

# Avaliação de Redes de Segmentação de Deep Learning para Segmentar Melanoma

Lucas B. M. de Souza<sup>1</sup> Samuel Pedro B. D. Lélis<sup>1</sup>, Romuere R. V. Silva<sup>1</sup>

<sup>1</sup>Sistemas de Informação, Universidade Federal do Piauí (UFPI)  
Picos – PI – Brazil

{lucasbezerra527, samuellelis88}@gmail.com, romuere@ufpi.edu.br

**Abstract.** *Melanoma is the most serious cause of most deaths among skin cancer types. In addition, its incidence is increasing worldwide, thus showing the importance of cancer detection systems through medical imaging to obtain a faster diagnosis. One of the steps is segmentation, which deals with isolating the injured region. In this research, segmentation methods were compared using different backbones with the U-Net and FPN neural networks. Using the image databases PH<sup>2</sup> and DermIS, 0.66 and 0.56 of Dice values were obtained, respectively. Thus, this set can favor the achievement of results closer to state of the art.*

**Resumo.** *Dentre os cânceres de pele, o melanoma é o mais grave e o motivo da maioria das mortes dentre os tipos de câncer de pele. Além disso, sua incidência está aumentando cada vez mais pelo mundo, assim mostrando a importância dos sistemas de detecção do câncer através de imagens médicas para a obtenção de um diagnóstico mais rápido. Uma das etapas é a segmentação, que trata do isolamento da região lesionada. Nesta pesquisa realizou-se a comparação de resultados utilizando diferentes backbones com as redes neural U-Net e FPN. Com a utilização das bases PH<sup>2</sup> e DermIS, foram obtidos 0,66 e 0,56 de valores de Dice, respectivamente. Assim acredita-se que esse conjunto pode favorecer a obtenção de resultados mais próximos ao estado da arte.*

## 1. Introdução

Em 2020 a estimativa para novos casos de melanoma foi de 324,6 milhões em todo mundo, sendo 173,8 milhões homens e 150,8 mulheres, já para mortes a estimativa é de 57,0 milhões, 32,4 milhões homens e 24,7 mulheres. Para não melanoma a estimativa foi de 1198,1 milhões para novos caso, com 722,3 milhões homens e 475,7 milhões mulheres, a estimativa de mortes foi de 63,7 milhões, 37,6 milhões homens e 26,1 milhões mulheres [Ferlay et al. 2021].

O melanoma é o motivo da maioria das mortes por câncer de pele, sendo fatores genéticos e radiação ultravioleta as principais causas envolvidas [Zhang et al. 2020]. Dentre os tipos de câncer de pele, o melanoma é o mais perigoso pois pode provocar metástase (disseminação do câncer para outros órgãos), ele se origina nos melanócitos (células produtoras de melanina). Os seus principais sintomas são manchas pruriginosas, sinais ou pintas que mudam de tamanho, forma ou cor e feridas que não cicatrizam em até 4 semanas [Ministério da Saúde 2021].

Em trabalhos como [Nazi and Abir 2020], são propostos algoritmos que conseguem classificar a lesão em benigna ou maligna. Com o objetivo de melhorar a classificação, foi feita a segmentação da região de interesse (ROI), ou seja, foi removida da imagem a porção que não constitui lesão. Conforme é afirmado no trabalho [Cicerone and Camp 2019], as abordagens de *Computer-Aided-Detection* CAD podem ajudar o patologista com o fornecimento de métricas de imagem para calcular vários diagnósticos, sendo uma delas a identificação de potenciais ROIs, pois proporciona uma economia de tempo significativa para o fornecimento do diagnóstico.

Dessa forma, o objetivo deste trabalho é desenvolver um algoritmo capaz de segmentar imagens de lesões. Para essa finalidade serão testados modelos de redes neurais com *backbones* de outras redes, visando trabalhar com o melhor conjunto para segmentação em trabalhos futuros. Após identificar o melhor conjunto, serão aplicados métodos de *machine learning* para melhorar o resultado da segmentação obtida. Foram utilizados no experimento 2 modelos, sendo eles da U-Net [Ronneberger et al. 2015] e *Feature Pyramid Networks* (FPN) [Lin et al. 2017]. Já os utilizados *backbones* são dos tipos VGG, ResNet, SE-ResNet, ResNeXt, DenseNet, Inception e MobileNet. Para a efetividade dos conjuntos de modelos e *backbones*, estes foram testados em duas bases de dados, sendo elas PH<sup>2</sup> [Mendonça et al. 2013] e DermIS [DermIS 2021].

