

Giovanni Martins de Sá Júnior

**Trabalho de Conclusão de Curso**  
**Monitoramento de Isoladores com Câmera**  
**Térmica e técnicas de Reconhecimento de**  
**Padrão**

Belo Horizonte

2024



Giovanni Martins de Sá Júnior

**Trabalho de Conclusão de Curso**  
**Monitoramento de Isoladores com Câmera Térmica e**  
**técnicas de Reconhecimento de Padrão**

Monografia apresentada durante o Seminário dos Trabalhos de Conclusão do Curso de Graduação em Engenharia Elétrica da UFMG, como parte dos requisitos necessários à obtenção do título de Engenheiro Eletricista.

Universidade Federal de Minas Gerais – UFMG

Escola de Engenharia

Curso de Graduação em Engenharia Elétrica

Orientador: Prof. Hermes Aguiar Magalhães

Belo Horizonte

2024



# Lista de ilustrações

Figura 1 – Isolador de Vidro para Linha de Transmissão . . . . .	13
Figura 2 – Arquitetura da Lenet 5 . . . . .	14
Figura 3 – Entrada de uma Rede Neural Convolucional . . . . .	15
Figura 4 – Aplicação da técnica de Max Pooling em uma sub-região . . . . .	16
Figura 5 – YOLO sendo aplicado para difentes conjuntos de imagens . . . . .	17
Figura 6 – Técnica de detecção de falha em Isolador . . . . .	18
Figura 7 – Monitoramento de linhas de transmissão com o auxílio de drones . . . .	19
Figura 8 – Técnica de pré-processamento da imagem do isolador . . . . .	20
Figura 9 – iCam LW-200 . . . . .	22



# Lista de tabelas

Tabela 1 – Especificações técnicas da câmera iCam LW-200 . . . . .	22
--	----





# Lista de abreviaturas e siglas

CNN	Convolutional Neural Network
DL	Deep Learning
FDD	Forceful Steel Defect Detector
IoT	Internet of Things
NLP	Natural Language Processing
R-CNN	Region Based Convolutional Neural Network
RNA	Rede Neural Artificial
RPN	Region Proposal Network
VC	Visão Computacional
ViT	Vision Transformer
YOLO	You only look once



# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>13</b>
<b>2.1</b>	<b>Nivelamento</b>	<b>13</b>
2.1.1	Isoladores	13
2.1.2	Redes Neurais Convolucionais (CNNs)	14
2.1.2.1	Entradas	14
2.1.2.2	Camadas Convolucionais	14
2.1.2.3	Funções de ativação	15
2.1.2.4	Camada de Pooling	15
<b>2.2</b>	<b>Estado da Arte</b>	<b>16</b>
<b>2.3</b>	<b>Trabalhos Correlacionados</b>	<b>17</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>21</b>
<b>3.1</b>	<b>Definição da Base de Dados</b>	<b>21</b>
3.1.1	Descrição da População	21
3.1.2	Aquisição de Dados	21
3.1.2.1	Descrição da Câmera utilizada	21
3.1.3	Rotulagem	22
<b>3.2</b>	<b>Pré-processamento dos Dados</b>	<b>23</b>
3.2.1	Normalização e Redimensionamento	23
3.2.2	Aumento da Base de Dados	23
3.2.3	Remoção de Ruídos	23
3.2.4	Segmentação	23
<b>3.3</b>	<b>Planejamento Experimental</b>	<b>23</b>
3.3.1	Divisão do Conjunto de Dados	23
3.3.2	Seleção de Algoritmos	24
3.3.3	Treinamento do Modelo	24
3.3.4	Métricas de Avaliação do Modelo	24
	<b>REFERÊNCIAS</b>	<b>25</b>



# 1 Introdução

O início da década de 1940 trouxe uma grande revolução técnica e científica em diferentes campos durante o desenvolvimento da Segunda Guerra Mundial, e dentre elas, o campo da matemática e da computação. Dentre os trabalhos desenvolvidos, o artigo desenvolvido pelos cientistas americanos Warren McCulloch e Walter Pitts ([MCCULLOCH; PITTS, 1943](#)), denominado "*A Logical Calculus of the Ideas Immanent in Nervous Activity*", foi considerado um trabalho muito importante para a modelagem matemática de um neurônio artificial.

