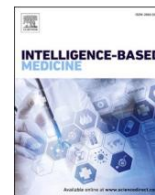


Listas de conteúdos disponíveis em [ScienceDirect](#)

Medicina Baseada em Inteligência

página inicial da revista: www.sciencedirect.com/journal/intelligence-based-medicine

DFU_MultiNet: Uma abordagem de rede neural profunda para detectar úlceras de pé diabético por meio da fusão de recursos em várias escalas usando o conjunto de dados DFU

Shuvo Biswas^a, Rafid Mostafiz b,^{*}, Bikash Kumar Paulo a, c, Khandaker Mohammad Mohi Uddin^d, MD Masudur Rahman^e, FNU Shariful^f

^a Departamento de Tecnologia da Informação e Comunicação, Universidade de Ciência e Tecnologia Mawlana Bhashani, Bangladesh

^b Instituto de Tecnologia da Informação, Universidade de Ciência e Tecnologia Noakhali, Bangladesh

^c Departamento de Engenharia de Software, Daffodil International University, Bangladesh

^d Departamento de Ciência da Computação e Engenharia, Universidade Internacional de Dhaka, Bangladesh

^e Departamento de Engenharia de Informação e Comunicações, Universidade de Ciência e Tecnologia Noakhali, Bangladesh

^f Escola de Computação, Universidade do Norte da Flórida, Jacksonville, FL, EUA

INFORMAÇÕES DO ARTIGO

Palavras-chave:

Úlceras de pé diabético (DFU)

VGG19

NasNetMobile

DenseNet201

Fusão de recursos

Diagnóstico auxiliado por computador (CAD)

ABSTRATO

A úlcera de pé diabético (UPD) é um problema comum entre pessoas com diabetes que pode resultar na amputação do membro afetado. Os métodos modernos de tratamento e diagnóstico de DFU são caros e demorados. Hoje, o desenvolvimento do método de diagnóstico auxiliado por computador (CAD) permite que os patologistas diagnostiquem a DFU com mais rapidez e precisão. Isso levou a um aumento no interesse em abordagens de aprendizagem profunda (DL) baseadas em CAD. Neste estudo, apresentamos uma nova estrutura chamada "DFU_MultiNet", que se concentra na abordagem de aprendizagem por transferência para classificar imagens de pele saudável e com úlceras usando repositórios disponíveis publicamente. A estrutura proposta é desenvolvida para oferecer um método eficiente e robusto para classificação DFU que determina a distinção entre pele saudável e pele ulcerada. A abordagem proposta extrai características de amostras de pés usando três modelos CNN pré-treinados bem conhecidos: VGG19, DenseNet201 e NasNetMobile. Finalmente, esses resultados extraídos são mesclados por meio de uma camada de soma para criar uma poderosa rede híbrida. Através da obtenção de exatidão impressionante (99,06%), precisão (100,00%), recall (98,18%), especificidade (100,00%), pontuação F1 (99,08%) e AUC (99,09%) a estrutura "DFU_MultiNet" proposta possui grande potencial como ferramenta de diagnóstico em ambientes clínicos e de saúde.

1. Introdução

As úlceras do pé diabético (DFUs) são uma exposição final do diabetes, identificadas por lesões nos pés. Segundo relatórios, a população global de indivíduos diabéticos era de 151 milhões no ano 2000, aumentou para mais de 422 milhões em 2014 e agora é estimada em cerca de 537 milhões em 2021 [1]. Nas últimas duas décadas, houve um aumento significativo de 10,5% na prevalência de diabetes entre adultos com 18 anos ou mais [2].

É importante notar que cerca de 80% destes pacientes diabéticos residem em países em desenvolvimento, que muitas vezes carecem de instalações e recursos de saúde adequados, levando a um menor nível de conscientização sobre as condições de saúde dos pacientes [3]. Uma proporção significativa de pacientes diabéticos (variando entre 15% e 25%) sofre de DFUs, que podem evoluir para um quadro grave

estágio, necessitando de amputação de membros inferiores, hospitalização e até morte se não tratada [4,5]. A infecção por DFUs é uma causa comum de amputação de membros ou pés [6].

Reduz as taxas de sobrevivência dos pacientes e diminui o poder da vida humana, impactando sua capacidade de ganhar a vida e participar de atividades sociais [7]. Condições como gangrena e morte de tecidos por doença podem levar à necessidade de amputação, e espera-se que o problema das DFUs aumente no futuro [8]. O pé é uma parte vital do corpo humano, mas, infelizmente, aproximadamente um milhão de pacientes com níveis elevados de açúcar no sangue perderão esta parte vital todos os anos. Após análise de vários estudos, observamos que a cada 20 segundos um pé diabético é operado.

Os médicos precisam de informações completas para um diagnóstico correto e um melhor tratamento da DFU. As técnicas de diagnóstico tradicionais exigem muito trabalho manual e estão sujeitas a erros. Diagnóstico assistido por computador

* Autor correspondente.

Endereços de e-mail: it21620@mbstu.ac.bd (S. Biswas), rafid.iit@nstu.edu.bd (R. Mostafiz), bikash.k.paul@ieee.org (BK Paul), jilanicsejnu@gmail.com (KM Mohi Uddin), masudur@nstu.edu.bd (MM Rahman), n01501509@unf.edu (FNU Shariful).

<https://doi.org/10.1016/j.ibmed.2023.100128> Recebido

em 23 de junho de 2023; Recebido em formato revisado em 2 de outubro de 2023; Aceito em 22 de novembro de 2023.

Disponível online em 28 de novembro de 2023.

2666-5212/© 2023 O(s) Autor(es). Publicado por Elsevier BV Este é um artigo de acesso aberto sob a licença CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(CAD) melhoram o desempenho com despesas mínimas. Além disso, avanços recentes em tecnologias móveis e de saúde vestíveis usadas para tratar complicações relacionadas ao diabetes. Eles podem melhorar a qualidade de vida do paciente e prolongar a remissão, detectando e controlando a inflamação prejudicial e a pressão no pé [9]. Às vezes, sensores e dispositivos ajudam a detectar vários tipos de sinais, como sinais físicos, químicos e biológicos. Eles podem registrar e medir esses sinais em um processo especializado e amplamente utilizado nos sistemas médicos atuais. Quando novos sensores e tecnologias relacionadas a sensores forem desenvolvidos esquematicamente, as áreas não médicas irão ajustá-los para utilização em seus setores de negócios. A evolução de novos sensores médicos aponta para uma aplicação mais ampla destes dispositivos no setor de saúde [10]. As imagens médicas [11–13] desempenham um papel crítico no diagnóstico e tratamento de várias condições médicas. O desempenho das abordagens de aprendizado de máquina (ML) e aprendizado profundo (DL) neste campo depende fortemente do uso de métodos avançados de extração e seleção de recursos que podem capturar com precisão características visuais importantes, como cor, forma e tamanho.

Estudos anteriores utilizando abordagens de ML e CNN alcançaram resultados notáveis no diagnóstico preciso de DFUs. No entanto, ainda há necessidade de mais pesquisas para garantir a eficácia dessas técnicas em ambientes do mundo real com diversas funções. O diagnóstico preciso e o manejo adequado das DFUs desempenham um papel crucial na melhoria do prognóstico do paciente. O manejo da DFU [14] envolve vários procedimentos, como ressecção vascular, remoção de feridas e tratamento de infecções. O regime de tratamento DFU depende do tipo e condição da ferida no pé. Os desafios associados ao atendimento DFU envolvem uma série de desafios educacionais que exigem a avaliação e comparação de várias técnicas de classificação, detecção e segmentação [15] para determinar as técnicas pré-treinadas e suas aplicações potenciais [16].

A aprendizagem por transferência refere-se à reutilização de um modelo CNN pré-treinado. Esta técnica é altamente popular em DL devido à sua capacidade de treinar redes CNN profundas com um pequeno conjunto de dados. Esta técnica funciona melhor para problemas com grandes conjuntos de dados. Contudo, a maioria dos problemas do mundo real consiste em grandes conjuntos de dados (ou seja, o sector médico) que não podem ser resolvidos pela aprendizagem por transferência tradicional. Para resolver esta questão importante, a técnica de aprendizagem por transferência multiescala (MTL) é aplicada para resolver esses problemas.

O MTL aprimora o aprendizado por transferência incorporando vários modelos CNN pré-treinados na rede neural, enquanto o aprendizado por transferência tradicional usa apenas um modelo CNN pré-treinado. Esses modelos processam amostras em várias escalas e capturam recursos em diferentes níveis de detalhe. Depois que os recursos são extraídos de cada modelo, eles são combinados harmoniosamente para criar uma representação abrangente.

O MTL é particularmente benéfico para os cientistas de dados extrair recursos dos vários modelos CNN pré-treinados em várias escalas. Tanto o MTL quanto o aprendizado por transferência podem ser empregados na classificação de imagens médicas, planejamento de tratamento, diagnóstico de doenças e identificação de anomalias. No setor da saúde, estas técnicas podem ser aplicadas para melhorar as ferramentas de diagnóstico e aumentar a sua precisão e fiabilidade. O processamento de imagens médicas em múltiplas escalas aumenta a capacidade de capturar eficazmente as características finas das lesões. O MTL contribui para a detecção precoce em imagens médicas, facilitando intervenções imediatas e, em última análise, melhorando os resultados dos pacientes. Além disso, esses modelos são excelentes na análise rápida de grandes quantidades de imagens médicas, permitindo que os médicos tomem decisões mais rápidas e precisas.

Neste estudo, propusemos um sistema de categorização automática de DFU baseado em estrutura híbrida "DFU_MultiNet" que é capaz de distinguir entre pele normal e ulcerada do pé a partir do conjunto de dados DFU. Os dados DFU foram primeiro pré-processados e particionados em conjuntos de teste de trem. Em seguida, o conjunto de treinamento é inserido no modelo de aprendizagem por transferência multiescala (MTL). Três modelos populares de CNN pré-treinados, nomeadamente DenseNet201, NasNetMobile e VGG19, compõem este modelo MTL, que é aplicado para extrair características de amostras de pele de pés. Depois disso, integramos todos os resultados extraídos por meio de uma camada de soma e ajustamos aqueles usando duas camadas de eliminação, duas camadas densas e duas camadas de normalização de lote (BN). Finalmente, a camada densa final é empregada para o DFU

tarefa de classificação.

Os pontos cruciais deste trabalho de pesquisa são descritos a seguir.

1. A estrutura "DFU_MultiNet" propõe uma técnica de fusão de recursos em múltiplas escalas que supera os modelos existentes.
2. A estrutura proposta demonstra um desempenho notável na classificação de todas as úlceras e imagens saudáveis com taxa de classificação de erros omissos.
3. "DFU_MultiNet" fornece uma técnica de extração de recursos sem segmentação em um grande conjunto de dados DFU e mostra a proficiência na aquisição de resultados elevados no conjunto de dados DFU. Nesta técnica, nenhum método manual de extração de características é necessário, o que a distingue das abordagens convencionais de ML. Por outro lado, captura diretamente características pertinentes de toda a imagem, eliminando a necessidade de um processo de segmentação distinto.

