



ESCOLA
POLITÉCNICA

RELATÓRIO PARCIAL

**DEEP LEARNING PARA A SEGMENTAÇÃO DE ÓRGÃOS EM
IMAGENS MÉDICAS UTILIZANDO SELEÇÃO DINÂMICA E
AUMENTO DE DADOS**

CURITIBA

02/2023

Daniela dos Santos Lima
Ciência da Computação
Escola Politécnica

**DEEP LEARNING PARA A SEGMENTAÇÃO DE ÓRGÃOS EM
IMAGENS MÉDICAS UTILIZANDO SELEÇÃO DINÂMICA E
AUMENTO DE DADOS**

Relatório Parcial apresentado à
Pontifícia Universidade Católica do Paraná,
Programa Institucional de Bolsas de
Iniciação Científica.

Orientador: Prof. André Hochuli

CURITIBA

02/2023

SUMÁRIO

- 1 INTRODUÇÃO 4
- 2 OBJETIVO(S) 5
- 3 MATERIAIS E MÉTODOS 6
 - 3.1 DETALHES DE IMPLEMENTAÇÃO 6
 - 3.2 BANCO DE DADOS..... 7
 - 3.3 PRÉ-PROCESSAMENTO 7
 - 3.3.1 Conversão de formato 7
 - 3.4 AUMENTO DE DADOS..... 8
 - 3.5 DADOS DE TREINAMENTO, VALIDAÇÃO E TESTE 9
 - 3.6 ARQUITETURA 9
 - 3.6.1 U-Net 10
- 4 RESULTADOS PARCIAIS11
 - 4.1 TREINAMENTO E VALIDAÇÃO 11
 - 4.2 RESULTADOS VISUAIS..... 13
- 5 ETAPAS FUTURAS.....14
- 6 REFERÊNCIAS BIBLIOGRÁFICAS15

1 INTRODUÇÃO

O tratamento mais comum para o câncer é a radioterapia, porém, por se tratar do uso de radiação, os órgãos adjacentes ao órgão alvo ficam expostos, correndo o risco de terem seus tecidos saudáveis afetados. Para isso, são feitas, por um especialista, segmentações manuais em imagens do tumor no órgão alvo. Uma tarefa trabalhosa e que consome muito tempo, além de demandar de profissionais altamente qualificados. Por essas razões, uma abordagem automática pode ser essencial para melhorar e simplificar a segmentação dos órgãos de risco.

Neste projeto, nosso foco constitui-se em criar algoritmos de Deep Learning (aprendizado profundo) e selecionar, de forma dinâmica, aqueles que apresentam os melhores resultados na segmentação dos órgãos presentes nas imagens médicas. Essas imagens, por sua vez, apresentam órgãos torácicos, que consistem no coração, na traqueia, no esôfago e na aorta. Essas partes possuem diferentes tamanhos e formas. Para segmentar esses órgãos de forma tridimensional, utilizamos uma base de dados chamada SegTHOR, Segmentation of THoracic Organs at Risk, que contém imagens que usamos para ensinar e testar nosso algoritmo. No entanto, mesmo com o avanço tecnológico nos permitindo lidar com grandes volumes de dados para análises, ainda existe como obstáculo a insuficiência de dados para garantir uma aplicação confiável.

Portanto, com o propósito de solucionar este óbice, inspirando-se em técnicas do estado da arte em Machine Learning (aprendizado de máquina) e Deep Learning, tornamos a geração de diagnóstico automatizada, bem como tratamos do aumento da base de dados e sua variabilidade, gerando imagens artificiais e rotuladas. Isto é, segmentamos e diagnosticamos esses órgãos através de inteligência artificial e aprendizado de máquina, com o objetivo de tornar esse processo mais eficiente.