## 2. Revisão de Literatura

Em [Nazi and Abir 2020], foi utilizada uma rede convolucional U-Net para segmentar imagens, obtendo uma média 0,87 com o Dice com a base ISIC 2018 [Codella et al. 2019]. Devido ao número de imagens, foram realizados os aumentos de dados: rotação aleatória, inversão aleatória, zoom aleatório, distorção gaussiana, brilho aleatório, cor aleatória, contraste aleatório, distorção elástica aleatória e equalização do histograma. Foi utilizada também a base PH<sup>2</sup> atingindo 0,93 de Dice.

No estudo de [Al-masni et al. 2018] foi proposto uma metodologia de segmentação com base em redes neurais convolucionais de resolução total [Long et al. 2015], sem a necessidade de pré ou pós-processamento para a remoção de artefatos ou do aprimoramento da lesão segmentada. Uma das bases de dados utilizadas para avaliar o método foi a PH<sup>2</sup>, na qual obteve-se 91,38% de coeficiente Dice.

Visando a resolução de problemas de desequilíbrio de dados e de desempenho no treinamento da U-Net tradicional, o trabalho proposto por [Abraham and Khan 2018] combinou a U-Net com uma nova variante da função de perda focal [Lin et al. 2020], obtendo uma rede mais adequada à segmentação de pequenas lesões. Foi usada a base de imagens ISIC 2018, na qual a segmentação atingiu 0,856 para o Dice.

Os autores de [Kim et al. 2017] obtiveram uma média de Dice na base de imagens PH<sup>2</sup> de 0,940. Para segmentação das imagens foi desenvolvida uma rede de aprendizado profundo iterativa de codificador-decodificador, que consiste em camadas de convoluções para codificador e camadas de deconvolução para decodificador. Esse método propõe a segmentação de várias imagens médicas, sendo testado na base PH<sup>2</sup> e na base de imagens de retina DRIVE [Staal et al. 2004].

No trabalho de [Santos et al. 2020] foram utilizadas 1380 imagens de quatro conjuntos de dados públicos contendo imagens de pele: o PH<sup>2</sup> e DermIS para estabelecer os

parâmetros do método, e ISIC 2016 [Gutman et al. 2016] e ISIC 2017 [Codella et al. 2018] para a avaliação de desempenho. O método proposto que tem o intuito de segmentar as imagens aplica o algoritmo SLICO [Nan and Mu 2014] para geração de super-pixels e extração de recursos de textura. Depois, o agrupador Seeded Fuzzy C-means [Veras et al. 2018] com base em regiões marcadas feitas por um especialista. A etapa final é o pós-processamento, onde são eliminados ruídos e suavizadas as bordas. Como resultado os autores atingiram uma acurácia superior a 95% para ambos os conjuntos de imagens (ISIC 2016 e ISIC 2017).

No trabalho de segmentação de melanoma [Aljanabi et al. 2018] é proposto uma metodologia baseada em artificial *bee colony algorithm* (ABC) [Karaboga 2005]. Na fase de pré-processamento é realizado uma remoção de ruídos, já na próxima etapa, é aplicado o algoritmo ABC para a detecção de melanoma. O método atingiu 95,24–97,61% e 83,56–85,25% para a média de accuracy e coeficiente de Jaccard, utilizando as bases de dados: PH<sup>2</sup>, ISBI 2016, ISBI 2017 e DermIS.

### 3. Metodologia

O método proposto tem o objetivo de realizar um estudo comparativo entre as arquiteturas de redes neurais para segmentação de imagens e utilizando diferentes backbones. Foram utilizados 14 *backbones* de 7 redes diferentes, em conjunto com as redes U-Net e FPN. Além disso, cada conjunto ser treinado e testado a rede com as bases PH<sup>2</sup> e DermIS. O detalhamento do método proposto está ilustrado na Figura 1.

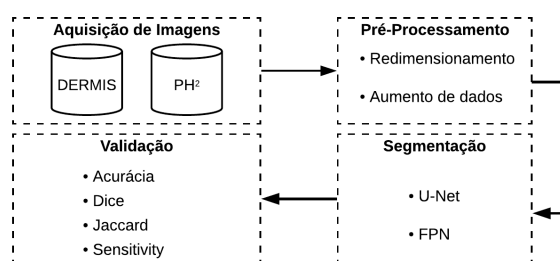
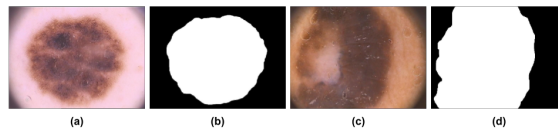


Figura 1. Fluxograma de etapas da metodologia.