As outras partes deste artigo foram estruturadas em diversas seções. Na Seção 2, descrevemos a revisão da literatura. Depois disso, a estrutura proposta "DFU_MultiNet" foi explicada em detalhes, incluindo pré-processamento de dados e estratégias de construção de modelo MTL na Seção 3. A configuração experimental e as consequências são apresentadas de forma abrangente na Seção 4. A Seção 5 contém uma análise comparativa de os resultados e sugestões para direção futura. As conclusões finais deste estudo estão resumidas na Seção 6.

2. Revisão da literatura

Numerosos estudos foram publicados sobre a classificação e detecção de imagens DFU para distinguir entre pele saudável e pele com úlcera. A maioria dos autores propõe abordagens de ML e processamento de imagens que analisam diferentes características como padrões, matizes e texturas morfológicas. O desempenho do trabalho proposto depende dos modelos e abordagens de formação adotados. Vários tipos de trabalhos relacionados ao DFU são brevemente descritos nesta seção.

Kaslimi et al. [17] (2022) introduziram uma revisão abrangente da investigação existente sobre a utilização de inteligência artificial (IA) na monitorização de DFUs, destacando as vantagens destes métodos, ao mesmo tempo que reconhecem os desafios na sua implementação eficaz para o atendimento remoto ao paciente. A análise deles se concentrou nos métodos de imagem e sensores ópticos empregados para detectar DFUs. O estudo contempla tanto as características dos sensores quanto os aspectos fisiológicos dos pacientes. A fonte de dados de imagem recomendou uma série de táticas de monitoramento, o que limita o uso de algoritmos de IA [18].

Thotad et al. [19] (2022) propuseram uma abordagem de EAD denominada EfficientNet, para previsão e detecção precoce de úlceras de pé diabético (DFU). O EfficientNet foi empregado para um conjunto de amostras cujo tamanho de imagem era de 844 pés, composto por úlceras diabéticas e pele saudável. Nesta abordagem, eles construíram uma rede robusta ajustando três propriedades cruciais (resolução, largura e profundidade) do modelo CNN para classificar pés diabéticos e normais. Seu método obteve excelentes resultados em comparação com algoritmos modernos como GoogleNet, VGGNet (VGG16 e VGG19) e AlexNet. Ele forneceu a máxima exatidão, recall, precisão e pontuação f1 de 98,97%, 98%, 99% e 98%, respectivamente.

Em outro método, Juan et al. [20] (2022) propuseram um novo classificador profundo (CNN) denominado DFU_VIRNet, para uma tarefa automática de classificação de pele DFU. Além disso, seu método focou em mapas de estimativa para identificar a probabilidade de áreas de risco responsáveis pelo desenvolvimento de DFU. Dois tipos de amostras, nomeadamente visíveis e invisíveis, foram inseridos no esquema proposto para fins de treinamento e teste. O DFU_VIRNet forneceu a pontuação AUC (0,99301) e ACC (0,97750) mais alta, superando os resultados recentes.

Doulamis et al. [21] (2021) propuseram um valioso dispositivo não invasivo que utiliza tecnologia de base fotônica para o tratamento de DFUs em pacientes diabéticos. Esse aparelho emprega imagens hiperespectrais e térmicas para avaliar o estado da úlcera e estimar os biomarcadores desoxihemoglobina e oxihemoglobina por meio da técnica de imagem.

Além disso, este novo dispositivo foi aprimorado incorporando métodos de processamento de sinal utilizando DL para melhorar a precisão dos pixels e reduzir o ruído usando abordagens de super-resolução.

Das et al. [22] (2021) propuseram uma estrutura única (DFU_SPNet), que foi construída a partir de camadas de convolução paralelas empilhadas (SP). DFU_SPNet empregou três módulos de tamanho de kernel distintos de camadas de convolução SP para extrair o mapa de recursos. Obtendo ACC (97,4%), o DFU_SPNet superou os trabalhos ultramodernos existentes após avaliar o conjunto de dados de teste DFU utilizando o otimizador (SGD) com uma taxa de aprendizado (1-e2).

Alzubaidi et al. [23] (2021) propuseram quatro modelos híbridos de CNN para classificar pele anormal versus pele normal. Os modelos que desenvolveram incorporaram camadas CNN tradicionais juntamente com camadas convolucionais paralelas (PCL). Cada modelo possui seis módulos de PCL, mas a faixa de ramificações do PCL é de 2 a 5. Todos os modelos extraíram os recursos da mesma entrada, utilizando PCL com tamanhos de kernel diferentes e, em seguida, mesclando os resultados extraídos. Entre os modelos testados, o modelo com quatro filiais obteve a maior pontuação F1 (95,8%).

Alzubaidi et al. [24] (2020) propuseram um estudo que utilizou um conjunto de dados DFU de 754 pés de pacientes com úlcera saudáveis e diabéticos. Para automatizar a categorização dessas imagens, foi proposta uma CNN profunda denominada DFU_QUTNet, que se diferenciava das CNNs convencionais. DFU_QUTNet era mais amplo, mas não necessariamente mais profundo. Esta abordagem melhorou o problema de propagação de gradiente, uma vez que os erros foram remetidos em vários canais diferentes.

Tan e Le et al. [25] (2019) demonstraram que otimizar a resolução, largura e profundidade da rede pode melhorar o desempenho e revelaram uma extensa descrição do dimensionamento do modelo. Eles introduziram um novo mecanismo de escala que utilizou um coeficiente composto para escalar uniformemente esses parâmetros. Ao aplicar esse mecanismo a um modelo básico, eles construíram a família EfficientNets de modelos CNN. Este estudo superou outros ConvNets existentes em precisão, eficácia e velocidade de inferência, ao mesmo tempo que obteve um ACC impressionante (84,3%) no ImageNet [26].

Manu Goyal et al. [27] (2017) endossaram o potencial dos recursos convencionais de visão computacional (CCV) como um método prático e econômico para detectar úlceras nos pés em pacientes diabéticos. Eles desenvolveram uma nova estrutura baseada em CNN chamada DFUNet para diferenciar entre DFUs e pele saudável, extraíndo um mapa de características de imagem. A DFUNet alcançou a AUC (0,962) aplicando uma abordagem de validação cruzada de 10 vezes. Demonstrou melhor desempenho do que utilizar abordagens tradicionais de ML e DL.

Wang et al. [28] (2017) empregaram uma caixa de captura especial para capturar imagens DFU e, por outro lado, identificaram a localização exata do DFU empregando uma abordagem de classificação SVM em dois estágios. Essas duas etapas foram: (i) a etapa de segmentação, que extrai superpixels; e (ii) a etapa de extração de características, que extrai características importantes da imagem.

Manu et al. [29] (2017) implementaram um método para segmentar DFUs a partir de imagens do pé inteiro. Embora o método apresentasse resultados promissores, a abordagem tinha limitações. A primeira limitação foi a impraticabilidade do uso de uma superfície de caixa para capturar imagens DFU devido a preocupações sobre potenciais riscos de infecção em um ambiente de saúde, o que enfatizou a importância de explorar abordagens alternativas para coleta e análise de imagens. A segunda foi a sua incapacidade de processar eficazmente um grande conjunto de dados DFU. A Tabela 1 descreve uma visão geral do artigo publicado.

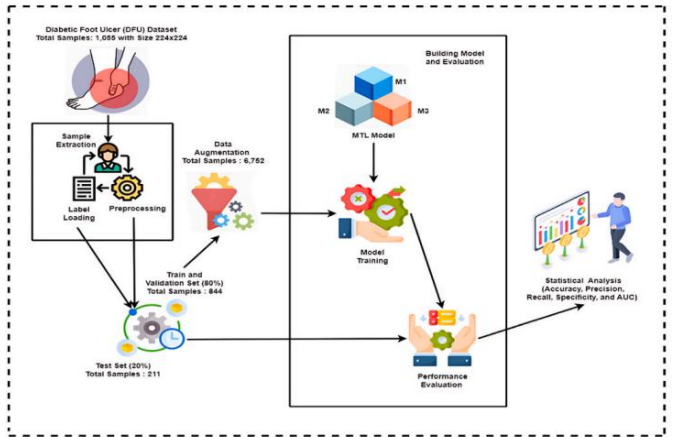
3. Metodologia

Desenvolver um sistema de diagnóstico preciso para classificar DFU é uma tarefa desafiadora para fins clínicos. Esta seção explica como propusemos uma estrutura híbrida "DFU_MultiNet" para o desafio de classificação DFU, mesclando vários modelos CNN pré-treinados. A Figura 1 mostra a estrutura da proposta "DFU_MultiNet". A estrutura realiza as seguintes operações: extrair amostras DFU, carregar rótulos,

Tabela

1 Uma tabela de verão de todas as abordagens.

Papel	Conjunto de dados tamanho	Abordagem	Desempenho
Thotad et al. [19]	1688	EficienteNet	Precisão = 98,97%, Precisão = 99%, Rechamada = 98%, pontuação F1 = 98%
Juan et al. [20]	13.200	DFU_VIRNet	Precisão = 0,9775, Sensibilidade = 0,98167, Especificidade = 0,97333, Precisão = 0,97355, Recall = 0,98167, pontuação F1 = 0,97759, AUC = 0,99301
Doulamis et al. [21]	-	Técnicas de imagem hiperespectral e térmica	Predição precoce e prognóstico de uma DFU, compreendendo o efeito dos biomarcadores na DFU
Das et al. [22]	3827	DFU_SPNet	Precisão = 0,964, Sensibilidade = 0,984, Especificidade = 0,951, Precisão = 0,926, Recall = 0,984, pontuação F1 = 0,954, AUC = 0,974
Alzubaidi et al. [23]	17.053	CNN híbrida	Precisão = 97,3%, recall = 94,5%, pontuação F1 = 95,8%
Alzubaidi e outros. [24]	17.053	DFU_QUTNet+ SVM	Precisão = 0,954, recall = 0,936, pontuação F1 = 0,945
Tan e Le e outros. [25]	ImageNet ConvNet		Precisão = 84,3%
Manu Goyal et al. [27]	22.777	DFUNet	Precisão = 0,925, Sensibilidade = 0,934, Especificidade = 0,911, Precisão = 0,945, Pontuação F1 = 0,939, AUC = 0,961
Wang et al. [28]	100	SVM	Sensibilidade = 73,3%, Especificidade = 94,6%
Manu et al. [29]	705	FCN-16	Dados para região da úlcera (UR) = 0,794 e pele circundante (SS) = 0,851, Especificidade para SS = 0,994, Sensibilidade para UR = 0,789 e SS = 0,874 e MCC para UR = 0,785 e SS = 0,852



1. Estrutura DFU_MultiNet proposta para tarefas de classificação DFU M1, M2 e M3 são três modelos CNN: DenseNet201, VGG19 e NasNetMobile, respectivamente. Esses três modelos construíram um modelo híbrido MTL (aprendizado de transferência multiescala).

pré-processamento de amostra, divisão do conjunto de dados DFU, aumento de dados, treinamento do modelo MTL e, finalmente, análise de parâmetros estatísticos utilizando o conjunto de teste. A tarefa geral de classificação do nosso framework é mostrada passo a passo no Algoritmo 1.

