

UNIVERSIDADE FEDERAL DO PARANÁ

JOÃO CLAUDIO NOGUEIRA

RECONHECIMENTO E CLASSIFICAÇÃO DE IMAGENS DE VOO
UTILIZANDO REDE NEURAL CONVOLUCIONAL

Curitiba
2023

JOÃO CLAUDIO NOGUEIRA

RECONHECIMENTO E CLASSIFICAÇÃO DE IMAGENS DE VOO
UTILIZANDO REDE NEURAL CONVOLUCIONAL

Dissertação apresentada como requisito à obtenção do título de Mestre, do programa de Pós-Graduação em Engenharia de Manufatura, Setor de Tecnologia da Universidade Federal do Paraná.

Orientador: Prof. Dr. Pablo Valle

Co-orientador: Prof. Dr. Alessandro Marques

Curitiba
2023

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA CIÊNCIA E TECNOLOGIA

Nogueira, João Claudio.

Reconhecimento e classificação de imagens de voo utilizando Rede Neural convolucional. / João Claudio Nogueira. – Curitiba, 2023.

1 recurso on-line : PDF.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-Graduação em Engenharia Manufatura.

Orientador: Prof. Dr. Pablo Valle.

Coorientador: Prof. Dr. Alessandro Marques.

1. Engenharia. 2. Rede neural. 3. Indústria 4.0. 4. Curitiba(PR). I. Valle, Pablo. II. Marques, Alessandro. III. Universidade Federal do Paraná. Programa de Pós-Graduação em Engenharia Manufatura. IV. Título.

Bibliotecário: Nilson Carlos Vieira Júnior CRB-9/1797



TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação ENGENHARIA DE MANUFATURA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **JOÃO CLAUDIO NOGUEIRA** intitulada: **RECONHECIMENTO E CLASSIFICAÇÃO DE IMAGENS DE VOO UTILIZANDO REDE NEURAL CONVOLUCIONAL**, sob orientação do Prof. Dr. PABLO DEIVID VALLE, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 15 de Dezembro de 2022.

Assinatura Eletrônica
18/12/2022 17:53:48.0
PABLO DEIVID VALLE
Presidente da Banca Examinadora

Assinatura Eletrônica
16/12/2022 13:52:18.0
FERNANDO DESCHAMPS
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
16/12/2022 14:33:14.0
DAVID MENOTTI GOMES
Avaliador Externo (PPG INFORMÁTICA - UFPR)

Assinatura Eletrônica
11/01/2023 15:34:58.0
ALESSANDRO MARQUES
Coorientador(a) (DEPARTAMENTO DE ENG. MECÂNICA UFPR)

RESUMO

Segundo Statista (2020) a corrosão em estruturas é uma falha estrutural que gera custos de trilhões de dólares em toda parte do mundo anualmente e ideias inovadoras com foco em manutenção preventiva podem reduzir em até 35% gastos relacionados a oxidações em telhados industriais. Neste estudo que foi conduzido, na maior parte de seu tempo, em uma unidade fabril de aproximadamente 89.000 m² de área construída com 60.000 m² de telhados, situada na cidade industrial de Curitiba/PR (Planta A), observou-se que este é um problema atual uma vez que houve, no decorrer dos últimos anos, um aumento exponencial no volume de falhas em coberturas, impactando na longevidade da estrutura e na sustentabilidade do negócio. Através de uma metodologia de análise das condições básicas de manutenção, constatou-se que a oxidação é responsável por uma grande parcela das falhas destas edificações. Neste sentido, o objetivo deste trabalho foi desenvolver um sistema inteligente capaz de detectar, classificar e alertar sobre falhas de oxidação em estruturas metálicas de telhados. Para isso, o conceito de transdisciplinaridade foi aplicado na integra através da união de várias ciências. Os dados foram obtidos utilizando diferentes drones com câmeras distintas para captura de imagem dinâmica. Experimento de Taguchi aliado a fatores de voo foram realizados para entender a correlação entre qualidade de imagem e aprendizado da IA. Diferentes maneiras de processamento dos dados de entrada foram avaliadas, bem como a associação de técnicas computacionais, análises da capacidade de generalização, *overfitting* e melhoria da performance foram discutidas. Escolheu-se pelo aprendizado de máquina para realizar o processamento de imagens e através da referência literária de Olaf Ronneberger (2015) idealizou-se o uso de duas redes neurais convolucionais do tipo U-NET, uma para a localização objetiva do telhado e a outra para a detecção assertiva da oxidação. A junção das redes formou a base da solução. Além disso através da segmentação dupla e frações de imagens de 640x640 pixels, tornou o processo de aprendizagem da IA e a localização da falha mais ágeis. Realizou-se a avaliação da eficácia da pesquisa através da aplicação da técnica em outra unidade fabril (Planta B), localizada na Cidade de Campinas/SP, e em conjunto com análises estatísticas observou-se resultados satisfatórios o que gerou maior interesse da indústria e associações de facility management na solução tecnológica proposta.

Palavras-chave: Manutenção. Industria 4.0. VANT. Redes neurais convolucionais.

ABSTRACT

According to Statista (2020) corrosion in structures is a structural failure that generates costs worldwide, worth around trillions of dollars per year and innovation ideas with focus on preventive maintenance can reduce up to 35% of the costs that come from oxidation in the industrial roofs. In this research, where most of the experiments were made in a manufacturing plant with around 89.000 m² of built area, located at the industrial city of Curitiba, Paraná (Plant A), an exponential increase in these failures was observed in the last few years and reducing the lifecycle of the roofs and business sustainability significantly. With the support of some analyses based on the Total Productive Maintenance methodology, oxidation was observed to be one of the most relevant aspects on these failures. In this work, an intelligent system was proposed to detect, classify oxidations and promotes alert upon failures in metal structures of roofs. For all that to happen, the transdisciplinary concept was totally applied through several sciences united. The data was obtained through different cameras mounted in different UAVs. Taguchi analysis regarding flight factors were accomplished to understand correlations between quality of image and AI learning. Different ways of input data processing were studied, such as integration of computational techniques, generalization capacity, overfitting and accuracy improvement were discussed. Decided to applied machine learning to make the images processing and following reference of Olaf Ronneberger (2015) two U-NET convolutional neural network were proposed, the first for localization of the roof and the second for oxidation detection. The concatenation of networks became the base of this solution. Moreover, due to dual segmentation applied in 640x640 image, the learning process and identification of the failures were most fast. To validate the efficiency upon this researching the technique was applied in another manufacturing plant (Plant B), located at the city of Campinas, São Paulo, and by using statistics analysis, good performance of the IA was noticed. Take into consideration the good results obtained, associations of facility management and industries raised the interest towards this solution.

Keywords: Maintenance. Industry 4.0. UAV. Convolutional neural network.

LISTA DE ILUSTRAÇÕES

Figura 1: (a) Componentes dos telhados que apresentam oxidações como falhas em suas estruturas. (b) Ações necessárias para a contenção de infiltrações nas áreas fabris. (c) Ações corretivas necessárias nas estruturas danificadas. Fonte: o autor (2023).....	13
Figura 2: Fluxo sumarizado da solução, desde o uso do drone até a notificação do time de manutenção. Fonte: o autor (2023).....	15
Figura 3: Eixos de pesquisa e seus respectivas palavras chave.....	18
Figura 4: Representação de um mapa de classificação obtido por uma rede neural da cidade de Washington, DC. Fonte: Adaptado de (Benediktsson, Palmason & Sveinsson, 2005).	24
Figura 5: Imagens dos corpos classificados e suas respectivas imagens térmicas. Fonte: (Rudol & Doherty, 2008.)	25
Figura 6: Dados aéreos coletados pelo modelo GA-ASI ALTUS® II em uma altitude de 945 metros, aproximadamente, sobre o incêndio controlado em El Miragem no dia 6 de setembro de 2001 na Califórnia. Os pequenos princípios de incêndio podem ser vistos pelos círculos na imagem. Fonte: (Ambrosia <i>et al</i> , 2003).....	25
Figura 7: Representação RGB. Fonte: (O Autor).....	28
Figura 8: Representação de neurônios biológicos conectados. Fonte: (Neto, 2014).....	29
Figura 9: Representação matemática de um neurônio. Fonte: (Mcculloch & Pitts, 1943).....	30
Figura 11: Representação de condições de posição de imagem para reconhecimento. Fonte: O autor (2023)	32
Figura 11: Representação de segmentação semântica. Fonte: O autor (2023).	33
Figura 12: Etapas do projeto. Fonte: o autor (2023).....	37
Figura 13: Imagem fonte, utilizada para o processo de anotação. Fonte: o autor (2023).	38
Figura 14: Imagem fonte carregada no <i>Labelimg</i> e passando pelo processo de apontamento de objetos. Fonte: o autor (2023).....	39
Figura 15: Exemplo de arquivo gerado ao final do processo de anotação de uma imagem. Fonte: o autor (2023)	39
Figura 16: Redução do tamanho do arquivo de entrada da rede neural. Fonte: o autor (2023).	39
Figura 17: Distribuição do <i>dataset</i> para a etapa de treinamento. Fonte: o autor (2023).	40

Figura 18: Arquitetura U-Net para segmentação de imagem. Fonte: Ronneberger, 2015 modificada.	42
Figura 19: Arquitetura do sistema desenvolvido. Fonte: o autor (2023).	43
Figura 20: Primeira RNC para localização do telhado. Fonte: o autor (2023).	43
Figura 21: Segunda RNC para identificação da oxidação. Fonte: o autor (2023).	44
Figura 22: Vista do <i>software</i> referente ao objeto de estudo. Fonte: o autor (2023).	45
Figura 23: Perfil de relevo obtido. Fonte: o autor (2023).	46
Figura 24: demonstrativo da segunda Laplaciana de uma imagem. Fonte: o autor (2023). ...	47
Figura 25: Processo geral para a avaliação das velocidades de voo. Fonte: o autor (2023). .	47
Figura 26: Imagem obtida para uma velocidade de 25 km/h.	48
Figura 27: Imagem obtida para uma velocidade de 45 km/h.	48
Figura 28: Comparativo de uma mesma área na qual as imagens foram capturadas em períodos distintos sendo eles respectivamente manhã, tarde e à noite. Fonte: o autor (2023).	51
Figura 29: Processo de avaliação da influência da iluminância na qualidade da imagem. Fonte: o autor (2023).	51
Figura 30: Histograma obtido referente aos três períodos analisados. Em azul, está representado a distribuição obtida para o período da manhã com aproximadamente 140.000 lx, em laranja está representado o período da tarde, com cerca de 40.000 lx, e em verde, está representado o período noturno, com cerca de 100 lx. Fonte: o autor (2023).	52
Figura 31: Experimento Taguchi obtido. Fonte: o autor (2023).	53
Figura 32: Ilustração do processo de cálculo para a determinação do intervalo de disparos da câmera. Fonte: o autor (2023).	54
Figura 33: Comparativo entre uma imagem original, retirada pelo VANT (esquerda), e uma mesma área gerada pelo ortomosáico. Fonte: o autor (2023).	55
Figura 34: Resultados obtidos da rede com validação em Ctp. Fonte: o autor (2023).	58
Figura 35: Resultados obtidos da rede com validação em Cap. Fonte: o autor (2023).	58
Figura 36: Resultados obtidos da rede. Fonte: o autor (2023).	59
Figura 37: Resultados obtidos da rede. Fonte: o autor (2023).	59

Figura 38: Ortomosaico com localização e intensidade das oxidações. Fonte: o autor (2023).	60
Figura 39: Página inicial de navegação do software gerenciado pela IA. Fonte: o autor (2023).	61
Figura 40: Relatório automatizado enviado por e-mail. Fonte: o autor (2023).	61

LISTA DE ACRÔNIMOS

MDCB	Manutenção das condições básicas
TPM	Manutenção preventiva total

SUMÁRIO

1. INTRODUÇÃO	11
1.1. CONTEXTUALIZAÇÃO.....	11
1.2. FORMULAÇÃO DO PROBLEMA.....	12
1.3. OBJETIVO GERAL.....	14
1.4. OBJETIVOS ESPECÍFICOS.....	15
1.5. CONTRIBUIÇÕES DO TRABALHO.....	16
2. FUNDAMENTAÇÃO TEÓRICA	17
2.1. INDÚSTRIA 4.0.....	18
2.2. VEÍCULOS AÉREOS NÃO TRIPULADOS PARA SENSORIAMENTO	23
2.3. REDES NEURAIS CONVOLUCIONAIS.....	26
3. MATERIAIS E MÉTODOS.....	35
3.1. PRÉ-ESTUDO DE TÉCNICAS PARA CONSTRUÇÃO RNC	36
3.2. PROCESSAMENTO E PREPARO DOS DADOS PARA O TREINAMENTO DA REDE 37	
3.3. BASE DE DADOS.....	40
3.4. YOU ONLY LOOK ONCE (YOLO)	41
3.5. U-NET	41
3.6. ARQUITETURA DO SISTEMA	42
4. RESULTADOS E DISCUSSÕES	45
4.1. ESTUDO DE RELEVO.....	45
4.2. CONDIÇÕES DE VOO	46
4.3. EXPERIMENTO TAGUCHI	52
4.4. ALGORÍTMO ORTOMOSÁICO	53
4.5. TREINAMENTO E VALIDAÇÃO RNA.....	55
4.6. CLASSIFICAÇÃO DAS OXIDAÇÕES	59
4.7. SOLUÇÃO TECNOLÓGICA – RELATÓRIO DE DIAGNÓSTICO PREVENTIVO ..	60
5. CONSIDERAÇÕES FINAIS	62
REFERÊNCIAS BIBLIOGRÁFICAS.....	65

1. INTRODUÇÃO

Nesta seção, serão abordados os pontos fundamentais para o entendimento e aplicação deste estudo. Será repassado uma breve fundamentação teórica por trás do problema fundamental. Em seguida, uma contextualização da situação problemática no objeto de estudo desta pesquisa. Após, será abordado qual os objetivos gerais e específicos. Por fim, será posto em evidência quais as principais contribuições deste trabalho para os âmbitos industrial e acadêmico.

1.1. CONTEXTUALIZAÇÃO

Corrosão é um grande problema estrutural, presente em todo o globo terrestre (Atha & Jahanshahi, 2018). Os custos globais relacionados à corrosão giram em torno de 2,5 trilhões de dólares/ano, o que corresponde à 2,95% do valor total do produto doméstico bruto (GDP) mundial (Statista, 2020).

É estimado que, fazendo uso das práticas de controle e prevenção de corrosão, pode-se economizar entre 15 a 35% dos custos totais desta falha (Koch et al, 2016). Estes custos não contabilizam valores indiretos, ligado principalmente à segurança de indivíduos e/ou ambiente. Através de quase-acidentes, incidentes, interrupções, acidentes e dentre outros fatores e consequências, diversas indústrias interpretaram que, de fato, o gerenciamento de corrosões pode ser custoso (Koch et al, 2016).

Na literatura, diversas metodologias para a detecção de corrosão foram propostas ao longo dos anos. Algoritmos que aplicam análises de textura e cor, baseados em abordagens estatísticas ou de filtros, são práticas comuns, podendo citar a que fazer o uso da transformada *wavelet* (Atha & Jahanshahi, 2018).

Matrizes de nível de cinza de coocorrência foram utilizadas por Haralick, onde estas matrizes representam a frequência de diferentes níveis de cinza em uma imagem, atuando como uma característica de textura estatística (Haralick, Shanmugam & Dinstein, 1973). Furuta e colaboradores desenvolveram uma rede neural baseada nos valores de saturação e matriz, afim de determinar o nível de deterioração de pontes para auxiliar a tomada de decisão de inspetores inexperientes (Furuta, Deguchi & Kushida, 1995). Livens e colaboradores utilizaram uma metodologia de decomposição de imagens via transformada *wavelet* e reconhecimento realizado através de uma rede de quantização vetorial de aprendizado para a classificação de corrosões (Livens et al., 1995). Gunatilake e outros propuseram uma metodologia de detecção de corrosões em

superfícies de aeronaves através de transformadas *wavelet* de sexta ordem (Gunatilake, Siegel & Jordan, 1997). Choi e Kim utilizaram os parâmetros de matriz, saturação e intensidade dos pixels para classificar múltiplos tipos de defeitos de corrosões (Choi & Kim, 2005). Lee e colaboradores classificaram ferrugens em revestimentos de aço de pontes com o auxílio de uma análise estatística dos canais de cores vermelho/verde/azul (Lee, Chang & Skibniewski, 2006). Medeiros e colaboradores aliaram as metodologias de matriz, saturação e intensidade e matrizes de níveis de cinza de coocorrência para detectar corrosões em tanques de aço carbono e em tubulações de uma refinaria de petróleo (Medeiros, Ramalho & Bento, 2010). Pakrashi e colaboradores propuseram um método fundamentado no contrastes óptico entre a região degradada comparada com seus arredores para a detecção e quantificação destas ocorrências em superfícies de alumínio (Pakrashi et al., 2010). Através de imagens de superfícies de pontes, revestidas em aço, e fazendo uso de uma transformada *wavelet* de Haar, Ghanta e outros detectaram defeitos nestas estruturas (Ghanta, Karp & Lee, 2011). Com o avanço da capacidade de processamento dos computadores, as redes neurais convolucionais vem ganhando, cada vez mais, força para atividades baseadas em imagens, visando a classificação, detecção ou segmentação de objetos ligados à corrosão (Atha & Jahanshahi, 2018; Yao et al., 2019; Holm et al., 2020; Hoskere et al., 2018).

1.2. FORMULAÇÃO DO PROBLEMA

No objeto de estudo deste trabalho, uma planta situada na cidade industrial de Curitiba, Paraná, com uma área total de 589.747 m² e 86.691 m² de área total construída, observou-se ao longo dos últimos anos um aumento no número de ocorrências de falhas relacionadas à infiltração de água em suas estruturas de cobertura. Dentre estas falhas, constatou-se uma grande parcela sendo referente à oxidação de elementos metálicos que compõem o sistema, como vigas calha, terminais de ventilação e rufos, como mostra a Figura 1(a). Além disto, fazendo uma análise do levantamento das manutenções de condições básicas (MDCB) destas estruturas, uma técnica amplamente utilizada no processo de manutenção preventiva total (TPM), notou-se que uma boa parte das falhas funcionais desta infraestrutura são provenientes de oxidações, conforme consta no Apêndice A, baseando-se no resultado do “Grau de impacto total (GIT)”, uma somatória de pontuações baseadas em aspectos relevantes para o planejamento e gestão de manutenção, como frequência de reparo, custos de

manutenção, variando entre 0 à 3, dependendo de sua criticidade conforme consta no Apêndice B.



Figura 1: (a) Componentes dos telhados que apresentam oxidações como falhas em suas estruturas. (b) Ações necessárias para a contenção de infiltrações nas áreas fabris. (c) Ações corretivas necessárias nas estruturas danificadas. Fonte: o autor (2023).

Por sua vez, estas ocorrências geram, em sua maioria, perdas de faturamento da planta, onde é necessário o isolamento do local para fins de manutenção e segurança dos colaboradores, Figura 1 (b) e (c).

Levando em consideração estes fatos, uma manutenção preventiva eficiente se faz necessário nestas estruturas. Adicionalmente, se ponderarmos alguns fatores como a idade média destes conjuntos ser de 25 anos, valor próximo ao seu ponto de falha estimado, um volume de 31 prédios distintos distribuídos ao longo de seus 589.747 m² de área total, os riscos envolvidos em trabalhos em altura (>2m de trabalho), a necessidade de mão de obra qualidade tanto para as manutenções preventivas como para análises dentre outros pontos, estima-se uma necessidade

anual, conforme listado na Tabela 1, de aproximadamente R\$ 148.800,00 para uma redução e prevenção de paradas de processos indesejadas e não planejadas.

Tabela 1: Comparativo entre custos de investimento total de uma manutenção preventiva entre o estado atual *versus* estado futuro, após o desenvolvimento desta pesquisa. Fonte: o autor (2023).

Dados relacionados	Estado atual	Futuro
Quantidade de prédios	31	31
Tempo médio por inspeção (hora/prédio)	02:00	00:02
Quantidade de horas trabalhadas por inspeção geral CtP	62:00	01:02
Quantidade de inspeções anuais (rotina de manutenção)	48	365
Custo por hora de inspeção (H/H)	R\$ 50,00	0
Gastos anuais com inspeções	R\$ 148.800,00	R\$15.000,00
Custo análise planejador/supervisor	R\$ 4.464,00	0
Custo análise/inspeção drone + AI (por hora)	R\$ 101,00	R\$ 41,09

Logo, este projeto propõe o desenvolvimento de uma rede neural convolucional para a detecção e classificação de falhas de oxidação nestas estruturas. As imagens foram captadas por meio de um veículo aéreo não-tripulado e um conjunto de câmeras, a fim de dar mais autonomia e consistência no processo de obtenção de dados. Estas informações serão processadas de maneira *offline* em um terminal e irão resultar em um relatório de levantamento, contendo a evolução de pontos críticos e o grau de cada ocorrência. Onde, ao fim disto, objetiva-se uma redução expressiva do custo de investimento anual para ações preventivas de inspeção, totalizando um valor de R\$ 41,05/hora, *versus* um valor inicial de R\$ 101,00/hora, resultando em uma redução de 40% por preventiva.

