

Listas de conteúdos disponíveis em [ScienceDirect](https://www.sciencedirect.com)

## Computadores em Biologia e Medicina

página inicial da revista: <http://www.elsevier.com/locate/combiomed>

## Métodos de classificação de imagens baseados em inteligência artificial para diagnóstico de câncer de pele: Desafios e oportunidades

Manu Goyal<sup>a,\*</sup>, Thomas Knackstedt<sup>b</sup>, Shaofeng Yan<sup>c</sup>, Saeed Hassanpour<sup>d</sup><sup>a</sup> Departamento de Ciência de Dados Biomédicos, Dartmouth College, Hanover, NH, EUA<sup>b</sup> Departamento de Dermatologia, Metrohealth System e Faculdade de Medicina, Case Western Reserve University, Cleveland, OH, EUA<sup>c</sup> Seção de Dermatopatologia, Departamento de Patologia e Medicina Laboratorial, Dartmouth-Hitchcock Medical Center, Geisel School of Medicine em Dartmouth, Libano, NH, EUA<sup>d</sup> Departamentos de Ciência de Dados Biomédicos, Ciência da Computação e Epidemiologia, Dartmouth College, Hanover, NH, EUA

## INFORMAÇÕES DO ARTIGO

## Palavras-chave:

Câncer de pele

Inteligência artificial

Aprendizado profundo

Dermatologistas

Diagnóstico auxiliado por computador

Dermatologia digital

## ABSTRATO

Recentemente, tem havido grande interesse no desenvolvimento de soluções de diagnóstico auxiliadas por computador com Inteligência Artificial (IA) para o diagnóstico de câncer de pele. Com a crescente incidência de cânceres de pele, a baixa sensibilização de uma população crescente e a falta de conhecimentos e serviços clínicos adequados, existe uma necessidade imediata de sistemas de IA para ajudar os médicos neste domínio. Um grande número de conjuntos de dados de lesões cutâneas está disponível publicamente, e os pesquisadores desenvolveram soluções de IA, particularmente algoritmos de aprendizagem profunda, para distinguir lesões cutâneas malignas de lesões benignas em diferentes modalidades de imagem, como imagens dermatoscópicas, clínicas e histopatológicas. Apesar das várias alegações de que os sistemas de IA alcançam maior precisão do que os dermatologistas na classificação de diferentes lesões cutâneas, estes sistemas de IA ainda estão nas fases iniciais de aplicação clínica em termos de estarem prontos para ajudar os médicos no diagnóstico de cânceres de pele. Nesta revisão, discutimos os avanços nas soluções de IA baseadas em imagens digitais para o diagnóstico de câncer de pele, juntamente com alguns desafios e oportunidades futuras para melhorar esses sistemas de IA para apoiar dermatologistas e aumentar sua capacidade de diagnosticar problemas de pele.

Câncer.

## 1. Introdução

De acordo com a Skin Cancer Foundation, a incidência global de câncer de pele continua a aumentar [1]. Em 2019, estima-se que 192.310 casos de melanoma serão diagnosticados nos Estados Unidos [2].

No entanto, as formas mais comuns de câncer de pele são as não melanocíticas, como o Carcinoma Basocelular (CBC) e o Carcinoma Espinocelular (CEC). O câncer de pele não melanoma é o câncer que ocorre mais comumente em homens e mulheres, com mais de 4,3 milhões de casos de CBC e 1 milhão de casos de CEC diagnosticados a cada ano nos Estados Unidos, embora esses números sejam provavelmente subestimados [3]. O diagnóstico precoce do câncer de pele é fundamental para melhorar os resultados e está correlacionado com 99% de sobrevivência global (SG). No entanto, uma vez que a doença progride para além da pele, a sobrevivência é fraca [4,5].

Na prática médica atual, os dermatologistas examinam os pacientes por inspeção visual com auxílio de ampliação de luz polarizada via dermatoscopia. O diagnóstico médico muitas vezes depende da história do paciente, etnia, hábitos sociais e exposição ao sol. Lesões preocupantes são

biopsiadas em consultório, submetidas ao laboratório, processadas como cortes permanentes de parafina e examinadas como lâminas de vidro representativas por um patologista para fazer um diagnóstico.

As soluções de diagnóstico auxiliado por computador (CAD) habilitadas para IA estão preparadas para revolucionar a medicina e os cuidados de saúde, especialmente em imagens médicas. Imagens médicas, incluindo ultrassonografia, tomografia computadorizada (TC) e ressonância magnética (RM), são amplamente utilizadas na prática clínica. No domínio dermatológico, a dermatoscopia ou, menos frequentemente, a microscopia confocal, permite uma visualização in vivo mais detalhada das características lesionadas e a estratificação de risco [6–10]. Em vários estudos, os algoritmos de IA igualam ou excedem o desempenho do médico na detecção de doenças em imagens médicas [11,12]. Recentemente, o aprendizado profundo forneceu várias soluções ponta a ponta na detecção de anormalidades, como câncer de mama, tumores cerebrais, câncer de pulmão, câncer de esôfago, lesões de pele e úlceras nos pés em múltiplas modalidades de imagem médica [13–17].

Na última década, os avanços na tecnologia levaram a uma maior acessibilidade a técnicas avançadas de imagem, como imagens de corpo inteiro em 3D.

\* Autor correspondente.

Endereço de e-mail: [manu.goyal@dartmouth.edu](mailto:manu.goyal@dartmouth.edu) (M. Goyal).<https://doi.org/10.1016/j.combiomed.2020.104065> Recebido

em 14 de agosto de 2020; Recebido de forma revisada em 15 de outubro de 2020; Aceito em 15 de outubro de 2020.

Disponível online em 27 de outubro de 2020.

0010-4825/© 2020 O(s) Autor(es). Publicado pela Elsevier Ltd. Este é um artigo de acesso aberto sob a licença CC BY (<http://creativecommons.org/licenses/by/4.0/>).

fotografia/digitalização, dermatoscopia, câmeras de alta resolução e scanners digitais de slides inteiros que são usados para coletar dados de câncer de pele de alta qualidade de pacientes em todo o mundo [18,19]. A International Skin Imaging Collaboration (ISIC) é uma força motriz que fornece conjuntos de dados digitais de imagens de lesões de pele com anotações especializadas para soluções CAD automatizadas para o diagnóstico de melanoma e outros cânceres de pele. Um amplo interesse de pesquisa em soluções de IA para diagnóstico de câncer de pele é facilitado pela Internet acessível e de alta velocidade, poder computacional e armazenamento seguro em nuvem para gerenciar e compartilhar conjuntos de dados de câncer de pele. Esses algoritmos podem ser escaláveis para vários dispositivos, plataformas e sistemas operacionais, transformando-os em instrumentos médicos modernos [20].

O objetivo desta revisão é fornecer ao leitor uma atualização sobre o desempenho de algoritmos de inteligência artificial utilizados para o diagnóstico de câncer de pele em diversas modalidades de conjuntos de dados de lesões cutâneas, especialmente em termos de estudos comparativos sobre o desempenho de algoritmos de IA. e dermatologistas/dermatopatologistas. Dedicamos subseções separadas para organizar esses estudos de acordo com os tipos de modalidade de imagem utilizada, incluindo fotografias clínicas, imagens de dermatoscopia e digitalização de patologia de lâmina inteira. Especificamente, procuramos discutir os desafios técnicos neste domínio e as oportunidades para melhorar as atuais soluções de IA para que possam ser utilizadas como uma ferramenta de apoio aos médicos para aumentar a sua eficiência no diagnóstico de cânceres de pele.

2. Inteligência artificial para câncer de pele

Os maiores avanços nesse campo vieram do trabalho de Esteva et al. [12] que usaram um algoritmo de aprendizagem profunda em um conjunto de dados de pele combinados de 129.450 imagens clínicas e dermatoscópicas, consistindo de 2.032 doenças de lesões cutâneas diferentes. Eles compararam o desempenho de um método de aprendizagem profunda com 21 dermatologistas certificados para classificação e diferenciação de carcinomas versus ceratoses seborreicas benignas; e melanomas versus nevos benignos. O desempenho da IA demonstrou estar no mesmo nível do desempenho dos dermatologistas na classificação do câncer de pele. Três tipos principais de modalidades são utilizadas para classificação e diagnóstico de lesões cutâneas nos trabalhos aqui descritos: imagens clínicas, imagens dermatoscópicas e imagens histopatológicas. Nesta seção, começamos com a análise de conjuntos de dados de lesões cutâneas disponíveis publicamente, depois fornecemos diferentes subseções dedicadas à solução de inteligência artificial relacionada a cada tipo de modalidade de imagem e breves detalhes dos desafios do ISIC.

2.1. Conjuntos de dados publicamente disponíveis para câncer de pele

- 1) Arquivo ISIC: A galeria do arquivo ISIC consiste em muitos conjuntos de dados clínicos e dermatoscópicos de lesões cutâneas de todo o mundo, como conjuntos de dados ISIC Challenges [21], HAM10000 [22] e BCN20000 [23].
- 2) Atlas Interativo de Dermatoscopia [24]: O Atlas Interativo de Dermatoscopia possui 1000 casos clínicos (270 melanomas, 49 ceratoses seborreicas), cada um com pelo menos duas imagens: dermatoscópica e clínica close-up. Está disponível para fins de investigação e tem um custo de 250€.
- 3) Biblioteca de Imagens Dermofit [25]: A Biblioteca de Imagens Dermofit consiste em 1300 imagens de alta resolução com 10 classes de lesões de pele. É necessário um acordo de licenciamento com uma taxa única de 75€ e está disponível uma licença acadêmica.
- 4) Conjunto de dados PH2 [26]: O conjunto de dados PH2 possui 200 imagens dermatoscópicas (40 casos de melanoma e 160 nevos). Ele está disponível gratuitamente após a assinatura de um breve formulário de registro online.
- 6) Conjunto de dados MED-NODE [27]: Consiste em 170 imagens clínicas (70 casos de melanoma e 100 nevos). Este conjunto de dados está disponível gratuitamente para download para pesquisa.
- 7) Asan Dataset [28,29]: É uma coleção de 17.125 imagens clínicas de 12 tipos de doenças de pele encontradas em povos asiáticos. O conjunto de dados de teste Asan (1276 imagens) está disponível para download para pesquisa.

