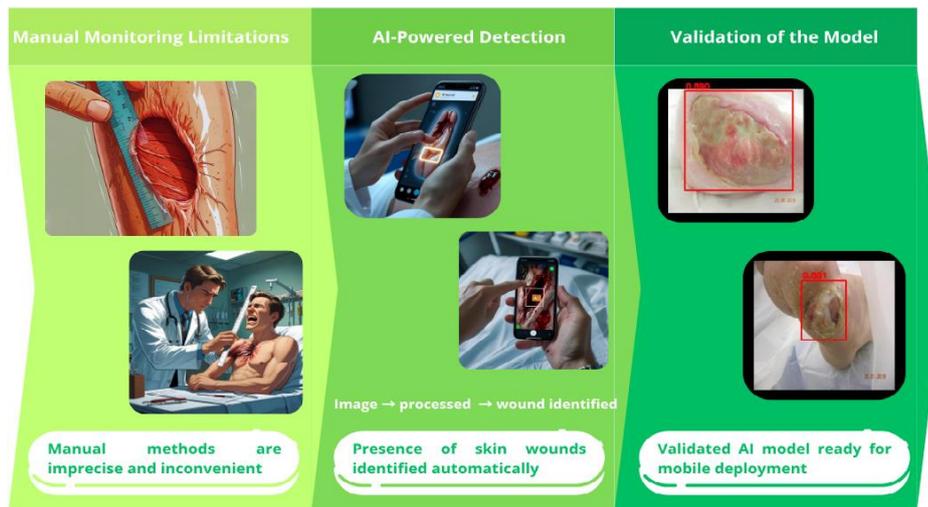


GRAPHICAL ABSTRACT

Traditional invasive methods are replaced by an AI-based system that detects skin wounds in images, functionally validated on real cases.

**DETECÇÃO DE FERIMENTOS CUTÂNEOS A PARTIR DE IMAGENS
COM O AUXÍLIO DE INTELIGÊNCIA ARTIFICIAL*****DETECTION OF SKIN WOUNDS USING ARTIFICIAL INTELLIGENCE
IN PHOTOGRAPHIC IMAGES OBTAINED FROM MOBILE DEVICES***

Anizio Moreira dos Santos Filho ¹*, André Luiz Andrade Rezende ¹

¹ Coordenação do Curso de Análise e Desenvolvimento de Sistemas, Instituto Federal de Educação, Ciência e Tecnologia Baiano – Campus Catu, R. Barão de Camaçari, 118 – Centro, Catu – BA, 48110-000, Brasil

* (anizmoreira@hotmail.com)

Resumo: Este trabalho surge no contexto do projeto Cicatríbio, iniciado em 2018, que investiga o uso do látex da mangabeira (*Hancornia speciosa*) no tratamento de ferimentos. Propõe-se uma solução automatizada para detectar ferimentos em imagens de um aplicativo móvel. Imagens de ferimentos foram coletadas e anotadas para criar um conjunto de dados de treinamento que teve sua quantidade de imagens aumentada pelo uso da técnica de data augmentation. Foi utilizado o algoritmo YOLO para treinar o modelo de detecção de ferimentos, visando a estimativa do tamanho da ferida no aplicativo. Os resultados mostram que o modelo obteve valores acima de 80% em precisão, *recall* e mAP50, demonstrando eficácia na detecção de ferimentos.

Palavras-chave: inteligência artificial; detecção de ferimentos; dispositivos móveis; planimetria digital; cicatrização de pele.

Abstract: This work emerges within the context of the Cicatríbio project, initiated in 2018, which investigates the use of latex from the mangabeira tree (*Hancornia speciosa*) in the treatment of wounds. An automated solution is proposed to detect wounds in images from a mobile application. Images of wounds were collected and annotated to create a training dataset, whose size was increased using data augmentation techniques. The YOLO algorithm was used to train the wound detection model, aiming to estimate the wound size in the application. The results show that the model achieved values above 80% in precision, *recall*, and mAP50, demonstrating effectiveness in wound detection.

Keywords: artificial intelligence; wound detection; mobile devices; digital planimetry; skin healing.

1 INTRODUÇÃO

Em 2020, a participação no Edital 05/2020, de Apoio ao Empreendedorismo Inovador do Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo (IFES), possibilitou a criação de um gel-creme regenerativo a partir do látex da Mangabeira (*Hancornia speciosa*), associado a um aplicativo operado por Inteligência Artificial (IA) para acompanhar o processo de cicatrização de lesões cutâneas em animais. A partir de conhecimentos da “sabedoria popular”, foram realizados testes *in vitro* e *in vivo* que comprovaram a eficácia cicatrizante do látex (CICATRIBIO, 2020; CICATRIBIO, 2021).

A participação no Edital, resultou ainda, em novembro de 2022, na fundação da *startup* Cicatribio, dedicada à produção de formulações farmacêuticas sustentáveis (CICATRIBIO, 2024). Atualmente, os aplicativos desenvolvidos para animais e humanos estão disponíveis na plataforma *Android* (GOOGLE PLAY, 2024a; GOOGLE PLAY, 2024b).

Inicialmente, a primeira versão do aplicativo, voltada para animais, apenas organizava registros dos ferimentos, acessíveis por tutores e veterinários. Atualmente, as versões para animais e humanos estão disponíveis na plataforma *Android*, ambas com foco na planimetria digital — registro e medição automatizados das lesões para estimar dimensões de forma não invasiva, evitando métodos manuais tradicionais. Um exemplo de método tradicional é o 'desenho da ferida', feito com acetato sobre a lesão para contorná-la.

Este procedimento apresenta dois problemas principais: riscos ao manuseio do ferimento e a possibilidade de inconsistências nas medições devido ao processo manual, suscetível a erros humanos, além do desconforto e o risco de infecções para os pacientes. Os resultados mostram que a planimetria digital por IA traz vantagens, permitindo medições mais precisas sem contato físico e evitando erros humanos.

Como mencionado, a planimetria digital por IA consiste em 2 etapas principais: identificar e medir os ferimentos. Para garantir medidas corretas, as imagens analisadas devem conter pelo menos um ferimento. Este trabalho propõe uma solução de IA para identificar com precisão imagens que contenham ferimentos, qualificando-as para o cálculo posterior das dimensões. Para isso, foram definidos objetivos específicos: pesquisar técnicas para identificar ferimentos, escolher o melhor método, testá-lo e integrá-lo aos aplicativos.

Ao longo do artigo, apresentaremos o Referencial Teórico, Processos Metodológicos, Resultados e Discussão, Conclusão e Perspectivas e, Agradecimentos aos apoiadores desse projeto.

