

# Segmentação de tireoide em imagens de ultrassom e classificação de nódulos

Artur Santos Nascimento  
Departamento de Computação  
Universidade Federal de Sergipe  
São Cristóvão, Sergipe, 49045-015  
Email: artur.nascimento@dcomp.ufs.br

Daniel Oliveira Dantas  
Departamento de Computação  
Universidade Federal de Sergipe  
São Cristóvão, Sergipe, 49045-015  
Email: ddantas@dcomp.ufs.br

**Resumo**—The identification and classification of malign thyroid nodules is fundamental to the thyroid cancer treatment. The TN-SCUI2020 challenge offers a dataset with 3644 ultrasound images for training and 910 ultrasound images as test dataset, adding up to 4554 ultrasound images. This work aims to use this dataset to develop a segmentation and classification task, segmenting the thyroid and classifying whether the thyroid is sick. Using several neural networks, it was possible to achieve an accuracy of 45.36% in the segmentation task using UNet-VGG16 neural network and 65.786% of accuracy in the classification using Mobilenet Neural Network.

## I. INTRODUÇÃO

A tireoide é uma glândula em formato de borboleta localizada na base do pescoço, em frente à traqueia. Ela produz hormônios responsáveis principalmente pelo metabolismo e crescimento do indivíduo, bem como interfere no peso, memória, concentração, humor, fertilidade e ciclo menstrual. Em alguns casos, a tireoide pode apresentar funcionamento inadequado, produzindo hormônios em quantidade maior (hipertiroidismo) ou menor (hipotiroidismo) [1].

O funcionamento inadequado da tireoide pode ser causado pelo desenvolvimento de nódulos. Na maioria das vezes, os nódulos são benignos. Todavia, em cerca de 5% dos casos esses nódulos são câncer. O tratamento para o câncer de tireoide se baseia na remoção da glândula, resultando na necessidade da administração de hormônios sintéticos que originalmente são produzidos pela tireoide. Cerca de 98% dos pacientes diagnosticados com câncer de tireoide sobrevivem ao tratamento e o diagnóstico precoce aumenta as chances de sobrevivência do paciente. Assim, para diagnosticar e tratar o câncer de tireoide nos estágios iniciais, é importante identificar os nódulos malignos com precisão. [1], [2].

O ultrassom (US) da tireoide é uma ferramenta fundamental para a avaliação dos nódulos. É um exame não-invasivo, em tempo real e que não expõe o paciente à radiação. Em contrapartida, é um exame de difícil interpretação e cujo reconhecimento da diferença sutil entre nódulos malignos e benignos também é difícil, fazendo com que o diagnóstico seja demorado e muito dependente do conhecimento e da experiência dos médicos envolvidos [2], [3].

Recentemente, diversos projetos de diagnóstico auxiliado por computador (CAD) têm sido usados para tentar suplan-

tar este problema. Entretanto, geralmente é difícil avaliar a eficácia de cada uma das abordagens, pois não havia nenhum parâmetro de referência disponível [3]. Com isso em mente, o desafio TN-SCUI2020 propõe que sejam desenvolvidos algoritmos automáticos para segmentar tireoides e classificar com precisão se os nódulos presentes nela são malignos ou benignos em um dataset com mais de 4 mil imagens de ultrassom da tireoide.

## II. TRABALHOS RELACIONADOS

Nesta seção, são descritos alguns trabalhos encontrados que ajudaram no desenvolvimento deste trabalho. A subseção II-A descreve as revisões sistemáticas encontradas e a subseção II-B descreve os trabalhos de segmentação e de definição de algoritmos que foram usados neste trabalho.

### A. Estudos Secundários

Um dos estudos encontrados foi a revisão sistemática sobre métodos para análise de imagens de US com *Deep Learning* a fim de identificar e classificar doenças a partir das imagens de US. Os estudos encontrados por essa revisão descrevem algoritmos que detectam, segmentam, registram ou classificam estruturas biológicas do corpo humano. No que se refere à segmentação, a revisão descreve segmentações de próstata, ventrículo esquerdo do coração, líquido amniótico de gestantes, nódulos de tireoide, segmentos de nervo, nódulos linfáticos, endométrio, entre outros [4]. Nessa revisão, foram encontrados dois estudos que descrevem o desenvolvimento de algoritmos para segmentação de nódulos da tireoide.

