

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS

FELIPE DE OLIVEIRA SANTOS

GABRIEL DE ANTONIO MAZETTO

GUSTAVO BARBOSA SILVA

LUCAS MAUAD SANT' ANNA

MATEUS PEREIRA ALVES

**WOUNDSCAN - CLASSIFICAÇÃO AUTOMATIZADA E SUPORTE AO
DIAGNÓSTICO DE FERIDAS CUTÂNEAS COM DEEP LEARNING**

CAMPINAS

2025

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS
CENTRO DE PROGRAMA DE GRADUAÇÃO**

FELIPE DE OLIVEIRA SANTOS

GABRIEL DE ANTONIO MAZETTO

GUSTAVO BARBOSA SILVA

LUCAS MAUAD SANT' ANNA

MATEUS PEREIRA ALVES

**WOUNDSCAN - CLASSIFICAÇÃO AUTOMATIZADA E SUPORTE AO
DIAGNÓSTICO DE FERIDAS CUTÂNEAS COM DEEP LEARNING**

Trabalho de Projetos Empreendedores A apresentado ao Corpo Docente da Escola Politécnica da Pontifícia Universidade Católica de Campinas, como parte dos requisitos necessário à obtenção do título de Bacharel em Ciência de Dados e Inteligência Artificial.

Orientador:

Prof. Me. Fernando Soares de Aguiar Neto

CAMPINAS

2025

Dedicatória:

Dedicamos este trabalho às nossas famílias e amigos, por sua compreensão e apoio incondicional. A todos que nos incentivaram nesta jornada.

AGRADECIMENTOS

Agradecemos ao nosso orientador, Prof. Me. Fernando Soares de Aguiar Neto, por todo o tempo e cuidado ao nos ajudar e orientar nesse trabalho. Expressamos nossa gratidão à equipe do Mescla, pelo valioso auxílio e mentoria. E por fim, nosso muito obrigado vai para aqueles heróis não celebrados que entenderam quando dissemos 'não posso, estou trabalhando no projeto de empreendedorismo'. Sua compreensão fez toda a diferença.

“Grandes realizações não são feitas por impulso, mas sim pela soma de esforços constantes e deliberados.”

(Adaptado de Vincent van Gogh)

RESUMO

Este trabalho de Projetos Empreendedores A apresenta o projeto WoundScan, focado no desenvolvimento de uma ferramenta digital para o acompanhamento e diagnóstico de feridas cutâneas utilizando modelos de aprendizado de máquina. Diante dos desafios impostos pelas feridas crônicas e agudas no sistema de saúde, que afetam milhões de pacientes e impõem uma significativa carga econômica, o WoundScan propõe otimizar avaliações e fornecer suporte padronizado à decisão clínica, avançando a Inteligência Artificial no setor médico. A pesquisa concentrou-se na análise exploratória das principais técnicas existentes na literatura para a classificação de imagens de feridas, com ênfase em Redes Neurais Convolucionais (CNNs) e *Transfer Learning*, e na implementação de modelos de *deep learning* para esse propósito. Utilizando um conjunto de dados inicial de 879 imagens classificadas como "Normal", "*Diabetic Wounds*", "*Pressure Wounds*" e "*Venous Wounds*", diferentes arquiteturas de modelos foram exploradas. Os resultados demonstram a eficácia da abordagem em estágios: o modelo binário (Normal vs. Ferida) atingiu 100% de acurácia, e o ResNet50 sem *data augmentation* obteve 85,26% para a classificação multi-classe das classes de feridas. A aplicação de um filtro de confiança a este modelo elevou a acurácia para 92,31% em amostras de alta certeza. A exploração do Derm Foundation Model também apresentou resultados competitivos. A análise qualitativa indicou que feridas venosas são geralmente mais fáceis de classificar, enquanto feridas diabéticas são as mais desafiadoras. Este projeto, portanto, reforça o potencial transformador da Inteligência Artificial no aprimoramento do diagnóstico e manejo de feridas, oferecendo uma contribuição significativa para o avanço da medicina digital.

Palavras-chave: Classificação de feridas, Inteligência Artificial, Redes Neurais Convolucionais, *Deep Learning*, *Transfer Learning*.

ABSTRACT

This Entrepreneurial Projects A final paper presents the WoundScan project, focused on developing a digital tool for monitoring and diagnosing skin wounds using machine learning models. Facing the challenges posed by chronic and acute wounds in the healthcare system, which affect millions of patients and incur significant economic costs, WoundScan proposes to optimize evaluations and provide standardized clinical decision support, advancing Artificial Intelligence in the medical sector. The research concentrated on the exploratory analysis of existing techniques in the literature for wound image classification, with emphasis on Convolutional Neural Networks (CNNs) and Transfer Learning, and on the implementation of deep learning models for this purpose. Using an initial dataset of 879 wound images classified as "Normal", "Diabetic Wounds", "Pressure Wounds", and "Venous Wounds", different model architectures were explored. Results demonstrate the effectiveness of the staged approach: the binary model (Normal vs. Wound) achieved 100% accuracy, and ResNet50 without data augmentation achieved 85.26% for multiclass classification of wound classes. Applying a confidence filter to this model increased accuracy to 92.31% on high-confidence samples. The exploration of the Derm Foundation Model also yielded competitive results. Qualitative analysis indicated that venous wounds are generally easier to classify, while diabetic wounds are the most challenging. This project, therefore, reinforces the transformative potential of Artificial Intelligence in improving wound diagnosis and management, offering a significant contribution to the advancement of digital medicine.

Keywords: Wound classification, Artificial Intelligence, Convolutional Neural Networks, Deep Learning, Transfer Learning.

Sumário

1. INTRODUÇÃO.....	11
2. PROPOSTA DE VALOR E APLICABILIDADE CLÍNICA	13
3. CONCEITOS FUNDAMENTAIS.....	14
3.1. O Desafio das Feridas Cutâneas no Cenário da Saúde.....	14
3.2. Inteligência Artificial e Visão Computacional para Análise de Imagens.....	15
3.3. Redes Neurais Convolucionais	15
3.4. Transfer Learning	16
3.5. Arquiteturas de Modelos Pré-treinados	17
3.5.1. ResNet50.....	17
3.5.2. Derm Foundation Model	17
3.6. Data Augmentation.....	18
4. ABORDAGENS EXISTENTES PARA CLASSIFICAÇÃO DE FERIDAS	20
5. METODOLOGIA	22
5.1. Métricas de Avaliação	22
5.2. Coleta e Curadoria do Conjunto de Dados.....	22
5.3. Divisão do Dataset para Experimentação	23
5.4. Estratégias de Modelagem e Treinamento	24
5.4.1 Modelo de Classificação Binária (Normal vs. Ferida)	24
5.4.2. Modelo de Classificação Multiclasse de Ferida	25
5.4.3. Aplicação de Data Augmentation Dinâmica (“on-the-fly”)	25
5.4.4. Exploração com Derm Foundation Model.....	26
6. RESULTADOS E DISCUSSÃO.....	27
6.1. Desempenho dos Modelos.....	27
6.1.1. Classificação Binária (Normal vs. Ferida).....	27
6.1.2. Classificação Multiclasse de Feridas (Resnet50).....	27

6.1.3. Classificação com Derm Foundation Model.....	29
6.2. Impacto do Filtro de Confiança	32
6.3. Comparativo de Desempenho entre os Modelos.....	34
6.4. Comparação com Outros Estudos.....	34
7. CONCLUSÃO	36
8. REFERÊNCIAS	38

1. INTRODUÇÃO

O avanço tecnológico tem transformado diversos setores, e a área da saúde se destaca como um campo promissor para a aplicação de soluções inovadoras. Feridas crônicas e agudas representam um desafio crescente para o sistema de saúde global, impactando anualmente mais de 13 milhões de pacientes, número que continua a crescer devido ao envelhecimento da população (Abubakar *et al.*, 2014). Segundo um estudo publicado na *Journal of Medical Economics*, que analisou dados do Medicare, revelou que os gastos com o tratamento foram de US\$ 22,5 bilhões em 2019 (Carter *et al.* 2023). O diagnóstico preciso e rápido dessas feridas é uma tarefa complexa, frequentemente marcada pela falta de padronização e pela subjetividade inerente à avaliação clínica tradicional (Romanelli *et al.*, 2011). Essa problemática é tão relevante que instituições de saúde e universidades estão ativamente buscando soluções; um exemplo notável é o evento "Motiv.se", um dos principais eventos de inovação da PUC-Campinas, em parceria com a Unimed Campinas, que em sua 11ª edição (2024) abordou o uso de tecnologias de Inteligência Artificial (IA) para solucionar desafios na saúde, incluindo a "Predição e Classificação de Feridas com Assistência de Prescrição", ressaltando a importância do manejo eficaz e reconhecimento do perfil da ferida para a recuperação do paciente¹.

