



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

FACULDADE DE ENGENHARIA ELÉTRICA

PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA BIOMÉDICA

OTÁVIO JOAQUIM TAVARES

Avaliação numérica da dose ocupacional em
procedimentos intervencionistas utilizando métodos
de *deep learning* em tempo real

UBERLÂNDIA, MG

2024

Otávio Joaquim Tavares

**Avaliação numérica da dose ocupacional em
procedimentos intervencionistas utilizando métodos
de *deep learning* em tempo real**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Biomédica da Universidade Federal de Uberlândia, como parte dos requisitos necessários à obtenção do grau de Mestre em Ciências.

Área de Concentração: Tecnologias em Radiações,
Imagens Médicas e Biológicas

Orientadora: Profa. Dra. Ana Paula Perini

UBERLÂNDIA, MG

2024

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

T231 2024	<p>Tavares, Otávio Joaquim, 1997- Avaliação numérica da dose ocupacional em procedimentos intervencionistas utilizando métodos de deep learning em tempo real [recurso eletrônico] / Otávio Joaquim Tavares. - 2024.</p> <p>Orientadora: Ana Paula Perini. Dissertação (Mestrado) - Universidade Federal de Uberlândia, Pós-graduação em Engenharia Biomédica. Modo de acesso: Internet. Disponível em: http://doi.org/10.14393/ufu.di.2024.124 Inclui bibliografia.</p> <p>1. Engenharia biomédica. I. Perini, Ana Paula, 1981-, (Orient.). II. Universidade Federal de Uberlândia. Pós- graduação em Engenharia Biomédica. III. Título.</p> <p style="text-align: right;">CDU: 62:61</p>
--------------	--

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Biomédica				
Defesa de:	Dissertação de Mestrado Acadêmico, 105, PPGEB				
Data:	vinte e três de fevereiro de dois mil e vinte e quatro	Hora de início:	14:00	Hora de encerramento:	15:32
Matrícula do Discente:	12112EBI014				
Nome do Discente:	Otávio Joaquim Tavares				
Título do Trabalho:	Avaliação numérica da dose ocupacional em procedimentos intervencionistas utilizando métodos de deep learning em tempo real				
Área de concentração:	Engenharia Biomédica				
Linha de pesquisa:	Tecnologias em Radiações, Imagens Médicas e Biológicas				
Projeto de Pesquisa de vinculação:	Avaliação numérica das doses ocupacionais em radiodiagnóstico usando redes neurais convolucionais				

Reuniu-se via pela plataforma Google Meet, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Biomédica, assim composta: Professores Doutores: William de Souza Santos - INFIS/UFU; Diego Castanon Galeano - HJUM-UFMT/EBSERH; Ana Paula Perini - PPGEB/UFU orientadora do candidato.

Iniciando os trabalhos a presidente da mesa, Dra. Ana Paula Perini, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir a senhora presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Ana Paula Perini, Professor(a) do Magistério Superior**, em 23/02/2024, às 15:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Diego Castanon Galeano, Usuário Externo**, em 23/02/2024, às 15:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **William de Souza Santos, Professor(a) do Magistério Superior**, em 23/02/2024, às 15:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5206996** e o código CRC **C8E611D3**.

Referência: Processo nº 23117.013754/2024-01

SEI nº 5206996



TERMO DE RESPONSABILIDADE DE AUTORIA

Eu, Otávio Joaquim Tavares, discente do Curso de Pós-Graduação em Engenharia Biomédica, matrícula nº 12112EBI014, declaro, para os devidos fins e efeitos, junto a Universidade Federal de Uberlândia, que, sob as penalidades previstas no artigo 299 do Código Penal Brasileiro, e/ou demais leis sobre o tema, os produtos apresentados nos componentes curriculares que constituem na minha DISSERTAÇÃO DE MESTRADO, TESE DE DOUTORADO, ARTIGOS PUBLICADOS OU QUALQUER OUTRO PRODUTO QUE TENHA SIDO ORIGINADO DO TRABALHO DE PESQUISA não contêm plágio, não foram escritos por outrem (incluindo tecnologias de Inteligência artificial ou similares), ou representam compra fraudulenta de documentos científicos.

Por ser verdade, e por ter ciência do referido artigo, firmo a presente declaração.

Uberlândia - MG, 29. de julho de 2024.

Assinatura gov.br

AGRADECIMENTOS

Agradeço primeiramente a Deus por me ajudar todos os dias e por me proporcionar todo conhecimento durante todos esses anos.

À minha mãe Maria Cristina Tavares, que me incentivou durante todos os dias e nunca desistiu de mim.

À minha namorada Paula que durante todo esse caminho ela sempre esteve comigo e me ajudou todos os dias sem exceção.

Ao meu sogro Luiz Carlos Libânio, que sempre me incentivava a continuar estudando e sempre me ouvia com muita atenção e carinho.

Aos meus amigos Jeferson com suas observações frequentes, Julian que me incentivou a estudar a área de ciência de dados, Rafael e Igor que sempre estiveram presentes durante esse processo.

Agradeço imensamente à minha orientadora Profa. Dra. Ana Paula Perini e ao Prof. Dr. Lucio Pereira Neves, por sempre acreditarem no meu potencial, por todo o carinho, atenção e calma.

Este trabalho foi realizado com o apoio financeiro da **Universidade Federal de Uberlândia (UFU)**, e pelas agências de fomento:

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

- Bolsa de Produtividade em Pesquisa 312124/2021-0 (Ana Paula Perini);
- Bolsa de Produtividade em Pesquisa 314520/2020-1 (Lucio Pereira Neves)
- Programa MAI DAI 403556/2020-1, juntamente à M.R.A. - Indústria de Equipamentos Eletrônicos LTDA;
- Projeto UNIVERSAL 407493/2021-2;
- Este trabalho faz parte do Instituto Nacional de Ciência e Tecnologia para Instrumentação Nuclear e Aplicações na Indústria e Saúde (INCT/INAIS), projeto CNPq 406303/2022-3.

Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG

- APQ-04215-22;
- APQ-01254-23;
- APQ-04348-23.

RESUMO

Este estudo aborda a avaliação da dose nos órgãos/tecidos, fundamentada na posição espacial do médico durante o procedimento de radiologia intervencionista. O objetivo principal foi analisar a capacidade de previsão, considerando diferentes distâncias em relação ao tubo de raios-X, em diferentes órgãos/tecidos. Os métodos aplicados incluíram um modelo de previsão de baixo custo de treinamento, porém, com limitações específicas, destinado a um tipo de procedimento considerando pacientes imóveis, gênero fixo e condições estáticas do equipamento de raios-X. Restringir o movimento do médico dentro da área dos marcadores foi crucial para reduzir deformações na imagem. Os resultados destacaram que, em órgãos pequenos, como a próstata, ocorreram as maiores diferenças entre os valores de dose preditos e os calculados com o MCNP6.2. Os resultados ressaltam a viabilidade da integração do modelo *Multilayer Perceptron* (MLP), utilizado para fazer a predição de dose em cada órgão, com o modelo YOLOv4, para identificação em tempo real das posições espaciais do médico. Essa integração apresenta a vantagem de ser de baixo custo, sendo, assim, um recurso crucial para monitorar os valores de doses da equipe médica durante procedimentos de radiologia intervencionista, embora ainda existam oportunidades de melhoria identificadas para ambos os modelos.

- **Palavras-chave:** Dose de órgãos. Previsão espacial. Exposição médica. Modelos de imagem. Monitoramento durante procedimentos. Inteligência Artificial.

ABSTRACT

This study addresses the assessment of dose to organs/tissues, based on the spatial position of the physician during the interventional radiology procedure. The main objective was to analyze the prediction capacity, considering different distances in relation to the X-ray tube, in different organs/tissues. The methods applied included a prediction model with low training cost, however, with specific limitations, intended for a type of procedure considering immobile patients, fixed gender and static conditions of the X-ray equipment. Restricting the physician's movement within the marker area was crucial to reducing image deformation. The results highlighted that, in small organs, such as the prostate, the greatest differences occurred between the predicted dose values and those calculated with MCNP6.2. The results highlight the feasibility of integrating the *Multilayer Perceptron* (MLP) model, used to predict doses in each organ, with the YOLOv4 model, for real-time identification of the doctor's spatial positions. This integration has the advantage of being low cost, thus being a crucial resource for monitoring the dose values of the medical team during interventional radiology procedures, although there are still opportunities for improvement identified for both models.

Keywords: Organ Dose. Spatial Prediction. Medical Exposition. Imaging Models. Procedure Monitoring. Artificial Intelligence.

LISTA DE FIGURAS

Figura 1 – Arquitetura da MLP demonstrando as camadas e número de neurônios	24
Figura 2 – Composição de um neurônio.	24
Figura 3 – Comportamento das funções de ativação não lineares: ReLu, Tahn e Sigmóide.	26
Figura 4 – Funcionamento do algoritmo backpropagation	27
Figura 5 – Exemplo de detecção usando YOLO	33
Figura 6 – Exemplo de uma imagem de 448×448 dividida em células de grade de $S = 3$, representadas pelas linhas verdes, onde os componentes da caixa delimitadora do passáro, são representados pelas linhas vermelhas. As duas tuplas $(220, 190)$ e $(224, 143)$ representam o ponto central da imagem detectada e a largura e altura da caixa delimitadora, respectivamente	34
Figura 7 – Arquitetura geral de um detector de objetos em um estágio, utilizada para implementação da YOLOv4.	35
Figura 8 – Resumo da arquitetura desse projeto	38
Figura 9 – Exemplo de uma simulação de um médico na posição 365 cm no eixo x e 35 cm no eixo y.	39
Figura 10 – Modelagem da sala durante o procedimento de RI	40
Figura 11 – Modelagem da Rede para o problema proposto neste trabalho	41
Figura 12 – Padrão de Calibração Xadrez	42
Figura 13 – Área que limita a região de deslocamento do médico durante o procedimento de RI	43
Figura 14 – Simulação da movimentação do médico durante um procedimento de RI	44
Figura 15 – Boxplot da incerteza dos valores de dose para cada órgão do médico no procedimento de RI avaliado neste trabalho	45
Figura 16 – Histogramas das variáveis	47
Figura 17 – Boxplot das variáveis antes da normalização	48
Figura 18 – Dispersão de doses de acordo com a movimentação do médico ao se aproximar do tubo de raios-x	49
Figura 19 – Boxplot das variáveis após normalização	49
Figura 20 – Treinamento da Rede Neural	50
Figura 21 – Gráfico de dispersão nos dados de treinamento da Rede Neural	51
Figura 22 – Gráfico de dispersão nos dados de teste da Rede Neural	51
Figura 23 – Desempenho do modelo entre MCNP6.2 e previsão para o pulmão	53
Figura 24 – Dose na Pele em Relação à Movimentação do Médico durante o procedimento de RI. A fonte radioativa está localizada conforme mostrado na Figura 9.	54

LISTA DE TABELAS

Tabela 1 – Estatística Descritiva do banco de dados	46
Tabela 2 – Resultados Redes Neurais	50
Tabela 3 – Diferenças percentuais entre os valores de dose calculados pelo MCNP6.2 e preditos neste trabalho, para o médico.	52

LISTA DE ABREVIATURAS E SIGLAS

MCNP	<i>Monte Carlo N-Particle</i>
MC	Monte Carlo
MLP	<i>Multilayer Perceptron</i>
PA	Pósterio-anterior
RI	Radiologia Intervencionista
MASH	<i>Male virtual anthropomorphic phantoms</i>
RNA	Rede Neural Artificial
RNAs	Redes Neurais Artificiais
DL	<i>Deep Learning</i> (Aprendizagem Profunda)
FPS	<i>Frames per Second</i> (Frames por segundo)
IA	Inteligência Artificial
YOLO	<i>You Only Look Once</i>

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Objetivos	22
1.1.1	Objetivos específicos	22
2	REVISÃO DA LITERATURA	23
2.1	Redes Neurais	23
2.1.1	Funções de ativação	26
2.1.2	Treinamento e Validação	27
2.1.3	Algoritmo de Otimização	27
2.1.4	Épocas e Taxa de Aprendizado	28
2.1.5	Validação Cruzada	28
2.2	Configuração da captação das imagens	28
2.2.1	Marcadores ArUco	29
2.2.2	Detecção de objetos	30
2.2.3	Detecção de objetos em vídeo	31
2.2.4	YOLO	32
3	MATERIAIS E MÉTODOS	37
3.1	Visão geral do método do cálculo em tempo real de doses equivalentes	37
3.1.1	Estimativa de dose	37
3.2	Conjunto de dados	38
3.2.1	Configuração de simulação	39
3.3	Pré-processamento dos dados para rede de predição	40
3.4	Configuração do modelo MLP de predição de dose	41
3.5	Configuração da visão computacional	42
3.5.1	Calibração da Câmera de Vídeo	42
3.5.2	Delimitação da Área de Movimentação através de Marcadores AruCo	43
3.5.3	Ajustes do modelo YOLOv4	43
4	RESULTADOS E DISCUSSÕES	45
4.1	Modelo de predição de Dose	45
4.1.1	Análise estatística dos dados de entrada	45
4.1.2	Estatística Multivariada	48
4.1.3	Rede Neural <i>Multilayer Perceptron</i>	50
5	CONCLUSÃO	55

REFERÊNCIAS 57

1 INTRODUÇÃO

A Organização Mundial da Saúde aponta que a principal causa de morte no mundo é devido a doenças cardiovasculares. Em 2017, foram registradas 384.000 mortes no Brasil devido a doenças cardiovasculares (FERES *et al.*, 2017). No mundo, o total de óbitos relacionados a essas enfermidades chegou a 15,2 milhões (SILVA *et al.*, 2019). Portanto, procedimentos como os de Radiologia Intervencionista (RI) têm grande importância no diagnóstico no tratamento de lesões e doenças cardiovasculares.

A utilização dos métodos de Imagem tem aumentado consideravelmente ao longo dos anos. Cerca de metade da população global realiza, pelo menos, um exame de imagem radiológica por ano. Esses procedimentos apresentam vantagens significativas, tais como a possibilidade de realizar intervenções complexas com cirurgias de menor invasividade, reduzindo assim a probabilidade de infecções. Além disso, há uma alta taxa de recuperação dos pacientes e uma consequente diminuição dos custos hospitalares (CANEVARO, 2009).

Até a década de 1950, o único método com o melhor resultado para observar diretamente uma imagem de raios X era o método de telas fluorescentes, denominada técnica de fluoroscopia. Neste exame o feixe de raios X, transmitido pelo paciente, é absorvido em uma tela especial, que converte a radiação X em luz visível. Com o avanço da tecnologia foi introduzido um sistema de intensificação e um monitor. A imagem que era gerada na tela fluorescente foi ampliada pelo intensificador de imagem, sendo processada até chegar no monitor, resultando em imagens em tempo real (LUZ *et al.*, 2007).

A RI é compreendida como a área médica que emprega radiações ionizantes em procedimentos diagnósticos e terapêuticos, utilizando imagens de fluoroscopia para guiar os cateteres e encontrar o local de tratamento. Meios de contraste são administrados ao paciente para facilitar a distinção de estruturas anatômicas com baixo número atômico (CANEVARO, 2009).

Dentre todos os procedimentos em radiodiagnóstico, a RI é o que apresenta as maiores doses de radiação na equipe médica. Os principais fatores que contribuem para estas altas doses são o posicionamento inadequado do tubo em relação aos profissionais, e longos tempos de exposição (LUZ *et al.*, 2007).

Devido à alta dose de radiação sobre os profissionais, deve-se realizar um monitoramento da dose sobre os órgãos e tecidos. Porém, como é impossível realizar experimentalmente o monitoramento da dose interna, é utilizado o Método Monte Carlo (MC) para a dosimetria interna de forma computacional. O Método MC vem se destacando em diversas áreas, inclusive em Física Médica. Este pode ser descrito como um método estatístico, que utiliza números pseudo-aleatórios para simular o transporte da radiação na matéria, e

permitindo o cálculo de grandezas como a dose absorvida (YORIYAZ, 2009).

