

Aumento de Dados Utilizando Firefly e Level Sets Aplicado à Segmentação de Imagens Médicas e Biológicas

Lucas P. Laheras^{1,*}, Paulo S. Rodrigues¹, Francisco J. P. Lopes²,
Ondina F. J. Palmeira², Alexandre X. Falcão³, Bárbara C. Benato³, Gilson A. Giraldi⁴

¹Departamento de Ciência da Computação
Centro Universitário FEI – São Bernardo – SP – Brasil

²Grupo de Biologia do Desenvolvimento e Sistemas Dinâmicos
Universidade Federal do Rio de Janeiro (UFRJ) – Duque de Caxias – RJ – Brasil

³Departamento de Sistemas de Informação
Universidade Estadual de Campinas (UNICAMP) – Campinas – SP – Brasil

⁴Departamento de Ciência da Computação
Laboratório Nacional de Computação Científica (LNCC) – Petrópolis – RJ – Brasil

*lucaslaheras@hotmail.com

Abstract. *It has been estimated that about 75 % of disease-associated genes in humans have homologs in *Drosophila melanogaster*. Despite recent advances in biological and medical image capture technologies, for counting collected cells in vivo, it is still necessary to make the manual supply process. On the other hand, convolutional networks have been advancing and obtaining good results, especially on bases with a large volume. This work proposes adapting a data augmentation technique, increasing a specific segmentation pipeline for confocal microscopy images, and uses the results in segmentation training using U-Net.*

Resumo. *Estima-se que cerca de 75 % dos genes associados a doenças em humanos tenham homólogos em *Drosophila melanogaster*. Apesar dos recentes avanços nas tecnologias de captura de imagens biológicas e médicas, para a contagem de células coletadas in vivo, ainda é necessário fazer o processo manual de segmentação. Por outro lado, redes convolucionais vêm avançando e obtendo bons resultados, especialmente em bases com grande volume. Esse trabalho propõe a adaptação de uma técnica de aumento de dados, incrementando um pipeline de segmentação específico para imagens de microscopia confocal, e utiliza os resultados no treinamento de segmentações utilizando uma rede neural U-Net.*

1. Introdução

Avanços significativos na geração de imagens das áreas de medicina e biologia têm gerado uma forte demanda pelo desenvolvimento de tecnologias relacionadas à digitalização automática, tanto em hardware quanto em software. Por conta disso, o desenvolvimento das áreas de processamento de imagens, inteligência artificial, aprendizagem profunda e ciência de dados é estratégico, particularmente para a área médica, uma vez que, atualmente tanto hardware quanto software têm influenciado em diagnósticos, tratamentos e planejamento cirúrgico [Doi 2007].

Por outro lado, o desenvolvimento e recentes avanços na microscopia confocal por varredura a laser, associado ao uso de marcadores fluorescentes, têm possibilitado a geração de dados sobre dos níveis de expressão gênica e localização proteica tanto em células fixadas quanto *in vivo* [Stephens and Allan 2003]. Estima-se que cerca de 75 % dos genes associados a doenças em humanos tenham homólogos em *Drosophila melanogaster* (mosca da fruta) [Pandey and Nichols 2011]. Essas grande semelhanças entre os genes em humanos e em *Drosophila*, bem como o seu rápido ciclo de desenvolvimento, têm contribuído para que esse organismo seja um dos mais importantes modelos biológicos nos mais diversos campos da biologia e medicina. Dentre os estudos se destacam Alzheimer [Higham et al. 2019], Câncer [Palmeira 2017, Sonoshita and Cagan 2017], doenças neurodegenerativas [Hirth 2010], vírus [Hughes et al. 2012], para citar alguns. A Figura 1 demonstra as imagens utilizadas, são células do ovário de *Drosophilas* produzidas no Grupo de Biologia do Desenvolvimento e Sistemas Dinâmicos da Universidade Federal do Rio de Janeiro (UFRJ). A análise dessas imagens (Figura 1) tem por objetivo a segmentação das diversas células marcadas com o objetivo de identificar as células-tronco embrionárias responsável pela formação do tecido.