A segunda revisão sistemática encontrada descreve diversos métodos de segmentação de tireoide e de nódulos de tireoide com uso de imagens de US. Para segmentação de tireoide, a revisão descreve 13 métodos, sendo que 4 segmentam a tireoide levando em consideração apenas o contorno e o formato da glândula; 2 estudos que usam a região em que a glândula está posicionada; e 7 métodos que usam *Machine Learning* e *Deep Learning* [5].

Para a segmentação de nódulos de tireoide, foram encontrados 17 estudos. Destes, 8 estudos usam apenas o contorno e o formato dos nódulos para segmentá-los; 2 estudos usam métodos baseados na região em que os nódulos são encontrados; 5 estudos usam *Machine Learning* ou *Deep Learning*;

e 2 estudos usam uma abordagem híbrida para identificar os nódulos [5].

### B. Estudos Primários

Dentre os estudos primários encontrados, foram encontradas propostas de redes neurais específicas para segmentação de imagens médicas (subsubseção II-B1); estudos que descrevem em específico a segmentação de imagens de tireoide ou de nódulos de tireoide (subsubseção II-B2); e estudos que descrevem uma melhora na acurácia ao alterar parâmetros de ativação, normalização e tamanho dos *batches* no treinamento das CNNs (subsubseção II-B3).

#### 1) Redes Neurais para segmentação de imagens médicas:

Dentre os estudos que utilizavam abordagem de rede neural convolucional (CNN), mais de um utilizou como base a U-Net. A U-Net é uma CNN desenvolvida especialmente para segmentação de imagens biomédicas pelo departamento de Ciência da Computação da Universidade de Freiburg, na Alemanha. É uma rede do tipo *Fully-Convolutional*. Como vantagens principais estão a necessidade de poucas imagens para treinamento e a velocidade para treinamento e teste [6].

Com base na U-Net foram desenvolvidas também as CNNs V-Net, que é uma especialização da U-Net para convolução volumétrica, utilizada especialmente em conjunto com imagens de ressonância magnética (MRI) [7]. Um dos estudos encontrados também utilizou uma especialização da U-Net chamada VGG16+UNet, uma CNN que utiliza a arquitetura da U-Net, mas com as camadas de convolução de uma VGG-16. Quando comparada com U-Net, a DC-UNet apresenta melhora que varia entre 1,5% e 11% na segmentação de imagens [8].

2) *Segmentação de tireoide ou de nódulos de tireoide:* Dos estudos encontrados, um utilizou uma CNN de função de base radial para classificar blocos convolucionais que são ou não da glândula e aplicando um método para completar a região da tireoide [9]. Outro, utilizou uma rede neural completamente diferente, também do tipo *Fully-Convolutional* para segmentar as imagens de tireoide. Entretanto, apenas algumas camadas da CNN são descritas, revelando poucos detalhes sobre a abordagem [10].

Outro trabalho compara o uso de redes bayesianas, KNN e SVN para classificação de nódulos em uma dada tireoide. Neste estudo, é descrito que a rede SVN apresenta melhor acurácia (84,52%) que redes KNN (46,15%) e Bayesianas (38,46%) [11].

3) *Alterações em redes neurais:* Alguns dos estudos encontrados se empenharam em alterar alguns parâmetros de diferentes redes neurais a fim de melhorar a acurácia delas. A normalização dos *batches* no treinamento melhorou o desempenho dos modelos *GoogLeNet (Inception)* na acurácia e no tempo do treinamento [12]. Com o mesmo objetivo, outro estudo mostra que é possível alterar a função de ativação para "*Scaled Exponential Linear Units*" (SELU) para alcançar resultados semelhantes [13].

## III. METODOLOGIA

O desenvolvimento do algoritmo de segmentação se dividiu em 5 etapas: busca por desafios semelhantes para observar

as ferramentas utilizadas; busca por artigos científicos que descrevessem o desenvolvimento de segmentação de tireoides; reprodução dos métodos encontrados; modificação dos métodos encontrados a fim de melhorar os resultados obtidos; registro dos resultados finais e escrita deste relatório.

