

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ - UTFPR
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA**

**DIEGO VITORIANO FREITAS
ISABELLE NAYARA ROMPATO DA SILVA
MARCELO GUEVARA DE SÁ**

**SEGMENTAÇÃO SEMÂNTICA DO HIPOCAMPO
EM IMAGENS DE RESSONÂNCIA MAGNÉTICA ESTRUTURAL
UTILIZANDO REDE NEURAL CONVOLUCIONAL**

TRABALHO DE CONCLUSÃO DE CURSO

**CURITIBA
2022**

DIEGO VITORIANO FREITAS
ISABELLE NAYARA ROMPATO DA SILVA
MARCELO GUEVARA DE SÁ

**SEGMENTAÇÃO SEMÂNTICA DO HIPOCAMPO
EM IMAGENS DE RESSONÂNCIA MAGNÉTICA ESTRUTURAL
UTILIZANDO REDE NEURAL CONVOLUCIONAL**

**SEMANTIC SEGMENTATION OF THE HIPPOCAMPUS
IN STRUCTURAL MAGNETIC RESONANCE IMAGING
USING CONVOLUTIONAL NEURAL NETWORK**

Trabalho de conclusão de curso de graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia de Controle e Automação do curso de Engenharia de Controle e Automação da Universidade Tecnológica Federal do Paraná (UTFPR)

Orientadora: Profa. Dra. Mariana Antonia Aguiar Furucho

CURITIBA

2022



Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

DIEGO VITORIANO FREITAS
ISABELLE NAYARA ROMPATO DA SILVA
MARCELO GUEVARA DE SÁ

**SEGMENTAÇÃO SEMÂNTICA DO HIPOCAMPO
EM IMAGENS DE RESSONÂNCIA MAGNÉTICA ESTRUTURAL
UTILIZANDO REDE NEURAL CONVOLUCIONAL**

Trabalho de conclusão de curso de graduação
apresentado como requisito para obtenção do
título de Bacharel em Engenharia de Controle e
Automação do curso de Engenharia de
Controle e Automação da Universidade
Tecnológica Federal do Paraná (UTFPR)

Curitiba, 10 de Junho de 2022.

Profa. Dra. Mariana Antonia Aguiar Furucho
Universidade Tecnológica Federal do Paraná

Prof. Dr. Marcelo De Oliveira Rosa
Universidade Tecnológica Federal do Paraná

Prof. Me. Rogerio Akira Furucho
Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

**CURITIBA
2022**

AGRADECIMENTOS

Gostaríamos de agradecer à orientadora Prof^a. Dra. Mariana Furucho pela dedicação, paciência, disponibilidade e incentivo durante todo o desenvolvimento do projeto. O seu apoio foi fundamental para que conseguíssemos atingir um bom resultado. Nesse processo adquirimos conhecimentos valiosos que vão nos acompanhar por toda a vida profissional.

RESUMO

Esse trabalho possui o objetivo de desenvolver um algoritmo de inteligência artificial capaz de segmentar o hipocampo em imagens médicas do cérebro. O hipocampo é uma estrutura do cérebro ligado à obtenção de memórias, por isso as suas características estão ligadas à presença ou não da Doença de Alzheimer. A região de interesse pode ser identificada em imagens de ressonância magnética por profissionais especializados da área médica. Utilizando um banco de dados aberto com 263 arquivos com máscaras foi possível separar 6627 imagens que possuíam o hipocampo, desse total foram usadas 90% para ensinar uma rede neural convolucional a identificar e segmentar o hipocampo em cada uma das imagens. O resultado dessa CNN foi satisfatório com uma acurácia superior a 99%, mostrando ser muito promissor a utilização dessa segmentação para aplicações médicas no auxílio ao diagnóstico da DA. A principal dificuldade encontrada durante esse projeto foi a obtenção e organização da base de dados a ser usada: foram considerados diferentes tipos de imagens e mesmo a mais apropriada necessitou intervenção para processos como adição do *zero-padding* e binarização.

Palavras-chaves: doença de Alzheimer; inteligência artificial; segmentação; aplicações médicas; rede neural convolucional

ABSTRACT

The main purpose of this paper is to develop an artificial intelligence algorithm capable of segmenting the hippocampus in medical brain images. The hippocampus is a brain structure connected to the memories obtention, therefore its characteristics are connected to the presence or not of Alzheimer's Disease. The region of interest can be identified on Magnetic Resonance Imaging by specialized medical professionals. Using an open database with 263 files with masks was possible to separate 6627 images that had the hippocampus, of which 90% were used to teach a convolutional neural network to identify and segment the hippocampus in each of the images. The outcome of this CNN was satisfactory with an accuracy higher than 99%, showing that the use of this segmentation for medical applications to aid in the diagnosis of AD is very promising. The main difficulty encountered during this project was obtaining and organizing the database to be used, different types of images were considered and even the most appropriate required intervention for processes such as zero-padding and binarization.

Keywords: Alzheimer's disease; artificial intelligence; segmentation; medical applications; convolutional neural network

LISTA DE FIGURAS

Figura 1 - Fases da DA e seus respectivos bioindicadores	18
Figura 2 - Equipamento de RM.....	20
Figura 3 - Alinhamento dos prótons durante uma RM.....	21
Figura 4 - Diagrama da sequência do spin eco	22
Figura 5 - Exemplo de imagem de RM.....	22
Figura 6 - Header DICOM	24
Figura 7 - Esquema da imagem NIfTI.....	25
Figura 8 - Segmentação do hipocampo por detecção de borda	27
Figura 9 - Segmentação semântica do hipocampo a partir de uma imagem de RM	28
Figura 10 - Componentes da estrutura de um neurônio biológico	30
Figura 11 - Modelo simplificado de um neurônio com xm entradas.....	30
Figura 12 - (a) Função de limiar (b) Função linear por partes (c) Função sigmóide	32
Figura 13 - Rede MLP com duas camadas escondidas	33
Figura 14 - Aplicação do comportamento do backpropagation.....	35
Figura 15 - Padding de tamanho 2.....	38
Figura 16 - Camada convolucional parcial utilizando um filtro 3x3.....	39
Figura 17 - Camada de pooling utilizando um kernel 2x2.....	40
Figura 18 - Operação de <i>flattening</i> dado um mapa de características	41
Figura 19 - Camada totalmente conectada	42
Figura 20 - Arquitetura simplificada de uma Rede Neural Convolucional ...	42
Figura 21 - Arquitetura U-Net	43
Figura 22 - Fluxograma do procedimento.....	47
Figura 23 - Imagem NIfTI tridimensional e cortes em três orientações.....	48
Figura 24 - Corte da primeira amostra do banco de imagens	50
Figura 25 - Interface do Colab.....	52
Figura 26 - Comparação entre a máscara real e a segmentação para treinamento com 15 épocas.....	57
Figura 27 - Comparação entre a máscara real e a segmentação para treinamento com 10 épocas.....	60
Figura 28 - Segmentação de imagens da base de testes para treinamento com 15 épocas.....	61

Figura 29 - Segmentação de imagens da base de testes para treinamento com 10 épocas.....62

LISTA DE GRÁFICOS

Gráfico 1 - Valores dos pixels para cada imagem de máscara	50
Gráfico 2 - Acurácia para treinamento com 15 épocas	56
Gráfico 3 - Perda para treinamento com 15 épocas	57
Gráfico 4 - Acurácia para treinamento com 10 épocas	59
Gráfico 5 - Perda para treinamento com 10 épocas	59

LISTA DE ABREVIATURAS E SIGLAS

ACR	<i>American College of Radiology</i>
CNN	Rede Neural Convolucional (<i>Convolutional Neural Network</i>)
DA	Doença de Alzheimer
DICOM	<i>Digital Imaging and Communications in Medicine</i>
MCI	Declínio Cognitivo Leve (<i>Mild Cognitive Impairment</i>)
MLP	Redes <i>Perceptron</i> Multicamadas (<i>Multilayer Perceptron</i>)
MSD	<i>Medical Segmentation Decathlon</i>
NEMA	<i>National Electrical Manufacturers Association</i>
NIfTI	<i>Neuroimaging Informatics Technology Initiative</i>
RF	Radiofrequência
RM	Ressonância Magnética Estrutural
RNA	Redes Neurais Artificiais
RNB	Redes Neurais Biológicas
TC	Tomografia Computadorizada

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Tema.....	12
1.1.1 Delimitação do tema	13
1.2 Problemas E Premissas.....	13
1.3 Objetivos.....	14
1.3.1 Objetivo geral.....	14
1.3.2 Objetivos específicos.....	14
1.4 Justificativa	15
2 REFERENCIAL TEÓRICO	16
2.1 Doença De Alzheimer.....	16
2.2.1 Diagnóstico.....	17
2.2 Ressonância Magnética.....	19
2.3 Formatos De Imagens.....	23
2.3.1 DICOM.....	23
2.3.2 NIfTI.....	25
2.3.3 JPG.....	26
2.4 Segmentação De Imagens.....	26
2.5 Redes Neurais	28
2.5.1 Modelo De Um Neurônio	29
2.5.2 Funções De Ativação.....	31
2.5.3 Redes <i>Perceptron</i>	32
2.5.4 Redes <i>Perceptron</i> Multicamadas.....	33
2.5.5 Backpropagation.....	33
2.6 Redes Neurais Convolucionais.....	35
2.6.1 Camadas Convolucionais	37
2.6.2 Redes U-Net.....	43
2.7 Métricas De Validação	44

3 MATERIAIS E MÉTODOS.....	46
3.1 Metodologia	46
3.1.1 Aquisição e preparação das imagens	47
3.1.2 Definição do ambiente de desenvolvimento	51
3.1.3 Desenvolvimento e treinamento do algoritmo.....	53
3.1.4 Testes e validação.....	54
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS	55
4.1 Treinamento E Validação.....	55
4.1 Predição	60
5 CONCLUSÃO.....	63
REFERÊNCIAS	65

1 INTRODUÇÃO

1.1 Tema

A Doença de Alzheimer (DA) é uma doença neurodegenerativa multifacetada conhecida como a principal causa de demência em pessoas de idade avançada. Alguns tratamentos atuais ajudam a retardar e amenizar os sintomas, mas a falta de conhecimento sobre o processo patogênico (como e porque ocorre o desenvolvimento da doença) dificulta a obtenção de um tratamento definitivo (KUMAR; SINGH; EKAVALI, 2015).

O quadro patológico da DA é representado por um processo lento de neuroinflamação, acúmulo de placas beta-amiloide e emaranhados neurofibrilares de proteína tau hiperfosforilada ligada a uma degeneração neural no cérebro que leva à morte neural, principalmente nas áreas do neocórtex e do hipocampo, o que prejudica muitas funções nos portadores como prejuízo de memória e déficits de cognição e da capacidade de raciocínio. Observa-se que o surgimento da demência ocorre por volta de 20 a 30 anos antes da apresentação dos sintomas, logo um diagnóstico precoce pode ser possível quando ainda há uma neurodegeneração inicial e, conseqüentemente, uma maior chance de um sucesso terapêutico (SWARBRICK et al., 2019).

Estima-se que, de 2011 até o ano 2040, oitenta milhões de pessoas sejam portadoras desta enfermidade, principalmente nos países subdesenvolvidos que aumentaram a expectativa de vida, representando um aumento de quase 300% em regiões como América Latina e Índia (HAMPEL et al. 2011). No Brasil, por sua vez, um estudo feito com 1656 indivíduos, acima dos 65 anos, realizado pela Escola de Medicina de Catanduva encontrou uma taxa de prevalência de 7,1% para os quadros demenciais, sendo o Alzheimer responsável por 55% desses diagnósticos (CLEMENTE; PINTO, 2017).

Caracterizada como uma doença crônica, degenerativa e fatal, a doença de Alzheimer é responsável por 60% - 70% dos diagnósticos de demência em todo o mundo e estima-se que afete 106,2 milhões de pessoas ou 1 a cada 85 pessoas até o ano de 2050 (AGUIAR-FURUCHO; PELÁEZ, 2020). Estima-se também que, em 2060, 13,85 milhões de pessoas acima dos 65 anos terão DA nos Estados Unidos,

enquanto 21,55 milhões terão um declínio cognitivo leve (*Mild Cognitive Impairment*, MCI) (RAJAN et al., 2021).

Até o momento não existe cura para a DA, sendo o tratamento direcionado para melhorar a qualidade de vida e aumentar a sobrevida do paciente. Essa ineficácia terapêutica se deve ao grau de complexidade da etiologia da enfermidade e ao diagnóstico tardio dos portadores (MOLINUEVO et al., 2018). Atualmente, as técnicas diagnósticas consistem em testes cognitivos, os quais não identificam os estágios iniciais da doença por serem geralmente realizados tardiamente; e nos biomarcadores de fluidos orgânicos e de imagem, que refletem os diferentes tipos de fisiopatologia do cérebro (SWARBRICK et al., 2019).

Por este motivo, diversas técnicas de IA são estudadas como forma de auxílio no processo de segmentação e análise da imagem. Entre elas as principais são o *Machine Learning* e o *Deep Learning*. Na utilização dessas técnicas a segmentação do hipocampo em imagens de ressonância magnética possuem desempenho melhor que as realizadas em imagens de tomografia computadorizada (HÄNSCH et al., 2020).

1.1.1 Delimitação do tema

Neste trabalho será desenvolvida uma ferramenta de *Deep Learning*, no modelo da Rede Neural Convolucional (*Convolutional Neural Network* - CNN), capaz de segmentar o hipocampo em imagens de ressonância magnética estrutural visando contribuir para o diagnóstico da doença de Alzheimer.

1.2 Problemas E Premissas

A memória, definida como habilidade de reter e consultar informações é, além de mecanismo biológico de sobrevivência, um marcador de quem nós somos. Deste modo, as nossas memórias ajudam a moldar quem somos e podem ser revisitadas para nos guiar em pensamentos e tomadas de decisão.

Baseando-se na duração das memórias podemos classificá-las como: memórias de curto prazo, sendo informações retidas por poucos segundos a alguns minutos; ou memórias de longo prazo, que permanecem por um longo período de

tempo, às vezes perdurando pela vida toda. Além disso, as memórias também podem ser distinguidas pela sua manifestação. Enquanto memórias explícitas são declarativas e fazem referência a fatos, pessoas e lugares, as memórias implícitas dizem respeito a processos, habilidades ou atividades (BISAZ; TRAVAGLIA; ALBERINI, 2014).

Desse modo, as memórias implícitas geralmente são recuperadas mediante demanda, ou seja, quando há necessidade de o indivíduo executar uma tarefa ou um procedimento aprendido anteriormente. Por outro lado, as memórias explícitas podem ser evocadas espontaneamente ou também sob alguma demanda específica, como a necessidade de recordar um evento ocorrido no passado. O hipocampo se encontra envolvido nos processos de armazenamento e recuperação das memórias explícitas como foi evidenciado em um experimento realizado por Scoville e Milner (1957).

Devido a essa estreita relação do hipocampo com a formação e recuperação das memórias explícitas, a identificação de sua atrofia é usada como diagnóstico e prognóstico da DA. Pacientes com MCI apresentam uma atrofia de 10-15% da estrutura, podendo ir de 15-30% no início da DA e até 50% em casos mais avançados (DHIKAV; ANAND, 2012).

1.3 Objetivos

1.3.1 Objetivo geral

Desenvolver uma ferramenta computacional baseada em técnicas de *Deep Learning* capaz de identificar o hipocampo em imagens de ressonância magnética.

1.3.2 Objetivos específicos

Para atingir o objetivo principal proposto, necessita-se cumprir os seguintes objetivos específicos:

- Estudar as principais características neuromorfológicas da doença de Alzheimer relacionadas com o hipocampo;
- Estudar a teoria e aplicações da Deep Learning, particularmente das Redes Neurais Convolucionais;

- Pesquisar aplicações da CNN para auxílio diagnóstico da DA;
- Estudar as bibliotecas de *Deep Learning* para *Python*;
- Pesquisar *datasets* abertos e gratuitos de imagens de Ressonância Magnética Estrutural (RM) para doença de Alzheimer da região do hipocampo;
- Elaborar a lógica para o diagnóstico de imagens de ressonância magnética, incluindo segmentação do hipocampo;
- Treinar o algoritmo com imagens que possuem o diagnóstico conhecido;
- Testar o algoritmo em imagens diferentes das utilizadas para o treinamento;
- Analisar os resultados obtidos e aperfeiçoar a lógica de programação;
- Obter o grau de precisão do resultado da técnica utilizada.

1.4 Justificativa

O diagnóstico precoce de uma doença permite obter uma vantagem vital na fase de tratamento, por representar a possibilidade de retardar ou até mesmo impedir seu avanço. Particularmente, para a doença de Alzheimer, o diagnóstico precoce poderia ser feito com até 30 anos antes da manifestação dos sintomas clínicos da doença (ASHRAF et al., 2015).

Exames clínicos que permitam analisar anormalidades em biomarcadores relevantes para a doença, como a quantidade de placas beta-amilóide e emaranhados neurofibrilares (ASHRAF et al., 2015), demandam técnicas de imageamento como a tomografia computadorizada e a ressonância magnética e costumam ser muito caros quando comparados com exames mais tradicionais. Além disso, há uma grande dificuldade na análise dessas imagens. As técnicas de IA permitem acelerar e facilitar este processo, aumentando o índice de acerto e agilizando os diagnósticos, o que contribui para a melhora da qualidade de vida dos pacientes que possuem a doença (AGUIAR-FURUCHO, 2017).

Além disso, as técnicas de Inteligência Artificial estão sendo utilizadas cada vez mais em diversos setores da sociedade. Por este motivo, este trabalho também visa agregar ao conhecimento e entendimento dos autores sobre as técnicas de IA que serão estudadas, permitindo aplicações futuras nas mais diversas áreas.

2 REFERENCIAL TEÓRICO

Para o desenvolvimento deste trabalho será realizado um estudo teórico sobre as principais características da DA, incluindo seus principais biomarcadores, especialmente os relacionados com alterações neuromorfológicas da região do hipocampo. Também será estudada a metodologia da *Deep Learning* para auxiliar o diagnóstico baseado em imagens médicas.

2.1 Doença De Alzheimer

O declínio cognitivo é um aspecto comum e temido do envelhecimento. O comprometimento cognitivo leve é considerado como o estágio inicial de demência, ou seja, o estágio em que aparecem os primeiros sintomas de um declínio cognitivo, não caracterizado como grave o suficiente para demandar ajuda nas atividades cotidianas.

Estudos mostram que muitas pessoas com sintomas de demência, especialmente aquelas em faixas etárias mais elevadas, têm anormalidades cerebrais associadas a mais de uma causa de demência (ALZHEIMER'S ASSOCIATION, 2017).