2 REFERENCIAL TEÓRICO

O conceito de ferida adotado é fundamental para o embasamento teórico. Ferida é uma obstrução dos tecidos da pele que compromete sua função protetora (ADNAN, 2023; PRASAD, 2020), causada por cirurgias, traumas, abrasões e queimaduras (LIU, 2021). Para acompanhar a cicatrização, quantificam-se características visuais dos ferimentos, usando métodos invasivos ou não invasivos. Os métodos mais utilizados calculam as dimensões do ferimento, como o *Minimum Bounding Rectangle* (MBR) que calcula a área de um retângulo formado pelo maior comprimento e largura da ferida. Outro método consiste em desenhar o contorno da ferida sobre uma superfície transparente sobreposta. Como envolvem manipulação manual, esses métodos podem causar desconforto e imprecisão nas medidas (CHINO, 2020; TOTTOLI, 2020).

Outro conceito importante é a Inteligência Artificial (IA), que é a capacidade de uma máquina conseguir desenvolver atividades que estão relacionadas a tomadas de decisões, reconhecimento de fala, resolução de

A partir desses conceitos, iniciou-se uma pesquisa por trabalhos acadêmicos que abordassem o problema e propusessem soluções. O foco inicial da pesquisa foi por trabalhos sobre o reconhecimento de ferimentos independente do método, técnica, tipo de ferida ou organismo.

Nessa linha, observou-se a recorrência de soluções baseadas em IA para dispositivos móveis, especialmente *smartphones*. Essa tendência foi impulsionada pelos avanços tecnológicos e pela facilidade de uso destes dispositivos (WANG, 2017). Outro fato é a presença das câmeras, que facilitam a captura de imagens e vídeos que podem ser exploradas nestas soluções (YILDIZ, 2024).

Na medicina, a IA visa apoiar profissionais da saúde nas tomadas de decisões, diagnóstico, prognósticos e em tarefas que exigem manipulação de dados e automatização de tarefas administrativas (RAMESH, 2004; CROSSNOHERE, 2022). Exemplos incluem o uso da detecção de objetos para auxiliar no diagnóstico de cáries dentárias (DHANAK, 2024), a adoção de *chatbot* para recomendação de especialistas médicos para pacientes (LEE, 2021), o uso de aprendizado profundo para prever o avanço da retinopatia diabética (ARCADU, 2019) e a análise de fezes para avaliar o preparo intestinal para colonoscopia (ZHU, 2023).

Observou-se o uso de IA em smartphones no contexto de ferimentos, como para detectar úlceras em pés de diabéticos (CASSIDY, 2023), na avaliação de ferida por pressão (LAU, 2022), e na detecção automatizada de feridas (ADNAN, 2023).

Esses estudos frequentemente utilizam o modelo YOLO (*You Only Look Once*), uma rede neural convolucional que trata a detecção de objetos como um problema de regressão. Antes, a detecção de objetos usava métodos mais complexos e lentos (janelas deslizantes ou propostas de regiões), com várias etapas: extração de características, geração e classificação de caixas delimitadoras e pós-processamento.

O YOLO simplifica tudo ao usar uma única rede para rever simultaneamente caixas delimitadoras e probabilidades de classe (REDMON, 2015). O Yolov8, lançado em janeiro de 2023, adiciona suporte para a segmentação estimativa de pose, rastreamento e classificação (JUAN, 2023).

As redes neurais convolucionais (CNNs), como as utilizadas no YOLO, têm a capacidade de identificar padrões em imagens de forma automática, dispensando a necessidade de extrair manualmente características relevantes. Essa extração ocorre ao longo de várias camadas, que vão desde a detecção de traços simples, como contornos e contraste, até formas mais complexas que permitem distinguir objetos específicos. Além de possibilitar resultados precisos, as CNNs têm flexibilidade para serem adaptadas a novas tarefas, sendo essenciais em áreas como reconhecimento facial e segmentação de imagens (ALAMRI, 2022).

O desempenho dos modelos é avaliado por métricas que indicam a qualidade das detecções. Precisão representa a proporção de acertos; *Recall*, a proporção em relação ao total de acertos possíveis. O mAP50 avalia a precisão média para detecções com IoU superior a 50%. Já o mAP50-95 considera detecções com IoU de 50% a 95%, fornecendo uma visão mais abrangente da performance.

Na construção dos modelos de IA, o ambiente computacional utilizado também é relevante. O Google Colab (GC) foi a infraestrutura escolhida como ambiente para executar o processo de treinamento devido ao acesso gratuito e aos recursos computacionais robustos. Com o GC, é possível desenvolver códigos diretamente do navegador, de forma interativa e colaborativa, sem a necessidade de instalação ou configuração local. O fator decisivo na escolha foi o acesso à GPU (Graphics Processing Unit), que no GC é disponibilizada gratuitamente, embora com algumas limitações. A GPU acelera significativamente o processo de

treinamento, diminuindo o tempo de execução de dias para algumas horas.

Além do uso de plataformas gratuitas como o Google Colab, existem também infraestruturas de alta performance dedicadas a projetos de inteligência artificial. Um exemplo é o Ògún, um supercomputador de alta performance (High Performance Computing - HPC), capaz de realizar 104 trilhões de operações por segundo, e que é o terceiro supercomputador do Centro de Supercomputação para Inovação Industrial do SENAI CIMATEC. O seu intuito é melhorar significativamente as simulações em áreas como cálculos de estruturas, dinâmica de fluidos, montagem de materiais, previsão de eventos e tecnologias de inteligência artificial (FIEB, 2018).

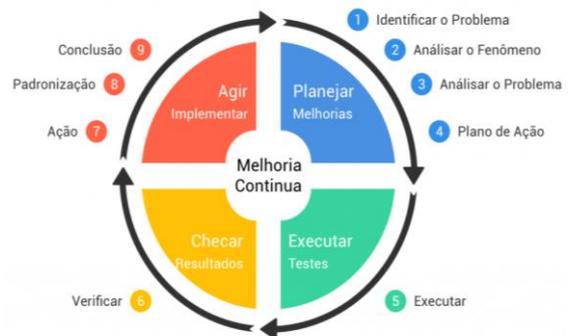
Para garantir resultados mais robustos e confiáveis, além do aprimoramento do ambiente computacional, é fundamental também trabalhar na qualidade e variedade dos dados. Nesse sentido, utilizou-se a técnica Data Augmentation. Essa técnica consiste em aplicar transformações nos dados existentes para gerar novos dados, ampliando a variedade e quantidade do conjunto original.

Além das técnicas específicas de detecção e reconhecimento de feridas, a organização do processo de desenvolvimento e avaliação dos modelos também é fundamental para alcançar resultados consistentes. Nesse sentido, o método PDCA (Plan-Do-Check-Act), concebido em 1950 por W. Edwards Deming, embora seja tradicionalmente utilizado como modelo de gestão da qualidade (ISNIAH, 2020; NGUYEN, 2020, HONG, 2020), também pode ser adaptado para projetos de inteligência artificial.

