de classificação, criando uma relação entre o resultado da classificação e a capacidade de generalização. Nessa abordagem do SVM há suavização das restrições do problema de otimização. As variáveis de folga relaxam as restrições impostas ao problema de otimização.

3.2.3 *k-Means*

O *k-Means* padrão (também referenciado como *standard* ou como básico) foi proposto em 1957 por Stuart Lloyd (LLOYD, 1957), em um documento interno à *Bell Telephone Laboratories*, como uma técnica para a modulação *pulse-code*, numa aplicação de quantização vetorial. Já o *k-Means* como algoritmo de agrupamento foi publicado em um artigo escrito por MACQUEEN (1967), sendo o precursor desse tipo de aplicação, cujo principal objetivo é o de agrupar um conjunto de dados em k conjuntos diferentes, onde cada observação venha a pertencer ao grupo mais próximo da média.

O algoritmo *k-Means* é do tipo não supervisionado, ou seja, que não trabalha com dados rotulados, e utiliza um método simples e eficiente para um agrupamento pleno. O objetivo desse algoritmo é encontrar similaridades entre os dados e agrupá-los conforme o número de clusters (grupos) passado pelo argumento k .

De acordo com MATTE (2020) esse algoritmo utiliza uma heurística que aplica uma abordagem gulosa para otimizar a função SQE (Soma dos Quadrados dos Erros), convergindo, assim, para um agrupamento ótimo local (mínimo local) e não global. O algoritmo *k-Means* busca uma partição do conjunto de instâncias que tenha k grupos, tal que a soma dos quadrados dos erros entre os centróides e as instâncias a eles atribuídas seja minimizada; o algoritmo, porém, não garante que o agrupamento induzido é o melhor agrupamento possível (JAIN, 2010).

MATTE e DO CARMO NICOLETTI (2019) afirmam que o algoritmo *k-Means* é reconhecidamente um dos algoritmos de Aprendizado de Máquina mais populares, dentre os algoritmos caracterizados como algoritmos de agrupamento.

Por ser um algoritmo que já vem sendo aplicado há muitos anos, vários pesquisadores estão tentando produzir novos métodos que sejam mais eficientes do que os métodos existentes e apresentem resultados melhores de classificação. Um desses trabalhos é apresentado por HAMERLY (2015), com objetivo de tornar o *k-Means* mais rápido, ele propõe métodos geométricos simples, baseados na desigualdade do triângulo em que são mantidos os limites nas distâncias computadas para reduzir o desperdício de computação e construir algoritmos mais eficientes que fornecem exatamente a mesma saída.

PUROHIT e JOSHI (2013) introduziram uma nova abordagem eficiente para o algoritmo de agrupamento *k-Means*. Eles propuseram um novo método para gerar o centro do cluster, reduzindo o erro quadrático médio do cluster final sem grande incremento no tempo de execução. Assim, é possível reduzir o erro quadrático médio sem sacrificar o tempo de execução. Muitas comparações foram feitas e pode-se concluir que a precisão é melhor para um conjunto de dados denso do que para um conjunto de dados esparsos.

3.2.4 BoVW

O método *Bag of Visual Words* (BoVW), proposto por CSURKA et al. (2004), é inspirado no método *Bag of Words* (DUMAIS et al., 1998 e BAEZA-YATES e RIBEIRO-NETO, 1999) da área de categorização de textos, em que um documento é classificado de acordo com a sua frequência de palavras. Em BoVW uma imagem é classificada de acordo com a sua frequência de palavras visuais. Esse é um método simples, flexível, robusto e com baixo custo computacional. Além de se apresentar invariante a escala, rotação e translação e parcialmente invariante a ruído, iluminação, ponto de vista, oclusão e *clutter* (LIMA, 2017).

Esse método também conhecido como *Bag of Features*, *Bag of Visual Features* ou *Bag of Keypoints*, tem se tornado, nas últimas décadas, uma abordagem bastante popular nas mais diversas áreas de visão computacional, como descritor de imagens, busca de vídeos, reconhecimento de textura, entre outras tarefas (PEDROSA, 2015).

Uma das importantes decisões quando se está aplicando o método BoVW se refere à escolha do detector de pontos de interesse que será utilizado, onde se destacam o SIFT e o SURF como os dois métodos mais populares para a construção do dicionário visual, que tem por objetivo determinar quais são as características e padrões que representam a estrutura de uma imagem. A seleção do detector depende de duas escolhas principais, sendo a primeira delas a escolha da quantidade de palavras visuais. Apesar de ser um fator importante, que interfere diretamente no desempenho da abordagem *bag of visual words*, geralmente é definida empiricamente pelo usuário. A outra escolha é do algoritmo de agrupamento, onde se destaca o algoritmo *k-Means*, que é um método simples e bastante utilizado na literatura (PEDROSA, 2015).

Como uma forma de melhor explicar a aplicação do método de BoVW será apresentado a seguir um exemplo de uma das formas em que este pode ser aplicado. Para esse exemplo será utilizado um banco de dados de imagens termográficas de um mancal de rolamento, conforme ilustrado na Fig. 16, já em tons de cinza, tendo em vista que o detector utilizado nesse exemplo será o SURF.

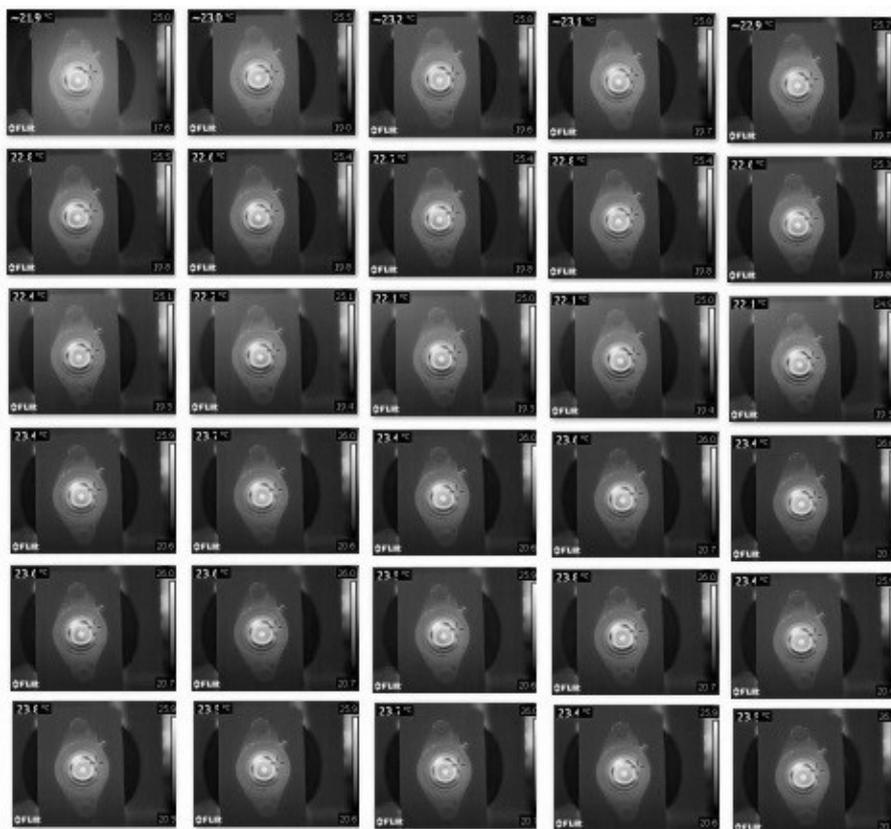


Figura 16 – Banco de dados de imagens termográficas de um mancal de rolamento (Fonte: autor).

A primeira etapa desse processo é a busca das “palavras visuais”, que pode ser entendido como as regiões de uma imagem que a caracterizam. Alguns dos atributos que servem para caracterizar essa imagem são os tons de cinza, cantos e bordas. Em geral, uma imagem possui um conjunto de pontos de interesse que são considerados relevantes para sua caracterização. Para efeito desse exemplo, um atributo relevante seria a variação dos tons de cinza, tendo em vista que a principal característica de uma imagem termográfica é justamente o gradiente de temperatura, que nesse caso é representado pela variação dos tons de cinza.

Através da Fig. 17, pode-se ter uma visão geral do processo de criação do vetor de características de um ponto de interesse da imagem, quando se está aplicando o método BoVW. Para cada ponto de interesse identificado, é definida uma região 16 x 16 ao seu redor, considerando sua orientação e escala. Em 17a está ilustrada a imagem termográfica em tons de cinza, para que possa se utilizar o detector SURF. Em 17b está apresentada a extração dos pontos de interesse da imagem com a aplicação do detector SURF, com base na variação das tonalidades da imagem. Em 17c está ilustrada a orientação e escala do ponto de interesse analisado. Em 17d é possível observar a região já dividida em grades 4 x 4, considerando as orientações da vizinhança na escala que o ponto foi detectado. Em 17e é calculado, para cada grade 4 x 4, um histograma de 8 direções levando em consideração as magnitudes dos gradientes da vizinhança. E por fim, os histogramas obtidos em cada grade são concatenados, formando o vetor de características daquele ponto de interesse.

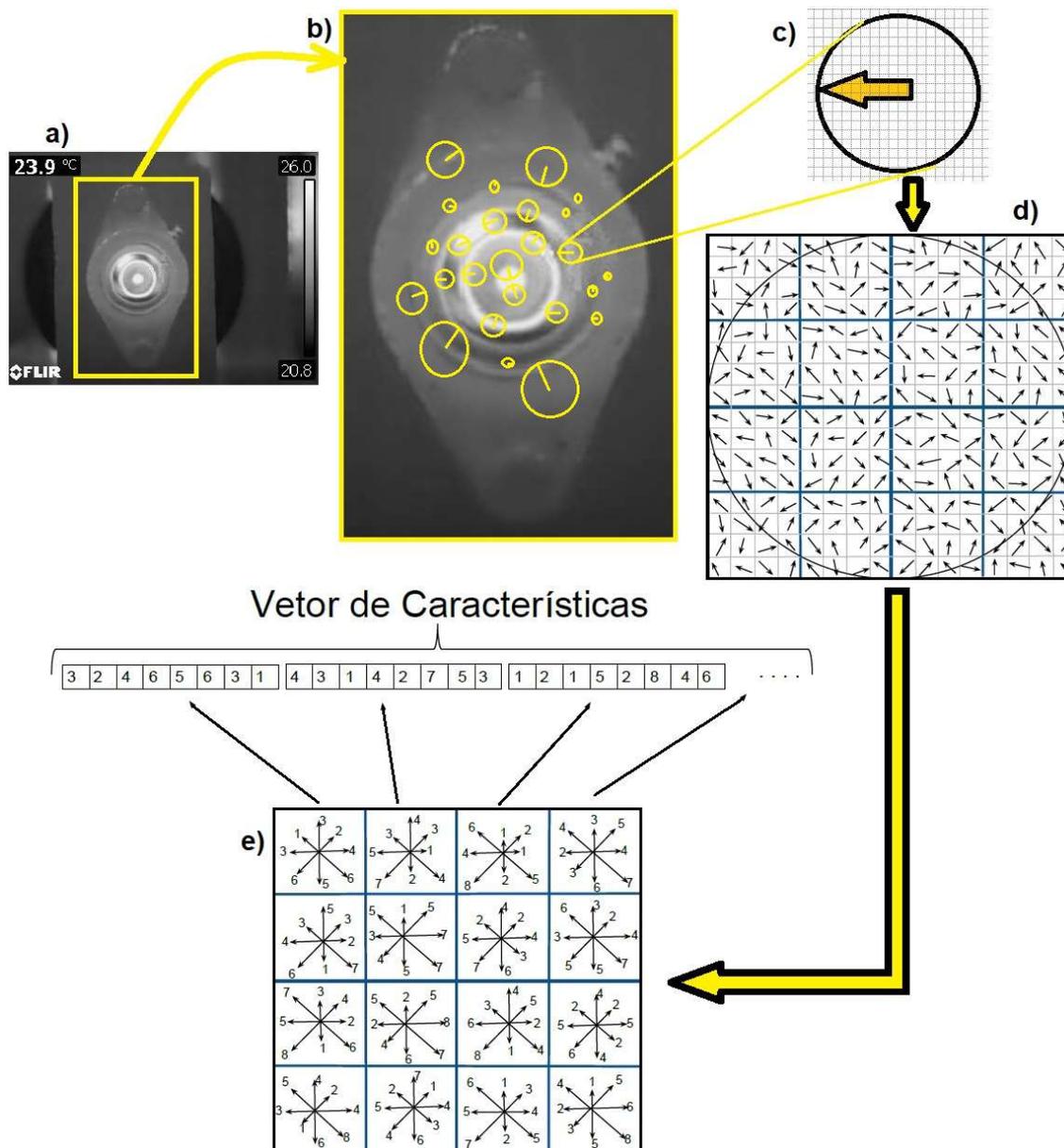


Figura 17 – Criação do vetor de características de um ponto de interesse da imagem (adaptado de SOARES, 2012).

Em seguida, é criado o vocabulário de palavras visuais, com base nos pontos de interesse já identificados e descritos. Nesse ponto, um dos algoritmos de agrupamento (*clustering*) mais utilizados e que pode ser aplicado é o *k-Means*, para criar esse vocabulário. O algoritmo *k-Means* consiste em gerar k grupos, onde cada grupo contém um representante, denominado centroide (SOARES, 2012).

Inicialmente, k centroides são aleatoriamente criados. Cada dado é atribuído a um grupo, de acordo com a sua distância ao centroide mais próximo, de forma que a cada iteração os centroides sejam recalculados e, conseqüentemente, é gerado um novo agrupamento. Isso se repete até que os valores dos centroides não se alterem mais ou um número máximo de iterações seja alcançado (SELIM AND ISMAIL, 1984).

No BoVW, cada centroide representa uma palavra visual no vocabulário de palavras visuais. Dessa forma, o tamanho do vocabulário de palavras visuais fica limitado ao valor de k , parâmetro do *k-Means*, que representa a quantidade de grupos a serem gerados ao final da execução do algoritmo (KANASHIRO, 2016).

Ainda de acordo com KANASHIRO (2016), a última etapa na aplicação do método BoVW consiste na geração do histograma das palavras visuais da imagem, em que os pontos de interesse foram rotulados pela etapa anterior. Esse histograma irá possuir o tamanho do vocabulário de palavras visuais e cada posição representa a frequência em que cada palavra visual ocorre na imagem, conforme ilustrado na Fig. 18. Assim, o histograma gerado possui importantes atributos que serão utilizados na etapa de classificação das imagens.

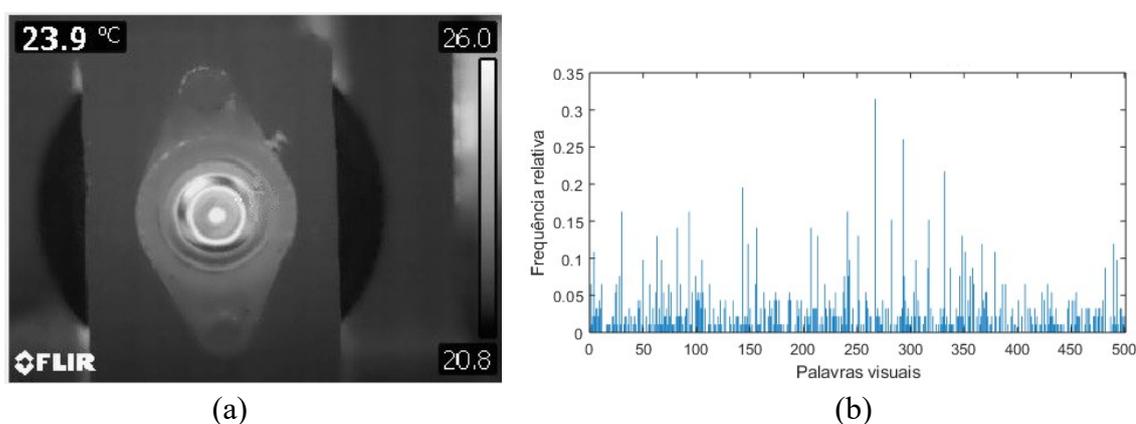


Figura 18: Exemplo de uma imagem em tons de cinza (a) e um histograma da frequência relativa de ocorrência de cada uma das 500 palavras visuais (b) (Fonte: autor).