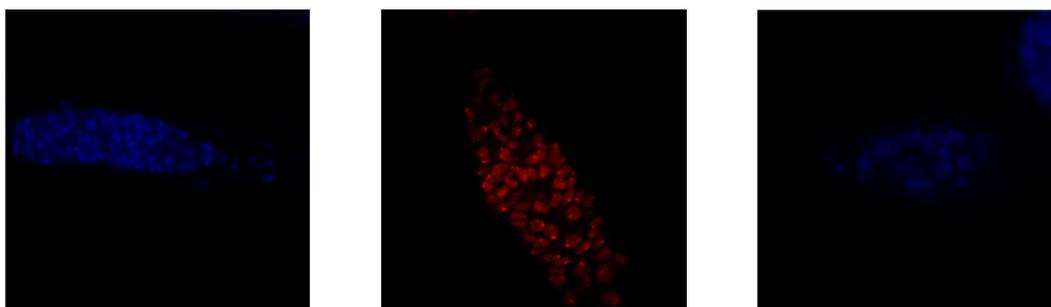


Figura 1. Exemplos de imagens de microscopia confocal do ovário de *Drosophila*. Marcação da estrutura de fusomos característica (cor azul, marcado DAPI) e para uma proteína Adducin (vermelho, marcação de membrana).

Também recentemente, a literatura de processamento digital de imagens e visão computacional tem demonstrado a crescente efetividade da aplicação de algoritmos bio-inspirados para segmentação de imagem [Wachs Lopes et al. 2018], bem como no método de Level Set [Malladi et al. 1994]. Por outro lado, métodos em aprendizagem profunda, baseados em rede neuronal convolucional (*Convolutional Neural Network - CNN*) foram também utilizados para segmentação de imagens, obtendo resultados muito promissores [Ronneberger et al. 2015]. Vale ressaltar que a área de aprendizagem profunda vem ganhado espaço em inúmeras aplicações em reconhecimento de padrões, jogos, veículos autônomos, além de outras [Naturalis 2018].

Porém, o treinamento de uma CNN para segmentação, em geral, necessita de grandes bancos de imagens que, na maioria das vezes, não estão disponíveis. Para acumular dados suficientes e melhorar o desempenho de modelos de aprendizado profundo em classificação de imagens, várias estratégias de aumento de dados foram propostas [Shorten and Khoshgoftaar 2019, Peixinho et al. 2018], no entanto, há ainda a necessidade de serem adaptadas para segmentação de imagens.

Sendo assim, este artigo propõe a adaptação de uma metodologia para a realização do aumento de dados em imagens segmentadas. Assim, para isso foi criado um pipeline para segmentação automática utilizando a metodologia de [Peixinho et al. 2018]. Essa metodologia combina técnicas de auto codificação, projeção 2D, visão computacional, Level Sets e algoritmos bio-inspirados. Os códigos relacionados a este artigo estão disponíveis no GitHub^{1 2}.

Este trabalho está organizado da seguinte forma: na Seção 2, será apresentada a teoria utilizada para o desenvolvimento deste trabalho. A Seção 3 detalha a metodologia proposta para o desenvolvimento deste trabalho, demonstrando em detalhes a abordagem. Na Seção 4, serão apresentados os resultados. Uma conclusão é apresentada na Seção 5.

2. Fundamentos

2.1. Firefly

O algoritmo *firefly* [Yang 2009] é uma meta-heurística inspirada pelo comportamento dos vaga-lumes, que são atraídos um pelo outro de acordo com sua luminescência natural. O algoritmo é utilizado para limiarização multi-níveis na imagem de entrada, auxiliando na segmentação de imagens de baixo contraste.

2.2. Algoritmo *Level Sets*

O conceito básico do algoritmo level set é delimitar a fronteira de uma região planar utilizando o conjunto de pontos no nível zero de uma função $\alpha : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$ dado pela seguinte expressão:

$$S(t) = \{(x, y) \in \mathbb{R}^2 \mid \alpha(x, y, t) = 0\} \quad (1)$$

Dada uma curva de nível zero $S(t)$ no instante t , consideram-se três possíveis valores para a função. Quando o valor é menor que 0 a coordenada (x, y) está dentro da curva S , quando o valor é maior que 0 a coordenada (x, y) está fora da curva, e quando o valor for igual a 0 a coordenada (x, y) pertence a S . O próximo passo é encontrar a formulação euleriana da evolução da frente, gerando uma equação diferencial parcial que descreve a evolução da curva de nível no tempo e no espaço [Malladi et al. 1994].

