

COMPARATIVO DE MODELOS DE REDES NEURAIAS CONVOLUCIONAIS NO DIAGNÓSTICO E CLASSIFICAÇÃO DE LESÕES DE CÂNCER DE PELE ATRAVÉS DE IMAGENS DERMATOSCÓPICAS

Wagner M. Schmitz¹, Marcelo Teixeira¹, Dalcimar Casanova¹, Jefferson Tales Oliva¹
e Marco A. Barbosa²

¹Programa de Pós Graduação em Engenharia Elétrica e de Computação

²Programa de Pós Graduação em Engenharia de Produção e Sistemas

Universidade Tecnológica Federal do Paraná (UTFPR), Campus Pato Branco, PR, Brasil

RESUMO

O uso das redes neurais convolucionais (RNC) é um dos métodos mais difundidos para a classificação de imagens médicas. No entanto, ainda não está claro qual variação dessa abordagem é mais eficaz para a análise de imagens dermatoscópicas. Por isso, neste trabalho, foram comparados diferentes tipos de RNC pré-treinadas, como ResNet50, MobileNet, VGG16, VGG19 e InceptionV3, visando compreender qual modelo se aplica melhor à análise de imagens dermatoscópicas, visando evitar erros no diagnóstico de lesões de pele. Para isso, foi utilizada a base de dados HAM10000, onde as imagens foram diferenciadas entre as classes Nevo melanocítico, melanoma, lesões benignas semelhantes à ceratose, carcinoma basocelular, queratoses actínicas, lesões vasculares, e dermatofibroma. Como resultado, os modelos resultantes atingiram acurácias entre 0,48 e 0,82. Após a aplicação de testes estatísticos de hipótese, constatou-se diferença estatisticamente significativa entre os modelos. Desse modo, é importante ressaltar que nenhum modelo atingiu o melhor desempenho para todas as classes, evidenciando que cada arquitetura tem suas próprias vantagens e limitações em contextos específicos relacionados a imagens dermatoscópicas.

PALAVRAS-CHAVES

Analytics and Data Science, Intelligent Data Analysis, Machine Learning, Artificial Neural Networks and Deep Learning

1. INTRODUÇÃO

O câncer de pele corresponde a 33% de todos os diagnósticos de câncer no Brasil, sendo registrado, pelo Instituto Nacional do Câncer (INCA), mais de 185 mil novos casos a cada ano (Sociedade Brasileira de Dermatologia 2024). Em estágios mais avançados, o melanoma, um tipo agressivo de câncer de pele, desenvolve lesões mais profundas e espessas, o que eleva o risco de metástase para outros órgãos e, conseqüentemente, reduz as chances de cura (SBD 2023).

A dermatoscopia é uma técnica aplicada por médicos para examinar a pele de modo detalhado e ampliado. Esse procedimento permite a visualização de características não perceptíveis a olho nu, como estruturas pigmentadas, vasos sanguíneos e padrões específicos (A.C 2023). Ela desempenha um papel fundamental na detecção precoce e no diagnóstico de diferentes condições, principalmente no contexto do câncer de pele.

A análise de exames dermatoscópicos é realizada por especialistas que avaliam padrões de pigmentação, tamanho e formato das lesões e manchas na pele. Esse processo pode ser caro e sujeito a erros devido à pouca experiência do examinador e ao seu treinamento limitado em critérios dermatoscópicos, cuja aplicação correta é essencial (Rezze 2006).

O exame dermatoscópico envolve a identificação de mais de cem estruturas diferentes (Frange et al. 2012). Assim, o uso de aprendizado de máquina e redes neurais convolucionais pode ser útil para auxiliar o médico especialista na análise rápida dessas informações, sem desconsiderar a importância do profissional na validação desse processo.

Nesse sentido, métodos de aprendizado de máquina têm sido amplamente aplicados em diversos estudos relacionados à área da saúde, como a classificação de imagens médicas (Nishi et al. 2015; Dougherty et al. 2017; Villiger et al. 2023). Abordagens de aprendizado profundo (deep learning), como as baseadas em redes neurais convolucionais (CNN – convolutional neural networks), podem auxiliar em tarefas como a detecção de enfermidades, auxiliando em diagnósticos médicos. A utilização das redes neurais convolucionais permite a identificação de características presentes nos dados e facilita a análise de imagens (Rguibi et al. 2021).

