



Aplicação de Redes Neuronais Convolucionais para a Classificação de Úlceras do Pé Diabético

MIGUELA TEIXEIRA SEQUEIRA

outubro de 2024

Aplicação de Redes Neurais Convolucionais para a Classificação de Úlceras do Pé Diabético

Miguela Teixeira Sequeira

**Dissertação para obtenção do Grau de Mestre em
Engenharia Biomédica**

Orientador: Luís Coelho

Júri:

Presidente: Natércia Maria Pereira Machado Lima

Vogais: Nuno Eduardo Dias Gueiral

[Nome do Vogal1, Categoria, Escola]

[Nome do Vogal2, Categoria, Escola] (até 4 vogais)

Resumo

Esta tese explora a aplicação de redes neurais convolucionais (CNNs) para a classificação de úlceras do pé diabético, utilizando as arquiteturas VGG16, VGG19 e MobileNetV2. O objetivo principal é desenvolver e comparar modelos de *deep learning* capazes de identificar com precisão áreas lesionadas em imagens clínicas de pés diabéticos, ajudando na prevenção e no tratamento eficaz das úlceras. Foi realizado um estudo com base num conjunto de dados de imagens anotadas, avaliando o desempenho dos modelos em termos da *accuracy*, precisão, *recall* e F1-Score. A VGG19 destacou-se com uma *accuracy* de 93%, evidenciando uma capacidade superior de localizar lesões em imagens complexas, com ativações mais focadas nas áreas relevantes. A MobileNetV2, por outro lado, apresentou um bom desempenho em termos de eficiência computacional, sendo adequada para dispositivos móveis e ambientes com restrições de hardware, embora tenha apresentado um valor de *accuracy* ligeiramente inferior em relação às VGGs. O estudo também discute as limitações de cada arquitetura, como a maior tendência ao *overfitting* nos modelos mais profundos e a menor capacidade de abstração de detalhes clínicos no MobileNetV2. Os resultados indicam que o uso de CNNs tem grande potencial no diagnóstico clínico assistido por imagem, especialmente para patologias como o pé diabético, onde a detecção precoce e precisa é crucial para evitar amputações.

Palavras-chave: Redes Neurais Convolucionais, Classificação de Imagens Médicas, Úlceras do Pé Diabético, *Deep Learning*

Abstract

This thesis investigates the application of convolutional neural networks (CNNs) for the classification of diabetic foot ulcers, using the VGG16, VGG19, and MobileNetV2 architectures. The primary objective is to develop and compare deep learning models capable of accurately identifying ulcerated regions in clinical images of diabetic feet, thereby aiding in the prevention and effective treatment of foot ulcers. A comprehensive study was conducted using an annotated dataset of medical images, evaluating the performance of the models in terms of accuracy, precision, recall, and F1-score. VGG19 achieved the highest accuracy at 93%, demonstrating superior ability to focus activations on relevant lesion areas in complex images. MobileNetV2, while slightly less accurate, excelled in computational efficiency, making it a suitable choice for mobile devices and environments with hardware constraints. The study also highlights the limitations of each architecture, such as increased risk of overfitting in deeper models and the lower capability of MobileNetV2 to capture fine clinical details. These findings suggest that CNNs hold significant potential in computer-aided clinical diagnosis, particularly in the early and precise detection of diabetic foot ulcers, where timely intervention is crucial to prevent amputations.

Keywords: Convolutional Neural Networks, Medical Image Classification, Diabetic Foot Ulcers, *Deep Learning*

Agradecimentos

Aos meus pais, que sempre foram o meu porto seguro. Obrigada por me darem asas para sonhar e raízes para me manter firme. Pelo incentivo constante, pela força nos momentos difíceis e por nunca hesitarem em dar-me todos os meios para eu alcançar os meus objetivos. A vossa confiança em mim fez toda a diferença. Esta conquista é vossa também.

À minha irmã, obrigada por seres a minha companhia em todas as horas, por estares sempre pronta a ajudar, e por me mostrares, nos momentos mais simples, o valor do apoio incondicional.

À minha madrinha, sempre à distância de uma chamada, com as palavras certas e o carinho que só quem nos conhece desde sempre consegue dar.

Ao meu namorado e companheiro de vida, que acreditou em mim até nos momentos em que eu mesma duvidava. Obrigada por seres a minha força constante, por veres em mim o que às vezes eu não conseguia ver, e por seres o meu apoio em todos os sentidos.

Aos amigos que fiz no ISEP um agradecimento especial. Esta jornada teria sido impossível sem as pessoas certas do meu lado, para partilhar não só os desafios, mas também as alegrias e conquistas. Com vocês construí memórias que vou levar para o resto da vida.

Aos amigos da Universidade Católica, apesar da distância física, nunca deixaram de estar ao meu lado. Mesmo longe, continuaram a ser um suporte inestimável, e não teria concluído a licenciatura, nem agora o mestrado, sem vocês. O que aprendi nesta faculdade, as experiências vividas, os momentos de praxe, são as memórias que vou levar para a vida.

Aos meus restantes amigos, agradeço todo o apoio dado, uma vez que continuam do meu lado ano após anos, desafio após desafio. A vida sem vocês realmente não faria sentido.

Ao meu professor e orientador de tese, Professor Luís, o meu mais sincero agradecimento. Obrigada por me ter mostrado uma cadeira tão cativante, que abriu as portas para o fascinante mundo do *machine learning*. Foi nesta descoberta que encontrei uma nova paixão e um novo horizonte para os meus estudos e sonhos.

Ao meu querido avô Zequinha e ao meu tio do Brasil, que já não estão fisicamente entre nós, mas que sinto presentes em cada passo. Espero que, onde quer que estejam, estejam orgulhosos de mim e das minhas conquistas. O legado de amor, ensinamentos e valores que deixaram continua a guiar-me e a dar-me força. Carrego-vos sempre comigo no coração, e este trabalho é também para vocês.

E, por fim, um agradecimento a mim mesma. Por ter a coragem de sonhar alto, de acreditar no meu próprio potencial e de seguir em frente, mesmo quando o caminho parecia incerto. Por ter lutado por cada conquista e por não ter desistido. Este trabalho é um reflexo da minha determinação, do meu desejo de voar mais alto e de não olhar para trás.

Índice

1	Introdução	1
1.1	Contexto	1
1.2	Problema	4
1.3	Perguntas de Investigação	5
1.4	Objetivo	6
1.5	Estrutura do Documento	6
2	Conceitos Teóricos	7
2.1	Fisiopatologia da Diabetes <i>Mellitus</i>	7
2.2	<i>Machine Learning</i>	10
	2.2.1 Desenvolvimento de um modelo de Machine Learning	10
	2.2.2 Aprendizagem Profunda	12
3	Estado da Arte	15
3.1	Modelos de Classificação	15
3.2	Bases de Dados	18
4	Análise e Desenvolvimento do Modelo	21
4.1	Ferramentas e Ambientes Utilizados	21
4.2	Modelos Utilizados	22
4.3	Preparação e Pré-Processamento de Dados	24
4.4	Configuração da Arquitetura dos Modelos	26
4.5	Treino e Validação	28
4.6	Análise com Grad-CAM	30
5	Resultados e Discussão	35
6	Conclusão	45
	Referências Bibliográficas	47

Lista de Figuras

Figura 1 - Prevalência da diabetes a nível mundial em 2021 (IDF Diabetes Atlas 10th Edition, 2021).....2

Figura 2- Prevalência da diabetes em Portugal entre 2019 e 2021 (Relatório Anual Do Observatório Nacional Da Diabetes-Edição de 2023 Diabetes Factos e Números, n.d.).....3

Figura 3- Complicações do pé diabético em comparação com o cancro (Armstrong et al., 2020).....5

Figura 4 - Imagens de pés diabéticos (a)–(d) Pé normal e saudável. (e)–(h) Pé afetado por uma úlcera diabética (Thotad et al., 2023).....9

Figura 5 - Classificação das feridas no pé diabético.....12

Figura 6 - Flowchart do algoritmo de machine learning relativo a aprendizagem supervisionada.....13

Figura 7- Componentes de uma CNN (Taye, 2023).14

Figura 8 - Arquiteturas das Redes VGG16 e VGG19 (Shadeed et al., 2020).....24

Figura 9 - Arquiteturas da Rede MobileNetV2 (Al-gaashani et al., 2022).....24

Figura 10 - Exemplo de grelha de imagens obtidas com Grad-Cam.....34

Figura 11 – Resultado da avaliação usando Grad-CAM na rede VGG16.....40

Figura 12 – Resultado da avaliação usando Grad-CAM na rede VGG19.....41

Figura 13 - Resultado da avaliação usando Grad-CAM na rede MobileNetV2.....42

Lista de Tabelas

Tabela 1- Sumário dos trabalhos encontrados na literatura.....	19
Tabela 2 - Resumo dos datasets encontrados	20
Tabela 3 - Resumos dos resultados das métricas obtidas em cada modelos	38
Tabela 4 -Resumos dos resultados das matrizes de confusão obtidas em cada modelos.....	40

Acrónimos e Símbolos

Lista de Acrónimos

CNN	Redes Neurais Convolucionais
DFU	<i>Diabetic Foot Ulcer</i>
DM	Diabetes <i>Mellitus</i>
FN	Falsos Negativos
FP	Falsos Positivos
HDL	High-Density Lipoprotein
IA	Inteligência Artificial
LDL	Low-Density Lipoprotein
ML	<i>Machine Learning</i>
PBE	Proposta Baseada em Evidências
TN	Verdadeiros Negativos
TP	Verdadeiros Positivos
VGG	<i>Visual Geometry Group</i>
WMC	<i>Wound Multimodality Classifier</i>

1 Introdução

Este capítulo introduz o trabalho contextualizando a alta prevalência de úlceras no pé diabético, uma condição responsável por um grande número de amputações, muitas das quais poderiam ser prevenidas com a detecção precoce e medidas adequadas. Diante desse problema, o estudo investiga a aplicação de redes neurais convolucionais (CNNs) como ferramenta para melhorar a identificação dessas lesões, procurando desenvolver modelos capazes de auxiliar no diagnóstico clínico.

1.1 Contexto

A Diabetes é uma doença metabólica e crônica representada por níveis elevados de glicose, mais conhecido como açúcar, na corrente sanguínea. Quando na presença de elevados níveis de glicose no sangue, denomina-se hiperglicemia e acontece devido a problemas de secreção de insulina ou da ação da mesma (Alberti & Zimmet, 1998) Com o passar do tempo e sem a devida medicação, pode provocar lesões graves no coração, nos vasos sanguíneos, nos olhos, nos rins, nos nervos, nos pés e entre outros (*Diabetes*, 2023).

Existem diversas causas que resultam no aparecimento da diabetes. Os diferentes tipos de diabetes dividem-se segundo o aparecimento desta. A Diabetes Tipo 1 pode surgir em qualquer idade, desde o nascimento até ao jovem adulto, sendo uma das doenças com uma maior incidência na juventude e consiste no processo de autodestruição das células responsáveis pela produção de insulina, resultando numa acumulação de glicose no sangue (Atkinson et al., 2014). No caso da Diabetes Tipo 2, esta surge predominantemente em adultos acima dos 40 anos e é frequentemente associada à resistência à insulina e à diminuição da capacidade do pâncreas de produzir insulina suficiente para superar essa resistência. Fatores genéticos e ambientais, como obesidade e estilo de vida sedentário, desempenham um papel importante no desenvolvimento da diabetes tipo 2 (Alberti & Zimmet, 1998).

Estimativas divulgadas na revista *The Lancet* sugerem que, em 2050, mais de 1.31 mil milhões de pessoas poderão estar a viver com diabetes em todo o mundo. Este número representa 1.31

mil milhões de indivíduos com uma condição que pode causar morbidade significativa, altas taxas de mortalidade e que interage de forma negativa com várias outras doenças. Este aumento, por sua vez, é causado pelo aumento da prevalência da obesidade e por mudanças demográficas. Em 2021, diabetes tipo 2 correspondia a 90 % de toda a prevalência de diabetes. Grande parte deste facto é atribuída a fatores de risco sociais, como um índice de massa corporal elevado, riscos alimentares, fatores ambientais e profissionais, tabagismo, consumo de álcool e baixa atividade física. Estes fatores prosperam na natureza obesogénica dos ambientes e na forma desigual como os nossos recursos são organizados (Ong et al., 2023).

Cerca de 422 milhões de pessoas em todo o mundo têm diabetes, sendo a maioria delas residente em países de baixo e médio rendimento. Anualmente, são atribuídas diretamente à diabetes 1,5 milhões de mortes. Tanto o número de casos como a prevalência da diabetes têm vindo a aumentar de forma constante ao longo das últimas décadas (*Diabetes*, 2023).

Acredita-se que existem 537 milhões de adultos entre os 20 e os 79 anos em todo o mundo, diagnosticados com diabetes, o que representa 10,5 % desse grupo etário. Prevê-se que em 2030 este número suba para 643 milhões e, em 2045, para 783 milhões. Deste modo, enquanto se estima um crescimento populacional mundial de 20 % durante este período, antecipa-se um aumento de 46% no número de pessoas com diabetes (Figura 1) (IDF Diabetes Atlas 10th Edition, 2021).

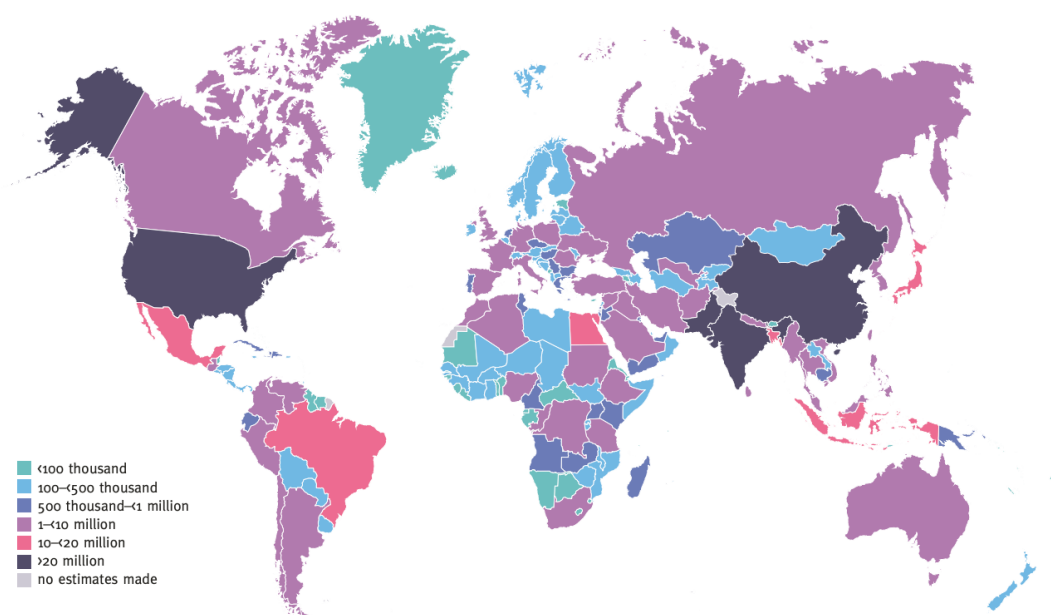


Figura 1 - Prevalência da diabetes a nível mundial em 2021 (IDF Diabetes Atlas 10th Edition, 2021)

Conforme mencionado pelo presidente da Sociedade Portuguesa de Diabetologia (SPD), Portugal encontra-se entre os três países europeus com maior prevalência de diabetes, juntamente com a Alemanha e a Espanha. No ano de 2021, a estimativa da prevalência da diabetes na população portuguesa foi de 14,1% entre os 20 e 79 anos (totalizando 7,8 milhões de pessoas). Isto significa que aproximadamente 1,1 milhões de portugueses nesse grupo etário são afetados pela doença. Em relação à composição da taxa de prevalência da diabetes, 56% dos indivíduos já tinham recebido o diagnóstico, enquanto 44% ainda não tinham sido diagnosticados (*Relatório Anual Do Observatório Nacional Da Diabetes-Edição de 2023 Diabetes Factos e Números, n.d.*).

Como observado na figura 2, evidencia-se uma diferença estatisticamente significativa na prevalência da Diabetes entre homens e mulheres. Além disso, observa-se um relevante aumento na prevalência desta doença com o avançar da idade. Notavelmente, mais de um quarto das pessoas na faixa etária dos 60-79 anos é afetado pela doença. Estes dados refletem a relevância de considerar fatores de género e idade ao abordar a incidência desta condição de saúde em Portugal (*Relatório Anual Do Observatório Nacional Da Diabetes-Edição de 2023 Diabetes Factos e Números, n.d.*).