- 8) Hallym Dataset [28]: Este conjunto de dados consiste em 125 imagens clínicas de casos de CBC.
- 9) Conjunto de dados SD-198 [30]: O conjunto de dados SD-198 é um conjunto de dados clínicos de lesões cutâneas contendo 6.584 imagens clínicas de 198 doenças de pele. Este conjunto de dados foi capturado com câmeras digitais e telefones celulares.
- 10) Conjunto de dados SD-260 [31]: Este conjunto de dados é um conjunto de dados mais equilibrado quando comparado ao conjunto de dados SD-198 anterior, pois controla a distribuição do tamanho da classe com preservação de 10 a 60 imagens para cada categoria. É composto por 20.600 imagens com 260 doenças de pele.
- 11) Dermnet NZ [32]: Dermnet NZ possui uma das maiores e mais diversas coleções de imagens clínicas, dermatoscópicas e histológicas de diversas doenças de pele. Essas imagens podem ser usadas para fins de pesquisa acadêmica. Eles têm imagens adicionais de alta resolução para compra.
- 12) Derm7pt [33]: Este conjunto de dados possui 1.011 imagens dermatoscópicas (252 casos de melanoma e 759 nevos), com critérios de lista de verificação de 7 pontos.
- 13) Atlas do Genoma do Câncer [34]: Este conjunto de dados é uma das maiores coleções de lâminas de lesões cutâneas patológicas, com 2.871 casos. Está disponível publicamente para uso da comunidade de pesquisa.

O resumo dos conjuntos de dados de lesões cutâneas mencionados acima, como modalidade, número total de imagens, número de imagens em cada classe, URL e custo, é apresentado na Tabela 1.

2.2. Inteligência artificial em imagens dermatoscópicas

A dermatoscopia é a inspeção/exame de lesões cutâneas com um aparelho dermatoscópico composto por uma lente de aumento de alta qualidade e um sistema de iluminação (polarizável). As imagens dermatoscópicas são capturadas com câmera reflex digital de lente única (DSLR) de alta resolução ou com acessórios de câmera de smartphone. O uso de imagens dermatoscópicas para algoritmos de IA está se tornando um campo de pesquisa muito popular desde a introdução de muitos grandes conjuntos de dados dermatoscópicos disponíveis publicamente, consistindo de diferentes tipos de lesões cutâneas benignas e cancerígenas, como mostrado na Fig. diagnóstico de lesões usando conjuntos de dados dermatoscópicos de lesões cutâneas, listados abaixo.

- 1) Codella et al. [35] desenvolveram um conjunto de algoritmos de aprendizagem profunda no conjunto de dados ISIC-2016 e compararam o desempenho desta rede com 8 dermatologistas para a classificação de 100 lesões cutâneas como benignas ou malignas. O método conjunto superou o desempenho médio dos dermatologistas, alcançando precisão de 76% e especificidade de 62% versus 70,5% e 59% alcançados pelos dermatologistas.
- 2) Haenssle et al. [36] treinaram um método de aprendizagem profunda InceptionV4 em um grande conjunto de dados dermatoscópicos que consiste em mais de 100.000 lesões benignas e imagens de melanoma e compararam o desempenho de um método de aprendizagem profunda com 58 dermatologistas. No conjunto de testes de 100 casos (75 lesões benignas e 25 casos de melanoma), os dermatologistas tiveram uma sensibilidade média de 86,6% e especificidade de 71,3%, enquanto o método de aprendizagem profunda alcançou uma sensibilidade de 95% e especificidade de 95%. 63,8%.
- 3) Brinker et al. [37] compararam o desempenho de 157 dermatologistas certificados em 12 hospitais universitários alemães com um método de aprendizagem profunda (ResNet50) para 100 imagens dermatoscópicas (MClass-D) consistindo de 80 nevos e 20 casos de melanoma. Os dermatologistas alcançaram uma sensibilidade geral de 74,1% e especificidade de 60,0% no conjunto de dados dermatoscópicos, enquanto um método de aprendizagem profunda alcançou uma especificidade de 69,2% e uma sensibilidade de 84,2%.
- 4) Tschandl et al. [38] usaram arquiteturas populares de aprendizagem profunda conhecidas como InceptionV3 e ResNet50 em um conjunto de dados combinado de 7.895 imagens dermoscópicas e 5.829 imagens de lesões em close-up para diagnóstico de câncer de pele não pigmentado. na experiência. Os algoritmos de aprendizagem profunda alcançaram precisão equivalente à humana

**Tabela**

1 Breves detalhes de conjuntos de dados de lesões cutâneas disponíveis publicamente com diferentes modalidades onde ade é Adenomas e Adenocarcinoma, adn é Adnexal, akiec é Queratose Actínica, bcc é Carcinoma Basocelular, bkl é Queratose Benigna, df é Dermatofibroma, hae é Haemangioma, ic é Carcinoma Intraepitelial, len é Lentigo, mel é melanoma, nv é Nevo, pg é Granuloma Píogênico, sk é Queratose Seborreica, scc é Carcinoma de Células Escamosas, unk é desconhecido, vasc é Lesão Vascular.

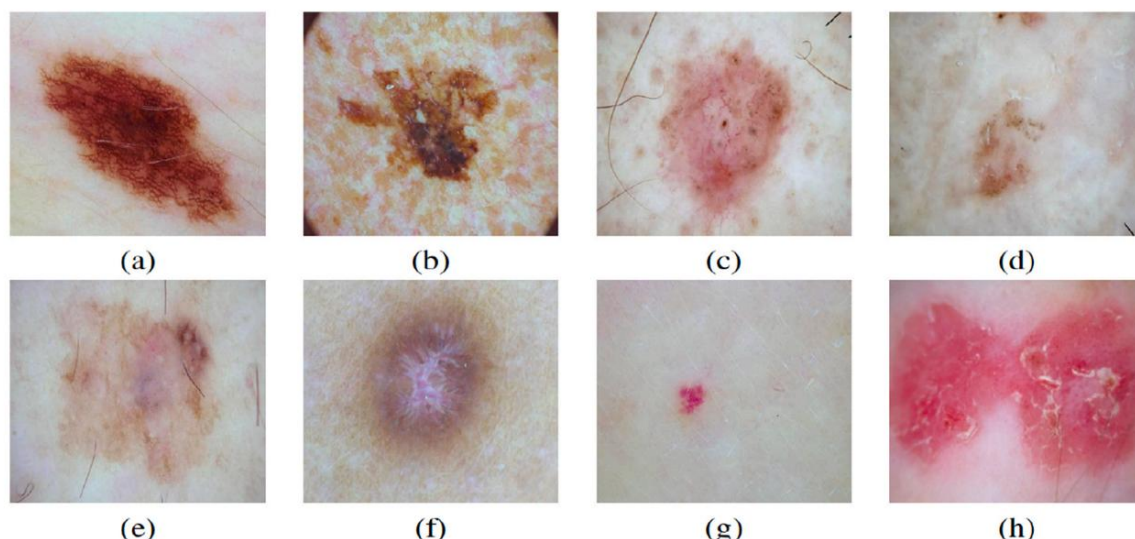
Conjunto de dados	Modalidade	Nº de imagens	Aulas e nº de Casos	URL	Custo
ISIC-2016	Dermoscópico	1279	mel – 248 nv - 1031	<a href="https://challenge.isic-archive.com/data">https://challenge.isic-archive.com/data</a>	Grátis para pesquisa
ISIC-2017	Dermoscópico	2750	mel - 521 nv – 1843 sk-386		
ISIC-2018 (HAM10000)	Dermoscópico	10015	akiec - 327 Cco - 514 bkl - 1099 df - 115 mel - 1113 nv - 6705 vasc – 142		
ISIC-2019 (HAM10000, BCN_20000 e MSK)	Dermoscópico	25331	akiec y 867 bcc - 3323 bkl - 2624 df - 239 mel - 4522 nv - 12875 scc - 628 vasc y 253		
Atlas interativo de Dermatoscopia	Dermoscópico & Clínico	1000	mel - 270 sk - 49 unk - 681	<a href="http://www.dermoscopia.org/atlas/default.asp">http://www.dermoscopia.org/atlas/default.asp</a>	250€
Biblioteca de imagens Dermofit	Dermoscópico	1300	akiec - 45 bcc - 239 df - 65 haha - 97 eu - 78 mel - 76 nv – 331 página - 24 sk - 257 scc - 88 mel - 40 nv - 160	<a href="https://licensing.edinburgh-innovations.ed.ac.uk/i/software/dermofit-image-library.html">https://licensing.edinburgh-innovations.ed.ac.uk/i/software/dermofit-image-library.html</a> _	£ 75
Conjunto de dados PH2	Dermoscópico	200	88 mel - 40 nv - 160	<a href="https://www.fc.up.pt/addi/ph2%20database.html">https://www.fc.up.pt/addi/ph2%20database.html</a>	Grátis para pesquisa
Conjunto de dados MED-NODE	Clínico	170	mel - 70 nv - 100	<a href="http://www.cs.rug.nl/%7Eimaging/databases/melano_ma_naevi/">http://www.cs.rug.nl/%7Eimaging/databases/melano_ma_naevi/</a> <a href="https://figshare.com/articles/Asan_and_Hallym_Datas_et_Thumbnails_/5406136">https://figshare.com/articles/Asan_and_Hallym_Datas_et_Thumbnails_/5406136</a>	
Conjunto de dados ASAN	Clínico	17125	akiec - 651 bcc - 1082 df - 1247 hae - 2715 ic - 918  lente - 1193 mel - 599 nv - 2706 pg - 375 scc - 1231 sk - 1423 verruca - 2985		
Conjunto de dados HALLYM	Clínico	125	Cco - 125		
SD-198	Clínico	6584	198 classes de pele	<a href="http://xiaopingwu.cn/assets/projects/sd-198/">http://xiaopingwu.cn/assets/projects/sd-198/</a>	
SD-260	Clínico	20660	260 classes de pele		
DermNet NZ	Clínico, Dermoscópico e Patologia	> 20.000 desconhecidos	Doenças	<a href="https://dermnetnz.org">https://dermnetnz.org</a>	Disponível para Comprar
Derm7pt	Dermoscópico	1011	mel-252 nv-759	<a href="http://derm.cs.sfu.ca/Welcome.html">http://derm.cs.sfu.ca/Welcome.html</a>	Grátis para pesquisa
O Atlas do Genoma do Câncer	Patologia	2871	ade - 207 adn – 57 bcc - 44 mel e nv - 2319 scc - 197 oth - 47	<a href="https://portal.gdc.cancer.gov/">https://portal.gdc.cancer.gov/</a>	

especialistas e superou os grupos humanos com avaliadores iniciantes e intermediários.

- 5) Maron et al. [39] compararam a sensibilidade e a especificidade de um método de aprendizagem profunda (ResNet50) com 112 dermatologistas alemães para classificação multiclasse de lesões cutâneas que incluem nevus,

melanoma, ceratose benigna, CBC e CEC (também ceratose solar e carcinoma intraepitelial). O método de aprendizagem profunda superou os dermatologistas em um nível significativo (p y 0,001).

