



UNIVERSIDADE FEDERAL DE GOIÁS  
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE  
COMPUTAÇÃO (EMC)  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA  
E DE COMPUTAÇÃO

PAULO VICTOR DOS SANTOS

**MODELO NÃO SUPERVISIONADO DE  
SEGMENTAÇÃO DE ESTRUTURAS EM  
EXAMES DE TOMOGRAFIA DE CRÂNIO**

[Goiânia - Goiás - Brasil]

21 de agosto 2024



UNIVERSIDADE FEDERAL DE GOIÁS  
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

## TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

### 1. Identificação do material bibliográfico

Dissertação     Tese     Outro\*: \_\_\_\_\_

\*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

### 2. Nome completo do autor

Paulo Victor dos Santos

### 3. Título do trabalho

“Modelo não supervisionado de segmentação de estruturas em exames de tomografia de crânio”

### 4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento  SIM     NÃO<sup>1</sup>

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

- a) consulta ao(a) autor(a) e ao(a) orientador(a);
  - b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.
- O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

**Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.**



Documento assinado eletronicamente por **Wesley Pacheco Calixto, Usuário Externo**, em 26/07/2024, às 09:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Paulo Victor Dos Santos, Discente**, em 30/07/2024, às 14:18, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4692266** e o código CRC **82ED7632**.

PAULO VICTOR DOS SANTOS

**MODELO NÃO SUPERVISIONADO DE  
SEGMENTAÇÃO DE ESTRUTURAS EM  
EXAMES DE TOMOGRAFIA DE CRÂNIO**

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica e de Computação, da Escola de Engenharia Elétrica, Mecânica e de Computação, da Universidade Federal de Goiás (UFG) como requisito para a obtenção do título de Doutor em Engenharia Elétrica e de Computação.

Área de concentração: Engenharia da Computação

Orientador: Professor Doutor Wesley Pacheco Calixto

Coorientadora: Professora Doutora Marcella Scoczynski Ribeiro Martins

[Goiânia - Goiás - Brasil]

21 de agosto 2024

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Santos, Paulo Victor Dos  
MODELO NÃO SUPERVISIONADO DE SEGMENTAÇÃO DE  
ESTRUTURAS EM EXAMES DE TOMOGRAFIA DE CRÂNIO  
[manuscrito] / Paulo Victor Dos Santos. - 2024.  
125 f.

Orientador: Prof. Dr. Wesley Pacheco Calixto; co-orientadora Dra. Marcella Scoczynski Ribeiro Martins.

Tese (Doutorado) - Universidade Federal de Goiás, Escola de Engenharia Elétrica, Mecânica e de Computação (EMC), Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Goiânia, 2024.

Bibliografia. Apêndice.

Inclui siglas, abreviaturas, símbolos, gráfico, tabelas, algoritmos, lista de figuras, lista de tabelas.

1. Aprendizado não supervisionado. 2. Segmentação. 3. Tomografia computadorizada de crânio. I. Calixto, Wesley Pacheco, orient. II. Título.

CDU 004



UNIVERSIDADE FEDERAL DE GOIÁS

ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

### ATA DE DEFESA DE TESE

Ata Nº **05** da sessão de Defesa de Tese de **Paulo Victor dos Santos** que confere o título de Doutor em **Engenharia Elétrica e de Computação**, na área de concentração em **Engenharia de Computação**.

Aos **dez dias do mês de julho de dois mil e vinte e quatro**, a partir das **14h00min.**, realizou-se a sessão pública de Defesa de Tese intitulada “**Modelo não supervisionado de segmentação de estruturas em exames de tomografia de crânio**”. Os trabalhos foram instalados pelo Orientador, Professor Doutor **Wesley Pacheco Calixto - (EMC/UFG)**, com a participação dos demais membros da Banca Examinadora: Professora Doutora **Marcella Scoczynski Ribeiro Martins - (ELE/UTFPR)** coorientadora, Professor Doutor **Flávio Henrique Teles Vieira - (EMC/UFG)** Membro Titular Interno, Professora Doutora **Cristhiane Gonçalves - (ELE/UTFPR)** Membro Titular Externo, Professor Doutor **Antônio Paulo Mendes Breda Dias Coimbra - (DEEC/UC-Portugal)** Membro Titular Externo, e Professora Doutora **Myriam Regattieri de Biase da Silva Delgado - (ELE/UTFPR)** Membro Titular Externo, Professora Doutora **Solange Amorim Nogueira (HIAE/BIOMED)** Membro Titular Externo, **cujas participações ocorreram através de videoconferência** através do link: <https://meet.google.com/pyz-vsru-ohu>. Durante a argüição os membros da banca **não fizeram** sugestão de alteração do título do **trabalho**. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Professor Doutor Wesley Pacheco Calixto Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos dez dias do mês de julho de dois mil e vinte e quatro.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Wesley Pacheco Calixto, Usuário Externo**, em 28/08/2024, às 08:24, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Myriam Regattieri De Biase da Silva Delgado, Usuário Externo**, em 28/08/2024, às 09:02, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Cristhiane Gonçalves, Usuário Externo**, em 28/08/2024, às 09:30, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcella Scoczynski Ribeiro Martins, Usuário Externo**, em 29/08/2024, às 07:44, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Paulo Victor Dos Santos, Discente**, em 29/08/2024, às 08:05, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Flavio Henrique Teles Vieira, Professor do Magistério Superior**, em 29/08/2024, às 08:34, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **SOLANGE AMORIM NOGUEIRA, Usuário Externo**, em 29/08/2024, às 09:19, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Antônio Paulo Mendes Breda Dias Coimbra, Usuário Externo**, em 29/08/2024, às 20:20, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



A autenticidade deste documento pode ser conferida no site [https://sei.ufg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4778055** e o código CRC **BCF6A747**.

---

Referência: Processo nº 23070.029170/2024-23

SEI nº 4778055

*"O impossível não é um fato, é apenas uma opinião."*

MUHAMMED ALI

*Dedico este trabalho a todas que fixeram parte da minha trajetória pessoal e profissional. Ao meu pai e minha mãe, especialmente, por serem meus alicerces.*

## AGRADECIMENTOS

Gostaria primeiramente de expressar minha gratidão a Deus por me conceder saúde e perseverança, fundamentais para alcançar meus objetivos. Ao Professor Wesley Pacheco Calixto, agradeço pela confiança depositada, pelo apoio em meu trabalho e pela paciência, além do dom excepcional de ensinar.

À Professora Marcella Scoczynski, agradeço pela simplicidade, pela calma, pela confiança e pela atenção dedicada à condução deste e outros trabalhos.

Sou imensamente grato também ao meu querido pai, Sebastião Batista dos Santos, por estar ao meu lado em todos os momentos. À minha querida mãe, Maria Fátima, que sempre me incentivou a ver o lado positivo da vida. Ao meu fiel irmão Carlos Andrerres, companheiro de muitas batalhas e conquistas. A Gleidiane Gomes de Sá, que me deu o maior presente desta vida, minha filha Maria Gabriela.

Agradeço profundamente às pessoas importantes que formam a base da minha jornada, minha família, composta por meus tios, tias, primos, primas e minha amada avó. Pessoas simples e que sempre me ajudaram a alcançar desde metas simples até as mais complexas.

Também expresso minha gratidão a todos os amigos e colegas que fazem parte da minha vida profissional e contribuíram significativamente para meu crescimento e amadurecimento.

À Escola de Engenharia Elétrica, Mecânica e Computação da Universidade Federal de Goiás, agradeço pela excelente estrutura disponibilizada, bem como a todos os docentes e colaboradores sempre prontos e dispostos a ajudar, especialmente ao Professor Flávio, à Professora Cacilda e o Professor João.

Por fim, meu agradecimento aos professores que tive a honra de conhecer durante as disciplinas que cursei: Ronaldo Martins da Costa, Cristhiane Gonçalves, Fabrízio Alphonsus, Gustavo Teodoro Laureano, Rogério Lopes Salvini, Telma Woerle, Anderson Soares, Sílvio Leão e Márcio Reis.

## RESUMO

Este trabalho propõe o desenvolvimento de método não supervisionado para a segmentação de imagens de tomografia computadorizada craniana. Na metodologia extrai-se características das imagens, aplicando restrições de similaridade e continuidade para criar mapas de segmentação de estruturas intracranianas e tecidos observáveis. Esta abordagem visa auxiliar especialistas no diagnóstico, identificando regiões com anomalias específicas. O método, aplicado ao conjunto de dados reais, utiliza função de avaliação da continuidade espacial relacionada à quantidade desejada de estruturas. Os resultados apresentam desempenho satisfatório, indicando abordagem simplificada e acessível que reduz a carga computacional, o tempo de treinamento e os custos financeiros, oferecendo potencial impacto na prática clínica ao agilizar a interpretação de exames anormais. A proposta serve como ferramenta aplicável na segmentação de imagens de tomografia computadorizada craniana, proporcionando contribuição significativa para a análise de imagens médicas em ambientes clínicos e diagnósticos.

**Palavras-chave:** Aprendizado não supervisionado; Segmentação; Tomografia computadorizado de crânio.

# UNSUPERVISED MODEL OF STRUCTURES SEGMENTATION IN BRAIN TOMOGRAPHY EXAMINATIONS

## ABSTRACT

This work proposes the development of an unsupervised method for segmenting cranial CT images. The methodology involves extracting image features and applying similarity and continuity constraints to create segmentation maps of intracranial structures and observable tissues. This approach aims to assist specialists in diagnosis by identifying regions with specific anomalies. Applied to real-world datasets, the method uses a spatial continuity evaluation function related to the desired number of structures. Results show satisfactory performance, indicating a simplified and accessible approach that reduces computational load, training time, and financial costs. This proposal serves as a practical tool for cranial CT image segmentation, providing significant contributions to the analysis of medical images in clinical and diagnostic settings.

**Keywords:** Unsupervised learning; Segmentation; Computerized tomography of cranial

## SUMÁRIO

Pág.

### LISTA DE FIGURAS

### LISTA DE TABELAS

### LISTA DE SÍMBOLOS

### LISTA DE ABREVIATURAS E SIGLAS

<b>CAPÍTULO 1 INTRODUÇÃO</b> . . . . .	<b>27</b>
<b>CAPÍTULO 2 TOMOGRAFIA COMPUTADORIZADA</b> . . . . .	<b>33</b>
2.1 Histórico da tomografia computadorizada . . . . .	33
2.2 Plano axial, coronal e sagital . . . . .	35
2.3 Imagens de tomografia computadorizada × ressonância magnética . . . . .	37
2.4 Cérebro humano . . . . .	37
2.4.1 Estruturas cerebrais . . . . .	37
2.4.2 Tecidos cerebrais . . . . .	40
2.5 Protocolos para exames de imagens . . . . .	42
2.6 Janelamento . . . . .	42
2.7 Anotação e rótulo em imagens médicas . . . . .	44
2.8 Imagens médicas . . . . .	44
2.9 Considerações finais . . . . .	45
<b>CAPÍTULO 3 INTELIGÊNCIA ARTIFICIAL APLICADA A DIAGNÓSTICO</b> . . . . .	<b>47</b>
3.1 Aprendizado de máquina . . . . .	47
3.2 Aprendizado supervisionado, não supervisionado e profundo . . . . .	48
3.3 Redes neurais convolucionais . . . . .	50
3.4 Segmentação supervisionada e não supervisionada . . . . .	52
3.5 Função de avaliação, de perda e entropia cruzada . . . . .	54
3.6 Otimização dos hiperparâmetros . . . . .	55
3.7 Considerações finais . . . . .	56

<b>CAPÍTULO 4</b>	<b>MODELOS DE SEGMENTAÇÃO NÃO SUPERVISIONADO . . . . .</b>	<b>57</b>
4.1	Modelos não supervisionados aplicados em segmentação de imagem . . . . .	57
4.2	Rede neural convolucional para segmentação de imagens . . . . .	58
4.3	Critérios e restrições para agrupamento de características diferenciáveis . . . . .	60
4.4	Extração de características, restrição na similaridade e número de rótulos . . . . .	60
4.5	Restrição na continuidade espacial e função de perda . . . . .	62
4.6	Aprendizagem por retropropagação e experimentos . . . . .	63
4.7	Considerações finais . . . . .	65
<b>CAPÍTULO 5</b>	<b>METODOLOGIA . . . . .</b>	<b>67</b>
5.1	Contextualização . . . . .	67
5.2	Fundamentos e critérios da metodologia . . . . .	67
5.2.1	Pré-Processamento . . . . .	68
5.3	Arquitetura do segmentador . . . . .	69
5.3.1	Extrator de características . . . . .	69
5.3.2	Otimização dos parâmetros e hiperparâmetros . . . . .	74
5.3.3	Validação do modelo proposto . . . . .	75
5.4	Considerações finais . . . . .	75
<b>CAPÍTULO 6</b>	<b>RESULTADOS . . . . .</b>	<b>77</b>
6.1	Base de dados e definição dos parâmetros . . . . .	77
6.2	Resultado do processo de otimização . . . . .	78
6.3	Comparação entre as metodologias de segmentação . . . . .	80
6.4	Comparação entre a volumetria do crânio . . . . .	87
6.5	Segmentação validada por especialista . . . . .	89
6.6	Discussão . . . . .	92
<b>CAPÍTULO 7</b>	<b>CONCLUSÃO . . . . .</b>	<b>95</b>
7.1	Contribuições do trabalho . . . . .	95
7.2	Trabalhos futuros . . . . .	97
<b>CAPÍTULO A</b>	<b>. . . . .</b>	<b>99</b>
<b>CAPÍTULO B</b>	<b>. . . . .</b>	<b>101</b>
<b>CAPÍTULO C</b>	<b>. . . . .</b>	<b>103</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>. . . . .</b>	<b>105</b>

## LISTA DE FIGURAS

	<u>Pág.</u>
2.1	Simplificação da tomografia computadorizada. . . . . 35
2.2	Plano de aquisição. . . . . 36
2.3	Planos de aquisição. . . . . 36
2.4	Exames de crânio. . . . . 37
2.5	Mapeamento cerebral. . . . . 39
2.6	Sistema de atlas cerebral. . . . . 41
2.7	Regiões no cérebro. . . . . 41
2.8	Representação hipotética da largura da janela. . . . . 43
2.9	Janelamento em tomografia de crânio. . . . . 44
3.1	Diagrama de alto nível de programas tradicionais. . . . . 47
3.2	Diagrama de alto nível de aprendizado de máquina. . . . . 47
3.3	Aprendizado supervisionado e não supervisionado. . . . . 49
3.4	Arquitetura do sistema de aprendizado profundo. . . . . 50
3.5	Operação matemática genérica de convolução. . . . . 51
3.6	Redução da dimensionalidade da imagem. . . . . 51
3.7	Arquitetura da rede neural convolucional. . . . . 52
3.8	Arquitetura da rede neural profunda U-net. . . . . 53
3.9	Expansão de escala para observação dos píxeis. . . . . 54
5.1	Fluxograma de validação da metodologia proposta. . . . . 68
5.2	Fluxograma da arquitetura proposta. . . . . 70
5.3	Plano axial e superfície: (a) imagem de tomografia computadorizada e (b) superfície da Figura 5.3(a). . . . . 71
6.1	Imagem com janelamento e segmentada: (a) janelamento de $40 \times 80$ [HU] e (b) gráfico de superfície de imagem segmentada. . . . . 81
6.2	Segmentação dos exames: (a) até (f) CT042 e de (g) até (l) CT195. . . . . 82
6.3	Segmentação dos exames: (a) até (f) CT200, de (g) até (l) CT299, de (m) até (r) CT418. . . . . 83
6.4	Segmentação sobreposta do exame CT195: (a) $c_1$ , (b) $c_2$ , (c) $c_3$ , (d) $c_4$ , (e) $c_5$ e (f) $c_6$ . . . . . 86
6.5	Volumetria do crânio para a classe $c_4$ do exame CT042. . . . . 88

6.6	Segmentação utilizando a metodologia proposta nos exames:	
	(a) CT042 $N_R = 7$ , (b) CT042 $N_R = 8$ rótulos, (c) CT195 $N_R = 7$ ,	
	(d) CT195 $N_R = 8$ , (e) CT200 $N_R = 7$ , (f) CT200 $N_R = 8$ ,	
	(g) CT299 $N_R = 7$ , (h) CT299 $N_R = 8$ , (i) CT418 $N_R = 7$ e	
	(j) CT418 $N_R = 8$ . . . . .	90
6.7	Comparação através de sobreposição . . . . .	91
C.1	Cálculo da taxa de continuidade. . . . .	103
C.2	Cálculo da entropia cruzada e da taxa de similaridade. . . . .	104

## LISTA DE TABELAS

	<u>Pág.</u>
1.1 Trabalhos relacionados à segmentação de imagens de ressonância magnética e tomografia computadorizada. . . . .	31
2.1 Escala <i>Hounsfield</i> em materiais para tomografia computadorizado. . . . .	34
2.2 Exemplo de <i>tags</i> utilizadas nos arquivos DICOM . . . . .	45
6.1 $\overline{D_c}$ , $\sigma$ e $t$ ao longo das doze simulações para cada $N_R$ e para cada classe $c_1$ , $c_2$ e $c_4$ . . . . .	79
6.2 Parâmetros e hiperparâmetros otimizados com base na quantidade de rótulos predefinidos. . . . .	80
6.3 Comparação de $D_c$ entre a metodologia proposta $\times$ metodologia de Kim et al. (2020). . . . .	84
6.4 Comparação dos valores da volumetria do crânio entre a metodologia proposta $\times$ CTSeg. . . . .	88



## LISTA DE SÍMBOLOS

$\alpha$	– Taxa de aprendizado ajustável
$\Delta(r'_{i,j})$	– Taxa de continuidade
$\lambda$	– Distância a ser encontrada
$\omega$	– Subconjunto das soluções factíveis
$\hat{\theta}$	– Valor mínimo
$\mu$	– Peso de equilíbrio
$\mu_m$	– Média do mapa de resposta
$\nu$	– Peso na função de perda de rabisco
$\sigma$	– Variância do mapa de resposta
$\infty$	– Infinito
$(\xi, \eta)$	– Valor do pixel
$\delta(t)$	– Rótulos pseudo-avos
$\partial$	– Gradiente
$\Sigma$	– Somatório
$\mathbb{R}^k$	– Conjunto dos números reais
$A_i$	– Altura da imagem
$b$	– Viés
$b_v$	– Valor da restrições do problema
$c$	– Classe
$c_n$	– Conjunto de rótulos
$Cont$	– Taxa de continuidade
$d_{cont}$	– Taxa de deslocamento de continuidade
$d_{sim}$	– Taxa de deslocamento de similaridade
$D$	– Número de variáveis
$D_c$	– Coeficiente <i>Dice</i>
$f$	– Função de ativação
$fh$	– Altura do campo receptivo
$filtro$	– Quantidade de neurônios
$fw$	– Largura do campo receptivo
$f(v)$	– Função objetivo
$f(x)$	– Função de avaliação
$F_{aval}$	– Função de avaliação de treinamento
$g$	– Função de mapeamento
$g(v)$	– Restrições do problema
$g(x_n)$	– Rótulos de agrupamento
$H$	– Altura da imagem
$H_n$	– Tamanho da camada intermediária <i>MLP</i>
$H_1$	– Camada intermediária <i>MLP</i>
$i$	– Iterador por linha
$input\_dim$	– Dimensão da imagem de entrada
$I$	– Representação de imagem

$I * K$	– Matriz resultante da convolução
$j$	– Iterador por coluna
$J$	– Conjunto de dados
$J_i$	– Instância dos dados
$K$	– Elemento estruturante
$k$	– Quantidade de agrupamentos
$l$	– Camada convolucional
$L$	– Valor da perda
$L_{con}$	– Função de continuidade espacial
$L_i$	– Largura da imagem
$L_j$	– Largura da janela
$L_{scr}$	– Função de perda de rabisco
$L_1$	– Função de perda da soma dos erros absolutos
$L_2$	– Função de perda da soma dos erros médios
$m$	– Dados do conjunto de treinamento
$M$	– Número de classes
$nChannel$	– Quantidade de filtros
$nConv$	– Quantidade de camadas
$N$	– Número de píxeis
$N_C$	– Número de camadas convolucionais
$N_{fatia}$	– Número de fatias
$N_F$	– Número de filtros convolucionais
$N_I$	– Número máximo de iterações
$N_j$	– Nível da janela
$N_R$	– Número de rótulos desejados
$O_{PT}$	– Otimizador
$pitch$	– Coeficiente de deslocamento por espessura
$P_n$	– Imagem processada
$p(x)$	– Distribuição de probabilidade real
$p(x_c)$	– Probabilidade verdadeira da classe $c$
$p_{i,x_c}$	– Probabilidade real de $i$ pertencer à classe $c$
$q$	– Mapa de características
$q'$	– Mapa de características normalizado
$q(x)$	– Distribuição de probabilidade predita
$q(x_c)$	– Probabilidade calculada da classe $c$
$q_{i,x_c}$	– Probabilidade prevista de $i$ pertencer à classe $c$
$r_n$	– Vetores de resposta
$r'_n$	– Vetores de resposta normalizados
$\mathbb{R}^p$	– Conjunto características
$\mathbb{R}^3$	– Conjunto de píxeis coloridos da imagem
$s_n$	– Rótulo de rabisco
$Sim$	– Taxa de similaridade

$S$	–	Similaridade
$T$	–	Quantidade de iterações
$Tag$	–	Chave do dicionário <i>DICOM</i>
$T_A$	–	Taxa de aprendizado
$u_n$	–	Píxel rabiscado
$v$	–	Variáveis do problema
$v_n$	–	Píxeis de imagem colorida
$w$	–	Vetor de pesos do neurônio
$W$	–	Largura da imagem
$W_c$	–	Classificador linear
$w_j$	–	Parâmetro do gradiente descendente
$w_n$	–	Tamanho do vetor de pesos do neurônio
$w_1$ a $w_3$	–	Peso do neurônio
$x$	–	Vetor de entrada do neurônio
$x_n$	–	Tamanho do vetor de entrada do neurônio
$x_1$ a $x_3$	–	Valor de entrada do neurônio
$y$	–	Saída do neurônio
$y_i$	–	Valor real
$\hat{y}_i$	–	Valor desejado
$Z$	–	Potencial de ativação
$\mathbb{Z}$	–	Conjunto dos números inteiros



## LISTA DE ABREVIATURAS E SIGLAS

AG	–	Algoritmo genético
ADAM	–	<i>Adaptive moment estimation</i>
AVC	–	Acidente vascular cerebral
BN	–	<i>Batch normalization</i>
DFC	–	<i>Differentiable feature clustering</i>
DICOM	–	<i>Digital imaging and communications in medicine</i>
EAM	–	Erro absoluto médio
EQM	–	Erro quadrático médio
FCN	–	<i>Fully connected network</i>
GPU	–	<i>Graphics processing unit</i>
IA	–	Inteligência artificial
IIC	–	<i>Invariant information clustering</i>
IoU	–	<i>Intersection over Union</i>
LCR	–	Líquido cefalorraquidiano
LSTM	–	<i>Long short-term memory</i>
MAE	–	<i>Mean average error</i>
MVS	–	Máquina de vetores de suporte
NAC	–	<i>Normalization across channels</i>
NLP	–	<i>Natural language processing</i>
MLP	–	<i>Multi layer perceptron</i>
OMS	–	Organização mundial da saúde
RELU	–	<i>Rectified linear unit</i>
RNA	–	Rede neural artificial
RNC	–	Rede neural convolucional
RNP	–	Rede neural profunda
RNR	–	Rede neural recorrente
RPR	–	Rede de pirâmide de recursos
RM	–	Ressonância magnética
RMSprop	–	<i>Root mean square propagation</i>
RX	–	raio-X
SBB	–	<i>Segmentation by backpropagation</i>
SGD	–	<i>Stochastic gradient descent</i>
SNC	–	Sistema nervoso central
TC	–	Tomografia computadorizada
HU	–	Unidade <i>Hounsfield</i>
keV	–	Quiloelétons-volt
MeV	–	Megaelétons-volt.
3D	–	Tridimensional



# CAPÍTULO 1

## INTRODUÇÃO

Distúrbios e doenças neurológicas constituem risco significativo não apenas para o desenvolvimento individual, mas também para a produtividade em geral. Entre os distúrbios neurológicos mais prevalentes estão as doenças neurodegenerativas, como Alzheimer e Parkinson, os transtornos mentais, como depressão e ansiedade, e os distúrbios do desenvolvimento neurológico, como autismo e epilepsia, que geram elevados custos econômicos (GOOCH et al., 2017; OWOLABI et al., 2023). Estes distúrbios envolvem anormalidades estruturais, bioquímicas ou elétricas que afetam o cérebro, a medula espinhal e os nervos (PATEL et al., 2021).

De acordo com dados da Organização Mundial da Saúde (OMS), estima-se que uma em cada três pessoas desenvolva um ou mais distúrbios neurológicos em algum momento de suas vidas. Estes distúrbios não são apenas a principal causa de incapacidade, mas também a segunda principal causa de morte. Diante destas estatísticas, torna-se necessário que a saúde pública e a sociedade como um todo priorizem a compreensão e o manejo eficiente dos distúrbios neurológicos. A identificação de estratégias de prevenção, diagnóstico precoce e tratamentos eficientes são fundamentais para minimizar o impacto adverso destas condições na saúde e no bem-estar dos indivíduos. Além disto, estas medidas podem aliviar o ônus econômico associado aos distúrbios neurológicos (ORGANIZATION, 2006). Segundo Steinmetz et al. (2024), cerca de 3,4 bilhões de indivíduos no mundo foram acometidos por distúrbios neurológicos no ano de 2021.

Com o aumento da incidência de distúrbios e doenças neurológicas e os avanços tecnológicos contínuos, o volume de exames neurodiagnósticos armazenados em repositórios tem crescido significativamente (LIMA et al., 2022; USMAN et al., 2022). Esta crescente quantidade de dados neurológicos disponíveis anualmente, especialmente dados de imagem, impulsiona a necessidade de utilizar a inteligência artificial (IA) para desenvolver modelos capazes de auxiliar os neurologistas no diagnóstico precoce e aprimorar a assistência médica nesta área (WAHL et al., 2018). A IA tem se apresentado como ferramenta importante na detecção, diagnóstico e até mesmo na prevenção de resultados irreversíveis decorrentes de distúrbios neurológicos (LANGEN et al., 2017). Por meio da análise de grandes volumes de dados, a IA pode identificar padrões complexos e sutis que podem passar despercebidos aos olhos humanos.

As técnicas de IA aplicadas ao processamento de imagens médicas têm se benefi-

ado do processo de anotação<sup>1</sup>, que cumpre ação importante no contexto de detecção, diagnóstico e prevenção. Com conjuntos de dados de imagens médicas corretamente rotuladas, organizadas e limpas, é possível desenvolver tecnologias computacionais que permitem a classificação eficiente e a obtenção de diagnósticos médicos preliminares de forma supervisionada (SAIT et al., 2021). Neste sentido, os algoritmos de aprendizado de máquina são treinados utilizando os rótulos previamente anotados, estabelecendo associações e inferências entre os dados e as categorias relevantes. Esta abordagem supervisionada viabiliza a generalização e inferência de novos exames com base em padrões previamente aprendidos. Entre os métodos de aprendizado supervisionado, as redes neurais artificiais (RNA) têm-se destacado como abordagem promissora (CARUANA; NICULESCU-MIZIL, 2006; SCHMIDHUBER, 2015). No entanto, o processo de anotação é oneroso em termos de tempo e recursos financeiros e várias vezes os conjuntos de dados disponíveis não possuem anotações adequadas ou quantidade suficiente para o treinamento dos modelos de IA (RISTER et al., 2020; TAJBAKSHI et al., 2020; YANG et al., 2017).

As RNA têm importante função no processamento de imagens médicas, sendo a segmentação<sup>2</sup> uma das tarefas desempenhadas por elas (ZHANG et al., 2019). Esta técnica é importante para diversas aplicações biomédicas, como a delimitação de estruturas anatômicas do corpo (NAZIR et al., 2023). Na área de neuroimagem, a segmentação de estruturas intracranianas é importante, possibilitando a classificação, visualização de substâncias cerebrais e identificação de anomalias, incluindo a localização de tumores (PHAM et al., 2000; MAHATA et al., 2018; KER et al., 2017). Para realizar a segmentação, as Redes Neurais Profundas (RNP) têm se apresentado eficientes, atuando como classificadores supervisionados e obtendo resultados mais precisos (KRIZHEVSKY et al., 2012; NAZIR et al., 2023; SZEGEDY et al., 2017).

Vários estudos têm aplicado técnicas de segmentação com base em aprendizado profundo supervisionado (APS). Monteiro et al. (2020) propõem o uso de APS para quantificar lesões cerebrais utilizando o conjunto de dados *CQ500* (CHILAMKURTHY et al., 2018). Ronneberger et al. (2015) e Li et al. (2020) constroem a rede neural profunda denominada *U-Net* para segmentação de hemorragias cerebrais, obtendo precisão de 98% na localização de sangramentos em dois conjuntos de dados diferentes. No entanto, estas metodologias estão limitadas ao uso de conjuntos de dados previamente anotados. A segmentação de imagens intracranianas é processo

---

<sup>1</sup>Identificação, marcação e rotulação de áreas específicas com lesões ou características relevantes, realizada por especialistas (CORDEIRO; CARNEIRO, 2020; GATIDIS et al., 2022).

<sup>2</sup>Divisão da imagem em regiões, objetos ou píxeis, com o objetivo de separar os objetos de interesse dos demais elementos presentes (SENTHILKUMARAN; RAJESH, 2009).

trabalhoso, devido à presença de ruído nas imagens e à necessidade de exames anotados para treinamento, especialmente em imagens de tomografia computadorizada (TC) (BALAFAR et al., 2010). Além disto, a segmentação supervisionada apresenta as características: i) é onerosa em tempo e financeiramente, pois depende de anotações fornecidas por especialistas, com viés de rotulagem, pois podem variar conceitualmente entre diferentes especialistas e ii) é custosa para a modelagem de diversos tipos de exames, exigindo alto custo computacional para treinar modelos para cada conjunto de dados específico.

Há na literatura outro método utilizado, a segmentação não supervisionada que tem se destacado por suas vantagens, como não depender de rótulos e lidar com o ruído presente nas imagens. Balafar et al. (2010) realizam revisões dos métodos de segmentação supervisionada e não supervisionada, sugerindo pesquisas futuras para melhorar a velocidade, precisão e combinação destes métodos. Atkins e Mackiewicz (1998) propõem o uso de segmentação não supervisionada utilizando filtros anisotrópicos e técnicas de processamento de imagem, incluindo remoção de ruído, para segmentar lesões cerebrais e gerar máscaras de contorno do cérebro. Este método automático é capaz de segmentar imagens de diferentes aparelhos de escâner e resoluções.

Lee et al. (2008) apresentam técnica que combinam algoritmos clássicos de agrupamento, como  $k$ -médias e *fuzzy c*-médias, para segmentar imagens em três regiões distintas. Os autores utilizam árvores de decisão para analisar a conexão entre os componentes das regiões normais e anormais em imagens de TC do crânio. Estas abordagens não supervisionadas têm apresentado resultados promissores na segmentação de imagens médicas, superando os desafios associados à necessidade de anotações especializadas e ao ruído presente nas imagens.

A segmentação não supervisionada em imagens cerebrais adquiridas por ressonância magnética (RM) tem sido abordada em estudos recentes, como Dalca et al. (2019) que propõem a combinação da inferência Bayesiana com a segmentação probabilística clássica com base em atlas cerebrais, utilizando técnicas de aprendizado profundo. Os autores desenvolvem redes neurais capazes de capturar, delinear e detectar anomalias em imagens cerebrais, permitindo o treinamento do modelo de segmentação para novos exames de RM sem a necessidade de anotação manual. Os experimentos demonstram que o método proposto alcança precisão na tarefa de segmentação, independentemente do contraste da RM.

Mahata et al. (2018) propõem técnica de segmentação *fuzzy* para imagens cerebrais

de RM. A metodologia integra a função gaussiana e informações contextuais locais para estabelecer associações entre píxeis vizinhos. A segmentação é realizada utilizando a função gaussiana para estimar a intensidade heterogênea em cada região de tecido, utilizando os gradientes locais da imagem. Os autores realizam simulações em dois bancos de dados de imagens cerebrais de RM e demonstram que o método proposto é eficiente em relação a outros algoritmos de agrupamento que têm base em lógica *fuzzy*. No entanto, a proposta se limita a estudos de RM.

[Khan et al. \(2021\)](#) propõem método para classificação de tumores cerebrais em imagens de RM, que consiste em três fases: i) pré-processamento, ii) segmentação do tumor cerebral utilizando a técnica de clusterização *k*-médias e iii) classificação dos tumores como benignos ou malignos. Para aumentar a precisão da classificação, os autores incorporam o conceito de aumento de dados sintéticos, expandindo o tamanho dos dados disponíveis para o treinamento do classificador. A abordagem proposta é avaliada nos conjuntos de dados *BraTS* de 2015 e os resultados apresentam a eficiência da estratégia, destacando a clusterização como técnica de segmentação.