Algorithm 1 Automated diabetic foot ulcer (DFU) classification and detection.

Input: DFU Training dataset γ_1 (70%), Validation dataset γ_2 (10%), and Testing dataset γ_3 (20%)

β = batch size

σ = epochs

λ = optimizer

η = learning rate

ε = the number of samples converted into mini-batch size

Output: Ω = weight of pre-trained CNN algorithms

Begin:

1: Convert each sample in the training dataset into a size of 224x224.

2: Apply the data augmentation method to enhance the sample size.

3: Extract the feature maps from the sample utilizing three pre-trained CNN algorithms, namely, VGG19, NasNetMobile, and DenseNet201.

4: Merge the extracted maps by applying the concatenation layer.

5: Initialize four fine-tuned CNN layers: batch normalization, dense, dropout, and softmax.

6: Set the training parameters: η , σ , λ , β , and ε .

7: Train the "DFU_MultiNet" framework and calculate the primary weights.

8: **for** $\sigma = 1$ to σ **do**

9: Select a mini-batch size ε .

10: Forward propagation and evaluation of the binary loss function.

11: Backpropagation and improving the weight Ω .

12: **end for**

End

3.1. Conjunto de dados DFU

Neste estudo, o conjunto de dados de trabalho denominado conjunto de dados DFU foi coletado de um repositório online de acesso público [30], que contém quatro pastas que são imagens originais, patches, conjunto de testes e imagens de aprendizagem por transferência. Para treinar e testar nosso modelo híbrido, selecionamos a pasta de patches dessas quatro pastas, que incluía um total de 1.055 patches de pele. Destas manchas, 512 foram identificadas como anormais (úlceras), enquanto as restantes 543 foram classificadas como normais (pele saudável). A Figura 2 mostra uma amostra de manchas na pele. O conjunto de dados DFU é dividido em dois conjuntos, trem (80%) e teste (20%), com a ajuda da função `train_test_split`. A função `train_test_split` é importada do pacote "sklearn.model_selection" em Python. Então, para o conjunto de validação, 10% dos dados são separados do conjunto de trem com a ajuda da função `train_test_split`. Finalmente, o conjunto de dados DFU (ou seja, 1.055 amostras) foi dividido em três fases distintas: 70% (ou seja, 760 amostras) para treinamento, 20% (ou seja, 211 amostras) para teste e 10% (ou seja, 84 amostras) para validação.

A Tabela 2 fornece mais detalhes sobre o conjunto de dados DFU.

Tabela 2

Informações detalhadas sobre dados de trabalho antes de aplicar técnicas de aumento.

Conjunto de dados	Rótulo	Treinamento	Validação	Teste
DFU	Pele saudável	390	43	110
	Pele de úlcera	370	41	101
	Total de amostras	760	84	211

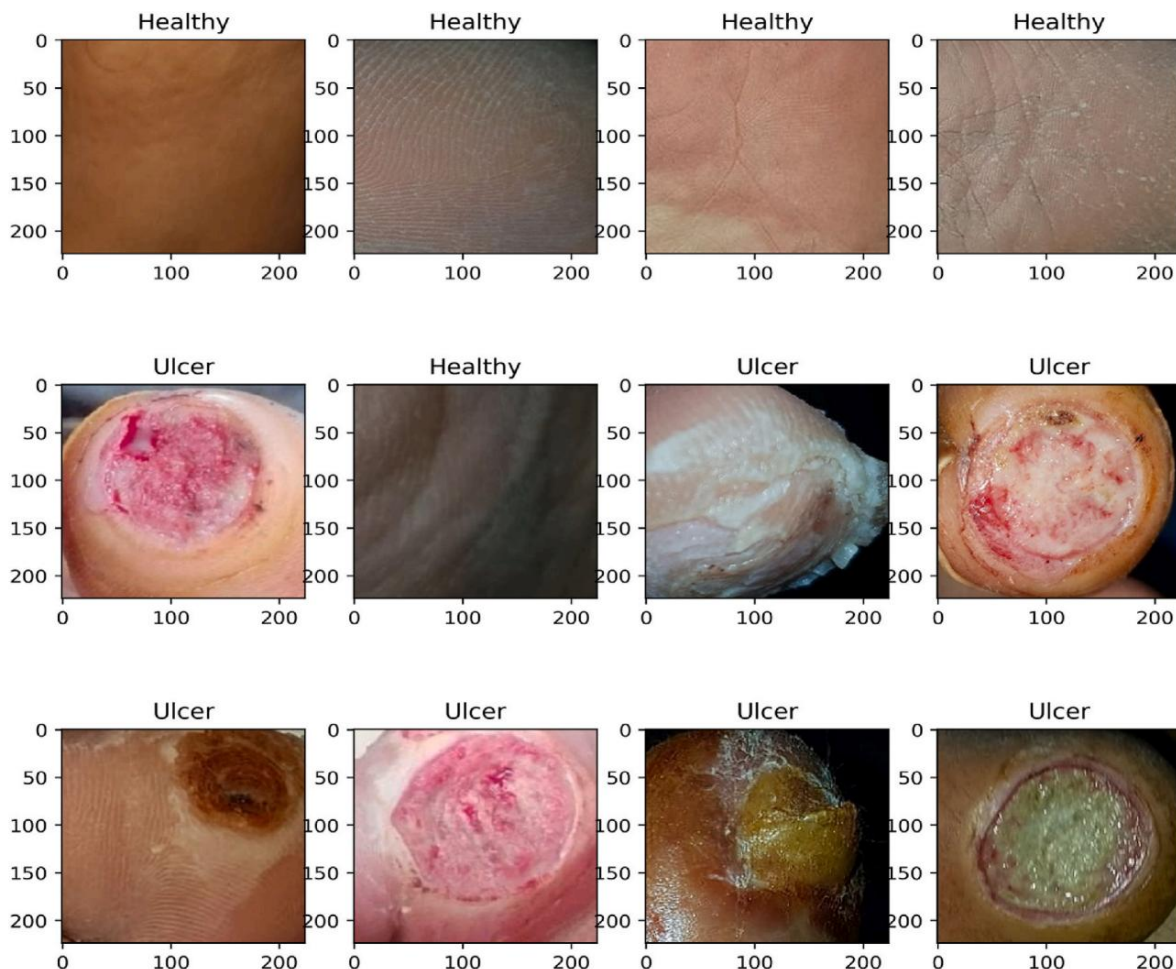


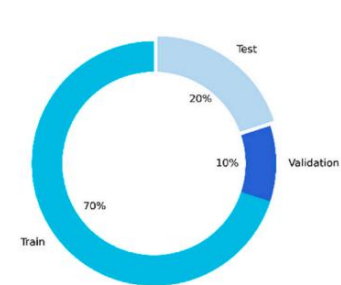
Figura 2. Amostras de imagens de pele (Saudável e Úlcera) do conjunto de dados DFU.

3.2. Pré-processamento de dados

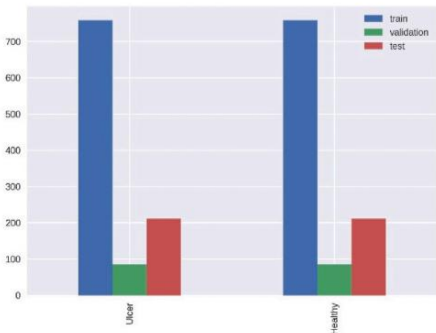
Antes de inserir as imagens DFU no modelo de aprendizagem por transferência multiescala (MTL), vários estágios de pré-processamento foram implementados. De acordo com os princípios de aprendizagem por transferência, cada imagem no conjunto de dados DFU estava no formato .jpg com tamanho de 224 x 224 pixels e canal RGB. As imagens foram convertidas em arrays Numpy para permitir um treinamento mais rápido e reduzir o uso de memória. Além disso, também embaralhamos o conjunto de dados para treinar amostras não ordenadas. O conjunto de dados DFU foi dividido em fases distintas (ver Fig. 3 (a)), com 70% (ou seja, 760 amostras) alocados para treinamento, 20% (ou seja, 211 amostras) para teste e 10% (ou seja, 84 amostras) para validação, respectivamente. O gráfico de barras (Fig. 3 (b)) indica o número de amostras após dividir os dados em três fases. Redes profundas requerem um grande número de amostras de treinamento devido aos seus numerosos parâmetros. Esse valor é aumentado com a aplicação de uma técnica poderosa chamada aumento de dados. Essa técnica pode servir a vários propósitos, como melhorar o desempenho da estrutura DFU_MultiNet, lidar com questões de overfitting e aumentar a robustez do modelo. Finalmente, aplicando esta técnica, aprimoramos o conjunto de dados DFU de 1055 para 6963. Em técnicas de aumento de dados envolvendo parâmetros não binários (verdadeiros ou falsos), como ângulo de rotação, deslocamento, zoom ou cisalhamento, esses valores de parâmetros são geralmente selecionados aleatoriamente de predefinidos. intervalos ou distribuições. Por exemplo, na rotação de imagem, o argumento rotation_range permite a seleção aleatória de qualquer grau entre 0 e 360. Ao aplicar zoom em uma imagem, ela normalmente é dimensionada dentro do intervalo de [1 - zoom_range, 1 + zoom_range]. O cisalhamento envolve a seleção de um valor de ponto flutuante de uma distribuição uniforme no intervalo entre 0 e 1. Em contraste, a inversão de uma imagem é binária, representada como Verdadeiro ou Falso. Todos os parâmetros de aumento usados neste estudo são fornecidos na Tabela 3. A Tabela 4 revela informações detalhadas sobre o conjunto de dados DFU após a abordagem de aumento. A Figura 4 mostra alguns exemplos de imagens aumentadas.

3.3. Construindo o modelo MTL

Há um interesse crescente em aplicar redes neurais pré-treinadas a uma ampla gama de tarefas além de seus domínios iniciais [31]. Isto é particularmente exigido na área médica, onde a obtenção de dados rotulados suficientes para o treinamento de redes DL pode ser um desafio [32]. Para resolver isso, os pesquisadores aproveitaram o poder das redes treinadas no ImageNet [33], um vasto banco de dados contendo mais de 14 milhões (M) de imagens em 20.000 categorias, mas usaram 1,2 M de imagens abrangendo mais de 1.000 classes para benchmarking. Essas classes podem ser conceitos abstratos, cenas, objetos e animais. Os modelos pré-treinados utilizados neste experimento foram pré-treinados neste vasto banco de dados e os pesos dessas redes são determinados por este banco de dados. Na EAD, essas redes pré-treinadas já adquiriram representações e recursos valiosos deste vasto banco de dados. Esses recursos aprendidos melhoram a proficiência de um modelo no tratamento de dados invisíveis, mas isso



(a): Data splitting ratio.