Apesar do BoVW utilizar técnicas conhecidas em suas etapas, o resultado final da informação (o descritor da imagem) é diferente da aplicação dessas técnicas de forma isolada, pois, simplifica a representação dessa imagem através de um vetor de características (palavras) e suas frequências, tendo em vista que o BoVW não é um classificador, e sim um descritor. O BoVW é um método que representa a imagem de uma forma diferente. Para classificar essa imagem pode ser utilizado qualquer classificador como o SVM, ou uma rede neural, por exemplo.

3.3 Desalinhamento de eixos

Pode-se afirmar que um eixo apresenta desalinhamento quando a linha de centro do eixo motor não está alinhada com o eixo movido. Tal desconformidade propicia desgaste prematuro das partes devido a um aumento no nível de vibração do conjunto.

O desalinhamento de eixos é largamente reconhecido como o segundo tipo de falha mais comum em máquinas rotativas, precedido apenas pelo desbalanceamento (MUSZYNSKA, 2005). Aproximadamente 70% dos problemas de vibrações em máquinas rotativas são causados por desalinhamento de eixo (BOGNATZ, 1995), que

leva ao tempo de inatividade de até 30% na indústria, pois, essa falha é um dos problemas mais comuns na operação de máquinas rotativas, que são o coração de qualquer indústria (HARIHARAN e SRINIVASAN, 2009).

De acordo com ARCO-VERDE (2008), existem dois tipos de desalinhamento: paralelo e angular. O desalinhamento paralelo é caracterizado quando as linhas de centro dos eixos das máquinas não coincidem, ou seja, não estão alinhadas, embora possam estar paralelas entre si. O desalinhamento angular ocorre quando as linhas de centro do eixo das máquinas se interceptam, formando um ângulo entre si, de forma que as linhas são concorrentes. Alguns autores consideram a combinação destes dois tipos de desalinhamento como um terceiro tipo de desalinhamento (EISENMANN e EISENMANN JR., 1997).

NEPOMUCENO (1989) explica que diferentemente do desbalanceamento, onde a vibração tem componentes máximos radiais, no desalinhamento angular por exemplo, as componentes com amplitude máxima são as axiais.

Os métodos de alinhamento de eixo que são utilizados em máquinas rotativas, conforme ilustrado da Fig. 19, são a régua com auxílio de calibrador de folga em 19a, considerado o mais simples e de menor precisão, o alinhamento com relógios comparadores em 19b e o que utiliza equipamento dotado de laser em 19c, que é considerado o de melhor precisão.

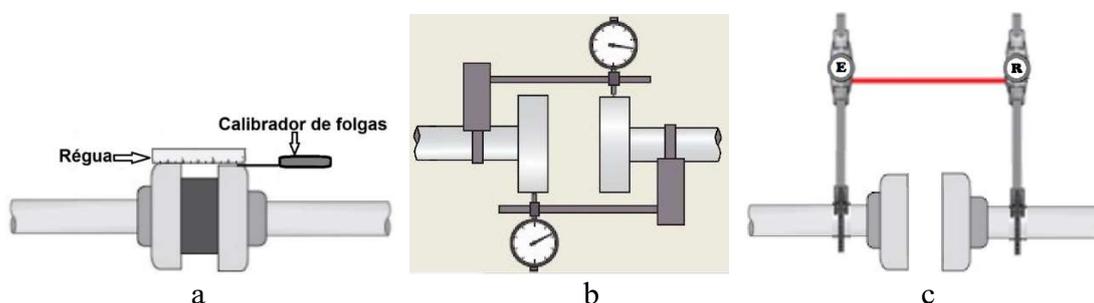


Figura 19 – a) Alinhamento com régua; b) Alinhamento com relógio comparador; c) Alinhamento a laser (adaptado de <http://www.flowcontrolnetwork.com/the-importance-of-shaft-alignment/>).

De acordo com PIOTROWSKI (2006), os efeitos secundários do desalinhamento podem se tornar críticos, pois danificam lentamente a máquina ao longo do tempo. Alguns dos sintomas mais comuns quando se tem desalinhamento de eixos em máquinas rotativas são os seguintes (PIOTROWSKI, 2006):

- Falhas prematuras de rolamento, vedação, eixo ou acoplamento;
- Temperaturas elevadas nos rolamentos ou próximo a eles ou altas temperaturas do óleo de descarga;
- Quantidade excessiva de vazamento de lubrificante nas vedações do rolamento;
- Certos tipos de acoplamentos flexíveis exibirão temperaturas mais altas do que o normal durante o funcionamento ou estarão com temperaturas elevadas imediatamente após o desligamento da unidade;
- Equipamentos semelhantes apresentam vidas úteis diferentes;

- Número excepcionalmente alto de falhas do acoplamento ou desgaste prematuro;
- Quebra do eixo (ou trinca) em regiões próximas aos rolamentos e próximas aos cubos do acoplamento;
- Quantidade excessiva de graxa (ou óleo) no interior da proteção do acoplamento;
- Parafusos de fixação da base na fundação soltos, normalmente causados por uma condição de contato inadequado entre a base e o piso, que são agravados pelo desalinhamento;
- Parafusos do acoplamento soltos ou quebrados. Isso ocorre frequentemente devido ao torque incorreto dos parafusos do acoplamento, o que é agravado pela condição do desalinhamento.

3.4 Mancais de rolamento

Os mancais de rolamento têm uma aplicação muito vasta no mundo moderno, pela sua importância de servir de apoio para eixos girantes de máquinas rotativas, oferecendo na maioria das vezes uma alta precisão e baixo atrito, permitindo assim altas velocidades de rotação.

Os rolamentos são, de uma forma geral, constituídos por anéis, corpos rolantes e gaiola, além de serem divididos em rolamentos radiais e rolamentos axiais. Mas em função do tipo de corpo rolante, são divididos principalmente em rolamentos de esferas e rolamentos de rolos (NSK, 2013), conforme ilustrado nas Figuras 20 e 21.

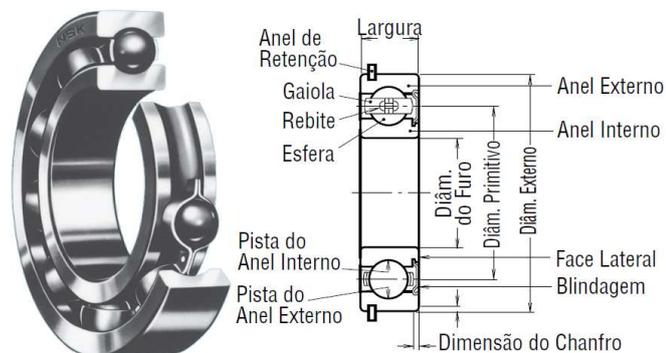


Figura 20 - Rolamento fixo de uma carreira de esferas (adaptado de NSK, 2013).

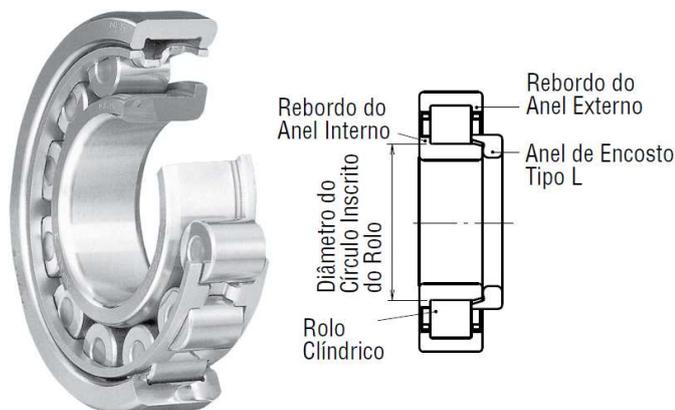


Figura 21 - Rolamento de rolos cilíndricos (adaptado de NSK – 2013).

3.4.1 Falhas em mancais de rolamento

Pode-se definir modo de falha como a maneira em que as falhas ocorrem, representando a perda da função requerida (IEC - do inglês *International Electrotechnical Commission*). Já a falha está associada à função do sistema, enquanto que o modo de falha está associado ao evento que provoca a transição para o estado de falha (SIQUEIRA, 2009).

Para CALLISTER (2018), a falha de materiais de engenharia é quase sempre um evento indesejável por várias razões, pois, podem colocar em risco vidas humanas, ocasionar perdas econômicas, além de causar interferência na disponibilidade de produtos e serviços. Embora as causas das falhas e o comportamento dos materiais possam ser conhecidos, existe uma grande dificuldade na prevenção dessas falhas.

Enquanto houver uma lubrificação adequada, a falha dos mancais de rolamento será por fadiga, dentro da vida útil do material. A falha será considerada prestes a ocorrer quando a pista ou as esferas (rolos) exibirem a primeira vala. Normalmente, quem falhará primeiro será a pista. O mancal de rolamento dará um aviso audível de que a cavitação começou emitindo ruído e vibração, mas poderá continuar funcionando além desse ponto, enquanto sua superfície continuará a se deteriorar até que ocorra sua fragmentação ou fratura (NORTON, 2013).

Para CALLISTER (2018), em maior ou menor grau, a maioria dos materiais apresenta algum tipo de interação com um grande número de ambientes diferentes. Essas interações ambientais podem comprometer a utilidade de um material como resultado da deterioração de suas propriedades mecânicas, físicas ou de sua aparência. Nos metais, existe uma efetiva perda de material, quer pela dissolução (corrosão) ou pela formação de uma incrustação ou filme não metálico (oxidação).

A corrosão pode ser definida como um ataque eletroquímico destrutivo e não intencional de um metal, começando normalmente na superfície. O problema da corrosão metálica é de proporções significativas, de forma que ao se analisar as questões econômicas, é estimado que aproximadamente 5% das receitas de uma nação industrializada são gastos na prevenção da corrosão e na manutenção ou substituição de produtos perdidos ou contaminados como resultado de reações de corrosão (CALLISTER, 2018).

De acordo com a NSK (2013) os rolamentos quando são utilizados de forma adequada e dentro de suas especificações para uso podem ser usados por um longo período, em geral, até a vida de fadiga, porém, há casos de ocorrências inesperadamente rápidas que não permitem a utilização continuada. Estas ocorrências prematuras em relação à vida de fadiga são os limites de uso, naturalmente denominadas de quebras ou acidentes, que na sua grande maioria têm como causas a falta de cuidados quanto a instalação, utilização e lubrificação; a penetração de partículas estranhas do exterior e a falta ao considerar a influência do calor no eixo e alojamento.

Dentre as diversas falhas que os rolamentos podem apresentar, destaca-se, na Fig. 22, as falhas por oxidação, em que se pode citar como causas principais os seguintes motivos (NSK, 2013):

- Entrada de gás corrosivo ou água;
- Lubrificante inadequado;
- Formação de partículas de água por condensação de umidade;
- Alta temperatura e alta umidade quando a peça está parada;
- Falha do protetivo contra corrosão durante o transporte e estocagem;
- Manuseio inadequado.



Figura 22 - Oxidação na face do rebordo, na superfície da pista e no espaçamento das esferas (adaptado de NSK – 2013).

De acordo com MENNA (2007), a oxidação de um rolamento ocorre em regiões dos anéis e elementos rolantes, que possuem condições eletroquímicas favoráveis à ocorrência da corrosão galvânica.

3.5 Acoplamentos

O acoplamento faz a ligação entre os eixos de máquinas rotativas acionadoras e acionadas e tem como principal função a transmissão de torque. Além dessa função, o acoplamento também compensa desalinhamentos, absorve choques e vibrações, além de atuar como fusível mecânico.

O acoplamento entre eixos motores e movidos tem sido estudado de forma bastante intensa, com objetivo de atingir resultados cada vez mais significativos, entretanto, de acordo com SEKHAR e PRABHU (1995) a condição de “alinhamento perfeito” não pode ser atingida.

De acordo com TUCKMANTEL (2018), os acoplamentos mecânicos são usualmente divididos em duas categorias, rígidos e flexíveis. Os acoplamentos rígidos, conforme ilustrado na Fig. 23, travam os dois eixos conjuntamente não permitindo movimento relativo entre eles, o que implica que na prática nenhum desalinhamento é permitido, embora algum ajuste axial seja possível durante a montagem.



Figura 23: Acoplamentos rígidos: (a) Flange, (b) Luva, (c) Luva de compressão, (d) *Quill shaft* (TUCKMANTEL, 2018).

Os acoplamentos flexíveis, conforme ilustrado na Fig. 24, têm como mecanismo de compensação os movimentos relativos entre os elementos que o compõem, formando elos, juntas e pivôs. Estas conexões móveis em seu interior são obtidas através de folga mecânica ou de rolamento ou deslizamento entre os elementos. Os acoplamentos com flexibilidade mecânica são os mais utilizados na indústria e, normalmente, o tempo de vida útil varia entre 3 e 5 anos. Geralmente esses acoplamentos requerem lubrificação, exceto quando uma das peças possui lubrificação oriunda do material. Devido ao seu mecanismo de operação, o modo de falha mais comum é o desgaste (TUCKMANTEL, 2018).

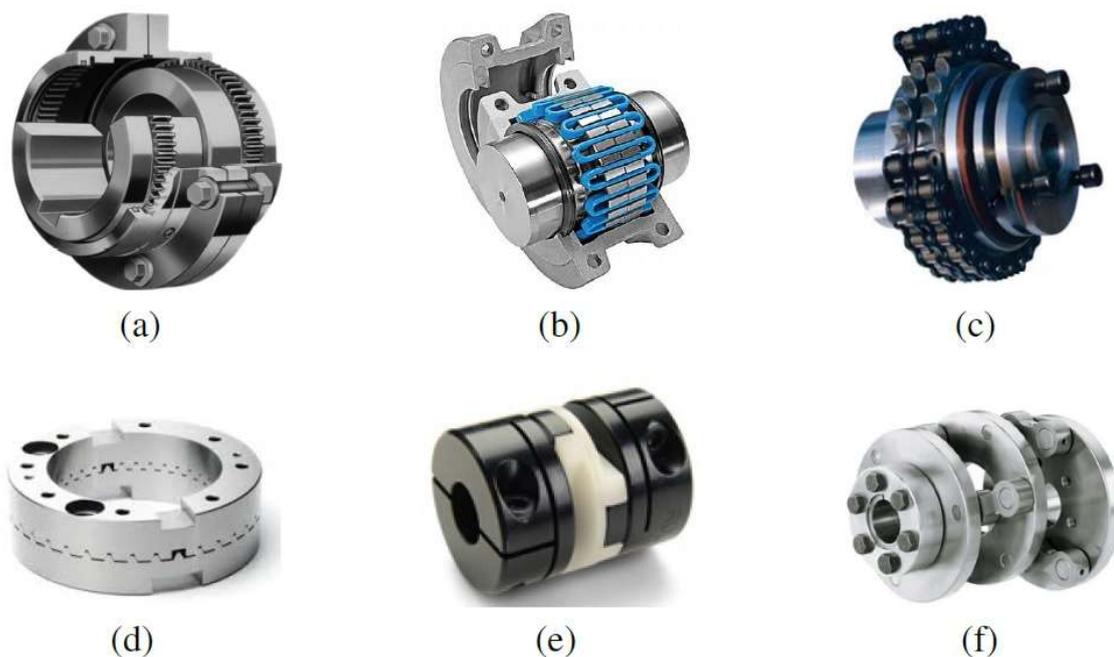


Figura 24: Acoplamentos com flexibilidade mecânica: (a) Engrenagens, (b) *Falk*, (c) Corrente de rolos, (d) *Hirth*, (e) *Oldham* e (f) *Schmidt* (TUCKMANTEL, 2018).

3.6 Desbalanceamento

De acordo com TUCKMANTEL (2018), o desbalanceamento de massa é reconhecidamente a mais significativa fonte de excitação em máquinas rotativas, é uma

força periódica com frequência de vibração síncrona. Devido ao centro de massa não coincidir com o centro de rotação, forças radiais de inércia são geradas no rotor.

