

UNIVERSIDADE ESTADUAL DA REGIÃO TOCANTINA MARANHÃO – UEMASUL  
CAMPUS AÇAILÂNDIA  
CURSO DE ENGENHARIA CIVIL BACHARELADO

**ÁTILA MARCONCINE DE SOUZA**

**APLICAÇÃO DE REDES NEURAS CONVOLUCIONAIS NA DETECÇÃO DE  
PATOLOGIAS DO TIPO PANELA EM PAVIMENTOS ASFÁLTICOS**

Açailândia  
2022

**ÁTILA MARCONCINE DE SOUZA**

**APLICAÇÃO DE REDES NEURAIAS CONVOLUCIONAIS NA DETECÇÃO DE  
PATOLOGIAS DO TIPO PANELA EM PAVIMENTOS ASFÁLTICOS**

Monografia apresentada ao Curso de Engenharia Civil da Universidade Estadual da Região Tocantina do Maranhão/Campus Açailândia, para a obtenção do grau de bacharelado em Engenharia Civil.

Orientador: Prof. Me. Lucas Manoel da Silva.  
Co-orientadora: Ma. Alcineide Dutra Pessoa de Sousa

Açailândia

2022

S729a

Souza, Átila Marconcine de

Aplicação de redes neurais convolucionais na detecção de patologias do tipo  
panela em pavimentos asfálticos / Átila Marconcine de Souza – Açailândia:  
UEMASUL, 2022.

66 f. : il.

Monografia (Curso de Bacharel em Engenharia Civil) – Universidade Estadual  
da Região Tocantina do Maranhão – UEMASUL, Açailândia, MA, 2022.

Orientador: Prof. Me. Lucas Manoel da Silva

1. Painela. 2. Pavimento asfáltico. 3. Rede neural convolucional. I. Título.

CDU 625.7:528.854(812.1)

**ÁTILA MARCONCINE DE SOUZA**

**APLICAÇÃO DE REDES NEURAIS CONVOLUCIONAIS NA DETECÇÃO DE  
PATOLOGIAS DO TIPO PANELA EM PAVIMENTOS ASFÁLTICOS.**

Monografia apresentada ao Curso de Engenharia Civil da Universidade Estadual da Região Tocantina do Maranhão/Campus Açailândia, para obtenção do grau de Bacharel em Engenharia Civil.

Aprovado em: 22, 12, 22.

**BANCA EXAMINADORA**

*Lucas Manoel da Silva.*

**Prof Meº Lucas Manoel da Silva (Orientador)**  
Mestrado Profissional em Engenharia de Barragens e Meio Ambiente  
Universidade Estadual da Região Tocantina do Maranhão

*Jéssica Almeida dos Santos*

**Prof Maº Jéssica Almeida dos Santos**  
Mestrado em Engenharia da Eletricidade  
Universidade Estadual da Região Tocantina do Maranhão

*Gean Carlos Lopes de Sousa*

**Prof Drº Gean Carlos Lopes de Sousa**  
Doutorado em Engenharia Elétrica  
Universidade Federal do Maranhão

## **AGRADECIMENTOS**

À minha mãe, Maria Aparecida Conceição Marconcini Prestes, e ao meu pai, Samuel Pereira de Souza, pelo incentivo, apoio e suporte, hoje e sempre.

À Universidade Estadual da Região Tocantina do Maranhão (UEMASUL) por através do seu Programa Institucional de Bolsas de Extensão (PIBEXT) ter apoiado financeiramente a pesquisa que resultou neste presente trabalho.

Ao meu orientador, Lucas Manoel da Silva, pelas orientações dadas ao decorrer dos dois anos de bolsa extensão.

À minha coorientadora, Alcineide Dutra Pessoa de Sousa e ao Dr. Gean Carlos Lopes de Sousa, pelas sugestões e por despertarem meu interesse na área da inteligência artificial.

Aos meus amigos, pelos momentos de descontração ao longo desses 5 anos de curso.

“A vida é perene como a infância”

(Silvério Pessoa)

## RESUMO

Um dos defeitos mais comuns em pavimentos asfálticos são os buracos (panelas), e a detecção desse tipo de patologia é um componente crítico da manutenção e reabilitação do pavimento. O presente trabalho visa o treinamento de uma Rede Neural Convolutiva (RNC) que seja capaz de identificar a patologia do tipo panela em pavimentos asfálticos. Para isso foi necessário a montagem de dois bancos de dados (DT1 e DT2), com imagens dessa patologia coletadas através da câmera GoPro Hero 7. Em seguida, foi realizado o pré-processamento das imagens coletadas, por meio do *software* Labellmg. Por fim, foi realizado o treinamento da RNC no ambiente de execução Google Colab, utilizando a linguagem Python, o algoritmo YoloV4 e o *framework* Darknet. Os resultados de treinamento obtidos foram satisfatórios para os dois bancos de dados utilizados. Entretanto, notou-se um desempenho melhor, em todas as métricas de avaliação do DT2, que continham a maior quantidade de imagens, com a métrica de precisão atingindo 0,9, o *recall* 0,89 o *F1-score* 0,88, o IoU 71,15% e o mAP 91,68%. Para validação dos resultados do treinamento foi realizada uma filmagem na cidade de Imperatriz-MA e, dessa forma, foi perceptível que a utilização na prática de um modelo neuronal para detecções de panelas se mostrou funcional, apesar de não obter resultados tão satisfatórios quanto o conjunto de treinamento, atingindo uma precisão de 73,29% no limiar 0,9 e um *recall* de 79,77% no limiar 0,5. Esse resultado em conjunto com as exposições dos dados de telemetria, facilitam a manutenção de pavimentos asfálticos, visto que ao final do processo é obtido um modelo que detecta buracos e indica a sua localização. Ressalta-se que é possível treinar a mesma RNC para detectar outras patologias, tornando assim o processo de avaliação de pavimentos mais completo.

**Palavras-chave:** Panela, Pavimento asfáltico, Rede Neural Convolutiva, YOLOv4.

## ABSTRACT

One of the most common defects in asphalt pavements are the potholes, and the detection of this type of distress is a critical component of pavement maintenance and rehabilitation. The present work has the objective of training a Convolutional Neural Network (CNN) that is capable of identifying potholes in asphalt pavements. For this, it was necessary making two databases (DT1 and DT2) with images of this pathology collected through GoPro Hero 7 camera. Then, the pre-processing of the collected images was done through Labelling software. Finally, the training of the CNN was carried out in Google Colab, using the programming language python, YOLOv4 algorithm and darknet framework. The training results obtained were excellent for the two databases used. However, a better performance was noted, in all evaluation metrics, of DT2, which contained the largest number of images, with the precision metric reaching 0.9, recall reaching 0.89, F1-score 0.88, IoU 71.15% and mAP 91.68%. To validate the results of the training, a video recording was made in the city of Imperatriz-MA and, thus, was noticeable that the practical use of a neuronal model for pothole detection proved to be functional, despite not obtaining as satisfactory results as the training set, reaching good precision (73.29%) at threshold 0.9 and good recall (79.77%) at threshold 0.5. This good functionality, with the telemetry data exhibition facilitate the maintenance of asphalt pavements, since at the end of the process a model is obtained that detects holes and indicates their location. It should be noted that it is possible to train the same CNN to detect other pavement distress, thus making the pavement evaluation process more complete.

**Keywords:** Pothole, Asphalt pavements, Convolutional Neural Network, YOLOv4

## LISTA DE ILUSTRAÇÕES

Figura 1 – Disposição das camadas de um pavimento.....	11
Figura 2 – Deflexão no pavimento causando tração e compressão.....	11
Figura 3 – Deformação elástica pela carga no pavimento flexível.....	12
Figura 4 – Deformação pela carga no pavimento rígido.....	13
Figura 5 – Camadas de um pavimento.....	14
Figura 6 – Termos aplicáveis a camadas de revestimento asfáltico.....	14
Figura 7 – Patologia do tipo panela.....	16
Figura 8 – Profundidade das panelas em pavimentos.....	17
Figura 9 – Curva de desempenho de um pavimento.....	18
Figura 10 – Relação entre IA, aprendizado de máquina e aprendizado profundo.....	20
Figura 11 – Estrutura de um neurônio.....	21
Figura 12 – Conexões sinápticas entre neurônios.....	22
Figura 13 – Neurônio de McCulloch & Pitts.....	22
Figura 14 – Neurônio biológico e Neurônio artificial.....	23
Figura 15 – Estrutura de uma rede neural artificial.....	23
Figura 16 – Rede neural biológica e artificial.....	24
Figura 17 – Rede neural simples e rede neural profunda.....	24
Figura 18 – Imagem sendo dividida em padrões locais.....	25
Figura 19 – Estrutura de uma RNC.....	25
Figura 20 – Tarefas da Visão Computacional.....	26
Figura 21 – Detecção de objetos aplicada na patologia panela.....	26
Figura 22 – Matriz de confusão.....	27
Figura 23 – Matriz de confusão aplicada em imagens com e sem buracos.....	28
Figura 24 – Ground-truth bounding box e bounding box prevista.....	29
Figura 25 – Determinação do IoU.....	30
Figura 26 – IoU de bounding boxes.....	30
Gráfico 1 – Exemplo de uma curva precisão/recall.....	31
Figura 27 – Passo a passo do TCC.....	32
Figura 28 – Câmera GoPro Hero 7 acoplada no capô do veículo.....	33
Figura 29 – Tecnologia HyperSmooth.....	33
Figura 30 – Frame de vídeo gravado com a câmera GoPro Hero 7.....	34
Figura 31 – Utilização do aplicativo GoPro Quik.....	34

Figura 32 – Informações de telemetria presentes em um vídeo.....	35
Figura 33 – Ferramenta Labellmg.....	36
Figura 34 – Bloco de notas contendo as informações das bounding boxes.....	36
Figura 35 – Elementos do arquivo de texto.....	36
Figura 36 – Representação dos elementos do arquivo de texto.....	37
Figura 37 – Exemplo de uma parte do banco de dados.....	37
Figura 38 – Especificações técnicas da nVidia Tesla T4.....	38
Gráfico 2 – Resultados: Precisão x Épocas.....	41
Gráfico 3 – Resultados: <i>Recall</i> x Épocas.....	42
Gráfico 4 – Resultados: <i>F1-score</i> x Épocas.....	42
Gráfico 5 – Resultados: IoU x Épocas.....	43
Gráfico 6 – Resultados: mAP x Épocas.....	44
Figura 39 – Comparação das <i>bounding boxes</i> .....	44
Figura 40 – Percurso da Gravação.....	45
Figura 41 – <i>Threshold</i> diferentes aplicados na mesma imagem.....	46
Figura 42 – Valores de <i>threshold</i> muito baixos (0,0001) .....	46
Figura 43 – Valores de <i>threshold</i> muito altos.....	47
Figura 44 – Métricas de avaliação para o limiar 0,5.....	48
Figura 45 – Painela com 53% de confiabilidade.....	48
Figura 46 – Mancha detectada erroneamente.....	49
Figura 47 – Objeto fora da pista detectado erroneamente.....	49
Figura 48 – Remendo detectado erroneamente.....	49
Figura 49 – Métricas de avaliação para o limiar 0,7.....	50
Figura 50 – Painela com confiabilidade de 89%.....	50
Figura 51 – Métricas de avaliação para o limiar 0,9.....	51
Gráfico 7 – Métricas de Avaliação no vídeo.....	51
Figura 52 – Pedra detectada com alta confiabilidade.....	52
Figura 53 – Mancha detectada com alta confiabilidade.....	52
Figura 54 – Poça de água detectada com alta confiabilidade.....	53
Figura 55 – Painela detectada com 99% de confiabilidade.....	53
Figura 56 – Painela não detectada no limiar 0,5.....	54

## LISTA DE EQUAÇÕES

Equação 1 – Precisão.....	28
Equação 2 – <i>Recall</i> .....	28
Equação 3 – <i>F1-score</i> .....	29
Equação 4 – <i>Average Precision (AP)</i> .....	30

## LISTA DE TABELAS

Tabela 1 – Divisão das imagens dos bancos de dados.....	39
Tabela 2 – Divisão das painelas dos bancos de dados.....	39
Tabela 3 – Matriz de confusão para o limiar 0.5.....	47
Tabela 4 – Matriz de confusão para o limiar 0.7.....	48
Tabela 5 – Matriz de confusão para o limiar 0.9.....	50

## LISTA DE ABREVIATURAS

AP – Average Precision

CNT – Confederação Nacional de Transportes

CNN – Convolutional Neural Network

CPU – Central Processing Unit

CUDA – Compute Unified Device Architecture

DNIT – Departamento Nacional de Infraestrutura de Transportes

DT1 – DataSet1

DT2 – DataSet2

FN – False Negative

FP – False Positive

GPU – Graphics Processing Unit

GPS – Global Positioning System

HD – Hard Disk

IA – Inteligência Artificial

IDE – Integrated Development Environment

IoU – Intersection over Union

mAP – mean Average Precision

RNA – Rede Neural Artificial

RNC – Rede Neural Convolutacional

TN – True Negative

TP – True Positive

YOLO – You Only Look Once

## SUMÁRIO

<b>RESUMO.....</b>	<b>III</b>
<b>ABSTRACT.....</b>	<b>IV</b>
<b>LISTA DE ILUSTRAÇÕES .....</b>	<b>V</b>
<b>LISTA DE EQUAÇÕES .....</b>	<b>VII</b>
<b>LISTA DE TABELAS .....</b>	<b>VIII</b>
<b>LISTA DE ABREVIATURAS.....</b>	<b>IX</b>
<b>1 INTRODUÇÃO.....</b>	<b>8</b>
<b>1.1 OBJETIVOS.....</b>	<b>9</b>
<b>1.1.1 Geral .....</b>	<b>9</b>
<b>1.1.2 Específicos.....</b>	<b>9</b>
<b>1.2 ESTRUTURA DO TRABALHO.....</b>	<b>9</b>
<b>2 FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>11</b>
<b>2.1 OS PAVIMENTOS .....</b>	<b>11</b>
<b>2.1.1 Camadas dos pavimentos .....</b>	<b>13</b>
<b>2.1.2 Patologias em pavimentos .....</b>	<b>15</b>
<b>2.1.3 Avaliações da condição de desempenho dos pavimentos.....</b>	<b>17</b>
<b>2.1.4 Normas do DNIT para avaliação funcionais de pavimentos .....</b>	<b>18</b>
<b>2.2 AS REDES NEURAIS.....</b>	<b>19</b>
<b>2.2.1 Redes neurais biológicas .....</b>	<b>19</b>
<b>2.2.2 Redes Neurais Artificiais .....</b>	<b>21</b>
<b>2.2.3 Redes Neurais Convolucionais .....</b>	<b>23</b>
<b>3 MATERIAIS E MÉTODOS .....</b>	<b>31</b>
<b>3.1 DESENHO DO ESTUDO .....</b>	<b>31</b>
<b>3.2 MONTAGEM DA BASE DE DADOS.....</b>	<b>31</b>
<b>3.3 PRÉ-PROCESSAMENTO DAS IMAGENS .....</b>	<b>34</b>
<b>3.4 PROGRAMAÇÃO DA REDE NEURAL .....</b>	<b>37</b>
<b>4 RESULTADOS E DISCUSSÕES.....</b>	<b>40</b>
<b>4.1 RESULTADOS DO TREINAMENTO .....</b>	<b>40</b>
<b>4.2 TESTE DA RNC EM UM VÍDEO .....</b>	<b>44</b>
<b>5 CONCLUSÃO .....</b>	<b>54</b>
<b>6 REFERÊNCIAS.....</b>	<b>56</b>

## 1 INTRODUÇÃO

De acordo com Balbo (2007, p.14) o homem, a fim de obter melhor acesso às áreas cultiváveis e às fontes de madeira, rochas, minerais e água, além do desejo de expandir sua área ou território de influência, criou o que chamamos de estradas, cujas lembranças mais remotas provêm da China – sendo provavelmente o país que as inventou. Mais tarde, os romanos as aperfeiçoaram, instalando pavimentos com drenagem, com o intuito de torná-los mais duradouros.

O principal objetivo da pavimentação, para a Confederação Nacional de Transportes, CNT (2017, p.61) é garantir a trafegabilidade em qualquer época do ano e sob quaisquer condições climáticas, além de proporcionar conforto no rolamento e na segurança aos usuários.

“Do ponto de vista do usuário, o estado da superfície do pavimento é o mais importante, pois os defeitos ou irregularidades nessa superfície são percebidos, uma vez que afetam seu conforto” (BERNUCCI et al. 2008, p.403).