1.3. OBJETIVO GERAL

Desenvolver uma técnica de reconhecimento e classificação de falhas de oxidação em estrutura metálica de telhados de uma unidade fabril, tendo como base as técnicas de rede neural convolucional e a técnica de segmentação de imagens por Olaf Ronneberger (2015).

1.4. OBJETIVOS ESPECÍFICOS

Estabelecer rotina MDCB para levantar todas as possíveis falhas e seus agentes responsáveis dos telhados a fim de validar a falha de oxidação como fator relevante de pesquisa.

Estabelecer uma sequência estruturada de pesquisa considerando três etapas fundamentais abaixo e visão geral da pesquisa, Figura 15:

- Etapa 1, elaborar e otimizar plano e configuração de voo do drone que será utilizado na captura das imagens. Realizar experimento Taguchi para validação da influência dos parâmetros na qualidade da imagem e consequentemente impactos na performance da rede neural.
- Etapa 2, confeccionar algoritmo ortomosáico para concatenar imagens adquiridas.
- Etapa 3, estabelecer metodologia de aquisição de imagens. Arquitetar e otimizar rede neural convolucional para o reconhecimento e classificação de falhas.

Propor um método transdisciplinar que associe técnicas de engenharia de manutenção, visão computacional, estatística e aprendizado de máquina relacionando-os as operações de inspeção de telhados industriais a fim de fornecer um produto tecnológico mais ágil, seguro e mais acessível, Figura 2.

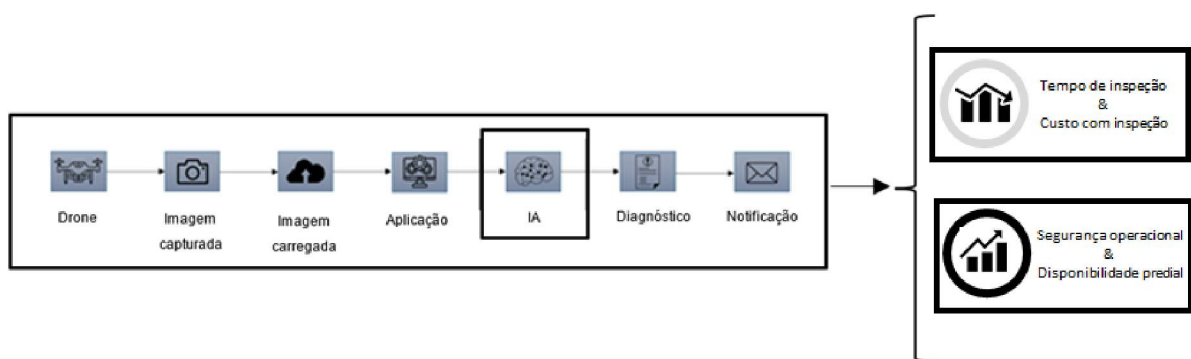


Figura 2: Fluxo sumarizado da solução, desde o uso do drone até a notificação do time de manutenção. Fonte: o autor (2023).

1.5. CONTRIBUIÇÕES DO TRABALHO

1.5.1. Contribuições para a academia

Para a academia, esta pesquisa objetiva-se aprofundar em um nicho o qual há grande relevância para a sociedade, porém, moderadamente discutido e explorado no segmento de coberturas fabris onde, boa parte dos estudos, direcionam-se para estruturas de, por exemplo, pontes e construções de grande escala urbana.

Este trabalho evidencia a influência dos fatores de voo na qualidade de imagens obtidas dinamicamente. Além disso o quanto os pixels são fundamentais para o aprendizado da IA. Uma relação significativa para redução do volume de imagens em treinamentos, validações e testes, o que aumenta a produtividade e reduz custos com processamentos em grande escala.

O estudo, também, busca fortalecer a importância da interdisciplinaridade, o qual este trabalho, aplica tal conceito através da conexão de diversas ciências com foco em obter uma nova contribuição tecnológica.

É mostrado como a associação de redes neurais pode superar o desafio de localizar características complexas de forma mais rápida e eficiente.

1.5.2. Contribuições para a indústria

O presente trabalho tem como principal contribuição, para a indústria, fornecer uma alternativa inteligente e rápida para a inspeção preventiva de estruturas que compõem a cobertura de barracões fabris. Esta, por sua vez, irá prover uma maior continuidade nas linhas de produção, uma vez que a ocorrência de infiltrações irá tender a zero. Bem como reduzir os custos de reparo onde, com o apoio de diagnósticos precoces, os valores da manutenção corretiva tendem a abaixar devido à uma menor gravidade e estado de degradação. Adicionalmente a solução contribui na sustentabilidade da indústria em termos de saúde e segurança do trabalho, pois sua aplicação proporciona redução significativa de trabalhos em altura para tarefas de inspeção de falhas em locais de difícil acesso.

2. FUNDAMENTAÇÃO TEÓRICA

Nesta seção, será abordado o estado da arte da literatura que embasa este trabalho.

No tópico 2.1, será tratada a evolução da indústria como um todo, desde os tempos primórdios das primeiras grandes revoluções até tendências da atualidade.

Em 2.2, será exposto o estado da arte das redes neurais convolucionais, desde sua idealização até as principais aplicações desta arquitetura.

Será abordado, em 2.3, a evolução e desenvolvimento das aplicações de veículos aéreos não-tripulados para diversos segmentos, incluindo para sensoriamento.

Para tal, foram definidos três grandes eixos de pesquisa, para delimitar o escopo de busca durante uma primeira etapa, de levantamento de uma base de artigos científicos. Levando em consideração que este projeto visa o desenvolvimento de uma inteligência artificial para a detecção e classificação de graus de oxidação em estruturas metálicas de telhados, com o auxílio de imagens geradas através de uma câmera embarcada em drone objetivando uma manutenção preditiva, escolheu-se como eixos delimitantes “Manutenção”, “Drone” e “Rede neural”.

Com estes três segmentos definidos, estabeleceu-se três palavras-chave para cada eixo, afim de obter um extenso leque de trabalhos científicos. Para o eixo de manutenção, foi selecionado “Manutenção preditiva”, “Telhado” e “Infraestrutura”. Para “Drone”, foram escolhidos os termos “Inspeção com drone”, “Drone orientado” e “Análise com drone”. Por fim, para “Inteligência artificial”, foram definidas como palavras-chave “Rede neural convolucional”, “*Deep-learning*” e “Processamento de imagem”.

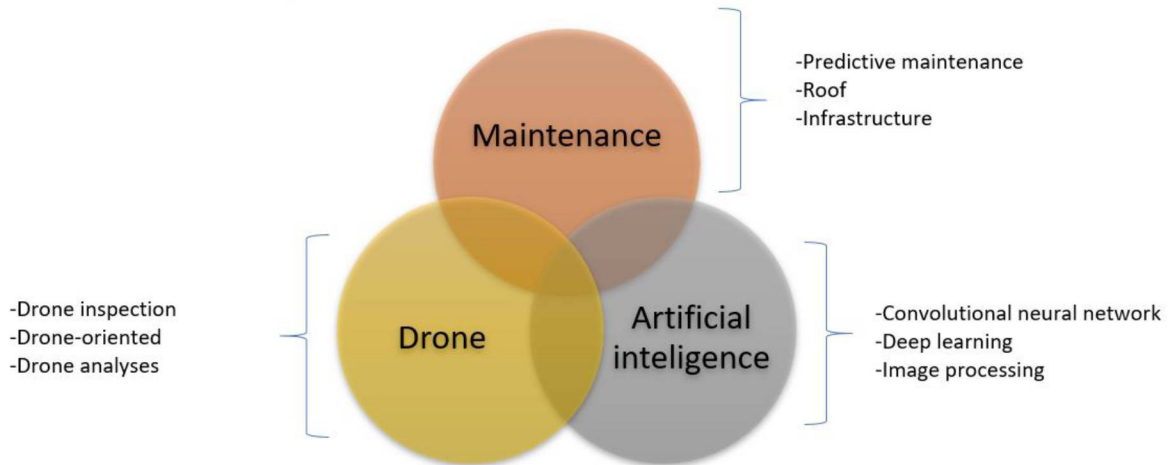


Figura 3: Eixos de pesquisa e seus respectivas palavras chave.

Em seguida, estas 9 palavras foram permutadas e foram utilizadas para realizar buscas em duas bases de pesquisa, Google Scholar e Crossref. Um total de 4.007 artigos iniciais foi obtido. Após esta triagem inicial, foram filtrados apenas os trabalhos publicados em inglês e no período dos últimos 10 anos (2011-2021), totalizando 3.154 trabalhos. Por fim, foram separados os títulos que não haviam entradas repetidas entre as permutações, resultando em 1.889 publicações únicas.

Esses 1.889 trabalhos foram ordenados de maneira decrescente por volumes de citação e foi realizada uma análise, para selecionar os artigos com o título alinhado com os eixos do trabalho, resultando em 130 publicações finais.

2.1. INDÚSTRIA 4.0

A manufatura é uma parte essencial da economia mundial. Sua tendência é se tornar cada vez mais “inteligente”, sendo capaz de providenciar uma rápida adaptação de seus aspectos físicos e intelectuais de infraestrutura para explorar mudanças em tecnologia, no passo que o processo de manufatura se torna cada vez mais ágil, mais responsiva a mudanças globais de mercado e mais alinhada as necessidades específicas dos clientes.

A produção industrial sofre mudanças desde sua ascensão. Estas mudanças, muitas vezes, tem um impacto tão grande que acaba aplicando-se o termo de revolução para descrever tal evento. Logo, o termo muito utilizado nos dias atuais, a “Indústria 4.0”, remete-se a existência de outras três grandes revoluções anteriores.

O termo de primeira revolução industrial remete a mudança do trabalho manual para a produção sendo realizada com maquinários (Wang, Dai, & Guo, 2016). Esta primeira grande mudança teve início entre os anos de 1760 e 1840, na Inglaterra, onde impactou diretamente em inúmeros processos, principalmente na indústria têxtil e de algodão (Schwab, 2017). O grande ápice desta revolução ocorreu em 1782 com o surgimento das máquinas à vapor, inventadas por James Watt, onde, desta data em diante, era possível ter energia disponível em qualquer local, e o foco deixou de ser o trabalho manual, dando início a produção mecânica (Spear, 2008). Os equipamentos emergentes desta primeira geração tinham como características principais sua simplicidade e superdimensionamento, em um ambiente industrial pouco mecanizado, com uma sistemática de manutenção voltada à ações reativas para a correção de falhas (Kardec & Nascif, 2012).

A segunda revolução industrial, iniciada no final do século XIX, foi caracterizada pelos princípios de racionalização expostos por F. W. Taylor (Vijai, Somayaji, Swamy, & Aital, 2017). Estes, eram baseados na divisão de trabalho, padronização, manufatura de precisão e construção de linhas de trabalho, onde, com o advento da eletricidade e das linhas de montagem, foi possível o início da produção em massa (Schwab, 2017). Também neste período, Henry Ford aplicou a primeira esteira transportadora para a produção do modelo-T e alcançou status de pioneiro na manufatura automobilística no início do século XX (Laird & Sherratt, 2010; Wang et al., 2016). Nesta janela temporal, observou-se o aumento da mecanização e complexidade dos processos, gerando, desta forma, a necessidade de uma maior disponibilidade, confiabilidade e produtividade do maquinário (Kardec & Nascif, 2012).

Já a terceira revolução industrial, datada por volta de 1960, teve seus alicerces no desenvolvimento de computadores e na tecnologia da informação, devido a ascensão dos dispositivos semicondutores, da computação em mainframe, da computação pessoal e da internet (Schwab, 2017). Esta quebra de paradigma teve como resultado maquinários controlados, como as máquinas de controle numérico e robôs industriais, nos quais podem ser modificados de acordo com as necessidades das instalações para terem um alto desempenho em uma automatização flexível, com um forte apelo em alta produtividade e flexibilidade (Kardec & Nascif, 2012). Com o crescimento da mecanização e automação dos processos fabris, os parâmetros de confiabilidade e disponibilidade tornaram-se pontos-chaves para manufatura, criando,

desta maneira, a necessidade de estabelecer padrões de qualidade de desempenho dos sistemas, com o foco em manutenções de caráter proativo (Kardec & Nascif, 2012).

Atualmente, a indústria mundial encontra-se à beira de uma quarta revolução. Tecnologias de informação e comunicação (ICT) estão em uma crescente e tendo inúmeros impactos em diversas ramificações da sociedade. O mundo real e o virtual estão crescendo lado a lado em largos passos.

Introduzido inicialmente em uma feira em Hannover, Alemanha, no ano de 2011, o termo “Indústria 4.0” foi utilizado para representar uma nova tendência na indústria, principalmente voltada às fábricas inteligentes, com uma visão de fortalecer a competitividade da indústria alemã (Hermann, Pentek, & Otto, 2015; Schwab, 2017). De maneira análoga, outros países também vem estabelecendo iniciativas para modelar o futuro da manufatura em seus territórios (Wang et al., 2016). Nos Estados Unidos, por exemplo, há a iniciativa conhecida por Smart Manufacturing Leadership Coalition (SMLC) (Davis et al., 2015). Já a China, adota uma estratégia de manufatura inteligente para aumentar sua capacidade industrial através de uma manufatura com bases em inovações, otimizando a estrutura da indústria chinesa, dando ênfase na qualidade sobre quantidade, treinamento e capacitação de talentos, também conhecido como “Made-in-China 2025 (Li, 2018).

A Indústria 4.0 é um termo que engloba tecnologias e conceitos de valor da cadeia de organização. A manufatura estará equipada com sensores, atuadores e sistemas autônomos capazes de monitorar o mundo real e transcrevê-lo para o virtual, possibilitando a tomada de decisões de maneira descentralizada através dos Cyber-Physical Systems (CPS). Através da Internet das Coisas (IoT), o CPS comunica e coopera com outros sistemas e indivíduos em tempo real enquanto através da Internet de Serviços (IoS), colaboradores e parceiros comerciais interagem através de serviços ofertados para todos os membros da cadeia de valor.

Algumas mudanças conceituais estão atreladas a I4.0. Uma delas é a individualização da demanda, que vem se tornando cada vez mais evidente, em um mercado onde quem dita o ritmo não é mais o vendedor e sim, o comprador (Lasi, Fettke, Kemper, Feld, & Hoffmann, 2014). Outra mudança é com relação a flexibilização no desenvolvimento de produtos, bem como períodos menores de desenvolvimento e inovação destes. Devido a esta flexibilização das especificações de produtos, um processo de tomada de decisões mais eficiente é necessário, sendo essencial a

descentralização das informações dos sistemas para deliberações mais rápidas, reduzindo, desta maneira, a hierarquia organizacional.

Para dar suporte a estes sistemas altamente automatizados, é fundamental a existência de planos de manutenção eficientes para fornecer processos com alta confiabilidade, disponibilidade, produtividade e continuidade.

A manutenção é a combinação de ações técnicas, administrativas e gerenciais durante o ciclo de vida de um equipamento, com o intuito de manter ou reestabelecer o estado no qual o mesmo possa desempenhar sua função (Viana, 2002). Existem, de maneira geral, duas abordagens distintas para a manutenção de um patrimônio baseadas no período de ação, a reativa e a proativa.

O plano de manutenção reativo ocorre apenas quando há o surgimento de uma falha no equipamento, visando recoloca-lo em condições de executar, novamente, uma função requerida (Kardec & Nascif, 2012; Susto, Schirru, Pampuri, McLoone, & Beghi, 2015). Esta é a abordagem mais simples dentre os planos de manutenção, com uma baixa eficiência associada devido aos custos de intervenção e tempos de parada da linha de produção não planejados (Kardec & Nascif, 2012).

Por outro lado, os planos de manutenção de caráter proativo buscam, através de intervalos predeterminados ou de acordo com critérios prescritos, reduzir a probabilidade de falhas aleatórias ou a degradação do equipamento (Kardec & Nascif, 2012).

Estes planos podem ser ramificados em duas outras categorias, a manutenção preditiva e a preventiva (Susto et al., 2015). Enquanto a manutenção preventiva é executada de acordo com um calendário planejado baseado em tempo ou iterações de processos, a manutenção preditiva é realizada baseada em parâmetros do equipamento (Krishnamurthy et al., 2005).

Em uma abordagem preventiva, as falhas dos equipamentos são usualmente evitadas porém, sua natureza traz muitos custos provenientes de ações corretivas, muitas vezes, desnecessárias, resultando em um uso ineficiente dos recursos e aumento dos custos de operação (Kardec & Nascif, 2012). Já para planos preditivos, estes permitem a garantia da qualidade do serviço desejado com base na aplicação sistemática de técnicas de análise, utilizando-se meios de supervisão realizado através de ferramentas de predição, métodos de inferência estatísticos e metodologias de engenharia afim de reduzir ao mínimo a manutenção preventiva e diminuir a manutenção de caráter corretivo (Susto et al., 2015).

No decorrer dos últimos anos, foram publicados na literatura diversos trabalhos voltados ao diagnóstico de falhas em equipamentos e sistemas industriais.

O diagnóstico e previsão de falhas em máquinas elétricas tem sido um tópico de extenso estudo, principalmente devido a sua grande influência na continuidade de inúmeros processos industriais (Awadallah & Morcos, 2003). O diagnóstico correto e a detecção de falhas destes equipamentos é crucial para curtas janelas de manutenção corretiva não planejadas nestes sistemas, resultando em uma redução no tempo de parada de linhas de produção e redução nas perdas financeiras. Falhas nos enrolamentos do estator (Bellini, Filippetti, Franceschini, & Tassoni, 2000; Cash, Habetler, & Kliman, 1998; Penman, Sedding, Lloyd, & Fink, 1994), barra do rotor (Bellini et al., 2000; Schoen & Habetler, 1995; Thomson & Fenger, n.d.), anéis (Awadallah & Morcos, 2003), rolamentos (Schoen & Habetler, 1995) e a excentricidade do rotor (S. Fruchtenicht; E. Pittius; H.O. Seinsch, 1989) foram estudadas ao longo dos anos.

Diferentes abordagens para o diagnóstico também foram exploradas. Redes neurais artificiais foram utilizadas extensivamente para o reconhecimento, estimação de padrões; treinamento baseado tanto no domínio da frequência como no tempo; diagnóstico não-supervisionado online e em tempo real; filtração de transientes, distúrbios e ruídos; previsão de faltas devido a anomalias de operação (Byung-Geun Hyun & Kwanghee Nam, 1995; Mo-Yuen Chow & Goode, 1993; Murray & Penman, 1997; Nejjari & Benbouzid, 2000; Rokouzzaman & Rahman, 1995). Sistemas Fuzzy também foram aplicados para lidar com estas tratativas onde suas aplicações compreendiam o levantamento de índices de desempenho; previsão de anomalias de operação e localização de faltas; utilizam de expertises humanas através de regras if-then; classificação de falhas e prognósticos (Altug, Mo-Yuen Chen, & Trussell, 1999; Consoli, Gennaro, Raciti, & Testa, 1998; Lasurt, Stronach, & Penman, 2000).

De maneira análoga, sistemas inteligentes também foram aplicados para tratar de outros problemas em um variado leque da indústria. Para sistemas a diesel, redes neurais artificiais foram aplicadas para o levantamento do plano de manutenção baseado nas condições do sistema, detecção de falhas e diagnósticos destes equipamentos através de um arranjo de termoacopladores, sensores de pressão, medidor de torque e medidor de fluxo de óleo (Basurko & Uriondo, 2015). Bem como sistemas Fuzzy foram aplicados para a previsão de faltas baseado nos parâmetros operacionais e ambientais do sistema (Diez-Olivan, Pagan, Sanz, & Sierra, 2017).

Para ferrovias, um modelo de otimização baseado em similaridades (SBM) foi elaborado para a estimação da desmagnetização de máquinas síncronas de ímã permanente e avaliação do desgaste do freio através de parâmetros de aceleração, corrente e força (Niu & Jiang, 2017). Outra abordagem explorada foi o uso de um sistema Fuzzy de inferências semânticas para a manutenção inteligente de EMUs (electric multiple unit) através de sinais de corrente e vibração (Niu & Li, 2017).

Na indústria aeroespacial, um ponto crítico é a falha dos rolamentos do sistema. Extensos estudos foram realizados para estimar a vida útil remanescente destes elementos através de redes neurais recorrentes baseadas nos sinais de vibração (Guo, Li, Jia, Lei, & Lin, 2017). Também para estimar a vida útil, porém de turbofans, foi aplicado um método de estimação Kaplan Meier com base no monitoramento de condições e indicadores baseados no número de falhas e suas covariâncias correspondentes (Ragab, Ouali, Yacout, & Osman, 2016). Há também o uso de uma abordagem Bayesiana para a predição da vida útil destes elementos (Mosallam, Medjaher, & Zerhouni, 2016).