O algoritmo empregado foi o Monte Carlo MCNP6.2 (*Monte Carlo N-Particle*), contudo, ele manifesta algumas limitações, restringindo-se à condução exclusiva de simulações estáticas. Porém, um procedimento de RI é demorado e as posições e parâmetros dos equipamentos são alterados durante o tempo. Para resolver esse problema, neste trabalho foi utilizado o aprendizado de máquina com redes neurais de forma inédita. O aprendizado de máquina com redes neurais está sendo um dos maiores avanços tecnológicos devido sua ampla área de aplicações, e dentro destes, as Redes Neurais Convolucionais apresentam os melhores resultados em relação a visão computacional, com reconhecimento em vídeos e imagens (VARGAS; PAES; VASCONCELOS, 2016).

1.1 Objetivos

O propósito primordial deste projeto consistiu em determinar em tempo real as doses absorvidas pelos profissionais em RI. Utilizando o método de Monte Carlo, foi criado um ambiente computacional que permite a obtenção de um banco de dados contendo todas as doses recebidas pelos profissionais. Por meio da aplicação da rede neural convolucional *You Only Look Once* (YOLO), foi possível identificar o médico durante o procedimento e utilizar essa informação de posicionamento como parâmetro de entrada para a rede de predição de doses. Esta rede foi treinada utilizando um banco de dados gerado computacionalmente, possibilitando previsões precisas das doses em tempo real (LI *et al.*, 2022).

1.1.1 Objetivos específicos

- Revisão de literatura de conceitos de inteligência artificial e método Monte Carlo;
- Construção do cenário utilizado no código de Monte Carlo MCNP6.2;
- Inserção dos objetos simuladores antropomórficos virtuais nos cenários do código MCNP6.2;
- Variação dos parâmetros da simulação;
- Criação de um banco de dados com as doses de radiação para a equipe médica;
- Criação de um algoritmo de determinação de doses em tempo real;
- Detecção da equipe médica com algoritmo YOLOv4;
- Comunicação em tempo real entre modelo YOLOv4 com o modelo de predição de dose;
- Apresentação das doses equivalentes em tempo real.

2 REVISÃO DA LITERATURA

2.1 Redes Neurais

Redes neurais artificiais são inspiradas no funcionamento do sistema nervoso humano e foram descritas inicialmente em (MCCULLOCH; PITTS, 1943) como um modelo capaz de representar funções não lineares.

Na representação da Figura 1, cada nó representa um neurônio, e cada aresta denota uma conexão entre neurônios, onde a saída de um neurônio à esquerda serve como entrada para o neurônio à direita. A determinação do número de neurônios na camada intermediária desempenha um papel crucial, influenciando na precisão das previsões, como destacado por (CABRERA; ELBULUK; HUSAIN, 1997).

As redes neurais são formadas por neurônios, onde cada neurônio é fundamentalmente uma função matemática que visa ponderar os dados que lhe são fornecidos como entrada. Estes dados são multiplicados por valores denominados de pesos sinápticos com o objetivo de individualizar o comportamento de cada neurônio. A Figura 2 ilustra a estrutura de um neurônio.

Na Figura 2 é possível observar os pesos associados a cada entrada, que serão usados para gerar os valores base para a função de transferência do neurônio. A função de transferência geralmente é linear, como por exemplo uma soma de fatores. A equação que define de forma geral o funcionamento de um neurônio é dada pela equação 2.1:

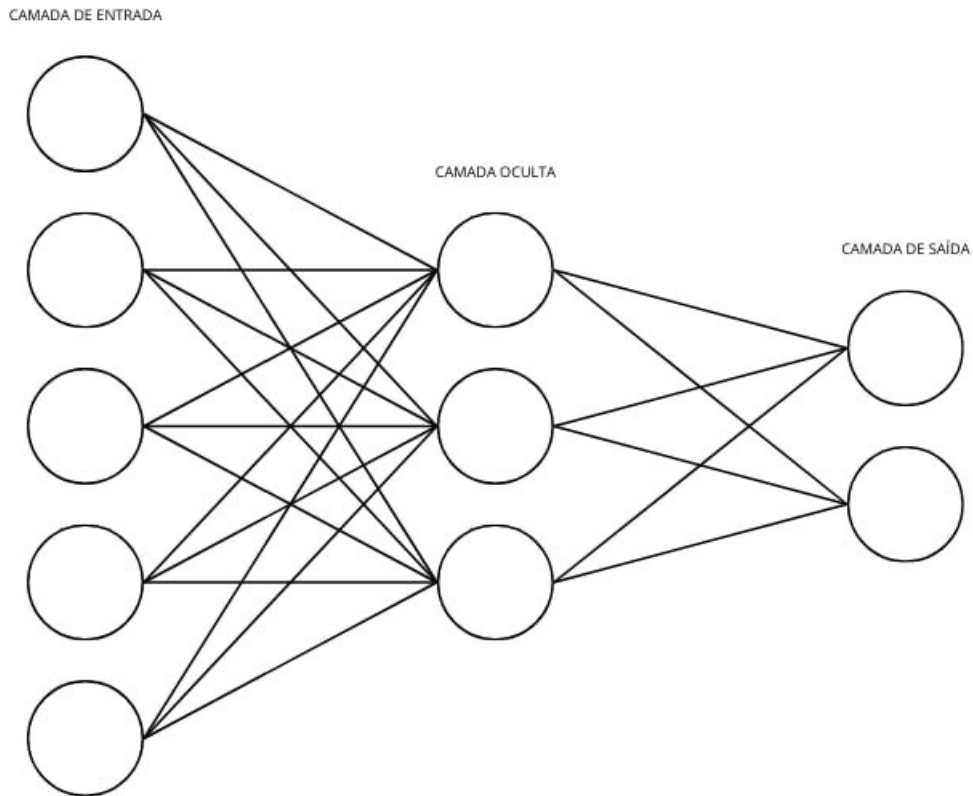
$$x = \sum_{n=1}^N u_n \cdot w_n \quad (2.1)$$

Onde: x é o resultado da soma da multiplicação dos pesos sinápticos, w , pelos valores de entrada da rede u .

Após a função de transferência existe a função de ativação que tem a propriedade de restringir os valores gerados pelo neurônio, essa função é importante para auxiliar na classificação dos sinais emitidos pela rede neural.

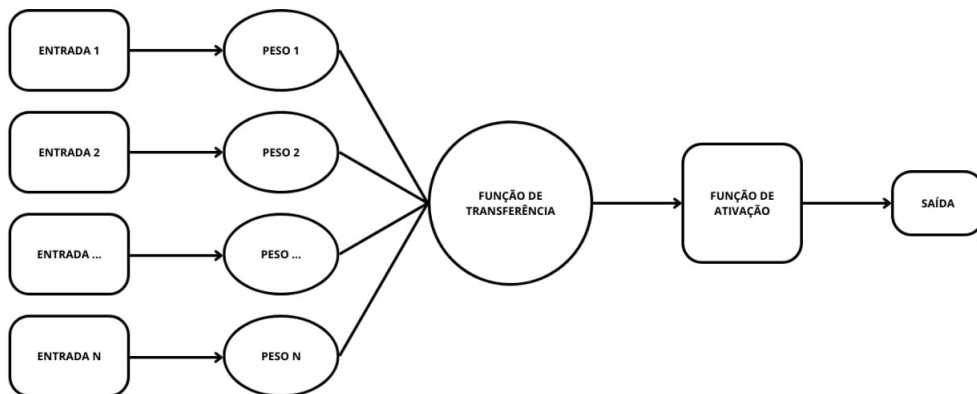
O desempenho das funções de ativação é medido principalmente pela capacidade da rede em classificar corretamente os dados em sua etapa final. É possível rastrear uma classificação errada devido a uma característica não representativa, que é propagada ao longo da rede. Algumas funções como a sigmóide, Equação 2.2, possuem a característica de zerar as saídas conforme seus gradientes são reduzidos, fazendo com que a rede não aprenda nessa faixa de valores. Para tentar contornar essa característica foram propostas funções como a Tanh na Equação 2.3, que varia seus valores entre -1 e 1, solucionando o

Figura 1 – Arquitetura da MLP demonstrando as camadas e número de neurônios



Fonte: Elaborada pelo autor.

Figura 2 – Composição de um neurônio.



Fonte: Elaborada pelo autor.

problema de zerar as saídas com valores muito baixos de gradiente.

Como essa função tem um papel fundamental no desempenho das redes neurais, diversas funções foram apresentadas com o objetivo de aprimorar a performance das redes (SHARMA; SHARMA; ATHAIYA, 2017). Dentre as funções comumente utilizadas vale destacar as seguintes:

- Sigmóide: a saída possui valores entre 0 e 1:

$$\sigma(y) = \frac{1}{1 + e^{-y}} \quad (2.2)$$

- Tangente hiperbólica (Tahn): a saída possui valores entre -1 e 1:

$$\tanh(y) = \frac{e^y - e^{-y}}{e^y + e^{-y}} \quad (2.3)$$

- Linear retificada (ReLU): a saída possui valores entre 0 e o máximo fornecido a função:

$$\text{ReLU}(y) = \begin{cases} ay, & \text{se } y > 0, \\ 0, & \text{caso contrário.} \end{cases} \quad (2.4)$$

Onde:

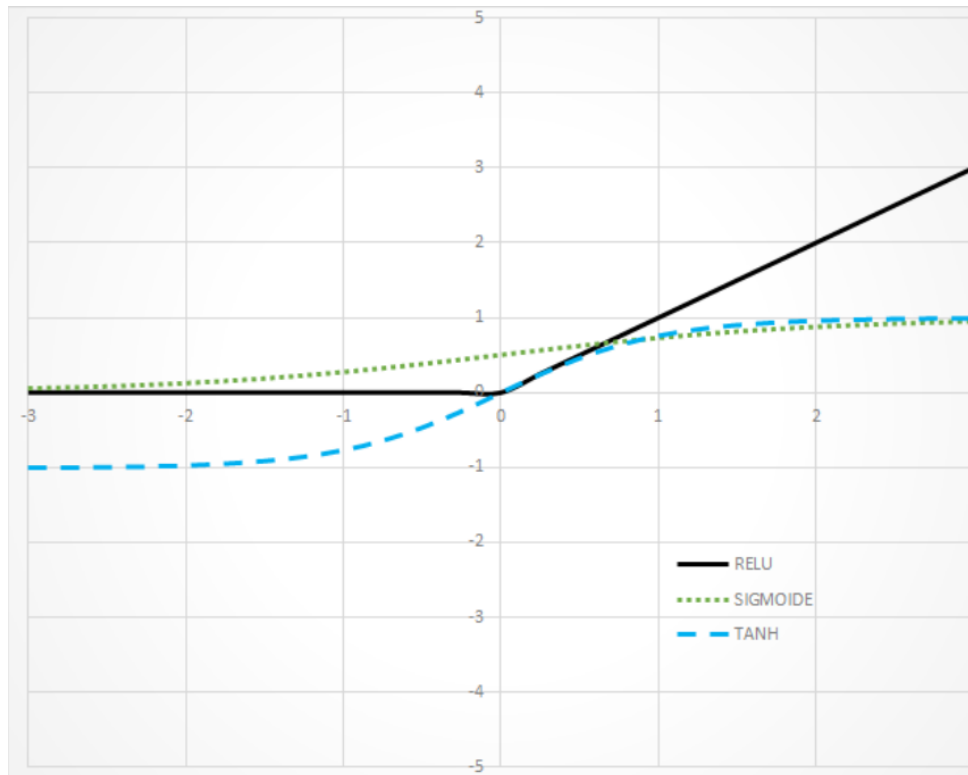
y é o valor calculado após a função de transferência do neurônio em questão e a corresponde ao parâmetro de inclinação da função, na função tangente hiperbólica Equação 2.3.

Segundo Neogi, Mohanta e Dutta (2014) uma função com bons resultados é a função linear retificada (ReLU), Equação 2.4, que possui a propriedade de tornar as conexões dos neurônios esparsas, dessa forma ativando apenas regiões significantes que podem ser responsáveis para a correta interpretação dos dados de entrada. O comportamento das funções de ativação ReLU (Equação 2.4), sigmóide (Equação 2.2) e tangente (Equação 2.3) pode ser observado na Figura 3.

Dentro do contexto de redes neurais existem aquelas denominadas de redes neurais de retro-propagação. Nas quais, através de um aprendizado supervisionado, os erros de classificação são utilizados para recalculer os valores dos pesos sinápticos através de uma função de correção. Um exemplo de rede neural de retro-propagação é a rede Perceptron (ROSENBLATT, 1958). A composição de uma rede Perceptron pode ser observada na Figura 1.

As redes Perceptron podem possuir inúmeras camadas entre a entrada de dados e a camada de saída, dependendo do objetivo e complexidade do projeto desenvolvido. Observa-se que redes mais profundas, com mais camadas internas, têm a capacidade de aprender conceitos e relações mais complexas entre os dados de entrada.

Figura 3 – Comportamento das funções de ativação não lineares: ReLu, Tahn e Sigmóide.



Fonte: (CHA; CHOI; BÜYÜKÖZTÜRK, 2017)

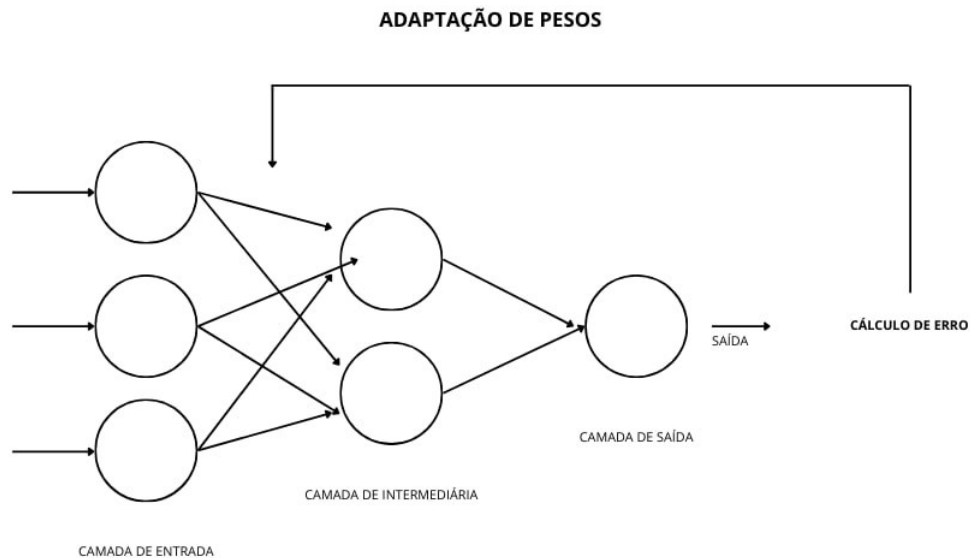
Para otimizar o treinamento de uma MLP, é essencial dividir os dados em conjuntos de treinamento, validação e teste. O conjunto de treinamento é empregado para ajustar os pesos da MLP, enquanto o conjunto de validação é crucial para mitigar o “*overfitting*”. Este fenômeno ocorre quando a MLP memoriza padrões de treinamento, mas falha em generalizar para novas situações. O conjunto de teste é reservado exclusivamente para avaliar a solução final e confirmar a capacidade preditiva da MLP, conforme discutido por Kartam, Flood e Garrett (1997).

O algoritmo de backpropagation destaca-se como uma das ferramentas mais utilizadas para o treinamento de MLPs. Sua eficácia reside na capacidade de ajustar os pesos da MLP de acordo com o comportamento da função de erro, conforme ilustrado na Figura 4.