O diagnóstico de MCI ocorre, predominantemente, em pacientes a partir de 65 anos, e o risco aumenta proporcionalmente à idade. Indivíduos com esse diagnóstico possuem um risco maior de desenvolver algum tipo de demência, como por exemplo, a doença de Alzheimer.

A doença de Alzheimer foi descoberta em 1906, mas cerca de 70 anos se passaram até que ela fosse reconhecida como uma causa comum de demência e uma das principais causas de morte. Apesar de muito ter sido descoberto através de inúmeras pesquisas, muito ainda se tem a desvendar sobre as exatas mudanças biológicas que causam a doença, o porquê da doença progredir mais em alguns indivíduos do que em outros e como prevenir, retardar ou interromper o seu avanço (ALZHEIMER'S ASSOCIATION, 2017).

Um dos primeiros sintomas a serem reportados pelos pacientes, é a dificuldade em se lembrar de fatos, eventos e nomes de pessoas próximas. Contudo, apenas essas queixas não são suficientes para afirmar um diagnóstico de DA. É

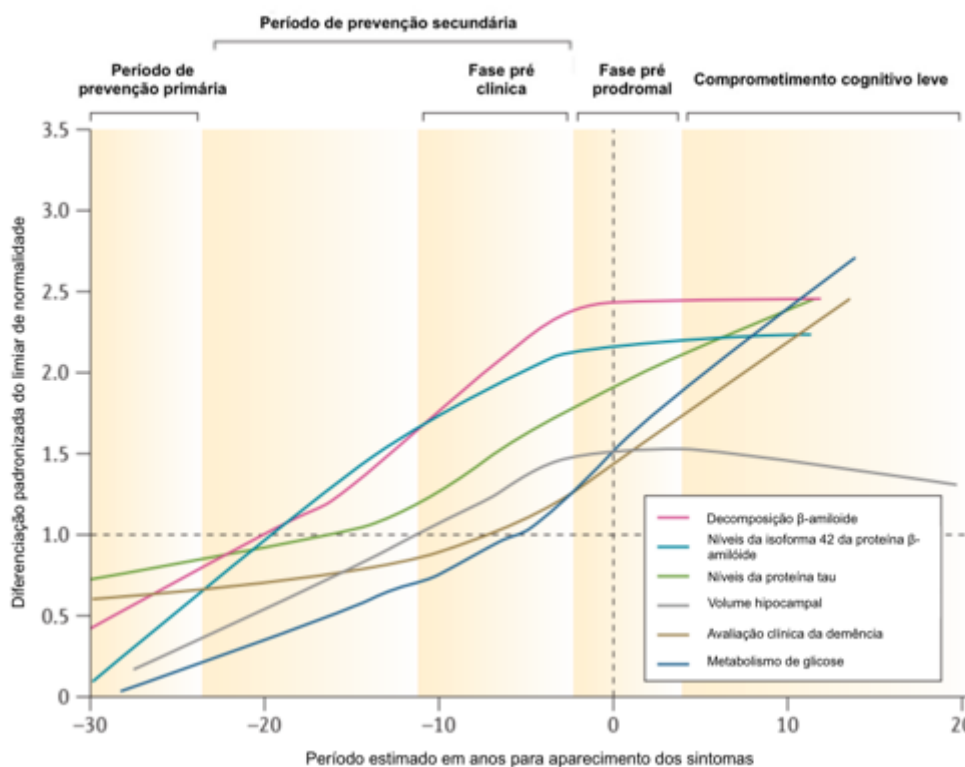
necessária a presença de outros problemas simultâneos, como confusão mental, apatia, dificuldade de comunicação e dificuldade em executar tarefas do cotidiano (ALZHEIMER'S ASSOCIATION, 2017).

2.2.1 Diagnóstico

Em termos patológicos, são características comuns da doença de Alzheimer o acúmulo das placas de proteína beta-amilóide ($A\beta$) na parte externa dos neurônios e também os emaranhados neurofibrilares formados pela proteína *tau*. As placas beta-amilóides interferem nas sinapses e contribuem para a morte celular, enquanto que os emaranhados neurofibrilares bloqueiam o transporte de neurotransmissores para a fenda sináptica e outras moléculas essenciais dentro dos neurônios. O cérebro de pessoas com a doença de Alzheimer mostra inflamação e atrofia da região do hipocampo, devido a perda de células e neurônios (ALZHEIMER'S ASSOCIATION, 2017).

De acordo com Masters et al. (2015), o comportamento desses biomarcadores pode ser verificado na Figura 1, na qual destacam-se as fases pré-clínica, prodromal e o comprometimento cognitivo leve. As alterações nos biomarcadores podem ser vistas em até 20 anos antes dos primeiros sintomas da doença. No início há um aumento na deposição das proteínas β -amilóide. Após isso os níveis da proteína tau no líquido cefalorraquidiano começam a aumentar em torno de 15 anos antes da fase de MCI, levando a um aumento do volume hipocampal que decai conforme a progressão da doença.

Figura 1 - Fases da DA e seus respectivos bioindicadores



Fonte: Adaptado de Masters et al. (2015)

Durante a fase prodromal acontecem as manifestações dos sintomas cognitivos da DA como demência e perda de memória, mas ainda é insuficiente para o diagnóstico da doença. Em termos dos biomarcadores eles já estão em níveis bem elevados indicando o início do comprometimento cognitivo.

No período de MCI a degradação das habilidades cognitivas passa a se diferenciar do envelhecimento comum do ser humano podendo afetar não apenas a memória, mas também outras funções cognitivas como a comunicação. Esse período pode ser considerado já como uma fase pré-clínica da DA (AGUIAR-FURUCHO, 2017).

Em termos de diagnóstico, não existe um teste ou um exame único capaz de fornecer um diagnóstico preciso da doença de Alzheimer. Portanto, médicos e especialistas, como neurologistas e geriatras, fazem uso de uma variedade de abordagens para definir um diagnóstico, que leva em consideração diversos fatores como: histórico médico familiar do paciente; testes cognitivos; testes físicos; exames neurológicos; exames de sangue; exames de imagens da área cerebral (ALZHEIMER'S ASSOCIATION, 2017).

Com o intuito de melhorar a acurácia diagnóstica da doença de Alzheimer, diversos estudos são realizados com o propósito de identificar um biomarcador (fator biológico que pode ser medido para indicar a presença ou ausência de uma determinada doença, ou o risco de desenvolvê-la) para a DA. Entre diversos fatores já estudados, os que mais se destacam como possíveis biomarcadores para a DA são a quantidade de placas beta-amilóides no cérebro, conforme é possível observar em exames de imagem por tomografia e de líquido cefalorraquidiano, a quantidade de emaranhados neurofibrilares, além de alterações fisiológicas na região do hipocampo. Diante disso, encontrar um meio de tornar o diagnóstico mais rápido e assertivo torna-se ideal para garantir um tratamento precoce e evitar o avanço da doença (ALZHEIMER'S ASSOCIATION, 2017).

Para o diagnóstico por imagens do hipocampo são utilizados os exames de RM e tomografia computadorizada (TC). Após a aquisição da imagem é feita a segmentação do hipocampo, que é a delimitação de seus contornos na imagem obtida, e em seguida a área delimitada é analisada buscando identificar possíveis anomalias fisiológicas, tanto em forma quanto em volume. Este é um processo que deve ser realizado por profissionais qualificados e costuma ser lento e custoso.

2.2 Ressonância Magnética

Enquanto o Raio X é amplamente usado em regiões rígidas, como a estrutura óssea do corpo humano, a Tomografia Computadorizada e a Ressonância Magnética são técnicas mais modernas que permitem o diagnóstico de tecidos por imagem.

A TC fornece informações precisas da anatomia e volume de órgãos, porém não apresenta grande contraste em tecidos moles. Em contraponto, a RM apresenta um ótimo contraste em tecidos moles (DECAZES et al., 2021), facilitando a identificação de regiões do cérebro. Por isso, serão usadas imagens referentes a esse último exame para o propósito deste trabalho. Na Figura 2 é possível ver o equipamento que realiza a RM.

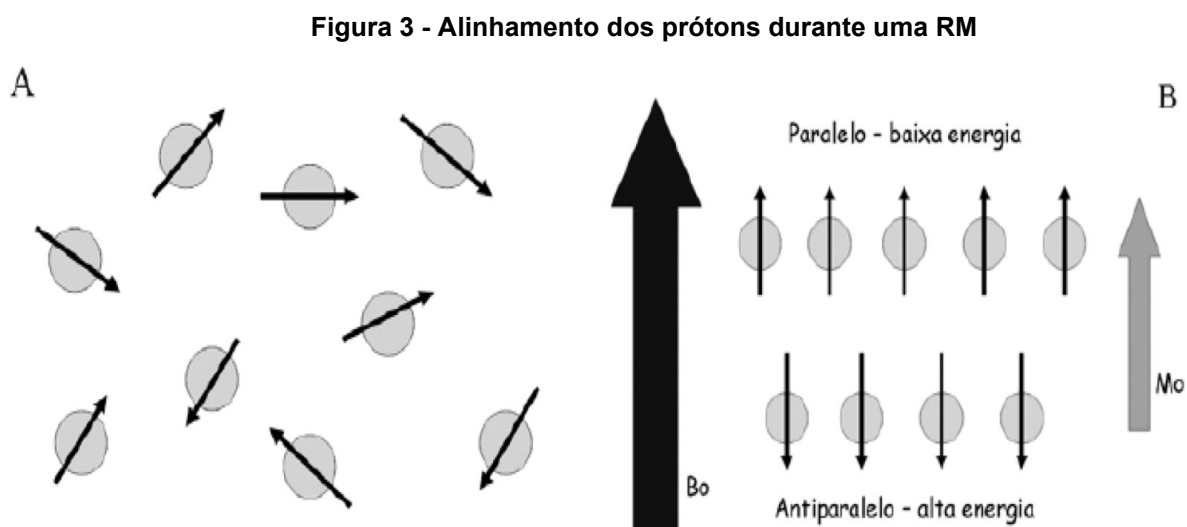
Figura 2 - Equipamento de RM

Fonte: Boletim UFMG (2012)

A RM é, em linhas gerais, a análise das ondas de rádio emitidas por células que foram excitadas por um campo magnético (HAGE; IWASAKI, 2009). Para entender um pouco melhor como a imagem resultante é formada, é preciso analisar como esse exame é realizado. Em um nível atômico, o núcleo de todo átomo possui um momento angular próprio, chamado de spin (PYKETT et al., 1982). Entretanto, dois prótons, ou dois nêutrons, se alinham anulando esse momento, sendo que só existirá um momento resultante caso exista um número ímpar de prótons ou de nêutrons. Além do spin, os prótons possuem a propriedade de um dipolo magnético, ou seja, podem ser considerados como um pequeníssimo ímã. Esse campo magnético resultante pode ser percebido por aparelhos, e assim formar a imagem a partir da RM (MAZZOLA, 2015).

O átomo de hidrogênio, além de ser o mais abundante no corpo humano, é o que possui elevado momento magnético, logo é o que pode ser mais bem detectado em uma RM (MAZZOLA, 2015). Em situações normais todos os dipolos magnéticos dos átomos estão aleatoriamente distribuídos, causando um campo magnético resultante nulo, por isso é necessário em um primeiro momento submeter o paciente a um campo magnético de forma a alinhar todos esses dipolos magnéticos (PYKETT et al., 1982).

Na Figura 3 percebe-se a aleatoriedade desses campos magnéticos, bem como o seu alinhamento ao serem expostos por um campo magnético.



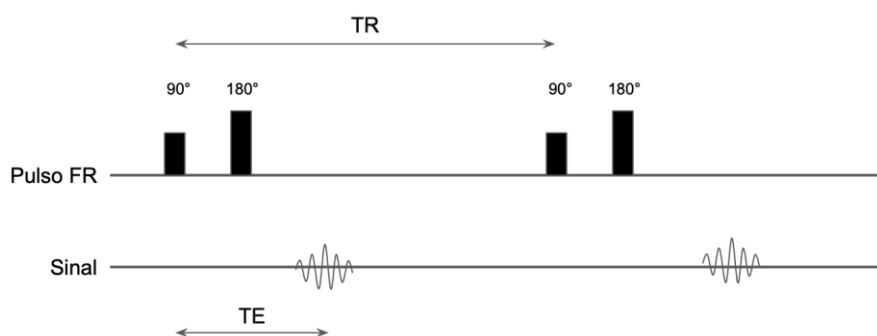
Nota: (A) Prótons distribuídos aleatoriamente, com um campo magnético resultante nulo. (B) prótons alinhados com campo magnético B_0 , resultando num campo magnético μ_0 de menor intensidade.

Fonte: Hage e Iwasaki (2009)

Frente a esse campo externo a intensidade de μ_0 é extremamente baixa tornando muito difícil sua detecção por aparelhos (HAGE; IWASAKI, 2009). Por isso, o tecido precisa ser submetido a pulsos de radiofrequência (RF) com o objetivo de deslocar um campo magnético em 90 graus.

Os pulsos de RF são emitidos com um intervalo chamado tempo de repetição. Após isso, será visto um eco do primeiro sinal aplicado, resultado do realinhamento dos momentos magnéticos (SILVA, 2011). O tempo entre esses sinais é chamado de tempo de eco, como pode ser visto na Figura 4.

Figura 4 - Diagrama da sequência do spin eco



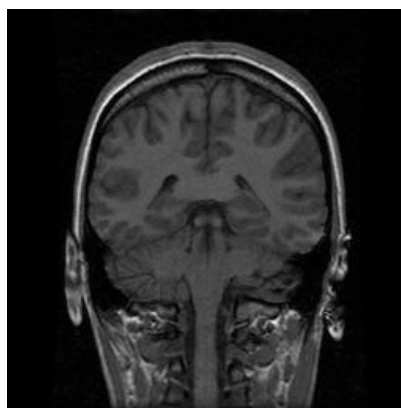
Nota: Sendo TR o tempo de repetição e TE o tempo de Eco.

Fonte: Adaptado de MAZOLLA (2015).

São por esses sinais de eco que se pode analisar os átomos do tecido a ser estudado, pois nesse processo chamado relaxação são emitidos sinais magnéticos sensíveis a uma bobina receptora. Esse processo pode ser modelado por duas constantes de tempo: T1 e T2. A primeira diz respeito a perda de energia dos spins com o sistema e átomos vizinhos (relaxação spin-rede), e o outro a perda resultante quando os spins interagem entre si (relaxação spin-spin) (SILVA, 2011).

A combinação entre tempo de repetição (TR) e tempo de eco (TE) é usada para contrastar tecidos com diferentes T1 e T2, produzindo imagens com áreas mais claras e áreas mais escuras. No caso de estruturas ósseas ou o ar, como existe uma baixa concentração de átomos de hidrogênio, não existe uma emissão de sinal relevante e serão mostradas como áreas pretas na imagem (HAGE; IWASAKI, 2009).

Figura 5 - Exemplo de imagem de RM



Fonte: DINCER (2021)

Uma imagem formada por RM permite ver as diferentes concentrações de prótons no tecido. Regiões com uma grande concentração de prótons apresentam uma grande intensidade na imagem e por isso ficam brancas, elas são chamadas de hipertensas. Do mesmo modo, a região vai escurecendo conforme diminui a concentração de prótons no tecido, até chegar em áreas pretas, chamadas de hipotensas. Esse resultado pode ser observado na Figura 5.

2.3 Formatos De Imagens

Existem diversas possibilidades para o armazenamento das imagens médicas. A seguir serão descritos os formatos DICOM e NifTI, muito utilizados na área médica, e o formato JPG, utilizado no tratamento das imagens médicas realizado neste trabalho.

2.3.1 DICOM

Nos últimos anos, houve um grande aumento no uso de imagens digitais em diagnósticos médicos. Com isso veio a necessidade de normalizar a comunicação e transmissão de dados referente a diagnósticos, procedimentos terapêuticos ou qualquer dado referente ao paciente ou ao exame (MILDENBERGER; EICHELBERG; MARTIN, 2001).

Por isso, em 1983, uma comissão formada pela *American College of Radiology* (ACR) e a *National Electrical Manufacturers Association* (NEMA) tentou criar um padrão que permitisse a transmissão de dados médicos sem o viés de qualquer fornecedor. Em 1985 ocorreu a publicação da primeira versão da ACR-NEMA. Conforme ocorreram atualizações, em 1993 aconteceu a publicação da terceira versão, agora com o nome DICOM 3.0 ou simplesmente DICOM - acrônimo para *Digital Imaging and Communications in Medicine* (GRAHAM et al., 2005).

Além de um protocolo de comunicação, DICOM é um formato digital de imagem (.dcm) que se diferencia das demais por trabalhar com conjuntos de dados, ou seja, um arquivo DICOM consiste em um *header* (cabeçalho) conjunto a uma imagem (VARMA, 2012). Esse *header* conta com informações do paciente,

parâmetros para aquisição daquela imagem, dimensões da imagem, informações do procedimento, entre outros dados (GRAHAM et al., 2005).