- 6) Haenssle et al. [40] compararam a arquitetura de aprendizagem profunda baseada no InceptionV4 (aprovado como dispositivo médico pela União Europeia) e



**Figura 1.** Ilustração de diferentes tipos de lesões cutâneas dermatoscópicas onde (a) Nevo (b) Melanoma (c) Carcinoma basocelular (d) Queratose actínica (e) Queratose benigna (f) Dermatofibroma (g) Lesão Vascular (h) Carcinoma de células escamosas [22].

dermatologistas em um conjunto de testes dermatoscópicos consiste em 100 casos (60 lesões benignas e 40 malignas). Este estudo foi realizado em dois níveis, ou seja, nível I: imagem dermatoscópica; nível II: imagens clínicas adicionais em close, imagem dermatoscópica e informações clínicas. O algoritmo de aprendizagem profunda alcançou pontuação de sensibilidade e especificidade de 95% e 76,7% respectivamente, enquanto sensibilidade e especificidade médias de 89% e 80,7% respectivamente alcançadas por dermatologistas no nível I. Com mais informações no nível II, a sensibilidade média dos dermatologistas aumentou para 94,1%, enquanto a especificidade média permaneceu a mesma.

- 7) Tschandl et al. [41] compararam o desempenho médio de ambos os algoritmos de IA (139 no total) participaram do desafio ISIC 2018 e 511 leitores humanos em um conjunto de teste de 1.511 imagens. Nos resultados, os algoritmos de IA alcançaram diagnósticos mais corretos do que os leitores humanos.

### 2.3. Inteligência artificial em imagens clínicas

Imagens clínicas são rotineiramente capturadas de diferentes lesões cutâneas com câmeras móveis para exame remoto e incorporação nos registros médicos do paciente, conforme mostrado na Fig. 2. Como as imagens clínicas são capturadas com diferentes câmeras com fundos, iluminação e cor variáveis, essas imagens fornecem diferentes insights para imagens dermatoscópicas.

- 1) Yang et al. [30] realizaram diagnóstico clínico de lesões cutâneas usando representação inspirada na regra ABCD no conjunto de dados SD-198. Eles compararam o desempenho dos métodos propostos com métodos de aprendizagem profunda e dermatologistas. Alcançou uma pontuação de 57,62% (precisão) em comparação com o aprendizado profundo de melhor desempenho

método (ResNet), que atingiu 53,35%. Quando comparados com os médicos, apenas os médicos seniores com experiência considerável em doenças de pele alcançaram uma precisão média de 83,29%.

- 2) Han et al. [28] treinaram uma arquitetura de aprendizado profundo (ResNet-152) para classificar as imagens clínicas de 12 doenças de pele em um conjunto de dados de treinamento Asan, um conjunto de dados MED-NODE e imagens de sites de atlas, e testaram em um conjunto de testes Asan e um Edinburgh Conjunto de dados (Dermofit). O desempenho do algoritmo esteve no mesmo nível da equipe de 16 dermatologistas em 480 imagens escolhidas aleatoriamente do conjunto de dados de teste Asan (260 imagens) e do conjunto de dados de Edimburgo (220 imagens), enquanto o sistema de IA superou os dermatologistas no diagnóstico de CBC.
- 3) Fujisawa et al. [42] testaram um método de aprendizagem profunda em 6.009 imagens clínicas de 14 diagnósticos, incluindo condições malignas e benignas. O algoritmo de aprendizagem profunda alcançou uma precisão diagnóstica de 76,5%, superior ao desempenho de 13 dermatologistas certificados (59,7%) e nove estagiários de dermatologia (41,7%) em um conjunto de dados de 140 imagens.
- 4) Brinker et al. [43] compararam o desempenho de 145 dermatologistas e um método de aprendizagem profunda (ResNet50) para o caso de teste de 100 imagens clínicas de lesões cutâneas (MClass-ND) consistindo em 80 casos de nevos e 20 casos de melanoma verificados por biópsia. Os dermatologistas alcançaram uma sensibilidade geral de 89,4%, uma especificidade de 64,4% e um AUROC de 0,769, enquanto um método de aprendizagem profunda alcançou a mesma sensibilidade e melhor pontuação de especificidade de 69,2%.

### 2.4. Inteligência artificial em imagens histopatológicas

Diagnóstico de câncer de pele é confirmado por dermatopatologistas



**2.** Ilustração de diferentes tipos de lesões clínicas da pele onde (a) Queratose Benigna (b) Melanoma (c) CBC (d) CEC [19].



com base na avaliação microscópica de uma biópsia de tecido. As soluções de aprendizagem profunda têm sido bem-sucedidas no campo da patologia digital com imagens de lâminas inteiras. Exemplos de imagens histopatológicas de lesões de pele são mostrados na Fig. 3. Essas técnicas são usadas para a classificação de amostras de tecido de biópsia para diagnosticar o número de cânceres, como pele, pulmão e mama. Nesta seção, exploramos os métodos de aprendizagem profunda usados na histopatologia digital específica para câncer de pele.

- 1) Heckler et al. [44] utilizaram um método de aprendizagem profunda (ResNet50) para comparar o desempenho de patologistas na classificação de melanoma e nevos. O modelo de aprendizagem profunda foi treinado em um conjunto de dados de 595 imagens histopatológicas (300 melanomas e 295 nevos) e testado em 100 imagens (melanoma/nevos = 1:1). A discordância total com o histopatologista foi de 18% para melanoma, 20% para nevos e 19% para o conjunto completo de imagens.
- 2) Jiang et al. [45] propuseram o uso de um algoritmo de aprendizagem profunda em imagens histopatológicas digitais (MOI) capturadas por smartphone para a detecção de CBC. Eles descobriram que o desempenho do algoritmo em MOI e Whole Slide Imaging (WSI) é comparável a uma pontuação AUC de 0,95. Eles introduziram uma rede de segmentação profunda para análise aprofundada dos casos difíceis para melhorar ainda mais o desempenho com pontuação de 0,987 (AUC), 0,97 (sensibilidade) e 0,94 (especificidade).
- 3) Cruz-Roa et al. [46] usaram uma arquitetura de aprendizagem profunda para discriminar entre CBC e padrões de tecido normal em 1.417 imagens de 308 regiões de interesse (ROI) de imagens histopatológicas da pele. Eles compararam o método de aprendizado profundo com o aprendizado de máquina tradicional com descritores de recursos, incluindo o conjunto de recursos, transformação wavelet canônica e baseada em Haar. A arquitetura de aprendizagem profunda provou ser superior às abordagens tradicionais, alcançando 89,4% em F-Measure e 91,4% em precisão balanceada.
- 4) Xie et al. [47] introduziram um grande conjunto de dados de 2.241 imagens histopatológicas de 1.321 pacientes de 2008 a 2018. Eles usaram duas arquiteturas de aprendizagem profunda, VGG19 e ResNet50, nos 9,95 milhões de patches gerados em 2.241 imagens histopatológicas para testar a classificação de melanoma e nevos em diferentes escalas de ampliação. Eles alcançaram alta precisão na distinção de melanoma de nevos com F1 médio (0,89), sensibilidade (0,92), especificidade (0,94) e AUC (0,98).

## 2.5. Desafios da ISIC

A International Skin Imaging Collaboration (ISIC) é uma força motriz ao fornecer conjuntos de dados de imagens digitais de lesões cutâneas com anotações de especialistas em todo o mundo para soluções CAD automatizadas para o diagnóstico precoce de melanoma e outros tipos de câncer. O principal objetivo desta comunidade é reduzir as mortes relacionadas ao câncer de pele e o número de biópsias, melhorando a precisão, a sensibilidade e a especificidade da detecção precoce do melanoma e de outros cânceres de pele. Esta comunidade também organiza os desafios anuais de lesões cutâneas para atrair uma participação mais ampla dos investigadores para melhorar o diagnóstico dos algoritmos CAD e difundir a conscientização sobre o cancro da pele. A cada ano aumenta o número de imagens e categorias de lesões cutâneas.

O desafio da segmentação da pele consiste na segmentação de

limites da lesão, independentemente dos tipos de lesões cutâneas. Embora a segmentação de lesões cutâneas seja crucial para apoiar os dermatologistas no fornecimento de informações sobre as diversas características e atributos das lesões cutâneas, como a regra ABCD, o escopo do nosso artigo é limitado às soluções de IA para diagnóstico de lesões. Portanto, não incluímos nenhum estudo relacionado à tarefa de segmentação de classe única.

Mais recentemente, na ISIC-2019, excluíram o desafio de segmentação da pele e focaram apenas nos desafios de classificação das lesões (apenas imagem e imagem com metadados). A Tabela 2 a seguir constitui breves detalhes do desafio de diagnóstico de lesões ISIC de cada ano. Outra coisa interessante sobre os desafios do ISIC, a partir de 2018, a organização decidiu não divulgar o conjunto de testes e configurar o servidor de avaliação online para verificar o desempenho do algoritmo.

A maioria dos participantes confiou em algoritmos de aprendizagem profunda para classificar as lesões cutâneas para diagnóstico de lesões. Semelhante a outros desafios de classificação de imagens médicas, a maioria dos algoritmos vencedores nos últimos anos são baseados em algoritmos de aprendizagem profunda em conjunto, conforme mostrado na Fig. 4. Notavelmente, com a introdução de mais categorias e imagens no desafio de 2019, o algoritmo vencedor a precisão caiu de 88,5% para 63,6%.

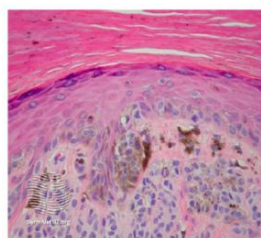
## 3. Desafios em inteligência artificial

Com os algoritmos de aprendizagem profunda ultrapassando os benchmarks dos conjuntos de dados populares de visão computacional em um curto período, a mesma tendência também poderia ser esperada no desafio do diagnóstico de lesões cutâneas. No entanto, à medida que exploramos mais o desafio do diagnóstico de lesões cutâneas, esta tarefa parece não ser simples como os desafios ImageNet, PASCAL-VOC, MS-COCO num domínio não médico [48,49]. Existem semelhanças intraclases e diferenças interclases em relação à cor, textura, tamanho, localização e aparência visual das lesões cutâneas. Algoritmos de aprendizagem profunda geralmente requerem uma quantidade substancial de dados de treinamento diversos, equilibrados e de alta qualidade que representam cada classe de lesões cutâneas para melhorar a precisão do diagnóstico. Para conjuntos de dados de lesões cutâneas de várias modalidades, há muito mais questões relacionadas ao diagnóstico de câncer de pele com soluções de IA, conforme discutido abaixo.

