

# Análise comparativa de redes para segmentação semântica

Pedro Lima  
Centro de Informática  
Universidade Federal de Pernambuco  
Recife, Brasil  
pbsl@cin.ufpe.br

Gabriel Vasconcelos  
Centro de Informática  
Universidade Federal de Pernambuco  
Recife, Brasil  
ghv@cin.ufpe.br

**Resumo**—Este documento propõe uma análise comparativa da aplicação de três redes convolucionais para a segmentação semântica, aplicadas ao dataset de ultrasonografias de câncer de mama. A proposta é aprofundar o entendimento das arquiteturas U-net, Residual U-net e FastFCN, em relação ao funcionamento geral delas e também nos resultados que obtivermos da aplicação em um problema real.

**Palavras-chave**—segmentação, imagem, convolucional

## I. INTRODUÇÃO

O câncer de mama é o tipo mais comum de câncer entre as mulheres, sendo o tipo de câncer com mais diagnósticos, sendo estimados mais de 2,5 milhões diagnósticos em mulheres, além responsável pela morte de aproximadamente 685 mil mulheres, tudo isso apenas no ano de 2020 [1]. Com o grande e rapidamente crescente número de diagnósticos e fatalidades, a necessidade por diagnósticos rápidos, eficientes e confiáveis também vêm aumentando; e uma possibilidade de tipos de diagnósticos que se encaixam nesses critérios são diagnósticos por meio de redes neurais artificiais. Com o campo da visão computacional apresentando cada vez mais modelos, se torna difícil acompanhar quais são os métodos mais precisos, ou com maior custo-benefício, para a detecção, classificação e segmentação de agentes patogênicos no corpo humano. Por meio da análise apresentada neste relatório, serão apresentadas as comparações entre os resultados da segmentação levando em consideração o custo computacional.

## II. OBJETIVO

O objetivo dessa análise é realizar uma comparação entre três modelos de aprendizagem profunda, voltadas para problemas de segmentação diferentes, a U-Net, a Residual U-Net e a FastFCN. A comparação com base em suas previsões e custo computacional será feita para analisar qual a viabilidade de cada uma em um cenário de diagnóstico médico.

## III. JUSTIFICATIVA

A motivação para o projeto foi inspirada pelo crescente número de pessoas diagnosticadas com câncer de mama, e consequentemente com a crescente demanda por diagnósticos mais rápidos, baratos, confiáveis e efetivos; que tem como uma opção o uso das redes neurais convolucionais. Com isso, também é fato de que o número de arquiteturas de

redes neurais tem crescido; o que implica a necessidade à busca por quais arquiteturas são eficientes, na segmentação, e classificação de células cancerígenas, assim como em seu treinamento e em seu gasto de recursos.

## IV. METODOLOGIA

O dataset escolhido foi o "Breast Ultrasound Images Dataset" [2] encontrado no Kaggle. Ele é um conjunto que possui imagens de ultrassom mamário de cerca de 600 mulheres, entre seus 25 e 75 anos. As ultrasonografias foram coletadas durante o ano de 2018, para o treinamento de modelos de *machine learning*, em tarefas de classificação e segmentação de tumores mamários. O conjunto de dados consiste em 780 imagens, no formato PNG com um tamanho médio de imagem de  $500 \times 500$  pixels. As imagens são categorizadas em três classes, que são normal, *benign* (benigno) e *malignant* (maligno). As 3 categorias são:

- Normal: Células normais mamárias do corpo humano.
- Benign: Classe representativa de tumores cancerígenos, porém benignas ao corpo humano.
- Malignant: Câncer maléfico que se espalha para outros sistemas do corpo.

### A. U-Net

A U-Net é uma rede neural totalmente convolucional (*Fully Convolutional Network* ou FCN) que foi projetada originalmente para segmentação de imagens da área da biomedicina [3]. Ela é constituída por três partes, o *encoder*, o *bottleneck* e o *decoder*, além dos *skips* que ligam o *encoder* ao *decoder*, como é mostrado na Fig. 2.

### B. Residual U-Net (ResU-net)

Uma variação da rede U-net. Enquanto o modelo original possui *skip connections*, conectando blocos do *encoder* com os blocos do *decoder*, na ResU-net todos os blocos de ambos *encoder*, *decoder* também são blocos residuais. Apesar do posicionamento exato das conexões residuais variar entre implementações do modelo, elas sempre conectam o início, ou o meio, do bloco convolucional-residual ao seu fim. Com blocos do *encoder*, a adição entre tensores ocorre logo antes da camada de *max-pooling*, para a diminuição da resolução da imagem. Já com blocos do *decoder*, a conexão residual sempre