Diante desse cenário, o projeto WoundScan propõe o desenvolvimento de uma ferramenta digital inovadora para o acompanhamento e diagnóstico de feridas na pele, utilizando modelos de aprendizado de máquina. O objetivo é otimizar avaliações, fornecer suporte padronizado à decisão clínica e contribuir para o avanço da Inteligência Artificial no setor médico, cuja expectativa é atingir um valor de mercado de aproximadamente US\$ 188 bilhões até 2030 (Grand View Research, 2022). A pesquisa concentra-se na análise exploratória das principais técnicas existentes na literatura para a classificação de imagens de feridas, com ênfase em Redes Neurais Convolucionais (CNNs) e *Transfer Learning*. Para esse propósito, foram desenvolvidos modelos de *deep learning*, explorando diferentes arquiteturas para identificar as abordagens mais eficazes em termos de

¹ PUC-CAMPINAS. 11ª edição do Motiv.se abre inscrições e terá foco em soluções de saúde em parceria com a Unimed. 2024. Disponível em: <https://www.puc-campinas.edu.br/11a-edicao-do-motivse-abre-inscricoes-e-tera-foco-em-solucoes-de-saude-em-parceria-com-a-unimed/>. Acesso em: 28 maio 2025.

precisão e aplicabilidade clínica em um conjunto de dados inicial de 879 imagens classificadas em "Normal", "*Diabetic Wounds*", "*Pressure Wounds*" e "*Venous Wounds*". O objetivo é padronizar e aumentar a precisão na classificação automatizada de feridas, mitigando erros clínicos e aprimorando a tomada de decisão médica. O desenvolvimento foi realizado utilizando a linguagem Python e o *framework* PyTorch, e os códigos-fonte estão disponíveis publicamente no GitHub².

Este documento está estruturado em capítulos que detalham o desenvolvimento do WoundScan. O Capítulo 2, "PROPOSTA DE VALOR E APLICABILIDADE CLÍNICA", explora o potencial de impacto do WoundScan e sua aplicabilidade prática. O Capítulo 3, "CONCEITOS FUNDAMENTAIS", apresenta os fundamentos teóricos necessários para a compreensão do projeto. O Capítulo 4, "ABORDAGENS EXISTENTES PARA CLASSIFICAÇÃO DE FERIDAS", discute outras técnicas e modelos relevantes na literatura. O Capítulo 5, "METODOLOGIA", detalha as métricas de avaliação, as etapas de coleta e preparação de dados, e as estratégias de modelagem e treinamento utilizadas. O Capítulo 6, "RESULTADOS E DISCUSSÃO", apresenta o desempenho dos modelos desenvolvidos e uma análise comparativa. Finalmente, o Capítulo 7, "CONCLUSÃO", resume os principais achados, incluindo os próximos passos da pesquisa.

² <https://github.com/WoundScan/WoundScan>

2. PROPOSTA DE VALOR E APLICABILIDADE CLÍNICA

O WoundScan se propõe a ser uma ferramenta digital de suporte à decisão clínica. A ideia central é que um profissional de saúde, utilizando a câmera de um smartphone, possa enviar uma imagem da ferida do paciente para o software. O sistema, então, processará a imagem e apresentará sugestões de diagnósticos (*Diabetic Wounds*, *Pressure Wounds*, *Venous Wounds*), ranqueadas de acordo com a confiança do modelo.

Além do diagnóstico inicial, o WoundScan será um sistema de acompanhamento longitudinal. Se um paciente realizar múltiplas consultas e avaliações, as sugestões do modelo para cada visita seriam salvas no prontuário digital. O médico teria a flexibilidade de adicionar comentários, observações e quaisquer informações clínicas relevantes sobre o caso específico do paciente em cada registro.

A implementação do WoundScan em hospitais e clínicas oferece múltiplos benefícios. Isso inclui a padronização do diagnóstico, que reduz a subjetividade das avaliações visuais, proporcionando um diagnóstico mais consistente entre profissionais. A ferramenta fornece suporte à decisão clínica, oferecendo aos médicos uma segunda opinião baseada em IA para decisões mais informadas e rápidas. O sistema permite o acompanhamento da evolução da ferida, registrando classificações e observações para um monitoramento preciso ao longo do tempo. Consequentemente, há uma otimização de recursos, pois um diagnóstico mais rápido e preciso pode levar a tratamentos mais eficazes, reduzindo tempo de internação e custos.

3. CONCEITOS FUNDAMENTAIS

Este capítulo aborda os conceitos essenciais para a compreensão do projeto WoundScan, desde a contextualização do problema das feridas cutâneas até as tecnologias de Inteligência Artificial e aprendizado de máquina empregadas.

3.1. O Desafio das Feridas Cutâneas no Cenário da Saúde

Feridas na pele representam uma preocupação significativa na área da saúde (Romanelli *et al.*, 2011). O manejo eficaz dessas condições é crucial para a recuperação dos pacientes e para a prevenção de complicações. A avaliação clínica tradicional, muitas vezes subjetiva e sem padronização, pode levar a variações no diagnóstico e na abordagem terapêutica (Romanelli *et al.*, 2011). Ferramentas digitais baseadas em Inteligência Artificial surgem como uma solução promissora, oferecendo suporte objetivo e preciso.

A categorização em Normal, *Diabetic Wounds*, *Pressure Wounds* e *Venous Wounds* é fundamental para o direcionamento das intervenções.

Este trabalho aborda três categorias principais de feridas cutâneas: diabéticas (neuropáticas), de pressão e venosas. Cada tipo apresenta causas e características clínicas distintas, descritas brevemente a seguir.

Feridas Diabéticas (Neuropáticas) ocorrem principalmente nos pés de pacientes diabéticos devido à neuropatia periférica e à má circulação, resultantes da hiperglicemia crônica. São geralmente indolores, com bordas bem definidas e profundidade variável, podendo atingir tecidos profundos e ossos (SACCO *et al.*, 2023). Uma complicação frequente é a infecção, podendo evoluir para osteomielite e levar a amputações (FERREIRA *et al.*, 2020).

Feridas de Pressão (Úlceras por Pressão) são lesões causadas por pressão prolongada sobre a pele, especialmente em pacientes acamados ou com mobilidade reduzida. Surgem sobre proeminências ósseas, como sacro e calcanhares. São classificadas em quatro estágios (I a IV), variando de vermelhidão superficial até exposição de músculos e ossos (NPIAP, 2019).

Feridas Venosas (Úlceras Venosas) resultam da insuficiência venosa crônica, geralmente associada a varizes ou histórico de trombose venosa profunda. São típicas na região dos tornozelos, superficiais, dolorosas, com formato irregular e grande quantidade de secreção. Frequentemente apresentam pele ao redor escurecida (hiperpigmentação) e inchaço (GODOY, 2021). As complicações incluem infecções recorrentes e cicatrização lenta com alta taxa de reincidência (SBACV, 2022).

Esses tipos de feridas exigem atenção clínica especializada devido ao impacto significativo na qualidade de vida dos pacientes, além da alta prevalência e custos associados ao tratamento (BRASIL, 2002).

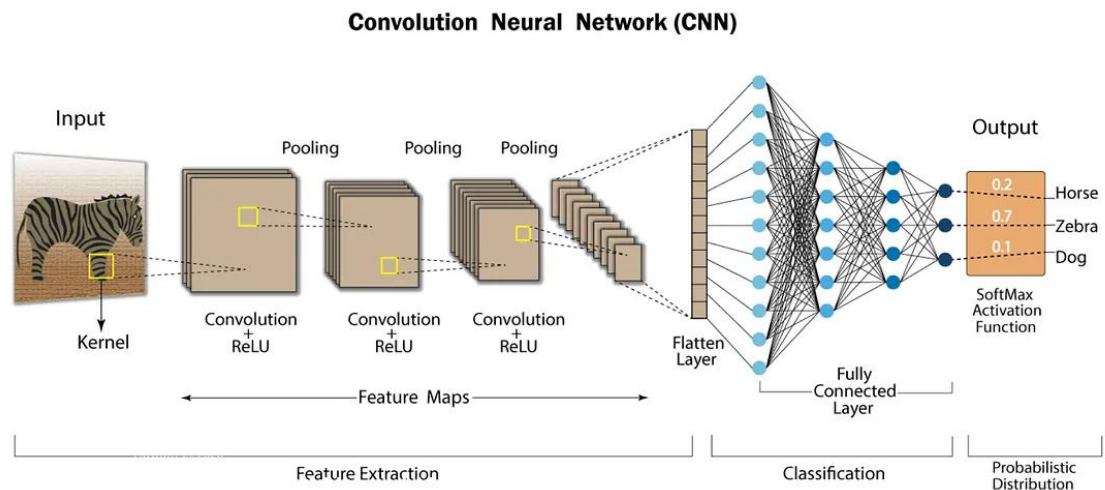
3.2. Inteligência Artificial e Visão Computacional para Análise de Imagens

A Inteligência Artificial (IA) tem se mostrado uma ferramenta poderosa para otimizar avaliações e fornecer suporte padronizado à decisão clínica, com um mercado em ascensão no setor médico (Kiraly *et al.*, 2024). Dentro da IA, a Visão Computacional é um campo que permite aos computadores "enxergarem" e interpretar imagens digitais, sendo fundamental para a classificação automatizada de feridas.