[Raja et al. \(2020\)](#) apresentam metodologia para classificação de tumores cerebrais em RM, denominada *autoencoder* profundo híbrido com abordagem de segmentação por agrupamento *fuzzy* bayesiano. Após a segmentação, são extraídas métricas de informação utilizando a transformada de dispersão e métodos de entropia. Este trabalho se destaca por combinar abordagens não supervisionadas com abordagens supervisionadas, mas está limitado a exames de RM. [Hua et al. \(2021\)](#) introduzem técnica aprimorada de clusterização com base no *fuzzy c*-médias para melhorar a precisão de segmentação de imagens cerebrais de RM. A metodologia utiliza mecanismo adaptativo de aprendizado de peso de visualização, permitindo que cada visualização obtenha o peso ideal com base na contribuição do seu grupo. O resultado da segmentação é obtido por meio da combinação das visualizações. Em comparação com outros algoritmos de agrupamento, o método desenvolvido pelos autores demonstra maior adaptabilidade e desempenho.

A segmentação de imagens médicas utilizando RM é amplamente explorada na literatura, com o emprego de abordagens que combinam técnicas de aprendizado supervisionado e não supervisionado. De acordo com [Lenchik et al. \(2019\)](#), o número de segmentações de crânios realizadas em 2019 apresentou uma proporção maior de uso de RM em comparação com a utilização de TC. Estima-se que haja aproximadamente 15,6 aparelhos de TC para cada milhão de habitantes no Brasil. Além disto, os aparelhos de RM apresentam custos de manutenção e depreciação mais elevados

em relação aos equipamentos de TC (ALMEIDA et al., 2019; SANTOS et al., 2017).

A Tabela 1.1 dispõe em ordem cronológica a síntese dos trabalhos relacionados à segmentação de imagens de RM e TC, incluindo o estudo de Kim et al. (2020). Neste estudo, os autores propõem abordagem de segmentação não supervisionada aplicável a diferentes tipos de imagens, não restrita ao contexto médico. A metodologia emprega arquitetura de rede neural profunda com filtros e processos diferenciáveis para agrupar características semelhantes, eliminando a necessidade de treinamento ou anotação manual. No entanto, é necessário adaptar esta metodologia específica de Kim et al. (2020) para a aplicação em exames de TC, levando em consideração as particularidades desta modalidade de exame.

Tabela 1.1 - Trabalhos relacionados à segmentação de imagens de ressonância magnética e tomografia computadorizada.

Proposta/Abordagem	Referência	Contexto/Descrição
Segmentação não supervisionado	Lee et al. (2008)	Imagens de TC cerebrais com abordagem de extração de anormalidades.
Segmentação supervisionada e clusterização	Mahata et al. (2018)	Imagens RM cerebrais com estimativa heterogênea de intensidade utilizando algoritmo de agrupamento <i>fuzzy</i> induzido por informações contextuais locais e função gaussiana.
Segmentação não supervisionada	Dalca et al. (2019)	Imagens RM cerebral com base em abordagem Bayesiana.
Segmentação supervisionada, classificação e clusterização	Raja et al. (2020)	Imagens de tumores cerebrais de RM utilizando <i>autoencoder</i> profundo híbrido com abordagem em agrupamento <i>fuzzy Bayesian</i> .
Segmentação não supervisionada e clusterização	Kim et al. (2020)	Imagens quaisquer com base em agrupamento diferenciável de características.
Segmentação supervisionada, classificação e clusterização	Khan et al. (2021)	Imagens de tumores cerebrais de RM utilizando o algoritmo de agrupamento <i>k</i> -médias, aprendizado profundo e dados sintéticos.
Segmentação supervisionada e clusterização	Hua et al. (2021)	Imagens RM cerebrais utilizando algoritmo de agrupamento <i>fuzzy c</i> -médias com múltiplas visualizações.

Portanto, este trabalho objetiva preencher a lacuna ao adaptar a metodologia de Kim et al. (2020) para a segmentação de exames de TC, utilizando número pré-determinado de máscaras e realizando calibração inicial da rede neural com imagens de referência. Esta abordagem permite a segmentação de imagens semelhantes sem a necessidade de treinamento ou anotação manual. A escolha de exames de TC é justificada pela ampla disponibilidade de equipamentos e maior volume de exames realizados, além dos desafios adicionais de lidar com imagens ruidosas e reduzir o tempo de emissão dos diagnósticos.

A hipótese central deste estudo é: se é possível construir modelo de segmentação não supervisionado utilizando abordagem de ponta a ponta para segmentar estruturas

intracranianas e otimizar os hiperparâmetros da rede, então é possível obter modelo aplicável ao contexto específico de exames de TC. Isto eliminaria as mudanças conceituais realizadas por especialistas, reduziria o custo computacional e facilitaria a identificação de regiões com anomalias, simplificando a tarefa de anotação. O objetivo principal é implementar arquitetura de rede neural profunda para segmentar estruturas intracranianas sem a necessidade de rótulos prévios, anotação manual ou supervisão. Os objetivos específicos são: i) avaliar diferentes técnicas de treinamento de redes neurais, ii) utilizar o processo de otimização para encontrar os melhores valores dos hiperparâmetros da rede, iii) controlar o número de máscaras utilizadas na segmentação e iv) avaliar o desempenho da abordagem proposta, validando os resultados das segmentações com especialistas na área e comparando-os com outros estudos da literatura.

Este trabalho apresenta como originalidade a metodologia não supervisionada de segmentação de imagens médicas, que se destaca pela flexibilidade na definição do número de máscaras e pela inclusão de métricas ajustáveis para cada segmento. A inovação está no processo de otimização dos hiperparâmetros da rede, eliminando a necessidade de rótulos prévios. Como razoabilidade, esta abordagem pretende reduzir custos e tempo na obtenção de resultados, possibilitando diagnósticos e intervenções clínicas mais rápidos e eficientes. A relevância está na busca pela simplificação dos modelos supervisionados de segmentação, que demandam anotação manual e alto poder computacional. Por fim, a aplicabilidade do modelo proposto tem potencial para ser estendida para outras áreas médicas.

A estrutura deste trabalho é dada por: no Capítulo 2 apresenta a história e a contextualização da tomografia computadorizada, além de abordar aspectos anatômicos intracraniano humano. O Capítulo 3 descreve métodos e tecnologias de inteligência artificial aplicadas. O Capítulo 4 descreve brevemente conceitos relacionados a otimização, sistemas, modelos e simulação. O Capítulo 5 descreve a metodologia proposta. O Capítulo 6 apresenta os resultados obtidos através da aplicação da metodologia proposta e finalmente o Capítulo 7 apresenta a conclusão e as sugestões de melhorias como trabalhos futuros.

## CAPÍTULO 2

### TOMOGRAFIA COMPUTADORIZADA

Este capítulo descreve as fundamentações necessárias para entendimento dos conceitos teóricos e físicos utilizados na geração de exames de tomografia, apresentando o histórico da tomografia computadorizada, os planos axial, coronal e sagital, o cérebro humano e suas estruturas com identificação dos tecidos cerebrais e por fim, os protocolos necessários para a realização de exames de imagens médicas.

#### 2.1 Histórico da tomografia computadorizada

Wilhelm Conrad Roentgen em 1895 descobre o raio-X (RX), que desde então é explorado em várias áreas. A medicina é a área que mais utiliza o RX, em aplicações como tratamentos de beleza, troca de cabelo, exame de imagens e várias outras (LIMA et al., 2009). Desde a década de 1890 os hospitais passam a ter departamentos de radiologia e os médicos são treinados para realizar a interpretação das imagens geradas pelos RX. Na década de 1936, o Brasil contribui com o diagnóstico e tratamento de tuberculose pulmonar a partir de imagens radiológicas (CAPA, 2009).

Em 1967 é desenvolvido o primeiro aparelho de tomografia computadorizado (TC), por Godfrey N. *Hounsfield*, que também determina o coeficiente de atenuação. Este coeficiente é a medida linear resultante do valor de radiação absorvido pelos tecidos, visto que o valor da intensidade do feixe de entrada é conhecido, basta conhecer o valor do feixe de saída captado pelos sensores do aparelho de tomografia para determinar o cálculo do coeficiente de atenuação. Assim, os valores de atenuação mais concentrados em determinados voxels<sup>1</sup> têm tons de cinza mais claros na imagem reconstruída e os menores valores de atenuação são exibidos em tons mais escuros (JÚNIOR; YAMASHITA, 2001; LOPES, 2016).

Desta forma, é definida a escala *Hounsfield* [HU], que passa a ser utilizada como referência, considerando o valor de  $-1.000$  HU para o ar, que é o material menos absorvente dos feixes de RX. Para os tecidos que mais absorvem RX, os valores designados são positivos iniciando na água que possui valor estabelecido em  $0$  HU. A escala varia de  $-1000$  HU até  $1000$  HU (LOPES, 2016; OSBORNE et al., 2016). A Tabela 2.1, adaptada de Kamalian et al. (2016), dispõe os valores de atenuação de alguns materiais e tecidos utilizados em tomografia computadorizada, em que  $\mu$

---

<sup>1</sup>Um *voxel* é equivalente a um píxel em 3D e representa o neologismo de *volume* e *píxel*. Cada píxel da imagem está associado a um *voxel*. O valor associado a cada píxel representa a média das atenuações do RX no volume interno do corpo correspondente ao *voxel*.

representa coeficiente de atenuação, para fótons de RX de  $60\text{ keV}$ , que é a diminuição da transmissão do feixe de RX ao atravessar os tecidos.

Tabela 2.1 - Escala *Hounsfield* em materiais para tomografia computadorizado.

Substância	[HU]	$[\mu(\text{cm}^{-1})]$
Ar	-1000	0,0004
Gordura	-100	0,185
Água	0	0,206
Substância Branca	20 - 30	0,213
Substância Cinza	36 - 46	0,212
Sangue	56 - 76	0,208
Osso	1000	0,528

No entanto, há alguns órgãos humanos que não absorvem radiação impossibilitando a geração da imagens. A solução encontrada é a ingestão de contraste, o que possibilita mapear nas imagens estes órgãos (LIMA et al., 2009). Na imagem digital codificada no computador a partir do exame de tomografia, cada píxel é transformado no número CT ou escala *Hounsfield*. O número CT dos tecidos é calculado por (2.1), em que a constante  $K = 1000$  é utilizada para que cada número CT seja equivalente a 0,1% da escala estabelecida por *Hounsfield*. Desta forma, os números CT são organizados com base na atenuação da água (LOPES, 2016).

$$CT = \left( \frac{\mu_{\text{tecido}} - \mu_{\text{água}}}{\mu_{\text{água}}} \right) \cdot K \quad (2.1)$$

Em 1971 é realizado o primeiro exame de crânio através de tomógrafo composto de sistemas de rotação associados a ampolas de anodos com feixes de RX e detectores, esta disposição física dos componentes permite a completa varredura do corpo do paciente em  $\approx 5[\text{min}]$ . A Figura 2.1(a), adaptada de Orth (2017), ilustra o funcionamento simplificado do equipamento de tomografia com rotação da fonte de RX e do sensor detector. A Figura 2.1(b), adaptada de Osborne et al. (2016), ilustra o objeto a ser examinado posicionado estático ao centro do aparelho e os feixes de RX atravessando o objeto (CARVALHO, 2007; MEDEIROS, 2009).

Na TC, feixes de RX atravessam o corpo e são detectados por sensores, que os convertem em sinais elétricos. Estes sinais são reconstruídos por computador, gerando imagens volumétricas em fatias à medida que o tubo de RX gira ao redor do corpo em ângulo de  $360^\circ$  (BAFARAJ et al., 2021). Equipamentos modernos normalmente produzem imagens com espessura média de  $0,5\text{mm}$ . Os avanços em imagens e tecnologias médicas levam ao desenvolvimento de protocolos específicos para diagnósticos pre-

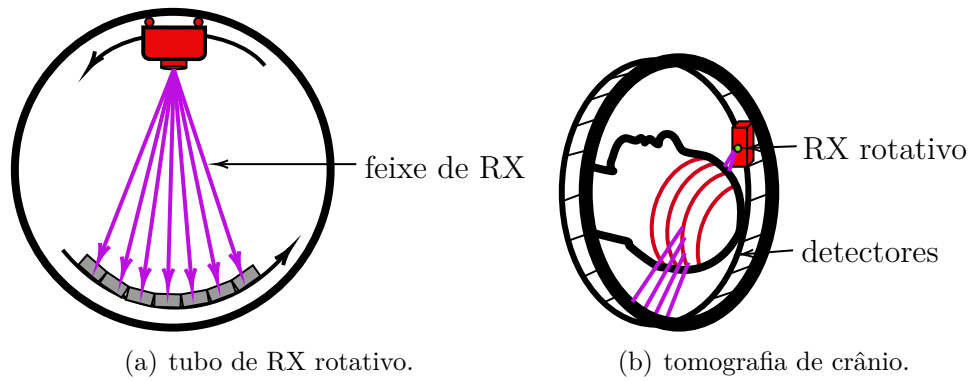


Figura 2.1 - Simplificação da tomografia computadorizada.

cisos, combinando equipamentos, protocolos e inteligência artificial (IA) (NOWAK; RÜGER, 2010). As imagens digitais produzidas por estes aparelhos são utilizadas em formatos padronizados, como o padrão Imagem Digital e Comunicações em Medicina (*Digital Imaging and Communications in Medicine - DICOM*), que estabelece diretrizes de armazenamento e comunicação, independentemente do fabricante do equipamento (MILDENBERGER et al., 2002).

## 2.2 Plano axial, coronal e sagital

As imagens geradas a partir de RX tradicional são projetadas, enquanto que as imagens obtidas na tomografia computadorizada são geradas a partir do tubo de RX, como ilustrado na Figura 2.1(a). Pela dinâmica da emissão dos RX em volta do corpo, pois o tubo gira em 360 graus, as imagens são obtidas em fatias. Posteriormente, técnicas de processamento digital de imagens são utilizadas para transformar as aquisições transversais no plano sagital, visto que fisicamente as imagens tomográficas são adquiridas no plano axial ou coronal. Além destas transformações, há a possibilidade da visualização do agrupamento das fatias no aspecto tridimensional (CARLOS, 2002; GARIB et al., 2007).

Segundo Garib et al. (2007), a aquisição do exame diz respeito quanto à posição do paciente no momento da obtenção dos dados para construção das imagens que compõem o exame, como ilustrado na Figura 2.2, adaptada de Tariq e Burney (2014), sendo os planos de aquisição: i) sagital, ii) coronal e iii) axial.

Plano sagital refere-se a imagens no sentido anteroposterior, divide o corpo em partes à direita e à esquerda em plano longitudinal. Toma-se o plano que separa o corpo em dois lados iguais, a esquerda e a direita, logo as imagens terão visões das laterais

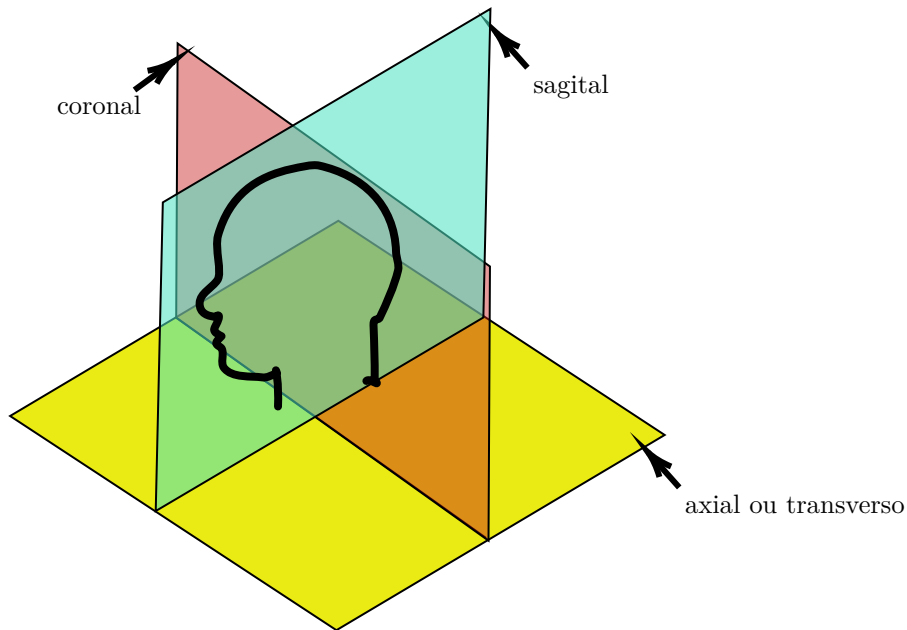


Figura 2.2 - Plano de aquisição.

do corpo, como apresentado na Figura 2.3, produzida por OpenAI (2023). O plano Coronal, também longitudinal, divide o corpo em partes anteriores e posteriores, ou seja, há a divisão que separa o corpo lateralmente, iniciando na orelha e terminando na outra. Já o plano Axial refere-se ao plano horizontal transversal que divide o corpo em parte inferior e superior, logo imagens neste plano anatômico têm a visão superior das partes examinadas do corpo humano (AL-JABBAR et al., 2022). Por fim, o plano anatômico possibilita ao especialista ter ampla noção para gerenciar e posicionar o paciente com precisão, a fim de executar as intervenções necessárias quanto aos diagnósticos, seja cirúrgicos ou radiológicos (CHEN, 1988).

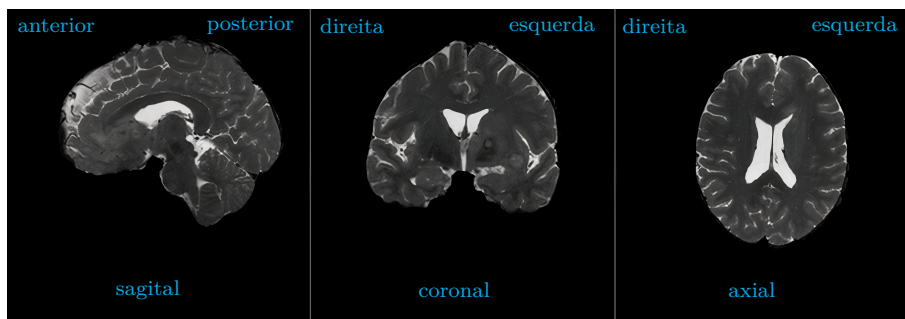


Figura 2.3 - Planos de aquisição.

### 2.3 Imagens de tomografia computadorizada × ressonância magnética

A TC e a RM são modalidades de imagem amplamente utilizadas para o diagnóstico de doenças neurológicas. A TC emprega RX para produzir imagens transversais detalhadas do corpo humano, enquanto a RM utiliza fortes campos magnéticos e ondas de radiofrequência para gerar imagens com resolução de alto contraste, permitindo a visualização de tecidos moles e a distinção entre diferentes tipos de tecidos com base em suas propriedades magnéticas (BAFARAJ et al., 2021; VINEGAR, 1986). Ao contrário da TC, a RM não envolve radiação ionizante, tornando-a mais segura em termos de exposição à radiação. No entanto, a TC apresenta vantagens distintas, como aquisição de imagem mais rápida e menor custo, tornando-a mais acessível em clínicas populares e instituições públicas (BAFARAJ et al., 2021). A Figura 2.4, adaptada de Le et al. (2020), ilustra as diferenças entre TC e a RM realizadas no mesmo indivíduo e na mesma visão transversal do cérebro.

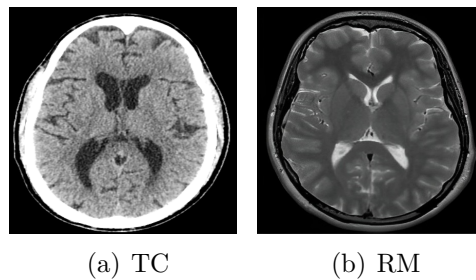


Figura 2.4 - Exames de crânio.

## 2.4 Cérebro humano

O cérebro humano é o principal órgão do sistema nervoso juntamente com a medula espinhal, e formam o sistema nervoso central (SNC). Enquanto a medula realiza a conexão dos outros órgãos do corpo humano com o cérebro. O sistema nervoso central recebe e envia informações para o cérebro, que através das células nervosas, neurônios, tem a capacidade de processar os diversos tipos de sinais e ações, sejam estas ações voluntárias ou involuntárias. Como por exemplo, a ação voluntária de caminhar e a ação involuntária de respirar (BEAR et al., 2002).

### 2.4.1 Estruturas cerebrais

O cérebro humano é composto por várias estruturas interconectadas que estão envolvidas em funções cognitivas superiores, formação de memória, processamento

sensorial, regulação autonômica e endócrina, coordenação motora e funções vitais. O cérebro é compreendido por diencéfalo, tronco cerebral (mesencéfalo, ponte e bulbo) e cerebelo e engloba os hemisférios cerebrais, os gânglios da base e os tratos de substância branca (ANDREASEN et al., 1993; MARTIN; CHAO, 2001). Maldonado e Alsayouri (2021) destacam que cada hemisfério cerebral é dividido em cinco lobos listados em ordem decrescente de tamanho: i) lobo frontal, ii) lobo temporal, iii) lobo parietal, iv) lobo occipital e v) lobo insular. O corpo caloso serve como a conexão entre os dois hemisférios, exhibe dobras características conhecidas como giros e sulcos, que aumentam sua área de superfície para processamento de informações (GAZZANIGA, 2000).

As localizações das estruturas cerebrais e seus funcionamentos demonstram desafio constante para a ciência. Até o final do século XVIII acreditava-se que o cérebro era composto de vários outros órgãos autônomos. A partir do século XIX, novas teorias se tornam referência anatômica e científica através de mapas das áreas cerebrais e da numeração das unidades funcionais. Finalmente, através de pesquisas não invasivas por meio de ressonância magnética comprova-se que as funções de cada área cerebral não são autônomas como acreditava-se, mas que possuem sistemas e funções interligados (AMTHOR, 2017; SANTOS, 2002).

Para Wenger et al. (2017), estudar e entender as estruturas cerebrais é importante pois estas regiões sofrem mudanças por diversos fatores, entre eles a aquisição de habilidades, visto que este processo faz com que haja expansão da estrutura da substância cinzenta e renormalização durante a prática continuada. Zeng et al. (2015) afirmam que as estruturas intracranianas de fetos com cardiopatia congênita são menores que as estruturas de fetos sem esta enfermidade. Kühn e Gallinat (2014) concluem que há variação volumétrica no estriado de indivíduos que fazem o uso frequente de pornografia. Estes trabalhos são exemplos de transformações que podem ocorrer nas regiões cerebrais. Portanto, estudar e entender as estruturas intracranianas é importante, dada a quantidade de fatos relevantes envolvendo o tema.

A Figura 2.5(a), adaptada de Amthor (2014), ilustra a superfície do cérebro. O neocórtex realiza as atividades cognitivas humanas e ele compreende todas as partes da figura, exceto o cerebelo. O encéfalo é a região do SNC que compreende: i) o cérebro, ii) tálamo, iii) mesencéfalo, iv) ponte cerebelo e v) bulbo. Paglioli (1936) demonstra que a parte externa do cérebro é composta por sulcos e circunvoluções com diferentes formatos e tamanhos e a parte superior desta região é constituída

por dois hemisférios, cada qual com quatro lobos: i) frontal, ii) parietal, iii) occipital e iv) temporal.

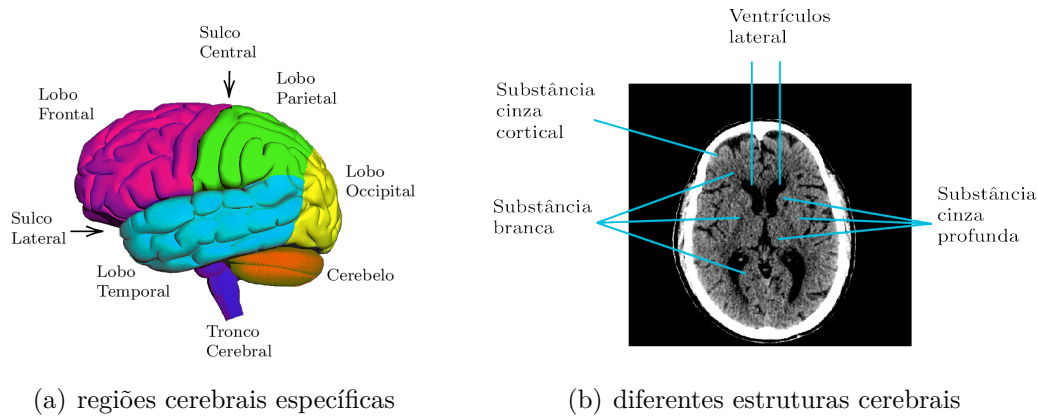


Figura 2.5 - Mapeamento cerebral.

Cada lobo refere-se a distintas atividades (SANTOS, 2002). E o lobo frontal desempenha as principais funções de coordenações motoras, ou seja, a motricidade tanto voluntária quanto automática: i) de coordenar os pensamentos, ii) a escrita e iii) fala, como também coordena funções relacionadas ao caráter humano e ao humor. O lobo temporal é responsável pelas funções: i) da memória, ii) da compreensão e iii) audição, além de estar relacionado com as funções psíquicas. O lobo temporal também é responsável pelo processamento dos eventos e por memorizá-los por caminhos emocionais. Enquanto que o lobo parietal é responsável por coordenar as sensações em geral, incluindo: i) algésica, ii) táctil e iii) térmica, conhecidas como sensibilidade superficiais ou exteroceptiva. O lobo parietal também é relacionado com as sensibilidades profundas, ou proprioceptiva, como: i) vibrações, ii) pressão, iii) dores nas articulações e iv) músculos (SANTOS, 2002).

Observa-se na Figura 2.5(a) as fissuras presentes no neocórtex, chamadas de sulcos e os cumes entre as fissuras são denominadas de giros. Os sulcos podem separar regiões cerebrais por funções: i) sulco central e ii) sulco lateral. O primeiro, separando o lobo frontal do lobo parietal e o segundo, separando o lobo frontal do lobo temporal (GUSMÃO et al., 2001). O cérebro é formado por dois lobos parecidos, o esquerdo e o direito, chamados de hemisfério esquerdo e hemisfério direito. O esquerdo é responsável por controlar o lado direito do corpo humano, enquanto o direito controla o lado esquerdo do corpo. Tradicionalmente, acredita-se que o hemisfério esquerdo é responsável pelo pensamento lógico e analítico, enquanto o direito é mais criativo

e intuitivo. No entanto, pesquisas recentes apresentam que várias funções cognitivas envolvem ambos os hemisférios trabalhando juntos (GÜNTÜRKÜN et al., 2020).

A Figura 2.5(b) ilustra a imagem axial de tomografia computadorizada cerebral, destacando diferentes estruturas. A substância cinza do cérebro é a que contém os corpos celulares neuronais. Dentro do cérebro, as duas principais localizações da substância cinza estão na superfície dos giros, conhecida como substância cinza cortical, e nos núcleos dos gânglios da base. Em contraste, a substância branca consiste em tratos de fibras compreendendo axônios neuronais. A substância cinza apresenta-se com densidade superior à da substância branca, manifestando-se com tonalidade levemente mais clara. Na Figura 2.5(b), a diferença na tonalidade possibilita a identificação das estruturas em exames de imagem (BUDDAY et al., 2015; MERCADANTE; TADI, 2020). Além disto, há no cérebro o sistema ventricular que consiste em ventrículos cheios de líquido cefalorraquidiano (LCR), sendo: i) dois ventrículos laterais, ii) o terceiro ventrículo na linha média e iii) o quarto ventrículo localizado entre a ponte do tronco encefálico e o cerebelo, visível somente em cortes baixos (BRINKER et al., 2014; DOMÍNGUEZ et al., 2016; JOHNSTONE et al., 1976; SAKKA et al., 2011; STRATCHKO et al., 2016). A segmentação das estruturas cerebrais contribui para o diagnóstico de doenças neurológicas (BEAR et al., 2020; KANDEL et al., 1991).

Outra importante ferramenta utilizada para melhorar a compreensão das estruturas cerebrais são os atlas de anatomia, atualmente gerados a partir de imagens médicas nas quais há a possibilidade de visualizar, identificar e atribuir as funções de pequenas estruturas. Os atlas são sistemas de coordenadas, como por exemplo o atlas de *Talairach*. Desenvolvido para facilitar o mapeamento das estruturas cerebrais independente do indivíduo, este sistema de grade é proporcional ao tamanho do cérebro (TALAIRACH; SZIKLA, 1980). Desta forma, o atlas é a referência ou gabarito para que imagens de outros cérebros sejam comparadas, como ilustrado na Figura 2.6, adaptada de Talairach e Szikla (1980), apresentando o sistema hipotético de atlas cerebral, com coordenadas abscissas que se repetem no meio da imagem, indicando a similaridades entre os dois hemisférios cerebrais. Facilitando assim, a identificação das posições dos tecidos, regiões e estruturas cerebrais (GONZALEZ, 2018).

#### 2.4.2 Tecidos cerebrais

Além das estruturas e regiões intracranianas, esta importante e complexa parte do corpo também é composta de tecidos específicos. A Figura 2.7 ilustra a fatia do exame de tomografia computadorizado de crânio e indica a localização dos tecidos:

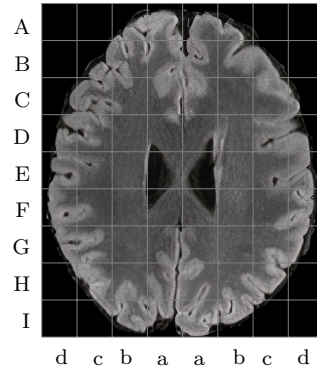


Figura 2.6 - Sistema de atlas cerebral.

i) substância branca e ii) substância cinza e iii) sistema ventricular. A substância branca, tecido branco ou massa branca é formado por axônios, comparados a fios que conduzem os sinais elétricos, através deles que há comunicação entre os neurônios e entre as regiões cerebrais intracranianas. A substância cinza, também conhecido como córtex cerebral é a camada fina e externa composta por neurônios e com elevada ocorrência de dobras, giros e sulcos. Este tecido é responsável por atividades funcionais motoras e desempenhar as funções cognitivas humanas (FILHO; DUNNINGHAM, 2013).

Segmentar a massa branca e cinza pode auxiliar na quantificação do volume cerebral, nos diagnósticos de doenças e anomalias (MILSZTAJN et al., 2003). Algumas abordagens apresentam que a massa cinza é mais dura que a massa branca, enquanto outros estudos dizem o contrário. E segundo os autores Budday et al. (2015), a massa branca é mais densa que a massa cinza do cérebro.

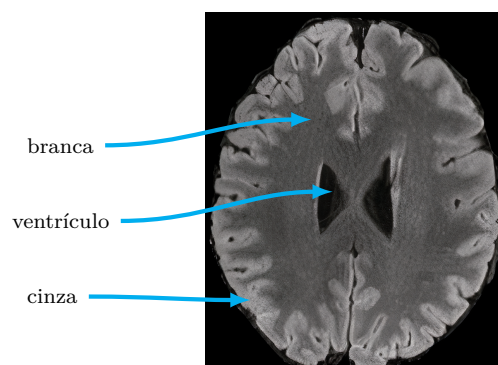


Figura 2.7 - Regiões no cérebro.

## 2.5 Protocolos para exames de imagens

Com o avanço da TC, várias doenças em diferentes partes do corpo podem ser diagnosticadas, contudo, há a necessidade de maior nitidez para observar as regiões vascularizadas e a solução é a utilização de contrastes, que podem ser ingeridos ou injetados (LUCAS et al., 2008). O contraste é uma das técnicas empregadas para melhorar a visualização das regiões durante o exame, enquanto que os equipamentos de TC dispõem de padrões e protocolos preparados pelos fabricantes dos equipamentos, a fim de que os profissionais possam ter maior precisão para cada diagnóstico e aquisição, como por exemplo os tipos de aquisição: i) axial e ii) helicoidal ou multicorte. Assim os padrões e protocolos existem e podem ser utilizados de forma a melhorar a qualidade do exame com o mínimo risco ao paciente.

A aquisição axial é obtida de forma sequencial pelo aparelho de tomografia, no qual cada fatia do exame é obtida a partir da movimentação da mesa. A aquisição helicoidal combina a rotação das ampolas detectoras com o movimento sequencial da mesa, realizando o exame no formato espiral. O aparelho de tomografia multicorte possui fileiras de detectores, adquirindo simultaneamente vários cortes, o que reduz o tempo do exame (AIHARA et al., 2003; MEURER et al., 2008). Dada a quantidade de enfermidades existentes, a quantidade de possíveis exames utilizando imagens e, também, a quantidade de possíveis diagnósticos, surge a demanda da padronização dos protocolos de vários tipos de exames para aquisição de imagens médicas. Para cada doença a ser examinada, há diferentes configurações a serem realizadas nos aparelhos. Outra importante função dos protocolos de exames por imagem é a diminuição da exposição dos pacientes a radiação (MOUNTFORD; TEMPERTON, 1992; ROMANO et al., 2015; SOUZA, 2016).

## 2.6 Janelamento

Nas TC, a imagem digital é representada por valores de intensidade em unidades *Hounsfield* [ $HU$ ], que é a escala numérica que atribui valores a diferentes substâncias e tecidos com base em sua atenuação radiológica (TOGA et al., 2002). As unidades *Hounsfield* são derivadas de transformações linear dos coeficientes de atenuação com base nas radiodensidades do ar com  $0 HU$  e da água pura com  $-1000 HU$  em temperatura e pressão padrão, nas quais cada píxel da imagem de TC recebe o valor de intensidade em  $HU$  (OSBORNE et al., 2016). O processo de visualização através da atribuição dos valores presentes na escala *Hounsfield* é chamado de janelamento.

