

Listas de conteúdos disponíveis em [ScienceDirect](#)

Computadores em Biologia e Medicina

página inicial da revista: www.elsevier.com/locate/complbiomed

Reconhecimento de isquemia e infecção em úlceras de pé diabético: conjunto de dados e técnicas

Manu Goyal ^{a, b}, Neil D. Reeves, Satyan Rajbhandari, Naseer Ahmad ^d, Chuan Wang ^e,
Moi Hoon Yap ^{a, y}

^a Centro de Ciências Computacionais Avançadas, Faculdade de Ciências e Engenharia, Universidade Metropolitana de Manchester, M1 5GD, Manchester, Reino Unido ^b Centro de Pesquisa para Ciências Musculoesqueléticas e Medicina Esportiva, Faculdade de Ciências e Engenharia, Universidade Metropolitana de Manchester, M1 5GD, Manchester, Reino Unido

^c Lancashire Teaching Hospital, PR2 9HT, Preston, Reino Unido

^d Universidade de Manchester e Manchester Royal Infirmary, M13 9WL, Manchester, Reino Unido

^e Departamento de Endocrinologia, Hospital Memorial Sun Yat-sen, Universidade Sun Yat-sen, Guangzhou 510120, PR China

INFORMAÇÕES DO ARTIGO

Palavras-chave:

Úlceras no pé diabético

Aprendizado profundo

Isquemia

Infecção

Aprendizado de máquina

ABSTRATO

O reconhecimento e análise de úlceras de pé diabético (DFU) usando métodos computadorizados é uma área de pesquisa emergente com a evolução de algoritmos de aprendizado de máquina baseados em imagens. A pesquisa existente que utiliza métodos visuais computadorizados concentra-se principalmente no reconhecimento, detecção e segmentação da aparência visual do DFU, bem como na classificação dos tecidos. De acordo com os sistemas de classificação médica da DFU, a presença de infecção (bactérias na ferida) e isquemia (fornecimento sanguíneo inadequado) tem implicações clínicas importantes para a avaliação da DFU, que são utilizadas para prever o risco de amputação. Neste trabalho, propomos um novo conjunto de dados e técnicas de visão computacional para identificar a presença de infecção e isquemia em DFU. Esta é a primeira vez que um conjunto de dados DFU com rótulos verdadeiros de isquemia e casos de infecção é introduzido para fins de pesquisa. Para a abordagem de aprendizado de máquina artesanal, propomos um novo descritor de recursos, denominado Superpixel Color Descriptor. Em seguida, usamos o modelo Ensemble Convolutional Neural Network (CNN) para um reconhecimento mais eficaz de isquemia e infecção. Propomos a utilização de um método natural de aumento de dados, que identifica a região de interesse nas imagens dos pés. Finalmente, avaliamos o desempenho das nossas técnicas propostas na classificação binária, ou seja, isquemia versus não isquemia e infecção versus não infecção. No geral, nosso método teve melhor desempenho na classificação de isquemia do que de infecção. Descobrimos que nossos algoritmos de aprendizado profundo Ensemble CNN propostos tiveram melhor desempenho para ambas as tarefas de classificação em comparação com algoritmos de aprendizado de máquina artesanais, com 90% de precisão na classificação de isquemia e 73% na classificação de infecção.

1. Introdução

As úlceras do pé diabético (DFUs) são uma complicação importante do diabetes que pode levar à amputação do pé ou membro. O tratamento de DFUs é um grande problema global de saúde, resultando em altos custos de cuidados e taxa de mortalidade. O reconhecimento da infecção e da isquemia é muito importante para determinar os fatores que predizem o progresso da cicatrização da UPD e o risco de amputação. A isquemia, a falta de circulação sanguínea, se desenvolve devido a complicações crônicas do diabetes. Isto pode resultar em gangrena da úlcera do pé diabético, que pode exigir a amputação de parte do pé ou da perna se não for reconhecida e tratada precocemente. O conhecimento detalhado da anatomia vascular da perna, e particularmente da isquemia, permite que os médicos especialistas tomem melhores decisões ao estimar a possibilidade de

Cura DFU, dado o suprimento de sangue existente [1]. Em estudos anteriores, estima-se que pacientes com isquemia crítica apresentam uma taxa de perda de membros em três anos de cerca de 40% [2]. Pacientes com DFU ativa, e particularmente aqueles com isquemia ou gangrena, devem ser examinados quanto à presença de infecção. Aproximadamente, 56% dos DFU são infectados e 20% das infecções por DFU levam à amputação de um pé ou membro [3–5]. Num estudo recente, 785 milhões de pacientes com diabetes nos EUA entre 2007 e 2013 sugeriram que a DFU e as infecções associadas constituem um poderoso factor de risco para visitas ao serviço de urgência e internamento hospitalar [6].

Existem vários sistemas de classificação DFU como Wagner, Universidade do Texas e sistemas de classificação SINBAD, que incluem

^y Autor correspondente.

Endereço de e-mail: M.Yap@mmu.ac.uk (MH Sim).

<https://doi.org/10.1016/j.complbiomed.2020.103616>

Recebido em 27 de outubro de 2019; Recebido em formato revisado em 9 de janeiro de 2020; Aceito em

9 de janeiro de 2020. Disponível on-

line em 10 de janeiro de 2020. 0010-4825/© 2020 The Authors. Publicado pela Elsevier Ltd.

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Este é um artigo de acesso aberto sob a licença CC BY-NC-ND

informações sobre o local da DFU, área, profundidade, presença de neuropatia, presença de isquemia e infecção [7–9]. SINBAD significa S (local), I (isquemia), N (neuropatia), B (infecção bacteriana), A (área), D (profundidade). Este artigo enfoca isquemia e infecção, que são definidas a seguir:

1. Isquemia: Suprimento sanguíneo inadequado que pode afetar a cicatrização da DFU. A isquemia é diagnosticada pela palpação dos pulsos dos pés e pela medição da pressão arterial no pé e nos dedos dos pés. O aspecto visual da isquemia pode ser indicado pela presença de má reperfusão no pé ou dedos pretos gangrenados (morte de tecidos em parte do pé). Do ponto de vista da visão computacional, estes podem ser indícios importantes da presença de isquemia na DFU.
2. Infecção bacteriana: A infecção é definida como infecção bacteriana de tecidos moles ou óssea na DFU, que se baseia na presença de pelo menos dois achados clássicos de inflamação ou purulência. É muito difícil determinar a presença de infecções no pé diabético a partir de imagens DFU, mas o aumento da vermelhidão dentro e ao redor da úlcera e a coloração purulenta podem fornecer indicações. No sistema médico, os exames de sangue são realizados como teste diagnóstico padrão-ouro. Além disso, no presente conjunto de dados, as imagens foram capturadas após o desbridamento de tecidos necróticos e desvitalizados que remove uma indicação importante da presença de infecção na DFU.

Em trabalho relacionado, Netten et al. [10] descobriram que os médicos alcançaram baixa validade e confiabilidade para avaliação remota de DFU em imagens de pés. Portanto, fica claro que analisar essas condições a partir de imagens é uma tarefa difícil para os médicos. Em várias aplicações de reconhecimento de imagem, como imagens médicas e tarefas de processamento de linguagem natural, os algoritmos de aprendizado de máquina tiveram melhor desempenho do que humanos qualificados, incluindo médicos [11–13].

O diagnóstico anterior de DFU baseado em imagem de última geração é composto de vários estágios, incluindo pré-processamento de imagem, segmentação de imagem, extração de recursos e classificação. Veredas et al. [14] propuseram o uso de recursos de cor e textura da área segmentada e rede neural multicamadas para realizar a classificação do tecido para distinguir entre tecido em cicatrização e pele para predição de cicatrização. Wannous et al. [15] realizaram classificação de tecidos a partir de descritores de região de cor e textura em um modelo 3-D para a ferida.