3.3. Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNNs) são uma classe de redes neurais artificiais amplamente utilizadas em tarefas de visão computacional, inspiradas no sistema visual biológico. Sua arquitetura é projetada para processar dados com uma topologia de grade, como imagens. A Figura 1 apresenta a arquitetura da rede neural convolucional.

Figura 1: Arquitetura da CNN



Fonte: SHAHRIAR, N. (2023)

As CNNs são compostas por camadas convolucionais, de *pooling* e totalmente conectadas, que permitem a extração automática de características hierárquicas das imagens, desde bordas e texturas até formas complexas e objetos completos. Essa capacidade de aprendizado de características torna as CNNs extremamente eficazes para tarefas como classificação de imagens, detecção de objetos e segmentação de imagens (Krizhevsky, Sutskever & Hinton, 2012). A exploração dessas técnicas de *deep learning*, com ênfase em CNNs, foi um dos focos principais desta pesquisa.

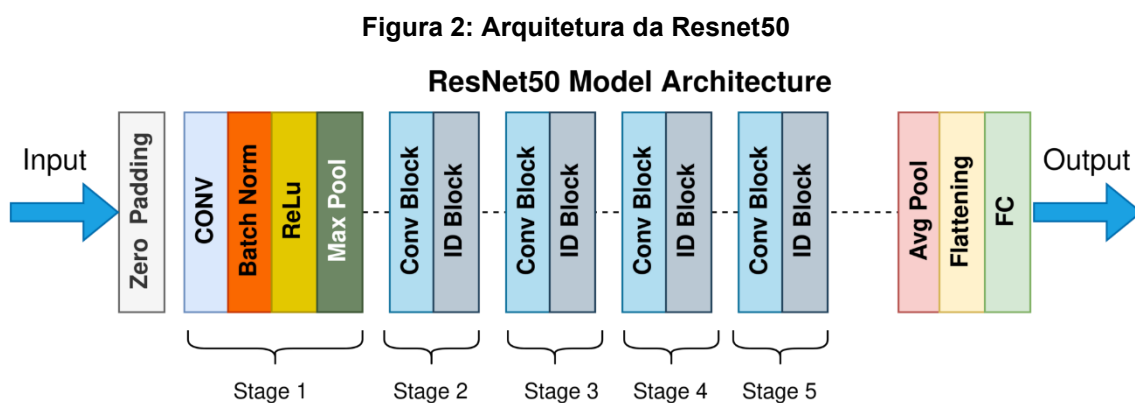
3.4. Transfer Learning

Transfer Learning é uma técnica de aprendizado de máquina onde um modelo pré-treinado em uma tarefa específica (geralmente com um grande conjunto de dados) é reutilizado como ponto de partida para uma nova tarefa relacionada. Em visão computacional, é comum utilizar modelos treinados em grandes *datasets* como ImageNet. A ideia é que o modelo já tenha aprendido características genéricas de imagens (bordas, texturas, padrões) que podem ser transferidas e adaptadas para um novo domínio com menos dados, como a classificação de feridas. Isso reduz significativamente o tempo e os recursos computacionais necessários para o treinamento, além de melhorar o desempenho em conjuntos de dados menores (Krizhevsky, Sutskever & Hinton, 2012). A aplicação de *Transfer Learning* foi fundamental nesta pesquisa.

3.5. Arquiteturas de Modelos Pré-treinados

3.5.1. ResNet50

ResNet50 é uma arquitetura de rede neural convolucional profunda que faz parte da família ResNet (*Residual Networks*). A Figura 2 ilustra a arquitetura da ResNet50.



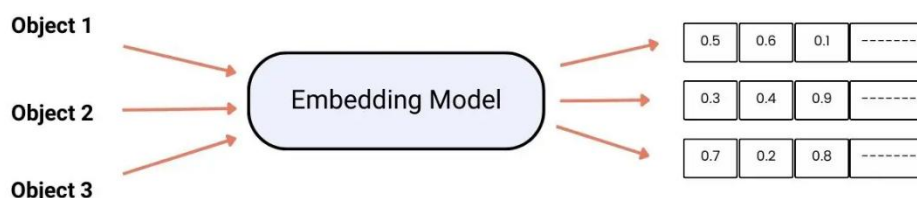
Fonte: MUKHERJEE, S. (2022)

Ela é conhecida por suas "conexões residuais" (ou *skip connections*), que permitem que o gradiente flua diretamente através de várias camadas, resolvendo o problema de *vanishing gradient* que afeta redes muito profundas. A ResNet50 possui 50 camadas e é amplamente utilizada em *Transfer Learning* devido à sua robustez e ao seu alto desempenho em tarefas de classificação de imagens (He *et al.*, 2015). Sua capacidade de aprender representações complexas a torna uma escolha valiosa para a análise de imagens médicas.

3.5.2. Derm Foundation Model

O Derm Foundation Model é um modelo de base (*foundation model*) desenvolvido especificamente para o domínio da dermatologia pelo Google (Kiraly *et al.*, 2024). Seu propósito é acelerar o desenvolvimento de IA para análise de imagens de pele em aplicações dermatológicas. A Figura 3 representa o processo de geração de embeddings.

Figura 3: Representação do Processo de Geração de Embeddings



Fonte: DATA SCIENCE DOJO (2024)

Ele é pré-treinado em grandes quantidades de imagens de pele rotuladas para produzir *embeddings* de 6144 dimensões, que são representações numéricas que capturam as características essenciais das imagens. Como resultado, os *embeddings* do Derm Foundation Model permitem o treinamento eficiente de modelos de IA com significativamente menos dados e computação do que os métodos tradicionais. Esse modelo foi treinado com um grande número de pares de imagens e textos públicos, e o componente de imagem foi então ajustado para classificação de condições e outras tarefas subsequentes, utilizando diversos conjuntos de dados clínicos. As imagens são provenientes de uma variedade de tipos de dispositivos, incluindo câmeras de smartphones, outras câmeras e dermatoscópios, e as imagens podem ter sido tiradas por médicos durante consultas ou auto capturadas pelos pacientes. Essa vasta diversidade de dados e fontes de aquisição contribui para que o Derm Foundation Model seja robusto a variações do ambiente real de uso, tornando-o promissor para aplicações práticas.

3.6. Data Augmentation

Data Augmentation é um conjunto de técnicas utilizadas para aumentar artificialmente o volume e a diversidade de um conjunto de dados de treinamento. Isso é feito gerando novas amostras de dados a partir das existentes

através de transformações como rotação, espelhamento horizontal e vertical, ajuste de contraste, brilho e saturação. Ao expandir a variabilidade dos dados de treinamento, o *Data Augmentation* ajuda os modelos a aprenderem representações mais robustas e a generalizarem melhor para dados não vistos, reduzindo o *overfitting* e melhorando a performance em cenários reais.

A utilização de *Data Augmentation* "*on-the-fly*", ou seja, em tempo real durante o treinamento, oferece vantagens significativas em comparação ao método tradicional de pré-processamento e armazenamento das imagens aumentadas. Essa abordagem elimina a necessidade de armazenar um volume massivo de imagens geradas, economizando espaço em disco. Além disso, promove uma maior variabilidade no treinamento, pois cada época pode apresentar diferentes transformações das imagens originais, evitando que o modelo "memorize" um conjunto fixo de imagens aumentadas. Isso contribui para uma melhor generalização do modelo a dados não vistos (Radecki & Wiatrak, 2022).