2.1.1 Funções de ativação

As MLPs incorporam o uso de funções de ativação para determinar a saída de um neurônio com base em entradas específicas. Existem dois tipos fundamentais de funções de ativação: a função de limiar e a função sigmóide, conforme mencionado por Haykin (HAYKIN, 2009). A função de ativação limiar estabelece um valor de limite, resultando em uma saída binária de 0 ou 1. Já a função de ativação tangente hiperbólica tem um intervalo de variação entre -1 e 1, enquanto a função de ativação sigmóide logística varia de 0 a 1. Devido à facilidade computacional de derivadas e à não linearidade, as funções

Figura 4 – Funcionamento do algoritmo backpropagation



Fonte: Elaborada pelo autor.

de ativação sigmóide são frequentemente adotadas em redes neurais, conforme observado por Han e Moraga (HAN; MORAGA, 1995). A equação 2.2 descreve a função sigmóide, a qual foi selecionada para este estudo, aplicável a todos os valores reais de x .

2.1.2 Treinamento e Validação

O sucesso de uma rede neural, como a MLP, depende significativamente do processo de treinamento e validação. O treinamento envolve ajustar os pesos e bias da rede com base em um conjunto de dados de treinamento para minimizar a discrepância entre as saídas desejadas e as saídas reais da rede. Já a validação é essencial para avaliar o desempenho do modelo em conjuntos de dados não vistos durante o treinamento, verificando se a rede generaliza bem para novos exemplos.

2.1.3 Algoritmo de Otimização

Durante o treinamento, um algoritmo de otimização é utilizado para ajustar os parâmetros da rede de forma a minimizar a função de custo. Um dos algoritmos amplamente utilizados é o Gradiente Descendente, que ajusta os pesos na direção oposta ao gradiente da função de custo em relação aos pesos. Existem variantes do Gradiente Descendente, como o Gradiente Descendente Estocástico (*Stochastic Gradient Descent* SGD) e o Gradiente Descendente em Lote (*Batch Gradient Descent* da sigla BGD), cada um com suas características específicas (HUH; SEJNOWSKI, 2018).

2.1.4 Épocas e Taxa de Aprendizado

O treinamento é realizado por meio de iterações chamadas épocas, em que todo o conjunto de treinamento é processado uma vez pela rede. A taxa de aprendizado é um hiperparâmetro que determina o tamanho do passo que os pesos são ajustados durante o treinamento. Um valor adequado de taxa de aprendizado é crucial, pois valores muito altos podem resultar em convergência inadequada, enquanto valores muito baixos podem prolongar excessivamente o processo de treinamento.

2.1.5 Validação Cruzada

Para avaliar a capacidade de generalização da MLP, é comum empregar técnicas de validação cruzada. A validação cruzada divide o conjunto de dados em partes, treinando o modelo em subconjuntos e avaliando-o em conjuntos distintos. Isso proporciona uma estimativa mais robusta do desempenho do modelo em dados não vistos.

2.2 Configuração da captação das imagens

Neste trabalho foi utilizado apenas uma única câmera para realizar o monitoramento da movimentação do médico na sala de RI, a câmera foi colocada em uma posição fixa, e calibrada com 4 pontos fixados na sala.

As câmeras pinhole mais baratas de hoje introduzem muita distorção nas imagens. Duas distorções principais são a distorção tangencial e a distorção radial.

Para corrigir essas irregularidades, Duane (1971) propôs duas estratégias: para as distorções radiais, ele sugeriu o uso de uma expansão em Série de Taylor de três termos k_1 , k_2 e k_3 , onde esses coeficientes representam as distorções; enquanto para as distorções tangenciais, foram introduzidos os coeficientes p_1 e p_2 .

Para resolver a distorção radial é resolvida da seguinte forma:

$$X_{\text{corrigido}} = x(1 + k_1r^2 + k_2r^4 + k_3r^6) \quad (2.5)$$

$$Y_{\text{corrigido}} = y(1 + k_1r^2 + k_2r^4 + k_3r^6) \quad (2.6)$$

Onde: k_1 , k_2 e k_3 são coeficientes de distorção radiais.

Da mesma forma, outra distorção é a distorção tangencial que ocorre porque a lente de captura de imagem não está alinhada perfeitamente paralelamente ao plano de imagem. Portanto, algumas áreas da imagem podem parecer mais próximas do que o esperado, sendo resolvida conforme as equações (2.8) e (2.9).

$$X_{\text{correto}} = x + [2p_1xy + p_2(r^2 + 2x^2)] \quad (2.7)$$

$$Y_{\text{correto}} = y + [p_1(r^2 + 2y^2) + 2p_2xy] \quad (2.8)$$

Onde: p_1 e p_2 são coeficientes de distorção tangenciais.

Resumindo, é necessário encontrar cinco parâmetros, conhecidos como coeficientes de distorção dados por:

$$\text{Coeficientes de distorção} = (k_1, k_2, p_1, p_2, k_3) \quad (2.9)$$

Além disso, precisa-se encontrar mais algumas informações, como parâmetros intrínsecos e extrínsecos de uma câmera. Os parâmetros intrínsecos são específicos de uma câmera, incluindo informações como distância focal (f_x, f_y), centros ópticos (c_x, c_y). Depende apenas da câmera, portanto, uma vez calculado, pode ser armazenado para fins futuros, e é expresso como uma matriz 3x3:

$$\text{Câmera matrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.10)$$

Os parâmetros extrínsecos correspondem a vetores de rotação e translação que traduzem as coordenadas de um ponto 3D em um sistema de coordenadas.

Para aplicações estéreo, estas distorções precisam ser corrigidas primeiro. Para encontrar todos esses parâmetros, deve-se fornecer algumas imagens de amostra de um padrão bem definido, conhecendo suas coordenadas na imagem. Com esses dados, algum problema matemático é resolvido em segundo plano para obter os coeficientes de distorção. Para melhores resultados foram realizadas 50 padrões de teste neste trabalho.

2.2.1 Marcadores ArUco

A estimativa de posição é uma tarefa amplamente empregada em várias aplicações, como visão computacional e realidade aumentada, entre outras. Corresponder pontos no mundo real com suas coordenadas de projeção na imagem 2D é uma tarefa complexa, geralmente abordada com o uso de modelos desenvolvidos especificamente para esse propósito, como os marcadores sintéticos ou fiduciais. Um exemplo comum são os marcadores quadrados binários, como os marcadores ArUco, concebidos por Garrido e Nicholson (2021).

Esses marcadores são quadrados sintéticos compostos por uma borda preta e uma matriz binária interna, responsável por determinar seu identificador (GARRIDO; NICHOLSON, 2021).

Os marcadores ArUco são frequentemente utilizados em pesquisas para implementar funcionalidades em algoritmos de visão computacional, explorando suas capacidades e limitações quando comparados à instrumentação convencional (PHAM *et al.*, 2021). Por exemplo, Pham *et al.* (2021). desenvolveram um sistema de medição de nivelamento entre superfícies planas usando marcadores ArUco. Essa abordagem tem aplicações práticas em indústrias automotivas para alinhamento de peças mecânicas.

Lupu *et al.* (2017) desenvolveram um sistema de realidade virtual para reabilitação pós-AVC (Acidente Vascular Cerebral) em membros superiores, empregando marcadores ArUco. Esse sistema proporciona um retorno visual ao paciente durante as atividades de reabilitação motora. A parte do corpo em reabilitação é monitorada por uma câmera que rastreia os movimentos dos marcadores fixados nessa região específica. As imagens são processadas em tempo real, apresentando um avatar do membro ao paciente.

Além disso, Roos-Hoefgeest, Garcia e Gonzalez (2021) propuseram um sistema para localização de robôs móveis. Esse sistema envolve o uso de um anel de câmeras, instalação de marcadores ArUco no ambiente e um algoritmo de mapeamento. A proposta visa substituir o uso de sensores a laser, devido ao seu alto custo, e podendo ser aplicada na indústria para inspeção e reparo de chapas de aço.

2.2.2 Detecção de objetos

A área de visão computacional engloba um conjunto diversificado de tarefas relacionadas à detecção de objetos em imagens digitais (FELZENSZWALB *et al.*, 2009). Essa técnica visa a identificação e localização de objetos em imagens ou vídeos, com aplicações amplas, tais como detecção de pedestres em ambientes urbanos (SUNG; POGGIO, 1998), reconhecimento facial (DOLLAR *et al.*, 2011), identificação de veículos em contextos de tráfego (PEPPA *et al.*, 2018; DU; ANG; RUS, 2017), e sistemas de contagem automática (CHEN *et al.*, 2018; SANTANA *et al.*, 2019).

Considerada um dos problemas fundamentais na visão computacional, a detecção de objetos provê informações cruciais para a compreensão semântica de imagens e vídeos, encontrando aplicabilidade em diversas áreas, tais como classificação de imagens (VEDALDI *et al.*, 2014; KRIZHEVSKY; SUTSKEVER; HINTON, 2012), análise do comportamento humano (CAO *et al.*, 2017) e reconhecimento facial (YANG; NEVATIA, 2016). Ao identificar e localizar objetos, essa técnica permite a contagem precisa de elementos em uma cena e o rastreamento de suas posições.

Devido à diversidade de pontos de vista, poses, oclusões e condições de iluminação,

a detecção de objetos, especialmente quando combinada com a localização, apresenta desafios significativos. Isso tem despertado considerável atenção e pesquisa nesse campo nos últimos anos (REDMON *et al.*, 2016; GIRSHICK *et al.*, 2014; GIRSHICK, 2015; REN *et al.*, 2015).

A detecção de objetos abrange a tarefa de localizar e categorizar cada objeto em uma imagem específica, combinando a identificação da localização e a classificação do objeto. Os modelos tradicionais de detecção de objetos geralmente seguem um *pipeline* composto por três estágios principais: seleção de região informativa, extração de características e classificação.

Na primeira etapa, conhecida como seleção de região informativa, os objetos podem surgir em diferentes posições e proporções na imagem. Para lidar com essa variação, é comum empregar uma análise da imagem completa usando uma técnica de "janela deslizante" em várias escalas. No entanto, essa abordagem torna-se onerosa computacionalmente devido ao grande número de janelas candidatas, gerando redundâncias. Mesmo ao utilizar um número fixo de modelos de "janela deslizante", podem surgir regiões inadequadas (COMASCHI *et al.*, 2013).

O estágio subsequente, a extração de características, visa reduzir uma imagem de tamanho variável a um conjunto fixo de atributos visuais, essenciais para contextualizar ou conferir significado à imagem. Tanto os modelos tradicionais de visão computacional quanto os métodos de aprendizagem profunda compartilham o objetivo de extrair atributos da imagem de entrada.

Nos sistemas de detecção de objetos, é comum utilizar modelos pré-treinados de classificação de imagens para a extração de características. Em seguida, o processo envolve a previsão da classe de um objeto em uma região específica da imagem principal. A detecção é responsável por determinar com precisão a localização do objeto na imagem, delimitando-o por uma caixa e classificando-o de acordo com as classes aprendidas durante o treinamento. Essa abordagem é crucial para a detecção instantânea em visão computacional, como será discutido adiante.

2.2.3 Detecção de objetos em vídeo

No início da década de 2010, os avanços pioneiros na visão computacional focaram principalmente no processamento de imagens, como classificação (JIA *et al.*, 2014; KRIZHEVSKY; SUTSKEVER; HINTON, 2012), detecção (LONG; SHELHAMER; DARRELL, 2015; CHEN *et al.*, 2016; RIBEIRO *et al.*, 2016; YANG; CHOI; LIN, 2016), e segmentação (LONG; SHELHAMER; DARRELL, 2015; JUNG; KIM, 2011; GIDARIS; KOMODAKIS, 2015; ZHU *et al.*, 2015). Em contraste, o processamento de vídeo recebeu menos atenção nesse período.

Isso se deve, em parte, à natureza do vídeo, essencialmente uma sequência de imagens chamadas de quadros (*frames* em inglês). Contudo, essa definição simplificada não abarca toda a complexidade do processamento de vídeo, que adiciona uma dimensão temporal ao problema.

Os vídeos não consistem apenas em sequências de imagens, mas sim em sequências de imagens inter-relacionadas. Essa distinção sutil abre oportunidades para explorar a dimensão temporal de maneira que não se aplicam a imagens individuais. Adicionalmente, devido à complexidade dos dados em vídeo, como o tamanho do conjunto de dados e as anotações relacionadas, e o alto custo de treinamento e inferência, avançar nesse campo tem sido desafiador.

Recentemente, no entanto, houve um aumento significativo de pesquisas relacionadas ao processamento de vídeo, especialmente com o lançamento do ImageNet VID (RUSSAKOVSKY *et al.*, 2015) e outros conjuntos de dados de vídeo na segunda metade da década de 2010. Surgiram abordagens que envolvem métodos de pós-processamento (LI *et al.*, 2017; MA; ZHANG; BULL, 2020; SANTAMARIA *et al.*, 2021), técnicas multi-frame (WEN; DING; XU, 2021; CHOI *et al.*, 2017) e uso de fluxo ótico (HUANG *et al.*, 2021b; TIAN *et al.*, 2020). Entre essas, destaca-se o modelo YOLO, proposto por Redmon *et al.* (2016), que, em sua simplicidade, prevê caixas delimitadoras e probabilidades de classe com uma única rede em uma única avaliação, permitindo previsões em tempo real.

2.2.4 YOLO

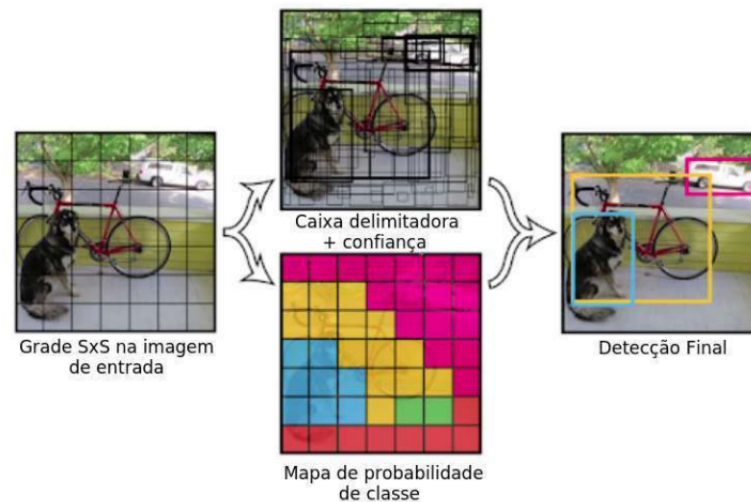
O YOLO (do inglês *You Only Look Once*) utiliza Redes Neurais Convolucionais (CNN) para a detecção de objetos, abordando essa tarefa como um único problema de regressão. Com isso, uma única olhada na imagem por meio da CNN permite prever as caixas delimitadoras e a confiança das classes dos objetos.

O diferencial desse método, como o nome sugere, está na abordagem de única visualização da imagem, tornando o YOLO um dos algoritmos mais rápidos para detecção já criados. Contudo, essa rapidez pode comprometer parte da precisão do algoritmo. Ainda assim, o YOLO é capaz de realizar detecções em tempo real a uma taxa de até 30 quadros por segundo (FPS) (REDMON *et al.*, 2016).

O processo de detecção ocorre ao dividir a imagem em uma grade de células $S \times S$. Essas células variam de tamanho de acordo com a versão do YOLO, podendo ser 3×3 , 5×5 ou 19×19 . Cada célula prevê N caixas delimitadoras possíveis e a confiança associada a cada uma delas. Isso resulta em um total de $S \times S \times N$ previsões de caixas delimitadoras. Para filtrar as detecções, o algoritmo define um limiar de confiança, descartando as caixas com valores abaixo desse limiar (JIANG *et al.*, 2022).

A probabilidade das classes é detectada para cada célula, independentemente do

Figura 5 – Exemplo de detecção usando YOLO



Fonte: (ARAÚJO, 2004)

número de caixas delimitadoras, permitindo apenas uma classe detectada por célula. Se o centro de um objeto estiver dentro de uma célula da grade, essa célula é responsável por detectar aquele objeto.