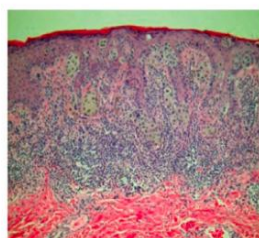
### Tabela

**2** Resumo do desafio de diagnóstico de lesão ISIC onde akiec é ceratose actínica, bcc é carcinoma basocelular, bkl é ceratose benigna, df é dermatofibroma, mel é melanoma, nv é Nevi, sk é ceratose seborreica, scc é carcinoma de células escamosas, e vasc é lesão vascular.

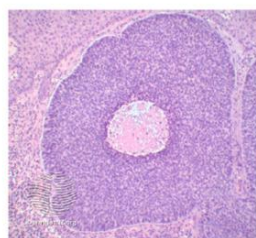
ISIC Desafio Ano	Aulas	Nº de imagens	Avaliação Métrica	Ganhando Algoritmo Pontuação
2016	mel e nv mel,	1279	Média AUC de	0,637
2017	nv e sk 2750 akiec, bcc, bkl,		precisão	0,874
2018	11720 df, mel, nv e		Multi balanceado	0,885
	Vasco		precisão de classe	
2019	akiec, bcc, bkl, df,	33569	Precisão multiclasse	0,636
	mel, nv, scc e vasc		balanceada	



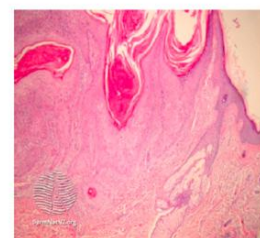
(a)



(b)



(c)



(d)

**3.** Ilustração de diferentes tipos de imagens histopatológicas onde (a) Nevo (b) Melanoma (c) Carcinoma Basocelular (d) Carcinoma Espinocelular [19].

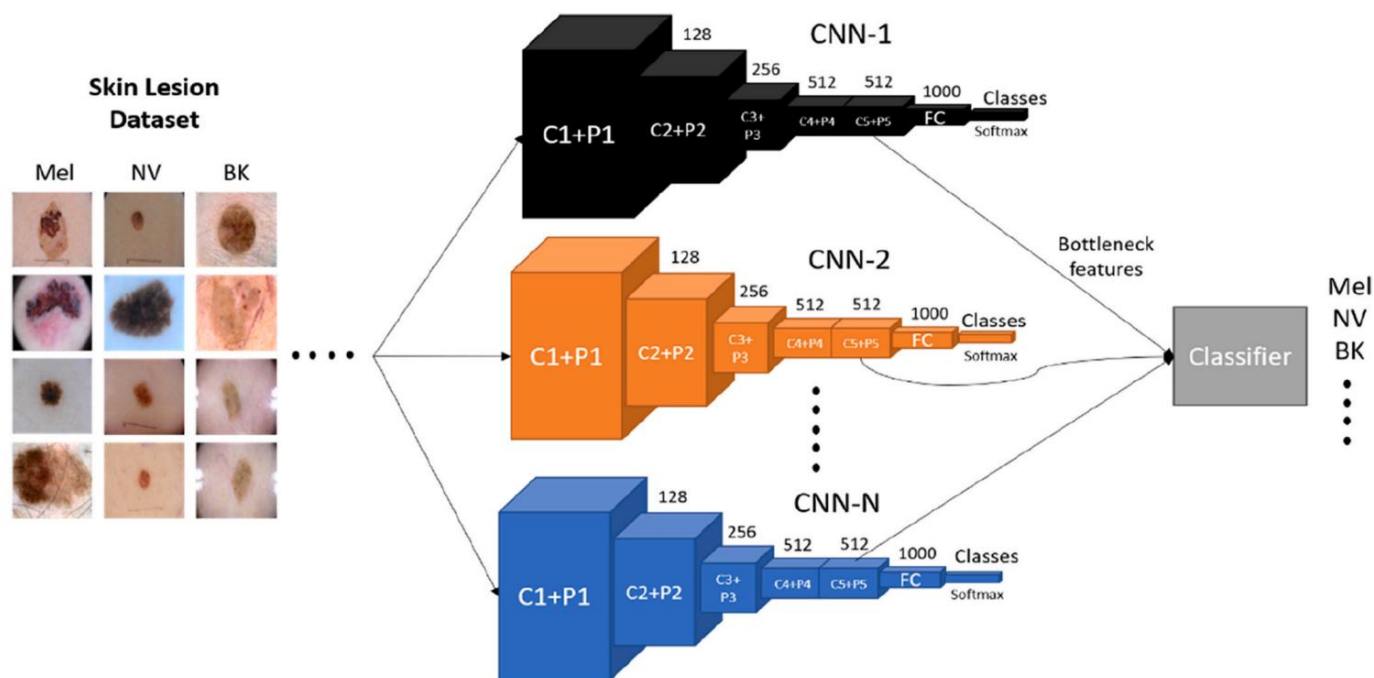


Figura 4. Abordagem Ensemble CNN para classificação de lesões cutâneas.

### 3.1. Desempenho de aprendizagem profunda e conjuntos de dados desequilibrados

O desempenho dos algoritmos de aprendizagem profunda depende principalmente da qualidade dos conjuntos de dados de imagens, em vez de ajustar os hiperparâmetros das redes, como é comumente visto nos diferentes conjuntos de dados de lesões cutâneas disponíveis publicamente. Geralmente há mais casos de lesões cutâneas benignas do que de lesões malignas. A maioria das arquiteturas de aprendizagem profunda são projetadas em um conjunto de dados balanceado, como o ImageNet, que consiste em 1.000 imagens por classe (1.000 classes) [48]. Consequentemente, o desempenho de um algoritmo de aprendizado profundo geralmente sofre com conjuntos de dados desequilibrados, apesar de usar truques de ajuste como uma penalidade para falsos negativos encontrados em classes de lesões cutâneas menores durante o treinamento usando funções de perda personalizadas.

### 3.2. Curioso caso de imagens histopatológicas/patologia digital

O tamanho das imagens em conjuntos de dados clínicos e dermatoscópicos de lesões cutâneas varia entre 1200 x 768 e 3648 x 2736, dependendo da câmera utilizada. A maioria dos algoritmos de aprendizagem profunda são geralmente desenvolvidos e validados em grandes conjuntos de dados de formação não médica. Esses algoritmos de aprendizagem profunda funcionaram muito bem em conjuntos de dados clínicos e dermatoscópicos de lesões cutâneas, ajustando os algoritmos por meio de técnicas de aprendizagem por transferência. Por outro lado, os exames histopatológicos consistem em milhões de pixels e suas dimensões são comumente maiores que 50.000 x 50.000. Consequentemente, existem muitos desafios técnicos para aprendizado profundo ou algoritmos de IA em patologia digital, como falta de dados rotulados, padrão infinito de diferentes tipos de tecidos, extração de recursos de alta qualidade, altos gastos computacionais e muito mais [50].

### 3.3. Histórico médico dos pacientes e metadados clínicos

O histórico médico, os hábitos sociais e os metadados clínicos dos pacientes são considerados ao fazer um diagnóstico de câncer de pele. É muito importante conhecer os metadados de diagnóstico, como histórico familiar e do paciente de câncer de pele, idade, etnia, sexo, localização anatômica geral, tamanho e estrutura da lesão cutânea, ao realizar uma inspeção visual de uma lesão cutânea suspeita com dermatoscopia. Consequentemente, apenas os algoritmos de aprendizagem profunda baseados em imagens usados para o diagnóstico de câncer de pele falham em aspectos-chave das informações clínicas e do paciente. Está comprovado em um estudo anterior [36]

que o desempenho dos dermatologistas 'iniciantes' e 'qualificados' é melhorado com a disponibilidade de informações clínicas e que eles tiveram um desempenho melhor do que os algoritmos de aprendizagem profunda. Infelizmente, tanto o histórico do paciente quanto os metadados clínicos estão faltando nos conjuntos de dados de lesões cutâneas mais disponíveis publicamente.

### 3.4. Regra ABCDE, lista de verificação de 7 pontos e conjuntos de dados de linha do tempo

No ambiente clínico, uma lesão suspeita é inspecionada visualmente com o auxílio da dermatoscopia. A regra ABCDE e a lista de verificação de 7 pontos são consideradas regras importantes para diferenciar manchas benignas (nevus) de melanoma. O ABCDE inclui se a lesão é assimétrica, tem bordas irregulares, apresenta múltiplas cores, se o diâmetro da lesão é maior que seis milímetros e se houve alguma evolução ou alteração na composição da lesão. A lista de verificação de 7 pontos inclui três critérios principais (rede pigmentar atípica, áreas cinza-azuladas e padrão vascular atípico) e quatro critérios menores (estrias, manchas, pontos e glóbulos irregulares e padrão de regressão) [51,52]. Apesar da disponibilidade de conjuntos de dados dermatoscópicos de lesões cutâneas, os algoritmos de aprendizagem profunda não funcionam da mesma maneira nem procuram um padrão semelhante à regra ABCDE em que os médicos confiam. Isso se deve principalmente à complexidade do reconhecimento de padrões para as características dos cânceres de pele em imagens médicas.

É por isso que, apesar das tentativas recentes dos investigadores para desmistificar o funcionamento dos algoritmos de aprendizagem profunda, tais esforços ainda são considerados uma abordagem de caixa negra, especialmente em imagens médicas. Como não existem conjuntos de dados dermatoscópicos de linha do tempo disponíveis publicamente, não é possível determinar a mudança nas características de uma lesão de acordo com a evolução da regra ABCDE.

### 3.5. A biópsia é obrigatória

Embora, em vários estudos, as soluções de IA tenham superado os especialistas humanos no diagnóstico de câncer de pele. Mas, mesmo que o cancro da pele seja confirmado por soluções de IA com uma elevada taxa de confiança, uma biópsia e um teste histológico ainda devem ser realizados para confirmar um diagnóstico, semelhante à prática clínica real de dermatologistas e clínicos gerais. A precisão do diagnóstico dos algoritmos de aprendizagem profunda também pode ser enganosa. Por exemplo, se um conjunto de testes consistir em 20 casos de melanoma e 80 casos de nevus, e

Se a precisão geral do diagnóstico for de 90% (100% em nevus e 50% em casos de melanoma), é perigoso usar um algoritmo de aprendizagem profunda neste caso como meio de fornecer um diagnóstico de melanoma. Como o diagnóstico incorreto de um paciente com câncer por um algoritmo de aprendizagem profunda pode causar risco de morte, uma biópsia deve ser realizada para garantir a segurança e confirmar o diagnóstico do algoritmo.