Os defeitos ou patologias de superfície dos pavimentos para a CNT (2022, p.48) são os danos ou as deteriorações que podem ser identificados a olho nu. Um desses defeitos são os buracos, que também podem ser chamados de panelas.

As panelas conforme Chen, Yao, Gu (2020, p.1, tradução nossa) são buracos em forma de tigela na superfície do pavimento e são um dos tipos mais comuns de patologias nas estradas. De acordo com Gupta *et al.* (2020, p.2, tradução nossa) a sua presença em pavimentos danificam o amortecedor e o sistema de suspensão dos veículos e, na pior das hipóteses, podem causar acidentes rodoviários fatais.

Em 2022 a pesquisa CNT de rodovias apontou que “70,2% do pavimento analisado no país possui algum tipo de defeito” (CNT. 2022, p.200). Essa constatação demonstra a gravidade do problema no que tange ao déficit de infraestrutura rodoviária do Brasil.

Chen, Yao, Gu (2020, p.1, tradução nossa) afirmam que a detecção de buracos é um componente crítico da manutenção e da reabilitação do pavimento devido ao seu efeito negativo na qualidade do mesmo e na segurança do motorista. No Brasil, hoje, esse processo de detecção é normatizado pelo Departamento Nacional de Infraestrutura de Transportes (DNIT), sendo realizado através de

avaliações subjetivas e objetivas, através das normas: DNIT 005/2003 - TER, DNIT 006/2003-PRO, DNIT 007/2003-PRO, DNIT 008/2003-PRO e DNIT 009/2003-PRO.

Esta pesquisa se justifica, pois atualmente os métodos normatizados pelo DNIT para o inventariamento de defeitos em pavimentos flexíveis são obsoletos, morosos (lentos) e ineficientes, além de dependerem de mão de obra humana.

Partindo desta explanação, este trabalho levanta o seguinte problema: Existe a possibilidade de utilizar uma Rede Neural Convolucional (RNC) para detectar os defeitos do tipo panela em pavimentos asfálticos? Essa metodologia proposta é mais precisa e eficiente que os métodos de avaliação prescritos nas normas atuais de serventia de pavimentos?

## **1.1 Objetivos**

### **1.1.1 Geral**

O presente trabalho visa o treinamento, o teste e a validação de uma Rede Neural Convolucional que seja capaz de identificar a patologia do tipo panela em pavimentos asfálticos utilizando a arquitetura YOLOv4 (You Only Look Once) e o *framework* Darknet.

### **1.1.2 Específicos**

- Estudar os principais conceitos de inteligência artificial, aprendizado de máquina e visão computacional;
- Criar um banco de imagens personalizado, retirando os frames de gravações realizadas através de uma câmera GoPro Hero 7;
- Treinar um modelo neuronal com o banco de imagens criado;
- Validar o modelo treinado em ruas distintas com o intuito de analisar sua precisão;

## **1.2 Estrutura do trabalho**

Os demais capítulos desta monografia abordarão os seguintes tópicos:

- Na Fundamentação teórica serão abordados, com base na literatura, os principais tópicos do tema proposto, ou seja, pavimentos e redes neurais.

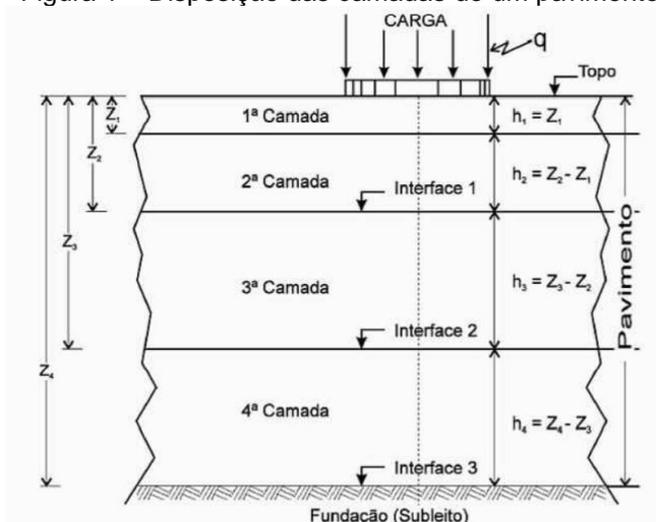
- Nos materiais e métodos será descrito o processo metodológico para a montagem da base de dados, pré-processamento das imagens, além da descrição do processo de treinamento e teste da RNC.
- Nos resultados e discussões serão apresentados os resultados para o treinamento da RNC, bem como a sua aplicação, em uma filmagem realizada na cidade de Imperatriz-MA.
- Por fim, no último tópico será discorrido sobre as conclusões obtidas a partir dos resultados.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 Os Pavimentos

O pavimento de acordo com Branco (2016, p.72) é uma estrutura composta de diversas camadas que apresentam espessuras finitas construída sobre uma superfície de fundação, intitulado subleito. Essa estrutura tem múltiplas camadas pois conforme Mallick e El- Korchi (2018, p.4, tradução nossa) a sua função é distribuir a carga na superfície e reduzir sua intensidade com a profundidade, de forma que a pressão no subleito seja muito menor que a pressão na superfície. A Figura 1 ilustra o sistema de diversas camadas que constituem os pavimentos.

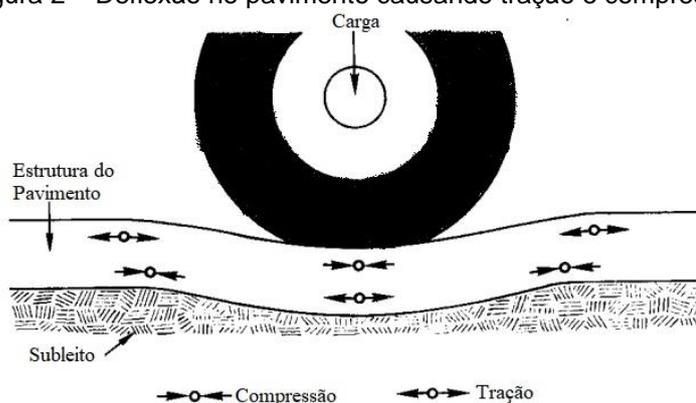
Figura 1 – Disposição das camadas de um pavimento



Fonte: Senço (2008, p.7)

A estrutura do pavimento, de acordo com Branco (2016, p.72) tem a função de resistir aos esforços provenientes do tráfego e do clima, além de propiciar aos usuários melhoria nas suas condições de utilização, provendo assim conforto, economia e segurança. A Figura 2 ilustra como o carregamento da roda deforma ligeiramente a estrutura do pavimento, causando tensões de tração e de compressão.

Figura 2 – Deflexão no pavimento causando tração e compressão

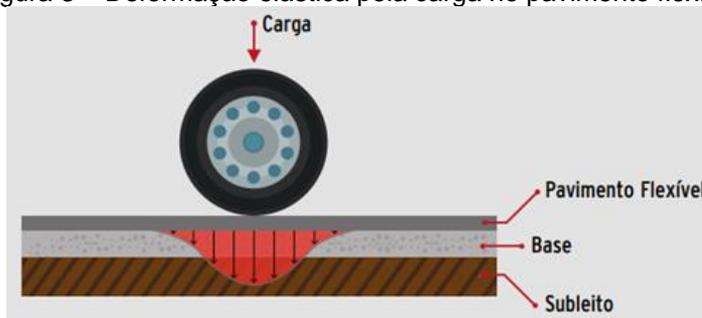


Fonte: Adaptado de Asphalt Institute (1989, p.21)

“Os tipos de pavimentos são classificados, geralmente, como pavimentos flexíveis ou rígidos, sendo que, na composição do primeiro, emprega-se usualmente material asfáltico e, na do segundo, concreto” (CNT, 2017, p.64). No Brasil, segundo a CNT (2022, p.87), “estima-se que cerca de 99% das rodovias são construídas em pavimento flexível, ou seja, o material principal do seu revestimento é o asfalto”.

Os pavimentos flexíveis são aqueles em que todas as camadas sofrem deformação elástica significativa sob o carregamento aplicado e, desse modo, a carga se distribui em parcelas equivalentes entre as camadas (DNIT, 2005, p.50). Balbo (2007, p.50) afirma que uma dada carga atuante no pavimento flexível impõe nessa estrutura um campo de tensões muito concentrado nas proximidades no ponto de aplicação da carga. Na Figura 3 é possível observar o efeito da carga da roda de um carro sobre o pavimento flexível.

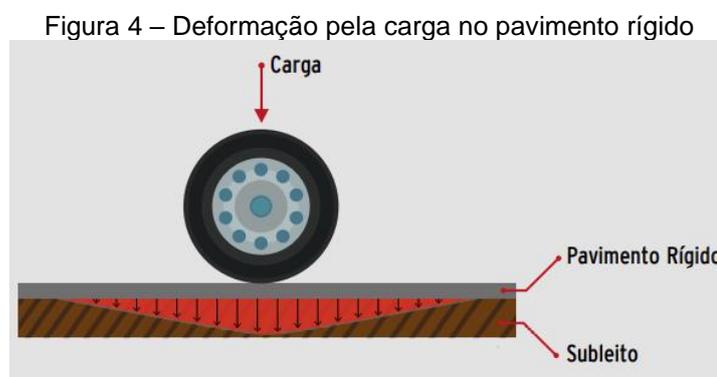
Figura 3 – Deformação elástica pela carga no pavimento flexível



Fonte: CNT (2017, p.25)

Já o pavimento rígido, de acordo com DNIT (2005, p.50) é aquele em que o revestimento tem uma rigidez maior quando comparado com às camadas inferiores, e, por consequência, absorve quase todas as tensões provenientes da carga aplicada. Existem diversos tipos de pavimentos rígidos, embora eles sejam comumente

associados ao concreto de cimento Portland (CNT, 2017, p.28). Na Figura 4 nota-se como esse tipo de pavimento se comporta quando aplicado um carregamento.



Fonte: CNT (2017, p.28)

Entretanto, Bernucci et al. (2008, p.9) afirma que existe uma tendência de utilizar a nomenclatura “pavimentos de concreto de cimento Portland”, ou apenas concreto-cimento, ao invés de “pavimento rígido” e “pavimentos asfálticos” ao invés de pavimentos flexíveis”.

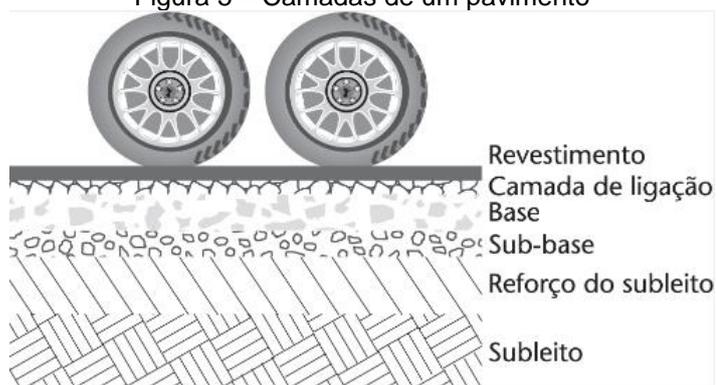
Os termos rígido e flexível dizem respeito às respostas estruturais dos materiais isoladamente, e mais importante ainda, às respostas estruturais do pavimento como um todo. Podendo assim, esses termos serem bastante inadequados para descrever ou classificar um pavimento quando se toma o revestimento como referência [...]. Quando se empregam as expressões “pavimento asfáltico” ou “pavimento de concreto”, delimita-se com precisão qual o tipo de material presente no revestimento (visualmente passível de classificação), não se depreende da expressão, contudo, qual seria o comportamento mecânico oferecido pela estrutura de pavimento (BALBO, 2007, p.62).

Para este trabalho o tipo de pavimentação abordada será a pavimentação asfáltica, não sendo utilizado o termo pavimentação flexível ao longo do texto, uma vez que não será avaliado o comportamento estrutural dos pavimentos.

### 2.1.1 Camadas dos pavimentos

A Figura 5 ilustra as camadas de um pavimento, que são: revestimento, camada de ligação, base, sub-base, reforço do subleito e subleito.

Figura 5 – Camadas de um pavimento



Fonte: Balbo (2007, p.40)

Balbo (2007, p.41) afirma que os revestimentos asfálticos são frequentemente subdivididos em duas ou mais camadas por razões técnicas, construtivas e de custo. Na Figura 6 é possível observar os termos aplicáveis a camadas de revestimento asfáltico e seus respectivos conceitos.

Figura 6 – Termos aplicáveis a camadas de revestimento asfáltico

DESIGNAÇÃO DO REVESTIMENTO	DEFINIÇÃO	ASSOCIAÇÕES
Camada de rolamento	É a camada superficial do pavimento, diretamente em contato com as cargas e com ações ambientais	Camada de desgaste, capa de rolamento, revestimento
Camada de ligação	É a camada intermediária, também em mistura asfáltica, entre a camada de rolamento e a base do pavimento	Camada de <i>binder</i> ou simplesmente <i>binder</i>
Camada de nivelamento	Em geral, é a primeira camada de mistura asfáltica empregada na execução de reforços (recapeamento), cuja função é corrigir os desníveis em pista, afundamentos localizados, enfim, nivelar o perfil do greide para posterior execução da nova camada de rolamento	Camada de reperfilagem ou simplesmente reperfilagem
Camada de reforço	Nova camada de rolamento, após anos de uso do pavimento existente, executada por razões funcionais, estruturais ou ambas	“Recape” e recapeamento são termos populares (usa-se também a expressão “pano asfáltico”, que muitas vezes parece comprometer menos)

Fonte: Balbo (2007, p.41)

A camada de base tem a função de resistir e de distribuir os esforços verticais provenientes do tráfego de veículos. É sobre a base que é construído o revestimento. (CNT, 2017, p.18). Já a sub-base, é a camada construída abaixo da base para complementá-la, ela apresenta as mesmas funções desta e é executada quando, por razões econômicas, for necessário reduzir a espessura da base (DNIT, 2017, p.288).

O reforço do subleito é a camada executada quando por motivos técnicos ou econômicos, se faz conveniente com reduzir a espessura da sub-base (DNIT

(2017, p.254). O subleito é a camada responsável por absorver de forma definitiva os esforços causados pelo tráfego, ela é constituída de material natural consolidado e compactado (CNT, 2017, p.16).

### **2.1.2 Patologias em pavimentos**

Existem diversas patologias em pavimentos asfálticos, segundo Mallick e El- Korchi (2018, p.265, tradução nossa) cada uma é ligada a um motivo específico (como dosagem ruim de materiais, construção ou condições ambientais) ou uma combinação deles, e a maioria ocorre como resultado do tráfego.

Os defeitos que ocorrem em pavimentos flexíveis e semirrígidos definidos pela norma DNIT 005/2003-TER são: as fendas, que podem ser divididas em fissuras, trincas, trincas isoladas e trincas interligadas, os afundamentos, que podem ser plástico e de consolidação, a ondulação, também conhecida como corrugação, o escorregamento, a exsudação, o desgaste, a panela (ou buraco) e os remendos, que são classificados em profundos e superficiais.

A detecção dos defeitos nos estágios iniciais é uma das tarefas mais importantes da manutenção de pavimentos (JÚNIOR; ODA; ZERBINI, 2006, p.3). Para que o pavimento mantenha um nível de serviço adequado quanto ao rolamento, à segurança e à capacidade de suporte do carregamento proveniente do tráfego, são indispensáveis a realização de intervenções de restauração (CNT, 2017, p.82).

As trincas e outras fraturas presentes no revestimento dos pavimentos, que inicialmente quase não são percebidas pelos usuários, podem evoluir de forma rápida para outras patologias e causar sérios problemas se não forem corrigidas (JÚNIOR; ODA; ZERBINI, 2006, p.3). Conforme Silva (2008, p.34) as trincas, afundamentos e desgastes podem evoluir para a patologia do tipo panela, popularmente conhecida como buraco. Esse tipo de defeito é o foco do presente trabalho.

#### **2.1.2.1 Painelas (ou buracos)**

As painelas, conforme notável na Figura 7, são cavidades que possuem dimensões e profundidades variadas. (DNIT, 2006, p.67). Esse tipo de patologia ocorre em qualquer parte do revestimento do pavimento, sobretudo nas trilhas de roda (DOMINGUES,1993, p.40).