A subseção III-A descreve como o *dataset* é obtido e como ele se constitui; a subseção III-B descreve quais os pré-processamentos realizados nas imagens de ultrassom; a subseção III-C descreve as métricas utilizadas para aferir a acurácia e *loss*; a subseção III-D descreve quais algoritmos foram usados para fazer a segmentação de tireoide; a subseção III-E descreve quais algoritmos foram usados para classificar a tireoide; a subseção III-F descreve o sistema em que os algoritmos foram desenvolvidos, treinados e testados; por fim, algumas configurações que foram utilizados em todos os treinamentos (subsubseção III-F1).

### A. Dataset

O desafio escolhido fornece dois *datasets* para o desafio: um para treinamento e um para teste. O *dataset* de treinamento consiste em 3644 imagens de ultrassom de tireoide, 3644 máscaras indicando onde a tireoide se encontra e um arquivo CSV indicando quais imagens possuem tireoides saudáveis e doentes.

As imagens do *dataset* foram obtidas a partir de equipamentos diferentes, em posições diferentes e em distâncias diferentes da glândula. As imagens também possuem resoluções diferentes, variando tanto o valor das resoluções quanto o aspecto das imagens. Além disso, algumas imagens de tireoide possuem anotações da máquina de ultrassom e outras não.

O *dataset* de teste é constituído por 910 imagens de US, com as mesmas características de heterogeneidade de resolução, aspecto e equipamento de obtenção. O *dataset* de teste não possui imagens de anotação e nem identificação sobre quais imagens mostram tireoides doentes e saudáveis.

### B. Pré-Processamento do Dataset

Para que seja possível utilizar as redes neurais, é necessário fazer com que todas as imagens tenham o mesmo tamanho. Para isso, é utilizada a função *resize* da biblioteca *skimage*.

Algumas imagens de US do *dataset* têm baixo contraste, como é o exemplo mostrado na Figura 2. Para melhorar um pouco este problema, é utilizado como pré-processamento, o processo de equalização de histograma conhecido como *Contrast-Limited Adaptive Histogram Equalization* (CLAHE) [14]. Foi escolhido por ter sido recomendado em outros projetos que processam imagens de US com baixo contraste.

### C. Métricas

Para aferir a assertividade do algoritmo de segmentação, os organizadores do desafio utilizaram a métrica de *Jaccard Score* (ou *Mean Intersection over Union*) para a segmentação da tireoide. A métrica *Mean Intersection over Union* (IoU) é uma métrica muito utilizada na segmentação de imagens. Ela mede a intersecção entre a área calculada pelo algoritmo

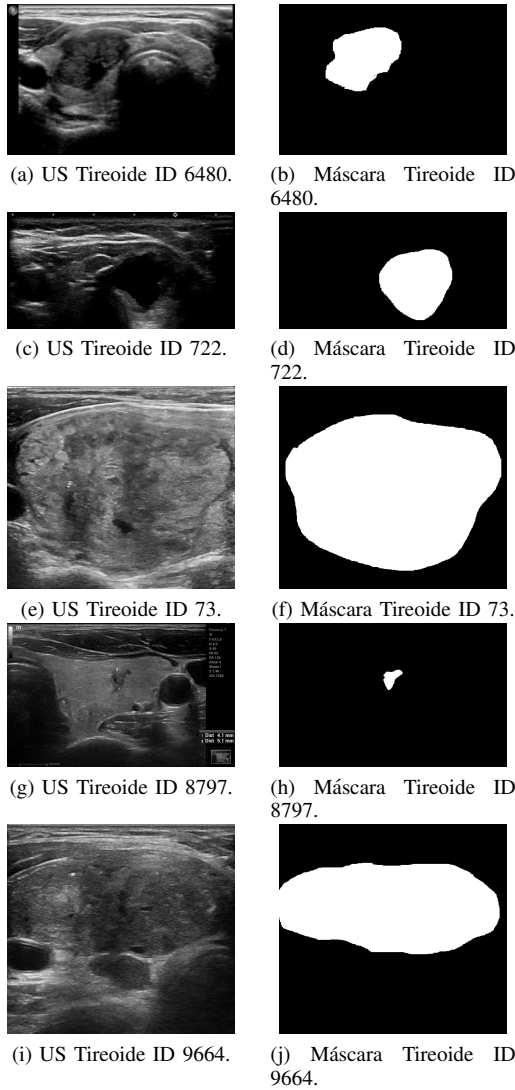


Figura 1. Exemplos de US de tireoide e suas respectivas máscaras do *dataset* de treinamento.