O processo de aprendizado e classificação do deep learning envolve camadas de neurônios artificiais que imitam o funcionamento dos neurônios biológicos. Essas camadas incluem a camada de entrada, que recebe os dados brutos; camadas ocultas, que processam os dados; e a camada de saída, que apresenta o resultado final. Durante o processo de treinamento, a rede ajusta os pesos dos neurônios visando minimizar o erro. Após o treinamento, a rede pode fazer a classificação dos dados, formando assim redes neurais que processam uma vasta gama de dados, incluindo documentos, imagens, áudios e vídeos (Goodfellow et al. 2016). Modelos baseados em CNN, por exemplo, são amplamente utilizados na avaliação de lesões, manchas e sinais na pele em imagens dermatoscópicas, podendo aumentar a precisão no diagnóstico de suspeitas de câncer de pele, como o melanoma (Goodfellow et al. 2016; Casagrande 2023).

Diversos modelos de arquitetura de CNN podem ser aplicados para auxiliar na análise de imagens de exames dermatoscópicos, sendo que uma tarefa desafiadora para os profissionais da área é decidir sobre qual modelo usar para cada caso em análise. Assim, este estudo tem o propósito de verificar qual desses modelos é mais acurado para a análise de imagens dermatoscópicas, minimizando erros no diagnóstico de lesões de pele.

Dessa forma, a adoção de um modelo robusto e bem estabelecido pode ser benéfico ao gerar resultados positivos na identificação de padrões em imagens, otimizando o trabalho dos profissionais e reduzindo a dependência de classificação manual que ocorre na maioria dos casos em exames dermatoscópicos, possibilitando o auxílio no diagnóstico de forma simples e acessível. A comparação de diferentes arquiteturas contribui para a análise e implementação do modelo, trazendo maior efetividade na compreensão das diferenças de desempenho entre as arquiteturas

2. REVISÃO DA LITERATURA

Diversos modelos de redes neurais têm sido utilizados para detecção de câncer de pele. Geralmente, esse estudo é feito utilizando algum modelo de CNN de forma isolada ou combinada. Esta sessão apresenta alguns trabalhos dessa década encontrados na literatura.

No trabalho (Gururaj et al. 2023), os autores utilizam técnicas de pré-processamento de dados e aprendizado por transferência com DenseNet169 e ResNet50 para melhorar a precisão da classificação. O estudo destaca a eficácia das CNNs em superar médicos especialistas na detecção precoce de câncer de pele. Como resultado, a técnica de subamostragem do DenseNet169 alcançou uma acurácia de 91,2%, enquanto a técnica de superamostragem do ResNet50 obteve uma acurácia de 83%.

A aplicação de várias técnicas de aprendizado por transferência pré-treinadas foi realizada no artigo (Islam and Panta 2024) para a classificação binária de câncer de pele, distinguindo entre estágios benignos e malignos. Os resultados experimentais mostraram que o modelo ResNet50 se destacou, alcançando uma acurácia de 93,5%.

Em (Alwakid et al. 2023), os modelos Inception-V3 e InceptionResNet-V2 foram utilizados como classificadores para detecção de melanomas em imagens dermatoscópicas com validação feita através de técnicas de data augmentation. Os resultados destacaram a eficácia do deep learning na identificação precoce de câncer de pele. Como resultado, o maior valor de acurácia alcançado foi de 91,26%.

No artigo de (Faghihi et al. 2024), o autor utiliza os modelos VGG16 e VGG19 para a classificação do câncer de pele. No estudo foi utilizada a otimização Adam e a validação cruzada para aprimorar os modelos, que alcançaram alta precisão na detecção de lesões de pele. Como resultado, acurácias entre 92,5% e 98,1% foram atingidas.

Os autores do artigo (Ahmed et al. 2023) empregaram técnicas de transferência com MobileNet e ResNet101 para buscar maior precisão na classificação de lesões de pele. Para isso, foram aplicadas técnicas híbridas que combinam características manuais e extraídas por CNN. O modelo alcançou um resultado de 98,4% de acurácia.