No modelo acima, os neurônios seriam unidades capazes de desenvolver operações lógicas simples, tais como somas e multiplicações. Dessa forma, os neurônios seriam capazes de se conectar uns aos outros por meio das sinapses que continham pesos. Assim, este trabalho fundamentou pesquisas que foram pausadas por um período considerável de tempo dadas as limitações tecnológicas e computacionais no momento em que estavam sendo desenvolvidas. Apesar dessas limitações, outras dificuldades também eram apresentadas naquele contexto, dentre elas, a obtenção de dados para treinamento das redes, e a falta de compreensão do mecanismo de aprendizagem dos modelos iniciais.

Em vista disso, as dificuldades que limitaram o desenvolvimento deste campo foram superadas na segunda metade do século XX, com o prenúncio da Terceira Revolução Industrial, como pode ser visto em ([MOHAJAN, 2021](#)). Essa revolução trouxe uma quebra de paradigma com a introdução dos primeiros supercomputadores, o que implicou em um considerável ganho de escala na capacidade de processamento. Tal avanço deu sustentação para o desenvolvimento de novos trabalhos, mas além deste aperfeiçoamento, a formação de grandes bases de dados para treinamento, que variavam desde textos até imagens e arquivos de áudio, que o treinamento das novas Redes Neurais em aprenderem tarefas mais complexas.

Dentre os trabalhos desenvolvidos no contexto de Terceira Revolução Industrial, podem ser listadas: as primeiras aplicações de reconhecimento de imagens, que puderam ser utilizadas para reconhecer determinados objetos em imagens, como placas, objetos e rostos; as aplicações de processamento de linguagem natural, e a robótica, em que as redes neurais puderam ser aplicadas no próprio controle de robôs, por exemplo.

Através desta perspectiva, o desenvolvimento de pesquisas e aplicações na área de Redes Neurais tornaram-na um dos principais indutores para a transição da Terceira para a Quarta Revolução Industrial, como pode ser visto em ([PATNAIK, 2020](#)). A ideia central fundamentava-se no processo de digitalização das indústrias e na automação dos processos industriais como um todo, trazendo o foco na otimização de processos, serviços e nos

produtos desenvolvidos. Ademais, um outro pilar também presente na Quarta Revolução Industrial, é o que entende-se por Internet das Coisas (*IoT - Internet of Things*, em inglês). A possibilidade de se coletar informações de máquinas e dispositivos em tempo real, fazendo com que estes dados sejam enviados para a nuvem, criam diferentes possibilidades de melhorias de processos internos.

As aplicações de Realidade Aumentada e Visão Computacional ganharam um grande destaque diante deste cenário, podendo ser utilizadas por exemplo, no auxílio de trabalhadores em linha de produção na montagem de produtos, na manutenção preditiva de máquinas e equipamentos identificando previamente potenciais problemas de desempenho, bem como em processos logísticos, orientando trabalhadores em armazens e depósitos, melhorando a cadeia de suprimentos de uma companhia.

Além do campo industrial, o setor de energia, campo no qual será tratado neste trabalho, é também outro potencial setor no contexto de possíveis aplicações. Com isso, os elementos fundamentais de uma infraestrutura elétrica, envolvidos no processo de geração, transmissão, distribuição de energia, emergem como candidatos promissores para a aplicação de técnicas de Visão Computacional, especialmente no contexto de supervisão e monitoramento dos equipamentos.

Um candidato em particular, são os isoladores, presentes em linhas de transmissão, que podem estar situados em condições de ambiente adversas, como altas temperaturas, e exposição constante a raios ultravioleta, chuva e poluição. Diante deste cenário, os algoritmos de Visão Computacional se fazem importantes para o monitoramento preventivo destes equipamentos, sobretudo em caso de detecção de sinais de desgastes, rachaduras e corrosões. Para isso, será desenvolvido neste trabalho um sistema de monitoramento térmico para isoladores com o auxílio destes algoritmos computacionais e câmeras térmicas acopladas em drones que farão o acompanhamento em tempo real destes dispositivos. Com essa metodologia, será possível efetuar o monitoramento preditivo com o objetivo final de evitar fenômenos indesejados, como as descargas elétricas.

Neste primeiro capítulo foi realizada uma discussão histórica dos avanços trazidos pelas Redes Neurais, passando pelas potenciais aplicações no Campo da Visual Computacional, e afunilando para o tema central deste trabalho. No segundo capítulo, será realizada a discussão da Revisão Bibliográfica, no qual será feito um nivelamento inicial, bem como a análise de alguns trabalhos clássicos, o estado arte e alguns trabalhos vinculados com o tema central. No terceiro capítulo, será desenvolvida a metodologia técnica e teórica aplicada em todo o trabalho, enquanto no quarto capítulo será discutido todos os resultados obtidos no desenvolvimento do trabalho. E por último, será evidenciado as conclusões obtidas a partir dos resultados.