(b): Number of samples after splitting.

Figura 3. (a) Taxa de divisão de dados. Figura 3 (b) Número de amostras após divisão.

Tabela

3 Técnicas e parâmetros de aumento de dados.

Número	Estratégias de técnicas de dados	Valores de parâmetros
1	Faixa de zoom	2
2	Faixa de rotação	90
3	Faixa de cisalhamento	0,4
4	Faixa de mudança de largura	0,2
5	Faixa de mudança de altura	0,2
6	Inversão horizontal	Verdadeiro
7	Inversão vertical	Verdadeiro

Tabela

4 Informações detalhadas sobre dados de trabalho após aplicação de técnicas de aumento.

Conjunto de dados	Rótulo	Treinamento	Validação	Teste
DFU	Pele saudável	3120	344	110
	Pele de úlcera	2960	328	101
	Total de amostras	6080	672	211

vantagem não é obtida no modelo retreinado. Por outro lado, o treinamento de modelos DL do zero requer recursos computacionais e hardware poderoso, mas esses recursos podem ser economizados usando modelos pré-treinados. Esses modelos são normalmente construídos sobre arquiteturas bem estruturadas que foram bastante ajustadas e testadas. Dadas essas vantagens, neste estudo optamos por modelos pré-treinados em vez de treiná-los novamente do zero.

Para construir o modelo MTL, primeiro três modelos pré-treinados DenseNet201, NasNetMobile e VGG19 são alimentados, cada um, com imagens coloridas DFU com tamanho de 224 x 224 como entrada para que cada modelo possa extrair os recursos das imagens separadamente. Em seguida, cada uma das três camadas Global-AveragePooling2D é aplicada separadamente a cada modelo, achatando as respectivas camadas em um vetor calculando a média dos recursos de cada entrada. Posteriormente, esses vetores individuais são combinados em um vetor unificado, empregando a camada concatenada. Depois disso, os recursos integrados são ajustados usando seis camadas CNN. Dentro dessas seis camadas, a primeira é a camada de abandono (taxa de abandono 0,4), a segunda é a camada de normalização de lote, a terceira é a camada densa (128 unidades e função de ativação ReLU), a quarta é a camada de abandono com uma taxa de abandono de 0,6, a quinta é a camada de normalização de lote e a camada final é a camada densa (2 unidades e função de ativação softmax). A Figura 5 representa o modelo MTL (aprendizado de transferência multiescala) para classificação de imagens DFU. O modelo MTL contém 43.077.398 parâmetros após a fusão de todos os recursos coletados, o que é aproximadamente 2, 2 e 10 vezes mais em comparação com os modelos VGG19, DenseNet201 e NasNetMobile. A explicação básica destes três modelos adotados e a abordagem de ajuste fino são fornecidas nas subseções seguintes.

3.3.1. DenseNet

Huang et al. (2017) [34], desenvolveram inicialmente o melhor pré-treinado

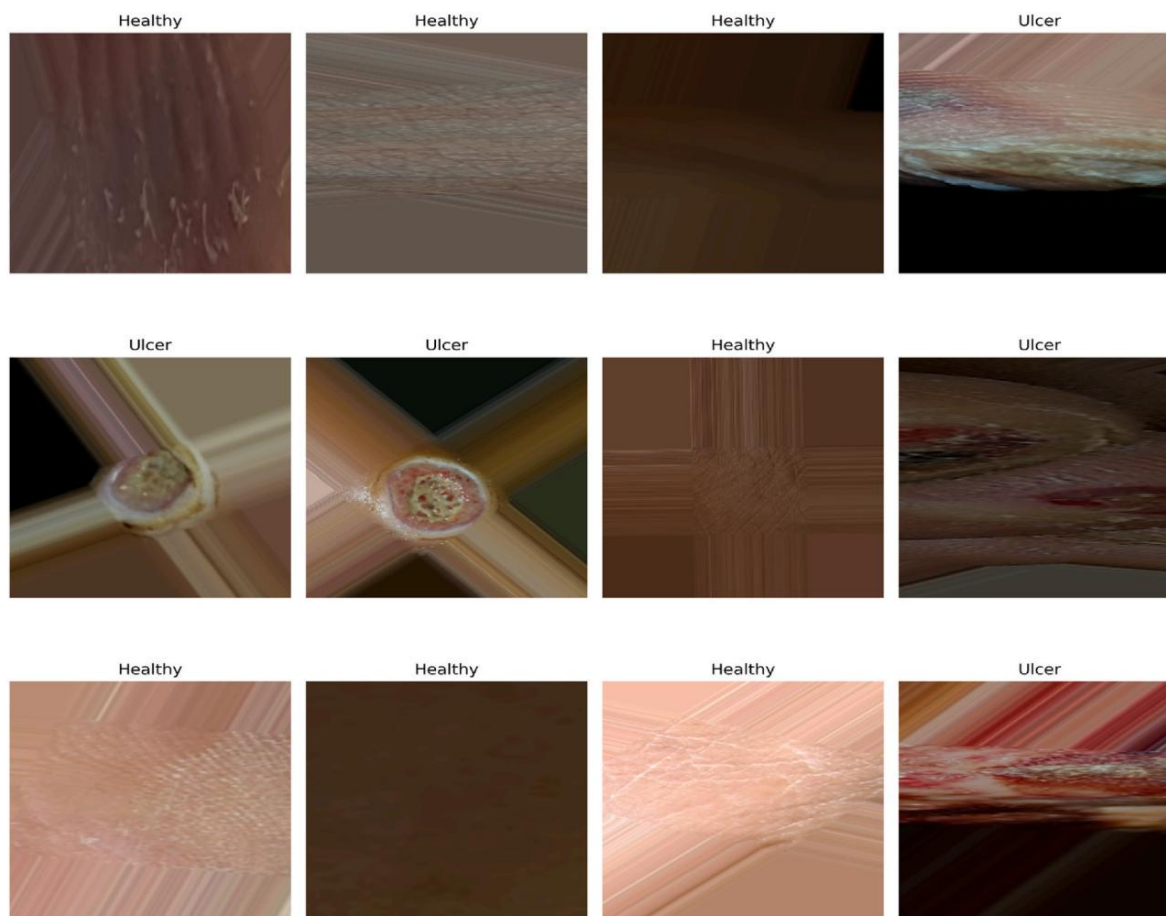
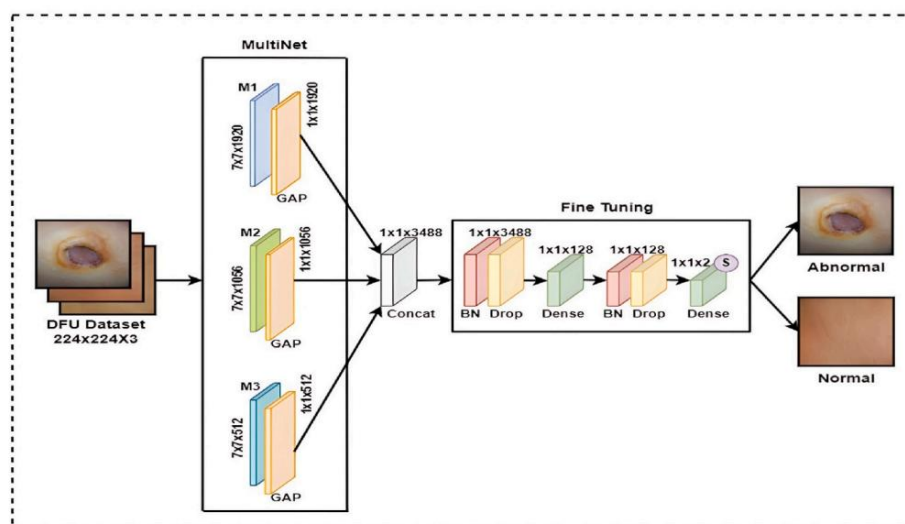


Figura 4. Exemplos de imagens aumentadas.



5. Modelo MTL (aprendizado de transferência multiescala) para extrair características de amostras e mesclá-las para diagnóstico de úlcera cutânea. M1, M2 e M3 são três modelos DenseNet201, VGG19 e NasNetMobile pré-treinados, respectivamente. GAP indica a camada de pooling média global, Concat indica a camada de concatenação, BN indica a camada de normalização em lote e Drop indica a camada de abandono.

Modelo de classificação CNN denominado DenseNet para obter a melhor precisão nos conjuntos de dados ImageNet, CIFAR10 e CIFAR-100. Este modelo foi criado usando um modelo de feedforwarding semelhante ao modelo ResNet. Tal conexão permite que a arquitetura troque dados cruciais dentro da rede, melhorando o desempenho do modelo e aumentando a eficácia do treinamento do modelo [35]. Neste estudo, empregamos

DenseNet201 como nossa primeira rede de detectores de recursos. Esta rede possui 201 camadas neurais profundas no total, cada uma das quais é construída para resolver dificuldades de overfitting ao trabalhar com amostras não ordenadas. Além disso, após treinar este modelo, ele contém um total de 18.321.984 parâmetros.

3.3.2. 3.3.2 NasNetMobile Zoph

et al. [36] (2018) propôs pela primeira vez o NasNetMobile, nosso segundo modelo pré-treinado de extrator de recursos. No conjunto de dados CIFAR-10, o modelo Nas-NetMobile alcançou uma taxa de erro de 2,4% usando um novo método de regressão chamado ScheduledDropPath. Neste estudo, este modelo é treinado e testado em imagens 224 x 224 DFU usando aproximadamente 5,3 M de parâmetros de treinamento. Segundo Saxena et al. [37] (2019), existe um modelo otimizado composto por blocos de construção básicos que foram otimizados por meio de aprendizagem por reforço. Esses blocos incorporam várias funções de pooling, convolução e convolução separável que melhoram a confiabilidade geral do modelo.

3.3.3. 3.3.3 VGGNet

Simonyan et al. [38] (2014) propuseram o VGGNet, que alcançou alto desempenho em localização e classificação de imagens, ficando em primeiro e segundo lugar, respectivamente, na competição ILSVRC. Comparado com a arquitetura AlexNet, o VGGNet exibiu uma taxa de erro impressionante (8,1%) que é muito melhor do que a arquitetura AlexNet. Para este estudo, utilizamos VGG19 como o último extrator de recursos para construir nosso framework DFU_MultiNet. O popular modelo VGG19 é organizado por 16 camadas de convolução e três camadas FC (totalmente conectadas). Os tamanhos de filtro para cada camada de convolução variam de 64 a 512, com tamanho de janela 3 x 3 para todos eles. Cinco blocos compõem este modelo, com as primeiras quatro camadas de convolução localizadas nos primeiros dois blocos e as doze restantes localizadas nos três seguintes. Uma camada MP (max pooling) após cada bloco com tamanho de janela 2 x 2 detecta os recursos mais importantes dos mapas de ativação modificados [39]. A função de ativação (ReLU) é aplicada a cada camada de convolução. Por fim, este modelo obteve um total de 20.024.384 parâmetros após o treinamento no conjunto de dados DFU, o que é mais que os outros dois modelos.