2.3. U-Net

A U-Net [Ronneberger et al. 2015] é um tipo de rede neural totalmente convolucional (*Fully Convolutional Network*) [Long et al. 2015]. Como observado na Figura 2, a arquitetura da U-Net consiste em duas etapas: contração e expansão. Na contração é aplicada repetidamente duas vezes a convolução 3x3 seguida pela ReLU. Em seguida, é aplicado a operação *max pool* 2x2 iterando com passo 2 para redução de dimensionalidade. Em cada camada de contração é dobrado o número de mapas de características. Na expansão é aplicada uma operação de *upsampling*, seguida por uma convolução 2x2 chamada de *up-convolution* que reduz pela metade o número de mapas de características. Em seguida ocorre a concatenação com a imagem correspondente e duas convoluções 3x3, ambas seguidas pela ReLU.

¹<https://github.com/LucasLaheras/CellDetectionImage>

²<https://github.com/LucasLaheras/dataaugmentation>

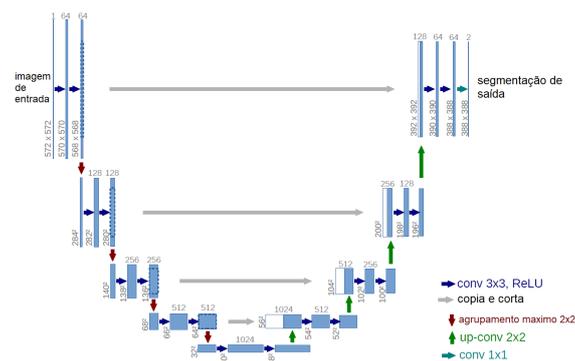


Figura 2. Arquitetura da U-Net. Modificado de Ronneberger et. al [Ronneberger et al. 2015]

2.4. Rede Auto-codificadora

Uma rede auto-codificadora que é um algoritmo de aprendizado de máquina não supervisionado que reproduz na saída os valores dos dados fornecidos como entrada. Esse método é composto por duas redes neurais (codificador e decodificador) sendo que a saída do codificador é a entrada do decodificador. Após o treinamento da rede, essa saída é interpretada como um vetor de características que representa o dado de entrada. Auto-codificadores são usados para reduzir a dimensionalidade dos dados [Peixinho et al. 2018].

3. Metodologia Proposta

A metodologia proposta neste trabalho é constituída de duas partes: pipeline específico para segmentação de imagens de microscopia confocal de óvulos de *Drosophilas* e aumento de dados utilizando o pipeline para novas imagens geradas.

3.1. Pipeline

O *pipeline* exibido na Figura 3 foi implementado em Python, sendo composto por 5 módulos principais, descritos em mais detalhes a seguir.



Figura 3. Pipeline utilizado para pre-segmentação das imagens.

O primeiro bloco (bloco 1, Figura 3) converte a imagem original para tons-de-cinza. No módulo correspondente ao bloco 2 da Figura 3, a imagem fornecida pelo bloco 1 é processada pelo algoritmo bio-inspirado *Firefly*.

No passo correspondente à construção da máscara (bloco 3, Figura 3) a imagem de saída do módulo 2 é composta por $N \geq 1$ regiões de interesse. Se a área de uma dessas regiões for maior que um limiar $L > 0, L \in N$, então essa região é definida como sendo uma máscara binária. Sendo assim, a saída desse módulo é o conjunto $S =$

$\{m_1, m_2, \dots, m_N\}$ de máscaras binárias. Em seguida, é aplicado três vezes a operação de dilatação elemento estruturante 3×3 em todo conjunto S gerado.

Assim, seleciona-se a máscara com a maior área, a qual é utilizada para a aplicação do histograma local (Bloco 4, Figura 3), sobre a imagem original. A razão da aplicação da equalização local é o fato de que, localmente, as regiões inomogêneas são reduzidas.

No módulo referente ao bloco 5 da Figura 3, para cada imagem de saída do módulo anterior, é utilizado o algoritmo *Level Set* (Seção 3) de forma a delimitar as regiões locais, para a obtenção de um resultado melhor na contagem de células.

3.2. Aumento de Dados

Para o aumento de dados foi utilizado um banco de imagens de microscopia confocal de câncer, composto por 250 imagens, que contém tanto as imagens segmentadas quanto o padrão-ouro da classificação entre tumor maligno e benigno.