Dividido em quatro etapas — planejar, fazer, verificar e agir — o PDCA organiza o processo de melhoria contínua. No planejamento (Plan), são definidos os objetivos e o caminho para atingi-los. Na fase de execução (Do), as ações são

realizadas de acordo com o plano traçado. Em seguida, a verificação (Check) consiste em analisar os resultados alcançados, comparando-os com os objetivos estipulados. Por fim, o agir (Act) envolve confirmar se os objetivos foram atingidos ou se há necessidade de ajustes e novos planos para otimizar o processo. Ao término dessa fase, o ciclo recomeça, permitindo que novas avaliações e aprimoramentos sejam continuamente realizados (ZAILANI, 2023).

Imagem 1: Representação visual do ciclo PDCA.



Fonte: (Vida de Produto, 2020)

3 PROCESSOS METODOLÓGICOS

A seguir, apresentamos as etapas da metodologia dívida nos ciclos 1 e 2. As ações específicas estão descritas no Quadro 1 e são aprofundadas no tópico “3.2 Desenvolvimento Metodológico”.

3.1 METODOLOGIA

As atividades deste projeto foram organizadas segundo as etapas do ciclo PDCA, facilitando a identificação de melhorias e ajustes contínuos. O Quadro 1 apresenta uma visão geral das atividades desenvolvidas em cada fase por ciclo

Quadro 1: Distribuição das atividades por etapa do Ciclo PDCA.

Ciclo 1	
Planejar	Construir um conjunto de dados com total próximo a 10 mil imagens de ferimentos.
	Os valores das variáveis de desempenho do modelo devem ser iguais ou superiores a 80%.
Fazer	Buscar por repositórios públicos de imagens de ferimentos.
	Buscar por ambientes de treinamento que atendessem aos critérios pré-estabelecidos
	Configurar o ambiente de treinamento.
	Treinar o modelo.
Verificar	Validar as implementações realizadas.
Agir	O conjunto de dados com total próximo a 10 mil imagens de ferimentos foi construído?
	Os valores das variáveis de desempenho do modelo foram iguais ou superiores a 80%?
Ciclo 2	
Planejar	Ampliar o conjunto de dados para no mínimo 50 mil imagens de ferimentos.
	Os valores das variáveis de desempenho do modelo devem ser iguais ou superiores a 80%.
Fazer	Aplicar a <i>Data Augmentation</i> .
	Treinar o modelo.
Verificar	Validar as implementações realizadas.

Agir	O conjunto de dados foi ampliado para no mínimo 50 mil imagens de ferimentos?
	Os valores das variáveis de desempenho do modelo foram iguais ou superiores a 80%?

Fonte: Elaborado pelo autor.

3.1.1 Planejar (*Plan*)

No primeiro ciclo de planejamento, foram definidos dois objetivos principais. O primeiro, criar um conjunto de imagens de ferimentos com cerca de 10 mil fotos. O segundo era estabelecer 80% como meta mínima para *Precisão*, *Recall*, *mAP50* e *mAP50-95*.

No segundo ciclo, o planejamento foi revisado, ampliando o conjunto de imagens para 50 mil fotos, aumentando a variedade de imagens. Essa ação buscou melhorar o desempenho do modelo, já que a meta mínima não foi atingida no primeiro ciclo.

3.1.2 Executar (*Do*)

Para alcançar o objetivo 1 do primeiro ciclo - *construir um conjunto de dados com total próximo a 10 mil imagens de ferimentos* - buscou-se repositórios públicos, evitando a coleta manual. Essa escolha foi feita pois identificar 10 mil imagens manualmente seria inviável no prazo do projeto.

Para o objetivo 2 - *Os valores das variáveis de desempenho do modelo devem ser iguais ou superiores a 80%* - foram estabelecidas três atividades. Determinar o ambiente de treinamento priorizando gratuidade e execução direta no navegador. Configurar o ambiente e executar o treinamento.

Para o objetivo 1 do segundo ciclo - *Ampliar o conjunto de dados para no mínimo 50 mil imagens de ferimentos* - utilizou-se a técnica *Data Augmentation*.

Para o objetivo 2 - *Os valores das variáveis de desempenho do modelo devem ser iguais ou superiores a 80%* - o

treinamento foi executado usando o novo conjunto de dados.

3.1.3 Verificar (*Check*)

Para o objetivo 1 do primeiro ciclo - *construir um conjunto de dados com total próximo a 10 mil imagens de ferimentos* - foram identificados repositórios públicos que atenderam à meta de fotos.

Para o objetivo 2 - *Os valores das variáveis de desempenho do modelo devem ser iguais ou superiores a 80%* - constatou-se que os valores de desempenho não atingiram o mínimo após o treinamento.

No segundo ciclo, o objetivo 1 - *Ampliar o conjunto de dados para no mínimo 50 mil imagens de ferimentos* - foi atendido.

Por fim, para o objetivo 2 - *Os valores das variáveis de desempenho do modelo devem ser iguais ou superiores a 80%* - foi verificado que as variáveis de desempenho do modelo atingiram o percentual definido.

3.1.4 Agir (*Act*)

No primeiro ciclo, os questionamentos foram: o conjunto de dados com cerca de 10 mil imagens de ferimentos foi construído? E as variáveis de desempenho atingiram 80%?

O conjunto de 10 mil imagens foi montado a partir de repositórios gratuitos. No entanto, as variáveis de desempenho não atingiram o mínimo esperado.

Diante disso, o segundo ciclo levantou novas questões: o conjunto foi ampliado para pelo menos 50 mil imagens? E as variáveis de desempenho atingiram 80%?

A quantidade de imagens foi ampliada para 50 mil usando *Data Augmentation*. Os valores das variáveis superaram o mínimo estabelecido: Precisão - 99,5%; *Recall* - 99,7%; MAP50 - 99,5% e o MAP50-95 - 87,7%. Com isso, não foi necessário um novo ciclo.

3.2 DESENVOLVIMENTO METODOLÓGICO

A construção do conjunto de dados iniciou-se no Roboflow (ROBOFLOW, 2024a). As palavras-chaves ‘*ulcer*’, ‘*injury*’, ‘*laceration*’ e ‘*wound*’ foram utilizadas no processo de pesquisa que ao ser finalizado proporcionou um total de 9753 imagens.

Uma vez reunidos, o processo de marcação (*labeling*) foi executado simultaneamente à revisão das imagens: enquanto as irrelevantes eram excluídas, as relevantes eram anotadas. Ao final, o conjunto totalizou 9180 imagens.

