

Ciência da Computação
Escola Politécnica PUC-PR

Daniela dos Santos Lima

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

CURITIBA

02/2023

Ciência da Computação
Escola Politécnica PUC-PR

Daniela dos Santos Lima

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

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

Orientador: Prof. Dr. André Hochuli

CURITIBA

02/2023

SUMÁRIO

1	INTRODUÇÃO	4
1.1	OBJETIVO(S)	5
1.2	ESTRUTURA DO TRABALHO.....	6
2	FUNDAMENTAÇÃO TEÓRICA	7
2.1	INTELIGÊNCIA ARTIFICIAL	7
2.2	APRENDIZADO DE MÁQUINA.....	7
2.3	OVERFITTING E UNDERFITTING	8
2.4	REDES NEURAIS ARTIFICIAIS	8
2.4.1	<i>Neurônio Artificial</i>	9
2.4.2	<i>Soma Ponderada (combinação das entradas - O "Net"):</i>	10
2.4.3	<i>Função de Ativação</i>	10
2.5	REDE NEURAL MULTICAMADAS	12
2.6	DEEP LEARNING.....	12
2.7	REDES NEURAIS CONVOLUCIONAIS	14
2.7.1	<i>Camada de convolução</i>	14
2.7.2	<i>Camada de pooling</i>	16
2.7.3	<i>Camada totalmente conectada</i>	16
2.8	SEGMENTAÇÃO SEMÂNTICA	17
3	MATERIAIS E MÉTODOS	19
3.1	DETALHES DE IMPLEMENTAÇÃO	19
3.2	BASE DE DADOS	20
3.3	PRÉ-PROCESSAMENTO	20
3.3.1	<i>Conversão de formato</i>	20
3.4	AUMENTO DE DADOS.....	21
3.5	DADOS DE TREINAMENTO, VALIDAÇÃO E TESTE	22
3.6	ARQUITETURA	23
3.6.1	<i>U-Net</i>	23
4	RESULTADOS FINAIS.....	24
4.1	TREINAMENTO E VALIDAÇÃO	24
4.2	ANÁLISES VISUAIS	26
5	CONSIDERAÇÕES FINAIS.....	27
5.1	ETAPAS FUTURAS.....	28

1 INTRODUÇÃO

A segmentação semântica é uma tarefa fundamental na área de processamento de imagens médicas, com o objetivo de identificar e delimitar regiões de interesse específicas, como órgãos ou estruturas anatômicas. No contexto do tratamento do câncer, a segmentação precisa dos órgãos torácicos é essencial para a radioterapia, visando proteger os tecidos saudáveis adjacentes ao órgão alvo. No entanto, a segmentação manual realizada por especialistas é uma tarefa laboriosa e que consome muito tempo, requerendo profissionais altamente qualificados. Portanto, a automação desse processo é de suma importância para aprimorar e simplificar a segmentação dos órgãos de risco.

Na área da medicina, o avanço da computação tem desempenhado um papel crucial, permitindo análises de grandes volumes de dados e o desenvolvimento de terapias e dispositivos médicos inovadores. A tecnologia atual possibilita que médicos tenham acesso a informações em larga escala, o que auxilia na identificação de padrões e anomalias em doenças anteriormente difíceis de serem exploradas. Além disso, técnicas de inteligência artificial, como redes neurais, têm possibilitado prognósticos precisos sobre a evolução clínica, como pode ser observado em estudos de detecção de lesões relacionadas ao Coronavírus (SILVA, 2021).

Neste projeto, o foco é desenvolver algoritmos de Deep Learning (aprendizado profundo) e selecionar, de forma dinâmica, aqueles que apresentam os melhores resultados na segmentação dos órgãos presentes em imagens médicas. As imagens utilizadas são de órgãos torácicos, incluindo o coração, traqueia, esôfago e aorta. Essas estruturas apresentam tamanhos e formas diferentes, tornando a segmentação tridimensional um desafio. Para treinar e testar o algoritmo, utilizamos o conjunto de dados SegTHOR (Segmentation of Thoracic Organs at Risk), descrito em detalhes no artigo de referência (LAMBERT et al., 2020). No entanto, mesmo com os avanços tecnológicos que nos permitem lidar com grandes volumes de dados, ainda enfrentamos o obstáculo da insuficiência de dados para garantir uma aplicação confiável.

Com o objetivo de abordar essa limitação, inspiramo-nos em técnicas de ponta de aprendizado de máquina e deep learning, para automatizar a segmentação dos órgãos e lidar com a escassez de dados. Desenvolvemos um modelo que utiliza inteligência artificial e aprendizado de máquina para segmentar esses órgãos, e abordamos o problema do aumento da base de dados e sua variabilidade, gerando imagens artificiais e rotuladas. Dessa forma, buscamos tornar o processo de segmentação mais eficiente e confiável.

Com base nas abordagens atuais em segmentação semântica e no uso do dataset SegTHOR, espera-se que os classificadores especialistas apresentem resultados promissores na segmentação dos órgãos torácicos. A combinação de diferentes classificadores, por meio de técnicas de fusão de informações, pode levar a uma melhoria significativa na precisão e robustez da segmentação. Além disso, a geração de imagens artificiais e rotuladas para aumento de dados pode permitir uma melhor generalização do modelo, superando a limitação da escassez de amostras rotuladas. No entanto, também é possível que sejam identificadas algumas limitações e desafios, como dificuldades na segmentação de estruturas com formas complexas e variações anatômicas.

1.1 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.

1.2 Estrutura do trabalho

O Capítulo 1 apresenta uma contextualização da motivação e dos objetivos do trabalho, além de apresentar a estrutura do trabalho. O Capítulo 2 abrange toda a fundamentação teórica necessária para compreender a problemática e os métodos utilizados no trabalho, adotando uma abordagem top-down. Ele aborda conteúdos mais amplos, como Inteligência Artificial, e vai especificando ao longo do capítulo até chegar aos tópicos tratados neste trabalho, como Aprendizado Profundo e Redes Neurais Totalmente Convolucionais. O Capítulo 3 descreve os materiais, métodos e técnicas utilizados para possibilitar a reprodução do experimento, com base em aspectos teóricos e práticos. O Capítulo 4 apresenta os resultados, discussões e análises obtidos, explorando possíveis limitações e melhorias a serem realizadas. Por fim, o Capítulo 5 apresenta as conclusões do trabalho e considerações sobre possíveis trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

O propósito deste capítulo é apresentar algumas das áreas da Inteligência Artificial que são empregadas neste projeto. Começando com uma breve introdução ao Aprendizado de Máquinas, apresentando algumas técnicas e termos relevantes. Em seguida, serão expostos os fundamentos teóricos da Aprendizagem Profunda e das Redes Neurais Convolucionais, abordando seus principais métodos, características e benefícios.

2.1 Inteligência artificial

Inteligência artificial pode ser definida como "o ramo da ciência da computação que se ocupa do estudo e desenvolvimento de algoritmos e sistemas capazes de realizar tarefas que, até então, somente poderiam ser executadas por seres humanos, tais como reconhecimento de fala, processamento de imagens, tomada de decisões, entre outras" (RUSSELL; NORVIG, 2010).