3.4. Processo de ajuste fino

A Figura 5 ilustra como integrar três modelos CNN pré-treinados para construir o modelo MTL para categorizar o conjunto de dados DFU, utilizando várias camadas FC (totalmente conectadas). Esses três modelos empregam Global-AveragePooling2D simultaneamente para nivelar em um vetor, o que é feito calculando o valor médio das imagens de entrada. Uma camada concatenada é então usada para mesclar cada vetor em um único vetor e alimentar seis camadas adicionais com uma função de ativação (softmax) para fins de ajuste fino para categorizar o conjunto de dados DFU. As explicações de cada camada DL são fornecidas abaixo.

O modelo DL enfrenta um problema significativo conhecido como overfitting, que acontece quando ele treina demais nos dados de treinamento e tem um desempenho ruim nos dados de teste [40]. Empregamos duas camadas de regularização de dropout para superar a situação de overfitting. Durante o treinamento da estrutura DFU_MultiNet, essas camadas excluíram 40% e 20% das amostras, ao mesmo tempo que melhoraram significativamente o tempo de treinamento. Além disso, tal procedimento acelera significativamente a tarefa de treinamento do conjunto de dados DFU [41].

Por outro lado, a inclusão de duas camadas BN (normalização em lote) [42] é crucial para o sucesso do nosso framework DFU_MultiNet. As principais operações desta camada são redimensionar e normalizar as amostras DFU, o que torna o modelo mais robusto e confiável.

A camada densa, também chamada de camada FC (totalmente conectada) [43], conecta todos os neurônios entre as duas camadas (anterior e atual). A principal tarefa desta camada especial é processar amostras de entrada e gerar o resultado da classificação. Em nossa abordagem, empregamos duas camadas FC, onde a primeira usa ReLU [44] e a segunda usa softmax como função de ativação. Esta camada final prevê a duração da classe e gera a previsão DFU. O softmax determina as características mais relevantes para prever a classe normal/úlcera, cujo valor de resultado varia de 0 a 1, e aciona o neurônio de acordo. Pode ser expresso como a seguinte equação:

$$\text{Softmax}(w)p = \frac{\exp(w p)}{\sum_{n=1}^N \exp(w p)} \tag{1}$$

Os resultados da combinação de vários modelos pré-treinados com camadas FC (totalmente conectadas) são relatados na tabela (consulte o Apêndice A). Esta tabela foi obtida durante a construção do framework "DFU_MultiNet" para classificação DFU. É por isso que, nesta estrutura, a camada FC final contém dois neurônios.

4. Descrição do conjunto de dados, métricas de desempenho e análise de resultados

Na configuração experimental, os hiperparâmetros utilizados no estudo e os resultados alcançados pela estrutura "DFU_MultiNet" do conjunto de dados DFU são apresentados nesta seção. Além disso, uma discussão comparativa abrangente entre a estrutura "DFU_MultiNet" e as estruturas CNN de última geração individuais é conduzida para avaliar o e

eficácia do quadro proposto.

4.1. Descrição do conjunto de dados

O conjunto de dados DFU, que está disponível online [30], foi usado para treinamento e teste. Este repositório é composto por quatro pastas. A pasta "imagens originais" contém 493 imagens de pés de diferentes pacientes com pés saudáveis e úlceras diabéticas. As fotos são do centro de diabetes do Hospital Nasiriyah, no sul do Iraque [24], e é digno de nota que o consentimento por escrito e a aprovação ética foram obtidos diligentemente de todos os pacientes relevantes e pessoas envolvidas no processo de coleta de dados. Essas fotos foram tiradas pelos especialistas com um iPad e um Samsung Galaxy Note 8 em diversas condições de iluminação e visualização. A pasta "patches" contém 543 manchas de pele normal (pele saudável) e 512 manchas de pele anormal (úlcera), que foram recortadas das amostras da pasta "imagens originais" com resolução de 224 x 224 pixels. Neste experimento, usamos amostras de pastas "patches" para treinar e testar nosso modelo.

4.2. Configuração experimental

O framework "DFU_MultiNet" proposto foi desenvolvido com a ajuda de Keras [45], para conectar Python [46] à NN (rede neural). A configuração experimental com parâmetros descritos na Tabela 5.

4.3. Métricas de desempenho

A avaliação de desempenho do framework "DFU_MultiNet" envolveu o uso de vários parâmetros estatísticos, como precisão (ACC), estatística Kappa, pontuação F1 (FS), precisão (PRE), coeficiente de correlação de Matthews (MCC), especificidade (SPE), sensibilidade (SEN) e recordação (REC). Esses parâmetros foram calculados com base nos valores de falso negativo (FN), verdadeiro positivo (TP), verdadeiro negativo (TN) e falso positivo (FP) na matriz de confusão.

Na avaliação do framework "DFU_MultiNet", **TP** refere-se à correta identificação da pele do pé positiva. **TN** refere-se à identificação precisa da pele negativa do pé. Por outro lado, **FP** denota identificação imprecisa de pele de pé positiva e **FN** refere-se à identificação incorreta de pele de pé negativa.

A seguir estão as métricas de desempenho da estrutura "DFU_MultiNet":

Precisão (ACC): ACC refere-se à relação entre amostras devidamente identificadas e o total de amostras do conjunto de dados DFU.

Tabela 5

Configuração experimental com parâmetros.

Item	Desempenho
Plataforma	Google Colab
GPU	Tesla K80
BATER	64 GB
CPU	Intel Core i5-12600K a 3700 MHz

$$Precis\tilde{a}o\ (ACC) = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{2}$$

Recall (REC): REC refere-se \u00e0 capacidade da estrutura "DFU_MultiNet" de identificar com sucesso amostras TP (verdadeiras positivas) calculando a propor\u00e7\u00e3o do total de amostras positivas no conjunto de dados DFU.

$$Recuperar\ (REC) = \tag{3}$$

Especificidade (SPE): SPE refere-se \u00e0 capacidade da estrutura "DFU_MultiNet" de identificar com sucesso amostras TN (verdadeiras negativas) calculando a divis\u00e3o do total de amostras negativas do conjunto de dados DFU, tamb\u00e9m chamada de TNR (taxa verdadeira negativa).

$$Especificidade\ (SPE) = \frac{TN}{TN + PF} \tag{4}$$

Precis\u00e3o (PRE): PRE refere-se \u00e0 capacidade da estrutura "DFU_MultiNet" de identificar com sucesso amostras TP (verdadeiramente positivas), calculando a propor\u00e7\u00e3o de amostras positivas previstas com sucesso para todos os exemplos positivos previstos do conjunto de dados DFU.

$$Precis\tilde{a}o\ (PRE) = \frac{PT}{TP + FP} \tag{5}$$

Pontua\u00e7\u00e3o F1 (FS): A m\u00e9dia harm\u00f4nica de REC e PRE \u00e9 conhecida como FS.

$$Pontua\tilde{c}\tilde{a}o\ F1(FS)=2\tilde{y}\frac{PRE\tilde{y}\ REC}{PRE + REC} \tag{6}$$

Coefficiente de correla\u00e7\u00e3o de Matthews (MCC): MCC \u00e9 um par\u00e2metro estat\u00edstico aplicado para rotulagem bin\u00e1ria. O valor \u00e9 limitado de \u00ed 1 (pior resultado) a 1 (melhor resultado). Pode ser expresso como a seguinte equa\u00e7\u00e3o:

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{7}$$

Coefficiente Kappa de Cohen (Kappa): Kappa \u00e9 utilizado para comparar as classes previstas da estrutura DFU_MultiNet com as classes reais nos dados DFU. O valor \u00e9 limitado de \u00ed 1 (pior resultado) a 1 (melhor resultado). Pode ser expresso como a seguinte equa\u00e7\u00e3o:

$$Kappa = \frac{ACC\ total - ACC\ aleat\acute{o}rio}{1 - ACC\ aleat\acute{o}rio} \tag{8}$$

4.4. Treinamento e otimiza\u00e7\u00e3o de par\u00e2metros

A Figura 6 apresenta os resultados da simula\u00e7\u00e3o do framework "DFU_-MultiNet" proposto, que s\u00e3o extra\u00eddos durante a fase de treinamento do framework. Para treinar o "DFU_MultiNet", foram empregados valores de hiperpar\u00e2metros espec\u00edficos, conforme descrito na Tabela 6. As fun\u00e7\u00f5es de otimiza\u00e7\u00e3o e perda de gradiente descendente s\u00e3o duas partes cruciais para a sele\u00e7\u00e3o

Tabela

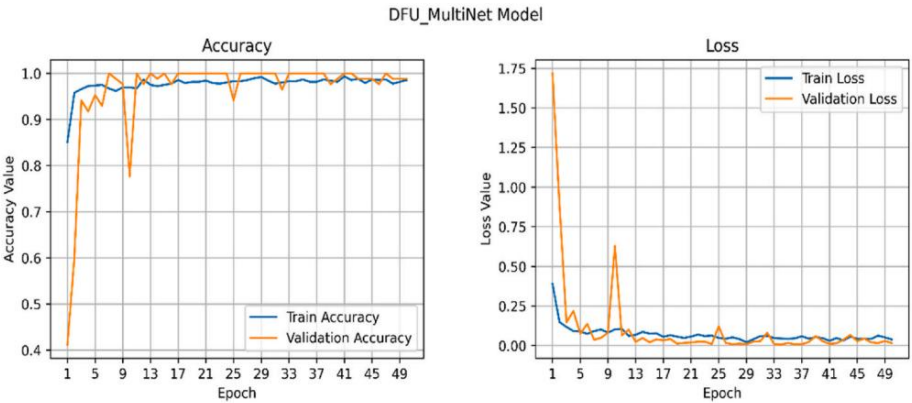
6 Par\u00e2metro de treinamento com valor para o framework "DFU_MultiNet".