### 3.6. Semelhanças entre classes (imitações de lesões cutâneas)

Várias lesões cutâneas podem mimetizar o câncer de pele tanto em ambientes clínicos quanto microscópicos, o que pode resultar em diagnósticos incorretos. Por exemplo, em imagens clínicas e dermatoscópicas, a ceratose seborreica pode mimetizar cânceres de pele, incluindo CBC, CEC e melanoma. Nas imagens histopatológicas, existem muitos imitadores histológicos do CBC, como CEC, tumores foliculares benignos, hamartoma folicular basaloide, um tumor do infundíbulo folicular, siringoma e carcinoma aneal microcístico [53]. Assim, os algoritmos de aprendizagem profunda, quando treinados em classes limitadas de lesões cutâneas num conjunto de dados, não distinguem de forma fiável os cânceres da pele dos seus imitadores conhecidos.

### 3.7. Dissimilaridades intraclasse

Diversas lesões cutâneas apresentam diferenças intraclasse em termos de cor, atributo, textura, tamanho, localização. Consequentemente, essas lesões cutâneas são posteriormente categorizadas em muitas subcategorias com base na aparência visual. Por exemplo, a cor da maioria dos melanomas é preta devido ao pigmento escuro da melanina. Mas certos melanomas apresentam cor de pele normal, aparência avermelhada e rosada. Da mesma forma, o CBC possui muitas subcategorias, como CBC nodular, CBC superficial, CBC morfeiforme, carcinoma basoescomoso, e sua aparência é completamente diferente uma da outra, variando da cor branca ao vermelho, como mostrado na Figura 5.

### 3.8. Barreira de comunicação entre IA e dermatologistas

Às vezes, até mesmo os especialistas em visão computacional têm dificuldade em compreender as decisões tomadas pelas estruturas de aprendizagem profunda. Por exemplo, se existir um algoritmo com 85% de precisão para o diagnóstico de cancro da pele, muitas vezes é muito difícil compreender porque é que o algoritmo está a fazer inferências erradas sobre os restantes 15% dos casos e como melhorar essas decisões. Esses algoritmos geralmente não são semelhantes ou representativos das maneiras pelas quais os médicos tomam tais decisões. Consequentemente, algoritmos de aprendizado profundo são frequentemente considerados como uma solução de caixa preta que não oferece uma explicação clara para sua conclusão e, muitas vezes, fornecem apenas uma saída em probabilidade de confiança variando de 0 a 1 para classificação de cada lesão de pele em um conjunto de testes. Atualmente, não está claro como os dermatologistas interpretariam os resultados dos modelos de aprendizagem profunda no diagnóstico do câncer de pele.

### 3.9. Dados barulhentos da vida real com fontes de dados heterogêneas

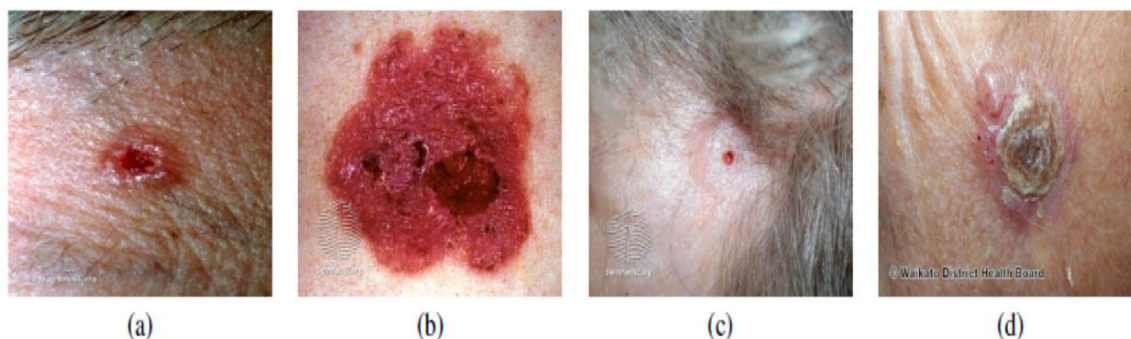
Nos atuais conjuntos de dados de lesões de pele, as imagens dermatoscópicas são capturadas com câmeras DSLR de alta resolução e em ambiente ideal de iluminação e distância de captura. Um algoritmo de aprendizagem profunda treinado nesses conjuntos de dados dermatoscópicos de alta qualidade, alcançando uma precisão diagnóstica razoável, poderia ser potencialmente escalado para aplicações de visão em smartphones. Quando este modelo é testado em múltiplas imagens capturadas por smartphones por diferentes câmeras em diferentes condições de iluminação e distâncias, é difícil obter a mesma precisão de diagnóstico. Os algoritmos de aprendizagem profunda são altamente sensíveis aos dispositivos de câmera usados para capturar os dados, e seu desempenho diminui se um tipo diferente de dispositivo de câmera for usado para teste. As imagens de pele autocapturadas fornecidas pelo paciente são frequentemente de baixa qualidade e não são adequadas para dermatologia digital [54,55].

### 3.10. Raça, etnia e população

A maioria dos casos nos conjuntos de dados atuais de lesões cutâneas pertence a indivíduos de pele clara, e não a pessoas pardas ou de pele escura. Embora o risco de desenvolver cancro da pele seja relativamente elevado entre a população de pessoas de pele clara, as pessoas com pele escura também podem desenvolver cancro da pele e são frequentemente diagnosticadas em fases posteriores [56]. O câncer de pele representa 4–5%, 2–4% e 1–2% de todos os cânceres em hispânicos, asiáticos e negros, respectivamente [57]. Consequentemente, estruturas de aprendizagem profunda validadas para o diagnóstico de câncer de pele em pessoas de pele clara apresentam um risco maior de diagnosticar erroneamente aquelas com pele mais escura [58]. Em um estudo recente, Han et al. [28] treinaram um algoritmo de aprendizado profundo em um conjunto de dados de treinamento Asan que consiste em lesões de pele de asiáticos. Eles relataram uma precisão de 81% no conjunto de testes asiáticos, enquanto relataram uma precisão de apenas 56% no conjunto de dados Dermofit, que consiste em lesões de pele de pessoas caucasianas. Portanto, esta queda na precisão significa uma falta de transferibilidade dos recursos aprendidos dos algoritmos de aprendizagem profunda entre conjuntos de dados que contêm pessoas de uma raça, etnia ou população diferente.

### 3.11. Câncer de pele raro e outras doenças de pele

CBC, CEC e melanoma representam coletivamente 98% de todos os cânceres de pele. No entanto, existem outros cânceres de pele, incluindo o carcinoma de células de Merkel (MCC), carcinomas de apêndice, linfoma cutâneo, sarcoma, sarcoma de Kaposi e secundários cutâneos, que são ignorados pela maioria dos algoritmos. Além desses cânceres de pele raros, existem outras doenças de pele, como úlceras, infecções de pele, neoplasias e granulomas não infecciosos, que podem imitar lesões de pele. Se os algoritmos de aprendizagem profunda forem treinados em conjuntos de dados que não apresentam casos adequados desses cânceres de pele raros e outras doenças de pele mencionadas, há um alto risco de diagnóstico incorreto quando testado nessas doenças de pele.



5. Ilustração de diferenças intraclasse no CBC (a) CBC nodular (b) CBC superficial (c) CBC morfeiforme (d) Carcinoma basoescomoso [32].

### 3.12. Pipeline de diagnóstico incompleto para inteligência artificial

No ambiente clínico, o diagnóstico do câncer de pele é feito pela inspeção da lesão cutânea com ou sem dermatoscopia, seguida de biópsia confirmatória e exame anatomopatológico. O principal problema com os atuais conjuntos de dados de lesões cutâneas disponíveis publicamente é que eles não possuem rótulos completos relacionados ao diagnóstico realizado por um dermatologista.

No entanto, a maioria dos rótulos de classificação para imagens dermatoscópicas de lesões cutâneas são determinados por exame anatomopatológico. Ainda assim, esses conjuntos de dados dermatoscópicos e clínicos de lesões cutâneas não possuem rótulos de classificação patológica correspondentes para desenvolver um pipeline completo de diagnóstico para IA.

## 4. Oportunidades

Os pesquisadores de IA invariavelmente afirmam que seus sistemas excedem o desempenho dos dermatologistas no diagnóstico de câncer de pele. Mas este quadro está longe da realidade, uma vez que estas experiências são realizadas em sistemas fechados com um conjunto definido de regras. Com os muitos desafios mencionados na seção acima, a natureza destas avaliações de desempenho relatadas não chega nem perto do diagnóstico real realizado pelos médicos que tratam do cancro da pele. Frequentemente, os algoritmos de aprendizado profundo são considerados opacos, pois aprendem apenas com valores de pixel de conjuntos de dados de imagem e não possuem nenhum conhecimento de domínio ou realizam inferências lógicas para estabelecer a relação entre diferentes tipos de lesões cutâneas [58]. Mas, no futuro, a aprendizagem profunda poderá ser muito boa para o diagnóstico do cancro da pele com as oportunidades listadas abaixo.

### 4.1. Conjunto de dados balanceado e seleção de casos

Um conjunto de dados balanceado é fundamental para o bom desempenho dos algoritmos de aprendizagem profunda usados para tarefas de classificação. Portanto, são necessários conjuntos de dados equilibrados com uma seleção de casos que representem completamente a categoria daquela lesão cutânea específica, e a contribuição de dermatologistas experientes pode ser muito útil para esta seleção.

### 4.2. Diagnóstico auxiliado por computador para patologia digital

A evolução da imagem de slides inteiros para análise de imagens digitais e clusters de GPU para poder computacional nos últimos anos atraiu o interesse de patologistas e sociedades de visão computacional para o desenvolvimento de diagnóstico auxiliado por computador em patologia digital. Uma das abordagens de IA mais populares para enfrentar o obstáculo da dimensionalidade é usar algoritmos de aprendizagem profunda como classificação de janela deslizante e agregar essas classificações para inferir padrões histológicos predominantes [15].