2.2 Aprendizado de Máquina

A definição de Aprendizado de Máquina consiste em "um campo da ciência da computação que visa desenvolver algoritmos e técnicas para permitir que os computadores aprendam com dados, sem necessidade de programação explícita" (ALPAYDIN, 2010).

Com base em informações prévias, uma máquina pode simular determinado comportamento, buscando explorar dados úteis e maximizar uma medida de desempenho, visando o treinamento de um modelo eficiente e auxiliando na análise de informações futuras (JORDAN; MITCHELL, 2015).

2.3 Overfitting e Underfitting

O *overfitting* e *underfitting* são dois problemas que surgem frequentemente no campo de aprendizado de máquina, e que têm um impacto significativo na capacidade dos modelos de generalizar para novos dados. O *overfitting* acontece quando o modelo é excessivamente complexo e se ajusta demasiado aos dados de treinamento, perdendo assim a capacidade de generalizar para dados que não tenha visto anteriormente. O *underfitting*, por outro lado, ocorre quando o modelo é muito simples e não consegue capturar a complexidade dos dados de treinamento, e, portanto, não consegue ajustar-se aos dados de treinamento ou generalizar para novos dados. Para evitar esses problemas, é importante ajustar os parâmetros do modelo ou modificá-lo para torná-lo mais adequado para o conjunto de dados, e dividir os dados em conjuntos de treinamento e teste para avaliar o desempenho do modelo (BISHOP, 2006).

Dessa forma, é possível avaliar se o modelo está sofrendo de *overfitting* ou *underfitting*, e fazer as correções necessárias para melhorar seu desempenho. Além disso, é importante escolher cuidadosamente as variáveis independentes ou camadas ocultas em redes neurais para garantir que o modelo possa lidar com a complexidade dos dados de treinamento e ainda assim generalizar para novos dados. Uma vez que esses problemas são resolvidos, o modelo pode ser mais eficaz na previsão de dados não vistos anteriormente e na identificação de padrões úteis para o usuário final (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

2.4 Redes Neurais Artificiais

Uma Rede Neural Convolutacional (Convolutional Neural Network - CNN) é um tipo de rede neural profunda especialmente projetado para o processamento de dados com estrutura em forma de matriz, como imagens, sinais de áudio e vídeos. Essa arquitetura é composta por camadas convolucionais, de pooling e de classificação, e

foi desenvolvida para extrair características dos dados de entrada por meio de filtros convolucionais e reduzir sua dimensionalidade por meio de camadas de pooling.

As redes neurais convolucionais foram amplamente adotadas em aplicações de visão computacional, incluindo reconhecimento de imagem, detecção de objetos, segmentação semântica e classificação de vídeo (GOODFELLOW; BENGIO; COURVILLE, 2016).

2.4.1 Neurônio Artificial

Um neurônio artificial é uma unidade básica de processamento em uma rede neural artificial. Ele é inspirado no neurônio biológico encontrado no cérebro humano. Assim como um neurônio biológico recebe sinais de entrada de outros neurônios, um neurônio artificial, Figura 3.1, recebe sinais de entrada de outras unidades de processamento, e processa esses sinais para produzir uma saída.

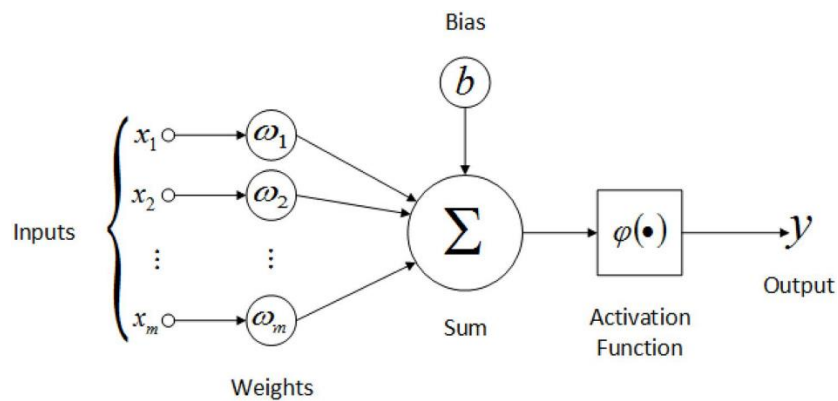


Figura 3.1: Neurônio Artificial (OLIVEIRA et al., 2017)

O funcionamento de um neurônio artificial é dividido em três partes principais: a soma ponderada, a função de ativação e a saída.

O processo de ajuste dos pesos sinápticos de um neurônio artificial é conhecido como treinamento. Durante o treinamento, a rede neural é exposta a um conjunto de exemplos de entrada e saída conhecidos, e os pesos sinápticos são ajustados de

forma a minimizar a diferença entre a saída produzida pelo neurônio e a saída desejada.

2.4.2 Soma Ponderada (combinação das entradas - O "Net"):

Na soma ponderada, Figura 3.2, cada entrada é multiplicada por um peso sináptico correspondente e os resultados são somados, dando origem ao chamado "net" de um neurônio.

$$net_i(t) = \sum_{j=1}^n w_{ij} x_j(t)$$

Figura 3.2: Somatório das Entradas - O "Net" (ZAMBIASI, 2017)

w_{ij} é um número real que representa a conexão sináptica da entrada **j**-ésimo neurônio. A conexão sináptica é conhecida como excitatória se **$w_{ij} > 0$** ou inibitória caso **$w_{ij} < 0$** .

A função de ativação é então aplicada à soma ponderada para produzir a saída do neurônio.

2.4.3 Função de Ativação

Existem várias funções de ativação diferentes que podem ser usadas em um neurônio artificial, incluindo a função sigmoide e a função ReLU.

A escolha da função de ativação depende do problema em questão e das características dos dados.

As funções de ativação são funções não-lineares conectadas ao final da estrutura de um neurônio artificial, definem a saída com base nos dados de entrada e o limiar de ativação.

A função sigmoide, Figura 3.3, é muito utilizada em problemas de classificação, onde tem como saída valores entre 0 e 1, e resulta na probabilidade de os dados de entrada estarem contidos na classe analisada.

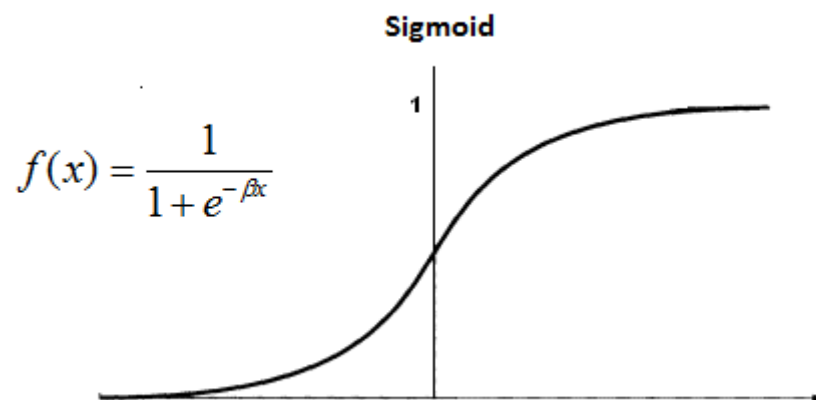


Figura 3.3: Sigmoide

A função ReLU, Figura 3.4, é uma função de ativação mais eficiente, pois não faz uso de expoentes, por este motivo é comumente utilizada em camadas ocultas.