Esta abordagem permite a separação de vários tecidos, com ossos aparecendo nas

regiões brancas e tecidos moles como músculos e órgãos com valores intermediários, visualizados em tons de cinza (TOGA et al., 2002). As regiões que contêm ar são exibidas em tons mais escuros. Ao aplicar filtros ou limites, tecidos específicos podem ser destacados para visualização. Geralmente, as TC usam imagens de 12 bits capazes de armazenar valores entre  $-1024 HU$  até  $3071 HU$  (GLIDE-HURST et al., 2013). A apresentação destes valores é determinada pela aplicação de configurações específicas de janelamento e nível, permitindo a melhor visualização das estruturas anatômicas em diferentes partes do corpo humano (BRODER, 2011; RAZI et al., 2014). Este janelamento adapta valores na escala de cinza em exames radiológicos aprimorando a visualização de estruturas de interesse. Envolve a seleção de intervalos específicos de densidade e a associação destes intervalos a faixas correspondentes na escala de cinza. O uso adequado deste procedimento resulta em melhorias na identificação das regiões (RADHIANA et al., 2013).

A Tabela 2.1 dispõe a relação entre a escala de diferentes substâncias, o que facilita a criação de protocolos de visualização para partes específicas. A Figura 2.8, adaptada de Tabakov et al. (2020), ilustra a largura de janela de  $400 HU$  com faixa de tons de cinza iniciando em  $-200 HU$  até  $200 HU$  e com o nível da janela com valor 0.

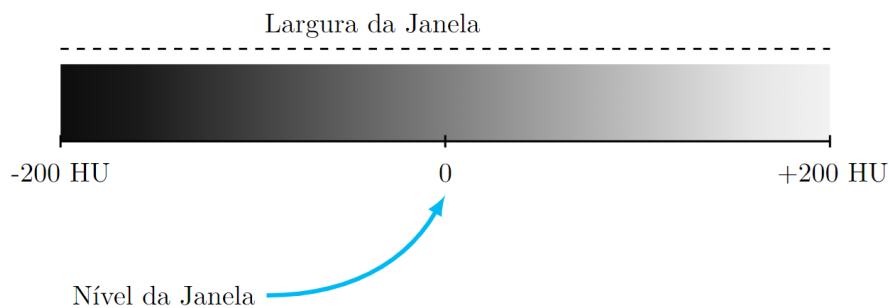


Figura 2.8 - Representação hipotética da largura da janela.

Esta técnica computacional é formada pelo contraste e pela densidade, respectivamente, largura da janela  $L_j$  e nível da janela  $N_j$ . Para  $L_j$  de  $-1000$  até  $1000$  números TC, o centro ou  $N_j$  será 0. Largura de janela com maior abrangência diminui a percepção de tecidos moles e a solução para aumentar a atenuação e ressaltar estas diferenças é a utilização de janelas com menor abrangência, ou seja, com faixas menores, pois partes indesejadas na visualização já são excluídas por estarem fora da largura da janela. A Figura 2.9 dispõe o janelamento em tomografia de crânio com largura da janela de  $640 HU$  a  $1280 HU$ , destacando o contraste presente na imagem (XUE et al., 2012).

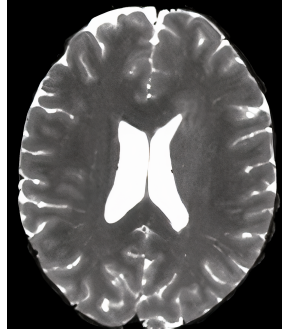


Figura 2.9 - Janelamento em tomografia de crânio.

## 2.7 Anotação e rótulo em imagens médicas

A anotação manual de imagens digitais é realizada por especialistas, que identificam e rotulam as regiões espaciais presentes na imagem, fornecendo descrições e comentários em forma de texto (ALJABRI et al., 2022). Na área médica, a anotação de imagens requer precisão, sendo comum a participação de vários especialistas que realizam anotações independentes (NOWAK; RÜGER, 2010). Após a revisão individual, as anotações são atualizadas e revisadas em conjunto para correções e atualizações adicionais. No entanto, este processo de anotação manual é financeiramente e temporalmente exigente, frequentemente envolvendo a participação de múltiplos profissionais (ALJABRI et al., 2022).

Diversas ferramentas computacionais têm sido desenvolvidas para permitir a anotação automática de imagens, abrangendo desde rótulos de classificação até segmentação píxel a píxel (DIAS et al., 2019). No contexto das imagens médicas, a anotação automática é aplicada para rotular imagens de RM e TC, a fim de treinar modelos de aprendizado de máquina (NOWAK; RÜGER, 2010). Estas ferramentas têm relevância significativa na segmentação de objetos em diferentes órgãos, como substâncias mamárias, nódulos pulmonares, vasos retinianos e tumores no fígado, cérebro, entre outros (ALJABRI et al., 2022).

## 2.8 Imagens médicas

O formato de arquivo em Comunicação de Imagens digitais na medicina (*Digital Imaging and Communications in Medicine* - DICOM) torna-se padrão na medicina, pois é concebido para tratar, manter e integrar informações médicas em protocolo único. Este arquivo agrupa normas para arquivar e manter a comunicação padronizada dos exames independente do fabricante do equipamento ou aparelho que fez a

aquisição (MILDENBERGER et al., 2002). Concebido em 1983, o padrão garante que não há incompatibilidade ou degradação da qualidade dos exames, além de permitir a manipulação do exame e fornece vários dados sobre o paciente, a doença, o hospital entre outros (MILDENBERGER et al., 2002).

Estes dados complementares, ou metadados, são organizados internamente no arquivo DICOM através de chaves, as quais possuem códigos e estão relacionadas a informações específicas. As chaves do dicionário de dados são chamadas de *Tags*. Portanto, há garantia da integridade dos dados do paciente, do exame e da comunicação (CARITÁ et al., 2004). A Tabela 2.2 dispõe algumas *tag* utilizadas no arquivo DICOM, assim como o dicionário de dados referente ao armazenamento da *tag*. Para manter a integridade alguns dados são obrigatórios e outros são opcionais.

Tabela 2.2 - Exemplo de *tags* utilizadas nos arquivos DICOM

Tag	Informação
0010:0010	Id do paciente
0010:0020	Nome do paciente
0010:0030	Data Nascimento do paciente
0020:0010	Id do estudo
0020:0011	Número de série do equipamento

## 2.9 Considerações finais

Este capítulo abordou a fundamentação teórica e conceitos relacionados aos exames de tomografia, principalmente de estruturas encefálicas. Também demonstrou-se a importância de desenvolvimento de protocolos de elaboração dos exames, bem como, do formato das imagens a serem manipuladas. O próximo capítulo descreverá a inteligência artificial, capaz de replicar funções cognitivas humanas.



## CAPÍTULO 3

### INTELIGÊNCIA ARTIFICIAL APLICADA A DIAGNÓSTICO

Este capítulo aborda os conceitos matemáticos e práticos sobre as técnicas de inteligência artificial, especificamente de aprendizado de máquina com ênfase em segmentação no contexto de imagens, a fim de embasar os conhecimentos necessários para entendimento da metodologia aplicada no desenvolvimento deste trabalho.

#### 3.1 Aprendizado de máquina

Segundo Moroney (2020), o aprendizado de máquina pode ser definido como a disciplina que capacita computadores a aprender a partir de dados. Neste contexto, os programas de computador são ensinados a aprimorar-se por meio de experiências, operando dentro de limites específicos para cada tarefa. Assim, é necessário medir a precisão dos resultados destes programas para garantir que eles melhorem com novas experiências. Além disto, o aprendizado de máquina pode ser descrito como algoritmos capazes de realizar abstrações matemáticas a partir de exemplos, diferentemente dos programas convencionais, que são criados a partir de regras definidas em linguagens de programação. Nestes casos tradicionais, a combinação de regras e dados pode produzir as respostas esperadas. As metodologias de aprendizado de máquina utilizam a combinação de dados e possíveis respostas para estabelecer regras, diferenciando-se, assim dos métodos programáticos tradicionais. A Figura 3.1 e Figura 3.2, adaptadas de Moroney (2020), ilustram o fluxo de componentes no caso tradicional com a combinação de regras e dados e o caso do aprendizado de máquina.



Figura 3.1 - Diagrama de alto nível de programas tradicionais.



Figura 3.2 - Diagrama de alto nível de aprendizado de máquina.

A estatística representa uma das ciências fundamentais no aprendizado de máquina. Esta área contribui para a elaboração de programas que analisam dados e tentam prever novos resultados, utilizando métricas específicas para avaliar a precisão. Os modelos de aprendizado de máquina, integrados ao campo da inteligência artificial, têm sua capacidade preditiva diretamente associada tanto à quantidade quanto à qualidade dos dados recebidos. Isto ocorre porque os algoritmos desenvolvem modelos matemáticos com base nos dados para identificar padrões comportamentais, possibilitando assim, previsões acertadas (GÉRON, 2022). Além disto, os algoritmos de aprendizado de máquina são categorizados em dois tipos principais: i) supervisionado e ii) não supervisionado.

### 3.2 Aprendizado supervisionado, não supervisionado e profundo

Existem duas tarefas clássicas no aprendizado supervisionado: classificação e regressão. Em ambas, a rede de aprendizado é treinada com o auxílio de supervisor, utilizando dados rotulados. O algoritmo é continuamente ajustado até conseguir gerar a saída desejada a partir dos dados de entrada rotulados (FLECK et al., 2016). Os classificadores são configurados para prever ou generalizar variáveis categóricas; por exemplo, uma rede neural pode aprender a identificar o rótulo da classe a partir de imagens. Por outro lado, os algoritmos de regressão lidam com variáveis contínuas e visam estabelecer a relação de dependência entre os dados de entrada e os valores preditos, tornando-os adequados para inferências precisas e próximas aos valores reais (HIRAN et al., 2021). Os algoritmos supervisionados aprendem com os exemplos durante o treinamento, definindo padrões matemáticos e estatísticos que relacionam a entrada com a saída. Entre os modelos mais empregados em aprendizado supervisionado estão as redes neurais artificiais (RNA) (HAYKIN, 1996).

Na abordagem de aprendizado não supervisionado, os dados não são rotulados, e o algoritmo identifica padrões para agrupar os dados com base em similaridades, sem a necessidade de supervisão ou definição prévia de respostas corretas. Este tipo de aprendizado é ilustrado na Figura 3.3, adaptada de Qian et al. (2020), contrastando com o aprendizado supervisionado que utiliza rótulos e fronteiras de decisão. Os algoritmos não supervisionados são principalmente utilizados em técnicas de agrupamento e associação. O agrupamento organiza dados não rotulados em grupos por similaridade, útil como pré-processamento para aprendizado supervisionado e em aplicações como a compressão de imagens via quantização de cores (ALPAYDIN, 2020). Os algoritmos de associação criam regras para identificar relações entre itens, sendo amplamente usados no comércio para analisar padrões de compra e prever

comportamentos de consumo (HIRAN et al., 2021).

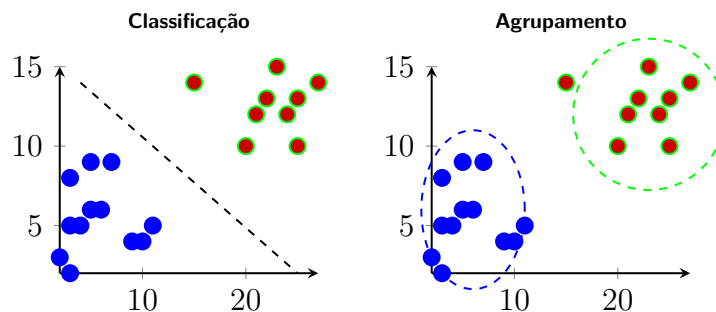


Figura 3.3 - Aprendizado supervisionado e não supervisionado.

O aprendizado profundo, subconjunto do aprendizado de máquina, tem base nas redes neurais complexas que incluem múltiplas camadas de neurônios virtuais. Estas camadas possibilitam o processamento não-linear de dados, com a profundidade das camadas contribuindo para a capacidade de representar e processar informações complexas (CHAGAS, 2019). A principal distinção entre as redes neurais profundas e as tradicionais reside na quantidade de camadas, que permitem abstração e aprendizado sofisticados. Embora as redes profundas sejam inspiradas no córtex visual humano, não se comprovou que o cérebro opere de maneira idêntica a estas estruturas, ainda que a inspiração seja de natureza neurobiológica (CAO et al., 2015; CHOLLET, 2021; MILLSTEIN, 2020).

O aprendizado profundo é amplamente aplicado em áreas como processamento de linguagem natural, visão computacional, medicina e biologia, além de ser utilizado em geração de imagens, jogos e finanças. Este método se diferencia dos modelos superficiais de aprendizado de máquina pela sua arquitetura complexa, que integra funções específicas para extrair características dos dados, calcular discrepâncias entre os resultados desejados e obtidos, e ajustar os pesos de forma otimizada até alcançar alta precisão nas previsões, como ilustra a Figura 3.4. O aprendizado profundo utiliza conjunto de algoritmos para aprender de forma eficiente, independentemente da complexidade, quantidade ou qualidade dos dados (PONTI; COSTA, 2017).

Na Figura 3.4, as técnicas de incorporação densa (*embedding denso*) em redes neurais profundas são utilizadas para transformar dados categóricos, tais como palavras ou identificadores de usuários, em vetores numéricos densos. Estes vetores facilitam a captura de relações semânticas complexas em aplicações como processamento de linguagem natural (Natural Language Processing - NLP), sistemas de recomendação

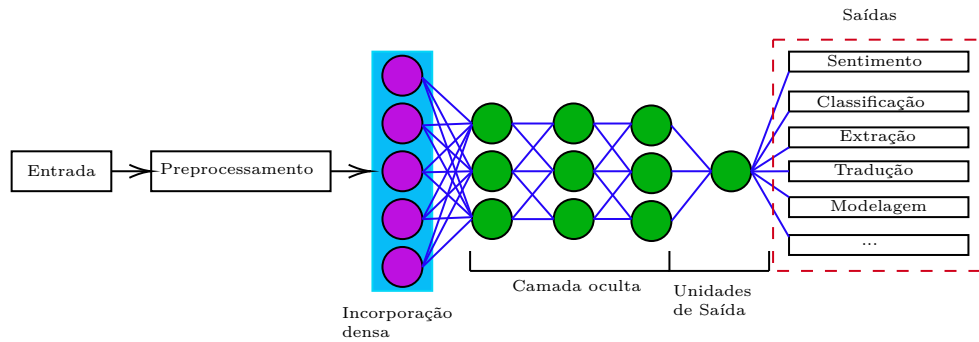


Figura 3.4 - Arquitetura do sistema de aprendizado profundo.

e análise de redes sociais. Redes como as CNN, em visão computacional, redes neurais recorrentes (RNR), em previsão de séries temporais, geralmente não recorrem à incorporação densa, operando diretamente com dados brutos ou características previamente extraídas. Em contextos nos quais a incorporação densa não é aplicada, outros métodos de pré-processamento são empregados para manipular dados categóricos, tais como a codificação *one-hot* ou a normalização, dependendo das características e do formato dos dados.

### 3.3 Redes neurais convolucionais

As redes neurais profundas apresentam variedade de arquiteturas além das convolucionais, incluindo redes densamente conectadas (*Fully Connected Network* – FCN), redes neurais recorrentes (RNR), memória de longo e curto prazo (*Long Short-Term Memory* – LSTM), e *transformers*. Esta diversidade reflete o amplo espectro da inteligência artificial. Particularmente, as redes neurais convolucionais (Convolutional Neural Network - CNN) são reconhecidas por sua eficiência em visão computacional. Nestas redes, as convoluções são operações fundamentais que envolvem a aplicação de filtros ou *kernels* matriciais sobre a imagem para destacar características como arestas, cantos e pontos. Durante o treinamento, os neurônios ajustam seus pesos e vieses para aprender a identificar atributos específicos em cada camada da rede, processo que permite a extração de características relevantes e a modificação significativa da entrada, como ilustrado na Figura 3.5.

A arquitetura da CNN é composta por várias camadas, incluindo camadas de agrupamento ou *pooling*, que têm a função de reduzir a dimensionalidade dos mapas de características. Esta redução é necessária para o processamento de grandes volumes de dados, pois diminui a quantidade de cálculos necessários e ajuda a prevenir o sobreajuste (*overfitting*). O agrupamento pode ser realizado de várias formas, sendo

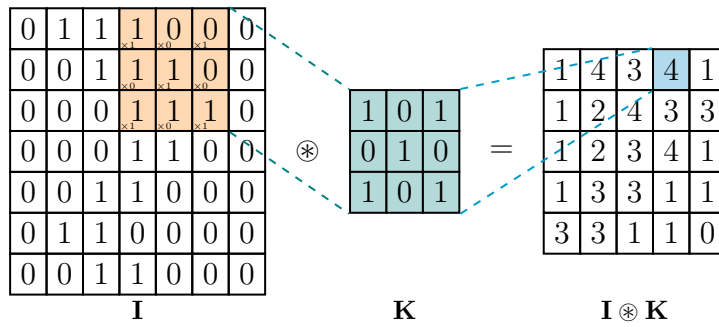


Figura 3.5 - Operação matemática genérica de convolução.

o agrupamento máximo um dos mais comuns, pois seleciona os píxeis de maior significância de cada região do mapa de características, contribuindo para a eficiência do modelo sem que haja perda considerável de informações. Esta capacidade de simplificar os dados de entrada mantendo apenas os aspectos mais relevantes é ilustrada na Figura 3.6 adaptada de (GÉRON, 2019). Ela auxilia na eficiência das redes neurais convolucionais em tarefas complexas de visão computacional (GÉRON, 2022). A combinação de convolução e agrupamento nas redes neurais convolucionais permite que elas sejam eficientes na análise e interpretação de imagens.

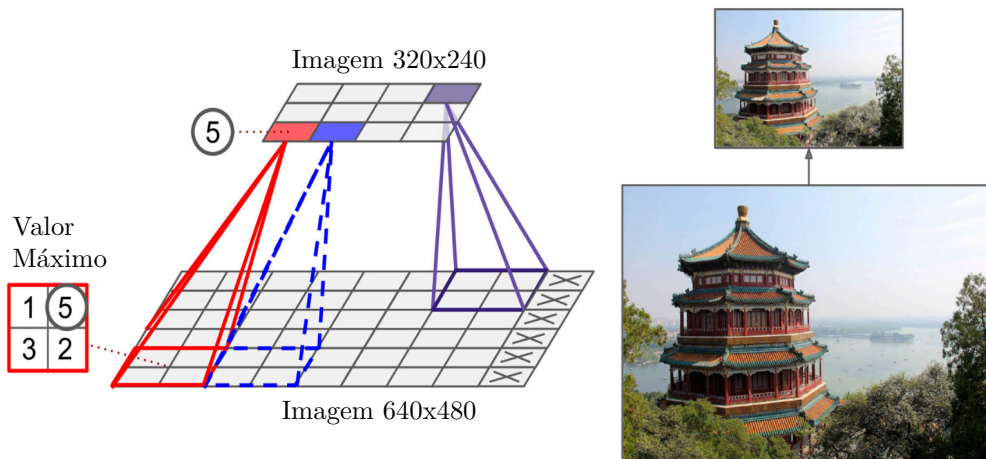


Figura 3.6 - Redução da dimensionalidade da imagem.

Após a extração de características com filtros convolucionais e a subsequente redução de dimensionalidade por meio de agrupamento, as CNN tipicamente empregam algoritmos de classificação nas camadas totalmente conectadas, como ilustrado na Figura 3.7, adaptada de LeCun et al. (1998). A arquitetura padrão destas redes ini-

cia com a conversão da entrada em matriz de píxeis que é processada pela primeira camada convolucional, por exemplo,  $l = 1$ , na qual filtros inicialmente aleatórios são aplicados. O resultado desta operação é encaminhado para a camada de subamostragem  $l = 2$ , que reduz as dimensões da imagem pela metade, por exemplo.

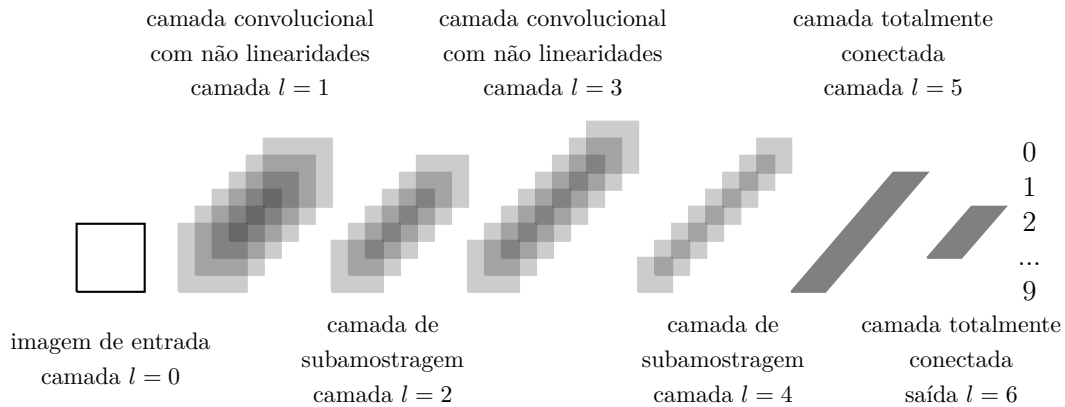


Figura 3.7 - Arquitetura da rede neural convolucional.

Este processo é repetido até que, na camada  $l = 5$ , as imagens processadas sejam preparadas para iniciar a classificação. As arquiteturas das CNN podem ser configuradas para produzir resultados em formatos binários, gerando saídas como 0 ou 1, ou para resultados multiclasse, que envolvem saídas como vetores, matrizes ou outras estruturas de dados mais complexas (LORENA et al., 2008). Assim, a arquitetura representada na Figura 3.7 pode ser adaptada para segmentar a imagem de entrada e gerar a matriz de píxeis como resultado, aprimorando a utilidade do modelo em diversas aplicações (BERSTAD et al., 2018).

### 3.4 Segmentação supervisionada e não supervisionada

A segmentação de imagens é o processo de dividir a imagem em regiões distintas e não sobrepostas, ou identificar objetos específicos contidos na imagem. O principal objetivo da segmentação é simplificar a representação visual, facilitando a análise ao dividir a imagem em regiões que isolam objetos ou bordas relevantes. Esta tarefa pode ser realizada por meio de técnicas que focam em características específicas, como bordas ou regiões (LEI; NANDI, 2022). A segmentação pode ser orientada para objetivos específicos ou executada em imagens sem rótulos predefinidos. Quando os rótulos estão disponíveis, algoritmos de inteligência artificial com aprendizado supervisionado podem ser aplicados para identificar e isolar precisamente os objetos

rotulados nas imagens. Esta abordagem permite que os modelos aprendam a partir de exemplos que deve ser detectado em contextos similares (HO, 2011; LIU et al., 2018; ZHU et al., 2019).

Na segmentação supervisionada, modelos são treinados com dados rotulados fornecidos por especialistas, que indicam as regiões de interesse nas imagens (HO, 2011; ZHU et al., 2019). Em contraste, a segmentação não supervisionada opera sem rótulos predefinidos, utilizando algoritmos para identificar automaticamente padrões, contrastes, texturas e outras características nos dados, agrupando píxeis ou regiões similares sem conhecimento prévio das categorias ou do número de objetos. O objetivo é descobrir a estrutura subjacente ou características distintivas da imagem de maneira autônoma. Exemplo de aplicação da segmentação não supervisionada é o modelo U-Net, proposto por Ronneberger et al. (2015) para segmentação semântica em imagens médicas, ilustrada na Figura 3.8.

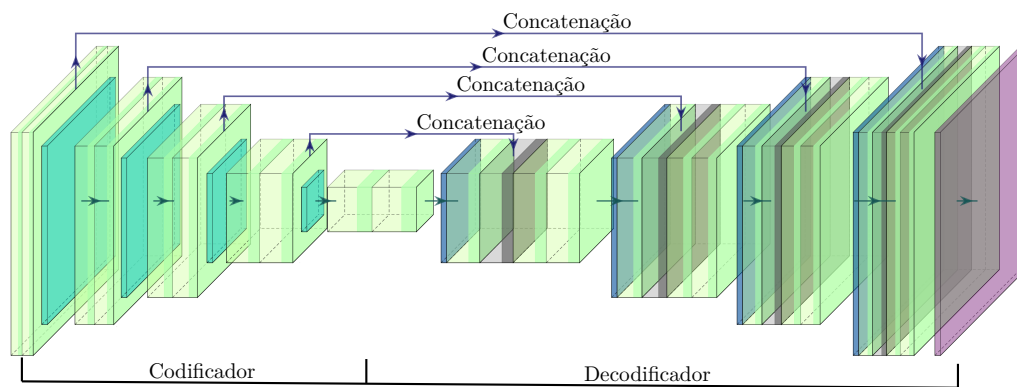


Figura 3.8 - Arquitetura da rede neural profunda U-net.

A abordagem de redes neurais profundas não supervisionadas tem sido amplamente utilizada para a segmentação de estruturas, utilizando a combinação de métodos estatísticos, comparativos e agrupamentos para alcançar resultados precisos (GONZÁLEZ-VILLÀ et al., 2016; BAUR et al., 2018; DALCA et al., 2019). Esta técnica minimiza a variabilidade dos processos de anotação e observação, facilitando a reprodutibilidade dos estudos (CLARK et al., 2017). Os métodos de segmentação não supervisionada aplicam filtros e modelos de busca que agrupam píxeis por similaridade, formando conjuntos nos quais os dados compartilham características como cor entre outras (CAPELLE et al., 2000; HALL et al., 1992; XUE et al., 2003). Na Figura 3.9, ilustra-se o processo de segmentação, no qual o objeto na cor vermelha

está contido no círculo na cor verde, indicando sua seleção através do agrupamento por similaridade.

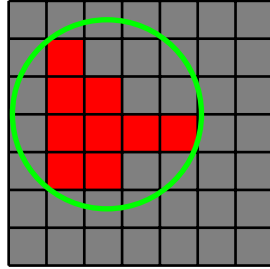


Figura 3.9 - Expansão de escala para observação dos píxeis.

### 3.5 Função de avaliação, de perda e entropia cruzada

Embora os termos função de custo ou de avaliação e função de perda sejam frequentemente usados de forma intercambiável em aprendizado de máquina, eles podem ter significados distintos dependendo do contexto. A função de custo é a média das funções de perda aplicadas a todo o conjunto de treinamento, representando o custo médio dos erros sobre todos os dados treinados. Por outro lado, a função de perda mede o erro para única instância de treinamento, quantificando o desvio da previsão em relação ao valor real (CHAI; DRAXLER, 2014; WILLMOTT; MATSUURA, 2005). Em contextos de aprendizado não supervisionado, no qual o modelo não conta com valores esperados ou dados rotulados, estas funções orientam o treinamento ao definir comportamentos desejados e monitorar erros nas previsões e no desempenho geral do modelo.

Exemplos de funções de perda incluem a função  $L_1$  dada por (3.1) e a função  $L_2$  dada por (3.2), sendo  $L_2$  particularmente sensível a valores discrepantes. A função de avaliação  $f(x)$  que pode ser representada pelo erro absoluto médio (MAE), é dada por (3.3). As variáveis  $y_i$  e  $\hat{y}_i$  representam, respectivamente, o valor real e o valor calculado nas expressões. A entropia cruzada, dada por (3.4), é alternativa ao erro quadrático médio (EQM), e frequentemente é utilizada em classificadores para medir a distância entre a distribuição de probabilidade calculada e a distribuição real, permitindo ajustes de peso mesmo em casos de saturação (GOLIK et al., 2013; KLINE; BERARDI, 2005; ZHANG; SABUNCU, 2018). A entropia cruzada multiclasse, dada por (3.5), é específica para classificações com múltiplas classes, calcula o desempenho do modelo somando o produto dos logaritmos das probabilidades previstas pelas etiquetas verdadeiras, auxiliando o modelo a se ajustar para otimizar os valores de

saída (KULMANOV et al., 2021; MITRA et al., 2002).

$$L_1 = \sum_{i=1}^D |y_i - \hat{y}_i| \quad (3.1)$$

$$L_2 = \sum_{i=1}^D (y_i - \hat{y}_i)^2 \quad (3.2)$$

$$f(x) = \frac{1}{D} \sum_{i=1}^D |y_i - \hat{y}_i| \quad (3.3)$$

$$H(p, q) = - \sum_{c=1}^M p(x_c) \log q(x_c) \quad (3.4)$$

$$H_m(p, q) = - \sum_{i=1}^N \sum_{c=1}^M p(x_{i,c}) \log q(x_{i,c}) \quad (3.5)$$

Nas expressões (3.4) e (3.5),  $M$  é o número de classes possíveis,  $p(x_c)$  é a probabilidade verdadeira da classe  $c$ ,  $q(x_c)$  é a probabilidade calculada (prevista) pelo modelo para a classe  $c$ ,  $N$  e  $D$  são os números totais de amostras,  $p(x_{i,c})$  é a probabilidade verdadeira da amostra  $i$  pertencer à classe  $c$  e  $q(x_{i,c})$  é a probabilidade calculada pelo modelo de que a amostra  $i$  pertence à classe  $c$  (RUBINSTEIN; KROESE, 2004). Em técnicas não supervisionadas de aprendizado profundo para segmentação de estruturas, a função de avaliação realiza comparações internas no algoritmo. Estas comparações facilitam as segmentações ao identificar diferenças detectadas, que são originadas das saídas dos filtros convolucionais, permitindo a análise de semelhanças entre píxeis, pontos de interesse e as bordas que definem as regiões ou objetos na imagem.

### 3.6 Otimização dos hiperparâmetros

O processo de otimização de modelos em RNP ocorre durante o treinamento, no qual os hiperparâmetros críticos são ajustados. Estes hiperparâmetros, que definem as configurações do modelo, não são aprendidos diretamente dos dados, mas escolhidos para influenciar a eficiência e comportamento do modelo (FEURER; HUTTER, 2019). Alguns dos otimizadores mais utilizados incluem o *Adaptive Moment Estimation* (Adam), que combina características do *Root Mean Square Propagation* (RMSprop)

e do *Stochastic Gradient Descent* (SGD), este último sendo eficiente em grandes conjuntos de dados (AGRAWAL, 2021; BOTTOU; BOUSQUET, 2007; HINTON, 2018; KINGMA; BA, 2014).

A otimização foca na minimização da função de perda, ajustando os hiperparâmetros para alcançar os melhores resultados (SRINIVAS; KATARYA, 2022). Ferramentas como `Optuna` auxiliam na otimização, utilizando técnicas como árvores de decisão e otimização bayesiana para navegar no espaço dos hiperparâmetros e identificar as melhores combinações, empregando o estimador de densidade de probabilidade de Parzen para modelar a distribuição de probabilidade dos hiperparâmetros (AKIBA et al., 2019; AGRAWAL, 2021; BERGSTRA et al., 2011). O `Optuna` facilita a integração com outras plataformas através de interface de programação de aplicativos, melhorando a interação entre diferentes softwares.

### 3.7 Considerações finais

Este capítulo explorou a hierarquia entre inteligência artificial, aprendizado de máquina e aprendizado profundo, destacando as diferenças entre o aprendizado supervisionado e não supervisionado. Foram abordados o funcionamento e a importância das redes neurais artificiais, que constituem a base do aprendizado profundo. Além disto, discutiram-se as redes neurais convolucionais na segmentação de imagens, a relevância das funções de avaliação e os métodos de otimização de hiperparâmetros. No próximo capítulo, será apresentado o modelo de segmentação de imagens que é a base metodologia proposta.

## CAPÍTULO 4

### MODELOS DE SEGMENTAÇÃO NÃO SUPERVISIONADO

Neste capítulo, são apresentados os conceitos e as representações relacionados a trabalhos que utilizam modelos de segmentação não supervisionados em imagens de diferentes tipos. Dois modelos descritos na literatura são brevemente introduzidos, com exposição de suas variáveis no contexto de sistemas de segmentação de imagens.

#### 4.1 Modelos não supervisionados aplicados em segmentação de imagem

Diversos modelos de aprendizado de máquinas não supervisionados são aplicados à segmentação de imagens. Exemplos incluem o agrupamento  $k$ -médias (*k-means clustering*), que segmenta os píxeis da imagem em grupos com base em suas características, sem a necessidade de rótulos pré-existentes (SINAGA; YANG, 2020). Outro exemplo são os modelos de mistura gaussiana (*Gaussian Mixture Models*), que utilizam a combinação de distribuições gaussianas para modelar a distribuição dos píxeis na imagem, agrupando-os sem supervisão (FIGUEIREDO; JAIN, 2002). O deslocamento da média (*Mean Shift*), tem base na densidade dos píxeis, encontra modos na distribuição dos dados para formar agrupamentos (MURE et al., 2015).

O agrupamento espacial com base em densidade (*Density-Based Spatial Clustering*) identifica regiões de alta densidade de píxeis, separando-as das regiões de baixa densidade (ESTER et al., 1996). O agrupamento espectral (*Spectral Clustering*) utiliza as propriedades do espectro da matriz de similaridade da imagem para agrupar os píxeis sem a necessidade de rótulos (LAW et al., 2017). Modelos como o agrupamento hierárquico (*Hierarchical Clustering*) formam árvores de agrupamentos sem rótulos, permitindo a segmentação ao cortar a árvore em nível específico (GOLDBERGER; ROWEIS, 2004). Os mapas auto-organizáveis (*Self-Organizing Maps*) mapeiam dados de alta dimensão para representação de menor dimensão, facilitando a segmentação (WICKRAMASINGHE et al., 2019).