2.2. VEÍCULOS AÉREOS NÃO TRIPULADOS PARA SENSORIAMENTO

Veículos aéreos não-tripulados (VANTs) vêm vivenciando uma grande expansão em seu uso para sensoriamento nos últimos anos (Pajares, 2015). Estes dispositivos trazem uma maior flexibilidade na maneira de aquisição de dados e uso de sensores, quando comparado com outras metodologias mais tradicionais (Yao, Qin & Chen, 2019), podendo ter aplicações tanto em meios ambientais, como urbanos (Nex & Remondino, 2014; Bhardwaj, Sam, Martín-Torres & Kumar, 2016). Esses, podem ser divididos em categorias de acordo com o grau de atributos que possuem, como peso, altura de vô, capacidade de carga, câmeras (Korchenko & Illyash, 2013).

Com o avanço e desenvolvimento de novas estruturas de VANTs, bem como o refinamento de diversos sensores disponíveis no mercado, tem aumentado o número de aplicações que objetivam a detecção e acompanhamento em tempo real de objetos, com o suporte de técnicas de visão computacional e aprendizado de máquinas (Moranduzzo & Melgani, 2014; Rodríguez-Canosa et al, 2012).

O uso de câmeras tipo RGB, integradas nestes veículos, aliado com outros sensores como um sistema de posicionamento global (GPS) e unidade de medição inercial (IMU) compõem a base de boa parte dos VANTs disponíveis (Yao, Qin & Chen,

2019; Eisenbeiss, 2009). Alguns parâmetros que influenciam diretamente a escolha destes dispositivos são lente da câmera, resolução, tamanho de pixel e proporção sinal-ruído (Nex & Remondino, 2014; Colomina & Molina, 2014).

Câmeras multiespectrais também são muito utilizadas na linha de sensoriamento remoto com VANTs (Yao, Qin & Chen, 2019). A possibilidade de levantar informações do espectro de absorção de objetos permite aplicações destes dispositivos em, por exemplo, análises de vegetações (Du et. Al, 2016; Gitelson, 2002).

Outros dispositivos que podem ser aplicados em VANTs são as câmeras hiperespectrais, capazes de registrar imagens com centenas de bandas de larguras entre 5-10 nm (Suomalainen et. Al, 2014). Embora esta possua algumas limitações (Rufino & Moccia, 2005), há grande aplicabilidade em diversos segmentos como mapeamento de vegetação (Adam, Mutanga & Rugege, 2010) e classificação de áreas urbanas (Benediktsson, Palmason & Sveinsson, 2005), como ilustra a Figura 4.

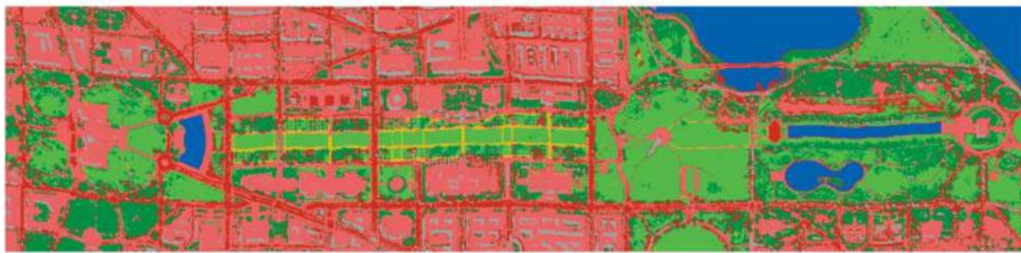


Figura 4: Representação de um mapa de classificação obtido por uma rede neural da cidade de Washington, DC. Fonte: Adaptado de (Benediktsson, Palmason & Sveinsson, 2005).

Os sensores térmicos infravermelhos são amplamente utilizados para medições de temperatura de superfície e emissão térmica (Prakash, 2000). Este, pode ser aplicado em seguimentos como na detecção e geolocalização de corpos humanos para resgates (Rudol & Doherty, 2008), monitoramento de incêndios (Ambrosia et. al, 2003) e manutenção preventiva de vazamentos em tubulações de usinas termoelétricas (Ibarguren et. Al, 2013).



Figura 5: Imagens dos corpos classificados e suas respectivas imagens térmicas. Fonte: (Rudol & Doherty, 2008.)

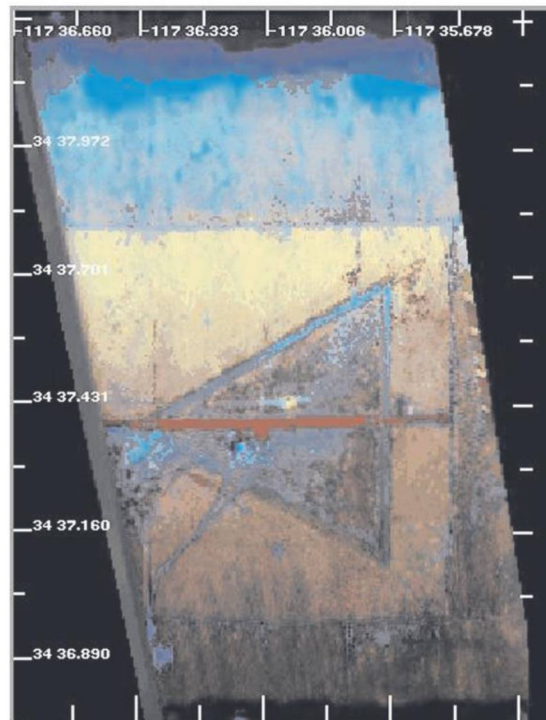


Figura 6: Dados aéreos coletados pelo modelo GA-ASI ALTUS® II em uma altitude de 945 metros, aproximadamente, sobre o incêndio controlado em El Miragem no dia 6 de setembro de 2001 na Califórnia. Os pequenos princípios de incêndio podem ser vistos pelos círculos na imagem. Fonte: (Ambrosia *et al*, 2003).

Outra ferramenta utilizada em VANTs são os sensores de detecção de luz e alcance (LiDAR), conhecidos na literatura como um dos meios mais precisos para aquisição de dados geométricos, utilizado na silvicultura, modelagem de informações de estruturas e aerotransporte (Yao, Qin & Chen, 2019). O seu princípio de funcionamento gira em torno da medição do tempo de vôo direto de sinais de lasers pulsado, com comprimento de onda na região do infravermelho. Algumas aplicações

destes sensores, aliados à VANTs, são a modelagem de enchentes (Li et. Al, 2021), extração de parâmetros de vegetação (Zhang et al, 2021) e controle de movimentos de VANTs, para evitar colisões (Park & Cho, 2020).

2.3. REDES NEURAS CONVOLUCIONAIS

Redes neurais convolucionais foram elaboradas baseado no modelo do cortex visual dos animais ((Hubel & Wiesel, 1968) onde, inicialmente, foi amplamente utilizada para reconhecimento de objetos e, hoje, tende a ser aplicada em outras atividades como estimação de movimentos humanos (Toshev & Szgedy, 2014), reconhecimento e detecção de textos escritos (Jaderber, Vedaldi & Zisserman, 2014), reconhecimento de ações, classificação de cenários (Farabet et al., 2013), entre outras aplicações (Nithin et al., 2015).

O Neocognitron (Fukushima, 1980), uma rede neural de modelo auto-organizador, utilizado para o reconhecimento de padrões, é considerado o predecessor das ConvNets. A LeNet é o trabalho pioneiro em redes neurais convolucionais, proposta por LeCun em 1990 (LeCun et al., 1990). Esta, foi especificamente desenvolvida para a classificação de dígitos escritos à mão e obteve sucesso no reconhecimento de padrões visuais diretamente da imagem de entrada, sem passar por qualquer etapa de pré-processamento (Aloysius & Geetha, 2017).

Krizhevsky e colaboradores propuseram um modelo de rede neural convolucional, em 2012, que obteve sucesso em reduzir a taxa de erro na competição ILSVRC (Krizhevsky et al., 2012). Ao longo dos últimos anos, esses trabalhos serviram de base para o desenvolvimento do campo de visão computacional, e foram utilizados por muito para o desenvolvimento das mais divers as arquiteturas de CNNs presentes hoje na literatura (Aloysius & Geetha, 2017).

Outras redes se fizeram populares na literatura ao longo da atualidade. Zeiler e Fergus introduziram um a DeConvNet, capaz de vizualizar a atividade dentre de uma CNN utilizando uma rede multicamadas deconvoluciona. A GoogleNet, baseada na arquitetura proposta por Szegedy (Szegedy et al., 2014), elaborada pelo grupo Google, introduziu uma nova arquitetura, nomeada Inception, na qual proporciona melhor utilização dos recursos computacionais na rede. Simonyan e Zisserman propuseram, através de uma análise no fator de profundidade de uma ConvNet, uma arquitetura mais precisa, nomeada VGGNet (Simonyan & Zisserman, 2014). Kaiming e

colaboradores, apresentaram um quadro de aprendizado residual, onde as camadas aprendem funções residuais, respeitando as entradas recebidas, sendo capaz, desta forma, de proporcionar uma melhor etapa de aprendizado para as redes, uma vez que redes residuais são mais fáceis de otimizar e ganham mais precisão.

2.3.1. Visão computacional

Rede neural convolucional é uma tarefa específica dentro do escopo da Visão computacional.

A visão computacional é uma subárea dentro da I.A que executa o processamento de vídeos e imagens. Essa tecnologia que busca colocar dentro do computador a capacidade de visualização e identificação de características assim como os seres humanos.

Neste processo a Rede Neural Convolucional é uma das técnicas da visão computacional que permite o reconhecimento de imagens e com benefícios em uma série de aplicações, tais como: Detecção de pessoas ou objetos, reconhecimento facial, comparação de características e até medições a distância.

Além disso a RNC é aplicável com grande performance na aprendizagem profunda (*Deep Learning*) a fim de performar a capacidade de aprendizado de uma IA e de forma isolada ou com a associação de outras técnicas promover maiores benefícios como diagnósticos, alertas, medições e carros autônomos.

2.3.2. Imagens

Uma imagem digital representa um espaço físico real. Boa parte destas imagens são armazenadas em formato de três canais de cores (Wan & Xie, 2016), conhecida usualmente por RGB (do inglês, representa uma abreviação de Vermelho, Verde e Azul). Esta representação usual diz respeito à uma matriz tridimensional onde os eixos x e y representam a área total desta enquanto o eixo z refere-se à um dos três canais, Figura 7. Por padrão, o valor de cada cor de cada pixel é dado em um número inteiro de 8 bits por canal (Andrade, 2017), onde a variação destes numerais cria a sensação de profundidade das imagens.

Outra maneira de representação de uma imagem digital é em uma escala de tons de cinza (do inglês, *grayscale*), no qual os três canais de uma representação em

RGB são unificados em uma única matriz bidimensional, na qual varia-se entre 0 (preto) e 255 (branco) (Andrade, 2017).

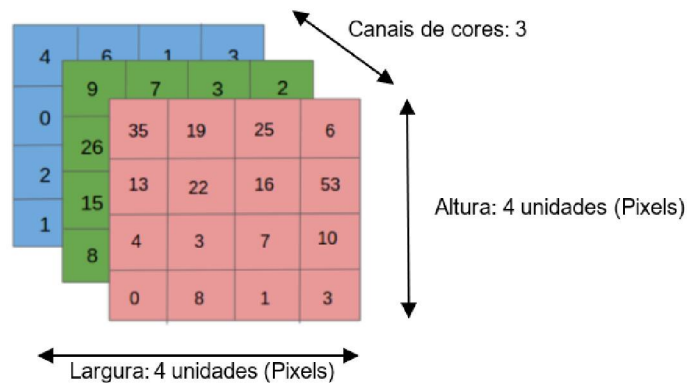


Figura 7: Representação RGB. Fonte: (O Autor).

2.3.3. Neurônio

2.3.3.1. *Biológico*

O neurônio tem uma série de regras para seu correto funcionamento. Os estímulos ou sinais elétricos, transmitidos pelo cérebro provocam sensações de dor, prazer e memória no comportamento humano. Este processo tem como principal elemento os neurônios biológicos que no cérebro humano chegam ao número aproximado de 100 bilhões com milhares de conexões entre si.

O neurônio tem como principais divisões o seu núcleo, os Dentritos, os Axions e suas terminações sinápticas. Seu princípio de funcionamento inicia-se pelo recebimento de sinapses vindas de outros neurônios. Os Dentritos executam a função de carregamento dos estímulos até o núcleo. Estes sinais elétricos sofrem reações no núcleo celular e posteriormente são transmitidos através de canais chamados de Axions e pelas terminações sinápticas até outros neurônios.

Em um ciclo de entrada, processamento de dados e saída de informações se repetem em conexão de neurônios até atingir o resultado proposto pela rede neural conforme indicado na *Figura 8*.

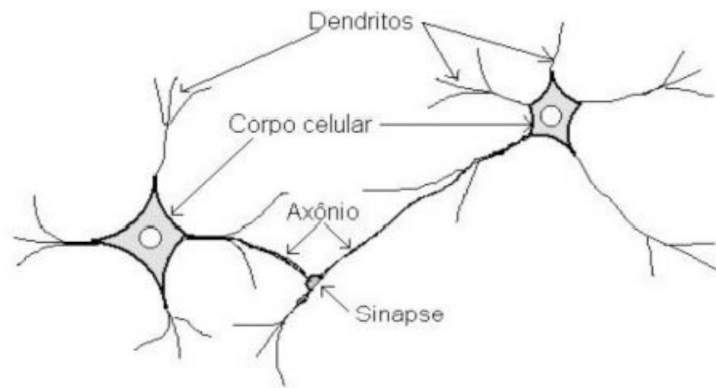


Figura 8: Representação de neurônios biológicos conectados. Fonte: (Neto, 2014).

2.3.3.2. Artificial

Seguindo o mesmo princípio do neurônio biológico, a rede neural artificial contém entradas, processamentos e saídas.

As sinapses representadas pelas variáveis “ X_1, X_2, \dots, X_n ”, vindas de outros neurônios possuem pesos atribuídos ao seu valor W_1, W_2, \dots, W_n , que representam o conhecimento e, então, geram uma saída composta por um potencial de ativação, u , que soma o produto das sinapses pelos seus respectivos pesos e conseqüentemente determina a função de ativação $f(u)$. Essa função pode se apresentar em diversos modelos e tem como objetivo gerar o sinal do neurônio a ser repassado para o próximo neurônio artificial.

$$f(u) = \sum_{i=1}^n X_i * W_i \quad (1)$$

Por outro lado, as redes não recorrentes, não possuem laços de realimentação e caracterizam-se por estarem dispostas em camadas (Minussi, 2003). Neste estudo o modelo de rede neural adotado será *Feedforward network* e *ConvNet*, uma vez que é utilizado em redes neurais convolucionais para a identificação e classificação de imagens.

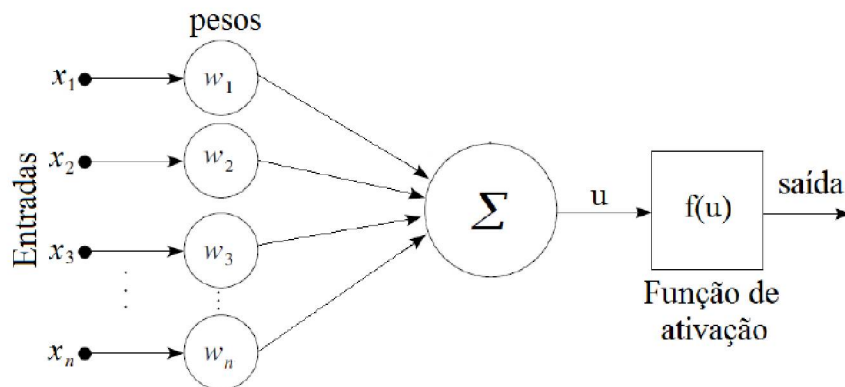


Figura 9: Representação matemática de um neurônio. Fonte: (Mcculloch & Pitts, 1943).

2.3.4. Inteligência artificial, *Machine learning* e *Deep Learning*

Inteligência artificial é a teoria e o desenvolvimento de sistemas para desempenhar atividades que exigem a inteligência humana para serem desempenhadas, como a necessidade de reconhecimento de fala, detalhes visuais e tomada de decisão.

Dentro deste segmento, o aprendizado de máquinas está inserido, onde inclui sistemas que buscam o aprendizado através de dados coletados, identificação de padrões e tomadas de decisões sem a intervenção humana.

E, dentro desta parcela da inteligência artificial, ramifica-se o *Deep Learning*, um dos tipos de aprendizado de máquinas baseado no aprendizado em representações de grupos de dados, onde cada nível deste segmento transforma os dados de entrada em modelos com abstrações de alto nível.

2.3.5. Funcionamento de RNC

Inspirada no córtex visual humano, as redes neurais convolucionais possuem em sua topologia um segmento com neurônios sensíveis às informações do campo visual humano (Andrade, 2017).

A primeira camada destas redes usualmente é uma camada de convolução. Nestas, usam-se *kernel's*, de dimensões inferiores às imagens em análise, para realizar filtros ao longo de toda a sua matriz, objetivando extrair características fundamentais para identificação ou classificação dos objetos (Deshpande, 2016).

Outras camadas subsequentes de uma CNN podem conter mais convoluções ao longo de seu processamento, porém, esta não é apenas composta de camadas

convolucionais. Após a aplicação dos filtros na matriz, é necessária uma camada de *flatten*, para analisar e gerar a classificação de objetos. Esta busca agrupar e transformar todas as matrizes geradas pelas camadas de convoluções em um único vetor, para servir de entrada para a rede analisar (Andrade, 2017).

2.3.6. Arquitetura

Uma Rede Neural Convolucional (ConvNet / Convolutional Neural Network / CNN) por definição é um algoritmo de aprendizado profundo que pode captar uma imagem de entrada, atribuir importância (pesos e vieses que podem ser aprendidos) a várias características da imagem e ser capaz de diferenciá-los (LeCun, 1988).

A arquitetura de uma ConvNet é análoga àquela do padrão de conectividade de neurônios no cérebro humano e foi inspirada na organização do córtex visual. Os neurônios individuais respondem a estímulos apenas em uma região restrita do campo visual conhecida como Campo Receptivo. Uma coleção desses campos se sobrepõe para cobrir toda a área visual.

Uma ConvNet é capaz de capturar com sucesso as dependências espaciais e temporais em uma imagem através da aplicação de filtros relevantes. A arquitetura executa um melhor ajuste ao conjunto de dados da imagem devido à redução no número de parâmetros envolvidos e à capacidade de reutilização dos pesos. Em outras palavras, a rede pode ser treinada para entender melhor a sofisticação da imagem.

Na Figura 7, temos uma imagem RGB (Red – Green – Blue) separada em seus três planos coloridos – Vermelho, Verde e Azul. Além destes existem vários espaços de cores nos quais existem imagens, como o graysacale, citado na sessão 2.3.2 Imagens.

No processamento de imagens o processo de computação se torna intenso principalmente com imagens de alta dimensões, como, 8K (7680 × 4320). A função da ConvNet, portanto, é reduzir as imagens para uma forma mais fácil de processar, sem perder recursos que são críticos para obter uma boa previsão. Isso é importante quando queremos projetar uma arquitetura que não seja apenas boa em recursos de aprendizado, mas que também seja escalável para conjuntos de dados massivos (Sumit, 2018).

2.3.7. Extração de características

Na aplicação de uma rede neural normal uma imagem é representada como uma matriz de valores de pixels na entrada do processamento e convertida em uma matriz $n \times n$ que por sua vez é convertida em um vetor $Z \times Y$ para ser alimentado para fins de classificação de um objeto simples. Uma técnica muito válida para os casos de imagens binárias extremamente básicas que se deseja localizar.

No entanto o método pode mostrar uma pontuação de precisão média durante a extração de características (Samuel L, 2018) e pode ter pouca ou nenhuma precisão quando se trata de imagens complexas com dependências de pixel por toda parte. Na prática a desvantagem de utilizar redes neurais normais com entrada de pixels de imagens resulta na dificuldade de localização do objeto. Na situação de uma imagem treinada em uma posição qualquer e ao utilizar localização por pixels em uma posição angular diferente, a mesma, pode não ser reconhecida, como mostrado na Figura WW. Adicionalmente a influência das cores das imagens, podem afetar negativamente na aprendizagem de máquina quando a qualidade é comprometida. Condições de tempo e iluminação são fatores relevantes em virtude da diferença entre pixels da situação de treino e situação reconhecimento de objetos e que podem ser sanados salvo houver uma base de dados extremamente grande para várias cores e posições de imagem.

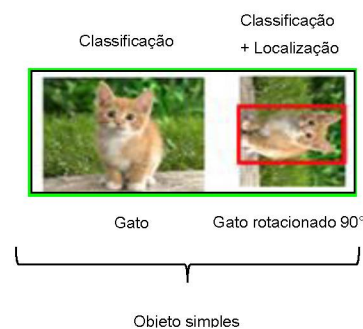


Figura 10: Representação de condições de posição de imagem para reconhecimento. Fonte: O autor (2023)

Para classificar imagens de forma mais objetiva a convolução, neste caso, traz uma grande vantagem para descobrir características únicas das imagens e classificá-la de forma assertiva, através da entrada de dados na rede neural (anotações) e através de coeficientes e pesos realiza cálculos para determinar os padrões da imagem (LeCun, 1988).