2 OBJETIVO(S)

O principal objetivo desse trabalho é desenvolver uma solução que auxilie e facilite o diagnóstico de um profissional da saúde ao visualizar uma tomografia computadorizada de um tórax com possíveis tumores, segmentando-as, e aferindo a porcentagem de determinado(s) órgão(os) afetado(s). Para se alcançar o resultado esperado, deve-se atingir as seguintes metas:

- Obter dados que representem a problemática apresentada;
- Revisar bibliografias para definir o modelo e arquitetura a serem usados como base que atenda os padrões do estado da arte;
- Propor e desenvolver melhorias no modelo de base;
- Realizar validações e testes no modelo proposto;
- Comparar resultados, mostrar limitações e propor possíveis melhorias.

3 MATERIAIS E MÉTODOS

Neste Capítulo são descritos os materiais e métodos utilizados para a solução do problema apresentado. A metodologia consiste no treinamento e validação de uma Rede Neural Convolucional para segmentação dos órgãos anteriormente citados a partir de tomografias torácicas. Para isto o experimento foi dividido em duas principais etapas: pré-processamento, aumento dos dados e concepção da arquitetura do modelo.

3.1 Detalhes de implementação

A rede neural convolucional e todo o pré-processamento dos dados foram implementados em Python, utilizando bibliotecas como TensorFlow, Keras, NumPy e OpenCV, no um sistema operacional Microsoft Windows 11.

Todas as execuções foram feitas em uma máquina com CPU Intel(R) Core(TM) i5- 12500H @ 2.50 GHz, GPU NVIDIA GeForce RTX 3050 Ti 4 GB VRAM e 16 GB 4800 MHz DDR5 de memória RAM. A fim de diminuir o tempo computacional de execução e otimizar as tarefas, fez-se o uso de processamento paralelo através da GPU, utilizando a biblioteca TensorFlow-GPU.

(Todas as execuções foram feitas na plataforma Google Colaboratory)

Todo código fonte necessário para reproduzir os experimentos pode ser encontrado em um repositório aberto na plataforma GitHub em (LIMA, 2022, Online).

3.2 Banco de Dados

As amostras utilizadas foram retiradas do dataset SegTHor [2], dedicado à segmentação de órgãos de risco no tórax.

Um total de 14.680 imagens foram utilizadas, sendo subdivididas igualmente entre 2 categorias: tomografias torácicas e máscaras com recorte de cada órgão.

3.3 Pré-processamento

O pré-processamento de dados é uma etapa integral do aprendizado de máquina, pois a qualidade dos dados e informações úteis que podem ser derivadas deles afetam diretamente a capacidade de aprendizado do modelo. Dessa maneira, é de suma importância trazê-los a um estado que o algoritmo possa facilmente analisá-los.

3.3.1 Conversão de formato

A primeira etapa do pré-processamento consiste em converter as imagens de entrada para um formato que se adéque aos tensores de entrada da rede convolucional. Inicialmente, as tomografias se encontram comprimidas no formato de arquivo Tagged Image File Format (.tiff).

As imagens foram transformados para o formato *numpy array* do tipo float64 e redimensionados para obtenção de imagens de 512 de largura por 512 de altura, apresentando apenas um canal de cor, contendo valores de intensidade de 0 a 255. Após a transformação, é possível visualizar as imagens, mostradas na Figura 1.