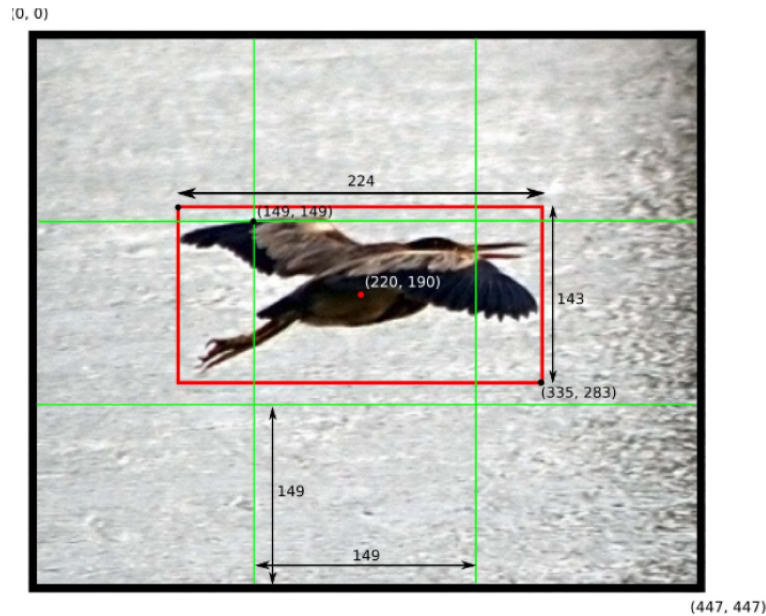
Para resolver o problema de várias detecções em uma única célula, é utilizado um mapa de probabilidades de classes. Esse mapa assegura que apenas uma classe seja detectada por célula, independentemente do número de caixas delimitadoras presentes.

Após a identificação das caixas sobrepostas, um processo de supressão é aplicado para eliminar possíveis objetos duplicados, resultando na retenção das detecções mais precisas. As caixas delimitadoras finais são determinadas a partir das extremidades definidas nas células do mapa de probabilidades, como demonstrado na figura 5 (REDMON *et al.*, 2016). No entanto, devido ao princípio de identificar apenas uma classe por célula da grade, a primeira versão da YOLO apresenta limitações quando objetos estão muito próximos na imagem, podendo falhar em detectar todos eles (REDMON *et al.*, 2016).

À medida que novas versões foram sendo lançadas, foram realizadas melhorias para lidar com as limitações da YOLO original. No entanto, essas atualizações introduziram novos desafios. Por exemplo, a YOLOv3 (REDMON; FARHADI, 2018), com previsões em múltiplas escalas, apresenta um desempenho geralmente superior, mas pode ter um desempenho inferior na detecção de objetos de tamanho médio e grande em relação ao tamanho total da imagem.

Cada caixa delimitadora gera diferentes métricas, como a pontuação de interseção sobre a união (IoU), que avalia a precisão da caixa, além de outros quatro elementos: (x, y, w, h) . O par (x, y) identifica o ponto central da caixa delimitadora, enquanto (w, h) denota sua largura e altura, conforme ilustrado na Figura 6. Esses quatro parâmetros

Figura 6 – Exemplo de uma imagem de 448×448 dividida em células de grade de $S = 3$, representadas pelas linhas verdes, onde os componentes da caixa delimitadora do passáreo, são representados pelas linhas vermelhas. As duas tuplas $(220, 190)$ e $(224, 143)$ representam o ponto central da imagem detectada e a largura e altura da caixa delimitadora, respectivamente



Fonte: (REDMON; FARHADI, 2018)

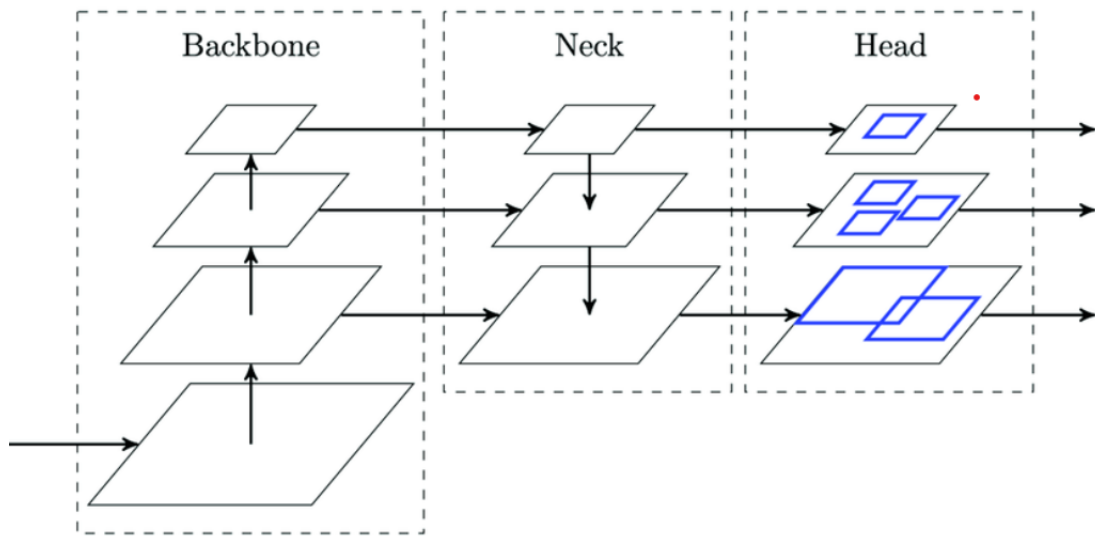
compõem as informações de posição e dimensões da caixa delimitadora identificada pela rede.

A arquitetura da YOLO é estruturada em três estágios: *Backbone*, *Neck* e *Head*, como ilustrado na Figura 7. O *Backbone* consiste principalmente em camadas de convolução em uma rede neural profunda, responsável por extrair os atributos fundamentais. A escolha do *Backbone* é crucial para aprimorar o desempenho da detecção de objetos, geralmente utilizando redes neurais pré-treinadas para esse fim (JIANG *et al.*, 2022).

O *Neck* tem a função crucial de coletar mapas de características de diferentes estágios. Por fim, o *Head*, realiza previsões densas, gerando um vetor com as coordenadas da caixa delimitadora detectada (centro, altura, largura), a pontuação de confiança e o rótulo (JIANG *et al.*, 2022).

A evolução dos algoritmos YOLO resultou em diversas versões desde o lançamento original em 2016. Desenvolvidas por Joseph Redmon, as versões YOLOv2 e YOLOv3 focaram em melhorar a precisão, localização e eficiência. A YOLOv2 introduziu mudanças na arquitetura, como normalização de batch em todas as camadas convolucionais e um classificador com maior resolução durante o treinamento (SANG *et al.*, 2018). A YOLOv3, por sua vez, prevê a pontuação de confiança para cada caixa delimitadora usando regressão logística e introduziu grades menores $S \times S$ nas imagens de entrada para melhorar a

Figura 7 – Arquitetura geral de um detector de objetos em um estágio, utilizada para implementação da YOLOv4.



Fonte: (SCIUS-BERTRAND *et al.*, 2021)

detecção (ZHAO; LI, 2020).

Versões posteriores, como YOLOv4, YOLOv5, YOLOv6, YOLOv7 e PP-YOLOv2, baseadas na arquitetura original, foram implementadas com modificações arquiteturais e implementações no framework PyTorch, resultando em melhorias no desempenho em termos de velocidade e precisão (LI *et al.*, 2022; WANG; BOCHKOVSKIY; YOLO, 2022; LONG *et al.*, 2020; HUANG *et al.*, 2021a).

A escolha da YOLOv4 para este trabalho se deu por representar o estado-da-arte em detecção de objetos no momento de sua elaboração. Embora, versões mais recentes tenham sido desenvolvidas, a mudança de modelo implicaria em desafios adicionais, como estudo dos modelos, configuração de parâmetros e comparação de métricas.

3 MATERIAIS E MÉTODOS

3.1 Visão geral do método do cálculo em tempo real de doses equivalentes

A metodologia empregada neste trabalho visa calcular com rapidez e precisão a dose de radiação nos órgãos da equipe médica durante procedimentos de RI. Foi proposto o desenvolvimento de uma rede neural *Multilayer Perceptron* (MLP) treinada com uma ampla variedade de simulações em MC, utilizando objetos simuladores voxelizados como dados de entrada e considerando as posições da equipe médica.

Os dados de entrada para este modelo incluem não apenas parâmetros das simulações, mas também a localização (x, y) do médico durante o procedimento. Adicionalmente, um modelo de rede neural convolucional foi empregado para identificar em tempo real a posição do médico durante o procedimento. Utilizando pontos de calibração posicionados no chão da sala, foi possível determinar as coordenadas da posição do médico. Essas coordenadas espaciais são então utilizadas como entrada na rede neural MLP para cálculo da dose.

A saída da rede neural MLP fornece a estimativa da dose de radiação nos órgãos do médico. É fundamental ressaltar que ambos os modelos têm a capacidade de operar em tempo real durante o procedimento, mantendo uma elevada precisão, validada por simulações em MC especialmente conduzidas para validar este sistema. A Figura 8 apresenta a arquitetura básica deste projeto.

3.1.1 Estimativa de dose

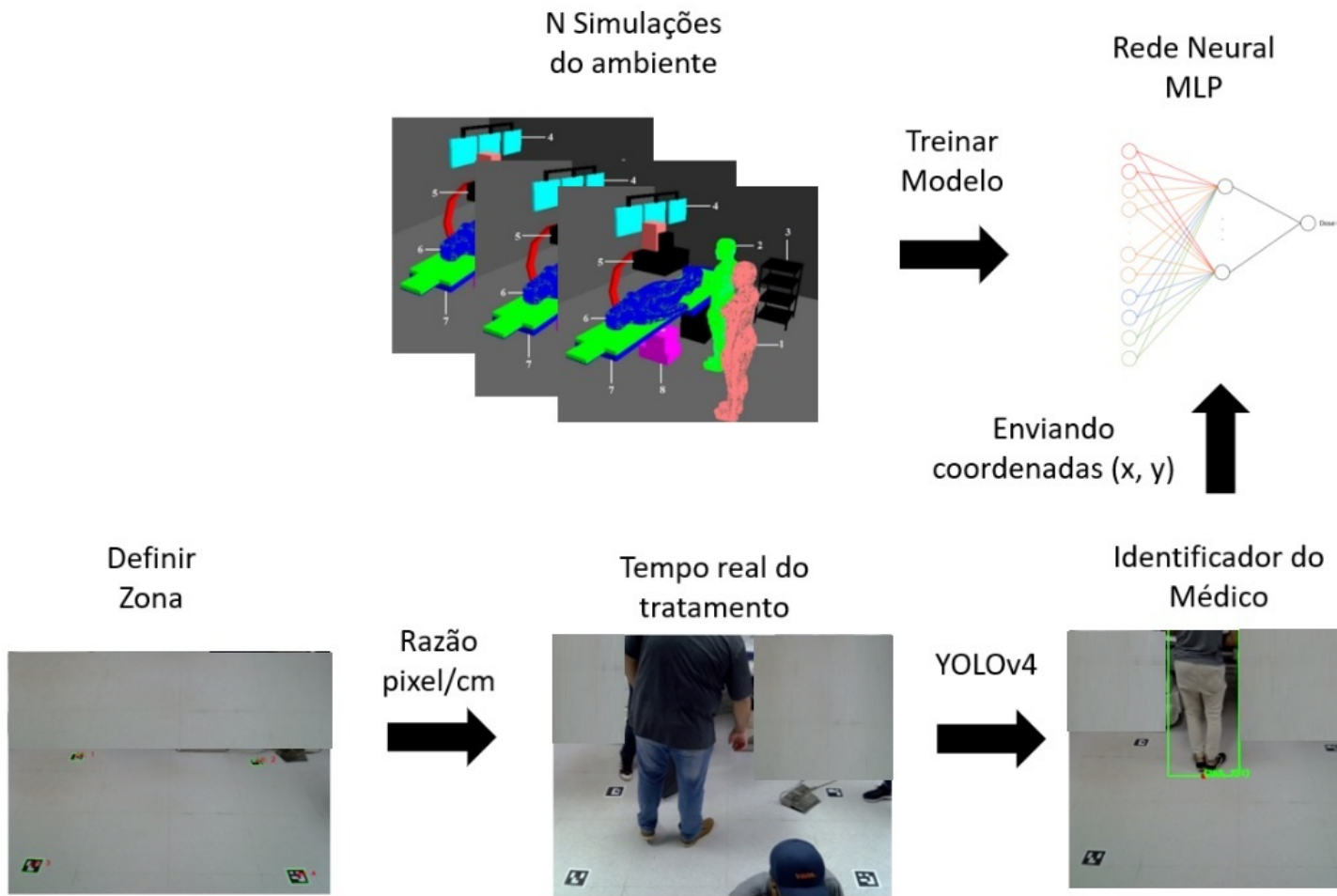
Considerando a natureza computacional do estudo, os coeficientes de conversão (CC), foram empregados para correlacionar os resultados da simulação às grandezas medidas experimentalmente, como Produto Kerma-Área (PKA). O $CC[H_t]$ representa a dose equivalente, estabelecendo uma relação entre o valor obtido computacionalmente e o valor experimental através da equação 3.1

$$CC[H_t] = \frac{H_t}{PKA} \quad (3.1)$$

Onde:

- H_t : Valor do tally $F6$ (obtido da simulação computacional) de cada órgão, multiplicado pelo fator de ponderação da radiação w_R , igual a 1 para raios-X.
- PKA : Valor do tally $F6$ para PKA (obtido da simulação computacional).

Figura 8 – Resumo da arquitetura desse projeto



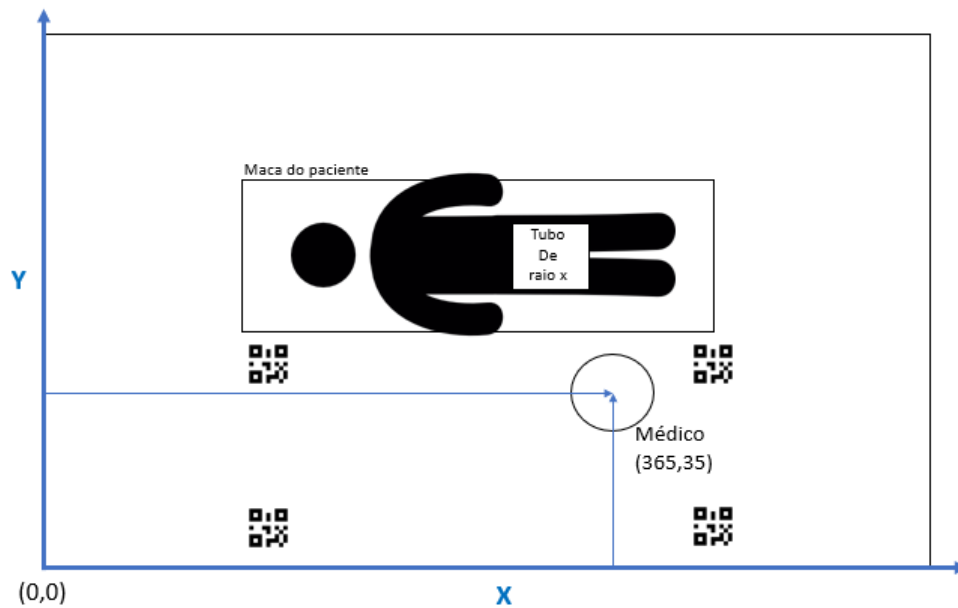
Fonte: Elaborada pelo autor.

3.2 Conjunto de dados

O conjunto de dados utilizado para a predição da dose foi composto por um total de 122 simulações geradas por meio do código de Monte Carlo MCNP6.2 (do inglês *Monte Carlo N-Particle*) (WERNER *et al.*, 2017) que é mantido pelo laboratório *Los Alamos*. Essas simulações foram especificamente conduzidas para simular o procedimento de embolização de mioma uterino, empregando um gantry fixo na posição pósterio-anterior (PA) e uma tensão constante no tubo de raios-X de 75 kVp, com um campo de $14,1 \times 14,1 \text{ cm}^2$. No entanto, vale ressaltar que a posição do médico foi variada dentro de um amplo intervalo, abrangendo desde 0 até 224 cm no eixo x, com um espaçamento de 10 cm, bem como de 0 a 96 cm no eixo y, com um espaçamento de 5 cm. Na Figura 9 é possível verificar um exemplo de como seria uma das 122 simulações realizadas com a perspectiva superior.

