



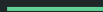
Construindo um ensemble de CNNs para classificação de imagens de lesões de pele

Classificação multiclasse de imagens dermatoscópicas

Autores

Pedro Lucas Silva Haga Torres

Roberta Barbosa Oliveira



Introdução

Caracterização do problema

- O câncer de pele é um dos tipos de câncer com maior incidência no Brasil [1]:
 - Podendo ser dividido em duas categorias: melanoma e não melanoma.
- O melanoma é a forma mais agressiva de câncer de pele, devido aos seus altos níveis de metástase;
- Existe um grande interesse mundial no desenvolvimento de sistemas de diagnóstico auxiliado por computador para processamento e análise de imagens de lesões de pele [2-4]:
 - Objetivando-se auxiliar a tomada de decisão dos dermatologistas.

Caracterização do problema

- Imagens dermatoscópicas são adquiridas a partir de técnicas não invasivas:
 - Vêm sendo amplamente utilizadas na literatura [6];
 - São adquiridas por meio de um dispositivo dermatoscópico que permite uma visualização mais detalhada dos padrões da lesão na superfície da pele;
 - Estudos anteriores demonstraram que métodos computacionais podem fornecer resultados adequados para a identificação de lesões de pele em imagem [7].
- Dentre as várias abordagens clínicas que utilizam imagens dermatoscópicas, a regra ABCD (Assimetria, Borda, Cor e estruturas Diferenciais) tem sido muito aplicada para extrair características para o diagnóstico automático (Abbas et al., 2011).

Caracterização do problema

- Algoritmos de aprendizagem de máquina (AM) se destacam na resolução de problemas envolvendo reconhecimento de padrões:
 - Entretanto, esses algoritmos necessitam de um grande volume de dados para alcançar uma solução satisfatória e generalizável.
- Em especial, arquiteturas de redes neurais convolucionais (CNN) são muito utilizadas em tarefas envolvendo imagens, dada sua capacidade de segmentação e reconhecimento de padrões [8];
- O International Skin Imaging Collaboration (ISIC) é um esforço internacional para aprimorar o diagnóstico de melanoma:
 - Em 2019 eles propuseram o desafio de classificar imagens dermatoscópicas em 9 classes diferentes:
 - Melanoma, nevo melanocítico, carcinoma basocelular, ceratose actínica, ceratose seborreica, dermatofibroma, lesão vascular e carcinoma de células escamosas, além de uma 9ª classe representando imagens que não pertencem a nenhuma das classes anteriores.

Metodologia

Base de dados

- Foi utilizada a base de dados de treinamento disponibilizada pelo ISIC [9-11], que consiste de mais de 25 mil imagens dermatoscópicas de lesões de pele:

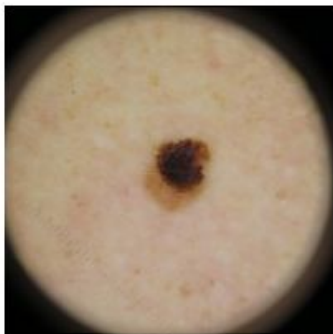
Diagnóstico	Número de amostras	Percentual do total
Melanoma	4522	17.85%
Nevo melanocítico	12875	50.83%
Carcinoma basocelular	3323	13.12%
Ceratose actínica	867	3.42%
Ceratose seborreica	2624	10.36%
Dermatofibroma	239	0.94%
Lesão vascular	253	1.00%
Carcinoma espinocelular	628	2.48%
N.D.A.	0	0.00%
Total	25331	100.00%

Base de dados

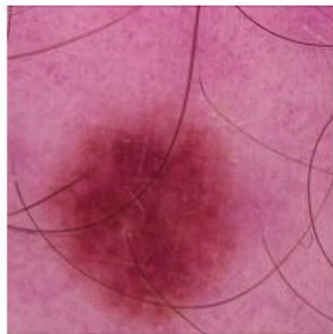
- Como é possível ver, a base está desbalanceada;
- Uma forma de se lidar com o desbalanceamento se dá pelo o uso de *data augmentation* [12]:
 - Pode-se aumentar o número de imagens que não pertencem a classe de nevo melanocítico;
 - Porém, o uso demasiado dessa técnica pode levar a uma baixa capacidade de generalização.
- Uma segunda estratégia para lidar com o desbalanceamento é utilizar diferentes pesos para cada uma das classes durante o cálculo da função de perda da rede [13]:
 - Utilizou-se a seguinte equação para calcular o peso das classes:

$$peso_{classe} = \frac{total_{imagens}}{num_{classes} * num_{amostras}}$$

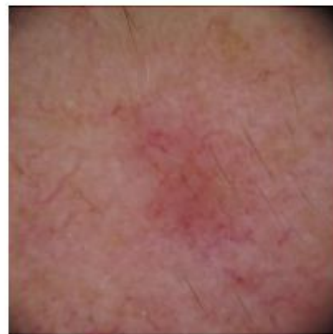
- Está se estudando como lidar com a 9ª classe (N.D.A.)