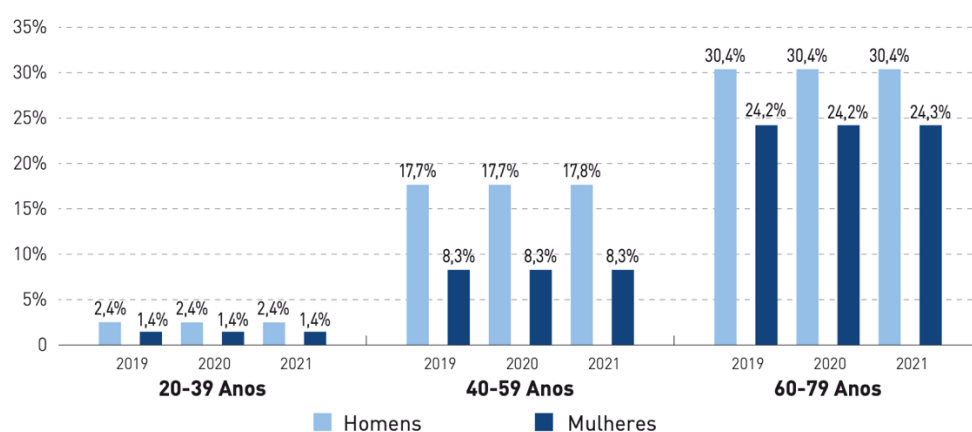


Figura 2- Prevalência da diabetes em Portugal entre 2019 e 2021 (*Relatório Anual Do Observatório Nacional Da Diabetes-Edição de 2023 Diabetes Factos e Números, n.d.*)

Um baixo valor de glicemia e tempo de doença são causa direta de várias complicações sistêmicas da Diabetes *Mellitus*. O pé diabético é um dos mais significantes, não apenas devido às consequências para a qualidade de vida do paciente, mas também ao alto custo social e económico. De acordo com a Organização Mundial de Saúde, um indicador do pé diabético é uma úlcera no pé que se desenvolve sob o maléolo em combinação, acompanhada de neuropatia autónoma e mais distúrbios isquémicos e infecciosos (Neves et al., 2013). Há evidências de que mais de 10% das pessoas com Diabetes *Mellitus* (DM) estão em risco de desenvolver úlceras nos pés ao longo da sua vida. Além disso, essa predisposição aumenta a probabilidade de lesões causadas pela neuropatia periférica em 80 a 90% dos casos, assim como pela doença vascular periférica e deformidades (Hassan et al., 2010).

A importância da avaliação adequada do pé diabético não pode ser subestimada. Cerca de 50% das amputações e ulcerações podem ser evitadas através de uma avaliação correta, que inclui a classificação do pé e a atribuição do grau de risco de ulceração. Isso permite a aplicação de estratégias preventivas eficazes. O papel dos enfermeiros é fundamental nesse processo, pois são eles que realizam o rastreio do pé diabético, promovem a saúde e capacitam os pacientes para o autocuidado. A literatura enfatiza a importância do envolvimento dos profissionais de saúde na prevenção de complicações e na manutenção da qualidade de vida dos pacientes diabéticos. Consultas regulares, acompanhamento contínuo e estratégias educacionais são essenciais para ajudar os pacientes e suas famílias a adotar comportamentos de autocuidado, contribuindo para uma melhor gestão e adaptação ao regime terapêutico (Teixeira do Nascimento et al., 2019).

Assim, torna-se imperativo a urgência na prevenção e sobretudo a antecipação do diagnóstico.

1.2 Problema

As complicações nos pés continuam a ser a principal razão para hospitalizações entre os indivíduos com diabetes nos países desenvolvidos. A probabilidade, ao longo da vida de um paciente com diabetes, desenvolver uma úlcera pode chegar a 30 %. Até 85 % de todas as amputações nos membros inferiores em casos de diabetes estão associadas a úlceras nos pés. Metade dos pacientes idosos com diabetes tipo 2 apresentam fatores de risco para problemas nos pés, sublinhando a importância de rastreios regulares por meio de exames clínicos cuidadosos. Foi demonstrado que até 50% das amputações e úlceras do pé na diabetes podem ser prevenidas através de uma identificação e educação eficazes (Boulton, 2019).

A figura 3 ilustra de forma clara a gravidade das úlceras do pé diabético (DFU), evidenciando-as como um problema sério e muitas vezes subestimado. Comparando a taxa de mortalidade em cinco anos entre diversas condições de saúde, as DFUs apresentam uma mortalidade de aproximadamente 30%, superando várias doenças graves, como o cancro de mama. Esse dado realça que as DFUs não são apenas uma complicação localizada, mas sim um fator de risco significativo para a saúde e sobrevivência dos pacientes diabéticos.

Além disso, o gráfico presente na figura mostra que as úlceras podem levar a amputações, que estão associadas a taxas de mortalidade ainda mais elevadas, reforçando a seriedade da condição. As amputações maiores, por exemplo, têm uma mortalidade que se aproxima dos 70%, um indicador bastante alarmante do impacto que essas complicações podem ter na vida dos pacientes.

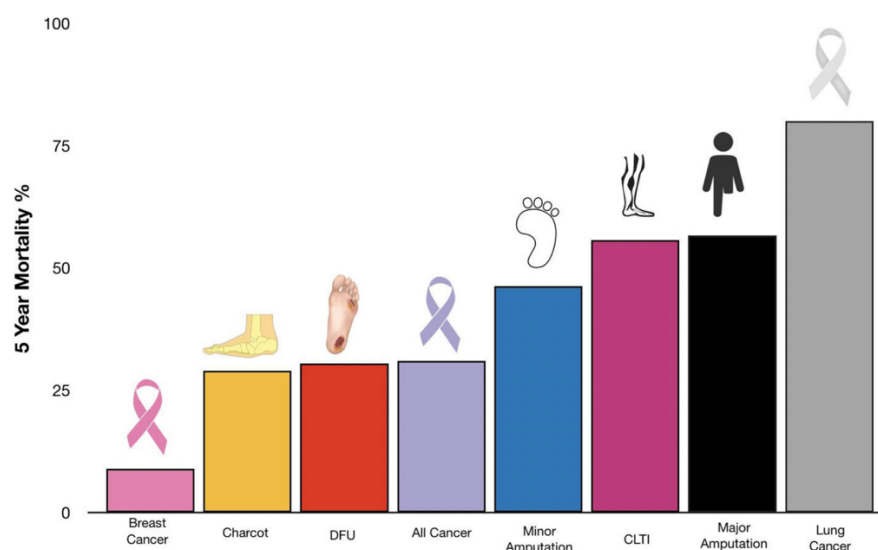


Figura 3- Complicações do pé diabético em comparação com o cancro (Armstrong et al., 2020).

1.3 Perguntas de Investigação

A Proposta Baseada em Evidências (PBE) sugere que os problemas clínicos encontrados na prática assistencial, de ensino ou pesquisa, sejam analisados de forma detalhada e organizados utilizando a estratégia PICO (Mamédio et al., 2007).

PICO é uma sigla que representa os elementos fundamentais de uma questão de pesquisa na Proposta Baseada em Evidências (PBE): Paciente (P), Intervenção (I), Comparação (C) e Resultados (R) (ou *Outcome*). Estes componentes são essenciais para formular perguntas que guiam a busca bibliográfica por evidências ('Popping the (PICO) Question in Research and Evidence-Based Practice', 2002). A estratégia PICO é versátil e pode ser aplicada para formular questões de pesquisa em diversas áreas, como na clínica, na gestão de recursos humanos e materiais, na procura de instrumentos de avaliação de sintomas, entre outras. Uma pergunta de pesquisa bem construída ajuda a identificar quais evidências são necessárias para responder à questão clínica, facilita a busca por evidências nas bases de dados, direciona o foco da pesquisa e evita buscas desnecessárias (Especial et al., 2004).

Relativamente a esta dissertação e seguindo a estratégia PICO, podemos colocar a seguinte pergunta de investigação: Em pacientes com úlceras no pé (P), como é que a utilização de visão computacional e algoritmos de aprendizagem de máquina (I) se compara a exatidão com os métodos tradicionais ou especialistas humanos (C), em termos da exatidão da classificação do tipo e gravidade das úlceras a partir de imagens (O)?

1.4 Objetivo

Na sequência do problema identificado, surge a necessidade de desenvolver um método eficaz para diagnóstico e classificação de úlceras no pé diabético. Atualmente, os métodos existentes apresentam limitações, tornando crucial o aprimoramento da abordagem para melhorar a precisão do diagnóstico, estratificar riscos e otimizar o tratamento. Isso visa a redução de complicações e a melhoria dos cuidados oferecidos aos pacientes diabéticos com úlceras nos pés.

A pergunta de investigação foca em comparar a utilização de visão computacional e algoritmos de aprendizagem automática com métodos tradicionais ou avaliações por especialistas humanos em pacientes com úlceras no pé diabético. O objetivo é verificar se a tecnologia moderna proporciona uma maior exatidão na classificação do tipo e gravidade das úlceras a partir de imagens. Assim, buscamos responder se a abordagem tecnológica resulta numa precisão superior na identificação das úlceras em comparação com os métodos e a experiência tradicionais.

Os objetivos específicos deste estudo são:

- **Comparar as métricas de desempenho de diferentes redes neurais:** Avaliar a eficácia de várias arquiteturas de redes neurais na classificação de úlceras do pé diabético.
- **Analisar a precisão na identificação e classificação das úlceras:** Verificar como os métodos de aprendizagem automática se comportam na tarefa de classificação de pés com úlceras e sem úlceras.
- **Identificar a abordagem mais eficaz:** Determinar qual técnica oferece os melhores resultados em termos de precisão e confiabilidade na identificação de úlceras do pé diabético.

1.5 Estrutura do Documento

A dissertação está dividida em seis partes. A primeira parte é representada pela introdução, onde se estabelece o contexto e os objetivos da pesquisa. A seguir, explora-se os conceitos teóricos relevantes para o estudo, fornecendo a base necessária para compreender o tema em questão. Em seguida, examina-se o estado da arte, analisando o que já foi feito na área e identificando lacunas ou áreas para aprimoramento. Posteriormente, foca-se na análise e no desenvolvimento do modelo proposto, detalhando o processo de criação e as decisões tomadas. Passa-se então para a implementação prática do modelo, onde se coloca em prática as teorias e conceitos discutidos anteriormente. Por fim, chega-se à conclusão, resumindo os principais resultados, destacando contribuições e delineando possíveis direções futuras para a pesquisa

2 Conceitos Teóricos

Este capítulo explora, primeiramente, a fisiopatologia da diabetes *mellitus*, destacando como esta condição afeta o metabolismo e contribui para complicações graves, como úlceras nos pés. Em seguida, é discutido o papel do *machine learning* no campo médico, com foco no desenvolvimento de modelos capazes de identificar padrões em grandes volumes de dados clínicos. A *deep learning*, um ramo avançado do *machine learning*, é abordada como uma técnica poderosa para a criação de modelos mais precisos e eficazes, particularmente no contexto de imagens médicas, permitindo melhorias significativas na detecção de úlceras no pé diabético.

2.1 Fisiopatologia da Diabetes *Mellitus*

A Diabetes *Mellitus* é uma condição crónica não transmissível para a qual não existe uma cura conhecida. Apesar de alguns pacientes diabéticos apresentarem sintomas, muitos não os têm, desconhecendo assim a presença da doença e mantendo os mesmos hábitos de vida (Casarin et al., 2022).

Esta é uma condição em que os níveis de glicose, um tipo de açúcar, estão elevados no sangue devido à falta de insulina, uma hormona produzida pelo pâncreas. A função principal da insulina é permitir que o açúcar proveniente dos alimentos entre nas células, onde é transformado em energia. No caso da diabetes, esse açúcar não consegue entrar nas células devido à deficiência de insulina, acumulando-se no sangue e resultando em hiperglicemia, que é o principal sintoma da doença (Ministério da Saúde, 2020).

A Diabetes *Mellitus* perdura ao longo da vida e pode resultar em múltiplas complicações ao longo do tempo. Se não for devidamente controlada, a doença pode apresentar sintomas graves, como dores de cabeça, inquietação, irritabilidade, palidez, suores, taquicardia, confusão mental, desmaios, convulsões e até mesmo a entrada em estado de coma. A falta de tratamento adequado aumenta significativamente o risco de complicações a longo prazo, como o desenvolvimento de problemas de coagulação sanguínea, dificultando a cicatrização de úlceras e potencialmente resultando em amputações. Além disso, a má circulação sanguínea nos olhos pode levar a complicações como

retinopatia diabética, glaucoma e cataratas, podendo resultar em perda parcial ou total da visão. Problemas renais também são comuns, incluindo insuficiência renal e nefropatia diabética, onde o mau funcionamento dos vasos sanguíneos renais pode levar a falhas na filtragem de impurezas e na retenção de nutrientes. Em termos de manifestações neurológicas, estas podem incluir perda sensorial e úlceras nos membros, enquanto as infecções são frequentes devido à redução da imunidade, tornando os pacientes diabéticos mais suscetíveis a infecções cutâneas, urinárias e bucais (Cole & Florez, 2020).

Clinicamente, esta pode ser dividida em dois tipos distintos: tipo 1 e tipo 2. A Diabetes *Mellitus* tipo 1, é caracterizada por um processo autoimune no qual não ocorre a destruição das células beta no pâncreas, que são responsáveis pela produção de insulina. Isso resulta numa deficiência parcial ou total na produção de insulina, levando a problemas metabólicos que se manifestam clinicamente através de sintomas como produção frequente de muco, sede excessiva, cetoacidose e perda de peso involuntária (Paula Vieira Carvalho et al., 2022).

A Diabetes *Mellitus* tipo 2, trata-se de uma condição variada que geralmente afeta indivíduos com mais de 40 anos, resultante de deficiências na secreção e ação da insulina, ambas influenciadas por fatores genéticos. O tipo 2 é a forma mais comum, presente em 90 % a 95 % dos pacientes, onde a secreção de insulina pelo pâncreas ocorre normalmente, mas há excesso de insulina e glicose no sangue, com pouca glicose nas células. A maioria dos pacientes apresenta excesso de peso ou obesidade. A característica principal deste tipo é a incapacidade de utilizar a glicose adequadamente para produzir energia, resultando num aumento dos níveis de glicose no sangue, que em alguns casos podem chegar a ser dez vezes mais elevados. Exames de diagnóstico nesses pacientes indicam elevados níveis de triglicéridos, aumento do colesterol LDL, o “colesterol mau” que contribui para a acumulação de placas nas artérias, e redução do colesterol HDL, o “colesterol bom”, que ajuda a remover o excesso de colesterol das artérias (Vidal, 2023).

Na Diabetes *Mellitus* tipo 1, como referido anteriormente, o pâncreas não produz insulina, o que significa que é necessário administrar a insulina por via injetável, já que se for tomada por via oral, será destruída pelo estômago. A administração da insulina é feita geralmente na camada de gordura sob a pele, utilizando uma seringa pequena e uma agulha fina, que são inseridas nos membros superiores, inferiores ou na parede abdominal, podendo variar de uma a quatro vezes por dia. Para aqueles que têm aversão a agulhas, existem dispositivos como a bomba de insulina, que injeta a insulina de forma subcutânea, ou a caneta de insulina, que contém um cartucho. A ação da insulina pode ser rápida e curta, intermediária ou prolongada, dependendo do tipo utilizado. O tratamento da Diabetes *Mellitus* deste tipo, procura regularizar o metabolismo, prevenir complicações agudas e crônicas, reduzir internações e problemas psicossociais, e educar sobre hábitos alimentares saudáveis. Além do uso de medicamentos, o tratamento envolve mudanças no estilo de vida, incluindo dieta e exercício físico regular (Vidal, 2023).

Como mencionado anteriormente, uma das preocupações dos pacientes com diabetes é o pé diabético. Este resulta de uma combinação de neuropatia periférica, doença vascular periférica e alterações biomecânicas que levam ao aparecimento de complicações, como ulcerações, infecções e, a um longo prazo, amputações. A neuropatia periférica afeta as fibras sensoriais, motoras e autónomas, o que resulta numa perda de sensibilidade nos pés, dificultando a percepção por parte do

paciente de lesões, pressão ou temperatura (de Souza Mendonça et al., 2011). Como consequência, pequenas lesões podem agravar-se significativamente. A par disso, a doença vascular periférica compromete a cicatrização das feridas ao dificultar o fluxo de sangue para os membros inferiores. Estas condições agravam-se ainda mais com alterações biomecânicas, que podem modificar a marcha do paciente, aumentando a pressão em determinadas áreas do pé e elevando o risco de ulceração (Fajardo, 2006).



Figura 4 - Imagens de pés diabéticos (a)–(d) Pé normal e saudável. (e)–(h) Pé afetado por uma úlcera diabética (Thotad et al., 2023).

Além disso, as úlceras podem ainda ser classificadas pela sua etiologia em neuropáticas ou isquêmicas. As úlceras neuropáticas ocorrem devido à perda de sensibilidade nos pés, apresentando calosidades, dedos em garra e pele quente e seca. Já as úlceras isquêmicas surgem como resultado de uma circulação insuficiente, caracterizando-se por pele fria, ausência de pêlos, unhas espessas e presença de gangrena (Duarte et al., 2011).

A avaliação do risco de ulceração é um passo essencial para evitar complicações graves. Este processo deve ser parte integrante do exame físico regular dos utentes diabéticos, devendo ser realizada de forma minuciosa. Para identificar a perda de sensibilidade associada à neuropatia, utiliza-se o monofilamento de 10g de Semmes-Weinstein, onde são efetuados toques ao pé do paciente de modo a avaliar a sua resposta a estes. Este pode ser complementado com testes de sensibilidade vibratória (diapasão de 128 Hz), sensibilidade tátil e reflexos patelares e aquilianos (Francisco Henrique Moura George, 2011).