Este montante foi dividido em 80% para treino e 20% para validação. Com o conjunto de dados pronto, configurou-se o ambiente e deu-se início ao treinamento.

Inicialmente, os resultados no Google Colab foram promissores, mas problemas surgiram com o aumento das épocas de treino. O tempo de execução ultrapassou uma hora, chegando ao limite do uso gratuito da GPU. Quando esse limite era atingido, a GPU ficava indisponível e o ambiente de execução era encerrado, fazendo com que todo o progresso fosse perdido.

Tal situação, impedia que o modelo ultrapassasse 80% de desempenho. Para resolver isso, buscou-se ampliar o conjunto para pelo menos 50 mil imagens, visando melhorar os resultados. A busca manual por novas imagens foi descartada devido ao tempo e retrabalho envolvidos. Após pesquisa em periódicos, optou-se por uma técnica automatizada especializada na geração de novas imagens.

A técnica utilizada foi *Data Augmentation*, que aplica transformações às imagens existentes para gerar novos dados, aumentando a variedade e o volume do conjunto. A proporção adotada foi 1:8, gerando oito imagens novas para cada original. Para isso, utilizou-se a biblioteca *Albumentations* (ALBUMENTATIONS, 2024).

Essa biblioteca oferece várias transformações; para este trabalho, foram

selecionadas: *Rotate*, *HorizontalFlip*, *VerticalFlip*, *RGBShift*, *ToGray*, *Blur*, *ColorJitter*.

Cada transformação modifica a imagem de forma específica: *Rotate* gira em ângulo aleatório; *HorizontalFlip* e *VerticalFlip* invertem nos eixos correspondentes; *RGBShift* ajusta aleatoriamente os canais de cor; *ToGray* converte para escala de cinza; *Blur* aplica desfoque; e *ColorJitter* altera brilho, contraste e saturação.

Antes da aplicação da *Data Augmentation*, o número total de imagens no conjunto de dados era de 9.180. Após o processo de aumento, foram criadas 73.440 novas, totalizando 82.620 imagens.

Com o aumento do conjunto de dados, esperava-se melhor desempenho em menos épocas de treinamento. No entanto, ao finalizar o treinamento, os resultados não corresponderam às expectativas. Apesar de ser previsto que um conjunto maior exigisse mais tempo para treinar, surpreendeu que em poucas épocas o tempo limite do *Google Colab* fosse excedido.

Para que fosse possível usufruir de todo o potencial do novo conjunto de dados, era preciso treinar por mais épocas. Como o Colab não atendia esta demanda, a partir de uma parceria com o Centro de Supercomputação do SENAI CIMATEC, os treinos foram transferidos para o Ògún, um ambiente mais robusto. A principal dificuldade nessa transferência foi preservar a estrutura já estabelecida. Após a migração e as configurações necessárias, os treinos no Ògún tiveram início.

4 RESULTADOS E DISCUSSÃO

Os resultados foram apresentados em duas etapas. A primeira mostra a evolução do modelo, com as métricas de cada cenário após o treinamento. A segunda apresenta o desempenho do melhor modelo, testado novamente com um conjunto inédito de imagens de diferentes tipos de feridas.

Foram realizados vários treinos em cada ambiente, mas apenas três cenários de cada foram destacados. Esses seis checkpoints resumem a progressão do desempenho das variáveis do modelo ao longo do desenvolvimento. O Quadro 2 apresenta a configuração dos cenários.

Quadro 2: Configuração dos cenários.

Cenário	Ambiente	N.º de imagens	N.º de épocas
I	Google Colab	7.344	005
II		7.344	050
III		7.344	100
IV	Ògún	66.090	005
V		66.090	050
VI		66.090	100

Fonte: Elaborado pelo autor.

A Tabela 1 apresenta as métricas obtidas após o treinamento do modelo de detecção de feridas nos cenários I, II e III.

Tabela 1: Métricas e resultados dos cenários I, II e III.

Ambiente Google Colab			
Métricas	Cenário I	Cenário II	Cenário III
N.º de imagens	7.344	7.344	7.344
N.º de épocas	5	50	100
Precisão	0,594	0,706	0,739
Recall	0,513	0,623	0,637
MAP50	0,516	0,624	0,659
MAP50-95	0,188	0,255	0,290

Fonte: Elaborado pelo autor.

Após analisar os resultados da Tabela 1, verificou-se que o desempenho do

modelo melhora conforme aumenta o número de épocas de treinamento. Os ganhos foram mais expressivos nas primeiras etapas, com uma evolução mais lenta nas épocas seguintes. Isso indica que o modelo atingiu um limite de aprendizado com os dados disponíveis, tornando pouco vantajoso prolongar o treinamento devido ao custo computacional envolvido.

Além disso, essa evolução reduzida das métricas nas últimas épocas sugere o risco de ocorrência de *overfitting*, um fenômeno comum em aprendizado de máquina. O *overfitting* ocorre quando o modelo começa a memorizar excessivamente os exemplos específicos do conjunto de treinamento, em vez de identificar padrões gerais úteis para novos dados. Isso significa que o modelo pode apresentar um desempenho muito bom nos dados em que foi treinado, mas falhar ao ser exposto a exemplos inéditos. Em outras palavras, o modelo se torna muito especializado no conjunto de treinamento e perde a capacidade de generalizar, o que vai contra o objetivo de criar um sistema robusto e confiável.

Dando continuidade, foram realizados os testes manuais como forma de simular uma situação real no contexto da detecção de feridas, utilizando um conjunto de imagens inéditas, não empregadas no processo de treinamento.

Nos testes, as imagens são submetidas ao modelo para possibilitar a avaliação visual da detecção. Quando o modelo identifica a presença da ferida, um quadrilátero delimita a região detectada, facilitando a análise do comportamento do modelo. Essa abordagem permite também identificar eventuais falhas e suas possíveis causas.

A seguir, serão apresentadas algumas imagens utilizadas nos testes manuais dos modelos gerados a partir dos cenários I, II e III.

Imagem 1: Imagens utilizadas nos testes manuais do cenário III.



Fonte:(Roboflow, 2024b)

Os testes manuais mostraram que o modelo do cenário III identificou corretamente o posicionamento das feridas, delimitadas por quadriláteros vermelhos. Contudo, em alguns casos, como na Imagem 1, segundo ferimento, foram observados comportamentos indesejados, como sobreposição, que afetaram negativamente as métricas de MAP. Isso indica a necessidade de aprimorar a precisão na localização das feridas, uma constatação válida para todos os cenários apresentados na Tabela 1.

Na Imagem 2, foi possível identificar os falsos negativos. Como exemplo, a imagem mostra duas feridas não detectadas pelo modelo, o que justifica o valor observado na métrica de *Recall*.