O algoritmo da bacia hidrográfica (*Watershed Algorithm*) utiliza a analogia de inundar a imagem e usa as linhas de contorno para segmentar diferentes regiões sem supervisão explícita (SANTANA et al., 2014). Atualmente, as redes neurais convolucionais (RNC) são predominantemente utilizadas em métodos supervisionados de segmentação de imagens e ultimamente têm sido usadas em segmentação não supervisionada. Exemplos incluem autocodificadores convolucionais (*Convolutional Autoencoders*), que são redes neurais que aprendem a representação compacta das

imagens de entrada e podem ser usadas para segmentação ao identificar características semelhantes nos dados não rotulados (RODRÍGUEZ, 2021). Os autocodificadores variacionais (*Variational Autoencoders*) são a extensão dos autocodificadores convolucionais que modelam a distribuição dos dados e são utilizados para segmentação não supervisionada (ZIMMERER et al., 2019).

As redes geradoras adversárias (*Generative Adversarial Networks*), embora mais comumente usadas em contextos supervisionados, possuem versões não supervisionadas ou semi-supervisionadas que podem ser adaptadas para segmentação de imagens (RADFORD et al., 2015). O aprendizado auto-supervisionado (*Self-Supervised Learning*) utiliza tarefas auxiliares para aprender representações úteis sem rótulos explícitos, que podem ser utilizados posteriormente para segmentação (MISRA; MATEN, 2020). O agrupamento profundo incorporado (*Deep Embedded Clustering*) combina redes neurais convolucionais com técnicas de agrupamento, aprendendo representações e realizando agrupamentos de forma conjunta (AJAY et al., 2022).

As redes de agrupamento com restrição espacial (*Spatially Constrained Clustering Networks*) integram convoluções com agrupamentos espaciais para segmentar imagens sem supervisão (DIPLAROS et al., 2007). A segmentação por retropropagação (*Segmentation by backpropagation* – SBB) utiliza retropropagação para otimizar diretamente as máscaras de segmentação na rede neural (KANEZAKI, 2018). A segmentação com base em agrupamento de características diferenciáveis (*Segmentation based on differentiable feature clustering* – DFC) combina a extração de características por RNC com agrupamento diferenciável para segmentação não supervisionada (KIM et al., 2020).

## 4.2 Rede neural convolucional para segmentação de imagens

A maioria dos modelos de segmentação não supervisionada utiliza imagens genéricas, sem critérios específicos que façam sentido ao escopo das imagens. As propostas não supervisionadas utilizando RNC podem ser entendidas como sistemas com critérios e restrições, como a SBB de Kanezaki (2018), que investiga a utilização de RNC para segmentação de imagens sem supervisão. A SBB gera rótulos aos píxeis que fazem parte do mesmo agrupamento ao qual o píxel pertence, e serve como base para a DFC de Kim et al. (2020), que emprega melhorias no processo de segmentação.

A DFC utiliza imagem-alvo para otimizar os rótulos dos píxeis em conjunto com as representações de características enquanto os parâmetros da rede são atualizados. O processo alterna entre a previsão de rótulos e a aprendizagem dos parâmetros da

rede, atendendo os critérios: i) píxeis com características semelhantes são atribuídos aos mesmos rótulos, ii) píxeis espacialmente contínuos são atribuídos aos mesmos rótulos e iii) rótulos únicos. O intuito da *DFC* é criar segmentos de imagem coerentes e distintos, aproveitando métodos de aprendizado profundo, introduzindo a função de perda da continuidade espacial e técnicas de agrupamento de características diferenciáveis sem a necessidade de dados rotulados. Diferente da *SBB*, a *DFC* inclui normalização na rede e função de agrupamento diferenciável (KIM et al., 2020).

A proposta da *DFC* é aplicada em imagens padrão para a avaliação de desempenho de segmentação, utilizando bases como BSD500<sup>1</sup>, COCO<sup>2</sup> e Pascal VOC<sup>3</sup>. A metodologia da *DFC* é descrita como: a) **extração de características** das imagens com representações de alto nível dos píxeis, b) **agrupamento diferenciável** que é a novidade, permitindo que o processo de agrupamento seja integrado ao treinamento da rede neural, inspirada no algoritmo *k*-médias, mas adaptada para possibilitar a retropropagação do erro durante o treinamento e c) **função de perda** projetada para minimizar a variação dentro dos segmentos e maximizar a variação entre eles, similar ao objetivo do *k*-médias, e considerando a continuidade espacial, incentivando para que píxeis próximos sejam agrupados no mesmo segmento (KIM et al., 2020).

Os resultados experimentais indicam que a abordagem da *DFC* consegue segmentar imagens de forma eficiente, criando segmentos que correspondem aos objetos visuais, apresentando desempenho competitivo em comparação com outros métodos não supervisionados e alguns métodos supervisionados. Em comparação com os resultados da *SBB*, a *DFC* apresenta solução para a atribuição de rótulos que equilibra os critérios de segmentação, com desempenho em conjuntos de dados de referência. A *DFC* além de confirmar a eficiência da abordagem da *SBB*, também expande a técnica ao introduzir funções de perda adicionais e extensões para casos de uso específicos, como segmentação com riscos e segmentação de imagens não vistas, resultando em precisão aprimorada em vários conjuntos de dados que contêm a avaliação para comparar (KIM et al., 2020).

---

<sup>1</sup>Berkeley Segmentation Data Set 500 é o padrão de referência para detecção de contornos e pode ser encontrado em: <https://github.com/BIDS/BSDS500> (MARTIN et al., 2001).

<sup>2</sup>Common Objects in Context que visa o reconhecimento de objetos no contexto do entendimento de cenas e pode ser encontrado em: <https://cocodataset.org/> (LIN et al., 2014).

<sup>3</sup>Pascal Visual Object Classes fornece conjuntos de dados de imagens padronizadas para reconhecimento de classes de objetos e pode ser encontrado em: <http://host.robots.ox.ac.uk/pascal/VOC/> (EVERINGHAM et al., 2010).

### 4.3 Critérios e restrições para agrupamento de características diferenciáveis

Para a *DFC* de Kim et al. (2020), dada a imagem colorida de entrada  $I = v_n \in \mathbb{R}^3$ , no qual  $N$  é o número de píxeis, define-se a função de extração de características  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^p$  e o conjunto de vetores de características  $x_n \in \mathbb{R}^p$ . Os rótulos dos agrupamentos  $c_n \in \mathbb{Z}$  são atribuídos a todos os píxeis por  $c_n = g(x_n)$ , na qual  $g : \mathbb{R}^p \rightarrow \mathbb{Z}$  é a função de mapeamento. Se  $f$  e  $g$  são treináveis e  $c_n$  são especificados, então a expressão pode ser entendida como problema de classificação supervisionada. No entanto, na *DFC* (KIM et al., 2020),  $c_n$  são desconhecidos e são previstos, enquanto os parâmetros de  $f$  e  $g$  são treinados de forma completamente não supervisionada.

Os critérios introduzidos na proposta da *DFC* são conflitantes e não são satisfeitos perfeitamente. A solução clássica seria aplicar o algoritmo  $k$ -médias em  $x_n$  para o primeiro critério, utilizar o algoritmo de corte de grafos para o segundo critério e determinar  $k$  no  $k$ -médias utilizando método não paramétrico para o último critério. No entanto, estes métodos clássicos são aplicáveis apenas a  $x_n$  constante, o que resulta em solução subótima. Assim, o algoritmo com base em RNC é proposto para resolver o problema, otimizando conjuntamente as funções de extração de características para  $x_n$  e  $c_n$  de forma a satisfazer todos os critérios. Para possibilitar o aprendizado fim-a-fim (*end-to-end*) da RNC, é proposta abordagem iterativa para prever  $c_n$  utilizando funções diferenciáveis. A estrutura da RNC é proposta juntamente com a função de perda para satisfazer os três critérios (KIM et al., 2020).

### 4.4 Extração de características, restrição na similaridade e número de rótulos

A imagem de entrada  $I$  é processada pela RNC para extrair características profundas  $x_n$  utilizando o módulo de extração de características. A camada convolucional unidimensional 1D calcula os vetores de resposta  $r_n$  nas características do espaço de agrupamento  $q$ -dimensional. Os vetores de resposta são normalizados através dos eixos do espaço de agrupamento utilizando a função de normalização em lote (*Batch Normalization* – BN). Em seguida, os rótulos de agrupamento  $c_n$  são determinados atribuindo identificadores de agrupamento aos vetores de resposta utilizando a função  $\text{argmax}^4$ . Então, os rótulos de agrupamento são considerados como pseudo-alvos para calcular a perda de similaridade de características. A perda de continuidade espacial, assim como a perda de similaridade de características, são calculadas e

---

<sup>4</sup> $\text{argmax}(\mathbf{x}) = \underset{i}{\text{argmax}} x_i = \{i \mid x_i = \max(x_1, x_2, \dots, x_n)\}$

retropropagadas na RNC (KIM et al., 2020).

Considerando o primeiro critério para atribuição do mesmo rótulo a píxeis com características semelhantes, a solução proposta da *DFC* é aplicar o classificador linear para classificar as características de cada píxel em  $q$  classes. O mapa de característica  $p$ -dimensional  $x_n$  é calculado a partir de  $v_n$ , através de componentes convolucionais  $M$ . Cada mapa é resultado de convolução bidimensional, função de ativação Unidade Linear Retificada (*Rectified Linear Unit* – ReLU) e função de normalização *BN*. O mapa de resposta  $r_n = W_c x_n$  é obtido através do classificador linear, no qual  $W_c \in \mathbb{R}^{q \times p}$ . O mapa de resposta é então normalizado para  $r'_n$  de forma que  $r'_n$  tenha média  $\mu_m = 0$  e variância  $\sigma^2 = 1$ . O rótulo de agrupamento  $c_n$  para cada píxel é obtido através da seleção da dimensão que possui o valor máximo em  $r'_n$  (KIM et al., 2020).

Na segmentação de imagens não supervisionada, geralmente não há indicação de quantos segmentos devem ser gerados pelas imagens. Portanto, o número de rótulos de agrupamentos únicos é adaptativo ao conteúdo da imagem. A estratégia da *DFC* classifica píxeis em números arbitrários  $q'$  com  $1 \leq q' \leq q$  de agrupamentos, no qual  $q$  é o valor máximo possível para  $q'$ . Valores elevados para  $q'$  indicam supersegmentação, enquanto  $q' \approx 1$  indica subsegmentação. Para treinar a RNC, é definido valor inicial máximo de rótulos de agrupamento  $q$ . No processo de atualização iterativa, píxeis semelhantes ou espacialmente próximos são integrados considerando restrições de similaridade de características e continuidade espacial. Este fenômeno leva à redução do número de rótulos de agrupamentos únicos  $q'$ , mesmo sem restrição explícita. A função de agrupamento da *DFC* tem base na classificação **argmax** e corresponde ao agrupamento de classe  $q'$ , no qual os âncoras  $q'$  correspondem ao subconjunto de  $q$ -pontos que  $\rightarrow \infty$  dos eixos (KIM et al., 2020).

O primeiro e o segundo critério facilitam o agrupamento dos píxeis, o que pode levar a simples solução de  $q' = 1$ . Para evitar esta falha de subsegmentação, o terceiro critério é introduzido, que é a preferência por valores maiores para  $q'$ . A solução proposta da *DFC* é inserir o processo de normalização intra-eixo (*Normalization Across Channels* – NAC), para o mapa de resposta  $r_n$  antes de atribuir rótulos de agrupamentos, utilizando a classificação **argmax**. Esta operação converte as respostas originais  $r_n$  para  $r'_n$ , no qual cada eixo tem  $\mu_m = 0$  e  $\sigma^2 = 1$ . Isto proporciona a cada  $r'_n, i$  com  $i = 1, 2, \dots, q$  a chance igual de ser o valor máximo de  $r'_n$  nos eixos. Embora esta operação não garanta que cada  $i$ -ésimo agrupamento atinja o valor máximo para qualquer  $n$  com  $n = 1, 2, \dots, N$ , vários índices de agrupamento alcançarão o valor

máximo para qualquer  $n$ . Consequentemente, este processo *NAC* permite que a *DFC* dê preferência por valores maiores de  $q'$  (KIM et al., 2020).

#### 4.5 Restrição na continuidade espacial e função de perda

A restrição da continuidade espacial envolve basicamente agrupar píxeis semelhantes em agrupamentos. No entanto, na segmentação de imagens, é preferível que os agrupamentos de píxeis sejam espacialmente contínuos. Esta restrição adicional favorece que os rótulos de agrupamentos sejam os mesmos dos píxeis vizinhos. Assim, a perda de continuidade espacial  $L_{con}$  é dada por (4.1), na qual  $W$  e  $H$  representam a largura e a altura da imagem de entrada, e  $r'_{\xi,\eta}$  representa o valor do píxel em  $\xi, \eta$  no mapa de resposta  $r'_n$  (KIM et al., 2020).

$$L_{con}(r'_n) = \sum_{\xi=1}^{W-1} \sum_{\eta=1}^{H-1} \left\| r'_{\xi+1,\eta} - r'_{\xi,\eta} \right\|_1 + \left\| r'_{\xi,\eta+1} - r'_{\xi,\eta} \right\|_1 \quad (4.1)$$

Na técnica de segmentação de imagens com informações de rabiscos, a *DFC* aborda a perda de rabisco  $L_{scr}$  como a entropia cruzada parcial, dada por (4.2), na qual  $u_n = 1$  se o  $n$ -ésimo píxel for o píxel rabiscado, caso contrário  $u_n = 0$ , e  $s_n$  denota o rótulo de rabisco para cada píxel. A função de perda proposta para a *DFC* consiste em duas restrições principais: i) similaridade de características e ii) continuidade espacial. A função de perda é dada por (4.3) (KIM et al., 2020), na qual  $\mu$  é o peso para equilibrar as duas restrições. A função de perda é modificada quando informações de rabiscos são utilizadas como entrada do usuário, adicionando peso  $\nu$  para a perda de informação de rabisco, como expresso por (4.4).

$$L_{scr}(r'_n, s_n, u_n) = \sum_{n=1}^N \sum_{i=1}^q -u_n \delta(i - s_n) \ln r'_{n,i} \quad (4.2)$$

$$L = L_{sim}(r'_n, c_n) + \mu L_{con}(r'_n) \quad (4.3)$$

$$L = L_{sim}(r'_n, c_n) + \mu L_{con}(r'_n) + \nu L_{scr}(r'_n, s_n, u_n) \quad (4.4)$$

Os rótulos de agrupamento  $c_n$  são obtidos aplicando a função **argmax** ao mapa de resposta normalizado  $r'_n$ . Estes rótulos são utilizados como pseudo-alvos. A perda de entropia cruzada entre  $r'_n$  e  $c_n$  é calculada como a restrição na similaridade de

características, dada por (4.5), na qual  $\delta(t) = 1$  se  $t = 0$ , e  $\delta(t) = 0$  caso contrário. O objetivo desta função de perda é aumentar a similaridade entre características semelhantes, atualizando os pesos da rede para facilitar a extração eficiente de características para o agrupamento (KIM et al., 2020).

$$L_{sim}(r'_n, c_n) = \sum_{n=1}^N \sum_{i=1}^q -\delta(i - c_n) \ln r'_{n,i} \quad (4.5)$$

#### 4.6 Aprendizagem por retropropagação e experimentos

A metodologia de Kim et al. (2020) para a *DFC* descreve o treinamento da RNC para segmentação de imagens sem supervisão. Quando a imagem-alvo é inserida, dois subproblemas são resolvidos: i) previsão de rótulos de agrupamento com parâmetros de rede fixos e ii) treinamento dos parâmetros da rede com os rótulos de agrupamentos fixos. O primeiro corresponde ao processo de avanço da rede, seguindo a arquitetura proposta. O segundo corresponde ao processo de retrocesso da rede com base no método de descida do gradiente. Em seguida, calcula-se e retropropaga-se a perda  $L$  para atualizar os parâmetros dos filtros convolucionais e do classificador. Utiliza-se a descida de gradiente estocástica com momento para a atualização dos parâmetros, e amostram-se valores da distribuição uniforme normalizada de acordo com o tamanho da camada de entrada e saída. Este processo de avanço-retrocesso é iterado  $T$  vezes para obter a previsão final dos rótulos de agrupamento  $c_n$ .

Como este processo iterativo requer tempo de computação, introduz-se o uso da *DFC* com uma ou várias imagens de referência. Se a imagem-alvo for semelhante às imagens de referência, os pesos de rede fixos treinados com estas imagens como pré-processamento podem ser reutilizados. A principal diferença da RNC proposta é a existência da camada de normalização *BN* entre a camada de convolução final e a camada de classificação *argmax*. Diferentemente do aprendizado supervisionado, no qual os rótulos-alvo são fixos, a normalização *BN* das respostas sobre os eixos é necessária para obter rótulos  $c_n$  razoáveis. Além disto, ao contrário do aprendizado supervisionado, existem múltiplas soluções de  $c_n$  com diferentes parâmetros de rede que alcançam a perda próxima de zero. O valor da taxa de aprendizado controla o equilíbrio entre as atualizações dos parâmetros e o agrupamento, o que leva a diferentes soluções de  $c_n$  (KIM et al., 2020).

Kim et al. (2020) realizam experimentos para validar a *DFC*, utilizando taxa de aprendizado de 0,1 com momento de 0,9. Eles propõem a perda de continuidade

espacial como alternativa e avaliam sua eficiência comparando-a com métodos clássicos. Os autores demonstram o uso da *DFC* com entrada de rabiscos e com imagens de referência. O número de camadas convolucionais  $M = 3$  é definido, e  $p = q = 100$  para todos os experimentos. Para a função de perda, diferentes valores de  $\mu$  são definidos para cada experimento, variando de acordo com o conjunto de dados e a seção correspondente. Os resultados são avaliados pela média da intersecção sobre união (*Intersection over Union* – IoU) de cada segmento do conjunto de verdade absoluta e o segmento estimado que tem o maior *IoU* com o segmento da verdade absoluta.

Observa-se que os rótulos de categoria de objeto no conjunto de dados PASCAL VOC são ignorados, e cada segmento, juntamente com a região de fundo, é tratado como segmento individual. Esta abordagem permite a avaliação focada na qualidade da segmentação em si, independentemente das categorias de objetos específicos. A perda de continuidade é avaliada nos conjuntos de validação de segmentação PASCAL VOC e BSD500. A variação do parâmetro  $\mu$  afeta a segmentação das imagens, por exemplo, com  $\mu = 5$  são apresentados os melhores resultados para segmentação não supervisionada e  $\nu = 0,5$  para segmentação com entrada do usuário. Comparando com métodos clássicos e o método de aglomeração de informação invariante (*Invariant Information Clustering* – IIC), a *DFC* alcança as melhores e as segundas melhores pontuações nos conjuntos de dados de validação PASCAL VOC e BSD500, respectivamente (KIM et al., 2020).

Os resultados de comparação de *IoU* para a segmentação não supervisionada apresenta que a normalização *BN* e a perda de continuidade, quando utilizadas em conjunto, melhoram o desempenho da segmentação (KIM et al., 2020). A aprendizagem supervisionada aprende com dados de treinamento e avalia o desempenho utilizando dados de teste. Em contraste, como a *DFC* que é aprendizagem não supervisionada, é necessário aprender os pesos da rede cada vez que nova imagem de teste é inserida para obter os resultados de segmentação. Além disto, o experimento de segmentação não supervisionada é conduzido com imagens de referência, sem anotação. A eficiência das redes com pesos fixos treinados em várias imagens como referência é avaliada para imagens de teste não vistas.

Os conjuntos de dados BSD500 e iCoseg<sup>5</sup> são utilizados no experimento, no qual a

---

<sup>5</sup>Interactive Co-Segmentation é o conjunto de dados utilizado principalmente para o processo de segmentar simultaneamente múltiplas imagens, identificando e isolando objetos comuns presentes em todas elas e pode ser encontrado em: <http://chenlab.ece.cornell.edu/projects/touch-coseg/> (BATRA et al., 2010).

*DFC* é treinada com quatro imagens no BSD500. Na fase de treinamento, a rede é atualizada para cada imagem de referência. Após o treinamento, os pesos da rede são fixados e as outras três imagens são segmentadas. As imagens de referência e as imagens de teste são selecionadas arbitrariamente de diferentes cenas na categoria natureza. Os resultados da segmentação apresentam que o método que utiliza imagens de referência detalha melhor que o método que parte do início (*from scratch*)<sup>6</sup>. Isto ocorre porque, quando se inicia sem prévio aprendizado, integram-se agrupamentos sob a influência da perda de continuidade ao treinar a imagem-alvo. O método que utiliza imagens de referência leva menos de 0.02s para a segmentação de cada imagem, enquanto o método que se inicia do começo leva aproximadamente 20s sob cálculo de unidade de processamento gráfico, (*Graphics Processing Unit – GPU*), *GeForce GTX TITAN X*, para obter os resultados de segmentação (KIM et al., 2020).

A *DFC* é treinada com quatro grupos do iCoseg, e como o iCoseg não distingue entre dados de treinamento e teste, duas imagens do grupo são selecionadas aleatoriamente para teste. Assim, a *DFC* é treinada nas imagens do grupo excluindo as imagens de teste amostradas, possibilitando a segmentação de imagens desconhecidas com pesos treinados não supervisionados em imagens de referência, desde que as imagens sejam semelhantes às imagens de referência (KIM et al., 2020).

#### 4.7 Considerações finais

Neste capítulo, foram abordados os conceitos e definições de redes de segmentação não supervisionada, com a descrição de alguns dos métodos encontrados na literatura. Posteriormente, detalhou-se a rede para segmentação com base em agrupamento de características diferenciáveis (*DFC*), apresentando critérios, restrições e função de perda. Por fim, foram apresentadas algumas validações da *DFC*, que será a base para a metodologia proposta.

---

<sup>6</sup>Refere-se ao processo de segmentação de imagens que começa sem qualquer conhecimento prévio ou referência. Isto é, iniciar a segmentação sem utilizar pesos ou parâmetros treinados anteriormente. Este método depende exclusivamente do algoritmo para aprender as características das imagens a partir do zero durante o treinamento, o que pode resultar em agrupamentos influenciados pela perda de continuidade.



## CAPÍTULO 5

### METODOLOGIA

Neste capítulo, detalham-se os métodos utilizados para o desenvolvimento e a implementação da segmentação não supervisionada aplicada a neuroimagens em exames de tomografia computadorizada. Primeiramente, é apresentada a contextualização do problema, seguida pela etapa de pré-processamento dos dados. Em seguida, define-se o segmentador projetado e descreve-se a arquitetura do algoritmo empregado. Explica-se o cálculo do erro como função de avaliação e, por fim, aborda-se a etapa de otimização dos hiperparâmetros da proposta.

#### 5.1 Contextualização

Na área da medicina diagnóstica, os exames de tomografia de crânio são ferramentas fundamentais para a detecção de diversas condições neurológicas. No entanto, a anotação minuciosa destas imagens representa desafio, exigindo que médicos especialistas identifiquem e marquem diferentes estruturas cerebrais em busca de anomalias. Neste trabalho, propõe-se abordagem que emprega técnicas de segmentação automatizada das regiões cerebrais, agilizando o processo de anotação e minimizando a intervenção manual. A estratégia envolve a implementação de rede neural profunda para segmentar estruturas intracranianas sem depender de rótulos pré-existentes, anotações manuais ou supervisão direta.

Adicionalmente, explora-se técnica de treinamento de redes neurais, a otimização dos hiperparâmetros da rede, a determinação do número ideal de máscaras para a segmentação e a avaliação do desempenho geral da abordagem. A utilização de modelo não supervisionado assegura que o algoritmo não aprenda o estilo de anotação dos médicos especialistas, evitando vieses, e garantindo que as estruturas intracranianas segmentadas possuam coerência. A proposta visa a separação destas estruturas por meio de cálculos de continuidade e similaridade das informações contidas em cada configuração encontrada. Os resultados das segmentações são validados em comparação com a opinião de especialistas médicos e contextualizados com estudos similares da literatura. A Figura 5.1 ilustra o fluxo de validação da metodologia proposta.

#### 5.2 Fundamentos e critérios da metodologia

Nesta abordagem proposta, uma vez fornecida a imagem alvo, os rótulos de píxeis e as representações de recursos são otimizados em conjunto, e seus parâmetros são atualizados por meio do algoritmo de descida do gradiente. Além disto, a função de

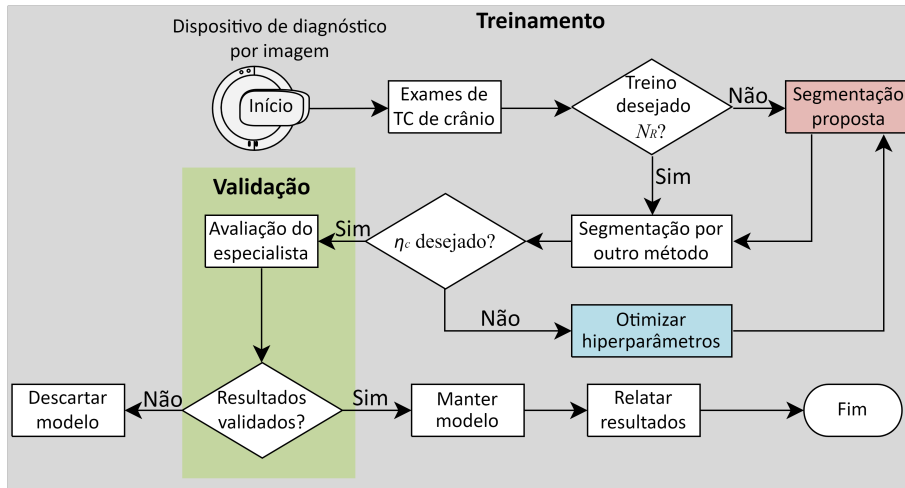


Figura 5.1 - Fluxograma de validação da metodologia proposta.

perda de continuidade espacial, Apêndice C, é otimizada para diminuir as limitações relacionadas aos limites dos segmentos fixos, como observado no trabalho de Kim et al. (2020). A arquitetura proposta emprega transformação linear para mapear as características de cada píxel em classes, seguida de normalização e classificação para determinar rótulos de agrupamentos.

A perda de similaridade de características, Apêndice C, é calculada com base na entropia cruzada (3.4), entre os mapas de resposta normalizados e os rótulos de agrupamentos previstos, enquanto a perda de continuidade espacial tem base na norma- $L_p$  das diferenças horizontais e verticais no mapa de resposta. Na Figura 5.1, O fluxo começa com a seleção da base de exames de TC de crânio. Verifica-se a existência do modelo pré-treinado capaz de produzir os rótulos desejados. Se não houver o modelo pré-treinado, inicia-se o treinamento de novo modelo. E então, os exames são submetidos à segmentação, gerando o número desejado de rótulos  $N_R$ .

### 5.2.1 Pré-Processamento

Para viabilizar o uso de exames de TC utilizando redes neurais profundas, é necessário padronizar os dados, assegurando que todos os exames no conjunto possuam o mesmo número de fatias e dimensões uniformes. Para atingir esta padronização, os exames são submetidos ao processo de interpolação das dimensões espaciais, que gera ou elimina píxeis com base nos píxeis adjacentes, permitindo o redimensionamento espacial das imagens de maneira conveniente. Assim, obtém-se o número predeterminado de fatias de imagens e resolução fixa de largura e altura.

Para eliminar informações irrelevantes nos exames, é aplicado o procedimento de janelamento, seguido pelo algoritmo que mapeia a escala de *Hounsfield* em nova gama de tons de cinza. Esta abordagem simplifica os cálculos durante o treinamento da rede neural, visto que trabalha com escala reduzida de valores, tornando o processo mais eficiente. Ao final da fase de pré-processamento, a estrutura de dados obtida apresenta dimensões  $[N_{fatia} \times L_i \times A_i]$ , onde  $N_{fatia}$  é o número de fatias,  $L_i$  é a largura e  $A_i$  é a altura da imagem em píxeis.

### 5.3 Arquitetura do segmentador

Na construção da arquitetura de segmentação utilizando redes neurais, é necessário filtrar os valores conhecidos da escala *Hounsfield* para separar os diversos tipos de tecidos e estruturas. A segmentação por limiar ou filtro falha ao assumir que os valores encontrados para único banco de dados são eficientes globalmente em outros bancos de dados. Técnicas de limiares dinâmicos, que se ajustam automaticamente em casos complexos para separar imagens em duas classes, não conseguem resultados satisfatórios em bases de exames extensas. Portanto, é necessária metodologia com capacidade de inferência semelhante à capacidade cognitiva humana, que observa a imagem e estabelece automaticamente as fronteiras dos objetos contidos nela.

Esta proposta é inspirada na pesquisa de Kim et al. (2020), que adota abordagem não supervisionada, integrando técnicas de agrupamento e métricas para avaliar a similaridade  $S$  e a taxa de continuidade  $\Delta(r'i, j)$  de forma estática. Um dos diferenciais desta proposta é o uso de valores dinâmicos para ajustar o cálculo de  $S$  e  $\Delta(r'i, j)$  com base no  $N_R$  desejado como saída da segmentação dos exames de TC. A arquitetura do segmentador proposto, indicada no bloco de cor cinza da Figura 5.1, é ilustrada na Figura 5.2, no qual o fluxo na cor preta representa a fase de aprendizado da segmentação, enquanto o fluxo na cor vermelha representa a fase de encerramento do processo iterativo quando o  $N_R$  desejado é atingido.

#### 5.3.1 Extrator de características

O extrator de características, ilustrado na Figura 5.2, utiliza camadas convolucionais configuráveis e funções de ativação. Nesta proposta, a extração de características é realizada por meio de uma rede neural convolucional, que codifica a imagem de entrada em representações de características. No âmbito deste estudo, as imagens com dimensões  $[L_i \times A_i]$  píxeis são processadas pela rede neural, resultando no tensor unidimensional  $[P_n]$ . Cada posição neste tensor armazena a imagem processada, mantendo as dimensões originais  $[L_i \times A_i]$ . A arquitetura do extrator de caracte-

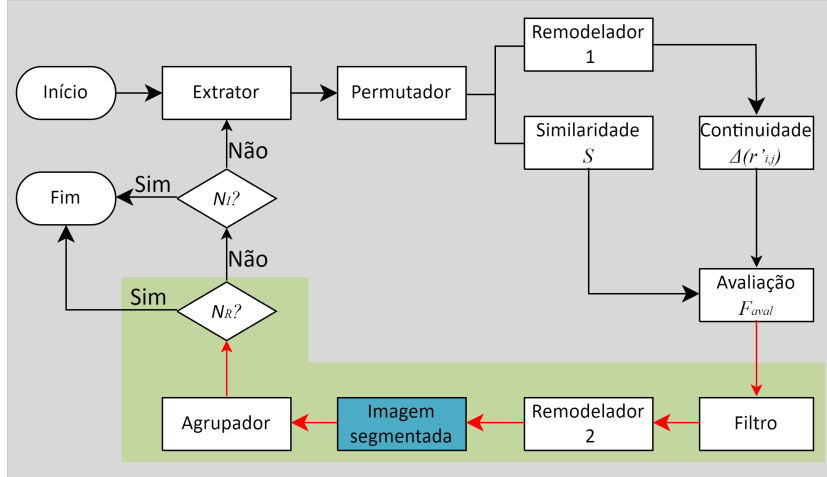


Figura 5.2 - Fluxograma da arquitetura proposta.

rísticas segue o formato sequencial padrão de camadas adjacentes, como ilustrado na Figura 3.7. No entanto, há dinamismo na largura e profundidade da rede, na qual a largura é representada pela quantidade de neurônios em cada camada e a profundidade pela quantidade de blocos de camadas adjacentes. Os blocos seguem o padrão de camadas convolucionais, seguidas por camadas de normalização e camada de função de ativação.

Após a extração de características, realiza-se a normalização utilizando o número específico de filtros convolucionais  $N_F$  para cada camada, pois agrupar píxeis com base na  $S$  exige a normalização dos píxeis na imagem. A Figura 5.3 ilustra a comparação entre a mesma imagem em tons de cinza na Figura 5.3(a) e a superfície na Figura 5.3(b), que ilustra picos e irregularidades no formato de visual. Esta etapa da arquitetura do modelo processa a normalização por meio da atualização dos filtros convolucionais de cada camada. O resultado da normalização é a estrutura tridimensional de dimensões  $[L_i \times A_i \times P_n]$ , composta por imagens processadas por diferentes filtros  $N_F$ .