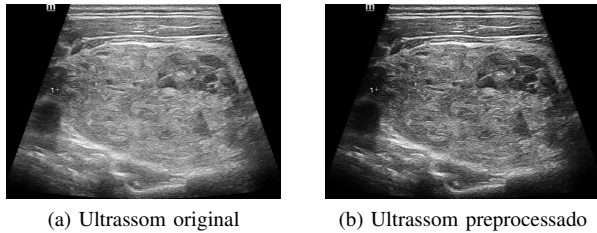


Figura 2. Exemplo de ultrassom em baixo contraste antes e depois de pré-processado com CLAHE.

de segmentação e a área da máscara assumida como *ground-truth*. A IoU pode ser definida pela equação abaixo, onde A é a máscara *ground-truth* e B é a máscara segmentada pelo algoritmo.

$$IoU(A, B) = \frac{A \cap B}{A \cup B}$$

Para aferir a assertividade do algoritmo de classificação, os

organizadores do desafio utilizaram a métrica de *F1 Score*. A *F1 Score*, expressa pela equação abaixo, pode ser interpretada como a média ponderada da precisão do algoritmo. Na equação, p é a precisão do algoritmo e r o *recall*.

$$F1 = 2 \cdot \frac{p \cdot r}{p + r}$$

Para manter coerência com o desafio, as mesmas métricas foram utilizadas para o treinamento e testes dos algoritmos de segmentação e classificação.

#### D. Segmentação

Foram buscados desafios anteriores que envolvessem a segmentação de algum objeto ou estrutura biológica para usar como ponto de partida para a segmentação da tireoide. Com isso, foram encontrados, entre outros, os estudos supracitados na seção II. Para este trabalho, foram utilizadas apenas CNNs criadas com a finalidade de segmentar imagens médicas, por serem modelos mais apropriados para o desafio enfrentado. Os seus respectivos desempenhos na acurácia foram comparados a fim de que seja possível escolher a melhor abordagem para o desafio.

Os algoritmos de segmentação das imagens de ultrassom foram produzidos usando os modelo U-Net [6] e outros modelos inspirados nele que obtiveram bons resultados em outros estudos [8], [15]–[17]. Estes estudos propõem alterações nas camadas; nos métodos de ativação; e de cálculo de *loss* e acurácia dessas redes neurais para proporcionar desempenho melhor nas segmentações. Alguns desses estudos utilizaram a ideia principal da U-Net, que é sua arquitetura, e substituíram as camadas originais por camadas de outras CNNs, como Resnet, e VGG16 [8], [17].

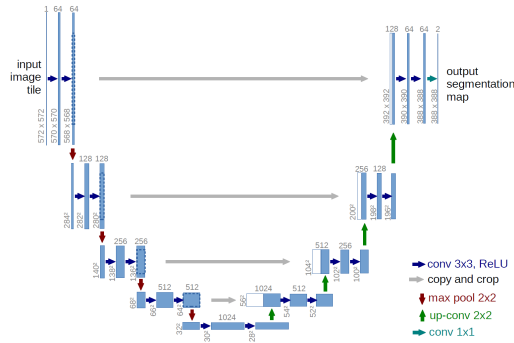
#### E. Classificação

Para classificar se a tireoide está ou não doente, foi empenhada uma metodologia semelhante a da segmentação. As CNNs VGG16 e VGG19 [18]; MobileNet e MobileNet-v2 [19], [20]; InceptionV3 e InceptionResNetV2 (ou InceptionV4) [21], [22]; e ResNet101V2 [23] foram utilizadas e suas acurácias foram comparadas com objetivo de aferir qual CNN tem melhor desempenho para esse desafio.

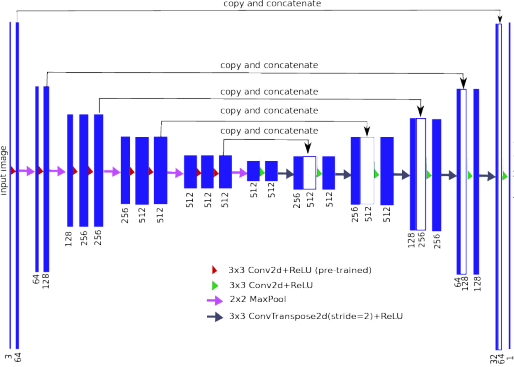
As implementações utilizadas das CNNs foram as inclusas na biblioteca Tensorflow. O Tensorflow possibilita a utilização de uma série de modelos de redes neurais com e sem pesos pré-treinados [24]. Para este trabalho, não foram utilizados pesos pré-treinados.