Imagem 2: Teste manual do cenário III com presença dos falsos negativos.



Fonte: (Roboflow, 2024b).

Os resultados dos testes manuais reforçaram a necessidade de aprimorar o modelo do cenário III. A partir desse momento, o treinamento, que anteriormente era realizado no ambiente Colab, passou a ser executado no Ògún, mantendo o número inicial de épocas.

Após o treinamento no Ògún, foi possível reavaliar o desempenho do modelo. A Tabela 2 apresenta os resultados das métricas.

Tabela 2: Métricas e resultados dos cenários IV, V e VI.

Ambiente Ògún			
Métricas	Cenário IV	Cenário V	Cenário VI
N.º de imagens (treino)	66090	66090	66090
N.º de épocas	5	50	100
Precisão	0,809	0,990	0,995
Recall	0,719	0,985	0,997
MAP50	0,815	0,994	0,995
MAP50-95	0,515	0,818	0,877

Fonte: Elaborado pelo autor.

Comparando os resultados dos cenários III e IV das Tabelas 1 e 2 respectivamente, foi observado um progresso de 7,0% na Precisão, 8,2% no *Recall*, 15,6% no MAP50 e 22,5% no MAP50-95. A inclusão de novas imagens no treinamento proporcionou a evolução do desempenho do modelo. Assim como nos resultados da Tabela 1, o avanço das métricas continuou a se evidenciar com o aumento do número de épocas, conforme mostrado na Tabela 2.

Ao expandir o número de épocas de 5 para 100, a Precisão avançou de 80,9% para 99,5%. O *Recall* progrediu de 71,9% para 99,7%, enquanto o MAP50 aumentou de 81,5% para 99,5%. Já o MAP50-95 foi elevado de 51,5% para 87,7%

Esses resultados permitem destacar três aspectos principais. O primeiro está relacionado ao desempenho apresentado entre o modelo do cenário IV (Tabela 2) e o do cenário III (Tabela 1). Mesmo com

noventa e cinco épocas adicionais para se desenvolver, o modelo do cenário III não superou as métricas do cenário IV, o que demonstra que a quantidade de imagens de treinamento teve influência direta na melhora dos índices.

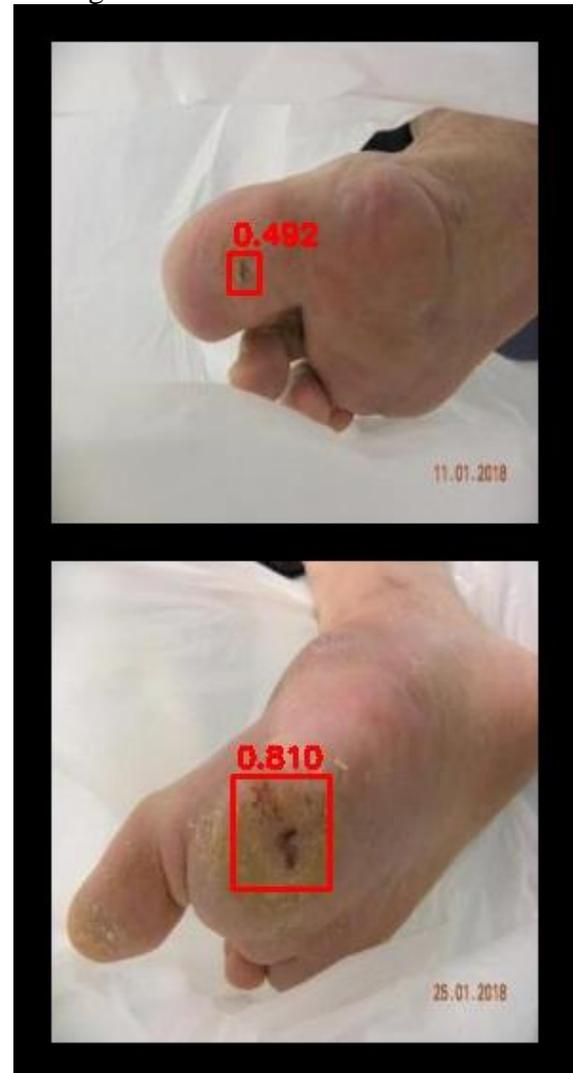
O segundo ponto refere-se à evolução observada ao aumentar o número de épocas para 50, como demonstrado no cenário V da Tabela 2. A Precisão, o *Recall* e o MAP50 ultrapassaram a marca dos 98%, indicando que o modelo alcançou um desempenho próximo da excelência na identificação da presença de feridas.

Por fim, o cenário VI apresentou um comportamento semelhante ao observado no cenário III. Houve um avanço mais modesto das métricas entre os cenários V e VI, confirmado pelos valores de 0,5%, 1,2%, 0,1% e 5,9% na Precisão, *Recall*, MAP50 e MAP50-95, respectivamente.

Cabe destacar que não há espaço para uma melhora acentuada, pois o cenário V já apresenta valores próximos a 100%, conforme ilustrado na Tabela 2. No cenário VI, com Precisão de 99,5%, observou-se que os falsos positivos regrediram de 26,1% (no cenário III da Tabela 1) para 0,5%. Essa mesma redução foi observada nos falsos negativos, que caíram de 36,6% para 0,3%, indicando que apenas 0,3% das feridas não são detectadas pelo modelo.

Em consonância com o percurso metodológico, ao término do processamento do modelo, foi realizada a validação manual utilizando as mesmas imagens do cenário III da Tabela 1. Nessa etapa, o modelo detectou feridas que anteriormente não haviam sido identificadas e conseguiu localizar precisamente o posicionamento delas, como mostrado na Imagem III.

Imagem 3: Teste manual do cenário VI.



Fonte:(Roboflow, 2024b).

Conforme definido no início desta seção, o próximo passo consistiu em testar o modelo com melhor desempenho em dois conjuntos de imagens inéditas. Assim, o modelo gerado no cenário VI da Tabela 2 foi submetido a esses novos conjuntos de imagens, conforme apresentado na Tabela 3

Tabela 3: Métricas e resultados de validação do cenário VI.

Novos conjuntos de dados		
Métricas	Conjunto n.º 1	Conjunto n.º 2
N.º de imagens	4061	2583
Precisão	0,612	0,744
Recall	0,540	0,463
MAP50	0,536	0,592
MAP50-95	0,233	0,348

Fonte: Elaborado pelo autor.

A Tabela 3 apresenta os resultados do teste do modelo do cenário VI em dois conjuntos de dados inéditos. Os resultados ficaram abaixo do valor mínimo definido na metodologia, com Precisão e Recall inferiores a 75% e 55%, respectivamente. O desempenho do modelo foi marcado por elevados índices de falsos positivos e falsos negativos.