As vibrações relacionadas ao desbalanceamento do rotor de uma máquina rotativa, extraem energia da rotação à vibração, portanto, a eficiência geral da máquina diminui. O desbalanceamento causa vibração e tensão alternada ou variável no próprio rotor e nos elementos de sua estrutura de suporte. A vibração pode resultar em desgaste excessivo nas buchas dos mancais, fusos, engrenagens, etc., reduzindo substancialmente sua vida útil (MUSZYNSKA, 2005).

A representação de um conjunto contendo eixo, rotor e mancais é ilustrado na Fig. 25, onde observa-se um modelo de como um mancal é representado. Também se considera que o mesmo rotor possui um certo desbalanceamento residual.

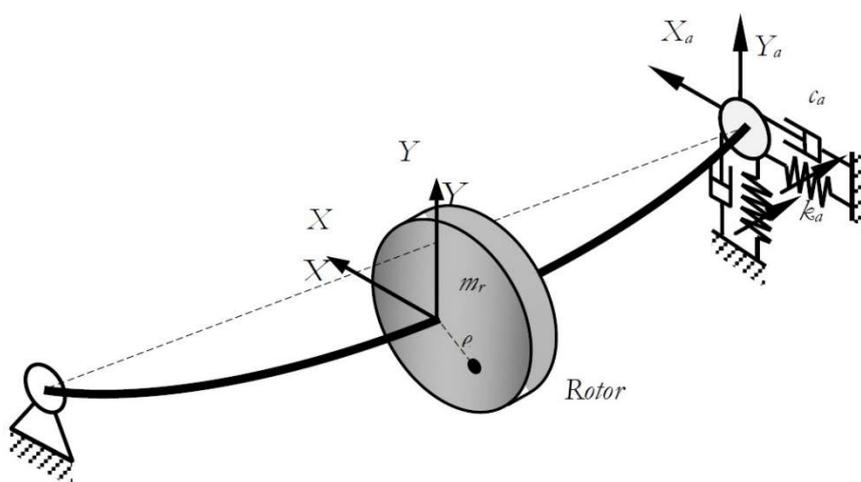


Figura 25: Sistema Rotor-Mancal (SENKO, 2013).

De acordo com LAMIM FILHO et al. (2006), referenciado por SENKO (2013), o fenômeno do desbalanceamento é caracterizado pela existência de desequilíbrios de massa em relação aos eixos de rotação. Esses desequilíbrios têm origem nas inevitáveis assimetrias, tolerâncias e desvios de forma, além das imperfeições da matéria-prima e da montagem. Quando o sistema está desbalanceado, pode gerar comportamentos indesejáveis, os quais podem ocasionar diversos danos a outros equipamentos que estão interligados (SENKO, 2013).

3.7 Validação cruzada

A técnica de validação cruzada pode ser feita de diferentes formas, sendo que os métodos mais usados são o *hold-out* (DEVROYE E WAGNER, 1979), *k-fold* (BURMAN, 1989), *Leaveuone-out* e *Boots-trap* (EFRON, 1983). Cada um desses métodos tem seus prós e contras, mas para efeito desse trabalho, serão detalhados a seguir os métodos *hold-out* e *k-fold*, que foram aplicados nessa pesquisa.

3.7.1 *Hold-out*

De acordo com DEVROYE E WAGNER (1979), o método *hold-out* também conhecido como validação simples, propõe que a amostra d seja dividida em duas partes, usando uma proporção p dela como amostra de validação. Formalmente, dado um conjunto de dados d , separa-se uma proporção p dos dados, criando-se a amostra de treino dt onde $t = n * (1 - p)$ e a amostra de validação dv , de tamanho $v = n * p$. O estimador *hold-out* de (3.7) é dado por:

$$hop^{-1} = \frac{1}{v} \sum_{i=1}^v L(y_i, \hat{g}_t(X_i)) \quad (3.6)$$

em que $\hat{g}_t(X)$ e preditor criado com a amostra de treino dt e a função de perda é avaliada em todos os pontos (y_i, x_i) da amostra de validação dv .

A equação 3.7 representa uma medida que leva em conta a perda em todos os pontos possíveis da função, sendo chamada de risco esperado (BORRA e DI CIACCIO, 2010). Ela leva em conta o preditor $\hat{g}(X)$ e a função $L(Y, \hat{g}(X))$, sendo definida como:

$$Err = R(\hat{g}(X), Y) = E_X E_{Y|X} [L(Y, \hat{g}(X)) | \hat{g}(X), X] \quad (3.7)$$

ou seja, foi calculado o erro esperado sob todos os possíveis valores do vetor de covariáveis X e Y para uma $g(X)$ fixa. Essa medida é sempre positiva, embora não seja necessariamente finita, pois a distribuição das variáveis pode não ter valor médio.

3.7.2 *K-fold*

O método *k-fold*, é baseado no método anterior, mas com maior utilidade quando o conjunto de dados é pequeno (Yang e Huang, 2014). Nesse caso, os dados totais são divididos em k subconjuntos, então aplicamos o método *hold-out* k vezes, cada vez usando um subconjunto diferente para validar o modelo treinado com os outros $k - 1$ subconjuntos (Jung e Hu, 2015).

De acordo com MCLACHLAN e DO (2005), na validação cruzada *k-fold*, a amostra original é dividida aleatoriamente em k subamostras de tamanhos iguais. Dessas k subamostras, l subamostras são retida como dados de validação para testar o modelo e as $k - l$ subamostras restantes são usadas como dados de treinamento. O processo de validação cruzada é então repetido k vezes, com cada uma das k subamostras usadas exatamente uma vez como dados de validação. Os k resultados podem então ser calculados para produzir uma única estimativa. A vantagem desse método sobre a subamostragem aleatória repetida é que todas as observações são usadas para treinamento e validação, e cada observação é usada para validação exatamente uma vez. A validação cruzada de 10 vezes é comumente usada, mas em geral k permanece um parâmetro não fixo (MCLACHLAN e DO, 2005).

Capítulo 4

Metodologia

Neste capítulo é descrita a metodologia utilizada para aquisição e processamento dos dados necessários para testar a hipótese de que um método que obteve bom desempenho para detecção de falhas em máquinas rotativas a partir de imagens termográficas com média ou alta resolução, também alcançará bom desempenho se forem utilizadas imagens adquiridas por câmeras termográficas de baixa resolução. Também será testada a hipótese de que os defeitos simulados e detectados pela técnica de análise de vibração também serão identificados e separados em categorias pela técnica de termografia, a partir do processamento das imagens adquiridas por câmeras termográficas de baixa resolução.

Dentre os métodos estudados, e mencionados no capítulo 2 – Revisão Bibliográfica, o método escolhido foi o BoVW (CSURKA et al., 2004), pois o mesmo alcançou desempenho acima de 99% em termos de acurácia e precisão para detecção de falhas em máquinas rotativas a partir de imagens termográficas, conforme apresentado no artigo de JIA et al. (2019), além de exigir menos recursos computacionais, se comparado a CNN, o que o torna mais acessível.

O diagrama de blocos da abordagem implementada neste trabalho está ilustrado na Fig. 26. De uma forma simplificada, é possível verificar que a partir da coleta das imagens termográficas, são identificadas e categorizadas as regiões de interesse que representam cada imagem com a aplicação do SURF. Em seguida, é criado o vocabulário de palavras visuais, com base nos pontos de interesse já identificados e descritos, o que gera o histograma de frequência das palavras visuais, que representa a aplicação do método de BoVW para tratamento dos dados e categorização das falhas estudadas.

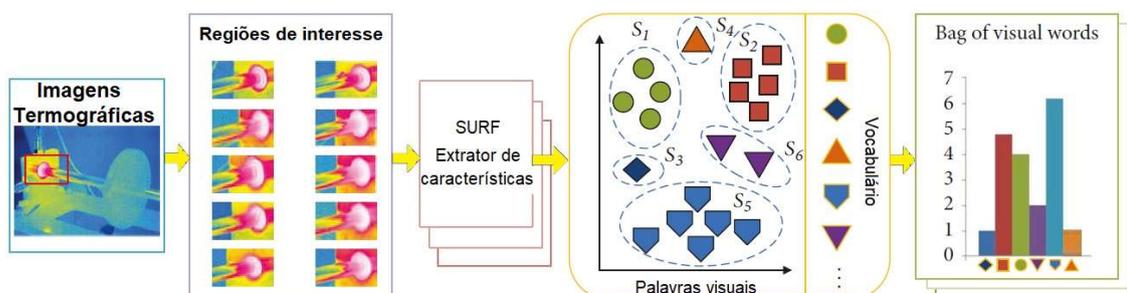


Figura 26 – Diagrama de fluxo - BoVW (adaptado de LI et al., 2019).

Na etapa de coleta das imagens termográficas, com um total de 150 imagens para cada uma das categorias de falhas estudadas, foi utilizado um simulador de falhas,

conforme ilustrado na Fig. 27. Após adquirir estabilidade operacional com 20 minutos de operação, as imagens termográficas do mancal de rolamento foram coletadas sequencialmente no mancal que fica na extremidade do simulador de falhas, com a câmera posicionada a 40 cm de distância desse mancal, em três velocidades de rotações diferentes (250 rpm, 500 rpm e 750 rpm) além da inserção de falhas prévias no simulador, como será detalhado adiante. O *layout* de instalação é ilustrado na Fig. 28.

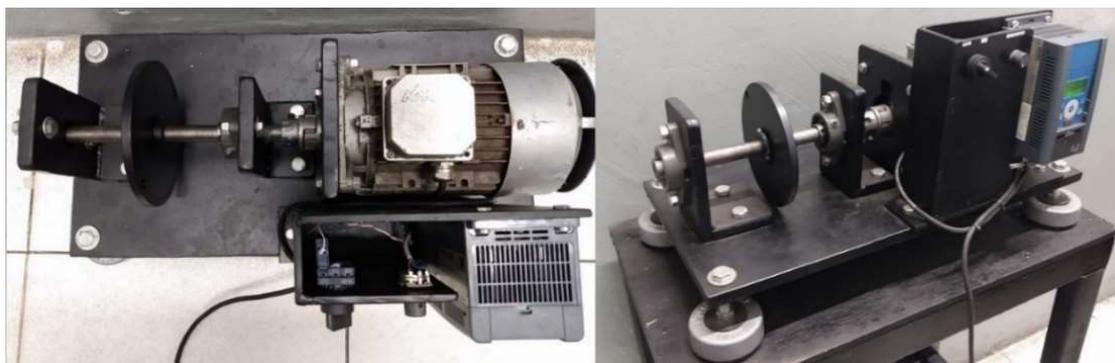


Figura 27 – Simulador de falhas.

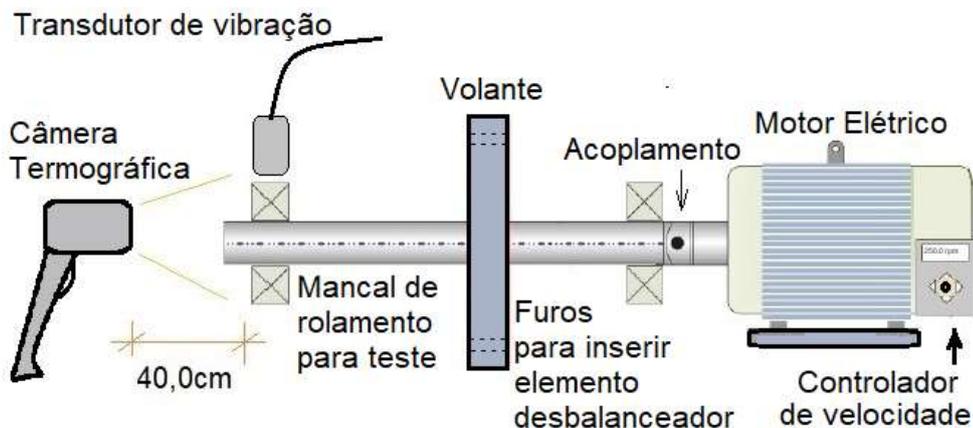


Figura 28 – Layout de instalação do simulador de falhas.

As câmeras termográficas utilizadas na coleta de dados foram uma Fluke VT02 e FLIR T420, conforme ilustradas na Fig. 29 e suas especificações na Tabela 1. Para padronização no método de coleta de imagens termográficas, foi consultada a Norma NBR 15572 para Termografia. A coleta de dados foi realizada em uma sala com temperatura ambiente estável em 25 °C e a iluminância mantida em 750 lux.



Figura 29– Câmeras termográficas (a) FLUKE VT02 – Resolução: 11025 pixels
(b) FLIR T420 – Resolução: 76800 pixels.

Tabela 1 – Especificações das câmeras (Fonte: Manual do usuário)

Especificações	FLUKE VT02	FLIR T420
Faixa de Medição de Temperatura	-10 °C a +250 °C	-20 °C a +1200 °C
Precisão da medição de Temperatura	±2 °C ou ±2 %	±2 °C ou ±2 %
Frequência de captura de imagem	8Hz	60Hz
Sensibilidade Térmica (NETD)	≤ 250 mK	45 mK
Banda espectral infravermelha	6.5 μm a 14 μm	7.5 μm a 13 μm
Resolução	105x105 (11025) pixels	320x240 (76800) pixels
Campo de visão (FOV)	20° x 20°	25° x 19°

A sensibilidade térmica (NETD - *Noise Equivalent Temperature Difference*) das câmeras representam a menor diferença de temperatura possível, que pode ser medida pela câmera termográfica, ou seja, quanto menor for este valor, melhor será a sensibilidade térmica de medição da câmera termográfica (TESTO, 2017). Para efeito dos experimentos realizados nesse trabalho, essa diferença de sensibilidade térmica entre as duas câmeras em referência pode ser desconsiderada, tendo em vista que, foi utilizada uma quantização nesses níveis de sensibilidade igual para as duas câmeras. Ou seja, as imagens adquiridas com as duas câmeras foram convertidas para 256 níveis de intensidade, antes de alimentarem a entrada do algoritmo.

Tabela 2 – Especificações do Simulador de Falhas

Especificações	Quantidade
Eixo em aço carbono: $\varnothing = 20,0$ mm / Comprimento = 300 mm	1
Rolamento FRM Y204 V22	2
Mancal de rolamento FRM FT 204	2
Volante em aço carbono: $\varnothing = 160,0$ mm / 4 furos a 90°	1
Acoplamento de alumínio Rotex KTR 19	1
Motor elétrico de indução trifásico 4 polos: 0,55 kW	1
Inversor de frequência WEG CFW 08	1

Os dados foram coletados no mancal de rolamento, conforme ilustrado na Fig. 30, a partir da extremidade do simulador de falhas, inicialmente com uma câmera FLIR T420. Na sequência, foi repetido o mesmo procedimento, com as mesmas condições, para coletar as imagens termográficas com as mesmas falhas simuladas, porém utilizando a câmera FLUKE VT02.



Figura 30 – Coleta de dados.

Para efeito desse trabalho, foram simuladas falhas em máquinas rotativas muito comuns no ambiente fabril, a saber: desalinhamento de eixo, rolamento oxidado, acoplamento com elemento elástico defeituoso e desbalanceamento.

A primeira coleta de dados foi realizada com o eixo desalinhado em $0,25^\circ$ para as três velocidades de rotações informadas anteriormente, conforme ilustrado na Fig. 31. Todas as simulações de falhas citadas a seguir também foram repetidas para as três velocidades de rotações informadas.

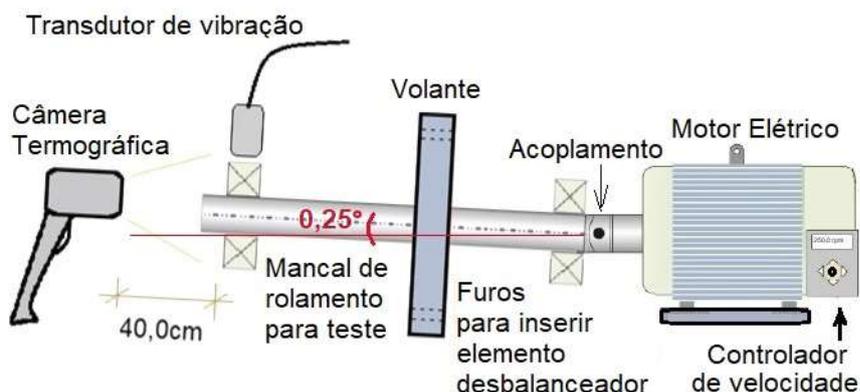


Figura 31 – Simulação de falha: eixo desalinhado.