Wang et al. [16] usaram um classificador em cascata de dois estágios para determinar os limites do DFU para determinação da área do DFU. Grandes progressos no campo do aprendizado de máquina baseado em imagens, especialmente algoritmos de aprendizado profundo, permitem o uso extensivo de dados de imagens médicas com modelos ponta a ponta para fornecer melhor diagnóstico, tratamento e previsão de doenças [17,18]. Modelos de aprendizado profundo para DFU, predominantemente liderados por trabalhos de nosso laboratório, alcançaram alta precisão no reconhecimento de DFUs com algoritmos de aprendizado de máquina [19–22].

As principais questões e desafios envolvidos na avaliação de DFU usando métodos de aprendizado de máquina a partir de imagens de pés são os seguintes: (1) uma grande carga de tempo envolvida na coleta de dados e na rotulagem especializada das imagens de DFU; (2) alta similaridade entre classes e variações intraclasses dependem das diferentes classificações de DFU; (3) não padronização do conjunto de dados DFU, como distância da câmera ao pé, orientação da imagem e condições de iluminação; (4) falta de metadados, como etnia do paciente, idade, sexo e tamanho do pé.

O diagnóstico preciso de isquemia e infecção requer o estabelecimento de uma boa história clínica, exame físico, exames de sangue, estudo bacteriológico e estudo Doppler dos vasos sanguíneos das pernas. Estes testes e recursos nem sempre estão disponíveis para médicos em todo o mundo. Daí a necessidade de uma solução para informar o diagnóstico, como a que propomos neste artigo. Os especialistas que trabalham na área da ulceração do pé diabético têm boa experiência em prever a presença de isquemia ou infecção subjacente simplesmente observando a úlcera. Nosso objetivo é replicar isso no aprendizado de máquina. Para aumentar a confiabilidade da anotação, dois especialistas prevêm a presença de isquemia e infecção por DFU

imagens. Devido aos altos riscos de infecção e isquemia na DFU, levando à admissão hospitalar do paciente e à amputação [23], o reconhecimento de infecção e isquemia na DFU com métodos de aprendizado de máquina econômicos é um passo muito importante para o desenvolvimento de um sistema computadorizado completo. Sistema de avaliação DFU para monitoramento remoto no futuro.

2. Conjunto de dados DFU e rotulagem especializada

Para classificação binária de isquemia e infecção em DFU, apresentamos um conjunto de dados de 1.459 imagens de pés de pacientes com DFU nos últimos cinco anos nos Hospitais Escolares de Lancashire, obtendo aprovação ética de todos os órgãos relevantes e consentimento informado por escrito do paciente. A aprovação foi obtida do Comitê de Ética em Pesquisa do NHS para usar essas imagens nesta pesquisa. Essas imagens DFU foram capturadas com diferentes câmeras (Kodak DX4530, Nikon D3300 e Nikon COOLPIX P100). O atual conjunto de dados que recebemos com a aprovação ética do NHS não continha quaisquer registros ou metadados sobre essas condições ou qualquer classificação médica.

Como não há metadados clínicos relativos a este conjunto de dados DFU, o experimento é realizado nas imagens com aprendizado de máquina tradicional artesanal e aprendizado profundo. Esta é a primeira vez que o reconhecimento de isquemia e infecção em DFU é realizado com base em imagens, portanto, não há nenhum conjunto de dados disponível publicamente. Aqui, apresentamos o primeiro conjunto de dados DFU com rótulos verdadeiros de casos de isquemia e infecção. A rotulagem especializada seguindo o popular sistema de classificação médica de cada condição DFU é particularmente importante para esta tarefa. A verdade básica foi produzida por dois profissionais de saúde (médicos consultores com especialização em pé diabético) sobre a inspeção visual de imagens DFU. Onde houve desacordo quanto à verdade básica, a decisão final foi tomada pelo médico mais experiente. Essas verdades básicas são usadas para a classificação binária de infecção e isquemia de DFU. Alguns exemplos de imagens de pés com DFU usadas para anotação binária especializada são mostrados na Fig. 1. O número completo de casos de anotação especializada de cada condição é detalhado na Tabela 1. O conjunto de dados, juntamente com seus rótulos de verdade, está disponível mediante solicitação em <http://www2.docm.mmu.ac.uk/STAFF/M.Yap/dataset.php>.

3. Metodologia

Esta seção descreve nossas técnicas propostas para o reconhecimento de isquemia e infecção do sistema de diagnóstico DFU. A preparação de um conjunto de dados balanceado, recursos artesanais e métodos de aprendizado de máquina (aprendizado de máquina artesanal e abordagens de aprendizado profundo) usados para classificação binária de isquemia e infecção são detalhados nesta seção.

3.1. Técnica natural de aumento de dados baseada em algoritmo de aprendizado profundo

Esta seção descreve nosso método de aumento de dados proposto, chamado Natural Data-augmentation, que é baseado no algoritmo de localização DFU profundo (Faster R-CNN).

No conjunto de dados DFU, as imagens (o tamanho varia entre 1600 × 1200 e 3648 × 2736 pixels) dependem das câmeras usadas para capturar os dados. No aprendizado profundo, o aumento de dados é visto como uma ferramenta importante para melhorar o desempenho dos algoritmos. Conforme mostrado na figura 2, aproximadamente 92% dos casos de DFU apresentam uma área entre 0% a 20% nas imagens dos pés. No aumento de dados, as técnicas comumente usadas são inversão, rotação, escala aleatória, corte aleatório, translação e ruído gaussiano. Como o DFU ocupa uma porcentagem muito pequena da área total das imagens dos pés, existe o risco de perder a região de interesse ao usar técnicas de aumento importantes, como escala aleatória, corte e tradução. Consequentemente, o aumento de dados naturais é mais adequado para a avaliação DFU do que o aumento de dados comum. Esta técnica de aumento ajuda a auxiliar os algoritmos da máquina a identificar a região de interesse (ROI) de imagens de pé DFU e focar na localização



Figura 1. Exemplos de imagens de pés com DFU usadas para anotações binárias especializadas para infecção e isquemia.

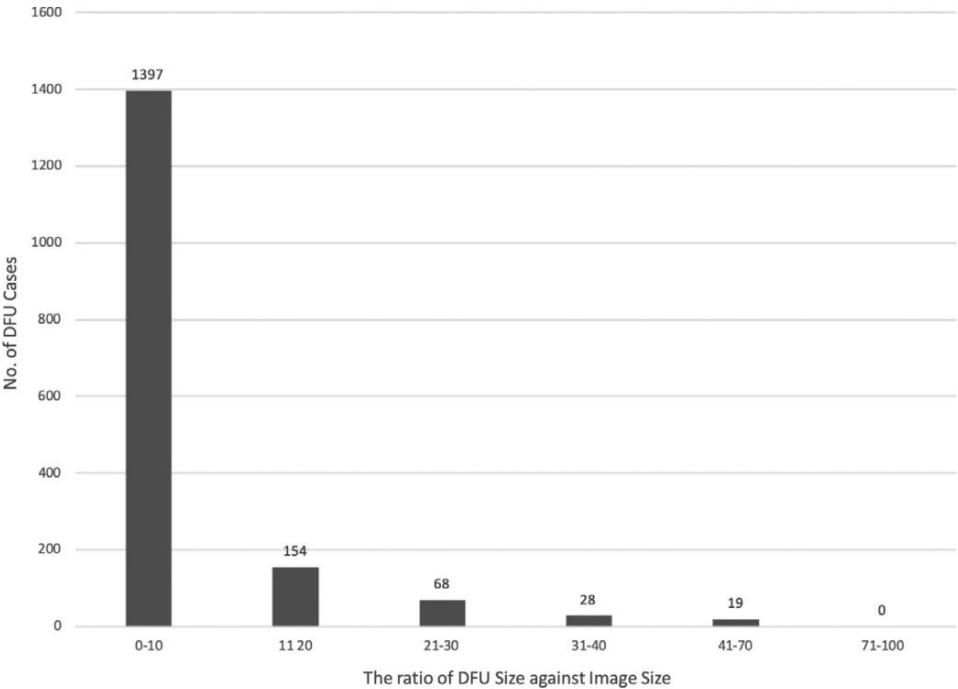


Figura 2. O número de casos de DFU de acordo com a área de DFU na imagem completa do pé do conjunto de dados DFU.

tabela 1

O número de casos de infecção e isquemia, número de manchas DFU e aumento patches usando aumento de dados naturais no conjunto de dados DFU.