No projeto a abordagem do tipo de IA deve detectar de forma assertiva a amostragem que proposta que é a oxidação. O desafio do modelo, além da identificação e classificação, foi a detecção da característica ideal em meio a diversos objetos utilizando a técnica de tratativa de imagens denominada segmentação semântica.

O objetivo dessa técnica é rotular cada pixel de uma imagem com uma classe correspondente do que está sendo representado (LeCun, 1988).

A segmentação semântica trabalha com escala de cinza e rotula pixel a pixel para definir a probabilidade deste pertencer àquela classe correspondente e que se deseja obter, no caso do projeto proposto, a oxidação. Desta forma o algoritmo fornece o percentual de zero a 100% daquele pixel ser de uma oxidação ou não, e baseado no contexto que ele está inserido Figura 11.

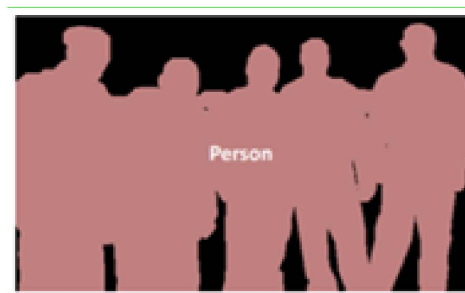


Figura 11: Representação de segmentação semântica. Fonte: O autor (2023).

2.3.8. Métricas de avaliação

Para avaliação qualitativa das redes neurais convolucionais, é necessário a utilização de algumas métricas, para ter-se uma base comparativa de desempenho com outras frentes de pesquisas.

Um dos parâmetros de avaliação é a precisão, na qual busca medir a acurácia das predições da rede neural convolucional que estão corretas. Para tal, esta faz uma proporção entre o total de verdadeiros-positivos encontrados pela rede e a somatória de verdadeiros-positivos e falsos-positivos.

Outra medida usual é a taxa de *recall*, na qual mensura o quão bem a rede consegue encontrar todas as predições positiva. Desta forma, esta medida é dada pela proporção entre os verdadeiros-positivos e a somatória entre verdadeiros-positivos e falsos-negativos.

A intersecção sobre uma união (do inglês, *intersection over union* – IoU) busca medir uma sobreposição entre dois limites. Ou seja, utiliza-se este parâmetro para verificar o quão próximo a rede conseguiu delimitar o objeto a ser encontrado comparado ao *ground truth* (a real área do objeto delimitada). Logo, a IoU é dada pela proporção entre a área de intersecção, entre a predição e o *ground truth*, e a união entre ambas as delimitações.

3. MATERIAIS E MÉTODOS

A metodologia utiliza os seguintes materiais e métodos:

- Método MDCB como ferramenta estatística para análise de relevância de falhas e definição da oxidação como característica de estudo.
- O levantamento das imagens a serem processadas é feito com o auxílio de dois VANTs. O primeiro, VANT A, é um drone Matrice 200, DJI, equipado com uma câmera Zenmuse x4S (20MP, sensor de 1-inch, lente 24 mm, 84° FOV). Já o segundo, VANT B, é um drone Mavic 2 Enterprise Advanced, DJI, equipado com uma câmera embarcada de 48 MP, sensor CMOS ½-inch, lente de 24 mm 84° FOV.
- Algoritmos de *Cany* e *Laplaciana* de imagem foram aplicados para a análise de nitidez e características de imagens obtidas em diferentes cenários.
- O *software* Minitab® e o método experimental de Taguchi foram aplicados na hipótese da influência da qualidade da imagem sobre o aprendizado da rede neural convolucional.
- Ferramentas e softwares estatísticos aplicados para mensurar resultados.
- Para a geração da base de dados foram considerados como objeto de estudo de caso, 78.000 m² de telhados de área fabril, totalizando 51 prédios com capacidade de cobertura entre 1.000 m² e 12.000 m².
- Trabalho envolve pesquisa de campo para coleta de dados, influência dos equipamentos de captura de imagem (Drones e câmeras) e análise de performance. Duas unidades fabris uma localizada na cidade de Curitiba/PR e a outra na de Campinas/SP foram utilizadas para captura de imagens de telhados. A primeira com foco no treinamento e testes da IA. A segunda como base para validação da eficácia.
- No total foram utilizadas 2767 amostras de imagens de diferentes classes de oxidações para a construção da rede.
- Anotações de características de treinamento da rede foi realizada através da ferramenta *LabelME*.
- O estudo considerou a rede YOLO como potencial para a construção da rede. No entanto o método aplicável se valeu do uso de duas redes

neurais convolucionais do tipo U-NET, sendo uma para localização do telhado e a segunda para localização da oxidação. Ambas através de convolução e “máx-pooling” para a segmentação das características de interesse.

- Atribuído como principal métrica para avaliar o desempenho da rede o coeficiente de pontuação *DSC* (Dice Score Coefficient).

Em síntese, a metodologia envolve procedimentos científicos para validação da influência da qualidade de imagem na performance da rede neural convolucional. Na sequência o estudo experimental adotado, tem como base, desenvolver a conexão de duas redes neurais com a técnica de segmentação para localizar, apenas, os objetos e características de interesse objetivando tornar a IA mais eficiente. O trabalho busca, também, interligar ao resultado da rede neural a um software estatístico capaz de classificar a densidade de características de oxidação e através de coordenadas GPS, informar via e-mail o time de manutenção de instalações.

Os objetos científicos deste trabalho exploram materiais, ferramentas e métodos de engenharia de manutenção, visão computacional, aprendizado de máquina aplicadas a reconhecimento de objetos e características específicas, interpretar dados coletados, identificar inconsistências a serem ajustadas e analisar resultados. As etapas de 3.1 a 3.6 descrevem a metodologia estruturada para a construção da solução aplicável a situação problemática.

3.1. PRÉ-ESTUDO DE TÉCNICAS PARA CONSTRUÇÃO RNC

Nesta pesquisa, com o objetivo de obter maior qualidade de captura de imagens para a base de dados mais compacta afim de otimizar a capacidade de processamento, aspecto de produtividade importante nas atividades de processamento de dados, foram definidos parâmetros de técnicas para serem avaliadas de maneira prévia, antecedendo ao treinamento e confecção da rede neural, Figura 12 etapa 1. Estas análises prévias compreendem valores ótimos de altura, velocidade e luminosidade, e se fazem necessárias para almejar bons resultados e um volume de retrabalho menor durante a etapa de construção da rede neural, Figura 12 etapa 3. Entre estas duas etapas, será confeccionado um algoritmo ortomoisáico, Figura 12

etapa 2, a fim de unificar a análise das estruturas em uma única foto, tornando este processo mais barato, computacionalmente falando.

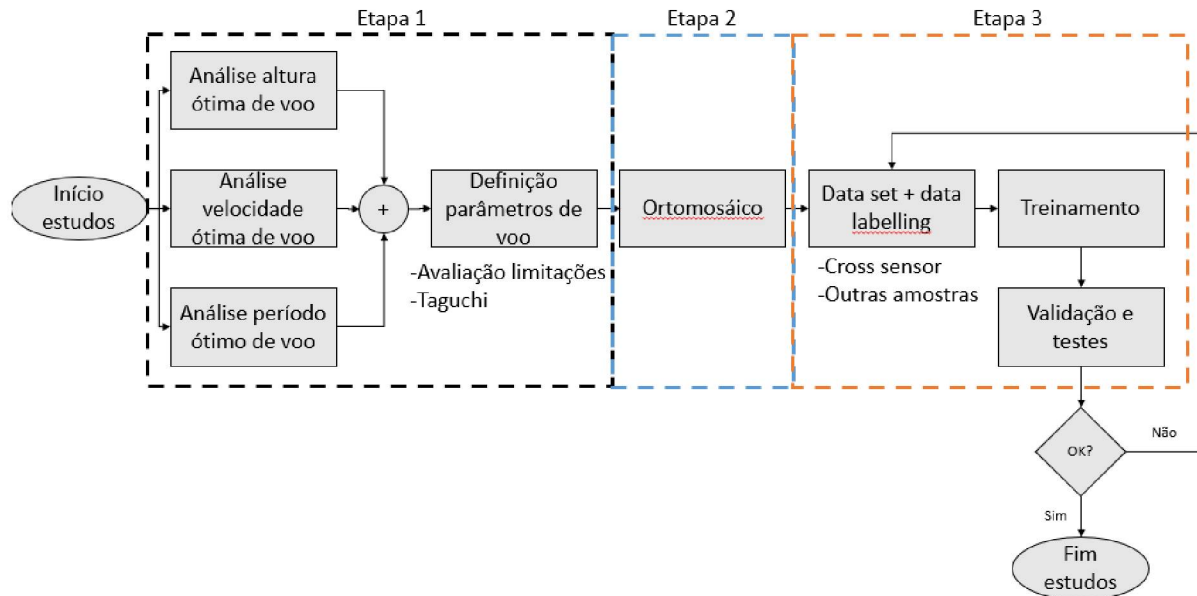


Figura 12: Etapas do projeto. Fonte: o autor (2023).

3.2. PROCESSAMENTO E PREPARO DOS DADOS PARA O TREINAMENTO DA REDE

Para o treinamento da rede neural a ser desenvolvida, as imagens obtidas, em formato “.JPEG”, passam por uma fase de anotações onde, com o auxílio de uma profissional com experiência no tema, é indicado em uma parcela dos dados obtidos exemplos de oxidações nas estruturas.

Para tal, foi utilizado a ferramenta *LabelMe*, um auxiliar de criação de anotações com uma interface gráfica, escrito em Python e que utilizada Qt para a criação de sua *graphical user interface* (GUI). Estas anotações são salvas em arquivos no formato “.json”.

Para exemplificar o processo, a imagem original, Figura 13, é carregada no *LabelMe* e, em seguida, o usuário cria em sua *interface* as classes nas quais é desejado apontar no arquivo. Após a criação destas classes, o usuário cria polígonos sobre cada objeto, Figura 14, classificando-as de acordo com as respectivas classes. Por fim, após todo o apontamento de objetos, é gerado um arquivo contendo as coordenadas de cada caixa criada, bem como a classe na qual o objeto está inserido, Figura 15. Este arquivo, irá alimentar a rede neural durante sua fase de treinamento.

As imagens rotuladas através da ferramenta citada anteriormente possuem um tamanho original de 5.472x3.078 pixels, objetivando um processamento em paralelo

mais rápido e considerando as limitações de entrada da rede neural utilizada nesta primeira seção de testes, é necessário a adequação de tamanho para o *input* da rede, 640x640. Logo, as imagens anotadas foram particionadas em segmentos menores, mantendo uma área de sobreposição entre os passos, Figura 16.

Em seguida, considerando a rota realizada durante a rotina de obtenção das imagens, as imagens que compõem o *dataset* foram distribuídas, previamente, em três grupos para a fase de treinamento da rede neural convolucional: imagens de validação, imagens de treinamento e imagens de *overlap*. A Figura 17 ilustra este processo. As imagens de treinamento correspondem aos dados que apenas serão utilizados para a etapa de treinamento da IA, o grupo de imagens para validação é composto pelos dados que apenas serão utilizados para a etapa de validação da rede e as imagens de *overlap* pertencem ao grupo de informações que serão utilizadas nas duas etapas, tanto de validação como teste.

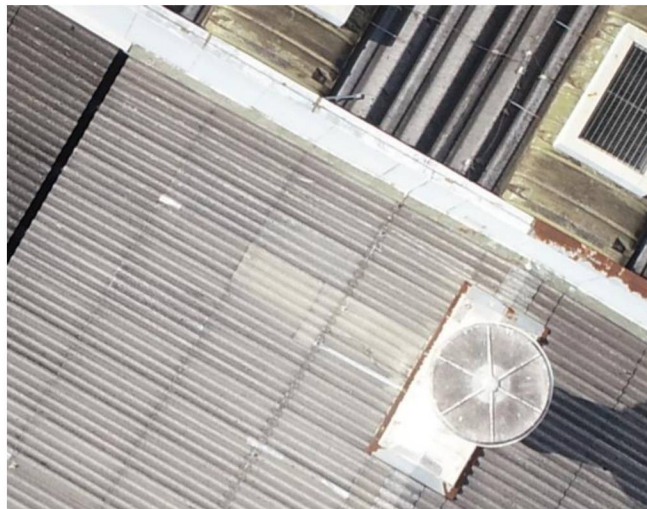


Figura 13: Imagem fonte, utilizada para o processo de anotação. Fonte: o autor (2023).

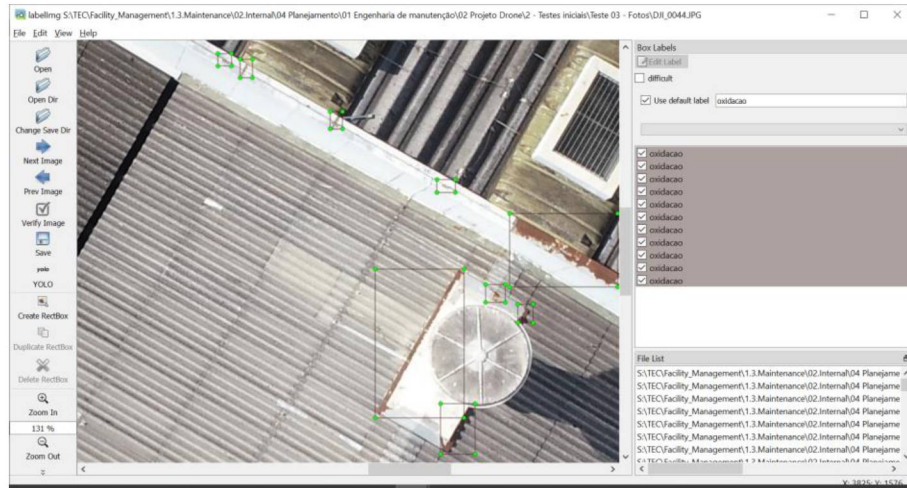


Figura 14: Imagem fonte carregada no *Labelling* e passando pelo processo de apontamento de objetos. Fonte: o autor (2023).

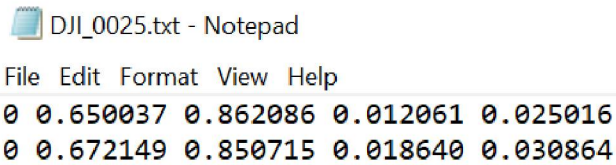


Figura 15: Exemplo de arquivo gerado ao final do processo de anotação de uma imagem. Fonte: o autor (2023)

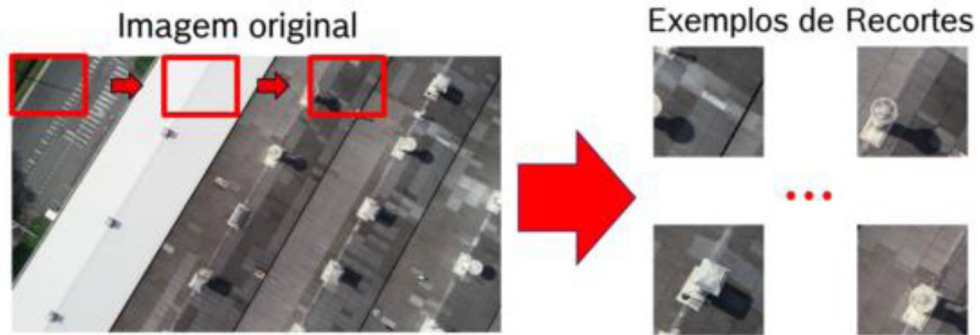


Figura 16: Redução do tamanho do arquivo de entrada da rede neural. Fonte: o autor (2023).

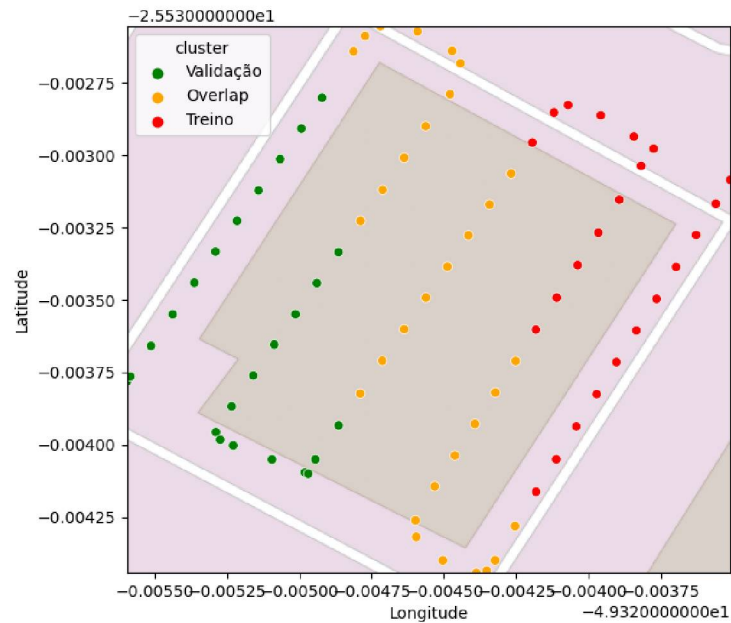


Figura 17: Distribuição do *dataset* para a etapa de treinamento. Fonte: o autor (2023).

3.3. BASE DE DADOS

Para esta pesquisa, dois *datasets* foram criados. Estes, se basearam em duas plantas industriais. A primeira, Planta A, tem seu foco em usinagem de sistemas à diesel, localizada em Curitiba, Paraná, Brasil, e possui um total de 86.691 m² de área construída ao longo de 31 prédios. Planta B, localizada em Campinas, São Paulo, Brasil, é uma planta diversificada, que possui manufatura de ferramentas elétricas, chassis de sistemas de controle, tecnologias prediais e a produção de alguns sistemas à diesel, com um total de 163.066 m² de área construída ao longo de 20 prédios.

Ao todos, foram 51 prédios, com um total de 249.757 m² de área construída, capturados em um ângulo perpendicular do VANT ao solo, em uma velocidade de 25 km/h durante o período da manhã, com níveis de iluminância próximo dos 110.000 lux, mantendo uma diferença de 37 metros entre o pé direito do prédio e o VANT.

O primeiro *dataset*, nomeado “Ctp”, consiste em 202 imagens da Planta A, obtidos pelo VANT A e 89 imagens da Planta A obtidos através do VANT B. Para a segunda base de dados, nomeado “Cap”, um total de 42 imagens da Planta B capturados através do VANT B foi utilizado.

Dentro destes *datasets*, foi gerado um total de 2.767 exemplos positivos contendo oxidações segmentadas. Estes exemplos foram particionados em amostras de 640x640 pixels cada, com sobreposições de 30% entre cada imagem subsequente,

visando um melhor aproveitamento das segmentações levantadas e evitando, desta maneira, a perda de dados para alguns casos específicos. Para os telhados, 305 amostras de exemplos positivos (*i.e.* imagens contendo telhados) foram obtidas, seguindo o mesmo padrão utilizado para as oxidações.

3.4. YOU ONLY LOOK ONCE (YOLO)

Nesta primeira fase de testes, foi utilizada a rede YOLO, proposta por Redmon e colaboradores inicialmente em 2016 (Redmon et al., 2016). Esta rede recebe a imagem de entrada e a passa por 24 camadas convolucionais, 4 camadas de levantamento de máximos locais e duas camadas totalmente conectadas. A redução dos canais é dada por uma convolução de 1×1 seguida de uma operação de 3×3 . A arquitetura apresenta uma função de ativação *Leaky ReLU*, baseada em uma unidade linear retificada (ReLU), porém com uma pequena inclinação para valores negativos, com algumas exceções em poucas camadas, onde utilizada uma ReLU.

Esta é uma rede pré-treinada na base de dados ImageNet-1000 e faz uso da soma de erros quadráticos como função de perda.

Alguns pontos positivos desta rede são sua alta velocidade de processamento, devido à mecânica de detecção baseada em um problema de regressão, análise global da imagem, no qual é levado em consideração os periféricos e o contexto que a imagem está inserida antes de realizar predições, e a capacidade de aprendizado de representações generalizadas de objetos (Redmon et al., 2016). No entanto a baixa capacidade de detecção dos periféricos e o contexto que a imagem está inserida antes de realizar predições é algo que foi levado em consideração para uma abordagem diferente de arquitetura de rede.

3.5. U-NET

A arquitetura de rede usada no projeto, com base na segmentação semântica, é a rede neural convolucional do tipo U-NET a qual foi criada em 2015 com foco na área biomédica para segmentar tecidos cancerígenos ou manchas de pele (Ronneberger, 2015). O princípio fundamental desta rede, Figura 18, é aplicar uma imagem colorida e como saída obter uma imagem binária em preto e branco, onde a cor branca é codificada com o número 1 e faz parte da nossa classe correspondente (oxidação) e o número zero classificado como a ausência de classe e identificado pela

cor preta. As variações da rede U-NET se comportam como o padrão de uma rede convolucional basicamente em duas etapas. A primeira através da entrada da imagem na rede, a qual sofre blocos de convolução e *pooling* para diminuir a sua dimensão, a fim de extrair características menos importantes das classes de interesse. Na segunda etapa ocorre o rótulo de intensificação de características da imagem onde a rede neural atua como um decodificador através de uma convolução transposta. Nessa etapa a ampliação das “*features*” são evidenciadas por um processo chamado *upsampling* ressaltando resolução da imagem até o valor inicial e consequentemente destacando contornos de classes de maior interesse segmentando o objeto.