#### F. Sistema de Execução

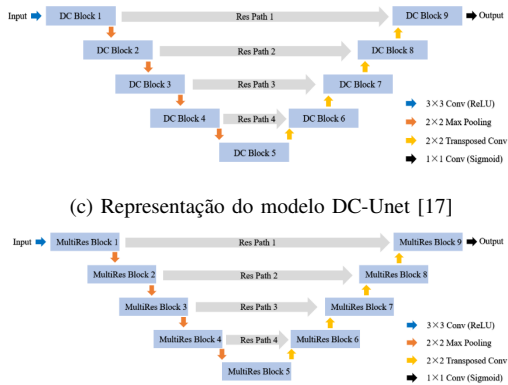
O sistema utilizado para desenvolvimento, treinamento e teste dos algoritmos supracitados está descrito na Tabela I. Como biblioteca para Redes Neurais e *Machine Learning* foi utilizado *Tensorflow2* por possuir comunidade ativa; documentação fácil de encontrar; oferecer todos os modelos utilizados na classificação já implementados, fazendo com que seja necessária apenas a preparação do dataset; e ter suporte para aceleração via GPU.



(a) Representação do modelo U-Net [6]



(b) Representação do modelo VGG16+UNet [8]



(c) Representação do modelo DC-UNet [17]

(d) Representação do modelo Multires-UNet [17]

Figura 3. Representação da arquitetura e camadas de alguns dos modelos utilizados.

Sistema Operacional	Ubuntu 20.02 LTS 64bits
Processador	AMD Ryzen 7 2700X @ 3.70GHz
Memória RAM	16GB DDR4 @ 2400MHz
Placa Mãe	Asus Prime B350-PLUS
Placa de Vídeo	GeForce GTX 1660 Super 6GB GDDR6
Python	Versão 3.8.5
Tensorflow	GPU, Versão 2.3.0

Tabela I

TABELA DE CONFIGURAÇÕES DO COMPUTADOR USADO PARA TREINAMENTO, TESTES E DESENVOLVIMENTO.

1) *Treinamento das redes*: Para treinamento das redes neurais, as seguintes configurações serão adotadas em todos os treinamentos:

- **Parada prematura**: parâmetros que devem ser atingidos para que o treinamento seja encerrado antes da quantidade de *epochs* definido anteriormente.
  - Paciência: se, em 10 *epochs*, o *loss* ou acurácia não melhorarem.
- **Compilador**
  - Otimizador *Adam*.
- **Treinamento**
  - A cada *epoch*, a sequência de itens usados para treinar a rede é randomizada para evitar qualquer tipo de vício que a rede possa estar sendo submetida. Essa é uma prática aconselhada pelo desafio. Isso é feito usando argumento *shuffle=True* na definição do treinamento da rede neural.
  - 10% do *dataset* de treinamento é separado para validação da rede neural. A validação é executada a cada *epoch*, e configurada usando o argumento *validation\_split=0.1* na definição do treinamento da rede neural.
- **Armazenamento do modelo**
  - Salvo usando evento interno do *Tensorflow* que é lançado ao fim de cada *epoch*.
  - Salva apenas quando a acurácia ou *loss* melhorarem.
  - Extensão utilizada é *.h5*.

#### IV. RESULTADOS

Nesta seção, estão apresentados os resultados obtidos pelo treinamento das redes neurais com uso do *dataset* fornecido.

##### A. Segmentação

Para segmentação, foram utilizadas as redes neurais *U-Net*, *VGG16+UNet*, *DC-UNet* e *Multires-UNet*. As figuras 4 a 7 mostram os gráficos que descrevem acurácia (*Mean IoU*) e *loss* obtidas. A Figura 8 mostra a comparação entre as métricas dessas redes neurais.