No que diz respeito à prevenção das ulcerações, a abordagem deve ser multidisciplinar, centrando-se na educação do utente e na vigilância regular dos pés. Todos os utentes diabéticos devem ser

submetidos a uma consulta anual de vigilância dos pés, onde são avaliados fatores como o uso de calçado adequado, a presença de calosidades, o estado das unhas e da pele dos pés, e a existência de deformidades ósseas. O uso de calçado apropriado é uma das medidas mais importantes para prevenir lesões no pé diabético. Este deve ser confortável, bem ajustado e feito de materiais flexíveis como couro. A educação para a saúde desempenha um papel fundamental na prevenção, uma vez que promove a adoção de comportamentos preventivos. Os utentes e suas famílias devem estar cientes dos cuidados diários necessários, como a inspeção regular dos pés, a higiene adequada, a hidratação da pele e o corte correto das unhas (Jesus Módulo São Pedro ICMa et al., 2015).

Quando na presença de úlceras, a sua classificação desempenha um papel crucial no tratamento adequado dos pacientes. O sistema de Meggitt-Wagner, como observado na figura 7, é amplamente utilizado como atenção primária, dividindo as úlceras em seis graus de severidade. O grau 0 caracteriza lesões pré-ulceração ou úlceras cicatrizadas, com a presença de deformidades ósseas; o grau 1 refere-se a úlceras superficiais sem envolvimento de tecido subcutâneo; o grau 2 inclui a penetração para além do tecido subcutâneo, com possível exposição de osso, tendão ou ligamento; o grau 3 envolve osteíte, abscesso ou osteomielite; o grau 4 refere-se à gangrena de um dedo, e o grau 5 descreve a gangrena de todo o pé (Nascimento et al., 2019).

Classificação das feridas no pé diabético



Figura 5 - Classificação das feridas no pé diabético.

O tratamento do pé diabético abrange a prevenção e a abordagem de úlceras, dividindo-se em cuidados para pés não ulcerados e tratamento de lesões ulceradas. Para os pés não ulcerados, o tratamento foca na prevenção, que inclui a remoção de calosidades, o tratamento das unhas e a correção de deformidades. Quando ocorrem úlceras, a abordagem varia: o alívio da pressão plantar

e a imobilização do membro são cruciais, assim como o controle da infecção, que pode exigir ação cirúrgica ou não cirúrgica (Ferreira & Carvalho, 2013).

2.2 Machine Learning

Machine Learning (ML) é uma área de inteligência artificial frequentemente associada ao desenvolvimento de ferramentas de diagnóstico. A presente secção pretende explorar a sua utilidade, em conjunto com os detalhes de implementação.

Por definição, ML representa um subconjunto de inteligência artificial (IA) que pretende estudar e compor algoritmos computacionais que simulam a inteligência humana ao aprender repetidamente com o ambiente em que estão inseridos (Mitchell, 2006). Esta aprendizagem, é uma parte da engenharia informática que procura reunir técnicas de matemática e estatística com algoritmos computacionais. A ML utiliza algoritmos baseados em IA e é aplicada em situações específicas onde se procuram padrões num conjunto de variáveis de modo a prever um resultado de interesse específico (Paixão et al., 2022).

2.2.1 Desenvolvimento de um modelo de Machine Learning

O desenvolvimento de um algoritmo de ML pode ser esquematicamente dividido em três fases: fase de pré-processamento, fase de treino e fase de avaliação do modelo. A primeira fase traduz-se na organização da base de dados, na definição das perguntas de investigação e na divisão dos dados para o treino e para o teste (Rajkomar et al., 2019).

Na fase de treino, o algoritmo pode aprender de uma forma supervisionada, ou de uma forma não supervisionada. As aprendizagens supervisionadas e não supervisionada divergem principalmente no algoritmo de treino.

Na aprendizagem supervisionada, o modelo possui conhecimento dos rótulos dos dados e o treino é baseado na comparação entre o resultado do modelo e os rótulos dos dados previamente classificados, de modo a minimizar o erro. O objetivo deste modo de aprendizagem divide-se em dois casos, os problemas de classificação e de regressão. O primeiro atribui os dados de teste a uma categoria específica enquanto o segundo pretende compreender e gerar uma função de relação entre as variáveis interdependentes (Nasteski, 2017).

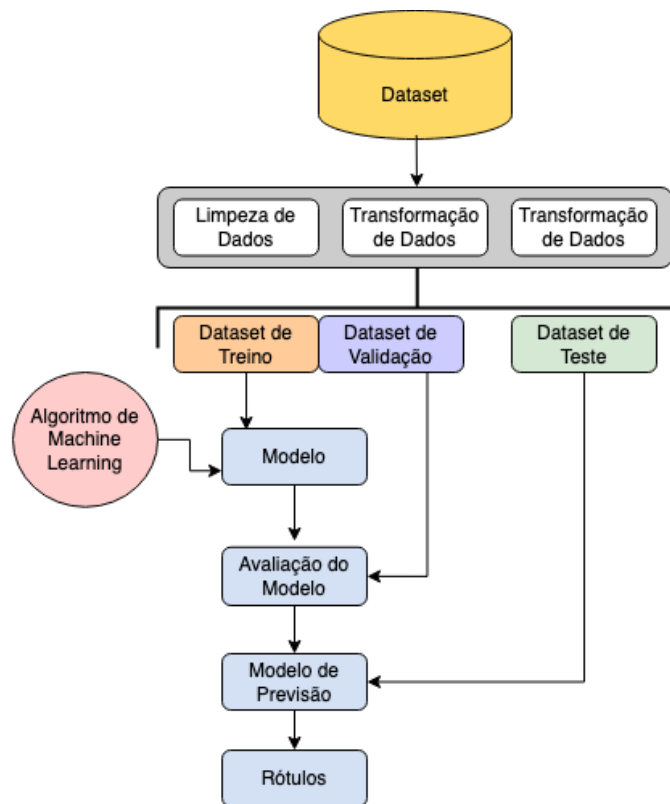


Figura 6 - Flowchart do algoritmo de *machine learning* relativo a aprendizagem supervisionada.

Por outro lado, na aprendizagem não supervisionada, o modelo de ML retira características diretamente dos dados e constrói uma representação sem conhecimento prévio dos rótulos. Tem como objetivo a descoberta autónoma de padrões e funcionalidades através de um conjunto de dados sem rotulação. (Nasteski, 2017).

Na terceira e última fase do desenvolvimento de um algoritmo de ML, na fase de avaliação, o modelo é comparado com os dados de teste e os resultados são gerados. Portanto, os algoritmos de ML aprendem através de repetidas observações e estabelecem um padrão de mapeamento com o intuito de rotular os dados e criar um modelo que generaliza as informações, de modo que novos dados (nunca analisados pelo algoritmo) possam ser rotulados com precisão e confiabilidade. É importante salientar que o processo de desenvolvimento de um algoritmo de ML deve ser realizado com uma base de dados consolidada e validada, caso contrário dados não consolidados podem gerar resultados enganosos (Rajkomar et al., 2019).

2.2.2 Deep Learning

Aprendizagem profunda, ou *deep learning*, é uma subárea de *machine learning* que constrói modelos complexos através de redes neuronais, utilizando várias camadas de processamento que aplicam transformações lineares e não lineares (I. Goodfellow et al., 2016).

Esta aprendizagem difere das metodologias convencionais, uma vez que para além da capacidade de prever os resultados, aprendem automaticamente o processo de classificação. No caso das metodologias tradicionais, há uma limitação na capacidade de processar dados sem pré-processamento ou rotulação, sendo assim menos vantajosas quando comparadas com modelos de *deep learning*.

Redes Neurais Convolucionais, são sistemas computacionais *de deep learning* que conseguem reconhecer e classificar características de imagens. A arquitetura destas redes foi inspirada pela organização e pelas funções do córtex visual, sendo projetada para se assemelhar às conexões entre os neurónios no cérebro humano. Estas operam ao receber uma imagem como entrada e atribuindo diferentes níveis de importância (através de pesos e parâmetros ajustáveis) aos vários objetos presentes na figura 6, permitindo assim distinguir uns dos outros. (Bhatt et al., 2021).

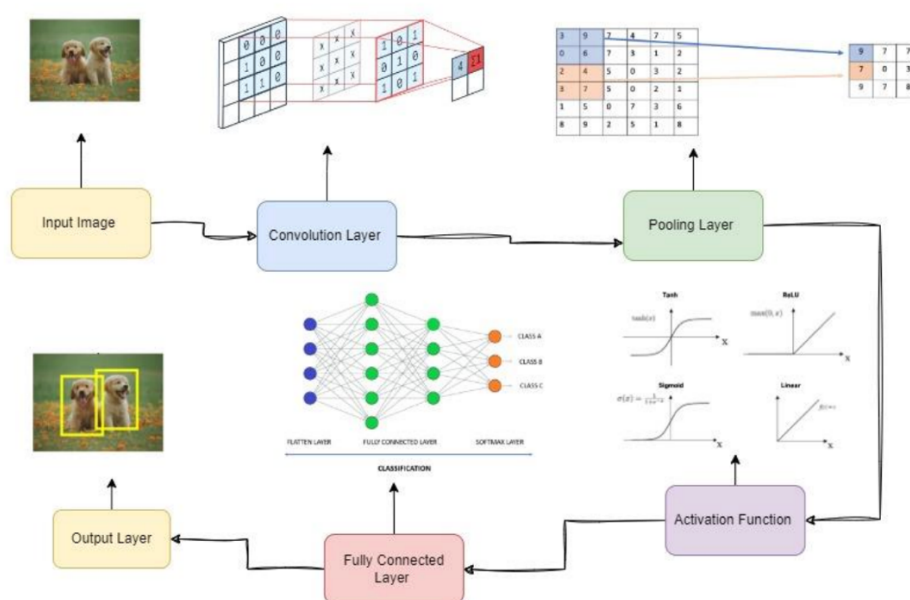


Figura 7- Componentes de uma CNN (Taye, 2023).

Uma imagem digital é formada por pixéis. Estes são blocos de construção que representam a forma binária dos dados visuais. Numa imagem, os pixéis apresentam valores entre 0 e 255 e estão organizados de forma sequencial numa estrutura semelhante a uma matriz. Quando os seres humanos visualizam uma imagem, o cérebro processa uma vasta quantidade de informação na primeira fração de segundo. Cada neurónio no cérebro humano possui um campo recetivo específico e está interligado a outros neurónios, permitindo cobrir a totalidade do campo visual. Este campo recetivo é uma pequena parte do campo visual, onde cada neurónio do sistema visual biológico responde a estímulos (Taye, 2023).

De forma semelhante, cada neurónio numa CNN analisa dados apenas na sua área recetiva. As camadas de uma CNN são programadas para identificar padrões mais simples, como linhas e curvas, antes de avançarem para padrões mais complexos, como rostos e objetos. Como ilustrado na figura

em cima, uma rede neuronal convolucional é constituída por cinco componentes que colaboram para processar as informações visuais, permitindo que a CNN extraia características significativas dos pixels e reconheça padrões complexos nas imagens.

A primeira camada é a camada de convolução, que utiliza filtros (ou *kernels*) para percorrer a imagem de entrada e extrair características importantes, como bordas e texturas. Estes filtros aplicam operações de multiplicação entre os pixels da imagem e os valores do filtro, gerando um mapa de características. À medida que a rede é treinada, os pesos dos filtros são ajustados para reconhecer padrões específicos da imagem.

Após a convolução, os dados passam pela *Pooling Layer*, cuja função é filtrar a dimensionalidade do mapa de características, mantendo apenas as informações mais relevantes. A operação de *pooling* mais comum é o *pooling* máximo, que retém o valor máximo de uma região do mapa de características, reduzindo assim a quantidade de dados e o esforço computacional, mas preservando as informações essenciais para o reconhecimento da imagem.

Em seguida, os dados passam pela função de ativação, como a ReLU (*Rectified Linear Unit*), que introduz uma não linearidade ao modelo. Esta etapa é crucial para que a rede consiga aprender e modelar relações complexas nos dados, permitindo uma melhor distinção entre diferentes características da imagem. A ReLU, por exemplo, transforma valores negativos em zero e mantém os positivos, acelerando o processo de treino.

Por fim, os dados são encaminhados para a camada totalmente conectada, onde o mapa de características é achatado e conectado a todos os neurónios desta camada. Essa camada combina as informações extraídas pelas camadas anteriores para realizar a classificação final da imagem, resultando na previsão da rede sobre a classe à qual a imagem pertence. A combinação dessas camadas permite então que a CNN aprenda e reconheça padrões complexos em imagens (Bhatt et al., 2021).

Estas redes neurais podem ter diferentes arquiteturas, que evoluíram com o avanço da tecnologia nesta área. A VGG-16 e a VGG-19 foram propostas por Simonyan em 2014, com o objetivo de investigar o impacto da profundidade da rede na precisão. Estas redes utilizam várias camadas convolucionais com filtros pequenos de 3x3, em vez dos grandes filtros usados em redes mais antigas. O modelo VGG (*Visual Geometry Group*) tem uma estrutura simples e eficaz, sendo facilmente transferível para outras redes, mas apresenta um grande número de parâmetros, tornando o seu ajuste complexo. A VGG-16 consiste em 16 camadas treináveis, enquanto a VGG-19 tem 19 camadas, o que lhes permite obter bons resultados em tarefas de reconhecimento de imagens, embora a grande quantidade de parâmetros possa causar dificuldades de *overfitting*, ou seja, quando um modelo se ajusta excessivamente aos dados do treino, e exigir mais poder computacional (Simonyan & Zisserman, 2015).

Outro exemplo seria a MobileNetV2. Esta surgiu como uma alternativa mais eficiente em termos de recursos computacionais. Lançada em 2018, esta arquitetura foi projetada para dispositivos móveis e cenários com limitações de processamento. A MobileNetV2 utiliza convoluções separáveis em profundidade, o que reduz significativamente o número de parâmetros e a complexidade

computacional, sem comprometer tanto a precisão. Uma característica central desta rede é o uso de blocos que permitem uma melhor propagação das características através da rede. Devido a essas inovações, a MobileNetV2 é amplamente usada em aplicações que exigem eficiência computacional, como em dispositivos móveis ou sistemas com restrições de hardware (Sandler et al., 2018).

As redes neurais convolucionais têm sido amplamente utilizadas em modelos de detecção de imagens médicas devido à sua capacidade de extrair automaticamente características relevantes de imagens complexas. Estas redes são eficazes na identificação de padrões em imagens médicas, como radiografias, ressonâncias magnéticas e tomografias, auxiliando no diagnóstico de doenças como cancro, lesões e anomalias diversas. CNNs, como VGG, ResNet e MobileNet, são particularmente úteis por conseguirem lidar com grandes volumes de dados e pela sua capacidade de generalizar para novos exemplos, o que melhora a precisão e a rapidez na detecção de patologias. A sua utilização tem permitido avanços significativos na área da radiologia, dermatologia e oftalmologia, entre outras, contribuindo para diagnósticos mais rápidos e precisos.

3 Estado da Arte

Neste capítulo, é apresentada a contextualização dos conceitos teóricos relativos aos modelos de detecção de úlceras no pé diabético, examinando as bases de dados utilizadas e as diferentes abordagens existentes. Além disso, é ilustrado o funcionamento dos modelos, explorando algoritmos e técnicas associadas ao processamento de imagens médicas e à capacidade de identificação precoce de úlceras em pacientes diabéticos.

3.1 Modelos de Classificação

Atualmente, o estudo da classificação de úlceras no pé diabético através de técnicas de *deep learning* (DL) pode englobar a detecção de úlceras e a sua posterior classificação.

Nos últimos anos foram desenvolvidos vários modelos com a finalidade de oferecer soluções automatizadas e precisas para o diagnóstico de úlceras no pé diabético.

Yap et al. (2021) realizou um estudo de classificação de úlceras do pé diabético onde o principal objetivo era a identificação de infecção e isquemia. Como *input*, foram utilizadas 15.683 imagens de pés de pacientes, anotadas com quatro classes: controlo, infecção, isquemia e ambas as condições. A rede neuronal escolhida incluiu vários modelos pré-treinados, como VGG16, ResNet101 e EfficientNet, testados com modelos de *deep learning*. O melhor desempenho foi obtido com a EfficientNetB0, que, após ajustes e aumento de dados, alcançou um *output* com uma precisão de 0.57, *recall* de 0.62 e F1-score de 0.55, destacando-se na classificação de isquemia. Além disso, o estudo aplicou técnicas de aumento de dados para equilibrar as classes menos representadas, como rotação, ruído gaussiano e ajustes de contraste. Isto melhorou a capacidade de generalização do modelo, especialmente para a classe de isquemia. O DenseNet121, sem aumento de dados, atingiu a

melhor precisão de 0.88, enquanto o EfficientNetB0 obteve melhorias quando foi aplicado o aumento de dados. Para compreender melhor as previsões, os autores utilizaram Grad-CAM para visualizar as áreas focadas pelo modelo durante a predição, mostrando que, em alguns casos, o modelo se concentrava incorretamente fora das regiões de úlcera, levando a previsões erradas.

Simonyan & Zisserman (2015) realizaram um estudo sobre reconhecimento de imagens em larga escala, com o objetivo de melhorar a classificação e localização de objetos. Como *input*, utilizaram imagens RGB de 224x224 pixels de um grande conjunto de dados, o ImageNet, composto por milhões de imagens classificadas em 1000 categorias. O estudo testou várias redes neurais profundas, como as redes VGG11, VGG13, VGG16 e VGG19, sendo que a rede VGG19 alcançou o melhor desempenho. Após otimizações e aumento de dados, a rede VGG19 obteve uma precisão de 92,7% e um erro de 7,3%. Aplicaram também técnicas de aumento de dados para melhorar a generalização. A combinação de múltiplos modelos permitiu reduzir o erro para 6,8%. O estudo demonstrou que redes mais profundas com pequenos filtros de 3x3, quando comparadas com arquiteturas mais rasas, proporcionam melhores resultados na classificação de imagens, sendo utilizadas em várias aplicações de visão computacional.