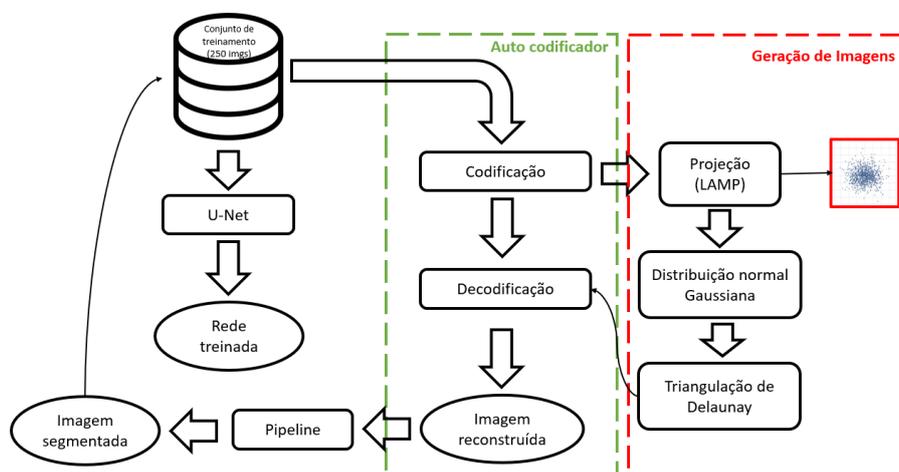


Figura 4. Arquitetura do sistema de aumento de dados.

Para a técnica de aumento de dados foi utilizado o fluxograma ilustrado na Figura 4, adaptada de [Peixinho et al. 2018]. A técnica consiste em, após o treinamento do auto-codificador, ele é utilizado para codificar o conjunto de imagens de entrada, transformando-as em vetores n -dimensionais. Em seguida, cada imagem é projetada em coordenadas bidimensionais utilizando a técnica *Local Affine Multidimensional Projection* (LAMP) [Joia et al. 2011]. Dado os pontos da projeção, aplica-se a triangulação de Delaunay. Em seguida, são criados novos pontos usando a distribuição normal gaussiana no plano. Os novos pontos serão considerados novas imagens codificadas, no espaço n -dimensional, somente se estiverem dentro de um triângulo que possuir todos os pontos com a mesma classificação. Após essa seleção, todos os novos pontos serão decodificados e transformados em imagens que antes não existiam, gerando assim a segmentação do pipeline utilizado.

4. Resultados

Na Figura 5 é mostrado um exemplo do processamento em cada etapa do *pipeline* proposto. Pode-se observar na Figura 5.(a) que, por ser uma imagem homogênea, a

visualização da área de interesse possui contraste muito baixo. Nas Figuras 5.(b) e 5.(c) são formadas as máscaras da região de interesse global, mas ainda não é possível identificar cada célula individualmente. A Figura 5.(d) observa-se a saída da etapa de segmentação local, onde pode-se observar uma melhora do contraste apenas dentro dessa região de interesse. A saída da etapa de segmentação local (Figura 5.(d)) é utilizada como entrada para a etapa final, Level Set. Considerando que esse processo final baseia-se fundamentalmente na utilização de gradiente, a melhora da segmentação local tem forte influência da qualidade da aplicação desse algoritmo. A Figura 5.(e) ilustra o resultado final.