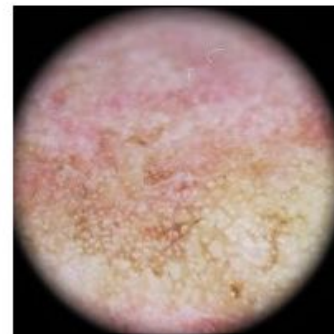
(a)



(b)



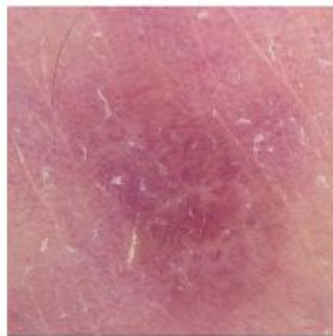
(c)



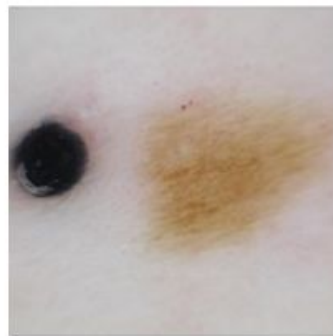
(d)



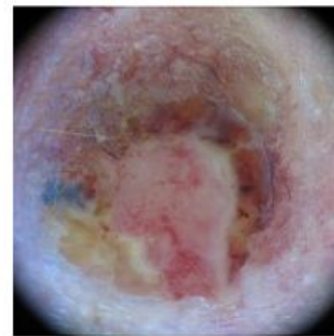
(e)



(f)



(g)



(h)

Exemplos de imagens dermatoscópicas: (a) melanoma, (b) nevo melanocítico, (c) carcinoma basocelular, (d) ceratose actínica, (e) ceratose seborreica, (f) dermatofibroma, (g) lesão vascular e (h) carcinoma espinocelular. Retiradas de [9-11].

Metodologia: Projeto de pesquisa do PIBIC

- Neste trabalho, se propõe o uso de uma arquitetura baseada no *ensemble* homogêneo de redes EfficientNet [14], utilizando diferentes conjuntos de transformações de *data augmentation* para treinar essas redes;
- **Ensemble:** conjunto de dois ou mais classificadores cujas decisões individuais são combinadas de alguma forma para gerar uma predição mais robusta [15];
- **Data augmentation:** método artificial para aumentar a quantidade de dados disponíveis para o conjunto de treinamento [12]:
 - No caso de imagens, utilizam-se transformações das imagens já disponíveis, como: rotação, alterações nos parâmetros de cor (como brilho, saturação e matiz), entre outros.