Para aumentar a confiabilidade das simulações e reduzir significativamente a incerteza associada aos resultados, cada simulação envolveu o rastreamento de $1\text{E}9$ histórias de partículas.

Figura 9 – Exemplo de uma simulação de um médico na posição 365 cm no eixo x e 35 cm no eixo y.



Fonte: Elaborada pelo autor.

3.2.1 Configuração de simulação

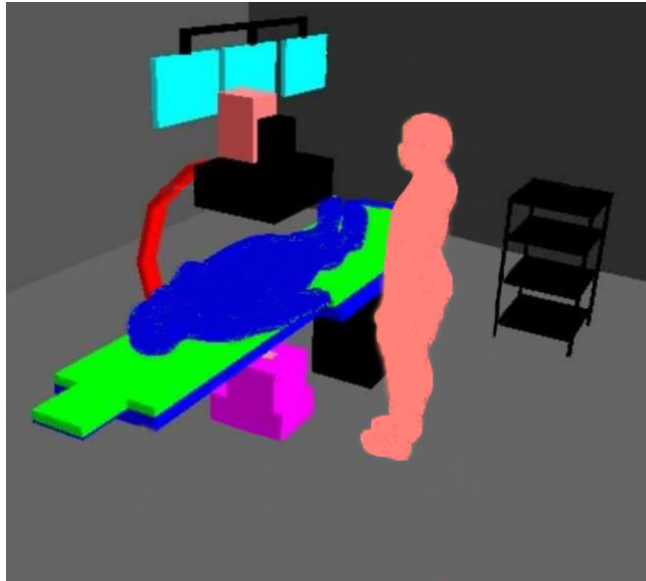
Para a simulação, foi desenvolvida uma representação simplificada de um cenário comumente encontrado em RI. Nesta representação, foram incluídos os equipamentos de proteção individual e coletiva, como avental de chumbo, óculos plumbíferos e protetor de tireoide.

A fonte de radiação foi configurada como um feixe pontual colimado, direcionado para formar um feixe cônico, estando a uma distância de 102 cm do paciente. O espectro que representa a fonte para a simulação foi obtido utilizando o software IPEM SRS 78 (EVANS, 2009).

A simulação contou com a utilização de dois modelos antropomórficos virtuais MESH (CASSOLA *et al.*, 2009), representando um médico e um paciente. O objetivo foi gerar simulações mais próximas da realidade, permitindo a avaliação da dose absorvida nos órgãos do médico. O paciente não foi objeto de estudo neste trabalho, sendo apenas o meio espalhador. O esquema da geometria criada com os dois modelos MESH está representado na Figura 10.

Uma vez obtidos os valores de $CC[H_t]$ as doses equivalentes foram calculadas multiplicando-se esses valores pelo PKA experimental (CASTILHO *et al.*, 2019). No estudo de Castilho *et al.* (2019), foram registrados valores de dose equivalente usando dosímetros posicionados estrategicamente em locais como olhos, tórax e tornozelo de profissionais de saúde durante procedimentos de RI. Além disso, o valor de PKA foi obtido

Figura 10 – Modelagem da sala durante o procedimento de RI



Fonte: Elaborada pelo autor.

considerando variações de tensão entre 80 kV e 110 kV, com campos de irradiação de 20 cm^2 , 31 cm^2 e 38 cm^2 . Foi calculada a média das doses quando o operador usou os modos contínuo e pulsado no equipamento de fluoroscopia, resultando em um PKA de $449,4 \text{ Gy} \cdot \text{cm}^2$ (CASTILHO *et al.*, 2019).

Para simular o medidor de PKA , foi utilizado um dispositivo com área de $199,82 \text{ cm}^2$, abrangendo o campo simulado. Este medidor foi posicionado perpendicularmente logo abaixo da saída do feixe no tubo de raios X.

3.3 Pré-processamento dos dados para rede de predição

Inicialmente, os dados referentes aos médicos foram submetidos a um processo de filtragem para extrair as informações relevantes. Para melhorar a comparabilidade entre diferentes órgãos, optou-se por normalizar os dados utilizando a técnica *Min-Max*, resultando em médias próximas a zero e desvios padrões aproximadamente igual a um para cada órgão. Essa abordagem foi escolhida para mitigar as disparidades nas escalas dos dados, garantindo que as características de diferentes órgãos pudessem ser comparadas de maneira mais equitativa.

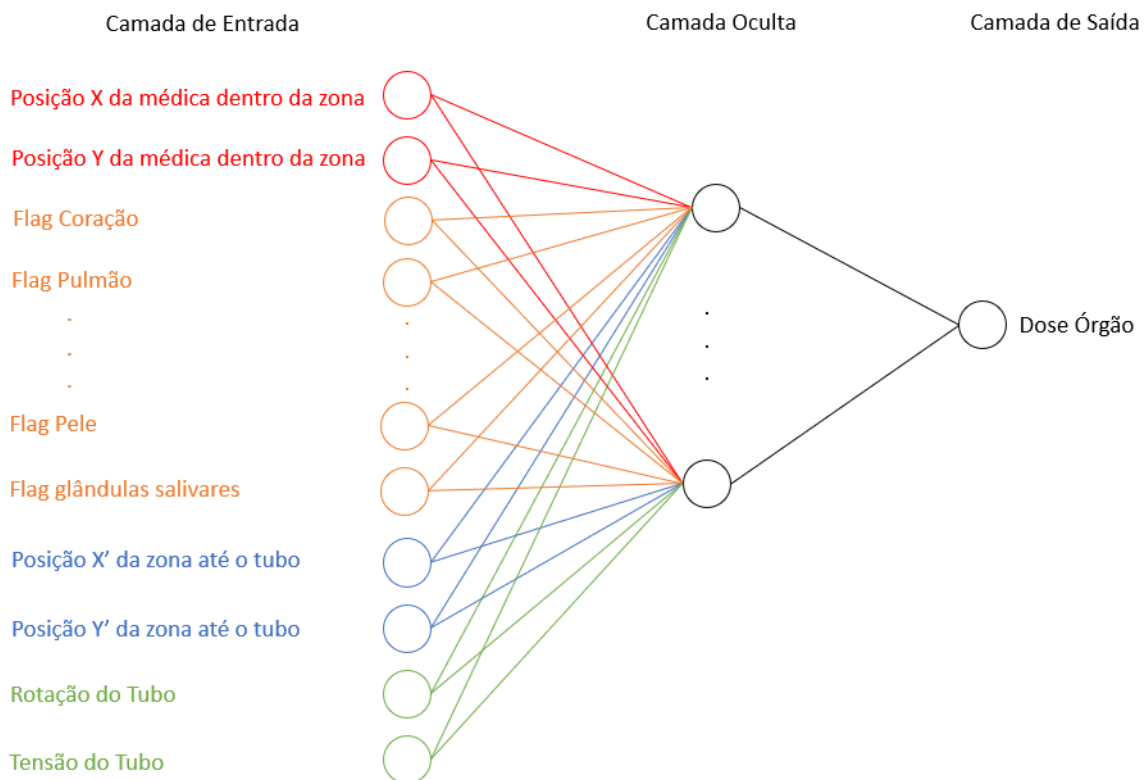
A variável categórica “Órgão” foi transformada utilizando a técnica de *one-hot encoding*. Essa abordagem foi adotada para converter a informação categórica em uma representação numérica mais apropriada para a modelagem estatística. Com a aplicação do one-hot encoding, criaram-se variáveis *dummy* para cada categoria presente na variável “Órgão”. Cada órgão passou a ser representado por um conjunto de *flags* binárias (0 ou 1), indicando a presença ou ausência do órgão para cada observação.

3.4 Configuração do modelo MLP de predição de dose

No decorrer desta pesquisa, foi optado por empregar uma rede neural artificial do tipo MLP (*Multilayer Perceptron*). A avaliação de seu desempenho foi conduzida mediante a divisão do conjunto de dados em conjuntos distintos de treinamento e teste, reservando 80% para o treinamento e 20% para o teste. Essa divisão foi realizada de forma aleatória, utilizando uma semente específica (42) para garantir a reprodutibilidade dos resultados. O processo de treinamento do modelo foi conduzido ao longo de 100 épocas, visando otimizar sua capacidade de aprendizado e generalização.

Na Figura 11 são apresentados os parâmetros de entrada para a modelagem da rede. Em vermelho tem-se a posição X e Y do médico dentro da zona delimitada nas simulações. As entradas laranjas são valores booleanos que indicam o órgão de interesse, essa técnica foi a transformação de uma variável categórica em *one hot encoding*. Isso ajuda a treinar melhor o modelo e conseguir obter as doses nos órgãos. As entradas azuis representam os valores X' e Y' sendo as distância da zona de simulação e o tubo de raios-X. As cores verdes são medidas sobre a rotação do gantry e a tensão no tubo. Neste trabalho os valores de posição da zona até o tubo de raios-x, rotação e tensão do tubo se mantiveram constantes.

Figura 11 – Modelagem da Rede para o problema proposto neste trabalho



Fonte: Elaborada pelo autor.

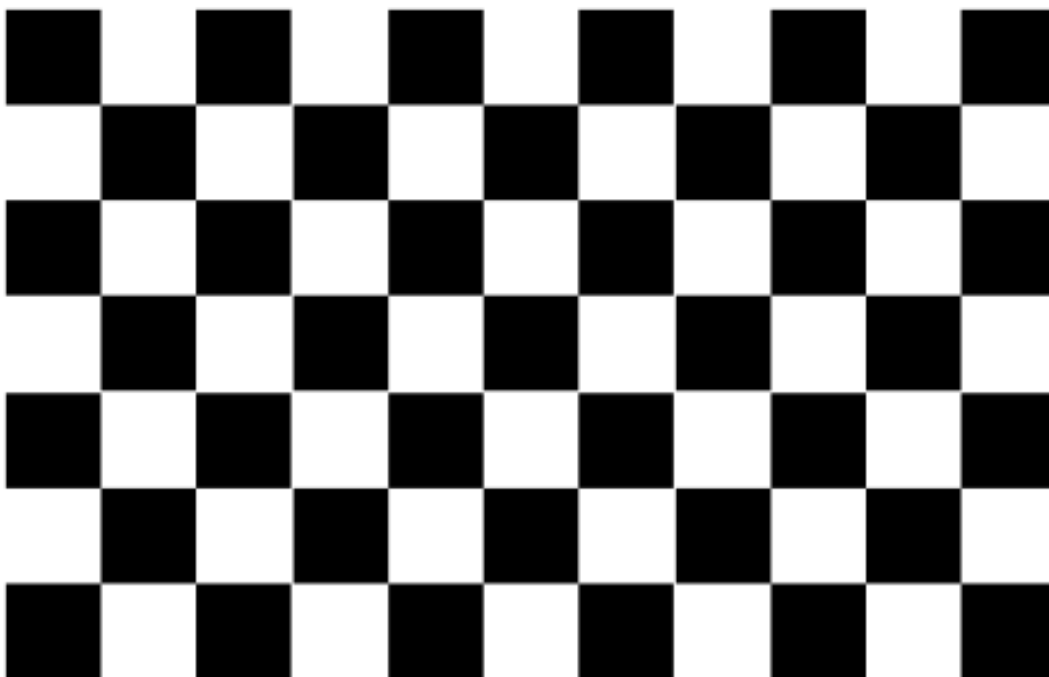
3.5 Configuração da visão computacional

3.5.1 Calibração da Câmera de Vídeo

A calibração de câmeras é uma etapa necessária em visão computacional, quando se trata de extração de informações métricas em imagens 2D (ZHANG, 2000). A biblioteca OpenCV possui a função *calibrateCamera*, que estima os parâmetros intrínsecos da câmera e os parâmetros extrínsecos de uma visualização, através de um algoritmo baseado nos métodos e implementações de Zhang (ZHANG, 2000).

No processo de calibração de câmeras, as coordenadas dos pontos do objeto 3D e suas projeções 2D precisam ser especificadas e, para isso, deve ser utilizado um objeto com geometria conhecida e pontos facilmente detectáveis, sendo este objeto denominado padrão de calibração.

Figura 12 – Padrão de Calibração Xadrez



Fonte: (LAUREANO, 2013)

Dentre os padrões de calibração disponíveis, diversos trabalhos citam a utilização do padrão do tipo xadrez, que pode ser observado na Figura 12.

O módulo OpenCV possui a implementação de detecção automática de padrões do tipo xadrez, por meio da função *findChessBoardCorners* (LAUREANO, 2013). Foram realizadas mais de 50 variações (inclinada, angulada) das imagens do tabuleiro, para que a função conseguisse perceber a distorção da lente.

Com a calibração realizada, é salva a matriz de distorção e utilizada em toda captação de imagens da YOLO.

Figura 13 – Área que limita a região de deslocamento do médico durante o procedimento de RI



Fonte: Elaborada pelo autor.

3.5.2 Delimitação da Área de Movimentação através de Marcadores Aruco

Antes de empregar o modelo YOLOv4 para a identificação do médico, foram estrategicamente posicionados quatro marcadores Aruco no solo, delineando meticulosamente a extensão total da área em que o profissional de saúde pode se deslocar durante o procedimento, conforme ilustrado na Figura 13. Cada marcador foi configurado e disposto em uma ordem predefinida, proporcionando uma referência espacial consistente.

Esta etapa inicial visou estabelecer uma base sólida para o mapeamento preciso da área de movimentação. A disposição ordenada dos marcadores foi cuidadosamente planejada para otimizar a captura de dados, assegurando uma representação fiel da realidade física. Após a implantação dos marcadores, procedeu-se à medição precisa das distâncias entre eles no ambiente real.