N\u00famero	Par\u00e2metro	Valor
1	Otimizador	Ad\u00e3o
2	Taxa de Aprendizagem	0,0001
3	Taxa M\u00ednima de Aprendizagem	1e-7
4	Decair	0,00001
5	Paci\u00eancia	5
6	Fator	0,2
7	Fun\u00e7\u00e3o de perda	bin\u00e1rio_crossentropy
8	M\u00e9tricas	precis\u00e3o 32
9	Tamanho do batch	
10	\u00c9pocas	50

hiperpar\u00e2metros durante o treinamento da estrutura. Ao selecionar a fun\u00e7\u00e3o otimizadora para nossa estrutura, optamos por Adam [42] devido \u00e0 sua capacidade de gerenciar efetivamente gradientes esparsos em grandes conjuntos de dados, combinando as propriedades desej\u00e1veis dos otimizadores RMSProp e AdaGrad. Como nosso modelo se concentra na classifica\u00e7\u00e3o bin\u00e1ria, consideramos uma fun\u00e7\u00e3o de perda denominada entropia cruzada bin\u00e1ria. Determinar uma taxa de aprendizagem apropriada \u00e9 crucial para minimizar a fun\u00e7\u00e3o de perda, o que \u00e9 muito desafiador. Na abordagem DL, uma pequena taxa de aprendizado torna o treinamento do classificador CNN mais lento; por este motivo, o peso das atualiza\u00e7\u00f5es do modelo \u00e9 m\u00ednimo. Para mitigar esses problemas, definimos taxa de aprendizagem = 0,0001, \u00e9pocas = 50 e tamanho do lote = 32 para demonstrar um modelo eficaz. A Figura 6 mostra claramente que ap\u00f3s o treinamento da 12\u00aa \u00e9poca, a estrutura "DFU_MultiNet" ofereceu atingir 98,68% de precis\u00e3o de treinamento, 97,65% de precis\u00e3o de valida\u00e7\u00e3o, 6,17% de perda de treinamento e 10,03% de perda de valida\u00e7\u00e3o, respectivamente. A Figura 6 (a) confirma ainda que o overfitting n\u00e3o foi observado durante o processo de treinamento DFU_MultiNet. A Figura 6 (b) confirmou que a curva mostrou uma r\u00e1pida diminui\u00e7\u00e3o no valor da perda. No entanto, ocorreram algumas flutua\u00e7\u00f5es ao selecionar o tamanho de lote estreito.

Usar um conjunto de valida\u00e7\u00e3o separado para ajuste de hiperpar\u00e2metros \u00e9 crucial, pois permite a sele\u00e7\u00e3o do modelo de melhor desempenho avaliando sistematicamente diferentes hiperpar\u00e2metros. Al\u00e9m disso, permite a busca eficiente de hiperpar\u00e2metros ideais e funciona como um indicador para detectar problemas de overfitting durante o processo de ajuste. Se o conjunto de testes for usado para ajuste de hiperpar\u00e2metros, poder\u00e1 ocorrer vazamento de dados, afetando potencialmente o desempenho do modelo. Al\u00e9m disso, tal ajuste pode resultar em overfitting do conjunto de teste. \u00c9 por isso que os hiperpar\u00e2metros n\u00e3o devem ser ajustados durante a avalia\u00e7\u00e3o do conjunto de testes.

Para ajuste de hiperpar\u00e2metro, o termo par\u00e2metro "Fator" significa um fator de escala que \u00e9 aplicado para ajustar o valor de um hiperpar\u00e2metro. N\u00e3o existe um intervalo fixo para o par\u00e2metro "Fator". Por exemplo, em DL, durante o ajuste das taxas de aprendizagem, o intervalo para este par\u00e2metro \u00e9 [0,1, 10]. Como integramos tr\u00eas modelos, selecionamos taxas de aprendizagem menores (0,0001) para induzir atualiza\u00e7\u00f5es mais precisas nos pesos durante cada itera\u00e7\u00e3o, o que ajuda



6. Melhoria do treinamento da estrutura DFU_MultiNet proposta: (a) gr\u00e1fico de precis\u00e3o de treinamento versus valida\u00e7\u00e3o (valores mais altos indicam melhor desempenho) e (b) gr\u00e1fico de perda de treinamento versus valida\u00e7\u00e3o (valores mais baixos indicam melhor desempenho).

para integrar modelos. Quando treinamos nosso modelo para poucas épocas (como 10), ocorre o subajuste; novamente, enquanto treinamos para muitas épocas (como 80), ocorre o sobreajuste. Por esse motivo, selecionamos o valor de época 50 com base na análise da precisão da validação e da perda de treinamento (ver Fig. 6). Em DL, o tamanho do lote (BS) de 32 é um bom ponto inicial e uma regra geral [47]. Ele atinge um equilíbrio entre eficiência computacional e precisão do modelo. BSs grandes podem agilizar o treinamento, mas podem correr o risco de overfitting e redução da precisão, enquanto BSs menores podem ser demorados e caros do ponto de vista computacional.

4.5. Análise de resultados

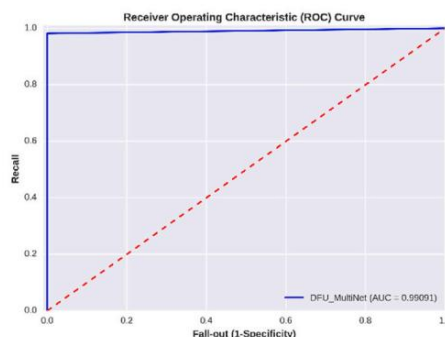
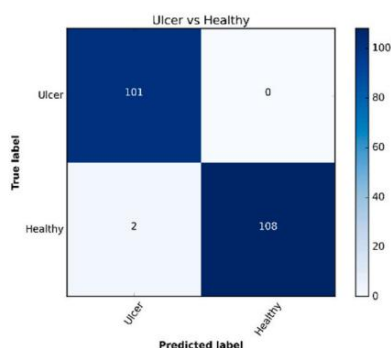
A Figura 7 apresenta as curvas CM (matriz de confusão) e ROC (característica de operação do receptor) para o conjunto de dados DFU, utilizando a estrutura "DFU_MultiNet". A estrutura fundiu três algoritmos de aprendizagem por transferência renomados, nomeadamente DenseNet201, VGG19 e NasNetMobile. Ao aproveitar os recursos de fusão, a abordagem proposta categoriza efetivamente se um pé diabético é ulcerado ou saudável. A Figura 7 (a) revela que a estrutura "DFU_MultiNet" proposta classifica com precisão 101 imagens de pele de úlcera e 108 imagens de pele saudável. Notavelmente, a estrutura classifica erroneamente apenas dois casos de pele saudável. Notavelmente, uma vantagem significativa da estrutura é o seu desempenho impecável na classificação incorreta de pele sem úlcera no conjunto de dados DFU. A estrutura "DFU_MultiNet" demonstra notável consistência e estabilidade, como evidenciado pela AUC alcançada (0,99091) na Fig. Esta alta pontuação AUC indica o forte desempenho do modelo. Além disso, as avaliações individuais de todos os modelos no conjunto de dados DFU aumentam ainda mais a robustez da estrutura "DFU_MultiNet". A Tabela 7 comprova a superioridade da estrutura "DFU_MultiNet" entre a estrutura "DFU_MultiNet" e cinco modelos CNN adicionais. Os resultados indicam que a estrutura "DFU_MultiNet" atinge métricas impressionantes como precisão de 1,00, recall de 0,982, pontuação f1 de 0,991, kappa de 0,981 e MCC de 0,981, superando todos os outros modelos ultramodernos. Notavelmente, DenseNet201 e VGG19 também demonstram forte desempenho com valores de precisão de 0,976 e 0,981, respectivamente. Enquanto isso, o VGG19 e o MobileNet apresentam desempenho comparável na capacidade de detectar a pele dos pés.

5. Discussão

A incidência de infecções do pé diabético (DFI) e problemas relacionados ao diabetes pode ser causada pela não obediência a uma dieta saudável e à falta de precauções de segurança adequadas entre os diabéticos afetados. Garantir orientação e cuidadores adequados aos pacientes diabéticos pode ser fundamental para enfrentar esses desafios. Além disso, novas técnicas de diagnóstico, terapia e previsão foram criadas como resultado da aplicação de tecnologia para controlar o diabetes. A estrutura proposta é baseada na arquitetura DL de conjunto paralelo heterogêneo que aprendeu recursos em paralelo a partir de amostras de entrada por meio de três modelos pré-treinados (ou seja, DenseNet201, VGG19 e NasNetMobile). Esta técnica poderosa pode ser aproveitada em outras máquinas clínicas

aplicações de aprendizagem. Ao utilizar esta estrutura, os profissionais de saúde podem tomar melhores decisões em relação à inscrição de pacientes em ensaios clínicos, otimizar os processos de desenvolvimento de medicamentos e integrar perfeitamente dados de diversas fontes. Como resultado, esta rede desempenhará um papel vital para aumentar a eficiência e a eficácia da investigação clínica e dos sistemas de saúde. Conforme mostrado na Tabela 7, os resultados da estrutura sugerida são mais confiáveis e robustos do que os dos modelos atuais. Este estudo sugere um novo método para classificar a pele dos pés afetados pelo diabetes, empregando uma estrutura híbrida "DFU_MultiNet" em imagens DFU. A Tabela 8 fornece um resumo do desempenho da estrutura "DFU_MultiNet" em comparação com pesquisas anteriores que empregaram o mesmo conjunto de dados, mas com várias estruturas, profundidades e parâmetros. Nossa abordagem, que combina previsões de vários modelos pré-treinados, oferece diversas vantagens para lidar com conjuntos de dados desequilibrados. Ele efetivamente atenua o overfitting e minimiza a possibilidade de ruído impactar o rótulo minoritário. Esta técnica também atribui maior peso à classe minoritária, resultando em melhor classificação para o conjunto de dados desequilibrado. Como usamos três modelos diferentes, se um modelo não consegue extrair recursos de alguns pontos de dados no conjunto de dados, outros extraem esses recursos, o que pode aumentar a capacidade de adaptação às mudanças no conjunto de dados desequilibrado. Além disso, treinar esses modelos em paralelo no mesmo conjunto de dados reduz significativamente o tempo geral de treinamento ao lidar com um conjunto de dados desequilibrado. A Tabela 8 torna óbvio que, quando comparado com os estudos anteriores, a estrutura sugerida oferece a maior precisão para o diagnóstico de úlceras do pé diabético (UPD). Notavelmente, a combinação de todos os algoritmos pré-treinados permite que a estrutura alcance uma precisão de classificação de 99% para o conjunto de dados DFU. A curva ROC apresentada na Fig. 8 compara o desempenho do DFU_MultiNet, um framework proposto, com diversas redes de aprendizagem por transferência, como DenseNet201, VGG16, VGG19, Nas-NetMobile, MobileNet e DFU_MultiNet, usando a mesma partição de dados. Os resultados demonstram que o DFU_MultiNet alcança um resultado excelente em relação às redes de aprendizagem por transferência padrão na classificação precisa de úlceras versus amostras saudáveis. A Fig. 9 exibe algumas das amostras de pele de pés afetados por diabéticos que são previstas com precisão com a ajuda da nova estrutura DFU_MultiNet.