## 2 Revisão Bibliográfica

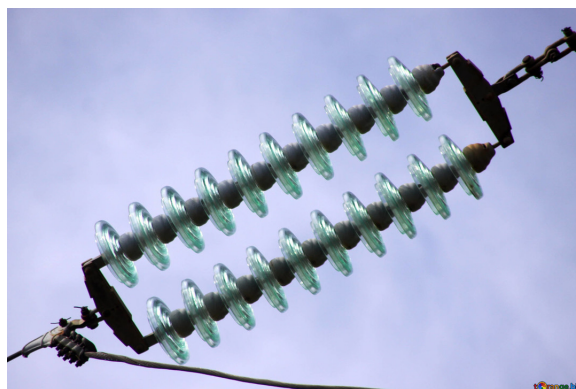
Neste capítulo, serão discutidos alguns dos principais artigos vinculados ao tema central deste trabalho, trazendo uma maior fundamentação técnica. Nesse sentido, a análise desses artigos será dividida em três momentos distintos. No primeiro deles, será realizado a introdução de alguns conceitos importantes para a fundamentação teórica do trabalho. Em um segundo momento, será feito o recorte de alguns dos mais recentes artigos correlacionados ao campo, trazendo as principais novidades. E por último, serão observados algumas das principais aplicações presentes no tema central do trabalho.

### 2.1 Nivelamento

#### 2.1.1 Isoladores

Isoladores são dispositivos fundamentais no contexto da transmissão de energia elétrica. Naturalmente detentores de uma alta rigidez dielétrica, os isoladores possuem o papel de impedir com que materiais condutores elétricos como cabos e linhas transmitam energia ao suporte no qual estão conectados. Assim, ele irá garantir com que o isolamento entre o terra e o material condutor funcione de uma maneira adequada(MATTEDE, 2023).

Figura 1 – Isolador de Vidro para Linha de Transmissão



Fonte: (SOUZA, 2019)

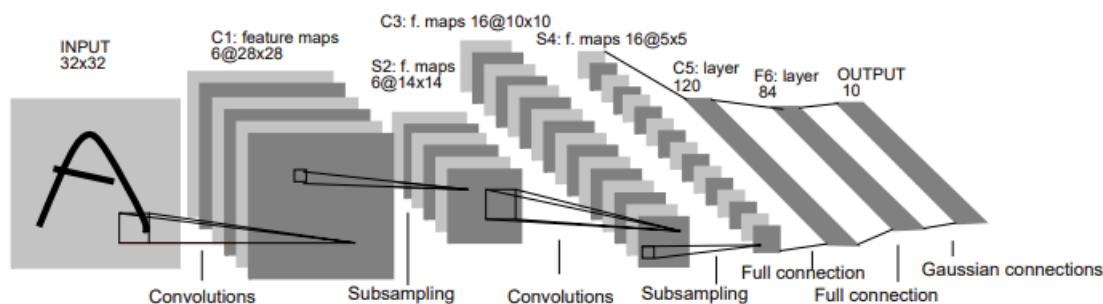
No contexto de uma linha de transmissão, os isoladores são os dispositivos mais frágeis, estando submetidos a esforços como variações de temperatura, poluição no entorno, pressões mecânicas por influência de ventos, atos de vandalismo e surtos de sobretenção(SOUZA, 2019).

### 2.1.2 Redes Neurais Convolucionais (CNNs)

A rede neural convolucional, é um tipo particular de rede neural que recebe por entrada imagens que serão utilizadas para o reconhecimento de padrões ou para a execução da classificação. Nesse sentido, elas foram inicialmente introduzidas por (LECUN et al., 1998) no qual ele tinha o desafio de executar o reconhecimento de documentos escritos em impressos.

Naquele momento os algoritmos de classificação não eram tão precisos quanto ele precisava. Nesse sentido, Yann LeCun propôs uma arquitetura de rede neural batizada de Lenet 5, para a extração das informações de imagens a serem analisadas. Assim, a rede era composta pelas camadas evolucionais, as camadas de pooling, e uma camada final que executa a classificação da imagem. Mais abaixo, serão detalhadas os principais tópicos acerca de uma CNN.