4. ABORDAGENS EXISTENTES PARA CLASSIFICAÇÃO DE FERIDAS

4.1. Classificadores Ensemble de CNNs

Com os avanços do *deep learning*, surgiram modelos baseados em redes neurais convolucionais (CNNs) dedicados à classificação de feridas. Um exemplo notável é o desenvolvimento de um classificador de feridas que combina duas CNNs (análise local por patches e global da imagem inteira) em um modelo *ensemble*, fundindo suas saídas por meio de um *perceptron* de múltiplas camadas para aprimorar o desempenho (Rostami *et al.*, 2020). Essa abordagem foi avaliada em feridas cirúrgicas, diabéticas e úlceras venosas, obtendo acurácia de até aproximadamente 96% em cenários binários e 92% em classificação de três categorias de feridas (Rostami *et al.*, 2020). Os resultados demonstram que modelos de CNN em *ensemble* podem servir como eficazes sistemas de apoio à decisão na classificação automática de feridas.

4.2. Abordagem Multimodal (Imagem e Localização Anatômica)

Pesquisas recentes exploram informações adicionais além da própria imagem da ferida. Foi proposta uma rede neural multimodal que integra imagens de feridas com dados de localização anatômica da lesão, melhorando a precisão da classificação (Patel *et al.*, 2024). Nessa arquitetura, características extraídas por redes convolucionais são combinadas a módulos de atenção (como Squeeze-and-Excitation e atenção axial) e a um *Multilayer Perceptron* (MLP) adaptativo para fundir os dados de imagem e localidade da ferida (Patel *et al.*, 2024). Ao categorizar quatro tipos de úlceras – diabéticas, de pressão, cirúrgicas e venosas – essa rede atingiu índices de acurácia superiores aos métodos tradicionais, variando de aproximadamente 74,8% até 100% em diferentes conjuntos de teste (Patel *et al.*, 2024). Os autores reportam que tal abordagem multimodal superou significativamente performances previamente publicadas na literatura, evidenciando o valor de incorporar metadados clínicos (como a localização do ferimento) no modelo de classificação.

4.3. Classificação Multiclasse com Modelo Multi-tarefa

Outra linha de investigação busca classificar múltiplas categorias de feridas simultaneamente. Foi desenvolvida a primeira ferramenta de *deep learning* capaz de classificar, de uma só vez, cinco condições principais de feridas: feridas profundas, infectadas, arteriais, venosas e úlceras por pressão (Huang *et al.*, 2023). O método adota um modelo de aprendizado multi-tarefa unificado que explora as relações entre essas diferentes condições para melhorar a generalização do classificador (Huang *et al.*, 2023). Notavelmente, o desempenho do modelo foi comparável ou superior ao de profissionais de saúde não especialistas em feridas, alcançando concordância (kappa de Cohen) e acurácia não inferiores às de médicos e enfermeiros na tarefa de classificação (Huang *et al.*, 2023). Esse resultado indica que o modelo proposto consegue igualar ou superar a precisão humana na identificação das cinco classes de feridas, representando um avanço importante no suporte automatizado ao diagnóstico.

4.4. Segmentação de Imagens como Pré-processamento

Uma abordagem complementar eficaz é aplicar segmentação de imagem antes da etapa de classificação, de modo a isolar a área da ferida e reduzir ruídos de fundo. Estratégias simples, como remover o fundo da fotografia para manter apenas a região lesionada, já se mostraram benéficas ao tornar a ferida o foco da análise do classificador. Trabalhos mais recentes incorporam segmentação automática ao pipeline: um estudo propôs um *framework* unificado que primeiro segmenta a ferida numa imagem (usando uma variante de U-Net com mecanismos de atenção) e depois classifica o tipo de lesão a partir das características da região segmentada (Wang *et al.*, 2023). Esse método integrado obteve cerca de 90% de acurácia na classificação multiclasse de feridas, ao mesmo tempo em que atingiu coeficientes Dice em torno de 86% na etapa de segmentação, demonstrando a eficácia de delimitar previamente a região da ferida (Wang *et al.*, 2023). Em suma, o pré-processamento por segmentação tende a melhorar a consistência da classificação ao fornecer à rede apenas os pixels relevantes da ferida, eliminando fundos complexos e outras estruturas não pertinentes.

5. METODOLOGIA

A metodologia para o desenvolvimento e avaliação do modelo de classificação de feridas no WoundScan seguiu uma abordagem sequencial, inspirada no modelo *Waterfall* (Pressman, 2016), devido à necessidade de requisitos bem definidos nas etapas iniciais. O foco principal foi a análise exploratória e a experimentação com diferentes técnicas de aprendizado de máquina e aprendizado profundo.

Para garantir que os modelos com o melhor desempenho fossem preservados ao longo do treinamento, a estratégia de *Model Checkpointing* foi aplicada em todos os experimentos. Essa técnica consiste em monitorar uma métrica de desempenho (neste caso, a acurácia no conjunto de validação) e salvar o modelo apenas quando essa métrica atinge um novo melhor valor. Isso assegura que, mesmo após múltiplas épocas de treinamento, o modelo com a melhor performance de validação seja o que será utilizado para as avaliações finais.

5.1. Métricas de Avaliação

A avaliação dos modelos foi focada principalmente na acurácia, considerando sua importância em diagnósticos médicos. Além da acurácia geral, foram utilizadas as seguintes métricas para uma avaliação mais abrangente, conforme observado em outros artigos na área (Rostami *et al.*, 2020): **Acurácia (Accuracy)**, **Precisão (Precision)**, **Recall (Sensibilidade)** e **F1-score**. No contexto médico, a sensibilidade é uma métrica de particular relevância. Um alto valor de sensibilidade significa que o modelo é capaz de identificar corretamente a maioria dos casos positivos (verdadeiros positivos), minimizando os falsos negativos. Isso é crucial em triagens e diagnósticos precoces de doenças, onde a falha em detectar a condição pode ter consequências graves para o paciente, como ressaltado em estudos sobre biomarcadores para o diagnóstico de doenças (Ferreira & Patino, 2017).

5.2. Coleta e Curadoria do Conjunto de Dados

O conjunto de dados inicial utilizado neste projeto, denominado '*Collected and Categorized Wound Images Dataset*', foi obtido da plataforma Kaggle (ibrahimfateen, 2024), originalmente composto por 2940 imagens distribuídas em 10

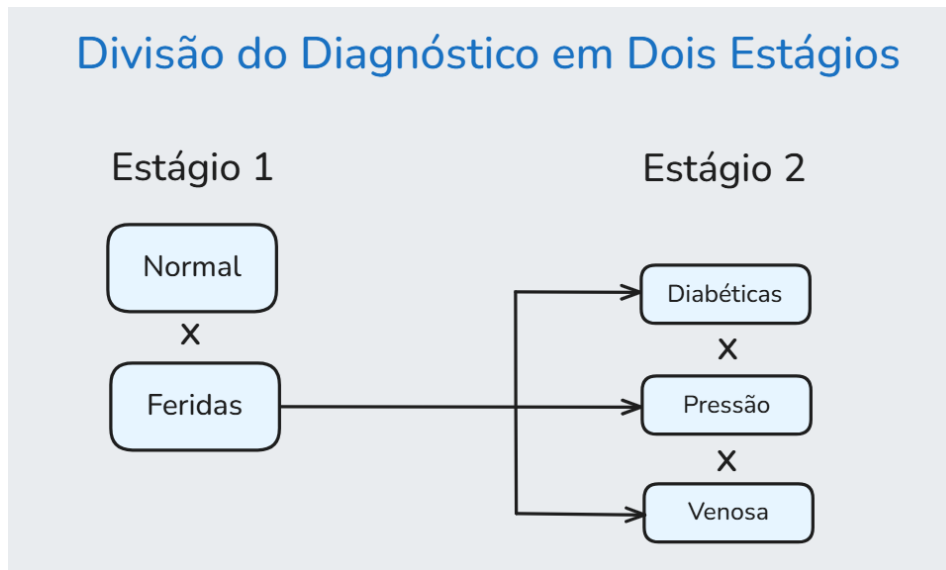
classes distintas. Para focar a pesquisa nas tipologias de feridas mais relevantes e recorrentes em estudos da área da saúde, mantivemos apenas quatro classes: Normal, Feridas Diabéticas, Feridas por Pressão e Feridas Venosas. Essa curadoria resultou em um conjunto refinado de 1758 imagens. Esses tipos específicos de feridas foram selecionados por serem frequentemente abordados em pesquisas médicas sobre classificação de feridas, dada sua significância clínica e prevalência (Huang *et al.*, 2023; Rostami *et al.*, 2020).

Foi identificado que o *dataset* original já incluía *data augmentation* pré-aplicado, especificamente a transformação de *flip* horizontal. Para prevenir o *Data Leaking* (vazamento de dados) entre os conjuntos de treinamento e teste, uma prática crucial para garantir a validade dos resultados do modelo, procedeu-se à remoção das imagens duplicadas geradas por essa técnica. Com essa filtragem, o conjunto de dados foi reduzido para 879 imagens únicas e originais. A distribuição dessas imagens entre as classes selecionadas é a seguinte: 301 imagens de *Pressure Wounds*, 247 de *Venous Wounds*, 231 de *Diabetic Wounds* e 100 de Normal.