Categoria	Definição	Casos	Patches DFU	Patches aumentados
Isquemia	Ausente	1249	1431	4935
	Presente	210	235	4935
Total de imagens		1459	1666	9870
Infecção bacteriana	Nenhum	628	684	2946
	Presente	831	982	2946
Total de imagens		1459	1666	5892

as fortes características que existem nesta área. Usamos o método de localização baseado em aprendizagem profunda, Faster-RCNN com InceptionResNetV2, para obter ROI das imagens de pé do DFU [24,25]. Dependendo do tamanho do DFU e da imagem, o aumento natural dos dados no DFU conjunto de dados com ampliação diferente é demonstrado na Fig. 3. Flexível parâmetros podem ser usados para escolher o número de fatores de ampliação (3 nesta classificação), bem como a distância de ampliação, que pode ser ajustado a partir de uma única imagem DFU por aumento natural. Depois ampliação, além disso, o aumento de dados é alcançado com a ajuda de ângulos, espelho, ruído gaussiano, contraste, nitidez, translação, cisalhamento usando nossos métodos propostos, conforme mostrado na Fig. 4.

Conforme mostrado na Tabela 1, o número de patches DFU gerados por cortar várias imagens DFU a pé e patches aumentados são gerado por aumento natural de dados (Fig. 3) e dados diferentes aumentos (Fig. 4). O número total de casos de isquemia e a não isquemia neste conjunto de dados DFU está desequilibrado (1249 casos vs 210 casos), enquanto infecção (628 casos) e não infecção (831 casos) são bastante equilibrado, conforme mostrado na Tabela 1. Realizamos classificação binária de isquemia e infecção com algoritmos de aprendizado de máquina porque para classificação multiclasse, este conjunto de dados DFU está desequilibrado, especialmente para casos (isquemia e sem infecção), conforme mostrado em 5.

3.2. Descritores de cores superpixel feitos à mão

Nós investigamos o uso de recursos de design humano com aprendizado de máquina tradicional na classificação binária de infecção e isquemia. Nossa primeira tentativa foi experimentar descritores de textura (Padrões Binários Locais e Histograma de Gradiente) e cores. descritores usados em trabalhos relacionados [19,21]. Contudo, conseguimos resultados muito ruins para esses problemas de classificação binária. Portanto, nós propor um novo Superpixel Color Descriptors (SPCD) para extrair o colore o ROI das imagens DFU que podem ser o elemento visual importante pistas para a identificação de isquemia e infecção em DFU. No primeiro passo, usamos uma técnica de superpixels SLIC para produzir superpixels supersegmentação de patches DFU com base na cor e intensidade do pixel valores [26]. A técnica de superpixels SLIC executa um k-means localizado otimização na cor 5-D CIELAB e espaço de imagem para agrupar pixels conforme descrito pelas Eqs. (1)–(4):

$$= \bar{y} \quad (1)$$

$$= + \quad (2)$$

$$= \bar{y}(\bar{y})^2 + (\bar{y})^2 + (\bar{y})^2 \quad (3)$$

$$= \bar{y}(-)^2 + (-)^2 \quad (4)$$

onde na Eq. (1), S é o tamanho aproximado de um superpixel, N é o número de pixels ek é o número de superpixels; na Eq. (2), D_s é a soma da distância lab (d_{lab}) e da distância do plano xy (d_{xy}); em

Eq. (3), I, a, b representam o espaço de cores do laboratório; e na Eq. (4), x e y representam as posições dos pixels.

Na segunda etapa, o valor médio da cor RGB de cada superpixel é calculado e aplicado a cada superpixel (S) denotado por:

$$= ((, ,), = 1, \dots, \quad (5)$$

onde na Eq. (5), $P(R, G, B)$ são os valores de pixel do canal R, G, B em cada i -ésima posição de S e k é o número total de superpixels na imagem.

Finalmente, com um número diferente de superpixels e valores limite de cada canal de cores, extraímos regiões de dois cores de interesse que são vermelho e preto dos patches DFU. Para essas tarefas de classificação, usamos o número de superpixels ($k = 200$) e valores limite (T1: 0,40, 0,45, 0,50, 0,055, 0,60; T2: 0,15, 0,20, 0,25, 0,30, 0,35) para extrair os recursos de cores de patches DFU de 256x256. Os valores limite são usados para restringir as intensidades de pixels vermelhos e pretos para serem utilizados como recursos artesanais. Portanto, nós utilizou um vetor de recursos de 10 com algoritmo SPCD junto com textura descritores (LBP, HOG) e recursos de cores (RGB, CIELAB) para treinar abordagens tradicionais de aprendizado de máquina. O pseudocódigo para o SPCD algoritmo é explicado no Algoritmo 1. O exemplo de extração de cores recursos usando nosso novo algoritmo SPCD são mostrados na Fig. 4.

Algoritmo 1 Pseudocódigo para os descritores de cores Superpixel Extração

```

1: A segmentação excessiva do patch DFU com superpixel SLIC é realizada;
2: O valor RGB médio de cada superpixel é calculado e aplicado;
3: Inicialize a variável S_Red & S_Black para 0
4: procedimento RedAndBlackRegion
5:   para  $\bar{y}(i)$  faça
6:     se  $eu(i) > 1$   $\bar{y}(i(i) + i(i) + i(i))$  então retorne S_Red=
       S_Vermelho + 1
7:     se  $eu(i) < 2$  &  $eu(i) < 2$  &  $eu(i) < 2$  então retorne
       S_Preto= S_Preto + 1
8:   = - ÷
9:   = - ÷

```

Para esses problemas de classificação, experimentamos vários de classificadores com hiperparâmetros padrão nesses recursos de cores. BayesNet, Random Forest e Multilayer Perceptron foram selecionados e alcançou a mais alta precisão entre outros aprendizados de máquina classificadores.

3.3. Abordagens de aprendizagem profunda

Para comparação com os recursos tradicionais, algoritmos de aprendizagem profunda são usados para realizar classificação binária para classificar (1) infecção e não infecção; e (2) classes de isquemia e não isquemia em Patches DFU. Para este trabalho, ajustamos (transferimos o aprendizado de modelos pré-treinados) os modelos CNN, ou seja, Inception-V3, ResNet50 e InceptionResNetV2 [27–29]. Para treinar as redes CNN, congelamos o pesos das primeiras camadas das redes pré-treinadas para comum recursos, como arestas e curvas. As camadas subsequentes de redes são descongelado para se concentrar no aprendizado de recursos específicos do conjunto de dados.

Além disso, utilizamos o método Ensemble CNN, que é um abordagem CNN muito eficaz para obter uma precisão muito boa em situações difíceis conjuntos de dados. O modelo Ensemble CNN combina os recursos de gargalo de vários modelos CNN (Inception-V3, ResNet50 e InceptionRes-NetV2) e usar o classificador SVM para produzir previsões, conforme mostrado em Figura 7.