Figura 6 - Header DICOM

Cho	Group	Elem	Value	Re	Name	Cho	Group	Elem	Value	Re	Name
0002	0002	UI	1.2.840.10008.5.1.4.1.1.1		MediaStorageSOPClassUID	0018	1150	IS	16		ExposureTime
0002	0003	UI	1.3.46.670589.26.992140021400		MediaStorageSOPInstanceUID	0018	1152	IS	15		Exposure
0002	0010	UI	1.2.840.10008.1.2		TransferSyntaxUID	0018	115E	DS	11,643		ImageAreaDoseProduct
0002	0012	UI	1.2.276.0.39.0.3.4.1.1		ImplementationClassUID	0018	1180	SH	0mmAl		FilterType
0002	0013	SH	'_20000505		ImplementationVersionName	0018	1184	DS	0,143		ImagerPixelSpacing
0008	0005	CS	ISO_IR 100		SpecificCharacterSet	0018	1186	CS	IN		Grid
0008	0008	CS	DERIVEDPRIMARY		ImageType	0018	1170	IS	50		GeneratorPower
0008	0016	UI	1.2.840.10008.5.1.4.1.1.1		SOPClassUID	0018	1180	DS	2		FocalSpots
0008	0018	UI	1.3.46.670589.26.992140021400		SOPInstanceUID	0018	1200	DA	09.05.2001		DateOfLastCalibration
0008	0020	DA	19.05.2001		StudyDate	0018	1201	TM	15:58:53		TimeOfLastCalibration
0008	0021	DA	19.05.2001		SeriesDate	0018	1260	SH	Flat-Panel 43x43		PlateType
0008	0023	DA	19.05.2001		ImageDate	0018	1400	LO	P=fullField.s: CD=1.20 G=2.50		AcquisitionDeviceProcessingC
0008	0030	TM	04:43:16		StudyTime	0018	5020	LO	8000,16460,0,7673,8290,14852		ProcessingFunction
0008	0031	TM	04:42:51		SeriesTime	0018	5021	LO	boneS_12D12G25C11_01		PostprocessingFunction
0008	0033	TM	04:42:51		ImageTime	0018	5101	CS	AP		ViewPosition
0008	0050	SH	010107119101		AccessionNumber	0018	6000	DS	630		Sensitivity
0008	0060	CS	CR		Modality	0019	0019	LO	DICD TO PCR 1.1		PrivateCreator
0008	0070	LO	Medical Systems		Manufacturer	0019	1923	??	UNKNOWN		Unknown Tag & Data
0008	0080	LO	Uniklinik Mainz		InstitutionName	0019	1924	??	UNKNOWN		Unknown Tag & Data
0008	0090	PN	AC AMB		ReferringPhysiciansName	0019	1925	??	UNKNOWN		Unknown Tag & Data
0008	1010	SH	Digital Diagnost		StationName	0019	1926	??	UNKNOWN		Unknown Tag & Data
0008	1030	LO	Abdomen		StudyDescription	0019	1927	??	UNKNOWN		Unknown Tag & Data
0008	1040	LO	Diagnostische Radiologie		InstitutionalDepartmentName	0019	1970	??	UNKNOWN		Unknown Tag & Data
0008	1090	LO	digital DIAGNOST		ManufacturersModelName	0019	1971	??	UNKNOWN		Unknown Tag & Data
0010	0010	PN	B Z		PatientsName	0020	00DD	UI	1.2.276.0.38.1.1.5891.20010519		StudyInstanceUID
0010	0020	LO	77060		PatientID	0020	000E	UI	1.3.46.670589.26.992140021400		SeriesInstanceUID
0010	0030	DA	22.02.19		PatientsBirthDate	0020	0010	SH	9731		StudyID
0010	0040	CS	M		PatientsSex	0020	0011	IS	9399		SeriesNumber
0018	0015	CS	ABDOMEN		BodyPartExamined	0020	0012	IS	9399		AcquisitionNumber
0018	0060	DS	81		kVp	0020	0013	IS	0		InstanceNumber

Fonte: MILDENBERGER et al., (2001)

Como é possível observar na Figura 6, cada informação em um arquivo DICOM é fortemente definido para não permitir ambiguidade. Além disso, os parâmetros são divididos em grupos dentro do cabeçalho, por exemplo o grupo 0010 contém informações pessoais do paciente, sendo o elemento 0010-0010 o nome do paciente e 0010-0030 a data de nascimento (VARMA, 2012). O grupo 0018 contém informações referentes a aquisição dessa imagem, como 0018-0060 a tensão usada em um exame de raio X e em 0018-1000 o número serial do dispositivo (MILDENBERGER et al., 2001). A imagem em si é armazenada no grupo 0028.

Essa riqueza de informação permite, por exemplo, registrar a sequência de pulsos de RF usados para capturar uma imagem de RM, possibilitando que um profissional recrie aquele exame para recuperar a mesma qualidade (VARMA, 2012).

Dada a quantidade de informação presente nesse tipo de arquivo é normal que eles ocupem um grande espaço de memória, como 35 megabytes para uma imagem de tomografia computadorizada do cérebro (GRAHAM et al., 2005). Por isso, esses arquivos sofrem compressão para facilitar o seu armazenamento e transferência. O processo de compressão pode ocorrer sem perdas, que é mais lento e computacionalmente trabalhoso, mas permite recuperar a imagem com a exata

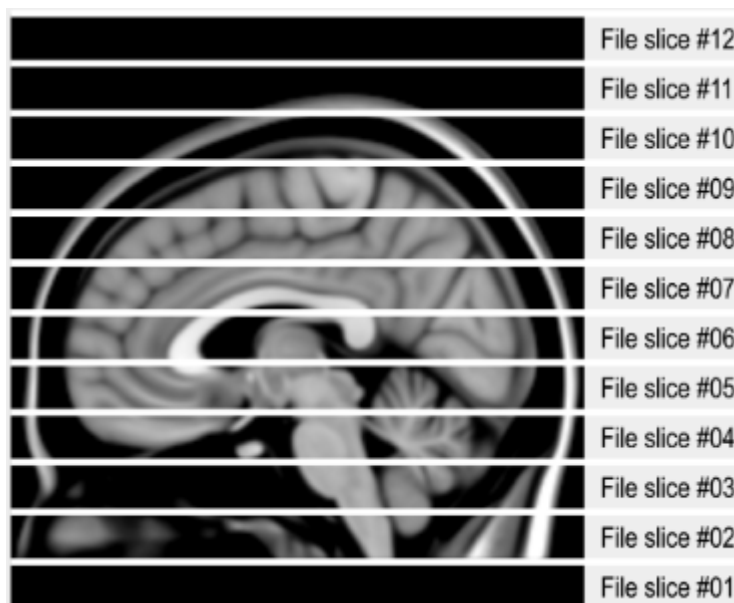
qualidade, ou em um processo chamado *lossy*, que permite a eliminação de dados redundantes sem comprometer a qualidade da imagem. Deve-se levar em consideração que uma compressão excessiva do arquivo pode resultar em perdas de qualidade (GRAHAM et al., 2005).

2.3.2 NIfTI

O formato NIfTI - acrônimo para *Neuroimaging Informatics Technology Initiative* - foi desenvolvido no início dos anos 2000 com a proposta de resolver alguns problemas que ocorriam em formatos mais antigos, principalmente a falta de informações associadas à imagem. Assim como o formato DICOM o NIfTI possui um *header*.

A diferença é que, enquanto o DICOM apresenta apenas um corte da RM, o NIfTI pode conter diversas camadas, ou seja, é uma representação inteira de uma parte do corpo humano, que contém todas as camadas obtidas pelo exame médico (PYCAD, 2021), conforme exposto na Figura 7.

Figura 7 - Esquema da imagem NIfTI



Fonte: Brainder (2012)

Essas imagens podem ser obtidas em um único arquivo (.nii) ou em um par com imagem e *header* separados (.hdr e .img). Em ambos os casos é possível a conversão em DICOM e vice-versa.

2.3.3 JPG

O formato de imagem JPEG, também conhecido como JPG, é o formato de imagem mais utilizado no ambiente digital. Por possuir uma taxa de compressão ajustável, é possível encontrar um equilíbrio entre qualidade e tamanho de imagem. Ou seja, para casos em que a taxa de compressão for mínima, o arquivo mesmo assim não ocupará um espaço grande de armazenamento.

Assim como para formato NIfTI, mencionado anteriormente, também é possível converter imagens DICOM para um formato mais familiar, como o JPG.

2.4 Segmentação De Imagens

O processo de segmentação consiste em dividir uma imagem em regiões ou objetos distintos. Essa divisão é feita de acordo com as características do objeto que determinam um conjunto de regras para o agrupamento dos *pixels*, como cor, brilho, contraste, textura e proximidade (MARENGONI; STRINGHINI, 2010). Essa técnica é utilizada em diversas aplicações como direção automatizada, vigilância por vídeo, visão de máquina e imagens médicas.

Uma das etapas fundamentais para o auxílio diagnóstico da DA é a identificação precisa do hipocampo e de eventuais anomalias nas imagens obtidas pelos exames. Esse processo costuma ser demorado e exige profissionais qualificados e experientes, o que eleva o custo do diagnóstico, e mesmo assim podem ocorrer imprecisões, especialmente quando é feita a análise de muitas imagens (ATALOGLOU et al., 2019).

A segmentação permite realizar de maneira rápida e automática a primeira etapa do diagnóstico, que é a identificação do hipocampo nas imagens médicas. Além disso, ela reduz a área da imagem que será analisada pelo algoritmo separando a região de importância para a análise e ajuda a extrair informações importantes como forma, volume, posição relativa e anormalidades (ANWAR et al., 2018). A segmentação do hipocampo em imagens de RM é uma linha de pesquisa em constante avanço devido à ligação entre a morfologia da estrutura e a doença (HÄNSCH et al., 2020).

As regras utilizadas e o nível de detalhamento da segmentação dependem do objetivo do procedimento e das propriedades das imagens que serão analisadas, como sua resolução, ruídos e o tamanho e posicionamento do objeto de estudo dentro da imagem. No caso de imagens de baixa resolução e excesso de ruídos é recomendado que a imagem seja previamente tratada, o que melhora a precisão do procedimento.

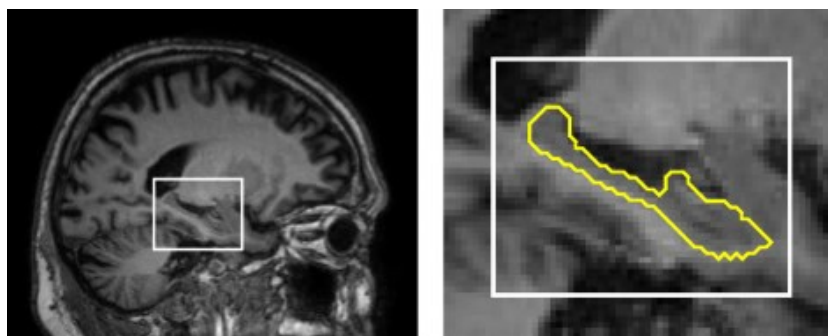
Algumas das técnicas de segmentação mais comuns são:

Segmentação por corte: os *pixels* são agrupados pelos seus valores de intensidade e um ponto de corte é definido. Pontos com intensidade menor que o valor de corte são agrupados em uma região e pontos com intensidade maior são agrupados em outra região criando uma imagem binária. Também é possível definir mais de duas regiões utilizando múltiplos pontos de corte.

Segmentação por detecção de borda: a detecção por borda se baseia na variação abrupta no nível de intensidade dos *pixels* que frequentemente ocorre nessas regiões. O algoritmo agrupa *pixels* próximos onde ocorre esse tipo de variação definindo assim uma região.

Para a identificação da variação de intensidade são utilizados operadores matemáticos baseados em derivadas de primeira e segunda ordem, além de operadores auxiliares de verificação que reduzem os pontos falsos (MARENGONI; STRINGHINI, 2021).

Figura 8 - Segmentação do hipocampo por detecção de borda



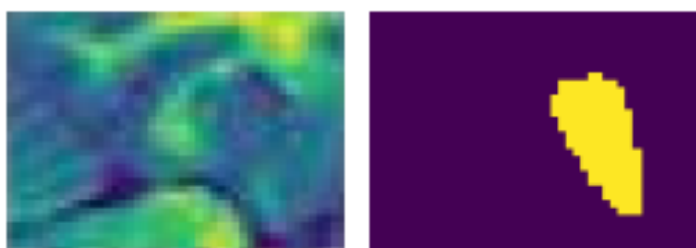
Fonte: ATALOGLOU et al. (2019)

Segmentação por crescimento de região: partindo de pontos de origem, o algoritmo analisa os *pixels* vizinhos e os agrupa caso possuam características similares de acordo com regras pré-definidas. Nessa técnica a escolha dos pontos de

origem é de extrema importância, assim como das regras de crescimento e critérios de parada.

Segmentação semântica: nesse modelo é atribuído um rótulo de classe para cada pixel da imagem, de forma que pixels que compartilhem do mesmo rótulo possuam características em comum (DATA SCIENCE ACADEMY, 2021). Dessa forma são delimitadas regiões que representam objetos na imagem. É possível definir diversas classes simultaneamente, permitindo por exemplo identificar numa imagem da região abdominal diversos órgãos, ou apenas uma região, como identificar o hipocampo numa imagem de RM do cérebro conforme exemplificado na Figura 9.

Figura 9 - Segmentação semântica do hipocampo a partir de uma imagem de RM



Fonte: Adaptado de Simpson et al. (2019)

O processo de segmentação também pode ser acelerado quando a posição relativa do objeto de interesse é conhecida na imagem. Por exemplo, se é de conhecimento que o hipocampo está localizado na região central, é possível desconsiderar as margens da imagem, o que reduz a quantidade de dados processados e torna o processo mais rápido e assertivo, principalmente na análise de muitas imagens.

2.5 Redes Neurais

Redes Neurais Artificiais (RNA), do inglês *Artificial Neural Networks*, são definidas como modelos computacionais que reproduzem o modelo do cérebro humano, e adquirem, armazenam e utilizam conhecimento experimental.

As RNA, foram criadas pelo Ph.D Frank Rosenblatt em 1958, quando se desenvolveu o *Perceptron*, um algoritmo para o reconhecimento de padrões baseado em uma rede neural computacional de duas camadas usando simples adição e subtração. As RNA e suas posteriores derivações englobam hoje as diversas técnicas

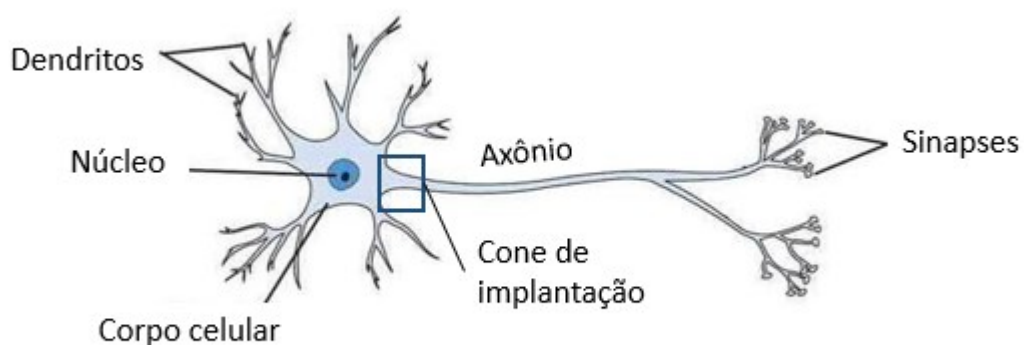
de Aprendizagem Profunda disponíveis e se baseiam no funcionamento dos neurônios no cérebro humano. Elas atraíram maior atenção da comunidade acadêmica e da indústria, nas últimas décadas, com o advento das Redes Neurais Profundas (*Deep Neural Networks*) e tem transformado os mais diversos campos, especialmente os campos de visão computacional, processamento de sinais, áudio e fala e processamento de linguagem natural (DATA SCIENCE ACADEMY, 2021).

As redes neurais artificiais utilizam modelos de neurônios e se inspiram no comportamento e/ou na ação humana, de forma a obter respostas automáticas e que possam ser consideradas racionais. Essa racionalidade é um critério sob o qual se espera que o resultado obtido por um agente computacional seja denominado correto, pensando nos limites das informações conhecidas quanto ao problema e ao ambiente envolvidos (NORVIG; RUSSELL, 2013).

2.5.1 Modelo De Um Neurônio

Os neurônios são considerados blocos básicos de construção em uma rede neural, seja essa uma Rede Neural Biológica (RNB) ou artificial. Na Figura 10, é possível observar de forma sintética as diferentes partes de um neurônio biológico. Os dendritos recebem sinais de outros neurônios, e os conduzem até o corpo da célula, onde ocorre uma soma ponderada dos sinais de entrada. Caso a somatória dos sinais de entrada atinja um limiar, ocorre a formação do impulso nervoso no cone de implantação, região entre o corpo celular e o axônio, percorrendo o axônio até o terminal axonal onde ocorrerá nova comunicação com outros neurônios através das sinapses (JAGREET, 2017).

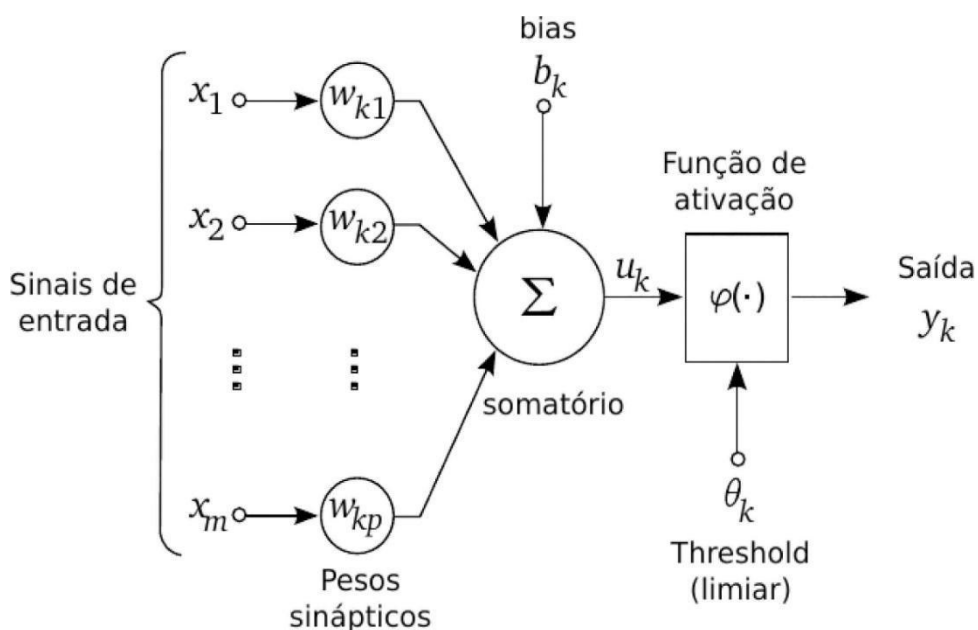
Figura 10 - Componentes da estrutura de um neurônio biológico



Fonte: Adaptado de Jagreet (2017)

O poder computacional está diretamente relacionado à quantidade de neurônios interligados entre si, portanto, um neurônio isolado não possui muita utilidade. Quanto maior for a quantidade de neurônios interligados entre si, maior será a capacidade de construir uma rede capaz de solucionar um determinado problema. A Figura 11 aponta o modelo de um neurônio responsável por constituir a base de uma RNA, ou seja, uma unidade de processamento.

Figura 11 - Modelo simplificado de um neurônio com x_m entradas



Fonte: Adaptado de Haykin (2004)

Um neurônio é composto por três principais partes:

(i) Um conjunto de sinapses, cada uma caracterizada por um peso próprio. Análogo à uma Rede Neural Biológica (RNB), existe um processo de multiplicação de um sinal de entrada e persistência desse sinal até o corpo da célula, caracterizando o funcionamento dos dendritos. De modo específico, um sinal x_j na entrada da sinapse j conectada ao neurônio k multiplica-se pelo peso sináptico w_{kj} , no qual o primeiro subíndice se refere ao neurônio e o segundo se refere ao terminal de entrada da sinapse à qual o peso se refere.

(ii) Um somador responsável por realizar a adição dos sinais de entrada, ponderados por seus respectivos pesos, o qual por analogia pode ser considerado como parte do corpo da célula em uma RNB.

(iii) Uma função de ativação, a qual possui objetivo de restringir a amplitude da saída do neurônio, e que também constitui parte da célula em uma RNB.