Já a segunda coleta de dados foi realizada com um rolamento que apresentava o defeito de oxidação. Para terceira coleta de dados foi utilizado um elemento elástico do acoplamento do eixo com defeito. Na quarta coleta de dados foi simulada a falha de desbalanceamento do conjunto, a partir de um contrapeso com massa pré-determinada adicionado ao volante, conforme ilustrado na Fig. 32. A relação de massa entre o volante e o elemento desbalanceador foi responsável por gerar a condição de desbalanceamento. Como guia para padronização e análise da simulação em referência, foi consultada a norma ISO 21940-11.



Figura 32 – Detalhes do contrapeso adicionado ao volante para simular a condição de desbalanceamento do conjunto.

Por fim, na quinta e última coleta de dados não foi inserido nenhum tipo de falha ao simulador, de forma que este experimento fosse utilizado na condição normal de funcionamento.

Pode-se afirmar que cada pixel de uma imagem (11025 pontos de medição de temperatura por foto para a câmera Fluke VT02 e 76800 pontos para câmera FLIR T420) representa um valor de temperatura. Com base no princípio físico de que o atrito gera calor e que o defeito indicado acima produz uma condição diferente da condição normal de operação, seja aumentando ou diminuindo o atrito, diferenças de temperatura são esperadas em pontos muito particulares do conjunto de mancal e rolamento, dentro de um gradiente de temperatura entre esses pontos. Assim, foi necessário aplicar um método

capaz de identificar características muito específicas nos grupos de imagens analisadas que contemplam essas variáveis, chamadas de defeitos.

No apêndice 3 foi apresentado um estudo dimensional do efeito da distância focal e do campo de visão (FOV) de cada câmera termográfica em relação a distância do rolamento, de forma a subsidiar trabalhos futuros no ajuste das distâncias de cada câmera, em relação ao rolamento, de forma a compensar o tamanho do pixel. Dessa forma, apenas seria considerada a distância focal de acordo com o campo de visão de cada câmera, tendo em vista que a resolução teria a mesma equivalência, devido ser compensada por distâncias diferentes.

As imagens termográficas coletadas por meio da Câmera Fluke VT02 foram convertidas de .is2 para .jpg, utilizando o aplicativo da Fluke, *Smatview* versão 4.3.154.0. Já as imagens coletadas pela câmera FLIR T420 foram capturadas diretamente em formato jpeg. A conversão de todas as imagens para tons de cinza foi realizada a partir do algoritmo desenvolvido para essa aplicação da técnica de BoVW.

A partir da quantidade mínima de imagens disponíveis em cada categoria, foi definida a proporção de 70% / 30% a partir da técnica de validação *Hold-Out* (DEVROYE E WAGNER, 1979), também conhecida como validação simples, utilizando 105 imagens para treinamento e 45 imagens para teste de cada uma das 15 categorias estudadas, tendo em vista que a quantidade mínima de imagens coletadas por categoria foi de 150 imagens.

Na etapa seguinte, foi realizado o processamento das regiões de interesse das imagens termográficas, com a extração das características, aplicando SURF (BAY et al., 2006) em todas as categorias de imagens, construindo um vocabulário de características, e reduzindo o número de pontos de interesse através da quantização do espaço, que é realizado por meio do algoritmo *k-Means*. A escolha do SURF para esse trabalho, se justifica por ser mais rápido na extração e comparação de características (BAY et al., 2006), em relação ao SIFT, além do que, como a câmera termográfica FLUKE VT02 tem a opção de salvar as imagens já em tons de cinza, optou-se por utilizar esse método que emprega apenas características de imagens de tons de cinza (JURASZEK, 2014).

O primeiro passo foi definir uma orientação reproduzível com base nas informações de uma região circular em torno do ponto de interesse. Em seguida, foi construída uma região quadrada, definida em uma matriz 8x8, alinhada com a orientação selecionada na qual o descritor SURF foi extraído, e que utilizou as imagens de forma integral para calcular as respostas *Haar Wavelet* nas direções horizontal e vertical. Essas respostas foram amostradas em uma grade regular e suas somas (com e sem sinal) serviram como descritores das imagens em estudo (BAY et al., 2006).

Neste trabalho, o *k-Means* (JAIN, 2010) foi utilizado para fornecer uma classificação automática sem a necessidade de supervisão humana, ou seja, sem qualquer pré-classificação existente, além de apresentar facilidade em sua implementação (JAIN et al., 1999) e ser um algoritmo que vem apresentando bons resultados ao fim que se propõe no meio científico ao longo de anos. O número de palavras visuais do vocabulário é justamente o valor de *k*, que significa o número de grupos que será usado no *k-Means*.

Para determinar o valor de k (*k-Means*), que é o número de centroides (pontos centrais dos grupos) que foram criados para encontrar a similaridade dos dados, três métodos podem ser usados, a saber, WSS (*Within-Cluster Sum of Squares*), que usa a soma dos quadrados para encontrar o número ideal de *clusters*; o método da silhueta, que calcula as distâncias de cada ponto em relação a todos os outros pontos da amostra; e, por fim, o método *gap*, que calcula as dispersões dentro de cada *cluster*. Porém, devido ao baixo tempo de processamento para altos valores de k e aos bons resultados de acurácia inicial, valores de $k = 300$, $k = 400$, $k = 500$ e $k = 1000$ foram definidos arbitrariamente.

Outro ajuste na configuração do algoritmo é o valor do *blockwidth*, dentro da aplicação do BoVW, que representa o tamanho da seleção utilizada na extração de características do detector SURF, conforme ilustrado na Fig. 33, em que pode ser definido entre as quatro opções de tamanhos possíveis, a saber, 32, 64, 96 e 128 pixels. Nesse trabalho foram considerados os valores do *blockwidth* de 64 e 128 pixels de forma empírica, tendo em vista que foram com esses valores que se obtiveram melhores resultados.

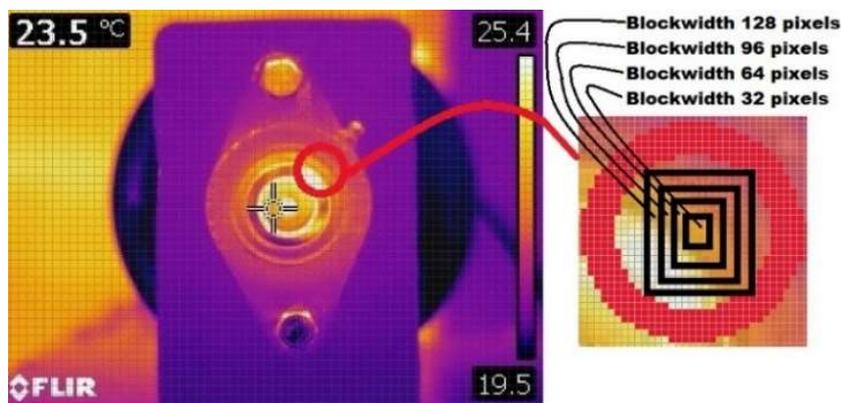


Figura 33 – Ilustração das opções de configuração do tamanho do *blockwidth*.

Ao final desse procedimento, cada imagem passou a ser representada por um vetor contendo a quantidade de cada uma das palavras visuais que a imagem possui. A etapa final consistiu em treinar um classificador SVM (CORTES E VAPNIK, 1995) para que aprendesse a distinguir as várias categorias de imagens a partir do vetor de palavras. A escolha do SVM como classificador para esse trabalho foi devido ao bom desempenho alcançado no trabalho de JIA et al. (2019) para classificação das 9 categorias de falhas em máquinas rotativas, utilizando imagens termográficas, tanto quando os dados de entrada foram BoVW como CNN.

Destacam-se a seguir algumas informações que foram extraídas do relatório que é gerado cada vez que o algoritmo é rodado, com o objetivo de quantificar as entradas e saídas das principais etapas do método. Na fase de treinamento foram usadas 105 imagens de cada uma das 15 categorias, totalizando 1575 imagens. As imagens foram adquiridas de forma sequencial, a intervalos de 20 s entre cada aquisição. De cada categoria com 105 imagens foram extraídos 126000 descritores pelo SURF. Portanto, considerando todas as imagens do grupo de treinamento tem-se um total de 1890000 descritores. Esses

descritores foram então direcionados para o algoritmo de agrupamento, o qual definiu quem foram as k palavras visuais (o algoritmo foi executado com $k = 300, 400, 500$ e 1000) presentes em cada uma das 1575 imagens disponíveis. Após definidas as k palavras, finalmente foi encontrada a frequência relativa de cada uma dessas palavras em cada uma das 1575 imagens utilizadas na fase de treinamento. Esses vetores de características, com k elementos representando as frequências relativas das palavras em cada uma das 1575 imagens foram então utilizados como entrada do SVM, juntamente com os rótulos de cada uma dessas imagens.

Também foi realizada uma sequência de testes utilizando o método de validação cruzada k -fold, conforme ilustrado na Fig. 34, em que foram retidas 3 subamostras para o teste de validação e as 7 restantes para treino, mantendo-se a mesma metodologia da proporção de 70% / 30% aplicada anteriormente.

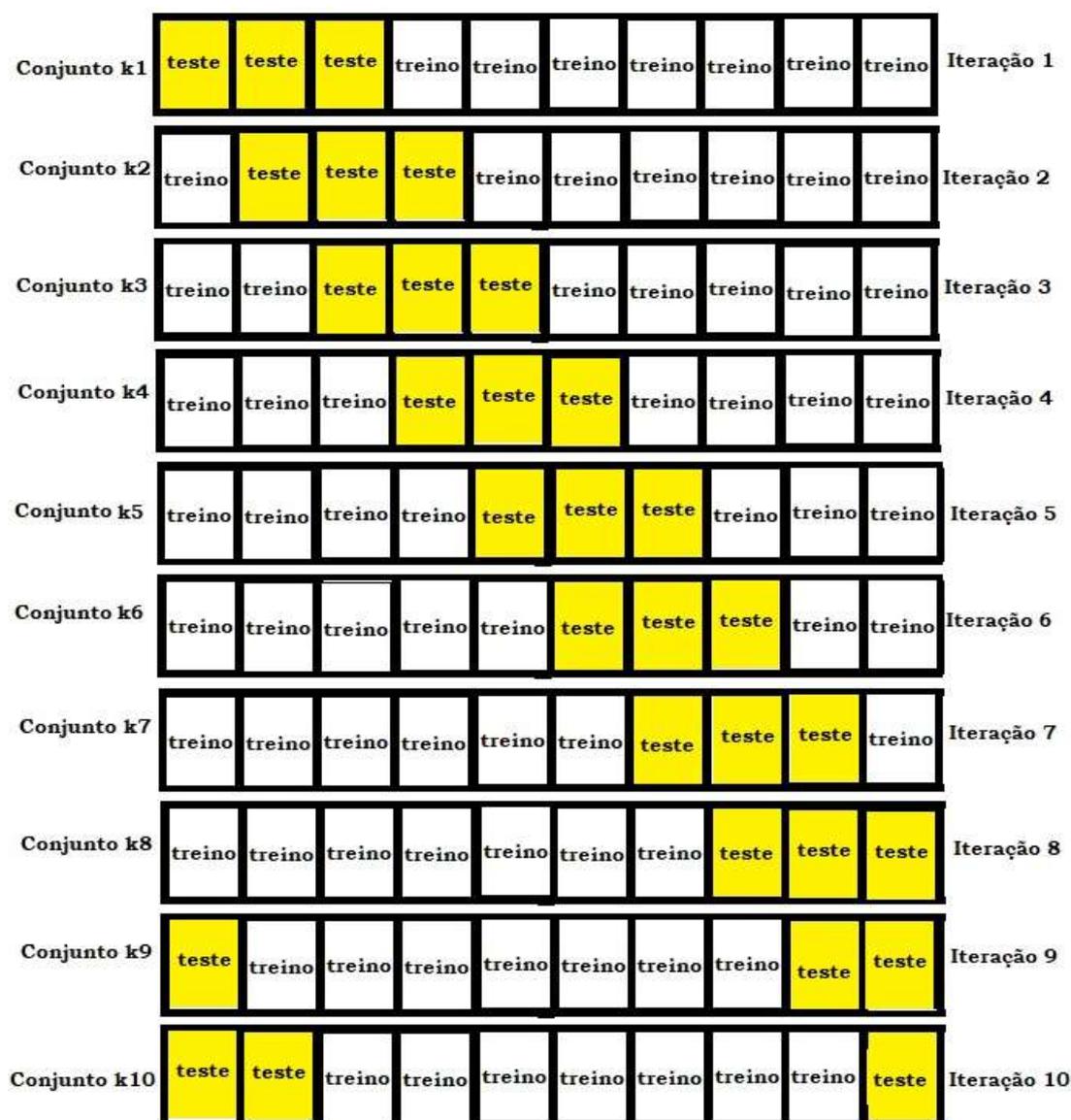


Figura 34 – Ilustração do método de validação cruzada k -fold com $k = 10$ (Fonte: autor)

Esses novos testes utilizando o método de validação cruzada k-fold foram realizados apenas para as imagens obtidas através da câmera Fluke VT02, com a configuração em que foram alcançados os melhores resultados da taxa de acerto, ou seja, tamanho das palavras visuais igual a 1000 e blockwidth igual a 128 pixels. Dessa forma, foi definido $k = 10$, devido essa definição ser comumente utilizada (MCLACHLAN e DO, 2005) resultando em uma validação cruzada de 10 vezes. Nesta validação cruzada de 10 vezes, o conjunto de dados foi aleatoriamente dividido em dez conjuntos, denominados de k_1, k_2, \dots, k_{10} , sendo 3 para validação e os outros 7 para treinamento, de modo que em cada iteração o subconjunto para teste de validação era diferente, assim como o subconjunto de treinamento também era diferente.

Na sequência, foi realizada uma nova etapa de testes utilizando o produto interno como forma de avaliar o nível de similaridade entre as imagens de uma mesma categoria em relação às demais. O objetivo deste teste foi de avaliar se um método muito mais simples que o BoVW seria capaz de categorizar com boa assertividade as falhas simuladas.

Esses novos testes utilizando o produto interno também foram realizados apenas para as imagens obtidas através da câmera Fluke VT02 com a configuração em que foram alcançados os melhores resultados da taxa de acerto. Vale destacar também que as matrizes 105×105 (105 linhas por 105 colunas para as imagens da câmera Fluke) foram normalizadas, de forma que se trabalhasse com valores de 0 à 1, ao invés de 0 à 255. Dessa forma, o conjunto de dados de cada uma das 15 categorias foi aleatoriamente dividido em um subconjunto para validação com 30% dos dados (45 imagens) e o outro subconjunto para teste com 70% dos dados (105 imagens). Foram separados 3 conjuntos distintos com esse arranjo, de modo que em cada iteração o subconjunto para validação era diferente, assim como o subconjunto de teste também era diferente.