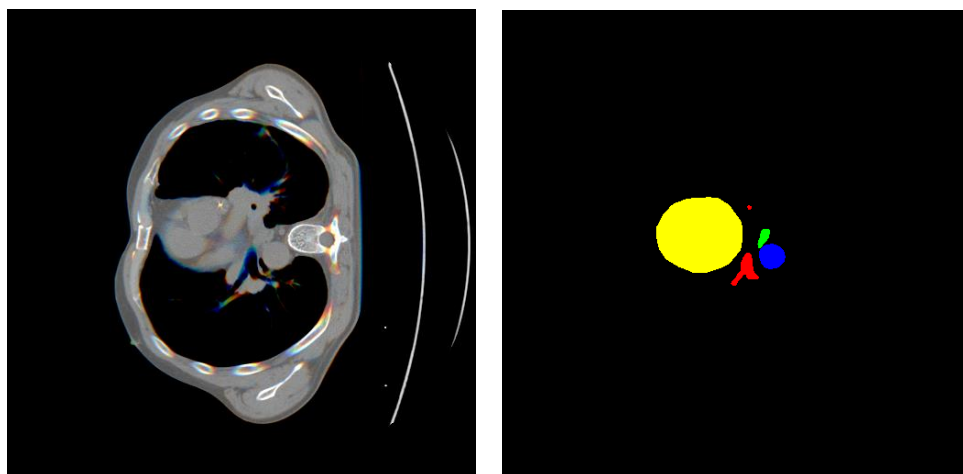


Figura 3.1: Dados de entrada: Tomografia e Máscara dos órgãos.

3.4 Aumento de dados

Data Augmentation, ou aumento de dados, é uma técnica utilizada para gerar novos exemplos de dados de treinamento com o objetivo de melhorar a generalização do modelo e reduzir o overfitting.

Uma rede neural convolucional possui a capacidade de classificar e segmentar dados de forma consistente, mesmo quando eles estão em orientações diferentes, devido à sua invariância em relação a escala, translação e outras transformações. No entanto, devido à grande quantidade de parâmetros que uma CNN utiliza, é necessário apresentar uma quantidade proporcionalmente grande de exemplos para obter uma boa generalização e desempenho do modelo. Como muitas vezes há uma baixa quantidade de amostras de entrada disponíveis em acervos públicos, é necessário recorrer ao aumento de dados para obter resultados satisfatórios.

Neste trabalho, a fim de aumentar a quantidade de dados, foram aplicadas as transformações que se seguem:

- Chance aleatória de virar a imagem verticalmente;
- Chance aleatória de virar a imagem horizontalmente;
- Chance aleatória de redimensionar o zoom da imagem, ampliando ou reduzindo por um fator de 30%;

- Chance aleatória de rotacionar em 40° para esquerda ou direita.

O resultado dessas transformações pode ser observado na Figura 3.2, onde para cada dado de entrada, foi gerado seu respectivo par transformado, dobrando a quantidade de exemplares de treinamento.