O modelo da Figura 11 inclui a utilização de um *bias* aplicado externamente (b_k), a fim de ampliar ou reduzir a entrada da função de ativação, o que dependerá do sinal ser positivo ou negativo, respectivamente.

2.5.2 Funções De Ativação

Responsável pelo processamento de cada neurônio, a escolha de uma determinada função de ativação de uma rede neural se dá baseada nas características dos dados de entrada. Segundo Haykin (2004) as funções de ativação estão divididas em três tipos básicos:

Função de limiar, ou função *Heaviside*. O primeiro modelo de neurônio a utilizar essa função ficou conhecido como neurônio de McCulloch e Pitts (1943). Neste caso, se o campo local induzido, também conhecido como potencial de ativação, for positivo a saída de um neurônio assumirá valor 1, e caso contrário assumirá valor 0.

$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases} \quad (1)$$

Função de ativação linear por partes, é uma função real definida nos números reais ou em um segmento, cujo gráfico é composto de seções em linha reta.

$$\varphi(v) = \begin{cases} 1, & v \geq 0,5 \\ v, & 0,5 > v > -0,5 \\ 0, & v \leq -0,5 \end{cases} \quad (2)$$

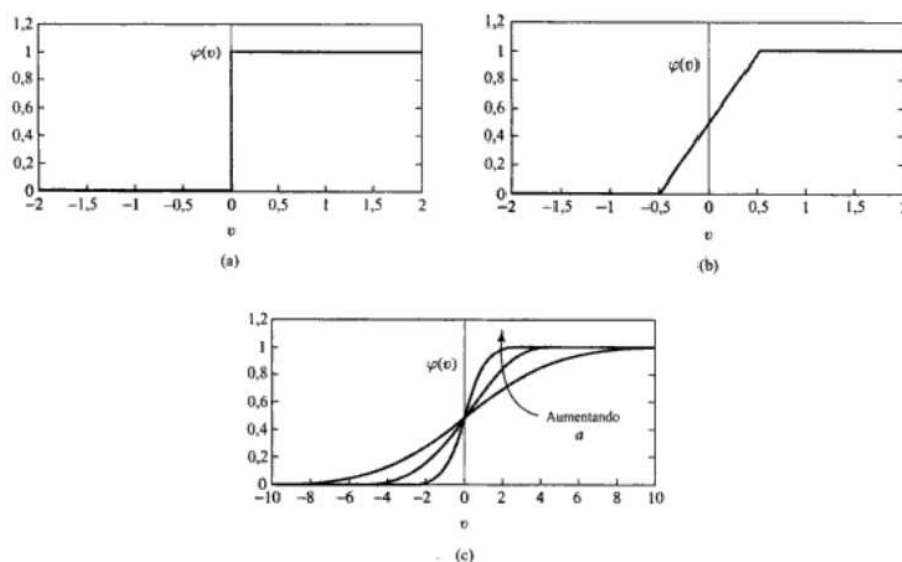
Função de ativação sigmóide é a forma mais utilizada em RNAs devido a sua forte inspiração biológica. É caracterizada como uma função rigorosamente crescente que exibe um balanceamento adequado entre comportamento linear e não-linear. Um exemplo de função sigmóide é a função logística definida por:

$$\varphi(v) = \frac{1}{1 + e^{-av}} \quad (3)$$

Sendo a definido como o parâmetro de inclinação da função sigmóide. Portanto, ao variar o a , sigmóides com diferentes inclinações serão obtidos.

Uma representação gráfica das funções supracitadas pode ser observada na Figura 12.

Figura 12 - (a) Função de limiar (b) Função linear por partes (c) Função sigmóide



Fonte: Adaptado de Haykin (2004)

2.5.3 Redes *Perceptron*

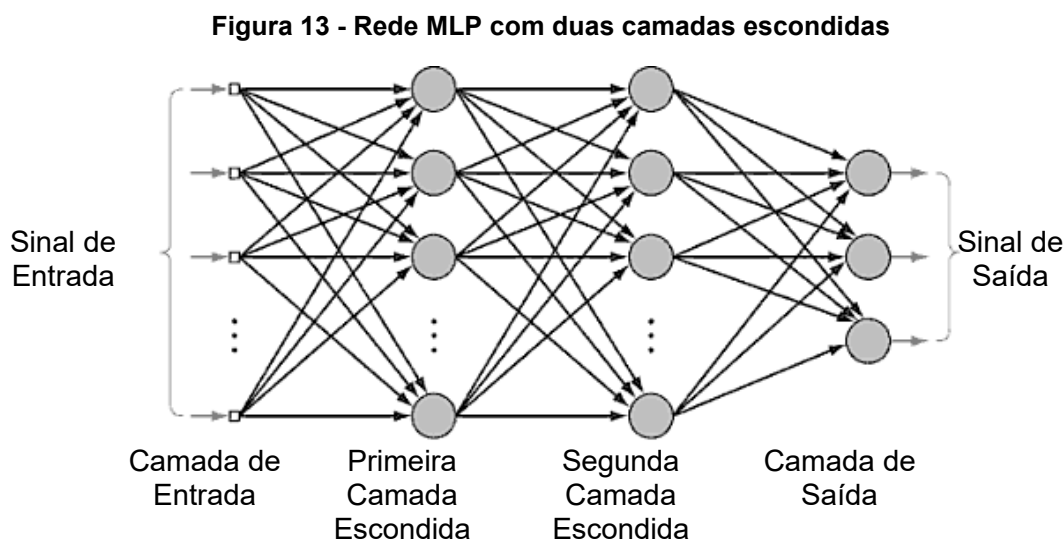
Single layer feed-foward neural network ou *Perceptron* é como se denomina uma rede neural com somente uma camada. Rosenblatt (1958) foi responsável por propor o modelo *Perceptron* como primeiro modelo para aprendizagem supervisionada. Sua ideia era criar uma rede que pudesse aprender a se ajustar sozinha.

Por ser constituída por apenas uma camada, a rede *Perceptron* é considerada a forma mais simples de uma rede neural. Sua arquitetura é usada para classificação

de padrões ditos linearmente separáveis (HAYKIN, 2004), em que as classes possíveis de classificação variam com base no número de neurônios encontrados na camada de saída.

2.5.4 Redes *Perceptron* Multicamadas

As Redes *Perceptron* Multicamadas (*Multilayer Perceptron* - MLP), são redes do tipo *feedforward* caracterizadas por conter camadas escondidas (*hidden layers*), situadas entre a camada de entrada e a de saída (NUNES; DANILO; FLAUZINO, 2010). A Figura 13 apresenta a estrutura de uma arquitetura MLP com duas camadas escondidas.



Fonte: Adaptado de Haykin (2004)

2.5.5 Backpropagation

Redes MLP utilizam um algoritmo de aprendizado supervisionado para realizar seu treinamento, o qual é denominado *backpropagation*. O objetivo desse algoritmo é otimizar os pesos para que a rede neural possa aprender a mapear corretamente as entradas arbitrárias para as respectivas saídas. As técnicas de propagação reversa popularizaram as redes densas e a MLP com algoritmo de treinamento *backpropagation* foi um dos mais utilizados modelos de RNA. Porém, quando o número de camadas crescia demasiadamente, essas redes eram incapazes de extrair com eficiência as informações durante o treinamento devido à quantidade

de cálculo necessário à nova camada para minimizar o custo ou o erro de resposta (SILVA; LENZ, 2019).

O aprendizado ocorre em duas etapas, a fase *forward* e a fase *backward*. Na fase *forward*, os sinais de entrada atravessam toda a rede em direção a camada de saída. Tais sinais são multiplicados pelos seus respectivos pesos, os quais iniciam de forma aleatória com valores entre 0 e 1. Para cada neurônio são somados os produtos das entradas com seu conjunto correspondente de pesos como mostra a Equação 4.

$$net = \sum_{i=1}^n x_i w_{ji} + b_i \quad (4)$$

A constante n_j simboliza o número de conexões na entrada do neurônio j , w_{ji} simboliza o peso da conexão entre a entrada x_i e o neurônio j e θ_i representa o peso do bias (CASANOVA, 2008).

Ao final dessa somatória a função de ativação é aplicada - isto ocorre até que um sinal de saída seja apresentado na última camada. O sinal de saída é dado pela Equação 5.

$$y_i = \varphi(net_j) \quad (5)$$

Após calculada a saída da rede, calcula-se o erro de saída para dar início a fase de retro-propagação, ou seja, a fase *backward*. Uma das maneiras de calcular o erro é dado pela Equação 6.

$$E_{total} = \frac{1}{2}(target - output) \quad (6)$$

A variável *output* caracteriza a saída da rede MLP, e a variável *target* simboliza a resposta esperada, o rótulo de saída a ser alcançado pela rede. Com o erro calculado, inicia-se a fase *backward*. O erro é retro-propagado com o intuito de atualizar o peso das suas sinapses, para que os valores *target* e *output* se aproximem, minimizando assim o erro de cada neurônio de saída e também da rede como um todo.

O objetivo nesta fase é minimizar o erro através da variação do peso. O que está expresso na Equação 7, sendo que η denominado como taxa de aprendizagem. A taxa de aprendizado é um hiperparâmetro que controla o quanto os pesos da rede serão ajustados, η deve ser um valor positivo entre 0 e 1. A escolha desse valor resultará na velocidade com que a rede irá convergir: caso o valor da taxa de aprendizado seja muito grande, o algoritmo irá exceder o custo global mínimo, e caso

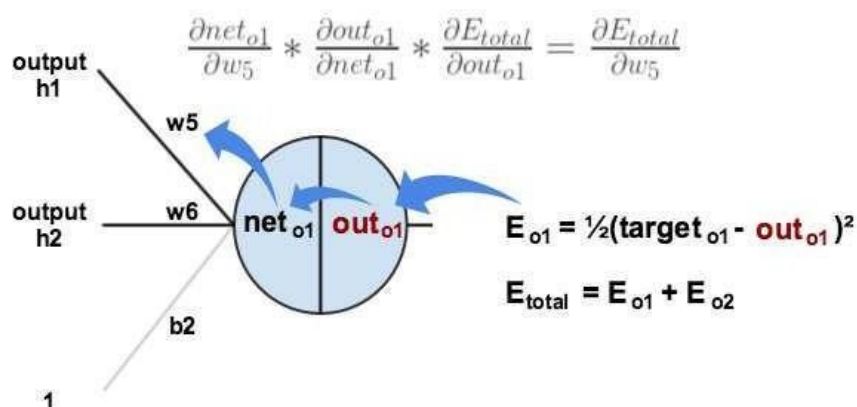
for muito pequeno, o algoritmo terá necessidade de mais épocas até a convergência, o que pode ocasionar lentidão para o aprendizado.

$$\Delta w_{ij} = -\eta \frac{\partial E_{total}}{\partial net_j} \quad (7)$$

Contudo, o erro dado por ∂E_{total} decorre da saída do neurônio net_j , o qual depende do peso das sinapses dada ∂w_{ij} . Desse modo, se faz necessária a utilização da regra da cadeia.

A Figura 14 exemplifica o comportamento mencionado previamente, em que out_{o1} equivale a y_i e w_5 representa uma sinapse dada por w_{ij} . Sendo assim, o E_{total} será derivado em relação à saída do neurônio y_i que por sua vez será derivado em relação à saída da função de ativação dada por net_j que por fim será derivada em relação ao peso do neurônio dado por w_{ij} .

Figura 14 - Aplicação do comportamento do backpropagation



Fonte: Adaptado de Mazur (2015)

2.6 Redes Neurais Convolucionais

No final do século passado, a proposta de novas redes para a resolução de problemas complexos, como a análise de imagens, carecia da concepção de novas arquiteturas de RNA, bem como de um maior poder computacional que permitisse o treinamento de redes com um maior número de camadas, ou mais profundas.

As redes neurais profundas surgiram no final da primeira década deste século. Dentre elas, as redes neurais convolucionais são as que têm aplicações mais difundidas. Elas trouxeram uma mudança fundamental na arquitetura das RNA, qual seja, o protagonismo de camadas de convolução. Ao invés da multiplicação de

grandes matrizes de peso pelas ativações de camadas anteriores, tem-se a convolução de pequenos filtros com um pequeno número dessas ativações, fazendo com que a memória utilizada para armazenar os dados da arquitetura seja significativamente reduzida. Além disso, inspirado no processamento do córtex visual dos mamíferos, é introduzida uma nova camada, a camada de amostragem. Essa última confere às redes a invariância em relação a translações de pequenos detalhes em uma imagem. Por fim, as redes convolucionais não mais carecem de uma etapa prévia de extração de características. Essa tarefa é agora delegada às primeiras camadas da arquitetura. As redes convolucionais foram consideradas um dos 10 maiores avanços tecnológicos do ano de 2013 (COSTA, 2017).

Uma Rede Neural Convolucional é denominada como um algoritmo de Aprendizado Profundo que pode captar uma imagem de entrada, atribuir importância (pesos e vieses que podem ser aprendidos) a vários aspectos da imagem e ser capaz de diferenciar um do outro. Uma CNN usa uma variação de *perceptrons* multi-camada desenvolvidos de modo a demandar o mínimo pré-processamento possível. Enquanto nos métodos primitivos os filtros são feitos à mão, com treinamento suficiente, as CNN têm a capacidade de aprender essas características (DATA SCIENCE ACADEMY, 2021).

As redes convolucionais são inspiradas nos processos biológicos. Nelas, o padrão de conectividade entre os neurônios é inspirado na organização do córtex visual dos animais. Neurônios corticais individuais respondem a estímulos apenas em regiões restritas do campo de visão conhecidas como campos receptivos. Os campos receptivos de diferentes neurônios se sobrepõem parcialmente de forma a cobrir todo o campo de visão (DATA SCIENCE ACADEMY, 2021).

Se comparada a outros algoritmos de classificação de imagens, a CNN tende a demandar um nível mínimo de pré-processamento. Isso significa que a rede "aprende" os filtros que em um algoritmo tradicional precisariam ser implementados manualmente. Essa independência de um conhecimento a priori e do esforço humano no desenvolvimento de suas funcionalidades básicas pode ser considerada a maior vantagem de sua aplicação.

Esse tipo de rede é utilizado principalmente em reconhecimento de imagens e processamento de vídeo. Por apresentar resultados muito precisos e comparáveis aos clínicos especializados, na área da saúde essa metodologia é utilizada com

algoritmos específicos, recorrendo a um grande número de imagens clínicas para o diagnóstico de determinadas doenças.

As CNN têm vantagens significativas sobre as demais técnicas quando processa sinais naturais, cuja diferença computacional necessária é expressiva se comparada às redes neurais densas. Elas são capazes de aprimorar o processamento subdividindo a análise em partes menores, por meio de filtros que se adaptam para encontrar características específicas em sons e imagens. Dessa forma, elas são divididas em camadas, cada qual com uma função claramente definida, para somente então entregar a classificação da imagem nos neurônios de saída com o auxílio de uma rede totalmente conectada similar às densas (SILVA; LENZ, 2019).

2.6.1 Camadas Convolucionais

Para completa compreensão do funcionamento das redes neurais convolucionais é necessário ter um conhecimento prévio de operações de convolução. Para entender o funcionamento da operação de convolução pode-se supor a existência de dois vetores x e w , sendo y apontado como a convolução entre eles, ou seja:

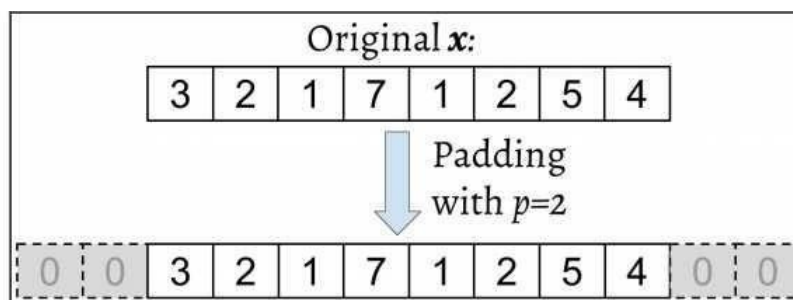
$$y = x * w \quad (8)$$

O vetor x é a entrada da camada convolucional, muitas vezes chamado de sinal, e o vetor w é o filtro de características ou filtro de kernel. A operação indicada na Equação 8 pelo símbolo $*$ simboliza a operação de convolução entre os dois vetores x e w . Em sua forma geral, a convolução pode ser descrita da seguinte maneira:

$$y = x * w : y[i] = \sum_{k=-\infty}^{+\infty} x[i - k]w[k] \quad (9)$$

Em aplicações de *Machine Learning* é comum trabalhar constantemente com vetores de características finitos. A título de exemplo, se o vetor de entrada x possui 10 características, terá seus índices de maneira tal a ser 0,1, 2,...,9, portanto os índices $-\infty:-1$ e $+\infty:10$ estão fora dos limites do sinal x . Esse processo é denominado como *Zero-padding* ou *Padding* (RASCHKA; MIRJALILI, 2017). É possível observar na Figura 15, em que o *padding* denotado por p é igual a 2.

Figura 15 - Padding de tamanho 2



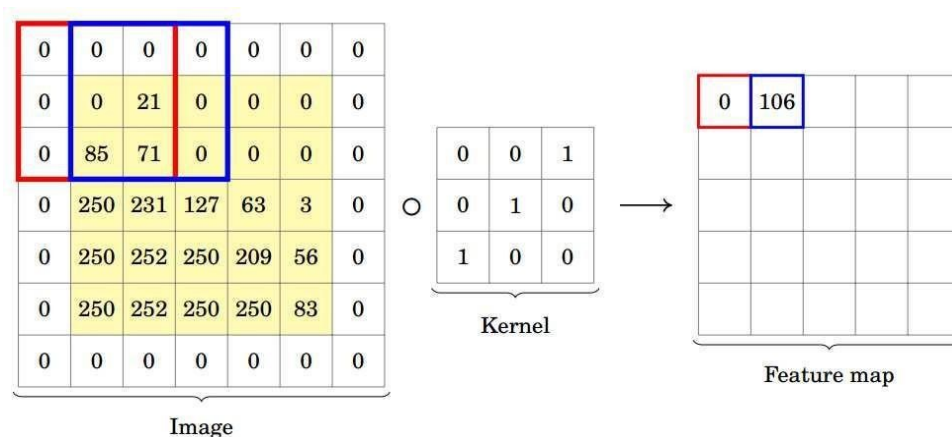
Fonte: Raschka e Mirjalili (2017)

As camadas convolucionais são compostas por um número " n " de filtros. Tais filtros possuem pesos previamente inicializados que, conforme a rede aprende a identificar regiões significantes a fim de minimizar a função custo, serão ajustados através do método de *backpropagation* (MASSUCATTO, 2018).