5.3. Divisão do Dataset para Experimentação

Para o processo de treinamento e validação dos modelos, o conjunto de dados foi dividido de forma estratificada. Isso garante que a proporção de imagens de cada classe seja mantida consistentemente em todas as partições (treinamento, validação e teste). Os testes dos modelos ocorreram em dois estágios. A Figura 4 demonstra o processo de divisão em estágios.

Figura 4: Processo de divisão em estágios



Fonte: Autoria própria (2025)

O primeiro estágio focou na tarefa de menor complexidade: Classificação Binária (Normal vs. Ferida), distinguindo entre imagens de pele "Normal" e "Ferida" (agrupando todas as categorias de feridas). O conjunto de dados foi dividido da seguinte forma: 562 imagens para treinamento, 141 para validação e 176 para teste. A expectativa é que, ao resolver essa distinção fundamental primeiro, os modelos subsequentes pudessem concentrar-se nas características mais sutis que diferenciam os tipos específicos de feridas. A decisão de separar a classificação binária da multiclasse é embasada na literatura, que aponta que a redução do número de classes pode simplificar o problema e melhorar o desempenho do modelo em cada tarefa específica (e.g., Thomsen *et al.*, 2020).

Para experimentos subsequentes, o conjunto de dados foi redividido, contendo apenas as imagens classificadas como "Ferida" no Estágio 1, abordando a tarefa de maior complexidade: Classificação Multiclasse (Tipos de Ferida). A divisão foi: 498 imagens para treinamento, 125 para validação e 156 para teste.

5.4. Estratégias de Modelagem e Treinamento

5.4.1 Modelo de Classificação Binária (Normal vs. Ferida)

Conforme a estratégia de divisão por estágios, o primeiro modelo desenvolvido focou na tarefa de classificação binária. O objetivo era distinguir entre

imagens de "Normal" e "Ferida" (englobando todas as categorias de feridas analisadas). Para essa tarefa, a arquitetura ResNet50 foi empregada como base, utilizando o conjunto de dados do Estágio 1. O treinamento foi realizado em um número limitado de épocas, especificamente 3 épocas, visando uma rápida validação da capacidade do modelo em realizar essa distinção fundamental. Dado o excelente desempenho do modelo para esta classificação binária (100% de acurácia), não foram exploradas arquiteturas adicionais para esta etapa, focando os esforços nas classificações multiclasse mais complexas. Similarmente, Ros-tami *et al.* (2020) também demonstraram facilidade na classificação da classe normal, obtendo alta acurácia em seus cenários de avaliação.

5.4.2. Modelo de Classificação Multiclasse de Ferida

Seguindo a metodologia de divisão por estágios, após a validação do modelo binário, o foco foi direcionado à tarefa de classificação multiclasse das três categorias específicas de feridas. Para este desafio mais complexo, a arquitetura ResNet50 foi novamente utilizada, mas com um treinamento mais extensivo. O modelo foi treinado por 20 épocas, utilizando o conjunto de dados do Estágio 2, que contém apenas as imagens de feridas. A escolha da ResNet50 é justificada por sua comprovada eficácia em tarefas de visão computacional e pela capacidade de aproveitar o *Transfer Learning*, o que contribui para um treinamento mais eficiente e resultados promissores.

5.4.3. Aplicação de Data Augmentation Dinâmica ("on-the-fly")

Adicionalmente ao pré-processamento, foram aplicadas técnicas de *Data Augmentation "on-the-fly"* (em tempo real durante o treinamento) para aumentar artificialmente a diversidade e o volume do conjunto de dados de treinamento. As transformações utilizadas incluíram *Random Horizontal Flip*, *ColorJitter* (ajustes de brilho, contraste e saturação), e *Gaussian Blur*. Essas técnicas visam expor o modelo a uma maior variabilidade de imagens, tornando-o mais robusto e capaz de generalizar melhor para dados não vistos em ambientes reais, reduzindo o *overfitting* (sobreajuste).

5.4.4. Exploração com Derm Foundation Model

Uma abordagem alternativa na metodologia envolveu o uso do Derm Foundation Model. Para esta etapa, os *embeddings* de alta dimensão (6144 dimensões) gerados por este modelo foram pré-computados para todas as imagens antes de serem passados para os classificadores subsequentes, otimizando o tempo computacional do treinamento. Devido ao tempo computacional de pré-processamento dos *embeddings* do Derm Foundation Model, foi possível treinar esses modelos com 50 épocas. Sobre esses *embeddings*, foram testadas três arquiteturas de classificação distintas.

A primeira abordagem consistiu na adição de uma única camada linear de saída aos *embeddings* do Derm Foundation Model, seguida de uma função de ativação *Softmax* para a classificação multiclasse.

Como uma alternativa mais complexa, foi implementado uma MLP composto por duas camadas ocultas, cada uma com 256 neurônios. Para promover a regularização e reduzir o *overfitting*, foi aplicado o valor padrão de *dropout* (0.25) após cada camada oculta, seguido de uma camada de saída para a classificação final.

Além das abordagens baseadas em redes neurais, uma *Support Vector Machine* (SVM) também foi testada para a classificação das feridas, utilizando os *embeddings* gerados pelo Derm Foundation Model como entrada. Para otimizar o desempenho do SVM, foi realizada uma busca em grade (*GridSearch*) para encontrar o melhor parâmetro de regularização C, testando valores como 0.1, 1, 10 e 100 no conjunto de validação. O objetivo foi identificar o valor de C que maximizasse a acurácia do classificador. A utilização de SVM sobre *embeddings* de modelos de *deep learning* é uma abordagem comum em tarefas de classificação de imagens na área médica, como observado em estudos (Abubakar *et al.*, 2019).

6. RESULTADOS E DISCUSSÃO

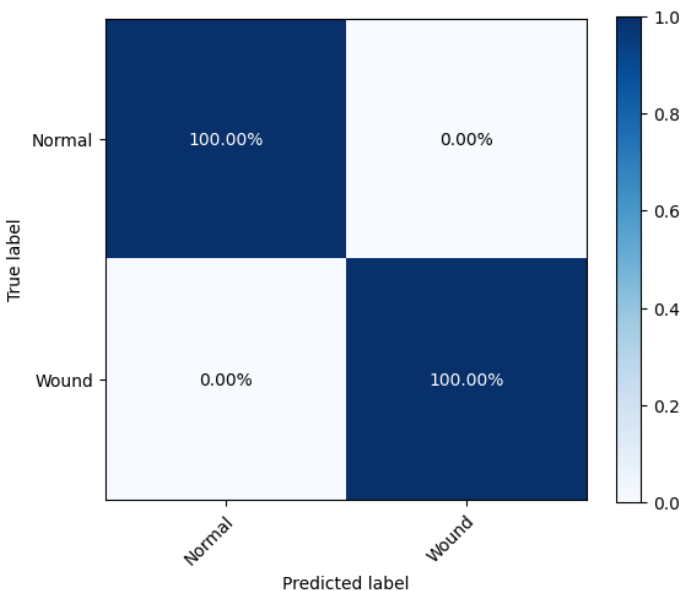
Este capítulo apresenta os resultados obtidos com os diferentes modelos de classificação desenvolvidos no projeto WoundScan, acompanhados de uma discussão sobre seu desempenho e implicações.

6.1. Desempenho dos Modelos

6.1.1. Classificação Binária (Normal vs. Ferida)

O modelo de classificação binária, treinado com a arquitetura ResNet50 no Estágio 1 do *dataset*, demonstrou um desempenho excelente. A Figura 5 apresenta a matriz de confusão no cenário binário de classificação entre normal e ferida.

Figura 5: Matriz de Confusão - Resnet50 (Normal vs. Ferida)



Fonte: Autoria própria (2025)

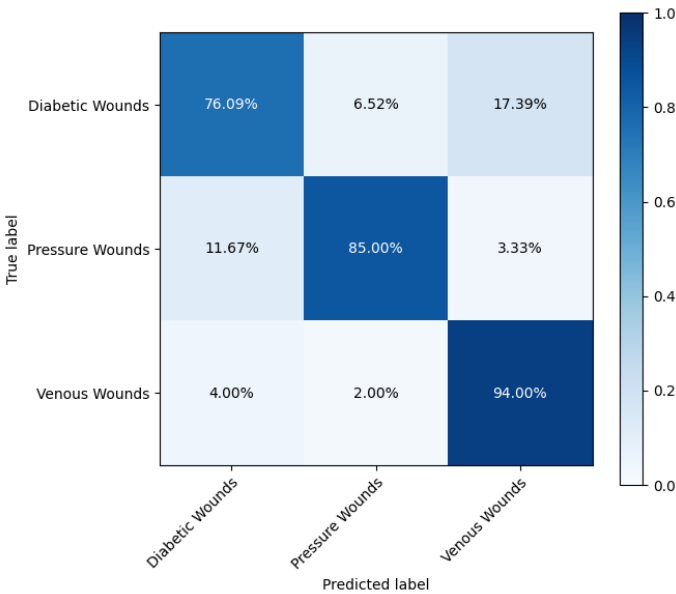
Este modelo atingiu 100% de acurácia tanto na validação quanto no teste, treinando com apenas 3 épocas. Este resultado indica que a distinção entre pele normal e ferida é uma tarefa bem definida e facilmente aprendida pela arquitetura ResNet50.