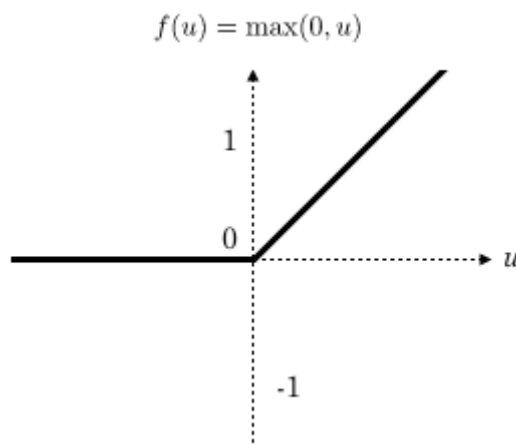


Figura 3.4: ReLU

2.5 Rede Neural Multicamadas

Redes Neurais Artificiais são modelos computacionais inspirados no sistema nervoso central humano. As redes neurais multicamadas (ou MLP, do inglês Multi-Layer Perceptron), Figura 3.5, são um tipo de rede neural que consiste em uma ou mais camadas de neurônios artificiais. Cada camada recebe entradas, realiza uma transformação linear das entradas por meio de um conjunto de pesos sinápticos e aplica uma função de ativação não linear aos resultados para gerar saídas. A saída da camada anterior é a entrada da camada seguinte, e assim por diante, até que a saída final da rede seja produzida. A rede MLP é treinada para aprender a relação entre as entradas e as saídas por meio do ajuste iterativo dos pesos sinápticos com base em um conjunto de dados de treinamento (BISHOP, 1995).

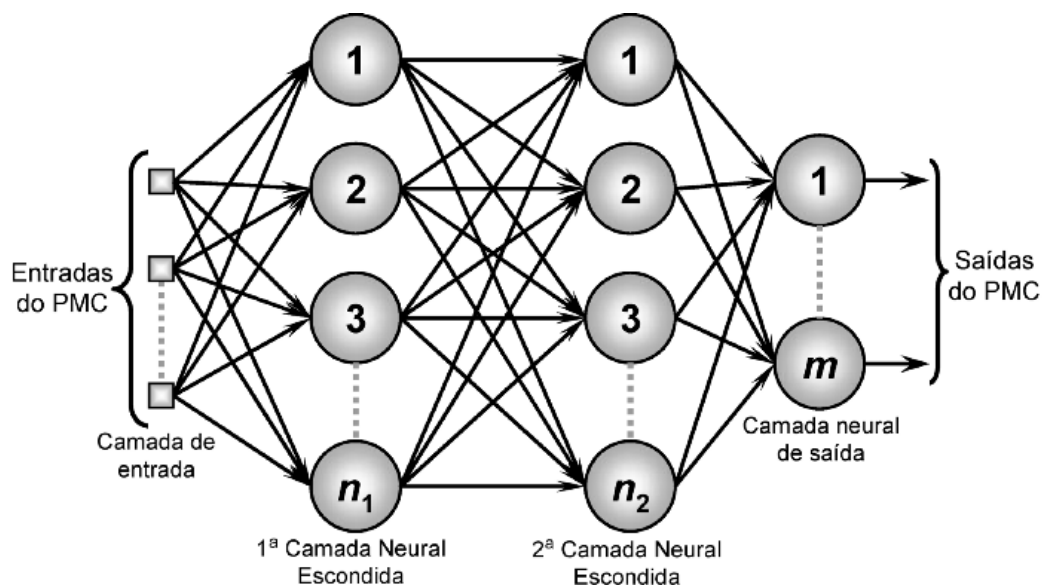


Figura 3.5: Multi-Layer Perceptron (OLIVEIRA et al., 2016)

2.6 Deep Learning

O *Deep Learning* (aprendizado profundo) é uma subárea do campo de inteligência artificial que envolve a construção e treinamento de modelos de redes neurais profundas para realizar tarefas de aprendizado automático. Esses modelos

são chamados de "profundos" porque possuem muitas camadas intermediárias entre a entrada e a saída da rede, permitindo que as informações sejam processadas de maneira hierárquica, como apresenta a Figura 3.6. O objetivo final do aprendizado profundo é criar modelos de inteligência artificial capazes de realizar tarefas complexas, como reconhecimento de fala, visão computacional, processamento de linguagem natural, entre outras (GOODFELLOW; BENGIO, 2015).

Uma das principais razões pelas quais o aprendizado profundo é tão poderoso é a capacidade das redes neurais profundas de aprender representações de alta qualidade das entradas que recebem. As camadas intermediárias da rede aprendem a reconhecer características abstratas nas entradas, e as camadas superiores da rede combinam essas características para realizar a tarefa de saída. Isso permite que os modelos de aprendizado profundo capturem nuances sutis nos dados de entrada e realizem tarefas complexas com precisão (GOODFELLOW; BENGIO, 2015).

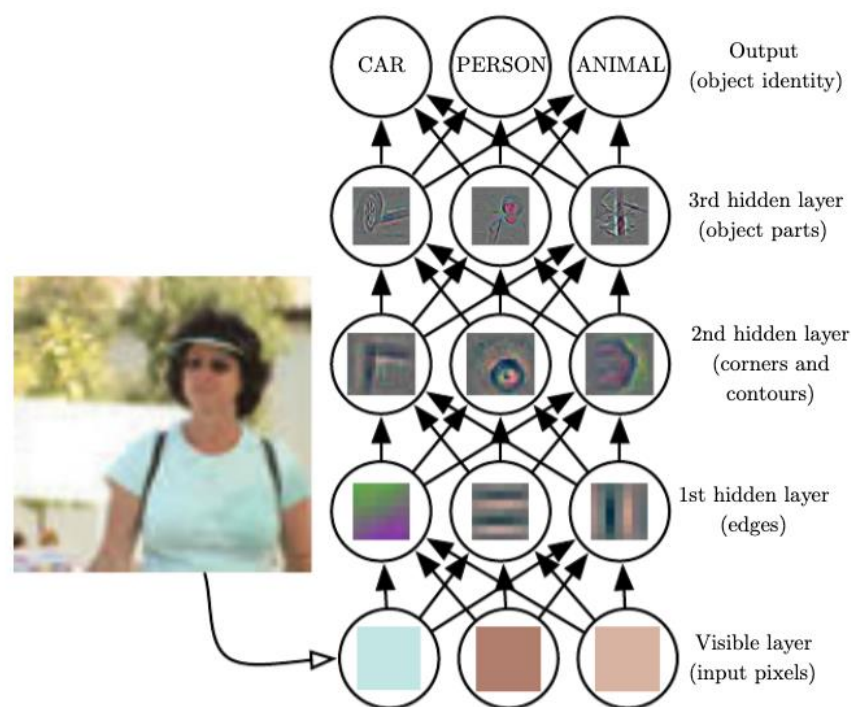


Figura 3.6: Modelo DL extraindo conceitos simples e abstratos.
(GOODFELLOW; BENGIO, 2015)

2.7 Redes Neurais Convolucionais

Uma rede neural convolucional (CNN) é um tipo de rede neural projetada especificamente para processar dados que possuem uma estrutura espacial, como imagens. As CNNs utilizam camadas convolucionais, que aplicam filtros ou kernels a pequenas regiões dos dados de entrada, buscando identificar padrões e características específicas, Figura 3.7. As camadas de pooling são comumente usadas para reduzir a dimensionalidade dos dados, preservando apenas as características mais importantes. As camadas totalmente conectadas são usadas para combinar as características extraídas pelas camadas convolucionais e produzir uma saída final (GOODFELLOW; BENGIO, 2015).