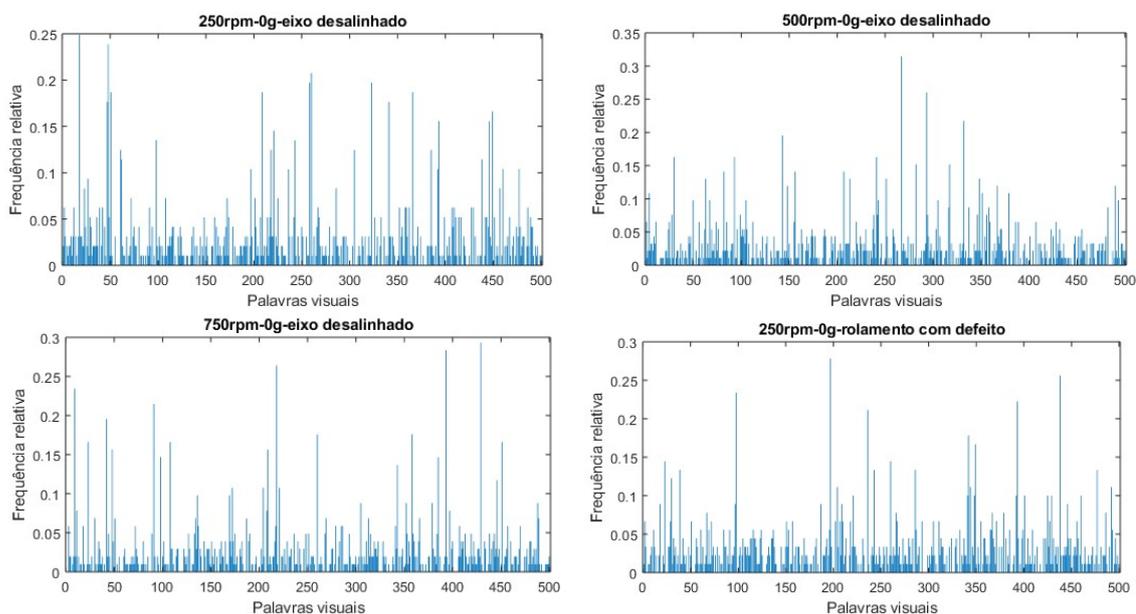
Vale ressaltar ainda que para aplicação do produto interno, inicialmente foi criado um modelo para cada uma das 15 categorias. Este modelo consistiu em se calcular a imagem média em tons de cinza, de forma que a imagem a ser verificada foi convertida em um vetor, o qual foi normalizado, dividindo-se cada um dos elementos do vetor por sua norma, resultando assim em um vetor com norma igual a 1. Já o produto interno foi obtido entre o vetor normalizado a partir da imagem que está sendo verificada e o vetor normalizado obtido a partir do modelo com o qual se busca comparar. Por fim, calculou-se o desempenho médio, somando o desempenho para cada uma das 15 categorias e dividindo o resultado por 15.

Como as imagens para cada categoria de falha foram adquiridas de forma sequencial e com a câmera posicionada em um ponto fixo, a tarefa de detecção e classificação das falhas é simplificado, propiciando uma taxa de acerto mais alta. Como forma de minimizar este problema, recomenda-se que as coletas para uma mesma categoria sejam realizadas em diferentes perspectivas de posicionamento da câmera, dias alternados, horários variados, etc.

Capítulo 5

Resultados e discussão

As Figuras 35 e 36 a seguir, mostram os histogramas característicos de cada uma das 15 categorias de dados estudadas para cada câmera termográfica. Nesses histogramas podem-se observar a frequência relativa de ocorrência de cada uma das 500 palavras visuais, que nada mais são do que pequenas porções da imagem que têm a capacidade de caracterizar a imagem como um todo. Nesse caso específico, as particularidades de cada imagem se diferenciam umas das outras, ou seja, apresentam diferenças para cada categoria estudada, pelo conjunto de temperaturas que representam essa porção da imagem. O objetivo do uso de uma câmera termográfica para levantamento desses dados consiste em avaliar as nuances de diferenças de temperaturas em pontos muito específicos de cada imagem, de forma que se possa agrupar em categorias as imagens que apresentam maior similaridades. Esses histogramas também foram gerados para 300, 400 e 1000 palavras visuais, quando definido esse valor de k , mas foram ilustradas a seguir apenas os histogramas com 500 palavras visuais.



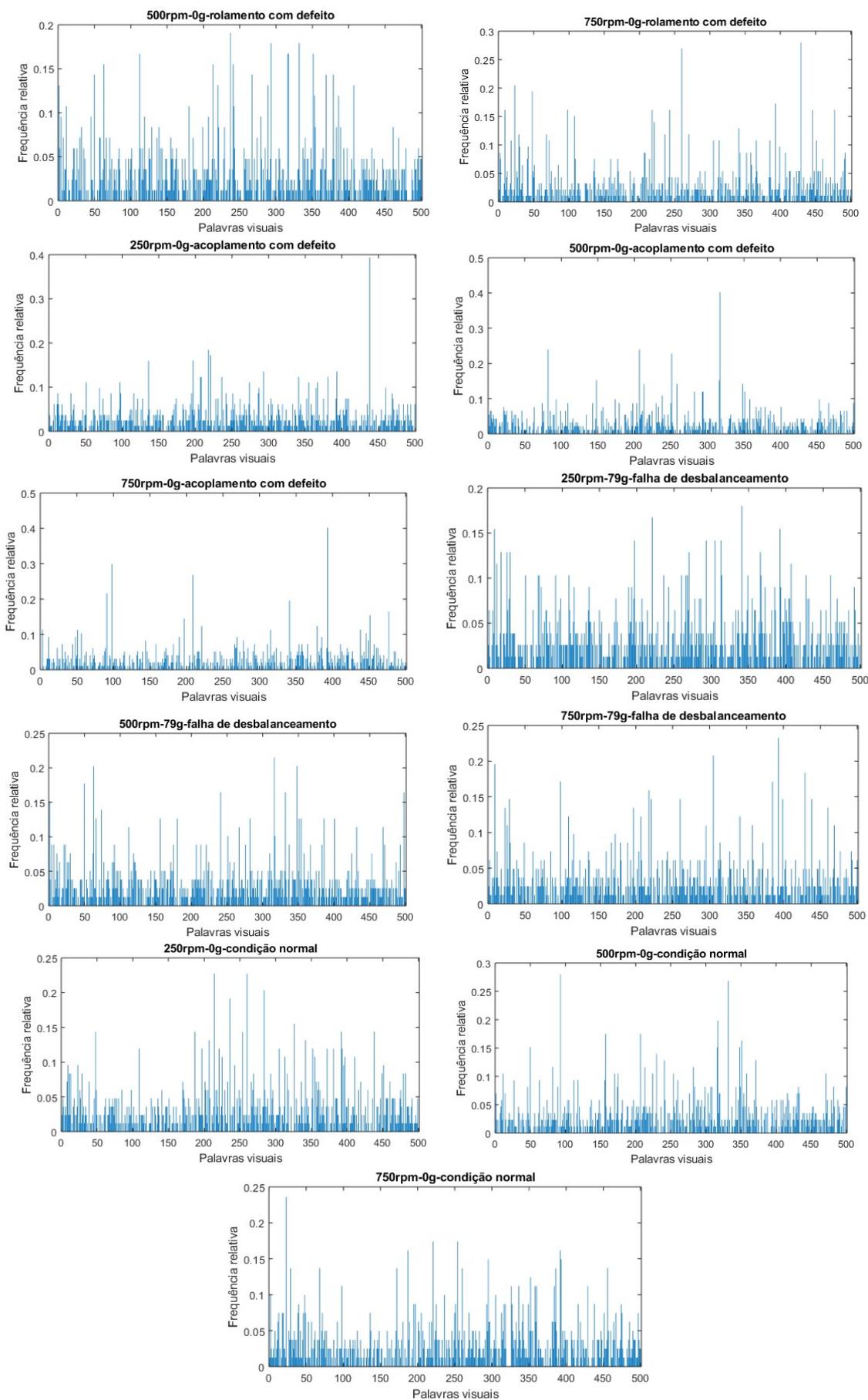
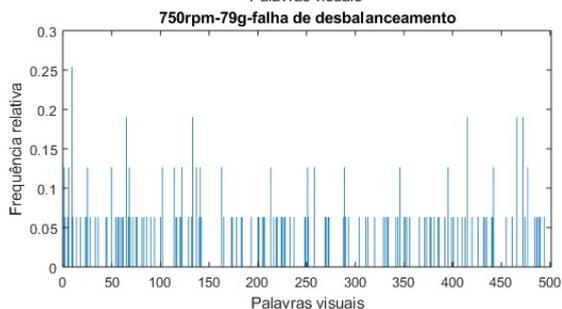
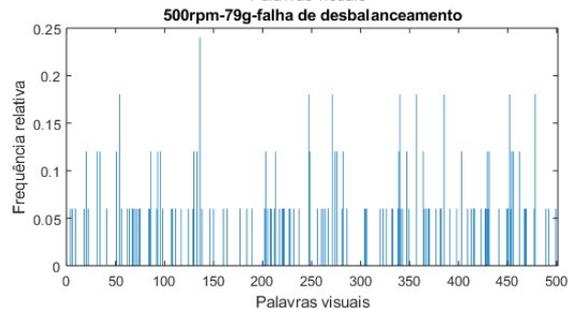
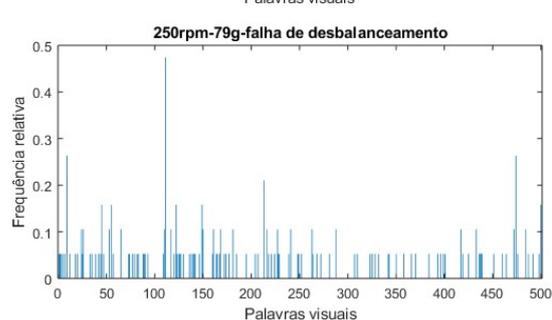
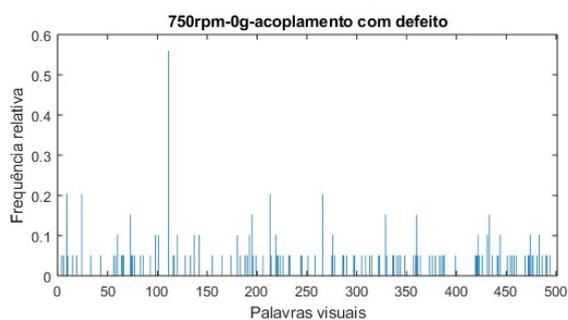
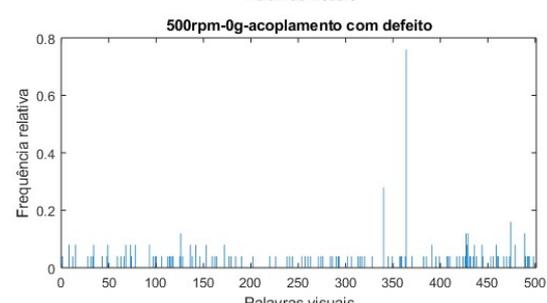
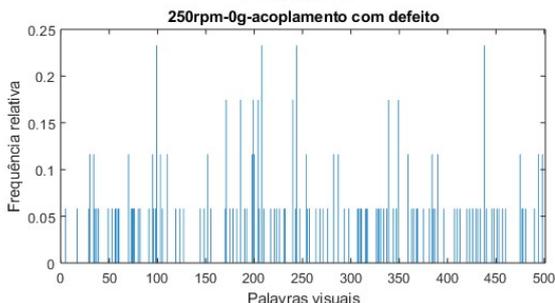
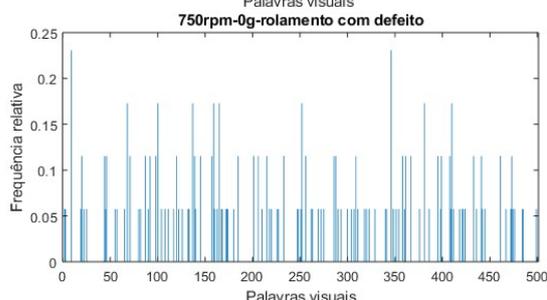
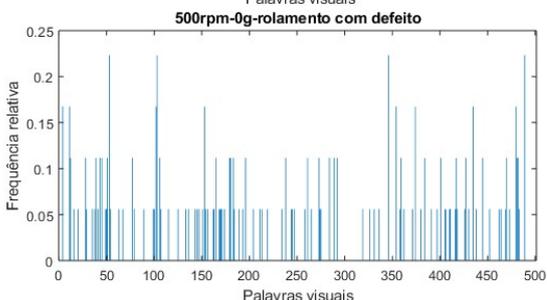
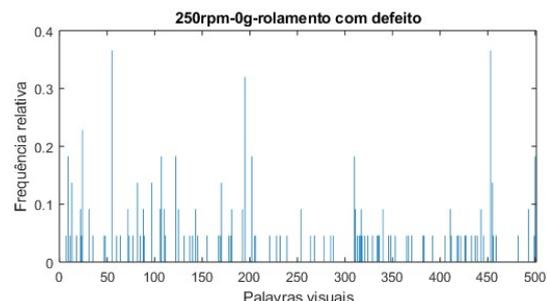
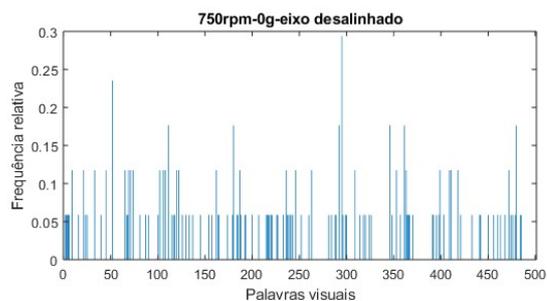
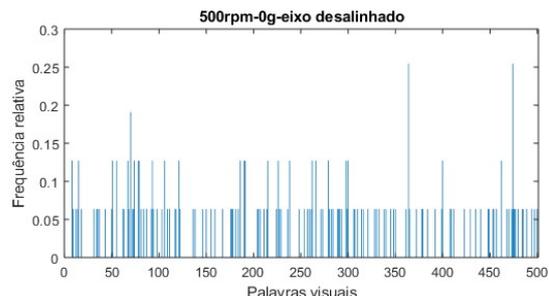
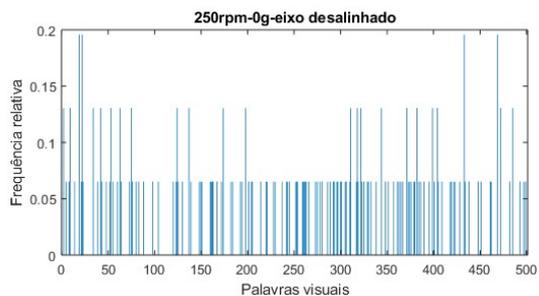


Figura 35 - Histograma das palavras visuais das imagens termográficas – FLIR T420.



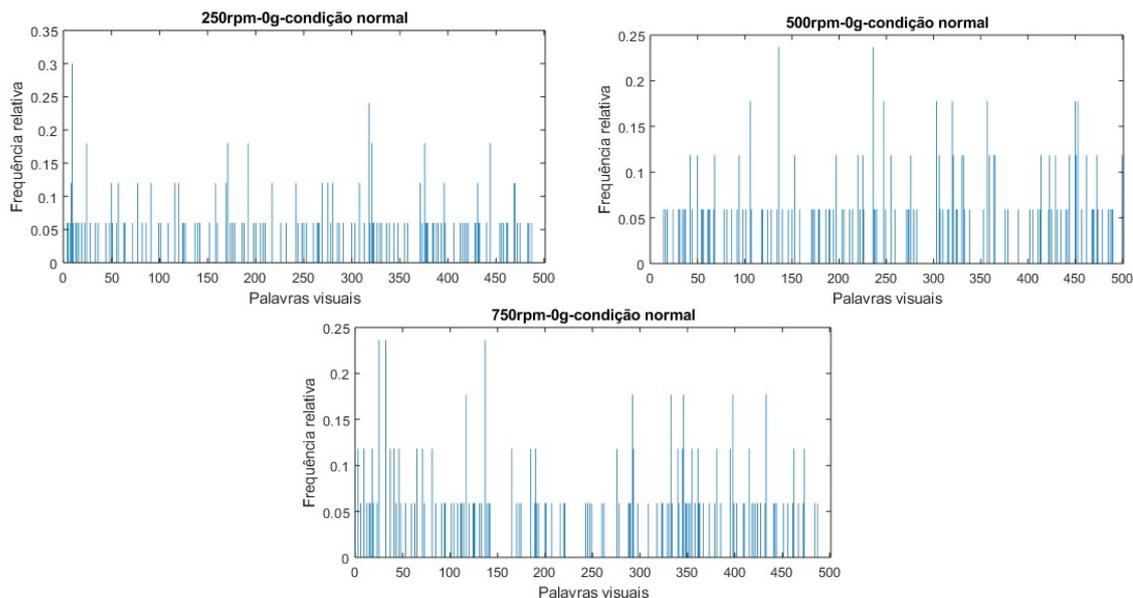


Figura 36 - Histograma das palavras visuais das imagens termográficas – FLUKE VT02.