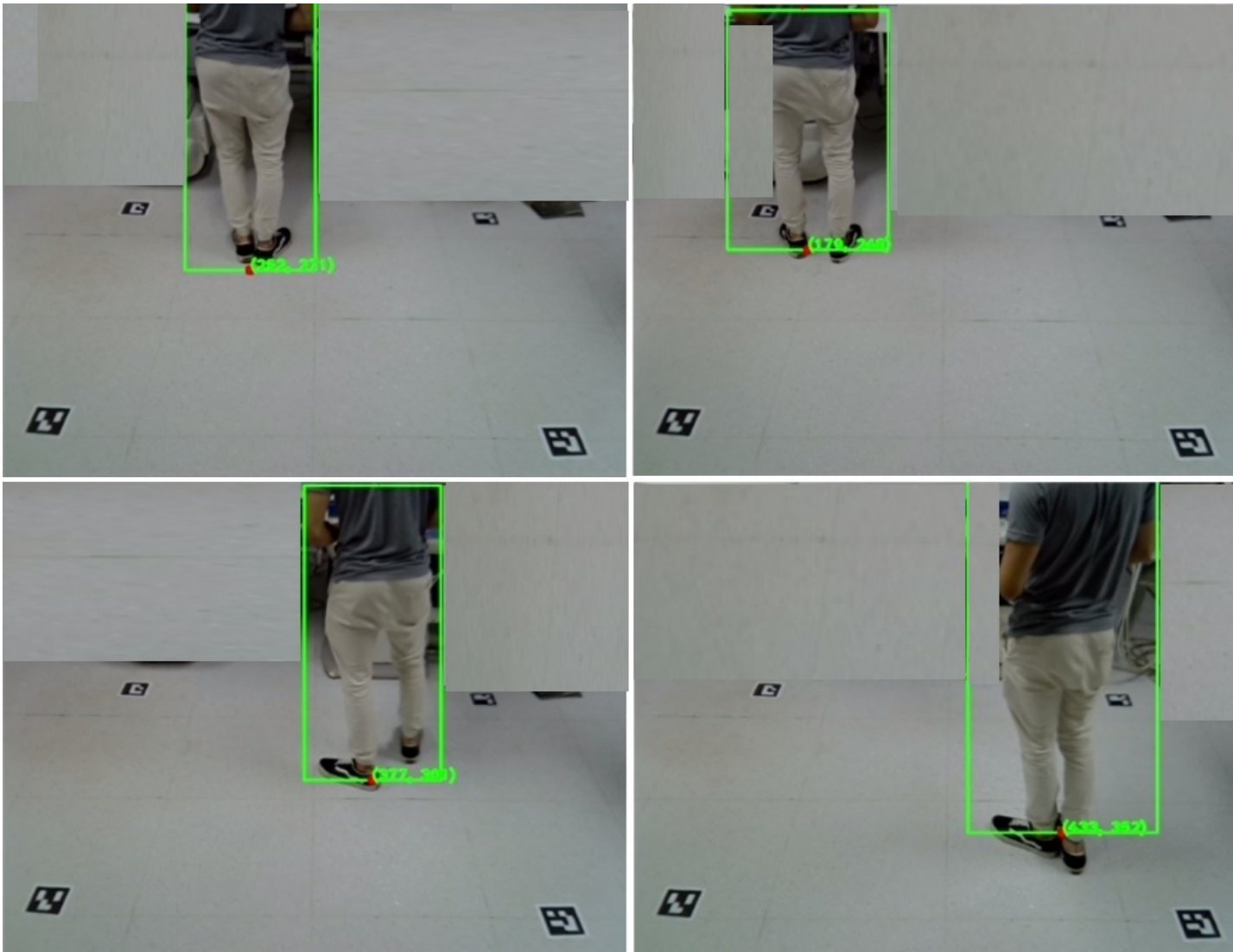
Essa abordagem permitiu a obtenção de uma relação pixel/cm, um fator crucial para a subsequente mensuração das coordenadas do médico durante a execução do procedimento. Ao converter as dimensões em pixels para unidades de medida física, foi possível estabelecer um sistema de coordenadas preciso, fornecendo uma base robusta para a aplicação efetiva do YOLOv4 na identificação e rastreamento do profissional de saúde em tempo real.

3.5.3 Ajustes do modelo YOLOv4

Após a conclusão do processo de calibração e o cálculo da relação pixel/cm, foi aplicado o modelo YOLOv4 treinado especificamente para a detecção de pessoas. Com a identificação bem-sucedida do médico, obteve-se o ponto central dos pés do profissional, como ilustrado na Figura 14, o qual desempenha um papel crucial para o cálculo preciso das coordenadas. Este ponto central dos pés serve como referência fundamental para monitorar e calcular as coordenadas do médico durante suas movimentações dentro da área delimitada pelos marcadores. Essa abordagem refinada não apenas melhora a precisão do rastreamento, mas também proporciona uma base sólida para análises espaciais detalhadas

durante o procedimento em questão.

Figura 14 – Simulação da movimentação do médico durante um procedimento de RI



Fonte: Elaborada pelo autor.

O processo de predição de dose para todos os órgãos foi conduzido por meio de um laço de repetição, permitindo a obtenção da dose para cada órgão na posição específica determinada pela localização do médico. Destaca-se que a eficiência desse sistema é notável, com o tempo total para identificação do médico e realização da predição de dose para todos os órgãos ocorrendo em apenas 0,1 segundo.

É relevante salientar que, desse tempo total, 50% é dedicado à etapa de captura e identificação da pessoa na imagem, evidenciando a agilidade do modelo YOLOv4 nesse processo. Importante ressaltar que o código foi concebido e testado exclusivamente utilizando recursos de CPU, em particular um processador i7 de 10ª geração com 8 GB de RAM, destacando a eficiência mesmo em ambientes computacionais mais convencionais.

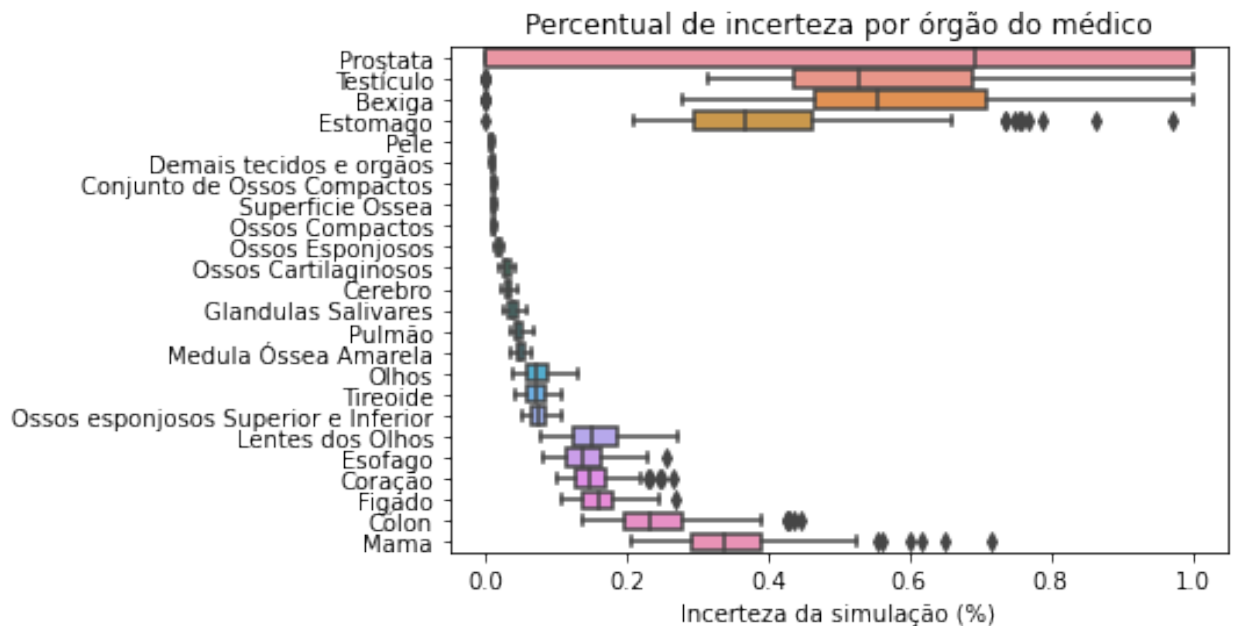
4 RESULTADOS E DISCUSSÕES

4.1 Modelo de predição de Dose

4.1.1 Análise estatística dos dados de entrada

As simulações em MCNP resultaram em uma incerteza média de 15,75% e uma mediana de 6,07%. É importante ressaltar que a presença de alguns casos atípicos distorceu a medida da incerteza média, impactando negativamente seu valor. Notavelmente, em órgãos menores, essa incerteza média tendeu a ser mais alta, o que pode ser atribuído à escala e à influência do tamanho dos órgãos nas simulações, conforme a Figura 15 demonstra. Para reduzir essa incerteza deve ser empregadas técnicas de redução de variância.

Figura 15 – Boxplot da incerteza dos valores de dose para cada órgão do médico no procedimento de RI avaliado neste trabalho



Fonte: Elaborada pelo autor.

O tempo de computação necessário para realizar essas simulações foi significativo, porém, o uso de um cluster baseado em computação paralela foi crucial.

No entanto, é importante contextualizar esse tempo de computação em relação à escala das simulações e à complexidade do modelo usado. Isso demonstra a necessidade e a justificativa da escolha do cluster para atender às demandas computacionais do projeto.

É fundamental realizar uma análise do banco de dados para determinar a aplicabilidade das técnicas testadas e garantir uma interpretação precisa dos resultados obtidos.

Com esse propósito, a Tabela 1 e a Figura 16 apresentam as estatísticas básicas do banco de dados.

Conforme mencionado anteriormente, as variáveis independentes numéricas do modelo são X e Y (coordenadas espaciais do médico), enquanto a variável categórica chamada “Órgão”, foi utilizada como um dado adicional após o rearranjo dos dados para a entrada dos modelos.

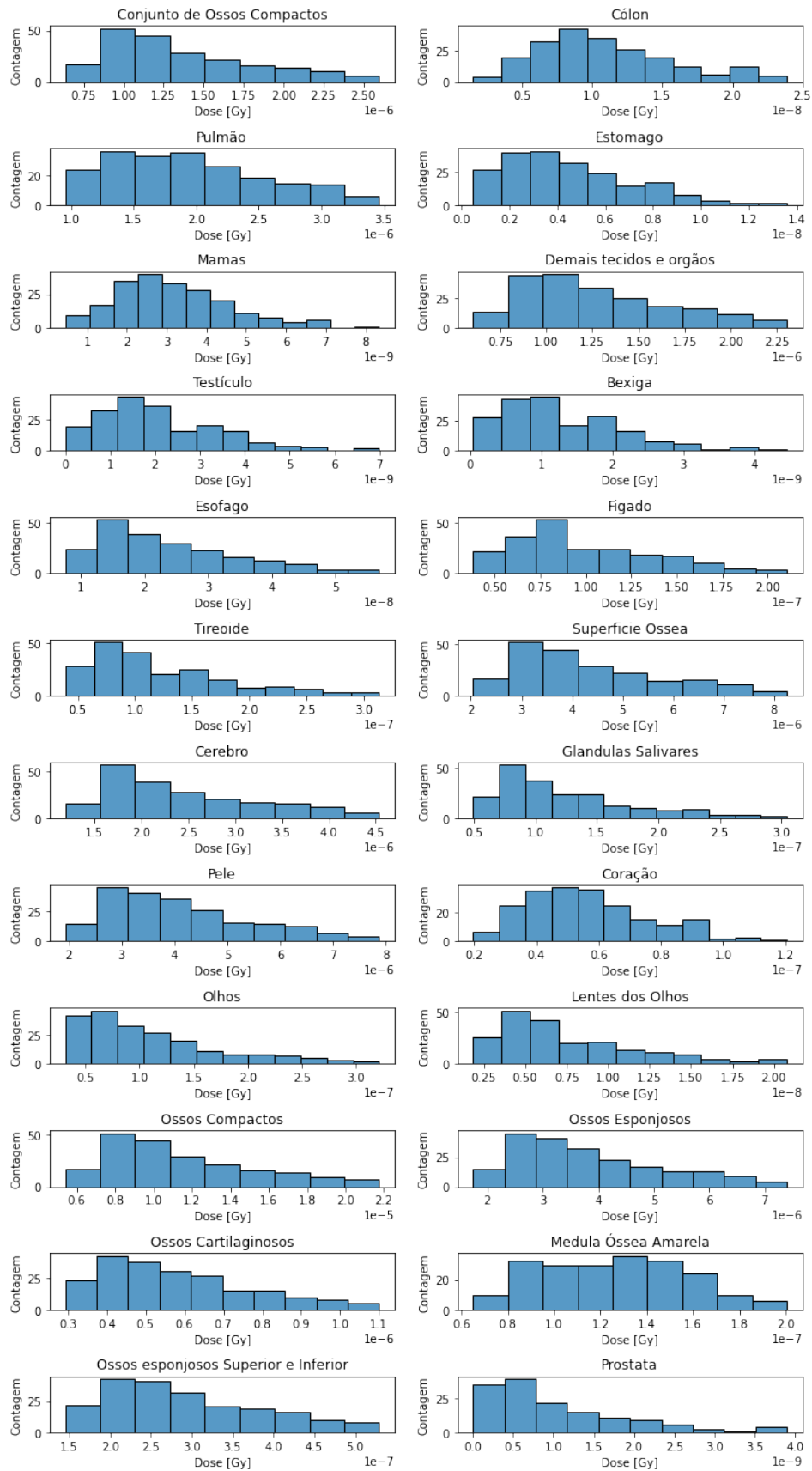
Neste estudo, a variável dependente é a dose recebida no órgão, após a variação no posicionamento do médico durante o procedimento. Na Tabela 1, é possível observar os valores de média, variância, mínimo, primeiro quartil, terceiro quartil e valor máximo para todas as variáveis mencionadas.

Tabela 1 – Estatística Descritiva do banco de dados

Estatística	Variável	Média	Variância	Mín	Q1	Q3	Máx
Parâmetros de Posicionamento	X[cm]	110,04	65,36	0	56	160	224
Parâmetros de Posicionamento	Y[cm]	54,30	27,71	0	32	80	96
Medidor de Dose	Dose[Gy]	1,37E3	2,71E-6	2,15E-12	1,03E-8	1,66E-6	2,17E-5

Fonte: Elaborada pelo autor.

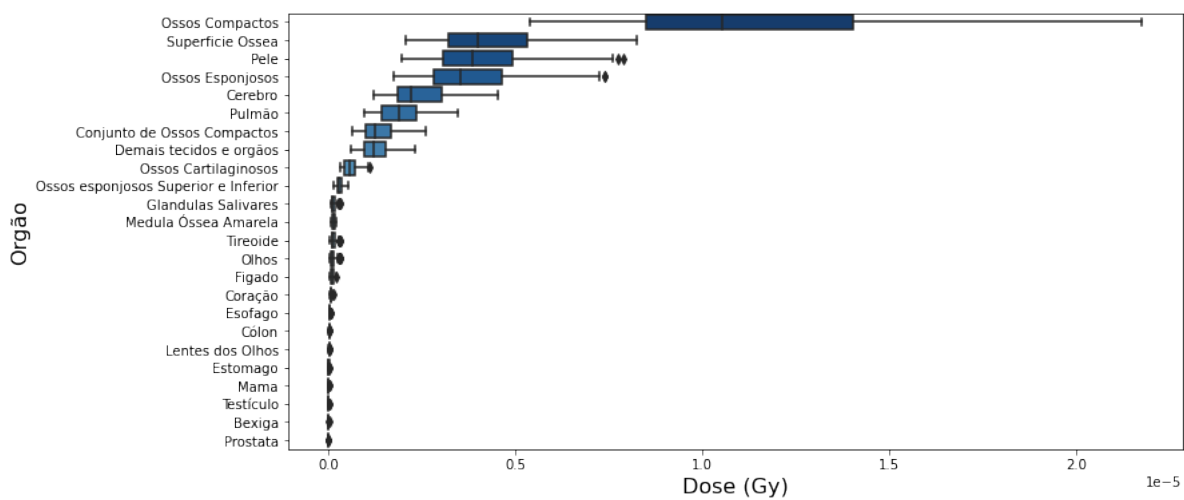
Figura 16 – Histogramas das variáveis



Fonte: Elaborada pelo autor.

Ao examinar os valores interquartis apresentados na Tabela 1, pode-se observar que a dose exibe variações significativas, embora não forneça muitos detalhes devido à quantidade de órgãos monitorados. Portanto, ao analisar o gráfico de caixa (Boxplot) representado na Figura 17, percebe-se que há altas variações para todos os órgãos e na pele. Esses valores elevados indicam que essas variáveis são menos homogêneas em comparação às demais, resultando em uma dispersão maior dos dados em relação à média. No contexto específico, valores elevados de variância são benéficos, pois possibilitam uma maior quantidade de informações a serem fornecidas ao modelo. Conseqüentemente, é necessário que o modelo seja suficientemente robusto para absorver essas informações e generalizá-las adequadamente.

Figura 17 – Boxplot das variáveis antes da normalização



Fonte: Elaborada pelo autor.