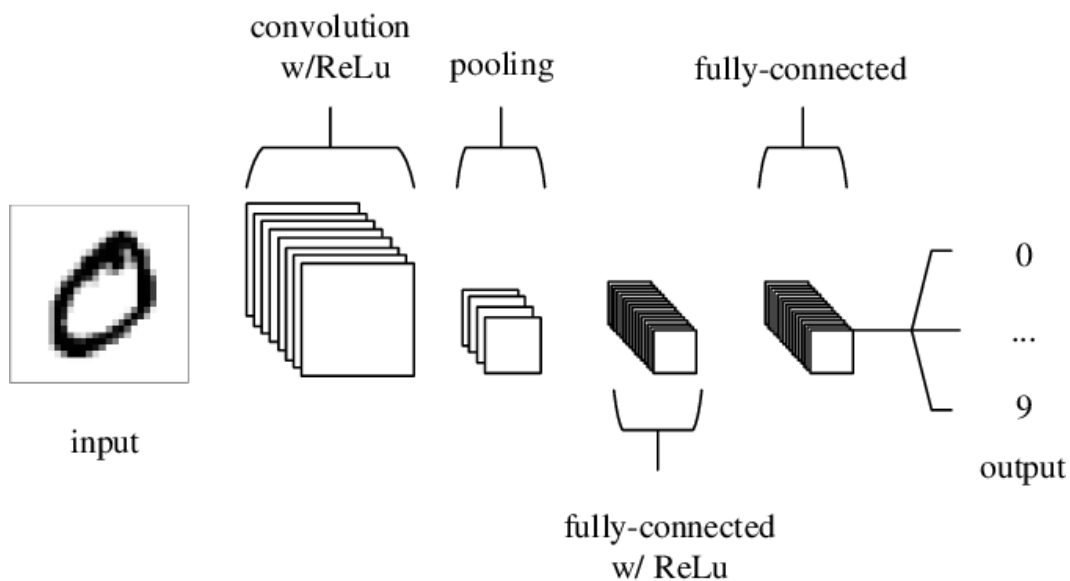


Figura 3.7: Arquitetura de uma Rede Neural Convolucional (GOODFELLOW; BENGIO, 2015)

2.7.1 Camada de convolução

Uma camada convolucional é uma das camadas fundamentais em uma rede neural convolucional (CNN). Esta camada é responsável por extrair características

importantes das imagens de entrada por meio da aplicação de operações de convolução em filtros de tamanho definido.

Segundo (GOODFELLOW et al., 2016), a camada convolucional é composta por vários filtros, cada um dos quais consiste em um pequeno conjunto de pesos que é aplicado a uma pequena região da imagem de entrada, produzindo um mapa de características para cada filtro, Figura 3.8. A convolução é aplicada repetidamente em diferentes regiões da imagem de entrada para produzir um conjunto de mapas de características que representam diferentes aspectos da imagem. Esses mapas de características são então passados para a próxima camada da rede neural para processamento adicional.

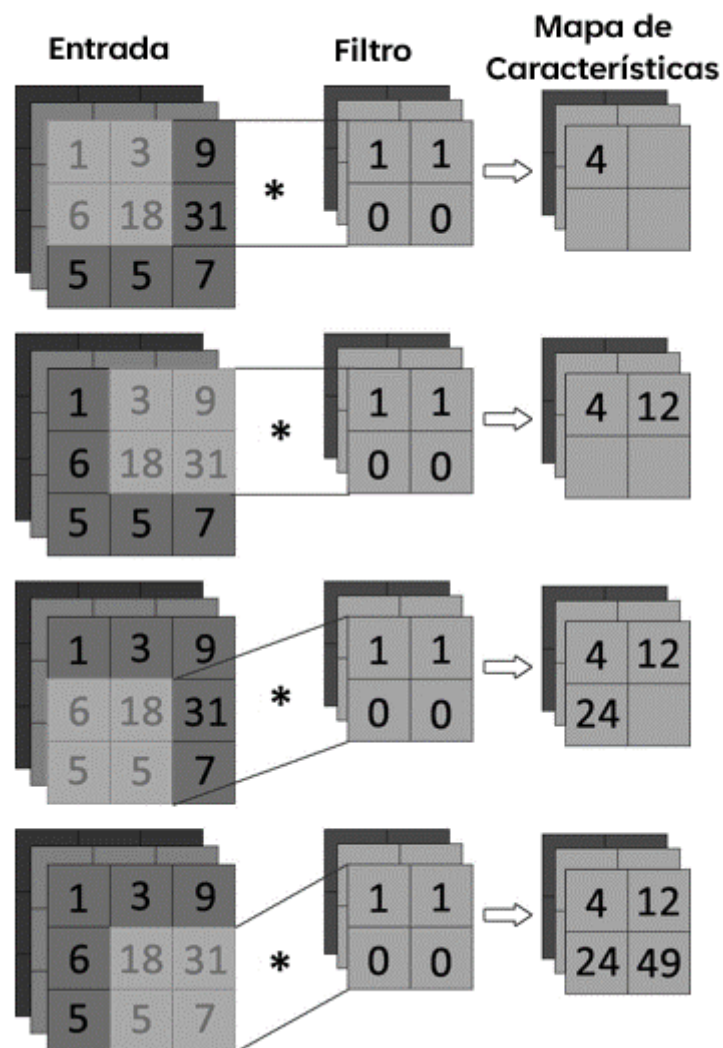


Figura 3.8: Representação visual da camada de convolução

2.7.2 Camada de pooling

De acordo com (GOODFELLOW et al., 2016), uma camada de pooling é usada principalmente para reduzir a dimensionalidade do conjunto de dados, simplificando o cálculo em camadas subsequentes e reduzindo o risco de *overfitting*. Na operação de pooling, a camada recebe um conjunto de características e aplica uma função matemática para reduzir o tamanho dos mapas de características de entrada. Isso é geralmente feito selecionando o valor máximo (max-pooling) ou a média (average-pooling) em uma região de entrada, Figura 3.9.