Além dos trabalhos anteriormente citados, diversos outros modelos de redes neurais convolucionais e técnicas já foram utilizados para a identificação de câncer de pele em imagens dermatoscópicas. Apesar da grande variedade de modelos CNN utilizados na detecção de câncer de pele utilizando imagens dermatoscópicas, grande parte dos trabalhos é direcionada para a classificação das imagens através de um só modelo. Neste trabalho, foram utilizados diversos modelos de CNN para a classificação de imagens da base de dados HAM10000, visando comparar os resultados apresentados dentre eles. Acredita-se que essa comparação possa aprimorar o entendimento sobre a eficácia das diferentes arquiteturas de CNN na detecção de câncer de pele e identificar quais modelos oferecem melhores resultados para a análise de imagens dermatoscópicas.

3. MATERIAIS E MÉTODOS

3.1 Base de Dados

Neste estudo, foi utilizada a base de dados HAM10000 (Human Against Machine with 10000 training images), desenvolvida pela Universidade de Viena e disponibilizada por Harvard (Tschandl et al. 2018). A base contém 10.015 imagens dermatoscópicas e inclui uma planilha com 8 colunas que identificam as características das imagens, tais como: identificador da lesão (lesion id); identificador da imagem (image id); classificação das lesões (dx); forma usada para a classificação das lesões (dx type); idade; sexo; e localização da lesão.

A categorização das lesões, ou coluna dx, conta com sete classes identificadas, tais como: Nevo melanocítico (nv); melanoma (mel); lesões benignas semelhantes à ceratose (bkl); carcinoma basocelular (bcc); queratoses actínicas (akiec); lesões vasculares (vas); e dermatofibroma (df).

Dentre as classes, mel e bcc são consideradas tumores malignos, que exigem maior atenção e tratamento médico, sendo essas o foco principal deste trabalho.

3.2 Pré-Processamento

Técnicas de análise de dados são fundamentais para o desenvolvimento de projetos, envolvendo o pré-processamento e a análise de informações. O pré-processamento inclui a limpeza de dados, que corrige informações incorretas aplicando, por exemplo, a mediana dos dados corretos para preencher os campos nulos; normalização de 0 a 1 para ajustar a escala dos dados numéricos; balanceamento, feito por sobreamostragem, onde cada classe foi ajustada para ter 500 amostras, sem a operação de aumento de dados; codificação, que transforma dados qualitativos em quantitativos; e divisão dos dados em conjuntos de treinamento, validação e teste, utilizando o método de holdout, com 25% dos dados reservados para o teste e 75% para o treinamento (de Almeida Prado Alves Batista 2003).

Ferramentas como Python, scikit-learn e TensorFlow são utilizadas para essas tarefas, começando com a leitura e padronização das imagens, seguida pela divisão dos dados de forma equilibrada para garantir um bom reconhecimento de padrões.

3.3 Aplicação dos Modelos Pré-treinados de Redes Neurais Convolucionais

Com a base de treino e teste devidamente separadas, foram aplicados os modelos de redes neurais convolucionais, que são modelos computacionais inspirados no sistema nervoso biológico, compostos por camadas de neurônios artificiais interconectados em um padrão feed-forward. As CNNs imitam a extração de características visuais do córtex, onde neurônios respondem a estímulos específicos e se sobrepõem para cobrir toda a área visual.

Por exemplo, na Figura 1, uma CNN recebe uma imagem de um carro e a propaga através da rede, passando por camadas convolucionais que extraem características visuais importantes. Em seguida, operações de *pooling* reduzem a dimensionalidade, e as camadas totalmente conectadas processam essas características para tomar a decisão final de classificação. O modelo então calcula o erro entre a saída prevista e a saída desejada, ajustando os pesos por retropropagação até atingir o desempenho desejado e aprendendo os padrões do conjunto de treinamento (Goodfellow et al. 2016).