#### 3.1. Aquisição de Imagens

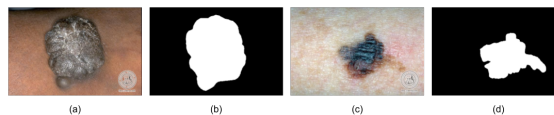
No processo de aquisição, são reunidas as imagens que contém as lesões cutâneas e as segmentadas manualmente por um médico (máscaras). Sendo elas as bases PH<sup>2</sup> e DermIS.

De acordo com [Mendonça et al. 2013], o conjunto de dados PH<sup>2</sup> foi desenvolvido para fins de pesquisa e *benchmarking*, a fim de facilitar estudos comparativos sobre algoritmos de segmentação e classificação de imagens dermatoscópicas. Tal base contém um total de 200 imagens, sendo 160 não melanomas e 40 melanomas, disponível em [Universidade do Porto 2021]. Na Figura 2, (a) tem-se um exemplo de uma imagem de lesão não melanoma e em (b) sua máscara, em (c) um exemplo de melanoma e (d) sua máscara da base PH<sup>2</sup>.



**Figura 2. Exemplo lesão e máscara PH<sup>2</sup>, sendo (a) não melanoma e (b) sua máscara, (c) melanoma e (d) máscara.**

A base DermIS é do serviço de informática DermIS.net, que se caracteriza por ser o maior serviço de informação dermatológica disponível na Internet [DermIS 2021]. Oferece atlas de imagens elaboradas completas com diagnósticos diferentes, relatos de casos e informações adicionais sobre doenças de pele. Ela contém 206 imagens, sendo 119 melanoma e 87 não melanoma. Na Figura 3, (a) tem-se um exemplo de uma imagem de lesão não melanoma e em (b) sua máscara, em (c) um exemplo de melanoma e (d) sua máscara da base DermIS.



**Figura 3. Exemplo lesão e máscara DermIS, sendo (a) não melanoma e (b) sua máscara, (c) melanoma e (d) máscara.**

### 3.2. Pré-processamento

Nessa etapa foram preparadas as imagens a serem utilizadas nas etapas posteriores, foi aplicado duas métricas o redimensionamento e o aumento de dados. Na aplicação da primeira métrica, as imagens foram redimensionadas de forma que a largura e comprimento fossem iguais, tanto da máscara quanto da imagem contendo a lesão. Após vários testes, para definir um tamanho de imagem possível para utilizar nas duas redes, a melhor opção foi o tamanho  $384 \times 384$ , um tamanho para o qual é possível redimensionar todas as imagens.

Na aplicação da segunda métrica, foi implementado um aumento de dados, uma das principais razões para isso é a quantidade de imagens que uma rede neural necessita para ser treinada. Assim como afirma [Ronneberger et al. 2015], esse processo é importante para ensinar à rede a invariância e propriedades de robustez da forma desejada quando apenas algumas amostras de treinamento estão disponíveis. Foram geradas deformações suaves usando vetores de deslocamento aleatórios em uma grade de  $3 \times 3$ . Os deslocamentos são amostras de uma distribuição gaussiana com padrão de 10 pixels de desvio. Os deslocamentos por pixel são então calculados usando interpolação bicúbica. As camadas de abandono no final do caminho de contração realizam aumentos implícitos de dados.

### 3.3. Segmentação

Nesse trabalho utilizamos uma biblioteca com redes neurais para segmentação de imagens [Pavel Yakubovskiy 2021], assim é possível escolher o modelo com pesos pré-treinados e *backbones* de outras redes. Os *backbones* utilizados são apresentados na Tabela 1, e os modelos utilizados são: U-Net, FPN.

**Tabela 1. Backbones utilizados nas redes neurais.**

Tipo	Nomes
VGG	vgg16, vgg19
ResNet	resnet18, resnet34, resnet50
SE-ResNet	seresnet18, seresnet34, seresnet50
ResNeXt	resnext50
DenseNet	densenet121
Inception	inceptionv3, inceptionresnetv2
MobileNet	mobilenet, mobilenetv2

A U-Net é uma rede neural convolucional, ela foi criada em um desafio para segmentação de estruturas neuronais em pilhas de imagens microscópicas eletrônicas. Os ganhadores propuseram essa rede que tem como característica o treinamento que se baseia no forte uso de aumento de dados para usar as amostras anotadas disponíveis de forma mais eficiente. A arquitetura consiste em um caminho de contração para capturar o contexto e um caminho de expansão simétrico que permite uma localização precisa. No trabalho [Ronneberger et al. 2015] é mostrado que tal rede pode ser treinada ponta a ponta a partir de muito poucas imagens.