Na Tabela 3 foi apresentada a legenda da simbologia utilizada nas Tabelas 4 e 5 para análise dos dados de vibração. Durante a simulação de cada uma das etapas apresentadas, foi realizada a medição da vibração do conjunto, para validação dos resultados, conforme apresentado nas Tabelas 4 e 5. Foi considerado o valor médio, após três medições da amplitude instantânea de vibração para cada uma das falhas simuladas. Portanto, espera-se identificar em que grau de incipiência, o defeito simulado e detectado pela técnica de análise de vibração, também será detectado pela técnica de termografia, a partir do processamento das imagens termográficas coletadas.

Tabela 3 – Legenda das Tabelas 4 e 5

Símbolo	Descrição
	Medição de vibração radial, horizontal do primeiro mancal
	Medição de vibração radial, vertical do primeiro mancal
	Medição de vibração radial, horizontal do segundo mancal
	Medição de vibração radial, vertical do segundo mancal

Tabela 4 – Medição de vibração durante os testes com a câmera FLUKE VT02

Número do teste	Condição	Rotação (rpm)	Nível de vibração (mm/s)			
						
1	Eixo desalinhado	250	0,2	0,2	0,1	0,2
2	Falha no rolamento	250	0,3	0,6	0,1	0,2
3	Falha no acoplamento	250	0,2	0,3	0,1	0,2

4	Desbalanceado	250	3,6	0,3	3,4	0,2
5	Normal	250	0,1	0,0	0,1	0,0
6	Eixo desalinhado	500	2,1	0,3	2,1	0,8
7	Falha no rolamento	500	0,6	0,9	0,3	0,6
8	Falha no acoplamento	500	0,6	0,6	0,6	0,4
9	Desbalanceado	500	8,2	1,8	8,5	1,1
10	Normal	500	0,4	0,1	0,3	0,1
11	Eixo desalinhado	750	0,5	0,9	0,6	0,7
12	Falha no rolamento	750	0,9	1,0	0,4	1,1
13	Falha no acoplamento	750	0,8	0,8	0,5	0,5
14	Desbalanceado	750	7,9	1,5	9,0	1,8
15	Normal	750	0,5	0,6	0,4	0,3

Tabela 5 – Medição de vibração durante os testes com a câmera FLIR T420

Número do teste	Condição	Rotação (rpm)	Nível de vibração (mm/s)			
						
1	Eixo desalinhado	250	0,5	0,3	0,5	0,3
2	Falha no rolamento	250	0,3	0,5	0,2	0,3
3	Falha no acoplamento	250	0,3	0,1	0,3	0,2
4	Desbalanceado	250	1,8	0,2	1,8	0,2
5	Normal	250	0,3	0,1	0,2	0,1
6	Eixo desalinhado	500	0,7	0,6	0,7	0,3
7	Falha no rolamento	500	0,8	0,8	0,3	0,5
8	Falha no acoplamento	500	0,5	0,1	0,5	0,2
9	Desbalanceado	500	10,2	1,7	10,8	1,7
10	Normal	500	0,4	0,2	0,3	0,3
11	Eixo desalinhado	750	0,4	0,4	0,7	0,5
12	Falha no rolamento	750	0,7	0,9	0,4	0,7
13	Falha no acoplamento	750	0,6	0,3	0,7	0,2
14	Desbalanceado	750	7,9	1,2	8,9	1,6
15	Normal	750	0,4	0,3	0,5	0,3

Ao analisar-se as Tabelas 4 e 5, com os dados das medições de vibração para cada uma das falhas simuladas, verifica-se que assim como a condição normal de funcionamento apresentou o menor nível de vibração, a condição da simulação da falha de desbalanceamento foi a que apresentou o maior nível de vibração. Apesar desse resultado já ser esperado, pôde-se observar que não só apenas para essas duas situações (condição normal e conjunto desbalanceado) foram observadas categorias de vibração diferentes, mas também para as outras três falhas simuladas, mesmo que essa diferença seja da ordem de 0,1mm/s. Vale ainda destacar que para um mesmo tipo de falha simulada, ou até mesmo na condição normal de funcionamento, houve diferença no nível de vibração entre as três velocidades de rotação analisadas.

Para elaborar a Tabela 6 foram utilizados os dados médios das três vezes que o algoritmo foi executado, com os resultados originados da taxa de acerto referente às imagens de validação. Cada vez que o algoritmo foi executado, foi gerado um relatório com diversas informações, dentre as quais se destacam os resultados das matrizes de confusão, que foram utilizados para compor essas tabelas em referência (disponíveis no relatório gerado durante o processamento do algoritmo e apresentados como apêndice 1 dessa dissertação). Os dados dessas matrizes de confusão foram tabulados e seus resultados apresentados nas tabelas do apêndice 2.

Tabela 6 – Resultado da taxa de acerto

k	<i>Blockwidth</i>	Taxa de acerto das imagens	
		FLIR T420	FLUKE VT02
300	64	99,60%	85,67%
400	64	99,67%	86,96%
500	64	99,87%	87,93%
1000	64	99,91%	89,02%
300	128	99,44%	89,33%
400	128	99,42%	89,93%
500	128	99,24%	90,31%
1000	128	99,64%	92,71%

De acordo com os resultados apresentados na Tabela 6, observa-se uma alta taxa de acerto, principalmente quando foram utilizadas 1000 palavras na configuração do *blockwidth* com 64 pixels, atingindo 99,91% de taxa de acerto para câmera FLIR T420 e 92,71% de taxa de acerto com a câmera FLUKE VT02, com 1000 palavras, mas na configuração do *blockwidth* com 128 pixels.

Ao analisar-se o tempo total de processamento na Tabela 7, verifica-se que o tempo de processamento do algoritmo com as imagens obtidas através da câmera FLUKE VT02 variaram de 4,8 (914,70s / 189,33s) até 9,2 (1196,1s / 129,66s) vezes mais rápido do que o processamento com as imagens obtidas através da câmera FLIR T420, o que pode ser

justificado pelo tamanho das imagens (76.800 pixels x 11.025 pixels). O tempo total de processamento teve como referência um notebook com processador *Intel Core I5* de sétima geração e 8 GB de memória RAM com sistema operacional *Windows 10* de 64 bits.

Tabela 7 – Tempo total de processamento dos resultados

k	Blockwidth	Tempo total de processamento (s)	
		FLIR T420	FLUKE VT02
300	64	575,79	117,77
400	64	622,96	109,15
500	64	939,01	104,01
1000	64	1196,10	129,66
300	128	861,01	155,00
400	128	1193,38	139,04
500	128	914,70	189,33
1000	128	1340,22	190,65

Com relação a uma melhor definição do número de palavras visuais do vocabulário, ainda não existe consenso relacionado ao seu tamanho ideal. Um vocabulário pequeno pode perder seu poder discriminativo, uma vez que dois detectores locais distintos podem ser atribuídos ao mesmo grupo, mesmo que não sejam semelhantes entre si. Por outro lado, um vocabulário grande e menos genérico, será menos tolerante a ruídos e gerará sobrecarga de processamento (JIANG, 2007).

Na Tabela 8 são apresentados os resultados da taxa de acerto para câmera Fluke VT02, utilizando a validação cruzada *k-fold*. Pode-se observar uma queda nos valores da taxa de acerto em relação ao método aplicado anteriormente (*hold-out*), reduzindo seu valor médio de 92,71% para 91,55%. Apesar de se utilizar um método mais rigoroso, não houve uma queda tão acentuada na taxa de acerto, o que demonstra uma boa robustez na qualidade dos dados.

Tabela 8 – Resultados da validação cruzada *k-fold*

Conjuntos utilizados na validação cruzada <i>k-fold</i>	Taxa de acerto	Conjuntos utilizados na validação cruzada <i>k-fold</i>	Taxa de acerto
K1	91,20%	K7	91,02%
K2	91,73%	K8	91,64%
K3	92,98%	K9	90,40%
K4	91,29%	K10	92,09%
K5	91,20%	Valor médio	91,55%
K6	91,91%		

Na Tabela 9 são apresentados os resultados da taxa de acerto para câmara Fluke VT02, utilizando o produto interno. Pode-se observar uma taxa de acerto média de 72,15% quando aplicado o teste de similaridade entre as imagens de cada categoria de falha em relação as demais, valor esse bem inferior quando comparado ao método BoVW com valores médios de 92,71% (*hold-out*) e 91,55% (*k-fold*), mas ainda assim muito elevado para uma classificação feita apenas com base em produto interno, o que de fato confirma que a forma sequencial de coleta das imagens simplifica muito a tarefa de detecção e classificação das falhas, mas, ainda assim, evidencia que o uso do BoVW melhora significativamente o resultado da classificação, justificando o seu uso.

Tabela 9 – Resultados dos testes de similaridade por produto interno

Categorias de falhas\ Testes	Taxa de acerto 1	Taxa de acerto 2	Taxa de acerto 3	Média
Falha no rolamento-250rpm	93,30%	97,78%	97,78%	96,29%
Falha no acoplamento-250rpm	91,10%	95,56%	93,33%	93,33%
Eixo desalinhado-250rpm	57,78%	66,67%	77,78%	67,41%
Normal-250rpm	91,11%	73,33%	86,67%	83,70%
Desbalanceado-250rpm	95,56%	97,78%	97,78%	97,04%
Falha no rolamento-500rpm	33,33%	73,33%	75,56%	60,74%
Falha no acoplamento-500rpm	75,56%	88,89%	91,11%	85,19%
Eixo desalinhado-500rpm	33,33%	75,56%	66,67%	58,52%
Normal-500rpm	62,22%	84,44%	82,22%	76,29%
Desbalanceado-500rpm	82,22%	48,89%	73,33%	68,15%
Falha no rolamento-750rpm	15,56%	15,56%	28,89%	20,00%
Falha no acoplamento-750rpm	48,89%	64,44%	62,22%	58,52%
Eixo desalinhado-750rpm	42,20%	40,00%	55,56%	45,92%
Normal-750rpm	75,56%	84,44%	88,89%	82,96%
Desbalanceado-750rpm	93,33%	77,78%	93,33%	88,15%
Valor médio	66,07%	72,30%	78,07%	72,15%

Capítulo 6

Conclusões

De acordo com os resultados obtidos verifica-se que o desempenho alcançado na identificação de falhas mecânicas em máquinas rotativas a partir de imagens termográficas capturadas com uma câmera de média resolução (FLIR T420) chega a ser de 99,91% enquanto que ao se utilizar uma câmera termográfica de baixa resolução (FLUKE VT02), se observou uma taxa de acerto de até 92,71%. Com isso, confirma-se a hipótese de que é possível utilizar câmeras termográficas de baixa resolução para identificação de falhas mecânicas em máquinas rotativas, ampliando assim a aplicabilidade dos métodos encontrados na literatura.

Vale destacar também, que ao se analisar a Tabela 7, verificou-se que o tempo de processamento do algoritmo com as imagens obtidas através da câmera de baixa resolução foi entre 4,8 e 9,2 vezes menor do que com o processamento das imagens obtidas através da câmera de média resolução. Isso se justifica devido ao tamanho menor da imagem obtida através da câmera de baixa resolução, e que é muito relevante quando se está analisando um grande número de imagens.

Observou-se também que mesmo na condição normal de funcionamento do simulador, ou seja, sem inserção de nenhum dos defeitos simulados, para diferentes velocidades de rotação (250 rpm, 500 rpm e 750 rpm), o método utilizado foi capaz de identificar e categorizar com boa assertividade cada uma dessas condições. Isso pode ser justificado pela diferença no nível de vibração em cada uma dessas rotações, conforme mostrado nas Tabelas 4 e 5, que gera níveis de atrito e, conseqüentemente, valores de temperatura muito particulares, em pontos específicos de cada imagem, levando a características diferentes entre elas. A partir dos dados apresentados nessas tabelas, foi possível identificar diferenças nos níveis de vibração, não só entre as falhas simuladas, mas também quando foi alterada a velocidade de rotação do conjunto, para cada uma das 15 categorias estudadas. Dessa forma, fica então confirmada a hipótese de que os defeitos simulados e detectados pela técnica de análise de vibração também foram apropriadamente identificados e separados em categorias pela técnica de termografia a partir do processamento das imagens termográficas coletadas.

Com relação aos métodos de validação cruzada, observou-se uma taxa de acerto menor quando aplicado o método de validação cruzada *k-fold*, o que pode ser justificado pelo método *k-fold* ter a vantagem de que todos os dados são usados para treinamento e validação, aplicando-se um embaralhamento dos conjuntos de dados, de forma que em cada iteração se utilize um conjunto amostral diferente, de modo que são esperados

resultados mais representativos. Já para o método *hold-out*, o processo é realizado de forma aleatória n vezes, o que não garante que os dados de treinamento e validação não sejam repetidos.

No que se refere ao teste de similaridade entre as imagens de cada categoria de falha em relação às demais, aplicando-se o produto interno para as imagens adquiridas através da câmera Fluke VT02, foi obtida uma taxa de acerto média de 72,15%. Apesar deste valor ser bem inferior quando comparado ao método BoVW em que se obtiveram valores médios da taxa de acerto de 92,71% (*hold-out*) e 91,55% (*k-fold*), ele mostra que há uma boa similaridade entre as imagens pertencentes a cada categoria de falha, o que facilita o processo de detecção e classificação de falhas e não corresponde ao que ocorre em aplicações reais, mas, por outro lado, justifica o uso do BoVW como forma de atingir melhores resultados de assertividade.

Ressalta-se ainda que o objetivo desse trabalho não era propor uma técnica mais apurada para reconhecimento de padrões em imagens termográficas, e sim, utilizar técnicas que já obtiveram sucesso nesse tipo de aplicação e replicar esse processo para outro tipo de câmera, especificamente nesse caso, de baixa resolução.

O classificador utilizado nesse trabalho se justifica pelos bons resultados alcançados por outros autores em outros trabalhos similares a este. No entanto, esses classificadores não são referências para o problema, portanto seria interessante testar outros tipos de classificadores que necessariamente não tenham sido desenvolvidos para esta aplicação. Inclusive, a utilização de classificadores mais simples pode até trazer resultados tão bons como os que foram atingidos neste trabalho.

Destaca-se ainda, uma possível oportunidade de melhoria na metodologia aplicada nesse trabalho, através do levantamento da hipótese de que, como as imagens para cada categoria de falha foram adquiridas de forma sequencial e com a câmera posicionada em um ponto fixo, poderia ter facilitado o trabalho do classificador, propiciando uma taxa de acerto mais alta. Portanto, fica o registro da importância de se aumentar a aleatoriedade durante o processo de coleta dos dados em futuros trabalhos, realizando as coletas para uma mesma categoria em diferentes perspectivas de posicionamento da câmera, dias alternados, horários variados, etc.

Como uma contribuição desse mestrado, foi disponibilizado o banco de dados com todas as imagens coletadas e utilizadas nessa pesquisa, na base de dados do Biochaves em que pode ser acessado através do *link*: <https://biochaves.com/downloads/banco-de-dados/>. O arquivo de nome *Thermography images*, está dividido em pastas, separadas pelo tipo de câmera termográfica e pela categoria de falha a que cada grupo de imagens pertence, a quem possa interessar, de forma a ser utilizado em novas pesquisas.