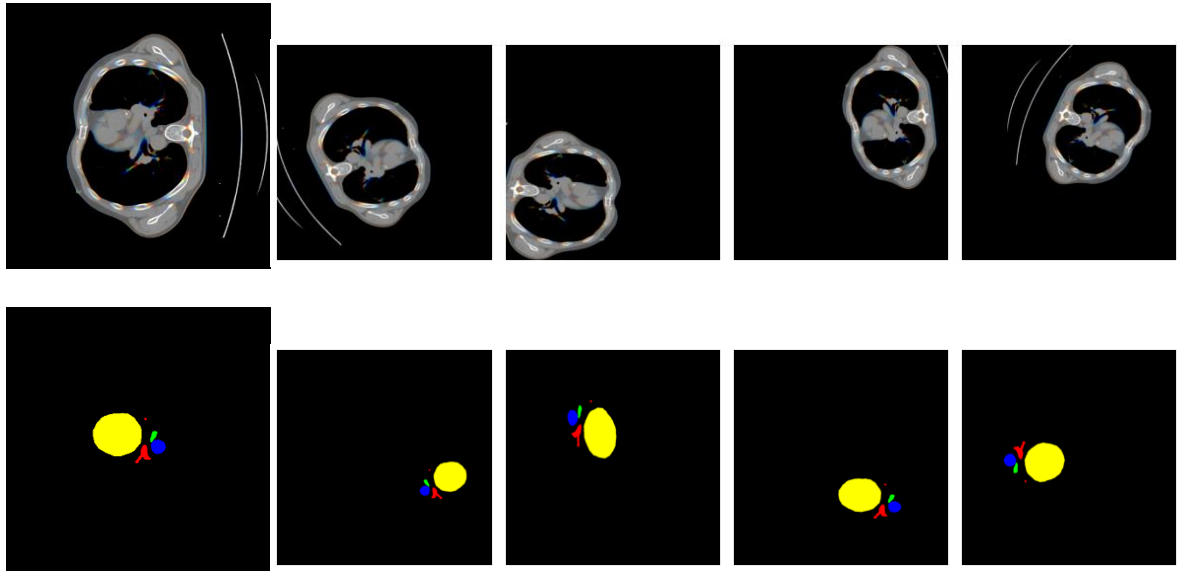


Figura 3.2: Resultado do aumento de dados aplicado em uma imagem.

3.5 Dados de treinamento, validação e teste

Com o aumento de dados, o número de imagens foi duplicado, resultando em 29.360 amostras. Logo em seguida, foram feitas separações aleatórias dos conjuntos em: 80% das imagens para treinamento, 10% para validação e 10% para teste.

3.6 Arquitetura

A rede neural convolucional escolhida como base para a realização dos experimentos foi a U-Net, devido a sua ampla utilização em tarefas de segmentação contendo baixa quantidades de dados de treinamento.

3.6.1 U-Net

Para ter sucesso no treinamento de redes profundas, é necessário ter dezenas de milhares de exemplos de treinamento. A Figura 3.3 apresenta a U-Net, uma rede e estratégia de treinamento que se destaca pelo uso intensivo de aumento de dados para tornar o melhor uso das amostras disponíveis. A U-Net é uma arquitetura derivada de uma Rede Neural Totalmente Convolutiva, mas que foi modificada e expandida para exigir poucas imagens de treinamento e, ao mesmo tempo, produzir segmentações mais precisas. (RONNEBERGER; FISCHER; BROX, 2015).