Munadi et al. (2022) realizaram um estudo sobre a classificação de imagens térmicas para a detecção de úlceras do pé diabético (DFU), utilizando um conjunto de dados composto por 2400 imagens. Essas imagens foram obtidas a partir de 142 pacientes diabéticos positivos (com úlceras) e 45 negativos (sem úlceras), com aumento de imagem para diversificação do *dataset*. O estudo recorreu ao uso de redes neurais convolucionais, especificamente MobileNetV2 e ShuffleNet, através de uma estratégia de fusão de decisões que combina as previsões de ambos os modelos. A fusão das decisões resultou num modelo com uma precisão de 1.000 em todas as métricas avaliadas, incluindo exatidão, sensibilidade, precisão e F-score, enquanto mantinha apenas 3,08 milhões de parâmetros de aprendizagem. Este desempenho superou outras abordagens tradicionais, evidenciando a eficácia do modelo proposto na identificação de complicações do pé diabético e sugerindo sua aplicabilidade em ambientes móveis devido à sua eficiência e alta precisão.

Abubakar et al. (2020) propuseram uma abordagem com a utilização da aprendizagem profunda de transferência para a classificação de queimaduras na pele, visando superar as limitações dos métodos tradicionais de avaliação visual, que são subjetivos e dependem de especialistas. O estudo utilizou dados de pacientes caucasianos e africanos, alcançando taxas de reconhecimento de até 99,3 % e 97,1 %, respetivamente. A pesquisa destacou a importância da inclusão de dados de diferentes etnias no treino para evitar viés racial, resultando num modelo mais robusto. Os autores utilizaram um modelo pré-treinado ResNet50, adaptado para discriminar entre pele queimada e saudável, e concluíram que essa abordagem oferece uma alternativa eficaz e acessível para a avaliação de queimaduras, especialmente em regiões com escassez de especialistas.

Em Anisuzzaman et al. (2022) focou-se na classificação de tipos de úlceras utilizando uma abordagem baseada em *deep learning*, destacando o uso do Wound Multimodality Classifier (WMC). Os dados analisados incluíram imagens de feridas e informações de localização, abrangendo classes como úlcera do pé diabético (D), úlcera venosa (V), úlcera de pressão (P), feridas cirúrgicas (S), pele normal (N) e fundo (BG). Para a classificação, foram utilizados modelos VGG16 e VGG19 em combinação com

MLP (Perceptron de Múltiplas Camadas) e LSTM (Memória de Longo Prazo) para lidar com as entradas de imagem e localização. Os melhores resultados foram alcançados com o WMC, demonstrando uma melhoria significativa nas taxas de precisão, recall e F1-score para diferentes classes. Por exemplo, o modelo VGG19 combinado com LSTM atingiu 100% de precisão e *recall* em algumas classificações, enquanto o modelo VGG16 + MLP alcançou um F1-score de até 95,31% em classificações específicas.

A Tabela 1 resume os trabalhos encontrados na literatura em termos de ano de publicação, técnica(s) de classificação utilizada(s), número de imagens utilizadas e os resultados alcançados, que podem ser *Accuracy* (A) e Precisão (P).

Tabela 1 - Sumário dos trabalhos encontrados na literatura.

Referência	Metodologia utilizada	Número de imagens	Performance (%)
(Goyal et al., 2020)	InceptionV3 ResNet50 InceptionResNetV2 SVM	1459	A Isq: 90,00 A Inf: 73,00
(Alzubaidi et al., 2020)	DFU-QUTNet SVM KNN	754	P: 95,40
(Wang et al., 2020)	MobileNetv2 VGG16	1109	F1: 94,05
(Anisuzzaman et al., 2022)	VGG16 VGG19	1088	P: 100
(Zhao et al., 2019)	VGG16	1639	A: 83,36
(Liu et al., 2022)	EfficientNet	58200	A Isq: 99,39 A Inf : 97,92

A tabela resume diferentes trabalhos da literatura que utilizam redes neurais para a classificação de úlceras do pé diabético. A arquitetura VGG16 é a mais utilizada, o que destaca a sua popularidade para esta tarefa. Outros modelos explorados incluem InceptionV3, ResNet50, InceptionResNetV2, MobileNetV2 e EfficientNet.

O número de imagens nos estudos varia bastante, de 754 até 58.200, com uma média de cerca de 1.675 imagens (excluindo o valor mais alto). Em termos de performance, os resultados são reportados em termos de *accuracy*, precisão e F1-score. O modelo EfficientNet apresenta os melhores resultados, com *accuracy* de até 99,39% para isquemia e 97,92% para infecção, enquanto a VGG16 tem um desempenho sólido, variando entre 83,36% e 100% de *accuracy* nos estudos analisados.

3.2 Bases de Dados

De modo a treinar o modelo para que ele desempenhe a sua função, é necessário usar *datasets*. Estes são os dados de treino referidos anteriormente.

Em Yap, Hachiuma, et al. (2021), aborda-se a disponibilidade de três tipos de conjuntos de dados de úlceras no pé diabético fornecidos pela Manchester Metropolitan University e pelo Lancashire Teaching Hospitals. Os conjuntos de dados utilizados incluem imagens de partes do pé para classificação de feridas, regiões de interesse para classificação de infeção e isquemia, e o conjunto de dados mais recente para deteção de úlceras no pé diabético. O conjunto de dados DFUC2020, publicado em abril de 2020, consiste em 2000 imagens de treino, 200 imagens de validação e 2000 imagens de teste, com um total de 2496 úlceras no conjunto de treino e 2097 no conjunto de teste. Além disso, algumas imagens no conjunto de teste não exibem úlceras, visando promover a robustez do modelo. As imagens foram redimensionadas para 640×480 pixéis para melhorar o desempenho dos métodos de *deep learning* e reduzir os custos computacionais. A disponibilidade desses conjuntos de dados visa incentivar a pesquisa e a realização de experiências reproduzíveis no campo da deteção de úlceras no pé diabético.

Em Thotad et al. (2023), vários estudos foram mencionados, incluindo o uso de *deep learning* e redes neurais convolucionais para classificar as imagens em pés saudáveis e pés diabéticos com úlceras. O conjunto de dados consiste num total de 844 imagens antes da aplicação de técnicas de aumento de dados, que aumentaram o número de imagens para 1688. Além disso, o artigo discute a importância de realizar uma compilação de dados devidamente identificados para treinar efetivamente os modelos e a necessidade de técnicas de aumento de dados para melhorar o desempenho do modelo.

Tabela 2 - Resumo dos *datasets* encontrados.

Nome	Imagens disponíveis	Condições de Recolha	Resolução (pixéis)	Open-Access (S/N)
DFUC 2020	4000 (2496 treino e 2.097 teste)	Capturadas por camara digital; Variações em distância, ângulo, orientação, iluminação, focagem e presença de objetos de fundo.	640 X 480	N
DFUC 2021	15683(3994 imagens não identificadas e 5734 imagens para o conjunto de teste)	Três câmeras usadas: Kodak DX4530, Nikon D3300 e Nikon COOLPIX P100.	640 X 480	S
Centro hospitalar diabético de Nasiriyah	754	Utilizado um Samsung Galaxy Note 8 e um iPad		N
Diabetic Foot Ulcer	2673	Captadas por câmara digital	224 x 224	S

O conjunto de dados utilizado neste trabalho consiste em imagens de úlceras do pé diabético capturadas por três tipos de câmeras (Kodak DX4530, Nikon D3300 e Nikon COOLPIX P100) após o desbridamento (remoção de tecidos necróticos e desvitalizados), coletadas ao longo de cinco anos no Lancashire Teaching Hospitals. Este conjunto de imagens, denominado Diabetic Foot Ulcer (DFU), é público e conta com anotações (*ground truth*) fornecidas por dois profissionais de saúde: um médico consultor e um especialista em pés diabéticos. Em casos de divergência de opinião entre os dois, a decisão final era tomada pelo médico mais experiente.

Foi utilizada uma combinação das bases de dados DFUC2021 e Diabetic Foot Ulcer. A base DFU, foca em duas classificações binárias: (1) úlceras com infecção versus úlceras sem infecção e (2) úlceras com isquemia versus úlceras sem isquemia. Além de conter diferentes classes em relação à base DFUC2021, a base DFU inclui imagens geradas através de aumento natural de dados. A classe de úlceras com infecção contém tanto a imagem original da úlcera quanto versões ampliadas da mesma

lesão. Já a classe de úlceras com isquemia inclui a imagem original, além de versões rotacionadas e espelhadas. Portanto, foi necessário um trabalho de adaptação da base DFU para alinhar suas classes às da base DFUC2021.

A base DFUC2021 está dividida em quatro classes: (1) "inexistente", que contém imagens de pele saudável, úlceras em processo de cicatrização ou sem infecção e isquemia; (2) "infecção", contendo imagens de úlceras apenas com infecção; (3) "isquemia", contendo imagens de úlceras apenas com isquemia; e (4) "ambas", que reúne imagens de úlceras que apresentam tanto infecção quanto isquemia.

No total, o modelo foi treinado com uma totalidade de 1039 imagens de pés com úlceras e 1185 imagens de pés com pele saudável.

4 Análise e Desenvolvimento do Modelo

Este trabalho propõe a definição de três modelos de CNN para a identificação de imagens de úlceras do pé diabético. Neste capítulo, foram abordadas as arquiteturas dos três modelos testados e o aumento de dados. Também foram discutidas as técnicas de regularização, como *dropout* e *batch normalization*, e as estratégias de aprendizagem adotadas.

4.1 Ferramentas e Ambientes Utilizados

Para o desenvolvimento deste trabalho, foram selecionadas ferramentas e ambientes que oferecem uma infraestrutura robusta e eficiente para o treino de redes neurais profundas. *TensorFlow* e *Keras* foram as bibliotecas principais utilizadas para a construção, treino e avaliação do modelo. O *TensorFlow*, amplamente reconhecido pela sua capacidade de realizar operações em larga escala, foi escolhido pela sua flexibilidade e capacidade de otimização de modelos em diferentes ambientes de execução, enquanto o *Keras*, sendo uma interface de alto nível para redes neurais, facilitou a implementação de modelos complexos com uma sintaxe simples e intuitiva.

Além disso, o ambiente de execução *Google Colab* foi selecionado como a plataforma ideal para o desenvolvimento das redes implementadas, dado que fornece acesso a recursos computacionais avançados, como GPUs e TPUs, essenciais para o treino intensivo de redes neurais profundas. O *Colab* permite não apenas a execução eficiente de processos de aprendizagem, mas também a integração contínua com bibliotecas como *TensorFlow*, acelerando o ciclo de desenvolvimento. Esta escolha viabilizou a exploração de modelos mais complexos, com a vantagem de uma infraestrutura acessível e adaptada às demandas computacionais do projeto.

Os modelos foram escritos na linguagem *Python*.

A integração com a *pipeline* de dados foi realizada utilizando a classe *ImageDataGenerator* do *Keras*. Este gerador facilitou o carregamento e a manipulação das imagens de úlceras, permitindo que os dados fossem preparados adequadamente para o treino.

```
def prepare_binary_data(split=0.3):
    train_datagen = ImageDataGenerator(
        rescale=1./255,
        rotation_range=40, # rotação
        width_shift_range=0.3, # deslocamento
        height_shift_range=0.3, # deslocamento
        shear_range=0.3, # cisalhamento
        zoom_range=0.3, # zoom
        brightness_range=[0.8, 1.2], # brilho
        horizontal_flip=True,
        fill_mode='nearest',
        validation_split=split # 30% dos dados para validação
    )
```

Código 1 – Extrato da função *prepare_binary_data* para o modelo da MobileNetV2.

O uso do *ImageDataGenerator* não apenas diversificou as amostras de treino, mas também ajudou a prevenir o *overfitting*, garantindo que o modelo tivesse uma base sólida para aprender a identificar características relevantes das úlceras do pé diabético.

4.2 Modelos utilizados

Devido à popularidade e bom desempenho, optou-se pelo uso destas três arquiteturas: VGG16, VGG19 e MobileNetV2.

Os modelos VGG-16 e VGG-19 são arquiteturas da rede VGGNet, desenvolvidas por Simonyan e Zisserman em 2015, que se destacam pela sua abordagem inovadora no uso de múltiplos filtros convolucionais de pequenas dimensões, especificamente 3x3, promovendo uma representação detalhada das características nas imagens. A arquitetura foi projetada para aumentar a profundidade da rede sem aumentar significativamente o número de parâmetros, e o seu sucesso está no uso eficiente de camadas convolucionais seguidas de *pooling* para captura de características progressivamente mais abstratas (Simonyan & Zisserman, 2015).

A VGG-16 é composta por 16 camadas treináveis, incluindo 13 camadas convolucionais organizadas em cinco blocos, intercaladas com camadas de *MaxPooling*, que reduzem progressivamente as dimensões espaciais das representações. As últimas camadas são formadas por duas camadas densas com 4.096 unidades cada, seguidas por uma camada de saída com a função de ativação *softmax*. Essa arquitetura é amplamente utilizada em tarefas de classificação de imagens devido à sua simplicidade e eficácia (Simonyan & Zisserman, 2015).

Por outro lado, a VGG-19 adiciona três camadas convolucionais extras, totalizando 19 camadas treináveis, o que aumenta a capacidade da rede de capturar detalhes mais finos em imagens

complexas. Essa arquitetura também mantém a organização em blocos de convolução e *pooling*, mas a profundidade extra torna-a ainda mais adequada para tarefas em que detalhes visuais são críticos, como o caso da classificação de úlceras do pé diabético.

Ambas as redes requerem uma quantidade significativa de capacidade computacional devido ao grande número de parâmetros: aproximadamente 138 milhões na VGG-16 e 143 milhões na VGG-19. Esta alta complexidade faz com que seja necessário o uso de técnicas de regularização para mitigar o risco de *overfitting*, especialmente em conjuntos de dados de menor quantidade.

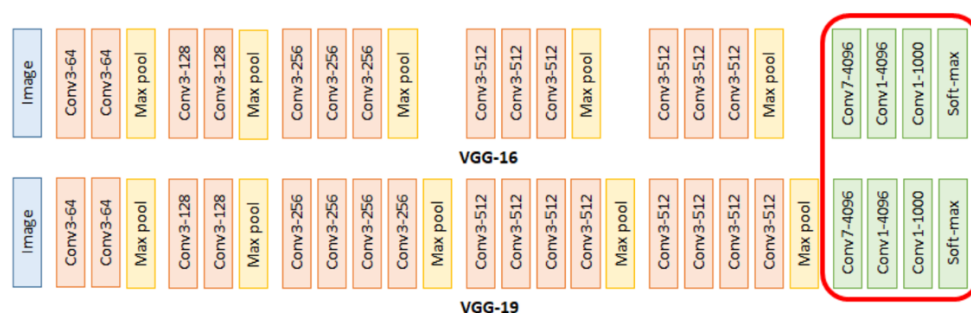


Figura 8 - Arquiteturas das Redes VGG16 e VGG19 (Shadeed et al., 2020).

Pelo contrário, a arquitetura MobileNetV2, foi projetada com foco em dispositivos móveis. Esta procura minimizar o uso de memória e operações de convolução sem comprometer a precisão do modelo. A principal inovação desta CNN reside no uso de um módulo que inverte o padrão das camadas residuais tradicionais, o que permite reduzir a quantidade de cálculos e de memória, mantendo a precisão do modelo. Esta abordagem processa a imagem numa representação compacta de baixa dimensão, que é expandida para uma dimensionalidade maior e convoluída de maneira leve, seguida de uma projeção para uma baixa dimensão. Essa eficiência estrutural permite que a MobileNetV2 seja frequentemente escolhida para implementação em dispositivos com recursos limitados, mantendo alta precisão, conforme descrito por Sandler et al. (2018).

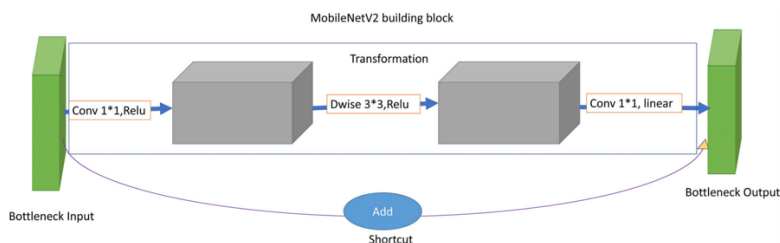


Figura 9 - Arquiteturas da Rede MobileNetV2 (Al-gaashani et al., 2022).

Ambas as arquiteturas (VGG e MobileNetV2) foram pré-treinadas no conjunto de dados ImageNet e posteriormente ajustadas para a tarefa específica de classificação binária de úlceras diabéticas. Esta metodologia é conhecida como "*transfer learning*", que consiste em aproveitar o conhecimento adquirido numa tarefa anterior (neste caso, a classificação de imagens no ImageNet) para acelerar e

melhorar o desempenho numa nova tarefa. Essa abordagem traz vantagens significativas em termos de performance, pois permite que os modelos comecem o treino já com representações robustas, exigindo menos dados e tempo para alcançar bons resultados na nova tarefa.