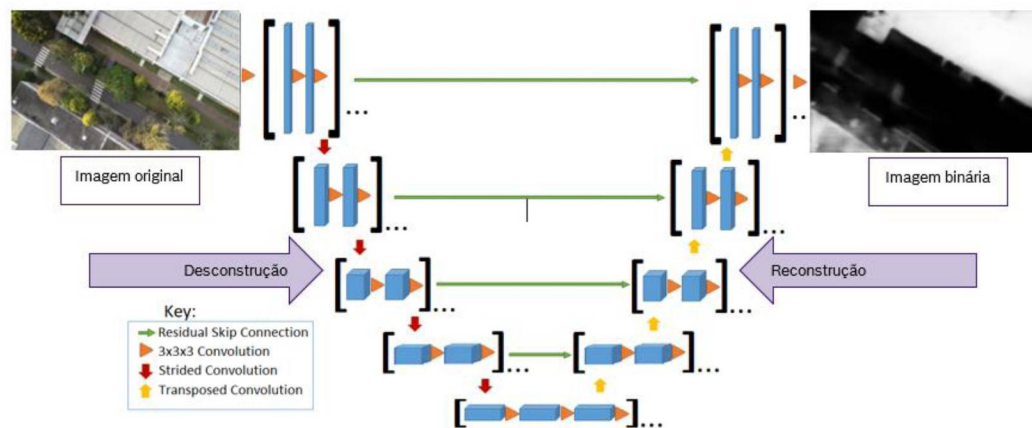


Figura 18: Arquitetura U-Net para segmentação de imagem. Fonte: Ronneberger, 2015 modificada.

3.6. ARQUITETURA DO SISTEMA

A Figura 18 ilustra o fluxograma do funcionamento do processo completo. O acionamento do drone para captura de imagens é feita manualmente, porém a rota é automatizada, ou seja, a partir do momento da decolagem todas as capturas e rotas são feitas de maneira automatizada. As N imagens capturadas pelo drone são armazenadas em um dispositivo de armazenamento externo de memória, que ao ser conectado ao computador faz com que o software (aplicação) seja automaticamente executado. A partir da execução, não há necessidade nenhuma de intervenção humana no processo. A aplicação realiza a importação das N imagens do dispositivo externo e executa o primeiro modelo de segmentação (primeira RNC), conforme Figura 20, para a extração dos telhados, poupando recursos computacionais e evitando que a segunda IA busque oxidações em regiões de não interesse (exemplo ruas, árvores, carros, entre outros). Conforme demonstrado na Figura 22, a imagem com o telhado extraído é então

subdividida em múltiplas imagens de 640x640 pixels, onde um segundo modelo extrator (segunda RNC), desta vez segue a mesma lógica da rede U-NET para segmentar e identificar as oxidações.

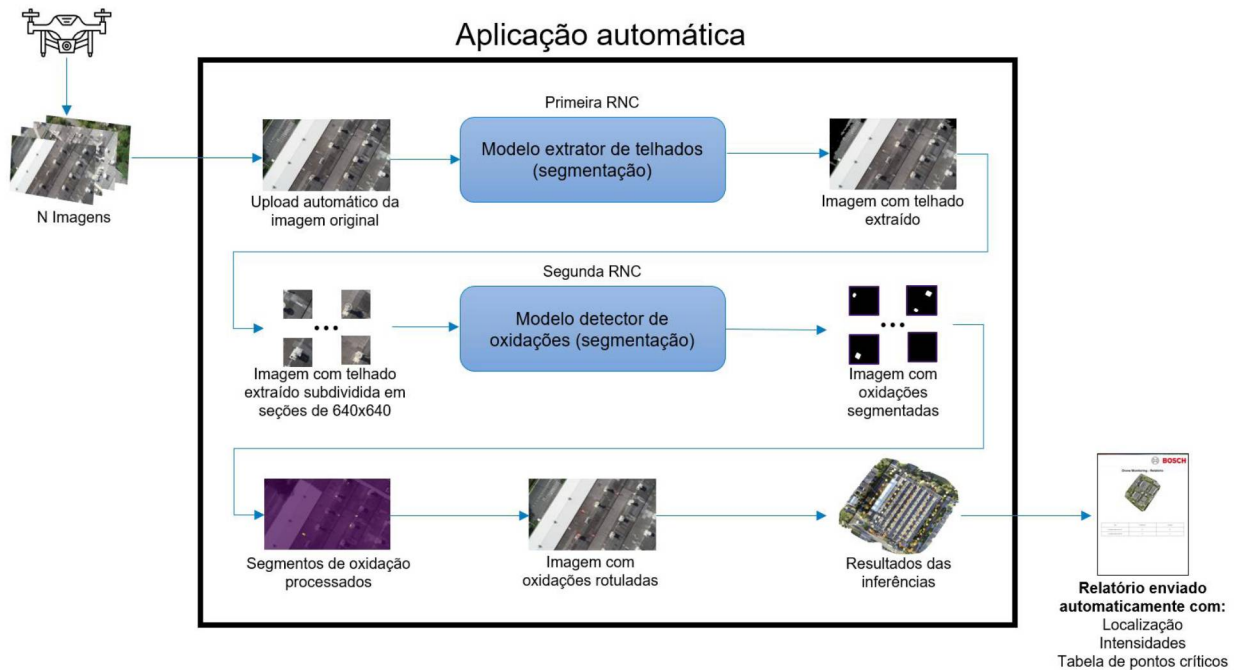


Figura 19: Arquitetura do sistema desenvolvido. Fonte: o autor (2023).

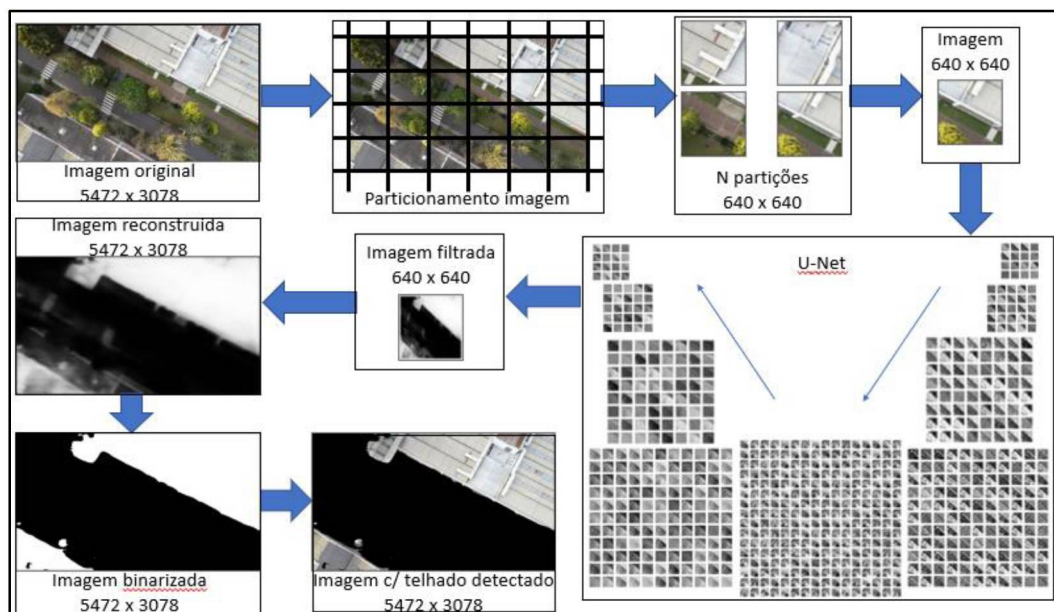


Figura 20: Primeira RNC para localização do telhado. Fonte: o autor (2023).

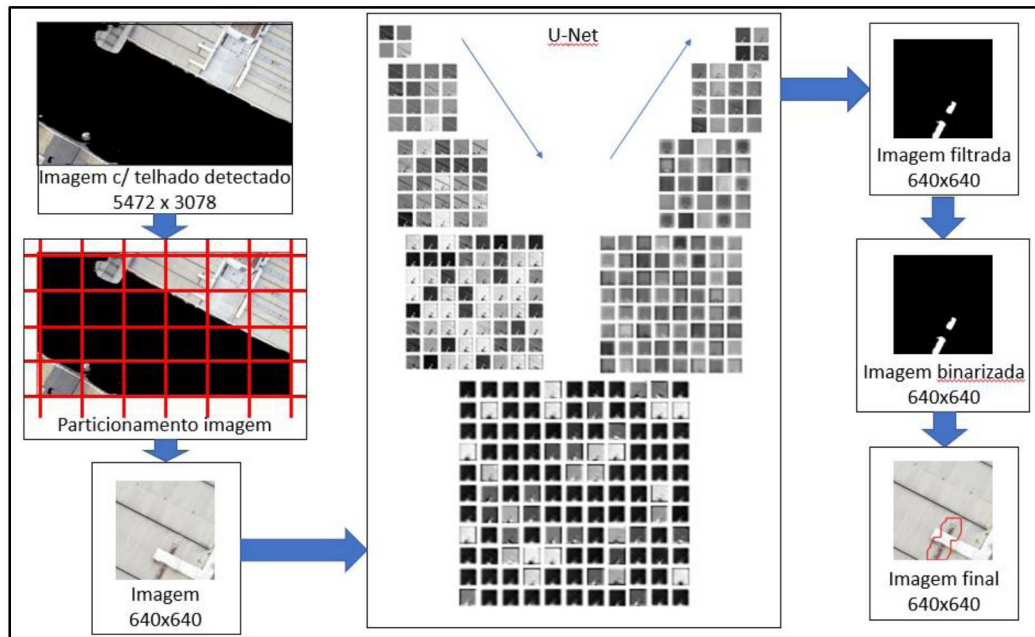


Figura 21: Segunda RNC para identificação da oxidação. Fonte: o autor (2023).



Figura 23: Perfil de relevo obtido. Fonte: o autor (2023).

4.2. CONDIÇÕES DE VOO

4.2.1. Determinação da velocidade de voo

Para determinar a velocidade de voo ótima, foi utilizado a técnica da segunda derivada da Laplaciana, Equação (1). Esta técnica consiste em encontrar em uma imagem de entrada, os valores nulos de pixels onde, tais valores, indicam fronteiras entre os objetos. Com tal técnica, objetiva-se avaliar qual a velocidade máxima na qual a aeronave pode percorrer sem causar distorções e variações na definição obtida na foto.

Conforme ilustra Figura 24, uma imagem inicial sofre, ao decorrer de toda a sua área, variações de intensidade de *pixel*, ao aplicar a primeira derivada na função que representa a dada foto, estas variações tornam-se pontos de máximo local. Em seguida, se aplicarmos a segunda derivada, estes pontos, anteriormente de máximos locais, tornam-se valores nulos. Desta forma, ao realizarmos uma análise da variância dos valores de pixel, pode-se determinar, de uma maneira concreta, qual velocidade gera fotos com uma resolução e discriminação de objetos maior. Uma visão geral deste processo é representada pela Figura 25.

$$Laplace(f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (1)$$

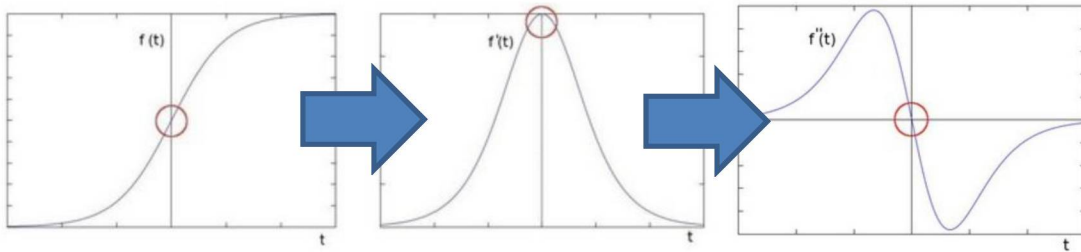


Figura 24: demonstrativo da segunda Laplaciana de uma imagem. Fonte: o autor (2023).

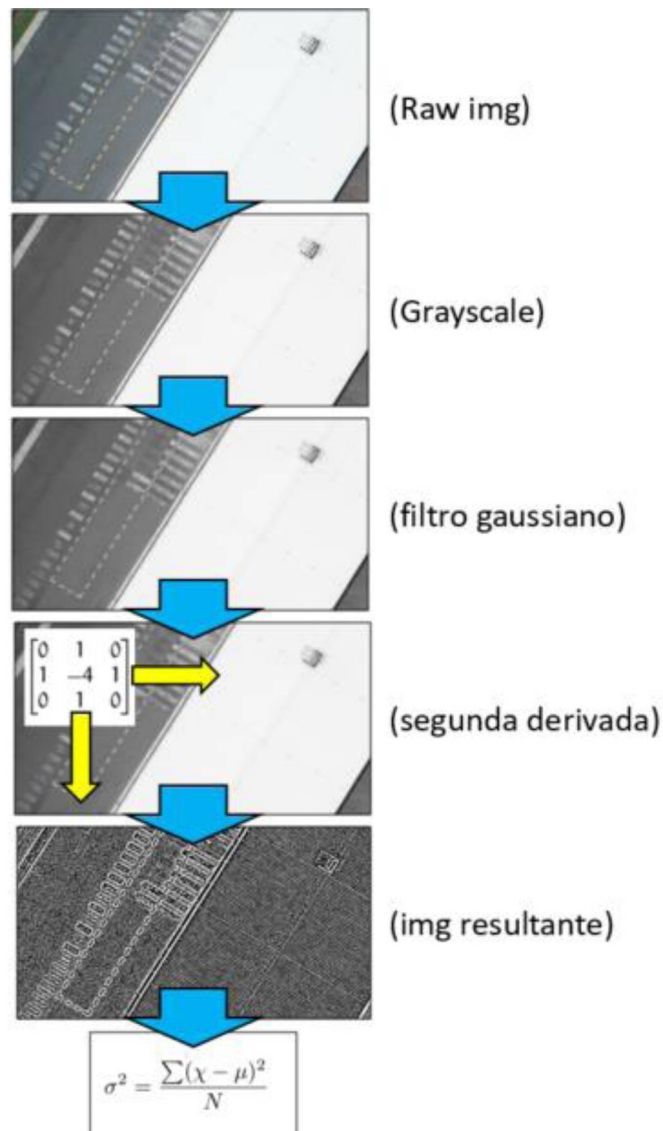


Figura 25: Processo geral para a avaliação das velocidades de voo. Fonte: o autor (2023).

Ao realizar este procedimento, obteve-se a Tabela 2, onde observa-se que a velocidade de voo tem um papel relevante no que tange a qualidade da foto. Para valores de até 25 km/h, há valores muito similares de variância observados e, ao aumentar este parâmetro para 30 km/h observa-se uma redução considerável em tal.

Tal redução se torna mais acentuada a partir de 35 km/h, onde observou-se um decréscimo das variâncias.

Logo, avaliando as médias destas variâncias de pixel, pode-se definir um valor de 25 km/h como sendo a velocidade máxima a ser utilizada, uma vez que para valores mais elevados, irá se observar uma imagem mais “embaçada”. Para fins comparativos, a Figura 26 ilustra uma foto capturada a 25 km/h, enquanto a Figura 27 representa uma mesma área, porém, obtida a uma velocidade de 45 km/h.

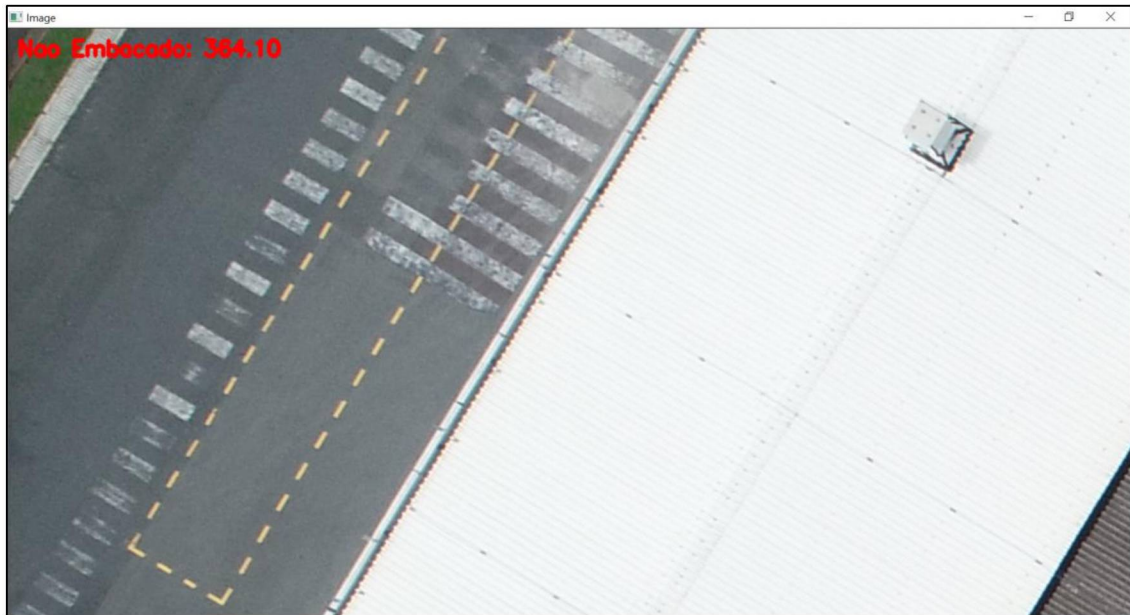


Figura 26: Imagem obtida para uma velocidade de 25 km/h.

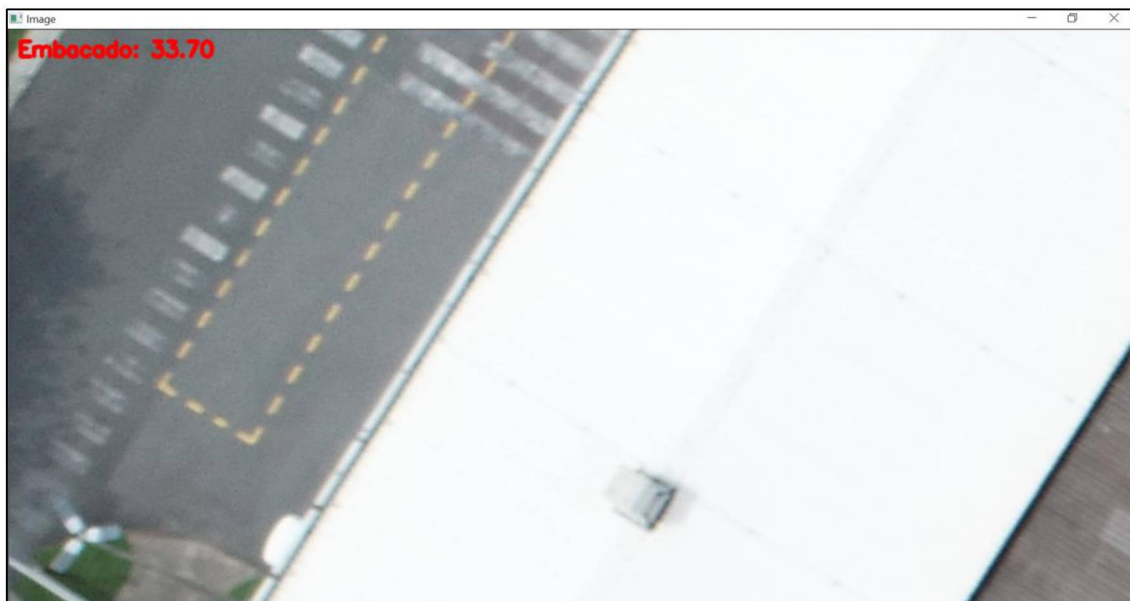


Figura 27: Imagem obtida para uma velocidade de 45 km/h.

Tabela 2: Resultados obtidos com a análise da variância das velocidade de voo. Fonte: o autor (2023).