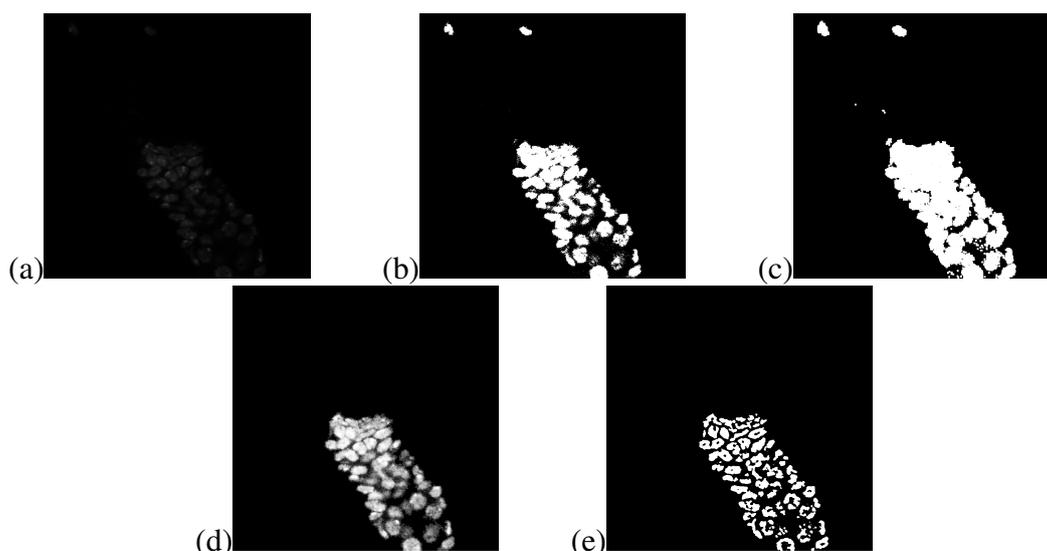


Figura 5. (a) Imagem original em escala de cinza. (b) Firefly aplicado na imagem após ajuste de contraste. (c) Máscaras delimitando as regiões obtidas. (d) Equalização do histograma aplicado localmente. (e) Resultado da aplicação do método de Level set.

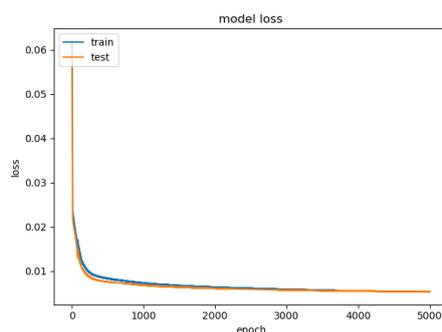


Figura 6. Histórico de perda durante o treinamento do auto-codificador, onde em azul representa a perda das imagens de treinamento e em laranja as imagens de teste

Com relação à metodologia para aumento de dados, o primeiro passo é o treinamento de uma rede auto-codificadora. Em seguida, os resultados do codificador são projetados no espaço bidimensional. Os resultados gerados após o treinamento do auto-codificador mostraram-se satisfatórios usando a função de perda do erro quadrático

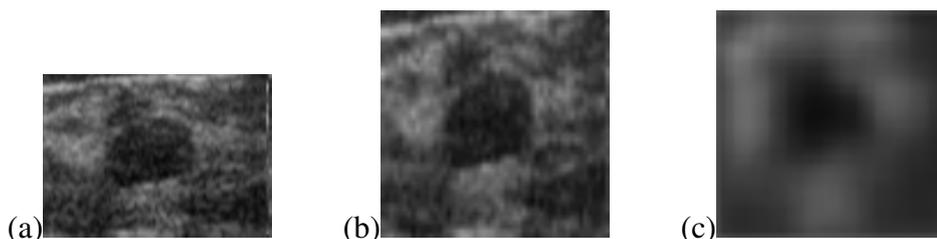


Figura 7. (a) Imagem original 99x71. (b) Imagem remodelada para 52x52. (c) Saída do auto-codificador.

médio, com erro de 00,529%. O codificador com três camadas de convolução utilizando a ativação ReLU seguidas pelo *pooling* máximo 2×2 com iteração 2 e o decodificador com a mesma configuração de convolução, seguido pelo *up-sampling* 2×2 , onde a entrada é a saída do codificador. Apesar de serem utilizadas 5000 épocas para o treinamento, a convergência próxima de 600 épocas apresentou resultados satisfatório, como pode-se observar na Figura 6.

A Figura 7.(a) ilustra um exemplo de uma imagem original da base de dados, a qual é redimensionada para 52×52 pixels por conta da padronização da rede (Figura 7.(b)). O resultado obtido na saída do auto-codificador, para a imagem de entrada da Figura 7.(b), está representada na Figura 7.(c). Apesar da perda de detalhes da imagem original, pode-se observar que o codificador gerou um vetor de características eficiente para esse exemplo.

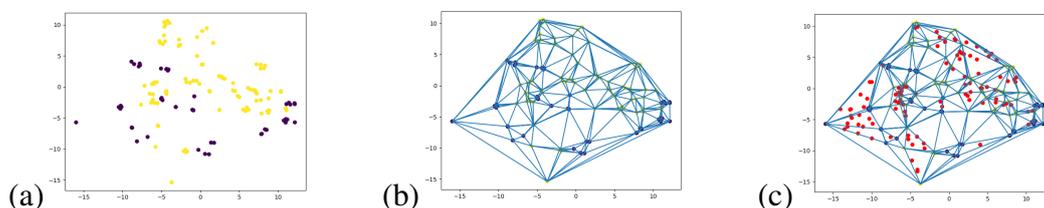


Figura 8. (a) Representação bidimensional das imagens. (b) Aplicação da triangulação de Delaunay. (c) Criação dos novos pontos em vermelho, que posteriormente serão convertidos em imagens.