Os filtros geralmente são matrizes pequenas, matrizes 5x5 nos 3 canais (R, G e B) de cor, composta por valores reais os quais serão convoluídos com os dados de entrada a fim de obter um mapa de características. Estes mapas indicam regiões nas quais as características extraídas pelos filtros se encontram (MASSUCATTO, 2018).

Os filtros, também chamados de campos receptivos ou detectores de características, possuem parâmetros principais importados da convolução unidimensional de dois vetores, que podem mudar o comportamento de cada camada. Alguns parâmetros podem ser pré-determinados, como o preenchimento da imagem e o passo com o qual o filtro irá percorrer toda a imagem de entrada. Na Figura 16 é possível observar uma camada convolucional parcial, utilizando um filtro 3x3.

Figura 16 - Camada convolucional parcial utilizando um filtro 3x3



Fonte: Pavlovsky (2017)

O tamanho da matriz resultante da camada de convolução será também a entrada da próxima camada da CNN, o qual pode ser obtido através da equação 10:

$$O = \lfloor \frac{(n + 2p - m)}{S} + 1 \rfloor \quad (10)$$

A qual:

O é o tamanho da matriz de saída em linhas x colunas.

n é o tamanho da imagem de entrada em linhas x colunas.

m é o tamanho do filtro.

p é o preenchimento.

S é o *stride*/passo.

A função piso (*floor*) denotada por $\lfloor . \rfloor$ retorna o maior inteiro igual ou menor que a entrada. Por exemplo:

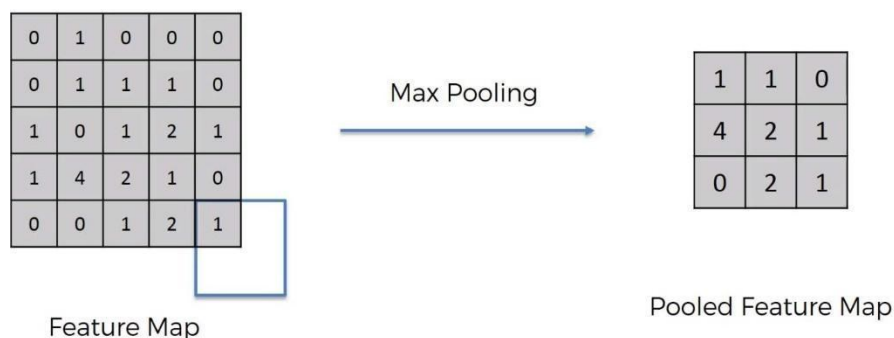
$$\text{floor}(1.77) = \lfloor 1.77 \rfloor = 1 \quad (11)$$

Como mostra a Figura 16, a matriz resultante possui um número menor de valores. O propósito é detectar as partes mais significantes da imagem. Diversos mapas de características são criados na primeira camada de convolução com o objetivo de preservar um maior número de informações. De forma geral, a camada de convolução informa a rede onde as características importantes da imagem estão localizadas.

2.6.1.1 Camadas De Agrupamento (*Pooling*)

A segunda camada em uma CNN é denominada camada de *pooling* e é utilizada com o intuito de diminuir o tamanho espacial das matrizes consequentes da convolução, diminuindo dessa forma os parâmetros a serem aprendidos na rede. As camadas de *pooling* são, portanto, uma forma de *down-sampling*, e contribuem para o controle de *overfitting* - fenômeno que ocorre quando existe um número grande de parâmetros livres e o classificador tende a se adaptar a detalhes específicos da base de treinamento, o que pode causar uma redução da taxa de acerto. Uma camada de *pooling* é como um filtro de tamanho e passo previamente determinado que estima qual é o máximo local de uma determinada região do mapa de atributos (MASSUCATTO, 2018). A Figura 17 mostra uma camada de *pooling* de tamanho 2x2 com um passo de 2.

Figura 17 - Camada de pooling utilizando um kernel 2x2



Fonte: SUPERDATASCIENCE (2017)

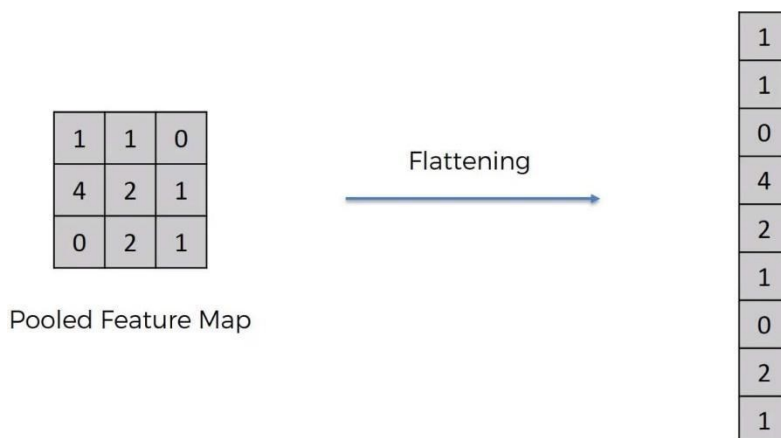
Dois tipos de operações de *pooling* bastante usadas são a *max pooling* e *mean-pooling* (também conhecida como *average pooling*). A função *max pooling* analisa o maior valor de uma região, portanto os valores máximos no mapa de características simbolizam onde estão localizadas as características, e ao aplicar a camada de *pooling* uma porcentagem de informação que não é característica é descartada. Mapa de características agrupadas (*pooled feature map*) é como pode ser chamada a matriz (imagem) resultante. Já a função *average-pooling* é responsável por estimar o valor médio de uma determinada região limitada pelo filtro (WU, 2017).

É comum que modelos de arquitetura CNN possuam diversos estágios de convolução, não-linearidade e camadas de *pooling* seguidos por mais camadas de convolução e camadas totalmente conectadas.

2.6.1.2 Camadas Totalmente Conectadas (*Fully Connected*)

De forma inicial, antes de chegar até as camadas totalmente conectadas é necessário converter as matrizes de características resultantes da camada de *pooling* em um vetor coluna o qual servirá de entrada para a RNA clássica, como é mostrado na Figura 18.

Figura 18 - Operação de *flattening* dado um mapa de características

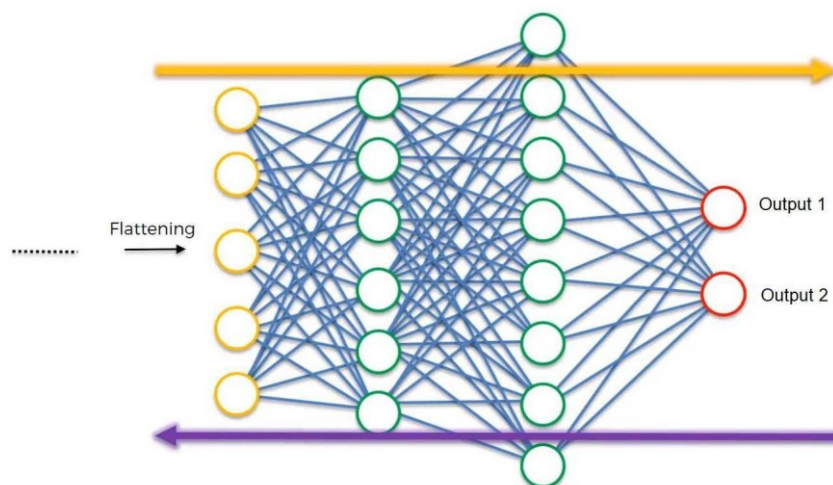


Fonte: SUPERDATASCIENCE (2017)

Normalmente as camadas totalmente conectadas são utilizadas como camadas finais. Ao contrário das camadas convolucionais, todos os pesos são totalmente conectados com a camada anterior.

Desta forma, uma RNA é adicionada ao final da CNN, em que cada elemento do vetor coluna criado pela camada de *flattening* se comporta como uma entrada para uma RNA clássica. Na sequência esses vetores serão ligados às camadas escondidas totalmente conectadas até a camada final, como observa-se na **Figura 19**.

Figura 19 - Camada totalmente conectada

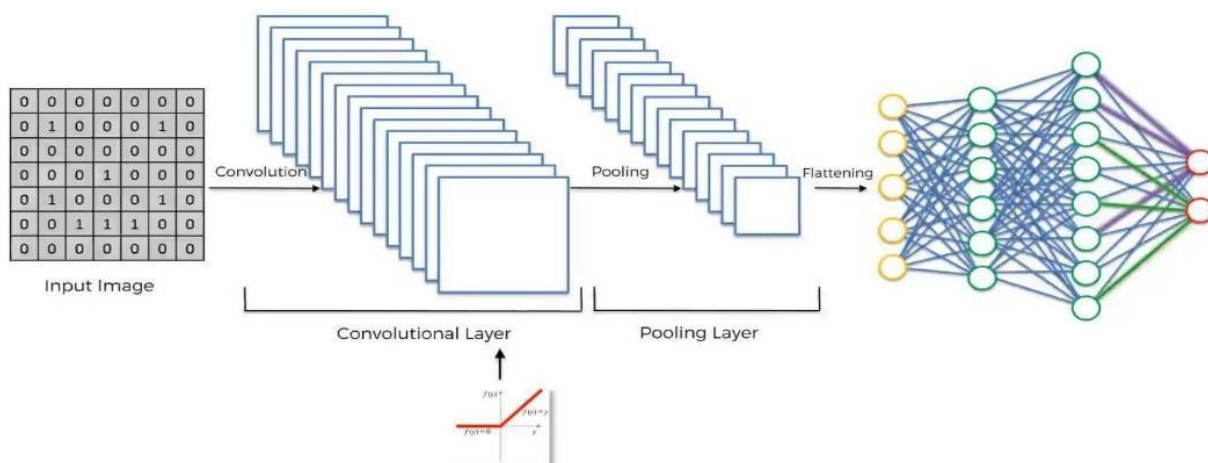


Fonte: SUPERDATASCIENCE (2017)

Posteriormente a uma operação de *flattening*, a finalidade de uma RNA é combinar as características em mais atributos para aprimorar o desempenho de previsão das classes. Calcula-se e propaga-se de volta o erro assim como em uma RNA comum, e alguns parâmetros como, por exemplo, os pesos da RNA (sinapses) e os filtros (detectores de características) são ajustados com o intuito de otimizar o desempenho. A princípio, os filtros procuram por certas características, mas tais características podem não ser adequadas, então os filtros são ajustados para que procurem por novas características (MASSUCATTO, 2018).

Por fim, na Figura 20 é possível observar uma Rede Neural Convolutiva desde sua imagem de entrada até a classe de saída.

Figura 20 - Arquitetura simplificada de uma Rede Neural Convolutiva



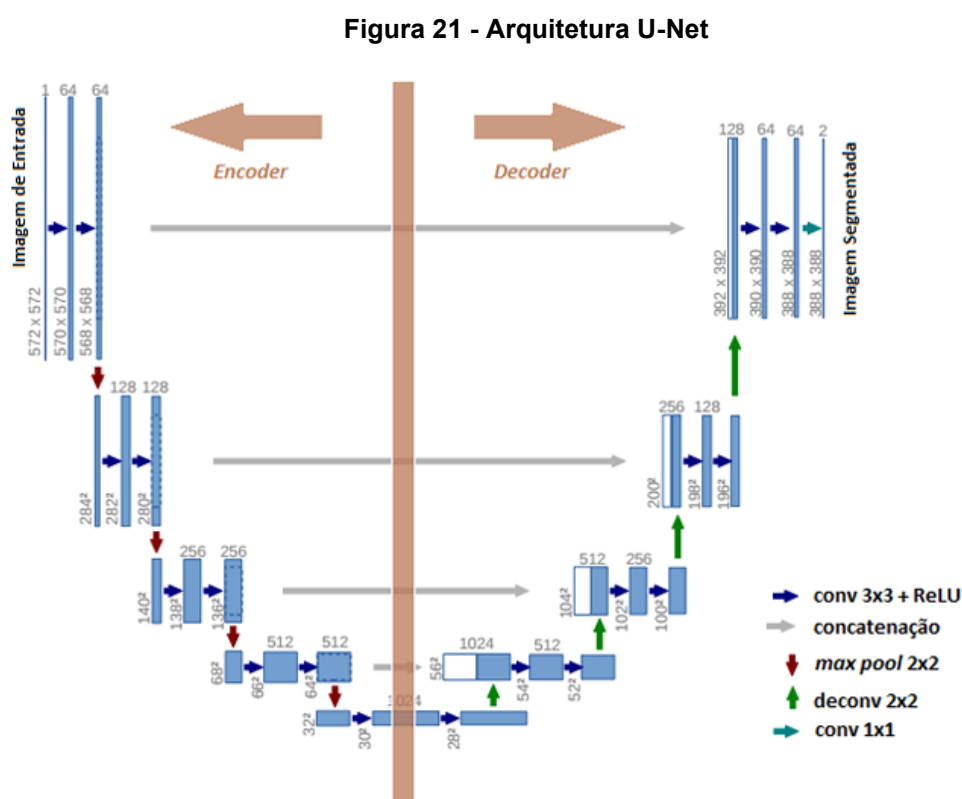
Fonte: SUPERDATASCIENCE (2017)

2.6.2 Redes U-Net

A abordagem de uma *fully convolutional network* consiste em transformar as camadas totalmente conectadas em camadas de convolução, fazendo com que seja possível a criação de mapas de características de segmentação para cada imagem, sendo de qualquer tamanho.

A Rede U-Net foi projetada e aplicada por Ronneberger, Fischer e Brox (2015) para o processamento de imagens biomédicas. Com uma abordagem superior a de uma rede neural convolucional convencional, a Rede U-Net não somente classifica a imagem, mas também localiza a área de anormalidade em casos biomédicos. Essa distinção de borda com localização de região é possível pelo fato da U-Net realizar a classificação em cada pixel, desta forma entrada e saída possuem o mesmo tamanho.

Com uma arquitetura em formato de “U”, como é possível observar na Figura 21 abaixo, a rede U-Net é simétrica e possui duas partes principais.



- **Encoder:** É a primeira metade da arquitetura do diagrama. Geralmente é uma rede de classificação pré-treinada onde são aplicados blocos de convolução

seguidos por um *maxpool downsampling* para codificar a imagem de entrada reduzindo assim sua resolução espacial.

- *Decoder*: O *decoder* é a segunda camada da arquitetura. Nesse caso, o objetivo é projetar semanticamente as características de discriminação (menor resolução) aprendidas pelo *encoder* no espaço de pixels (maior resolução) para obter uma classificação densa. O *decoder* consiste em *upsampling* (convolução transposta) e concatenação seguidos por operações de convolução.

O principal diferencial da U-Net se faz na camada de *decoder*, na qual existe um número grande de mapas de características, o que proporciona uma simetria em relação à camada de *encoder*. O aumento do número de mapas de características no *decoder* se dá pela concatenação da saída de uma camada de *unpooling* com a saída da camada de convolução do *encoder* correspondente.

2.7 Métricas De Validação

Diversas métricas podem ser utilizadas para determinar o desempenho do modelo da CNN implementada. As duas mais difundidas e utilizadas no estágio de treinamento da rede são a acurácia e a perda.

A perda é a soma dos erros do modelo durante as fases de treinamento e validação. Diversas imagens são utilizadas durante o treinamento, e para cada uma delas há um erro (uma diferença entre a imagem segmentada e a real). A soma dos erros de cada imagem utilizada durante o treinamento/validação é a perda do modelo. É importante que a perda diminua ao longo do treinamento, o que significa que o modelo está aprendendo e melhorando seus resultados. Se o valor da perda estiver apenas oscilando, pode ser que o modelo não esteja aprendendo. Caso esteja diminuindo no conjunto de treinamento, mas não no de validação, significa que o modelo está se tornando super ajustado, ou seja, se adequando apenas às imagens de treinamento e tornando-se inválido para imagens externas (RIVA, 2021).

A acurácia é um valor percentual que mede o quão bem um modelo prevê, comparando as previsões com o objeto real (RIVA, 2021). Por exemplo, uma acurácia de 0,8 significa que o modelo é capaz de prever corretamente 80% dos casos.

A acurácia e a perda podem apresentar um problema em casos onde o objeto a ser segmentado apresenta uma quantidade de *pixels* muito pequena em relação à imagem toda. Nessas situações podemos ter uma acurácia alta e perda baixa (resultado ideal) mas o modelo não necessariamente irá acertar a localização do objeto na imagem (HOFESMANN, 2020).

3 MATERIAIS E MÉTODOS

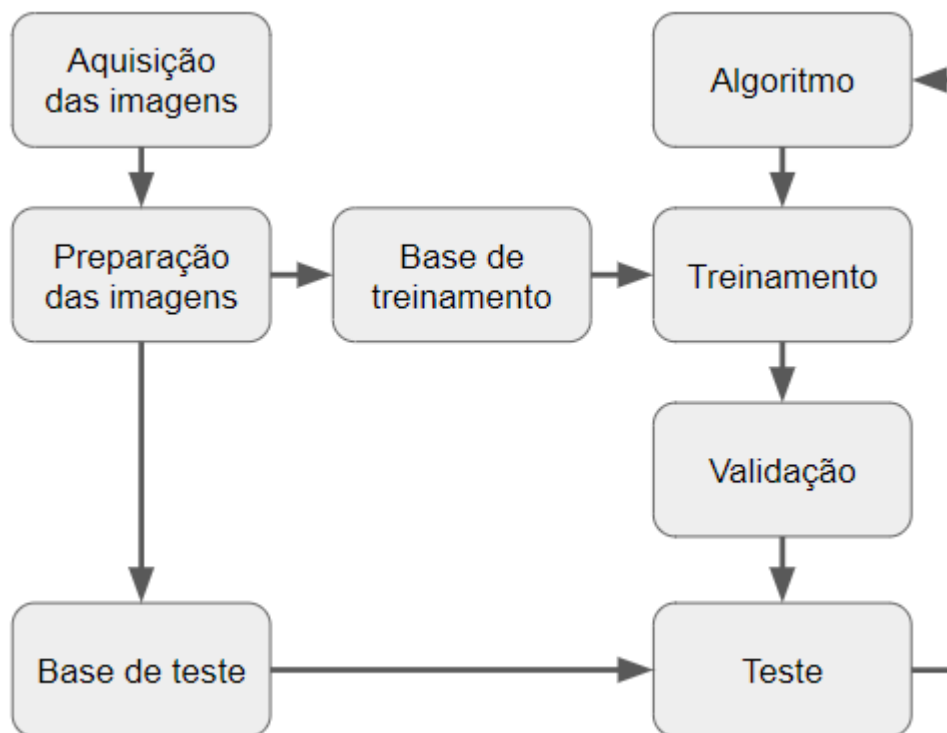
3.1 Metodologia

O trabalho consistiu no desenvolvimento de um algoritmo que permitisse identificar o hipocampo em imagens de RM utilizando uma técnica chamada segmentação semântica de imagens. A segmentação é o primeiro passo no processo que permite analisar o volume da região e classificá-la dentro de parâmetros de normalidade e de atrofia, sendo esta relacionada com o diagnóstico de DA.