4. Resultados e discussão

Os conjuntos de dados de infecção e isquemia foram divididos em 70% de treinamento, 10% de conjuntos de validação e 20% de testes e adotamos a técnica de validação cruzada de 5 vezes. Utilizamos a técnica natural de aumento de dados para conjuntos de treinamento e validação em abordagens tradicionais de aprendizado de máquina e aprendizado profundo. Portanto, neste conjunto de dados de isqu

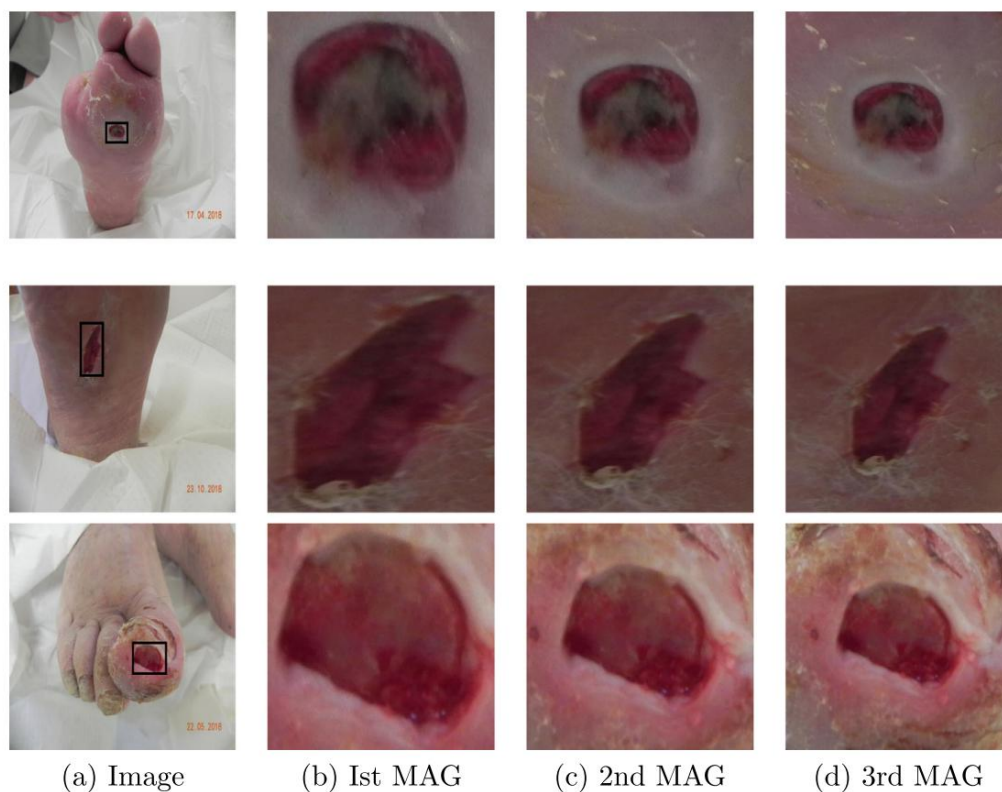


Figura 3. Aumento natural de dados produzido a partir da imagem original com diferentes ampliações (três ampliações neste experimento). MAG refere-se à ampliação.

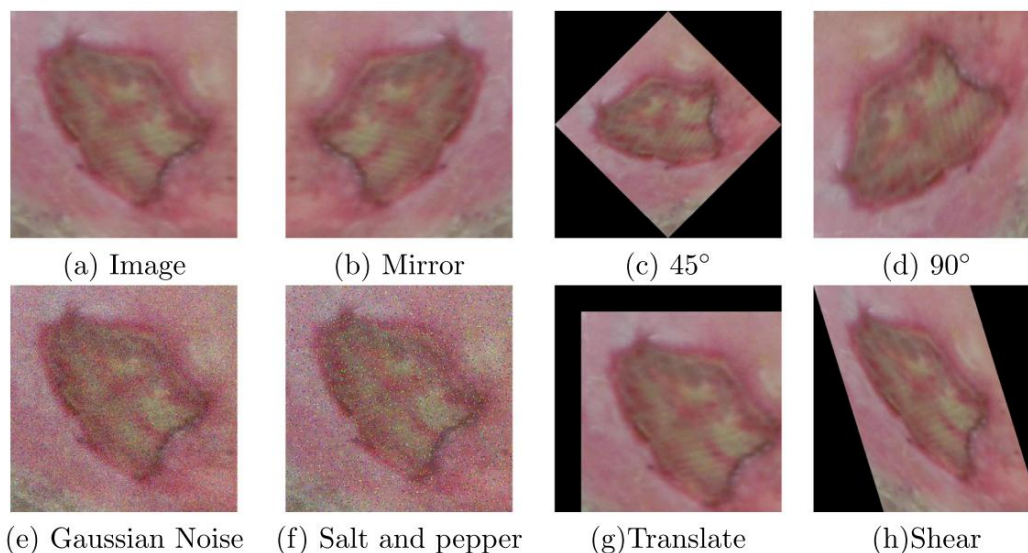


Figura 4. Após a ampliação, diferentes tipos de aumento de dados são alcançados pelo aumento de dados naturais proposto.

usamos aproximadamente 11.564 patches, 1.652 patches e 3.304 patches em conjuntos de treinamento, validação e teste, respectivamente, enquanto, no conjunto de dados de infecção, usamos 7.136 patches (treinamento), 1.019 patches (validação) e 2.038 patches (teste) de as 2.611 imagens originais de pés. Conforme mencionado anteriormente, usamos modelos tradicionais de aprendizado de máquina artesanais (doravante TML) e modelos CNN para realizar a tarefa de classificação e utilizamos imagens RGB 256×256 como entrada para TML e InceptionV3, AlexNet e ResNet50. Para InceptionResNetV2, redimensionamos o conjunto de dados para 299×299 . Para este experimento, o TensorFlow é usado para aprendizado profundo e o Matlab é usado para abordagens tradicionais de aprendizado de máquina.

Nas Tabelas 2 e 3, relatamos *Precisão*, *Sensibilidade*, *Precisão*, *Especificidade*, *Medida F*, *Coefficiente de Correlação de Matthew (MCC)* e *Área sob a curva ROC (AUC)* como nossas métricas de avaliação.

Ao comparar o desempenho dos métodos computadorizados e das técnicas propostas, as CNNs tiveram melhor desempenho na classificação binária de isquemia do que de infecção, apesar de dados mais desequilibrados no conjunto de dados de isquemia, devido a mais casos de não isquemia no conjunto de dados. O desempenho médio de todos os modelos em termos de precisão no conjunto de dados de isquemia foi de 83,3%, o que é notavelmente melhor do que a precisão média de 65,8% no conjunto de dados de infecção. Da mesma forma, a pontuação MCC e a pontuação AUC são consideradas medidas de desempenho aceitáveis para comparar os resultados da classificação. Obtemos

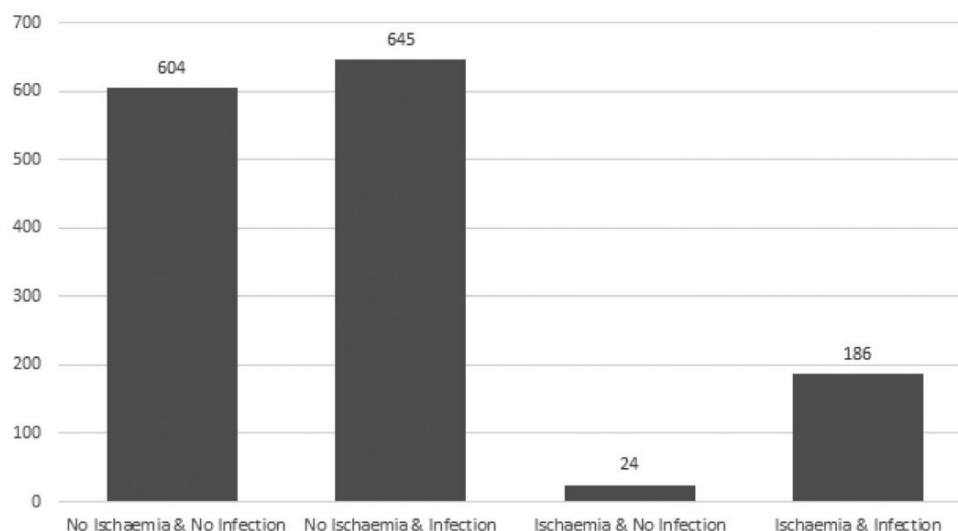
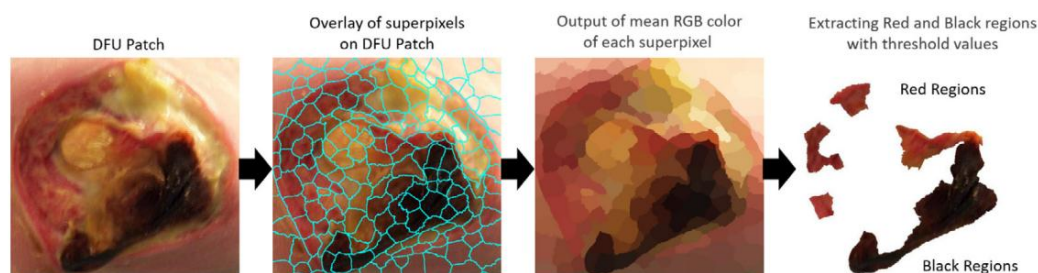
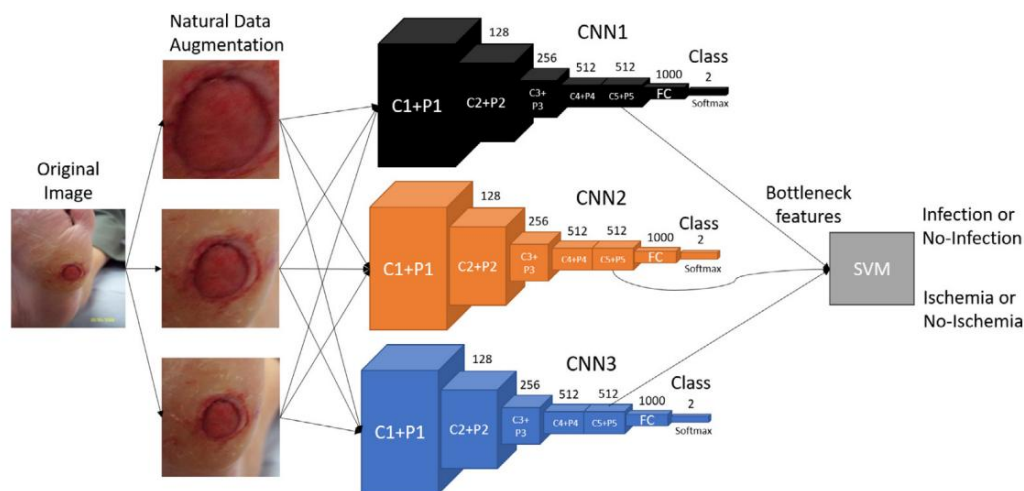