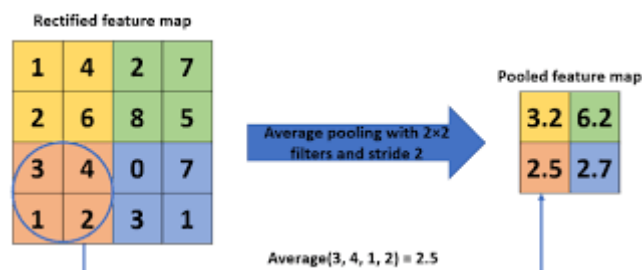


Figura 3.9: Representação visual da camada de pooling (GHOLAMALINEZHAD; KHOSRAVI, 2020)

2.7.3 Camada totalmente conectada

Uma camada totalmente conectada é uma camada em uma rede neural que conecta todos os neurônios de uma camada à camada seguinte, ou seja, cada neurônio na camada totalmente conectada é conectado a todos os neurônios na camada subsequente. Essa camada é frequentemente usada para a classificação final em uma rede neural, onde cada neurônio representa uma classe possível.

Segundo (GOODFELLOW et al., 2016), uma camada totalmente conectada é responsável por mapear as características extraídas pelas camadas convolucionais ou outras camadas intermediárias para as classes de saída, Figura 3.10. Cada

neurônio na camada totalmente conectada recebe como entrada a ativação de todos os neurônios na camada anterior, e produz uma saída que é passada para a próxima camada da rede neural.

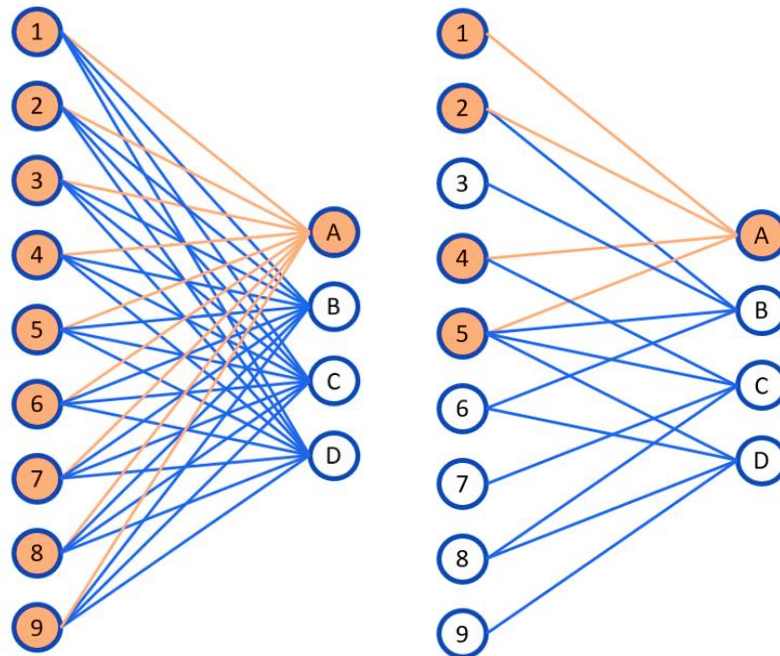


Figura 3.10: Representação visual da camada totalmente conectada (UNZUETA, 2021)

2.8 Segmentação Semântica

A segmentação semântica é um processo que envolve a atribuição de um rótulo a cada pixel de uma imagem com base em suas características, como cor, textura e forma, Figura 3.11. Essa tarefa é realizada por meio do uso de redes neurais convolucionais, que são capazes de aprender a representação de características das imagens de entrada e associá-las aos rótulos corretos (LONG et al., 2015).

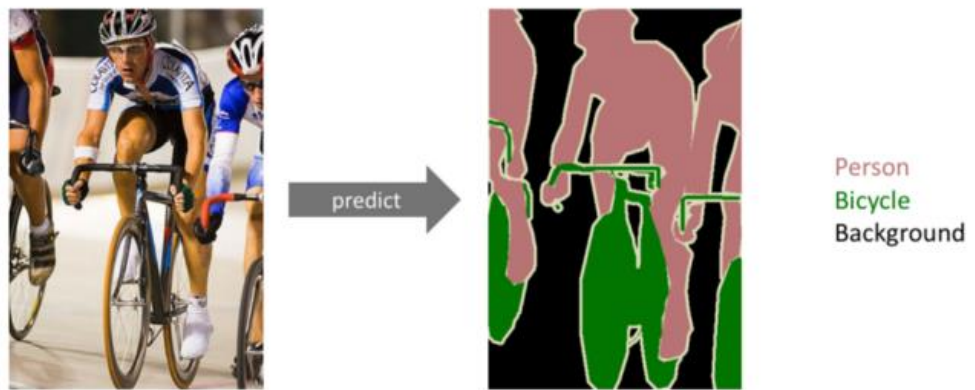


Figura 3.11: Segmentação Semântica (LAMBA, 2019)

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.

As execuções foram feitas em duas plataformas, sendo uma parte no Google Colaboratory e outra parte 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.

Todo código fonte necessário para reproduzir os experimentos pode ser encontrado em um repositório aberto na plataforma GitHub disponível em: <https://github.com/mrslima/pibic>.

3.2 Base de Dados

As amostras utilizadas foram retiradas do dataset SegTHor (LAMBERT et al., 2020), 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 essencial 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 transformadas 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 4.1.

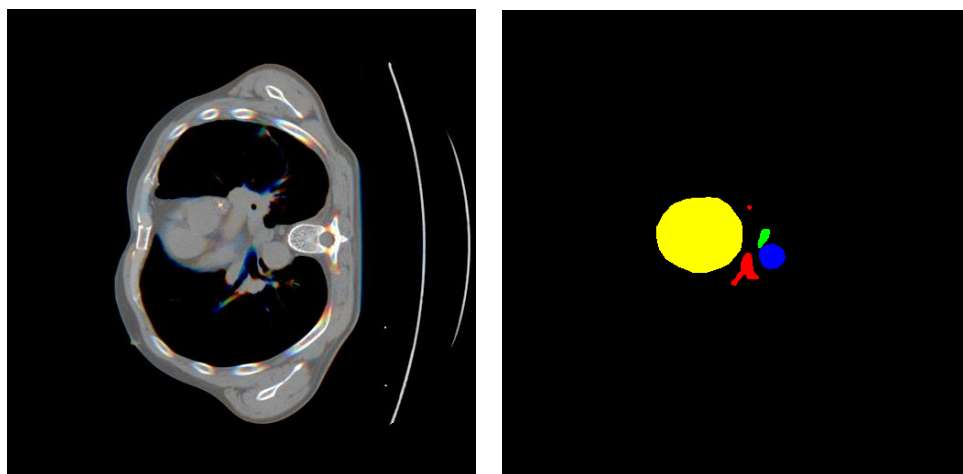


Figura 4.1: Dados de entrada: Tomografia e Máscara dos órgãos.

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 4.2, onde para cada dado de entrada, foi gerado seu respectivo par transformado, dobrando a quantidade de exemplares de treinamento.