A adaptação da abordagem "DFU_MultiNet" para uso clínico envolve várias etapas e considerações cruciais para garantir sua eficácia e segurança em um ambiente de saúde. Inicialmente, o conjunto de dados deve passar por anotação com rótulos verdadeiros para distinguir pés saudáveis daqueles com úlceras de pé diabético (DFUs). Para aumentar a transparência e a compreensão, várias técnicas explicáveis de IA (por exemplo, LIME e SHAP) são empregadas para interpretar a estrutura, fornecendo aos médicos informações sobre a utilização de recursos do modelo para classificação. Posteriormente, é imperativa a integração do modelo "DFU_MultiNet" em sistemas clínicos existentes ou ferramentas de diagnóstico utilizadas por profissionais de saúde. As previsões do modelo devem ser apresentadas num formato claro e prático, garantindo que os médicos possam compreender prontamente e agir de acordo com os resultados do diagnóstico. O monitoramento contínuo do desempenho do modelo é essencial, abrangendo precisão diagnóstica, tratamento



7. (a) Matriz de confusão (b) Curva ROC para a estrutura "DFU_MultiNet".

Tabela 7

Resultados obtidos a partir da estrutura "DFU_MultiNet" e cinco algoritmos individuais ajustados no conjunto de dados DFU.

Modelo	Precisão	Precisão	Lembrar	pontuação F1	Especificidade	CUA	Taxa de erro	MCC	Kappa
VGG19	0,867	1,00	0,745	0,854	1,00	0,873	0,133	0,764	0,737
VGG16	0,981	1,00	0,964	0,981	1,00	0,982	0,019	0,963	0,962
NasNetMobile	0,773	1,00	0,564	0,721	1,00	0,782	0,227	0,618	0,553
DenseNet201	0,976	1,00	0,955	0,977	1,00	0,977	0,024	0,954	0,953
MobileNet	0,867	0,894	0,845	0,869	0,891	0,868	0,133	0,736	0,735
DFU_MultiNet	0,991	1,00	0,982	0,991	1,00	0,991	0,009	0,981	0,981

Tabela 8

Comparação da estrutura "DFU_MultiNet" com técnicas existentes para conjunto de dados DFU.

Papel	Precisão (%)	Precisão (%)	Lembrar (%)	Pontuação F1 (%)	Método de Classificação	Treinamento e validação Amostra	Teste Amostra
Thotad et al. (2022) [18]	98,97	99	98	98	EficienteNet	1350	338
Juan et al. (2021) [19]	97,8	97,4	98,2	97,8	DFU_VIRnet	12.600	600
K. Das et al. (2021) [21]	96,4	92,6	98,4	95,4	DFU_SPNet	3491	336
Alzubaidi et al. (2021) [22]	-	97,3	94,5	95,8	CNN híbrida	16.731	322
Alzubaidi et al. (2020) [23]	-	95,4	93,6	94,5	DFU_QUTNet+ SVM	16.731	322
Goyal et al. (2017) [26]	92,5	94,5	-	93,9	DFUNet	22.605	172
Estrutura Proposta	99,1	100	98,2	99,1	DFU_MultiNet	6752	211

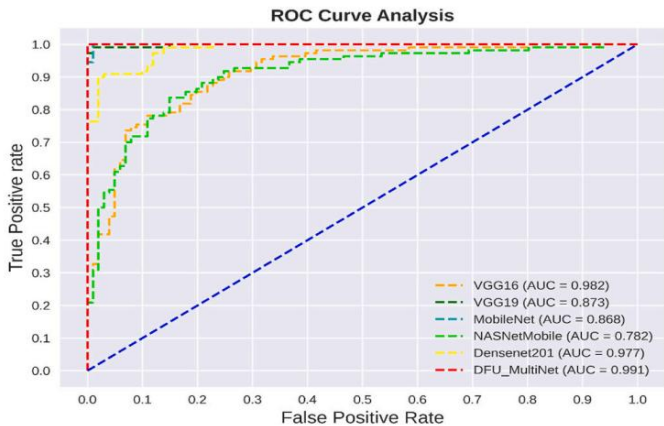


Figura 8. Curva ROC para vários algoritmos de aprendizagem por transferência.

planejamento e avaliação dos resultados dos pacientes em cenários clínicos reais.

Neste contexto clínico, os médicos podem empregar a estrutura como uma ferramenta prospectiva de diagnóstico auxiliado por computador (CAD). Eles podem inserir facilmente imagens DFU na interface da ferramenta CAD, onde a estrutura "DFU_MultiNet" realiza a extração de recursos analisando as imagens. Esses recursos extraídos são então amalgamados por meio de uma camada de soma, gerando uma representação abrangente da imagem para avaliação diagnóstica e tomada de decisão.

5.1. Limitações do Estudo

Neste quadro, é viável diagnosticar se a amostra conferida é uma úlcera ou apenas saudável. Não pode fornecer uma avaliação em tempo real da gravidade da dor ou dos níveis de complexidade.

Para maximizar o potencial desta estrutura em áreas clínicas com poder computacional limitado, trabalhos futuros devem priorizar o desenvolvimento de algoritmos de treinamento distribuído e aprendizagem federada, juntamente com a criação de sistemas de IA interpretáveis. Além disso, o desenvolvimento de modelos DL que operem eficientemente em dispositivos de ponta e plataformas móveis será crucial para áreas clínicas onde não há capacidade computacional significativa disponível.

Para testar o desempenho da abordagem DFU_MultiNet para classificação DFU, treinamos esta abordagem, bem como cinco

modelos pré-treinados (ou seja, DenseNet201, VGG19, VGG16, MobileNet e NasNetMobile) no mesmo conjunto de dados com a mesma taxa de divisão. Esses cinco modelos pré-treinados são treinados da seguinte forma: primeiro, imagens de entrada com tamanho 224 x 224 são alimentadas nesses modelos para produzir um conjunto de mapas de recursos de saída, extraindo recursos da imagem de entrada. Uma camada Global-AveragePooling2D é então aplicada após cada modelo para reduzir o mapa de recursos de saída a um vetor unidimensional. Isso é feito calculando a média de todos os valores no mapa de características. Depois disso, seis camadas de ajuste fino são adicionadas seguindo a camada GlobalAveragePooling2D, uma após a outra que segue na maneira dropout - batch_normalization - densa e novamente dropout - batch_normalization - densa. Essas seis camadas são descritas na Seção 3.3, que também é usada para construir o modelo MTL.

Embora o modelo proposto apresente bom desempenho neste conjunto de dados DFU, ele fornecerá um resultado melhor e mais preciso em um conjunto de dados maior. Durante o treinamento em um grande conjunto de dados, esta estrutura aprenderá mapas de recursos exclusivos de diferentes amostras, o que a torna uma poderosa ferramenta de detecção de úlceras. Nos cuidados de saúde e na imagiologia médica, esta poderosa ferramenta promete detectar doenças com mais precisão e mais cedo, permitindo um melhor planejamento do tratamento e melhores resultados para os pacientes. No futuro, esta estrutura será aplicável para quantificação precisa de diferentes parâmetros em amostras de DFU, como tamanho da úlcera, densidade do tecido, volume ou taxa de crescimento. Esses dados quantitativos facilitarão o monitoramento contínuo do paciente e melhorarão o planejamento do tratamento.

6. Conclusão

Exames regulares dos pés são essenciais para que os diabéticos identifiquem possíveis lesões, além de avaliações abrangentes de problemas arteriais periféricos e neuropatias, pois essas condições têm potencial para resultar na formação de úlceras ou feridas. A DFU pode ser prevenida com exames regulares dos pés, controle da glicose, ocultação do paciente, calçados adequados e tratamento oportuno para infecções pré-ulcerativas. Nesta tarefa, oferecemos uma estrutura híbrida inovadora chamada "DFU_MultiNet" para diagnosticar a pele afetada pelo diabetes a partir de amostras de pés com mais precisão e consistência. A estrutura "DFU_MultiNet" é baseada em uma abordagem de extração de recursos e ajuste fino que permite que vários algoritmos CNN pré-treinados extraiam e mesquem mapas de recursos em paralelo para classificação DFU. Ele é desenvolvido de forma equilibrada que pode controlar diversos conjuntos de dados DFU. Os resultados dos experimentos mostram que a estrutura "DFU_MultiNet", que supera os algoritmos pré-treinados da CNN separados e todos os outros algoritmos modernos relatados no trabalho publicado, atinge 99%



9. "DFU_MultiNet" testado com sucesso em algumas amostras DFU.

precisão. Esta estrutura híbrida superou os métodos de classificação DFU mais recentes. Considerando os resultados promissores, temos forte confiança no potencial da nossa estrutura "DFU_MultiNet" como uma excelente ferramenta para ajudar os médicos na detecção e diagnóstico eficiente da DFU. Além disso, tem um desempenho admirável na localização de úlceras cutâneas, melhorando a probabilidade de sobrevivência.

Em estudos futuros, esta estrutura híbrida deve ser expandida para detectar e classificar a DFU em isquemia, neuropatia, osteomielite ou artropatia de Charcot.

Disponibilidade de dados

Os dados que apoiam as conclusões deste estudo estão disponíveis no autor correspondente mediante solicitação razoável.

Financiamento

Nenhum.

Apêndice A

Detalhes da estrutura "DFU_MultiNet".

Camada (tipo)	Formato de saída	Parâmetro #	Conectado a
input_1	(224, 224, 3) (7,	0	
rede_densa201	7, 1920) (7, 7,	18321984	input_1[0][0]
NASNet	1056) (7, 7,	4269716	input_1[0][0]
vgg19	512) (1920)	20024384 0	input_1[0][0]
global_average_pooling2d	(1056)		Densenet201[0][0]
global_average_pooling2d	(512)	0	NASNet[0][0]
global_average_pooling2d	(3488)	0	vgg19[0][0]
concatenate_4		0	global_average_pooling2d[0][0] global_average_pooling2d_1[0][0]
abandono	(3488)	0	global_average_pooling2d_2[0][0]
batch_normalization_denso	(3488)	13952	concatenate_4[0][0]
	(128)	446592	dropout[0][0]
dropout_1	(Nenhum, 128)	0	batch_normalization [0][0] denso[0]
batch_normalization_1	(Nenhum, 128)	512	[0] dropout_1[0]
			[0]

(Continua na próxima página)

(contínuo)

Camada (tipo)	Formato de saída	Parâmetro #	Conectado a
denso_1	(Nenhum, 2)	258	batch_normalization_1[0][0]
Parâmetros totais: 43.077.398			
Parâmetros treináveis: 42.804.372			
Parâmetros não treináveis: 273.026			

Referências

[1] Diabetes [On-line]. Disponível: <https://www.who.int/health-topics/diabetes>. [Acessado em 29 de setembro de 2023].

[2] Wild S, Roglic G, Green A, Sicree R, King H. Prevalência global de diabetes. *Diabetes Care*, maio de 2004;27(5):1047–53. <https://doi.org/10.2337/diacare.27.5.1047>.

[3] Aguirre F, Brown A, Cho NH, Dahlquist G, Dodd S, Dunning T, Hirst M, Hwang C, Magliano D, Patterson C, Scott C. Atlas de diabetes IDF. 2013.