Figura 2 – Arquitetura da Lenet 5



Fonte: (LECUN et al., 1998)

#### 2.1.2.1 Entradas

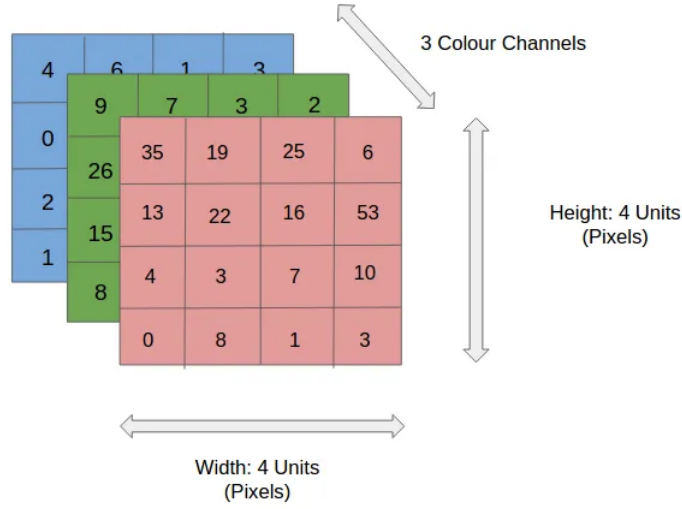
Assim como foi mencionado anteriormente, as Redes Neurais Convolucionais recebem imagens como fonte de entrada. No entanto, as imagens baseiam em matrizes tridimensionais, em que cada matriz representaria um canal do RGB, por exemplo. Ademais, a dimensão das matrizes define a resolução da imagem.

#### 2.1.2.2 Camadas Convolucionais

A camada receptora dos dados de entrada, é conhecida por Camada Convolucional, executa um processo de filtragem, que consiste em uma varredura gradual sobre os dados de entrada. Matematicamente, são executadas multiplicações escalares com a matriz de kernel, na qual será resultará em uma matriz de saída armazenará as principais características da referida entrada.



Figura 3 – Entrada de uma Rede Neural Convolutacional



Fonte: (ALVES, 2018)

### 2.1.2.3 Funções de ativação

A matriz resultante da convolucional passará por uma função de ativação, que tem como objetivo, introduzir a não linearidade à rede, fazendo com que ela possa aprender com maior facilidade. Funções como a tangente hiperbólica e sigmoidal são populares no contexto de redes neurais, contudo, a função ReLU (Rectified Linear Unit, em inglês), costuma ser a mais utilizada (ALBAWI; MOHAMMED; AL-ZAWI, 2017).

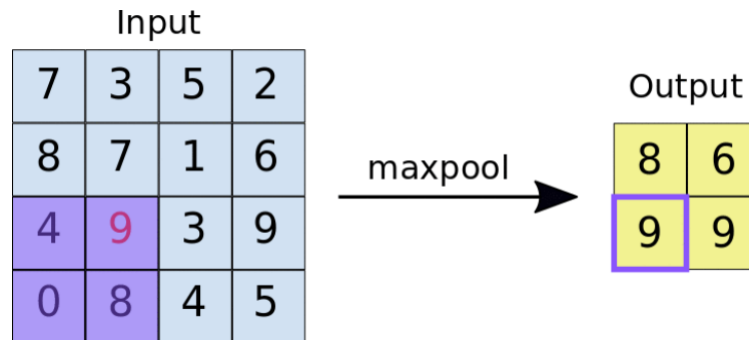
$$f(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ x & \text{se } x > 0 \end{cases} \quad (2.1)$$

### 2.1.2.4 Camada de Pooling

Após a aplicação da função de ativação, a Camada de Pooling aplica um processo de varredura, como a vista na seção 2.1.2.2, contudo, sem a utilização de um kernel. A ideia central nesta etapa é extrair o valor mais pertinente à sub-região analisada. Citando o método mais utilizado na etapa de *Pooling* (ALVES, 2018), o *Max Pooling* baseia-se no processo de varredura, selecionando os maiores valores de cada sub-região, como pode ser visto na figura abaixo.

Com isso, o processo de convolução, aplicação da função de ativação, e a execução do pooling é repetido sucessivamente, dependendo da profundidade da rede, até o momento em que é efetuado a classificação da imagem de entrada.

Figura 4 – Aplicação da técnica de Max Pooling em uma sub-região



Fonte: (ALVES, 2018)

## 2.2 Estado da Arte

Neste segundo momento, são discutidos alguns dos principais modelos arquiteturais que compõem o estado da arte do campo de Visão Computacional. Nesse sentido, o primeiro caso a ser analisado, é discutido em (REN et al., 2015), no qual ele propõe um novo *framework* para a detecção de objetos denominado Faster R-CNN (em inglês, Region Based Convolutional Neural Network), uma variação de uma R-CNN tradicional, proposta pelo mesmo autor. Fazendo-se uma comparação entre as R-CNNs e as CNNs tradicionais, as R-CNNs são uma pequena variação da primeira anteriormente citada, com a capacidade de gerar um conjunto de propostas de regiões. Assim, elas são definidas como caixas delimitadoras que conterão determinados objetos. Com isso, essas regiões alimentam a CNN para executar a classificação e localização.