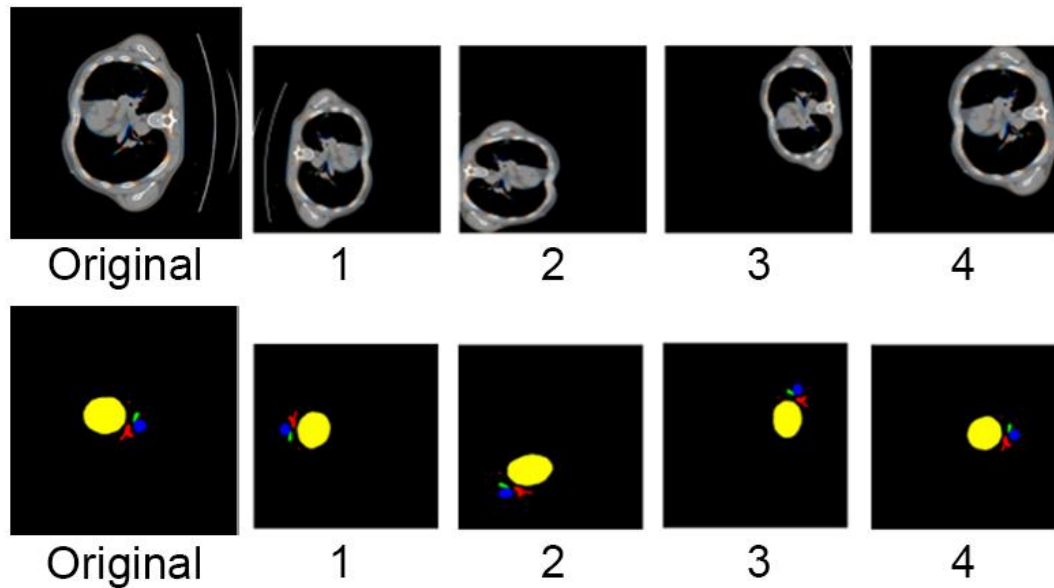


Figura 4.2: Resultado do aumento de dados aplicado em uma imagem.

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 4.3 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 Convolucional, mas que foi modificada e expandida para exigir poucas imagens de treinamento e, ao mesmo tempo, produzir segmentações mais precisas. (RONNEBERGER; FISCHER; BROX, 2015).

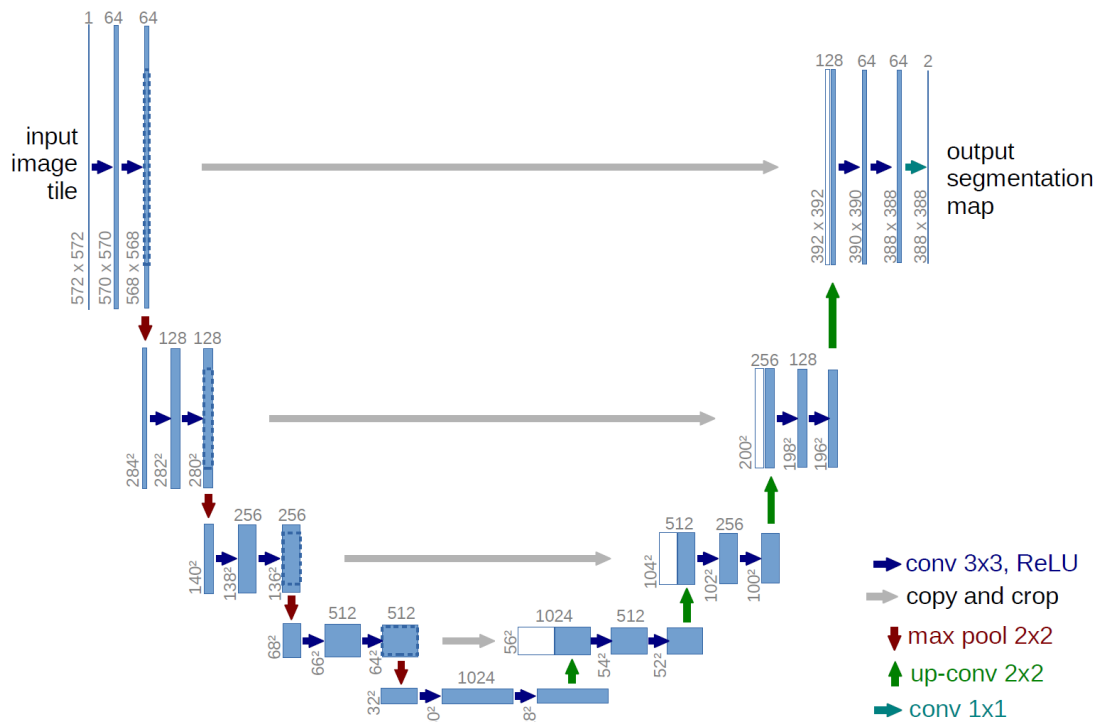


Figura 4.3: Arquitetura U-Net (RONNEBERGER; FISCHER; BROX, 2015)

4 RESULTADOS FINAIS

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, Figura 5.1, e erro, Figura 5.2, 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.