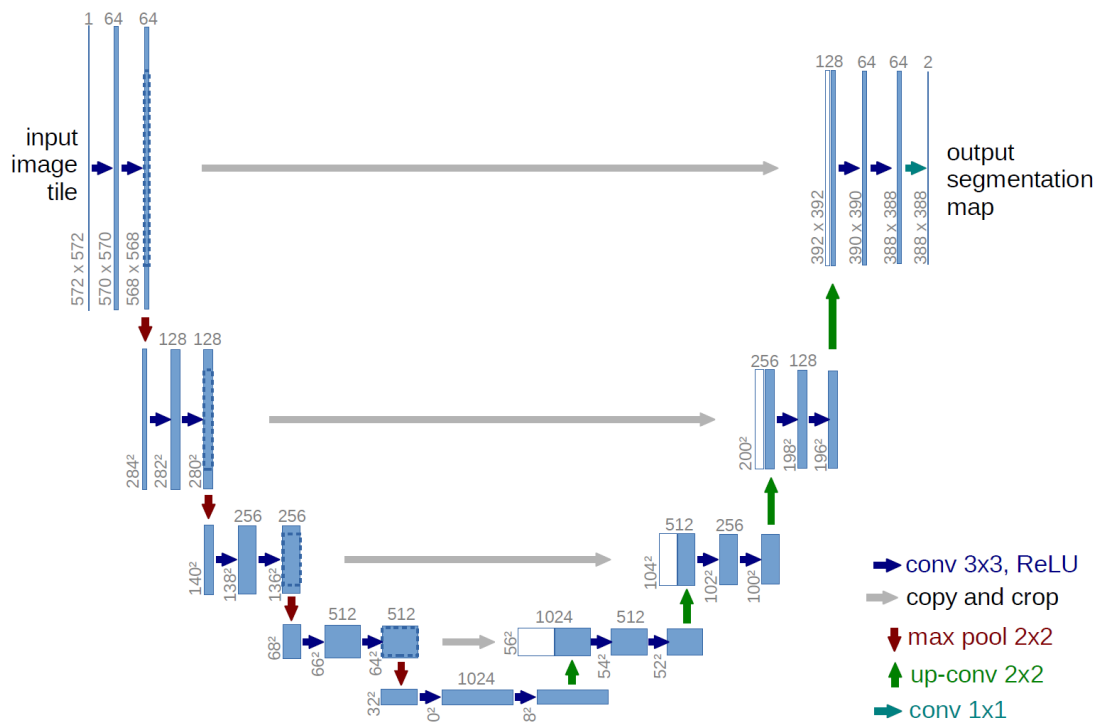


Figura 3.3: Arquitetura U-Net (RONNEBERGER; FISCHER; BROX, 2015)

4 RESULTADOS PARCIAIS

4.1 Treinamento e validação

Para realizar o treinamento da topologia, as etapas de pré-processamento, aumento de dados e separação do conjunto de dados foram executadas, conforme descrito no Capítulo 3. O treinamento foi realizado em 50 épocas e com um tamanho de lote de 16. Foram utilizados callbacks de redução da taxa de aprendizagem em platôs e parada precoce para evitar overfitting do modelo durante o treinamento.

Foi gerado um gráfico para cada época, demonstrando os resultados de acurácia e erro das segmentações pulmonares e das lesões para os conjuntos de treinamento e validação. Esses gráficos permitem uma análise qualitativa da influência da taxa de aprendizagem e do desempenho das topologias durante o treinamento.