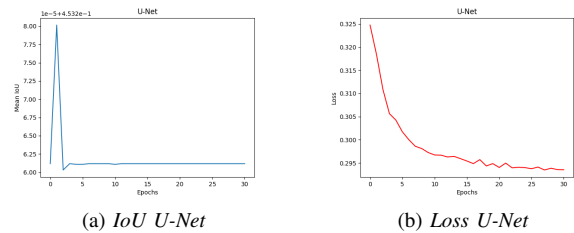
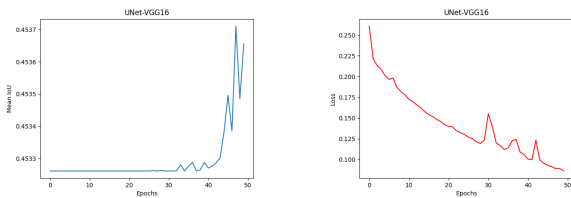


Figura 4. IoU e Loss usando U-Net.

É possível observar, portanto, que a rede neural UNet-VGG16 apresentou melhor desempenho em comparação com as demais. Entretanto, essa superioridade se reflete em aproximadamente 0,02%, diferença que pode ser desconsiderada.

##### B. Classificação

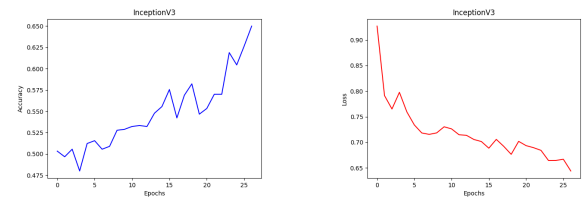
Para classificação, foram usados os modelos *InceptionResNetV2* (Figura 9); *InceptionV3* (Figura 10); *MobileNet* (Figura 11); *MobileNetV2* (Figura 12); *ResNet101V2* (Figura 13); *VGG16* (Figura 14); e *VGG19* (Figura 15). A Figura 16



(a) IoU *UNet-VGG16*

(b) Loss *UNet-VGG16*

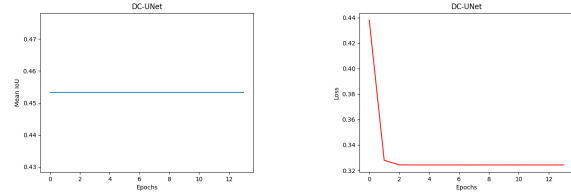
Figura 5. IoU e Loss usando *UNet-VGG16*.



(a) Acurácia *InceptionV3*

(b) Loss *InceptionV3*

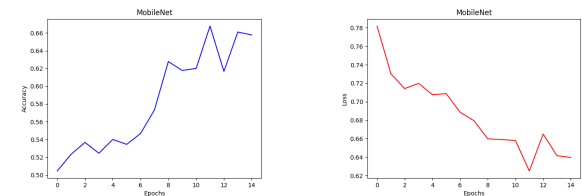
Figura 10. Acurácia e Loss usando *InceptionV3*.



(a) IoU *DC-UNet*

(b) Loss *DC-UNet*

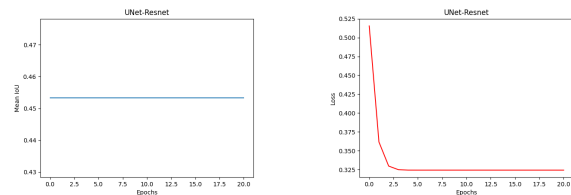
Figura 6. IoU e Loss usando *DC-UNet*.



(a) Acurácia *MobileNet*

(b) Loss *MobileNet*

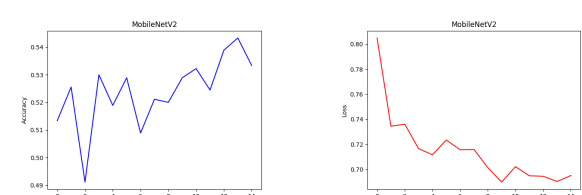
Figura 11. Acurácia e Loss usando *MobileNet*.



(a) IoU *UNet-ResNet*

(b) Loss *UNet-ResNet*

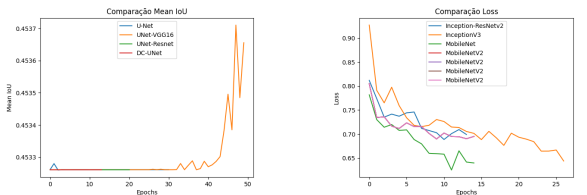
Figura 7. IoU e Loss usando *UNet-ResNet*.