6.1.2. Classificação Multiclasse de Feridas (Resnet50)

O modelo de classificação multiclasse das três categorias de feridas (*Diabetic Wounds*, *Pressure Wounds*, *Venous Wounds*), utilizando ResNet50 com

20 épocas de treinamento e sem *data augmentation* dinâmica, obteve os seguintes resultados. A Figura 6 apresenta a matriz de confusão gerada pelo modelo ResNet50 no cenário de classificação multiclasse sem utilização de *data augmentation*.

Figura 6: Matriz de Confusão - Resnet50 (Multiclasse sem data aug.)

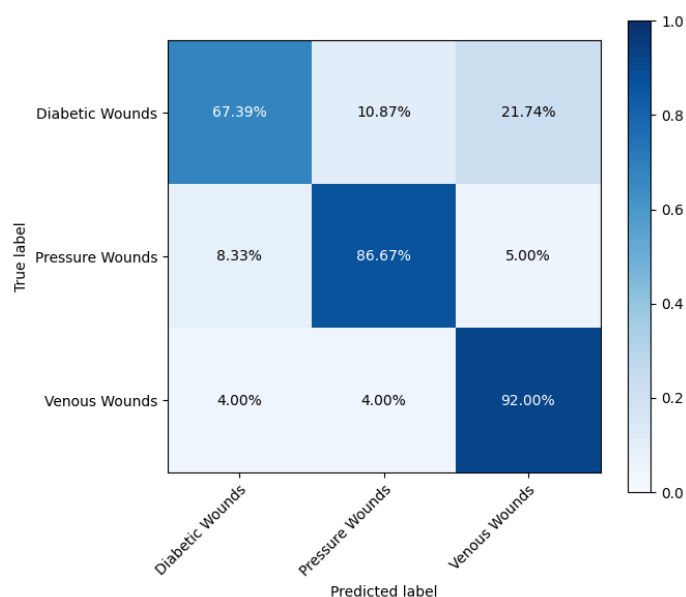


Fonte: Autoria própria (2025)

A acurácia geral do modelo Estágio 2 no teste foi de **85.26%**. Observa-se que Feridas Venosas tiveram a maior taxa de acerto (94.00%), enquanto Feridas Diabéticas tiveram a menor (76.09%), com algumas sendo erroneamente classificadas como Feridas Venosas. No entanto, ao aplicar o filtro de confiança (seção 6.3), o desempenho melhorou significativamente.

Em seguida, para investigar a efetividade do *data augmentation* dinâmico ("*on-the-fly*") neste cenário de classificação multiclasse, essa técnica foi incorporada durante o treinamento do modelo ResNet50. Os resultados detalhados são apresentados na Figura 7.

Figura 7: Matriz de Confusão - Resnet50 (Multiclasse com data aug.)



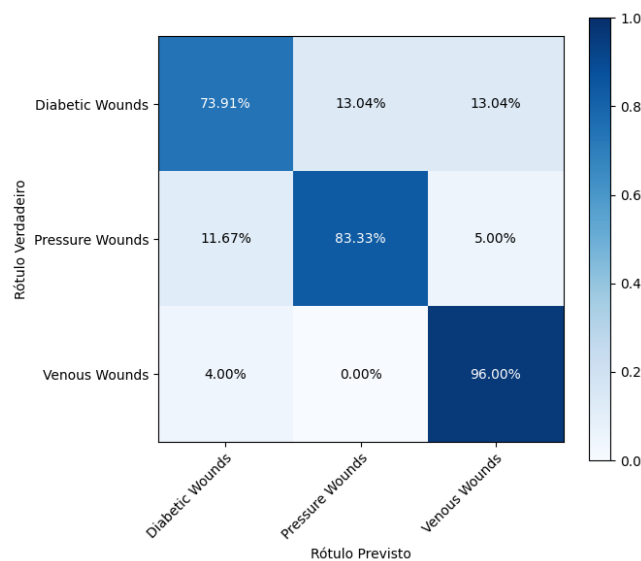
Fonte: Autoria própria (2025)

A acurácia geral para este modelo foi de **82.69%**. Comparado ao modelo sem *data augmentation*, este apresentou uma pequena melhora nas Feridas de Pressão (86.67% vs 85.00%), mas uma perda considerável de acurácia nas Feridas Diabéticas (67.39% vs 76.09%) e nas Feridas Venosas (92.00% vs 94.00%). Isso resultou em um desempenho geral pior em todas as métricas em comparação ao modelo sem *data augmentation*, sugerindo que, neste caso, o *data augmentation* não se demonstrou útil para a classificação multiclasse.

6.1.3. Classificação com Derm Foundation Model

A exploração do Derm Foundation Model, utilizando seus *embeddings* pré-computados, resultou nos seguintes desempenhos para as arquiteturas de classificação. A Figura 8 mostra a matriz de confusão obtida a partir do Derm Foundation Model com camada de saída simples.

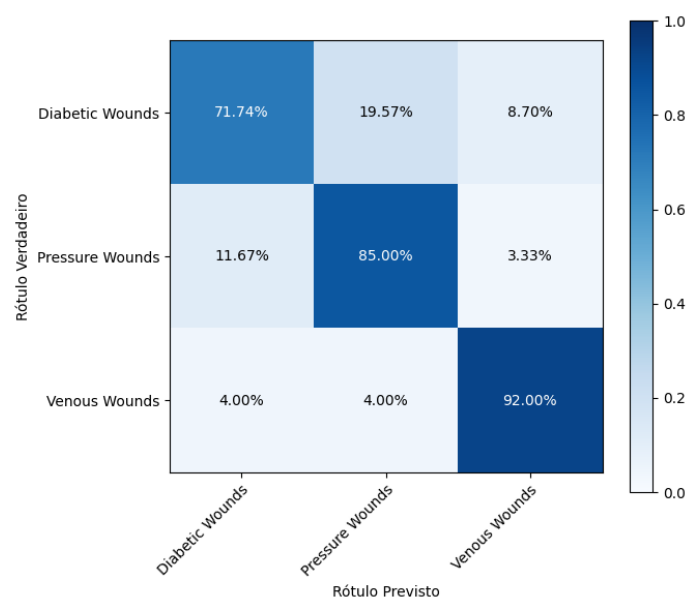
Figura 8: Matriz de Confusão - Derm Foundation Model + Camada de Saída



Fonte: Autoria própria (2025)

O modelo com camada linear e *Softmax* obteve uma acurácia geral de **84.62%**. Este modelo demonstrou um desempenho ligeiramente melhor nas feridas venosas (96.00%) em comparação ao ResNet50 sem *data augmentation* (94.00%), mas apresentou pior acurácia nas feridas diabéticas (73.91% vs 76.09%) e nas feridas de pressão (83.33% vs 85.00%). Seu desempenho geral ficou próximo do ResNet50 sem *data augmentation*, mas ligeiramente pior, como ilustra a Figura 9.

Figura 9: Matriz de Confusão - Derm Foundation Model + MLP

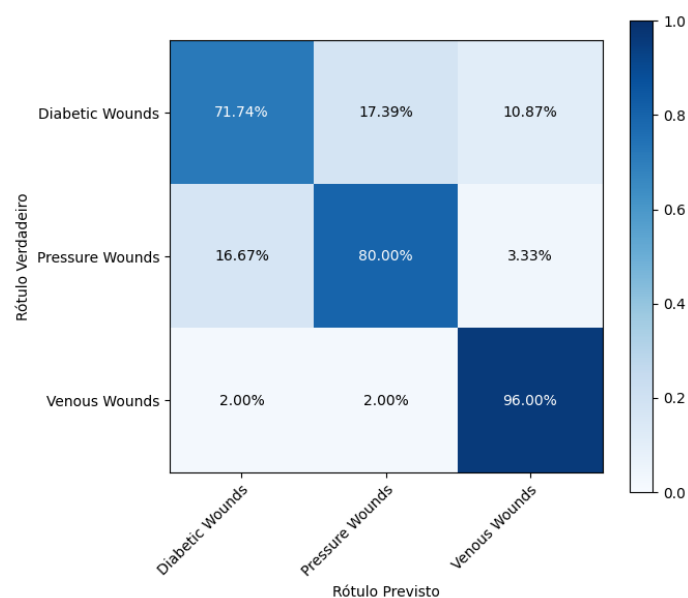


Fonte: Autoria própria (2025)

O modelo com MLP obteve uma acurácia geral de **83.33%**. Ele não demonstrou melhora em relação ao ResNet50 sem *data augmentation* e obteve métricas gerais piores do que o modelo que utilizou apenas a camada de saída linear.