O permutador atua na saída do extrator de características, que consiste na construção do tensor  $[P_n]$ . Esta estrutura fornece dados para o processo de avaliação do aprendizado da segmentação, representado pelo fluxo na cor preta da Figura 5.2, e alimenta a estrutura necessária para a aplicação do filtro unidimensional de valores máximos, Apêndice B, representado pelo fluxo na cor vermelha da Figura 5.2. Cada elemento de  $[P_n]$  corresponde a uma imagem processada por um  $N_F$  específico. Assim, considerando  $N$  como o número de píxeis de cada imagem, esta estrutura de dados é transformada em novo tensor com dimensões  $[N \times P_n]$  em píxeis. Após

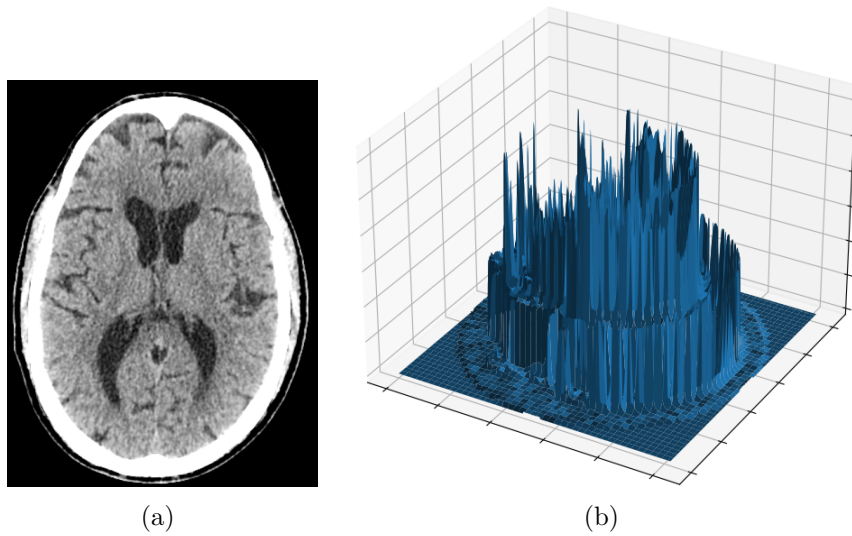


Figura 5.3 - Plano axial e superfície: (a) imagem de tomografia computadorizada e (b) superfície da Figura 5.3(a).

a permutação, Apêndice A, são realizados dois procedimentos de reconfiguração: i) Remodelador 1 e ii) Remodelador 2. O objetivo destes procedimentos é reconstruir a estrutura de dados nas dimensões apropriadas. O Remodelador 1 opera na estrutura de dados proveniente do permutador, que está dimensionada em  $[N \times P_n]$ , gerando como resultado nova estrutura com dimensões  $[A_i \times L_i \times P_n]$ . Por outro lado, o Remodelador 2 assume o tensor  $[N]$  com elementos de entrada e reconstitui esta informação no tensor  $[A_i \times L_i]$  em píxeis.

O cálculo de  $S$  produz valores positivos para características similares, enquanto valores negativos são sempre ajustados para zero. Esta abordagem normaliza os valores para o intervalo desejado, resultando na criação do mapa de similaridade. Este mapa garante que características semelhantes sejam agrupadas ou tenham domínios próximos, ao mesmo tempo que diferencia comportamentos distintos, dependendo do tipo de imagem ou exame processado. Neste trabalho, para calcular  $S$ , utiliza-se a função de entropia cruzada, dada em 3.4, na qual  $p(x_c)$  representa a probabilidade da classe real obtida a partir do mapa de referência para o píxel  $x_c$ , e  $q(x_c)$  representa a probabilidade estimada pela rede neural a partir do mapa de similaridade para o mesmo píxel. A entropia cruzada serve como indicador da concordância entre a saída do modelo e os dados de referência. Quando aplicada à segmentação de imagens médicas, a entropia cruzada é a métrica que avalia a correspondência entre a segmentação produzida pela rede neural e a segmentação desejada, permitindo a quantificação da qualidade da segmentação obtida.

O cálculo da perda de continuidade espacial  $\Delta(r'_{i,j})$ , realiza a delimitação precisa dos objetos presentes nas imagens. Este processo tem base na observação de que, ao examinar a continuidade em determinadas regiões, espera-se que os valores dos píxeis sejam semelhantes, e qualquer variação abrupta pode indicar os limites das regiões ou dos objetos em questão. Assim, a  $\Delta(r'_{i,j})$  analisa o tensor tridimensional  $[A_i \times L_i \times P_n]$ . Durante este processo, ocorre a comparação entre todos os elementos do tensor. Por exemplo, se parte da imagem estiver contida em algum elemento, a técnica calcula a diferença de possíveis relevos entre os outros elementos, medindo a continuidade destas características. A  $\Delta(r'_{i,j})$  resulta na delimitação das regiões presentes nas imagens, indicando as diferenças entre os tipos de objetos existentes.

A  $\Delta(r'_{i,j})$  pode ser medida utilizando imagens filtradas sobrepostas à imagem original e deslocando-as em diferentes direções em relação à imagem original. O píxel em determinada posição aparece deslocado nas imagens sobrepostas, o que permite calcular sua distância em relação ao mesmo píxel nas outras imagens. Valores de distância menores indicam continuidade espacial, sugerindo que não houve mudanças abruptas no objeto. Por outro lado, valores de distância maiores indicam a interrupção do objeto presente nas duas imagens, sinalizando a descontinuidade espacial. O método proposto por Kim et al. (2020) avalia a distância absoluta entre a imagem original e seu mapa de resposta em várias direções, identificando discrepâncias na continuidade dos píxeis em relação aos seus vizinhos. Esta distância absoluta é quantificada pelo erro absoluto médio  $\overline{E_{abs}}$ .

Neste trabalho, são sugeridas alterações na abordagem de Kim et al. (2020), como a função  $\overline{E_{abs}}$ , que os autores consideram a distância  $\lambda$  constante e, para todos os casos  $\lambda = 1$ . Diferentemente, nesta proposta considera-se o  $\overline{E_{abs}}$  como  $\Delta(r'_{i,j})$  e a distância  $\lambda$  é determinada com base no valor desejado de  $N_R$ , sendo tratada como parâmetro a ser otimizado. Para a região específica na imagem considerada como ponto central e a correspondente região na outra imagem após a aplicação de diferentes  $N_F$ , o  $\Delta(r'_{i,j})$  mensura a diferença entre estas regiões. Em coordenadas cartesianas, o  $\Delta(r'_{i,j})$  tende a ser mais pronunciado nos pontos mais extremos.

Estes ajustes incluem a introdução de valores dinâmicos para comparação da imagem central e seu mapa de resposta, bem como valores específicos para determinar o  $N_R$  desejado. Semelhante a Shibata et al. (2017), considera-se a norma- $\Delta(r'_{i,j})$  das diferenças horizontais e verticais do mapa de resposta  $r'_{i,j}$  como restrição espacial. Assim, o processo pode ser implementado através do operador diferencial, definindo a perda de continuidade espacial em (4.1) e rescrevendo-a na forma:

$$\Delta(r'_{i,j}) = \sum_{i=1}^{w-1} \sum_{j=1}^{h-1} \|r'_{i+\lambda,j} - r'_{i,j}\| \lambda + \|r'_{i,j+\lambda} - r'_{i,j}\| \lambda, \quad (5.1)$$

na qual  $\lambda$  é a distância a ser encontrada com base no  $N_R$  desejado,  $r'_{i,j}$  refere-se ao mapa de resposta,  $r'_{i+\lambda,j}$  e  $r'_{i,j+\lambda}$  representam os valores dos píxeis nas posições  $(i, j)$  do mapa de resposta e  $w$  e  $h$  correspondem ao comprimento e largura da imagem, respectivamente. A escolha do valor de  $\lambda$  é necessária na segmentação de estruturas em exames médicos. Valores menores de  $\lambda$  tendem a negligenciar regiões menores nos exames segmentados, enquanto valores maiores de  $\lambda$  conseguem segmentar estruturas pequenas e desconexas. Após avaliar a estrutura de dados resultante da permutação, com dimensão  $[N \times P_n]$ , ela é submetida à aplicação de filtro unidimensional (fluxo de cor vermelha da Figura 5.2).

Isto a transforma em tensor de dimensão  $[N]$ , no qual são destacados os valores máximos. O critério de seleção dos principais píxeis no filtro unidimensional tem base na intensidade, visando preservar estruturas contínuas enquanto suprime artefatos que possam interromper a coesão. Após a aplicação do filtro unidimensional, o Remodelador 2 reconstitui o tensor, contribuindo para realçar as bordas e fronteiras de cada objeto, permitindo a obtenção da imagem segmentada parcialmente. O agrupador, Apêndice B, tem a finalidade de consolidar os píxeis, unificando valores repetidos em elementos singulares. Neste processo, realiza-se a contagem dos rótulos presentes na imagem. A partir do vetor unidimensional obtido na etapa do filtro, seus elementos são identificados, classificados como únicos e organizados. Esta etapa resulta na unificação dos píxeis na imagem, e a quantidade de píxeis únicos pode ser quantificada, correspondendo à quantidade de estruturas identificadas na segmentação não supervisionada.

Este processo compara a estrutura de dados resultante da etapa de permutação com os valores máximos obtidos na etapa do filtro. A abordagem de segmentação não supervisionada se orienta automaticamente com base nos valores gerados pela  $\Delta(r'i, j)$  e o cálculo da  $S$ , na qual a  $\Delta(r'i, j)$  direciona o processo por meio de operações matemáticas que delineiam os contornos dos objetos ou estruturas presentes na imagem. Paralelamente, o cálculo de  $S$  avalia o grau de semelhança entre os píxeis vizinhos. A união destas duas métricas viabiliza a identificação, delimitação e segmentação de regiões específicas na imagem. A  $\Delta(r'i, j)$ , o cálculo da  $S$  e o  $N_R$  são empregados como critérios de parada do algoritmo de segmentação. Assim, quando  $N_R$  é atingido **ou** quando o número máximo de iterações  $NI$  é alcançado, o

algoritmo cessa seu fluxo.

### 5.3.2 Otimização dos parâmetros e hiperparâmetros

Alguns parâmetros e hiperparâmetros do modelo proposto são otimizados, como ilustrado na Figura 5.1, bloco de cor azul. Para isto, adota-se a ferramenta de busca **Optuna** Akiba et al. (2019), que atua durante o treino do modelo para determinar os valores dos parâmetros e hiperparâmetros, incluindo como parâmetros: i) otimizador  $O_{PT}$ , ii) distância  $\lambda$ , iii) termo de regularização de similaridade  $d_{sim}$  e iv) termo de regularização de continuidade  $d_{cont}$ , com  $N_F, N_C, N_I, \lambda \in \mathbb{Z}^*$ , e  $d_{sim}, d_{cont}, T_A \in \mathbb{R}^*$ . Como hiperparâmetro, inclui: i) número de filtros convolucionais  $N_F$ , ii) número de camadas convolucionais  $N_C$ , iii) número máximo de iterações  $N_I$  e iv) taxa de aprendizado  $T_A$ .

A escolha do  $O_{PT}$  envolve a avaliação de diferentes algoritmos para determinar quais deles se apresentam mais eficientes na identificação do  $N_R$  desejado. A variável  $d_{sim}$  controla o valor de  $S$ , garantindo que píxeis próximos pertençam ao mesmo grupo. Em imagens com intensidades de píxeis semelhantes, os valores de  $S$  podem ser próximos de zero, mesmo que as imagens sejam distintas. Assim,  $d_{sim}$  otimizado controla a relação ótima ou otimizada entre  $S$  e  $N_R$ . A configuração dos  $N_F$  refere-se à determinação do número de  $N_C$  do extrator de características. A definição da quantidade apropriada de  $N_C$  na arquitetura do extrator de características objetiva atingir a eficiência ótima e a captura de detalhes precisos. A otimização destas quantidades tem como objetivo minimizar a função de avaliação dado o  $N_R$  desejado.

A otimização do hiperparâmetro  $\lambda$  é realizada em faixa de valores pré-definida, na qual valores maiores desconsideram regiões menores nos exames segmentados e valores menores segmentam estruturas pequenas e desconexas. A otimização do  $N_F$  é necessária, pois objetiva acompanhar o progresso da  $F_{aval}$  ao longo do treinamento, permitindo comparações entre diferentes configurações de parâmetros e hiperparâmetros, possibilitando controlar o desempenho do modelo durante a segmentação de estruturas cerebrais. A eficiência das características geradas pelo extrator está diretamente ligada a  $T_A$ . Valores baixos de  $T_A$  podem levar a rede a estagnar em mínimos locais reduzindo as perdas lentamente, enquanto valores altos podem fazer com que a rede continue a treinar, mesmo após ter alcançado o ponto ótimo ou otimizado.

### 5.3.3 Validação do modelo proposto

O processo de validação do modelo segue duas etapas: a primeira compreende a comparação com outra metodologia de segmentação estabelecida na literatura, como ilustrado na cor verde da Figura 5.1, utilizando o coeficiente de Dice Sørensen  $D_c$ , medida de similaridade amplamente utilizada em avaliações e validações de segmentação (BERTELS et al., 2019; EELBODE et al., 2020; THADA; JAGLAN, 2013). O  $D_c$  fornece avaliação quantitativa da qualidade da segmentação gerada pela metodologia proposta em relação à abordagem existente.

Na segunda etapa, a validação é realizada por médicos especialistas em imagens intracranianas, abrangendo avaliação detalhada que examina as segmentações produzidas pelo método proposto. Os especialistas avaliam critérios como: i) precisão, ii) delineamento de contornos e iii) habilidade em identificar características ou regiões relevantes. Esta avaliação qualitativa complementa a análise quantitativa, validando a eficiência da nova metodologia. Se os resultados destas métricas estiverem dentro do limiar aceitável predefinido  $\eta_c$ , então as segmentações são validadas. Caso contrário, se forem reprovadas, o modelo é descartado.

## 5.4 Considerações finais

Este capítulo descreveu a metodologia proposta para segmentação não supervisionada de exames de tomografia de crânio. Destacaram-se as principais etapas e as formulações desenvolvidas, possibilitando implementar o aprendizado de máquina. O próximo capítulo dispõe os resultados encontrados a partir da aplicação da metodologia proposta.



## CAPÍTULO 6

### RESULTADOS

Neste capítulo, são apresentados os resultados da aplicação da metodologia proposta. Isto inclui a descrição da base de dados utilizada, a otimização dos parâmetros e hiperparâmetros, a etapa de segmentação e a subsequente análise estatística. São realizadas comparações entre metodologias de segmentação, avaliação da volumetria do crânio e validação dos resultados com especialistas. Por fim, constrói-se a discussão com base nos resultados encontrados.

#### 6.1 Base de dados e definição dos parâmetros

Neste trabalho é utilizado o banco de dados *Computed Tomography Quality 500*<sup>1</sup> (CQ500). O CQ500 é o conjunto de dados públicos que contém 500 exames de tomografia computadorizada de crânio, realizados em diferentes hospitais da Índia. Este banco inclui imagens de qualidade, as quais foram devidamente anonimizadas (CHILAMKURTHY et al., 2018). Além das imagens, o CQ500 fornece relatórios clínicos emitidos por três radiologistas com experiências de 8, 12 e 20 anos na interpretação de tomografias de crânio. A organização dos arquivos dos exames tem a estrutura: i) identificação do paciente, ii) identificação do estudo e iii) identificação do exame, contendo os arquivos no formato DICOM. Os exames são identificados por códigos que iniciam com CQ500CT00x, nos quais 00x representa o número do exame. Neste trabalho, optou-se por remover o prefixo CQ500, mantendo apenas o sufixo CT00x.

Na Figura 5.1, a segmentação por outro método é realizada utilizando a ferramenta CTSeg, que é empregada para segmentar e normalizar imagens de TC do crânio. Esta ferramenta utiliza a segmentação unificada, incluindo prioridades nos parâmetros do modelo de mistura gaussiana e atlas aprendido a partir de RM e TC. Define-se o limiar aceitável de  $\eta_c = 0,8$ , indicando que, nas comparações entre segmentações, valores superiores a este limiar denotam maior similaridade entre as imagens. Seguindo o princípio de Pareto (ARNOLD, 2014; SANDERS, 1987), selecionam-se 100 exames de referência a partir do conjunto de 500 exames provenientes do CQ500. Dentre estes 100 exames, escolhe-se apenas um para ajustar os parâmetros e hiperparâmetros, na tentativa de minimizar a função de avaliação. Para os experimentos, selecionam-se aleatoriamente cinco exames dos 400 restantes. Nenhuma anomalia ou condições de saúde fora do padrão normal foi identificada nos cinco exames escolhidos.

---

<sup>1</sup><http://headctstudy.qure.ai/dataset>

Os exames incluídos no conjunto de experimentos são: i) CT047, utilizado como referência para ajuste dos parâmetros e hiperparâmetros, e ii) CT042, CT195, CT200, CT299 e CT418, usados para validação. As imagens têm resolução de  $A_i = 512, L_i = 512$  píxeis, totalizando  $N = A_i \cdot L_i = 262144$  píxeis. Após o pré-processamento, a estrutura de dados resultante possui dimensões  $[N_{fatia} = 256, L_i = 512, A_i = 512]$ . No contexto deste estudo, as imagens são submetidas ao processo de extração de características, resultando em estrutura de dados de dimensão  $P_n = 100$ , em que cada posição contém a imagem processada mantendo as dimensões originais de  $[A_i \times L_i]$  píxeis.

Para o treinamento, o exame de referência CT047 é utilizado para cada rótulo  $N_R^2$  desejado, permitindo a segmentação dos outros exames e a validação dos resultados com CT042, CT195, CT200, CT299 e CT418. A similaridade entre as fatias adjacentes dos exames de TC de crânio permite que único exame forneça informações suficientes para o treinamento da rede neural, reduzindo significativamente a necessidade de recursos computacionais e tempo.

## 6.2 Resultado do processo de otimização

A busca pelos valores otimizados dos parâmetros e hiperparâmetros que controlam o comportamento do extrator de características durante o treinamento e das funções responsáveis pela segmentação, considera: como parâmetros, i) o otimizador  $O_{pt}$  Adam, RMSprop e SGD, ii) distância  $\lambda$  variando de 1 a 9, iii) termos de regularização de similaridade  $d_{sim}$  e de continuidade  $d_{cont}$  variando de 0, 1 a 5 em incrementos de 0, 1 e como hiperparâmetros, i) número de filtros convolucionais  $N_F$  variando de 15 a 150, ii) número de camadas convolucionais  $N_C$  variando de 1 a 9, iii) número máximo de iterações  $N_I$  variando de 1 a 10 e iv) taxa de aprendizado  $T_A$  variando de 0,001 a 0,1. As três técnicas de otimização são aplicadas devido à sua popularidade no treinamento de redes neurais e aprendizado de máquina (ZAHEER; SHAZIYA, 2019; ZHANG, 2018). Os intervalos para otimização dos parâmetros e hiperparâmetros foram selecionados com base no conhecimento empírico dos pesquisadores. Para realizar esta otimização, utiliza-se a ferramenta de busca `Optuna` (AKIBA et al., 2019).

Executam-se doze simulações para  $N_R$ , cobrindo o intervalo com rótulos  $N_R \in \mathbb{Z}^*$ ,  $3 \leq N_R \leq 8$ . A Tabela 6.1 dispõe as médias  $\overline{D_c}$  e desvio padrão  $\sigma$  para  $D_c$  ao

---

<sup>2</sup>Para garantir consistência e uniformidade nos resultados, adotou-se a terminologia **rótulos**. Isto ocorre porque as segmentações geradas pela abordagem proposta serão comparadas às segmentações da ferramenta CTSeg, permitindo a avaliação de similaridade e estabelecendo que a segmentação comparada pertence ao rótulo correspondente gerado pela CTSeg.

longo das doze simulações para cada  $N_R$  e para cada classe: substância cinza  $c_1$ , substância branca  $c_2$  e crânio  $c_4$ . Na Tabela 6.1 são dispostos os tempos  $t$  gastos em cada segmentação para cada  $N_R$ . As segmentações de referência utilizadas como padrão-ouro, são obtidas pela ferramenta CTSeg e o cálculo de  $D_c$  é realizado pelo 3D Slicer, plataforma de software de código aberto para processamento e análise de imagens médicas (PINTER et al., 2012).

Tabela 6.1 -  $\overline{D_c}$ ,  $\sigma$  e  $t$  ao longo das doze simulações para cada  $N_R$  e para cada classe  $c_1$ ,  $c_2$  e  $c_4$ .

Exame	$N_R$	$\overline{D_c}$ e $\sigma$ para $c_1$	$\overline{D_c}$ e $\sigma$ para $c_2$	$\overline{D_c}$ e $\sigma$ para $c_4$	$t$ [min] Proposta	$t$ [min] CTSeg
CT042	3	–	–	$0,80 \pm 0,03$	1,50	46
	4	–	–	$0,59 \pm 0,06$	1,19	
	5	–	–	$0,55 \pm 0,02$	1,20	
	6	–	$0,55 \pm 0,06$	$0,64 \pm 0,03$	1,90	
	7	–	$0,39 \pm 0,05$	$0,69 \pm 0,04$	1,15	
	8	–	$0,40 \pm 0,07$	–	1,80	
CT195	3	$0,50 \pm 0,03$	–	$0,74 \pm 0,01$	1,20	120
	4	$0,49 \pm 0,05$	–	$0,70 \pm 0,03$	1,70	
	5	$0,44 \pm 0,05$	$0,25 \pm 0,06$	–	1,50	
	6	$0,57 \pm 0,05$	$0,30 \pm 0,05$	$0,65 \pm 0,03$	1,30	
	7	$0,57 \pm 0,04$	$0,40 \pm 0,05$	$0,70 \pm 0,03$	1,16	
	8	$0,45 \pm 0,05$	$0,35 \pm 0,07$	–	1,35	
CT200	3	$0,43 \pm 0,05$	–	$0,70 \pm 0,05$	1,78	61
	4	$0,43 \pm 0,03$	–	$0,72 \pm 0,03$	1,85	
	5	$0,42 \pm 0,02$	–	$0,45 \pm 0,06$	1,60	
	6	$0,45 \pm 0,04$	$0,42 \pm 0,08$	$0,75 \pm 0,04$	1,10	
	7	$0,58 \pm 0,03$	$0,35 \pm 0,10$	$0,80 \pm 0,02$	1,48	
	8	$0,46 \pm 0,04$	–	$0,45 \pm 0,06$	1,80	
CT299	3	$0,42 \pm 0,05$	–	–	1,12	57
	4	$0,40 \pm 0,04$	–	–	1,20	
	5	$0,48 \pm 0,03$	–	–	1,19	
	6	$0,45 \pm 0,06$	$0,58 \pm 0,05$	–	1,36	
	7	$0,50 \pm 0,02$	$0,60 \pm 0,06$	–	1,16	
	8	$0,35 \pm 0,06$	–	–	1,49	
CT418	3	$0,45 \pm 0,08$	–	$0,70 \pm 0,03$	1,60	35
	4	$0,42 \pm 0,09$	–	$0,65 \pm 0,07$	1,30	
	5	$0,40 \pm 0,05$	–	–	1,69	
	6	$0,42 \pm 0,03$	$0,55 \pm 0,05$	$0,72 \pm 0,05$	1,83	
	7	$0,60 \pm 0,03$	$0,40 \pm 0,05$	$0,82 \pm 0,04$	1,12	
	8	$0,45 \pm 0,04$	–	–	1,60	

Observa-se na Tabela 6.1, a ausência de alguns valores do  $\overline{D_c}$ . Estas omissões são decorrentes de discrepâncias entre a segmentação do modelo testado e a segmentação de referência, CTSeg. As metodologias não supervisionadas não garantem resultados idênticos, pois não envolvem o treinamento de pesos e classificadores. Por exemplo, ao utilizar as segmentações da ferramenta CTSeg como referência, a segmentação do modelo em teste pode apresentar variações nos tecidos e estruturas cerebrais, resultando em divergências entre os métodos. Estas diferenças resultam em falta de correspondência ao comparar com o  $D_c$ . A melhor configuração entre as doze simulações é aquela que apresenta a melhor média  $D_c$  entre as três classes  $c_1$ ,  $c_2$  e

$c_4$ .

A Tabela 6.2 dispõe os resultados do processo de otimização dos parâmetros e hiperparâmetros, apresentando a configuração mais eficiente para cada valor de  $N_R$  conforme o valor do  $D_c$ , além do tempo de processamento  $t$  [min] para cada processo de otimização. Na coluna  $N_R$ , é indicada a quantidade predefinida de segmentações desejadas. Observa-se aumento na expansão das redes neurais, tanto em largura quanto em profundidade, para encontrar a quantidade desejada de  $N_R$ . Além disto, utilizam-se diferentes  $T_A$  para cada otimizador. São conduzidas duas simulações de otimização para cada valor de  $N_R$ , com o objetivo de reduzir a  $T_A$  ou alcançar o  $N_F$  máximo. Ao todo, são doze simulações, abarcando o intervalo com  $N_R \in \mathbb{Z}^*$ ,  $3 \leq N_R \leq 8$  rótulos.

Tabela 6.2 - Parâmetros e hiperparâmetros otimizados com base na quantidade de rótulos predefinidos.

$Opt$	$N_F$	$N_C$	$\lambda$	$N_I$	$d_{sim}$	$d_{cont}$	$T_A$	$N_R$	$F_{aval}$	$t$ [min]
RMSprop	125	5	1	9	0,2	0,1	0,02071	3	0,0240	120
SGD	55	3	1	6	1,8	0,1	0,07809	4	0,0776	22
SGD	85	1	3	6	1,8	5,0	0,06904	5	0,1666	44
RMSprop	125	3	1	10	1,4	1,9	0,00532	6	0,0932	48
SGD	45	3	1	3	0,4	3,4	0,08104	7	0,2999	100
ADAM	65	1	7	3	1,6	1,6	0,04715	8	0,2212	60

### 6.3 Comparação entre as metodologias de segmentação

O janelamento proposto na metodologia, com centro em 40 [HU] e largura de janela de 80 [HU], visa garantir a presença dos tecidos de interesse, como a substância branca e a substância cinza, enquanto exclui artefatos físicos indesejados presentes no exame, conforme as práticas médicas estabelecidas (EE et al., 2016; HO et al., 2012). No entanto, este filtro pode resultar em interferências, como a presença de píxeis isolados ou pequenos grupos de píxeis com valores extremamente baixos (ruído tipo sal) ou valores extremamente altos (ruído tipo pimenta) (AZZEH et al., 2018; SIVAKUMAR, 2007; TOH; ISA, 2009). Estes pontos isolados ou pequenos grupos de píxeis com valores discrepantes distorcem as informações das imagens dos exames, como observado na Figura 5.3, dificultando a interpretação e o uso clínico. A Figura 6.1(a) apresenta o resultado obtido da aplicação do janelamento. Observa-se que a região contendo o tecido cerebral apresenta ruído do tipo sal e pimenta, havendo também distinção da estrutura ventricular, incluindo calcificação, que ocorre devido ao depósito de cálcio no plexo coroide (SAADE et al., 2019). Como resultado final desta proposta, espera-se que as imagens segmentadas apresentem definições claras entre

as regiões anatômicas e que píxeis próximos estejam agrupados no mesmo rótulo, como a comparação hipotética da Figura 5.3 e o resultado esperado na Figura 6.1(b).

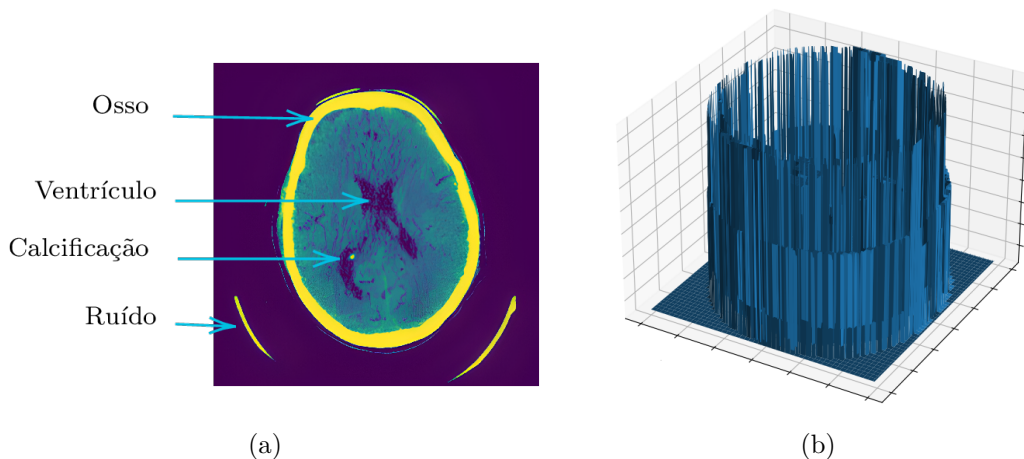


Figura 6.1 - Imagem com janelamento e segmentada: (a) janelamento de  $40 \times 80$  [HU] e (b) gráfico de superfície de imagem segmentada.

Ao utilizar a ferramenta CTSeg, não é necessário especificar  $N_R$ , pois as classes/rótulos são identificados automaticamente durante o processamento, reconhecendo as estruturas anatômicas humanas. Geralmente, os rótulos incluem substância cinza, substância branca, líquido cefalorraquidiano (LCR), osso do crânio, tecido mole e o plano de fundo, que não corresponde a nenhum dos anteriores. Na validação da metodologia proposta, para evitar viés favorável, adotam-se as três classes que apresentam melhor desempenho em termos de segmentação em comparação com a ferramenta CTSeg: i) substância cinza  $c_1$ , ii) substância branca  $c_2$  e ii) crânio  $c_4$ . Os resultados da segmentação são obtidos tanto com a ferramenta CTSeg quanto com a metodologia proposta, utilizando os parâmetros e hiperparâmetros otimizados da Tabela 6.2, ajustados para o mesmo  $N_R$  da ferramenta CTSeg. A Figura 6.2 e Figura 6.3 apresentam estes resultados para os exames CT042, CT195, CT200, CT299 e CT418, respectivamente.

Na Figura 6.2 e Figura 6.3, as imagens em preto e branco apresentam os resultados da segmentação obtidos pela ferramenta CTSeg, enquanto as imagens multicoloridas apresentam os resultados do modelo proposto. Neste estudo, as imagens estão relacionadas a três classes distintas de tecido: i) a primeira coluna representa a classe  $c_4$ , ii) a segunda coluna corresponde à classe  $c_1$  e iii) a terceira coluna à classe  $c_2$ . Os resultados obtidos pela metodologia proposta, utilizando os parâmetros otimizados

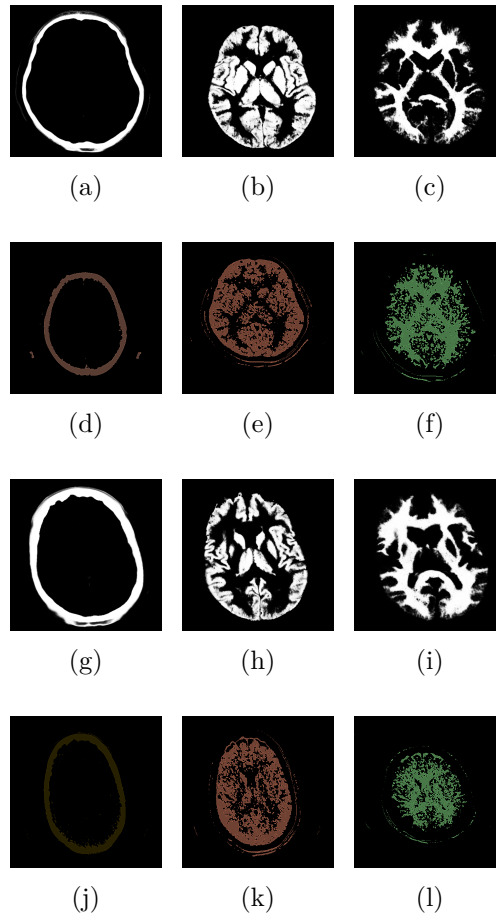


Figura 6.2 - Segmentação dos exames: (a) até (f) CT042 e de (g) até (l) CT195.

dispostos na Tabela 6.2, são dispostos na Tabela 6.3 e comparados com a abordagem proposta por (KIM et al., 2020). A avaliação dos resultados é expressa em termos do  $D_c$ . As segmentações de referência utilizadas para calcular o  $D_c$  são provenientes da ferramenta CTSeg, e o cálculo do  $D_c$  é realizado pela Plataforma 3D Slicer (PINTER et al., 2012). O valor de  $D_c$  pode ser encontrado entre  $0 \leq D_c \leq 1$ , no qual 0 indica nenhuma sobreposição e 1 representa 100% de sobreposição.