[4] Mariam TG, Alemayehu A, Tesfaye E, Mequannt W, Temesgen K, Yetwale F, Limenih MA. Prevalência de úlcera no pé diabético e fatores associados entre pacientes diabéticos adultos que frequentam a clínica de acompanhamento de diabéticos no Hospital de Referência da Universidade de Gondar, Noroeste da Etiópia, 2016: estudo transversal de base institucional. *J Diabetes Res* 2017. 2017.

[5] Almobarak AO, Awadalla H, Osman M, Ahmed MH. Prevalência de pé diabético ulceração e fatores de risco associados: um antigo e ainda importante problema de saúde pública em Cartum, Sudão? *Ann Transl Med* 2017;5(17).

[6] Singh G, Gupta S, Chanda A. Modelagem biomecânica de úlceras de pé diabético: um estudo computacional. 127:110699.

[7] Pourkazemi A, Ghanbari A, Khojiamli M, Balo H, Hemmati H, Jafaryparvar Z, Motamed B. Cuidados com o pé diabético: conhecimento e prática. *Transtorno Endocr BMC* 2020; 20:1–8.

[8] Ghosh P, Valia R. Carga de úlceras de pé diabético na Índia: cenário de evidências da literatura publicada. *Valor Saúde* 2017;20(9):A485.

[9] Najafi B, Reeves ND, Armstrong DG. Aproveitar tecnologias inteligentes para melhorar o tratamento de úlceras de pé diabético e prolongar os dias sem úlceras em remissão. Pesquisa e revisões sobre diabetes/metabolismo 2020;36:e3239.

[10] Formica D, Schena E. Sensores inteligentes para aplicações médicas e de saúde. *Sensores* 2021;21(2):543.

[11] Veredas FJ, Luque-Baena RM, Martín-Santos FJ, Morilla-Herrera JC, Morente L. Avaliação de imagens de feridas com aprendizado de máquina. *Neurocomputação* 2015;164: 112–22.

[12] Alzubaidi L, Fadhel MA, Al-Shamma O, Zhang J, Santamaría J, Duan Y. Robusto aplicação de novas ferramentas de aprendizagem profunda: um estudo experimental em imagens médicas. *Ferramentas e Aplicativos Multimídia*; 2022. pág. 1–29.

[13] Mostafiz R, Rahman MM, Uddin MS. Classificação de pólipos gastrointestinais através de decomposição em modo empírico e características neurais. *SN Appl Sci* junho. 2020;2(6): 1143. <https://doi.org/10.1007/s42452-020-2944-4>.

[14] Vas PR, Kavarthapu V. Tratamento da doença do pé diabético. Em: *Diabético neuropatia*. Elsevier; 2022. pág. 235–58.

[15] Cassidy B, Kendrick C, Reeves ND, Pappachan JM, O’Shea C, Armstrong DG, Sim, MH. Grande desafio da úlcera no pé diabético 2021: avaliação e resumo. In: Grande desafio das úlceras do pé diabético: segundo desafio, DFUC 2021, realizado em conjunto com MICCAI 2021, Estrasburgo, França, 27 de setembro de 2021, procedimentos. Cham: Springer International Publishing; 2022. pág. 90–105.

[16] Yap MH, Hachiuma R, Alavi A, Brüngel R, Cassidy B, Goyal M, Zhu H, Rückert J, Olshansky M, Huang X, Saito H. Aprendizado profundo na detecção de úlceras de pé diabético: uma avaliação abrangente. *Comput Biol Med* 2021;135:104596.

[17] Najafi B, Mohseni H, Grewal GS, Talal TK, Menzies RA, Armstrong DG. Um tecido inteligente baseado em fibra óptica (meias inteligentes) para gerenciar fatores de risco biomecânicos associados à amputação do pé diabético. *J Diabetes Sci Technol* 2017;11(4): 668–77.

[18] Kaselimi M, Protopapadakis E, Doulamis A, Doulamis N. Uma revisão de sensores não invasivos e modelos de inteligência artificial para monitoramento de pés diabéticos. *Front Physiol* 2022:2230.

[19] Thotad PN, Bharamagoudar GR, Kallur SS. Aprendizado de máquina baseado em boosting abordagens para previsão de diabetes usando dados demográficos e de saúde da Índia - dados de 2021. 2023.

[20] Reyes-Luévano J, Guerrero-Viramontes JA, Romo-Andrade JR, Funes-Gallanzi M. DFU_VIRnet: uma nova CNN de infravermelho visível para melhorar a classificação de úlceras de pé diabético e detecção precoce de zonas de risco de úlcera. 2022.

[21] Doulamis A, Doulamis N, Angeli A, Lazaris A, Luthman S, Jayapala M, Silbermangel G, Napp A, Lazarou I, Karalis A, Hoveling R. Um dispositivo não invasivo baseado em fotônica para monitoramento de úlceras de pé diabético: componentes arquitetônicos/sensoriais e especificações técnicas. *Invenções* 2021;6(2):27.

[22] Das SK, Roy P, Mishra AK. DFU_SPNet: uma CNN baseada em camadas de convolução paralelas empilhadas para melhorar a classificação de úlceras de pé diabético. *TIC Express* 2022;8(2):271–5.

[23] Alzubaidi LAITH, Abbood AA, Fadhel MA, Al-Shamma OMRAN, Zhang JIN. Comparação de modelos de redes neurais convolucionais híbridas para classificação de úlceras de pé diabético. *J Eng Sci Technol* 2021;16(3):2001–17.

[24] Alzubaidi L, Fadhel MA, Olewi SR, Al-Shamma O, Zhang J. DFU_QUTNet: classificação de úlcera de pé diabético usando nova rede neural convolucional profunda. *Multimed Tool Appl* 2020;79(21–22):15655–77.

[25] Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. Um convnet para a década de 2020 (2022. 2022. URL: <https://arxiv.org/abs/2201.03545>.

[26] Tan M, Le Q. Efficientnet: repensando o escalonamento do modelo para neural convolucional redes. In: *Conferência Internacional sobre Aprendizado de Máquina*. PMLR; 2019, maio. pág. 6105–14.

[27] Goyal M, Reeves ND, Davison AK, Rajbhandari S, Spragg J, Yap MH. Dfunet: redes neurais convolucionais para classificação de úlceras de pé diabético. *Transações IEEE sobre Tópicos Emergentes em Inteligência Computacional* 2018;4(5):728–39.

[28] Wang L, Pedersen PC, Agu E, Strong DM, Tulu B. Determinação de área de imagens de úlcera de pé diabético usando uma classificação baseada em SVM em cascata de dois estágios. *IEEE (Inst Electr Electron Eng) Trans Biomed Eng* 2016;64(9):2098–109.

[29] Goyal M, Yap MH, Reeves ND, Rajbhandari S, Spragg J. Redes totalmente convolucionais para segmentação de úlcera de pé diabético. In: *Conferência internacional IEEE 2017 sobre sistemas, homem e cibemética (SMC)*. IEEE; 2017, outubro. pág. 618–23.

[30] Conjunto de dados: úlcera do pé diabético (DFU). Link disponível: <https://www.kaggle.com/laithji/ulcera-do-pe-diabetico-dfu>.

[31] Pan SJ, Yang Q. Uma pesquisa sobre aprendizagem por transferência. *IEEE Trans Knowl Data Eng* outubro de 2010;22(10):1345–59.

[32] Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Redes neurais convolucionais profundas para detecção auxiliada por computador: arquiteturas CNN, características de conjuntos de dados e aprendizagem por transferência. *IEEE Trans Med Imag* maio de 2016;35(5):1285–98.

[33] Russakovsky O, et al. Desafio de reconhecimento visual em grande escala do ImageNet. arXiv; Janeiro. 29, 2015 [On-line]. Disponível em <http://arxiv.org/abs/1409.0575>. [Acessado em 23 de setembro de 2023].

[34] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densamente conectado redes convolucionais. In: *Anais da conferência IEEE sobre visão computacional e reconhecimento de padrões*; 2017. pág. 4700–8.

[35] Hao W, Zhang Z. Rede de conectividade densa destilada espaço-temporal para reconhecimento de ação de vídeo. *Reconhecimento de padrões* 2019;92:13–24.

[36] Zoph B, Vasudevan V, Shlens J, Le QV. Aprendendo arquiteturas transferíveis para reconhecimento de imagem escalonável. In: *Anais da conferência IEEE sobre visão computacional e reconhecimento de padrões*; 2018. pág. 8697–710.

[37] Saxen F, Werner P, Handrich S, Othman E, Dinges L, Al-Hamadi A. Detecção de atributos faciais com mobilenetv2 e nasnet-mobile. In: 11º Simpósio Internacional de Processamento e Análise de Imagens e Sinais (ISPA) 2019. IEEE; 2019, setembro. pág. 176–80.

[38] Simonyan K, Zisserman A. Redes convolucionais muito profundas para reconhecimento de imagens em grande escala. 2014. *pré-impressão arXiv arXiv:1409.1556*.

[39] Scherer D, Müller A, Behnke S. Avaliação de operações de pooling em arquiteturas convolucionais para reconhecimento de objetos. In: *Redes neurais artificiais – ICANN 2010: 20ª conferência internacional, Thessaloniki, Grécia, 15 a 18 de setembro de 2010, procedimentos, Parte III* 20. Springer Berlin Heidelberg; 2010. pág. 92–101.

[40] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Abandono: a maneira simples de evitar o overfitting das redes neurais. *J Mach Learn Res* 2014;15 (1):1929–58.

[41] Ioffe S, Szegedy C. Normalização em lote: acelerando o treinamento profundo da rede por reduzindo a mudança interna de covariáveis. In: *Conferência Internacional sobre Aprendizado de Máquina*. pmlr; 2015, junho. pág. 448–56.

[42] Koushik J. Compreendendo redes neurais convolucionais. 2016. *pré-impressão arXiv arXiv:1605.09081*.

[43] Dahl GE, Sainath TN, Hinton GE. Melhorando redes neurais profundas para LVCSR usando unidades lineares retificadas e dropout. In: *Conferência internacional IEEE 2013 sobre acústica, fala e processamento de sinais*. IEEE; 2013, maio. pág. 8609–13.

[44] Chollet F. "Keras. Biblioteca de aprendizagem profunda para theano e tensorflow 2015;7(8):T1. URL: <https://keras.io/k>.

[45] Kerkar N. Aprendizado profundo com Python. Berkeley, CA: Apress; 2017. <https://doi.org/10.1007/978-1-4842-2766-4>.

[46] Kingma Diederik P, Ba Jimmy. Adam: Um Método para Otimização Estocástica. 2014. *pré-impressão arXiv arXiv:1412.6980*.

[47] Mostafiz R, Uddin MS, Uddin KMM, Rahman MM. COVID-19 juntamente com outros diagnósticos de infecção torácica usando R-CNN mais rápido e rede adversária generativa. *ACM Trans. Sistema de Algoritmos Espaciais*. Setembro de 2022;8(3):1–21. <https://doi.org/10.1145/3520125>.