Figura 7 – Patologia do tipo panela

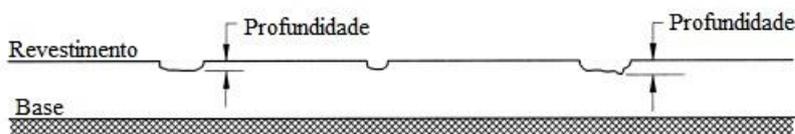


Fonte: Autoria própria

As principais causas desse defeito, de acordo com o CNT (2017, p.93) são: trincas de fadiga, desintegração localizada na superfície do pavimento, deficiência na compactação, umidade excessiva em camadas de solo e falha na imprimção.

“A profundidade do buraco é a profundidade máxima abaixo da superfície do pavimento” (MILLER; BELLINGER, 2014, p.8, tradução nossa). Essa profundidade pode ser visualizada na Figura 8.

Figura 8 – Profundidade das panelas em pavimentos



Fonte: Adaptado de Miller e Bellinger (2014, p.8)

Miller e Bellinger (2014, p.8, tradução nossa) classificam os buracos em 3 níveis de severidade, que são: baixo (com menos de 25mm de profundidade), moderado (entre 25 e 50mm de profundidade) e alto (acima de 50mm).

Os buracos, de acordo com Júnior, Oda, Zerbini (2006, p.55), devem ser imediatamente reparados, pois comprometem a segurança e o conforto e aumentam os custos operacionais.

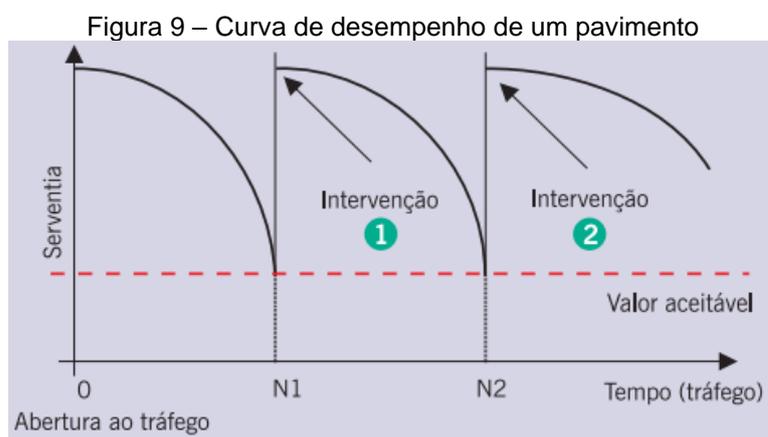
Para reparar esse tipo de patologia são utilizados os remendos que, de acordo Júnior, Oda, Zerbini (2006, p.55), é o método de reparo mais utilizado na manutenção de rodovias e ruas porque todos os pavimentos, uma hora ou outra, vão apresentar buracos. Porém, apesar de ser uma atividade de conservação, o remendo é considerado uma patologia, pelo fato de apontar um local de fragilidade do

revestimento e também por afetar o conforto do usuário (BERNUCCI et al., 2008, p.416)

### 2.1.3 Avaliações da condição de desempenho dos pavimentos

“A capacidade de um pavimento servir satisfatoriamente ao tráfego durante um dado período é o seu desempenho, que pode ser interpretado como a variação da serventia com o tempo e/ou tráfego” (JÚNIOR; ODA; ZERBINI, 2006, p.8).

Na figura 9 é notável que a serventia decresce conforme o tempo e/ou o tráfego aumentam, sendo necessária a realização de intervenções para que a serventia aumente novamente e fique acima de um valor aceitável para trafegabilidade.



Fonte: Bernucci et al (2008, p.442)

Desta forma, se faz necessário realizar avaliações para averiguar a condição de desempenho dos pavimentos. Essas avaliações podem ser funcionais ou estruturais e ambas são fundamentais para orientar sobre as intervenções necessárias, sobretudo quando utilizadas em conjunto (CNT, 2017, p.61)

Na avaliação funcional é verificada a condição da superfície do pavimento, através do levantamento e da análise de defeitos superficiais, além das irregularidades longitudinais (BERNUCCI et al, 2008, p.463).

Já a avaliação estrutural, de acordo com DNIT (2017, p.36) “tem como objeto avaliar a estrutura do pavimento. Ela é realizada com os equipamentos Viga Benkelman e/ou Falling Weight Deflectometer (FWD) que registram as deflexões geradas por diferentes níveis de cargas”.

Tendo em vista que o objeto de estudo desse presente trabalho é o defeito do tipo panela, o foco será nas avaliações funcionais. Existem duas técnicas para esse tipo de avaliação, as subjetivas e as objetivas.

As subjetivas são o tipo de avaliação que permite estabelecer o estado de degradação da superfície do pavimento através de conceitos qualitativos, enquanto as avaliações objetivas têm a mesma função, entretanto esta utiliza a quantificação numérica e a distribuição da severidade dos defeitos presentes na camada de revestimento (CNT, 2017, p.62).

#### **2.1.4 Normas do DNIT para avaliação funcionais de pavimentos**

As avaliações funcionais (objetivas e subjetivas) são normatizadas pelo DNIT, através das normas: DNIT 005/2003 - TER, DNIT 006/2003-PRO, DNIT 007/2003-PRO, DNIT 008/2003-PRO e DNIT 009/2003-PRO.

A norma DNIT 005/2003 – TER (2003a, p.1), para DNIT (2006, p.59) tem o objetivo de definir os termos empregados em patologias que ocorrem na superfície dos pavimentos asfálticos através de uma codificação assumida e uma classificação adotada conforme a gravidade de ocorrência.

A Norma DNIT 006/2003-PRO (DNIT, 2003b), conforme DNIT (2006, p.69) tem o objetivo de apresentar a sistemática de cálculo do Índice de Gravidade Global (IGG), que consiste em um índice combinado de falhas. Para tal, é utilizada uma treliça de alumínio padronizada e materiais auxiliares, como: trena, giz, caneta, formulários, dentre outros.

Bernucci et al. (2008, p.424) afirma que a norma DNIT 007/2003 – PRO (DNIT, 2003c) geralmente é empregada para sistemas de gerência de manutenção. Já com a utilização da norma DNIT 008/2003 – PRO (DNIT, 2003d), segundo DNIT (2006, p.74) é possível a determinação da porcentagem de área com trincamento e de outros defeitos e do número de painéis e remendos, além da determinação dos índices: ICPF – Índice da Condição do Pavimento Flexível, IGGE – Índice de Gravidade Global Expedito e IES – Índice do Estado de Superfície do Pavimento

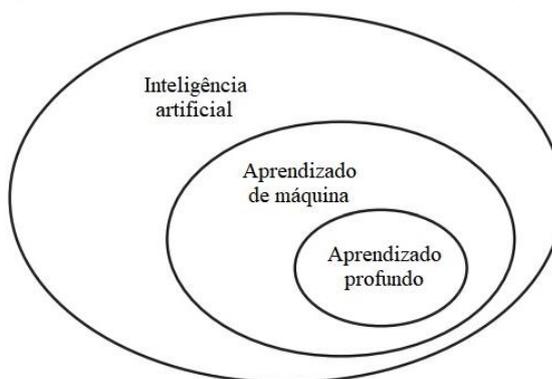
Por fim, a norma DNIT 009/2003 – PRO (DNIT, 2003e) é um procedimento para a avaliação subjetiva que conforme DNIT (2006, p.76) determina o VSA (Valor da Serventia Atual) que afere, de forma subjetiva, as qualidades de suavidade e conforto e rolamento proporcionado pelo pavimento existente.

## 2.2 As Redes Neurais

Taylor (2017, p.13, tradução nossa) afirma que as redes neurais fazem parte do *Deep Learning* (aprendizado profundo), que é um ramo do *Machine Learning* (aprendizado de Máquina) que se mostrou valioso para resolver problemas difíceis, como reconhecer objetos em imagens e processamento de linguagem.

“Já o aprendizado de máquina é um subcampo da inteligência artificial (IA) que estuda a capacidade de melhorar o desempenho com base na experiência”. (RUSSELL; NORVIG, 2021, p.31, tradução nossa). Na Figura 10 é possível observar de forma sucinta como a inteligência artificial, o aprendizado de máquina e o aprendizado profundo se relacionam.

Figura 10 – Relação entre IA, aprendizado de máquina e aprendizado profundo



Fonte: Adaptado de Chollet (2018, p.4)

“A inteligência artificial nasceu na década de 1950, quando os pioneiros do recém-nascido campo da Ciência da Computação começaram a se perguntar se os computadores poderiam ser feitos para ‘pensar’” (CHOLLET, 2018, p.4, tradução nossa).

O aprendizado de máquina, de acordo com Chollet (2018, p.6, tradução nossa), começou a “florescer” apenas na década de 1990, e rapidamente se tornou o subcampo de IA mais popular e bem-sucedido, uma tendência impulsionada pela disponibilidade de *hardwares* mais rápidos e conjuntos de dados maiores.

A área do *Deep Learning*, conforme Hollerweger (2019, p.13), tem origens antigas, assim como o aprendizado de máquina de forma geral, porém tal área se tornou popular na última década devido ao advento de algoritmos mais eficientes para o treinamento de redes neurais.

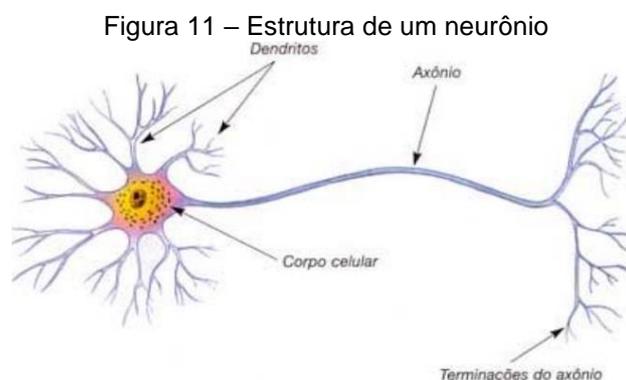
### 2.2.1 Redes neurais biológicas

Os modelos de redes neurais, conforme Müller, Reinhardt, Strickland (1995, p.13) são algoritmos para tarefas cognitivas, como aprendizagem e otimização, que são, em amplo sentido, baseados em conceitos derivados de pesquisas sobre a natureza do cérebro.

Embora as redes neurais apresentem algumas semelhanças com o cérebro humano, elas não pretendem modelá-lo. O cérebro humano é muito mais complexo e, infelizmente, uma grande parte das suas funções cognitivas é ainda hoje desconhecida (BOTERO, 2018, p.15).

“O cérebro humano contém em torno de 100 bilhões ( $10^{11}$ ) de neurônios, sua célula fundamental. Cada um destes neurônios processa informações e se comunica com milhares de outros continuamente e em paralelo” (BRAGA; LUDERMIR; CARVALHO, 2000, p.4).

O neurônio é a célula elementar do sistema nervoso cerebral, ele tem a função de conduzir impulsos sob determinadas condições de operações (SILVA; SPATTI; FLAUZINO, 2010, p.29). Como é possível observar na Figura 11, um neurônio é composto por dendritos, corpo celular e axônio.



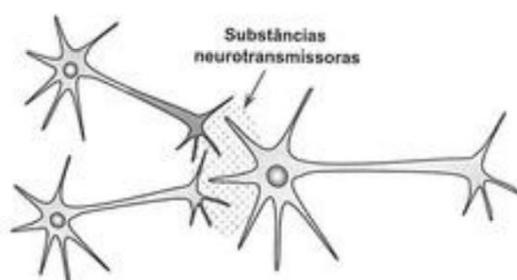
Fonte: Paula e Silva (2010, p.53)

A principal função dos dendritos para Silva, Spatti, Flauzino (2010, p.29), é em captar, de forma contínua, os impulsos de outros neurônios ou do meio externo onde eles podem estar em contato. O corpo celular, que também pode ser chamado de soma, conforme Paula e Silva (2010, p.52), “é responsável pelo ‘processamento’ dos sinais de entrada do neurônio”. Já o axônio, de acordo com Finocchio (2014, p.7), “é responsável por transmitir estímulos para outras células”. “O axônio transmite em ambas as direções, mas o fluxo de informações parte dos dendritos para o axônio” (BREGA, 1996, p.21).

O ponto de contato entre as terminações do axônio de um neurônio e os dendritos de outro neurônio é chamado de sinapse. É através das sinapses que os

neurônios se unem funcionalmente, formando as redes neurais (BRAGA; LUDERMIR; CARVALHO, 2000, p.6). É mencionado por Silva, Spatti, Flauzino (2010, p.30) que “inexiste contato físico entre os neurônios na junção sináptica, sendo que elementos neurotransmissores liberados são os responsáveis por ponderar a transmissão de impulsos elétricos de um neurônio para o outro”. Na Figura 12 é possível observar as conexões sinápticas entre neurônios.

Figura 12 – Conexões sinápticas entre neurônios



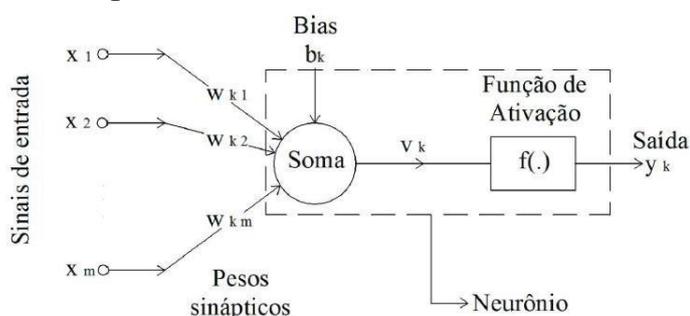
Fonte: Silva, Spatti, Flauzino (2010, p.30)

## 2.2.2 Redes Neurais Artificiais

Conforme Graupe (2013, p.1, tradução nossa) as Redes Neurais Artificiais (RNA) são, como o próprio nome indica, redes computacionais que tentam simular, de grosso modo, o processo de decisão em redes de neurônios do sistema nervoso central biológico.

Silva, Spatti, Flauzino (2010, p.33) afirmam que o modelo de neurônio mais simples e que engloba as principais características de uma rede neural biológica foi proposto em 1943 por McCulloch & Pitts, e até hoje é o modelo mais utilizado nas diferentes arquiteturas de RNA's. O neurônio de McCulloch & Pitts pode ser observado na Figura 13.

Figura 13 – Neurônio de McCulloch & Pitts



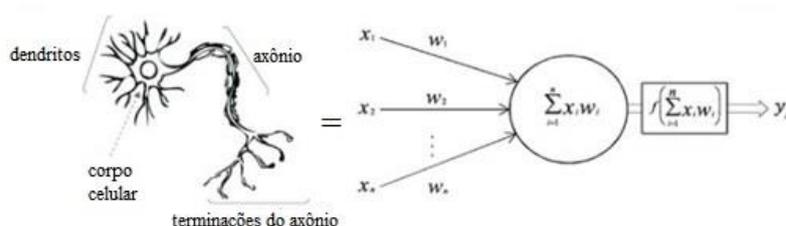
Fonte: Araújo, Neto, Filho (2015, p.39)

De acordo com Braga, Ludermir, Carvalho (2000, p.8), a descrição matemática da figura consiste em modelo com  $m$  terminais de entrada  $X_1, X_2, \dots, X_m$  (que representam os dendritos) e apenas um terminal de saída  $y$  (representando o axônio).

Para emular o comportamento das sinapses, os terminais das mesmas têm pesos acoplado  $w_1, w_2, \dots, w_m$ , cujos valores podem ser positivos ou negativos (BRAGA; LUDERMIR; CARVALHO, 2000, p.8). Sismoto (2021, p.29) ressalta que “comumente há uma entrada constante especial associada a um peso próprio denominada bias ( $b_k$ )”.

Na Figura 14 é possível notar a semelhança do neurônio biológico com o neurônio artificial de McCullloch & Pitts.

Figura 14 – Neurônio biológico e Neurônio artificial

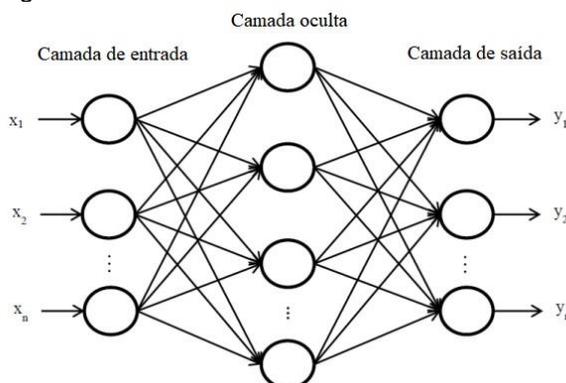


Fonte: Adaptado de Sumari, Syed, Abualigah (2021, p.2002)

“Grande parte das redes neurais são formadas pela organização dos neurônios em camadas ou campos. Uma camada deve agrupar neurônios com funções ou propriedades semelhantes” (BREGA, 1996, p.23).