Após o treinamento, o auto-codificador é utilizado para a geração de novas imagens, como descrito na Seção 3. A Figura 8.(a) ilustra a representação de cada imagem no plano bidimensional após passar pelo codificador. A Figura 8.(b) mostra a triangulação de Delaunay no plano, que é utilizada para selecionar os novos pontos baseado na classificação das imagens originais em benignas ou malignas. Os pontos em vermelho na Figura 8.(c) serão usados para gerar, por interpolação, novos vetores de características no espaço do codificador. Em seguida, estes vetores são processados pelo decodificador a fim de gerar novas imagens. A Figura 9.(a) mostra uma nova imagem gerada neste processo e a Figura 9.(b) mostra sua segmentação obtida com a metodologia proposta neste trabalho. Apesar do efeito de suavização as novas imagens geradas são consideradas satisfatórias para o aumento de dados no treinamento da U-Net, uma vez que é possível visualizar o tumor modificado das imagens da base de dados.

Nos testes de assertividade durante o treinamento foram utilizados três conjuntos de imagens para treinamento: (A) 100 imagens da base original; (B) 400 imagens, sendo

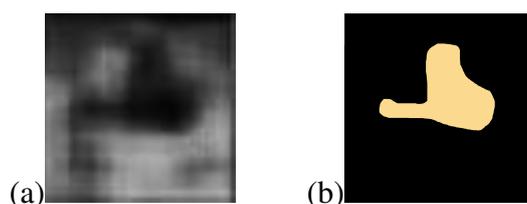


Figura 9. (a) Imagem gerada pelo aumento de dados. (b) Segmentação realizada pelo *pipeline*.

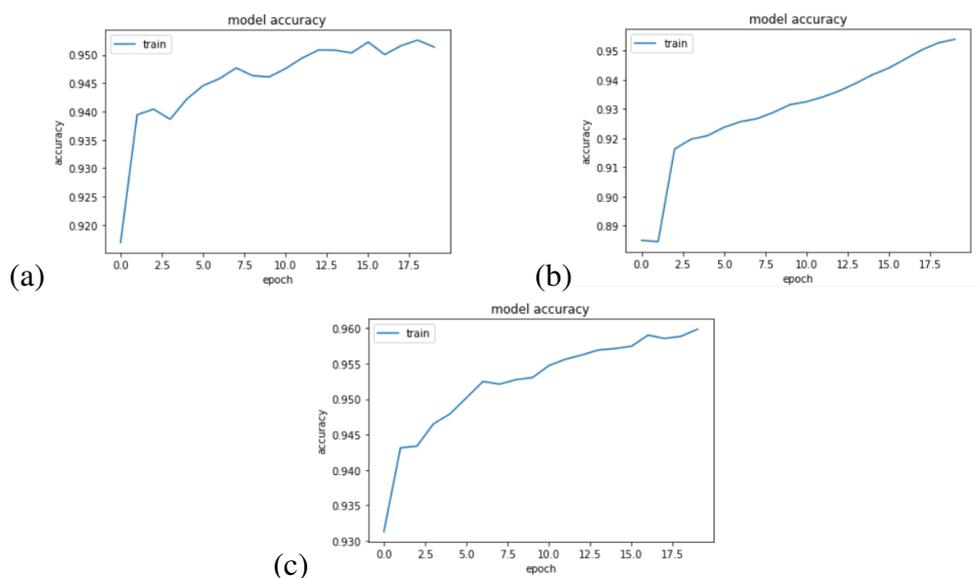


Figura 10. Gráfico da acurácia da U-Net (a) No conjunto sem aumento de dados. (b) No conjunto com aumento de dados padrão. (c) No conjunto com aumento de dados implementado e o padrão.