Como sugestões para trabalhos futuros, pode-se indicar:

- Utilizar o método de CNN para comparar seu desempenho com o que foi alcançado utilizando-se BoVW;

- Simular outros defeitos de rolamentos, comuns no ambiente industrial, como por exemplo: superaquecimento por lubrificação inadequada e escamamento, devido fadiga do aço, para avaliar se essas falhas também serão detectadas com a aplicação da metodologia em referência;
- Aplicar outras velocidades de rotação diferentes das que foram estudadas para verificar se haverá alteração no desempenho;
- Investigação da aplicabilidade da utilização desse método com a coleta das imagens termográficas em ambientes menos controlados, como o posicionamento manual da câmera termográfica (por exemplo) em um ambiente industrial, de forma a permitir que um operador execute o processo de aquisição de imagens para verificação;

Referências Bibliográficas

AHMADIPOUR, Foroughi, NEMATOLLAHI, SotoudehBahreini. Detection of Pillow Block Bearings Defect Using Infrared Thermography: A Case Study in Golgohar Mining and Industrial Co. Quantitative InfraRed Thermography Journal, 2015;

ARCO-VERDE, Melquesedeque Melo. Identificação de falhas em sistemas rotativos empregando técnicas não-lineares. 2008. Dissertação de Mestrado. Universidade Federal do Rio de Janeiro–RJ;

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 15572: Ensaios não destrutivos — Termografia — Guia para inspeção de equipamentos elétricos e mecânicos. São Paulo, 2013;

AZAD D., RAMJI K. Process plant reliability through infrared thermography. International Journal of Applied Engineering Research, Vol. 6 - 2011;

BAEZA-YATES, R.; RIBEIRO-NETO, B. Modern information retrieval addison-wesley longman. Reading MA – USA. 1999;

BAY, Herbert; TUYTELAARS, Tinne; VAN GOOL, Luc. Surf: Speeded up robust features. In: European conference on computer vision. Springer, Berlin, Heidelberg, 2006. p. 404-417;

BEZERRA, Roberto de Araújo. Detecção de falhas em rolamentos por análise de vibração. Tese (doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica, Campinas, SP - 2004;

BISHOP, Christopher M. Pattern recognition. Machine learning, v. 128, n. 9, 2006;

BÔAS, N. V. Métodos de aprendizagem: terminologia. São Paulo: [s.n.], 1987;

BOGNATZ, S. Alignment of critical and non critical machines. Orbit, pp. 23–25, 1995;

BORTH, Marcelo Rafael. RIBAS, Lucas C.. PISTORI, Hemerson. GONÇALVES, Wesley Nunes. JUNIOR, A. A. C.. "Classificação de Espécies de Peixe usando Inferência Gramatical no Reconhecimento de Padrões em Problemas de Visão Computacional" in Em: Memórias do X WTA 2016, EPUSP, São Paulo, 2016;

BORRA, Simone; DI CIACCIO, Agostino. Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. Computational statistics & data analysis, v. 54, n. 12, p. 2976-2989, 2010;

BOUSDEKIS, A., MAGOUTAS, B., APOSTOLOU, D., & MENTZAS, G.. A proactive decision making framework for condition-based maintenance. Industrial Management & Data Systems. Emerald Group Publishing Limited Copyright © - 2015;

BROWN, M.; LOWE, D. G. Invariant features from interest point groups. In: Electronic Proceedings of the 13th British Machine Vision Conference. College Park, Washington DC, EUA: [s.n.], 2002;

BURGES, Christopher JC. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, v. 2, n. 2, p. 121-167, 1998;

BURMAN, Prabir. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. *Biometrika*, v. 76, n. 3, p. 503-514, 1989;

CALADO, Flávio Augusto Rezende. Sistema de visão multiespectral para plataforma robótica de inspeção termográfica em subestações de energia. Dissertação de Mestrado – UFMG, 2013;

CALLISTER, William D.; RETHWISCH, David G. *Ciência e engenharia de materiais: uma introdução*. 8ª. Edição. Rio de Janeiro-RJ: LTC, 2018;

CANAHUIRE CABELLO, Frank Alexis. Contribuições à detecção automática de pontos quentes em postes de energia elétrica usando o algoritmo Bag of Visual Words e métodos de Kernel. Tese (doutorado) – UEC - Faculdade de Engenharia Elétrica e de Computação, Campinas, SP - 2016;

CARAMALHO, Alberto. 25 anos em termografia. Cap. 7-8. Ed. 1º, Portugal, 2012.

CARVALHO JÚNIOR, José Regivaldo de. Processamento digital de imagens para a identificação automática de falhas em rolos dos transportadores de correias. Dissertação de Mestrado - Universidade Federal de Ouro Preto, Ouro Preto, 2018;

CROW, F. Summed-area tables for texture mapping. In *Proceedings of SIGGRAPH*, volume 18 (3), pages 207-212. 1984;

CSURKA, Gabriella et al. Visual categorization with bags of keypoints. In: *Workshop on statistical learning in computer vision, ECCV*. 2004. p. 1-2;

DEVROYE, Luc; WAGNER, Terry. Distribution-free performance bounds for potential function rules. *IEEE Transactions on Information Theory*. 1979. Vol. 25, n. 5. pp. 601-604;

DUMAIS, Susan et al.. Inductive learning algorithms and representations for text categorization. In: *Proceedings of the seventh international conference on Information and knowledge management*. 1998. p. 148-155;

EFRON, Bradley. Estimating the error rate of a prediction rule: improvement on cross-validation. *Journal of the American statistical association*, v. 78, n. 382, p. 316-331, 1983;

EISENMANN, Robert C. Sr. and EISENMANN, Robert C. Jr., *Machinery Malfunction Diagnosis and Correction: vibration analysis and troubleshooting for the process industries*, New Jersey, Prentice-Hall, Inc., 1997;

GARG, A., & DESHMUKH, S. G.. Maintenance management: Literature review and directions. *Journal of Quality in Maintenance Engineering* - 2006;

HAMERLY, Greg; DRAKE, Jonathan. Accelerating Lloyd's algorithm for k-Means clustering. In: *Partitional clustering algorithms*. Springer, Cham, 2015. p. 41-78;

HARIHARAN, V. and SRINIVASAN, P. Vibration analysis of misaligned shaft-ball bearing system. *India J Sci Technol* 2009; 2: 45–50;

HESS, D. P., SOOM, A. Normal Vibrations and Friction under Harmonic Loads: Part I - Hertzian Contacts. *Journal of Tribology*, vol.113, 1991;

HIDAYAT, Agung Yudhistira. WIDODO , Achmad. HARYADI, Gunawan Dwi. Fault Diagnostic System Bearing Centrifugal Pump Using k-Means Method for Thermography Image and Signal Analysis Vibrations. Paper - MATEC Web of Conferences 159 – 2018;

HOWELL, J.R.; SIEGEL, R.; MENGUC, M.P. *Thermal Radiation Heat Transfer*, 5th ed.; Taylor and Francis: New York, NY, USA, 2010;

HUO, Zhiqiang. ZHANG, Yu. SATH, Richard, SHU, Lei. Self-adaptive Fault Diagnosis of Roller Bearings using Infrared Thermal Images. *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*;

INTERNATIONAL ELECTROTECHNICAL COMMISSION. Failure mode. Disponível em <<http://www.electropedia.org/>>. Acesso em: 24 out. 2017;

ISO - International Organization for Standardization. ISO 21940-11: 2016 - Mechanical vibration — Rotor balancing — Part 11: Procedures and tolerances for rotors with rigid behaviour;

JAIN, Anil K.; MURTY, M. Narasimha; FLYNN, Patrick J. Data clustering: a review. *ACM computing surveys (CSUR)*, v. 31, n. 3, p. 264-323, 1999;

JAIN, A. K.. Data clustering: 50 years beyond k-Means, *Pattern Recognition Letters*, v. 31, no. 8, pp. 651–666, 2010;

JIA, Zhen. JIU, Zhenbao, VONG, Chi-Man. PECHT, Michael. A Rotating Machinery Fault Diagnosis Method based on Feature Learning of Thermal Images. *IEEE Access (Volume: 7)* - Institute of Electrical and Electronics Engineers - 2019;

JIANG, Yu-Gang; NGO, Chong-Wah; YANG, Jun. Towards optimal bag-of-features for object categorization and semantic video retrieval. In: *Proceedings of the 6th ACM international conference on Image and video retrieval*. 2007. p. 494-501;

JUNG, Yoonsuh; HU, Jianhua. AK-fold averaging cross-validation procedure. *Journal of nonparametric statistics*, v. 27, n. 2, p. 167-179, 2015;

JURASZEK, Guilherme Defreitas. Reconhecimento de produtos por imagem utilizando palavras visuais e redes neurais convolucionais. Dissertação (mestrado) – UESC - Joinville – SC – 2014;

KAPLAN, H. Practical applications of infrared thermal sensing and imaging equipment. 2. ed. Washington: The International Society for Optical Engineering, 1999;

KASPERS, Anne, — Blob Detection Biomedical Image Sciences, Image Sciences Institute, UMC Utrecht, 2011;

KANASHIRO, Wesley Eiji Sanches. Histograma de palavras visuais para caracterização de texturas e cenas dinâmicas. Dissertação de Mestrado. FACION – UFMS. 2016;

LAMIM FILHO, P. C. M.; PEDERIVA, R.; BRITO, J. N. "Detecção de Desbalanceamento em máquinas rotativas através da lógica fuzzy". Conem 2006, 2006, Recife - PE. IV Congresso Nacional de Engenharia Mecânica;

LEEMANS, Vincent. DESTAIN, Marie-France. KILUNDU, Bovic. DEHOMBREUX, Pierre. Evaluation of the Performance of Infrared Thermography for on-Line Condition Monitoring of Rotating Machines. Published Online October 2011 (<http://www.SciRP.org/journal/eng>);

LEONARDI, Danilo Montilha. Medida de deslocamentos e vibrações utilizando uma câmera digital como sensor sem contato. Dissertação de Mestrado - UTFPR. 2015;

LI, Yongbo et al. A new intelligent fault diagnosis method of rotating machinery under varying-speed conditions using infrared thermography. Complexity, v. 2019, 2019;

LIMA, Geovana Veloso Loureiro de et al. Proposta de bag-of-visual-words por meio de redes complexas. 2017. Dissertação de Mestrado. Universidade Tecnológica Federal do Paraná;

LLOYD, S. P. Least square quantization in PCM. Bell Telephone Laboratories Paper. Published in journal much later: Lloyd, SP: Least squares quantization in PCM. IEEE Trans. Inform. Theor.(1957/1982), v. 18, 1957;

LOWE, David G. Distinctive image features from scale-invariant keypoints. International journal of computer vision, v. 60, n. 2, p. 91-110, 2004;

LOWE, David G. Object recognition from local scale-invariant features. In: Proceedings of the seventh IEEE international conference on computer vision. Ieee, p. 1150-1157. 1999;

MACQUEEN, James et al. Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967. p. 281-297;

MAHIAGITE, X., MOORE, P.. Nondestructive Testing Handbook — Infrared and Thermal. s.l. : American Society for Nondestructive Testing. Vol. 3 – 2001;

MALDAGUE, X. Theory and Practice of Infrared Technology for Nondestructive Testing. Wiley-Interscience. 2001;

MATTE, Marcelo Kuchar; DO CARMO NICOLETTI, Maria. Revisão de Estratégias para a Aceleração do Algoritmo k-Means. In: Anais do Workshop em Computação da FACCAMP. 2019. p. 1-6;

MATTE, Marcelo Kuchar. Impacto do Uso da Desigualdade Triangular para Acelerar o Algoritmo k-Means. Dissertação de Mestrado. UNIFACCAMP. 2020;

MCLACHLAN, Geoffrey J.; DO, Kim-Anh; AMBROISE, Christophe. Analyzing microarray gene expression data. 2005;

MENNA, Alexandre Ribeiro. Detecção de falhas em mancais de rolamento por análise de vibrações em banda larga: um caso pratico de aplicação em uma população de equipamentos rotativos. Porto Alegre - RS. Dissertação de Mestrado - UFRS. 2007;

MEOLA, C. Origin and Theory of Infrared Thermography. In Infrared Thermography: Recent Advances and Future Trends; Meola, C., Ed.; Bentham Science: New York, NY, USA, 2012;

MOLLMANN, K; KARSTADT, D.; PINNO, F.; VOLLMER, M. Selected critical applications for thermography: Convections in fluids, selective emitters and highly reflecting materials. In Proceedings of the Infrared Camera Calibration Conference, Las Vegas, NV, USA, Volume 6, pp. 161—173, October 2005;

MONTANARI, Raphael. Detecção e classificação de objetos em imagens para rastreamento de veículos. Tese de Doutorado. Universidade de São Paulo. São Paulo, 2016;

MOUSSA, Wael. Thermography-assisted bearing condition monitoring. Tese de Doutorado, Ottawa-Carleton Institute for Mechanical and Aerospace Engineering. University of Ottawa – Canada. 2014;

MUSZYNSKA, Agnieszka. Rotordynamics. CRC press, 2005;

NORTON, Robert L.. Projeto de máquinas: uma abordagem integrada; [tradução: Konstantinos Dimitriou Stavropoulos ... et al.]. – 4ª. ed. – Dados eletrônicos. – Porto Alegre: Bookman, 2013;

NSK Bearing Doctor. Diagnóstico Rápido de Ocorrências em Rolamentos. Catálogo da NSK Brasil Ltda.. 2013;

NUNEZ, J. A. et al. Low cost thermographic analysis for bearing fault detection in induction motors. *Journal of Scientific & Industrial Research* - Vol. 75. 2016;

PEDRINI, H.; SCHWARTZ, W. R.. *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*. São Paulo: Thomson Learning, p. 508, 2008;

PEDROSA, Glauco Vitor et al. Making an image worth a thousand visual words. In: *Workshop de Visao Computacional*, XI. 2015;

PENG, Y., DONG, M., & ZUO, M. J.. Current status of machine prognostics in condition-based maintenance: A review. *The International Journal of Advanced Manufacturing Technology* - 2010;

PIOTROWSKI, J. *Shaft alignment handbook*. 3^a ed. New York: Taylor & Francis Group, 2006;

PUROHIT, Pallavi; JOSHI, Ritesh. A new efficient approach towards k-Means clustering algorithm. *International Journal of Computer Applications*, v. 65, n. 11, p. 7-10, 2013;

RISSÁ, Henrique Sidney. *Identificação de deslocamentos de baixa amplitude e frequência utilizando uma câmera digital*. Dissertação de Mestrado - UTFPR. 2017;

SAKAMOTO, T. Normal Displacement and Dynamic Friction Characteristics in a Stick-Slip Process. *Journal Article Tribology International*, vol. 20, 1987;

SAKAMOTO, T., TSUKIZOE, T. Metal Transfer in the Frictional Contact of a Rough Hard Surface. *Wear*, vol. 47, ed. 2, 1978;

SENKO, Richard et al. *Estudo do comportamento dinâmico de rotores e controle dos níveis de vibração sob desbalanceamento rotativo*. Campina Grande, PB, 2013;

SEO, Jin Ju, CHOI, Nam Ryoung, KIM, Won Tae, HONG, Dong Pyo. Infrared Thermography Quantitative Diagnosis in Vibration Mode of Rotational Mechanics - *Journal of the Korean Society for Nondestructive Testing* Vol. 32, No. 3 (2012. 6);

SEKHAR, A. S.; PRABHU, B. S. Effects of coupling misalignment on vibrations of rotating machinery. *Journal of Sound and Vibration*, v. 185, n. 4, 1995;

SELIM, Shokri Z.; ISMAIL, Mohamed A. k-Means-type algorithms: A generalized convergence theorem and characterization of local optimality. *IEEE Transactions on pattern analysis and machine intelligence*, n. 1, p. 81-87, 1984;

SILVA, Francisca Cibele da. *Estudo termográfico de auxílio ao diagnóstico de falhas mecânicas em pistas de rolamento: o caso dos rolamentos rígidos de uma carreira de esferas sem blindagem*. Dissertação de Mestrado – Campina Grande, 2016;

SINGH, R. C.. Analysis of ball bearings under dynamic loading using non-destructive technique of thermography. International Journal of Advance Research and Innovation, Vol. 2 – 2014;

SIQUEIRA, Y. P. D. S. Manutenção centrada na confiabilidade: manual de implementação. 1ª (Reimpressão). Rio de Janeiro: Qualitymark, 2009;

SOARES, Robson de Carvalho et al.. Extensão do bag-of-visual-features para incorporar informação espacial na descrição de características de acordo com a percepção visual humana. Dissertação de Mestrado. Ciência da Computação. Universidade Federal de Uberlândia – MG. 2012;

TESTO SE & Co. KGa. Thermography, Pocket guide. (Link: <https://static-int.testo.com/media/1d/b7/21fc65abbea1/Pocket-Guide-Thermography-EN.pdf>). Lenzkirch – Germany, 2017;

TUCKMANTEL, Felipe Wenzel da Silva et al. Análise de modelo numérico de acoplamento de disco flexível. Tese (doutorado). Universidade Estadual de Campinas – Campinas - São Paulo, 2018;

USAMENTIAGA, Rubén et al. Infrared thermography for temperature measurement and non-destructive testing. Sensors, v. 14, n. 7, p. 12305-12348, 2014;

WAEYENBERGH, G., & PINTELON, L.. CIBOCOF: A framework for industrial maintenance concept development. International Journal of Production Economics – 2009;

YADAV, A.; YADAV, P.. Digital Image Processing. New Delhi: University Science Press, 2009;

YANG, Yuqing; HUANG, Shongming. Suitability of five cross validation methods for performance evaluation of nonlinear mixed-effects forest models—a case study. Forestry: An International Journal of Forest Research, v. 87, n. 5, p. 654-662, 2014;

ZEIDAN, Ahmad Al. Fotodetectores de radiação infravermelha baseados em pontos quânticos de submonocamada. Disertação de Mestrado. Universidade de São Paulo. São Paulo-SP, 2017.