Na Tabela 6.3, observa-se a ausência de alguns valores de  $D_c$ , assim como na Tabela 6.1, devido a falta de correspondência entre as segmentações. Nos exames dispostos na Tabela 6.3, os resultados são referentes a 256 fatias após o pré-processamento. No geral, os valores dispostos na Tabela 6.3 apresentam assertividade superior a 65% para a metodologia proposta, em comparação com a assertividade de  $\approx 33\%$  para a metodologia de Kim et al. (2020). A otimização resulta em aumento significativo na assertividade da metodologia de Kim et al. (2020) quando utilizando o método proposto. Ao analisar os resultados da segmentação da classe  $c_1$  para o

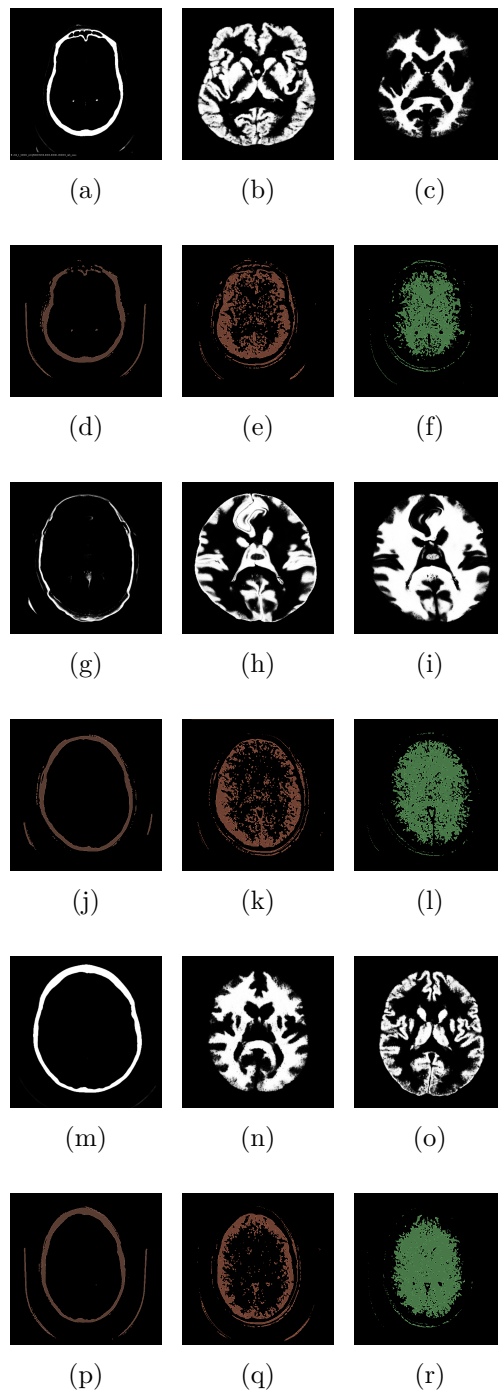


Figura 6.3 - Segmentação dos exames: (a) até (f) CT200, de (g) até (l) CT299, de (m) até (r) CT418.

exame CT042, observa-se que a segmentação do CTSeg não apresenta correlação significativa com a segmentação da metodologia proposta. Em contraste, ao comparar com a segmentação utilizando a metodologia de [Kim et al. \(2020\)](#), a falta de associação é observada para  $N_R = 3$  e  $N_R = 4$ , com  $0,58 \leq D_c \leq 0,74$ . Considerando todos

Tabela 6.3 - Comparação de  $D_c$  entre a metodologia proposta  $\times$  metodologia de Kim et al. (2020).

exame	arquivo	$N_R$	$D_c$ para $c_1$		$D_c$ para $c_2$		$D_c$ para $c_4$	
			Kim <i>et. al.</i>	Proposta	Kim <i>et. al.</i>	Proposta	Kim <i>et. al.</i>	Proposta
CT042	3		–	–	–	–	–	<b>0,83</b>
	4		–	–	–	–	–	0,67
	5		0,69	–	0,31	–	<b>0,82</b>	0,57
	6		0,58	–	0,36	<b>0,61</b>	0,81	0,67
	7		0,73	–	0,43	0,44	0,78	0,73
	8		<b>0,74</b>	–	<b>0,60</b>	0,47	0,78	–
CT195	3		–	0,53	–	–	–	<b>0,75</b>
	4		–	0,48	–	–	–	0,72
	5		0,61	0,49	–	0,31	<b>0,85</b>	–
	6		0,51	0,59	–	0,35	0,76	0,68
	7		0,65	<b>0,61</b>	–	<b>0,47</b>	0,83	0,72
	8		<b>0,67</b>	0,49	–	0,42	0,83	–
CT200	3		–	0,48	–	–	0,84	0,74
	4		–	0,46	–	–	–	0,75
	5		0,53	0,44	0,26	–	<b>0,86</b>	0,50
	6		0,54	0,49	–	<b>0,50</b>	0,73	0,79
	7		0,55	<b>0,61</b>	0,40	0,45	0,81	<b>0,82</b>
	8		<b>0,57</b>	0,50	<b>0,44</b>	–	0,80	0,50
CT299	3		–	0,47	–	–	–	–
	4		–	0,44	–	–	–	–
	5		0,49	0,51	0,44	–	0,25	–
	6		0,44	0,49	0,47	0,63	0,24	–
	7		<b>0,50</b>	<b>0,52</b>	0,53	<b>0,66</b>	<b>0,26</b>	–
	8		0,49	0,41	<b>0,59</b>	–	0,23	–
CT418	3		–	0,53	–	–	<b>0,85</b>	0,73
	4		–	0,50	–	–	–	0,72
	5		<b>0,58</b>	0,45	0,34	–	0,83	–
	6		0,56	0,45	0,37	<b>0,60</b>	0,72	0,77
	7		–	<b>0,62</b>	0,43	0,45	0,79	<b>0,86</b>
	8		–	0,49	<b>0,50</b>	–	0,81	–

os exames para a classe  $c_1$ , a metodologia proposta alcançou  $0,41 \leq D_c \leq 0,62$ , enquanto a metodologia de Kim et al. (2020) obteve  $0,44 \leq D_c \leq 0,74$ .

Com exceção do exame CT042, a segmentação com o CTSeg apresenta associação em comparação com o método proposto em todos os outros exames, incluindo todos os rótulos da classe  $c_1$ . A metodologia proposta obteve resultados superiores em quatro exames, em comparação à segmentação de Kim et al. (2020), que segmentou todos os exames, mas alcançou desempenho superior apenas em dois deles. Para as classes  $c_1$  e  $c_2$ , a metodologia proposta apresenta resultados superiores em termos de  $D_c$  em comparação com a abordagem de Kim et al. (2020). A otimização proporcionou melhorias médias de 4% para a classe  $c_2$  e 5% para a classe  $c_4$  em comparação com a metodologia de Kim et al. (2020). Estes resultados indicam que a proposta deste trabalho, com a otimização dos parâmetros e hiperparâmetros, consegue realizar segmentações mais precisa da substância branca em comparação com a abordagem de Kim et al. (2020).

A substância branca é o tecido cerebral responsável por estabelecer conexões entre

diferentes regiões do cérebro, sendo necessária sua segmentação precisa para análise do funcionamento cognitivo e motor. A segmentação precisa da substância branca é importante no diagnóstico e tratamento de doenças neurológicas como esclerose múltipla, Alzheimer e Parkinson (GHRIBI et al., 2019; PAGNOZZI et al., 2019; TRAN et al., 2022). Em relação aos resultados gerais para a segmentação da classe  $c_4$ , a abordagem de Kim et al. (2020) obteve valores de  $0,23 \leq D_c \leq 0,86$ , enquanto na metodologia proposta alcançaram-se valores de  $0,50 \leq D_c \leq 0,86$ . Ao comparar ambas as metodologias, observa-se que, das três classes e cinco exames analisados, a metodologia proposta apresenta desempenho superior em dez exames, enquanto a metodologia de Kim et al. (2020) supera em cinco exames, considerando os resultados em termos de  $D_c$ . Estes resultados indicam que a metodologia proposta consegue segmentar os tecidos nos exames de tomografia computadorizada de forma mais precisa, produzindo segmentos que se aproximam dos obtidos pela ferramenta CTSeg.

O processo de otimização demonstra eficiência, não resultando em saturação para a classe  $c_4$ , o que indica desempenho satisfatório da  $F_{aval}$ . Contudo, ao considerar apenas o melhor resultado de  $D_c$  para cada exame na comparação entre as metodologias, observa-se que a metodologia proposta apresenta desempenho inferior nas segmentações da classe  $c_4$ . Embora a metodologia proposta gere segmentações para  $3 \leq N_R \leq 8$ , isto não implica necessariamente que a classe  $c_4$  no CTSeg encontrará segmentações semelhantes para todos os rótulos gerados pela proposta. Pode haver ou não a presença da classe em qualquer rótulo. Ao analisar as dificuldades da segmentação, observa-se que, das 30 segmentações realizadas na classe  $c_4$ , onze não apresentam associação entre a metodologia proposta e as segmentações do CTSeg. Além disto, no exame CT299, a abordagem de Kim et al. (2020) para a classe  $c_4$  apresenta melhor resultado com  $D_c = 0,26$ . Estes resultados indicam a complexidade da segmentação da classe  $c_4$  e destaca os desafios específicos encontrados pela metodologia proposta em comparação com a abordagem de Kim et al. (2020).

Ao analisar detalhadamente os resultados da segmentação do exame CT299 pelo CTSeg, como apresentado na Figura 6.3(h) e Figura 6.3(i), observam-se distorções assimétricas que podem indicar falhas no processo de segmentação. Estas distorções são particularmente observadas na classe  $c_4$ , como apresentado na Figura 6.3(g), na qual ocorrem rupturas no segmento ósseo, variações na espessura óssea e morfologia distorcida. Em contraste, os resultados da segmentação pela metodologia proposta, apresentados na Figura 6.3(j), não apresentam tais distorções na classe  $c_4$ . Em vez disto, observa-se a formação óssea com formato ovalado, como esperado para a ana-

tomia do crânio humano. Além disto, na Figura 6.3(k) e Figura 6.3(l), as regiões correspondentes às classes  $c_1$  e  $c_2$  mantêm a integridade dos tecidos, sem distorções assimétricas. Esta comparação indica que, em alguns casos, a segmentação realizada pela metodologia proposta pode superar a metodologia de referência, a ferramenta CTSeg, proporcionando resultados mais precisos e anatomicamente mais corretos.

A Figura 6.4 apresenta a sobreposição da segmentação do exame CT195 realizada pela ferramenta CTSeg e a metodologia proposta. Ao associar os seis rótulos utilizados pelo CTSeg aos rótulos obtidos pela metodologia proposta, observa-se a sobreposições nas classes  $c_1$ ,  $c_2$  e  $c_4$ , bem como nas classes específicas: líquido cefalorraquidiano classe  $c_3$ , tecidos moles classe  $c_5$  e plano de fundo classe  $c_6$ . Na Figura 6.4(a) até Figura 6.4(f), observa-se parcial correspondência entre as segmentações da ferramenta CTSeg e da metodologia proposta.

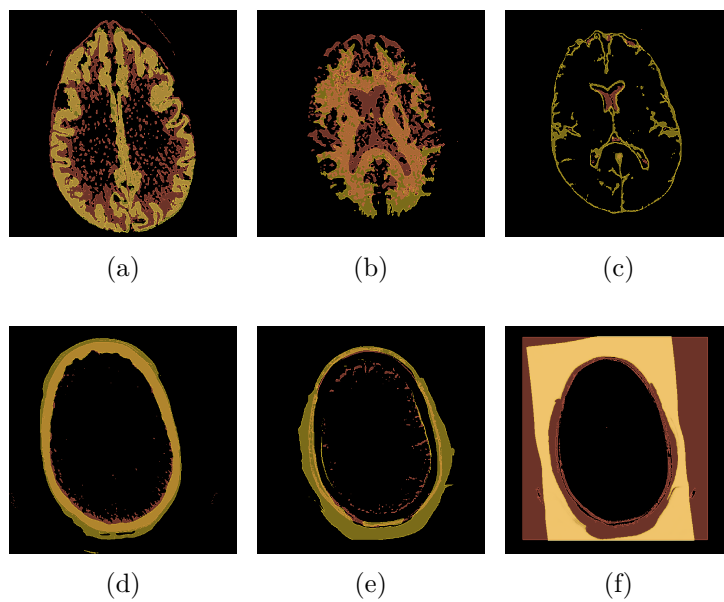


Figura 6.4 - Segmentação sobreposta do exame CT195: (a)  $c_1$ , (b)  $c_2$ , (c)  $c_3$ , (d)  $c_4$ , (e)  $c_5$  e (f)  $c_6$ .

Nas segmentações das classes  $c_1$  e  $c_2$ , Figura 6.4(a) e Figura 6.4(b), a metodologia proposta é representada em tom escuro, enquanto a ferramenta CTSeg, em tom claro. Em contraste, nas classes  $c_3$ ,  $c_4$  e  $c_5$  apresentadas na Figura 6.4(c), Figura 6.4(d) e Figura 6.4(e), ocorre interseção, com CTSeg como conjunto e a metodologia proposta como subconjunto. Na classe  $c_6$ , Figura 6.4(f), a ferramenta CTSeg erroneamente agrupou as classes  $c_4$ ,  $c_5$  e  $c_6$ , resultando em segmentação imprecisa, enquanto a me-

metodologia proposta segmentou corretamente. Os resultados na Figura 6.4 destacam a importância da segmentação precisa em exames de imagem médica, ressaltando a necessidade de revisão e otimização dos processos para garantir a confiabilidade dos diagnósticos.

#### 6.4 Comparação entre a volumetria do crânio

O cálculo da volumetria do crânio é importante em avaliações neurológicas, como tumores cerebrais e atrofia cerebral, sendo necessária para planejamento cirúrgico. Amplamente utilizada em pesquisas, a volumetria é especialmente relevante para investigar variações no tamanho do cérebro e acompanhar doenças neurodegenerativas, contribuindo para estudos sobre o desenvolvimento cerebral em pediatria entre outros (GIORGIO; STEFANO, 2013). O processo de cálculo envolve a segmentação de imagens para identificar e delimitar estruturas intracranianas, resultando na criação de modelo tridimensional. A determinação do volume ocorre multiplicando a área da segmentação pela espessura da fatia, obtendo o volume de cada fatia. Posteriormente, os volumes de todas as fatias são somados para determinar o volume total do crânio. Softwares de reconstrução tridimensional, como o 3D Slicer, são empregados para criar modelos volumétricos a partir das imagens segmentadas.

Neste estudo, a comparação da medida do volume do crânio é realizada utilizando as segmentações obtidas pela ferramenta CTSeg e pelo método proposto. As segmentações de ambas as abordagens são utilizadas na Plataforma 3D Slicer para gerar o sistema tridimensional e calcular o volume do crânio nos exames selecionados com os maiores valores de  $D_c$  da Tabela 6.3. Os resultados desta avaliação estão dispostos na Tabela 6.4, em que  $V_{c_4}$  representa os valores da volumetria do crânio para a classe  $c_4$ . Observa-se que em 62,5% dos casos, os valores de  $V_{c_4}$  são superiores na metodologia proposta. Isto ocorre devido à aplicação de técnicas de pós-processamento pela ferramenta CTSeg, que aprimoram os contornos da segmentação (BRUDFORS et al., 2020; BRUDFORS, 2020), reduzindo o valor de  $V_{c_4}$ .

Na Tabela 6.4,  $\overline{V_{c_4}}$  representa a média dos valores, enquanto  $\sigma^{V_{c_4}}$  indica o erro percentual entre os valores obtidos por ambas metodologias, utilizando como referência os valores obtidos pela ferramenta CTSeg. A Figura 6.5 apresenta a visualização do volume ósseo da classe  $c_4$  do exame CT042 com  $N_R = 3$ , com a segmentação utilizando a metodologia proposta em tons escuros e a ferramenta CTSeg em tons claros. Como observado na Tabela 6.4, o volume resultante da segmentação da classe  $c_4$  pela metodologia proposta é maior que o obtido pela ferramenta CTSeg. A Figura 6.5 apresenta a visualização no plano tridimensional, visto que as segmentações são rea-

Tabela 6.4 - Comparação dos valores da volumetria do crânio entre a metodologia proposta  $\times$  CTSeg.

Exame		$V_{c_4}$ [ $cm^3$ ]		$\overline{V}_{c_4}$ [ $cm^3$ ]	$\sigma^{V_{c_4}}$ [%]
Arquivo	$N_R$	CTSeg	Proposta		
CT042	3	600,85	679,10	639,97	11,5
	5	600,85	664,67	632,76	9,6
CT195	3	789,14	756,35	772,74	-4,3
	5	789,14	712,20	750,67	-10,8
CT200	5	739,56	762,06	750,81	2,9
	6	739,56	992,81	866,18	25,5
CT418	3	596,89	624,79	610,84	4,4
	7	596,89	593,89	595,39	-0,5

lizadas em plano bidimensional. Com a organização dos dados, cada segmentação é incorporada a matriz resultante, a qual é interpretada como o volume, possibilitando a projeção e a subsequente visualização.

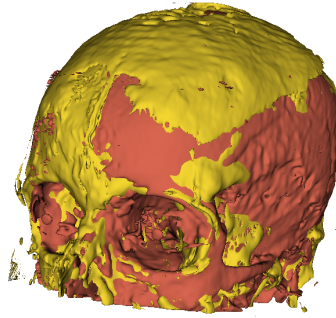


Figura 6.5 - Volumetria do crânio para a classe  $c_4$  do exame CT042.

Ao quantificar o volume das estruturas segmentadas em paciente específico, torna-se viável a comparação com outros pacientes da mesma base de dados, facilitando a identificação de valores discrepantes. Esta abordagem possibilita o uso do modelo de segmentação automática para triagem e priorização de casos suspeitos, agilizando o diagnóstico de condições críticas em estágios iniciais. Além disto, destaca-se a eficiência e economia proporcionadas pelo método proposto. Por se tratar de modelo não supervisionado, a rápida e eficiente capacidade de processamento computacional reduz os custos associados ao tratamento de dados, tornando a tecnologia mais acessível, especialmente em contextos com recursos limitados. É importante observar que, embora as imagens de tomografia computadorizada do crânio sejam mais acessíveis em países em desenvolvimento, modelos de aprendizado profundo para segmentação deste tipo de exame são menos comuns do que aqueles destinados à ressonância magnética. Este trabalho tem representação significativa na redução

desta disparidade na prática clínica.

## 6.5 Segmentação validada por especialista

Os resultados encontrados a partir da aplicação da metodologia proposta são submetidos à análise de especialistas, os quais destacam que a análise comparativa da segmentação de exames de crânio por TC proporciona percepções significativas sobre a eficiência das metodologias utilizadas, indicando diferentes aspectos relevantes para a interpretação clínica. Observa-se que a base de dados CQ500 apresenta variedade de suportes para a fixação da cabeça dos pacientes, sugerindo a utilização de diferentes aparelhos. Esta diversidade é benéfica, introduzindo variabilidade na metodologia proposta e possibilitando avaliações melhores diante de diferentes configurações de tomógrafos. A avaliação abarcou diversas variáveis, desde a influência da diversidade de equipamentos até a otimização com configuração do número de rótulos, com faixa de valores de  $3 \leq N_R \leq 8$ , que, segundo os especialistas, destaca-se como intervalo desnecessário em alguns exames. Esta amplitude sobrepõe a maioria dos tecidos cranianos, considerando a escala *Hounsfield*, proporcionando menor representação e detalhamento das estruturas anatômicas. A escolha apropriada desta faixa é necessária para evitar perda de informação ou excesso de detalhes que possam prejudicar os resultados, garantindo, assim, a segmentação coerente dos tecidos e estruturas.

Os especialistas discutem a análise visual dos resultados, e destacam a relevância da escolha de  $N_R$  na segmentação. A preferência por  $N_R = 6$  é identificada como visualmente agradável, pois oferece distinção entre substância branca e substância cinza. A sugestão é reduzir a faixa de rótulos para  $3 \leq N_R \leq 6$ , respaldada pela observação de que esta faixa proporciona maior clareza e discernimento anatômico. Os resultados dos exames CT042, CT195, CT200, CT299 e CT418 para  $N_R = 7$  e  $N_R = 8$ , apresentados na Figura 6.6, corroboram as observações sobre a redução da faixa de  $N_R$ . No entanto, a segmentação com  $N_R = 6$  oferece riqueza de detalhes e delimitação precisa de estruturas anatômicas. Ao analisar as configurações com  $N_R = 7$  e  $N_R = 8$ , na Figura 6.6, observa-se a combinação de tecidos distintos, como gordura e ar. Este fenômeno é explicado pela natureza da escala *Hounsfield*, apresentando complexidade na segmentação ao ampliar variedade de rótulos. É observado o aumento do *speckle*<sup>3</sup> em regiões sólidas, particularmente com  $N_R = 8$ .

---

<sup>3</sup>Refere-se a pequenos artefatos ou ruídos indesejados que podem aparecer nas imagens devido a várias razões, como imperfeições nos equipamentos de imagem ou condições de captura da imagem. Estes geralmente aparecem como pequenos pontos ou padrões granulares na imagem que não representam informação relevante para a segmentação ou análise clínica.

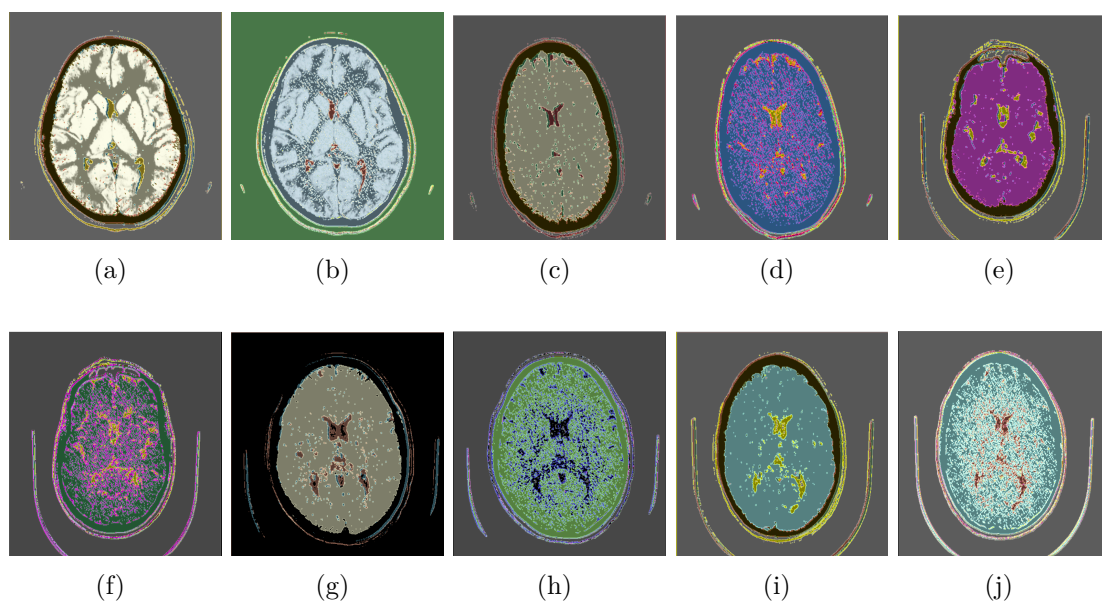


Figura 6.6 - Segmentação utilizando a metodologia proposta nos exames: (a) CT042  $N_R = 7$ , (b) CT042  $N_R = 8$  rótulos, (c) CT195  $N_R = 7$ , (d) CT195  $N_R = 8$ , (e) CT200  $N_R = 7$ , (f) CT200  $N_R = 8$ , (g) CT299  $N_R = 7$ , (h) CT299  $N_R = 8$ , (i) CT418  $N_R = 7$  e (j) CT418  $N_R = 8$ .

Na análise sequencial da quantidade de  $N_R$  aplicados no mesmo exame, apresentado pelo CT195, Figura 6.4, destaca-se a capacidade da metodologia proposta em separar ou agrupar tecidos de acordo com a quantidade desejada. Observa-se clareza e definição das regiões nas segmentações com  $N_R = 6$ , indicando a eficiência da metodologia proposta em delimitar estruturas anatômicas. Ao avaliar a ferramenta CTSeg para o exame CT299, Figura 6.3(g) até Figura 6.3(i) em comparação com a metodologia proposta, Figura 6.3(j) até Figura 6.3(l), observam-se deformações assimétricas nas segmentações. Em contrapartida, a metodologia proposta se apresenta como solução alternativa, indicando avanços significativos na segmentação de exames de crânio por tomografia computadorizada e promovendo melhorias em alguns exames nos quais ferramentas amplamente adotadas têm dificuldades.

A Figura 6.7 apresenta as imagens de segmentação nos cortes axial, coronal e sagital com o método proposto, tendo na cor azul, na cor verde e na cor vermelha a correspondência das classe  $c_1$ , classe  $c_2$  e classe  $c_4$ , respectivamente, com a sobreposição da cor branca que representa os resultados do CTSeg, para os exames: (a) (b) e (c) CT042, (d) (e) e (f) CT195, (g) (h) e (i) CT200, (j) (k) e (l) CT299, (m) (n) e (o) CT418. Observa-se que a Figura 6.7(j) e a Figura 6.7(m) para os exames CT299 e CT418 apresentam distorções assimétricas que indicam erros de segmenta-

ção em comparação com CTSeg. Estas distorções são perceptíveis na classe  $c_1$ . Além disto, no exame CT299, há distúrbios no segmento ósseo, como apresentado na Figura 6.7(l), acompanhados de variações na espessura óssea e morfologia distorcida. Comparado aos resultados de segmentação de CT042, CT195, CT200 e CT418, leves distorções são observadas na classe  $c_4$ .

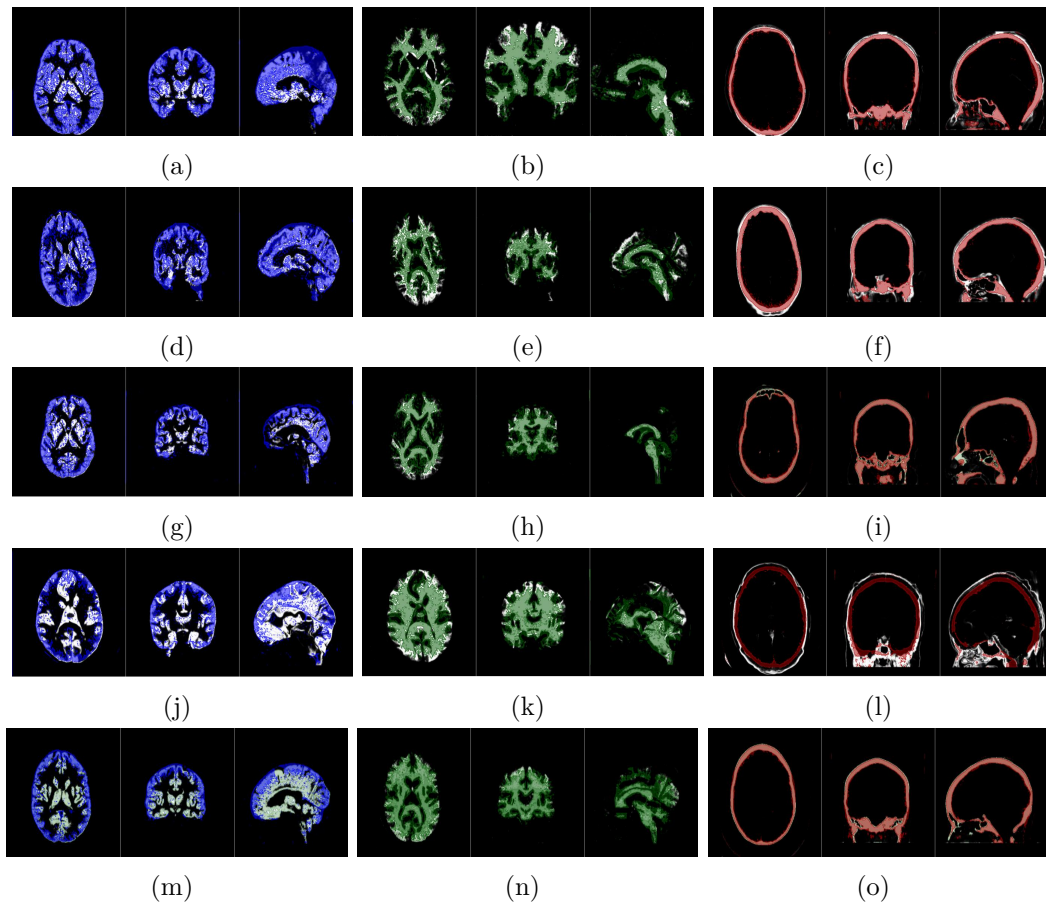


Figura 6.7 - Comparação através de sobreposição

A análise dos médicos especialistas nos resultados da segmentação de exames de TC de crânio, apresentados na Figura 6.7, destacou a importância de considerar vários fatores para a eficiência do processo e a interpretação clínica. A diversidade do conjunto de dados, refletindo diferentes configurações de tomógrafos computadorizados, introduziu variabilidade benéfica para avaliações mais resilientes. A avaliação abarca a diversidade de dispositivos, a otimização e o impacto na qualidade da segmentação. Os especialistas observam a diferenciação mais clara entre a substância branca, substância cinza e o osso do crânio.

## 6.6 Discussão

Os resultados encontrados da aplicação da metodologia proposta são de segmentação de exames de TC de crânio, empregando abordagem não supervisionada. Esta metodologia foi aplicada ao banco de dados CQ500, com otimização de parâmetros e hiperparâmetros, realizado por meio de técnicas de otimização disponíveis na ferramenta *Optuna* (AKIBA et al., 2019). Os experimentos conduzidos forneceram valores de  $D_c$  que variam para diferentes classes de tecido. Estes resultados foram comparados com as segmentações obtidas pela ferramenta CTSeg e com a metodologia proposta por Kim et al. (2020), apresentando convergência nos valores, com alguns casos destacando a metodologia proposta. Estes resultados coadunam com achados na literatura (RAJINI; BHAVANI, 2014; TUSTISON et al., 2015; HENRY et al., 2021) que utilizam técnicas de segmentação com base em modelos supervisionados.

O método proposto é não supervisionado, sua relevância é significativa, visto que o processo não supervisionado não demanda grandes volumes de dados para o treinamento, aliviando a carga computacional na etapa de treinamento. As observações comparativas entre a metodologia proposta e a ferramenta CTSeg, embora apresentem semelhanças em diversos casos, destacam a importância da proposta, considerando a ampla adoção da ferramenta CTSeg em vários centros de imagiologia médica e sua extensa aplicação na literatura (BRUDFORS et al., 2020; BRUDFORS, 2020).

A maioria das comparações entre a metodologia proposta e a abordagem de Kim et al. (2020), a metodologia proposta superou em desempenho para imagens médicas. A metodologia de Kim et al. (2020) não foi originalmente concebida para imagens médicas no formato DICOM. No que diz respeito aos valores de  $D_c$ , em certos casos, não foi possível estabelecer correspondência direta entre os rótulos gerados pela metodologia proposta e aqueles obtidos pela ferramenta CTSeg. No entanto, esta falta de correspondência não reflete necessariamente na qualidade da segmentação da metodologia proposta. O cálculo do valor de  $D_c$  é realizado pelo 3D Slicer (PINTER et al., 2012), que utiliza a segmentação da ferramenta CTSeg como referência, que inclui pós-processamento para aprimorar contornos, reduzir volumes e, por vezes, gera distorções assimétricas nas segmentações.

Este trabalho contribui para a otimização de parâmetros e hiperparâmetros na segmentação de imagens de TC de crânio, utilizando rede neural profunda não supervisionada de forma inovadora, juntamente com a criação da função de avaliação ajustável para diferentes quantidades de rótulos. Apesar dos desafios remanescentes

na metodologia proposta, como a dificuldade em correlacionar os rótulos gerados com os rótulos pré-definidos, as melhorias observadas na segmentação da substância cinza e substância branca em comparação com métodos existentes destacam sua relevância para a prática clínica. Esta dificuldade em correlacionar os rótulos pode indicar a necessidade de futuros estudos sobre técnicas de pós-processamento de dados, enquanto os acertos demonstram a capacidade da metodologia proposta de ser aplicada em diferentes tipos de segmentação.

A segmentação manual realizada por especialistas treinados ainda é considerada o padrão-ouro para a identificação de estruturas cranianas. No entanto, esta abordagem apresenta limitações, como alto custo e consumo oneroso de tempo, tornando-se impraticável para lidar com grandes volumes de dados. Além disto, está sujeita a vieses de interpretação intraobservador e interobservador. Apesar de as segmentações geradas pelo modelo automático não supervisionado se restringirem a algumas estruturas, elas podem servir como base para validação subsequente por especialistas, reduzindo significativamente a carga de trabalho manual. Se validado e integrado com sucesso no fluxo de trabalho de radiologia em tempo real, este modelo tem o potencial de se tornar ferramenta valiosa para triagem e priorização de exames na rotina do radiologista.

A rápida interpretação de exames de TC de crânio anormais pode potencializar o cuidado ao paciente. A abordagem não supervisionada não adiciona carga ao tempo de processamento, uma vez que não requer anotação manual de dados para treinamento, prática comum em métodos supervisionados. O método proposto analisa os dados disponíveis e identifica padrões ou estruturas sem depender da atribuição de rótulos para cada entrada. A aprendizagem não supervisionada tem a capacidade de apresentar percepções e padrões não óbvios nos dados, facilitando a identificação de relações intrínsecas entre variáveis ou características específicas dos dados.

Ao aplicar redes neurais não supervisionadas a conjuntos de dados clínicos reais e anteriormente não auditados, surgem desafios, como a variabilidade na qualidade da imagem na resolução, ruído, iluminação e artefatos, que afetam a capacidade de aprendizagem e generalização da rede. Artefatos introduzidos durante a aquisição, processamento ou transmissão de imagens podem distorcer as informações, enquanto variações anatômicas entre os pacientes dificultam a identificação de padrões (AM-BELLAN *et al.*, 2019). A obtenção de dados de qualidade para treinamento também pode ser outro desafio. Estratégias como pré-processamento, aumento de dados e a colaboração de especialistas clínicos são necessárias para enfrentar estas limitações,

garantindo a interpretabilidade e a relevância clínica dos resultados.