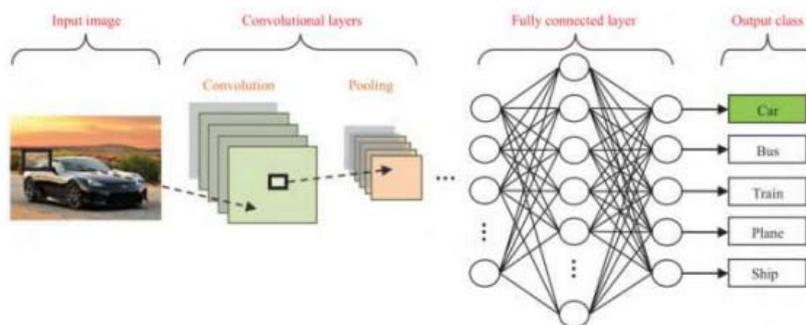


Figura 1. Operação Convolutacional
Fonte: (Rawat and Wang 2017)

Os modelos de redes neurais convolucionais utilizados foram o ResNet50, MobileNet, VGG16, VGG19 e InceptionV3. O procedimento começou com a preparação dos dados, que envolve o pré-processamento do conjunto de imagens de entrada. Em seguida, cada modelo foi treinado e avaliado individualmente quanto à sua precisão e desempenho.

O treinamento dos classificadores começou com a definição do modelo base, que é a seleção da arquitetura de rede neural que serve de ponto de partida para o desenvolvimento do modelo. Em seguida, foram adicionadas camadas à arquitetura CNN para extrair características mais avançadas das imagens. Especificamente, as camadas convolucionais foram incorporadas com base em sua capacidade de aplicar filtros sobre a imagem para gerar uma série de mapas de características, que servem para identificar e aprender os padrões espaciais das imagens (Souza et al. 2020), essencial para o processamento de imagens dermatoscópicas.

Além disso, foram implementadas camadas de dropout com uma taxa de 30% em duas fases do modelo para reduzir a complexidade e prevenir o overfitting. Especificamente, o dropout foi aplicado antes da camada de batch normalization e da camada densa com 256 neurônios. O uso dessas camadas auxilia no descarte aleatório de neurônios, ajudando a melhorar a generalização do modelo (Bademian e França 2020).

Além de incluir quatro camadas totalmente conectadas adicionais após o modelo base, o classificador é composto por uma camada densa com 256 neurônios e duas camadas de batch normalization, seguidas de uma camada de saída com 7 neurônios e a função de ativação softmax. A escolha de quatro camadas conectadas foi motivada pela necessidade de balancear a capacidade do modelo e a complexidade computacional.

Embora uma busca extensiva de hiperparâmetros possa ser realizada para otimizar esses valores, optou-se por manter a configuração padrão para reduzir o custo computacional, baseado na evidência de que parâmetros default podem oferecer resultados razoáveis (Amancio et al. 2014).

Em seguida, as camadas totalmente conectadas combinam as características extraídas em uma representação que pode ser usada pelo modelo para a classificação das imagens. Por fim, para o treinamento de modelos, foram usadas as seguintes redes CNN pré-treinadas:

- ResNet50: esse modelo permite o treinamento de redes neurais com mais de 150 camadas. A ResNet50 foi projetada para lidar com o problema de degradação do desempenho à medida que a rede se torna mais profunda, através de conexões de atalho para facilitar o fluxo de informação (Kaiming He 2015). Os mesmos hiperparâmetros e processos foram utilizados como base para o desenvolvimento dos modelos e análise dos respectivos resultados de treinamento.
- MobileNet: a arquitetura MobileNet foi especialmente concebida para eficiência em dispositivos móveis, incorporando convoluções separáveis e a capacidade de ajuste na largura. Essas características são estratégicas para garantir um equilíbrio adequado entre eficiência e desempenho

em contextos com recursos restritos, como em smartphones e sistemas embarcados (Howard et al. 2017). Os mesmos hiperparâmetros e processos utilizados na ResNet50 foram utilizados como base para o desenvolvimento de todos os modelos de treinamento.