Como é possível observar na Figura 15, as redes neurais artificiais são divididas basicamente em 3 partes, ou camadas: a camada de entrada, a camada oculta e a camada de saída.

Figura 15 – Estrutura de uma rede neural artificial

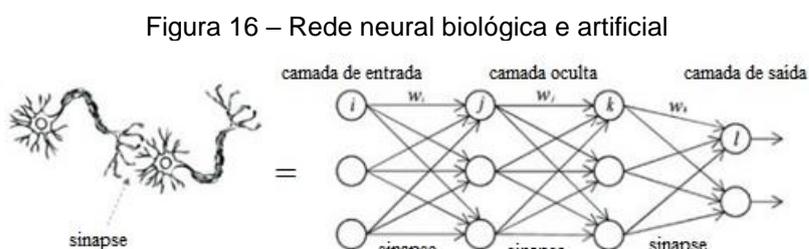


Fonte: Adaptado de Ahmadian, Khanteymooori (2015, p.4)

Conforme Silva, Spatti, Flauzino (2010, p.46), “a camada de entrada é responsável pelo recebimento de informações (dados), sinais, características ou medições advindas do meio externo”.

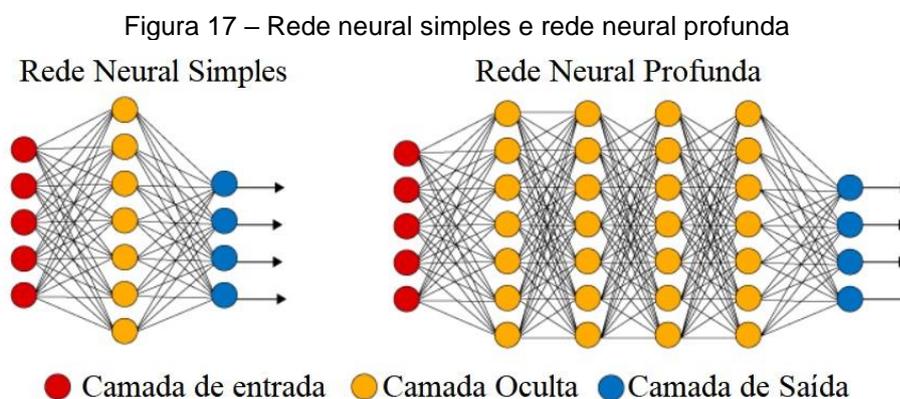
A camada oculta, também conhecida como intermediária e/ou escondida é responsável por grande parte do processamento e, por meio das conexões ponderadas, ela pode ser considerada como uma camada extratora de características (FINOCCHIO, 2014, p.20),

Já a camada de saída, para Silva, Spatti, Flauzino (2010, p.46), “é responsável pela apresentação dos resultados da rede, os quais são advindos dos processamentos realizados pelos neurônios das camadas anteriores”. Na Figura 16 é notável a semelhança entre as redes neurais biológicas e as redes neurais artificiais.



Fonte: Adaptado de Sumari, Syed, Abualigah (2021, p.2002)

Entretanto, as redes neurais que apresentam uma grande quantidade de camadas ocultas, conforme Hollerweger (2019, p.13), são chamadas de redes neurais profundas, e representam um avanço significativo que levou a resultados impressionantes obtidos recentemente com inteligência artificial. Na Figura 17 é possível visualizar a diferença entre uma RNA simples e uma rede neural profunda.



Fonte: Adaptado de Vásquez (2017, online)

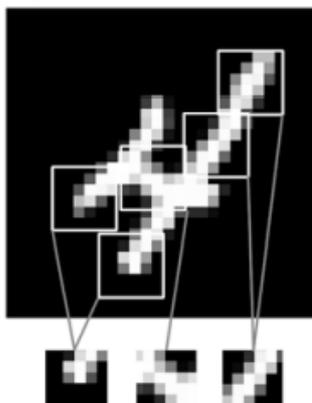
### 2.2.3 Redes Neurais Convolucionais

De acordo com Yamashita et al. (2018, p.611, tradução nossa), as Redes Neurais Convolucionais (RNC) são uma classe de redes neurais artificiais que se tornaram dominante em várias tarefas de visão computacional.

A diferença fundamental entre uma camada oculta de uma rede neural e uma camada de convolução, conforme Chollet (2018, p.122, tradução nossa), é que as camadas das redes profundas aprendem padrões globais em seu espaço de recursos de entrada, enquanto as camadas de convolução aprendem padrões locais.

Na Figura 18 é possível observar que as imagens podem ser divididas em padrões locais, como bordas, texturas e assim por diante.

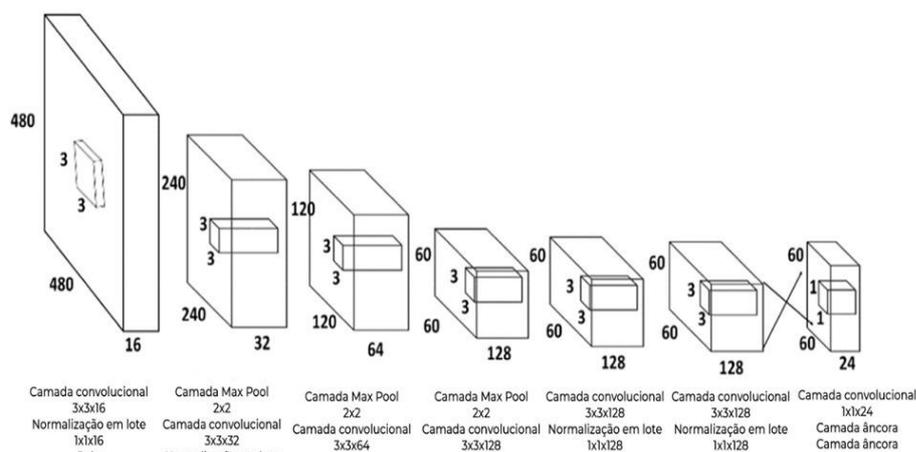
Figura 18 - Imagem sendo dividida em padrões locais



Fonte: Chollet (2018, p.122)

A camada de convolução identifica e extrai os melhores padrões da imagem de entrada e preserva as informações genéricas em uma matriz. A representação matricial da imagem de entrada é multiplicada por elementos com filtros e somada para produzir um mapa de características (KATEGARU, 2020, online). Na figura 19 é possível observar a estrutura de uma RNC.

Figura 19 - Estrutura de uma RNC



Fonte: Adaptado de Barreiros et al. (2021)

### 2.2.3.1 Visão computacional

Visão Computacional é a ciência que estuda e desenvolve tecnologias que permitem que máquinas enxerguem e extraiam características do meio, através de imagens capturadas por diferentes tipos de sensores e dispositivos. Essas informações extraídas permitem reconhecer, manipular e processar dados sobre os objetos que compõem a imagem capturada (BALLARD, BROWN, 1982 apud BARELLI, 2019, p.1).

Algumas das tarefas da visão computacional podem ser visualizadas na Figura 20, são elas: segmentação semântica, classificação e localização, detecção de objeto e segmentação de instância.



Fonte: Adaptado de Li, Johnson, Yeung (2017, online)

A tarefa de visão computacional abordada neste trabalho é a de detecção de objetos. Conforme Surga (2018, online, tradução nossa), a detecção de objetos consiste em prever a localização do objeto junto com o rótulo da classe. Para identificar a posição e o tipo de objetos são utilizadas “*bounding boxes* (caixas delimitadoras).

Uma *Bounding Box* é um retângulo que envolve um objeto, que especifica sua posição, classe (por exemplo: carro, pessoa) e confiança (qual a probabilidade de estar naquele local) (SUBRAMANYAM, 2021, online, tradução nossa).

No caso deste trabalho, a classe de objeto será a patologia do tipo panela e o retângulo delimitará onde esse objeto se localiza na imagem, além de expor sua confiabilidade, conforme é demonstrado na Figura 21.

Figura 21 – Detecção de objetos aplicada na patologia panela



Fonte: Autoria própria

### 2.2.3.2 Treinamento de uma RNC

Conforme Hollerweger (2019, p.31), o treinamento de um algoritmo de inteligência artificial é dividido em épocas, sendo cada época um período de treinamento em que o algoritmo tenha percorrido todo o conjunto de treinamento pelo menos uma vez.

As redes neurais geralmente são treinadas utilizando 3 conjuntos de dados, são eles: conjunto de treinamento, de validação e de teste. De acordo com Taylor (2017, p.16, tradução nossa), o conjunto de dados de treinamento é usado para ajustar os pesos da RNC, o de validação é usado para minimizar o *overfitting* e o de teste é usado como um teste final para avaliar a precisão com que a rede foi treinada.

O *overfitting* ocorre quando uma rede funciona bem com um conjunto de treinamento específico e minimiza o erro, mas quando apresentado a um novo conjunto de treinamento a taxa de erro é muito maior [...]. Isso geralmente ocorre porque a rede apenas parece ter aprendido com sucesso, mas na realidade apenas “memorizou” o conjunto de treinamento inicial (TAYLOR, 2017, p.288, tradução nossa).

### 2.2.3.3 Métricas de Avaliação

As métricas de avaliação são utilizadas após o treinamento de uma rede neural, visando avaliar o seu desempenho e a necessidade da mesma ser treinada por mais épocas, caso o seu resultado não seja satisfatório.

Existem diversas métricas de avaliação; as que serão abordadas nesse presente trabalho são: matriz de confusão, precisão, *recall*, *f1-score*, IoU (*Intersection over Union*) e mAP (*mean Average Precision*).

A matriz de confusão, conforme nota-se na Figura 22, consiste em uma matriz 2x2, onde os elementos da diagonal principal são os acertos: verdadeiros positivos (*True Positive* - TP) e verdadeiros negativos (*True Negatives* - TN), enquanto os elementos da diagonal secundária representam os erros: falsos positivos (*False Positives* - FP) e falsos negativos (*False Negatives* - FN).

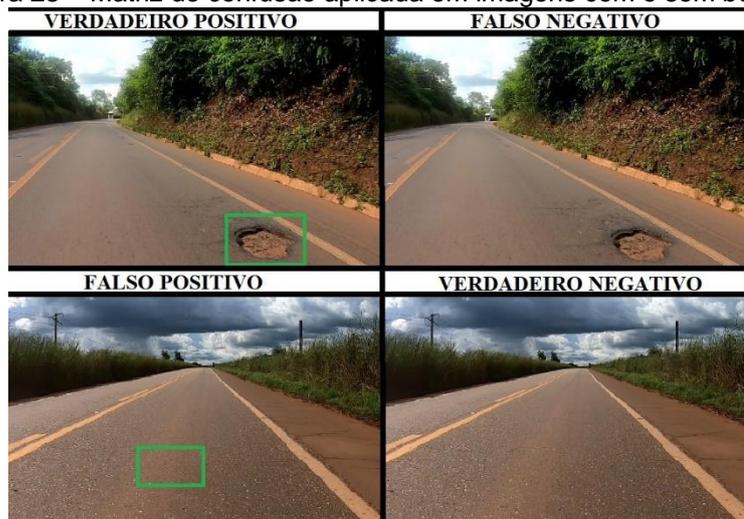
Figura 22 – Matriz de confusão

		Valor Predito	
		Sim	Não
Real	Sim	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (TN)

Fonte: Nogare (2020, online)

Na Figura 23 é possível observar, de forma sucinta, cada um desses elementos aplicados em imagens que contêm e não contêm a patologia estudada.

Figura 23 – Matriz de confusão aplicada em imagens com e sem buracos



Fonte: Autoria própria

Observa-se na Figura 23 que o verdadeiro positivo (TP) ocorre quando na imagem há um buraco e a RNC previu corretamente que havia um buraco. O falso negativo (FN) ocorre quando na imagem há um buraco e a RNC previu incorretamente

que não há. O falso positivo (FP) acontece quando na imagem não há nenhum buraco, mas a RNC identifica como se houvesse um. Já no verdadeiro negativo (TN) a imagem não apresenta nenhum buraco e a RNC previu corretamente isso. Vale ressaltar que para a detecção de objetos o verdadeiro negativo não é levado em consideração.

As métricas de avaliação: precisão, *recall* e *f1-score* são calculadas a partir de elementos da matriz de confusão.

A precisão, como pode-se notar na equação 1, é a razão entre a quantidade de verdadeiros positivos e a soma dos verdadeiros positivos com os falsos positivos. O resultado é um valor entre 0 e 1 e, quanto maior, mais precisa é a RNC.

$$Precisão = \frac{TP}{TP + FP} \quad (1)$$

O *recall*, conforme observa-se na equação 2, é a razão entre a quantidade de verdadeiros positivos e a soma dos verdadeiros positivos com os falsos negativos. O resultado é um valor entre 0 e 1 e, quanto maior, mais a RNC classificou corretamente as imagens.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

O *f1-score*, de acordo com Korstanje (2021, online, tradução nossa), tem o objetivo de combinar as métricas precisão e recall em uma única métrica e sua fórmula, apresentada na equação 3, consiste na média harmônica da precisão e do recall. O resultado é um valor entre 0 e 1 e, quanto maior, melhor.

$$F1 \text{ score} = 2 \times \frac{Precisão \times Recall}{Precisão + Recall} \quad (3)$$

“*Intersection over Union (IoU)* é uma métrica de avaliação usada para medir a precisão de um detector de objetos com um conjunto de dados específico” (ROSEBROCK, 2016, online, tradução nossa).

Para calcular essa métrica é necessário duas *bounding boxes*: a *ground-truth* e a *bounding box* prevista pela RNC. Na Figura 24 é possível observar um exemplo dessas duas *bounding boxes* aplicadas em um buraco.

Figura 24 – *Ground-truth bounding box* e *bounding box* prevista

Fonte: De autoria própria

A *ground-truth*, em verde, é uma *bounding box* marcada manualmente, através de uma ferramenta de anotação, durante a montagem da base de dados da RNC e a *bounding box* prevista, em vermelho, é a *bounding box* que foi prevista pela RNC após o seu treinamento. O IoU consiste na razão entre a área de sobreposição das *bounding boxes* e a área de união das mesmas, como pode ser visualizado na Figura 25.

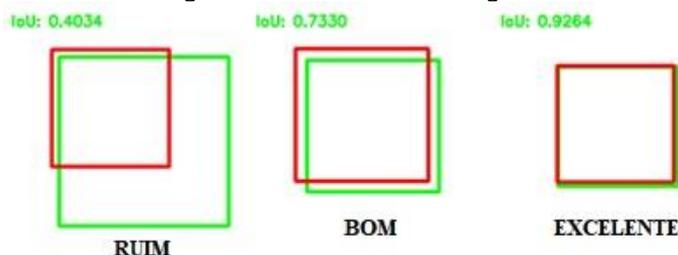
Figura 25 – Determinação do IoU

$$\text{IoU} = \frac{\text{área de sobreposição}}{\text{área de união}} = \frac{\text{área de sobreposição}}{\text{área de união}}$$

Fonte: Adaptado de Padilla, Netto, Silva (2020, p.2)

A partir deste cálculo é obtido um valor que vai de 0 a 1. Rosebrock (2016, online, tradução nossa) afirma que é extremamente improvável que as coordenadas da *bounding box* prevista correspondam exatamente às coordenadas da *ground-truth bounding box*, dessa forma, quanto mais próximo de 1 for o valor da métrica, melhor, conforme pode ser notado na Figura 26.