Contudo, apesar de serem consideravelmente mais precisas que o modelo original, as R-CNNs são consideravelmente mais lentas por requererem a execução de duas etapas distintas (geração de regiões e classificação do modelo / localização dos objetos). Assim, para a atenuação deste problema, os autores propõem uma variação do modelo, que introduz o conceito de RPN, denominado Rede de Proposta de Região. Além disso, o artigo introduz a utilização de unidades de processamento gráfico para o treinamento dessas novas redes, resultando em um aumento substancial na velocidade de treinamento. Por fim, a introdução dessa nova arquitetura de rede se tornou vigente no campo da visão computacional.

A segunda arquitetura a ser comentada, é a de Vision Transformers (ViTs). Ela se baseia em um modelo de RNA arquitetura de um Transformer, inicialmente utilizada para o Processamento de Linguagem Natural, sendo então adaptadas para o contexto de VC. De acordo com (BHALERAO, 2023), a grande diferença entre uma ViT e uma CNN tradicional se dá na troca camadas convolucionais por uma pilha de camadas de atenção

própria.

Diante desse cenário, o trabalho de (RAGHU et al., 2021) traz uma análise comparativa entre a capacidade de classificação das ViTs frente a CNN, uma arquitetura mais tradicional. O observado foi um desempenho similar ao visto na CNN, apesar de possuírem uma arquitetura diferente e que até então não era compatível com a resolução deste tipo de problema. Além disso, as ViTs apresentaram uma taxa de aprendizado maior a medida que o dataset de treinamento utilizado aumentava de tamanho.

A terceira arquitetura a ser comentada, é a conhecida como YOLO (REDMON et al., 2016), acrônimo de *You only look once*. Assim como as outras arquiteturas, ela também é capaz de executar a detecção de objetos em imagens e vídeos, mas podendo executar classificação em tempo real, dependendo do hardware utilizado. Uma característica peculiar dessa arquitetura, é a forma como o processo de detecção é implementado, trocando a classificação pela regressão. Nesse sentido, o estudo recente de (JIANG et al., 2022) também traz uma análise comparativa da performance das diferentes versões da arquitetura YOLO frente aos modelos tradicionais de CNNs. A conclusão denota uma melhora na velocidade no processo de identificação e na precisão dos modelos, havendo ainda uma margem para a melhoria dos modelos futuros.

Figura 5 – YOLO sendo aplicado para difentes conjuntos de imagens



Fonte: (REDMON et al., 2016)

## 2.3 Trabalhos Correlacionados

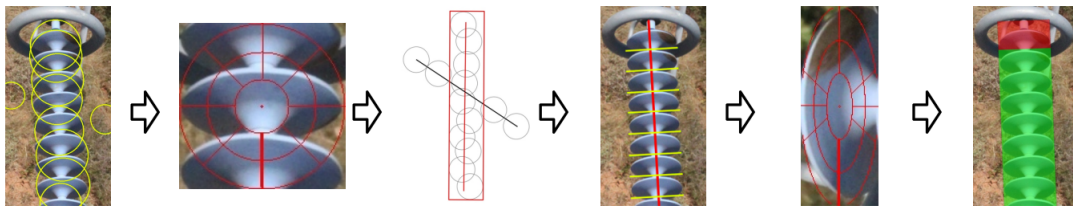
Neste momento, trazemos alguns estudos orientados para a aplicação em mundo real e que de certa forma demonstram alguma correlação com o tema de estudo deste trabalho. O primeiro a ser mencionado é o de (OBERWEGER; WENDEL; BISCHOF, 2014), no qual foi proposta uma abordagem de detecção de isoladores com imagens aéreas, e analisar as imagens automaticamente em busca de potenciais falhas.

Para a metodologia, foi implementado um treinamento discriminativo de descritores locais baseados em gradientes e um esquema de votação subsequente para a localização.

Assim, a abordagem foi baseada com o auxílio de técnicas de reconhecimento de padrão, incluindo uma abordagem de agrupamento baseada em RANSAC <sup>1</sup>(do inglês, *Random Sample Consensus*). Com esse método o processo de identificação do isolador seria facilitado ainda que estivesse em ambientes desfavoráveis, independente do seu formato, tamanho, ou posicionamento na imagem.