FPN é um modelo que por conta da sua estrutura em pirâmide, é utilizado uma pirâmide de imagens, em que é possível avaliar a imagem original em escalas diferentes, mantendo a informação em todos os níveis a um custo computacional alto. Além do uso de apenas a última camada convolucional permitir um processamento muito mais rápido, porém com um alcance de escalas limitado [Lin et al. 2017].

### 3.4. Validação

As métricas utilizadas para a avaliação do acerto na segmentação foram: Acurácia (ACC), Dice, *Jaccard* (JAC)[Cohen 1968], *Sensitivity* (SEN). Para calcular essas métricas é utilizada a matriz de confusão [Provost and Kohavi 1998] na geração dos valores: Falso Negativo (FN), Falso Positivo (FP), Verdadeiro Positivo (VP) e Verdadeiro Negativo (VN).

## 4. Resultados e discussão

Os experimentos consistem em mesclar cada *backbone* com as redes e as bases, em busca de definir o melhor conjunto de treinamento para as duas bases de imagens utilizadas, os melhores resultados de cada conjunto de treinamento estão presente na Tabela 2. Utilizando a base PH<sup>2</sup> foi obtido 0,65 ou 0,66 de Dice para a maioria dos *backbones* utilizando a rede U-Net, com exceção dos *backbones* do tipo VGG que foram baixos.

**Tabela 2. Melhor média resultante dos conjunto de treinamento.**

Base	Rede	Backbone	Dice	SEN	ACC	JAC
PH <sup>2</sup>	U-Net	seresnet34	0,66±0,02	0,68±0,03	0,77±0,02	0,49±0,03
DERMIS	U-Net	seresnet34	0,56±0,22	0,65±0,29	0,90±0,01	0,39±0,15
PH <sup>2</sup>	FPN	mobilenetv2	0,45±0,05	0,73±0,21	0,44±0,08	0,29±0,04
DERMIS	FPN	mobilenetv2	0,16±0,02	0,76±0,29	0,30±0,23	0,09±0,01

Utilizando a rede U-Net e a base DermIS, a maioria dos *backbones* obteve uma média de Dice próxima, sendo a menor, 0,45, para o *backbone* mobilenetv2 e os demais ficaram entre 0,52 e 0,55.

Os resultados obtidos com a utilização da FPN com os *backbones* não foram satisfatórios, tanto na base PH<sup>2</sup> como na DermIS, pois foram obtidos resultados mais baixos do que com a U-Net, o melhor resultado obtido no conjunto de dados PH<sup>2</sup> foi 0,45 para a métrica Dice.

Utilizando o *dataset* DermIS e a FPN, não foram obtidos resultados maiores que 0,20 para a métrica dice, assim apenas FPN com esses *backbones* não são adequados para segmentação, pois a sua segmentação não se caracteriza como uma lesão de câncer de pele, nesse conjunto de dados.

Ao decorrer do experimento devido a aquisição dos melhores resultados foi obtido o melhor resultado com a utilização do *backbone* seresnet34 em conjunto com a rede U-Net, sendo este de 0,66 para a métrica Dice na base de imagens PH<sup>2</sup>, já utilizando a DermIS é atingido uma média de 0,56. São resultados menores do que os presente trabalhos relacionados, mas com isso será possível definir pré e pós-processamento para melhorar este conjunto que já obteve resultados medianos para segmentação.

Dentre as dificuldades, a maior delas é o tempo para o processamento e recursos, pois necessita de um computador com GPU e memória dedicadas. Além do tempo que dificultou o teste em outras base, não foi possível realizar processamentos com outras redes e outros *backbones*. Também não foi realizado a busca de parâmetros para definir uma melhor configuração da rede para segmentação. Poderia ainda aplicar outros métodos de pré ou pós-processamento como remoção de ruídos, fechamento de regiões e limiarização. Assim poderia melhorar a segmentação, porém o uso desses métodos requerem uma análise detalhada das segmentações para ser identificado qual técnica aplicar.

## 5. Conclusão

Este trabalho propõe definir o melhor conjunto de treinamento para a segmentação de imagens de lesões de câncer de pele avaliando vários algoritmos para segmentação, diferente dos trabalhos relacionados que propuseram o desenvolvimento de um algoritmo específico para segmentação. Apesar da metodologia incompleta devido à ausência outros métodos de pré ou pós-processamento, pois ainda não foi trabalhado para melhorar os resultados do melhor conjunto, obteve-se 0,66 para a métrica Dice utilizando a base PH<sup>2</sup> no modelo da U-Net com o *backbone* seresnet34, assim acredita-se que esse conjunto pode favorecer a obtenção de resultados mais próximos ao estado da arte.