A análise detalhada do uso de métodos automatizados para segmentação de imagens médicas em práticas cirúrgicas e médicas indica benefícios e desafios. A automação oferece vantagens, como eficiência e precisão na interpretação de imagens médicas, mas enfrenta limitações, como dependência da qualidade dos dados e a necessidade de confiança nos resultados. Por outro lado, a segmentação manual, embora precisa, é demorada e sujeita a variações interobservador e intraobservador. A combinação de segmentação automatizada e revisão humana surge como solução promissora, permitindo validação e ajustes por especialistas, garantindo decisões clínicas precisas e confiáveis (FURRIEL et al., 2024; ZHONG et al., 2023; LÉVÊQUE et al., 2021). O campo da segmentação de imagens médicas está avançando rapidamente, especialmente no uso de técnicas de aprendizagem profunda para aumentar a precisão da segmentação.

Este trabalho demonstra que a aplicação de RNC em cenários não supervisionados produz resultados promissores, sugerindo potencial maior quando combinada com abordagens supervisionadas e semi-supervisionadas. Esta integração visa capturar de forma mais eficiente as complexidades inerentes às imagens médicas, melhorando assim a precisão do diagnóstico e a personalização do tratamento. Além disto, a incorporação de dados multimodais de diversas modalidades de imagem, como ressonância magnética, tomografia computadorizada e ultrassom, é reconhecida como um avanço necessário. A combinação destas diversas fontes de dados pode enriquecer os modelos de segmentação, aumentando sua resiliência e precisão ao fornecer visões mais abrangentes e detalhadas dos aspectos clínicos a serem analisados.

Uma direção importante para o futuro da segmentação de imagens médicas é o desenvolvimento de soluções capazes de operar em tempo real, especialmente em contextos clínicos nos quais decisões rápidas são importantes. A capacidade de realizar segmentações precisas rapidamente poderá revolucionar os procedimentos cirúrgicos e de diagnóstico, permitindo intervenções imediatas com base em informações detalhadas e confiáveis. Isto implica em desafios tanto no desenvolvimento de algoritmos altamente eficientes quanto no avanço da infraestrutura de hardware para garantir a viabilidade destas tecnologias em ambientes clínicos.

## CAPÍTULO 7

### CONCLUSÃO

Este trabalho desenvolveu metodologia de segmentação não supervisionada de tomografias computadorizadas cranianas com base na abordagem de (KIM et al., 2020). Os objetivos foram plenamente alcançados ao implementar a arquitetura de rede neural profunda para segmentar estruturas intracranianas sem necessidade de rótulos prévios, anotação manual ou supervisão. Três técnicas de treinamento foram comparadas, utilizando otimização para encontrar valores de parâmetros, hiperparâmetros e determinar o número de máscaras de segmentação. A metodologia foi avaliada por especialistas e comparada a outras ferramentas. A hipótese principal foi confirmada, demonstrando eficiência ao treinar a rede neural em único exame aleatório, reduzindo recursos, tempo de treinamento e custos, aliviando o trabalho do médico radiologista. Em comparação com a abordagem de Kim et al. (2020), a metodologia proposta alcançou assertividade superior a 65%. Os resultados indicaram melhor desempenho na segmentação da substância branca e resultados similares ou superiores à ferramenta CTSeg na volumetria craniana. Os especialistas recomendaram a faixa de  $3 \leq N_R \leq 6$  para obter resultados visualmente melhores.

Este estudo contribui significativamente para o avanço na precisão da segmentação de tecidos cerebrais em imagens médicas, com potencial impacto em aplicações clínicas e diagnósticas. Representa progresso na segmentação automática não supervisionada em tomografias computadorizadas cranianas, oferecendo abordagem simplificada e acessível, reduzindo disparidades na prática clínica e facilitando a identificação precoce de condições críticas. Portanto, conclui-se que, a metodologia proposta apresenta-se aplicável na segmentação de tomografias computadorizadas cranianas, contribuindo significativamente para a análise de imagens médicas, tanto na clínica quanto no diagnóstico.

#### 7.1 Contribuições do trabalho

As contribuições podem assim serem descritas: i) adequação da metodologia não supervisionada de segmentação para aplicação em imagens médicas, ii) desenvolvimento de função de custo para segmentação não supervisionada que relaciona similaridade e continuidade, iii) implementação de processo de otimização dos parâmetros e hiperparâmetros da rede e iv) aplicação em base de exames pública CQ500.

**Artigos em revistas:**

Santos, P. V.; Martins, M. S. R.; Nogueira S. A.; Gonçalves C.; Loureiro R. M.; Calixto W. P.. Unsupervised model for structure segmentation applied to brain computed tomography. PLoS One. Fator de Impacto(2022 JCR): 3,7000, 2024

Nogueira S, Luz F, Ortiz de Camargo T F, dos Santos P V, Tornieri D, Calixto W. Myocardial perfusion scintigraphy image database (version 1.0.1). PhysioNet. 2024.

Silva L O, Bueno da Silva M C, Ribeiro G A S, Ortiz de Camargo T F, dos Santos P V, Mendes G S, de Paiva J P Q, Soares A S, Reis M R C, Loureiro R M, Calixto W. Artificial intelligence-based pulmonary embolism classification: development and validation using real-world data. Plos One. 2024.

Nogueira, S. A.; Luz, F. A. B.; Camargo, T. F. O.; Oliveira, J. C. S.; Neto, G. C. C.; Carvalhaes, F. B. F.; Reis M. R. C.; Santos, P. V.; Mendes G. S.; Loureiro, R. M.; Tornieri, D.; Pacheco, V. M. G.; Coimbra, A. P.; Calixto W. P.. Artificial intelligence applied in identifying left ventricular walls in myocardial perfusion scintigraphy images: pilot study. Plos One. 2024.

Pacheco, V. M. G.; Paiva, J. P. Q.; Furriel, B. C. R. S.; Santos, P. V.; Ferreira Junior, J. R.; Reis, M. R. C.; W. P. ; Tornieri, D.; Ribeiro, G. A. S.; Silva, L. O.; Nogueira, S. A.; Loureiro, R. M.; Calixto, W. P.. Pilot deployment of a cloud-based universal medical image repository in a large public health system: a protocol study. PLoS One, 2024.

### **Artigos em congressos:**

Simm, Vinicius Sylvestre.; Lima, Jader Tavares.; De Souza, Joao Pedro Busnardo.; Mello, Murilo Moreira.; Costa, Mateus Dalla.; Seixas, Monique Borges.; De Figueiredo, Maria Fernanda Oliveira.; Premebida, Sthefanie Monica.; Dos Santos, Paulo Victor.; Pacheco, Wesley.; Dos Santos Lima, Heron Oliveira.; Scoczynski, Marcella. Polyp Segmentation in Colonoscopy Images for Detection and Diagnosis of Colorectal Cancer. In: 2023 15th IEEE International Conference on Industry Applications (INDUSCON), 2023, São Bernardo do Campo. 2023 15th IEEE International Conference on Industry Applications (INDUSCON), 2023. p. 989.

De Souza, João Pedro Busnardo.; Lima, Jader Tavares.; Simm, Vinicius Sylvestre.; Mello, Murilo Moreira.; Costa, Mateus Dalla.; Seixas, Monique Borges.; De Figueiredo, Maria Fernanda Oliveira.; Premebida, Sthefanie Monica.; Dos Santos, Paulo Victor.; Pacheco, Wesley.; Dos Santos Lima, Heron Oliveira.; Scoczynski, Marcella. Polyp Detection in Colonoscopy Images Using a Vision Transformer Classifier. In:

2023 15th IEEE International Conference on Industry Applications (INDUSCON), 2023, São Bernardo do Campo. 2023 15th IEEE International Conference on Industry Applications (INDUSCON), 2023. p. 627.

Camargo, T. F. O.; Santos, P. V.; Premevida, S. M.; Ribeiro, G. A. S.; Soares, V. R.; Olombrada, M.; Golcalves, C.; Baroncini, V. H.; Correa, F. C.; Calixto, W. P.; Barbosa, R. M.; Scoczynski, R. M.. A practical Deep Learning approach to assist COVID-19 detection based on Chest X-ray images.. In: 2021 IEEE LA-CCI, 2021, Tenucco. Proceedings, 2021.

Seixas, M. B.; Souza, J. P. B.; Mello, M. M.; Lima, J. T.; Simm, V. S.; Costa, M. D.; Santos, P. V.; Figueiredo, M. F. O.; Premevida, S. M.; Lima, H. S.; Calixto, W. P.; Martins, Marcella. Segmentação de Polipos em imagens de colonoscopia para detecção e diagnostico de câncer colorretal. In: XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR, 2023, Ponta Grossa. Anais do XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR. Goiânia - GO: Zoe. Ed. Metropolitan Tokyo, 2023.

Seixas, M. B.; Souza, J. P. B.; Mello, M. M.; Lima, J. T.; Figueiredo, M. F. O.; Premevida, S. M.; Lima, H. S.; Simm, V. S.; Costa, M. D.; Calixto, W. P.; Santos, P. V.; Martins, Marcella. Classificador Vision Transformer para detecção de pólipos em imagens de colonoscopia. In: XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR, 2023, Ponta Grossa. Anais do XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR. Goiânia - GO: Zoe. Ed. Metropolitan Tokyo, 2023. v. 1.

## 7.2 Trabalhos futuros

- Aplicar a metodologia desenvolvida nas mesmas imagens de [Kim et al. \(2020\)](#) para analisar o desempenho da proposta em outros grupos de imagens.
- Aplicar o estudo agrupando rótulos semelhantes.
- Aplicar o estudo de segmentação em outros tipos de exames.
- Aplicar o estudo de segmentação em grupos de patologia.
- Aplicar o estudo de segmentação em imagens de ressonância magnética.

- Realizar testes em outras funções de perda específica para o tecido cerebral.
- Realizar testes e comparações com metodologias semi-supervisionadas.

## APÊNDICE A

### Processo de permutação

Considere a estrutura de dados representada em (A.1), na qual a representação dimensional (2, 3, 4). Sendo a matriz (A.1) composta por duas matrizes  $[3 \times 4]$  de valores aleatórios.

$$\left[ \begin{array}{c} \begin{pmatrix} -1.7872 & 0.1316 & -0.8408 & -1.0083 \\ 2.3570 & 1.3677 & 0.3449 & 0.3241 \\ 0.3122 & 0.9390 & 1.5273 & -0.2647 \end{pmatrix} \\ \begin{pmatrix} -1.2692 & 0.3952 & -1.5439 & 0.9820 \\ -0.9603 & 0.7267 & -0.2874 & -1.1066 \\ -1.1911 & 1.9282 & 1.1305 & 0.1571 \end{pmatrix} \end{array} \right] \quad (\text{A.1})$$

A permutação dos índices estabelecidos em ordem sequencial da estrutura em (A.1) deverá obedecer a troca dos valores dos índices iniciando por zero. Por exemplo, para gerar nova estrutura de dados de dimensão (3, 2, 4) a partir dos dados existentes, basta a troca dos índices de (0, 1, 2) para (1, 0, 2). A saída gerada terá a ordenação, como dados em (A.2), sendo a matriz (A.2) agora composta por três matrizes  $[2 \times 4]$ .

$$\left[ \begin{array}{c} \begin{pmatrix} -1.7872 & 0.1316 & -0.8408 & -1.0083 \\ -1.2692 & 0.3952 & -1.5439 & 0.9820 \end{pmatrix} \\ \begin{pmatrix} 2.3570 & 1.3677 & 0.3449 & 0.3241 \\ -0.9603 & 0.7267 & -0.2874 & -1.1066 \end{pmatrix} \\ \begin{pmatrix} 0.3122 & 0.9390 & 1.5273 & -0.2647 \\ -1.1911 & 1.9282 & 1.1305 & 0.1571 \end{pmatrix} \end{array} \right] \quad (\text{A.2})$$

Após o processo de permutação dos dados é aplicado novo formato de visualização a fim de reduzir a estrutura dos dados e garantir proximidade entre os valores similares independente da posição nos diferentes aspectos. Após estabelecido o dimensionamento da estrutura, conforme exemplo em (A.1) e (A.2), continua-se com o total de vinte e quatro valores. Portanto, a visualização final no formato de única matriz deverá obedecer a quantidade de elementos na matriz inicial, como a nova matriz de dados dada por:

$$\begin{bmatrix} -1.7872 & 0.1316 & -0.8408 & -1.0083 \\ 2.3570 & 1.3677 & 0.3449 & 0.3241 \\ 0.3122 & 0.9390 & 1.5273 & -0.2647 \\ -1.2692 & 0.3952 & -1.5439 & 0.9820 \\ -0.9603 & 0.7267 & -0.2874 & -1.1066 \\ -1.1911 & 1.9282 & 1.1305 & 0.1571 \end{bmatrix} \quad (\text{A.3})$$



## APÊNDICE B

### Processo do filtro de valores máximos

Considere a matriz (A.3) com valores aleatórios, o resultado obtido do filtro de valores máximos será a matriz (B.1). Tais características, fornecem solução unidimensional para os dados, denotando o filtro de valores máximos como filtro unidimensional.

$$[0.1316 \quad 2.3570 \quad 1.5273 \quad 0.9820 \quad 0.7267 \quad 1.9282] \quad (\text{B.1})$$

### Operação realizada no agrupador

Considere a matriz (B.2) com valores aleatórios e repetidos submetidos a função *Unique*, presente na ferramenta *Numpy* de Walt et al. (2011). O resultado será o vetor unidimensional dado por (B.3).

$$\begin{bmatrix} 1 & 0 & 3 & -1 \\ 2 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 \\ 4 & 0 & 4 & -1 \\ -1 & 1 & 1 & 0 \end{bmatrix} \quad (\text{B.2})$$

$$[-1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4] \quad (\text{B.3})$$



## APÊNDICE C

### Cálculo de avaliação do modelo proposto

Considere a Figura (C.1), que após o processo de convolução, a imagem será achatada horizontal e verticalmente, gerando duas novas imagens, denominadas Imagem- $Y$  e Imagem- $Z$ . Estas imagens serão submetidas à função de avaliação do erro absoluto médio (EAM) juntamente com a matriz de valores zerados. Quanto menor o valor encontrado, melhor a taxa de continuidade, uma vez que o EAM entre as duas estruturas de dados terá o comportamento similar ao da regressão linear. A Figura (C.2) ilustra a discrepância nas dimensões das matrizes após a etapa de permutação, juntamente com o vetor de valores máximos.

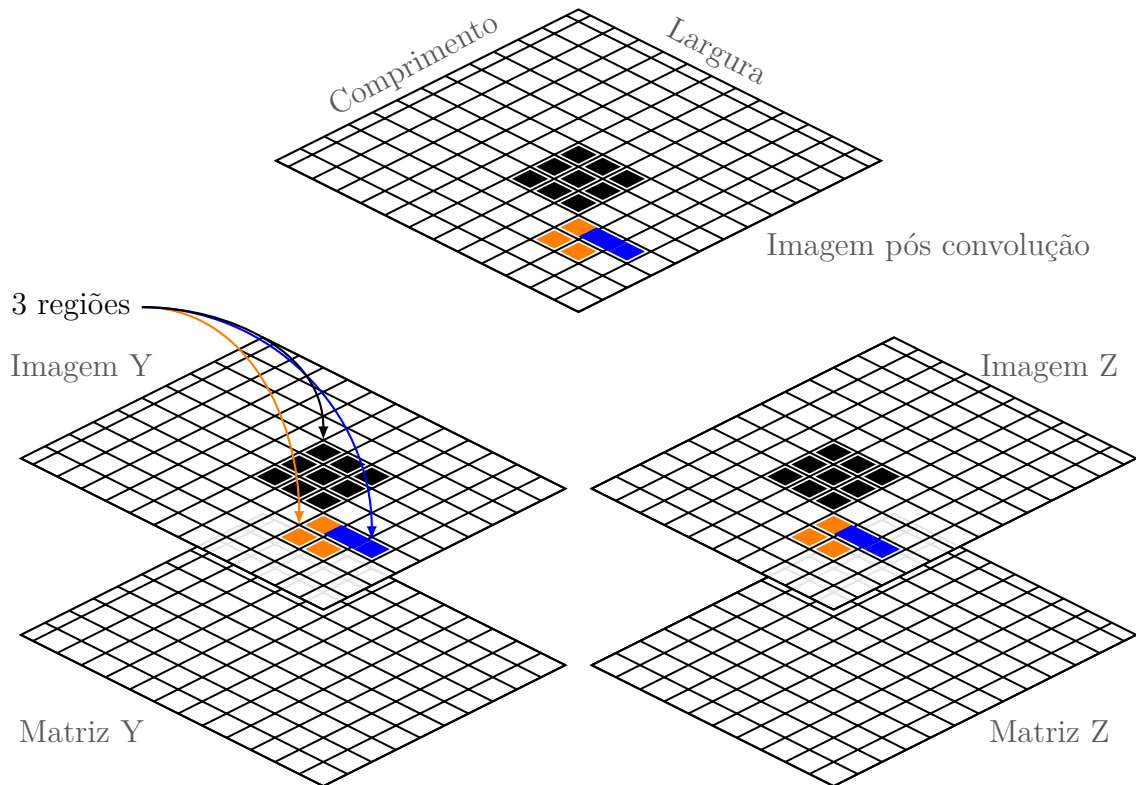
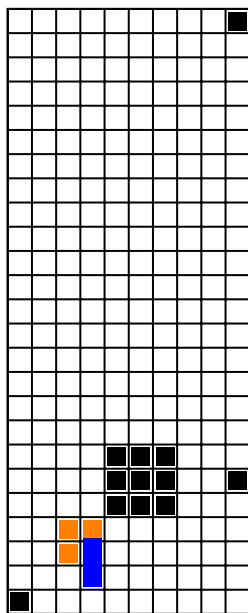


Figura C.1 - Cálculo da taxa de continuidade.

Dados pós permutação



Valores Máximos



Figura C.2 - Cálculo da entropia cruzada e da taxa de similaridade.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AGRAWAL, T. Hyperparameter optimization in machine learning. **Apress Berkeley: Berkeley, CA, USA**, p. 81–108, 2021. 56
- AIHARA, A. Y.; FERNANDES, A. d. R. C.; VIERTLER, C. M.; NATOUR, J. Tomografia multi-slice no sistema músculo-esquelético. **Revista Brasileira de Reumatologia**, SciELO Brasil, v. 43, p. 372–376, 2003. 42
- AJAY, P.; NAGARAJ, B.; KUMAR, R. A.; HUANG, R.; ANANTHI, P. et al. Unsupervised hyperspectral microscopic image segmentation using deep embedded clustering algorithm. **Scanning**, Hindawi, v. 2022, 2022. 58
- AKIBA, T.; SANO, S.; YANASE, T.; OHTA, T.; KOYAMA, M. Optuna: A next-generation hyperparameter optimization framework. In: **ACM. Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining**. Anchorage AK USA, 2019. p. 2623–2631. 56, 74, 78, 92
- AL-JABBAR, E. Y. A.; FATHEL, W. R.; QASIM, M. A.; NOORI, M. S.; ABDULJABAR, A. Q. Study axial plane with artificial intelligence for the purpose of classifying brain tasks. **Journal of Optoelectronics Laser**, v. 41, n. 4, p. 433–439, 2022. 36
- ALJABRI, M.; ALAMIR, M.; ALGHAMDI, M.; ABDEL-MOTTALEB, M.; COLLADO-MESA, F. Towards a better understanding of annotation tools for medical imaging: A survey. **Multimedia tools and applications**, Springer, v. 81, n. 18, p. 25877–25911, 2022. 44
- ALMEIDA, J. F. d. F.; PINTO, L. R.; CONCEIÇÃO, S. V.; CAMPOS, F. C. C. d. Medical centers location and specialists' allocation: a healthcare planning case study. **Production**, SciELO Brasil, v. 29, 2019. 31
- ALPAYDIN, E. **Introduction to machine learning**. Boston, MA: MIT press, 2020. 48
- AMBELLAN, F.; LAMECKER, H.; TYCOWICZ, C. von; ZACHOW, S. **Statistical shape models: understanding and mastering variation in anatomy**. New York: Springer, 2019. 93
- AMTHOR, F. **Neurobiology For Dummies**. New York: Wiley, 2014. (For dummies). ISBN 9781118689318. 38

\_\_\_\_\_. **Neurociência Para Leigos**. Sao Paulo: Alta Books, 2017. (Para Leigos). ISBN 9788550801735. 38

ANDREASEN, N. C.; FLAUM, M.; SWAYZE, V.; O'LEARY, D. S.; ALLIGER, R.; COHEN, G.; EHRHARDT, J.; YUH, W. T. et al. Intelligence and brain structure in normal individuals. **American Journal of Psychiatry**, American Psychiatric Association, v. 150, p. 130–130, 1993. 38

ARNOLD, B. C. Pareto distribution. **Wiley StatsRef: Statistics Reference Online**, Wiley Online Library, p. 1–10, 2014. 77

ATKINS, M. S.; MACKIEWICH, B. T. Fully automatic segmentation of the brain in mri. **IEEE transactions on medical imaging**, IEEE, v. 17, n. 1, p. 98–107, 1998. 29

AZZEH, J.; ZAHRAN, B.; ALQADI, Z. Salt and pepper noise: Effects and removal. **JOIV: International Journal on Informatics Visualization**, v. 2, n. 4, p. 252–256, 2018. 80

BAFARAJ, S. M. et al. Evaluation of neurological disorder using computed tomography and magnetic resonance imaging. **Journal of Biosciences and Medicines**, Scientific Research Publishing, v. 9, n. 02, p. 42, 2021. 34, 37

BALAFAR, M. A.; RAMLI, A. R.; SARIPAN, M. I.; MASHOHOR, S. Review of brain mri image segmentation methods. **Artificial Intelligence Review**, Springer, v. 33, n. 3, p. 261–274, 2010. 29

BATRA, D.; KOWDLE, A.; PARIKH, D.; LUO, J.; CHEN, T. icoseg: Interactive co-segmentation with intelligent scribble guidance. In: IEEE. **2010 IEEE computer society conference on computer vision and pattern recognition**. San Francisco, CA, USA, 2010. p. 3169–3176. 64

BAUR, C.; WIESTLER, B.; ALBARQOUNI, S.; NAVAB, N. Deep autoencoding models for unsupervised anomaly segmentation in brain mr images. In: SPRINGER. **International MICCAI brainlesion workshop**. Granada, Spain, 2018. p. 161–169. 53

BEAR, M.; CONNORS, B.; PARADISO, M. A. **Neuroscience: exploring the brain, enhanced edition: exploring the brain**. Burlington, MA 01803 USA: Jones & Bartlett Learning, 2020. 40

BEAR, M. F.; CONNORS, B. W.; PARADISO, M. A. **Neurociências: desvendando o sistema nervoso**. Porto Alegre: Artmed editora, 2002. 37

BERGSTRA, J.; BARDENET, R.; BENGIO, Y.; KÉGL, B. Algorithms for hyper-parameter optimization. **Advances in neural information processing systems**, v. 24, 2011. 56

BERSTAD, T. J. D.; RIEGLER, M.; ESPELAND, H.; LANGE, T. de; SMEDSRUD, P. H.; POGORELOV, K.; STENSLAND, H. K.; HALVORSEN, P. Tradeoffs using binary and multiclass neural network classification for medical multidisease detection. In: IEEE. **2018 IEEE International Symposium on Multimedia (ISM)**. Taichung, Taiwan, 2018. p. 1–8. 52

BERTELS, J.; EELBODE, T.; BERMAN, M.; VANDERMEULEN, D.; MAES, F.; BISSCHOPS, R.; BLASCHKO, M. B. Optimizing the dice score and jaccard index for medical image segmentation: Theory and practice. In: SPRINGER. **Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22**. Switzerland AG, 2019. p. 92–100. 75

BOTTOU, L.; BOUSQUET, O. The tradeoffs of large scale learning. **Advances in neural information processing systems**, v. 20, 2007. 56

BRINKER, T.; STOPA, E.; MORRISON, J.; KLINGE, P. A new look at cerebrospinal fluid circulation. **Fluids and Barriers of the CNS**, BioMed Central, v. 11, n. 1, p. 1–16, 2014. 40

BRODER, J. **Diagnostic Imaging for the Emergency Physician E-Book: Expert Consult - Online and Print**. St. Louis, MO 63043 USA: Elsevier Health Sciences, 2011. ISBN 9781437735871. 43

BRUDFORS, M. **Generative Models for Preprocessing of Hospital Brain Scans**. Tese (Doutorado) — UCL (University College London), 2020. 87, 92

BRUDFORS, M.; BALBASTRE, Y.; FLANDIN, G.; NACHEV, P.; ASHBURNER, J. Flexible bayesian modelling for nonlinear image registration. In: SPRINGER. **International Conference on Medical Image Computing and Computer-Assisted Intervention**. New York, NY 10004-1562 USA: Springer, 2020. p. 253–263. 87, 92

BUDDAY, S.; NAY, R.; ROOIJ, R. de; STEINMANN, P.; WYROBEK, T.; OVAERT, T. C.; KUHL, E. Mechanical properties of gray and white matter brain tissue by indentation. **Journal of the mechanical behavior of biomedical materials**, Elsevier, v. 46, p. 318–330, 2015. 40, 41

CAO, C.; LIU, X.; YANG, Y.; YU, Y.; WANG, J.; WANG, Z.; HUANG, Y.; WANG, L.; HUANG, C.; XU, W. et al. Look and think twice: Capturing top-down visual attention with feedback convolutional neural networks. In: **Proceedings of the IEEE international conference on computer vision**. Santiago, Chile: IEEE Computer Society, 2015. p. 2956–2964. 49

CAPA, N. Wilhelm Röntgen e a criação dos raios x. **J Bras Patol Med Lab**, SciELO Brasil, v. 45, n. 1, 2009. 33

CAPELLE, A.-S.; ALATA, O.; FERNANDEZ, C.; LEFÈVRE, S.; FERRIE, J. Unsupervised segmentation for automatic detection of brain tumors in mri. In: IEEE. **Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)**. Vancouver, BC, Canada, 2000. v. 1, p. 613–616. 53

CARITÁ, E. C.; MATOS, A. L. M.; AZEVEDO-MARQUES, P. M. d. Ferramentas para visualização de imagens médicas em hospital universitário. **Radiologia Brasileira**, SciELO Brasil, v. 37, p. 437–440, 2004. 45

CARLOS, M. T. Tomografia computadorizada: Formação da imagem e radioproteção. **LNMRI, IRD/CNEN**, 2002. 35

CARUANA, R.; NICULESCU-MIZIL, A. An empirical comparison of supervised learning algorithms. In: ACM. **Proceedings of the 23rd international conference on Machine learning**. Carnegie, 2006. p. 161–168. 28

CARVALHO, A. C. P. História da tomografia computadorizada. **Revista Imagem**, v. 29, n. 2, p. 61–66, 2007. 34

CHAGAS, E. T. D. O. Deep learning e suas aplicações na atualidade. **Revista Científica Multidisciplinar Núcleo do Conhecimento. Ano**, v. 4, p. 05–26, 2019. 49

CHAI, T.; DRAXLER, R. R. Root mean square error (rmse) or mean absolute error (mae). **Geoscientific Model Development Discussions**, v. 7, n. 1, p. 1525–1534, 2014. 54

CHEN, G. T. Y. Ct in radiation therapy treatment planning. In: WITHERS, H. R.; PETERS, L. J. (Ed.). **Innovations in Radiation Oncology**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1988. p. 191–200. 36

CHILAMKURTHY, S.; GHOSH, R.; TANAMALA, S.; BIVIJI, M.; CAMPEAU, N. G.; VENUGOPAL, V. K.; MAHAJAN, V.; RAO, P.; WARIER, P. Deep

learning algorithms for detection of critical findings in head ct scans: a retrospective study. **The Lancet**, Elsevier, v. 392, n. 10162, p. 2388–2396, 2018. 28, 77

CHOLLET, F. **Deep learning with Python**. Nova Iorque: Simon and Schuster, 2021. 49

CLARK, M. C.; HALL, L. O.; GOLDFOG, D. B.; VELTHUIZEN, R.; MURTAGH, R.; SILBINGER, M. S. Unsupervised brain tumor segmentation using knowledge-based fuzzy techniques. In: **Fuzzy and neuro-fuzzy systems in medicine**. Boca Raton, FL: CRC Press, 2017. p. 137–170. 53

CORDEIRO, F. R.; CARNEIRO, G. A survey on deep learning with noisy labels: How to train your model when you cannot trust on the annotations? In: IEEE. **2020 33rd SIBGRAPI conference on graphics, patterns and images (SIBGRAPI)**. Recife, Brazil: IEEE, 2020. p. 9–16. 28

DALCA, A. V.; YU, E.; GOLLAND, P.; FISCHL, B.; SABUNCU, M. R.; IGLESIAS, J. E. Unsupervised deep learning for bayesian brain mri segmentation. In: SPRINGER. **International Conference on Medical Image Computing and Computer-Assisted Intervention**. Vancouver, Canada, 2019. p. 356–365. 29, 31, 53

DIAS, P. A.; SHEN, Z.; TABB, A.; MEDEIROS, H. Freelabel: a publicly available annotation tool based on freehand traces. In: IEEE. **2019 IEEE Winter Conference on Applications of Computer Vision (WACV)**. Waikoloa Village, HI, USA: IEEE, 2019. p. 21–30. 44

DIPLAROS, A.; VLASSIS, N.; GEVERS, T. A spatially constrained generative model and an em algorithm for image segmentation. **IEEE Transactions on Neural Networks**, IEEE, v. 18, n. 3, p. 798–808, 2007. 58

DOMÍNGUEZ, M. G.; HERNÁNDEZ, C.; RUISOTO, P.; JUANES, J. A.; PRATS, A.; HERNÁNDEZ, T. Morphological and volumetric assessment of cerebral ventricular system with 3d slicer software. **Journal of medical systems**, Springer, v. 40, p. 1–8, 2016. 40

EE, C.; SIM, K.; TEH, V.; TING, F. Estimation of window width setting for ct scan brain images using mean of greyscale level to standard deviation ratio. In: IEEE. **2016 International Conference on Robotics, Automation and Sciences (ICORAS)**. Melaka, Malaysia: IEEE Computer Society, 2016. p. 1–6. 80

EELBODE, T.; BERTELS, J.; BERMAN, M.; VANDERMEULEN, D.; MAES, F.; BISSCHOPS, R.; BLASCHKO, M. B. Optimization for medical image segmentation: theory and practice when evaluating with dice score or jaccard index. **IEEE Transactions on Medical Imaging**, IEEE, v. 39, n. 11, p. 3679–3690, 2020. 75

ESTER, M.; KRIEGEL, H.-P.; SANDER, J.; XU, X. Density-based spatial clustering of applications with noise. In: **Int. Conf. knowledge discovery and data mining**. German: AAAI Press, 1996. v. 240, n. 6. 57

EVERINGHAM, M.; GOOL, L. V.; WILLIAMS, C. K.; WINN, J.; ZISSERMAN, A. The pascal visual object classes (voc) challenge. **International journal of computer vision**, Springer, v. 88, p. 303–338, 2010. 59

FEURER, M.; HUTTER, F. Hyperparameter optimization. Springer, Cham, p. 3–33, 2019. 55

FIGUEIREDO, M. A. T.; JAIN, A. K. Unsupervised learning of finite mixture models. **IEEE Transactions on pattern analysis and machine intelligence**, Ieee, v. 24, n. 3, p. 381–396, 2002. 57

FILHO, A. de S. A.; DUNNINGHAM, W. A. Valorada substância esnobada por décadas. **Revista Brasileira de Neurologia e Psiquiatria**, v. 17, n. 1, 2013. 41

FLECK, L.; TAVARES, M. H. F.; EYNG, E.; HELMANN, A.; ANDRADE, M. d. M. Redes neurais artificiais: Princípios básicos. **Revista Eletrônica Científica Inovação e Tecnologia**, v. 1, n. 13, p. 47–57, 2016. 48

FURRIEL, B. C.; OLIVEIRA, B. D.; PRÔA, R.; PAIVA, J. Q.; LOUREIRO, R. M.; CALIXTO, W. P.; REIS, M. R.; GIAVINA-BIANCHI, M. Artificial intelligence for skin cancer detection and classification for clinical environment: a systematic review. **Frontiers in Medicine**, Frontiers, v. 10, p. 1305954, 2024. 94

GARIB, D. G.; JR, R. R.; RAYMUNDO, M. V.; RAYMUNDO, D. V.; FERREIRA, S. N. Tomografia computadorizada de feixe cônico (cone beam): entendendo este novo método de diagnóstico por imagem com promissora aplicabilidade na ortodontia. **Revista Dental Press de Ortodontia e Ortopedia Facial**, SciELO Brasil, v. 12, p. 139–156, 2007. 35

GATIDIS, S.; HEPP, T.; FRÜH, M.; FOUGÈRE, C. L.; NIKOLAOU, K.; PFANNENBERG, C.; SCHÖLKOPF, B.; KÜSTNER, T.; CYRAN, C.; RUBIN,