### 4.3. Constância de cores para iluminação e fontes de dados heterogêneas

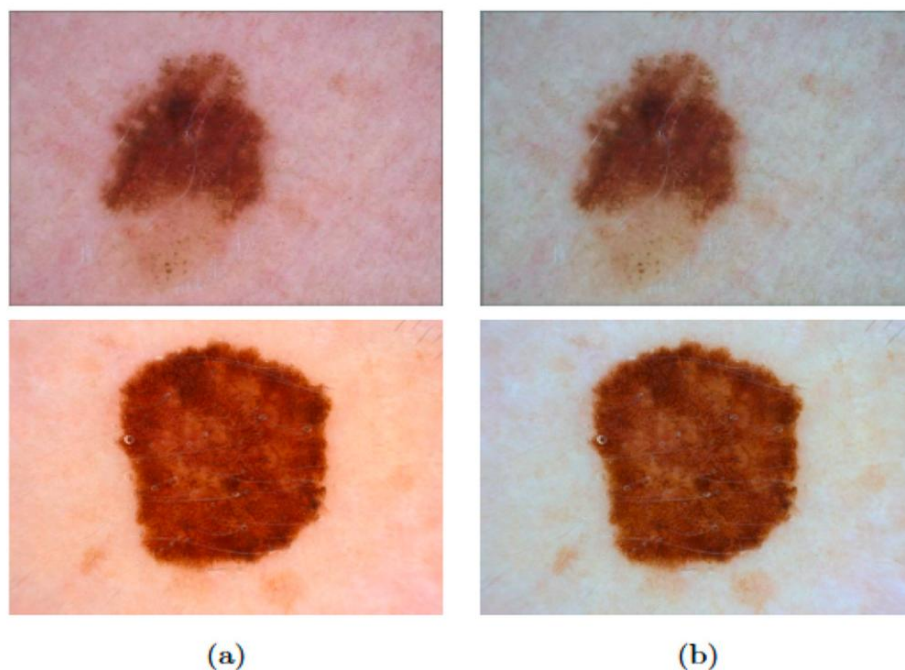
Em conjuntos de dados clínicos e dermatoscópicos disponíveis publicamente, as imagens de lesões cutâneas são adquiridas com diferentes configurações de iluminação e dispositivos de aquisição que podem reduzir o desempenho dos sistemas de IA. Está comprovado em muitos estudos que algoritmos de constância de cores como Shades of Grey, max-RGB podem ser usados para melhorar o desempenho de algoritmos de IA para a classificação de imagens multifonte [59,60]. O exemplo do algoritmo Shadow of Grey como método de pré-processamento para normalizar a iluminação e o efeito de iluminação em imagens dermatoscópicas de lesões cutâneas, conforme mostrado na Figura 6.

### 4.4. Conjuntos de dados diversos

As redes de aprendizagem profunda são frequentemente criticadas por preconceitos sociais devido à maioria dos dados de imagem pertencerem a pessoas de pele clara. Os conjuntos de dados de lesões cutâneas precisam ter diversidade racial, ou seja, devem adicionar casos de lesões cutâneas distribuídos igualmente de pessoas de pele clara e de pele escura para reduzir preconceitos sociais ou étnicos em modelos de aprendizagem profunda. A mesma preocupação pode ser estendida à idade, especialmente quando o grau de envelhecimento da pele ou os danos solares circundantes podem influenciar o conjunto de dados e a tomada de decisões.

### 4.5. Aumento de dados

As técnicas de aumento de dados podem mitigar muitas limitações dos conjuntos de dados, como dados desequilibrados entre as classes de lesões cutâneas e fontes heterogêneas de dados, adicionando amostras aumentadas com



**Figura 6.** Exemplos de pré-processamento com algoritmo Shades of Grey. (a) Imagens originais com diferentes cores de fundo; e (b) Imagens pré-processadas com cores de fundo mais consistentes.



diferentes transformações de imagem, como rotação, corte aleatório, inversão horizontal e vertical, translação, cisalhamento, instabilidade de cor e espaço de cores. Está comprovado em muitos estudos que o aumento de dados melhorou o diagnóstico do câncer de pele [61,62]. No conjunto de dados HAM10000, [22], as imagens das lesões cutâneas foram capturadas em diferentes ampliações ou ângulos ou com diferentes câmeras, um processo conhecido como aumento natural de dados. Notavelmente, Goyal et al. [63] usaram uma arquitetura de aprendizagem profunda chamada Faster R – CNN para desenvolver um algoritmo para gerar cópias aumentadas semelhante ao método natural de aumento de dados usado para outros conjuntos de dados de lesões de pele.

#### 4.6. Redes adversárias generativas

Redes Adversárias Generativas (GAN) são arquiteturas de aprendizagem profunda que estão atraindo interesse na comunidade de imagens médicas. O GAN é usado principalmente para gerar dados de imagens falsos de alta qualidade para superar um conjunto de dados limitado [64–66]. Para câncer de pele, o GAN pode ser usado para gerar imagens sintéticas realistas de lesões de pele para superar a falta de dados anotados [67]. A distribuição de lesões cutâneas em conjuntos de dados disponíveis publicamente é fortemente distorcida pela prevalência de cada classe entre os pacientes, e o GAN pode ser usado para gerar dados de imagem para classes de lesões cutâneas sub-representadas ou classes raras de câncer de pele, como MCC, carcinoma sebáceo ou sarcoma de kaposi.

#### 4.7. Identificando subcategorias

Pode haver muitas diferenças visuais intraclasse na aparência das lesões cutâneas em termos de textura, cor e tamanho. Na maioria dos conjuntos de dados disponíveis publicamente, a coleção de lesões cutâneas pertence a cada superclasse, em vez de dividi-las em subcategorias. Lidando com muitas diferenças intraclasse e semelhanças entre classes (imitações de uma lesão cutânea) no conjunto de dados de lesões cutâneas, é um desafio para algoritmos de aprendizagem profunda classificar ou diferenciar tais lesões. Como possível solução para lidar com este problema, as subcategorias de cada lesão cutânea devem ser tratadas como diferentes classes no conjunto de dados utilizado para treinar algoritmos de aprendizagem profunda. No entanto, isto exigirá um maior volume de imagens de treinamento e também será mais desafiador traduzir para a prática clínica. Portanto, a subcategorização exigiria um certo grau de suspeita ou uma probabilidade pré-teste razoável para ajudar adequadamente o médico na escolha do algoritmo.

#### 4.8. Explicação semântica da previsão

Para ajudar os médicos, os algoritmos de aprendizagem profunda precisam fornecer uma explicação semântica, em vez de apenas uma pontuação de confiança para a previsão de lesões cutâneas. Uma solução possível poderia ser redes de aprendizagem profunda usando conjuntos de dados longitudinais para fornecer uma explicação semântica da previsão das redes de acordo com os critérios ABCDE (assimetria, borda, cor, diâmetro, evolução) ou lista de verificação de malignidade de lesão cutânea de 7 pontos (pigmento rede, estruturas de regressão, pigmentação, estruturas vasculares, estrias, pontos e glóbulos, véu azul esbranquiçado) [33].

#### 4.9. Múltiplos modelos para diagnóstico de câncer de pele

Em vez de depender de uma única solução de IA para o diagnóstico do cancro da pele, vários modelos de aprendizagem profunda podem avaliar diferentes características ou aspectos das lesões cutâneas, submeter previsões e gerar uma conclusão final. Nesse sentido, o poder computacional e o armazenamento em nuvem estão se tornando mais acessíveis e será possível hospedar múltiplos modelos para auxiliar dermatologistas no diagnóstico de câncer de pele, em todo o mundo, em paralelo (ou em sincronia).

#### 4.10. Combinando informações clínicas e recursos de imagem

Os metadados clínicos e o histórico do paciente são considerados clinicamente importantes no diagnóstico do câncer de pele. Essas informações podem fornecer

insights além dos recursos de imagem usados por algoritmos de aprendizado profundo. Portanto, há necessidade de desenvolver algoritmos de fusão de dados que possam combinar recursos compostos por informações clínicas com recursos de imagem de modelos de aprendizagem profunda para fornecer previsões finais do diagnóstico de câncer de pele. Em estudo recente, Pacheco et al. [68] combinaram modelos de aprendizagem profunda (imagens clínicas) e informações clínicas do paciente para alcançar uma melhoria de aproximadamente 7% na previsão equilibrada precisão.

#### 4.11. Solução multimodal: pipeline completo de diagnóstico

Se dados histopatológicos correspondentes para lesões cutâneas dermatoscópicas estivessem disponíveis em conjuntos de dados, poderíamos desenvolver uma solução completa de IA semelhante ao pipeline de diagnóstico de um dermatologista. Na primeira etapa, uma solução de IA é usada para classificar lesões cutâneas dermatoscópicas, com um algoritmo de aprendizagem profunda treinado no conjunto de dados dermatoscópicos. Para casos suspeitos, o algoritmo de aprendizagem profunda pode ser desenvolvido num conjunto de dados patológicos para determinar se a lesão é cancerosa ou não.

#### 4.12. Validação clínica rigorosa

É um facto bem conhecido, tanto para médicos como para investigadores de IA, que os erros podem informar futuras tomadas de decisão. Uma vez que não podemos permitir diagnósticos errados por parte da tecnologia, é melhor manter as soluções de IA em segundo plano para uma validação rigorosa de dados ruidosos provenientes de pacientes reais e para melhorar as previsões destes sistemas tecnológicos até à data, até que sejam finalmente validados para fornecer resultados úteis. insights sobre o diagnóstico de câncer de pele e auxiliar os médicos em ambientes hospitalares e remotos.

## 5. Conclusão

A pesquisa envolvendo IA está fazendo progressos encorajadores no diagnóstico do câncer de pele. Apesar das várias alegações de que algoritmos de aprendizagem profunda superam o desempenho dos médicos no diagnóstico de câncer de pele, há muito mais desafios enfrentados por esses algoritmos para se tornarem um sistema de diagnóstico completo. Como tais experimentos são realizados em ambientes controlados, os algoritmos nunca são testados no diagnóstico real de pacientes com câncer de pele. O processo de diagnóstico no mundo real requer levar em consideração a etnia, a pele, a cor do cabelo e dos olhos do paciente, a ocupação, a doença, os medicamentos, os danos solares existentes, o número de nevos e os hábitos de vida (como exposição ao sol, tabagismo e consumo de álcool). ), história clínica, resposta a tratamentos anteriores e outras informações do prontuário médico do paciente. No entanto, os atuais modelos de aprendizagem profunda dependem predominantemente apenas dos dados de imagem dos pacientes.

Além disso, tais sistemas correm frequentemente o risco de erros de diagnóstico sempre que são aplicados a lesões ou condições cutâneas que não estão presentes no conjunto de dados de treino. Este artigo explora ainda mais oportunidades para construir algoritmos robustos para auxiliar os médicos no diagnóstico de câncer de pele. As sociedades de visão computacional e de dermatologistas precisam trabalhar juntas para melhorar as soluções atuais de IA e aumentar a precisão diagnóstica dos métodos utilizados para o diagnóstico do câncer de pele. A IA tem o potencial de proporcionar uma mudança de paradigma no diagnóstico do cancro da pele e, portanto, uma solução de cuidados de saúde económica, acessível remotamente e precisa.