Para trabalhos futuros objetiva-se melhorar os resultados obtidos com o melhor conjunto e utilizar neste a base ISIC 2018, que possui uma quantidade de imagens maior. Pretende-se aplicar mapas de textura, tais como LBP, LPQ e GLCM, antes da segmentação e aplicar métodos na etapa de pós-processamento como remoção de pequenas regiões, fechamento de buracos. Além disso, pretende-se realizar busca de parâmetros para definir uma melhor configuração da rede para segmentação.

## Referências

Abraham, N. and Khan, N. M. (2018). A novel focal tversky loss function with improved attention U-Net for lesion segmentation.

- Al-masni, M. A., Al-antari, M. A., Choi, M.-T., Han, S.-M., and Kim, T.-S. (2018). Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Computer Methods and Programs in Biomedicine*, 162:221 – 231.
- Aljanabi, M., Özok, Y. E., Rahebi, J., and Abdullah, A. S. (2018). Skin lesion segmentation method for dermoscopy images using artificial bee colony algorithm. *Symmetry*, 10(8):347.
- Cicerone, M. T. and Camp, C. H. (2019). 21 - potential roles for spectroscopic coherent raman imaging for histopathology and biomedicine. In Alfano, R. R. and Shi, L., editors, *Neurophotonics and Biomedical Spectroscopy*, Nanophotonics, pages 547 – 570. Elsevier.
- Codella, N., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., Kittler, H., and Halpern, A. (2019). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC).
- Codella, N. C. F., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., and Halpern, A. (2018). 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).
- Cohen, J. (1968). Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70(4):213.
- DermIS (2021). Dermatology information system. Disponível em: <https://www.dermis.net/dermisroot/en/home/index.htm>. Acesso em 29 de maio de 2021.
- Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., and Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International Journal of Cancer*.
- Gutman, D., Codella, N. C. F., Celebi, E., Helba, B., Marchetti, M., Mishra, N., and Halpern, A. (2016). Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic).
- Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization. Technical report, Citeseer.
- Kim, J. U., Kim, H. G., and Ro, Y. M. (2017). Iterative deep convolutional encoder-decoder network for medical image segmentation. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 685–688.
- Lin, T., Goyal, P., Girshick, R., He, K., and Dollár, P. (2020). Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2):318–327.
- Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017). Feature pyramid networks for object detection.

- Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440.
- Mendonça, T., Ferreira, P., Marques, J., Marçal, A., and Rozeira, J. (2013). Ph2 - a dermoscopic image database for research and benchmarking. *Conference proceedings:2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2013:5437–5440.
- Ministério da Saúde (2021). Câncer de pele: o que é, causas, sintomas, tratamento e prevenção. Disponível em: <https://antigo.saude.gov.br/saude-de-a-z/cancer-de-pele>. Acesso em 15 de outubro de 2020.
- Nan, B. and Mu, Z. (2014). Slic0-based superpixel segmentation method with texture fusion. *Chinese Journal of Scientific Instrument*, 35(3):527–534.
- Nazi, Z. A. and Abir, T. A. (2020). Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with U-Net and dcnn-svm. In Uddin, M. S. and Bansal, J. C., editors, *Proceedings of International Joint Conference on Computational Intelligence*, pages 371–381, Singapore. Springer Singapore.
- Pavel Yakubovskiy (2021). Segmentation models. Disponível em: <https://segmentation-models.readthedocs.io/en/latest/tutorial.html#quick-start>. Acesso em 30 de maio de 2021.
- Provost, F. and Kohavi, R. (1998). Glossary of terms. *Journal of Machine Learning*, 30(2-3):271–274.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation.
- Santos, E., Veras, R., Miguel, H., Aires, K., Claro, M. L., and Junior, G. B. (2020). A skin lesion semi-supervised segmentation method. In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 33–38.
- Staal, J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., and van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509.
- Universidade do Porto (2021). Ph2 database. Disponível em: <https://www.fc.up.pt/addi/ph2%20database.html>. Acesso em 15 de outubro de 2021.
- Veras, R., Aires, K., Britto, L., et al. (2018). Medical image segmentation using seeded fuzzy c-means: A semi-supervised clustering algorithm. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE.
- Zhang, N., Cai, Y.-X., Wang, Y.-Y., Tian, Y.-T., Wang, X.-L., and Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. *Artificial Intelligence in Medicine*, 102:101756.