Figura 26 – IoU de bounding boxes



Fonte: Adaptado de Rosebrock (2016, online)

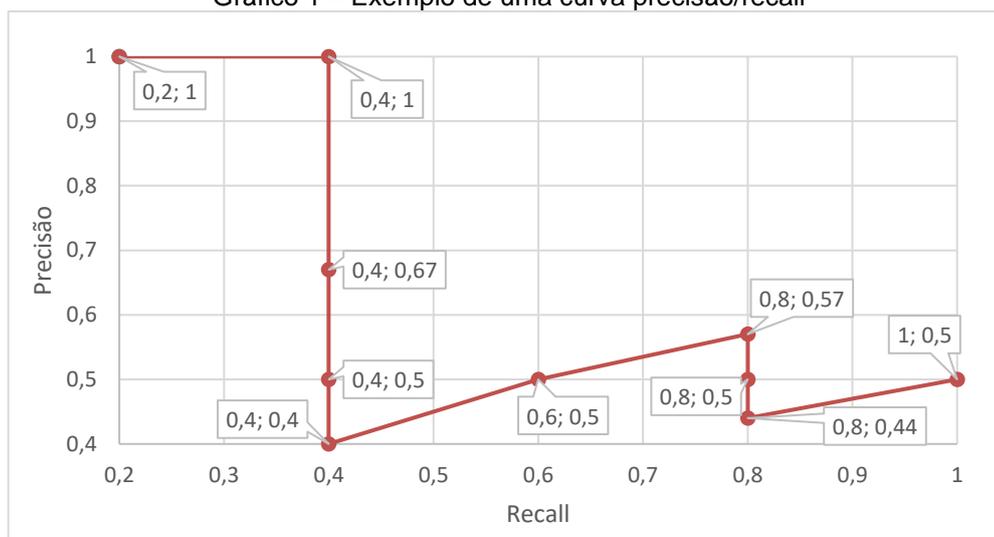
Por fim, temos a métrica de avaliação mAP (*mean Average Precision*), que consiste na média aritmética de classes da AP (*Average Precision*) para cada classe. Como no presente trabalho será utilizado somente uma classe, mAP e AP terão o mesmo valor.

De acordo com Padilla, Netto e Silva (2020, p.1, tradução nossa), AP é uma métrica popular para avaliar a precisão de detectores de objetos. Hui (2018, online, tradução nossa) afirma que a definição geral para a AP é encontrar a área sob a curva de precisão/recall, através da equação 4.

$$AP = \int_0^1 p(r)dr \quad (4)$$

A curva precisão/recall consiste em um gráfico formado por pares ordenados, na qual o eixo das ordenadas é a precisão e o eixo das abscissas é o recall. Um exemplo de uma curva precisão/recall pode ser observada no Gráfico 1.

Gráfico 1 – Exemplo de uma curva precisão/recall

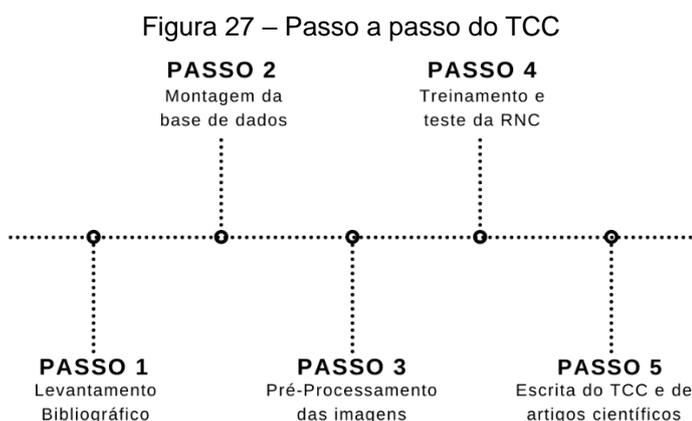


Fonte: Adaptado de Hui (2018, online)

### 3 MATERIAIS E MÉTODOS

#### 3.1 Desenho do estudo

Para execução do presente trabalho, os métodos de pesquisa foram executados em 5 etapas, conforme pode ser visualizado na Figura 27.



Fonte: Autoria própria

De início foi realizado o levantamento bibliográfico, através da leitura de livros, artigos e monografias pertinentes ao tema estudado. Após isso foi montada a base de dados, através da coleta de imagens de patologias com a câmera GoPro Hero 7. O terceiro passo consistiu no pré-processamento das imagens coletadas, por meio do *software* Labellmg.

O quarto passo consistiu no treino e teste da Rede Neural Convolutiva no ambiente de execução Google Colab, utilizando a linguagem Python, o algoritmo YoloV4 e o *framework* Darknet. Por fim, foi realizada a escrita da monografia, que posteriormente resultará na publicação de artigos científicos.

#### 3.2 Montagem da Base de Dados

Para montar a base de dados utilizada para treinamento do modelo neuronal foi necessário recolher diversas de imagens de painéis em pavimentos asfálticos, com a finalidade de se montar uma base de dados robusta.

Apesar de na internet já existirem diversas imagens disponíveis da patologia estudada, optou-se pela montagem de um conjunto de imagens exclusivas. Para tal, foi acoplada, como é exposto na Figura 28, uma GoPro Hero 7 no capô de um veículo, com a finalidade de gravar diversas ruas com revestimento asfáltico na

cidade de Açailândia-MA e posteriormente recolher os frames dos vídeos que apresentam buracos.

Figura 28 – Câmera GoPro Hero 7 acoplada no capô do veículo



Fonte: Autoria própria

Foi utilizada essa câmera em específico devido à sua excelente qualidade de imagem, possuindo 12 megapixels e conseguindo gravar vídeos em HD e 4K, além de possuir a tecnologia de estabilização de imagem HyperSmooth, que conforme Kislevitz (2018, online, tradução nossa), elimina a maioria dos tiques e solavancos das filmagens através da combinação de frames usando dados de acelerômetro e giroscópio. Na figura 29 é notável o funcionamento desta tecnologia.

Figura 29 – Tecnologia HyperSmooth



Fonte: Kislevitz (2021, online)

Esta tecnologia foi demasiadamente vantajosa para o presente trabalho, pois a câmera realizou as filmagens no capô de um carro em movimento, estabilizando assim os frames trêmulos devido à movimentação do carro e às irregularidades das pistas. Na Figura 30 observa-se um exemplo de um frame de vídeo gravado com a câmera GoPro Hero 7.

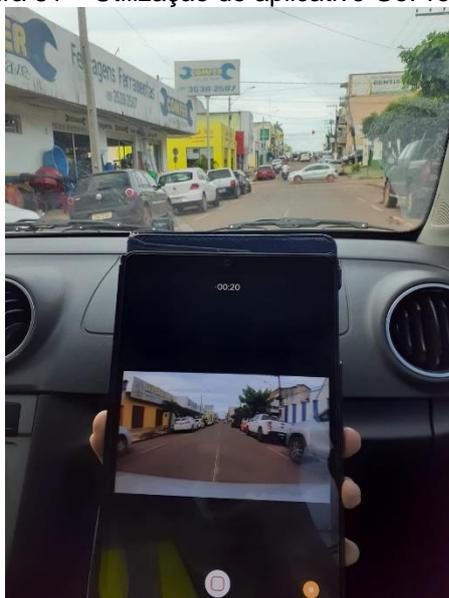
Figura 30 – Frame de vídeo gravado com a câmera GoPro Hero 7



Fonte: Autoria própria

Salienta-se que outra vantagem da utilização da câmera GoPro Hero 7 é a conexão Bluetooth com o aplicativo GoPro Quik, disponível para smartphones com sistema operacional Android e iOS, que possibilita a conexão remota com a câmera no capô do veículo - vide Figura 31 - sendo possível iniciar, pausar e interromper gravações sem a necessidade de sair do veículo.

Figura 31 – Utilização do aplicativo GoPro Quik



Fonte: Autoria própria

Ademais, a câmera utilizada também possui a capacidade de armazenar informações de telemetria, devido ao seu GPS, dentre elas: velocidade do veículo, altitude, trajeto, distância percorrida e coordenadas geográficas (latitude e longitude).

Estes dados foram inseridos em uma das gravações realizadas, e esta gravação foi submetida a RNC após seu treinamento, visando testar a capacidade de detecção das painelas presentes na filmagem. Para extrair esses dados foi usado o

software Telemetry Extractor e para inseri-los no vídeo foi utilizado o Dashware. Na Figura 32 verifica-se as informações de telemetria presentes em uma das gravações realizadas.



Fonte: Autoria própria

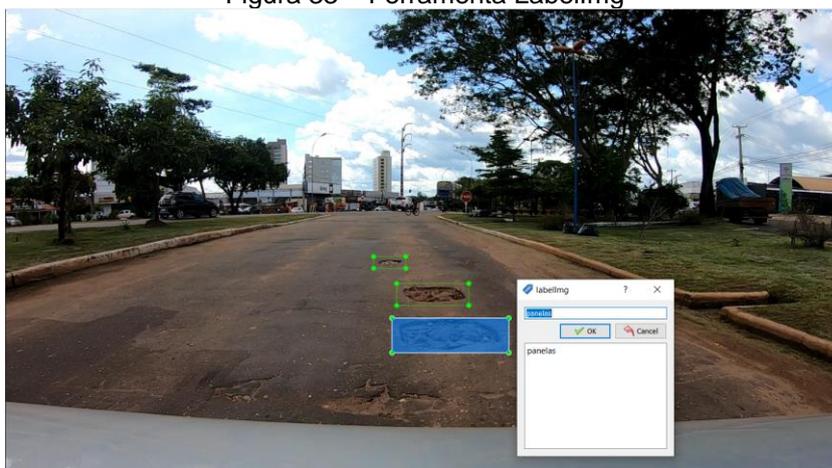
Dentre esses dados de telemetria o mais interessante para o presente trabalho são as coordenadas geográficas, fornecendo a localização exata das placas nas filmagens dos pavimentos. Essa localização possibilita uma facilidade, ao poder público, por exemplo, de executar medidas corretivas no defeito em questão.

### 3.3 Pré-Processamento das imagens

Após realizadas as filmagens foram recolhidos os frames que apresentaram a patologia do tipo placa das filmagens, através do software reprodutor de vídeos VLC media player, e posteriormente inseridos no Labellmg.

O Labellmg, segundo Nelson (2020, online), é uma ferramenta gratuita escrita em Python, com código aberto, cuja função é rotular imagens graficamente, sendo uma maneira fácil de rotular algumas centenas de imagens para incluir em projetos de detecção de objetos. Na Figura 33 é possível visualizar como a ferramenta funciona, sendo possível criar e nomear *bounding boxes* personalizadas.

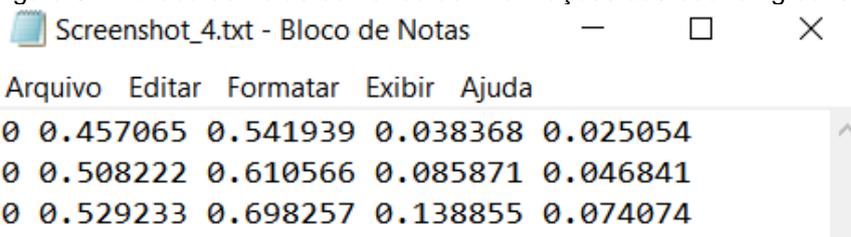
Figura 33 – Ferramenta LabelImg



Fonte: Autoria própria

Vale ressaltar que ao final do processo as informações das *bounding boxes* anotadas manualmente, as *ground-truth*, são salvas em um arquivo de bloco de notas, em formato .txt, conforme o conteúdo mostrado na Figura 34.

Figura 34 – Bloco de notas contendo as informações das bounding boxes



Fonte: Autoria própria

Observa-se que na Figura 34 existem 3 linhas, cada uma representando uma panela presente na imagem anotada, que nesse caso foi a Figura 33. Nota-se também que em cada linha existem 5 elementos. Cada elemento representa um determinado parâmetro, e eles estão representados na Figura 35.

Figura 35 – Elementos do arquivo de texto

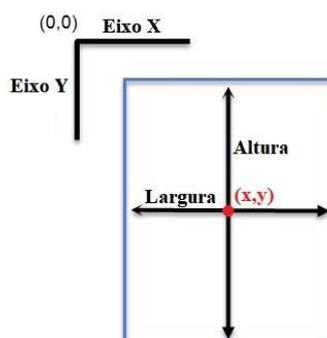
0	0.457065	0.541939	0.038368	0.025054	
0	0.508222	0.610566	0.085871	0.046841	
0	0.529233	0.698257	0.138855	0.074074	
<b>CLASSE</b>	<b>ID</b>	<b>X</b>	<b>Y</b>	<b>LARGURA</b>	<b>ALTURA</b>

Fonte: Autoria própria

O primeiro elemento, como pode ser notado, é comum para todas as linhas pelo número 0, isso denota que todos os elementos compactuam a mesma classe, ou seja, panela, visto que é a única patologia que será detectada.

Os elementos 2 e 3 representam as coordenadas centrais dos eixos X e Y da bounding box, respectivamente. Já os elementos 4 e 5 representam a largura e a altura da bounding box. Na Figura 36 é possível observar de forma objetiva os elementos citados acima representados em uma bounding box.

Figura 36 – Representação dos elementos do arquivo de texto

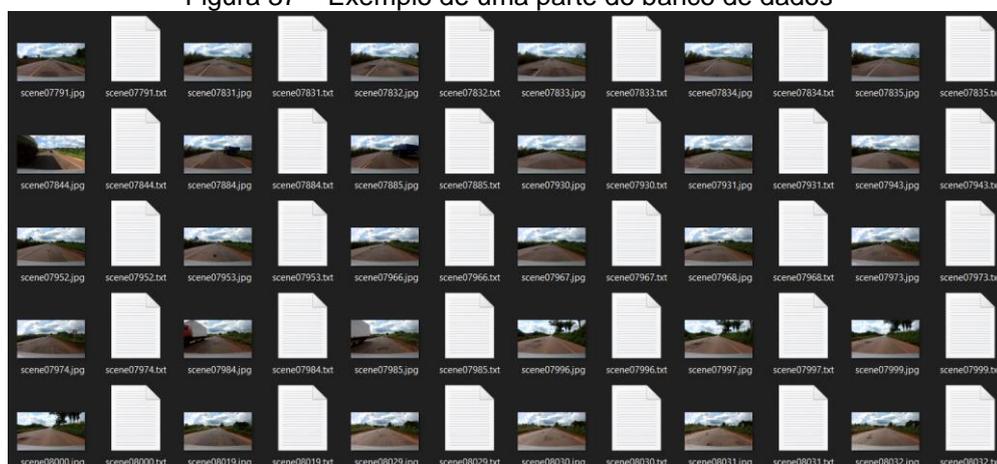


Fonte: Adaptado de Vijayvergiya (2021, online)

Esse processo de criação de *ground-truth bounding boxes* foi repetido para todos os frames de imagens que contêm a patologia estudada, então, ao final foram obtidos dois bancos de dados, um com 1376 imagens, que será chamado de DT1 (DataSet1) e o outro com 2000 imagens, que será chamado de DT2 (DataSet2). Ressalta-se que o DT2 é composto das 1376 imagens do DT1 somado a outras 624 imagens inéditas que foram recolhidas utilizando a presente metodologia.

Nota-se na Figura 37 que existem sempre dois arquivos de mesmo nome, um arquivo de imagem com as placas (em .jpg) e um arquivo de texto, com as anotações dos elementos das *bounding boxes* (em .txt).

Figura 37 – Exemplo de uma parte do banco de dados



Fonte: Autoria própria

É relevante destacar que nos bancos de dados utilizados existem algumas imagens similares, isto ocorreu por causa das imagens utilizadas nos bancos de dados terem sido retiradas de frames de vídeos de um carro em movimento, esse fato pode influenciar nos resultados obtidos.

Destaca-se também que todas as filmagens foram realizadas no período matutino e vespertino, possibilitando uma melhor visualização, devido à claridade, das painelas presentes nos vídeos, algo que não aconteceria caso as gravações fossem efetuadas no período noturno.

### 3.4 Programação da rede neural

Para a programação da RNC empregada no presente trabalho foi utilizada a linguagem de programação Python, através do ambiente Google Colaboratory, também conhecido como Colab, que possui diversas vantagens, como funcionar diretamente do browser, sem ser necessária a instalação, além de oferecer, gratuitamente, a unidade de processamento gráfico (GPU) nVidia Tesla T4.

As GPUs são ideais para processamento de algoritmos de aprendizado de máquina pois, conforme Matos (2016, online), “elas são otimizadas para processar aplicações e realizar a mesma operação diversas vezes, de forma bastante rápida”. Na Figura 38 é possível visualizar a especificações técnicas da nVidia Tesla T4.

Figura 38 – Especificações técnicas da nVidia Tesla T4

NVIDIA-SMI 460.32.03			Driver Version: 460.32.03			CUDA Version: 11.2		
GPU	Name	Persistence-M	Bus-Id	Disp.A	Volatile	Uncorr.	ECC	
Fan	Temp	Perf	Pwr:Usage/Cap	Memory-Usage	GPU-Util	Compute	M.	MIG M.
0	Tesla T4	Off	00000000:00:04.0	Off			0	
N/A	52C	P0	27W / 70W	284MiB / 15109MiB	0%	Default	N/A	

Fonte: Autoria própria

Dentro do próprio Colab é possível importar o algoritmo de detecção de objetos You Only Look Once (YOLO), que em português significa “você olha apenas uma vez”, pois o algoritmo precisa “olhar” a imagem apenas uma vez para enviá-la à rede neural convolucional.