#### 5.1. Estratégia de pesquisa e critérios de seleção

Usamos o Google Scholar e o PubMed para encontrar manuscritos relevantes usando as diretrizes Preferred Reporting Items for a Systematic Review and Meta- analyze of Diagnostic Test Accuracy Studies (PRISMA-DTA) para artigos envolvendo IA e câncer de pele. Restringimos nossa pesquisa a artigos publicados em inglês entre 1º de janeiro de 2012 e 11 de janeiro de 2020. Usamos os seguintes termos em diferentes combinações: “câncer de pele”, “lesões de pele”, “câncer de pele raro”, “aprendizado profundo”, “inteligência artificial”, “dermatologistas”, “imagens clínicas”, “dermoscópica”, “histopatologia”, “preconceito social”, “inteligência artificial e câncer de pele”, “câncer de pele e

aprendizagem profunda”, “dermatologistas e aprendizagem profunda”, “conjuntos de dados de câncer de pele”, “diagnóstico de lesões e aprendizagem profunda”, “informações clínicas, aprendizagem profunda e câncer de pele”, “aumento de dados e câncer de pele”, “GAN e pele Câncer”. Para serem incluídos, os estudos precisavam mencionar IA, câncer de pele, uma comparação de desempenho entre algoritmos de IA, médicos e dermatologistas em diferentes modalidades de imagens de pele ou conjuntos de dados de lesões/câncer de pele ou relatar medidas de resultados relevantes em relação à precisão do algoritmo de IA. 53 de 4.869 artigos foram incluídos após a aplicação dos critérios de elegibilidade.

Contribuições do autor

Todos os autores leram e aprovaram o manuscrito, e cada autor participou suficientemente no desenvolvimento do projeto e do manuscrito. MG contribuiu na revisão da literatura e análise do estudo e redação do manuscrito. TK e SY contribuíram com aspectos clínicos de desafios e oportunidades técnicas e na redação do manuscrito. SH supervisionou os detalhes técnicos e a análise do estudo. TK e SH revisaram o manuscrito.

Financiamento

Esta pesquisa foi apoiada em parte pelas bolsas do Instituto Nacional de Saúde R01LM012837 e R01CA249758.

Declaração de interesse concorrente

Nenhum declarado.

Referências

[1] SC Foundation, “Skin Cancer Facts and Statistics”, Online, janeiro de 2017. [Online]. Disponível: <https://www.skincancer.org/skin-cancerinformation/skin-cancer-fact-s/general>.

[2] W. Street, Fatos e números sobre câncer 2019, American Cancer Society, Atlanta, GA, EUA, 2019.

[3] HW Rogers, MA Weinstock, SR Feldman, BM Coldiron, estimativa de incidência de câncer de pele não melanoma (carcinomas de queratinócitos) na população dos EUA, JAMA Dermatol. 151 (10) (2012) 1081–1086, 2015.

[4] F. Bray, J. Ferlay, I. Soerjomataram, RL Siegel, LA Torre, A. Jemal, Estatísticas globais de câncer 2018: estimativas globocanas de incidência e mortalidade em todo o mundo para 36 tipos de câncer em 185 países, Ca - Câncer J. Clin . 68 (6) (2018) 394–424.

[5] Z. Apalla, A. Lallas, E. Sotiriou, E. Lazaridou, D. Ioannides, Tendências epidemiológicas no câncer de pele, Dermatol. Pratique. Conceito. 7 (2) (2017) 1.

[6] A. Hosny, C. Parmar, J. Quackenbush, LH Schwartz, HJ Aerts, Inteligência artificial em radiologia, Nat. Rev. 18 (8) (2018) 500–510.

[7] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, D. Liang, Acelerando a ressonância magnética via aprendizagem profunda, em: 2016 IEEE 13º Simpósio Internacional em imagens biomédicas (ISBI), IEEE, 2016, pp.

[8] MH Yap, M. Goyal, FM Osman, R. Mart’ y, E. Denton, A. Juette, R. Zwiggelaar, Reconhecimento de lesões por ultrassom mamário: abordagens de aprendizagem profunda de ponta a ponta, J. Med. Imagem. 6 (1) (2018), 011007.

[9] M. Goyal, “Inteligência artificial em dermatologia”, DermNet NZ – Tudo sobre a pele — DermNet NZ. [Online]. Disponível: <https://www.dermnetnz.org/topics/artificial-intelligence/>.

[10] T. Wurfl, FC Ghesu, V. Christlein, A. Maier, Tomografia computadorizada de aprendizagem profunda, em: Conferência Internacional sobre Computação de Imagens Médicas e Intervenção Assistida por Computador, Springer, 2016, pp.

[11] X. Liu, L. Faes, AU Kale, SK Wagner, DJ Fu, A. Bruynseels, T. Mahendiran, G. Moraes, M. Shandas, C. Kern, et al., Uma comparação do desempenho de aprendizagem profunda com profissionais de saúde na detecção de doenças por imagens médicas: uma revisão sistemática e meta-análise, The Lancet Digital Health 1 (6) (2019) e271–e297.

[12] A. Esteve, B. Kuprel, RA Novoa, J. Ko, SM Swetter, HM Blau, S. Thrun, Classificação de câncer de pele em nível de dermatologista com redes neurais profundas, Nature 542 (7639) (2017) 115–118.

[13] MH Yap, M. Goyal, F. Osman, E. Ahmad, R. Mart’ y, E. Denton, A. Juette, R. Zwiggelaar, Reconhecimento de lesões de ultrassom mamário de ponta a ponta com uma abordagem de aprendizagem profunda , em: Medical Imaging 2018: Aplicações Biomédicas em Imagens Moleculares, Estruturais e Funcionais, vol. 10578, Sociedade Internacional de Óptica e Fotônica, 2018, p. 1057819.

[14] M. Havai, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Segmentação de tumor cerebral com neural profundo redes, Med. Imagem Anal. 35 (2017) 18–31.

[15] JW Wei, LJ Tafe, YA Linnik, LJ Vaickus, N. Tomita, S. Hassanpour, Classificação de padrões histológicos em nível de patologista em lâminas de adenocarcinoma pulmonar ressecadas com redes neurais profundas, Sci. Rep. 9 (1) (2019) 3358.

[16] N. Tomita, B. Abdollahi, J. Wei, B. Ren, A. Suriawinata, S. Hassanpour, Redes neurais profundas baseadas em atenção para detecção de tecido canceroso e pré-canceroso do esôfago em lâminas histopatológicas, JAMA Netw. Aberto 2 (11) (2019) e1914645(e1914645).

[17] M. Goyal, N. Reeves, S. Rajbhandari, MH Yap, Métodos robustos para tempo real detecção e localização de úlcera no pé diabético em dispositivos móveis, IEEE J. Biomed. Informações de saúde (2018).

[18] HA Daanen, FB Ter Haar, scanners de corpo inteiro 3D revisitados, Programas 34 (4) (2013) 270–275.

[19] T. Ching, DS Himmelstein, BK Beaulieu-Jones, AA Kalinin, BT Do, GP Way, E. Ferrero, P.-M. Agapow, M. Zietz, MM Hoffman, et al., Oportunidades e obstáculos para aprendizagem profunda em biologia e medicina, JR Soc. Interface 15 (141) (2018) 20170387.

[20] Y. Bengio, I. Goodfellow, A. Courville, Aprendizagem profunda. CiteSeer 1 (2017).

[21] NC Codella, D. Gutman, ME Celebi, B. Helba, MA Marchetti, SW Dusza, A. Kallou, K. Liopyris, N. Mishra, H. Kittler, et al., Análise de lesões cutâneas para detecção de melanoma: um desafio no simpósio internacional de imagens biomédicas de 2017 (isbi), organizado pela colaboração internacional de imagens de pele (isic), em: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE, 2018, pp. .

[22] P. Tschandl, C. Rosendahl, H. Kittler, O conjunto de dados HAM10000, uma grande coleção de imagens dermatoscópicas de múltiplas fontes de lesões cutâneas pigmentadas comuns, Sci. Dados 5 (2018) 180161, <https://doi.org/10.1038/sdata.2018.161>.

[23] M. Combaila, NC Codella, V. Rotemberg, B. Helba, V. Vilaplana, O. Reiter, A. C. Halpem, S. Puig, J. Malvey, Bcn20000: lesões dermatoscópicas na natureza, 2019 pré-impressão arXiv arXiv:1908.02288.

[24] G. Argenziano, HP Soyer, V. De Giorgio, D. Piccolo, P. Carli, M. Delfino, A. Ferrari, R. Hofmann-Wellenhof, D. Massi, G. Mazzocchi et al., Atlas interativo de dermatoscopia.

[25] Uma prótese cognitiva para auxiliar no diagnóstico de lesões cutâneas focais. URL <https://páginas.iniciais.inf.ed.ac.uk/rbfi/DERMOFIT/>.

[26] T. Mendonça, PM Ferreira, JS Marques, AR Marcal, J. Rozeira, Ph 2-a banco de dados de imagens dermatoscópicas para pesquisa e benchmarking, em: Engineering in Medicine and Biology Society (EMBC), 2013 35ª Conferência Internacional Anual do IEEE, IEEE, 2013, pp.

[27] I. Giotis, N. Molders, S. Land, M. Biehl, MF Jonkman, N. Petkov, Med-node: um sistema de diagnóstico de melanoma assistido por computador usando imagens não dermatoscópicas, Expert Syst. Apl. 42 (19) (2015) 6578–6585.

[28] SS Han, MS Kim, W. Lim, GH Park, I. Park, SE Chang, Classificação do imagens clínicas para tumores cutâneos benignos e malignos usando um algoritmo de aprendizagem profunda, J. Invest. Dermatol. 138 (7) (2018) 1529–1538.

[29] SS Han, GH Park, W. Lim, MS Kim, J. Im Na, I. Park, SE Chang, Redes neurais profundas mostram um desempenho equivalente e muitas vezes superior aos dermatologistas no diagnóstico de oncomicos: construção automática de conjuntos de dados de oncomicos por rede neural profunda convolucional baseada em região, PLoS One 13 (1) (2018), e0191493.

[30] J. Yang, X. Sun, J. Liang, PL Rosin, Diagnóstico clínico de lesões cutâneas usando representações inspiradas em critérios dermatologistas, em: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp.