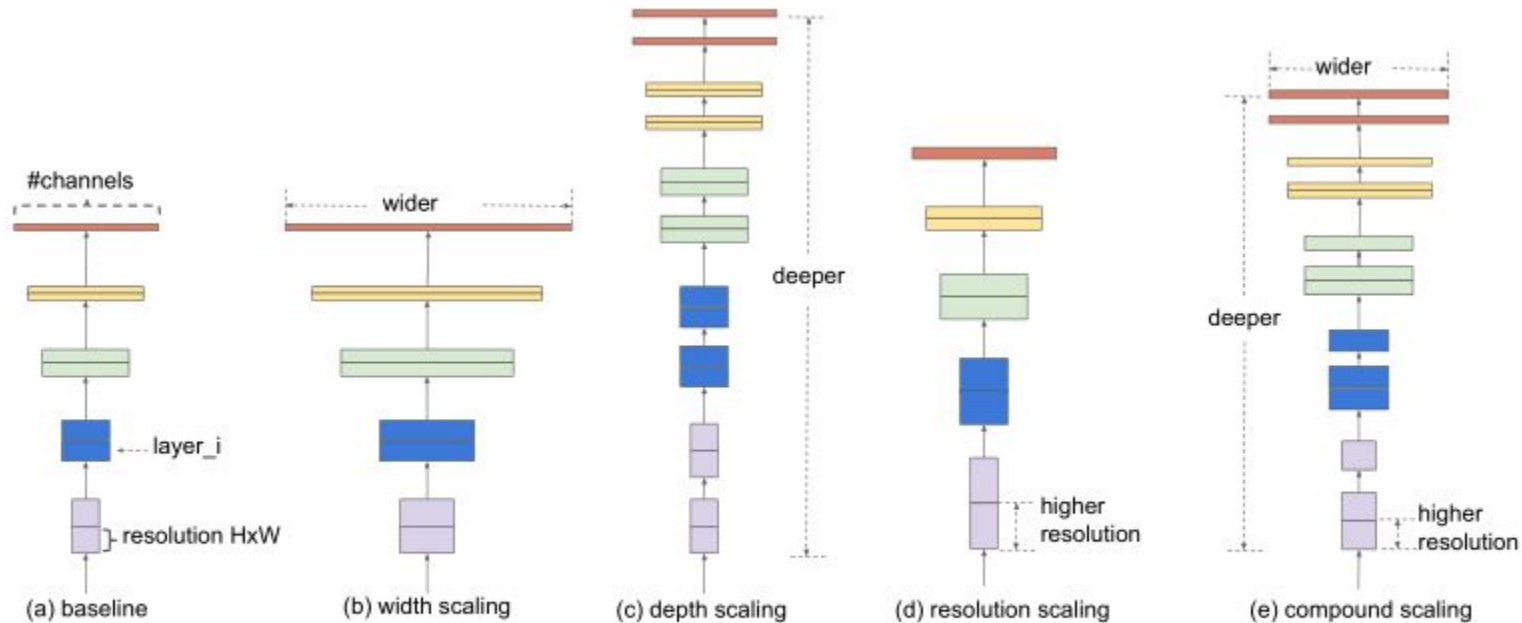
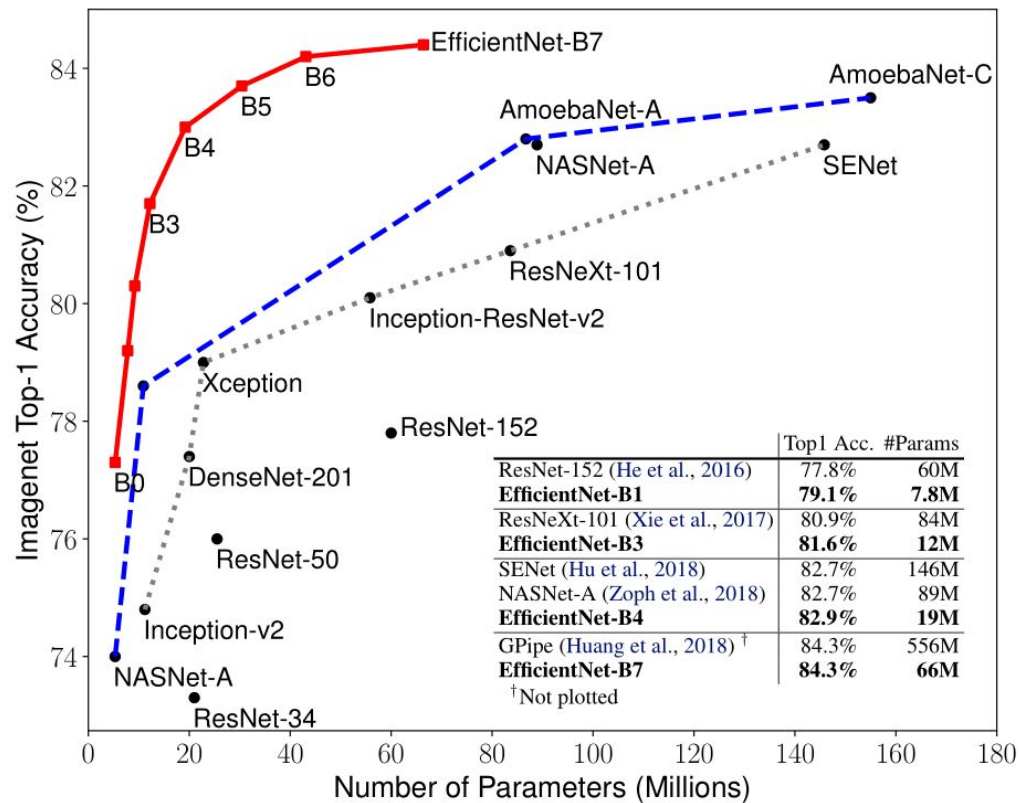


Figure 2. Model Scaling. (a) is a baseline network example; (b)-(d) are conventional scaling that only increases one dimension of network width, depth, or resolution. (e) is our proposed compound scaling method that uniformly scales all three dimensions with a fixed ratio.

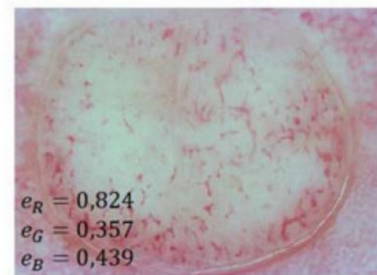
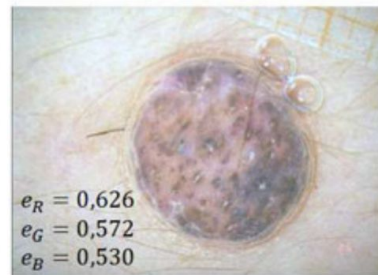
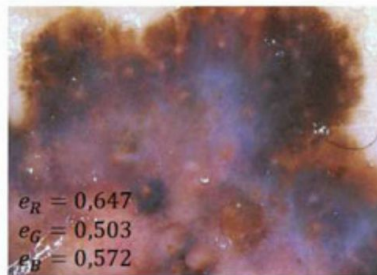
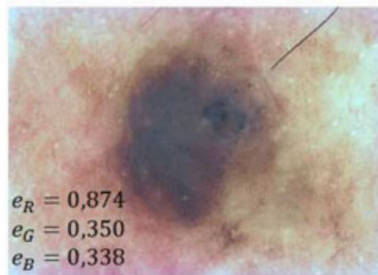
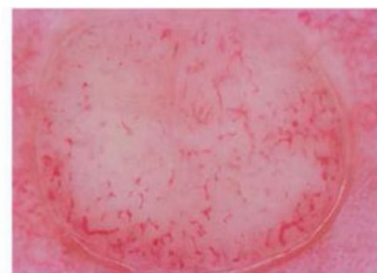
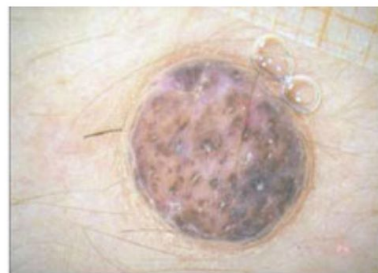
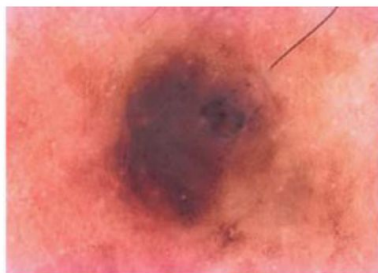
Fonte: M. Tan e Q. V. Le [14].