O YOLO treina em imagens inteiras e otimiza diretamente o desempenho de detecção. Esse modelo unificado tem vários benefícios em relação aos métodos tradicionais de detecção de objetos (REDMON et al. 2016, p.1, tradução nossa).

O YOLO utiliza uma arquitetura de rede neural profunda chamada Darknet, mesmo nome do *framework* utilizado para implementá-la. Redmon (2016, online, tradução nossa) afirma que o Darknet é escrito em linguagem C e CUDA e ele apresenta as vantagens de ser rápido, fácil de instalar e suportar computação de CPU e GPU.

Esse algoritmo apresenta diversas versões; a que foi utilizada nesse trabalho é a quarta versão, chamada de YOLOv4 e que, de acordo com Bochkovskiy, Wang, Liao (2020, p.10, tradução nossa), é superior aos outros detectores que são considerados os mais rápidos e precisos em termos tanto de velocidade quanto de precisão.

Atrela-se ao exposto que para o treinamento da rede neural para detecção de placas foi necessário importar o Darknet e o YOLOv4 no ambiente de execução do Google Colab, para dessa forma inserir os dois bancos de dados montados anteriormente. Os bancos de dados foram divididos em conjunto de treino (80% do total de imagens) e de teste (20% do total de imagens), conforme pode ser observado na Tabela 1.

Tabela 1 – Divisão das imagens dos bancos de dados

Banco de dados	Total de imagens	Conjunto de treinamento	Conjunto de teste
DT1	1376	1100	276
DT2	2000	1600	400

Fonte: Autoria própria

Salienta-se que diversas imagens dos bancos de dados continham mais de uma patologia do tipo placa, sendo necessária assim a confecção da Tabela 2, onde é possível observar a quantidade de placas totais e por conjunto.

Tabela 2 – Divisão das placas dos bancos de dados

Banco de dados	Total de imagens	Conjunto de treinamento	Conjunto de teste
DT1	2223	1842	381
DT2	3123	2592	531

Fonte: Autoria própria

Isto posto, foram realizados dois treinamentos de 6000 épocas, um para cada banco de dados. No início dos dois treinamentos foi utilizado o método chamado *transfer learning* (transferência de aprendizagem). De acordo com Ribani e Maregoni (2019, p.47, tradução nossa), esse é um método utilizado para transferir conhecimento adquirido em uma tarefa para resolver outra; esse procedimento pode ajudar, por exemplo, a melhorar a precisão ou reduzir o tempo de treinamento.

Para a RNC deste trabalho, a transferência de aprendizagem consistiu em utilizar um conjunto de pesos iniciais de uma outra RNC que já havia aprendido as informações mais básicas presentes em uma imagem, como por exemplo, bordas, linhas e cores, não sendo necessário assim ela aprender essas características padrões presentes em todas as imagens desde o início.

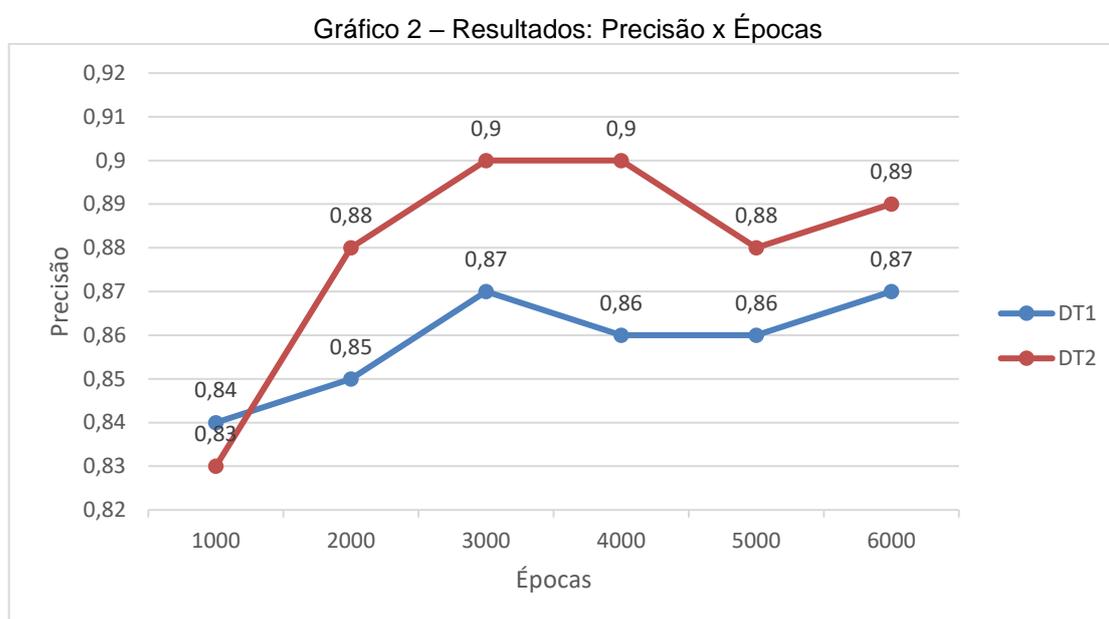
Após isso foi necessário apenas aguardar os resultados dos dois treinamentos, visto que eles são realizados de forma automática pela RNC.

## 4 RESULTADOS E DISCUSSÕES

### 4.1 Resultados do treinamento

Utilizando a metodologia apresentada anteriormente foram obtidos diversos resultados para os treinamentos realizados. Visando facilitar a visualização, os mesmos serão apresentados em gráficos e separados por métricas de avaliação.

O primeiro resultado pode ser observado no Gráfico 2, no qual é mostrada a precisão das duas bases de dados utilizadas.

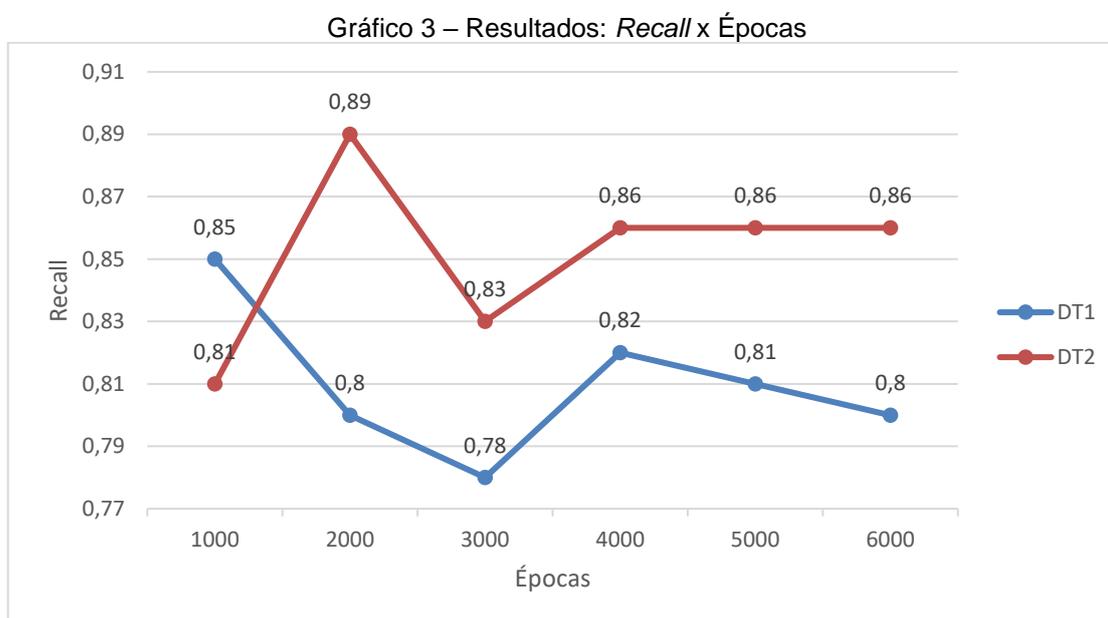


Fonte: Autoria própria

Nota-se que ambos os bancos de dados obtiveram resultados satisfatórios, porém o DT2, em vermelho, teve um desempenho superior quando comparado ao DT1, em azul, exceto pela época de número 1000 onde a precisão do DT1 foi um pouco maior. Entretanto, após isso DT2 conseguiu aumentar sua precisão, chegando a alcançar o valor de 0,9 nas épocas 3000 e 4000, o que é excelente, visto que o valor máximo que pode ser atingido é 1. A superioridade do DT2 pode ser explicada, pois o mesmo tem mais imagens que o DT1.

Verifica-se também que a precisão de ambos os modelos aumentou conforme as épocas iam aumentando, apesar da precisão de DT2 diminuir na época 5000 é notável que ela voltou a aumentar na época 6000, sendo assim necessário treinar por mais épocas para saber se o mesmo voltaria a aumentar e alcançar resultados maiores.

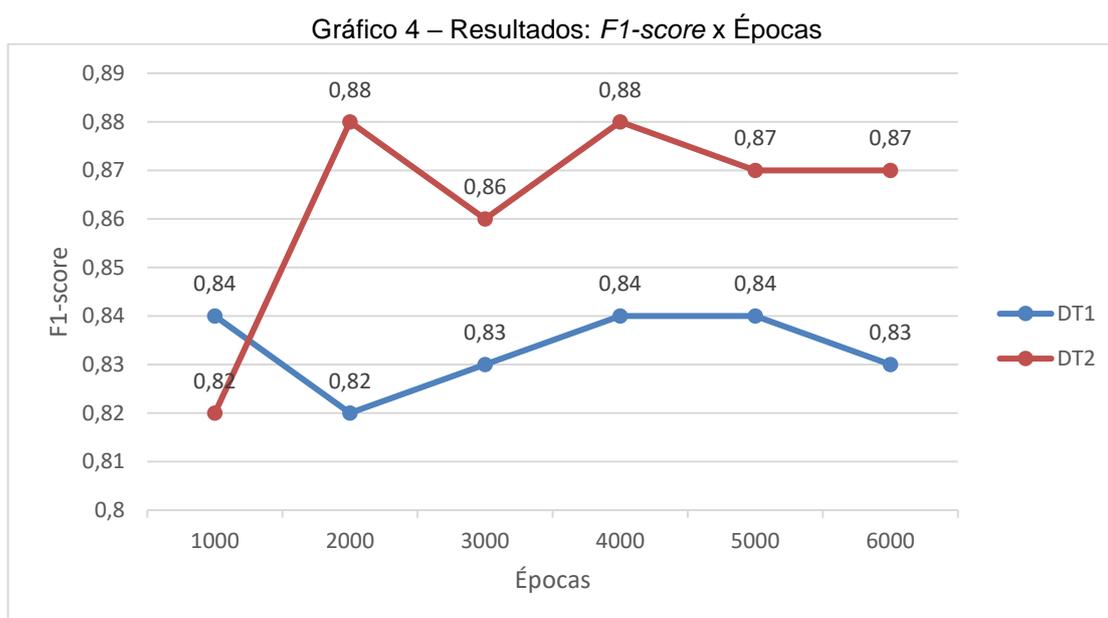
O próximo resultado obtido diz respeito à métrica de avaliação Recall e pode ser visualizado no Gráfico 3.



Fonte: Autoria própria

Novamente, é perceptível a superioridade do DT2 ao DT1, onde o maior resultado de *recall* foi de 0,89, na época 2000. Apesar do resultado ter caído para 0,83 na época seguinte, o recall se estabilizou em 0,86 a partir da época 4000, fato que não ocorreu no DT1, onde os valores de recall decresceram ao passar das épocas.

No Gráfico 4, é possível visualizar os resultados da métrica *F1-score*.

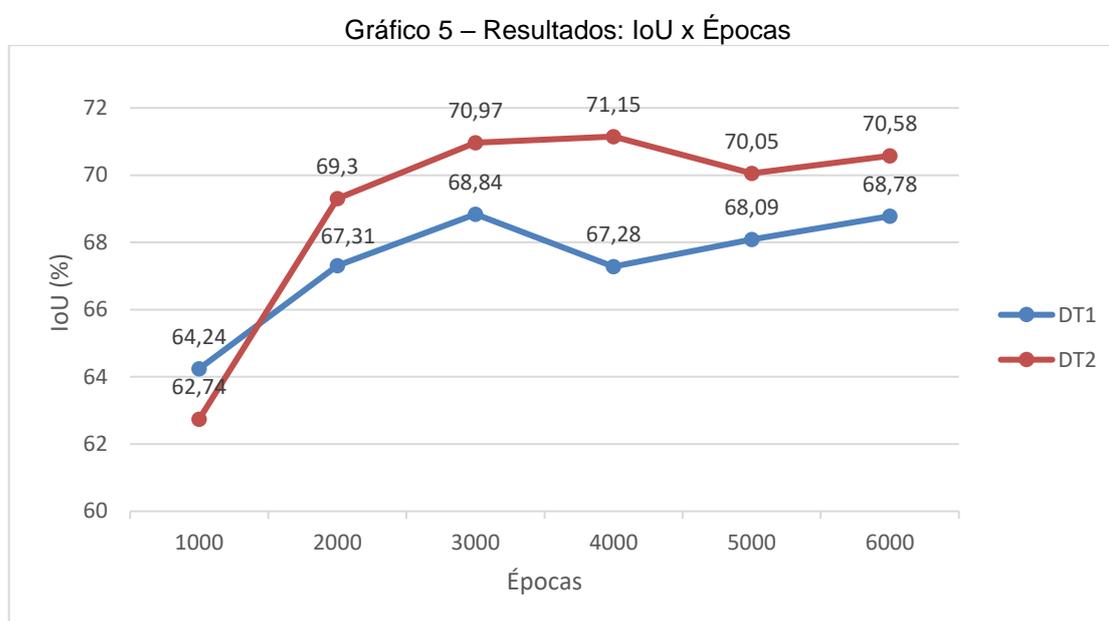


Fonte: Autoria própria

A terceira métrica avaliada foi o *F1-score*, onde logicamente DT2 foi superior, visto que essa métrica consiste na média harmônica da precisão e do *recall*, estas que DT2 também obteve um melhor desempenho.

No Gráfico notamos que para DT2 o maior valor obtido foi de 0,88, em duas épocas diferentes, 2000 e 4000, e teve o valor estabilizado de 0,87 a partir da época 5000. Quanto ao DT1 nota-se um bom desempenho, quando analisado isoladamente, oscilando entre os valores de recall de 0,82, 0,83 e 0,84.

O penúltimo resultado obtido consiste no IoU e pode ser observado na Gráfico 5.

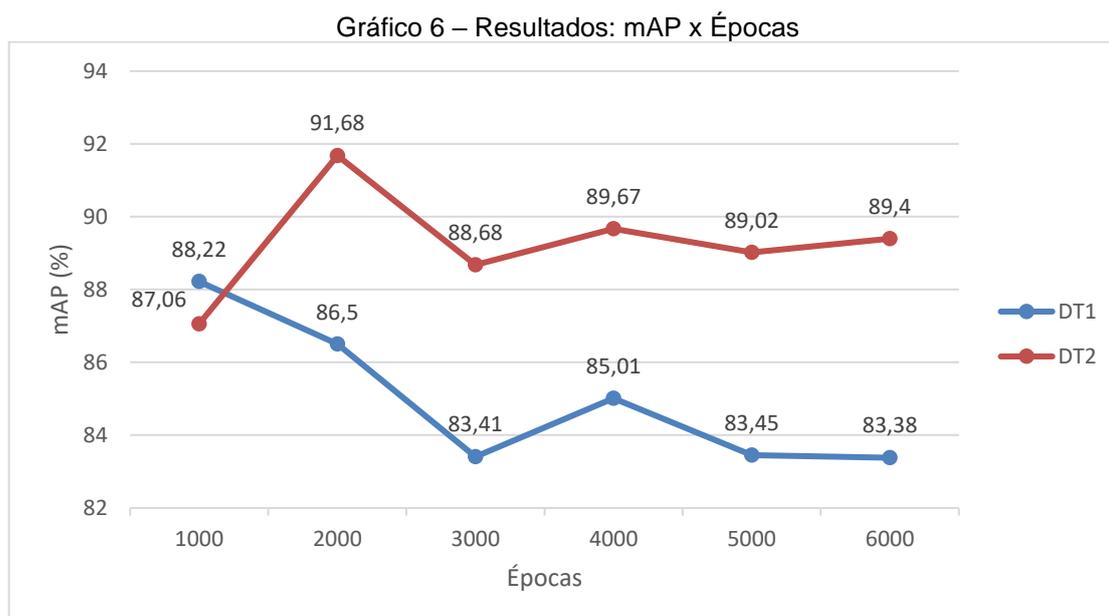


Fonte: Autoria própria

Para a métrica IoU novamente o DT2 apresentou valores superiores, tendo atingido 71,15% na época 4000, enquanto o maior valor do DT1 foi na época 3000, com 68,84%.