Figura 5. Distribuição dos casos de isquemia e infecção como problema de classificação multiclasse.



6. Exemplo de extração de regiões vermelhas e pretas do patch DFU com o algoritmo Superpixel Color Descriptor proposto, que foi então usado para informar a identificação de isquemia e infecção. O valor k de 200 para o algoritmo superpixel supersegmentou efetivamente os patches DFU.



7. Extração de recursos de gargalo de CNNs e alimentação no classificador SVM para realizar classificação binária de isquemia e infecção, onde C1 – C5 são camadas convolucionais, P1 – P5 são camadas de pooling e FC é camada totalmente conectada. Nota: As CNNs nesta figura são apenas representações da arquitetura geral das CNNs e não representam as arquiteturas CNN originais de Inception-V3, ResNet50 e InceptionResNetV2.

Pontuação e pontuação AUC para classificação de isquemia de 67,1% e 83,2% respectivamente, em comparação com a classificação de infecção de 32,3% e 65,8%, respectivamente. As curvas ROC para todos os algoritmos, incluindo TML e CNNs para classificação binária de isquemia e infecção, são mostradas nas Figs. 8 e 9. Ao comparar o desempenho na classificação de isquemia de TML e CNNs, as CNNs (86,5%) tiveram melhor desempenho que os modelos TML (79%). Da mesma forma, na classificação de infecções, a precisão das CNNs (68,4%) teve melhor desempenho do que a TML (62,1%) com

uma margem de 6,3%. Notavelmente, o método Ensemble CNN alcançou a pontuação mais alta em todas as medidas de desempenho na classificação de isquemia e infecção.

Sensibilidade e Especificidade são consideradas importantes medidas de desempenho em imagens médicas. O método ensemble rendeu alta *sensibilidade* para o conjunto de dados de isquemia com uma margem de 6,9% do segundo algoritmo de melhor desempenho, perceptron multicamadas. Curiosamente, um perceptron multicamadas teve o pior desempenho na *Especificidade* com uma pontuação

Tabela 2

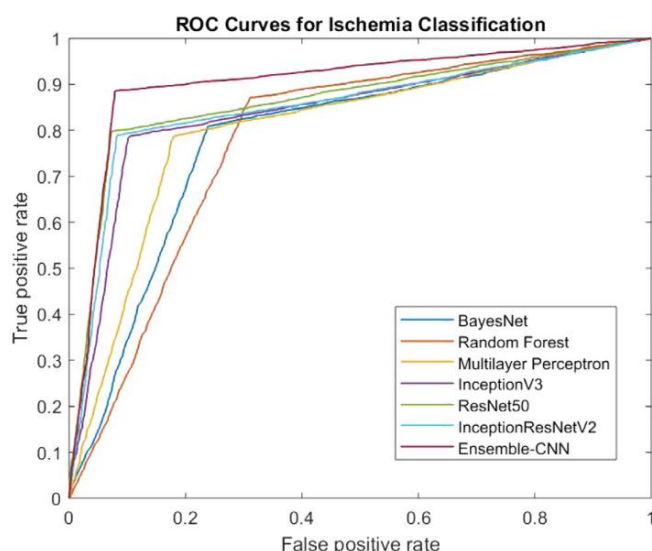
As medidas de desempenho da classificação binária da isquemia por meio de nossas propostas de aprendizado de máquina tradicional artesanal e abordagens CNN.

	Precisão	Sensibilidade	Precisão	Especificidade	Medida F	Pontuação MCC	Pontuação AUC
BayesNet	0,785 ± 0,022	0,774 ± 0,034	0,809 ± 0,034	0,800 ± 0,027	0,790 ± 0,020	0,572 ± 0,044	0,783
Floresta aleatória	0,780 ± 0,041	0,739 ± 0,049	0,872 ± 0,029	0,842 ± 0,034	0,799 ± 0,033	0,571 ± 0,078	0,780
Perceptron multicamadas	0,804 ± 0,022	0,817 ± 0,040	0,787 ± 0,046	0,795 ± 0,031	0,800 ± 0,023	0,610 ± 0,045	0,804
InícioV3 (CNN)	0,841 ± 0,017	0,784 ± 0,045	0,886 ± 0,018	0,898 ± 0,022	0,831 ± 0,021	0,688 ± 0,031	0,840
ResNet50 (CNN)	0,862 ± 0,018	0,797 ± 0,043	0,917 ± 0,015	0,927 ± 0,017	0,852 ± 0,022	0,732 ± 0,032	0,865
InícioResNetV2 (CNN)	0,853 ± 0,021	0,789 ± 0,054	0,906 ± 0,017	0,917 ± 0,019	0,842 ± 0,027	0,714 ± 0,039	0,851
Conjunto (CNN)	0,903 ± 0,012	0,886 ± 0,035	0,918 ± 0,019	0,921 ± 0,021	0,902 ± 0,014	0,807 ± 0,022	0,904

Tabela 3

As medidas de desempenho da classificação binária de infecção por nossas propostas de aprendizado de máquina tradicional artesanal e abordagens CNN.

	Precisão	Sensibilidade	Precisão	Especificidade	Medida F	Pontuação MCC	Pontuação AUC
BayesNet	0,639 ± 0,036	0,619 ± 0,018	0,653 ± 0,039	0,660 ± 0,015	0,622 ± 0,079	0,290 ± 0,070	0,643
Floresta aleatória	0,605 ± 0,025	0,608 ± 0,025	0,607 ± 0,037	0,601 ± 0,069	0,606 ± 0,012	0,211 ± 0,051	0,601
Perceptron multicamadas	0,621 ± 0,026	0,680 ± 0,023	0,622 ± 0,057	0,570 ± 0,023	0,627 ± 0,074	0,281 ± 0,055	0,619
InícioV3 (CNN)	0,662 ± 0,014	0,693 ± 0,038	0,653 ± 0,015	0,631 ± 0,034	0,672 ± 0,019	0,325 ± 0,029	0,662
ResNet50 (CNN)	0,673 ± 0,013	0,692 ± 0,051	0,668 ± 0,023	0,654 ± 0,051	0,679 ± 0,019	0,348 ± 0,028	0,673
InícioResNetV2 (CNN)	0,676 ± 0,015	0,688 ± 0,052	0,672 ± 0,015	0,664 ± 0,039	0,680 ± 0,024	0,352 ± 0,031	0,678
Conjunto (CNN)	0,727 ± 0,025	0,709 ± 0,044	0,735 ± 0,036	0,744 ± 0,050	0,722 ± 0,028	0,454 ± 0,052	0,731