Fonte: M. Tan e Q. V. Le [14].

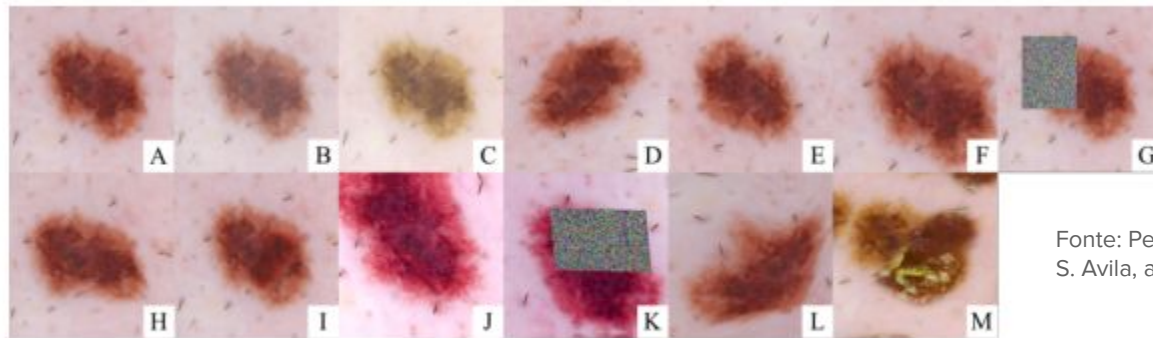
Metodologia

- Aplicou-se o algoritmo Shades of Gray [16] para garantir a consistência de cores entre todas as imagens:

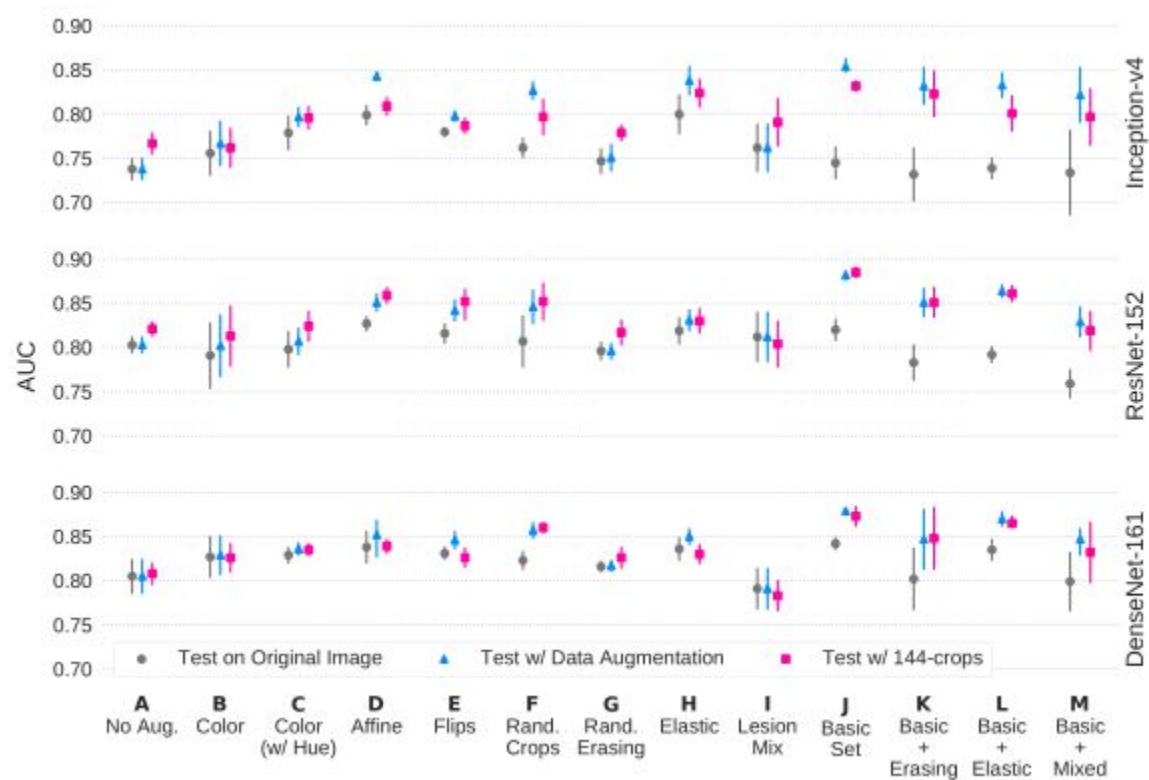


Metodologia

- As transformações de *data augmentation* foram feitas somente no conjunto de treinamento, de maneira aleatória;
- Conforme proposta por Perez et al. [18]:
 1. Cortes aleatórios;
 2. *Affine*;
 3. Inversão horizontal e/ou vertical;
 4. Alterações nos parâmetros de saturação, contraste e brilho.