O algoritmo foi treinado utilizando-se um conjunto de imagens (base de treinamento) que possuem uma máscara associada, que é uma segmentação manual da região de interesse realizada por especialistas e profissionais com experiência na análise de imagens de RM. Na sequência o algoritmo foi aplicado a outro conjunto de imagens (base de teste) composto por imagens distintas do primeiro, mas oriundos da mesma fonte.

O desempenho do algoritmo foi avaliado pela comparação entre a segmentação gerada pelo processamento dos dados e a segmentação manual fornecida pela base. Foram realizados diversos testes que levaram ao aperfeiçoamento do algoritmo até a obtenção de índices satisfatórios de acurácia (porcentagem de elementos classificados corretamente) e perda (acúmulo de erros durante o processo de treinamento e validação). O processo está ilustrado na Figura 22.

Figura 22 - Fluxograma do procedimento



Fonte: Autoria própria (2021)

3.1.1 Aquisição e preparação das imagens

O trabalho teve início pela identificação de bancos de imagens médicas de RM utilizados em outros estudos sobre auxílio diagnóstico para a DA. Foram identificados diversos bancos, mas apenas alguns atendiam aos critérios necessários para este estudo, que eram:

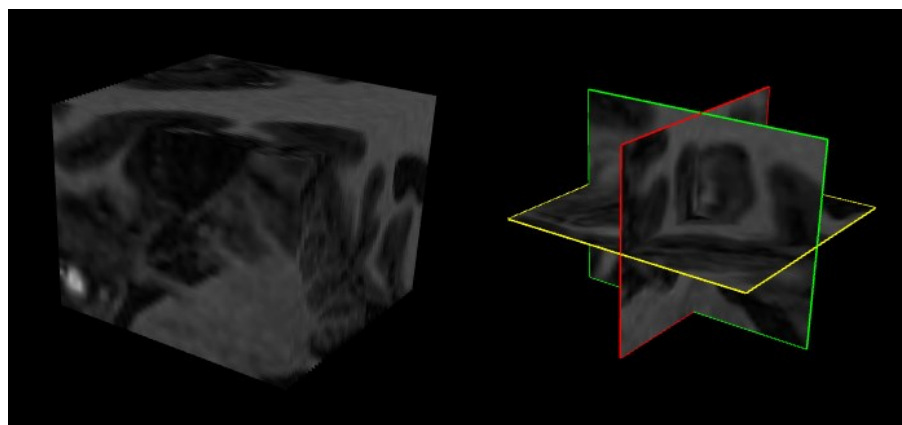
- As imagens deveriam conter o hipocampo;
- Preferencialmente o banco deveria ser o mesmo já utilizado em outros estudos e análises referentes a DA.
- As imagens deveriam possuir qualidade suficiente e um formato adequado;
- As imagens deveriam conter uma máscara representando a segmentação manual realizada por especialistas;
- O banco deveria possuir uma quantidade suficiente de imagens para que o treinamento do algoritmo fosse satisfatório;
- O banco deveria ser público ou ser possível obter a autorização da entidade proprietária para uso e divulgação.

Alguns dos conjuntos de imagens encontrados inicialmente possuíam características indesejáveis, como marcas d'água e fontes incertas. Outros atendiam aos critérios de qualidade, mas não representavam a região de interesse da forma adequada, todos esses foram logo descartados.

Dentre as melhores opções escolheu-se o banco disponibilizado pela *Medical Segmentation Decathlon (MSD)*, um desafio internacional de programação na área de inteligência artificial aplicada à medicina que buscava a obtenção de um método capaz de ter um bom desempenho em várias tarefas diferentes, como por exemplo a identificação de tumores em diversas partes do corpo, identificação de órgãos na região abdominal, identificação de vasos sanguíneos e identificação do hipocampo (ANTONELLI et al., 2021). Este banco possui uma licença que permite que os dados sejam distribuídos, compartilhados e melhorados.

O banco de imagens para o hipocampo foi fornecido para a MSD pelo *Vanderbilt University Medical Center*, uma entidade sem fins lucrativos vinculada à Universidade de Vanderbilt, situada em Nashville, Estados Unidos, e é composto por 394 arquivos no formato NIfTI, conforme ilustra a Figura 23, gerada no visualizador *online Brain Viewer*¹, sendo 263 com máscara e 131 sem máscara (SIMPSON et al., 2019).

Figura 23 - Imagem NIfTI tridimensional e cortes em três orientações



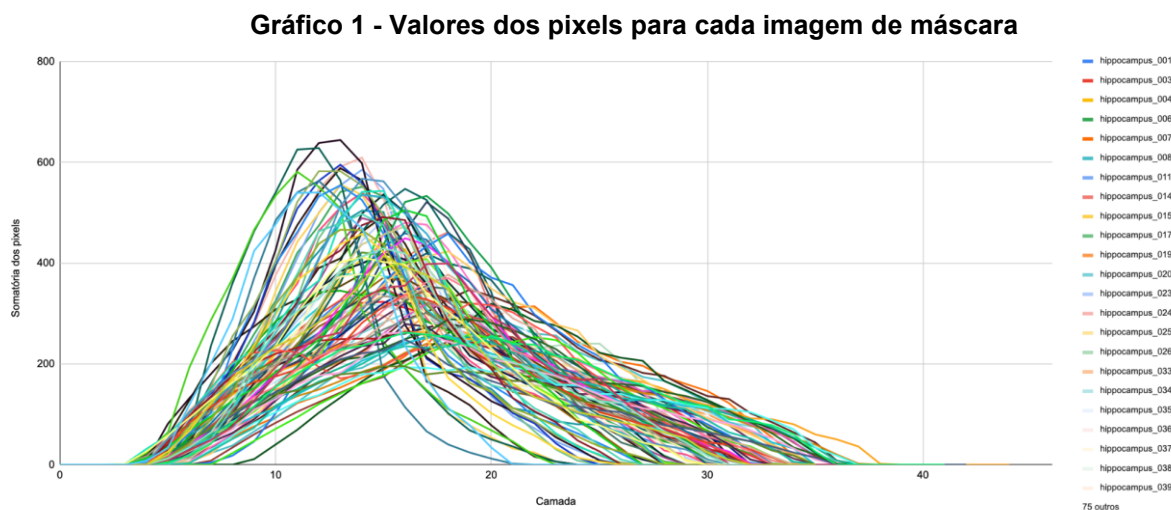
Fonte: Autoria própria (2022)

¹ Site do visualizador Brain Viewer: <https://socr.umich.edu/HTML5/BrainViewer/>

No primeiro momento selecionou-se apenas os arquivos com máscara, visto que elas são necessárias para o treinamento e para a validação dos resultados, mas foram identificados três problemas com o conjunto de 263 arquivos.

O primeiro problema é que alguns arquivos das máscaras continham apenas uma imagem totalmente preta, sem a delimitação do hipocampo. O segundo era que as imagens do cérebro não apresentavam o órgão inteiro, mas apenas uma pequena região que continha o hipocampo, e apresentavam baixíssimas resoluções na ordem de 50×30 *pixels* e diferindo entre si. O último era que as imagens possuíam três camadas distintas, uma para o fundo e duas para diferenciar as áreas do hipocampo. Na realidade, a presença dessas camadas pode ser uma grande vantagem, pois elas contêm informações que podem ser utilizadas, mas no caso deste projeto isso foi um problema devido à necessidade de conversão das imagens “.nii” para “.jpg”. O processo de conversão exigiu um tratamento especial e ocasionou a perda de informações contidas nas camadas originais.

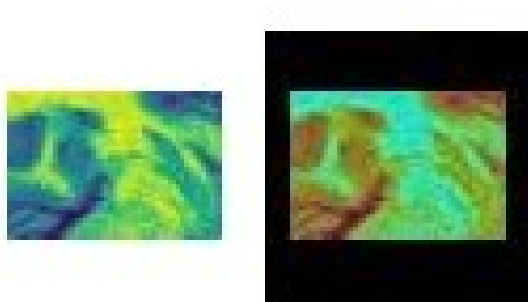
Para resolver o primeiro problema foi desenvolvido um algoritmo capaz de identificar quais arquivos de máscara estavam vazios e removê-los do conjunto. Como as imagens sem máscara eram totalmente pretas e *pixels* com essa cor possuem valor zero foi possível determinar quais imagens seriam removidas pela soma do valor de todos os *pixels*, excluindo aquelas cuja soma fosse zero. O Gráfico 1 apresenta o valor de cada *pixel* (eixo y) pela identificação do mesmo (eixo x) para as diversas imagens das máscaras (curvas). No Gráfico 1 as imagens vazias possuem linhas horizontais de valor zero.



Fonte: Autoria própria (2022)

Ao final foram removidas 2647 imagens, mantendo 6627 que realmente possuíam o hipocampo, resolvido esse problema das imagens sem máscaras foi necessário resolver o problema da resolução e formato irregular das imagens. Primeiramente as imagens NifTI foram divididas em camadas 2D em formato JPG. Em seguida foi adicionada uma borda preta ao redor do corte, mantendo o mesmo centralizado, como exemplifica a Figura 24, onde uma imagem com 51×35 pixels foi aumentada para 64×64 . A borda foi necessária para aumentar o tamanho das imagens e deixá-las com a mesma quantidade de pixels nas duas dimensões, ou seja, torná-las matrizes quadradas, além de dimensioná-las para uma potência de dois e padronizá-las, já que estes eram pré-requisitos do código utilizado como base (GRANATYR, 2022).

Figura 24 - Corte da primeira amostra do banco de imagens



Fonte: Autoria própria (2022)

Outra abordagem seria redimensionar a imagem, porém o redimensionamento poderia implicar na deformação dos padrões da imagem original levando a erros durante os processos de treinamento e validação. Já o preenchimento com *pixels* pretos não causa deformações, não causa perdas de acurácia e acelera os cálculos em relação a outra abordagem, resultando em melhor eficiência computacional. Porém a adição de valores zero na matriz pode resultar em diminuição do aprendizado nessas regiões de borda, já que alguns processos convolucionais podem não iniciar (HASHEMI, 2019).

Por fim, tivemos que unificar as classes de *pixels* das imagens para corresponder ao esperado pela CNN, isso inclui binarizar os arquivos. O processo de binarização consiste em marcar como zero valores abaixo de determinado limiar, e um os que estiverem acima. Essa divisão binária permite que a rede facilmente diferencie o fundo e a parte a ser segmentada.

Do conjunto final de imagens tratadas, 90% foram selecionadas de forma aleatória para o treinamento da CNN e os 10% restantes reservadas para a teste.

3.1.2 Definição do ambiente de desenvolvimento

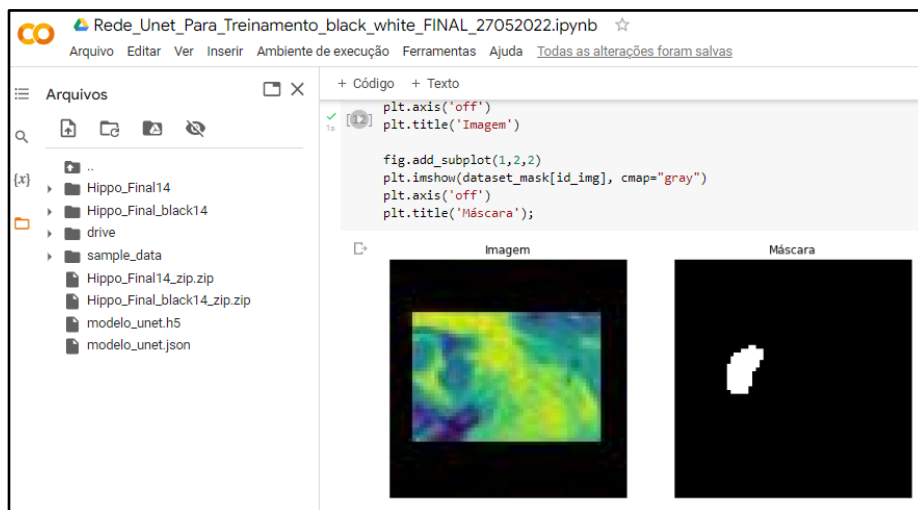
Em paralelo à busca por bancos de imagens foi realizada também a procura por códigos de *Deep Learning* utilizando CNN aplicados à área da saúde que serviriam como exemplos e modelos ao trabalho a ser desenvolvido. Foram encontrados diversos códigos abertos à comunidade, principalmente nas linguagens MATLAB e *Python*.

Entre as duas linguagens optou-se pela utilização do *Python*. Apesar do MATLAB ser uma ferramenta muito poderosa para o desenvolvimento de projetos que envolvem grande carga computacional e matemática, ele necessita de licença e de um computador com boa configuração, o que poderia impossibilitar o uso do código desenvolvido em determinados casos. Enquanto isso o *Python* é aberto e acessível a toda a comunidade e permite a utilização de ambientes de desenvolvimento que utilizam servidores remotos para o processamento.

A escolha do *Python* permitiu a utilização do Colab, um produto da *Google Research* que permite a escrita e execução de códigos *Python* pelo navegador. Além

disso, ele permite a utilização de arquivos salvos no *Google Drive* (serviço de armazenamento na nuvem) e todo o processamento é realizado em uma máquina virtual configurada em um servidor remoto (GOOGLE RESEARCH, 2022).

Figura 25 - Interface do Colab



Fonte: Autoria própria (2022).

Isso permitiu a fácil utilização do banco de imagem por todos os membros da equipe, além de agilizar os testes e treinamentos do algoritmo. Em alguns casos a execução do algoritmo foi acelerada em até 6 vezes quando comparada à execução local, visto a diferença de configuração entre os computadores particulares e o servidor remoto. Outra vantagem é que o Colab não requer nenhuma instalação referente ao ambiente e já possui as principais bibliotecas utilizadas pelo *Python* disponibilizadas.

Vale destacar a importante biblioteca Keras, que é uma interface de programação de aplicações executada na plataforma de *Machine Learning* TensorFlow (KERAS, 2022). A Keras disponibiliza diversas funções voltadas ao aprendizado de máquina que foram desenvolvidas e aprimoradas pela comunidade ao longo do tempo, entre elas as principais funções relacionadas à convolução utilizadas no desenvolvimento da CNN apresentada neste trabalho.

3.1.3 Desenvolvimento e treinamento do algoritmo

Tendo como ponto de partida os modelos obtidos foi desenvolvido um algoritmo de rede neural artificial em linguagem *Python*. O algoritmo inicia pela comunicação com o *Google Drive* e aquisição de arquivos compartilhados, seleção dos arquivos adequados e sua organização em base de treinamento e base de testes, aplica o tratamento de imagens descrito anteriormente e realiza a segmentação para a identificação do hipocampo.

Como optou-se por utilizar os dados de forma colaborativa, a primeira etapa do algoritmo era responsável por acessar os dados contidos e compartilhados na plataforma *Google Drive* e organizá-los no servidor remoto.

As imagens recuperadas estavam no formato “.nii”, o que nos permite o processamento de cada camada, excluir as imagens sem hipocampo - completamente pretas - e por fim salvar cada camada em um arquivo “.jpg”. Na sequência a borda preta foi adicionada a cada arquivo, mantendo a imagem original centralizada. O conjunto final foi ordenado e agrupado em listas (uma para as imagens do exame, uma para as imagens das máscaras, e outra para as imagens de teste) de forma que cada elemento na lista das imagens do exame ficasse na posição correspondente a sua máscara na lista de máscaras. Do conjunto de treinamento 90% são reservadas para o treinamento em si e 10% para a validação.

Com as listas de imagens prontas e ordenadas, foram montadas as funções de *encoder* e *decoder* da rede U-Net, que utilizam funções fornecidas pela biblioteca *Keras*. A rede U-Net, adaptada do modelo proposto por Granatyr (2022), foi organizada com quatro camadas de *encoder*, uma de convolução em sua base, e quatro de *decoder*, retornando uma imagem de saída com resolução igual às imagens de entrada. As listas de imagens também foram formatadas para se adequar ao formato de entrada de dados da CNN e então aplicadas à rede para iniciar o treinamento.

Durante o treinamento o modelo adequa seus parâmetros internos ao conjunto de imagens da base de treinamento. Feito isso, ele utiliza os novos parâmetros calculados para segmentar cada imagem de validação e então comparar

suas segmentações à suas máscaras. A diferença entre segmentação e máscara real gera um erro, que acumulado representa a perda do modelo.

Após o treinamento são gerados os gráficos de acurácia e perda, e exibidas as imagens das máscaras e segmentação para fins de comparação e análise visual do resultado da rede. Por fim, o modelo é aplicado às imagens da lista de teste, que não possuem máscaras conhecidas no momento, simulando uma aplicação real e exibindo na tela o resultado da segmentação.

3.1.4 Testes e validação

Durante a etapa de treinamento e validação foram obtidos valores de perda e acurácia. Sempre que o resultado era insatisfatório, com altos valores de perda e baixos valores de acurácia, o algoritmo era ajustado manualmente e uma nova fase de treinamento era iniciada. Além disso, caso as imagens de predição fossem visivelmente incoerentes, ajustes eram realizados independente dos valores das métricas. O procedimento foi repetido até que o diagnóstico obtido pela rede neural fosse satisfatório.

Após treinado, o modelo utilizou um segundo conjunto de imagens da mesma fonte daquelas utilizadas para seu treinamento. Com isso foi possível a comparação visual entre a imagem que temos com a delimitação feita por um profissional e a imagem gerada pela rede.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

4.1 Treinamento E Validação

O algoritmo desenvolvido obteve bons valores para a acurácia e perda durante o processo de treinamento e validação quando utilizados 15 ciclos de treinamento (épocas). Os dados obtidos para um desses treinamentos estão apresentados na Tabela 1. O modelo atualiza seus parâmetros internos apenas quando há uma redução no valor da perda durante a validação. Isso significa que para este treinamento específico a última atualização ocorreu na época 12, já que nas iterações seguintes o valor da perda na validação aumentou.