- VGG16: a arquitetura VGG (Visual Geometry Group) é uma família de CNN desenvolvida pelo Visual Geometry Group, da Universidade de Oxford (Simonyan e Zisserman 2014). A VGG16 é composta por 16 camadas, incluindo camadas convolucionais e totalmente conectadas. Cada bloco convolucional consiste em camadas convolucionais 3x3 com a função de ativação ReLU, que introduz não linearidade e melhora a eficiência computacional em redes neurais, seguidas por camadas de pooling máximas (Mascarenhas e Agarwal 2021).
- VGG19: a VGG19 é uma extensão da VGG16, com 19 camadas. A principal diferença é a presença de camadas convolucionais adicionais, proporcionando uma maior profundidade à rede. Essas camadas extras ajudam a aprender características mais complexas (Mascarenhas e Agarwal 2021).
- InceptionV3: o InceptionV3 é uma arquitetura de rede neural convolucional que foi treinada no conjunto de dados ImageNet, um banco de dados de imagens anotadas a partir de uma grande escala de categorias de objetos (Russakovsky et al. 2015). Ela alcançou um resultado com precisão superior a 78,1% nesse conjunto de dados. O objetivo principal do InceptionV3 é aprimorar o desempenho e a eficiência computacional em relação às iterações anteriores da arquitetura Inception (Developers 2023).

Durante o treinamento, os modelos de rede neural convolucional foram configurados com uma camada convolucional composta por 256 filtros. Essa camada foi aplicada a imagens tridimensionais com dimensões de 112x112 pixels e 3 canais de cor (RGB). Além disso, para prevenir o sobreajuste, os modelos incluíram uma camada de dropout com uma taxa de 0,3, o que significa que 30% dos neurônios da camada anterior foram desativados aleatoriamente durante o treinamento.

O treinamento e a validação de cada modelo foram realizados ao longo de sessenta épocas para otimizar a eficiência do processo. Esse período foi projetado para garantir que os modelos tivessem tempo suficiente para aprender de forma eficaz, minimizando o risco de sobreajuste. Durante o treinamento, a acurácia e a perda de validação foram avaliadas a cada época.

3.4 Avaliação dos Modelos

Para a avaliação do desempenho dos modelos, foram consideradas as seguintes medidas: acurácia e taxa de acerto por classe (também denominado sensibilidade). Em seguida, foi aplicado um teste estatístico de hipótese não paramétrico, como o de Friedman, para verificar se existem diferenças estatisticamente significativas nos resultados de experimentos repetidos (Siegel 2006). Caso seja encontrada diferença estatisticamente significativa por meio do teste de Friedman, o pós-teste de Nemenyi deverá ser conduzido para verificar quais modelos são significativamente diferentes dos outros.

4. RESULTADOS

Na Tabela 1 são apresentados os principais resultados deste estudo.

Tabela 1. Sensibilidade por classe e Acurácia

Classe	ResNet50	MobileNet	VGG16	VGG19	InceptionV3
Acurácia	0,73	0,82	0,48	0,76	0,78
Actinic keratoses (akiec)	0,78	0,83	0,26	0,73	0,81
Basal cell carcinoma (bcc)	0,59	0,86	0,06	0,74	0,69
Benign keratosis-like lesions (bkl)	0,56	0,72	0,29	0,74	0,68
Dermatofibroma (df)	0,90	0,98	0,83	0,92	0,98
Melanoma (mel)	0,54	0,63	0,23	0,59	0,79
Melanocytic nevi (nv)	0,73	0,72	0,75	0,57	0,49
Vascular lesions (vas)	0,94	0,98	1,00	0,99	1,00

Conforme apresentado na Tabela 1, o modelo MobileNet obteve a maior acurácia geral de validação de 82%. Além disso, MobileNet alcançou a maior sensibilidade para a maioria das classes: Actinic keratoses (akiec) com 83%, Basal cell carcinoma (bcc) com 86%, Dermatofibroma (df) com 98%, Melanoma (mel) com 63%, e Melanocytic nevi (nv) com 72%. Para as classes em que o MobileNet não obteve a maior sensibilidade, a diferença em relação ao modelo com a maior sensibilidade foi de apenas 0,02. As classes onde MobileNet não teve a maior sensibilidade foram Benign keratosis-like lesions (bkl), na qual o modelo VGG19 obteve melhores resultados, e Vascular lesions (vas), na qual VGG16 e InceptionV3 obtiveram melhores resultados. Essa diferença pode estar relacionada à complexidade das características das imagens dessas classes, o que pode ter tornado o MobileNet menos eficiente para elas.