		Velocidade (km/h)						
		15	20	25	30	35	40	45
Variação da laplaciana		251,09	35,72	225,77	208,65	34,49	31,33	30,89
		269,35	32,43	214,7	190,64	31,98	34,42	33,92
		226,82	30,28	202,95	180,94	31,99	33,91	30,48
		209,92	33,98	204,78	197,87	32,16	32,67	30,95
		210,26	31,2	209,9	197,15	32,45	34,02	32,73
		233,64	40,37	242,79	209,78	36,82	37,15	36,36
		267,3	36,48	272,72	215,18	36,94	40,56	37,82
		278,12	35,85	253,31	206,81	31,61	34,19	31,79
		348,45	32,75	291,47	194,24	30,84	34,79	36,4
		343,13	36,57	313,05	214,89	32,11	36,71	34,01
		351,39	291,55	307,2	190,33	31,94	37,52	36,2
		376,53	275,59	297,01	217,98	30,32	33,67	36,58
		332,96	298,12	339,6	245,38	31,47	34,34	38,1
		321,34	316,31	341,01	240,04	31,42	34,98	37,43
		320,14	346,62	351,6	284,41	33,2	34,86	37,26
		339,65	351,1	354,33	299,54	32,8	35,11	50,63
		362,31	361,44	364,1	316,85	41,22	33,7	49,81
		363,33	378,38	363,62	326,39	32,7	33,84	34,82
		407,01	398,5	387,44	338,15	35,14	35,11	39,55
		423,71	421,3	402,96	346,14	37,97	40,16	44,63
		411,78	391,03	386,07	339,66	37,45	41,4	40,81
		394,35	385,48	366,77	321,32	38,1	41,55	41,58
		378,26	353,7	343,37	296,45	40,32	39,58	43,32
		338,18	328,35	327,27	276,41	38,28	43,32	41,36
		368,05	327,44	321,27	268,55	41,37	42,13	40,24
		340,49	365,11	328,25	286,92	41,45	42,55	42,99
		337,73	335,06	324,4	252,15	37,46	38,46	39,18
		347,62	327,3	348,34	272,85	37,45	40,02	41,56
	339,53	374	350,58	300,69	38,34	41,53	42,4	
	345,48	352,68	351,56	266,09	39,44	38,56	40,54	
	372,94	341,92	337,92	277,55	38,3	37,75	38,14	
	346,62	339,65	320,68	204,64	36,93	39,67	38,6	
	Média	329,9213	250,1956	313,9622	255,77	35,45188	37,17375	38,47125

4.2.2. Determinação período de voo

Outro fator relevante para uma imagem de entrada de qualidade é a iluminância do ambiente. Esta variável também irá afetar a maneira na qual a rede neural irá ler uma área. Dependendo do período, pode-se obter variações na uniformidade de dados de uma mesma área, como exemplifica a Figura 28. Logo, uma avaliação dos impactos que este parâmetro tem sob a qualidade das imagens fez-se necessário.

Para tal, foi-se levantado um histograma contendo a distribuição da intensidade de pixels da imagem em uma escala de cinza. Desta maneira, objetivou-se analisar o comportamento destas curvas em diferentes períodos de um mesmo dia, para verificar quais os impactos o fato iluminância teria sob. A Figura 29 ilustra de maneira objetiva, as etapas deste processo.

Avaliando o histograma obtido, representado pela Figura 30, observa-se que a curva obtida para o período da manhã (~140.00 lx), representada pela cor verde, demonstra um número elevado de pontos de máximo local ao longo de todo o eixo horizontal, com dois picos próximos ao 240, dois picos de máximo local em 120 e outros dois próximos do 0. Quando comparado com a curva laranja, na qual representa uma distribuição de escala de cinza para valores de uma iluminância ambiente de 40.000 lx, nota-se um deslocamento para a esquerda de todo o espectro obtido. Além deste deslocamento, observa-se que os dois pontos de máximos locais próximos ao 240 acabaram unificando-se em um único pico, o que sugere uma menor diferenciação de pixels de tons próximos. No mais, nota-se um achatamento destes pontos de máximos locais na região abaixo do 100, o que indica uma maior uniformidade destas intensidades. Por fim, ao avaliar a curva verde, a qual diz respeito à uma imagem obtida a noite, com uma iluminância próxima de 100 lx, verifica-se um maior deslocamento da distribuição das densidades de pixel para a esquerda. Além disto, percebe-se apenas três pontos de máximo local, onde todos estão situados após a faixa de 100, com o pico de máximo global próximo de 20, indicando uma densidade maior de pixels pretos nestas fotos.

Portanto, sugere-se que períodos “mais claros”, ou seja, com intensidades luminosas maiores, obtém-se capturas mais definidas, com uma maior diferenciação de detalhes e, de acordo com a redução deste parâmetro, observa-se uma imagem “mais escura” com alguns detalhes reduzidos, devido a disparidade de intensidades menor.

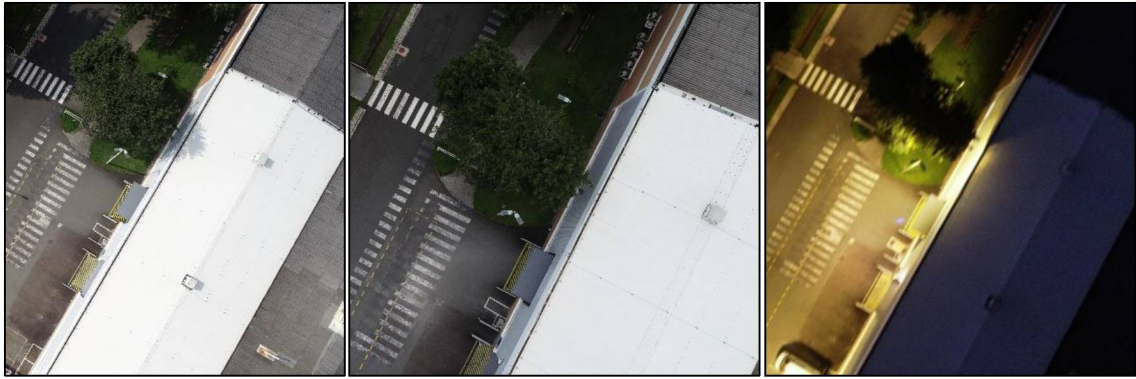


Figura 28: Comparativo de uma mesma área na qual as imagens foram capturadas em períodos distintos sendo eles respectivamente manhã, tarde e à noite. Fonte: o autor (2023).

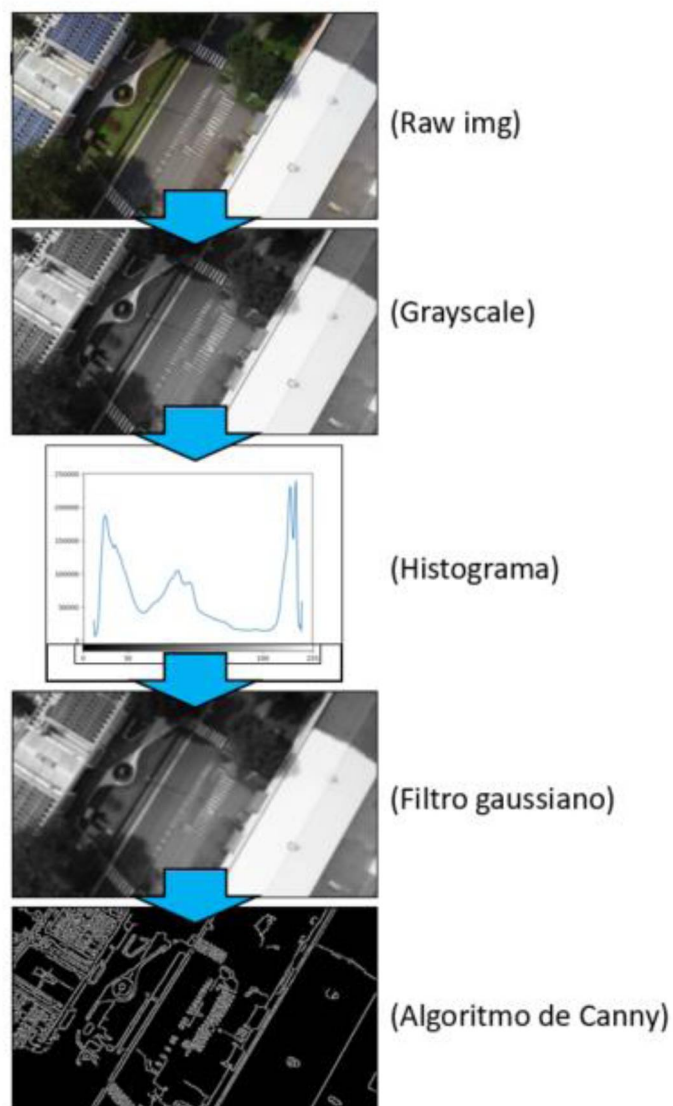


Figura 29: Processo de avaliação da influência da iluminância na qualidade da imagem. Fonte: o autor (2023).

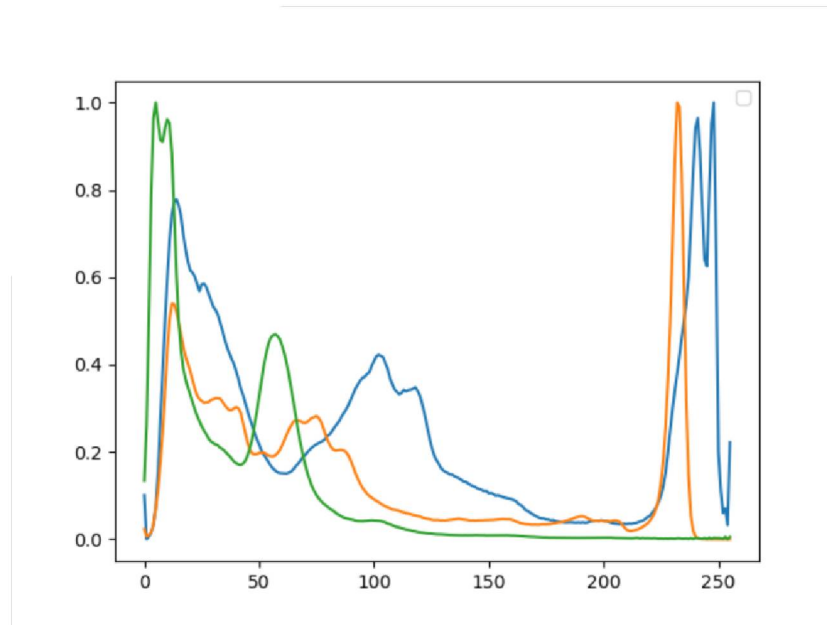


Figura 30: Histograma obtido referente aos três períodos analisados. Em azul, está representado a distribuição obtida para o período da manhã com aproximadamente 140.000 lx, em laranja está representado o período da tarde, com cerca de 40.000 lx, e em verde, está representado o período noturno, com cerca de 100 lx. Fonte: o autor (2023).

4.3. EXPERIMENTO TAGUCHI

Com estas três experiências realizadas, utilizou-se da ferramenta de análise Taguchi para avaliar o impacto de cada parâmetro sob o conjunto de dados obtidos. Para tal, estas variáveis foram processadas com o auxílio do *software* Minitab® sob um perfil de análise L⁹, onde foram permutados os valores de altura de 30, 50 e 80m, velocidades de 15, 25 e 45 km/h e uma luminosidade de ambiente de 140.000, 40.000 e 10 lx. A Figura 31 demonstra os resultados obtido com tal experimento.

Observa-se que a altura exibiu uma influência de perfil linear com coeficiente angular negativo, onde 30m exibiu melhores médias. Com relação ao parâmetro de velocidade, observa-se um valor de média constante para a faixa entre 15 à 25 km/h e, para valores mais elevados, neste caso 45 km/h, há uma redução significativa no valor médio. Por fim, observou uma redução no valor médio obtido referente à luminosidade do ambiente onde uma redução linear de coeficiente negativo é observada entre 140.000 à 40.000 lx e esta sofre uma redução ainda mais acentuada para a transição para 10 lx.

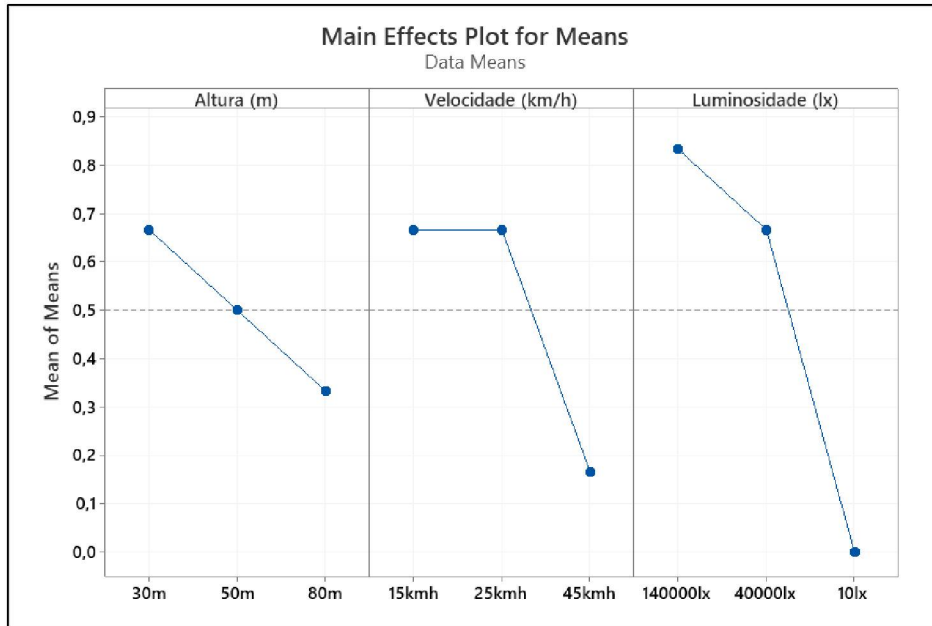


Figura 31: Experimento Taguchi obtido. Fonte: o autor (2023).

4.4. ALGORÍTMO ORTOMOSÁICO

Um algoritmo ortomosáico foi desenvolvido durante o estudo. Este, através de imagens de entrada sobrepostas umas das outras, objetiva criar uma única imagem, de resolução maior e um alto detalhamento, de um todo.

Para tal, considerando uma altura efetiva, $H_{efetiva}$, de 37 metros como sendo a diferença entre a altura de voo, H_{voo} , e a altura do telhado, $H_{telhado}$, conforme equação (2):

$$H_{efetiva} = H_{voo} - H_{telhado} \quad (2)$$

E que o campo de visão da câmera do VANT como sendo de 84° , temos, conforme equação (3), a determinação da distância de cobertura, $D_{cobertura}$, de meia foto capturada:

$$\tan \alpha = \frac{D_{cobertura}}{H_{efetiva}} \rightarrow D_{cobertura} = 33,31m \quad (3)$$

Logo, um total de 66,63 metros de cobertura são capturados por foto. Então, considerando a necessidade de uma sobreposição entre capturas de, ao menos, 70% para garantir uma reconstrução fiel as imagens originais, é necessário uma foto a cada 19,98m. Considerando uma velocidade de voo ótima de 25 km/h, conforme levantado

anteriormente em 4.2.1, é preciso, ao menos, uma foto a cada 2,88 segundos. A Figura 32 demonstra a disposição das medidas utilizadas para este cálculo.

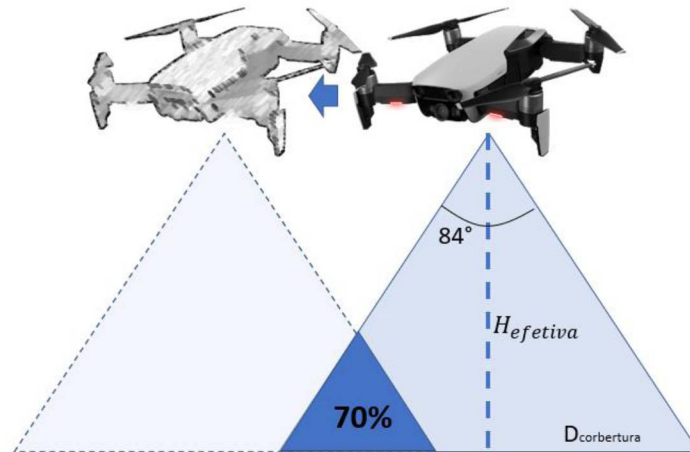


Figura 32: Ilustração do processo de cálculo para a determinação do intervalo de disparos da câmera. Fonte: o autor (2023).

A Figura 33 ilustra um comparativo entre uma imagem capturada com o VANT e a mesma área obtida após a geração da imagem ortomosáica.

Nota-se algumas diferenças significativas entre elas onde, a principal, diz respeito à perda de perspectiva das estruturas. Esta perda se dá devido à uma característica própria da técnica, onde a partir de uma certa altura das estruturas presentes nas capturas, a sobreposição gera inconsistências durante o processamento. Esta perda de perspectiva acaba “mascarando” parte das falhas, principalmente as situadas nas bases de estruturas metálicas, como nos terminais de ventilação, conforme Figura 33. Adicionalmente, observou-se uma grande dificuldade do algoritmo em reconstruir uma versão “achatada” de elementos com altura, o que resultou, adicionalmente, em uma imagem ortomosáica com deformações.

Desta maneira, após extensas avaliações, optou-se por não seguir em frente com a ideia de treinar a imagem e, posteriormente, alimentá-la com as figuras ortomosáicas, uma vez que a rede não seria capaz de encontrar parte das falhas e a necessidade de otimizar o processamento do algoritmo. No entanto a imagem ortomosáica foi utilizada como base para a geração do relatório preventivo, contendo a classificação e localização das oxidações e que de acordo com o método é enviado automaticamente por e-mail ao time de manutenção.

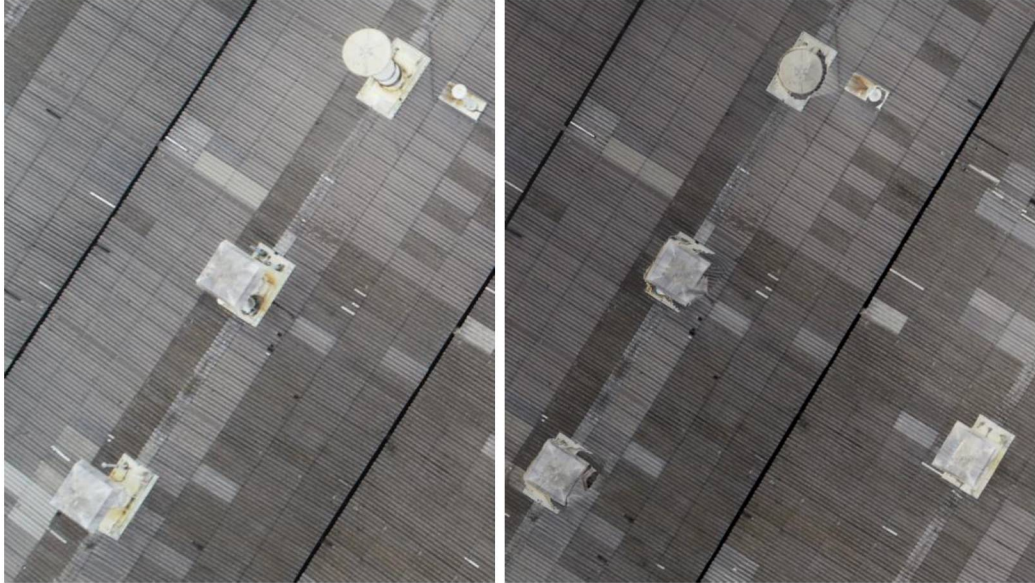


Figura 33: Comparativo entre uma imagem original, retirada pelo VANT (esquerda), e uma mesma área gerada pelo ortomosáico. Fonte: o autor (2023).

4.5. TREINAMENTO E VALIDAÇÃO RNA

Logo, após a definição dos parâmetros ótimos para captura das imagens, realizou-se a etapa de treinamento da rede neural convolucional, conforme processo descrito em 3.2.

Devido ao fato de que a rede neural proposta neste trabalho contém duas RNCs distintas, uma para detecção de telhados e outra para segmentação de oxidações, e para esta pesquisa foram utilizados dois *datasets* distintos, quatro arranjos diferentes para o treinamento e avaliação do sistema foram realizados, em duas rodadas diferentes. Esta “permutação” foi necessária para avaliar o desempenho de cada parâmetro (telhado e oxidação) e a sua influência na rede como um todo.

A principal métrica utilizada para avaliar o desempenho da rede em questão foi o coeficiente de pontuação Dice (DSC), o qual é um índice de sobreposição espacial e uma métrica de validação de reprodutibilidade. Este, também conhecido como Sorensen-Dice coeficiente, foi inicialmente introduzido no campo de botânica, em 1948, e sua formula original, Equação (3), tinha uma aplicação em dados discretos onde, dado dois *datasets*, X e Y, o coeficiente era resultado do dobro da intersecção destes grupos de dados, dividido pela somatória destes.

$$DSC = \frac{2 |X \cap Y|}{|X| + |Y|} \quad (3)$$

Quando aplicado dados Booleanos, fazendo-se uso das definições de Verdadeiros Positivos (VP), Falsos Positivos (FP) e Falsos Negativos (FN), tem-se:

$$DSC = \frac{2xTP}{(TP + FP) + (TP + FN)} \quad (4)$$

Os valores obtidos a partir do DSC variam entre 0, indicando nenhuma sobreposição espacial entre dois conjuntos de resultados de segmentação binária, à 1, indicando sobreposição completa. Esta métrica é o principal parâmetro utilizado para mensurar a capacidade de acerto de redes de segmentação, onde valores acima de 0,5 podem ser considerados bons resultados. Este parâmetro tem como principal característica a priorização dos acertos da rede (Verdadeiros Positivos) e uma penalização não tão severa quando comparado com as taxas de *Intersection over Unit* (IoU) utilizadas durante a qualificação desta pesquisa.