Inicialmente, essa queda no desempenho poderia ser interpretada como um indicativo de *overfitting*, uma vez que o modelo apresentou bons resultados durante o treinamento, mas não generalizou bem para novos dados. Contudo, uma análise mais detalhada revelou a causa do problema: as *labels* do conjunto de dados apresentavam inconsistências ou estavam ausentes, o que impactou negativamente a métrica de precisão. A Imagem 4 ilustra essa situação: na fotografia superior, um quadrilátero indica a localização do ferimento de acordo com o gabarito, enquanto na fotografia inferior são destacados os ferimentos detectados pelo modelo. Quando o modelo identificava um ferimento que não estava devidamente

etiquetado, essa detecção era considerada falso positivo.

Imagem 4: Origem da precisão reduzida.

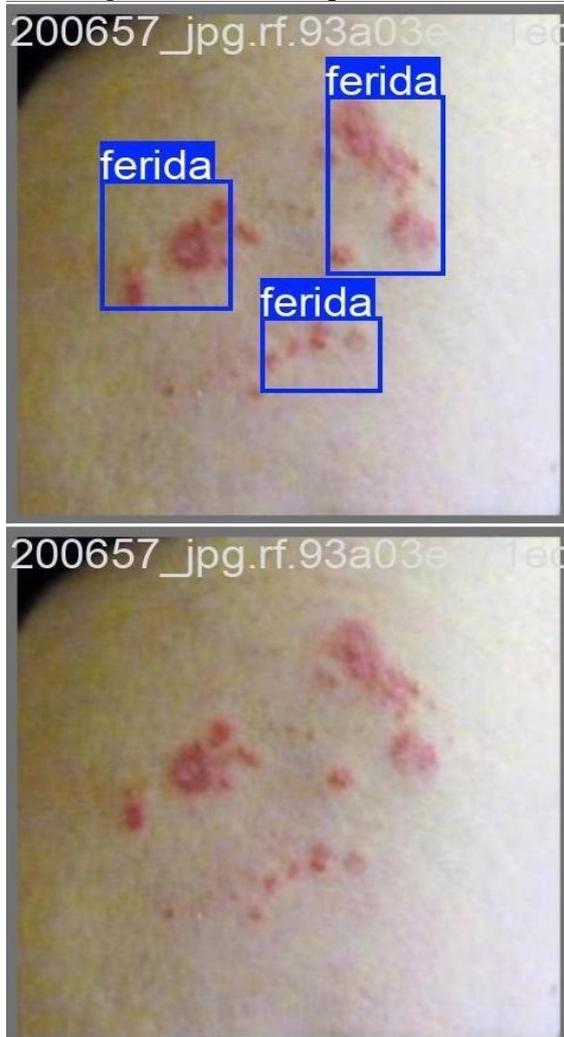


Fonte: (Roboflow, 2024b).

O Recall enfrentou um desafio distinto. Enquanto a Precisão foi prejudicada pela ausência de marcações, o Recall sofreu com o excesso delas. Algumas imagens apresentavam marcações que indicavam a presença de feridas, mas que correspondiam a outras condições dermatológicas. Nesses casos, o modelo não reconheceu tais marcações como feridas, e, por isso, essas detecções foram contabilizadas como falsos negativos, o que comprometeu a métrica de Recall. Na Imagem 5, é possível observar esse tipo de

situação: a primeira fotografia mostra as marcações presentes no conjunto de dados e a segunda imagem exhibe a detecção realizada pelo modelo.

Imagem 5: Fonte da queda do *recall*.



Fonte: (Roboflow, 2024b).

Note-se que, na segunda fotografia da Imagem 5, nenhuma marcação foi criada, o que indica que, após a análise, o modelo não identificou nenhuma região como ferida. Embora esses casos representem apenas uma pequena fração do conjunto de dados, foram suficientes para distorcer a percepção do desempenho do modelo. Por isso, tornou-se evidente a necessidade de realizar um novo teste, desta vez com o conjunto de dados revisado e limpo, garantindo a coerência das marcações.

Após essa limpeza, realizada apenas no conjunto número 1, um novo teste foi

feito com o mesmo conjunto, gerando resultados que refletem melhor a capacidade do modelo. Os valores obtidos foram 80,07%, 89,5%, 88% e 56,2% para Precisão, *Recall*, MAP50 e MAP50-95 respectivamente. Ao invés de mostrar imagens deste teste, apresento exemplos de feridas que não são úlceras de pé, para avaliar a capacidade do modelo de generalizar e identificar corretamente feridas em contextos diferentes daqueles em que foi treinado. Essa análise ajuda a entender o potencial do modelo para aplicações mais amplas.

Imagem 6: Ferimento em perna humana.



fonte: (Roboflow, 2025).

Imagem 7: Ferimento em braço humano.



fonte: (Roboflow, 2025).

As imagens 6 e 7 demonstram que o modelo conseguiu identificar corretamente feridas que não são de úlceras de pé, indicando uma boa capacidade de generalização. Isso sugere que o modelo não está simplesmente decorando padrões específicos, mas aprendendo características relevantes das feridas em geral.

5 CONCLUSÃO E PERSPECTIVAS

Este trabalho propôs uma solução automatizada para a detecção de ferimentos cutâneos em imagens, com a finalidade de alcançar desempenho satisfatório nas métricas de Precisão, Recall, MAP50 e MAP50-95, conforme definido na metodologia.

Os experimentos demonstraram que o aumento da quantidade de dados de treinamento e do número de épocas contribuiu para a melhoria dos resultados, especialmente ao utilizar o ambiente Ògún com um conjunto ampliado de imagens. O modelo do cenário VI apresentou o melhor desempenho nas avaliações internas.

Nos testes realizados com conjuntos inéditos de imagens, os resultados iniciais foram impactados por inconsistências nas anotações, prejudicando as métricas de Precisão e Recall. Após a revisão e limpeza do conjunto número 1, o modelo atingiu

80,07% em Precisão, 89,5% em Recall, 88% em MAP50 e 56,2% em MAP50-95.

Dessa forma, a solução proposta atingiu ou superou a meta estabelecida para as métricas de Precisão, Recall e MAP50, enquanto o MAP50-95 indicou a necessidade de aprimoramento para garantir maior robustez em diferentes limiares de confiança.

Os resultados evidenciam o potencial da abordagem automatizada para auxiliar na identificação precisa de feridas, ressaltando também a importância da qualidade e consistência das anotações para a avaliação do desempenho do modelo.

Como direcionamento para trabalhos futuros, sugere-se a ampliação da solução para incluir a classificação dos ferimentos por tipo ou estágio, o que poderia aprimorar ainda mais o suporte clínico e o diagnóstico.