APÊNDICES

APÊNDICE 1

TABULAÇÃO DOS DADOS

1 Tabulação dos dados

Na Tabela 1 são apresentadas as categorias das falhas que foram simuladas para esse trabalho, e o número correspondente a cada categoria foi utilizado nas tabelas de 2 a 9, devendo ser considerado na análise dos seus dados.

Tabela 1 – Categorias de falhas simuladas

Número	Categoria
1	Eixo desalinhado – 250rpm
2	Rolamento com defeito – 250rpm
3	Acoplamento com defeito – 250rpm
4	Falha de desbalanceamento – 250rpm
5	Condição normal - 250rpm
6	Eixo desalinhado – 500rpm
7	Rolamento com defeito – 500rpm
8	Acoplamento com defeito – 500rpm
9	Falha de desbalanceamento – 500rpm
10	Condição normal - 500rpm
11	Eixo desalinhado – 750rpm
12	Rolamento com defeito – 750rpm
13	Acoplamento com defeito – 750rpm
14	Falha de desbalanceamento – 750rpm
15	Condição normal - 750rpm

Depois da extração das características, três repetições para cada experimento foram realizadas, alcançando os seguintes resultados médios dessas três vezes que o algoritmo foi rodado, conforme apresentados nas tabelas de 2 a 9.

Tabela 2 – Resultado da taxa de acerto
Blockwidth = 64 e *k* = 300

Número	Taxa de acerto das imagens	
	FLIR T420	FLUKE VT02
1	100%	77,67%
2	100%	98,67%
3	100%	92,67%
4	100%	81,33%
5	100%	100%
6	99,33%	74,67%
7	100%	83%
8	99,33%	87,33%
9	97,33%	86,67%
10	99,33%	86,67%
11	98,67%	88%
12	100%	53,33%
13	100%	83%
14	100%	96,67%
15	100%	95,33%
TOTAL	99,60%	85,67%

Tabela 3 – Resultado da taxa de acerto
Blockwidth = 64 e *k* = 400

Número	Taxa de acerto das imagens	
	FLIR T420	FLUKE VT02
1	100%	80%
2	100%	100%
3	100%	99,67%
4	100%	82,67%
5	100%	99,33%
6	99,33%	76,33%
7	100%	83,67%
8	100%	89,67%
9	97%	85%
10	99,33%	88,33%
11	100%	91,67%
12	99,33%	59,33%
13	100%	86,67%
14	100%	95,67%
15	100%	96,33%
TOTAL	99,67%	86,96%

Tabela 4 – Resultado da taxa de acerto
Blockwidth = 64 e *k* = 500

Número	Taxa de acerto das imagens	
	FLIR T420	FLUKE VT02
1	100%	82,33%
2	100%	100%
3	100%	94%
4	100%	80,67%
5	100%	100%
6	100%	76,33%

Tabela 5 – Resultado da taxa de acerto
Blockwidth = 64 e *k* = 1000

Número	Taxa de acerto das imagens	
	FLIR T420	FLUKE VT02
1	100%	80,67%
2	100%	100%
3	100%	94%
4	100%	85%
5	100%	99,33%
6	100%	84,33%

7	100%	86,67%
8	99,33%	89%
9	98,67%	89,67%
10	100%	90,33%
11	100%	90,33%
12	100%	63%
13	100%	83%
14	100%	96,67%
15	100%	97%
TOTAL	99,87%	87,93%

7	100%	86,67%
8	100%	90,67%
9	100%	87%
10	99,87%	94,33%
11	100%	88,67%
12	100%	69,67%
13	100%	80%
14	100%	98%
15	100%	97%
TOTAL	99,91%	89,02%

Tabela 6 – Resultado da taxa de acerto
Blockwidth = 128 e *k* = 300

Número	Taxa de acerto das imagens	
	FLIR T420	FLUKE VT02
1	100%	86,67%
2	100%	99,33%
3	100%	95%
4	100%	92,67%
5	100%	100%
6	99,33%	77,33%
7	100%	95%
8	98,67%	90,33%
9	96,33%	87,33%
10	98,67%	91%
11	99,33%	80,67%
12	99,33%	66%
13	100%	85,67%
14	100%	98%
15	100%	95%
TOTAL	99,44%	89,33%

Tabela 7 – Resultado da taxa de acerto
Blockwidth = 128 e *k* = 400

Número	Taxa de acerto das imagens	
	FLIR T420	FLUKE VT02
1	100%	89,67%
2	100%	99,33%
3	100%	93,67%
4	100%	92,67%
5	100%	100%
6	100%	77,67%
7	100%	95,67%
8	99,33%	87,67%
9	97%	86%
10	98%	91%
11	100%	78,67%
12	97%	77%
13	100%	89,33%
14	100%	97,33%
15	100%	93,33%
TOTAL	99,42%	89,93%

Tabela 8 – Resultado da taxa de acerto
Blockwidth = 128 e *k* = 500

Número	Taxa de acerto das imagens	
	FLIR T420	FLUKE VT02
1	100%	88,67%
2	100%	100%
3	100%	97,33%
4	100%	88,33%
5	100%	100%
6	99,33%	81,33%
7	100%	93,33%
8	100%	94%
9	97,67%	88,67%
10	96%	91,67%
11	98,67%	79,33%
12	97%	76,33%
13	100%	83%
14	100%	95%
15	100%	97,67%
TOTAL	99,24%	90,31%

Tabela 9 – Resultado da taxa de acerto
Blockwidth = 128 e *k* = 1000

Número	Taxa de acerto das imagens	
	FLIR T420	FLUKE VT02
1	100%	85%
2	100%	99,33%
3	100%	98,67%
4	100%	92,33%
5	100%	100%
6	98,67%	82,67%
7	100%	95%
8	98,67%	91%
9	99,33%	91%
10	98%	94,33%
11	100%	93%
12	100%	85%
13	100%	92%
14	100%	93,33%
15	100%	98%
TOTAL	99,64%	92,71%

APÊNDICE 2

**ESTUDO DIMENSIONAL DO EFEITO DA
DISTÂNCIA FOCAL E DO CAMPO DE VISÃO
(FOV) EM RELAÇÃO A DISTÂNCIA DO
ROLAMENTO**

2 Estudo dimensional do efeito da distância focal e do campo de visão (FOV) em relação a distância do rolamento

Neste apêndice é apresentado um estudo dimensional do efeito da distância focal e do campo de visão (FOV) de cada câmera termográfica em relação à distância do rolamento, de forma a subsidiar trabalhos futuros no ajuste das distâncias de cada câmera em relação ao rolamento, compensando assim o tamanho do pixel. Dessa forma, apenas seria considerada a distância focal de acordo com o campo de visão de cada câmera, tendo em vista que a resolução teria a mesma equivalência, devido ser compensada por distâncias diferentes.

Foi ilustrado na Fig. 1 o rolamento utilizado nesse experimento, com suas dimensões de acordo com o catálogo do fabricante FRM (Link: <https://www.frm.ind.br/y200.pdf>).



Figura 1 – Rolamento FRM Y204 V22 (adaptado do catálogo da FRM)

Pode-se observar na Fig. 1, que foi formada uma coroa circular com largura de 27,0mm (47,0mm – 20,0mm). Essa foi a área de interesse para avaliação termográfica, e resta agora avaliar quantos pixels de cada câmera termográfica estão inseridos nessa área da coroa circular em referência. O cálculo da área formada pela coroa circular foi dado pela subtração da área maior (A_1) pela área menor (A_2). Portanto:

$$A_1 - A_2 = \pi r_1^2 - \pi r_2^2 = 3,14.23,5^2 - 3,14.10,0^2 = 1420,06mm^2 \quad (1)$$

Para calcular o tamanho do pixel da câmera FLIR T420, que apresenta um FOV de $25^\circ \times 19^\circ$, pode ser aplicado os conceitos da trigonometria para se encontrar a largura e altura de cada pixel que forma a imagem. Inicialmente foi dividido $25^\circ/2$, para calcular a tangente e depois multiplicado por 0,4 metros, que foi a distância utilizada para coleta das imagens do rolamento no simulador de falhas. Esse resultado foi multiplicado por 2, sendo obtido o resultado de 0,17736. Esse valor em seguida foi dividido por 320, que vem a ser a quantidade de linhas na resolução dessa câmera (320x240 pixels) chegando à largura do pixel de $5,5 \times 10^{-4}$ m. Na sequência, para encontrar a altura do pixel, é feito o mesmo procedimento, agora considerando o 19° do FOV. Chegando à altura do pixel de $11,5 \times 10^{-4}$ m. Ao se multiplicar esses dois resultados chega-se à área de cada pixel, no valor de $63,25 \times 10^{-8} \text{ m}^2$, que é igual a $0,63 \text{ mm}^2$. Para melhor ilustrar esse entendimento, foi apresentada na Fig. 2 o tamanho do pixel de acordo com a distância que a câmera foi posicionada.

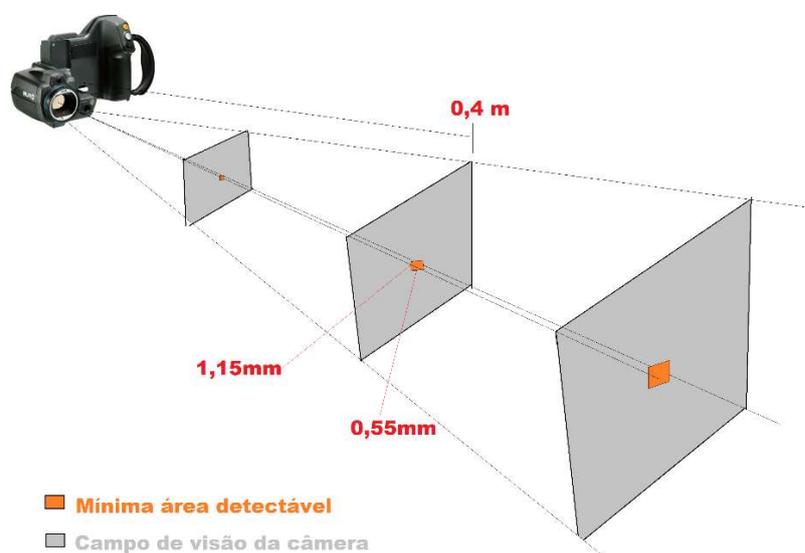


Figura 2 – Tamanho do pixel em relação a área de visão da câmera FLIR T420 (Fonte: autor).

Para calcular o tamanho do pixel da câmera Fluke VT02, que apresenta um FOV de $20^\circ \times 20^\circ$, foi dividido $20^\circ/2$, para calcular a tangente e depois multiplicado por 0,4 metros, que foi a distância utilizada para coleta das imagens do rolamento no simulador de falhas. Esse resultado foi multiplicado por 2, sendo obtido o resultado de 0,14106. Esse valor foi dividido por 105, que vem do número de linhas da resolução dessa câmera (105x105 pixels) chegando ao tamanho do pixel de $13,4 \times 10^{-4}$ m. Em função da mesma dimensão da resolução, ou seja, número de linhas igual ao número de colunas, campo de visão com mesmo ângulo (20°), para calcular a área basta multiplicar esse resultado encontrado por ele mesmo, chegando à área do pixel de $179,56 \times 10^{-8} \text{ m}^2$, que é igual a $1,8 \text{ mm}^2$. Para melhor ilustrar esse entendimento, foi apresentada na Fig. 3 o tamanho do pixel de acordo com a distância que a câmera foi posicionada.

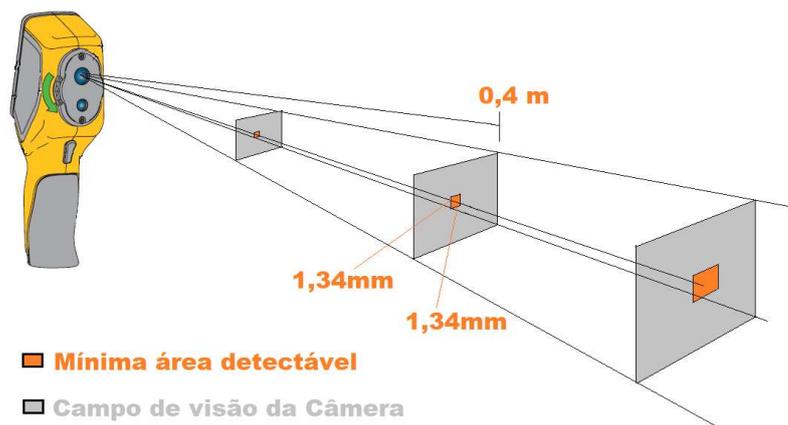


Figura 3 – Tamanho do pixel em relação a área de visão da câmera FLUKE VT02 (Fonte: autor).

Para que possa se fazer uma análise de qual distância a câmera FLIR deveria ser posicionada para manter o mesmo tamanho de pixel da câmera FLUKE, deve-se calcular quantos pixels de cada câmera cobrem a área formada pela coroa circular do rolamento. Vale destacar que a câmera FLUKE já está no limite máximo de proximidade ao objeto, portanto, só resta afastar a câmera FLIR do objeto de forma a compensar o tamanho do pixel. Para câmera FLIR, ao dividir-se a área de $1420,06\text{mm}^2$ por $0,63\text{mm}^2$ chega-se ao resultado de 2254 pixels cobrindo essa área da coroa circular do rolamento. Com relação a câmera FLUKE, ao dividir-se a área de $1420,06\text{mm}^2$ por $1,8\text{mm}^2$ chega-se ao resultado de 789 pixels cobrindo essa área. Ao realizar uma regra de três simples, chega-se ao resultado de que a câmera FLIR deveria ser mantida a uma distância de 1,14 metros, para que as duas câmeras ficassem com o mesmo tamanho de pixel, correspondente a uma área de $1,8\text{mm}^2$. Dessa forma, fica esse experimento como sugestão para trabalhos futuros, de forma a avaliar se o desempenho das duas câmeras ficaria mais próximos no caso de mantida a mesma resolução, que seria compensada a partir de distâncias diferentes das câmeras em relação ao objeto.

De acordo com TESTO (2017), para que se possa obter uma medição precisa, o objeto de medição deve ser três vezes maior do que o menor objeto identificável. A partir desse conceito, e de acordo com o estudo apresentado anteriormente, o tamanho do pixel da câmera FLIR de $1,15\text{mm} \times 0,55\text{mm}$ passa a ser $3,45\text{mm} \times 1,65\text{mm}$ e o tamanho do pixel da câmera FLUKE de $1,34\text{mm} \times 1,34\text{mm}$ passa a ser $4,02\text{mm} \times 4,02\text{mm}$. Portanto, pode-se concluir que esses tamanhos de pixel podem perfeitamente cobrir a área formada pela coroa circular do rolamento, como destacado no terceiro parágrafo desse apêndice, que tem a maior dimensão de $47,0\text{mm}$. Resultado esse que vem a validar a capacidade das duas câmeras acessarem todo contorno da coroa circular do rolamento com vários pontos (pixels) de medição. Além do que, a transmissão do calor por condução irá permitir que todo entorno da área de troca térmica, formado especialmente pelo mancal e pelo eixo, também apresentarão nuances de temperatura específicas para cada tipo de defeito simulado. Sendo um indicativo de que a área que apresenta o comportamento térmico a ser estudado é bem superior a área inicialmente calculada de $1420,06\text{mm}^2$ (coroa circular do rolamento).