Verifica-se também que na época 6000 o valor de IoU do DT2 volta a aumentar, algo que ocorre também na DT1, porém na época 5000. Isso pode indicar que, caso a rede neural fosse treinada por mais tempo, os resultados de IoU seriam maiores para os dois bancos de dados.

Por fim, no Gráfico 6 são mostrados os resultados para o mAP.



Fonte: Autoria própria

E por fim, mais uma vez, DT2 obteve os melhores resultados, alcançando 91,68% na época 2000, um resultado excelente para um detector de objetos. Os resultados das outras épocas também foram muito bons, com nenhum mAP abaixo de 87%.

Já o DT1, novamente inferior, apresentou o seu melhor resultado (88,02%) na época 1000, e que foi decrescendo ao passar das épocas e apesar de ter aumentado na época 4000 para 85,01%, decresceu de novo após isso.

Com o propósito de ilustrar o desempenho dos dois bancos de dados, elaborou-se a Figura 39. Esta figura compara duas *bounding boxes* de uma imagem de panela presente no conjunto de teste, a *ground-truth bounding box*, em verde, e a *bounding box* prevista juntamente com a confiabilidade, em rosa. Essa comparação leva em consideração o número de épocas e o banco de dados.

Figura 39 – Comparação das bounding boxes

	1000	2000	3000	4000	5000	6000
DT1	panela: 0.95	panela: 0.99	panela: 0.99	panela: 1.00	panela: 1.00	panela: 1.00
DT2	panela: 0.97	panela: 0.99	panela: 0.99	panela: 0.99	panela: 1.00	panela: 1.00

Fonte: Autoria própria

Analisando essa figura é possível perceber os resultados satisfatórios de confiabilidade obtidos, tanto para o DT1 quanto para o DT2, além disso observa-se que a confiabilidade aumenta conforme o número de épocas também aumenta. Esse

fato indica que por quanto mais tempo (mais épocas) a RNC for treinada, maior será a confiabilidade das *bounding boxes* previstas. Também se percebe que o DT1 atingiu a confiabilidade máxima na época 4000, enquanto a para o DT2 essa confiabilidade máxima foi atingida na época 5000.

Em contrapartida percebe-se que as *bounding boxes* do DT2 se adaptaram melhor às *ground-truth bounding boxes* dos que às *bounding boxes* do DT1. Isso acontece pois, como verificado na Figura 39, a métrica de avaliação IoU teve resultados melhores para o DT2.

## 4.2 Teste da RNC em um vídeo

A fim de testar o modelo treinado em um vídeo, realizou-se uma filmagem na cidade de Imperatriz-MA, com a câmera GoPro Hero 7, percorrendo o trecho de 7km partindo do IFMA (Instituto Federal de Educação, Ciência e Tecnologia do Maranhão) – Campus Imperatriz até a Posto da Polícia Rodoviária Federal (PRF), na saída da cidade. O percurso pode ser observado na Figura 40.



Fonte: Google Maps (2022)

Após realizada a gravação, optou-se submeter o vídeo na RNC com a DT2 treinada por 4000 épocas, visto que nessas condições a RNC atingiu os melhores desempenhos nas métricas de avaliação, sendo a maior precisão (0,9), o segundo maior *recall* (0,86), o maior *F1-score* (0,88), maior IoU (71,15%) e o segundo maior mAP (89,67%).

Antes da RNC começar a detectar as painelas presentes na gravação foi necessária a definição de um limiar (*threshold*). Esse limiar funciona como um limite, em um limiar de 0,5 por exemplo, as detecções realizadas abaixo dessa porcentagem não são detectadas pela RNC.

Na Figura 41 é possível observar uma detecção realizada com 2 limiares diferentes (0,1 e 0,2), onde no limiar de 0,1 a RNC tem 13% de certeza que o objeto dentro da *bounding box* é uma painela, quando na realidade não é, ou seja, é um falso positivo. Já no limiar de 0.2 essa predição não ocorre, pois foi definido que a RNC não preveja nada abaixo de 20% de certeza.

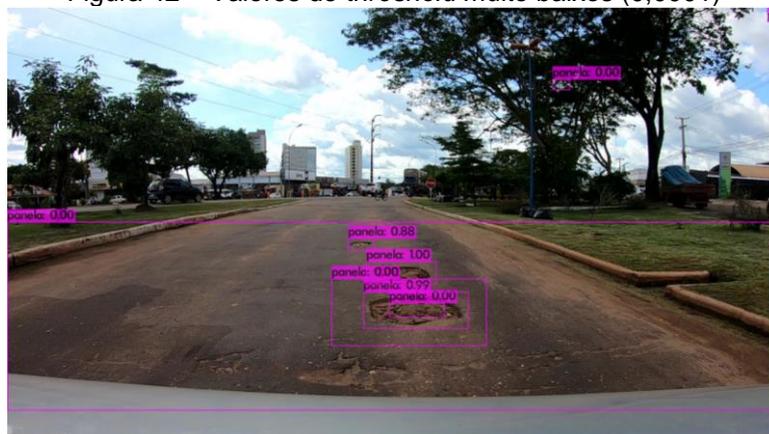
Figura 41 – Threshold diferentes aplicados na mesma imagem  
**threshold 0.1** **threshold 0.2**



Fonte: Autoria própria

Conforme pode ser notado na Figura 41, o objetivo de utilizar esse parâmetro é reduzir a quantidade de FP (Falsos Positivos). Entretanto é necessária cautela na hora de definir os limiares, visto que não é recomendado usar valores de *threshold* muito baixos, pois a RNC irá detectar diversos falsos positivos, como pode se observar na Figura 42, onde foi definido um limiar de 0,0001.

Figura 42 – Valores de *threshold* muito baixos (0,0001)



Fonte: Autoria própria

Também não é interessante utilizar *threshold* com valores muito altos ou então irá ocorrer a ausência de detecção, ou seja, falsos negativos. Isto pode ser observado na Figura 43.

Figura 43 – Valores de threshold muito altos



Fonte: Autoria própria

Sendo assim, para a análise da filmagem realizada foram utilizados três limiares, 0,5, 0,7 e 0,9, com a finalidade de observar o comportamento das detecções realizadas pela RNC utilizada.

Para o limiar de 0,5 a RNC detectou 229 *bounding boxes*, sendo 142 delas verdadeiros positivos e 87 falsos positivos.

Também foi imprescindível a contagem dos falsos negativos, para isso foi necessário a análise humana do vídeo, sendo realizada a contagem das painelas presentes no vídeo e que não foram detectadas pela RNC. No caso do limiar de 0,5 36 painelas deixaram de ser detectadas. A matriz de confusão para o limiar 0,5 pode ser observada na Tabela 3. Com base nos resultados obtidos na Tabela 1 e nas equações 1, 2 e 3, é possível calcular as métricas de avaliação: precisão, *recall* e *F1-score*; seus resultados são respectivamente 62,01%, 79,77% e 69,78%

Tabela 3 – Matriz de confusão para o limiar 0,5

Limiar 0,5					
Precisão = 62,01%		Recall = 79,77%		F1-score = 69,78%	
Matriz de confusão		Detecção			
		Sim	Não		
Buraco	Sim	142	36		
	Não	87	-		

Fonte: Autoria própria

O cálculo e os valores das métricas para esse limiar podem ser vistos na Figura 44.

Figura 44 – Métricas de avaliação para o limiar 0,5

$$Precisão = \frac{142}{142 + 87} \times 100 \rightarrow Precisão = 62,01\%$$

$$Recall = \frac{142}{142 + 36} \times 100 \rightarrow Recall = 79,77\%$$

$$F1\ score = 2 \times \frac{62,01 \times 79,77}{62,01 + 79,77} \rightarrow F1\ score = 69,78\%$$

Fonte: Autoria própria

Para o limiar de 0,7 foram detectadas 195 *bounding boxes*, sendo 132 verdadeiros positivo, 63 falsos positivos errados e 46 falsos negativos. A matriz de confusão para esse limiar pode ser observada na Tabela 4.

Tabela 4 – Matriz de confusão para o limiar 0,7

Limiar 0,7					
Precisão = 67,69%		Recall = 74,16%		F1-score = 70,78%	
Matriz de confusão		Detecção			
		Sim	Não		
Buraco	Sim	132	46		
	Não	63	-		

Fonte: Autoria própria

Percebe-se, com este resultado, que 10 verdadeiros positivos do limiar 0,5 agora se tornaram falsos negativos; isso ocorreu por causa do aumento do limiar utilizado, pois essas placas que deixaram de ser detectadas no limiar 0,7 tinham a confiabilidade abaixo de 70%, como é o caso da placa da Figura 45, que tinha a confiabilidade de 53%. Esse valor baixo de confiabilidade ocorre provavelmente pelo fato do buraco da figura em questão ser pequeno.

Figura 45 – Placa com 53% de confiabilidade



Fonte: Autoria própria

Quanto aos falsos positivos nota-se que 24 deixaram de ser detectados. Dentre eles encontram-se manchas (Figura 46), objetos fora da pista (Figura 47) e remendos (Figura 48).

Figura 46 – Mancha detectada erroneamente



Fonte: Autoria própria

Figura 47 – Objeto fora da pista detectado erroneamente



Fonte: Autoria própria

Figura 48 – Remendo detectado erroneamente



Fonte: Autoria própria

O cálculo das métricas de avaliação para o limiar 0,7 se encontram na Figura 49, e com esses valores é possível notar um aumento da precisão, visto que o número de falsos positivos diminuiu e houve um decréscimo do recall, já que o número de falsos negativos aumentou.

Figura 49 – Métricas de avaliação para o limiar 0,7

$$Precisão = \frac{132}{132 + 63} \times 100 \rightarrow Precisão = 67,69\%$$

$$Recall = \frac{132}{132 + 46} \times 100 \rightarrow Recall = 74,16\%$$

$$F1\ score = 2 \times \frac{67,69 \times 74,16}{67,69 + 74,16} \rightarrow F1\ score = 70,78\%$$

Fonte: Autoria própria

Por fim, para o limiar de 0,9 foram obtidas 146 *bounding boxes*, sendo 107 verdadeiros positivos e 39 falsos positivos, além de 71 falsos negativos. A matriz de confusão do limiar 0,9 se encontra na Tabela 5.

Tabela 5 – Matriz de confusão para o limiar 0,9

Limiar 0,9			
Precisão = 73,29%		Recall = 60,11%	
		F1-score = 66,05%	
Matriz de confusão		Detecção	
		Sim	Não
Buraco	Sim	107	71
	Não	39	-

Fonte: Autoria própria

É perceptível um grande aumento do número de falsos negativos para este limiar, isso ocorre pois o mesmo restringe a confiabilidade a 90%, fazendo assim a quantidade de verdadeiros positivos diminuir e a de falsos negativos aumentar.

Assim, diversas *bounding boxes* com confiabilidades consideradas boas (entre 70 e 89%) deixaram de serem detectadas, como por exemplo a panela presente na Figura 50 que apresenta uma confiabilidade de 89%.

Figura 50 – Panela com confiabilidade de 89%



Fonte: Autoria própria

A Figura 51 apresenta os resultados para as métricas de avaliação do limiar 0,9, e com ele é possível notar uma queda no recall quando comparado aos outros

limiares, justamente pelo fato supracitado (decréscimo dos verdadeiros positivos e aumento do número de falsos negativos).

Também é notável que a precisão aumenta mesmo com o decréscimo dos verdadeiros positivos, isso ocorre porque os falsos positivos diminuíram em uma maior quantidade.

Figura 51 – Métricas de avaliação para o limiar 0,9

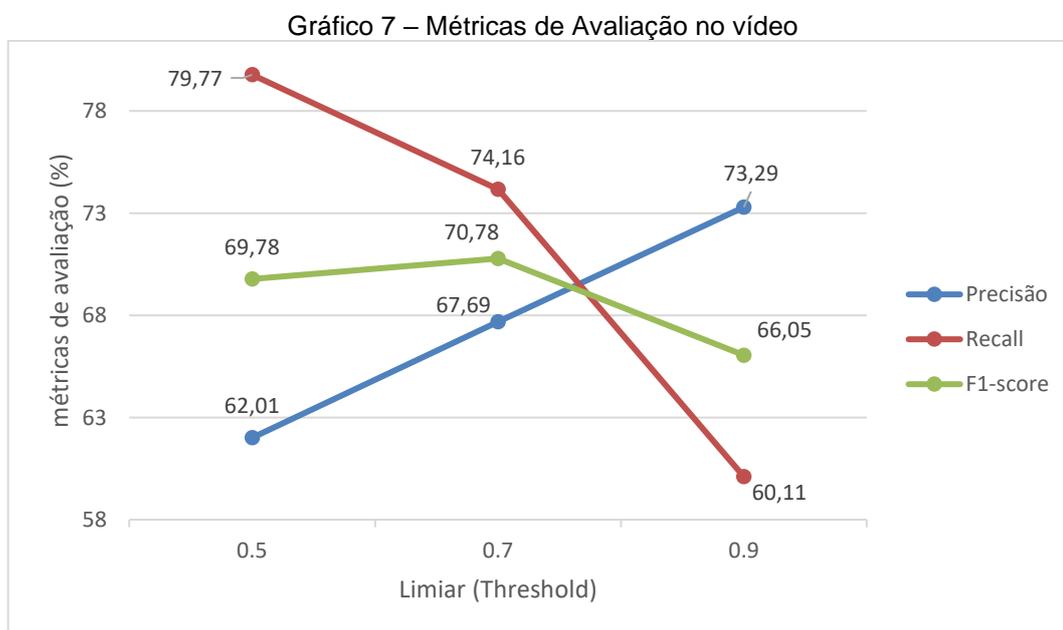
$$Precisão = \frac{107}{107 + 39} \times 100 \rightarrow Precisão = 73,29\%$$

$$Recall = \frac{107}{107 + 71} \times 100 \rightarrow Recall = 60,11\%$$

$$F1\ score = 2 \times \frac{73,29 \times 60,11}{73,29 + 60,11} \rightarrow F1\ score = 66,05\%$$

Fonte: Autoria própria

Visando facilitar a visualização dos resultados foi confeccionado o Gráfico 7, onde é possível verificar quais métricas de avaliação diminuíram e aumentaram, ao passo que o limiar foi mudando.



Fonte: Autoria própria

Ao analisar o gráfico é possível visualizar a queda do recall conforme o limiar aumenta. Esse fato acontece, pois as métricas verdadeiro positivo e falso negativo se relacionam, conforme os verdadeiros positivos decrescem os falsos negativos crescem. Ressalta-se que isso também ocorreu porque o valor de recall

para a época dessa base de dados utilizada foi de 0,86, caso esse valor fosse maior mais painéis seriam detectados.

Também é notável que ao passo que o limiar aumenta a precisão também aumenta; isso ocorre pois geralmente as confiabilidades dos falsos positivos são baixas, reduzindo a sua quantidade conforme a restrição da confiabilidade. Entretanto mesmo assim algumas detecções de objetos que não são patologias de pavimentos ocorreram com uma confiabilidade muito alta, como por exemplo pedras (Figura 52), manchas (Figura 53) e poças de água fora da estrada (Figura 54). Essas detecções errôneas influem diretamente na precisão do modelo.

Figura 52 – Pedra detectada com alta confiabilidade



Fonte: Autoria própria

Figura 53 – Mancha detectada com alta confiabilidade



Fonte: Autoria própria

Figura 54 – Poça de água detectada com alta confiabilidade



Fonte: Autoria própria

Uma possível explicação para essas detecções erradas é de que esses objetos se assemelham devido ao seu formato e a sua cor com a patologia estudada. Para seres-humanos a diferença pode ser óbvia, mas para uma máquina, não. Uma possível correção para diminuir os números de falsos positivos seria isso o aumento da base de dados para que dessa forma a RNC conseguisse detectar apenas painéis.