AGRADECIMENTOS

Os autores gostariam de expressar seu agradecimento ao Centro de Supercomputação para Inovação Industrial do SENAI CIMATEC, pela cooperação, fornecimento e operação das instalações computacionais, bem como pelo compromisso contínuo em investir em pesquisa, desenvolvimento e inovação, beneficiando tanto a indústria quanto a sociedade na totalidade.

REFERÊNCIAS

ALBUMENTATIONS. Biblioteca para aumento de dados em visão computacional. Disponível em: <<https://albumentations.ai/>>. Acesso em: 10 out. 2024.

ADNAN, M. et al. **An Automatic Wound Detection System Empowered by Deep Learning**. Journal of Physics: Conference Series. **Anais...Institute of Physics**, 2023. DOI: 10.1088/1742-6596/2547/1/012005. Disponível em:

<https://iopscience.iop.org/article/10.1088/1742-6596/2547/1/012005>. Acesso em: 01 out. 2024.

ALAMRI, N. M. H.; PACKIANATHER, M.; BIGOT, S. Deep Learning: Parameter Optimization Using Proposed Novel Hybrid Bees Bayesian Convolutional Neural Network. **Applied Artificial Intelligence**, v. 36, n. 1, 2022. Disponível em: <https://www.tandfonline.com/doi/epdf/10.1080/08839514.2022.2031815?needAccess=true>. Acesso em: 01 jun. 2025

ALDOSERI, A.; AL-KHALIFA, K. N.; HAMOUDA, A. M. **Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges**. **Applied Sciences (Switzerland)**MDPI, , 1 jun. 2023. DOI: **10.3390/app13127082**. Disponível em: <https://www.mdpi.com/2076-3417/13/12/7082>. Acesso em: 09 out. 2024.

ARCADU, F. et al. Deep learning algorithm predicts diabetic retinopathy progression in individual patients. **npj Digital Medicine**, v. 2, n. 1, 1 dez. 2019. DOI: 10.1038/s41746-019-0172-3. Disponível em: <https://www.nature.com/articles/s41746-019-0172-3>. Acesso em: 05 set. 2024.

CASSIDY, B. et al. Artificial intelligence for automated detection of diabetic foot ulcers: A real-world proof-of-concept clinical evaluation. **Diabetes Research and Clinical Practice**, v. 205, 1 nov. 2023. DOI: 10.1016/j.diabres.2023.110951. Disponível em: [https://www.diabetesresearchclinicalpractice.com/article/S0168-8227\(23\)00714-3/fulltext](https://www.diabetesresearchclinicalpractice.com/article/S0168-8227(23)00714-3/fulltext). Acesso em: 24 set. 2024.

CHINO, D. Y. T. et al. Segmenting skin ulcers and measuring the wound area using deep convolutional networks. **Computer Methods and Programs in Biomedicine**, v. 191, 1 jul. 2020. DOI: 10.1016/j.cmpb.2020.105376. Disponível

em:

<https://www.sciencedirect.com/science/article/pii/S016926071931404X?via%3Dihub>. Acesso em: 09 out. 2024

CICATRIBIO. Patente BR1020200143875. Instituto Nacional da Propriedade Industrial – INPI, 2020.

CICATRIBIO. Patente BR1020210054719. Instituto Nacional da Propriedade Industrial – INPI, 2021.

CICATRIBIO. Página oficial. Disponível em: <<https://www.cicatribio.com.br>>. Acesso em: 10 out. 2024.

CROSSNOHERE, N. L. et al. **Guidelines for Artificial Intelligence in Medicine: Literature Review and Content Analysis of Frameworks**. **Journal of Medical Internet Research**JMIR Publications Inc., , 1 ago. 2022. DOI: 10.2196/36823. Disponível em: <https://www.jmir.org/2022/8/e36823>. Acesso em: 11 set. 2024.

DHANAK, N. et al. Artificial intelligence enabled smart phone app for real-time caries detection on bitewing radiographs. **Bioinformatics**, v. 20, n. 3, p. 243–247, 31 mar. 2024. DOI: 10.6026/973206300200243. Disponível em: <https://www.bioinformatics.net/020/973206300200243.htm>. Acesso em: 01 out. 2024.

FIEB. SENAI CIMATEC amplia capacidade em supercomputação e deverá ser potência no Hemisfério Sul. Publicado em: 25 out. 2018. Disponível em: <https://www.senaicimatec.com.br/noticias/senai-cimatec-amplia-capacidade-em-supercomputacao-e-devera-ser-potencia-no-hemisferio-sul/>. Acesso em: 28 ago. 2024.

GOOGLE PLAY. Cicatribio Vet. Disponível em:

<https://play.google.com/store/apps/details?id=com.cicatribioVet&hl=pt>. Acesso em: 10 out. 2024.

GOOGLE PLAY. Cicatribio Skin. Disponível em: https://play.google.com/store/apps/details?id=com.cicatribioSKIN&hl=pt_BR. Acesso em: 10 out. 2024.

HONG, Y. et al. Continuous improvement on the rationality of prophylactic injectable PPIs usage by a clinical pharmacist-led guidance team at a Chinese tertiary teaching hospital. **Journal of International Medical Research**, v. 48, n. 10, 2020. DOI: 10.1177/0300060520954729. Disponível em: <https://journals.sagepub.com/doi/10.1177/0300060520954729>. Acesso em: 15 set. 2024.

HULSEN, T. **Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare.** AI (Switzerland) Multidisciplinary Digital Publishing Institute (MDPI), , 1 set. 2023. DOI: **10.3390/ai4030034**. Disponível em: <https://www.mdpi.com/2673-2688/4/3/34>. Acesso em: 09 out. 2024

HUSSAIN, M. **YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection.** **Machines** Multidisciplinary Digital Publishing Institute (MDPI), , 1 jul. 2023. DOI: **10.3390/machines11070677**. Disponível em: <https://www.mdpi.com/2075-1702/11/7/677>. Acesso em: 03 set. 2024.

ISNIAH, S.; HARDI PURBA, H.; DEBORA, F. Plan do check action (PDCA) method: literature review and research issues. **Jurnal Sistem dan Manajemen Industri**, v. 4, n. 1, p. 72–81, 31 jul. 2020. DOI: 10.30656/jsmi.v4i1.2186. Disponível em: <https://ejournal.lppmunsera.org/index.php/JSMI/article/view/2186>. Acesso em: 20 ago. 2024.

LAU, C. H. et al. An artificial intelligence-enabled smartphone app for real-time pressure injury assessment. **Frontiers in Medical Technology**, v. 4, 2022. DOI: 10.3389/fmedt.2022.905074. Disponível em: <https://www.frontiersin.org/journals/medical-technology/articles/10.3389/fmedt.2022.905074/full>. Acesso em: 09 out. 2024.