[31] J. Yang, X. Wu, J. Liang, X. Sun, M.-M. Cheng, PL Rosin, L. Wang, Aprendizagem de equilíbrio individualizado para reconhecimento clínico de doenças de pele, IEEE Trans. Rede Neural. Aprender. Sist. (2019).

[32] Dermnet nz.” [Online]. Disponível: <https://www.dermnetnz.org/>.

[33] J. Kawahara, S. Daneshvar, G. Argenziano, G. Hamarneh, Lista de verificação de sete pontos e classificação de lesões cutâneas usando redes neurais multimodais multitarefa, IEEE J. Biomédica. Informações de saúde 23 (2) (2018) 538–546.

[34] O programa atlas do genoma do câncer.” [Online]. Disponível: <https://www.cancer.gov/about-nci/organization/ccg/research/structuralgenomics/tcga> .

[35] NC Codella, Q.-B. Nguyen, S. Pankanti, DA Gutman, B. Helba, AC Halpem, J. R. Smith, Conjuntos de aprendizagem profunda para reconhecimento de melanoma em imagens de dermatoscopia, IBM J. Res. Dev. 61 (4/5) (2017), 5–1.

[36] HA Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kallou, ABH Hassen, L. Thomas, A. Enk, et al., Homem contra máquina: diagnóstico desempenho de uma rede neural convolucional de aprendizagem profunda para reconhecimento dermatoscópico de melanoma em comparação com 58 dermatologistas, Ann. Oncol. 29 (8) (2018) 1836–1842.

[37] TJ Brinker, A. Hekler, AH Enk, J. Klode, A. Hauschild, C. Berking, B. Schilling, S. Haferkamp, D. Schadendorf, T. Holland-Letz, et al., Aprendizado profundo superado 136 de 157 dermatologistas em uma tarefa de classificação de imagens dermatoscópicas de melanoma frente a frente , Eur. J. Canc. 113 (2019) 47–54.

[38] P. Tschandl, C. Rosendahl, BN Akay, G. Argenziano, A. Blum, RP Braun, H. Cabo, J.-Y. Gourhant, J. Kreusch, A. Lallas, et al., Diagnóstico de nível especializado de câncer de pele não pigmentado por redes neurais convolucionais combinadas, JAMA Dermatol. 155 (1) (2019) 58–65.

[39] RC Maron, M. Weichenthal, JS Utikal, A. Hekler, C. Berking, A. Hauschild, A. H. Enk, S. Haferkamp, J. Klode, D. Schadendorf, et al., Desempenho superior sistemático de 112 dermatologistas na classificação multiclasse de imagens de câncer de pele por redes neurais convolucionais, Eur. J. Canc. 119 (2019) 57–65.

[40] HA Haenssle, C. Fink, F. Toberer, J. Winkler, W. Stolz, T. Deinlein, R. Hofmann- Wellenhof, A. Lallas, S. Emmert, T. Buhl, M. Zutt, Homem contra máquina recarregada: desempenho de uma rede neural convolucional aprovada pelo mercado na classificação de um amplo espectro de lesões cutâneas em comparação com 96 dermatologistas trabalhando sob condições menos artificiais, Ann. Oncol. 31 (1) (2020) 137–143.

[41] P. Tschandl, NC Codella, BN Akay, G. Argenziano, RP Braun, H. Cabo, D. Gutman, et al., Comparação da precisão de leitores humanos versus leitores de máquina

- algoritmos de aprendizagem para classificação de lesões cutâneas pigmentadas: um estudo de diagnóstico internacional aberto, baseado na web, *Lancet Oncol.* 20 (7) (2019) 938–947.
- [42] Y. Fujisawa, Y. Otomo, Y. Ogata, Y. Nakamura, R. Fujita, Y. Ishitsuka, R. Watanabe, N. Okiyama, K. Ohara, M. Fujimoto, aprendizagem profunda baseada em computador **classificador** auxiliado desenvolvido com um pequeno conjunto de dados de imagens clínicas supera dermatologistas certificados no diagnóstico de tumores de pele, *Br. J. Dermatol.* 180 (2) (2019) 373–381.
- [43] T.J. Brinker, A. Hekler, A.H. Enk, J. Klode, A. Hauschild, C. Berking, B. Schilling, S. Haferkamp, D. Schadendorf, S. Frohling, et al., Uma rede neural convolucional treinada com imagens dermatoscópicas realizadas em paridade com 145 dermatologistas em uma tarefa de classificação de imagens clínicas de melanoma, *Eur. J. Canc.* 111 (2019) 148–154.
- [44] A. Hekler, J.S. Utikal, A.H. Enk, C. Berking, J. Klode, D. Schadendorf, P. Jansen, C. Franklin, T. Holland-Letz, D. Krah, et al., Classificação de nível de patologista de imagens histopatológicas de melanoma com redes neurais profundas, *Eur. J. Canc.* 115 (2019) 79–83.
- [45] Y. Jiang, J. Xiong, H. Li, X. Yang, W. Yu, M. Gao, X. Zhao, Y. Ma, W. Zhang, Y. Guan, et al., Reconhecendo o carcinoma basocelular em imagens histopatológicas digitais capturadas por smartphone com rede neural profunda, *Br. J. Dermatol.* (2019).
- [46] A.A. Cruz-Roa, J.E.A. Ovalle, A. Madabhushi, F.A.G. Osorio, Um aprendizado profundo arquitetura para representação de imagens, interpretabilidade visual e detecção automatizada de câncer de carcinoma basocelular, em: Conferência Internacional sobre Computação de Imagens Médicas e Intervenção Assistida por Computador, Springer, 2013, pp.
- [47] P. Xie, K. Zuo, Y. Zhang, F. Li, M. Yin, K. Lu, Interpretable Classification from Skin Cancer Histology Slides using Deep Learning: a Retrospective Multicenter Study, 2019 arXiv pré-impressão arXiv:1904.06156.
- [48] A. Krizhevsky, I. Sutskever, G.E. Hinton, classificação Imagenet com profundidade redes neurais convolucionais, em: Advances in Neural Information Processing Systems, 2012, pp.
- [49] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, Microsoft coco: objetos comuns no contexto, em: Conferência Europeia sobre Visão Computacional, Springer, 2014, pp.
- [50] H.R. Tizhoosh, L. Pantanowitz, Inteligência artificial e patologia digital: desafios e oportunidades, *J. Pathol. Inf.* 9 (2018).
- [51] D.S. Rigel, R.J. Friedman, A.W. Kopf, D. Polsky, Abcde - um conceito em evolução na detecção precoce de melanoma, *Arch. Dermatol.* 141 (8) (2005) 1032–1034.
- [52] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, M. Delno, Microscopia de epiluminescência para o diagnóstico de lesões cutâneas melanocíticas duvidosas: comparação da regra abcd da dermatoscopia e uma nova lista de verificação de 7 pontos baseada na análise de padrões, *Arch. Dermatol.* 134 (12) (1998) 1563–1570.
- [53] L.M. Stanoszek, G.Y. Wang, P.W. Harms, imitadores histológicos do carcinoma basocelular, *Arquivos de patologia e medicina laboratorial* 141 (11) (2017) 1490–1502.
- [54] J. Weingast, C. Scheibb, "EM Wurm, E. Ranharter, S. Porkert, S. Dreiseitl, C. Posch, M. Binder, Um estudo prospectivo de telefones celulares para dermatologia em um ambiente clínico, *J. Telemed. Teleassistência* 19 (4) (2013) 213–218.
- [55] K. Hogan, J. Cullan, V. Patel, A. Rajpara, D. Aires, Overcalling a teledermatologia selfie: uma nova reviravolta em um campo em crescimento, *Dermatol. On-line J.* 21 (6) (2015).
- [56] H.M. Gloster Jr., K. Neal, Câncer de pele em peles de cor, *J. Am. Acad. Dermatol.* 55 (5) (2006) 741–760.
- [57] S. Hu, R.M. Soza-Vento, D.F. Parker, R.S. Kirsner, Comparação do estágio no diagnóstico de melanoma entre pacientes hispânicos, negros e brancos no condado de miami-dade, Flórida, *Arch. Dermatol.* 142 (6) (2006) 704–708.
- [58] G. Marcus, E. Davis, Reinicialização de IA, Construindo Inteligência Artificial em que Podemos Confiar, 2019. Pantheon.
- [59] C. Barata, M.E. Celebi, J.S. Marques, Melhorando a classificação de imagens dermatoscópicas usando constância de cores, *IEEE J. Biomed. Informações de saúde* 19 (3) (2014) 1146–1152.
- [60] J.H. Ng, M. Goyal, B. Hewitt, M.H. Yap, O efeito dos algoritmos de constância de cores na segmentação semântica de lesões cutâneas, em: Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging, vol. 10953, Sociedade Internacional de Óptica e Fotônica, 2019, p. 109530R.
- [61] F. Perez, C. Vasconcelos, S. Avila, E. Valle, Aumento de dados para lesão cutânea análise, em: salas de operação com reconhecimento de contexto OR 2.0, endoscopia robótica assistida por computador, procedimentos baseados em imagens clínicas e análise de imagens da pele, Springer, 2018, pp.
- [62] T.-C. Pham, C.-M. Luong, M. Visani, V.-D. Hoang, Deep cnn e aumento de dados para classificação de lesões cutâneas, em: Conferência Asiática sobre Sistemas Inteligentes de Informação e Banco de Dados, Springer, 2018, pp.
- [63] M. Goyal, S. Hassanpour, M.H. Yap, Detecção de região de interesse em imagens dermatoscópicas para aumento de dados naturais, pré-impressão arXiv de 2018 arXiv:1807.10711.
- [64] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Redes adversárias gerativas, em: Avanços no processamento de informações **neurais** Sistemas, 2014, pp.
- [65] X. Yi, E. Walia, P. Babyn, Rede adversária generativa em imagens médicas: uma revisão, em: Análise de Imagens Médicas, 2019, p. 101552.
- [66] J. Wei, A. Suriawinata, L. Vaickus, B. Ren, X. Liu, J. Wei, S. Hassanpour, Tradução generativa de imagens para aumento de dados em imagens de histopatologia colorretal, em: Proceedings of Machine Learning for Health Workshop no NeurIPS, 2019.
- [67] A. Bissoto, F. Perez, E. Valle, S. Avila, Síntese de lesão cutânea com gerador redes adversárias, em: salas de operação com reconhecimento de contexto OR 2.0, endoscopia robótica assistida por computador, procedimentos baseados em imagens clínicas e análise de imagens da pele, Springer, 2018, pp.
- [68] A.G. Pacheco e R.A. Krohling, "O impacto das informações clínicas do paciente sobre detecção automatizada de câncer de pele", *Comput. Biol. Med.*