A escolha dessas arquiteturas baseia-se tanto na eficiência de extração de características visuais importantes para a classificação de imagens médicas quanto no balanceamento entre capacidade computacional disponível e desempenho esperado.

4.3 Preparação e Pré-processamento de Dados

A preparação e o pré-processamento de dados são etapas cruciais em qualquer *pipeline* de *deep learning*, influenciando diretamente o desempenho e a capacidade de generalização do modelo. Neste trabalho, utilizamos um conjunto de dados que continha imagens de úlceras do pé diabético, que foram divididas em conjuntos de treino e validação. Optou-se por uma divisão de 70% para o conjunto de treino e 30% para o conjunto de validação. Esta escolha foi feita para garantir que o modelo tivesse um conjunto substancial de dados para aprender, enquanto ainda permitia uma avaliação eficaz do desempenho do modelo através do conjunto de validação. A validação é essencial para ajustar hiperparâmetros e evitar *overfitting*. Embora não tenhamos utilizado um conjunto de teste separado, a avaliação no conjunto de validação fornece uma indicação do desempenho do modelo em dados não vistos durante o treino.

A fim de assegurar que os modelos fossem capazes de lidar com a diversidade visual dessas imagens, implementamos uma série de técnicas de aumento de dados e normalização.

A função de preparação de dados foi implementada separadamente para os modelos VGG e MobileNetV2, com algumas diferenças nas técnicas de aumento aplicadas, refletindo as características distintas de cada arquitetura.

Para os modelos VGG16 e VGG19, o aumento de dados foi moderado, dando ênfase em transformações que preservam as características visuais gerais das úlceras. As técnicas aplicadas incluíram:

- Rotação das Imagens: Variando os ângulos de rotação em até 30 graus, o que ajuda o modelo a reconhecer úlceras independentemente da orientação da imagem.
- Deslocamento Horizontal e Vertical (*Width e Height Shifts*): Simulando pequenas variações de posição do objeto de interesse na imagem, aumentando a robustez do modelo em relação à localização da úlcera.
- Zoom: Aplicado com um intervalo de até 50%, permitindo que o modelo concentrasse tanto em detalhes menores quanto em aspetos mais amplos das imagens.

- Modificação do Brilho: O brilho das imagens foi ajustado num intervalo de 20% para cima ou para baixo, o que simula diferentes condições de iluminação e torna o modelo mais robusto a essas variações.
- *Shear* (Cisalhamento): Introduzindo pequenas distorções geométricas, o que força o modelo a aprender variações mais complexas nas formas das úlceras.

No extrato de código 2, encontra-se o extrato de código em *Python* da função usado no aumento de dados para as referidas redes:

```
def prepare_binary_data():
    train_datagen = ImageDataGenerator(
        rescale=1./255,
        rotation_range=30, # rotação
        width_shift_range=0.2, # translation
        height_shift_range=0.2,
        shear_range=0.2,
        zoom_range=0.5, # de zoom
        brightness_range=[0.8, 1.2], # brilho
        channel_shift_range=30.0, # cor
        horizontal_flip=True,
        fill_mode='nearest',
        validation_split=0.3
    )
```

Código 2 – Extrato da função *prepare_binary_data* para os modelos da VGG16 e VGG19.

Para o MobileNetV2, as transformações de aumento de dados foram aplicadas de forma mais agressiva, refletindo a necessidade de maior variabilidade em modelos mais leves e profundos:

- Rotação: A rotação foi aumentada para 40 graus, ampliando a capacidade do modelo de lidar com variações de orientação mais extremas.
- Deslocamento Horizontal e Vertical: Aumentados para 30%, permitindo variações mais significativas na posição das úlceras.
- Zoom e Cisalhamento: Ajustados em 30%, o que força o modelo a se adaptar a maiores variações de escala e forma das úlceras.
- Modificação do Brilho: Mantida igual à da VGG16, ajustando o brilho em 20% para garantir a robustez em condições de iluminação variadas.

Em baixo encontra-se o extrato de código em *Python* da função usado no aumento de dados para a referida rede:

```
def prepare_binary_data(split=0.3):
    train_datagen = ImageDataGenerator(
```

```

        rescale=1./255,
        rotation_range=40, # rotação
        width_shift_range=0.3, # deslocamento
        height_shift_range=0.3, # deslocamento
        shear_range=0.3, # cisalhamento
        zoom_range=0.3, # zoom
        brightness_range=[0.8, 1.2], # brilho
        horizontal_flip=True,
        fill_mode='nearest',
        validation_split=split # 30% dos dados para validação
    )

```

Código 3 – Extrato da função *prepare_binary_data* para o modelo da MobileNetV2.

Além das técnicas de aumento de dados, a normalização dos pixels foi aplicada de maneira semelhante para ambos os modelos. Utilizamos o parâmetro `rescale=1./255` (presente nos extratos de Código 1 e Código 2) que converte os valores de intensidade dos pixels de uma escala de 0 a 255 para uma escala de 0 a 1. Esse processo é fundamental para garantir que o modelo não seja influenciado por variações absolutas de intensidade nas imagens, tornando o treino mais estável e eficaz.

4.4 Configuração da Arquitetura dos Modelos

Especialmente para tarefas complexas como a identificação de úlceras do pé diabético, é comum utilizar redes neurais pré-treinadas em grandes conjuntos de dados, como o *ImageNet*, já referido anteriormente. No entanto, há a necessidade de ajustarmos estas redes pré-treinadas ao novo problema. Para isso são aplicadas técnicas como o *fine-tuning* (ajuste fino) e a *Batch Normalization*, ambas fundamentais para adaptar as redes a um conjunto de dados menor e mais específico, evitando problemas como *overfitting* e instabilidade durante o treino.

O *fine-tuning* é uma técnica derivada do *transfer learning*, em que um modelo já treinado num conjunto de dados grande (neste caso, o *ImageNet*) é adaptado para um novo conjunto de dados, mais específico, mantendo as características gerais aprendidas anteriormente. Durante o *fine-tuning*, partes do modelo original são congeladas enquanto outras partes são ajustadas para o novo conjunto de dados. Na prática, as camadas mais baixas da rede são responsáveis por capturar características visuais genéricas, como bordas e texturas. Estas características são úteis em quase qualquer tarefa de visão computacional e, portanto, podem ser mantidas, ou seja, serão as camadas denominadas congeladas. Já as camadas mais altas da rede aprendem padrões mais complexos e específicos do conjunto de dados original e precisam ser ajustadas (camadas descongeladas) para se adaptarem às características das úlceras do pé diabético.

No caso da VGG16 e VGG19 as primeiras oito camadas foram congeladas, preservando as características genéricas que o modelo já aprendeu. As camadas a partir da oitava camada foram descongeladas, permitindo que o modelo se ajustasse aos padrões específicos das úlceras. Dessa

forma, evitou-se a necessidade de treinar todas as camadas desde o início, o que seria computacionalmente mais caro e poderia resultar em *overfitting*.

No extrato de código abaixo, realiza-se então este ajuste:

```
for layer in base_model.layers[:8]: # mais camadas
    layer.trainable = False # Manter as primeiras camadas congeladas
for layer in base_model.layers[8:]:
    layer.trainable = True # Descongelar as camadas superiores para ajuste fino
```

Código 4 – Extrato da aplicação do fine-tuning nos modelos VGG16 e VGG19.

No caso da MobileNetV2, por ser uma rede mais leve e compacta, utiliza uma abordagem diferente. Todas as camadas da MobileNetV2 foram descongeladas para permitir um ajuste fino completo, uma vez que as características aprendidas pelo modelo no *ImageNet* podem não ser suficientes para captar as nuances mais subtis das úlceras. Como a MobileNetV2 tem menos camadas e é otimizada para eficiência computacional, descongelar todas as camadas permite que o modelo se adapte totalmente ao novo conjunto de dados, sem o risco de se tornar muito pesado.

```
for layer in base_model.layers:
    layer.trainable = True
```

Código 5 – Extrato da aplicação do fine-tuning no modelo para a MobileNetV2.

A *Batch Normalization* foi introduzida por Sergey Ioffe e Christian Szegedy em 2015 e tornou-se numa técnica padrão em redes neuronais profundas. Tem como principal objetivo estabilizar e acelerar o treino, reduzindo o problema da variação nas ativações ao longo das camadas de uma rede neural. Na prática, a *Batch Normalization* normaliza as ativações de uma camada de acordo com a média e o desvio padrão da mini-batch, ajustando estes valores de modo que a rede tenha ativações mais estáveis e bem escaladas (Ioffe & Szegedy, 2022)

Esta técnica resolve um problema comum em redes neurais: a explosão ou desaparecimento dos gradientes, que ocorre quando os valores de ativação nas camadas tornam-se muito grandes ou muito pequenos, dificultando a aprendizagem do modelo. Ao garantir que os valores estejam sempre controlados, a *Batch Normalization* acelera a convergência do modelo e permite o uso de taxas de aprendizagem mais altas, reduzindo a necessidade de um ajuste manual detalhado de hiperparâmetros.

No modelo das redes VGG, a *Batch Normalization* foi aplicada após a saída da base convolucional, estabilizando as ativações antes de passar para as camadas densas. Isso foi especialmente útil para descongelar as camadas superiores durante o *fine-tuning*, pois normalizou as ativações e impede que

mudanças nas camadas superiores causem grandes flutuações no treino. Esta ação pode ser vista no extrato de código apresentando a seguir:

```
def build_model():  
    base_model = VGG16(weights='imagenet', include_top=False,  
input_shape=(IMAGE_HEIGHT, IMAGE_WIDTH, 3))  
    x = base_model.output  
    x = BatchNormalization()(x)  
    x = GlobalAveragePooling2D()(x)
```

Código 6 – Extrato da aplicação do *Batch Normalization* no modelo para a VGG16 e VGG19.

No caso da MobileNetV2, que possui uma arquitetura diferente baseada em convoluções mais leves e eficientes, a *Batch Normalization* também foi aplicada logo após a saída da base convolucional. Isso assegura que o modelo mantenha a estabilidade durante o treino, especialmente ao descongelar todas as camadas, onde o risco de instabilidade nos gradientes é maior. Esta abordagem é ilustrada no extrato de código apresentado a seguir.

```
def build_model():  
    base_model = MobileNetV2(weights='imagenet', include_top=False,  
input_shape=(IMAGE_HEIGHT, IMAGE_WIDTH, 3))  
    x = base_model.output  
    x = BatchNormalization()(x)  
    x = GlobalAveragePooling2D()(x)
```

Código 7 – Extrato da aplicação do *Batch Normalization* no modelo para a MobileNetV2.

O uso de *Batch Normalization*, em ambos os casos, desempenhou um papel fundamental para garantir um treino mais eficiente e estável, prevenindo problemas de instabilidade que poderiam surgir com o ajuste fino. A aplicação do *Batch Normalization* provém da biblioteca *Keras*, que será abordada com maior profundidade mais à frente neste trabalho.

4.5 Treino e Validação

O processo de treino e validação é uma fase crítica no desenvolvimento destes modelos, pois determina a capacidade do modelo de caracterizar os novos dados. Para otimizar esse processo, foram implementadas diversas técnicas e estratégias, incluindo a definição de uma taxa de aprendizagem apropriada (*learning rate*), a utilização de *callbacks* para monitorizar o desempenho

do modelo durante o treino e a aplicação de pesos de classe para lidar com a desbalanceamento no conjunto de dados.

A taxa de aprendizagem (*learning rate*) é um hiperparâmetro fundamental que controla a magnitude das atualizações dos pesos do modelo durante o treino. Uma taxa de aprendizagem adequada é crucial para garantir uma convergência eficiente e eficaz. Se a taxa de aprendizagem for muito alta, o modelo pode fazer atualizações excessivas nos pesos, o que resulta em instabilidade e dificuldade em encontrar a melhor solução. Por outro lado, se for muito baixa, o treino será lento e o modelo pode encontrar uma solução que não é a melhor possível, conhecida como "mínimo local".

Para ambos os modelos, VGG16, VGG19 e MobileNetV2, foi utilizada uma taxa de aprendizagem de $1e-5$, que foi escolhida após várias tentativas de ajuste para encontrar a melhor configuração. Todos os modelos foram configurados para usar o otimizador Adam, conhecido por sua eficiência e adaptabilidade. O otimizador Adam é um algoritmo de otimização que ajusta automaticamente as taxas de aprendizagem para cada parâmetro, o que permite uma convergência mais rápida e estável durante o treino do modelo.

Outra estratégia utilizada foi o uso de *callbacks*. Estes são funções que permitem realizar ações em determinados pontos durante o treino. São essenciais para monitorar o desempenho do modelo e realizar ajustes dinâmicos, melhorando a eficiência do processo de treino.

- *Early Stopping*: Esta função de *callback* é utilizada para interromper o treino quando a métrica de validação não apresenta melhorias após um certo número de épocas, conhecido como paciência. Este método garante que o modelo não continue a treinar desnecessariamente após atingir o seu melhor desempenho. Foi configurado com uma paciência de 10 épocas para ambos os modelos. A aplicação desta função pode ser verificada no seguinte extrato de código:

```
early_stopping = EarlyStopping(monitor='val_loss', patience=10, restore_best_weights=True)
```

Código 8 – Extrato da aplicação do *Early Stopping* nos modelos.

- *Reduce Learning Rate on Plateau*: Esta função de *callback* reduz a taxa de aprendizagem quando a métrica de validação para de melhorar, permitindo que o modelo faça ajustes mais finos à medida que se aproxima da convergência. Esta abordagem é especialmente útil para evitar oscilações nas últimas fases do treino e pode ser verificada no extrato de código em baixo:

```
reduce_lr = ReduceLROnPlateau(monitor='val_loss', factor=0.2, patience=5, min_lr=1e-6)
```

Código 9 – Extrato da aplicação de *Reduce Learning Rate on Plateau* nos modelos.

- *Model Checkpoint*: Esta função guarda o modelo durante o treino sempre que a métrica de validação (neste caso, a *loss*) melhora. Isso garante que a melhor versão do modelo seja guardada. A aplicação desta função pode ser verificada no seguinte extrato de código:

```
checkpoint = ModelCheckpoint('best_binary_model_mobilenetv2.keras', monitor='val_loss',
save_best_only=True, mode='min', verbose=1)
```

Código 10 – Exemplo de extrato do código da aplicação do *Model Checkpoint* no modelo da rede MobileNetV2.

Quando falamos de conjuntos de dados desbalanceados, ou seja, onde uma classe pode ter significativamente mais exemplos do que outra, a utilização de pesos de classe torna-se uma estratégia eficaz para ajudar o modelo a aprender de forma mais equilibrada. Neste trabalho, após observar que os modelos VGG16 e VGG19 estavam a tornar-se tendenciosos em relação à classe "normal" (sem úlcera) durante os treinos, foi decidido ajustar os pesos de classe. A classe "normal" foi configurada com um peso maior, enquanto a classe "com úlcera" recebeu um peso menor. Para isso, utilizou-se a função de *Class Weights*, que permitiu configurar um dicionário de pesos de classe, onde a classe "normal" recebeu um peso de 2.0 e a classe 'com úlcera' um peso de 1.0. Adicionar pesos a uma classe significa que a rede dá mais importância a essa classe durante o treino, ajudando a corrigir o desbalanceamento entre as classes. Isso indica à rede que a classe "normal" é menos comum e, portanto, precisa de mais atenção na atualização dos pesos, resultando num modelo que se torna mais sensível a essa classe e melhora a sua capacidade de classificação. Esta abordagem assegura que os modelos aprendam a reconhecer adequadamente ambas as classes, mitigando o viés em relação à classe mais prevalente. Por outro lado, no modelo MobileNetV2, não foram utilizados pesos de classe, pois a arquitetura leve e adaptativa foi suficiente para lidar com o desbalanceamento do conjunto de dados, permitindo que o modelo se ajustasse sem necessidade de compensação adicional. Esta abordagem é ilustrada no extrato de código apresentado a seguir:

```
class_weight = {0: 2.0, 1: 1.0}
```

Código 11 – Extrato de código da aplicação de *Class Weights* nos modelos das redes VGG16 e VGG19.

4.6 Análise com Grad-CAM

A técnica Grad-CAM (*Gradient-Weighted Class Activation Mapping*) é uma ferramenta eficaz para interpretar e visualizar as decisões tomadas pelos modelos, especialmente em redes neurais convolucionais. Esta permite identificar quais partes de uma imagem influenciaram mais a decisão do modelo em tarefas de classificação, oferecendo uma compreensão mais profunda sobre como o modelo está a ver e analisar os dados.

O Grad-CAM utiliza os gradientes da camada de saída em relação às ativações de uma camada convolucional específica para criar um mapa de calor que destaca as regiões mais relevantes da imagem. O processo inicia com o cálculo dos gradientes durante a retropropagação, onde são calculados os gradientes da saída do modelo em relação às ativações da camada convolucional escolhida. Em seguida, os gradientes são agrupados, geralmente pela média, para obter pesos que indicam a importância de cada canal da camada convolucional. Com estes pesos, as ativações da camada convolucional são ponderadas, resultando num mapa de calor que destaca as regiões da imagem que mais contribuíram para a decisão do modelo. Por fim, esse mapa de calor é sobreposto à imagem original, permitindo visualizar claramente as áreas que foram mais significativas para a classificação.