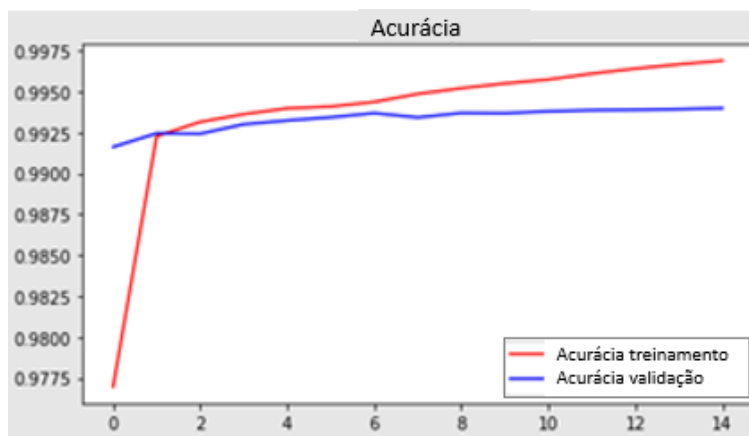
Tabela 1 - Dados de treinamento com 15 épocas

Época	Treinamento		Validação	
	Perda	Acurácia	Perda	Acurácia
1	0.1332	0.9770	0.0660	0.9916
2	0.0483	0.9923	0.0373	0.9924
3	0.0306	0.9931	0.0287	0.9924
4	0.0227	0.9936	0.0225	0.9930
5	0.0187	0.9940	0.0196	0.9932
6	0.0168	0.9941	0.0177	0.9934
7	0.0151	0.9944	0.0164	0.9937
8	0.0134	0.9949	0.0164	0.9934
9	0.0122	0.9952	0.0159	0.9937
10	0.0113	0.9955	0.0163	0.9937
11	0.0106	0.9957	0.0158	0.9938
12	0.0096	0.9961	0.0157	0.9939
13	0.0088	0.9964	0.0163	0.9939
14	0.0082	0.9967	0.0165	0.9939
15	0.0077	0.9969	0.0166	0.9940

Fonte: Autoria própria (2022)

Como podemos observar no Gráfico 2, a acurácia do treinamento é uma curva crescente, enquanto a acurácia da validação cresce a um ritmo menor e com algumas oscilações.

Gráfico 2 - Acurácia para treinamento com 15 épocas

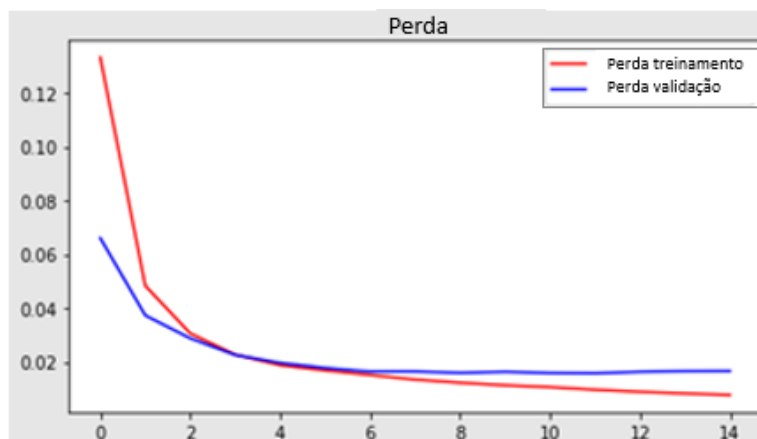


Fonte: Autoria própria (2022)

Isso significa que a cada nova época o modelo se ajusta melhor ao conjunto de imagens de treinamento, mas isso não se reflete da mesma forma nas imagens de validação, o que pode significar que está havendo um super ajuste às imagens de treinamento. Mesmo assim, a partir da segunda época o modelo já apresentou uma acurácia de mais de 99%.

As perdas do modelo durante o treinamento possuem uma curva sempre decrescente, conforme apresentado na Gráfico 3, indicando que ele está aprendendo a cada nova época. Já para as imagens de validação, entre a quarta e sexta época a perda estabiliza na faixa de 0,02, indicando que talvez a execução de mais etapas de treinamento além dessas possam ser desnecessárias.

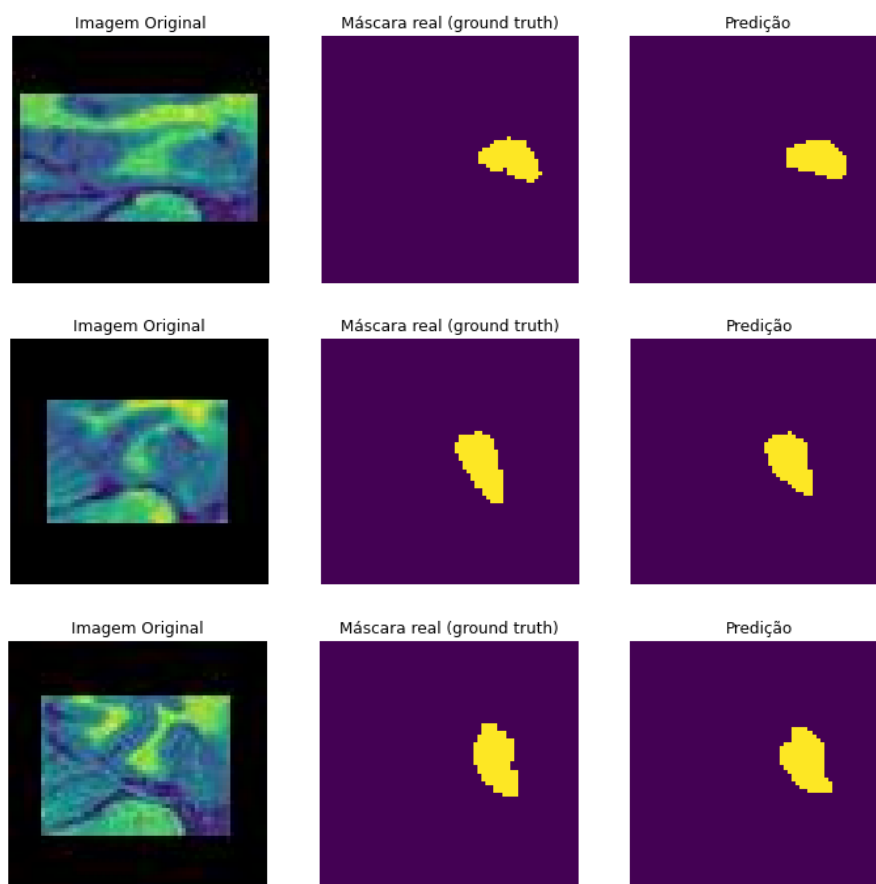
Gráfico 3 - Perda para treinamento com 15 épocas



Fonte: Autoria própria (2022)

Visualmente as imagens geradas pela segmentação do conjunto de imagens de validação foram muito similares às máscaras reais, como podemos ver nos exemplos mostrados na Figura 26.

Figura 26 - Comparação entre a máscara real e a segmentação para treinamento com 15 épocas



Fonte: Autoria própria (2022)

Em muitos casos notou-se que as imagens segmentadas possuíam até contornos mais bem definidos e corrigiam a falta de pixels internos presentes nas máscaras reais, ocasionada pelo processo de binarização da imagem.

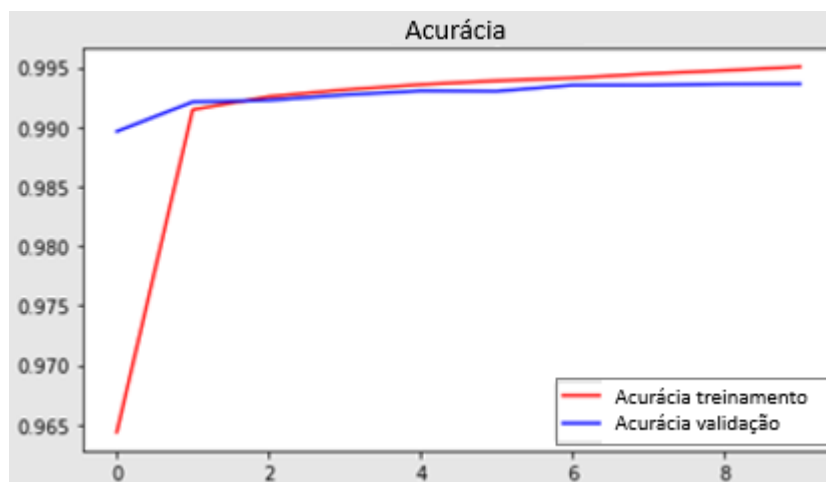
Como no treinamento anterior não houve melhorias após a época 12, foi executado para fins de comparação o modelo para 10 épocas, cujo resultado está apresentado na Tabela 2. Há uma certa aleatoriedade no treinamento, e mesmo que o modelo seja iniciado do mesmo ponto de partida o resultado muito provavelmente será diferente para o mesmo número de ciclos.

Tabela 2 - Dados de treinamento com 10 épocas

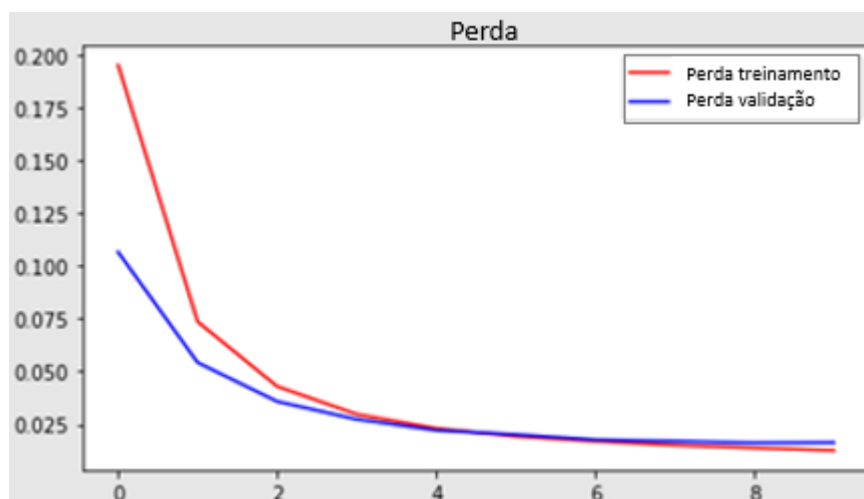
Época	Treinamento		Validação	
	Perda	Acurácia	Perda	Acurácia
1	0.1952	0.9644	0.1067	0.9896
2	0.0736	0.9915	0.0543	0.9921
3	0.0430	0.9926	0.0358	0.9922
4	0.0298	0.9931	0.0274	0.9927
5	0.0231	0.9936	0.0222	0.9930
6	0.0193	0.9939	0.0202	0.9930
7	0.0170	0.9941	0.0177	0.9935
8	0.0151	0.9945	0.0170	0.9935
9	0.0138	0.9947	0.0162	0.9936
10	0.0127	0.9951	0.0164	0.9936

Fonte: Autoria própria (2022)

Da mesma forma que para o treinamento anterior, o Gráfico 4 mostra que para este treinamento a acurácia é crescente para os dados de treinamento e ligeiramente menor para os dados de validação, mantendo-se praticamente constante após as primeiras épocas. A perda foi decrescente tanto para o treinamento quanto para a validação, conforme o Gráfico 5.

Gráfico 4 - Acurácia para treinamento com 10 épocas

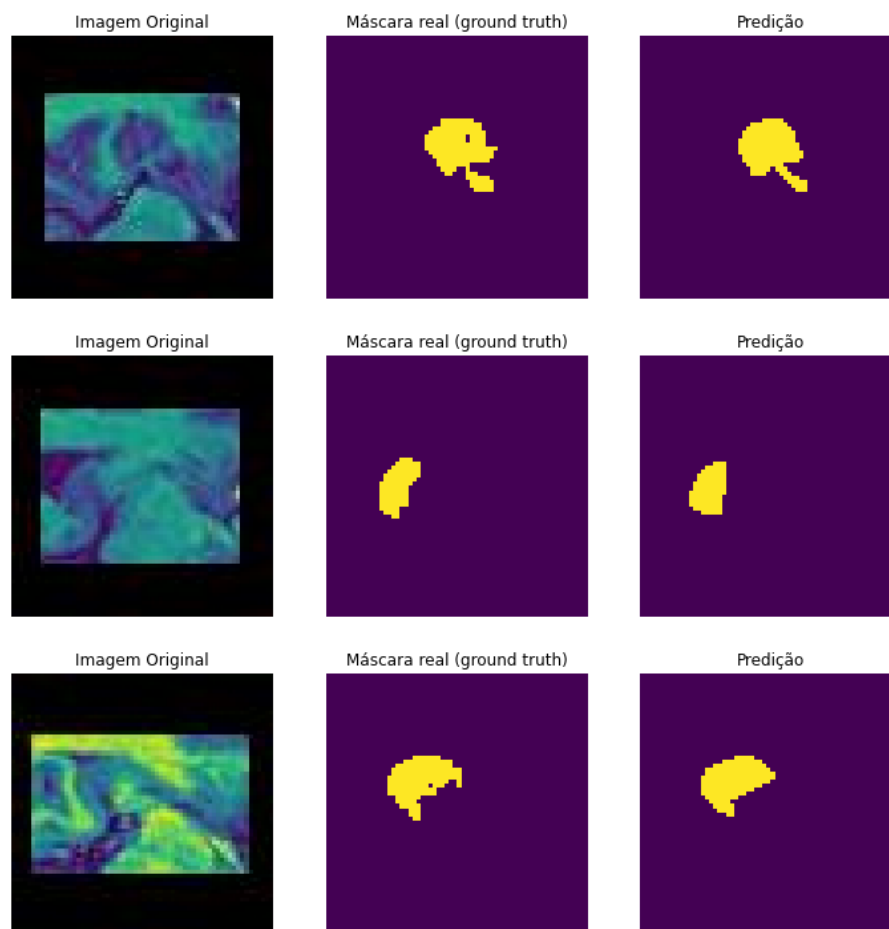
Fonte: Autoria própria (2022)

Gráfico 5 - Perda para treinamento com 10 épocas

Fonte: Autoria própria (2022)

Os gráficos para o treinamento com 10 épocas são similares em forma ao do treinamento com 15 épocas, tanto numa análise visual comparativa entre as máscaras e a validação (Figura 27) quanto qualitativa pelos valores de perda e acurácia (perda na validação de 0.0164 contra 0.0163 para a décima época do primeiro treinamento). Isso significa que a redução do número de épocas não interferiu nos resultados de forma significativa, mas outras métricas seriam necessárias para uma avaliação mais precisa.

Figura 27 - Comparação entre a máscara real e a segmentação para treinamento com 10 épocas

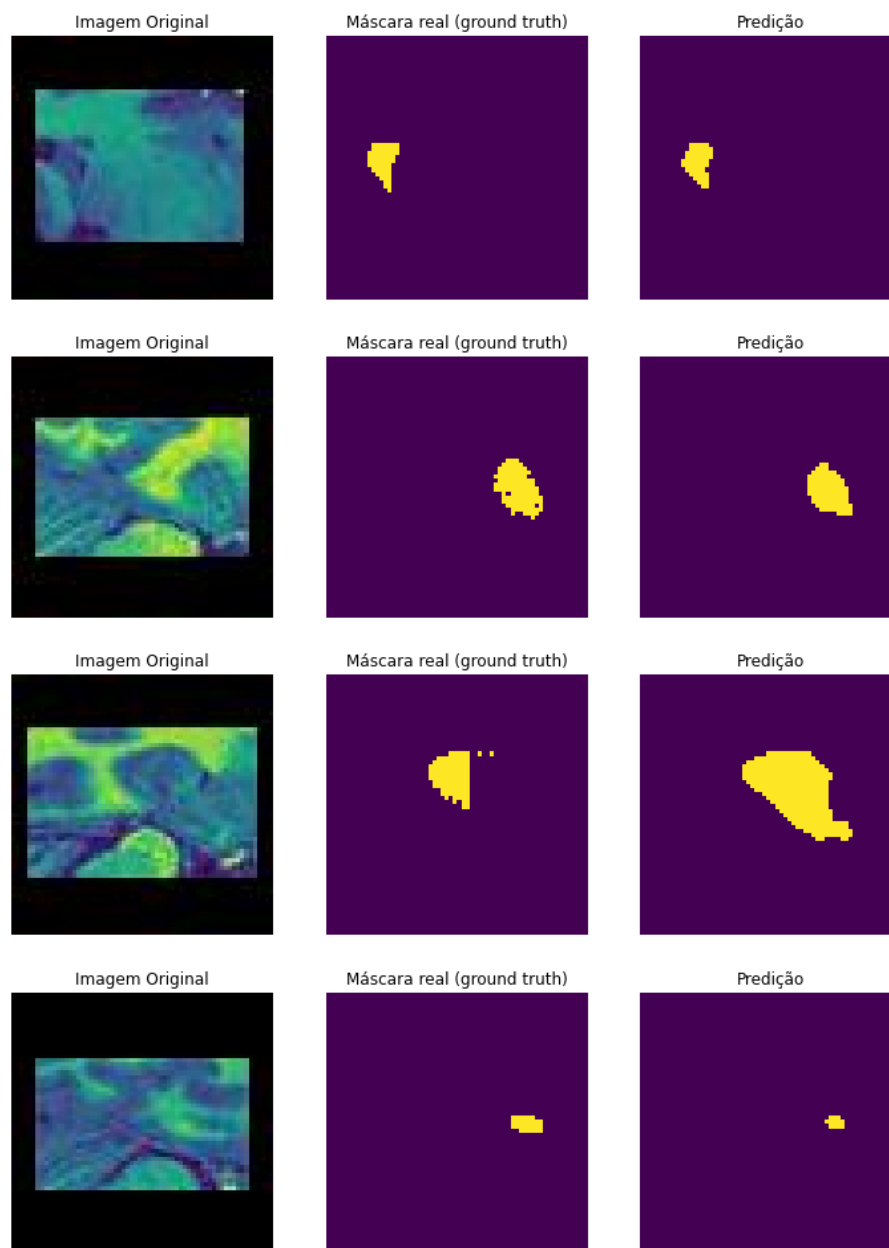


Fonte: Autoria própria (2022)

4.1 Predição

Após treinado, o modelo foi aplicado à imagens inéditas ao código. Considerando a acurácia obtida podemos ter um certo grau de confiança no resultado obtido, mas mesmo assim visualmente a segmentação do hipocampo foi bem sucedida. Vale lembrar que a segmentação é apenas a primeira etapa num processo de diagnóstico que envolve muito mais detalhes e é realizado por profissionais altamente qualificados que passam anos estudando o tema.