LEE, H.; KANG, J.; YEO, J. Medical specialty recommendations by an artificial intelligence chatbot on a smartphone: Development and deployment. **Journal of Medical Internet Research**, v. 23, n. 5, 1 maio 2021. DOI: 10.2196/27460. Disponível em: <https://www.jmir.org/2021/5/e27460>. Acesso em: 13 set. 2024.

LIU, D. S. et al. Perceptions of US Medical Students on Artificial Intelligence in Medicine: Mixed Methods Survey Study. **JMIR Medical Education**, v. 8, n. 4, 1 dez. 2022. DOI: 10.2196/38325. Disponível em: <https://mededu.jmir.org/2022/4/e38325>. Acesso em: 09 out. 2024

LIU, W. et al. Bioactive antiinflammatory antibacterial hemostatic citrate-based dressing with macrophage polarization regulation for accelerating wound healing and hair follicle neogenesis. **Bioactive Materials**, v. 6, n. 3, p. 721–728, 1 mar. 2021. DOI: 10.1016/j.bioactmat.2020.09.008. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2452199X20302176?via%3Dihub>. Acesso em: 15 set. 2024 .

MA, N. et al. Wheat Seed Detection and Counting Method Based on Improved YOLOv8 Model. **Sensors**, v. 24, n. 5, 1 mar. 2024. DOI: **10.3390/s24051654**. Disponível em: <https://www.mdpi.com/1424-8220/24/5/1654>. Acesso em: 28 ago. 2024

- NGUYEN, V. et al. Practical application of plan-do-check-act cycle for quality improvement of sustainable packaging: A case study. **Applied Sciences (Switzerland)**, v. 10, n. 18, 1 set. 2020. DOI: **10.3390/app10186332**. Disponível em: <https://www.mdpi.com/2076-3417/10/18/6332>. Acesso em: 09 set. 2024.
- PEREIRA, I. B. S.; SOUZA, M. V. S. Cicatrício: Produção de formulações farmacêuticas e desenvolvimento de software para tratamento de ferimentos cutâneos. In: **MOSTRATEC, 2022, Novo Hamburgo**. Anais eletrônicos da Mostratec. Novo Hamburgo: Fundação Liberato, 2022. Disponível em: <https://virtual.mostratec.com.br/projeto/cicatrício-producao-de-formulacoes-farmacêuticas-e-desenvolvimento-de-software-para-tratamento-de-ferimentos-cutâneos/>. Acesso em: 01 set. 2024.
- PRASAD, Y. S. et al. An injectable self-healing anesthetic glycolipid-based oleogel with antibiofilm and diabetic wound skin repair properties. **Scientific Reports**, v. 10, n. 1, 1 dez. 2020. DOI: 10.1038/s41598-020-73708-7. Disponível em: <https://www.nature.com/articles/s41598-020-73708-7>. Acesso em: 09 out. 2024.
- RAMESH, A. N. et al. **Artificial intelligence in medicine**. **Annals of the Royal College of Surgeons of England**, set. 2004. DOI: 10.1308/147870804290. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1964229/pdf/15333167.pdf>. Acesso em: 13 set. 2024.
- REDMON, J. et al. You Only Look Once: Unified, Real-Time Object Detection. 8 jun. 2015. DOI: 10.48550/arXiv.1506.02640. Disponível em: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CV_PR_2016_paper.pdf. Acesso em: 20 set. 2024.
- ROBOFLOW. Plataforma de gerenciamento de dados para visão computacional. Disponível em: <https://roboflow.com/>. Acesso em: 10 out. 2024.
- ROBOFLOW. Detect Ulcer Dataset. 15 fev. 2024. Disponível em: https://universe.roboflow.com/detectulcer/detect_ulcer/dataset/2. Acesso em: 24 jun. 2025.
- ROBOFLOW. Wound Detect Dataset. 17 mar. 2025. Disponível em: <https://universe.roboflow.com/howard-tqye6/wound-detect-aiykj/dataset/1>. Acesso em: 24 jun. 2025.
- TANG, G. et al. A Survey of Object Detection for UAVs Based on Deep Learning. **Remote Sensing**, v. 16, n. 1, 1 jan. 2024. DOI: **10.3390/rs16010149**. Disponível em: <https://www.mdpi.com/2072-4292/16/1/149>. Acesso em: 10 out. 2024.
- TOTTOLI, E. M. et al. **Skin wound healing process and new emerging technologies for skin wound care and regeneration**. **Pharmaceutics** MDPI AG, , 1 ago. 2020. DOI: **10.3390/pharmaceutics12080735**. Disponível em: <https://www.mdpi.com/1999-4923/12/8/735>. Acesso em: 09 out. 2024
- VIDA DE PRODUTO. Tecnologia e Produto. 25 ago. 2020. Disponível em: <https://vidadeproduto.com.br/ciclo-pdca/>. Acesso em: 10 out. 2024.
- WANG, S. C. et al. Point-of-care wound visioning technology: Reproducibility and accuracy of a wound measurement app. **PLoS ONE**, v. 12, n. 8, 1 ago. 2017. DOI: 10.1371/journal.pone.0183139. Disponível em: <https://journals.plos.org/plosone/article?id>

=10.1371/journal.pone.0183139. Acesso em: 01 out. 2024.

WANG, X. et al. Small Object Detection Based on Deep Learning for Remote Sensing: A Comprehensive Review. **Remote Sensing**, v. 15, n. 13, 1 jul. 2023. DOI: **10.3390/rs15133265**. Disponível em: <https://www.mdpi.com/2072-4292/15/13/3265>. Acesso em: 10 out. 2024

YILDIZ, M. et al. Segmentation and classification of skin burn images with artificial intelligence: Development of a mobile application. **Burns**, v. 50, n. 4, p. 966–979, maio 2024. DOI: 10.1016/j.burns.2024.01.007. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0305417924000135?via%3Dihub>. Acesso em: 29 ago. 2024.

ZAILANI, Q. N. N. et al. PLAN-DO-CHECK-ACT CYCLE: A METHOD TO IMPROVE CUSTOMER SATISFACTION AT A MUNICIPAL COUNCIL IN MALAYSIA. **International Journal of Professional Business Review**, v. 8, n. 4, 2023. DOI: 10.26668/businessreview/2023.v8i4.931. Disponível em: <https://openaccessojs.com/JBReview/article/view/931/645>. Acesso em: 01 out. 2024.

ZHU, Y. et al. Improving bowel preparation for colonoscopy with a smartphone application driven by artificial intelligence. **npj Digital Medicine**, v. 6, n. 1, 1 dez. 2023. DOI: 10.1038/s41746-023-00786-y. Disponível em: <https://www.nature.com/articles/s41746-023-00786-y>. Acesso em: 13 set. 2024.