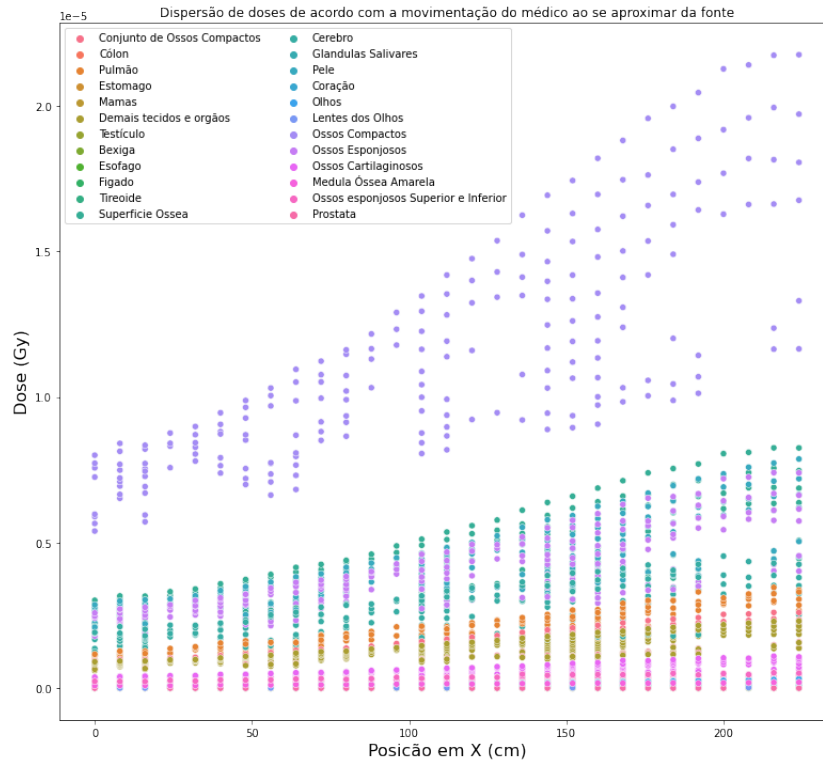
Outro ponto a ser analisado, na Figura 16, os histogramas aparecem enviesados para a esquerda indicando uma mediana menor que a média.

4.1.2 Estatística Multivariada

Ao analisar a Figura 18, que apresenta um gráfico de dispersão e os respectivos coeficientes de correlação entre as variáveis, fica evidente uma correlação significativa entre as variáveis independentes de posicionamento e o aumento da dose. No entanto, essa correlação é observável apenas nos órgãos de maior volume, conforme demonstrado na Figura 18. Isso está diretamente relacionado à dose recebida, já que quanto maior o volume do órgão, maior é o fluxo de radiação sobre ele.

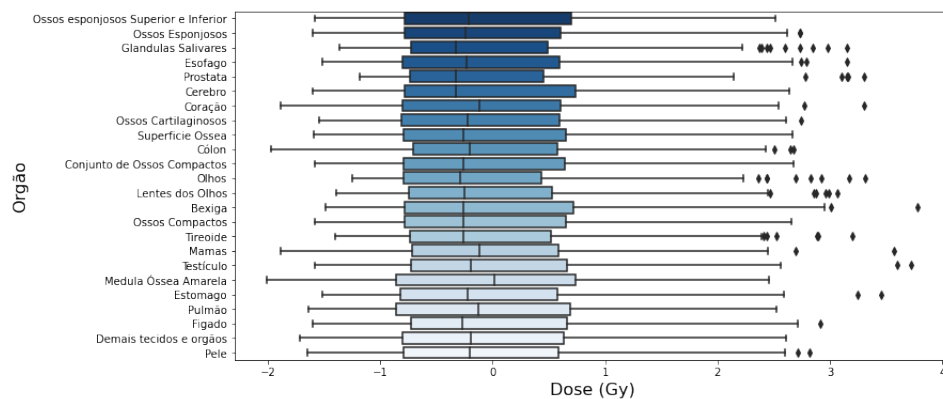
Após a compreensão das correlações existentes entre as variáveis, o boxplot apresentado na Figura 17 destaca, entre outras informações, a discrepância de escala entre essas variáveis. A fim de evitar que esse efeito prejudique o modelo, é necessário realizar a normalização das variáveis utilizando a abordagem *Z-Score*, ilustrada na Figura 19.

Figura 18 – Dispersão de doses de acordo com a movimentação do médico ao se aproximar do tubo de raios-x



Fonte: Elaborada pelo autor.

Figura 19 – Boxplot das variáveis após normalização



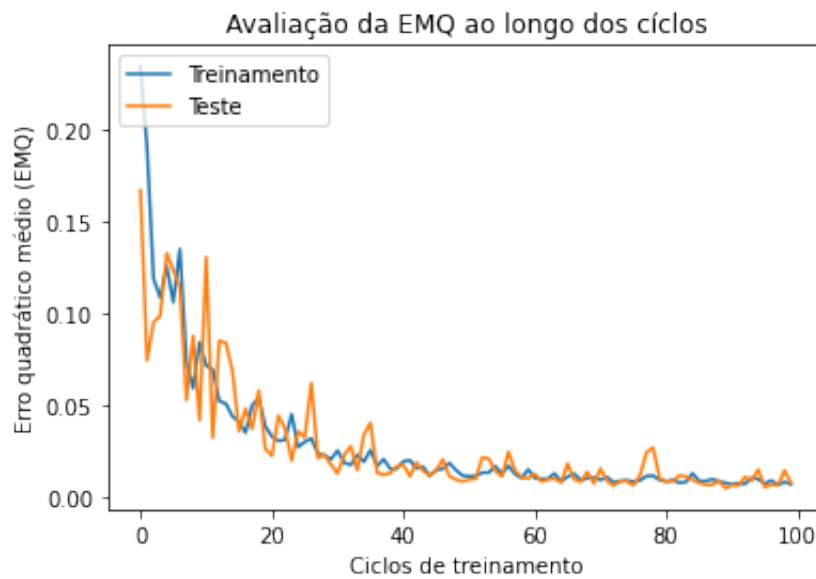
Fonte: Elaborada pelo autor.

Uma vez entendida a complexidade das relações entre as variáveis, parte-se agora para o estudo dos modelos com as técnicas propostas. Na sequência, as técnicas abordadas foram: redes neurais artificiais, *random forest* e Regressão de vetores de suporte. Acrescenta-se ao fim a comparação dos modelos.

4.1.3 Rede Neural *Multilayer Perceptron*

Por meio da Figura 20, pode-se observar um ajuste satisfatório da rede neural aos dados de treinamento, sem indícios de *overfitting* quando comparados aos dados de teste.

Figura 20 – Treinamento da Rede Neural



Fonte: Elaborada pelo autor.

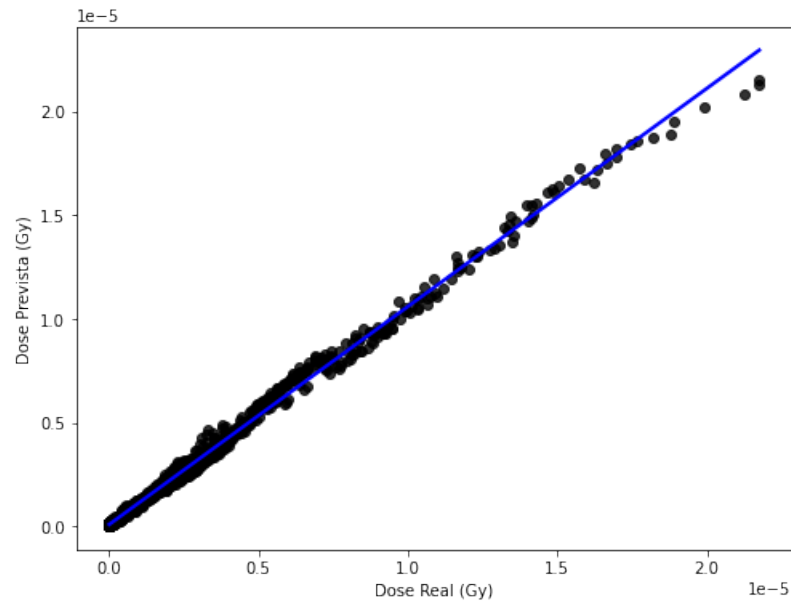
Na Tabela 2, são apresentados os valores da raiz quadrada do erro médio (RMSE), coeficiente de correlação (R^2) e erro médio absoluto (MAE) para a validação de treino e teste da rede neural. Nas figuras 21 e 22, são exibidos os gráficos de dispersão da validação no treinamento e teste, respectivamente. Observa-se que, na validação do teste, a rede neural obteve um (R^2) de 0.99 o que indica um ótimo desempenho.

Tabela 2 – Resultados Redes Neurais

Dados	R^2	MAE	RMSE
Treinamento	0,98	1,80E-7	7,86E-14
Teste	0,99	1,80E-7	7,61E-14

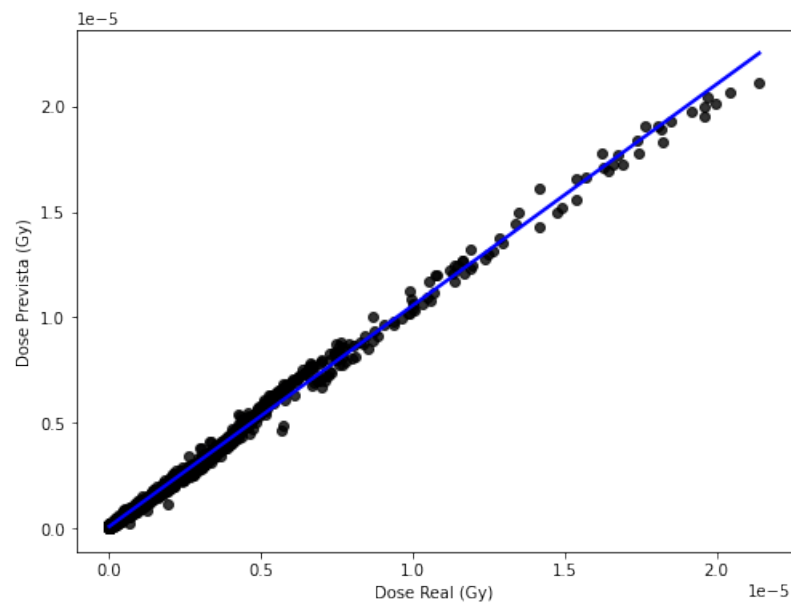
Fonte: Elaborada pelo autor.

Figura 21 – Gráfico de dispersão nos dados de treinamento da Rede Neural



Fonte: Elaborada pelo autor.

Figura 22 – Gráfico de dispersão nos dados de teste da Rede Neural



Fonte: Elaborada pelo autor.

A Tabela 3 apresenta uma comparação entre as doses previstas e calculadas com o MCNP6.2 e a sua diferença percentual. A diferença percentual da próstata foi de 121871% enquanto que a pele teve uma diferença de 4%. As maiores diferenças ocorreram em órgãos de menor volume devido a própria dificuldade das simulações do MCNP6.2 conforme visto anteriormente na Figura 15 que acabaram refletindo no aprendizado do modelo MLP.

Tabela 3 – Diferenças percentuais entre os valores de dose calculados pelo MCNP6.2 e preditos neste trabalho, para o médico.

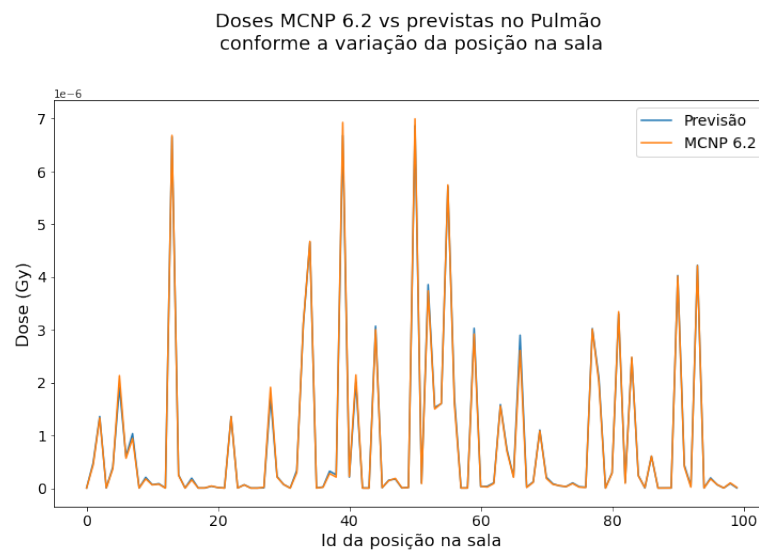
Orgão	Dose - MCNP	Dose - Predita	Posição - X	Posição - Y	Diferença (%)
Bexiga	5,29E-11	2,62E-09	48	88	4856
Cerebro	1,19E-06	1,55E-06	16	0	31
Conjunto de Ossos Compactos	6,65E-07	3,84E-07	0	8	42
Coração	5,41E-08	8,96E-08	144	88	66
Cólon	7,75E-09	3,04E-08	80	88	292
Demais tecidos e órgãos	6,49E-07	4,35E-07	0	8	33
Esôfago	1,54E-08	4,34E-08	80	88	181
Estômago	1,30E-08	2,62E-09	160	88	80
Fígado	1,78E-07	1,24E-07	144	64	30
Glândulas Salivares	2,35E-07	2,66E-07	176	88	13
Lentes dos Olhos	1,30E-08	2,35E-08	208	72	82
Mamas	7,08E-09	2,62E-09	200	72	63
Medula Óssea Amarela	1,77E-07	2,32E-07	200	88	31
Olhos	3,13E-07	2,81E-07	224	80	10
Ossos Cartilagosos	6,85E-07	7,72E-07	136	80	13
Ossos Compactos	1,96E-05	1,83E-05	176	88	7
Ossos Esponjosos	6,53E-06	6,21E-06	176	88	5
Ossos esponjosos Superior e Inferior	2,69E-07	3,28E-07	216	24	22
Pele	6,96E-06	6,71E-06	184	88	4
Próstata	2,15E-12	2,62E-09	0	24	121871
Pulmão	1,42E-06	1,72E-06	64	48	21
Superfície Óssea	8,06E-06	7,64E-06	200	88	5
Testículo	6,99E-09	2,62E-09	216	88	62
Tireoide	1,98E-07	2,38E-07	160	88	20

Fonte: Elaborada pelo autor.

Além disso, é relevante notar que os pontos mais distantes da fonte de raio-X exibem um aumento no erro em comparação com as distâncias menores. Essa observação indica uma correlação clara entre a distância relativa à fonte de raios-X e a precisão das previsões, ressaltando a importância da proximidade em relação à fonte para reduzir as incertezas e aprimorar a precisão das estimativas de dose para as áreas de menor volume.

A Figura 23 apresenta um exemplo do desempenho do modelo MLP no pulmão, mas existe um comportamento semelhante com todos os órgãos, tecidos e regiões que foram treinadas no modelo. Essa análise visual mostra a concordância entre as doses previstas e calculadas pelo MCNP6.2 para o órgão do pulmão nos dados de teste. A linha azul representa a dose prevista pelo modelo de rede neural MLP, enquanto a linha laranja retrata a dose calculada. A proximidade entre essas duas linhas demonstra uma correspondência consistente e encorajadora entre as previsões do modelo e a realidade. Esse alinhamento entre as doses previstas e calculadas sugere uma capacidade significativa do modelo em capturar e estimar com precisão os níveis de dose para o pulmão, o que é essencial para sua aplicabilidade e confiabilidade em contextos clínicos e de pesquisa.