(a) Acurácia *MobileNetV2*

(b) Loss *MobileNetV2*

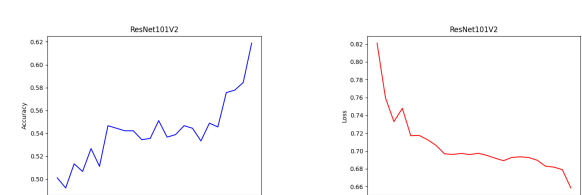
Figura 12. Acurácia e Loss usando *MobileNetV2*.



(a) Comparação IoU CNNs

(b) Comparação Loss CNNs

Figura 8. Comparação IoU e Loss.

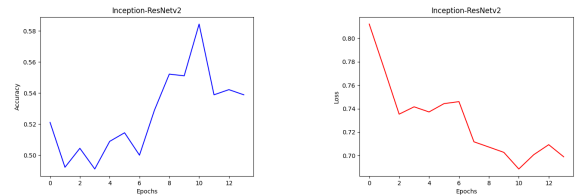


(a) Acurácia *ResNet101V2*

(b) Loss *ResNet101V2*

Figura 13. Acurácia e Loss usando *ResNet101V2*.

compara as métricas de Acurácia e *loss* entre os modelos utilizados.

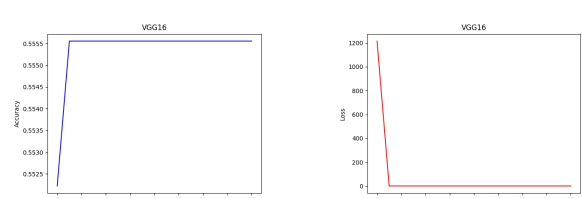


(a) Acurácia *InceptionResNetV2*

(b) Loss *InceptionResNetV2*

Figura 9. Acurácia e Loss usando *InceptionResNetV2*.

Neste caso, é possível observar que o modelo *MobileNet* se saiu melhor (65.786% de acurácia) seguido pelo modelo *InceptionV3*, com 64,99% de acurácia e o *ResNet101V2*, com



(a) Acurácia *VGG16*

(b) Loss *VGG16*

Figura 14. Acurácia e Loss usando *VGG16*.

61,89% de acurácia. Os dois primeiros têm menos de 1% de diferença, mas a distância do terceiro para o quarto (*VGG16* e *VGG19*, empatados com 55% de acurácia) é maior que 5%, mostrando uma diferença considerável entre os três primeiros e o restante dos modelos empregados para classificação.

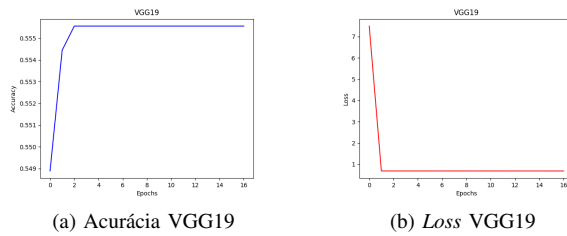


Figura 15. Acurácia e Loss usando VGG19.

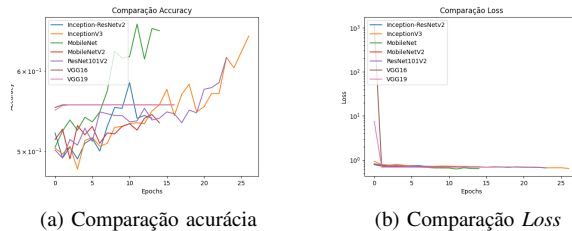


Figura 16. Acurácia e Loss na classificação.

## V. CONCLUSÃO

A Figura 1 mostra o quão heterogêneo é o *dataset* disponibilizado pelo desafio. As diferenças de resolução e proporção das capturas, bem como a posição de visualização da tireoide e os níveis completamente diferentes da doença fazem com que a tarefa de segmentar e classificar sejam difíceis para uma rede neural, pois as redes neurais se empenham em obter aspectos semelhantes nas imagens a fim de executar sua tarefa e um *dataset* heterogêneo como esse dificulta a segmentação da tireoide e a classificação dela.