8. Curva ROC para todos os métodos TML e CNN para classificação de isquemia.

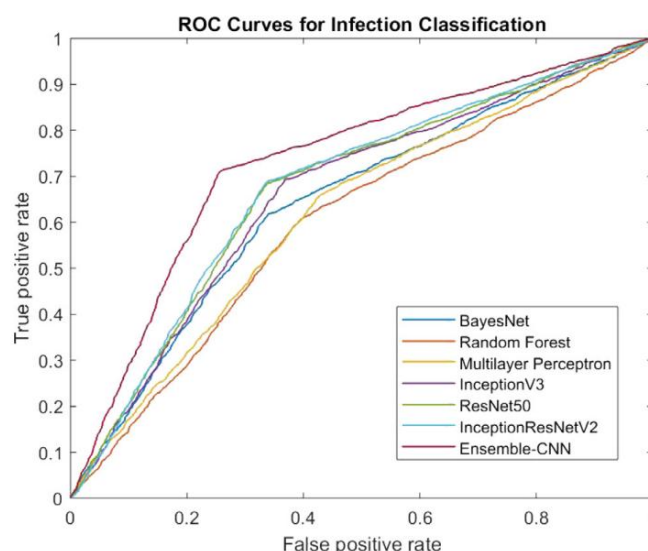


Figura 9. Curva ROC para todos os métodos TML e CNN para classificação de infecção.

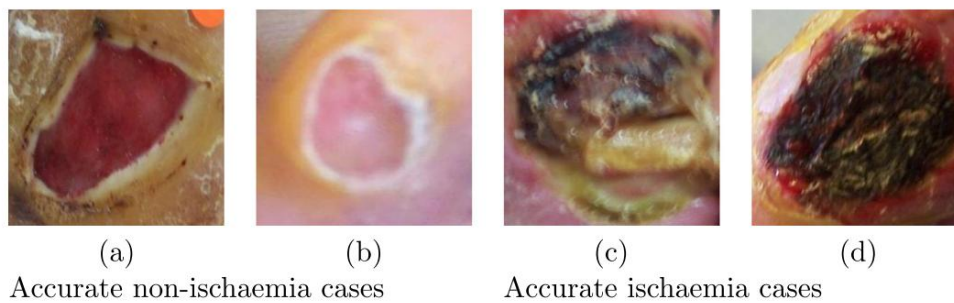
de 79,5%. Para *especificidade* no conjunto de dados de isquemia, o método ensemble obteve novamente a pontuação mais alta de 92,9%, o que é ligeiramente melhor do que ResNet50 (92,7%).

Na classificação da infecção, os métodos TML e CNN receberam pontuações moderadas nas medidas de desempenho. Novamente, os métodos da CNN teve melhor desempenho do que os métodos TML, alcançando a pontuação mais alta em todas as medidas de desempenho. O método Ensemble CNN teve melhor desempenho do que outros classificadores CNN, especialmente para *Especificidade*, com uma pontuação de 74,4% na classificação de infecção com uma margem notável de 8% do que o segundo algoritmo com melhor desempenho, InceptionResNetV2 (66,4%). Para *Sensibilidade*, todas as CNNs tiveram um desempenho marginalmente bom com o Ensemble alcançando a pontuação mais alta de 70,9%. Ao comparar o desempenho dos métodos TML, o Multilayer Perceptron (68,0%) apresentou bem em *Sensibilidade*, enquanto BayesNet (66%) melhor em *Especificidade*.

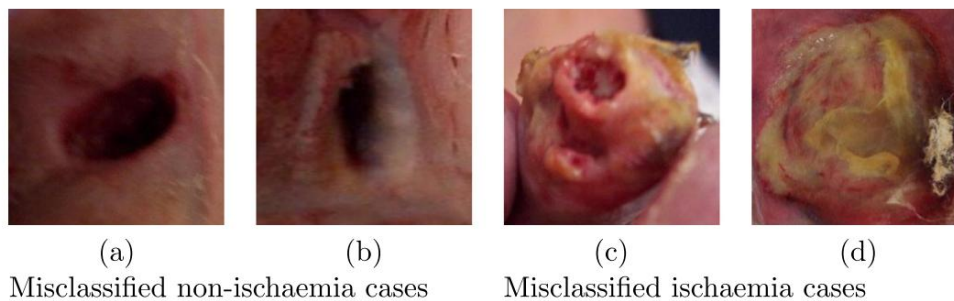
4.1. Análise experimental e discussão

A avaliação da DFU com métodos informatizados é muito importante para apoiar os sistemas globais de saúde através da melhoria da triagem e monitorar procedimentos e reduzir o tempo hospitalar para pacientes e médicos. Este experimento preliminar é focado automaticamente identificando as condições importantes de isquemia e infecção de DFU.

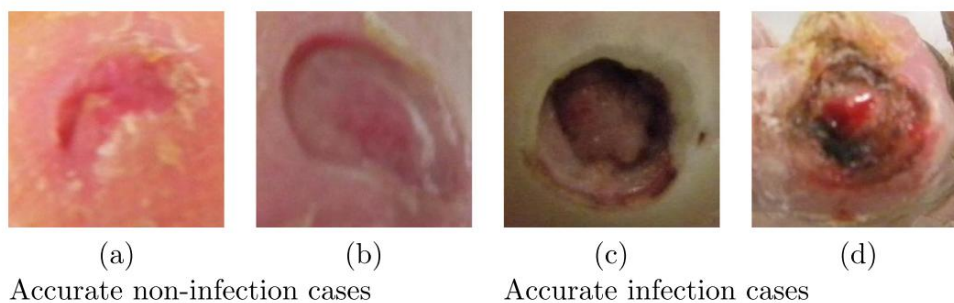
O objetivo principal foi identificar isquemia e infecção a partir de imagens dos pés usando aprendizado de máquina. Ilustramos exemplos de casos classificados correta e incorretamente em ambas as classificações binárias de isquemia (Figs. 10 e 11) e infecção (Figs. 12 e 13). Como para os casos mal classificados, existem enormes diferenças intraclass e semelhanças interclasses entre (1) infecção e não infecção; (2) casos de isquemia e não isquemia na DFU que tornam os classificadores difíceis de prever a classe correta. Além disso, existem outros fatores que influenciam na classificação dessas condições, como condições de iluminação, marcas e tom de pele. Em casos mal classificados de não isquemia, como mostrado na Fig. 11, os casos (a) e (b) são dificultados por uma condição de iluminação (sombra), respectivamente, enquanto em (c) e (d) casos de isquemia classificados incorretamente, as características da isquemia podem ser muito sutis ser reconhecido a partir das imagens pelo algoritmo. Alternativamente, é provavelmente precisávamos de uma medida objetiva mais sensível da verdade básica a partir de avaliações vasculares. Descobrimos que as sombras são particularmente problemático porque algoritmos de aprendizado de máquina podem ser enganados por sombras, especialmente na determinação de condições importantes, como isquemia. Na Figura 13, casos mal classificados de não infecção, a presença de sangue no caso (a), enquanto o caso (b) pertence a um dos raros casos com presença de isquemia e não infecção. Em classificação incorreta casos de infecção, os indicadores visuais de infecção eram provavelmente muito sutis,



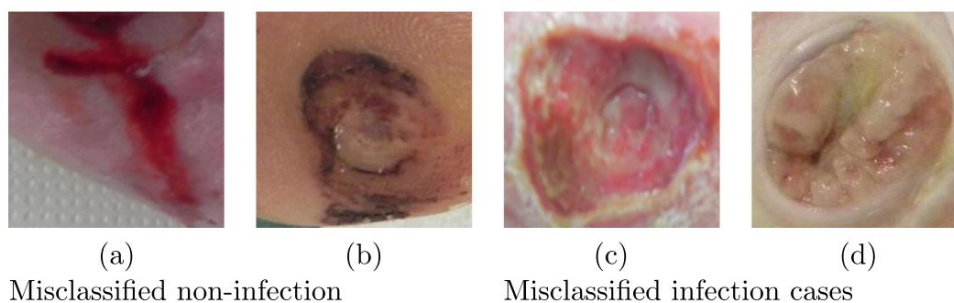
10. Exemplos de casos classificados corretamente pelo Ensemble-CNN no conjunto de dados de isquemia. (a) e (b) representam casos sem isquemia. (c) e (d) representam casos de isquemia.