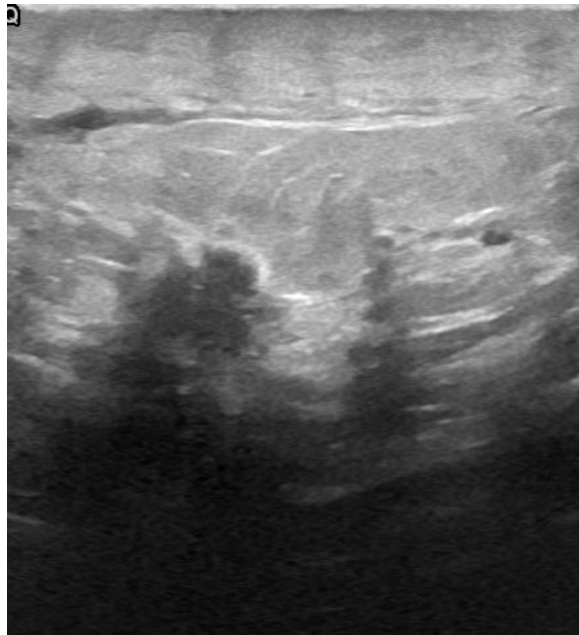


Figura 1. Exemplo de imagem com tumor maligno

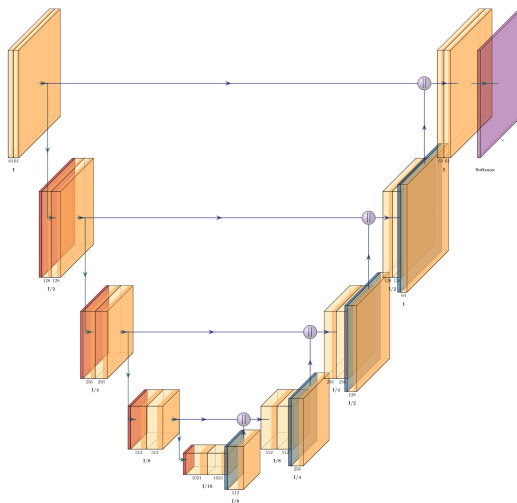


Figura 2. Representação da arquitetura U-net. Adquirida em <https://github.com/HarisIqbal88/PlotNeuralNet.git>

ocorre ligando a concatenação do resultado do *upsample* e a *skip connection*, ao resultado das convoluções do bloco, logo antes do próximo *upsample*.

A principal premissa das conexões residuais é a de permitir que as informações passem diretamente de uma camada convolucional para outra, evitando que possíveis informações importantes se percam ao longo das convoluções. Além disso, também ajudam a mitigar o problema do desaparecimento do gradiente e a proporcionar um treinamento mais rápido do modelo.

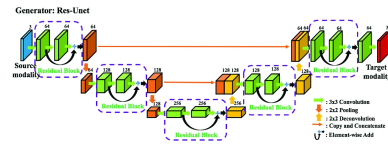


Figura 3. Exemplo de arquitetura Residual U-net. Adquirida em [https://link.springer.com/chapter/10.1007/978-3-030-00889-5\\_18](https://link.springer.com/chapter/10.1007/978-3-030-00889-5_18)

### C. FastFCN

A FastFCN é uma versão melhorada da FCN para problemas de visão computacional. Sua arquitetura permite reduzir o custo computacional em até 3 vezes sem que haja perda significativa de desempenho. Essa melhoria é realizada trocando as convoluções para extrair mapas de características de alta resolução, que ocorre na U-net por exemplo, por uma etapa Joint Pyramid Upsampling (JPU). A JPU recebe os últimos 3 mapas de características da camada de convolução downsampling, sendo concatenados e passando por novas convoluções com taxas de dilatação diferentes. Ao final desta etapa os resultados são concatenados novamente e passam por uma camada final de convolução, dessa forma atingindo uma performance melhor.

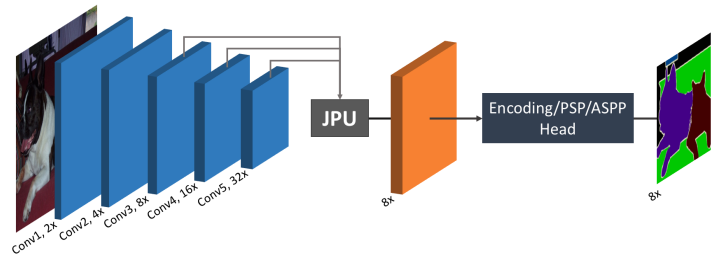


Figura 4. Representação da arquitetura fastfcn. Adquirida em <https://github.com/wuhuikai/FastFCN/blob/master/images/Framework.png>

### D. Análise comparativa

Como mencionado, o objetivo do projeto é analisar a eficiência dos modelos escolhidos comparativamente a fim de identificar pontos positivos e negativos na utilização de cada um deles. Dessa forma, definimos métricas de avaliação para que haja uma forma objetiva de realizar análises correlacionais alheias a subjetividade.

1) *Pixel Accuracy*: Calcula a proporção dos pixels corretamente classificados em relação ao total de pixels na tela. Métrica para avaliação geral do desempenho.

2) *Intersection over Union*: O Intersection over Union (IoU) mede a diferença entre a máscara de segmentação predita e a máscara esperada.

3) *Avaliação de observação*: Avaliação da qualidade da segmentação por meio da comparação com a imagem original. Dessa forma, será possível enxergar as falhas na segmentação do modelo avaliado, e assim entender as capacidades e limitações do reconhecimento de formas e objetos na cena.

## REFERÊNCIAS

- [1] “Breast Cancer - Metastatic: Statistics” - <https://www.cancer.net/cancer-types/breast-cancer-metastatic/statistics>.
- [2] AL-DHABYANI, Walid; GOMAA, Mohammed; KHALED, Hussien; FAHMY, Aly. Dataset of breast ultrasound images. Cairo, 21 nov. 2019. DOI: <https://doi.org/10.1016/j.dib.2019.104863>. Disponível em: <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset/>
- [3] R. Olaf et al., “U-Net: Convolutional Networks for Biomedical Image Segmentation”2015, doi: <https://doi.org/10.48550/arXiv.1505.04597>.
- [4] Wu, Huikai, et al. "Fastfcn: Rethinking dilated convolution in the backbone for semantic segmentation."arXiv preprint arXiv:1903.11816 (2019)..

## CRONOGRAMA

<b>Data</b>	<b>Descrição</b>
20/12	Entrega da proposta do projeto
23/12-30/01	Recesso Escolar
31/01-06/02	Entendimento e pré-processamento da base
07/02-13/02	Treinamento com as redes U-net, R2U-net e FastFCN
14/02-21/02	Ajustes e mudanças de parâmetros para ver o desempenho
21/02-28/02	Validação e considerações/mudanças finais
28/02-04/03	
04/02-11/03	