A utilização de SVM sobre os *embeddings* do Derm Foundation Model, após otimização do hiperparâmetro C via *GridSearch*, apresentou os seguintes resultados. A busca pelo melhor parâmetro C no conjunto de validação indicou que **C=1** foi o que obteve o melhor resultado de acurácia no conjunto de validação. A Figura 10 mostra a matriz de confusão do Derm Foundation Model combinado com um classificador SVM, permitindo comparar o desempenho com abordagens baseadas em aprendizado profundo.

Figura 10: Matriz de Confusão - Derm Foundation Model + SVM



Fonte: Autoria própria (2025)

A acurácia geral no conjunto de teste para o melhor SVM ($C=1$) foi de **82.69%**. Comparando com o ResNet50 sem *data augmentation* (acurácia geral de 85.26%), o SVM apresentou um desempenho ligeiramente superior na classificação de Feridas Venosas (96.00% vs 94.00%), mas foi inferior nas Feridas Diabéticas (71.74% vs 76.09%) e, de forma mais acentuada, nas Feridas de Pressão (80.00% vs 85.00%), resultando em uma acurácia geral mais baixa.

Embora os modelos utilizando o Derm Foundation Model não tenham alcançado a maior acurácia absoluta neste conjunto de dados, é crucial ressaltar seu potencial. Por ter sido treinado com uma vasta gama de imagens de dermatologia, provenientes de diversos dispositivos e cenários, o Derm Foundation Model pode ser mais robusto a variações do ambiente real de uso, tornando-o promissor para aplicações práticas.

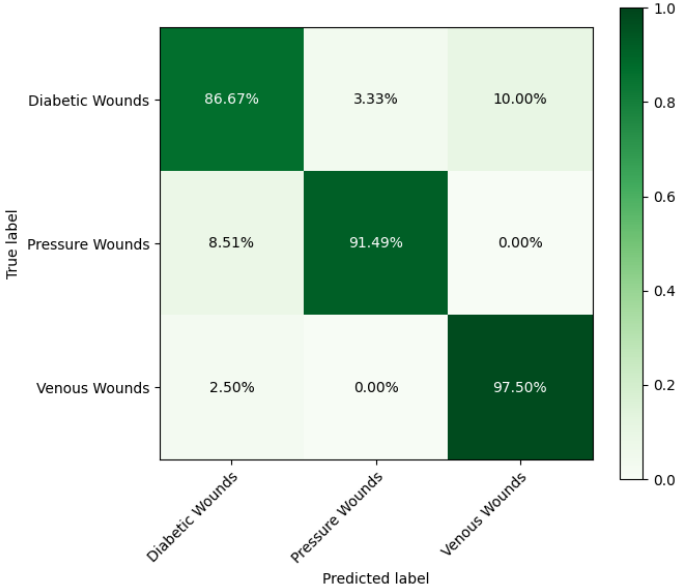
6.2. Impacto do Filtro de Confiança

Para simular um cenário clínico onde a alta confiança no diagnóstico é fundamental, um filtro de confiança foi aplicado às saídas dos modelos, aceitando apenas previsões com nível de confiança dado pelo modelo entre 95% e 100%. A análise desse filtro permitiu verificar como a restrição a previsões de

alta confiança impactava a acurácia do sistema e a proporção de amostras que poderiam ser automaticamente classificadas com um alto grau de certeza.

No modelo multiclasse ResNet50 sem *data augmentation*, que foi o modelo com melhor desempenho, o filtro de confiança demonstrou um impacto positivo. Com este filtro, a acurácia atingiu **92.31%** em 117 amostras que passaram pelo critério de confiança, representando 75% do total de amostras. A Figura 11 traz a matriz de confusão do modelo ResNet50 sem *data augmentation*, considerando apenas predições com confiança entre 95% e 100%. Essa filtragem permite avaliar a confiabilidade das predições mais seguras do modelo.

Figura 11: Matriz de Confusão - Resnet50 (Multiclasse sem data aug. Confiança [95%-100%])



Fonte: Autoria própria (2025)

Este resultado é particularmente relevante para um cenário de aplicação real. Em situações clínicas, a acurácia é de extrema importância. Ao invés de apresentar um diagnóstico com baixa confiança, o sistema pode indicar ao profissional de saúde que a sugestão do modelo não possui alta certeza, talvez devido à confusão entre classes, e que uma avaliação humana mais aprofundada é necessária. Essa abordagem pode aumentar a confiança dos médicos na ferramenta, utilizando a IA como um suporte à decisão e não como um substituto.

6.3. Comparativo de Desempenho entre os Modelos

Para facilitar a comparação entre as diferentes abordagens testadas, a Tabela 1 resume as acurácias gerais e métricas ponderadas dos modelos de classificação multiclasse.

Tabela 1. Comparativo de Desempenho dos Modelos (com classificação multiclasse)

Modelo	Acurácia Geral (%)	Precisão (WA)	Recall (WA)	F1-Score (WA)	Acurácia com Filtro (95-100%)	Amostras com Filtro
ResNet50 (sem <i>Data Augmentation</i>)	85.26	0.8555	0.8526	0.8521	92.31	117
ResNet50 (com <i>Data Augmentation</i>)	82.69	0.8294	0.8269	0.8243	-	-
Derm Foundation Model + Camada Linear	84.62	0.8483	0.8462	0.8450	-	-
Derm Foundation Model + MLP	83.33	0.8316	0.8333	0.8318	-	-
Derm Foundation Model + SVM (C=1)	82.69	0.8248	0.8269	0.8249	-	-

O modelo ResNet50 sem *data augmentation* alcançou o melhor desempenho em todas as métricas avaliadas, incluindo a maior acurácia geral (85.26%). Além do mais, ao considerar o impacto do filtro de confiança, esse modelo demonstrou um aumento significativo na acurácia para 92.31% em 117 amostras, que representam 75% das amostras que passaram pelo filtro. Isso sugere que a aplicação do filtro de confiança pode ser uma estratégia valiosa para aumentar a acurácia em cenários práticos.

6.4. Comparação com Outros Estudos

A acurácia alcançada pelo nosso melhor modelo (ResNet50 sem *data augmentation*, com 85.26%) e, especialmente, a acurácia de 92.31% nas amostras de alta confiança do mesmo modelo, são resultados promissores no campo da classificação automatizada de feridas. Para contextualizar esses resultados, é importante compará-los com o que tem sido reportado em outros estudos na literatura. Por exemplo, Rostami *et al.* (2020) em seus experimentos de classificação de 3 classes, utilizando o conjunto de dados AZH (Anisuzzaman *et al.*

2022), obtiveram uma acurácia máxima de 91.9% e uma acurácia média de 87.7%. Já Patel *et al.* (2024) obtiveram uma acurácia de 87.30% para a classificação de 3 classes (diabética, pressão e venosa) utilizando imagens inteiras. As abordagens detalhadas utilizadas por Rostami *et al.* (2020) e Patel *et al.* (2024), que envolvem a combinação de classificadores e a análise de imagens e localização, respectivamente, já foram exploradas no Capítulo 4, "Abordagens Existentes para Classificação de Feridas".

Durante a análise dos resultados do nosso estudo, notou-se que, de maneira geral, as feridas venosas tendiam a ser mais simples de serem classificadas pelos modelos desenvolvidos, frequentemente apresentando as maiores acurácias em trabalhos relacionados (Huang *et al.*, 2023; Rostami *et al.*, 2020). Em contraste, as feridas diabéticas demonstraram ser as mais desafiadoras para a classificação precisa em nosso estudo.

Nossa pesquisa demonstra que abordagens baseadas em *deep learning* e *transfer learning*, combinadas com estratégias de filtro de confiança, podem oferecer desempenho competitivo e relevante para aplicações clínicas, destacando a importância de considerar não apenas a acurácia geral, mas também a confiança das previsões em cenários de alta sensibilidade como o diagnóstico médico.