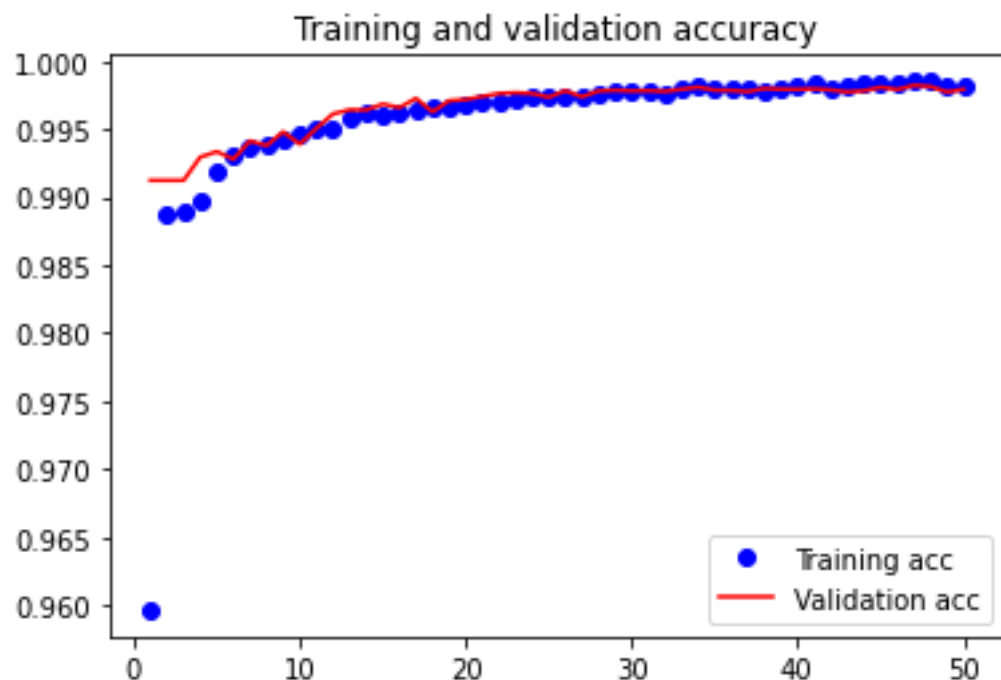


Figura 4.1: Gráfico de acurácia da rede U-Net

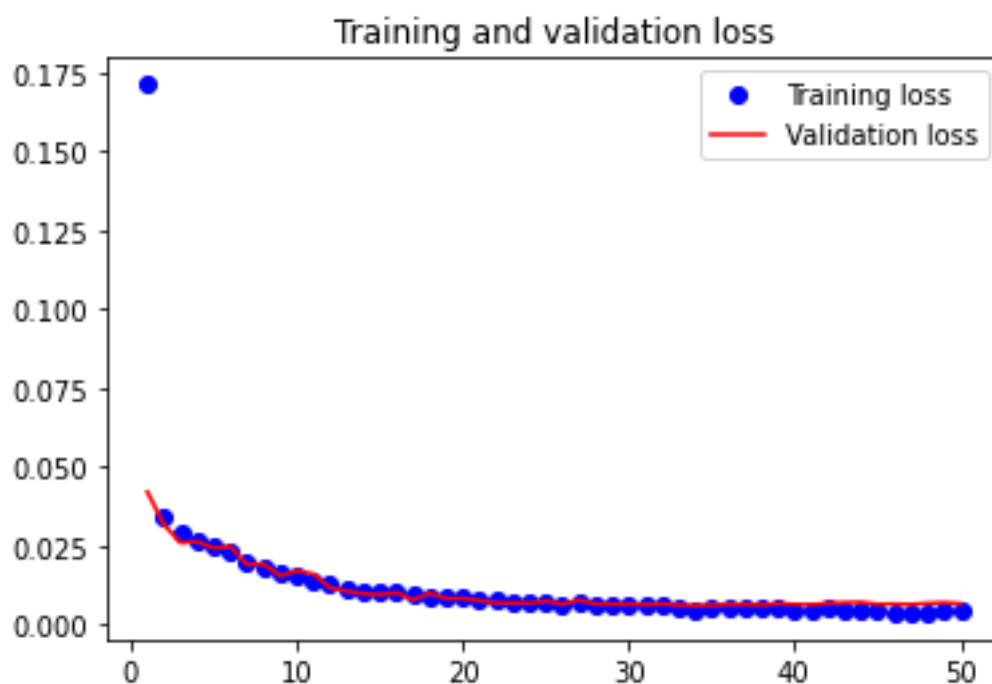


Figura 4.2: Gráfico de custo da rede U-Net

	Modelo
Acurácia	99.804%
Custo	00.056%
IoU Score (Jaccard Index)	99,495%
Dice Coefficient (F1 Score)	99,746%

Tabela 4.1: Resultados do conjunto de teste da U-Net.

Conforme o previsto, essa estrutura possui um desempenho satisfatório para situações de segmentação semântica em que a quantidade de dados de treinamento é limitada. Verifica-se que não houve excesso (overfitting) ou falta (underfitting) de ajuste do modelo, o que gerou uma boa capacidade de generalização e, consequentemente, um resultado adequado nas amostras de teste.

4.2 Resultados visuais

Com o objetivo de validar a precisão dos resultados de desempenho, foram criados exemplos de saída e comparados com as máscaras de treinamento correspondentes. A Figura 4.5 apresenta uma comparação entre as máscaras de treinamento e as geradas pelo modelo, permitindo observar a semelhança entre elas.