Um resultado satisfatório foi o de limiar 0,9, onde 107 painéis ainda se mantiveram detectáveis com alta confiabilidade. Na Figura 55 é possível visualizar um desses verdadeiros positivos, que foi detectado com 99% de confiabilidade.

Figura 55 – Painel detectado com 99% de confiabilidade



Fonte: Autoria própria

Em contrapartida, a ausência de algumas detecções foi insatisfatória, como é o caso da placa da Figura 56, não detectada nem no limiar de 0,5. A presença desses falsos negativos influi diretamente no valor de recall.

Figura 56 – Painel não detectada no limiar 0.5



Fonte: Autoria própria

Uma possível explicação para a ausência de algumas detecções é o *overfitting*. Conforme já foi exposto anteriormente, o *overfitting* ocorre quando uma RNC funciona bem somente com o conjunto na qual ela foi treinada, já quando ela é submetida a outro conjunto a sua quantidade de erros é bem maior.

Possíveis soluções para esse problema são o aumento da base de dados, a utilização de um conjunto de imagens para validação e a utilização de um algoritmo menos complexo, como o yolov4-tiny, por exemplo.

Com a análise da filmagem foi possível perceber que os resultados das métricas de avaliação dos vídeos foram inferiores aos obtidos pelo conjunto de treinamento. Mesmo assim, destaca-se a facilidade da utilização da metodologia do presente trabalho, onde em nenhum momento foi necessário descer do veículo para realizar o levantamento de patologias, evitando assim a necessidade de interromper o trânsito, a exposição dos avaliadores ao sol e a possíveis acidentes, bem como o cansaço.

Além disso, a utilização de Redes Neurais Convolucionais foi muito vantajosa, pois contamos com um algoritmo que “pensa” como humano, mas não tem as suas limitações físicas e biológicas, tornando assim o processo de levantamento de patologias em pavimentos mais rápido, eficiente e moderno.

Por fim, destaca-se também as exposições das informações de telemetria presentes na filmagem, principalmente as coordenadas de latitude e longitude. Com isso é possível a localização da patologia em questão, sendo possível executar as medidas corretivas com mais facilidade.

## 5 CONCLUSÃO

Através da utilização da arquitetura de detecção de objetos YOLOv4 (You Only Look Once), juntamente com o *framework* Darknet, confirmou-se que é possível realizar o treinamento de uma RNC para detectar a patologia do tipo panela em pavimentos asfálticos.

Com a análise dos gráficos dos resultados do treinamento foram notáveis os excelentes resultados para os dois bancos de dados utilizados (DT1 e DT2). Entretanto, notou-se um desempenho melhor, em todas as métricas de avaliação, do DT2, com a precisão atingindo (0,9) nas épocas 3000 e 4000, o *recall* 0,89 na época 2000, o *F1-score* 0,88 nas épocas 2000 e 4000, o IoU 71,15% na época 4000 e o mAP 91,68% na época 2000.

Esse resultado melhor para o DT2 ocorre devido à quantidade maior de imagens presentes nesse banco de dados. Conclui-se então que quanto mais imagens um banco de dados tiver, melhor a RNC será.

Também é possível concluir que para um banco de dados com apenas uma classe de imagem o tempo de treinamento (épocas) não necessariamente influi no resultado das métricas de avaliação, visto que nenhum dos melhores resultados obtidos ocorreu na época 6000.

Entretanto, seria interessante treinar a rede neural por mais milhares de épocas para observar o comportamento das métricas de precisão IoU e mAP, uma vez que o valor dessas métricas voltou a aumentar na época 6000, fato que não ocorreu nas métricas *recall* e *F1-score*, que se estabilizaram.

Para o resultado, o treinamento utilizando uma filmagem torna perceptível que a utilização na prática de um modelo neuronal para detecções de painéis se mostrou funcional, apesar de não obter resultados tão satisfatórios quanto o conjunto de treinamento, atingindo uma boa precisão (73,29%) no limiar 0.9 e um bom *recall* (79,77%) no limiar 0,5.

Essa boa funcionalidade, em conjunto com as exposições dos dados de telemetria, facilitam a manutenção de pavimentos asfálticos por parte do poder público, por exemplo, visto que ao final do processo é obtido um modelo que detecta buracos e indica em forma de coordenadas geográficas a sua localização.

Ademais, existe a possibilidade de melhorar os resultados das métricas de avaliação, principalmente a de *recall*, que decresceu demasiadamente conforme o

limiar aumentava, a partir da expansão do banco de dados utilizando técnicas como o *data augmentation*, e a utilização de outras arquiteturas de RNC, como, por exemplo, a yolov4-tiny, que possui uma arquitetura mais simples, e devido a isso é a mais recomendada para resolver problemas de uma só classe.

Também é importante ressaltar que é possível treinar a mesma RNC para detectar outras patologias que comumente são encontradas em pavimentos, como trincas e remendos, tornando assim o processo de correção de patologias muito mais completo, visto que a RNC conseguiria detectar outras patologias presentes em pavimentos asfálticos.

Por fim, destaca a facilidade de utilização da metodologia do presente trabalho, visto que em nenhum momento foi necessário descer do veículo para realizar o levantamento de patologias.

## 6 REFERÊNCIAS

- AHMADIAN, Sajad; KHANTEYMOORI, Ali Reza. **Training back propagation neural networks using asexual reproduction optimization**. In: 2015 7th Conference on Information and Knowledge Technology (IKT). IEEE, 2015. p. 1-6.
- ARAÚJO, Carla Beatriz Costa de; NETO, Silvrano Adonias Dantas; FILHO, Francisco de Assis Souza. **Previsão de vazões para o açude Orós/CE a partir de dados hidrometeorológicos utilizando perceptrons**. 2015.
- ASPHALT INSTITUTE. **The asphalt handbook**: manual series no. 4. Asphalt Institute, 1989.
- BALBO, José Tadeu. **Pavimentação asfáltica**: materiais, projeto e restauração. São Paulo: Oficina de Textos, 2007.
- BARREIROS, M. DE O. et al. Zebrafish tracking using YOLOv2 and Kalman filter. **Scientific Reports**, v. 11, n. 1, p. 1–14, 2021.
- BARELLI, Felipe. **Introdução a visão computacional**: uma abordagem prática com Python e Opencv. Casa do Código, 2019.
- BERNUCCI, Liedi Légi Bariani *et al.* **Pavimentação asfáltica**: formação básica para engenheiros. ABEDA, 2008.
- BOCHKOVSKIY, Alexey; WANG, Chien-Yao; LIAO, Hong-Yuan Mark. **Yolov4**: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, 2020.
- BOTERO, Eduardo Montoya. **Metodologia para aplicação de redes neurais artificiais para sistemas de alerta de escorregamentos deflagrados por chuvas em regiões montanhosas**. 2018.
- BRAGA, Antônio de Pádua; LUDERMIR, Teresa Bernarda ; CARVALHO, André Carlos Ponce de Leon Ferreira. **Redes neurais artificiais**: teoria e aplicações. LTC editora, 2000.
- BRANCO, Luiz Henrique Castelo. **MANIAC**: uma metodologia para o monitoramento automatizado das condições dos pavimentos utilizando VANTs. 2016.
- BREGA, José Remo Ferreira. **A utilização de redes neurais artificiais em um sistema de gerência de pavimentos**. 1996.
- CHEN, Hanshen; YAO, Minghai; GU, Qinlong. **Pothole detection using location-aware convolutional neural networks**. International Journal of Machine Learning and Cybernetics, v. 11, n. 4, p. 899-911, 2020.
- CHOLLET, François. **Deep learning with Python**. Manning Publications, 2018.
- CNT - CONFEDERAÇÃO NACIONAL DE TRANSPORTES. **Pesquisa CNT de rodovias 2022**. CNT: SEST SENAT, 2022.

\_\_\_\_\_. **Transporte rodoviário:** por que os pavimentos das rodovias do Brasil não duram? CNT, 2017.

DOMINGUES, Felipe Augusto Aranha. **MID – Manual para identificação de defeitos de revestimentos asfálticos de pavimentos.** 1993

DNIT - Departamento Nacional de Infraestrutura de Transportes. **DNIT 005/2003-TER:** defeitos nos pavimentos flexíveis e semi-rígidos: terminologia. Rio de Janeiro: IPR, 2003.

\_\_\_\_\_. **DNIT 006/2003 – PRO:** avaliação objetiva da superfície de pavimentos flexíveis e semirrígidos: procedimento. Rio de Janeiro: IPR, 2003.

\_\_\_\_\_. **DNIT 007/2003 – PRO:** levantamento para avaliação da condição de superfície de subtrecho homogêneo de rodovias de pavimentos flexíveis e semirrígidos para gerência de pavimentos e estudos e projetos: procedimento. Rio de Janeiro: IPR, 2003.

\_\_\_\_\_. **DNIT 008/2003 – PRO:** levantamento visual contínuo para avaliação da superfície de pavimentos flexíveis e semirrígidos: procedimento. Rio de Janeiro: IPR, 2003.

\_\_\_\_\_. **DNIT 009/2003 – PRO:** avaliação subjetiva da superfície de pavimentos flexíveis e semirrígidos: procedimento. Rio de Janeiro: IPR, 2003.

\_\_\_\_\_. **DNIT – IPR 700:** Glossário de termos técnicos rodoviários. Rio de Janeiro: DNIT, 2017.

\_\_\_\_\_. **DNIT – IPR 710:** Manual de conservação rodoviária. Rio de Janeiro: DNIT, 2005.

\_\_\_\_\_. **DNIT – IPR 720:** Manual de restauração de pavimentos asfálticos. Rio de Janeiro: DNIT, 2006.

FINOCCHIO, Marco Antonio Ferreira. **Noções de redes neurais artificiais.** 2014.

GRAUPE, Daniel. **Principles of artificial neural networks.** World Scientific, 2013.

GUPTA, Saksham *et al.* **Detection and localization of potholes in thermal images using deep neural networks.** Multimedia tools and applications, v. 79, n. 35, p. 26265-26284, 2020.

HOLLERWEGER, Marcos Meyer. **Aplicação de Visão Computacional no Auxílio ao Levantamento de Defeitos em Pavimento Rodoviário.** 2019.

HUI, Jonathan. **mAP (mean Average Precision) for Object Detection.** Jonathan Hui Medium, 6 mar. 2018. Disponível em: <https://jonathan-hui.medium.com/map-mean-average-precision-for-object-detection-45c121a31173>. Acesso em: 11 out. 2022.

JÚNIOR, José Leomar Fernandes; ODA, Sandra; ZERBINI, Luiz Francisco. **Defeitos e atividades de manutenção e reabilitação em pavimentos asfálticos.** 2006.

KATEGARU, Swapna. **Convolution Neural Networks**. Developers Breach, 21 ago. 2020. Disponível em: <https://developersbreach.com/convolution-neural-network-deep-learning/>. Acesso em: 9 ago. 2022.

KISLEVITZ, Abe. **How To Fix Lowlight HyperSmooth on your GoPro**. Abe Kislevitz, 13 jan. 2021. Disponível em: <https://abekislevitz.com/fix-lowlight-hypersmooth-on-your-gopro/>. Acesso em: 6 ago. 2022.

KISLEVITZ, Abe. **Insider's Guide to GoPro HERO7 Black**. Abe Kislevitz, 1 out. 2018. Disponível em: <https://abekislevitz.com/gopro-hero7-black/>. Acesso em: 6 ago. 2022.

KORSTANJE, Joos. **The F1 score**. Towards Data Science, 31 ago. 2021. Disponível em: <https://towardsdatascience.com/the-f1-score-bec2bbc38aa6>. Acesso em: 6 out. 2022.

LI, Fei-Fei; JOHNSON, Justin; YEUNG, Serena. **Lecture 11: Detection and Segmentation**, Stanford, 10 mai. 2017. Disponível em: [http://cs231n.stanford.edu/slides/2017/cs231n\\_2017\\_lecture11.pdf](http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf). Acesso em 9 ago. 2022.

MALLICK RAJIB, B.; EL-KORCHI, T. **Pavement Engineering: Principles and Practice**. 2018.

MATOS, David. **GPU e Deep Learning**. Ciência e dados, 5 set, 2016. Disponível em: <https://www.cienciaedados.com/gpu-e-deep-learning/>. Acesso em 9 ago. 2022

MILLER, John S; BELLINGER, William Y. **Distress identification manual for the long-term pavement performance program**. Federal Highway Administration. Office of Infrastructure Research and Development, 2014.

MÜLLER, B; REINHARDT, J; STRICKLAND, M T. **Neural networks: an introduction**. Durham: Springer Science & Business Media, 1995.

NELSON, Joseph. **Labelimg for Labeling Object Detection Data**. Roboflow, 16 mar, 2020. Disponível em: <https://blog.roboflow.com/labelimg/>. Acesso em 2 ago. 2022.

NOGARE, Diego. **Performance de Machine Learning – Matriz de Confusão**. Diego Nogare: Inteligência Artificial & Machine Learning, 27 Abr, 2020. Disponível em: <https://diegonogare.net/2020/04/performance-de-machine-learning-matriz-de-confusao/>. Acesso em 6 out. 2022.

PADILLA, Rafael; NETTO, Sergio L.; DA SILVA, Eduardo AB. A survey on performance metrics for object-detection algorithms. In: **2020 international conference on systems, signals and image processing (IWSSIP)**. IEEE, 2020. p. 237-242.

PAULA E SILVA, Marília. **Aplicação de Redes Neurais Artificiais no Diagnóstico de Falhas de Turbinas a Gás**. 2010.

REDMON, Joseph *et al.* **You only look once**: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 779-788.

REDMON, Joseph. **Darknet: Open Source Neural Networks in C**, 2016. Disponível em: <https://pjreddie.com/darknet/>. Acesso em 9 ago. 2022.

RIBANI, Ricardo; MARENGONI, Mauricio. A survey of transfer learning for convolutional neural networks. In: **2019 32nd SIBGRAPI conference on graphics, patterns and images tutorials (SIBGRAPI-T)**. IEEE, 2019. p. 47-57.

ROSEBROCK, Adrian. **Intersection over Union (IoU) for object detection**. PyImageSearch, 7 nov. 2016. Disponível em: <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>. Acesso em: 10 out. 2022.

RUSSELL, Stuart; NORVIG, Peter. **Artificial intelligence a modern approach**. Pearson Education, 2021.

SENÇO, Wlastermiler de. **Manual de Técnicas de Pavimentação**. PINI, 2008.

SILVA, Ivan Nunes da; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. **Redes neurais artificiais para engenharia e ciências aplicadas**. 2010.

SISMOTO, Pedro Afonso Barbosa. **Estudo e aplicação da visão computacional em imagens aéreas para detecção de edificações situadas a jusante de barragens**. 2021.

SUBRAMANYAM, Vineeth S. **Basics of Bounding Boxes**. Analytics Vidhya, 16 jan. 2021. Disponível em: <https://medium.com/analytics-vidhya/basics-of-bounding-boxes-94e583b5e16c>. Acesso em: 10 ago. 2022.

SUMARI, Putra; SYED, Saqib Jamal; ABUALIGAH, Laith. **A novel deep learning pipeline architecture based on CNN to detect Covid-19 in chest X-ray images**. Turkish Journal of Computer and Mathematics Education (TURCOMAT), v. 12, n. 6, p. 2001-2011, 2021.

SURGA, Koteshwar. **Introduction to Object Detection in Deep Learning**. Koteshwar Surga Medium, 5 nov. 2018. Disponível em: <https://medium.com/@skoteshwar930/introduction-to-object-detection-in-deep-learning-5f0c3b9862f0>. Acesso em: 10 ago. 2022.

TAYLOR, Michael. **Neural Networks: A visual introduction for beginners**. Blue Windmill Media, 2017.

VASQUEZ, Fávio. **Deep Learning made easy with Deep Cognition**. Becoming Human: Artificial Intelligence Magazine, 21 dez. 2017. Disponível em: <https://becominghuman.ai/deep-learning-made-easy-with-deep-cognition-403fbe445351>. Acesso em: 23 jul. 2022.

VIJAYVERGIYA, Gaurav. **Custom Image dataset for YOLO object detection using Open image dataset**. Gaurav Vijayvergiya Medium, 5 dez. 2021. Disponível em: <https://g-vj.medium.com/custom-image-dataset-for-yolo-object-detection-using-open-image-dataset-41c128b9231b>. Acesso em: 20 set. 2022.

YAMASHITA, Rikiya et al. **Convolutional Neural Networks: an overview and application in radiology.** Insights into imaging, v. 9, n. 4, p. 611-629, 2018.