7. CONCLUSÃO

O projeto WoundScan abordou o complexo desafio do diagnóstico e acompanhamento de feridas cutâneas, visando mitigar a subjetividade e a falta de padronização na avaliação clínica tradicional. Por meio da aplicação de modelos de *deep learning*, com ênfase em Redes Neurais Convolucionais (CNNs) e *Transfer Learning*, este trabalho demonstrou o potencial da Inteligência Artificial em aprimorar os cuidados com feridas.

Os resultados alcançaram uma notável acurácia de 100% na classificação binária (Normal vs. Ferida) utilizando a arquitetura ResNet50, validando a capacidade de distinguir com segurança a presença de uma lesão cutânea. Para a classificação multiclasse das classes de feridas, o ResNet50 sem *data augmentation* obteve 85.26% de acurácia. A aplicação de um filtro de confiança a este modelo elevou a acurácia para 92.31% em amostras de alta certeza. A exploração do Derm Foundation Model também apresentou resultados competitivos, reforçando o valor de modelos pré-treinados em domínios específicos. A análise qualitativa confirmou que feridas venosas são geralmente mais fáceis de classificar, enquanto feridas diabéticas representam o maior desafio.

Contudo, este estudo possui limitações. O *dataset* utilizado é de tamanho restrito e pode não capturar toda a diversidade das apresentações clínicas. Além disso, os desafios persistentes na classificação de feridas diabéticas indicam a necessidade de refinar abordagens para características visuais sobrepostas.

Essas limitações, embora presentes, direcionam claramente os próximos passos da pesquisa. Para aprimorar e expandir o WoundScan, futuras direções incluem a investigação de técnicas de segmentação de imagens como pré-processamento para refinar a análise da ferida, a implementação e comparação com outras arquiteturas de *deep learning* e modelos de *ensemble*, e a exploração de modelos mais leves para a classificação binária, visando um menor custo de inferência. Crucialmente, a expansão e diversificação do conjunto de dados, com a utilização de *datasets* como o WoundsDB, que apesar de ter apenas 188 casos, oferece imagens multimodais (termovisão, estereovision e percepção de profundidade) além das coloridas, permitirá avaliar o ganho em acurácia e

generalização com a fusão dessas informações para uma compreensão mais completa da ferida. A investigação contínua do Derm Foundation Model e a combinação com outros modelos também permanecem promissoras. O WoundScan, portanto, estabelece uma base sólida para a colaboração entre tecnologia e expertise clínica, prometendo um futuro de cuidado mais preciso e eficiente aos pacientes.

8. REFERÊNCIAS

ABUBAKAR, A.; ALI, A.; UGAIL, H. Mobile metadata assisted community database of chronic wound images. **Proceedings of the 6th International Conference on Digital Image Processing (ICDIP 2014)**, 91590F, 2014. Disponível em: https://www.researchgate.net/publication/280266251_Mobile_metadata_assisted_community_database_of_chronic_wound_images. Acesso em: 28 maio 2025.

CARTER, M. J.; DAVANZO, J. E.; HAUGHT, R.; NUSGART, M.; CARTWRIGHT, D.; FIFE, C. E. Chronic wound prevalence and the associated cost of treatment in Medicare beneficiaries: changes between 2014 and 2019. *Journal of Medical Economics*, v. 26, n. 1, p. 894–901, 2023. Disponível em: <https://doi.org/10.1080/13696998.2023.2232256>. Acesso em: 4 jun. 2025.

DATA SCIENCE DOJO. Choosing the Right Vector Embedding Model. [S.l.], 2024. Disponível em: <https://datasciencedojo.com/blog/choosing-the-right-vector-embedding-model/>. Acesso em: 30 maio 2025.

Ferreira, J. C., & Patino, C. M. (2017). Entendendo os testes diagnósticos. Parte 1. *Jornal Brasileiro de Pneumologia*, 43(5), 330-330. <http://dx.doi.org/10.1590/S1806-37562017000000330>. Acesso em: 28 maio 2025.

GRAND VIEW RESEARCH. **Artificial Intelligence in Healthcare Market Size, Share & Trends Analysis Report By Component (Hardware, Software, Services), By Application, By Technology, By End-use, And Segment Forecasts, 2022 - 2030**. San Francisco: Grand View Research, 2022. Disponível em: <https://www.grandviewresearch.com/press-release/global-artificial-intelligence-healthcare-market>. Acesso em: 28 maio 2025.

GUO, Y.; LU, H.; ZHANG, J.; SHANG, K.; LI, S. Multi-Class Wound Classification via High and Low-Frequency Guidance Network. **PubMed** (PMID: 38135976), 2023. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/38135976/>. Acesso em: 28 maio 2025.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep Residual Learning for Image Recognition. **arXiv preprint arXiv:1512.03385**, 2015. Disponível em: <https://arxiv.org/abs/1512.03385>. Acesso em: 28 maio 2025.

HUANG, Z.; TAN, J.; LIU, Y.; LI, Y. Development of a deep learning-based tool to assist wound classification. **arXiv preprint arXiv:2303.16522**, 2023. Disponível em: <https://ar5iv.labs.arxiv.org/html/2303.16522>. Acesso em: 28 maio 2025

IBRAHIMFATEEN. **Collected and Categorized Wound Images Dataset**. Kaggle, 2024. Disponível em: <https://www.kaggle.com/datasets/ibrahimfateen/wound-classification>. Acesso em: 28 maio 2025.

JUSZCZYK, J. M. et al. Wound 3D Geometrical Feature Estimation Using Poisson Reconstruction. **IEEE Access**, v. 9, p. 7894-7907, 2021. Disponível em: <https://ieeexplore.ieee.org/document/9245494>. Acesso em: 28 maio 2025.

KIRALY, A. P. et al. Health AI Developer Foundations. **arXiv preprint arXiv:2411.15128v2**, 2024. Disponível em: <https://arxiv.org/pdf/2411.15128v2>. Acesso em: 28 maio 2025.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In: **Advances in Neural Information Processing Systems (NIPS 2012)**, 2012. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf. Acesso em: 28 maio 2025.

MUKHERJEE, S. The Annotated ResNet-50. **Towards Data Science**, 18 ago. 2022. Disponível em: <https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758>. Acesso em: 30 maio 2025.

PATEL, Y.; SHAH, T.; DHAR, M. K.; ZHANG, T.; NIEZGODA, J.; GOPALAKRISHNAN, S.; YU, Z. Integrated image and location analysis for wound classification: a deep learning approach. **Scientific Reports**, v. 14, n. 1, 56626, 2024. Disponível em: <https://www.nature.com/articles/s41598-024-56626-w.pdf>. Acesso em: 28 maio 2025

RADECKI, M.; WIATRAK, D. Data Augmentation On-the-fly and Active Learning in Data Stream Classification. **arXiv.org**, 2022. Disponível em: <https://arxiv.org/abs/2210.06873>. Acesso em: 28 maio 2025.

ROMANELLI, M. et al. Standardized Assessment and Documentation of Wounds: A Comprehensive Review. **Advances in Wound Care**, v. 20, n. 4, p. 182-194, 2011. Disponível em: <https://www.magonlinelibrary.com/doi/abs/10.12968/jowc.2020.29.3.162>. Acesso em: 28 maio 2025.

ROSTAMI, B.; ANISUZZAMAN, D. M.; WANG, C.; GOPALAKRISHNAN, S.; NIEZGODA, J.; YU, Z. Multiclass Wound Image Classification using an Ensemble Deep CNN-based Classifier. **arXiv preprint arXiv:2010.09593v1**, 2020. Disponível em: <https://arxiv.org/pdf/2010.09593>. Acesso em: 28 maio 2025.

SEN, C. K. et al. Human skin wounds: A major and snowballing threat to public health and the economy. **Wound Repair and Regeneration**, v. 17, n. 6, p. 763–771, 2009. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC2810192/pdf/nihms165224.pdf>. Acesso em: 28 maio 2025.

SHAHRIAR, N. What Is Convolutional Neural Network (CNN)? Deep Learning. **Medium**, 2023. Disponível em: <https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5>. Acesso em: 30 maio 2025.

THOMSEN, K. et al. Deep Learning for Diagnostic Binary Classification of Multiple-Lesion Skin Diseases. **Frontiers in Medicine**, v. 7, p. 574329, 2020. Disponível em: <https://www.frontiersin.org/articles/10.3389/fmed.2020.574329/full>. Acesso em: 28 maio 2025.

WANG, C.; ZHANG, H.; LI, K.; LV, X.; ZHANG, T. Unified wound diagnostic framework for wound segmentation and classification. **Computers in Biology and Medicine**, v. 167, 107615, 2023. Disponível em: <https://www.peeref.com/works/84840261>. Acesso em: 28 maio 2025.