Adicionalmente, o modelo VGG16 alcançou o menor índice de acurácia, sendo esses 48%. Mesmo assim, alcançando uma taxa de acerto de 100% dos exemplos da classe Vascular lesions (vas). Por outro lado, esse modelo atingiu o pior desempenho na classificação de instâncias do tipo Actinic keratoses (akiec) (26%), Basal cell carcinoma (bcc) (06%), Benign keratosis-like lesions (bkl) (29%) e Melanoma (mel) (23%). Em outras palavras, dentre essas classes, o VGG16 foi o modelo que atingiu o pior desempenho para a detecção de tumores malignos (bcc e mel). A complexidade do VGG16 pode reduzir sua eficiência na captura de padrões gerais, resultando em um desempenho inferior nas outras classes. O ótimo desempenho de todos os modelos na classe Vascular lesions (vas) pode estar relacionado ao número relativamente baixo de imagens dessa classe, que é apenas 142. Após a reamostragem para 500 imagens, os modelos conseguiram identificar as características da classe “vas” de forma bastante precisa.

A análise dos dados mostra que os modelos ResNet50, MobileNet, VGG19 e InceptionV3 alcançaram uma acurácia de treino final próxima de 1, enquanto o VGG16 teve uma acurácia de 0,7 com o mesmo número de épocas. Além disso, os modelos ResNet50, MobileNet e InceptionV3 exibiram um crescimento mais acentuado no início, seguido pela estabilização. Já VGG16 e VGG19 mostraram um crescimento incremental, com estabilização no VGG19. Para alcançar um melhor desempenho no modelo VGG16, uma possibilidade seria aplicar mais épocas de treinamento, dando mais tempo para ele se estabilizar.

Por fim, o teste Friedman, considerando um nível de significância de 95%, foi utilizado para verificar se existem diferenças estatisticamente significativas entre os experimentos repetidos que foram realizados (Siegel 2006). Como resultado, foi obtido um p-valor $< 0,0001$, evidenciando uma diferença estatística extremamente significativa entre os modelos. Desse modo, para identificar as comparações entre os modelos onde foram registradas diferenças estatisticamente significativas, foi aplicado o pós-teste de Nemenyi, cujos resultados são apresentados na Tabela 2.

Tabela 2. Teste de Nemenyi

	InceptionV3	MobileNet	ResNet50	VGG16	VGG19
InceptionV3	1,00000	0,00001	0,00927	0,00000	0,31546
MobileNet	0,00001	1,00000	<0,00001	<0,00001	<0,00001
ResNet50	0,00927	<0,00001	1,00000	0,11777	0,69471
VGG16	0,00000	<0,00001	0,11777	1,00000	0,00142
VGG19	0,31546	<0,00001	0,69471	0,00142	1,00000

A partir da análise dos resultados do teste de Nemenyi, conclui-se que os modelos MobileNet e InceptionV3 apresentaram diferenças estatisticamente significativas em relação a todos os outros modelos testados, exceto ao modelo VGG19. Em contrapartida, o modelo ResNet50 não mostrou diferenças significativas em relação aos modelos VGG16 e VGG19, mas teve diferenças significativas em relação aos modelos MobileNet, InceptionV3 e VGG19. O modelo VGG16 também apresentou diferenças significativas em relação ao modelo VGG19.

5. CONCLUSÃO

A aplicação de diferentes arquiteturas de redes neurais convolucionais na base de dados HAM10000 levou à constatação de que a seleção dos modelos pode levar a variações significativas nos resultados, exemplificado pelos modelos VGG16 e VGG19. Devido às limitações do dispositivo utilizado para o treinamento dos modelos, foram necessárias algumas modificações, como a diminuição do tamanho das imagens e a restrição

do número de épocas para o treinamento e teste. Com um poder computacional mais elevado, seria possível comparar um maior número de classificadores.

O uso de técnicas de aumento de dados representa uma perspectiva promissora para aprimorar o projeto em trabalhos futuros, possibilitando um estudo mais detalhado que poderia levar à melhoria do desempenho dos modelos.

A comparação entre os modelos revelou um caminho potencial para pesquisa adicional, visando encontrar o modelo de CNN mais eficaz para a detecção de melanoma em imagens dermatoscópicas.