Na segmentação, a diferença de 0,1% entre as abordagens utilizadas faz com que interpretemos o resultado obtido como empate entre os modelos utilizados. Já na tarefa de classificação, consideramos que os modelos *MobileNet* e *InceptionV3* estão empatados (com menos de 1% de diferença entre si) com o modelo *ResNet101V2* com uma acurácia aproximadamente 3% menor.

Como trabalhos futuros, faz-se interessante uma Revisão Sistemática relativa as técnicas de segmentação em *datasets* heterogêneos como o disponibilizado pelo desafio; e a aplicação das técnicas encontradas para as imagens de US do TN-SCUI2020.

## REFERÊNCIAS

- [1] T. E. of Encyclopaedia Britannica, *Thyroid gland*. Encyclopædia Britannica, inc., Aug 2019. [Online]. Available: <https://www.britannica.com/science/thyroid-gland>
- [2] C. E. Board, *Thyroid Cancer*. Cancer.net, Jan 2020. [Online]. Available: <https://www.cancer.net/cancer-types/thyroid-cancer/statistics>
- [3] H. Gireesha and S. Nanda, "Thyroid nodule segmentation and classification in ultrasound images," *International Journal of Engineering Research and Technology*, 2014.
- [4] S. Liu, Y. Wang, X. Yang, B. Lei, L. Liu, S. X. Li, D. Ni, and T. Wang, "Deep learning in medical ultrasound analysis: A review," *Engineering*, vol. 5, no. 2, pp. 261 – 275, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2095809918301887>
- [5] J. Chen, H. You, and K. Li, "A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasound images," *Computer Methods and Programs in Biomedicine*, vol. 185, p. 105329, 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169260719308454>
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015. [Online]. Available: <http://arxiv.org/abs/1505.04597>
- [7] F. Milletari, N. Navab, and S. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," *CoRR*, vol. abs/1606.04797, 2016. [Online]. Available: <http://arxiv.org/abs/1606.04797>
- [8] A. Lou, S. Guan, and M. Loew, "Dc-unet: Rethinking the u-net architecture with dual channel efficient cnn for medical images segmentation," *arXiv preprint arXiv:2006.00414*, 2020.
- [9] C.-Y. Chang, Y.-F. Lei, C.-H. Tseng, and S.-R. Shih, "Thyroid segmentation and volume estimation in ultrasound images," *IEEE transactions on biomedical engineering*, vol. 57, no. 6, pp. 1348–1357, 2010.
- [10] J. Ma, F. Wu, T. Jiang, Q. Zhao, and D. Kong, "Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks," *International journal of computer assisted radiology and surgery*, vol. 12, no. 11, pp. 1895–1910, 2017.
- [11] N. Singh and A. Jindal, "A segmentation method and comparison of classification methods for thyroid ultrasound images," *International Journal of Computer Applications*, vol. 50, no. 11, pp. 43–49, 2012.
- [12] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-normalizing neural networks," in *Advances in neural information processing systems*, 2017, pp. 971–980.
- [13] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.
- [14] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355 – 368, 1987. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0734189X8780186X>
- [15] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," *CoRR*, vol. abs/1807.10165, 2018. [Online]. Available: <http://arxiv.org/abs/1807.10165>
- [16] —, "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation," 2019.
- [17] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation," *CoRR*, vol. abs/1802.06955, 2018. [Online]. Available: <http://arxiv.org/abs/1802.06955>
- [18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [19] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *CoRR*, vol. abs/1704.04861, 2017. [Online]. Available: <http://arxiv.org/abs/1704.04861>
- [20] M. Sandler, A. G. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation," *CoRR*, vol. abs/1801.04381, 2018. [Online]. Available: <http://arxiv.org/abs/1801.04381>
- [21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *CoRR*, vol. abs/1512.00567, 2015. [Online]. Available: <http://arxiv.org/abs/1512.00567>
- [22] C. Szegedy, S. Ioffe, and V. Vanhoucke, "Inception-v4, inception-resnet and the impact of residual connections on learning," *CoRR*, vol. abs/1602.07261, 2016. [Online]. Available: <http://arxiv.org/abs/1602.07261>
- [23] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," *CoRR*, vol. abs/1603.05027, 2016. [Online]. Available: <http://arxiv.org/abs/1603.05027>
- [24] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals,

P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng,  
“TensorFlow: Large-scale machine learning on heterogeneous systems,”  
2015, software available from tensorflow.org. [Online]. Available:  
<https://www.tensorflow.org/>