100 imagens da base com utilização do aumento de dados padrão rotacionando a imagem em 90, 180 e 270 graus; (C) sendo as imagens do conjunto (B) com as imagens geradas pelo aumento de dados. Os resultados do treinamento estão ilustrados no gráfico da Figura 10.(a), 10.(b) e 10.(c). Foi observado que a rede treinada com o conjunto (A) alcançou uma acurácia de 94,82% e perda de 11,94%, o conjunto (B) teve acurácia de 95,38% e perda de 11,11%, e por fim, o conjunto (C) atingiu acurácia de 96,01% e perda de 9,54%. Apesar da semelhança entre as acurácias, os testes preliminares mostram que o treinamento usando o conjunto (A) gerou uma rede hiper-especializada no conjunto de treinamento, o que afeta o desempenho de novas amostras de imagens, de forma que, não identifica corretamente o contorno das regiões.

5. Conclusão

Nesse trabalho, foi apresentada uma adaptação de aumento de dados para segmentação de imagens aplicada na rede neural U-Net. Para tanto, foi utilizado um pipeline de segmentação que combina técnicas de visão computacional, algoritmo *Firefly* e Level set. As imagens de microscopia foram coletadas por biólogos da Universidade Federal do Rio de Janeiro (UFRJ). Nota-se que, na base sem o aumento de dados ocorre *overfitting* e ambos os métodos de aumento de dados melhoram a rede, com vantagem de 0,63% maior de acurácia usando ambos os métodos. Como trabalhos futuros, pretende-se estudar esse

tipo de procedimento em outras técnicas de aumento de dados para realizar comparações e utilizar outras métricas para validação.

6. Agradecimentos

Agradecimentos ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico - Projeto 155130/2019-6), à FAPERJ (Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro - Projeto E-26 010.001877/2015 e E-26/200.675/2017), assim como à FEI (Fundação Educacional Inaciana), ao LNCC (Laboratório Nacional de Computação Científica) e à Universidade Federal do Rio de Janeiro (UFRJ), pelo suporte deste trabalho.

Referências

- Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*, 31(4-5):198–211.
- Higham, J. P., Malik, B. R., Buhl, E., Dawson, J. M., Ogier, A. S., Lunnon, K., and Hodge, J. J. (2019). Alzheimer’s disease associated genes ankyrin and tau cause shortened lifespan and memory loss in drosophila. *Frontiers in cellular neuroscience*, 13.
- Hirth, F. (2010). *Drosophila melanogaster* in the study of human neurodegeneration. *CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders)*, 9(4):504–523.
- Hughes, T. T., Allen, A. L., Bardin, J. E., Christian, M. N., Daimon, K., Dozier, K. D., Hansen, C. L., Holcomb, L. M., and Ahlander, J. (2012). *Drosophila* as a genetic model for studying pathogenic human viruses. *Virology*, 423(1):1–5.
- Joia, P., Paulovich, F., Coimbra, D., Cuminato, J., and Nonato, L. (2011). Local affine multidimensional projection. *IEEE transactions on visualization and computer graphics*, 17:2563–71.
- 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.
- Malladi, R., Sethian, J. A., and Vemuri, B. C. (1994). Shape modeling with front propagation: A level set approach.
- Naturalis, P. (2018). 30 amazing applications of deep learning.
- Palmeira, O. F. (2017). Efeitos da irradiação em células-tronco germinativas de *drosophila melanogaster* e influência da dieta na recuperação dessas células e tecidos ovarianos submetidos ao tratamento com cisplatina. Monografia apresentada ao Programa de Graduação em Ciências Biológicas: Biotecnologia, do Polo Avançado de Xerém, Universidade Federal do Rio de Janeiro. UFRJ.
- Pandey, U. B. and Nichols, C. D. (2011). Human disease models in *drosophila melanogaster* and the role of the fly in therapeutic drug discovery. *Pharmacological reviews*, 63(2):411–436.

- Peixinho, A. Z., Benato, B. C., Nonato, L. G., and Falcão, A. X. (2018). Delaunay triangulation data augmentation guided by visual analytics for deep learning. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 384–391. IEEE.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.
- Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):60.
- Sonoshita, M. and Cagan, R. (2017). Modeling human cancers in drosophila. In *Current topics in developmental biology*, volume 121, pages 287–309. Elsevier.
- Stephens, D. J. and Allan, V. J. (2003). Light microscopy techniques for live cell imaging. *science*, 300(5616):82–86.
- Wachs Lopes, G. A., Beltrame, F. S., Santos, R. M., and Rodrigues, P. S. (2018). Comparison of bio-inspired algorithms from the point of view of medical image segmentation. In *2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI)*, pages 1–7.
- Yang, X.-S. (2009). Firefly algorithms for multimodal optimization. In *SAGA*.