Na Figura 27, observam-se os resultados obtidos para os quatro treinamentos com a sua fase de validação no *dataset* Ctp, representando a primeira rodada de testes. Nesta primeira rodada, a rede foi treinada com dados de telhado e oxidação de ambas as bases de dados, como observa-se no eixo x desta figura. Uma pequena redução nos valores de Q1 e Q3 foram observadas quando houve a introdução do *dataset* Cap durante a fase de treinamento. Avaliando a Tabela 4, a maior alteração foi observada quando a fase de treinamento foi realizada com dados do telhado e oxidação de Cap, onde os valores de mediana sofreram uma redução de 0,55 para 0,51. Esta alteração indica um possível *overfitting* da rede quando o treinamento e validação foi realizado com o *dataset* Ctp. Porém, considerando o tamanho do banco de dados, a complexidade do problema e o desvio padrão obtido, mais evidências são necessárias para uma conclusão mais sólida.

Com isto em mente, um segundo sprint de testes foi realizado, validando a rede no *dataset* Cap, mantendo as permutações para o treinamento da RNC. A Figura 28 ilustra os resultados obtidos nesta rodada. Dado as diferentes características arquitetônicas, quando comparando as duas plantas industriais, a IA teve o seu pior *score* quando treinada apenas com as informações provenientes do *dataset* Ctp,

Tabela 5, devido à sua incapacidade de generalização para os diferentes prédios. Com a introdução de dados de Cap à rede, durante a fase de treinamento, um aumento no desempenho desta foi observado. Quando comparando os arranjos 1B e 2B, o primeiro quartil não sofreu uma alteração considerável, assim como o valor de Q2. Quando os dados de oxidação do dataset Cap foram introduzidos durante a fase de treino, a IA atingiu maiores pontuações, devido à sua maior capacidade de generalização resultando em um aumento no valor máximo, terceiro quartil e na janela interquartil. O fato de que a IA não foi treinada em 2B com dados do telhado de Cap, interferiu em seu desempenho, devido às divergências nas construções. Esta interferência foi observada na janela do interquartil onde houve uma maior dispersão das pontuações obtidas. Esta rede não estava apta a detectar estruturas de telhados, o que impactou diretamente na segmentação das oxidações.

Ao examinar o arranjo 3B, o valor Q1 sofreu um aumento, quando comparado com 1B e 2B, o que indica um aumento no desempenho na rede ao adicionar dados de telhado do *dataset* Cap ao treino da IA. Enquanto o primeiro quartil aumentou, uma redução no terceiro quartil ocorreu. Quando comparando 3B com 2B, essa redução era esperada, uma vez que a rede era capaz de detectar telhados, porém, não era capaz de detectar oxidações devido às diferentes estruturas e padrões arquitetônicos.

No mais, no arranjo 4B, quando os dados de Cap foram introduzidos para ambos, telhado e oxidação, durante a fase de treinamento, o valor Q1 manteve-se parelho com os valores obtidos em 3B enquanto o terceiro quartil sofreu um aumento. Este aumento da janela do interquartil aliado à melhora no valor de mediana, expressa uma grande diferença não só dos padrões de infraestrutura do prédio, mas também das oxidações, as quais variam de uma planta para outra.

Figura 29 ilustra um exemplo de desempenho para o melhor modelo treinado com dados de Cap e Ctp. Neste exemplo, onde obteve-se um DSC de 0,706 para uma imagem do *dataset* Cap, a RNC conseguiu segmentar a maior parte das áreas relevantes de oxidação. Figura 30 demonstra outra imagem de teste, onde a rede obteve um DSC de 0,525. Mesmo que este resultou em uma pontuação baixa, quando comparado com a imagem anterior, a rede foi capaz de sementar uma grande área de oxidação indicada, e até mesmo, criar uma melhor máscara do que a realizada manualmente por um especialista, como observado no topo direito da Figura 30.

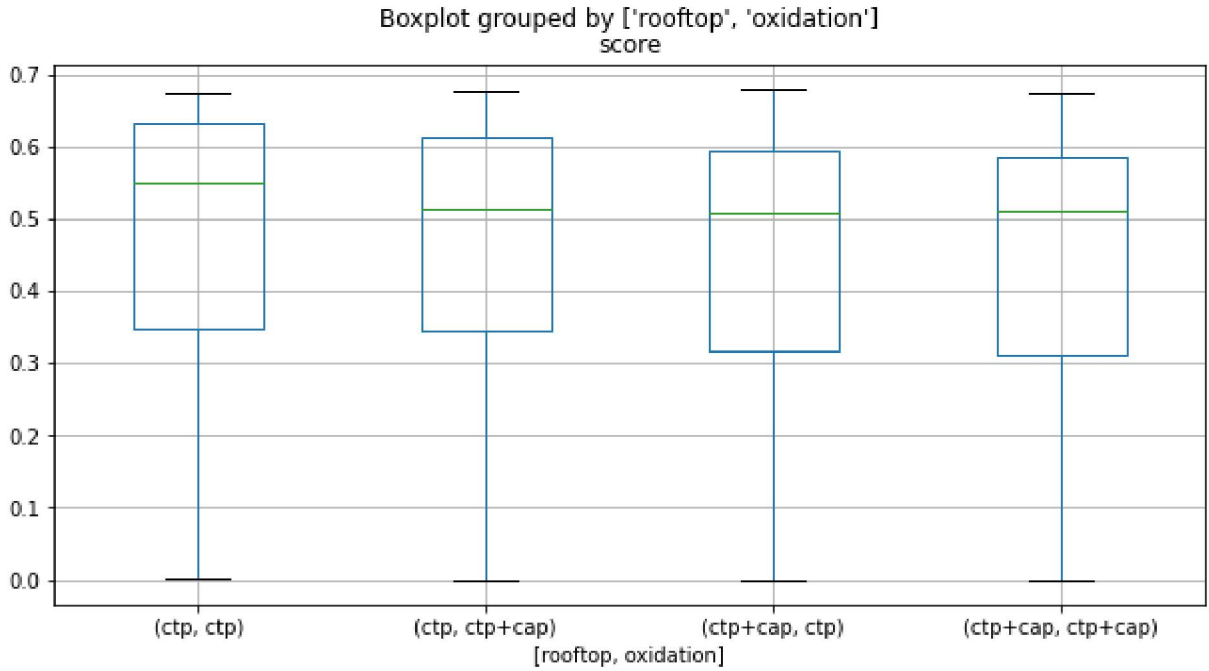


Figura 34: Resultados obtidos da rede com validação em Ctp. Fonte: o autor (2023).

Tabela 3: Valores obtidos durante a primeira rodada de testes. Fonte: o autor (2023).

Grupo	Telhado	Oxidação	Média	Mediana	Desvio padrão
1A	Ctp	Ctp	0.458	0.549	0.230
2A	Ctp	Ctp + Cap	0.442	0.512	0.230
3A	Ctp + Cap	Ctp	0.429	0.509	0.232
4A	Ctp + Cap	Ctp + Cap	0.433	0.510	0.219

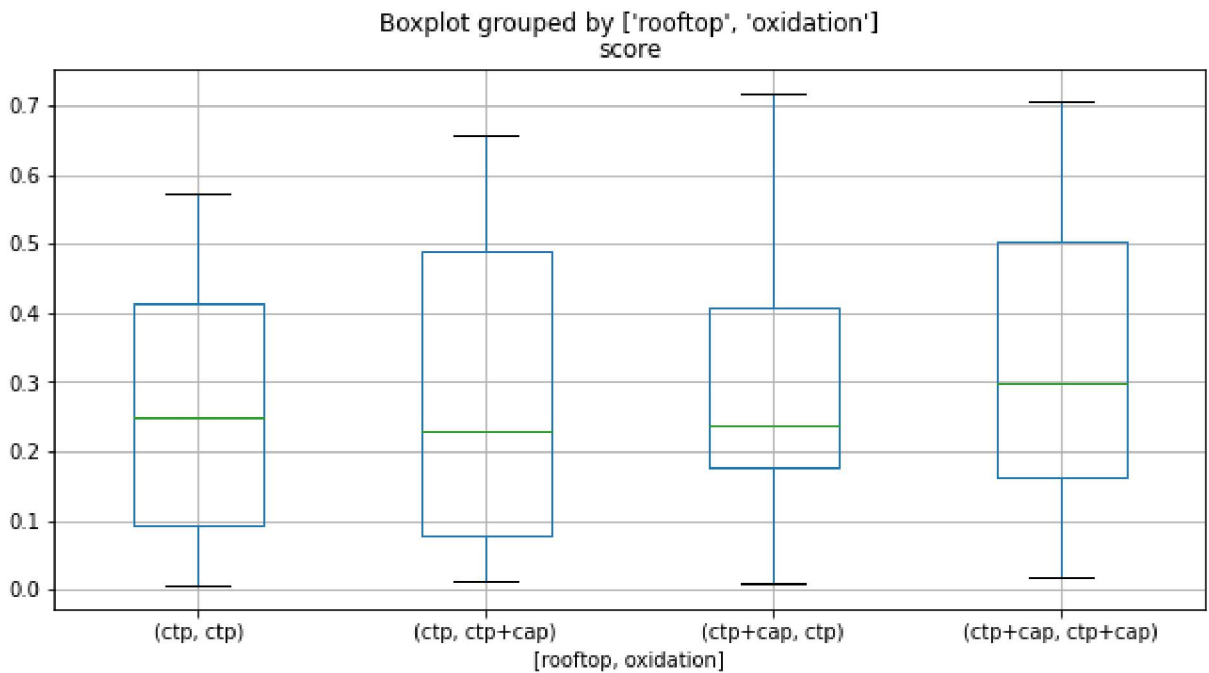


Figura 35: Resultados obtidos da rede com validação em Cap. Fonte: o autor (2023).

Tabela 4: Resultados obtidos durante a segunda rodada de testes. Fonte: o autor (2023).

Grupo	Telhado	Oxidação	Média	Mediana	Desvio padrão
1B	Ctp	Ctp	0.257	0.249	0.185
2B	Ctp	Ctp + Cap	0.284	0.229	0.223
3B	Ctp + Cap	Ctp	0.295	0.238	0.221
4B	Ctp + Cap	Ctp + Cap	0.331	0.298	0.222

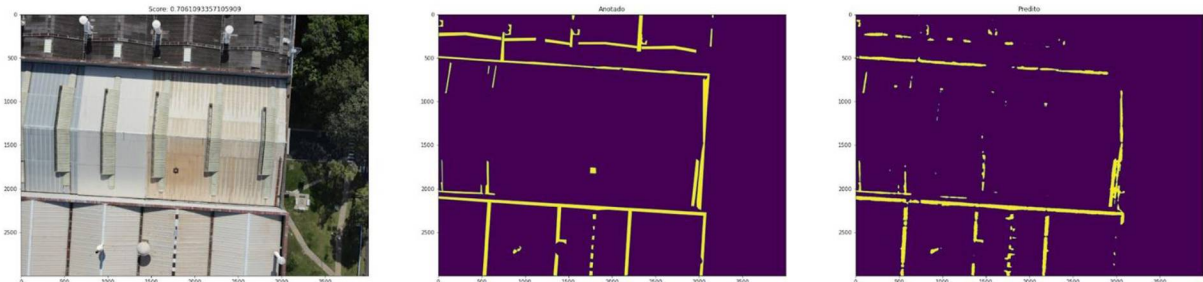


Figura 36: Resultados obtidos da rede. Fonte: o autor (2023).

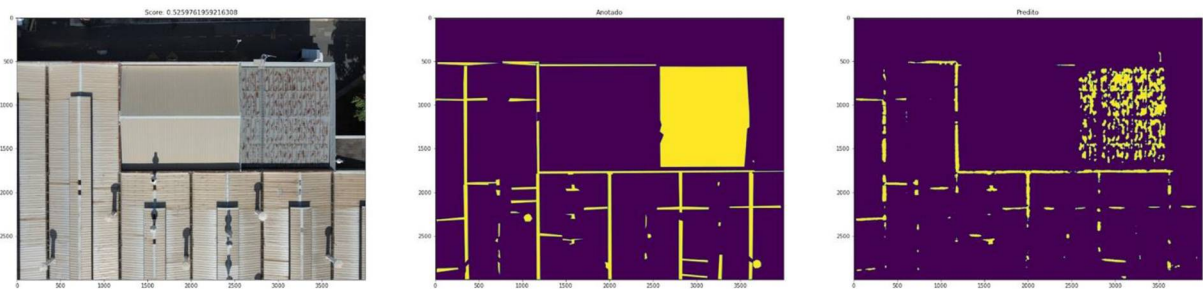


Figura 37: Resultados obtidos da rede. Fonte: o autor (2023).

4.6. CLASSIFICAÇÃO DAS OXIDAÇÕES

Após localizada a falha de oxidação as imagens subdivididas são então remontadas a fim de gerar uma imagem binarizada (preta e branca) do tamanho da imagem original.

A imagem remontada que contém os segmentos de oxidação processados são utilizados sobrepondo a imagem original que contém os metadados de latitude e longitude, automaticamente registrados pela VANT no momento da captura, indicando de maneira precisa a localização das oxidações.

Então de posse dos resultados das inferências as oxidações são classificadas em diferentes cores em um ortomosaico previamente gerado. Estas cores variam de branco a marrom, Figura 38, onde branco indica a ausência ou número muito baixo de oxidações encontradas e marrom indica os locais onde há maiores quantidades (densidade) de oxidações encontradas.



Figura 38: Ortomosaico com localização e intensidade das oxidações. Fonte: o autor (2023).

4.7. SOLUÇÃO TECNOLÓGICA – RELATÓRIO DE DIAGNÓSTICO PREVENTIVO

A arquitetura inicialmente proposta para o funcionamento da solução, idealizada na Figura 19, por fim, é materializada através de uma interface, Figura 39, onde a inteligência executa as tarefas de identificação e classificação das falhas de oxidação em estruturas metálicas dos telhados. Dando continuidade, de maneira automática um software, que é concatenado a IA, gera um relatório que é enviado via *e-mail* aos funcionários cadastrados. Um relatório com a localização precisa das oxidações (destacado em cores, com coordenadas GPS e tabela de intensidade de priorização dos pontos críticos) confere um diagnóstico preventivo inteligente conforme ideia proposta, Figura 40.

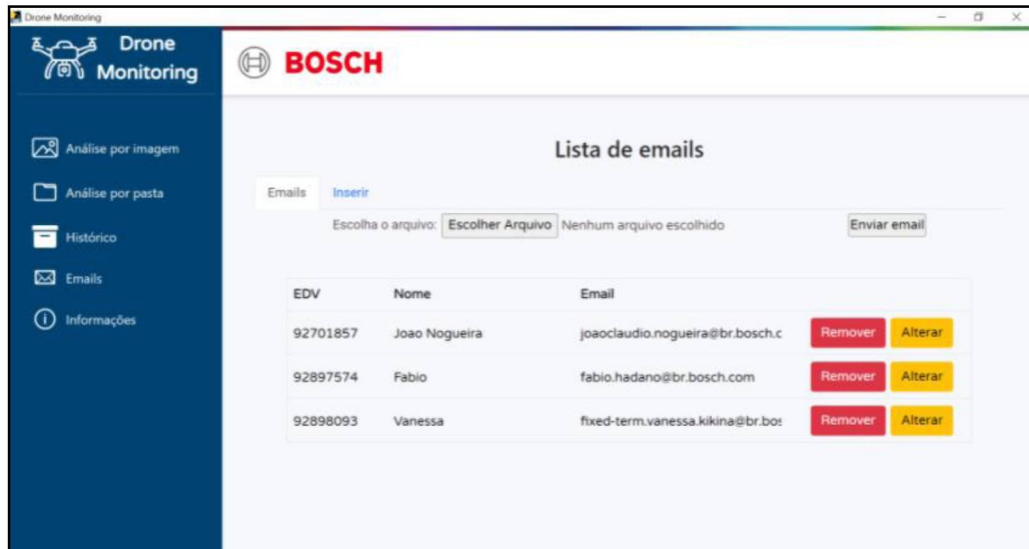


Figura 39: Página inicial de navegação do software gerenciado pela IA. Fonte: o autor (2023).

The screenshot shows an automated report titled 'Drone Monitoring - Relatório'. At the top right is the Bosch logo. Below the title is a 3D aerial view of a building complex. Below the image is a table with the following data:

Nome	Densidade (%)	Detecções
D:\DCIM\100\MEDIA\QJL_0984.JPG	0.11	5
D:\DCIM\100\MEDIA\QJL_0980.JPG	0.09	7
D:\DCIM\100\MEDIA\QJL_0980.JPG	0.08	7
D:\DCIM\100\MEDIA\QJL_0984.JPG	0.08	8
D:\DCIM\100\MEDIA\QJL_0986.JPG	0.08	7
D:\DCIM\100\MEDIA\QJL_0986.JPG	0.08	7
D:\DCIM\100\MEDIA\QJL_0988.JPG	0.06	4
D:\DCIM\100\MEDIA\QJL_0988.JPG	0.05	4
D:\DCIM\100\MEDIA\QJL_0982.JPG	0.04	5
D:\DCIM\100\MEDIA\QJL_0986.JPG	0.04	5

Figura 40: Relatório automatizado enviado por e-mail. Fonte: o autor (2023).

5. CONSIDERAÇÕES FINAIS

Neste trabalho, no qual o foco é a detecção e classificação de oxidações em estruturas metálicas que compõem a cobertura de edificações fabris, notou-se, baseado em estudos prévios na literatura que a oxidação neste cenário é um problema moderno e que assola o setor industrial. Com esse aprendizado foi possível então desenvolver uma ferramenta de diagnóstico preventivo inteligente, ágil e de baixo custo em comparação com processos manuais. Sendo assim, uma opção sustentável para operações de manutenção predial e para a indústria em geral.

Os objetivos específicos foram alcançados em sua totalidade. Com o auxílio da ferramenta de MDCB, evidenciou-se, que estas falhas são de elevado grau de relevância e envolvem custos adicionais diretos e indiretos ao processo de manutenção e manufatura.

Logo, o desenvolvimento de técnicas inovadoras com segurança e de baixo custo faz-se necessário para manter a competitividade e produtividade da indústria em seu segmento. Tecnologias já existentes na literatura, aliado às tendências da I4.0 como, por exemplo, IA, IoT, IoP e IoS podem gerar a próxima geração de alternativas para esta nova década e com base em referências importantes este trabalho foi conduzido.

Com o auxílio de técnicas quantitativas e qualitativas já comumente conhecidas na literatura, como segunda Laplaciana e o algoritmo de Canny, foi possível determinar aspectos fundamentais para a geração de imagens de qualidade e com alto grau de detalhamento, onde observou-se uma velocidade máxima de 25 km/h, uma altura de voo de 50 m e o período da manhã (~140.000 lx) como sendo as variáveis adequadas para maior produtividade no treinamento rede neural.

No ponto de vista de condensação de informações, o algoritmo ortomosaico, cujo objetivo era reduzir o custo computacional da rede, demonstrou resultados satisfatórios, porém, devido à característica de “achatar” imagens cujo objetos demonstram uma perspectiva (3D), devido à altura, não resultou em uma alternativa factível, uma vez que este iria “mascarar” parte das falhas estruturais. Por outro lado, a imagem ortomosaica em 2D foi considerada como parte da construção do software que gera o relatório preventivo e é enviado automaticamente ao time de manutenção com a indicação e classificação dos pontos de oxidação a serem monitorados.

A rede neural em si, mostrou-se capaz de detectar as falhas em questão, demonstrando resultados satisfatórios com relação ao DSC. Devido à descoberta em que o estilo arquitetônico bem como as oxidações variam, a depender do local fornecido para o dataset, esta rede atingiu valores razoáveis de desempenho, no entanto, sendo necessário um maior volume de dados referente à Planta B. A ferramenta foi implementada e validada através de simulações e a comprovação da eficácia do projeto resultou na entrada de registro de patente pela planta A.

Valido ressaltar também que no decorrer do projeto de mestrado o artigo ***“Pattern recognition and oxidation classification in metal structures of industrial roofs using artificial intelligence”*** foi publicado pelo ISTE 2022 (*International Society of Transdisciplinary Engineering 2022*) e apresentado oralmente no MIT-Massachusetts Institute of Technology, em Boston nos EUA, em julho de 2022 (Publicação referenciada com doi 10.3233/ATDE28). O projeto também foi premiado pela Associação Brasileira de Facility Management (ABRAFAC) em outubro de 2022 e reconhecido como um dos três projetos de referência de inovação industrial, em novembro de 2022, pela Associação Brasileira de Internet Industrial (ABII).

Por fim a rede neural de segmentação semântica, inicialmente usada na área biomédica foi essencial como referência para identificação de características complexas. A associação de duas redes neurais convolucionais do tipo U-NET, conceito adotado neste trabalho, foi um diferencial para desenvolver uma IA com aspectos de aprendizado rápido e identificação precisa de características de interesse com baixo volume de dados de entrada.