O primeiro passo para utilizarmos o Grad-Cam é carregar os modelos guardados anteriormente com a função *Model Checkpoint*, como podemos observar no extrato de código a seguir:

```
model = load_model('best_binary_model_mobilenetv2.keras')
model = load_model('best_binary_model_vgg16.keras')
model = load_model('best_binary_model_vgg19.keras')
```

Código 12 – Extratos de código do carregamento dos modelos guardados anteriormente para a utilização no Grad-Cam.

Numa segunda fase, procedeu-se à identificação da última camada de cada modelo, sendo nesta onde o Grad-CAM será utilizado. Esta foi implementada da seguinte forma:

```
last_conv_layer_name = 'block5_conv3' # VGG19
last_conv_layer_name = 'block5_conv3' # VGG16
last_conv_layer_mobilenet = 'Conv_1' # MobileNetV2
```

Código 13 – Extratos de código da identificação das camadas para a utilização do Grad-Cam.

Posteriormente, a função *generate_gradcam_single* é responsável por aplicar o Grad-CAM a uma única imagem. Primeiro, a imagem é carregada e pré-processada para adequar-se à entrada do modelo, utilizando a função de pré-processamento apropriada. Em seguida, um modelo gradiente é criado para mapear as ativações da última camada convolucional e as previsões. O *tf.GradientTape* é utilizado para calcular os gradientes da previsão em relação às ativações, permitindo a identificação das áreas com maior influência na classificação. Os gradientes são então agrupados para obter os pesos que refletem a importância de cada canal da camada convolucional. A média das ativações ponderadas gera o mapa de calor, que é normalizado e redimensionado para coincidir com as dimensões da imagem original. Este mapa de calor é uma representação visual que destaca as regiões da imagem que mais contribuíram para a decisão do modelo, facilitando a interpretação de como a rede neuronal chegou à sua conclusão. Por fim, o mapa de calor é sobreposto à imagem original,

criando uma visualização que destaca as regiões significativas para a decisão do modelo. Esta função foi aplicada de forma semelhante para diferentes modelos, como VGG16, VGG19 e MobileNetV2, cada um utilizando a camada convolucional correspondente, como referido anteriormente. Esta ação pode ser vista no extrato de código apresentando a seguir:

```
def generate_gradcam_single(img_path, model, last_conv_layer_name, pred_index=None):
    # Carregar a imagem e pré-processá-la
    img = image.load_img(img_path, target_size=(IMAGE_WIDTH, IMAGE_HEIGHT))
    img_array = image.img_to_array(img)
    img_array = np.expand_dims(img_array, axis=0)
    img_array = preprocess_input(img_array)

    # modelo que mapeia a imagem com a ativação da última camada convolucional e a
    # previsão
    grad_model = tf.keras.models.Model(
        [model.inputs], [model.get_layer(last_conv_layer_name).output, model.output]
    )

    with tf.GradientTape() as tape:
        conv_outputs, predictions = grad_model(img_array)

        if pred_index is None:
            pred_index = tf.argmax(predictions[0]).numpy().item()

        class_channel = predictions[0][pred_index]

    grads = tape.gradient(class_channel, conv_outputs)

    # Calcular a média dos gradientes para cada filtro na última camada convolucional
    pooled_grads = tf.reduce_mean(grads, axis=(0, 1, 2))

    # Multiplicar cada mapa de ativação pelo peso médio dos gradientes
    conv_outputs = conv_outputs[0].numpy() # Converter para NumPy
    pooled_grads = pooled_grads.numpy()
    for i in range(pooled_grads.shape[-1]):
        conv_outputs[:, :, i] *= pooled_grads[i]

    # Calcular o mapa de ativação da classe e redimensioná-lo para o tamanho da
    # imagem de entrada
    heatmap = np.mean(conv_outputs, axis=-1)
    heatmap = np.maximum(heatmap, 0) / np.max(heatmap)
    heatmap = cv2.resize(heatmap, (img.size[0], img.size[1]))

    # Aplicar o Grad-CAM sobre a imagem original
    img = cv2.imread(img_path)
    heatmap = np.uint8(255 * heatmap)
    heatmap = cv2.applyColorMap(heatmap, cv2.COLORMAP_JET)

    # Combinar a imagem original com o heatmap
    superimposed_img = cv2.addWeighted(img, 0.6, heatmap, 0.4, 0) # Usar addWeighted
    # para combinar

    return img, superimposed_img
```

Código 14 – Extratos de código com a aplicação do Grad-Cam em todos os modelos utilizados.

Por fim, após gerar o Grad-CAM, é importante então visualizar os resultados obtidos. A função *generate_gradcam_multiple* pode ser utilizada para aplicar o Grad-CAM em várias imagens e exibir os resultados lado a lado, sendo aplicado o seguinte extrato de código:

```
def generate_gradcam_multiple(img_paths, model, last_conv_layer_name, num_images=9):
    plt.figure(figsize=(15, 15))

    for i, img_path in enumerate(img_paths[:num_images]):
        original_img, gradcam_img = generate_gradcam_single(img_path, model,
last_conv_layer_name)

        # Mostrar imagem original
        plt.subplot(3, 6, 2 * i + 1)
        plt.imshow(cv2.cvtColor(original_img, cv2.COLOR_BGR2RGB))
        plt.title("Original")
        plt.axis('off')

        # Mostrar imagem com Grad-CAM
        plt.subplot(3, 6, 2 * i + 2)
        plt.imshow(cv2.cvtColor(gradcam_img, cv2.COLOR_BGR2RGB))
        plt.title("Grad-CAM")
        plt.axis('off')

    plt.tight_layout()
    plt.show()
```

Código 15 – Extratos de código para a visualização dos resultados obtidos com Grad-Cam.

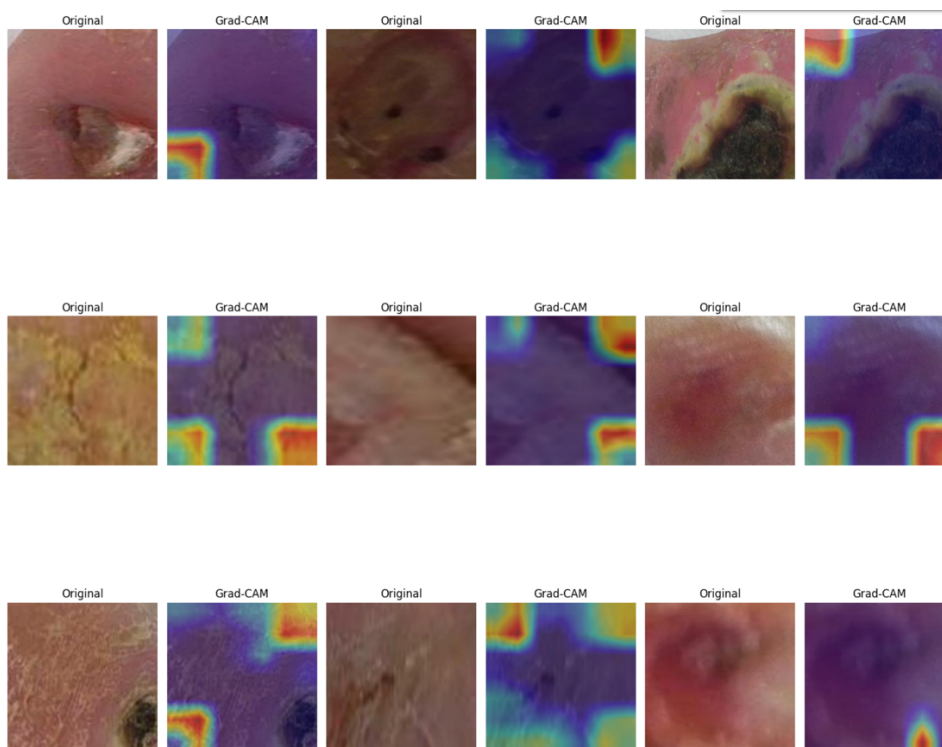


Figura 10 - Exemplo de grelha de imagens obtidas com Grad-Cam.

5 Resultados e Discussão

Neste capítulo, abordamos as ferramentas e os ambientes utilizados para o desenvolvimento do modelo, explicamos a integração com a *pipeline* de dados e detalhamos o processo de exportação, testes e avaliação do modelo. Além disso, discutimos possíveis melhorias e trabalhos futuros que poderão otimizar ainda mais a aplicação clínica do modelo.

Para avaliar os modelos VGG16, VGG19 e MobileNetV2, foi utilizado um conjunto de dados separado para validação, que não foi utilizado durante o treino. Isto garantiu que as métricas de desempenho sejam calculadas com base em dados que os modelos não tinham tido contacto antes, proporcionando uma avaliação mais precisa das suas capacidades. Os modelos foram testados num número significativo de imagens de validação, e as previsões foram comparadas com as classes reais para calcular as métricas.

A *accuracy*, foi uma das métricas avaliadas. Esta é uma das principais métricas utilizadas para avaliar o desempenho dos modelos de classificação. Para os modelos VGG16, VGG19 e MobileNetV2, a *accuracy* foi calculada utilizando a função *evaluate* do *Keras*, que fornece a proporção de previsões corretas em relação ao total de previsões realizadas. Esta função usa o *validation_generator* que contém as imagens de validação, e calcula a perda (*val_loss*) e a *accuracy* (*val_accuracy*) do modelo nesse conjunto.

```
def evaluate_model(model, validation_generator):  
    val_loss, val_acc = model.evaluate(validation_generator)  
    print(f"Validation Loss: {val_loss}, Validation Accuracy: {val_acc}")
```

Código 16 – Extrato da função *evaluate_model* para o modelo da MobileNetV2.

A *accuracy* foi calculada da seguinte forma:

$$Accuracy = \frac{\text{Número de Previsões Corretas}}{\text{Número Total de Previsões}} \quad (1)$$

Em termos do modelo do VGG16, a *accuracy* alcançada foi de 92%, indicando que o modelo classificou corretamente 92% das imagens de úlceras do pé diabético no conjunto de validação. O modelo VGG19 apresentou uma *accuracy* maior, alcançando 93%. Este resultado sugere que o VGG19, tendo uma arquitetura mais profunda, foi capaz de identificar padrões relevantes nas imagens com maior eficácia. Apesar de ser otimizado para eficiência e uso em dispositivos móveis, como referido anteriormente, o MobileNetV2 obteve uma *accuracy* de 81%. Isso demonstra que, apesar da sua leveza, o modelo ainda conseguiu realizar uma classificação eficaz, embora com um desempenho inferior em comparação com os modelos VGG.

Estes resultados ressaltam a eficácia dos modelos em identificar corretamente as características das úlceras, com o VGG19 destacando-se como o mais robusto entre eles. A *accuracy*, como métrica, oferece uma visão inicial do desempenho dos modelos, servindo como um ponto de partida para análises mais detalhadas e a exploração de outras métricas de desempenho, como precisão, *recall* e *F1-score*, que fornecem uma compreensão mais abrangente do comportamento dos modelos em diferentes condições.

Uma segunda métrica calculada foi a precisão. A precisão é uma métrica crucial na avaliação de modelos de classificação, especialmente em contextos médicos onde o peso de falsos positivos pode ser significativo. Um falso positivo neste contexto seria, por exemplo, o modelo classificar uma imagem como se tivesse úlceras presentes mas na verdade não ter. Esta mede a proporção de verdadeiros positivos em relação ao total de casos classificados como positivos, fornecendo uma indicação de quão confiáveis são as previsões positivas feitas pelo modelo.

$$Precisão = \frac{VP \text{ (Verdadeiros Positivos)}}{VP \text{ (Verdadeiros Positivos)} + FP \text{ (Falsos Positivos)}} \quad (2)$$

Num cenário de classificação de úlceras do pé diabético, é essencial minimizar os falsos positivos, pois pode levar a diagnósticos incorretos e a tratamentos desnecessários para pacientes. Portanto, uma alta precisão indica que quando o modelo classifica uma imagem como "abnormal", ou seja, com úlcera, há então uma alta probabilidade de que a imagem realmente apresente uma úlcera.

Os resultados de precisão para cada modelo foram analisados separadamente para as classes "abnormal" (com úlcera) e "normal" (sem úlcera). O modelo MobileNetV2 obteve uma precisão de 82% para a classe "abnormal", o que indica uma boa probabilidade de que as classificações de úlceras estejam corretas, enquanto a precisão de 79% para a classe "normal" sugere que ele tende a classificar incorretamente algumas imagens normais como anormais. O VGG16 apresentou uma precisão equilibrada de 92% para ambas as classes, demonstrando um desempenho consistente e confiável em todas as classificações. O VGG19, por sua vez, destacou-se com uma precisão de 97%

para a classe "abnormal", mostrando uma excelente capacidade na identificação de imagens com úlceras, embora a precisão de 90% para a classe "normal" indique que o modelo pode ainda classificar algumas imagens normais de forma incorreta.

Recall, também conhecido como sensibilidade ou taxa de acerto, é uma métrica que mede a capacidade do modelo de identificar corretamente os verdadeiros positivos (ou seja, os casos reais de uma determinada classe). Em termos simples, o *recall* avalia a proporção de instâncias positivas que foram corretamente classificadas pelo modelo.

$$Recall = \frac{VP (Verdadeiros Positivos)}{VP (Verdadeiros Positivos) + FN(Falsos Negativos)} \quad (3)$$

Um valor alto de *recall* indica que o modelo está a identificar a maior parte dos casos positivos corretamente, ou seja, há poucos falsos negativos. Neste caso um falso negativo seria o modelo considerar uma imagem como normal mas na verdade essa imagem ter uma úlcera num pé diabético. Num cenário onde um modelo usa dados clínicos, como a detecção de úlceras, o objetivo é ter um alto *recall*, pois significa que o modelo está a identificar a maioria dos casos de úlceras corretamente. No entanto, uma percentagem de *recall* alta pode vir às custas de um aumento nos falsos positivos, o que pode ser capturado pela precisão.

Para o modelo MobileNetV2, o *recall* para a classe "abnormal" foi de 74%, significando que o modelo identificou corretamente 74% das imagens com úlceras, mas deixou de reconhecer 26%, resultando em falsos negativos. Para a classe "normal", o *recall* foi de 86%, indicando uma maior eficácia na identificação de imagens sem úlceras. O VGG16 apresentou um *recall* de 93% para a classe "abnormal", demonstrando uma excelente capacidade de identificar corretamente as imagens com úlceras, e um *recall* de 91% para a classe "normal", evidenciando um desempenho igualmente robusto em ambas as classes. Estes valores equilibrados sugerem que o VGG16 tem uma boa sensibilidade tanto na detecção de úlceras quanto na confirmação de imagens normais. O VGG19 obteve um *recall* impressionante de 97% para a classe "abnormal", quase sempre identificando corretamente as imagens com úlceras. No entanto, o *recall* de 88% para a classe "normal" indica que o modelo falhou em identificar corretamente algumas imagens sem úlceras, resultando em possíveis falsos positivos. Apesar disso, o alto desempenho na detecção de úlceras torna o VGG19 uma excelente escolha para cenários em que a prioridade é a identificação de anomalias.

A última métrica avaliada foi a *F1-Score*. Esta é uma métrica amplamente utilizada em *machine learning*, especialmente em cenários onde há um desequilíbrio entre classes ou onde tanto a precisão quanto o *recall* são importantes. O *F1-score* é a média harmónica entre a precisão e o *recall*, combinando estas duas métricas numa única pontuação que reflete o equilíbrio entre a exatidão das previsões positivas e a capacidade do modelo de identificar corretamente as instâncias positivas.

$$F1Score = 2 \times \frac{Precisão \times Recall}{Precisão + Recall} \quad (4)$$

Esta métrica pode ser crucial em aplicações clínicas, como na classificação de úlceras do pé diabético, onde tanto a capacidade de detectar corretamente as úlceras (*recall*) quanto a exatidão das previsões positivas (precisão) têm implicações diretas no tratamento e diagnóstico.

O modelo MobileNetV2 obteve um *F1-Score* de 78% para a classe "abnormal" e 82% para a classe "normal". Estes resultados indicam um equilíbrio razoável entre precisão e *recall*, especialmente na classe "normal", onde o modelo demonstrou um bom desempenho. No entanto, o *F1-Score* mais baixo para a classe "abnormal" sugere que o modelo teve mais dificuldade em identificar corretamente todas as imagens de úlceras. O VGG16 apresentou um *F1-Score* de 93% para a classe "abnormal" e 91% para a classe "normal", mostrando um desempenho equilibrado em ambas as classes. Estes resultados indicam que o modelo foi capaz de alcançar um bom compromisso entre precisão e *recall*, sendo altamente confiável tanto na identificação de imagens com úlceras quanto na classificação correta de imagens normais. Já o VGG19 obteve um *F1-Score* de 94% para a classe "abnormal" e 92% para a classe "normal", sendo o modelo que apresentou o melhor equilíbrio entre as duas classes. Com um excelente desempenho em ambas, especialmente na classe "abnormal", o VGG19 destaca-se como o modelo mais eficaz na identificação correta de úlceras, minimizando tanto os falsos positivos quanto os falsos negativos.

As métricas foram calculadas utilizando a função *classification_report* da biblioteca *Scikit-learn*, que fornece um resumo das principais métricas de desempenho para cada classe, além de médias ponderadas.