Dessa forma, os resultados do trabalho destacam não só as diferenças observadas entre as arquiteturas de CNN testadas, mas também abrem novas possibilidades para investigações futuras. Essas descobertas podem contribuir significativamente para o avanço do campo, estimulando o desenvolvimento de modelos mais eficazes na análise de imagens dermatoscópicas e o aprimoramento das técnicas de detecção de câncer de pele.

REFERÊNCIAS

- A.C. Camargo Cancer Center, (2023). *Digital health*. São Paulo. Available at: <https://accamargo.org.br/sobre-o-cancer/medicina-diagnostica/dermatoscopia> (Accessed: 17 September 2024).
- Ahmed, I.A., Senan, E.M., Shatnawi, H.S.A., Alkhraisha, Z.M., and Al-Azzam, M.M.A., (2023). Multi-Models of Analyzing Dermoscopy Images for Early Detection of Multi-Class Skin Lesions Based on Fused Features. *Processes*, 11(3), Article 910. Available at: <https://www.mdpi.com/2227-9717/11/3/910> (Accessed 8 Oct. 2024).
- Alwakid, G., Gouda, W., Humayun, M. and Jhanjhi, N.Z., (2023). Diagnosing melanomas in dermoscopy images using deep learning. *Diagnostics*, 13(10), article number 1815. Available at: <https://www.mdpi.com/2075-4418/13/10/1815> (Accessed: 08 October 2024).
- Amancio, D.R., Comin, C.H., Casanova, D., Travieso, G., Bruno, O.M., Rodrigues, F.A., et al., (2014). A Systematic Comparison of Supervised Classifiers. *PLoS ONE*, 9(4), Article e94137. Available at: <https://doi.org/10.1371/journal.pone.0094137> (Accessed 8 Oct. 2024).
- Bademian, G.B.B. and França, C.A.P., (2020). Classificação de Placas de Trânsito com Redes Neurais para Automação de Veículos. *Departamento de Engenharia Elétrica da Universidade Federal de São Carlos*, São Carlos, SP, Brasil. Monografia.
- Batista, G.E.A.P.A., (2003). Pré-processamento de dados em aprendizado de máquina supervisionado. Tese (Doutorado em Ciências de Computação e Matemática Computacional). Instituto de Ciências Matemáticas e de Computação, University of São Paulo, São Carlos. Available at: <https://doi.org/10.11606/T.55.2003.tde-06102003-160219> (Accessed: 17 July 2024).
- Campos, G.M., (2007). Estatística Prática para Docentes e Pós-Graduandos. Editora XYZ.
- Casagrande, J., (2023). *Dermatoscopia, dermatoscopia digital e microscopia confocal, os exames que ajudam a diagnosticar o câncer de pele*. A.C.Camargo Cancer Center, São Paulo. Available at: <https://accamargo.org.br/sobre-o-cancer/noticias/dermatoscopia-dermatoscopia-digital-e-microscopia-confocal-os-exames-que> (Accessed: 24 September 2024).
- Costa, M.G.F., (2017). Redes neurais convolucionais na saúde. *Journal of Health Informatics*, 9, pp. i-ii.
- Dougherty, M., Yuan, Z., Chen, T.-M., Wakefield, D., and Iott, B., (2017). Machine learning for social services: A study of prenatal case management in Illinois. *American Journal of Preventive Medicine*, 53(5), pp. 686-694. Elsevier.
- Faghihi, A., Fathollahi, M., and Rajabi, R., (2024). Diagnosis of Skin Cancer Using VGG16 and VGG19 Based Transfer Learning Models. *arXiv*. Available at: <https://arxiv.org/abs/2404.01160> (Accessed 8 Oct. 2024).
- Frange, V.M.N., Arruda, L.H.F., and Daldon, P.É.C., (2012). Dermatoscopia: importância para a prática clínica. *Revista de Ciências Médicas*, 18(4). Available at: <https://periodicos.puc-campinas.edu.br/cienciasmedicas/article/view/637> (Accessed: 08 October 2024).
- Glorot, X., Bordes, A., and Bengio, Y., (2011). Deep sparse rectifier neural networks. arXiv preprint arXiv:1106.1813.
- Goodfellow, I., Bengio, Y., and Courville, A., (2016). *Deep Learning*. MIT Press, Rio de Janeiro. Available at: <http://www.deeplearningbook.org>.
- Google Developers, (2023). Guia avançado do Inception v3. Available at: <https://cloud.google.com/tpu/docs/inception-v3-advanced?hl=pt-br> (Accessed: 08 October 2024).
- Gururaj, H.L., Manju, N., Nagarjun, A., Aradhya, V.N.M. and Flammini, F., (2023). DeepSkin: A deep learning approach for skin cancer classification. *IEEE Access*, 11, pp. 50205-50214. DOI: 10.1109/ACCESS.2023.3274848.
- He, K., Zhang, X., Ren, S., and Sun, J., (2015). Deep Residual Learning for Image Recognition. Available at: <https://arxiv.org/abs/1512.03385> (Accessed: 08 October 2024).

- Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H., (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *CoRR*, volume abs/1704.04861. Available at: <http://arxiv.org/abs/1704.04861> (Accessed: 08 October 2024).
- Islam, M.S. and Panta, S., (2024). Skin Cancer Images Classification using Transfer Learning Techniques. *arXiv*. Available at: <https://arxiv.org/abs/2406.12954> (Accessed 8 Oct. 2024).
- Mascarenhas, S. and Agarwal, M., (2021). A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for image classification. In: *2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON)*, volume 1, pp. 96-99. DOI: 10.1109/CENTCON52345.2021.9687944.
- Nishi, D., Matsuoka, Y., Kim, Y.-I., Yonetani, H., Hu, C.-K., and Koenen, K.C., (2015). How well can post-traumatic stress disorder be predicted from pre-trauma risk factors? An exploratory study in the WHO World Mental Health Surveys. *World Psychiatry*, 14(3), pp. 306-313. Wiley.
- Rawat, W. and Wang, Z., (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 29(9), pp. 2352–2449. MIT Press.
- Rezze, S. and Neves, (2006). Dermatoscopia: o método de análise de padrões. *An Bras Dermatol*, 3, pp. 261-268.
- Rguibi, Z., Hajami, A., and Dya, Z., (2021). Explaining deep neural networks in medical imaging context. In *Proceedings of the 2021 IEEE/ACS 18th International Conference on Computer Systems and Applications (AICCSA)*. pp. 1-2.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satkin, L., Ma, S., and Fei-Fei, L., (2015). ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), pp. 211--252. DOI: 10.1007/s11263-015-0816-y.
- Siegel, S. and Castellan Jr., N.J., (2006). *Estatística não-Paramétrica para Ciências do Comportamento*. Grupo A.
- Simonyan, K. and Zisserman, A., (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. Available at: http://www.robots.ox.ac.uk/~vgg/research/very_deep/ (Accessed: 08 October 2024).
- Sociedade Brasileira de Dermatologia, (2023). *Câncer da pele*. Available at: <https://www.sbd.org.br/doencas/cancer-da-pele/> (Accessed: 17 September 2024).
- Sociedade Brasileira de Dermatologia, (2024). *Câncer de Pele*. Available at: <https://www.sbd.org.br/doencas/cancer-da-pele/> (Accessed: 12 July 2024).
- Souza, V., Silva, L., Araújo, L., and Santos, A., (2020). Análise Comparativa de Redes Neurais Convolucionais no Reconhecimento de Cenas. *XI Computer on the Beach*, Balneário Camboriú, SC, Brasil, pp. 2-4 de Setembro.
- Tschandl, P., Rosendahl, C., and Kittler, H., (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 5(1). Available at: <http://dx.doi.org/10.1038/sdata.2018.161> (Accessed: 08 October 2024).
- Vargas, A.C.G., Carvalho, A.M.P., and Vasconcelos, C.N., (2016). Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In *Proceedings of the SIBGRAPI Conference on Graphics, Patterns and Images*. Available at: <http://gibis.unifesp.br/sibgrapi16/eproceedings/wuw/7.pdf> (Accessed: 08 October 2024).
- Villiger, B., Gambardella, L., and Franco, M., (2023). Machine learning para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil. *Cadernos de Saúde Pública*, 39(1), e00000023. SciELO.
- World Health Organization, (2023). *Digital health*. Available at: https://www.who.int/health-topics/digital-health#tab=tab_1 (Accessed: 17 September 2024).