Sobre o processo de detecção da falha do isolador, o dispositivo seria destacado em torno de um eixo principal e dividido em partes individuais ao longo do eixo. Com isso, cada uma das partes será avaliada por um descritor elíptico, no qual irá definir uma pontuação para o potencial nível da falha. Assim, o processo de segmentação traz uma maior precisão para o processo de identificação da falha. Por conseguinte, o método trouxe uma taxa 95% de verdadeiros positivos, e uma taxa de 12% para falsos positivos, mostrando ser melhora significativa aos métodos até então disponíveis.

Figura 6 – Técnica de detecção de falha em Isolador



Fonte: (OBERWEGER; WENDEL; BISCHOF, 2014)

Um segundo trabalho correlacionado (JENSSEN; ROVERSO et al., 2018) com o tema central deste trabalho começa introduzindo alguns dos principais desafios enfrentados na questão do monitoramento e inspeção de dos isoladores em linhas de transmissão. O monitoramento, que anteriormente era feito de maneira manual e era considerado um procedimento lento e perigoso, passou por um momento de transição com a introdução dos veículos aéreos não tripulados para a realização dessa tarefa.

Nesse sentido, o artigo detalha de maneira clara os difentes métodos e tarefas de inspeção, e sugere a utilização de múltiplos conjuntos de dados para a inspeção visual. Dentre as possíveis formas para a montagem destes conjuntos de dados, os autores sugerem por exemplo, a utilização de câmeras térmicas e câmeras sensíveis à radiação ultravioleta, sobretudo para a detecção de fenômenos que não são visíveis, como o efeito Corona.

Além disso, os autores ressaltam o potencial papel das técnicas de aprendizado profundo e das redes neurais convolucionais (CNN's) como uma abordagem avançada para a inspeção autônoma baseada em visão.

<sup>1</sup> Segundo (DERPANIS, 2010), RANSAC pode ser definido como uma técnica de estimação de parâmetros para um número considerável de *outliers* em um dado conjuntos de dados de entrada, sendo então um método popular no campo de Visão Computacional.



Figura 8 – Técnica de pré-processamento da imagem do isolador



Fonte: ([CORSO et al., 2021](#))

multicamadas , a Árvore de decisão e a Máquina de vetor de suporte. Como resultado, o método se mostrou o mais eficaz diante das metodologias, sendo um potencial candidato para a resolução de problemas do mesmo gênero.

## 3 Metodologia

Neste capítulo, será discutido a metodologia aplicada para o monitoramento dos isoladores com câmeras térmicas e a utilização das técnicas de reconhecimento de padrão para a identificação das quebras e trincas. Nesse sentido, serão discutidos três etapas principais na metodologia. Na primeira etapa, será explicado e justificado a maneira como a base de dados será implementada, na segunda, será implementado o pré-processamento da base de dados, discutindo algumas metodologias para a melhoria da visualização das imagens, implementando algumas técnicas de filtragem, por exemplo. Por último, será analisado o planejamento experimental do trabalho, incluindo as arquiteturas utilizadas, e métricas de avaliação do trabalho.

### 3.1 Definição da Base de Dados

#### 3.1.1 Descrição da População

A base de dados será criada, a partir da filmagem de isoladores presentes em linhas de transmissão e subestações, sendo divididos em isoladores de apoio ou de suspensão por exemplo, podendo ser poliméricos, de porcelana, ou vidro.

#### 3.1.2 Aquisição de Dados

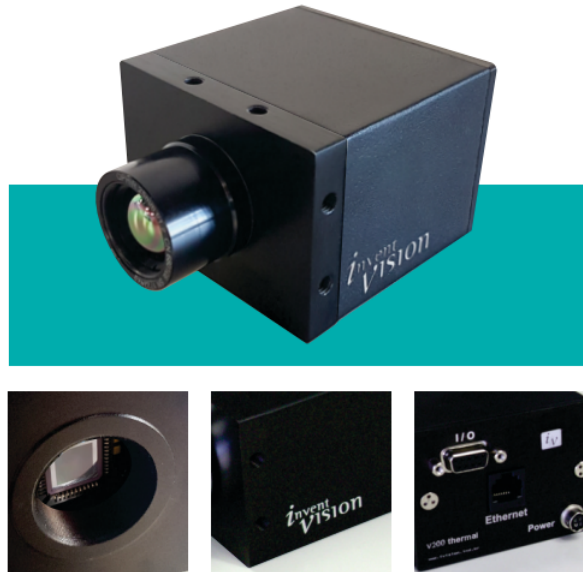
A obtenção de dados será realizada com o auxílio de uma câmera térmica, disponibilizado pela empresa *Invent Vision*, modelo *iCam LW 200*. Com ela, serão feitas filmagens dos isoladores, com o auxílio ou não de veículos aéros não tripulados, popularmente conhecidos como *drones*. Abaixo são descritas algumas informações importantes a respeito da câmera.