11. Exemplos de casos classificados incorretamente pelo Ensemble-CNN no conjunto de dados de isquemia. (a) e (b) representam casos não isquêmicos. (c) e (d) representam casos de isquemia.



12. Exemplos de casos classificados corretamente pelo Ensemble-CNN no conjunto de dados de infecção. (a) e (b) representam casos de não infecção. (c) e (d) representam casos de infecção.



13. Exemplos de casos classificados incorretamente pelo Ensemble-CNN no conjunto de dados de infecção. (a) e (b) representam casos de não infecção. (c) e (d) representam casos de infecção.

ou precisávamos de informações objetivas mais sensíveis fornecidas por meio de análises de sangue.

Neste trabalho, usamos o aumento natural de dados proposto com a ajuda da localização DFU para criar patches DFU a partir de imagens de pés em tamanho real. Esses adesivos são úteis para focar mais na localização de indicadores visuais de fatores importantes de DFU, como infecção e isquemia. Em seguida, investigamos o uso de TML e CNNs para determinar essas condições como classificação binária. Neste experimento, obtivemos um desempenho muito bom em termos de classificação correta da isquemia, apesar dos casos desequilibrados no conjunto de dados DFU. Porém, no caso da infecção, os classificadores não tiveram o mesmo desempenho, uma vez que o

a condição de infecção é difícil de reconhecer a partir das imagens dos pés, mesmo por especialistas médicos experientes especializados em DFU e, portanto, provavelmente requer a determinação da verdade por meio de exames de sangue objetivos para identificar a infecção bacteriana.

A pesquisa atual concentra-se no reconhecimento de isquemia e infecção em sistemas de classificação médica, exigindo a orientação de médicos especialistas especializados em DFU. Para desenvolver uma ferramenta assistida por computador para médicos especialistas em análise remota do pé, ou seja, um sistema de diagnóstico remoto DFU, é necessário enfrentar os seguintes desafios:

1. Reconhecimento da isquemia e infecção com algoritmos de aprendizado de máquina como um importante estudo de prova de conceito para pés

classificação de patologias. Análises adicionais de cada patologia nas imagens dos pés são necessárias de acordo com os sistemas de classificação médica, como a Classificação DFU da Universidade do Texas [8] e o Sistema de Classificação SINBAD [9]. Isto requer uma estreita colaboração com médicos especialistas especializados em DFU.

- Algoritmos de aprendizagem profunda precisam de conjuntos de dados substanciais para obter uma precisão muito boa, especialmente para imagens médicas. Este experimento incluiu um conjunto de dados DFU desequilibrado (imagens de 1.459 pés) para condições de isquemia e infecção. No futuro, se estes algoritmos fossem treinados com um número maior de um conjunto de dados mais equilibrado, poderia melhorar o reconhecimento de isquemia e infecção.
- Um estudo do desempenho de algoritmos em diferentes tipos de dispositivos de captura é um aspecto importante de trabalhos futuros. Este experimento avalia o desempenho de algoritmos de aprendizado de máquina no conjunto de dados DFU coletados com diferentes câmeras (fontes heterogêneas de dados). Isto leva a uma maior variabilidade das características da imagem, uma vez que os algoritmos têm que lidar com padrões e características mais heterogêneas que não são intrínsecos à patologia em si. Nesta experiência, embora saibamos que foram utilizados três tipos de dispositivos, não temos a informação sobre a associação de imagens e o tipo de dispositivos.
- A verdade actual baseia-se apenas na inspecção visual efectuada por especialistas e não é apoiada por notas médicas ou testes clínicos (avaliação vascular para isquemia e análises ao sangue para identificar a presença de qualquer infecção bacteriana). Além disso, as imagens DFU foram desbridadas antes de serem capturadas. Consequentemente, o desbridamento da DFU remove importantes indicadores visuais de infecção, como o exsudato colorido. Portanto, a sensibilidade e a especificidade destes algoritmos poderiam ser melhoradas no futuro, alimentando-se de testes clínicos, como avaliações vasculares (isquemia) e exames de sangue (para identificar a presença de qualquer infecção bacteriana).
- A prática clínica actual obtém a foto do pé utilizando diferentes modelos de câmeras, poses e iluminação. É um grande desafio para um algoritmo computacional prever a profundidade e o tamanho da ferida com base em imagens não padronizadas. Conjunto de dados padronizado, como o método de coleta de dados proposto por Yap et al. [30] aumentará a precisão do sistema de diagnóstico DFU.
- A anotação do conjunto de dados é um processo trabalhoso, especialmente para médicos especialistas rotularem as patologias dos pés em 16 classes, de acordo com o sistema de classificação da Universidade do Texas. Para reduzir a carga sobre os especialistas médicos no delineamento e anotação do conjunto de dados, há uma necessidade urgente de se concentrar no desenvolvimento de técnicas de aprendizagem automática não supervisionadas ou auto-supervisionadas.
- A recolha do conjunto de dados cronológicos é crucial para a detecção precoce das principais patologias. Isto permitirá o monitoramento longitudinal da saúde e das alterações dos pés, onde médicos especialistas e algoritmos de computador poderão aprender os primeiros sinais de DFU. A longo prazo, o sistema de diagnóstico de DFU pode ser capaz de prever o processo de cicatrização de úlceras e prevenir a DFU antes que ela ocorra.
- Um aplicativo para smartphone poderia ser desenvolvido para triagem e monitoramento remoto de DFU. Para ampliar o sistema de diagnóstico DFU, o aplicativo precisaria ser executado em vários dispositivos, independentemente da plataforma e/ou sistema operacional.

5. Conclusão

Neste trabalho, treinamos vários classificadores baseados em algoritmos tradicionais de aprendizado de máquina e CNNs para discriminar as condições de: (1) isquemia e não isquemia; e (2) infecção e não infecção relacionada a uma determinada DFU. Encontramos medidas de alto desempenho na classificação binária de isquemia, em comparação com desempenho moderado dos classificadores na classificação de infecção. É vital compreender as características de ambas as condições em relação à DFU (isquemia e infecção) a partir de uma perspectiva de visão computacional. Determinando estes

condições, especialmente infecção de imagens de pés não padronizadas, são muito desafiadoras devido a: (1) altas diferenças visuais intraclasse e semelhanças interclasses entre classes; (2) os indicadores visuais de infecção e isquemia são potencialmente muito sutis na DFU; (3) são necessários exames médicos objetivos para o suprimento vascular e infecção bacteriana para fornecer informações mais objetivas e melhorar ainda mais a classificação dessas condições; e (4) outros fatores como condições de iluminação, marcas e tom de pele são importantes para incorporar na previsão.

Com um conjunto de dados mais equilibrado e uma melhor captura de dados do DFU, o desempenho desses métodos poderá ser melhorado no futuro. A otimização adicional em hiperparâmetros de métodos de aprendizado profundo e de aprendizado de máquina tradicional também poderia melhorar o desempenho dos algoritmos neste conjunto de dados. Verdades básicas aprimoradas por testes clínicos para isquemia e infecção podem fornecer informações adicionais e melhorias adicionais dos algoritmos, mesmo quando não há nenhum indicador visual aparente a olho nu. No caso de infecção, mesmo após o desbridamento, a verdade informada por exames de sangue para detecção de infecção pode produzir melhorias na sensibilidade e especificidade, mesmo na ausência de indicadores visuais abertamente óbvios. Este trabalho tem o potencial da tecnologia para transformar o reconhecimento e o tratamento das úlceras do pé diabético e levar a uma mudança de paradigma no cuidado clínico do pé diabético.

Declaração de interesse concorrente

Os autores declaram que não têm interesses financeiros concorrentes ou relações pessoais conhecidas que possam ter influenciado o trabalho relatado neste artigo.