Fonte: Perez, C. Vasconcelos, S. Avila, and E. Valle [18].

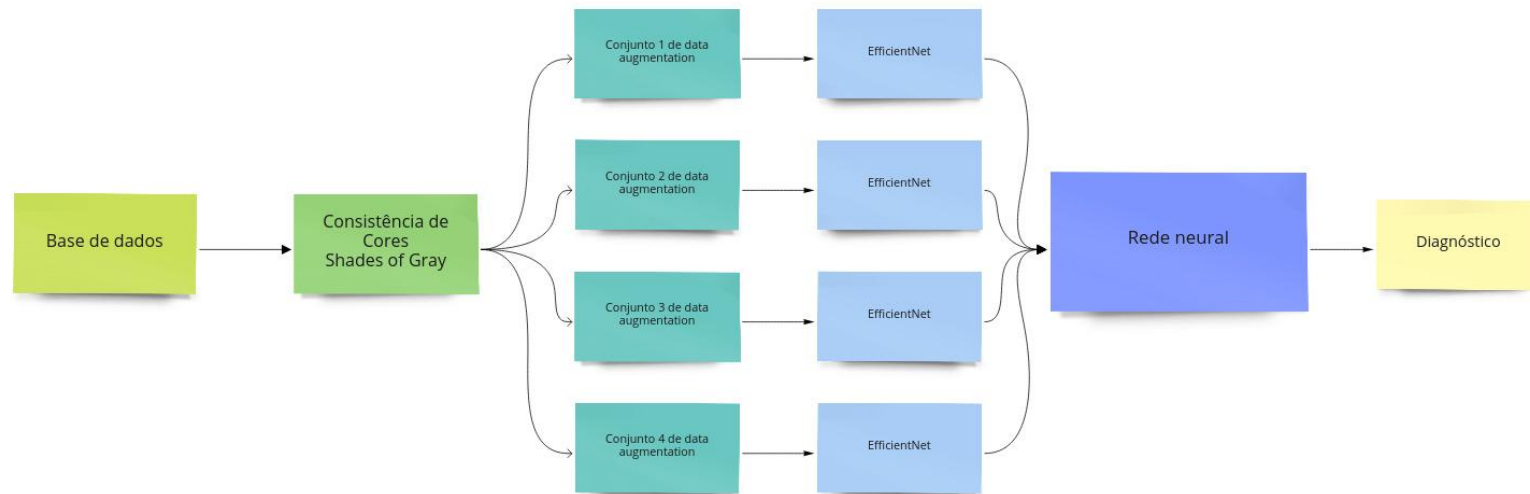


Fonte: Perez, C. Vasconcelos, S. Avila, and E. Valle [18].

Metodologia: Projeto de pesquisa (FSI)

- Foram avaliadas as performances de classificação de dois otimizadores: SGD com gradiente acelerado de Nesterov [19] e Adam [20]. Esses otimizadores são complementares:
 - O Adam converge para uma solução rapidamente e a qualidade da solução é satisfatória:
 - Entretanto, ele pode ter problemas com alguns conjuntos de dados [21].
 - O SGD converge mais lentamente, entretanto a qualidade de sua solução é superior.
- Ambos foram treinados e avaliados sobre o mesmo conjunto de dados:
 - Isso é, as mesmas imagens originais somadas as mesmas imagens geradas pela técnica de *data augmentation*.
- E treinados sobre a arquitetura B2 da EfficientNet (dadas as limitações de hardware disponível):
 - Utilizando *transfer learning* com os pesos pré-treinados da ImageNet [22].