##### 3.1.2.1 Descrição da Câmera utilizada

A LW 200 é uma câmera térmica capaz de realizar filmagens em ambientes abertos e fechados com grande qualidade. O processamento das imagens é feito em tempo real por meio de uma arquitetura FPGA (*Field Programmable Gate Array*, em inglês), e a transmissão de vídeo é feita por meio de uma interface Ethernet([INVENT VISION, 2023b](#)). Além disso, a câmera também oferece um SDK (*Software Development Kit*, em inglês) para a captura e visualização das imagens pelo usuário. Abaixo, é mostrada imagens da referida câmera, bem como dados importantes do datasheet([INVENT VISION, 2023a](#)).



Figura 9 – iCam LW-200



Fonte: ([INVENT VISION, 2023a](#))

Tabela 1 – Especificações técnicas da câmera iCam LW-200

Resoluções	384x288 ou 640x480
Taxas de Quadros por Segundo	30 ou 50 FPS
Faixa Espectral	8 a 14 $\mu\text{m}$
Temperatura de Operação	-20 a 60°C
Lentes Compatíveis	9 mm - 12,8 mm - 25 mm - 50 mm - 100 mm
Saída de Vídeo	TCP/IP
Interface de Ethernet	10/100 Mbits/s
Peso	290g
Dimensões	78 x 58 x 82 mm
Alimentação	6 Vcc

### 3.1.3 Rotulagem

A etapa de rotulagem será executada com o auxílio de especialistas em campo indicando a presença ou não dos defeitos nos isoladores. Outras informações importantes dos componentes analisados serão salvos, como o tipo de material, contexto de utilização, e a qualidade da iluminação, por exemplo. Informações estas que serão fundamentais para a próxima etapa, na qual será descrita o processo de pré-processamento das imagens.



## 3.2 Pré-processamento dos Dados

### 3.2.1 Normalização e Redimensionamento

No processo de normalização e redimensionamento, será executada uma padronização na resolução das imagens, apesar da câmera utilizada ter a capacidade de executar as filmagens em diferentes resoluções. Assim, a etapa de processamento subsequente será facilitada.

### 3.2.2 Aumento da Base de Dados

Com o objetivo de aumentar o tamanho do *dataset*, as imagens dos isoladores serão capturadas sobre ângulos, orientações e distâncias.

### 3.2.3 Remoção de Ruídos

Na etapa de remoção de ruídos, será implementada a técnica do filtro de mediana, particularmente útil para a remoção de ruído de sal e pimenta. Além disso, ela também é muito importante para a preservação de bordas e detalhes enquanto reduz possíveis interferências.

### 3.2.4 Segmentação

A segmentação se baseará na região de interesse dos isoladores, na qual será destacada a região em que o componente apresentará a trinca ou quebra, é mais provável de acontecer. Para isso, utiliza-se a técnica de segmentação do threshold, aplicando a técnica de limiar para fazer a distinção entre a região de interesse com o fundo.

## 3.3 Planejamento Experimental

### 3.3.1 Divisão do Conjunto de Dados

Após a execução do tratamento da base de dados descritos na seção anterior, passamos agora a discutir a forma como o *dataset* é manipulado. Inicialmente, a base de dados será dividida em três conjuntos distintos: o conjunto de treinamento, o conjunto de validação e o conjunto de teste.

Para o conjunto de treinamento, serão destinados cerca de 70% das amostras do *dataset*. Além deste percentual, outros 10% do espaço amostral serão destinados para o conjunto de validação, que será fundamental para o ajuste de hiperparâmetros e evitar um potencial *overfitting* dos modelos. E os últimos 20% do conjunto inicial, serão aplicados para a avaliação de desempenho dos modelos.

Naturalmente, será tomado o cuidado para com que não ocorra o desbalanceamento da base de dados, no sentido que em ocorra a presença de poucas imagens de isoladores danificados em algum dos conjuntos citados, por exemplo.

### 3.3.2 Seleção de Algoritmos

Para a etapa de seleção de algoritmos, serão selecionadas algumas arquiteturas pertinentes para a etapa de classificação, dentre elas, algumas listadas na seção de "Estado da Arte", do capítulo anterior.