Reconhecimentos

Os autores expressam a sua gratidão aos Hospitais Escolares de Lancashire e aos especialistas clínicos pelo seu amplo apoio e contribuição na realização desta investigação. Gostariamos de agradecer ao Kim's English Corner (<https://kimsenglishcorner.com>) para revisão.

Referências

- JD Santilli, SM Santilli, Isquemia crítica crônica de membros: diagnóstico, tratamento e prognóstico., Amer. Física de Família. 59 (7) (1999) 1899–1908.
- M. Albers, AC Fratezi, N. De Luccia, Avaliação da qualidade de vida de pacientes com isquemia grave como resultado de doença oclusiva arterial infrainguinal, J. Vasco. Surg. 16 (1) (1992) 54–59.
- L. Prompers, M. Huijberts, J. Apelqvist, E. Jude, A. Piaggies, K. Bakker, et al., Alta prevalência de isquemia, infecção e comorbidade grave em pacientes com pé diabético na Europa. resultados iniciais do estudo eurodiale, Diabetologia 50 (1) (2007) 18–25.
- BA Lipsky, AR Berendt, PB Cornia, JC Pile, EJ Peters, DG Armstrong, et al., Diretriz de prática clínica da Sociedade de Doenças Infecciosas da América de 2012 para o diagnóstico e tratamento de infecções do pé diabético, Clin. Infectar. Dis. 54 (12) (2012) e132–e173.
- LA Lavery, DG Armstrong, RP Wunderlich, J. Tredwell, AJ Boulton, Síndrome do pé diabético: avaliando a prevalência e incidência de patologia do pé em mexicanos-americanos e brancos não-hispanicos de uma coorte de gerenciamento de doenças de diabetes, Diabetes Care 26 (5) (2003) 1435–1438.
- GH Skrepnek, JL Mills, LA Lavery, DG Armstrong, Serviço de saúde e resultados entre cerca de 6,7 milhões de casos de pé diabético em atendimento ambulatorial nos EUA, Diabetes Care 40 (7) (2017) 936–942.
- FW Wagner, O pé diabético, Ortopedia 10 (1) (1987) 163–172.
- LA Lavery, DG Armstrong, LB Harkless, Classificação de feridas no pé diabético, J. Cirurgia do Tornozelo do Pé. 35 (6) (1996) 528–531.
- P. Ince, ZG Abbas, JK Lutale, A. Basit, SM Ali, F. Chohan, S. Morbach, et al., Uso do sistema de classificação sinbad e pontuação na comparação dos resultados do tratamento de úlceras nos pés em três continentes , Diabetes Care 31 (5) (2008) 964–967.
- JJ van Netten, D. Clark, PA Lazzarini, M. Janda, LF Reed, A validade e confiabilidade da avaliação remota de úlcera de pé diabético usando imagens de telefone celular, Sci. Rep. 7 (1) (2017) 9480.
- V. Gulshan, L. Peng, M. Coram, MC Stumpe, D. Wu, A. Narayanaswamy, et al., Desenvolvimento e validação de um algoritmo de aprendizagem profunda para detecção de retinopatia diabética em fotografias de fundo de retina, JAMA 316 (22) (2016) 2402–2412.

- [12] A. Esteve, B. Kuprel, RA Novoa, J. Ko, SM Swetter, HM Blau, et al., Classificação de câncer de pele em nível de dermatologista com redes neurais profundas, *Nature* 542 (7639) (2017) 115–118.
- [13] A. Krizhevsky, I. Sutskever, GE Hinton, classificação Imagenet com redes neurais convolucionais profundas, em: *Advances in Neural Information Processing Systems*, 2012, pp.
- [14] F. Veredas, H. Mesa, L. Morente, Classificação binária de tecidos em imagens de feridas com redes neurais e classificadores bayesianos, *IEEE Trans. Med. Imagens* 29 (2) (2009) 410–427.
- [15] H. Wannous, Y. Lucas, S. Treuillet, Avaliação aprimorada do processo de cicatrização de feridas por classificação precisa de tecidos multivisualização, *IEEE Trans. Med. Imagens* 30 (2) (2010) 315–326.
- [16] L. Wang, P. Pedersen, E. Agu, D. Strong, B. Tulu, Determinação de área de imagens de úlcera de pé diabético usando uma classificação baseada em SVM de dois estágios em cascata, *IEEE Trans. Biomédica. Eng.* (2016).
- [17] MH Yap, M. Goyal, FM Osman, R. Martí, E. Denton, A. Juette, et al., Reconhecimento de lesões por ultrassom mamário: abordagens de aprendizagem profunda ponta a ponta, *J. Med. Imagem* 6 (1) (2018) 011007.
- [18] E. Ahmad, M. Goyal, JS McPhee, H. Degens, MH Yap, Segmentação semântica do músculo quadríceps da coxa humana em imagens de ressonância magnética, 2018, pré-impressão arXiv [arXiv:1801.00415](https://arxiv.org/abs/1801.00415).
- [19] M. Goyal, ND Reeves, AK Davison, S. Rajbhandari, J. Spragg, MH Yap, Dfunet: redes neurais convolucionais para classificação de úlceras de pé diabético, *IEEE Trans. Emergir. Principal. Computação. Intel.* (2018) 1–12, <http://dx.doi.org/10.1109/TETCI.2018.2866254>.
- [20] M. Goyal, MH Yap, ND Reeves, S. Rajbhandari, J. Spragg, Redes totalmente convolucionais para segmentação de úlcera de pé diabético, em: *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp 618–623, <http://dx.doi.org/10.1109/SMC.2017.8122675>.
- [21] M. Goyal, ND Reeves, S. Rajbhandari, MH Yap, Métodos robustos para detecção e localização de úlceras no pé diabético em tempo real em dispositivos móveis, *IEEE J. Biomed. Informações de saúde (ISSN: 2168-2194)* 23 (4) (2019) 1730–1741, <http://dx.doi.org/10.1109/JBHI.2018.2868656>.
- [22] C. Wang, X. Yan, M. Smith, K. Kochhar, M. Rubin, SM Warren, et al., Uma estrutura unificada para segmentação e análise automática de feridas com redes neurais convolucionais profundas, em: *Engenharia em Medicina e Biology Society (EMBC), 2015 37ª Conferência Internacional Anual do IEEE, IEEE, 2015*, pp.
- [23] JL Mills Sr, MS Conte, DG Armstrong, FB Pomposelli, A. Schanzer, AN Sidawy, et al., Sistema de classificação de membros ameaçados de membros inferiores da sociedade de cirurgia vascular : estratificação de risco baseada em ferida, isquemia e infecção no pé (wifi), *J. Vasc. Surg.* 59 (1) (2014) 220–234.
- [24] M. Goyal, MH Yap, Detecção de região de interesse em imagens dermatoscópicas para aumento de dados naturais, 2018, pré-impressão arXiv [arXiv:1807.10711](https://arxiv.org/abs/1807.10711).
- [25] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, et al., Trade-offs de velocidade/precisão para detectores de objetos convolucionais modernos, 2016, pré-impressão arXiv [arXiv:1611.10012](https://arxiv.org/abs/1611.10012).
- [26] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, *Slic Superpixels*, Tech. rep., 2010.
- [27] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Repensando a arquitetura inicial para visão computacional, em: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. –2826.
- [28] C. Szegedy, S. Ioffe, V. Vanhoucke, Inception-v4, inception-resnet e o impacto das ligações residuais na aprendizagem, 2016, CoRR [abs/1602.07261](https://arxiv.org/abs/1602.07261). [arXiv:1602.07261](https://arxiv.org/abs/1602.07261). URL <http://arxiv.org/abs/1602.07261>.
- [29] K. He, X. Zhang, S. Ren, J. Sun, Aprendizagem residual profunda para reconhecimento de imagem, em: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp.
- [30] MH Yap, KE Chatwin, C.-C. Ng, CA Abbott, FL Bowling, S. Rajbhandari, et al., Um novo aplicativo móvel para padronizar imagens de pés diabéticos, *J. Diabetes Sci. Tecnologia*. 12 (1) (2018) 169–173.