- D. A whole-body fdg-pet/ct dataset with manually annotated tumor lesions. **Scientific Data**, Nature Publishing Group UK London, v. 9, n. 1, p. 601, 2022. 28
- GAZZANIGA, M. S. Cerebral specialization and interhemispheric communication: does the corpus callosum enable the human condition? **Brain**, Oxford University Press, v. 123, n. 7, p. 1293–1326, 2000. 38
- GÉRON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. Sebastopol, CA 95472 USA: O'Reilly Media, 2019. ISBN 9781492032618. 51
- \_\_\_\_\_. **Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow**. 2nd. ed. Sebastopol, CA 95472 USA: "O'Reilly Media, Inc.", 2022. ISBN 9781098122461. 48, 51
- GHRIBI, O.; MAALEJ, A.; SELLAMI, L.; SLIMA, M. B.; MAALEJ, M. A.; MAHFOUDH, K. B.; DAMMAK, M.; MHIRI, C.; HAMIDA, A. B. Advanced methodology for multiple sclerosis lesion exploring: Towards a computer aided diagnosis system. **Biomedical Signal Processing and Control**, Elsevier, v. 49, p. 274–288, 2019. 85
- GIORGIO, A.; STEFANO, N. D. Clinical use of brain volumetry. **Journal of Magnetic Resonance Imaging**, Wiley Online Library, v. 37, n. 1, p. 1–14, 2013. 87
- GLIDE-HURST, C.; CHEN, D.; ZHONG, H.; CHETTY, I. Changes realized from extended bit-depth and metal artifact reduction in ct. **Medical physics**, Wiley Online Library, v. 40, n. 6Part1, p. 061711, 2013. 43
- GOLDBERGER, J.; ROWEIS, S. Hierarchical clustering of a mixture model. **Advances in neural information processing systems**, v. 17, 2004. 57
- GOLIK, P.; DOETSCH, P.; NEY, H. Cross-entropy vs. squared error training: a theoretical and experimental comparison. In: ISCA. **INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association, Lyon, France, August 25-29, 2013**. Lyon, France: ISCA, 2013. v. 13, p. 1756–1760. 54
- GONZALEZ, L. F. P. **Segmentação automática de tecidos cerebrais em imagens de ressonância magnética do tipo fluid-attenuated inversion recovery**. Brasil. Tese (Doutorado) — Pontifícia Universidade Católica do Rio

Grande do Sul, Programa de Pós-Graduação em Ciência da Computação, abril 2018. 40

GONZÁLEZ-VILLÀ, S.; OLIVER, A.; VALVERDE, S.; WANG, L.; ZWIGGELAAR, R.; LLADÓ, X. A review on brain structures segmentation in magnetic resonance imaging. **Artificial intelligence in medicine**, Elsevier, v. 73, p. 45–69, 2016. 53

GOOCH, C. L.; PRACHT, E.; BORENSTEIN, A. R. The burden of neurological disease in the united states: A summary report and call to action. **Annals of neurology**, v. 81, n. 4, p. 479–484, 2017. 27

GÜNTÜRKÜN, O.; STRÖCKENS, F.; OCKLENBURG, S. Brain lateralization: a comparative perspective. **Physiological reviews**, American Physiological Society Bethesda, MD, 2020. 40

GUSMÃO, S.; RIBAS, G. C.; SILVEIRA, R. L.; TAZINAFFO, U. Localização dos sulcos e giros da face súpero-lateral do cérebro na tomografia computadorizada e na ressonância magnética. **Arquivos de neuro-psiquiatria**, SciELO Brasil, v. 59, p. 65–70, 2001. 39

HALL, L. O.; BENSALID, A. M.; CLARKE, L. P.; VELTHUIZEN, R. P.; SILBINGER, M. S.; BEZDEK, J. C. A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. **IEEE transactions on neural networks**, IEEE, v. 3, n. 5, p. 672–682, 1992. 53

HAYKIN, S. **Adaptive Filter Theory**. Upper Saddle River, NJ 07458 USA: Prentice Hall, 1996. (Prentice-Hall information and system sciences series). ISBN 9780133227604. 48

HENRY, T.; CARRÉ, A.; LEROUSSEAU, M.; ESTIENNE, T.; ROBERT, C.; PARAGIOS, N.; DEUTSCH, E. Brain tumor segmentation with self-ensembled, deeply-supervised 3d u-net neural networks: a brats 2020 challenge solution. In: SPRINGER. **Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part I 6**. Paris, France, 2021. p. 327–339. 92

HINTON, G. Deep learning—a technology with the potential to transform health care. **Jama**, American Medical Association, v. 320, n. 11, p. 1101–1102, 2018. 56

HIRAN, K.; JAIN, R.; LAKHWANI, K.; DOSHI, R. **Machine Learning: Master Supervised and Unsupervised Learning Algorithms with Real Examples (English Edition)**. Seattle, WA 98109-5210 USA: BPB Publications, 2021. (ITpro collection). ISBN 9789391392352. 48, 49

HO, M.-L.; ROJAS, R.; EISENBERG, R. L. Cerebral edema. **American Journal of Roentgenology**, American Roentgen Ray Society, v. 199, n. 3, p. W258–W273, 2012. 80

HO, P.-G. **Image segmentation**. Erfurt Germany: BoD–Books on Demand, 2011. 53

HUA, L.; GU, Y.; GU, X.; XUE, J.; NI, T. A novel brain mri image segmentation method using an improved multi-view fuzzy c-means clustering algorithm. **Frontiers in Neuroscience**, Frontiers Media SA, v. 15, p. 662674, 2021. 30, 31

JOHNSTONE, E.; FRITH, C.; CROW, T.; HUSBAND, J.; KREEL, L. Cerebral ventricular size and cognitive impairment in chronic schizophrenia. **The Lancet**, Elsevier, v. 308, n. 7992, p. 924–926, 1976. 40

JÚNIOR, E. A.; YAMASHITA, H. Aspectos básicos de tomografia computadorizada e ressonância magnética. **Brazilian Journal of Psychiatry**, SciELO Brasil, v. 23, p. 2–3, 2001. 33

KAMALIAN, S.; LEV, M. H.; GUPTA, R. Computed tomography imaging and angiography – principles. **Handbook of Clinical Neurology**, Elsevier B.V., v. 135, p. 3–20, 12 2016. ISSN 22124152. 33

KANDEL, E. R.; SCHWARTZ, J. H.; JESSELL, T. M. (Ed.). **Principles of Neural Science**. Third. New York: Elsevier, 1991. 40

KANEZAKI, A. Unsupervised image segmentation by backpropagation. In: **IEEE. 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)**. Calgary, AB, Canada: IEEE, 2018. p. 1543–1547. 58

KER, J.; WANG, L.; RAO, J.; LIM, T. Deep learning applications in medical image analysis. **Ieee Access**, IEEE, v. 6, p. 9375–9389, 2017. 28

KHAN, A. R.; KHAN, S.; HAROUNI, M.; ABBASI, R.; IQBAL, S.; MEHMOOD, Z. Brain tumor segmentation using k-means clustering and deep learning with synthetic data augmentation for classification. **Microscopy Research and Technique**, Wiley Online Library, v. 84, n. 7, p. 1389–1399, 2021. 30, 31

- KIM, W.; KANEZAKI, A.; TANAKA, M. Unsupervised learning of image segmentation based on differentiable feature clustering. **IEEE Transactions on Image Processing**, IEEE, v. 29, p. 8055–8068, 2020. 19, 31, 58, 59, 60, 61, 62, 63, 64, 65, 68, 69, 72, 82, 83, 84, 85, 92, 95, 97
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 2014. 56
- KLINE, D. M.; BERARDI, V. L. Revisiting squared-error and cross-entropy functions for training neural network classifiers. **Neural Computing & Applications**, Springer, v. 14, n. 4, p. 310–318, 2005. 54
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, v. 25, 2012. 28
- KULMANOV, M.; SMAILI, F. Z.; GAO, X.; HOEHNDORF, R. Semantic similarity and machine learning with ontologies. **Briefings in bioinformatics**, Oxford University Press, v. 22, n. 4, p. bbaa199, 2021. 55
- KÜHN, S.; GALLINAT, J. Brain Structure and Functional Connectivity Associated With Pornography Consumption: The Brain on Porn. **JAMA Psychiatry**, v. 71, n. 7, p. 827–834, 07 2014. ISSN 2168-622X. 38
- LANGEN, K.-J.; GALLDIKS, N.; HATTINGEN, E.; SHAH, N. J. Advances in neuro-oncology imaging. **Nature Reviews Neurology**, Nature Publishing Group, v. 13, n. 5, p. 279–289, 2017. 27
- LAW, M. T.; URTASUN, R.; ZEMEL, R. S. Deep spectral clustering learning. In: PMLR. **International conference on machine learning**. Toronto, Canada: PMLR, 2017. p. 1985–1994. 57
- LE, H.; NGUYEN, M.; YAN, W. Q. A web-based augmented reality approach to instantly view and display 4d medical images. In: PALAIAHNAKOTE, S.; BAJA, G. Sanniti di; WANG, L.; YAN, W. Q. (Ed.). **Pattern Recognition**. Cham: Springer International Publishing, 2020. p. 691–704. ISBN 978-3-030-41299-9. 37
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, Ieee, Paris, France, v. 86, n. 11, p. 2278–2324, 1998. 51

- LEE, T. H.; FAUZI, M. F. A.; KOMIYA, R.; HAW, S.-C. Unsupervised abnormalities extraction and brain segmentation. In: **IEEE. 2008 3rd International Conference on Intelligent System and Knowledge Engineering**. Xiamen, China: IEEE, 2008. v. 1, p. 1185–1190. 29, 31
- LEI, T.; NANDI, A. **Image Segmentation: Principles, Techniques, and Applications**. Hoboken, NJ 07030-5774 USA: John Wiley & Sons, 2022. ISBN 9781119859000. 52
- LENCHIK, L.; HEACOCK, L.; WEAVER, A. A.; BOUTIN, R. D.; COOK, T. S.; ITRI, J.; FILIPPI, C. G.; GULLAPALLI, R. P.; LEE, J.; ZAGUROVSKAYA, M. et al. Automated segmentation of tissues using ct and mri: a systematic review. **Academic radiology**, Elsevier, v. 26, n. 12, p. 1695–1706, 2019. 30
- LÉVÊQUE, L.; OUTTAS, M.; LIU, H.; ZHANG, L. Comparative study of the methodologies used for subjective medical image quality assessment. **Physics in Medicine & Biology**, IOP Publishing, v. 66, n. 15, p. 15TR02, 2021. 94
- LI, L.; WEI, M.; LIU, B.; ATCHANEYASAKUL, K.; ZHOU, F.; PAN, Z.; KUMAR, S. A.; ZHANG, J. Y.; PU, Y.; LIEBESKIND, D. S. et al. Deep learning for hemorrhagic lesion detection and segmentation on brain ct images. **IEEE journal of biomedical and health informatics**, IEEE, v. 25, n. 5, p. 1646–1659, 2020. 28
- LIMA, A. A.; MRIDHA, M. F.; DAS, S. C.; KABIR, M. M.; ISLAM, M. R.; WATANOBÉ, Y. A comprehensive survey on the detection, classification, and challenges of neurological disorders. **Biology**, MDPI, v. 11, n. 3, p. 469, 2022. 27
- LIMA, R. d. S.; AFONSO, J. C.; PIMENTEL, L. C. F. Raios-x: fascinação, medo e ciência. **Química nova**, SciELO Brasil, v. 32, p. 263–270, 2009. 33, 34
- LIN, T.-Y.; MAIRE, M.; BELONGIE, S.; HAYS, J.; PERONA, P.; RAMANAN, D.; DOLLÁR, P.; ZITNICK, C. L. Microsoft coco: Common objects in context. In: **SPRINGER. European conference on computer vision**. Zurich, Switzerland: Springer, 2014. p. 740–755. 59
- LIU, S.; QI, L.; QIN, H.; SHI, J.; JIA, J. Path aggregation network for instance segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. Salt Lake City, UT, USA: IEEE Computer Society, 2018. p. 8759–8768. 53

LOPES, C. I. A. **Estudo de materiais utilizando imagens de tomografia computadorizada para o desenvolvimento de objetos simuladores**. 2016. 33, 34

LORENA, A. C.; CARVALHO, A. C. D.; GAMA, J. M. A review on the combination of binary classifiers in multiclass problems. **Artificial Intelligence Review**, Springer, v. 30, n. 1, p. 19–37, 2008. 52

LUCAS, E. Marco de; SANCHEZ, E.; GUTIERREZ, A.; MANDLY, A. G.; RUIZ, E.; FLOREZ, A. F.; IZQUIERDO, J.; ARNAIZ, J.; PIEDRA, T.; VALLE, N.; BANALES, I.; QUINTANA, F. Ct protocol for acute stroke: Tips and tricks for general radiologists. **Radiographics**, Radiological Soc North America, OAK BROOK, v. 28, n. 6, p. 1673–1687, 2008. ISSN 0271-5333. 42

MAHATA, N.; KAHALI, S.; ADHIKARI, S. K.; SING, J. K. Local contextual information and gaussian function induced fuzzy clustering algorithm for brain mr image segmentation and intensity inhomogeneity estimation. **Applied Soft Computing**, Elsevier, v. 68, p. 586–596, 2018. 28, 29, 31

MALDONADO, K. A.; ALSAYOURI, K. Physiology, brain. In: **StatPearls [Internet]**. Treasure Island, FL 33706 USA: StatPearls Publishing, 2021. 38

MARTIN, A.; CHAO, L. L. Semantic memory and the brain: structure and processes. **Current opinion in neurobiology**, Elsevier, v. 11, n. 2, p. 194–201, 2001. 38

MARTIN, D.; FOWLKES, C.; TAL, D.; MALIK, J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: IEEE. **Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001**. Vancouver, BC, Canada, 2001. v. 2, p. 416–423. 59

MEDEIROS, J. **Qualidade de imagem versus dose em tomografia computadorizada: otimização dos protocolos de crânio**. Portugal. Tese (Doutorado) — Universidade de Coimbra, Faculdade de Ciências e Tecnologia, abril 2009. 34

MERCADANTE, A. A.; TADI, P. Neuroanatomy, gray matter. 2020. 40

MEURER, M. I.; MEURER, E.; SILVA, J. V. L. d.; BÁRBARA, A. S.; NOBRE, L. F.; OLIVEIRA, M. G. d.; SILVA, D. N. Aquisição e manipulação de imagens por tomografia computadorizada da região maxilofacial visando à obtenção de

protótipos biomédicos. **Radiologia Brasileira**, SciELO Brasil, v. 41, p. 49–54, 2008. 42

MILDENBERGER, P.; EICHELBERG, M.; MARTIN, E. Introduction to the dicom standard. **European radiology**, Springer, v. 12, n. 4, p. 920–927, 2002. 35, 45

MILLSTEIN, F. **Convolutional Neural Networks In Python: Beginner's Guide To Convolutional Neural Networks In Python**. New York, USA: Frank Millstein, 2020. 49

MILSZTAJN, F.; GEUS, K. de; ENERGIA, C.-C. P. de. Segmentação de tecidos cerebrais em imagens tridimensionais do cérebro. 2003. 41

MISRA, I.; MAATEN, L. v. d. Self-supervised learning of pretext-invariant representations. In: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**. Seattle, WA, USA: IEEE, 2020. p. 6707–6717. 58

MITRA, P.; MURTHY, C.; PAL, S. K. Unsupervised feature selection using feature similarity. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 24, n. 3, p. 301–312, 2002. 55

MONTEIRO, M.; NEWCOMBE, V. F.; MATHIEU, F.; ADATIA, K.; KAMNITSAS, K.; FERRANTE, E.; DAS, T.; WHITEHOUSE, D.; RUECKERT, D.; MENON, D. K. et al. Multiclass semantic segmentation and quantification of traumatic brain injury lesions on head ct using deep learning: an algorithm development and multicentre validation study. **The Lancet Digital Health**, Elsevier, v. 2, n. 6, p. e314–e322, 2020. 28

MORONEY, L. **AI and Machine Learning for Coders: A Programmer's Guide to Artificial Intelligence**. London, United Kingdom: O'Reilly, 2020. ISBN 9781492078197. 47

MOUNTFORD, P.; TEMPERTON, D. Recommendations of the international commission on radiological protection (icrp) 1990. **European Journal of Nuclear Medicine**, v. 19, n. 2, 1992. ISSN 0340-6997. 42

MURE, S.; GRENIER, T.; MEIER, D. S.; GUTTMANN, C. R.; BENOIT-CATTIN, H. Unsupervised spatio-temporal filtering of image sequences. a mean-shift specification. **Pattern Recognition Letters**, Elsevier, v. 68, p. 48–55, 2015. 57

NAZIR, S.; DICKSON, D. M.; AKRAM, M. U. Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks.

**Computers in Biology and Medicine**, Elsevier, p. 106668, 2023. 28

NOWAK, S.; RÜGER, S. How reliable are annotations via crowdsourcing: a study about inter-annotator agreement for multi-label image annotation. In: ACM.

**Proceedings of the international conference on Multimedia information retrieval**. Philadelphia, Pennsylvania, USA: ACM, 2010. p. 557–566. 35, 44

OPENAI. **Planos de aquisição**. Maio 2023. Imagem gerada pelo DALL-E. 36

ORGANIZATION, W. H. **Neurological disorders: public health challenges**.

Geneva Switzerland: World Health Organization, 2006. 27

ORTH, D. **Essentials of Radiologic Science**. New York, NY: Wolters Kluwer, 2017. (M - Medicine Series). ISBN 9781496317278. 34

OSBORNE, T.; TANG, C.; SABARWAL, K.; PRAKASH, V. How to interpret an unenhanced ct brain scan. part 1: Basic principles of computed tomography and relevant neuroanatomy. **South Sudan Medical Journal**, v. 9, n. 3, p. 67–69, 2016. 33, 34, 42

OWOLABI, M. O.; LEONARDI, M.; BASSETTI, C.; JAARSMA, J.; HAWROT, T.; MAKANJUOLA, A. I.; DHAMIJA, R. K.; FENG, W.; STRAUB, V.; CAMARADOU, J. et al. Global synergistic actions to improve brain health for human development. **Nature Reviews Neurology**, Nature Publishing Group UK London, v. 19, n. 6, p. 371–383, 2023. 27

PAGLIOLI, E. Cortes cerebrais: noções praticas de anatomia cerebral. **Revista dos Cursos da Faculdade de Medicina de Porto Alegre**, v. 22, n. 22, p. 61–73, 1936. 38

PAGNOZZI, A. M.; FRIPP, J.; ROSE, S. E. Quantifying deep grey matter atrophy using automated segmentation approaches: A systematic review of structural mri studies. **Neuroimage**, Elsevier, v. 201, p. 116018, 2019. 85

PATEL, U. K.; ANWAR, A.; SALEEM, S.; MALIK, P.; RASUL, B.; PATEL, K.; YAO, R.; SESHADRI, A.; YOUSUFUDDIN, M.; ARUMAITHURAI, K. Artificial intelligence as an emerging technology in the current care of neurological disorders. **Journal of neurology**, Springer, v. 268, p. 1623–1642, 2021. 27

- PHAM, D. L.; XU, C.; PRINCE, J. L. A survey of current methods in medical image segmentation. **Annual review of biomedical engineering**, v. 2, n. 3, p. 315–337, 2000. 28
- PINTER, C.; LASSO, A.; WANG, A.; JAFFRAY, D.; FICHTINGER, G. Slicerrt: radiation therapy research toolkit for 3d slicer. **Medical physics**, Wiley Online Library, v. 39, n. 10, p. 6332–6338, 2012. 79, 82, 92
- PONTI, M.; COSTA, B. Tópicos em gerenciamento de dados e informações. **SBC, Uberlândia**, p. 63–88, 2017. 49
- QIAN, B.; SU, J.; WEN, Z.; JHA, D. N.; LI, Y.; GUAN, Y.; PUTHAL, D.; JAMES, P.; YANG, R.; ZOMAYA, A. Y. et al. Orchestrating the development lifecycle of machine learning-based iot applications: A taxonomy and survey. **ACM Computing Surveys (CSUR)**, ACM New York, NY, USA, v. 53, n. 4, p. 1–47, 2020. 48
- RADFORD, A.; METZ, L.; CHINTALA, S. Unsupervised representation learning with deep convolutional generative adversarial networks. **arXiv preprint arXiv:1511.06434**, 2015. 58
- RADHIANA, H.; SYAZARINA, S.; AZURA, M. S.; HILWATI, H.; SOBRI, M. Non-contrast computed tomography in acute ischaemic stroke: a pictorial review. **Med J Malaysia**, PubMed, v. 68, n. 1, p. 93–100, 2013. 43
- RAJA, P. S. et al. Brain tumor classification using a hybrid deep autoencoder with bayesian fuzzy clustering-based segmentation approach. **Biocybernetics and Biomedical Engineering**, Elsevier, v. 40, n. 1, p. 440–453, 2020. 30, 31
- RAJINI, N. H.; BHAVANI, R. Automatic classification of computed tomography brain images using ann, k-nn and svm. **AI & society**, Springer, v. 29, p. 97–102, 2014. 92
- RAZI, T.; NIKNAMI, M.; GHAZANI, F. A. Relationship between hounsfield unit in ct scan and gray scale in cbct. **Journal of dental research, dental clinics, dental prospects**, Tabriz University of Medical Sciences, v. 8, n. 2, p. 107, 2014. 43
- RISTER, B.; YI, D.; SHIVAKUMAR, K.; NOBASHI, T.; RUBIN, D. L. Ct-org, a new dataset for multiple organ segmentation in computed tomography. **Scientific Data**, Nature Publishing Group, v. 7, n. 1, p. 1–9, 2020. 28

RODRÍGUEZ, A. A. Deep convolutional autoencoders for reconstructing magnetic resonance images of the healthy brain. Universitat Oberta de Catalunya (UOC), 2021. 58

ROMANO, R. F. T.; SALVADORI, P. S.; TORRES, L. R.; BRETAS, E. A. S.; BEKHOR, D.; CALDANA, R. P.; MEDEIROS, R. B.; D'IPPOLITO, G. Readequação de protocolos de exames de tomografia computadorizada de abdome em um hospital universitário: impacto na dose de radiação. **Radiologia Brasileira**, SciELO Brasil, v. 48, p. 292–297, 2015. 42

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. **International Conference on Medical image computing and computer-assisted intervention**. Munich, Germany: Springer International Publishing, 2015. p. 234–241. 28, 53

RUBINSTEIN, R. Y.; KROESE, D. P. **The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation, and machine learning**. New York: Springer, 2004. 55

SAADE, C.; NAJEM, E.; ASMAR, K.; SALMAN, R.; ACHKAR, B. E.; NAFFAA, L. Intracranial calcifications on ct: an updated review. **Journal of radiology case reports**, EduRad, v. 13, n. 8, p. 1, 2019. 80

SAIT, U.; KV, G. L.; SHIVAKUMAR, S.; KUMAR, T.; BHAUMIK, R.; PRAJAPATI, S.; BHALLA, K.; CHAKRAPANI, A. A deep-learning based multimodal system for covid-19 diagnosis using breathing sounds and chest x-ray images. **Applied Soft Computing**, Elsevier, v. 109, p. 107522, 2021. 28

SAKKA, L.; COLL, G.; CHAZAL, J. Anatomy and physiology of cerebrospinal fluid. **European annals of otorhinolaryngology, head and neck diseases**, Elsevier, v. 128, n. 6, p. 309–316, 2011. 40

SANDERS, R. The pareto principle: its use and abuse. **Journal of Services Marketing**, MCB UP Ltd, v. 1, n. 2, p. 37–40, 1987. 77

SANTANA, E. F.; BATISTA, L. V.; SILVA, R. Marques da; SANTOS, C. A. G. Multispectral image unsupervised segmentation using watershed transformation and cross-entropy minimization in different land use. **GIScience & remote sensing**, Taylor & Francis, v. 51, n. 6, p. 613–629, 2014. 57

- SANTOS, R.; PIRES, A.; ALMEIDA, R.; PEREIRA, W. Computed tomography scanner productivity and entry-level models in the global market. **Journal of healthcare engineering**, Hindawi, v. 2017, 2017. 31
- SANTOS, R. O. Estrutura e funções do córtex cerebral. 2002. 38, 39
- SCHMIDHUBER, J. Deep learning in neural networks: An overview. **Neural networks**, Elsevier, v. 61, p. 85–117, 2015. 28
- SENTHILKUMARAN, N.; RAJESH, R. Image segmentation-a survey of soft computing approaches. In: **IEEE. 2009 International Conference on Advances in Recent Technologies in Communication and Computing**. Kottayam, Kerala, India: IEEE Computer Society, 2009. p. 844–846. ISBN 978-0-7695-3845-7. 28
- SHIBATA, T.; TANAKA, M.; OKUTOMI, M. Misalignment-robust joint filter for cross-modal image pairs. In: **Proceedings of the IEEE International Conference on Computer Vision**. Venice, Italy: IEEE, 2017. p. 3295–3304. 72
- SINAGA, K. P.; YANG, M.-S. Unsupervised k-means clustering algorithm. **IEEE access**, IEEE, v. 8, p. 80716–80727, 2020. 57
- SIVAKUMAR, R. Denoising of computer tomography images using curvelet transform. **ARNP Journal of Engineering and Applied Sciences**, v. 2, n. 1, p. 21–26, 2007. 80
- SOUZA, S. A. S. d. Descrição de protocolo de tórax em tomografia computadorizada. Universidade Estadual Paulista (Unesp), 2016. 42
- SRINIVAS, P.; KATARYA, R. hyoptxg: Optuna hyper-parameter optimization framework for predicting cardiovascular disease using xgboost. **Biomedical Signal Processing and Control**, Elsevier, v. 73, p. 103456, 2022. 56
- STEINMETZ, J. D.; SEEHER, K. M.; SCHIESS, N.; NICHOLS, E.; CAO, B.; SERVILI, C.; CAVALLERA, V.; COUSIN, E.; HAGINS, H.; MOBERG, M. E. et al. Global, regional, and national burden of disorders affecting the nervous system, 1990–2021: a systematic analysis for the global burden of disease study 2021. **The Lancet Neurology**, Elsevier, v. 23, n. 4, p. 344–381, 2024. 27
- STRATCHKO, L.; FILATOVA, I.; AGARWAL, A.; KANEKAR, S. The ventricular system of the brain: anatomy and normal variations. In: **ELSEVIER. Seminars in Ultrasound, CT and MRI**. Bethesda, MD, USA, 2016. v. 37, n. 2, p. 72–83. 40

- SZEGEDY, C.; IOFFE, S.; VANHOUCKE, V.; ALEMI, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. In: **Thirty-first AAAI conference on artificial intelligence**. San Francisco, California, USA: AAAI Press, 2017. v. 31. 28
- TABAKOV, S.; MILANO, F.; SPRAWLS, P. **Encyclopaedia of Medical Physics**. New York, USA: CRC Press, 2020. ISBN 9781482224849. 43
- TAJBAKHS, N.; JEYASEELAN, L.; LI, Q.; CHIANG, J. N.; WU, Z.; DING, X. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. **Medical Image Analysis**, Elsevier, v. 63, p. 101693, 2020. 28
- TALAIRACH, J.; SZIKLA, G. Application of stereotactic concepts to the surgery of epilepsy. In: **Advances in Stereotactic and Functional Neurosurgery 4**. Vienna: Springer, 1980. p. 35–54. ISBN 978-3-7091-8592-6. 40
- TARIQ, H.; BURNEY, A. Brain mri literature review for interdisciplinary studies. **Journal of Biomedical Graphics and Computing**, v. 4, n. 4, p. 41–53, 2014. 35
- THADA, V.; JAGLAN, V. Comparison of jaccard, dice, cosine similarity coefficient to find best fitness value for web retrieved documents using genetic algorithm. **International Journal of Innovations in Engineering and Technology**, Citeseer, v. 2, n. 4, p. 202–205, 2013. 75
- TOGA, A. W.; MAZZIOTTA, J. C.; MAZZIOTTA, J. C. **Brain mapping: the methods**. Los Angeles, California: Academic press, 2002. 42, 43
- TOH, K. K. V.; ISA, N. A. M. Noise adaptive fuzzy switching median filter for salt-and-pepper noise reduction. **IEEE signal processing letters**, IEEE, v. 17, n. 3, p. 281–284, 2009. 80
- TRAN, P.; THOPRAKARN, U.; GOURIEUX, E.; SANTOS, C. L. D.; CAVEDO, E.; GUIZARD, N.; COTTON, F.; KROLAK-SALMON, P.; DELMAIRE, C.; HEIDELBERG, D. et al. Automatic segmentation of white matter hyperintensities: validation and comparison with state-of-the-art methods on both multiple sclerosis and elderly subjects. **NeuroImage: Clinical**, Elsevier, v. 33, p. 102940, 2022. 85
- TUSTISON, N. J.; SHRINIDHI, K.; WINTERMARK, M.; DURST, C. R.; KANDEL, B. M.; GEE, J. C.; GROSSMAN, M. C.; AVANTS, B. B. Optimal symmetric multimodal templates and concatenated random forests for supervised

brain tumor segmentation (simplified) with ants. **Neuroinformatics**, Springer, v. 13, p. 209–225, 2015. 92

USMAN, M. B.; OJHA, S.; JHA, S. K.; CHELLAPPAN, D. K.; GUPTA, G.; SINGH, S. K.; DUA, K.; ROYCHOUDHURY, S.; KUMAR, N.; KHAN, F. A. et al. Biological databases and tools for neurological disorders. **Journal of Integrative Neuroscience**, IMR Press, 2022. 27

VINEGAR, H. X-ray ct and nmr imaging of rocks. **Journal of petroleum technology**, SPE, v. 38, n. 03, p. 257–259, 1986. 37

WAHL, B.; COSSY-GANTNER, A.; GERMANN, S.; SCHWALBE, N. R. Artificial intelligence (ai) and global health: how can ai contribute to health in resource-poor settings? **BMJ global health**, BMJ Specialist Journals, v. 3, n. 4, p. e000798, 2018. 27

WALT, S. V. D.; COLBERT, S. C.; VAROQUAUX, G. The numpy array: a structure for efficient numerical computation. **Computing in science & engineering**, IEEE, v. 13, n. 2, p. 22–30, 2011. 101

WENGER, E.; BROZZOLI, C.; LINDENBERGER, U.; LöVDÉN, M. Expansion and renormalization of human brain structure during skill acquisition. **Trends in Cognitive Sciences**, v. 21, n. 12, p. 930–939, 2017. ISSN 1364-6613. 38

WICKRAMASINGHE, C. S.; AMARASINGHE, K.; MANIC, M. Deep self-organizing maps for unsupervised image classification. **IEEE Transactions on Industrial Informatics**, IEEE, v. 15, n. 11, p. 5837–5845, 2019. 57

WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. **Climate research**, v. 30, n. 1, p. 79–82, 2005. 54

XUE, J.-H.; PIZURICA, A.; PHILIPS, W.; KERRE, E.; WALLE, R. V. D.; LEMAHIEU, I. An integrated method of adaptive enhancement for unsupervised segmentation of mri brain images. **Pattern Recognition Letters**, Elsevier, v. 24, n. 15, p. 2549–2560, 2003. 53

XUE, Z.; ANTANI, S.; LONG, L. R.; DEMNER-FUSHMAN, D.; THOMA, G. R. Window classification of brain ct images in biomedical articles. In: AMERICAN MEDICAL INFORMATICS ASSOCIATION. **AMIA Annual Symposium Proceedings**. Illinois, USA: AMIA, 2012. v. 2012, p. 1023. 43

YANG, L.; ZHANG, Y.; CHEN, J.; ZHANG, S.; CHEN, D. Z. Suggestive annotation: A deep active learning framework for biomedical image segmentation. In: SPRINGER. **Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III 20**. Quebec City, QC, Canada: Springer, 2017. p. 399–407. 28

ZAHEER, R.; SHAZIYA, H. A study of the optimization algorithms in deep learning. In: IEEE. **2019 third international conference on inventive systems and control (ICISC)**. Coimbatore, India, 2019. p. 536–539. 78

ZENG, S.; ZHOU, Q. C.; ZHOU, J. W.; LI, M.; LONG, C.; PENG, Q. H. Volume of intracranial structures on three-dimensional ultrasound in fetuses with congenital heart disease. **Ultrasound in Obstetrics & Gynecology**, v. 46, n. 2, p. 174–181, 2015. 38

ZHANG, J.; ZHAO, X.; CHEN, Z.; LU, Z. A review of deep learning-based semantic segmentation for point cloud. **IEEE Access**, IEEE, v. 7, p. 179118–179133, 2019. 28

ZHANG, Z. Improved adam optimizer for deep neural networks. In: IEEE. **2018 IEEE/ACM 26th international symposium on quality of service (IWQoS)**. Montreal, Qubec, Canada, 2018. p. 1–2. 78

ZHANG, Z.; SABUNCU, M. Generalized cross entropy loss for training deep neural networks with noisy labels. **Advances in neural information processing systems**, v. 31, 2018. 54

ZHONG, N.-N.; WANG, H.-Q.; HUANG, X.-Y.; LI, Z.-Z.; CAO, L.-M.; HUO, F.-Y.; LIU, B.; BU, L.-L. Enhancing head and neck tumor management with artificial intelligence: Integration and perspectives. In: ELSEVIER. **Seminars in Cancer Biology**. Hubei Province, China, 2023. 94

ZHU, W.; HUANG, Y.; ZENG, L.; CHEN, X.; LIU, Y.; QIAN, Z.; DU, N.; FAN, W.; XIE, X. Anatomynet: deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy. **Medical physics**, Wiley Online Library, v. 46, n. 2, p. 576–589, 2019. 53

ZIMMERER, D.; ISENSEE, F.; PETERSEN, J.; KOHL, S.; MAIER-HEIN, K. Unsupervised anomaly localization using variational auto-encoders. In: SPRINGER. **Medical Image Computing and Computer Assisted**

**Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV 22.** Magdeburg, Germany, 2019. p. 289–297. [58](#)