Com o olhar de escalabilidade, este trabalho possui potencial para aplicação em diversos campos, não apenas o da indústria. Devido ao grande potencial de tecnologias de aprendizado de máquina aplicadas ao uso da rede U-NET em curto espaço de tempo, novas revisões e algoritmos são desenvolvidos. Portanto para articular melhorias nessa área, recomenda-se abordagens mais recentes como as propostas por Saeid Asgari Taghanaki (2020), uma revisão sobre segmentação semântica e a de Nabil Ibtehaz (2020), que traz uma reflexão sobre melhoria de performance em arquitetura de rede do tipo U-NET.

REFERÊNCIAS BIBLIOGRÁFICAS

- Adam, Elhadi, Onesimo Mutanga, and Denis Rugege. "Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review." *Wetlands Ecology and Management* 18.3 (2010): 281-296.
- Altug, S., Mo-Yuen Chen, & Trussell, H. J. (1999). Fuzzy inference systems implemented on neural architectures for motor fault detection and diagnosis. *IEEE Transactions on Industrial Electronics*, 46(6), 1069–1079. <https://doi.org/10.1109/41.807988>
- Ambrosia, Vincent G., et al. "Demonstrating UAV-acquired real-time thermal data over fires." *Photogrammetric engineering & remote sensing* 69.4 (2003): 391-402.
- Atha, Deegan J., and Mohammad R. Jahanshahi. "Evaluation of deep learning approaches based on convolutional neural network"
- Asgari Taghanaki, S., Abhishek, K., Cohen, J.P. *et al.* Deep semantic segmentation of natural and medical images: a review. *Artif Intell Rev* **54**, 137–178 (2021). <https://doi.org/10.1007/s10462-020-09854-1>.
- Awadallah, M. A., & Morcos, M. M. (2003). Application of AI tools in fault diagnosis of electrical machines and drives-an overview. *IEEE Transactions on Energy Conversion*, 18(2), 245–251. <https://doi.org/10.1109/TEC.2003.811739>
- Basurko, O. C., & Uriondo, Z. (2015). Condition-Based Maintenance for medium speed diesel engines used in vessels in operation. *Applied Thermal Engineering*, 80, 404–412. <https://doi.org/10.1016/j.applthermaleng.2015.01.075>
- Bellini, A., Filippetti, F., Franceschini, G., & Tassoni, C. (2000). Closed-loop control impact on the diagnosis of induction motors faults. *IEEE Transactions on Industry Applications*, 36(5), 1318–1329. <https://doi.org/10.1109/28.871280>
- Benediktsson, Jón Atli, Jón Aevar Palmason, and Johannes R. Sveinsson. "Classification of hyperspectral data from urban areas based on extended morphological profiles." *IEEE Transactions on Geoscience and Remote Sensing* 43.3 (2005): 480-491.
- Benešová, A., & Tupa, J. (2017). Requirements for Education and Qualification of People in Industry 4.0. *Procedia Manufacturing*, 11, 2195–2202. <https://doi.org/10.1016/j.promfg.2017.07.366>

- Bhardwaj, Anshuman, et al. "UAVs as remote sensing platform in glaciology: Present applications and future prospects." *Remote sensing of environment* 175 (2016): 196-204.
- Burkart, Andreas, et al. "A novel UAV-based ultra-light weight spectrometer for field spectroscopy." *IEEE sensors journal* 14.1 (2013): 62-67.
- Byung-Geun Hyun, & Kwanghee Nam. (1995). Faults diagnoses of rotating machines by using neural nets: GRNN and BPN. *Proceedings of IECON '95 - 21st Annual Conference on IEEE Industrial Electronics*, 2, 1456–1461. <https://doi.org/10.1109/IECON.1995.484165>
- Cash, M. A., Habetler, T. G., & Kliman, G. B. (1998). Insulation failure prediction in AC machines using line-neutral voltages. *IEEE Transactions on Industry Applications*, 34(6), 1234–1239. <https://doi.org/10.1109/28.738983>
- Choi, Ki-Young, and S. S. Kim. "Morphological analysis and classification of types of surface corrosion damage by digital image processing." *Corrosion Science* 47.1 (2005): 1-15.
- Consoli, A., Gennaro, F., Raciti, A., & Testa, A. (1998). Fuzzy logic application to pre-fault diagnoses of induction motors. *Proceedings of the 1998 Second IEEE International Caracas Conference on Devices, Circuits and Systems. ICCDCS 98. On the 70th Anniversary of the MOSFET and 50th of the BJT. (Cat. No.98TH8350)*, 249–254. <https://doi.org/10.1109/ICCDCS.1998.705843>
- D. H. Hubel and T. N. Wiesel, Receptive fields and functional architecture of monkey striate cortex, *The Journal of physiology*, 1968.
- Davis, J., Edgar, T., Graybill, R., Korambath, P., Schott, B., Swink, D., ... Wetzel, J. (2015). Smart Manufacturing. *Annual Review of Chemical and Biomolecular Engineering*, 6(1), 141–160. <https://doi.org/10.1146/annurev-chembioeng-061114-123255>
- Diez-Olivan, A., Pagan, J. A., Sanz, R., & Sierra, B. (2017). Data-driven prognostics using a combination of constrained K-means clustering, fuzzy modeling and LOF-based score. *Neurocomputing*, 241, 97–107. <https://doi.org/10.1016/j.neucom.2017.02.024>
- Du, Shouji, et al. "Building change detection using old aerial images and new LiDAR data." *Remote Sensing* 8.12 (2016): 1030.
- Farabet, C. Couprie, L. Najman, and Y. LeCun, Learning hierarchical features for scene labeling, *PAMI*, 2013

- Ghanta, Sindhu, Tanja Karp, and Sangwook Lee. "Wavelet domain detection of rust in steel bridge images." 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2011.
- Gitelson, Anatoly A., et al. "Novel algorithms for remote estimation of vegetation fraction." *Remote sensing of Environment* 80.1 (2002): 76-87.
- Gunatilake, Priyan, et al. "Image understanding algorithms for remote visual inspection of aircraft surfaces." *Machine Vision Applications in Industrial Inspection V*. Vol. 3029. International Society for Optics and Photonics, 1997.
- Guo, L., Li, N., Jia, F., Lei, Y., & Lin, J. (2017). A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing*, 240, 98–109. <https://doi.org/10.1016/j.neucom.2017.02.045>
- Haralick, Robert M., Karthikeyan Shanmugam, and Its' Hak Dinstein. "Textural features for image classification." *IEEE Transactions on systems, man, and cybernetics* 6 (1973): 610-621.
- Hecklau, F., Galeitzke, M., Flachs, S., & Kohl, H. (2016). Holistic Approach for Human Resource Management in Industry 4.0. *Procedia CIRP*, 54, 1–6. <https://doi.org/10.1016/j.procir.2016.05.102>
- Hermann, M., Pentek, T., & Otto, B. (2015). *Design Principles for Industrie 4.0 Scenarios: A Literature Review* (No. 01/2015). Retrieved from http://www.iim.mb.tu-dortmund.de/cms/de/forschung/Arbeitsberichte/Design-Principles-for-Industrie-4_0-Scenarios.pdf
- Holm, Egil, et al. "Classification of corrosion and coating damages on bridge constructions from images using convolutional neural networks." *Twelfth International Conference on Machine Vision (ICMV 2019)*. Vol. 11433. International Society for Optics and Photonics, 2020.
- IBTEHAZ, Nabil; RAHMAN, M. Sohel. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural networks*, v. 121, p. 74-87, 2020.
- Hoskere, Vedhus, et al. "Vision-based structural inspection using multiscale deep convolutional neural networks." *arXiv preprint arXiv:1805.01055* (2018).
- Ibarguren, Aitor, et al. "Thermal tracking in mobile robots for leak inspection activities." *Sensors* 13.10 (2013): 13560-13574.
- K. Fukushima, Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biological cybernetics*, 1980.
- Kardec, A., & Nascif, J. (2012). *Manutenção. Função Estratégica* (4th ed.). QualityMark.

- Koch, Gerhardus, et al. "International measures of prevention, application, and economics of corrosion technologies study." *NACE international* (2016): 216.
- Korchenko, A. G., and O. S. Ilyash. "The generalized classification of unmanned air vehicles." 2013 IEEE 2nd International Conference Actual Problems of Unmanned Air Vehicles Developments Proceedings (APUAVD). IEEE, 2013.
- Krishnamurthy, L., Adler, R., Buonadonna, P., Chhabra, J., Flanigan, M., Kushalnagar, N., ... Yarvis, M. (2005). Design and deployment of industrial sensor networks. *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems - SenSys '05*, 64. <https://doi.org/10.1145/1098918.1098926>
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems* 25 (2012): 1097-1105.
- Laird, R. A., & Sherratt, T. N. (2010). The economics of evolution: Henry Ford and the Model T. *Oikos*, 119(1), 3–9. <https://doi.org/10.1111/j.1600-0706.2009.17613.x>
- Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industrie 4.0. *WIRTSCHAFTSINFORMATIK*, 56(4), 261–264. <https://doi.org/10.1007/s11576-014-0424-4>
- Lasurt, I., Stronach, A. F., & Penman, J. (2000). A fuzzy logic approach to the interpretation of higher order spectra applied to fault diagnosis in electrical machines. *PeachFuzz 2000. 19th International Conference of the North American Fuzzy Information Processing Society - NAFIPS (Cat. No.00TH8500)*, 158–162. <https://doi.org/10.1109/NAFIPS.2000.877411>
- Lee, Sangwook, Luh-Maan Chang, and Mirosław Skibniewski. "Automated recognition of surface defects using digital color image processing." *Automation in Construction* 15.4 (2006): 540-549.
- Li, Bingyao, et al. "Application of LiDAR UAV for High-Resolution Flood Modelling." *Water Resources Management* 35.5 (2021): 1433-1447.
- Li, L. (2018). China's manufacturing locus in 2025: With a comparison of "Made-in-China 2025" and "Industry 4.0." *Technological Forecasting and Social Change*, 135, 66–74. <https://doi.org/10.1016/j.techfore.2017.05.028>
- Livens, Stefan, et al. "Classification of corrosion images by wavelet signatures and LVQ networks." *International Conference on Computer Analysis of Images and Patterns*. Springer, Berlin, Heidelberg, 1995.

- M. Jaderberg, A. Vedaldi, and A. Zisserman, Deep features for text spotting, in ECCV, 2014
- Medeiros, Fátima NS, et al. "On the evaluation of texture and color features for nondestructive corrosion detection." *EURASIP Journal on Advances in Signal Processing* 2010 (2010): 1-7.
- Moranduzzo, Thomas, and Farid Melgani. "Detecting cars in UAV images with a catalog-based approach." *IEEE Transactions on Geoscience and remote sensing* 52.10 (2014): 6356-6367.
- Mosallam, A., Medjaher, K., & Zerhouni, N. (2016). Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing*, 27(5), 1037–1048. <https://doi.org/10.1007/s10845-014-0933-4>
- Mo-Yuen Chow, & Goode, P. V. (1993). The advantages and challenges of machine fault detection using artificial neural network and fuzzy logic technologies. *Proceedings of 36th Midwest Symposium on Circuits and Systems*, 708–711. <https://doi.org/10.1109/MWSCAS.1993.342948>
- Murray, A., & Penman, J. (1997). Extracting useful higher order features for condition monitoring using artificial neural networks. *IEEE Transactions on Signal Processing*, 45(11), 2821–2828. <https://doi.org/10.1109/78.650108>
- Nejjari, H., & Benbouzid, M. E. H. (2000). Monitoring and diagnosis of induction motors electrical faults using a current Park's vector pattern learning approach. *IEEE Transactions on Industry Applications*, 36(3), 730–735. <https://doi.org/10.1109/28.845047>
- Nithin, D. Kanishka and Sivakumar, P. Bagavathi, Generic Feature Learning in Computer Vision, Elsevier, Vol.58, Pages202-209, 2015
- Niu, G., & Jiang, J. (2017). Prognostic control-enhanced maintenance optimization for multi-component systems. *Reliability Engineering & System Safety*, 168, 218–226. <https://doi.org/10.1016/j.ress.2017.04.011>
- Niu, G., & Li, H. (2017). IETM centered intelligent maintenance system integrating fuzzy semantic inference and data fusion. *Microelectronics Reliability*, 75, 197–204. <https://doi.org/10.1016/j.microrel.2017.03.015>
- Olaf Ronneberger, et al. "U-Net: Convolutional Networks for Biomedical Image Segmentation", Springer International Publishing Switzerland 2015.

- Pajares, Gonzalo. "Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs)." *Photogrammetric Engineering & Remote Sensing* 81.4 (2015): 281-330.
- Pakrashi, Vikram, et al. "ROC dependent event isolation method for image processing based assessment of corroded harbour structures." *Structures & Infrastructure Engineering* 6.3 (2010): 365-378.
- Park, Jongho, and Namhoon Cho. "Collision avoidance of hexacopter UAV based on LiDAR data in dynamic environment." *Remote Sensing* 12.6 (2020): 975.
- Penman, J., Sedding, H. G., Lloyd, B. A., & Fink, W. T. (1994). Detection and location of interturn short circuits in the stator windings of operating motors. *IEEE Transactions on Energy Conversion*, 9(4), 652–658. <https://doi.org/10.1109/60.368345>
- Prakash, Anupma. "Thermal remote sensing: concepts, issues and applications." *International Archives of Photogrammetry and Remote Sensing* 33.B1; PART 1 (2000): 239-243.
- Ragab, A., Ouali, M.-S., Yacout, S., & Osman, H. (2016). Remaining useful life prediction using prognostic methodology based on logical analysis of data and Kaplan–Meier estimation. *Journal of Intelligent Manufacturing*, 27(5), 943–958. <https://doi.org/10.1007/s10845-014-0926-3>
- Rodríguez-Canosa, Gonzalo R., et al. "A real-time method to detect and track moving objects (DATMO) from unmanned aerial vehicles (UAVs) using a single camera." *Remote Sensing* 4.4 (2012): 1090-1111.
- Rokonuzzaman, M., & Rahman, M. A. (1995). Neural network based incipient fault detection of induction motors. *Proceedings of IEEE/IAS International Conference on Industrial Automation and Control*, 199–202. <https://doi.org/10.1109/IACC.1995.465853>
- Rudol, Piotr, and Patrick Doherty. "Human body detection and geolocalization for UAV search and rescue missions using color and thermal imagery." 2008 IEEE aerospace conference. Ieee, 2008.
- Rufino, Giancarlo, and Antonio Moccia. "Integrated VIS-NIR hyperspectral/thermal-IR electro-optical payload system for a mini-UAV." *Infotech@ Aerospace*. 2005. 7009.
- S. Fruchtenicht ; E. Pittius ; H.O. Seinsch. (1989). A diagnostic system for three-phase asynchronous machines. *Fourth International Conference on Electrical Machines and Drives Conf.*

- Schoen, R. R., & Habetler, T. G. (1995). Effects of time-varying loads on rotor fault detection in induction machines. *IEEE Transactions on Industry Applications*, 31(4), 900–906. <https://doi.org/10.1109/28.395302>
- Schwab, K. (2017). *The Fourth Industrial Revolution*. Crown.
- Spear, B. (2008). James Watt: The steam engine and the commercialization of patents. *World Patent Information*, 30(1), 53–58. <https://doi.org/10.1016/j.wpi.2007.05.009>
- Sumit Saha (2018). A Comprehensive Guide to Convolutional Neural Networks. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- Suomalainen, Juha, et al. "A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles." *Remote Sensing* 6.11 (2014): 11013-11030.
- Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015). Machine Learning for Predictive Maintenance: A Multiple Classifier Approach. *IEEE Transactions on Industrial Informatics*, 11(3), 812–820. <https://doi.org/10.1109/TII.2014.2349359>
- Thomson, W. T., & Fenger, M. (n.d.). Industrial application of current signature analysis to diagnose faults in 3-phase squirrel cage induction motors. *Conference Record of 2000 Annual Pulp and Paper Industry Technical Conference (Cat. No.00CH37111)*, 205–211. <https://doi.org/10.1109/PAPCON.2000.854227>
- Toshev A. and C. Szegedy, Deep -pose: Human pose estimation via deepneural networks, in CVPR, 2014.
- Viana, H. R. G. (2002). *PCM. Planejamento e Controle da Manutenção* (1st ed.; QualityMark, ed.).
- Vijai, J. P., Somayaji, G. S. R., Swamy, R. J. R., & Aital, P. (2017). Relevance of F.W. Taylor's principles to modern shop-floor practices. *Benchmarking: An International Journal*, 24(2), 445–466. <https://doi.org/10.1108/BIJ-02-2015-0019>
- Wang, K., Dai, G., & Guo, L. (2016). Intelligent Predictive Maintenance (IPdM) for Elevator Service- Through CPS, IOTaS and Data Mining. *Proceedings of the 6th International Workshop of Advanced Manufacturing and Automation*. <https://doi.org/10.2991/iwama-16.2016.1>
- Y. Wan and Q. Xie, "A Novel Framework for Optimal RGB to Grayscale Image Conversion," 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2016, pp. 345-348, doi: 10.1109/IHMSC.2016.201

Yan LeCun, et al. "Gradient-Based Learning Applied to Document Recognition". PROC. OF THE IEEE (1998).

Yao, Huang, Rongjun Qin, and Xiaoyu Chen. "Unmanned aerial vehicle for remote sensing applications—A review." Remote Sensing 11.12 (2019): 1443.

Yao, Yuan, et al. "Artificial intelligence-based hull structural plate corrosion damage detection and recognition using convolutional neural network." Applied Ocean Research 90 (2019): 101823.

Zhang, Xiang, et al. "Using UAV LiDAR to Extract Vegetation Parameters of Inner Mongolian Grassland." Remote Sensing 13.4 (2021): 656.

APÊNDICE A: MDCB das coberturas.

MODO DE FALHA			APARENTE/FUNCIONAL	PRODUÇÃO	PROPAGAÇÃO	QUALIDADE	SEGURANÇA	CUSTO	FREQUENCIA	GRAU TOTAL
COMPONENTE	PROBLEMA	CAUSA								
Viga calha	Obstruído	Acúmulo de sujidade	A							0
Viga calha	Transbordo	Acúmulo de sujidade	A							0
Viga calha	Quebrada	Sobreuso	F	0	3	0	0	2	1	6
Viga calha	Oxidação	Sobreuso	F	0	3	0	0	2	1	6
Viga calha	Infiltração	Agentes externos	A							0
Viga calha	Infiltração	Sobreuso	F	3	3	1	2	2	1	12
Dreno Vertical	Vedação deteriorada	Sobreuso	F	0	3	0	0	2	1	6
Dreno Vertical	Rachado	Sobreuso	F	0	3	0	0	2	1	6
Dreno Vertical	Rachado	Agentes externos	A							0
Dreno Vertical	Obstruído	Acúmulo de sujidade	A							0
Dreno Vertical	Vedação deteriorada	Agentes externos	A							0
Dreno Vertical	Obstruído	Falta de componentes	F	0	3	0	0	2	1	6
Flanges do dreno	Vazamento	Sobreuso	F	0	3	0	0	2	1	6
Grade do dreno	Transbordo	Acúmulo de sujidade	A							0
Terminais de ventilação	Infiltração	Colarinho corroído	F	3	3	2	1	2	1	12
	Infiltração	Má vedação	F	3	3	2	1	1	1	11
	Solto	Má fixação	F	0	3	0	3	0	1	7
	Solto	Oxidação	F	0	3	0	3	0	1	7
Rufo	Quebrado	Chapa deteriorada	F	0	3	0	3	2	1	9
	Oxidação	Sobreuso	F	2	3	0	0	2	2	9

APÊNDICE B: Classificação de priorização das frentes avaliadas.

<u>CRITÉRIOS DE PRIORIZAÇÃO</u>		
PRODUÇÃO		
G	0	Nenhum efeito ou consequência
R	1	Parada do processo \leq 2 horas
A	2	2 horas < Parada do processo \leq 8 horas
U	3	Parada do processo > 8 horas
PROPAGAÇÃO		
G	0	Nenhum efeito ou consequência - Baixo custo (\leq R\$500,00)
R	1	Parada do processo \leq 3 horas - Médio custo (R\$500<Custo \leq R\$5.000)
A	2	3 horas<Parada do processo \leq 8 horas - Alto custo (R\$5.000<Custo \leq
U	3	Parada do processo> 8 horas - Altíssimo custo (>R\$40.000)
QUALIDADE		
G	0	Nenhum efeito ou consequência
R	1	Quantidade de produtos defeituosos \leq 5%
A	2	Quantidade de produtos defeituosos >5%
U	3	Defeito percebido pelo cliente final
SEGURANÇA E MEIO AMBIENTE		
G	0	Nenhum efeito ou consequência
R	1	Danos leves \leq R\$20.000
A	2	R\$20.000 < Danos graves \leq R\$60.000
U	3	Danos gravíssimos > R\$60.000
CUSTO DE REPARO		
G	0	Baixo custo (\leq R\$400)
R	1	Médio custo (R\$400<Custo \leq R\$4.000)
A	2	Alto custo (R\$4.000<Custo \leq R\$40.000)
U	3	Altíssimo custo (>R\$40.000)
FREQUÊNCIA DA OCORRÊNCIA		
GRAU	1	Ocasional (1<Falhas por ano \leq 4)
	2	Provável (4<Falhas por ano \leq 12)
	3	Frequente (>12 Falhas por ano)