Na tabela 3 temos um resumo dos resultados das métricas de desempenho para cada modelo:

Tabela 3 - Resumos dos resultados das métricas obtidas em cada modelos.

Modelo	Accuracy	Precisão	Recall	F1-Score
MobileNetV2	81 %	Ab: 82 % N: 79 %	Ab: 74 % N: 86 %	Ab: 78 % N: 82 %
VGG16	92 %	Ab: 92 % N: 92 %	Ab: 93 % N: 91 %	Ab: 93 % N: 91 %
VGG19	93 %	Ab: 97 % N: 90 %	Ab: 97 % N: 88 %	Ab: 94 % N: 92 %

Estes resultados indicam que o VGG19 apresentou a melhor *accuracy* e F1 score, sugerindo que é o modelo mais robusto para esta tarefa. O MobileNetV2, embora eficiente, apresentou um desempenho ligeiramente inferior, o que pode ser um *trade-off* esperado dado seu design leve e otimizado para dispositivos móveis.

As métricas de desempenho calculadas são fundamentais para avaliar a eficácia dos modelos e determinar a sua aplicabilidade clínica. Através dessa avaliação, é possível identificar não apenas o desempenho geral, mas também as áreas em que cada modelo pode ser aprimorado, contribuindo para a melhoria contínua do sistema de triagem automática de pacientes com úlceras do pé diabético.

Além das métricas tradicionais, foram calculadas as matrizes de confusão de cada modelo. A matriz de confusão é uma ferramenta essencial para a avaliação de modelos de classificação, especialmente em problemas de aprendizagem supervisionado, como a detecção de úlceras do pé diabético. Esta matriz permite a visualização clara do desempenho do modelo ao fornecer uma comparação entre os valores previstos e os reais. Em classificações binárias, como o caso deste trabalho, a matriz é composta por quatro elementos principais:

- Verdadeiros Positivos (TP): amostras da classe positiva que foram corretamente classificadas pelo modelo.

$$TP = Recall \times \text{Número Total de Positivos} \quad (5)$$

- Falsos Negativos (FN): amostras da classe positiva que foram incorretamente classificadas como pertencentes à classe negativa.

$$FN = \text{Número Total de Positivos} \quad (6)$$

- Falsos Positivos (FP): amostras da classe negativa que foram incorretamente classificadas como pertencentes à classe positiva.

$$FP = \frac{TP}{\text{Precisão}} - TP \quad (7)$$

- Verdadeiros Negativos (TN): amostras da classe negativa que foram corretamente classificadas.

$$TN = \text{Número Total de Negativos} \quad (8)$$

A matriz de confusão permite uma análise mais detalhada do desempenho do modelo, indo além de métricas agregadas, como a *accuracy*. Por exemplo, a precisão e o *recall* podem ser derivados da

matriz de confusão para avaliar a taxa de falsos positivos e falsos negativos, respetivamente, fornecendo *insights* importantes sobre a capacidade do modelo em minimizar erros críticos.

No contexto da deteção de úlceras do pé diabético, a correta identificação de úlceras (classe abnormal) é particularmente crucial. Um modelo com uma alta taxa de falsos negativos pode deixar de detetar uma úlcera num estágio inicial, o que pode resultar em consequências clínicas graves. Por outro lado, um número elevado de falsos positivos pode levar a tratamentos desnecessários.

Nos resultados da tabela 4 a seguir, são apresentadas as matrizes de confusão para os modelos MobileNetV2, VGG16 e VGG19, fornecendo uma visão detalhada sobre o desempenho de cada um na classificação de imagens normais e anormais.

Tabela 4 - Resumos dos resultados das matrizes de confusão obtidas em cada modelos.

Modelo	Classe	TP	FN	FP	TN
MobileNetV2	Abnormal	769	270	169	1016
	Normal	1019	166	271	768
VGG16	Abnormal	966	73	84	1101
	Normal	1078	107	94	945
VGG19	Abnormal	1008	31	31	1154
	Normal	1043	142	116	923

Os resultados das matrizes de confusão oferecem uma visão mais profunda sobre o desempenho de cada modelo, complementando as métricas apresentadas anteriormente.

O modelo MobileNetV2 apresentou uma performance inferior, especialmente em relação à classe abnormal, onde classificou corretamente 769 das 1039 imagens, mas falhou em identificar 270 casos, resultando numa quantidade significativa de falsos negativos. Isso reflete o *recall* relativamente baixo observado para essa classe. A consequência clínica desse desempenho é preocupante, já que falhas na deteção de úlceras podem atrasar o tratamento adequado. Além disso, o modelo também classificou erradamente 166 imagens da classe normal, o que pode ter impacto no aumento de falsos positivos. No entanto, considerando o design leve do MobileNetV2, focado em eficiência computacional, esta perda de precisão é esperada. O modelo mostrou-se eficiente, mas com limitações claras no que diz respeito à robustez necessária para uma aplicação clínica precisa.

O VGG16 apresentou uma melhora considerável. A matriz de confusão mostra que o modelo classificou corretamente 966 das 1039 imagens da classe abnormal, reduzindo significativamente o número de falsos negativos para 73. Esse resultado está em consonância com o *recall* mais elevado

observado para essa classe. A redução dos falsos negativos é crucial, pois significa que o modelo está a identificar melhor as úlceras, minimizando o risco de falhas na detecção de lesões importantes. O modelo também obteve um desempenho melhor na classificação da classe normal, com 1078 imagens corretamente classificadas e com apenas 94 falsos positivos. Estes números refletem a melhoria no equilíbrio entre as classes e mostram que o VGG16, com uma maior profundidade em comparação ao MobileNetV2, consegue capturar características mais detalhadas das imagens, resultando numa detecção mais precisa.

Finalmente, o VGG19 apresentou os melhores resultados em termos de desempenho. O modelo foi capaz de classificar corretamente 1008 das 1039 imagens da classe abnormal, o que significa que houve apenas 31 falsos negativos, o menor número entre os modelos testados. Isto alinha-se ao excelente *recall* obtido, indicando que o VGG19 é capaz de detectar quase todas as úlceras com precisão, minimizando significativamente o risco de não detecção de lesões. Na classe normal, o modelo classificou corretamente 1043 imagens, cometendo 142 erros, o que é uma ligeira piora em relação ao VGG16. No entanto, o desempenho geral do VGG19, especialmente em termos de detecção de úlceras, o coloca como o modelo mais robusto. A sua profundidade maior em relação ao VGG16 permitiu uma captura ainda mais refinada das características das imagens, garantindo um equilíbrio superior entre precisão e *recall*.

Essa análise das matrizes de confusão confirma a robustez do VGG19, que se destaca como o melhor modelo para a tarefa de detecção de úlceras do pé diabético. O VGG16 também mostrou resultados satisfatórios e pode ser uma alternativa viável dependendo do contexto, enquanto o MobileNetV2 demonstrou limitações, sendo mais adequado para situações onde a eficiência computacional é uma prioridade, mas com o custo de uma menor precisão. Num cenário clínico, onde a exatidão da detecção é crítica, os resultados indicam que o VGG19 seria a escolha mais apropriada.

Em seguida, a avaliação com o Grad-CAM, entra como uma análise qualitativa, ajudando-nos a ter uma análise visual dos resultados obtidos anteriormente.

Avaliando os resultados obtidos pela aplicação da Grad-CAM, podemos afirmar que a rede VGG16, foca, em grande parte, nas regiões mais relevantes para a detecção de anomalias, como áreas com lesões, úlceras ou sinais visíveis de infecção e isquemia nos pés. As ativações evidenciadas pelo Grad-CAM indicam que o modelo está a aprender a identificar corretamente padrões associados a condições patológicas. Isso é um sinal positivo de que o VGG16 consegue correlacionar as características visuais mais pertinentes das imagens com a classificação final.

Como podemos observar na figura 11, na primeira linha, as ativações não se concentram nas áreas lesionadas, mostrando dispersão em regiões irrelevantes. Na segunda linha, há alguma melhora, mas as ativações ainda estão espalhadas, incluindo áreas não patológicas. Na terceira linha, o padrão de dispersão continua, com poucas ativações corretas nas lesões e muita atenção em regiões sem patologia

Em suma, em algumas imagens, as áreas de maior ativação estão centralizadas sobre as partes mais danificadas do tecido, sugerindo que o modelo compreende a gravidade e a localização de certas

lesões. Isso reflete uma boa correspondência entre a decisão do modelo e a importância clínica das áreas identificadas.

No geral, as visualizações do Grad-CAM fornecem uma validação visual importante de que o modelo VGG16 está a focar corretamente nas regiões de maior interesse clínico, reforçando a eficácia do modelo na tarefa de classificação de pés diabéticos.

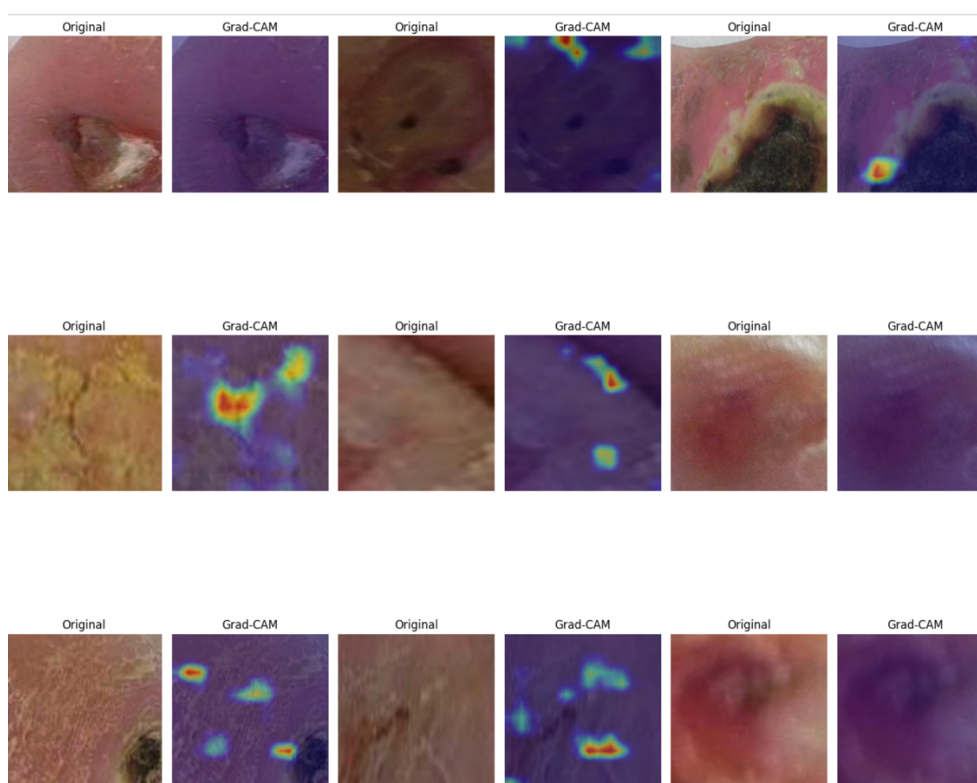


Figura 11 – Resultado da avaliação usando Grad-CAM na rede VGG16.

Nos resultados do Grad-CAM aplicados à VGG19, observa-se um desempenho superior em comparação à VGG16, especialmente na identificação das áreas patológicas em pés diabéticos. As ativações da VGG19 estão mais concentradas nas regiões relevantes, como ulcerações e lesões, com menor ativação em áreas não patológicas.

Como observado na figura 12, na primeira linha, as regiões lesionadas são claramente destacadas pelo Grad-CAM, embora em algumas imagens ainda haja ativação em áreas adjacentes não afetadas. Na segunda linha, as lesões menores e significativas são bem identificadas, com poucas ativações errôneas. Na terceira linha, o modelo continua mostrando boa correspondência entre as regiões ativadas e as lesões, com menos ativação dispersa.

De uma maneira resumida, a VGG19 demonstra maior precisão na identificação de lesões, com ativações mais focadas e menos dispersas, o que reflete uma maior capacidade de capturar detalhes clínicos relevantes em comparação com a VGG16.

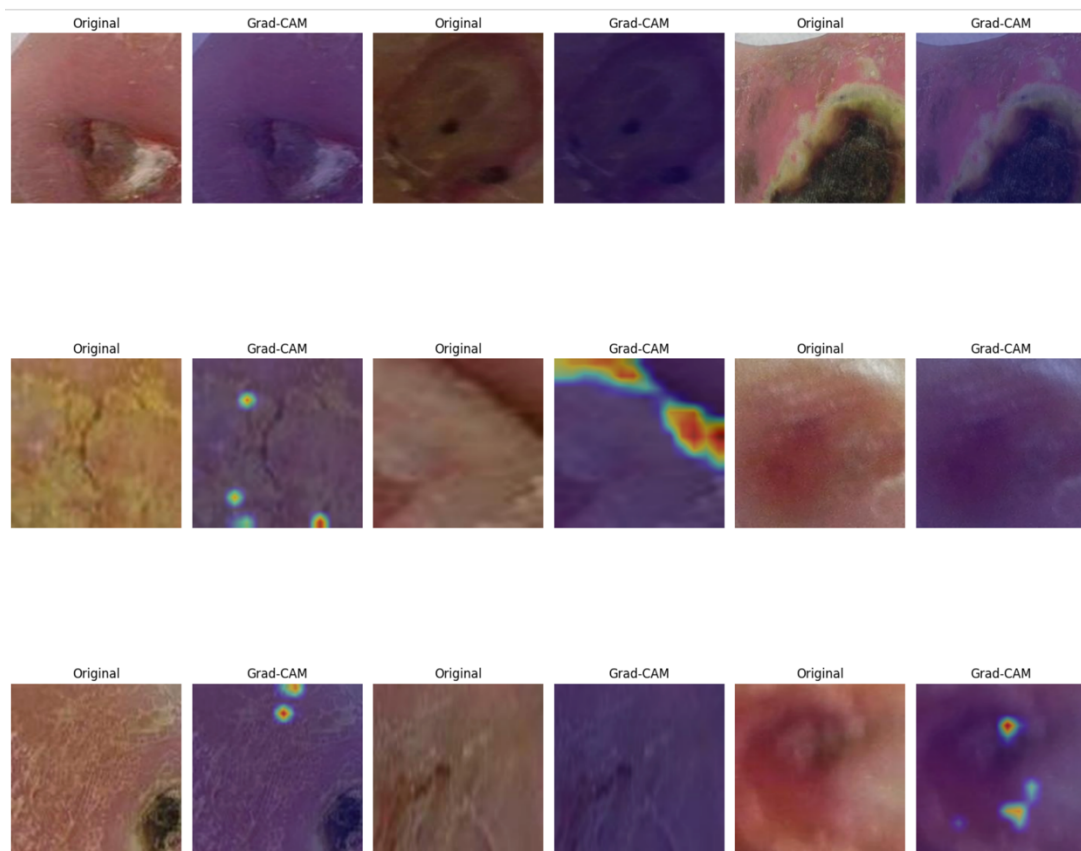


Figura 12 – Resultado da avaliação usando Grad-CAM na rede VGG19.

Nos resultados do Grad-CAM aplicados à MobileNetV2, observa-se um desempenho consistente na identificação das áreas patológicas em úlceras de pés diabéticos. As ativações da MobileNetV2 estão mais espalhadas em algumas imagens, mas ainda conseguem destacar áreas importantes associadas às lesões, embora com menos precisão em comparação a arquiteturas mais profundas, como a VGG19.

Na primeira linha, o Grad-CAM identifica as regiões lesionadas, mas em algumas imagens há ativação em áreas adjacentes que não estão diretamente relacionadas às lesões, indicando uma dispersão maior nas ativações. A precisão é satisfatória, mas um pouco mais difusa, com sinais em regiões próximas às áreas patológicas, como podemos observar na figura 13.

Na segunda linha, as ativações continuam a destacar as lesões, porém há uma certa sobreposição com áreas que não apresentam danos evidentes. Em algumas imagens, as ativações parecem ser mais generalizadas, não focando diretamente nas ulcerações ou nas bordas das lesões, o que pode ser um reflexo da arquitetura mais leve da MobileNetV2.

Na terceira linha, as ativações são mais concentradas, indicando um desempenho um pouco melhor em algumas imagens ao focar nas regiões lesionadas. No entanto, ainda há ativação em partes não afetadas, sugerindo que o modelo pode estar menos refinado na discriminação de detalhes clínicos mais sutil em comparação a arquiteturas mais profundas.

Resumindo, a MobileNetV2 apresenta um desempenho satisfatório, com ativações coerentes com as áreas patológicas, mas com uma leve tendência a ativar regiões não lesionadas. O foco das ativações é mais difuso, o que reflete uma menor capacidade de capturar detalhes clínicos finos, especialmente em comparação com a VGG19.