### 3.3.3 Treinamento do Modelo

Após a escolha das arquiteturas, serão efetuados os treinamentos dos modelos, incluindo o ajuste dos hiperparâmetros, taxa de aprendizado, número de camadas, visando evitar um potencial *overfitting*.

### 3.3.4 Métricas de Avaliação do Modelo

Na etapa de avaliação do modelo, serão definidas algumas métricas para a avaliação do modelo sobre difentes aspectos. Dentre as algumas métricas recomendadas ([RODRIGUES, 2019](#)), são consideradas:

- Cálculos das Acurácias com os respectivos desvios-padrões, variando parâmetros da arquitetura;
- Recall(revoação): analisando as situações de classe Positivo como valor esperado, quantas estão corretas.
- F1-score: Média Harmônica entre a acurácia e revocação.
- Matriz de confusão: Tabela que indica os erros e acertos do modelo, comparado com o resultado esperado.

# Referências

- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. In: IEEE. *2017 international conference on engineering and technology (ICET)*. [S.l.], 2017. p. 1–6. Citado na página 15.
- ALVES, G. Entendendo redes convolucionais (cnns). 2018. Disponível em: <<https://medium.com/neuronio-br/entendendo-redes-convolucionais-cnns-d10359f21184>>. Citado 2 vezes nas páginas 15 e 16.
- BHALERAO, C. Vision transformers [vit]: A very basic introduction. 2023. Disponível em: <<https://medium.com/data-and-beyond/vision-transformers-vit-a-very-basic-introduction-6cd29a7e56f3>>. Citado na página 16.
- CORSO, M. P. et al. Classification of contaminated insulators using k-nearest neighbors based on computer vision. *Computers*, MDPI, v. 10, n. 9, p. 112, 2021. Citado 2 vezes nas páginas 19 e 20.
- DERPANIS, K. G. Overview of the ransac algorithm. *Image Rochester NY*, v. 4, n. 1, p. 2–3, 2010. Citado na página 18.
- INVENT VISION. *Datasheet da Câmera iCam LW-200*. [S.l.], 2023. Disponível em: <<http://www.inventvision.com.br/products/lw200/index.html>>. Citado 2 vezes nas páginas 21 e 22.
- INVENT VISION. *Manual da Câmera iCam LW-200*. [S.l.], 2023. Disponível em: <<http://www.inventvision.com.br/products/lw200/index.html>>. Citado na página 21.
- JENSSEN, R.; ROVERSO, D. et al. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. *International Journal of Electrical Power & Energy Systems*, Elsevier, v. 99, p. 107–120, 2018. Citado 2 vezes nas páginas 18 e 19.
- JIANG, P. et al. A review of yolo algorithm developments. *Procedia Computer Science*, Elsevier, v. 199, p. 1066–1073, 2022. Citado na página 17.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, Ieee, v. 86, n. 11, p. 2278–2324, 1998. Citado na página 14.
- MATTEDE, H. Isoladores para cabos elétricos. 2023. Disponível em: <<https://www.mundodaeletrica.com.br/isoladores-para-cabos-eletricos/>>. Citado na página 13.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, p. 115–133, 1943. Citado na página 11.
- MOHAJAN, H. Third industrial revolution brings global development. 2021. Citado na página 11.

OBERWEGER, M.; WENDEL, A.; BISCHOF, H. Visual recognition and fault detection for power line insulators. In: *19th computer vision winter workshop*. [S.l.: s.n.], 2014. p. 1–8. Citado 2 vezes nas páginas 17 e 18.

PATNAIK, S. New paradigm of industry 4.0. *Switzerland: Springer*, Springer, 2020. Citado na página 11.

RAGHU, M. et al. Do vision transformers see like convolutional neural networks? *Advances in Neural Information Processing Systems*, v. 34, p. 12116–12128, 2021. Citado na página 17.

REDMON, J. et al. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 779–788. Citado na página 17.

REN, S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks. In: CORTES, C. et al. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2015. v. 28. Disponível em: <[https://proceedings.neurips.cc/paper\\_files/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf)>. Citado na página 16.

RODRIGUES, V. Métricas de avaliação: acurácia, precisão, recall... quais as diferenças? 2019. Disponível em: <<https://vitorborbarodrigues.medium.com/métricas-de-avaliação-acurácia-precisão-recall-quais-as-diferenças-c8f05e0a513c>>. Citado na página 24.

SOUZA, A. C. V. D. Isoladores de linhas de transmissão. 2019. Disponível em: <<https://www.linkedin.com/pulse/isoladores-de-linhas-transmissão-antonio-carlos-vieira-de-souza/?originalSubdomain=pt>>. Citado na página 13.