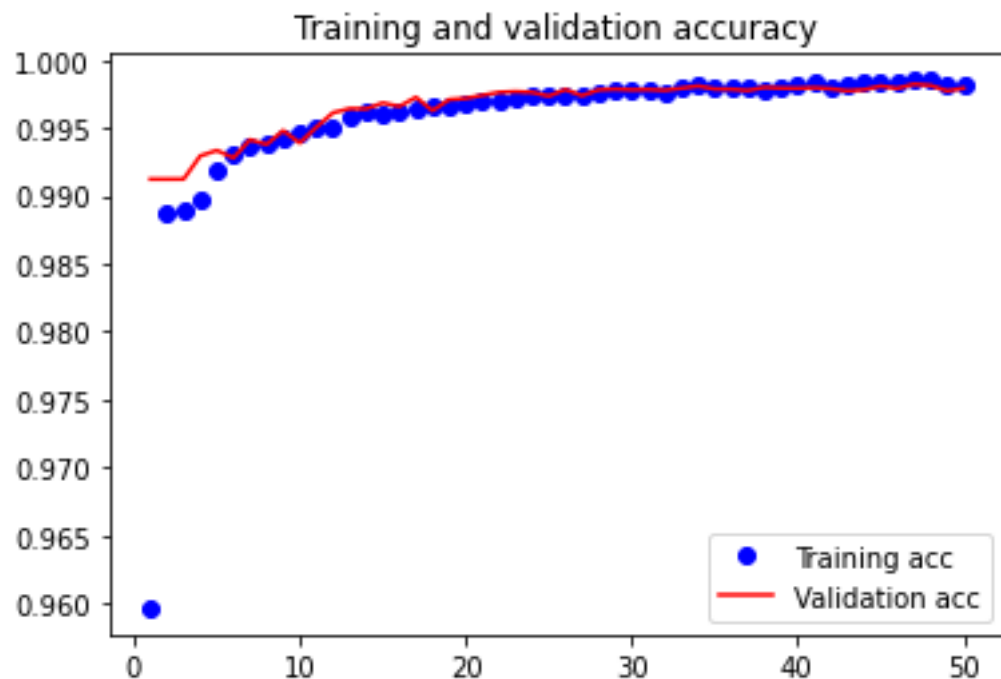


Figura 5.1: Gráfico de acurácia da rede U-Net

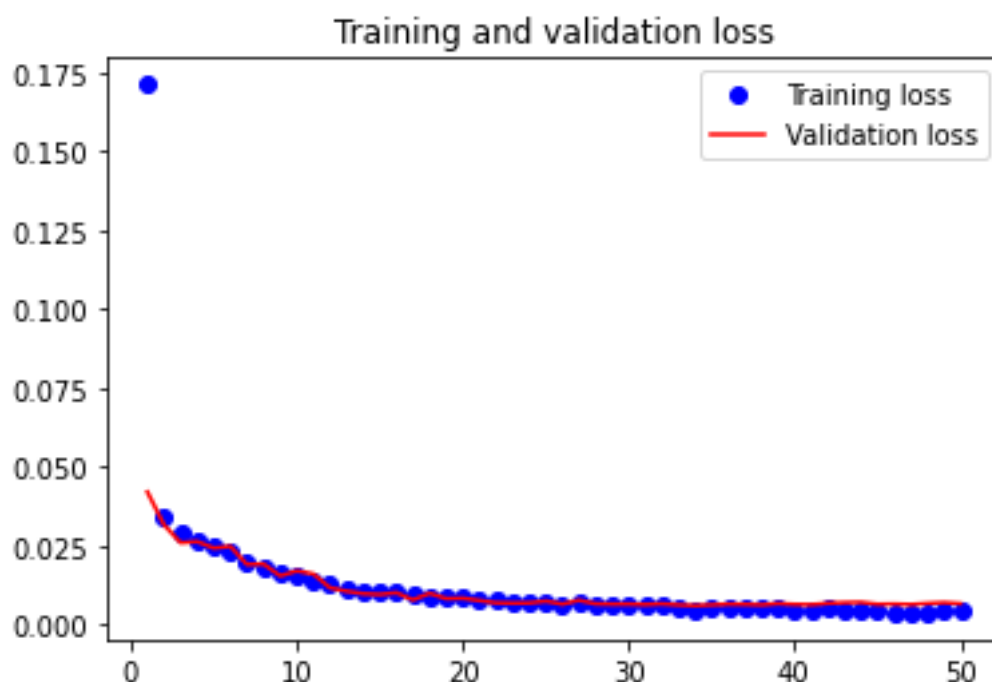


Figura 5.2: Gráfico de custo da rede U-Net

	Modelo
Acurácia	99,804%
Custo	00,056%
IoU Score (Jaccard Index)	99,495%
Dice Coefficient (F1 Score)	99,746%

Tabela 5.1: Resultados do conjunto de teste da U-Net.

Conforme os resultados da Tabela 5.1, essa estrutura apresentou um desempenho excelente. No entanto, durante a análise, identificamos que há uma falha na forma como as métricas são calculadas, o que nos gera questionamentos. Foi observado que o modelo captura uma classe indesejada, o background, o que pode levar a uma resposta tendenciosa, uma vez que essa classe não deve ser considerada. Por isso, estamos trabalhando para corrigir esse problema. Infelizmente, devido a essa falha, não podemos verificar se ocorre excesso (overfitting) ou falta (underfitting) de ajuste do modelo.

4.2 Análises 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 5.5 e a Figura 5.6 apresentam uma comparação entre as máscaras de treinamento e as geradas pelo modelo, permitindo observar a semelhança entre elas.

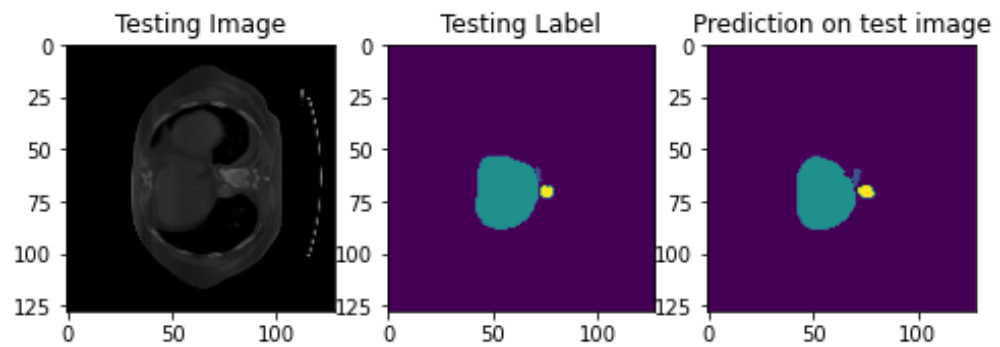


Figura 5.5: Saída prevista pelo modelo: Segmentação dos órgãos aparentes

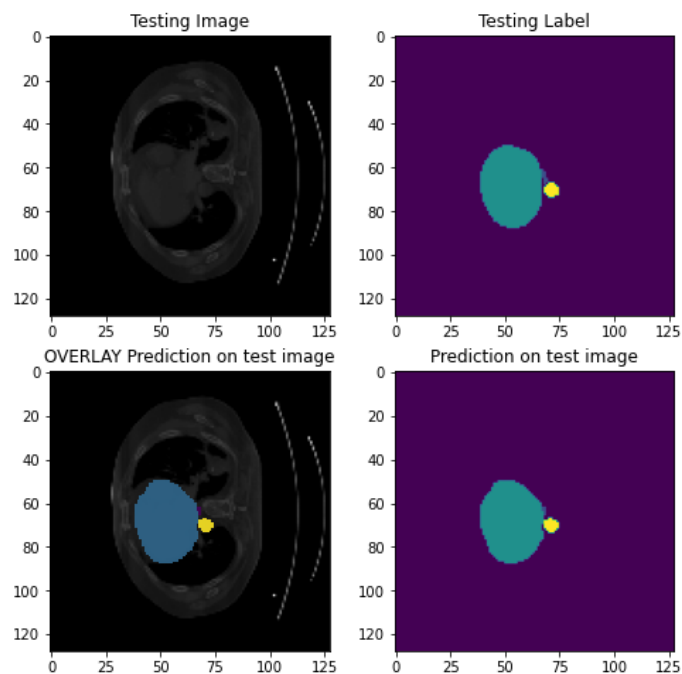


Figura 5.6: Sobreposição dos órgãos previstos com a tomografia original

5 CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo principal desenvolver um modelo de segmentação de órgãos torácicos em imagens médicas por meio de aprendizado profundo. Após a realização dos experimentos e análise dos resultados obtidos, verificamos que a abordagem de deep learning aplicada apresentou excelentes resultados na segmentação dos órgãos de interesse, proporcionando alta acurácia e sensibilidade na identificação dos tecidos-alvo.

Os resultados obtidos confirmam que o modelo desenvolvido pode ser uma ferramenta promissora para o auxílio ao diagnóstico e tratamento de doenças oncológicas, oferecendo uma solução automatizada e mais eficiente para a segmentação de órgãos em imagens médicas.