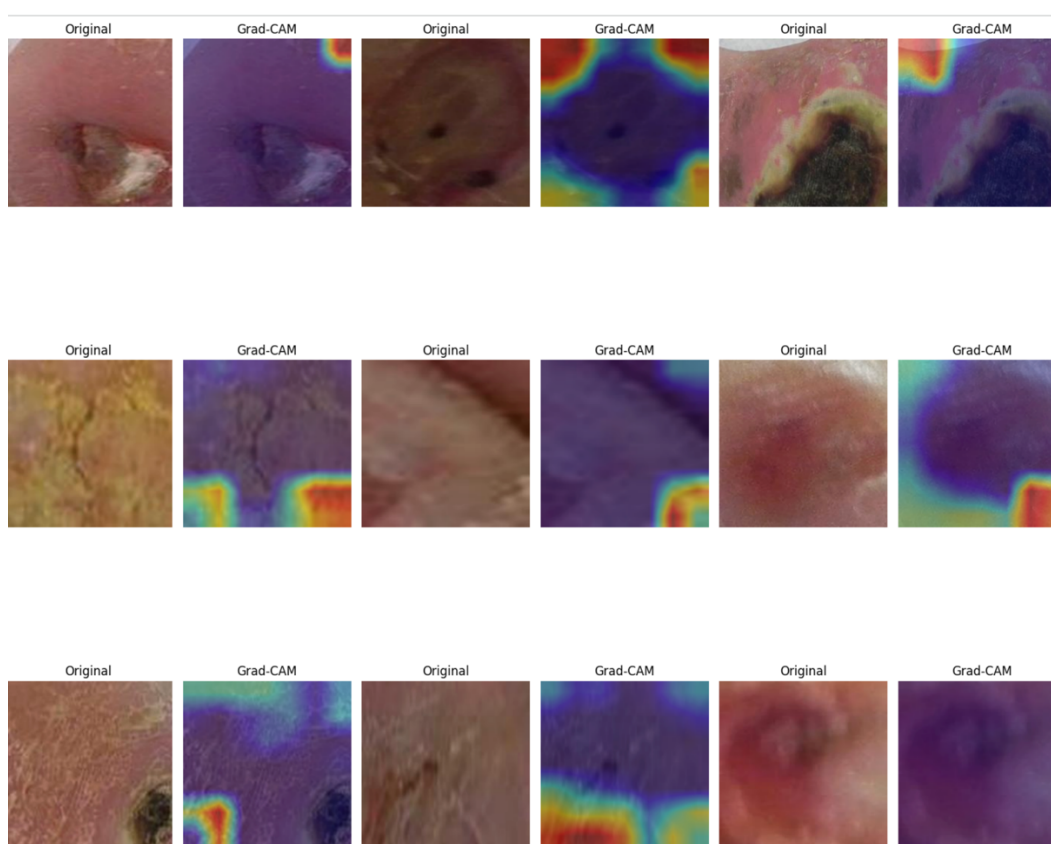


Figura 13 – Resultado da avaliação usando Grad-CAM na rede MobileNetV2..

6 Conclusões

Neste trabalho, diversas arquiteturas de Redes Neurais Convolucionais foram refinadas e avaliadas para a classificação de úlceras do pé diabético, com foco em três modelos principais: MobileNetV2, VGG16 e VGG19. A análise concentrou-se na identificação de imagens em duas classes: "abnormal" (imagens com úlceras) e "normal" (imagens sem úlceras).

Os resultados obtidos demonstraram a eficácia das abordagens de *machine learning* para a classificação de imagens médicas. O modelo MobileNetV2, otimizado para eficiência, alcançou uma *accuracy* de 81 %, com uma precisão de 82% para a classe "abnormal" e 79 % para a classe "normal". O recall foi de 74 % para a classe "abnormal", sugerindo que o modelo deixou de identificar 26 % das imagens com úlceras, enquanto para a classe "normal" o recall foi de 86 %, evidenciando maior eficácia na identificação de imagens saudáveis.

Em contraste, o modelo VGG16 apresentou resultados mais robustos, com uma *accuracy* de 92 % e uma precisão equilibrada de 92 % para ambas as classes. O recall de 93% para "abnormal" e 91 % para "normal" reflete a confiabilidade do modelo em identificar corretamente tanto imagens com úlceras quanto saudáveis. Estes resultados destacam o VGG16 como uma opção viável para aplicações clínicas, onde a precisão é crucial para evitar diagnósticos errôneos.

O VGG19 destacou-se como o modelo mais eficaz, alcançando uma *accuracy* de 93 % e uma precisão de 97 % para a classe "abnormal", indicando uma excelente capacidade de identificar imagens com úlceras. O recall de 97 % para "abnormal" sugere alta eficiência na detecção de imagens com úlceras, embora o recall de 88 % para "normal" aponte a possibilidade de alguns falsos positivos.

A análise comparativa revela que, enquanto o MobileNetV2 apresentou limitações, os modelos VGG16 e VGG19 mostraram desempenho superior em termos de precisão e recall. A adição de

camadas de dropout e batch normalization contribuiu significativamente para a melhoria do desempenho geral, ressaltando a importância dos ajustes finos nas arquiteturas de modelos.

Um dos principais fatores que podem contribuir para o aprimoramento do desempenho dos modelos de aprendizagem é a disponibilidade de mais recursos computacionais, especialmente unidades de processamento gráfico (GPUs). O treino de redes neurais profundas, como as utilizadas neste trabalho, é um processo intensivo e demorado, especialmente quando se trabalha com grandes volumes de dados. A utilização de múltiplas GPUs permitiria acelerar significativamente o processo de treino, permitindo que modelos mais complexos sejam treinados num tempo razoável. Isto não apenas melhoraria a eficiência do desenvolvimento, mas também possibilitaria a realização de experiências mais extensas e a exploração de arquiteturas de modelos alternativas, que podem exigir mais poder computacional.

Além disso, a disponibilidade de conjuntos de dados mais amplos e diversificados é fundamental para o aprimoramento da capacidade de generalização dos modelos. Um maior número de exemplos, especialmente em contextos médicos, pode ajudar a mitigar problemas de overfitting e melhorar a robustez das previsões. A inclusão de mais imagens de diferentes fontes, condições de iluminação e variabilidades de úlceras ajudaria os modelos a aprender características mais abrangentes, resultando num desempenho superior em cenários do mundo real. A criação de colaborações com instituições de saúde para acesso a conjuntos de dados mais amplos pode ser uma estratégia eficaz para este propósito.

Uma das direções mais promissoras para o futuro é a incorporação dos modelos desenvolvidos na prática médica. A utilização de sistemas de inteligência artificial para a triagem automática de pacientes com úlceras do pé diabético pode proporcionar uma ferramenta valiosa para profissionais de saúde, permitindo diagnósticos mais rápidos e precisos. A integração desses modelos em plataformas clínicas pode facilitar a identificação precoce de complicações, ajudando a melhorar os resultados dos pacientes e a otimizar o uso dos recursos de saúde.

Por fim, futuras pesquisas podem explorar o desenvolvimento de modelos que não apenas identifiquem úlceras, mas também prevejam o risco de desenvolvimento de novas lesões com base em dados históricos dos pacientes, como condições de saúde preexistentes e histórico de tratamento. Esta abordagem proativa pode ser transformadora, permitindo intervenções mais direcionadas e melhorando a qualidade do cuidado ao paciente.

Referências Bibliográficas

- Abubakar, A., Ugail, H., & Bukar, A. M. (2020). Assessment of Human Skin Burns: A Deep Transfer Learning Approach. In *Journal of Medical and Biological Engineering* (Vol. 40, Issue 3, pp. 321–333). Springer. <https://doi.org/10.1007/s40846-020-00520-z>
- Alberti, K. G. M. M., & Zimmet, P. Z. (1998). *Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications Part 1: Diagnosis and Classification of Diabetes Mellitus Provisional Report of a WHO Consultation ORIGINAL ARTICLES* (Issue 8).
- Al-gaashani, M. S. A. M., Shang, F., Muthanna, M. S. A., Khayyat, M., & Abd El-Latif, A. A. (2022). Tomato leaf disease classification by exploiting transfer learning and feature concatenation. *IET Image Processing*, 16(3), 913–925. <https://doi.org/10.1049/ipr2.12397>
- Alzubaidi, L., Fadhel, M. A., Olewi, S. R., Al-Shamma, O., & Zhang, J. (2020). DFU_QUTNet: diabetic foot ulcer classification using novel deep convolutional neural network. *Multimedia Tools and Applications*, 79(21–22), 15655–15677. <https://doi.org/10.1007/s11042-019-07820-w>
- Anisuzzaman, D. M., Patel, Y., Rostami, B., Niezgod, J., Gopalakrishnan, S., & Yu, Z. (2022). Multi-modal wound classification using wound image and location by deep neural network. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-022-21813-0>
- Armstrong, D. G., Swerdlow, M. A., Armstrong, A. A., Conte, M. S., Padula, W. V., & Bus, S. A. (2020). Five year mortality and direct costs of care for people with diabetic foot complications are comparable to cancer. In *Journal of Foot and Ankle Research* (Vol. 13, Issue 1). BioMed Central Ltd. <https://doi.org/10.1186/s13047-020-00383-2>
- Atkinson, M. A., Eisenbarth, G. S., & Michels, A. W. (2014). Type 1 diabetes. In *The Lancet* (Vol. 383, Issue 9911, pp. 69–82). Elsevier B.V. [https://doi.org/10.1016/S0140-6736\(13\)60591-7](https://doi.org/10.1016/S0140-6736(13)60591-7)
- Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K., & Ghayvat, H. (2021). Cnn variants for computer vision: History, architecture, application, challenges and future scope. In *Electronics (Switzerland)* (Vol. 10, Issue 20). MDPI. <https://doi.org/10.3390/electronics10202470>
- Boulton, A. J. (2019). *The diabetic foot*.
- Casarin, D. E., Donadel, G., Dalmagro, M., Oliveira, P. C. de, Ceranto, D. de C. F. B., & Zardeto, G. (2022). Diabetes mellitus: causas, tratamento e prevenção / Diabetes

mellitus: causes, treatment and prevention. *Brazilian Journal of Development*, 8(2), 10062–10075. <https://doi.org/10.34117/bjdv8n2-107>

Diabetes. (2023, April 5). World Health Organization . https://www.who.int/health-topics/diabetes#tab=tab_1

Duarte, N., Gonçalves, A., & De Revisão, A. (2011). Pé diabético. In *Angiologia e Cirurgia Vasculiar* | (Vol. 7).

Especial, A., Marques Bernardo, W., Roberto Cuce Nobre, M., & Biscegli Jatene, F. (2004). A prática clínica baseada em evidências. Parte II - Buscando as evidências em fontes de informação. In *Rev Assoc Med Bras* (Vol. 50, Issue 1). <http://www.biosis.org/>

Francisco Henrique Moura George. (2011). *Orientação da Direção-Geral Da Saúde*. www.dgs.pt

Goyal, M., Reeves, N. D., Rajbhandari, S., Ahmad, N., Wang, C., & Yap, M. H. (2020). Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques. *Computers in Biology and Medicine*, 117. <https://doi.org/10.1016/j.combiomed.2020.103616>

Hassan, N., De Andrade, S., Dal, K., Li, S.-M., Regina De Souza, C., Vi, T., Lúcia, M., Vii, Z., & Vi, E. (2010). *Pacientes com diabetes mellitus: cuidados e prevenção do pé diabético em atenção primária à saúde*.

I. Goodfellow, Y. Bengio, & A. Courville. (2016). *Deep Learning*. MIT Press.

IDF Diabetes Atlas 10th edition. (2021). www.diabetesatlas.org

Ioffe, S., & Szegedy, C. (2022). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*.

Liu, Z., John, J., & Agu, E. (2022). Diabetic Foot Ulcer Ischemia and Infection Classification Using EfficientNet Deep Learning Models. *IEEE Open Journal of Engineering in Medicine and Biology*, 3, 189–201. <https://doi.org/10.1109/OJEMB.2022.3219725>

Mamédio, C., Santos, C., Andrucioli De Mattos Pimenta, C., Roberto, M., & Nobre, C. (2007). *A Estratégica PICO para a construção da pergunta de pesquisa e busca de evidências*. www.eerp.usp.br/rlaeArtigodeAtualização

Mitchell, T. M. (2006). *The Discipline of Machine Learning*.

Munadi, K., Saddami, K., Oktiana, M., Roslidar, R., Muchtar, K., Melinda, M., Muharar, R., Syukri, M., Abidin, T. F., & Arnia, F. (2022). A Deep Learning Method for Early

Detection of Diabetic Foot Using Decision Fusion and Thermal Images. *Applied Sciences (Switzerland)*, 12(15). <https://doi.org/10.3390/app12157524>

Nascimento, R., França, S., Andrade, C., Alencar, M., & Costa, L. (2019). *Avaliação de Risco de Ulceração em Usuários Diabéticos* (Issue 3).

Nasteski, V. (2017). An overview of the supervised machine learning methods. *HORIZONS.B*, 4, 51–62. <https://doi.org/10.20544/horizons.b.04.1.17.p05>

Neves, Matias, Formiga, Cabete, A., Moniz, J., & Sampaio. (2013). *O pé diabético com infecção aguda: tratamento no Serviço de Urgência em Portugal*.

Ong, K. L., Stafford, L. K., McLaughlin, S. A., Boyko, E. J., Vollset, S. E., Smith, A. E., Dalton, B. E., Duprey, J., Cruz, J. A., Hagins, H., Lindstedt, P. A., Aali, A., Abate, Y. H., Abate, M. D., Abbasian, M., Abbasi-Kangevari, Z., Abbasi-Kangevari, M., ElHafeez, S. A., Abd-Rabu, R., ... Vos, T. (2023). Diabetes: a defining disease of the 21st century. *The Lancet*, 402(10397), 203–234. [https://doi.org/10.1016/S0140-6736\(23\)01301-6](https://doi.org/10.1016/S0140-6736(23)01301-6)

Paixão, G. M. de M., Santos, B. C., de Araujo, R. M., Ribeiro, M. H., de Moraes, J. L., & Ribeiro, A. L. (2022). Machine Learning in Medicine: Review and Applicability. In *Arquivos Brasileiros de Cardiologia* (Vol. 118, Issue 1, pp. 95–102). Arquivos Brasileiros de Cardiologia. <https://doi.org/10.36660/abc.20200596>

Paula Vieira Carvalho, A., Caldas Neiva, L., Paes Ribeiro de Souza, M., & Hudson Interaminense Mendes de Araújo, A. (2022). Desafios no enfrentamento da diabetes mellitus tipo 1 em crianças e adolescentes: revisão de literatura Challenges in dealing with type 1 diabetes mellitus in children and adolescents: a literature review. *Estudos Acadêmicos*, 5(11). <https://doi.org/10.5281/zenodo.7382427>

Popping the (PICO) question in research and evidence-based practice. (2002). In *Applied Nursing Research* (Vol. 15, Issue 3, pp. 197–198). W.B. Saunders. <https://doi.org/10.1053/apnr.2002.34181>

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, 380(14), 1347–1358. <https://doi.org/10.1056/nejmra1814259>

Relatório Anual do Observatório Nacional da Diabetes-Edição de 2023 Diabetes Factos e Números. (n.d.). www.spd.pt

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks*.

Shadeed, G. A., Tawfeeq, M. A., & Mahmoud, S. M. (2020). Automatic medical images segmentation based on deep learning networks. *IOP Conference Series: Materials Science and Engineering*, 870(1). <https://doi.org/10.1088/1757-899X/870/1/012117>

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks For Large-Scale Image Recognition*. <http://www.robots.ox.ac.uk/>

Taye, M. M. (2023). Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. In *Computation* (Vol. 11, Issue 3). MDPI. <https://doi.org/10.3390/computation11030052>

Teixeira do Nascimento, M., De Oliveira Silva, N. I., Castelo Branco de Andrade Brito, F., De Lima Fontes, F. L., Freire Oliveira, A., Alves Oliveira, J. V., Costa Soares, J., Ferreira de Brito, S. G., Dos Santos, M. O., Matias Facundes, D., Sousa Dias, I. T., Soares de Sousa Ponciano, R. C., Crispim Alencar, E. J., Ferreira de Lira, S., & Ferreira da Luz, R. (2019). Fatores de risco associados ao desenvolvimento do pé diabético e ações executadas na Atenção Primária à Saúde para prevenção do agravo. *Revista Eletrônica Acervo Saúde*, 33, e1371. <https://doi.org/10.25248/reas.e1371.2019>

Thotad, P. N., Bharamagoudar, G. R., & Anami, B. S. (2023). Diabetic foot ulcer detection using deep learning approaches. *Sensors International*, 4. <https://doi.org/10.1016/j.sintl.2022.100210>

Vidal, K. C. N. (2023). DIABETES MELLITUS: O PAPEL DA ENFERMAGEM NA ASSISTÊNCIA AO PACIENTE. *Revista Ibero-Americana de Humanidades, Ciências e Educação*, 9(6), 503–513. <https://doi.org/10.51891/rease.v9i6.8738>

Wang, C., Anisuzzaman, D. M., Williamson, V., Dhar, M. K., Rostami, B., Niezgodna, J., Gopalakrishnan, S., & Yu, Z. (2020). Fully automatic wound segmentation with deep convolutional neural networks. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-78799-w>

Yap, M. H., Cassidy, B., Pappachan, J. M., O'Shea, C., Gillespie, D., & Reeves, N. D. (2021). Analysis towards classification of infection and ischaemia of diabetic foot ulcers. *BHI 2021 - 2021 IEEE EMBS International Conference on Biomedical and Health Informatics, Proceedings*. <https://doi.org/10.1109/BHI50953.2021.9508563>

Zhao, X., Liu, Z., Agu, E., Wagh, A., Jain, S., Lindsay, C., Tulu, B., Strong, D., & Kan, J. (2019). Fine-Grained Diabetic Wound Depth and Granulation Tissue Amount Assessment Using Bilinear Convolutional Neural Network. *IEEE Access*, 7, 179151–179162. <https://doi.org/10.1109/ACCESS.2019.2959027>