Visão geral

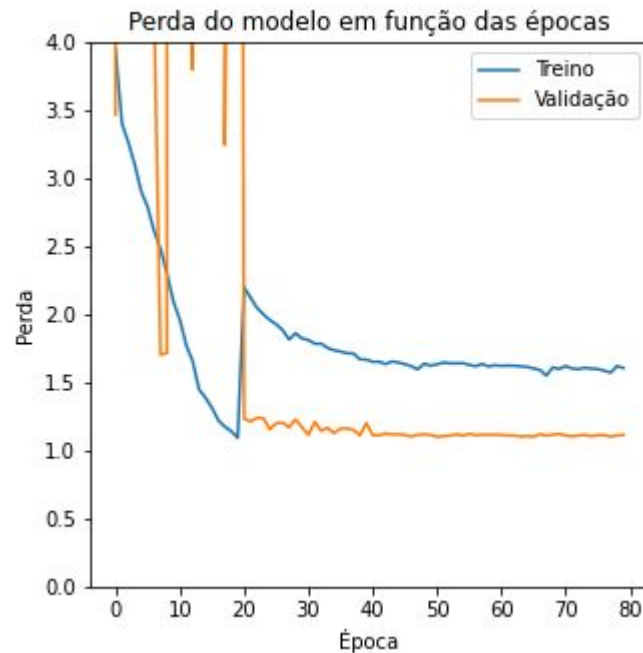


Resultados

Resultados de treinamento - perda (loss) da rede

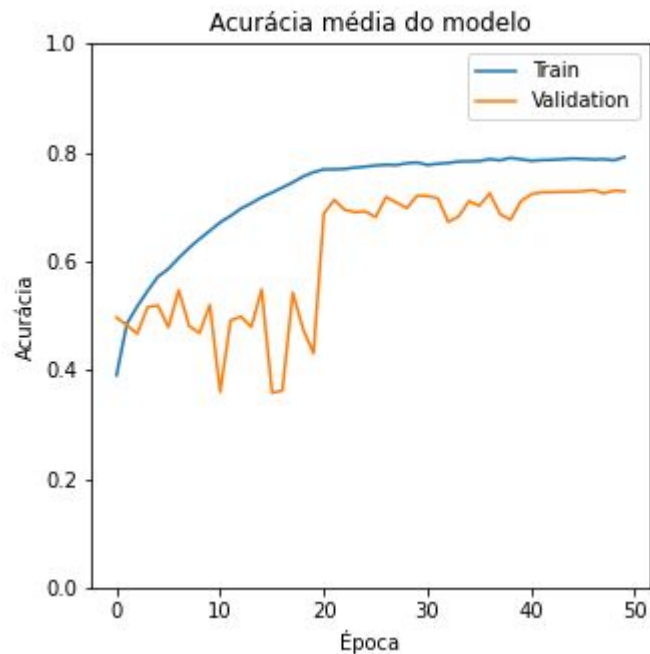


SGD + Nesterov

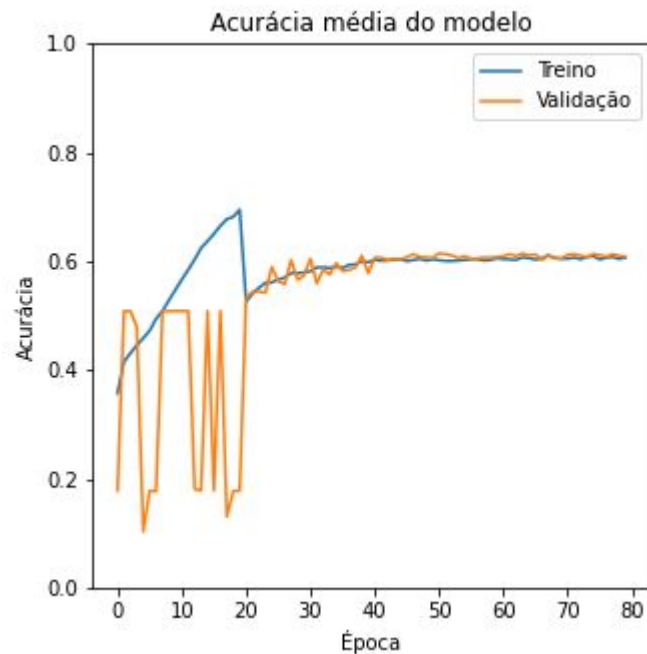


Adam

Resultados de treinamento - acurácia

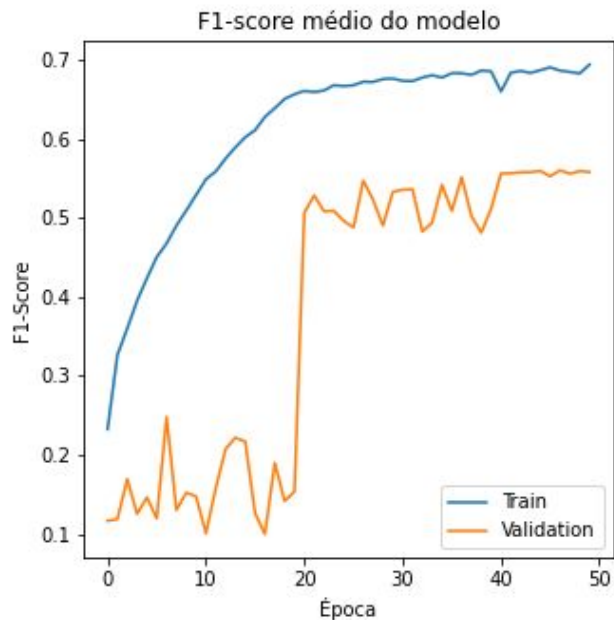


SGD + Nesterov

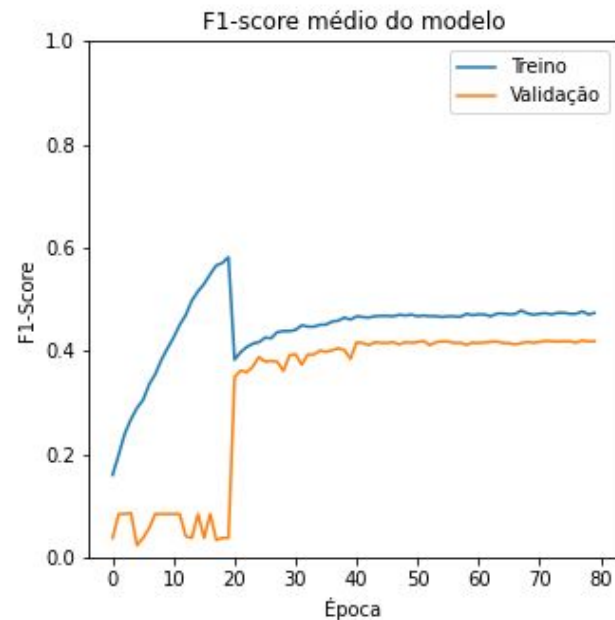


Adam

Resultados de treinamento - Macro F1-score