Apesar dos excelentes resultados, também identificamos algumas limitações em relação à segmentação do background e o risco de viés na avaliação das métricas. Sendo assim, sugerimos como trabalhos futuros a correção desses problemas, bem como a ampliação da base de dados e a realização de novos experimentos para validar e aprimorar ainda mais o modelo desenvolvido.

Em suma, a aplicação de deep learning na segmentação de órgãos torácicos em imagens médicas mostrou-se uma abordagem promissora, com potencial para contribuir significativamente para o avanço da medicina e oferecer benefícios práticos na identificação e tratamento de doenças oncológicas.

Além disso, como uma estudante que iniciou este projeto no segundo período do curso de ciência da computação e sem qualquer experiência com inteligência artificial, pude ter um desenvolvimento expressivo nesta área, graças à esse programa e o tempo para a fundamentação teórica deste.

5.1 Etapas Futuras

As etapas realizadas até agora foram relacionadas principalmente no aprendizado teórico e 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 com a seleção dinâmica dos modelos e avaliar possíveis resultados.

O algoritmo de seleção dinâmica, Figura 6.1, é responsável por identificar a melhor arquitetura de rede neural para segmentar uma classe específica em uma imagem aleatória. Para isso, ele passará por alguns modelos diferentes e, com base nos melhores resultados de métricas, o algoritmo determinará qual arquitetura apresenta o melhor desempenho para segmentar uma classe específica contida na imagem de entrada. Em seguida, o algoritmo irá gerar uma imagem de saída que combina as classes segmentadas com o modelo mais adequado para cada uma delas, garantindo resultados mais confiáveis.

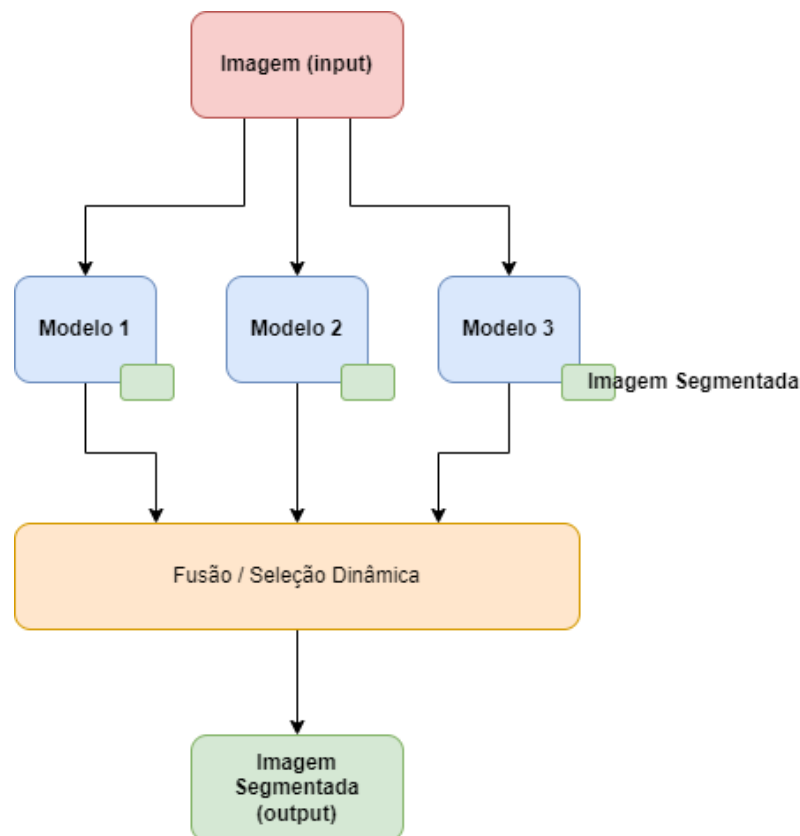


Figura 6.1: Seleção Dinâmica

Portanto, para que todos os resultados sejam atingidos, elaboramos o cronograma abaixo:

Tarefas	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out
Seleção dinâmica dos modelos	X	X	X					
Experimentações Finais			X	X				
Produção do Relatório Final SEMIC				X	X			
								X

Tabela 6.1: Cronograma

Inicialmente, abordaremos os problemas descritos na seção cinco e, em seguida, procederemos com a construção e treinamento dos algoritmos de seleção dinâmica dos modelos. Posteriormente, realizaremos as experimentações finais nos modelos para produzir o relatório final. Por fim, apresentaremos o projeto no SEMIC.

REFERÊNCIAS BIBLIOGRÁFICAS

LAMBERT, Z., PETITJEAN, C., DUBRAY, B., & KUAN, S. SegTHOR: segmentation of thoracic organs at risk in CT images. 2020.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: convolutional networks for biomedical image segmentation. 2015.

RUSSELL, Stuart J.; NORVIG, Peter. Artificial Intelligence: A Modern Approach. Prentice Hall, 2010.

ALPAYDIN, Ethem. Introduction to Machine Learning. MIT Press, 2010.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives and prospects. 2015.

BISHOP, C. M. Pattern recognition and machine learning (Vol. 4). springer. 2006.

HASTIE, T., TIBSHIRANI, R., & Friedman, J. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media, 2009.

GOODFELLOW, I., BENGIO, Y., & COURVILLE, A. Deep Learning. MIT Press, 2016.

OLIVEIRA, R. M. et al. A system based on artificial neural networks for automatic classification of hydro-generator stator windings partial discharges. 2017.

ZAMBIASI, Saulo P. O Neurônio Artificial. GSIGMA UFSC, 2022. Disponível em: <https://www.gsigma.ufsc.br/~popov/aulas/rna/neuronio_artificial/index.html>. Acesso em: 25 fev, 2023.

BISHOP, C. M. Neural networks for pattern recognition. Oxford University press, 1995.

OLIVEIRA, R. A. et al. Modelo de previsão de demanda energética utilizando redes neurais artificiais: um estudo de caso para a região Nordeste do Brasil. Revista Brasileira de Métodos Quantitativos. 2016.

GHOLAMALINEZHAD, Hossein; KHOSRAVI, Hossein. Pooling Methods in Deep Neural Networks, a Review. 2020.

UNZUETA, D. Convolutional Layers vs Fully Connected Layers. Towards Data Science, 2021. Disponível em: <<https://towardsdatascience.com/convolutional-layers-vs-fully-connected-layers-364f05ab460b>>. Acesso em: 25 fev, 2023.

LAMBA, H. Understanding Semantic Segmentation with UNET. Towards Data Science, 2019. Disponível em: <<https://towardsdatascience.com/understanding-semantic-segmentation-with-unet-6be4f42d4b47>>. Acesso em: 25 fev, 2023.

LONG, J., SHELHAMER, E., & DARRELL, T. Fully convolutional networks for semantic segmentation. 2015.

SILVA, Thiago Araújo da. Uma rede UNet modificada para segmentação e identificação de lesões pulmonares em tomografias. 2021. 48 f. Trabalho de Conclusão de Curso (Graduação em Engenharia de Computação) – Curso de Engenharia de Computação – Universidade de Brasília, Brasília, 2021.