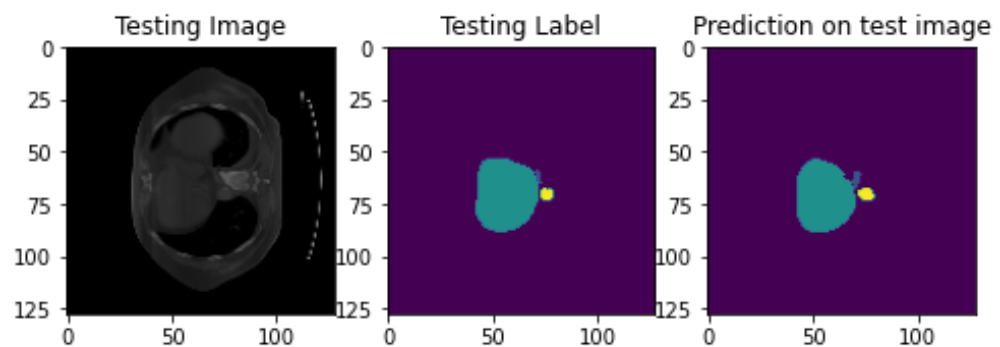


Figura 4.5: Saída prevista pelo modelo: Segmentação dos órgãos aparentes

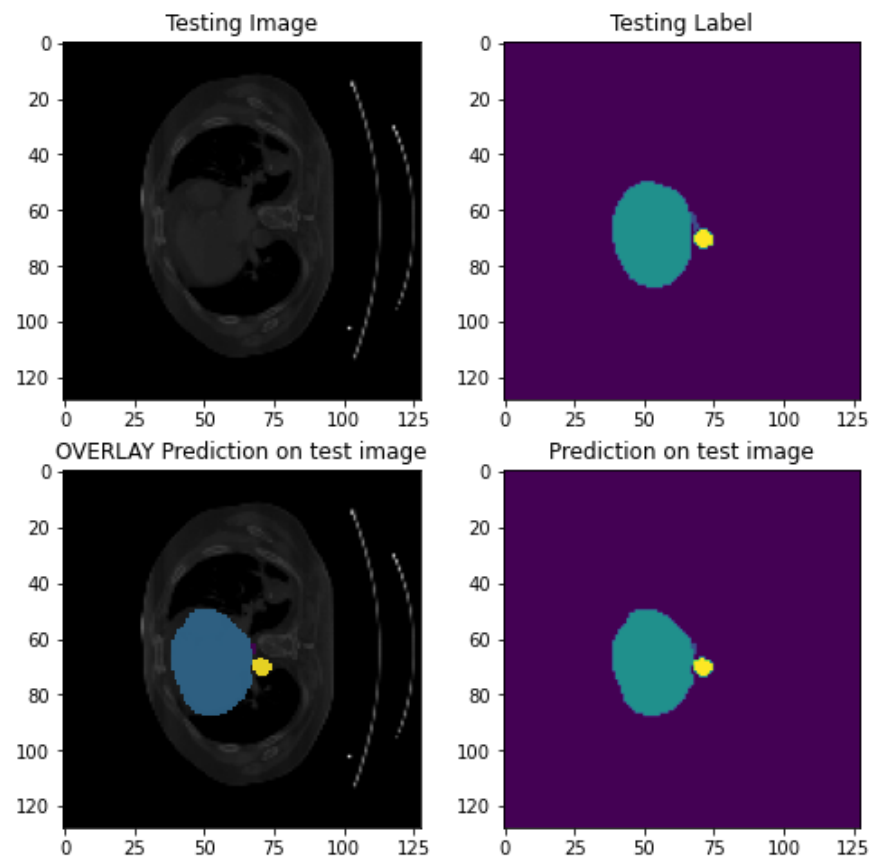


Figura 4.6: Sobreposição dos órgãos previstos com a tomografia original

5 ETAPAS FUTURAS

As etapas realizadas até agora foram relacionadas principalmente no aprendizado teórico/prático de aplicações de inteligência artificial para manipulação de imagens. Após a entrega deste relatório parcial, o objetivo é aplicar as teorias pesquisadas de forma prática e avaliar possíveis resultados.

Nossas etapas futuras serão:

- Realizar benchmark das nossas aplicações
- Exibir os resultados obtidos por nossa aplicação final

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 LIMA, Daniela. Source code GitHub repository. Disponível em: <<https://github.com/mrslima/pibic>>.
- 2 LAMBERT, Z., PETITJEAN, C., DUBRAY, B., & KUAN, S. (2020, November). **SegTHOR**: segmentation of thoracic organs at risk in CT images. In 2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1-6). IEEE.
- 3 RONNEBERGER, O.; FISCHER, P.; BROX, T. **U-net**: convolutional networks for biomedical image segmentation.