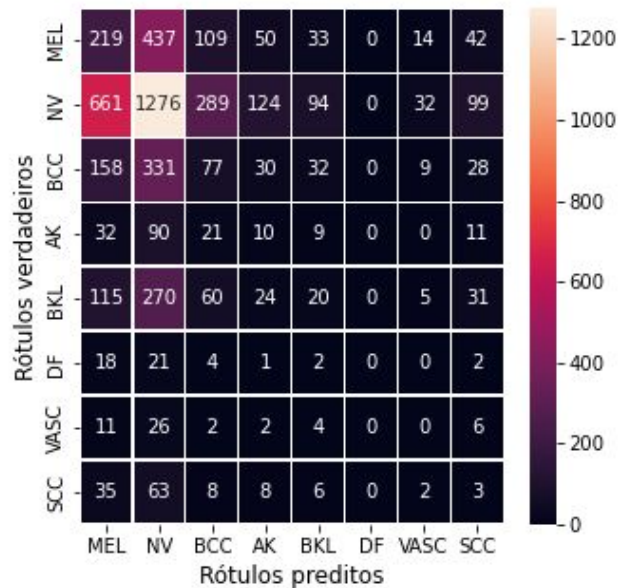


SGD + Nesterov

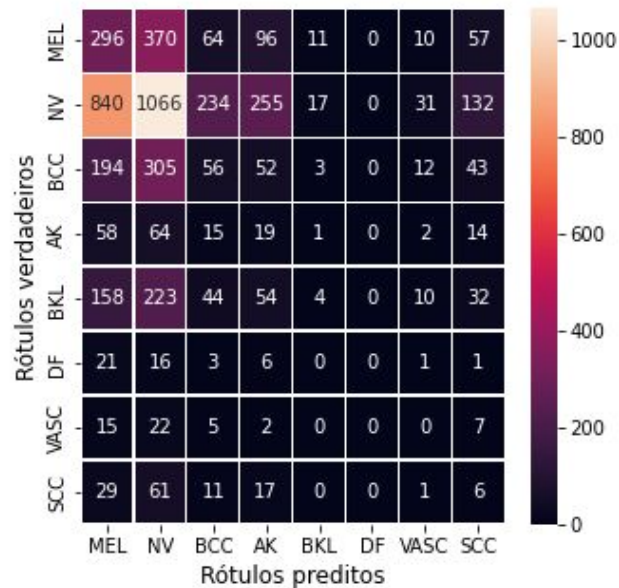


Adam

Resultados de validação - matriz de confusão

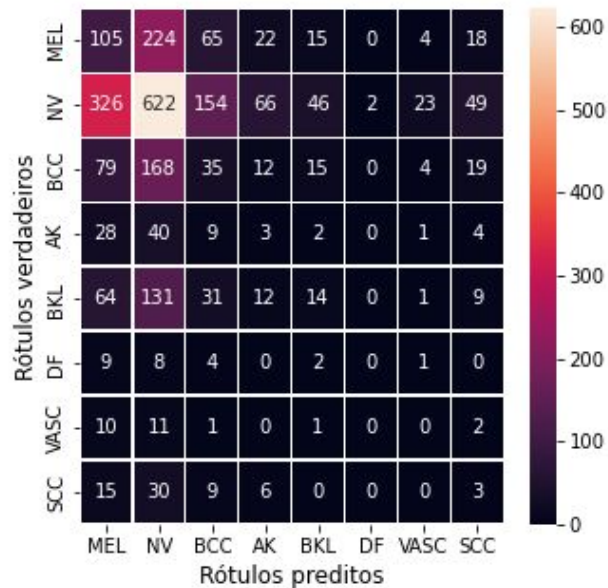


SGD + Nesterov

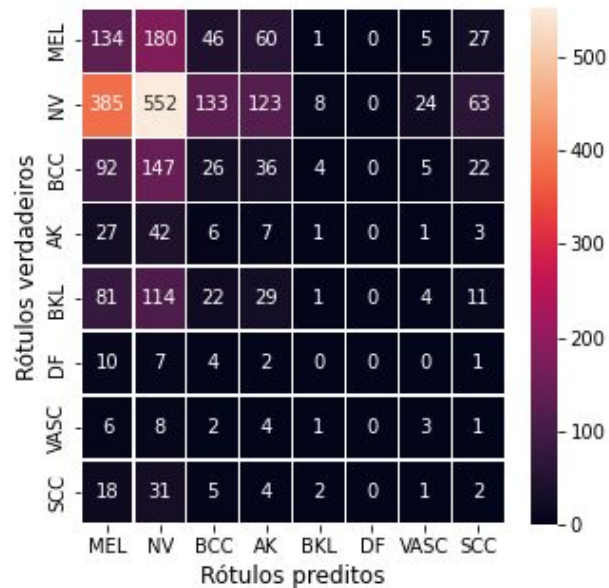


Adam

Resultados de teste - matriz de confusão



SGD + Nesterov



Adam

Resultados resumidos de validação e teste

Média aritmética entre as classes:

Métrica	SGD + Nesterov		Adam	
	Validação	Teste	Validação	Teste
Revocação	12,66%	12,16%	12,39%	12,98%
Especificidade	87,47%	87,58%	87,45%	87,84%
Precisão	12,67%	12,62%	12,30%	12,03%
F1-score	12,67%	12,12%	12,34%	12,49%
Acurácia	82,91%	82,91%	82,14%	82,15%

Conclusão

Conclusão

- Métricas de avaliação bastante próximas;
- Entretanto, o comportamento da curva de perda do modelo treinado com otimizador SGD + Nesterov é mais próxima do desejável:
 - Converge mais rápido e para uma solução melhor.
- Dessa forma, para o treinamento dos demais modelos a compor o *ensemble* usar-se-á SGD + Nesterov.

Referências bibliográficas

- [1] Instituto Nacional de Câncer José Alencar Gomes da Silva, Estimativa 2020: incidência de câncer no Brasil. INCA, 2019. [Online]. Available: <https://www.inca.gov.br/publicacoes/livros/estimativa-2020-incidencia-de-cancer-no-brasil>
- [2] A. Mahbod, G. Schaefer, I. Ellinger, R. Ecker, A. Pitiot, and C. Wang, “Fusing fine-tuned deep features for skin lesion classification,” Computerized Medical Imaging and Graphics, vol. 71, pp. 19– 29, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S089561118306050>
- [3] M. S. Ali, M. S. Miah, J. Haque, M. M. Rahman, and M. K. Islam, “An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models,” Machine Learning with Applications, vol. 5, p. 100036, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666827021000177>
- [4] I. Iqbal, M. Younus, K. Walayat, M. U. Kakar, and J. Ma, “Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images,” Computerized Medical Imaging and Graphics, vol. 88, p. 101843, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0895611120301385>
- [5] M. E. Celebi, T. Mendonca, and J. S. Marques, Dermoscopy image analysis. CRC Press, 2015, vol. 10.
- [6] Silveira, M., Nascimento, J. C., Marques, J. S., Marca, A. R. S., Mendonca, T., Yamauchi, S., Maeda, J., & Rozeira, J. (2009). Comparison of Segmentation Methods for Melanoma Diagnosis in Dermoscopy Images. IEEE Journal of Selected Topics in Signal Processing, 3, 35-45.
- [7] Abbas, Q., Fondón, I., & Rashid, M. (2011). Unsupervised skin lesions border detection via two-dimensional image analysis. Computer Methods in Biomedicine, 104, e1-e15.

Referências bibliográficas

- [8] W. Wang, Y. Yang, X. Wang, W. Wang, and J. Li, “Development of convolutional neural network and its application in image classification: a survey,” *Optical Engineering*, vol. 58, no. 4, pp. 1 – 19, 2019. [Online]. Available: <https://doi.org/10.1117/1.OE.58.4.040901>
- [9] P. Tschandl, C. Rosendahl, and H. Kittler, “The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions,” *Scientific data*, vol. 5, no. 1, pp. 1–9, 2018.
- [10] N. C. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler et al., “Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic),” in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018, pp. 168–172.
- [11] M. Combalia, N. C. Codella, V. Rotemberg, B. Helba, V. Vilaplana, O. Reiter, C. Carrera, A. Barreiro, A. C. Halpern, S. Puig et al., “Bcn20000: Dermoscopic lesions in the wild,” *arXiv preprint arXiv:1908.02288*, 2019.
- [12] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, 2019, cited By :516. [Online]. Available: <https://doi.org/10.1186/s40537-019-0197-0>
- [13] J. M. Johnson and T. M. Khoshgoftaar, “Survey on deep learning with class imbalance,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–54, 2019.
- [14] M. Tan and Q. V. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” *CoRR*, vol. abs/1905.11946, 2019. [Online]. Available: <http://arxiv.org/abs/1905.11946>

Referências bibliográficas

- [15] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach. Upper Saddle River, NJ, USA: Pearson, 2010, vol. 3, ch. 18, p. 748.
- [16] G. D. Finlayson and E. Trezzi, “Shades of gray and colour constancy,” Color and Imaging Conference, vol. 2004, no. 1, pp. 37–41, 2004. [Online]. Available: <https://www.ingentaconnect.com/content/ist/cic/2004/00002004/00000001/art00008>
- [17] C. Barata, M. E. Celebi, and J. S. Marques, “Improving dermoscopy image classification using color constancy,” IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 3, pp. 1146–1152, 2015.
- [18] F. Perez, C. Vasconcelos, S. Avila, and E. Valle, “Data augmentation for skin lesion analysis,” in OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis. Springer, 2018, pp. 303–311.
- [19] Y. Nesterov, “A method for unconstrained convex minimization problem with the rate of convergence $o(1/k^2)$,” Doklady AN USSR, vol. 269, pp. 543–547, 1983. [Online]. Available: <https://ci.nii.ac.jp/naid/20001173129/en/>
- [20] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization”. 2017.
- [21] A. C. Wilson, R. Roelofs, M. Stern, N. Srebro, and B. Recht, “The marginal value of adaptive gradient methods in machine learning,” 2018.
- [22] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255.