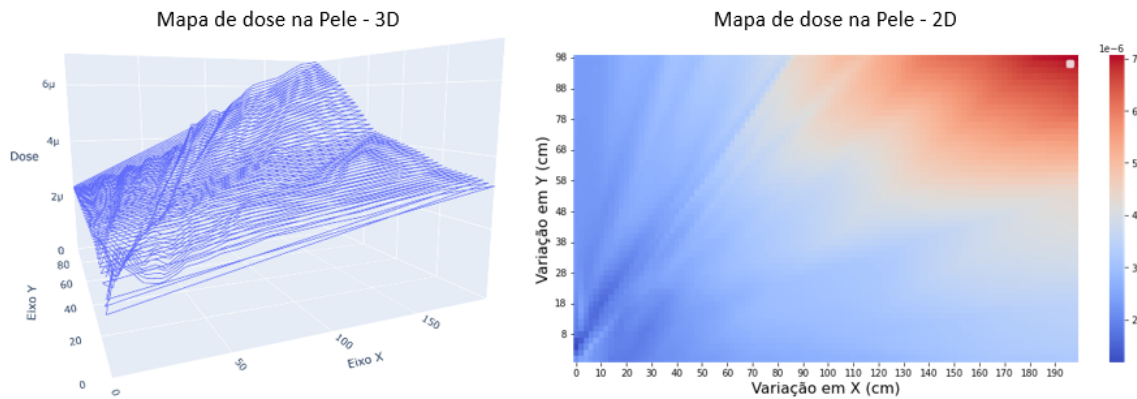
Figura 23 – Desempenho do modelo entre MCNP6.2 e previsão para o pulmão



A partir do treinamento do modelo, pode-se explorar a capacidade de gerar mapas de dose individuais para diferentes órgãos ou regiões específicas. A Figura 24 é um exemplo para a pele do médico. Esses mapas representam uma ferramenta valiosa para identificar e visualizar áreas com variações significativas de doses de radiação. A análise desses mapas oferece informações cruciais ao destacar as regiões com doses mais elevadas e aquelas com doses mais baixas de radiação. Essa informação pode auxiliar muito a prática clínica, permitindo a identificação de áreas críticas que possam exigir uma atenção especial ou modulação de doses, bem como áreas que possam estar menos expostas aos efeitos adversos

da radiação. Essa abordagem não apenas fornece uma representação visual intuitiva das variações de doses, mas, também serve como um indicador valioso para a tomada de decisões informadas, contribuindo significativamente para a otimização e personalização dos procedimentos de RI.

Figura 24 – Dose na Pele em Relação à Movimentação do Médico durante o procedimento de RI. A fonte radioativa está localizada conforme mostrado na Figura 9.



Fonte: Elaborada pelo autor.

5 CONCLUSÃO

Em suma, este estudo alcançou com êxito a previsão de doses em órgãos/tecidos com base nas posições espaciais do médico durante procedimentos de RI.

No entanto, ao realizar previsões em órgãos que apresentaram incertezas altas na simulação do MCNP 6.2, acabou refletindo altas taxas de erro durante o modelo MLP. Esse comportamento está relacionado a regiões pequenas, como próstata, bexiga e cólon. Para melhorar as previsões do modelo MLP é necessário fazer alterações nas simulações do MCNP, usando métodos de redução de variância para melhorar a incerteza dos órgãos e tecidos que apresentaram altas variações.

Comparando as doses previstas e calculadas com o MCNP6.2, observou-se uma diferença percentual de 121871% para a próstata, enquanto que a pele teve uma diferença de 4%. As maiores diferenças ocorreram em áreas de menor volume devido a própria dificuldade das simulações do MCNP6.2.

É relevante destacar que o modelo MLP apresentou um custo de treinamento baixo, graças à facilidade de compreender as diversas simulações, baseadas na diminuição da radiação com o quadrado da distância. As limitações deste modelo de previsão de dose estão restritas a um único tipo de procedimento, envolvendo pacientes imóveis, com gênero fixo, tubo de radiação em projeção constante (PA), tensão no tubo também constante e cenário estático, sem movimentação dos equipamentos ao redor.

Embora, o modelo de captura de imagem seja adequado, é necessário restringir o movimento do médico dentro da região dos marcadores, pois as regiões fora desses marcadores estão sujeitas a maiores deformações da imagem, mesmo com a utilização de calibração intrínseca e extrínseca. O modelo é capaz de identificar pessoas mesmo se a imagem estiver cortada, permitindo a identificação do médico mesmo na presença de obstruções, como um carrinho com equipamentos médicos entre o médico e a câmera. Entretanto, é crucial ajustar o modelo para distinguir entre médico e enfermeiro. Para obter a caixa delimitadora do médico, o modelo demorou 0,5 segundos, sendo importante ressaltar que o tempo está diretamente relacionado à capacidade da máquina que está realizando o procedimento. O modelo YOLOv4 possibilita uma taxa de captura de até 30 FPS.

Finalmente, embora este método apresente oportunidades significativas para melhorias, a integração do MLP e YOLO apresentou resultados muito promissores, de baixo custo e se revela como um importante fator de monitoramento para a equipe médica durante procedimentos de RI.

REFERÊNCIAS

ARAÚJO, R. M. d. Aprendizado de máquina em sistemas complexos multiagentes: estudo de caso em um ambiente sob racionalidade limitada. 2004.

CABRERA, L. *et al.* Tuning the stator resistance of induction motors using artificial neural network. **IEEE Transactions on Power Electronics**, v. 12, n. 5, p. 779–787, 1997.

CANEVARO, L. Aspectos físicos e técnicos da radiologia intervencionista. **Revista Brasileira de Física Médica**, v. 3, n. 1, p. 101–115, 2009.

CAO, Z. *et al.* Realtime multi-person 2d pose estimation using part affinity fields. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2017. p. 7291–7299.

CASSOLA, V. *et al.* Fash and mash: female and male adult human phantoms based on polygon mesh surfaces: I. development of the anatomy. **Physics in Medicine & Biology**, IOP Publishing, v. 55, n. 1, p. 133, 2009.

CASTILHO, A. V. B. *et al.* A study of radiation doses to the patient and medical team at embolization procedures. **Journal of Radiation Protection and Research**, The Korean Association for Radiation Protection, v. 44, n. 3, p. 110–117, 2019.

CHA, Y.-J. *et al.* Deep learning-based crack damage detection using convolutional neural networks. **Computer-Aided Civil and Infrastructure Engineering**, Wiley Online Library, v. 32, n. 5, p. 361–378, 2017.

CHEN, D. *et al.* Supervised transformer network for efficient face detection. In: SPRINGER. **Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V 14**. [S.l.], 2016. p. 122–138.

CHEN, Z. *et al.* Differentially private location protection with continuous time stamps for vanets. In: SPRINGER. **Algorithms and Architectures for Parallel Processing: 18th International Conference, ICA3PP 2018, Guangzhou, China, November 15-17, 2018, Proceedings, Part IV 18**. [S.l.], 2018. p. 204–219.

CHOI, D. Y. *et al.* Cnn-based pre-processing and multi-frame-based view transformation for fisheye camera-based avm system. In: IEEE. **2017 IEEE International Conference on Image Processing (ICIP)**. [S.l.], 2017. p. 4073–4077.

COMASCHI, F. *et al.* Rasw: a run-time adaptive sliding window to improve viola-jones object detection. In: IEEE. **2013 Seventh International Conference on Distributed Smart Cameras (ICDSC)**. [S.l.], 2013. p. 1–6.

DOLLAR, P. *et al.* Pedestrian detection: An evaluation of the state of the art. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 34, n. 4, p. 743–761, 2011.

DU, X. *et al.* Car detection for autonomous vehicle: Lidar and vision fusion approach through deep learning framework. In: IEEE. **2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)**. [S.l.], 2017. p. 749–754.

DUANE, C. B. Close-range camera calibration. **Photogramm. Eng.**, v. 37, n. 8, p. 855–866, 1971.

EVANS, S. Catalogue of diagnostic x-ray spectra and other data. **Journal of Radiological Protection**, v. 18, 07 2009.

FELZENSZWALB, P. F. *et al.* Object detection with discriminatively trained part-based models. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 32, n. 9, p. 1627–1645, 2009.

FERES, F. *et al.* Diretriz da sociedade brasileira de cardiologia e da sociedade brasileira de hemodinâmica e cardiologia intervencionista sobre intervenção coronária percutânea. **Arquivos Brasileiros de Cardiologia**, SciELO Brasil, v. 109, p. 1–81, 2017.

GARRIDO, S.; NICHOLSON, S. **Detection of aruco markers. opencv: Open source computer vision**. 2021.

GIDARIS, S.; KOMODAKIS, N. Object detection via a multi-region and semantic segmentation-aware cnn model. In: **Proceedings of the IEEE international conference on computer vision**. [S.l.: s.n.], 2015. p. 1134–1142.

GIRSHICK, R. Fast r-cnn. In: **Proceedings of the IEEE international conference on computer vision**. [S.l.: s.n.], 2015. p. 1440–1448.

GIRSHICK, R. *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2014. p. 580–587.

HAN, J.; MORAGA, C. The influence of the sigmoid function parameters on the speed of backpropagation learning. In: SPRINGER. **International workshop on artificial neural networks**. [S.l.], 1995. p. 195–201.

HAYKIN, S. **Neural networks and learning machines**. 3. ed. [S.l.]: Pearson Education India, 2009.

HUANG, X. *et al.* Pp-yolov2: A practical object detector. **arXiv preprint arXiv:2104.10419**, 2021.

HUANG, Y. *et al.* Learning optical flow with r-cnn for visual odometry. In: IEEE. **2021 IEEE International Conference on Robotics and Automation (ICRA)**. [S.l.], 2021. p. 14410–14416.

HUH, D.; SEJNOWSKI, T. J. Gradient descent for spiking neural networks. **Advances in neural information processing systems**, v. 31, 2018.

JIA, Y. *et al.* Caffe: Convolutional architecture for fast feature embedding. In: **Proceedings of the 22nd ACM international conference on Multimedia**. [S.l.: s.n.], 2014. p. 675–678.

-
- JIANG, P. *et al.* A review of yolo algorithm developments. **Procedia computer science**, Elsevier, v. 199, p. 1066–1073, 2022.
- JUNG, C.; KIM, C. A unified spectral-domain approach for saliency detection and its application to automatic object segmentation. **IEEE Transactions on Image Processing**, IEEE, v. 21, n. 3, p. 1272–1283, 2011.
- KARTAM, N. *et al.* Artificial neural networks for civil engineers: Fundamentals and applications. In: . USA: American Society of Civil Engineers, 1997. ISBN 0784402256.
- KRIZHEVSKY, A. *et al.* ImageNet classification with deep convolutional neural networks. **Neural Information Processing Systems**, v. 25, p. 1106–1114, 01 2012.
- LAUREANO, G. T. **Detecção topológica de padrões xadrez para calibração de câmeras**. 2013. Tese (Doutorado) — Universidade de São Paulo, 2013.
- LI, C. *et al.* Yolov6: A single-stage object detection framework for industrial applications. **arXiv preprint arXiv:2209.02976**, 2022.
- LI, C. *et al.* Cnn based post-processing to improve hevc. In: IEEE. **2017 IEEE International Conference on Image Processing (ICIP)**. [S.l.], 2017. p. 4577–4580.
- LONG, J. *et al.* Fully convolutional networks for semantic segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2015. p. 3431–3440.
- LONG, X. *et al.* Pp-yolo: An effective and efficient implementation of object detector. **arXiv preprint arXiv:2007.12099**, 2020.
- LUPU, R. G. *et al.* Virtual reality system for stroke recovery for upper limbs using aruco markers. In: IEEE. **2017 21st International Conference on System Theory, Control and Computing (ICSTCC)**. [S.l.], 2017. p. 548–552.
- LUZ, E. d. S. *et al.* A importância do controle de qualidade em serviços de hemodinâmica e cardiologia intervencionista. **Radiologia Brasileira**, SciELO Brasil, v. 40, p. 27–31, 2007.
- MA, D. *et al.* Mfrnet: a new cnn architecture for post-processing and in-loop filtering. **IEEE Journal of Selected Topics in Signal Processing**, IEEE, v. 15, n. 2, p. 378–387, 2020.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, p. 115–133, 1943.
- NEOGI, N. *et al.* Review of vision-based steel surface inspection systems. **EURASIP Journal on Image and Video Processing**, SpringerOpen, v. 2014, n. 1, p. 1–19, 2014.
- PEPPA, M. V. *et al.* Urban traffic flow analysis based on deep learning car detection from cctv image series. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, Copernicus Publications Göttingen, Germany, v. 42, p. 499–506, 2018.
- PHAM, L. H. *et al.* A mobile vision-based system for gap and flush measuring between planar surfaces using aruco markers. In: IEEE. **2021 International Conference on Electronics, Information, and Communication (ICEIC)**. [S.l.], 2021. p. 1–4.

- REDMON, J. *et al.* You only look once: Unified, real-time object detection. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 779–788.
- REDMON, J.; FARHADI, A. Yolov3: An incremental improvement. **arXiv preprint arXiv:1804.02767**, 2018.
- REN, S. *et al.* Faster r-cnn: Towards real-time object detection with region proposal networks. **Advances in neural information processing systems**, v. 28, 2015.
- RIBEIRO, D. *et al.* A real-time pedestrian detector using deep learning for human-aware navigation. **arXiv preprint arXiv:1607.04441**, 2016.
- ROOS-HOEFGEEST, S. *et al.* Mobile robot localization in industrial environments using a ring of cameras and aruco markers. In: IEEE. **IECON 2021–47th Annual Conference of the IEEE Industrial Electronics Society**. [S.l.], 2021. p. 1–6.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958.
- RUSSAKOVSKY, O. *et al.* Imagenet large scale visual recognition challenge. **International journal of computer vision**, Springer, v. 115, p. 211–252, 2015.
- SANG, J. *et al.* An improved yolov2 for vehicle detection. **Sensors**, MDPI, v. 18, n. 12, p. 4272, 2018.
- SANTAMARIA, M. *et al.* Content-adaptive convolutional neural network post-processing filter. In: IEEE. **2021 IEEE International Symposium on Multimedia (ISM)**. [S.l.], 2021. p. 99–106.
- SANTANA, C. J. de *et al.* A solution for counting aedes aegypti and aedes albopictus eggs in paddles from ovitraps using deep learning. **IEEE Latin America Transactions**, IEEE, v. 17, n. 12, p. 1987–1994, 2019.
- SCIUS-BERTRAND, A. *et al.* Transcription alignment of historical vietnamese manuscripts without human-annotated learning samples. **Applied Sciences**, MDPI, v. 11, n. 11, p. 4894, 2021.
- SHARMA, S. *et al.* Activation functions in neural networks. **Towards Data Sci**, v. 6, n. 12, p. 310–316, 2017.
- SILVA, L. C. C. *et al.* Evolução da mortalidade por doenças do aparelho circulatório em um município mineiro. **Rev. enferm. atenção saúde**, p. 17–26, 2019.
- SUNG, K.-K.; POGGIO, T. Example-based learning for view-based human face detection. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 20, n. 1, p. 39–51, 1998.
- TIAN, L. *et al.* Unsupervised learning of optical flow with cnn-based non-local filtering. **IEEE Transactions on Image Processing**, IEEE, v. 29, p. 8429–8442, 2020.
- VARGAS, A. C. G. *et al.* Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: SN. **Proceedings of the xxix conference on graphics, patterns and images**. [S.l.], 2016. v. 1, n. 4.

-
- VEDALDI, A. *et al.* Convolutional architecture for fast feature embedding. **Cornell University**, 2014.
- WANG, C. *et al.* Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors., 2022. DOI: <https://doi.org/10.48550/arXiv>, v. 2207, 2022.
- WEN, L. *et al.* Multiframe detection of sea-surface small target using deep convolutional neural network. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 60, p. 1–16, 2021.
- WERNER, C. J. *et al.* **MCNP User’s Manual Code Version 6.2**. Los Alamos, NM, USA, 2017. Disponível em: <http://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-17-29981>.
- YANG, F. *et al.* Exploit all the layers: Fast and accurate cnn object detector with scale dependent pooling and cascaded rejection classifiers. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 2129–2137.
- YANG, Z.; NEVATIA, R. A multi-scale cascade fully convolutional network face detector. In: **IEEE. 2016 23rd International Conference on Pattern Recognition (ICPR)**. [S.l.], 2016. p. 633–638.
- YORIYAZ, H. Método de monte carlo: princípios e aplicações em física médica. **Revista Brasileira de Física Médica**, v. 3, n. 1, p. 141–149, 2009.
- ZHANG, Z. A flexible new technique for camera calibration. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 22, n. 11, p. 1330–1334, 2000.
- ZHAO, L.; LI, S. Object detection algorithm based on improved yolov3. **Electronics**, MDPI, v. 9, n. 3, p. 537, 2020.
- ZHU, Y. *et al.* segdeepm: Exploiting segmentation and context in deep neural networks for object detection. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2015. p. 4703–4711.