Figura 28 - Segmentação de imagens da base de testes para treinamento com 15 épocas

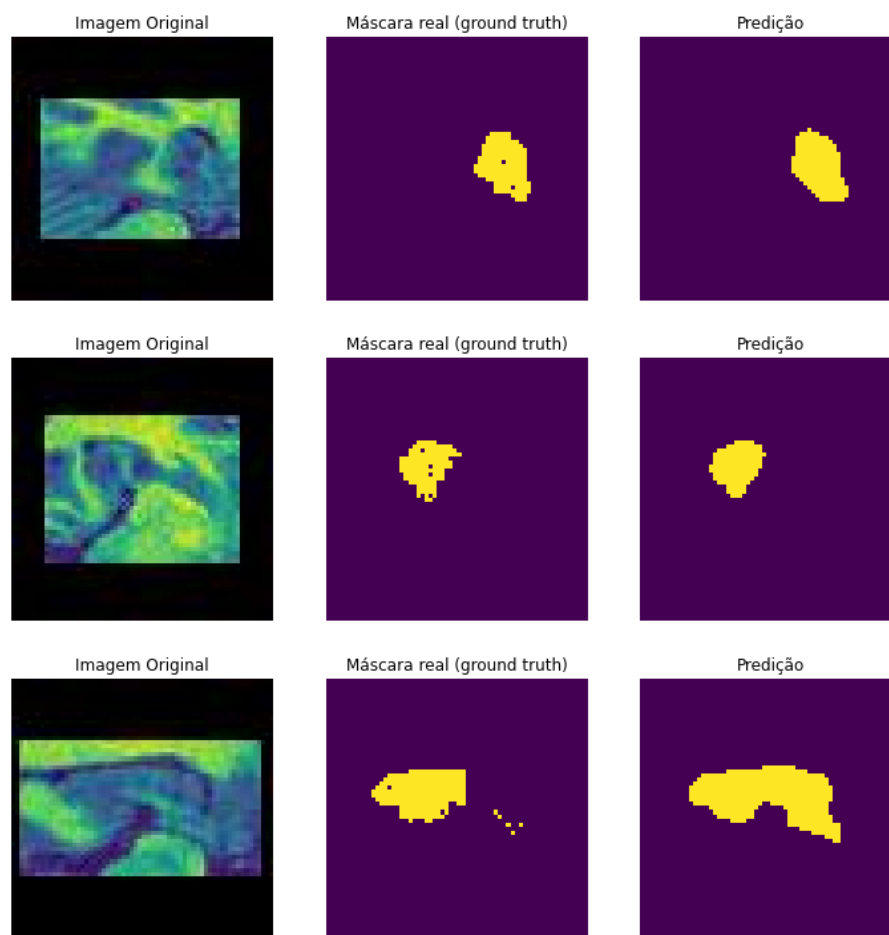


Fonte: Autoria própria (2022)

Analisando as segmentações apresentadas na Figura 28 e comparando com as imagens utilizadas para gerá-las, podemos identificar nas imagens que há uma estrutura na posição indicada pela segmentação. Em algumas imagens, como na última da figura, isso pode ser difícil de observar, mas como a RM fornece imagens de diversas camadas do cérebro deveria-se aplicar o modelo a todas as camadas disponíveis, aumentando ainda mais as chances de identificar o hipocampo.

Aplicando o modelo treinado para 10 épocas podemos notar que o resultado visual (Figura 29) é semelhante ao obtido pelo treinamento com 15 épocas, confirmando a análise feita com base nas métricas de perda.

Figura 29 - Segmentação de imagens da base de testes para treinamento com 10 épocas



Fonte: Autoria própria (2022)

Um treinamento com menos épocas pode economizar tempo, principalmente quando há um grande volume de dados. Apesar disso, o ponto de parada ideal deve ser analisado com muito cuidado, visto que a utilização de poucos ciclos de treinamento pode levar a um resultado insatisfatório. No caso do modelo apresentado, a quantidade ideal aparenta ser em torno de 12 a 14 épocas para o banco de imagens e parâmetros utilizados. Qualquer alteração dos dados de entrada ou parâmetros da CNN podem alterar o treinamento e a quantidade ideal de épocas para um aprendizado satisfatório da rede.

5 CONCLUSÃO

O objetivo inicial deste estudo era elaborar uma rede neural capaz de classificar diferentes estágios de atrofia do hipocampo e associar esse volume à Doença de Alzheimer, mesmo que num período inicial. Para isso seria necessário obter uma base de imagens aberta com dados e informações que permitisse realizar a segmentação e a classificação das imagens nos diferentes estágios do desenvolvimento da DA.

Durante o trabalho, houve muita dificuldade em encontrar uma base que atendesse aos requisitos mínimos, muitas foram consideradas como candidatas nesta pesquisa, porém apresentaram erros e precisaram ser descartadas. Entre essas desqualificadas encontrou-se bases com um número reduzido de imagens ou bases sem a imagem da máscara.

Para ampliar o número de bases a serem consideradas, foram analisados diferentes tipos de arquivos usados em imagens médicas. Nesse momento encontrou-se a base fornecida pela MSD que atendia a maior parte dos requisitos. Ela era composta por arquivos ".nii" que possibilitam a extração das camadas e montagem de uma base auxiliar de imagens ".jpg". Apesar disso, a base da MSD possuía arquivos de máscara, mas não de classificação, o que limitou o escopo do trabalho apenas à segmentação.

Com essa base intermediária montada foi possível iniciar os testes com os modelos base da CNN, porém, como as funções de convolução fornecidas pelas bibliotecas padrão de *Deep Learning* exigem um formato específico, foi necessário um tratamento adicional das imagens. As imagens precisavam ser quadradas. Inicialmente cogitou-se cortar a imagem, porém isso poderia eliminar o hipocampo da máscara e logo a possibilidade foi descartada. Foi então tentado o redimensionamento, mas as distorções causadas por esse tipo de manipulação foram prejudiciais ao aprendizado da rede. Por fim optou-se pela adição da margem preta centralizando a imagem. Essas alterações nos permitiram montar a base que usamos para os resultados finais deste trabalho.

Após o correto tratamento dos dados de entrada o modelo obteve ótimos resultados, com uma acurácia de mais de 99% e perdas que indicam que a rede é capaz de aprender e otimizar seus parâmetros internos. A maior parte das

segmentações obtidas após um treinamento adequado eram muito similares às máscaras fornecidas. Esses resultados transmitem otimismo para o uso dessa estratégia de segmentação no auxílio diagnóstico para a DA, pois a partir da identificação do hipocampo é possível iniciar a análise de suas características, como forma e volume, e classificar um eventual nível de atrofia da estrutura.

Como projetos futuros sugere-se a expansão do modelo para análise de volume, forma, classificação, da rede, aperfeiçoamento do desempenho computacional, alterações na forma da CNN em busca de melhores resultados e adição de outras métricas de validação. Sabe-se que a acurácia e perda são métricas para avaliação da CNN, mas que no caso de segmentação há outras métricas mais apropriadas como a *Intersection over Union* (IoU), que avalia a quantidade de sobreposição de pixels da imagem resultante da segmentação com a máscara original.

O tratamento das imagens mostrou-se como a etapa mais trabalhosa desse processo, mas mesmo assim a equipe se beneficiou muito com a realização desse estudo. Atualmente, imagens médicas não são a área de atuação de nenhum dos integrantes, mas o trabalho com CNN e o tratamento dessas imagens é algo que ultrapassa esse tema e pode ser replicado em diversas áreas.

Considerando todas as dificuldades, o projeto apresentou resultados satisfatórios e espera-se que ele possa contribuir com o maior entendimento da Doença de Alzheimer e no combate a esse mal que atinge cada vez um número maior de pessoas.

REFERÊNCIAS

- AGUIAR-FURUCHO, M.A., PELÁEZ, F.J.R. **Alzheimer's Disease as a Result of Stimulus Reduction in a GABA-A-Deficient Brain: A Neurocomputational Model.** *Neural Plasticity*, vol. 2020, Article ID 8895369, 26 pages, 2020. <https://doi.org/10.1155/2020/8895369>
- AGUIAR-FURUCHO, Mariana. **Modelo neurocomputacional dos estágios iniciais da doença de Alzheimer.** Tese (Doutorado - Programa de Pós-Graduação em Neurociências e Comportamento) - Instituto de Psicologia, Universidade de São Paulo, 2017.
- ALZHEIMER'S ASSOCIATION. 2017 **Alzheimer's Disease Facts and Figures.** Disponível em: <https://www.alz.org/media/images/2017-facts-and-figures.pdf>. Acesso em: 09 ago. 2021.
- ANTONELLI, M., REINKE, A., BAKAS, S. et al. **The Medical Segmentation Decathlon.** *Nature Digital Medicine* (2021). <https://doi.org/10.48550/arXiv.2106.05735>
- ANWAR, S.M., MAJID, M., QAYYUM, A. et al. **Medical Image Analysis using Convolutional Neural Networks: A Review.** *J Med Syst* 42, 226 (2018). <https://doi.org/10.1007/s10916-018-1088-1>
- ASHRAF, A. et al. **Cortical hypermetabolism in MCI subjects: a compensatory mechanism?** *European journal of nuclear medicine and molecular imaging*, v. 42, n. 3, p. 447-458, 2015.
- ASSESSORIA DE COMUNICAÇÃO DO HC, ed. **HC Ganha Setor De Ressonância Magnética.** *Boletim da UFMG*. mar, 2012 Acesso em: 20 nov 2021. <https://www.ufmg.br/boletim/bol1766/5.shtml>.
- ATALOGLU, D., DIMOU, A., ZARPALAS, D. et al. **Fast and Precise Hippocampus Segmentation Through Deep Convolutional Neural Network Ensembles and Transfer Learning.** *Neuroinform* 17, 563–582 (2019). <https://doi.org/10.1007/s12021-019-09417-y>
- BISAZ, R., TRAVAGLIA, A., ALBERINI, C. M. (2014). **The Neurobiological Bases of Memory Formation: From Physiological Conditions to Psychopathology.** *Psychopathology*, 47(6), 347–356.
- BRAINER. **The NIFTI file format.** 2012. Disponível em: <https://brainer.org/2012/09/23/the-nifti-file-format/>. Acesso em: 20 jun. 2022.
- CASANOVA, D. **Identificação de espécies vegetais por meio da análise de textura foliar.** Tese (Doutorado) — Universidade de São Paulo, 2008.
- CAVALCANTE COUTINHO, M. **Segmentação de Imagens de lesões de pele usando a rede neural convolucional U-net.** 2019. Projeto de Graduação.
- CLEMENTE, C.E.T.; PINTO, R.A.S.R. **Revisão e atualização dos biomarcadores no líquido para doença de alzheimer.** *Revista Brasileira de Neurologia e Psiquiatria*, v. 21, p.1-18, 2017
- COSTA, M. G. F.; **Redes Neurais Convolucionais na Saúde.** *Journal of Health Informatics*. Disponível em: <http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/viewFile/615/323>Acesso em: 20. nov. 2021.
- DATA SCIENCE ACADEMY. **Deep Learning Book**, 2021. Disponível em: <https://www.deeplearningbook.com.br/>. Acesso em: 10 outubro. 2021
- DATA SCIENCE ACADEMY. **Segmentação de Imagens Médicas com Deep Learning**, 2021. Disponível em: <https://blog.dsacademy.com.br/segmentacao-de-imagens-medicadas-com-deep-learning/>. Acesso em: 31 maio. 2022

- DINCER, B. **MRI Hippocampus-Alzheimer Segmentation/ALL Process**. ago, 2021. Disponível em: <https://www.kaggle.com/brsdincer/mri-hippocampus-alzheimer-segmentation-all-process/notebook>. Acesso em: 12 dez. 2021.
- DECAZES P, HINAULT P, VERESEZAN O, THUREAU S, GOUEL P, VERA P. **Trimodality PET/CT/MRI and Radiotherapy: A Mini-Review**. *Front Oncol*. 2021;10:614008. Published 2021 Feb 4. doi:10.3389/fonc.2020.614008
- DHIKAV, V., ANAND, K. (2012). **Hippocampus in health and disease: An overview**. *Annals of Indian Academy of Neurology*, 15(4), 239.
- GOOGLE RESEARCH. **Perguntas frequentes**. 2022. Disponível em <https://research.google.com/colaboratory/intl/pt-BR/faq.html>. Acesso em: 25 de junho de 2022.
- GRAHAM, R. N. J., PERRISS, R. W., SCARSBROOK, A. F. (2005). **DICOM demystified: A review of digital file formats and their use in radiological practice**. *Clinical Radiology*, 60(11), 1133–1140. doi:10.1016/j.crad.2005.07.003
- GRANATYR, J. **Segmentação de Imagens com Python de A à Z: Segmentação Semântica - Microscópicas** (Programa de computador), 2022.
- HAGE, M. C. F. N. S., IWASAKI, M. (2009). **Imagem por ressonância magnética: princípios básicos**. *Ciência Rural*, 39(4), 1275–1283. doi:10.1590/s0103-84782009005000041
- HAMPEL, H.; PRVULOVIC, D.; TEIPEL, S.; JESSEN, F.; LUCKHAUS, C.; FRÖLICH, L. **The future of alzheimer's disease: the next 10 years**. *Progress in Neurobiology*, v. 95, p.718-728, 2011
- HASHEMI, M. **Enlarging smaller images before inputting into convolutional neural network: zero-padding vs. interpolation**. *Journal of Big Data* 6, 98. 2019. <https://doi.org/10.1186/s40537-019-0263-7>
- HAYKIN, S. **A comprehensive foundation**. *Neural Networks*, v. 2, n. 2004, p. 41, 2004.
- HÄNSCH, A; MOLTZ, J. H.; GEISLER, B; ENGEL, C.; KLEIN, J.; GENGGHI, A.; SCHREIER, J.; MORGAS, T.; HAASB, B. **Hippocampus segmentation in CT using deep learning: impact of MR versus CT-based training contours**. *Journal of Medical Imaging*. v. 7. nov/dez 2020.
- HOFESMANN, E. **IoU a better detection evaluation metric**. *Towards Data Science*, 2020. Disponível em: <https://towardsdatascience.com/iou-a-better-detection-evaluation-metric-45a511185be1>.
- JAGREET. **Overview of Artificial Neural Networks and its Applications**. 2017. Disponível em <<https://www.xenonstack.com/blog/overview-of-artificial-neural-networks-and-its-applications>>. Acesso em: 20 de novembro de 2021.
- KERAS. **About Keras**. 2022. Disponível em <<https://keras.io/about/>>. Acesso em: 20 de junho de 2022.
- KUMAR, A., SINGH, A., EKAVALI. (2015). **A review on Alzheimer's disease pathophysiology and its management: an update**. *Pharmacological Reports*, 67(2), 195–203. doi:10.1016/j.pharep.2014.09.004
- LI, H., HABES, M., WOLK, D., FAN, Y. **A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal MRI**. Philadelphia, PA. 2019.
- MARENGONI, M.; STRINGHINI, S. **Tutorial: Introdução à Visão Computacional usando OpenCV**. *Revista de Informática Teórica e Aplicada*, v. 16, n. 1, p. 125–160, mar. 2010. ISSN 21752745. Disponível em:

https://seer.ufrgs.br/rita/article/view/rita_v16_n1_p125. Acesso em: 20 de novembro de 2021.

MASSUCATTO, J.D.P.; **Aplicação dos conceitos de Redes Neurais Convolucionais na classificação de imagens de folhas**. Pato Branco, Brasil. 2018.

MASTERS, C. L. et al. **Alzheimer's disease**. Nature Reviews Disease Primers, v. 1, art. n. 15056, 2015.

MAZZOLA, A. A. (2015). **Ressonância magnética: princípios de formação da imagem e aplicações em imagem funcional**. Revista Brasileira De Física Médica, 3(1), 117–129. <https://doi.org/10.29384/rbfm.2009.v3.n1.p117-129> Acesso em: 27 out. 2021.

MAZUR, M. **Step by step Backpropagation Example**. 2015. <https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/>. Acesso em: 20 nov. 2021.

MCCULLOCH, W. S.; PITTS, W. **A logical calculus of the ideas immanent in nervous activity**. The bulletin of mathematical biophysics, Springer, v. 5, n. 4, p. 115–133, 1943.

MILDENBERGER, P., EICHELBERG, M., MARTIN, E. (2001). **Introduction to the DICOM standard**. European Radiology, 12(4), 920–927. doi:10.1007/s003300101100

MOLINUEVO, J. L.; AYTÓN, S.; BATRLA, R.; BEDNAR, M.; M., BITTNER, T.; CUMMINGS, J.; BLENNOW, K. **Amyloid blood biomarker detects alzheimer's disease**. Embo Molecular Medicine, v. 10, p.1-24, 2018, nov. 2021.

MORENO, O., GONZÁLEZ, J. Y BALLESTEROS J. (2005). **Transformación de archivos DICOM a formatos XML, JPEG y PNG**. Revista Científica,21, 71-80.

NUNES, I.; DANILO, S.; FLAUZINO, R. **Redes Neurais Artificiais para engenharia e ciências aplicadas**. [S.l.]: ArtLiber, 2010.

PAVLOVSKY, V. **Introduction To Convolutional Neural Networks**. 2017. Disponível em: <https://www.vaetas.cz/blog/intro-convolutional-neural-networks>. Acesso em: 07 nov 2021.

PYCAD. **What is the difference between Dicom and Nifti images?** 2021. Disponível em: <https://pycad.co/what-is-the-difference-between-dicom-and-nifti-images/>. Acesso em: 20 jun. 2022.

PYKETT, I. L., NEWHOUSE, J. H., BUONANNO, F. S., BRADY, T. J., GOLDMAN, M. R., KISTLER, J. P., POHOST, G. M. (1982). **Principles of nuclear magnetic resonance imaging**. Radiology, 143(1), 157–168. doi:10.1148/radiology.143.1.7038763

RAJAN, K. B., WEUVE, J., BARNES, L. L., MCANINCH, E. A., WILSON, R. S., EVANS, D. A. (2021). **Population estimate of people with clinical Alzheimer's disease and mild cognitive impairment in the United States (2020–2060)**. Alzheimer's & Dementia. doi:10.1002/alz.12362

RASCHKA, S.; MIRJALILI, V. **Python Machine Learning**. [S.l.]: Packt Publishing Ltd, 2017. Acesso em: 18. nov. 2021.

RIVA, M. **Interpretation of Loss and Accuracy for a Machine Learning Model**. 2021. Disponível em: <<https://www.baeldung.com/cs/ml-loss-accuracy>>. Acesso em: 28 jun. 2022.

RONNEBERGER, O.; FISCHER, P; BROX, T. U-Net: **Convolutional networks for biomedical image segmentation**. In: International Conference on Medical

image computing and computer-assisted intervention. Springer, Cham, p. 234-241, 2015.

ROSENBLATT, F. **The perceptron: A probabilistic model for information storage and organization in the brain.** *Psychological review*, American Psychological Association, v. 65, n. 6, p. 386, 1958

RUSSELL, Stuart; NORVIG, Peter. **Inteligência Artificial.** 3. ed. Rio de Janeiro: Elsevier, 2013. 1320 p.

SCOVILLE, W. B., MILNER, B. (1957). **Loss Of Recent Memory After Bilateral Hippocampal Lesions.** *Journal of Neurology, Neurosurgery & Psychiatry*, 20(1), 11–21.

SILVA, E.C. **Estudo Da Técnica De Ressonância Magnética Para Algumas Aplicações Voltadas À Neurologia E À Neurociência.** 37 f. Trabalho de conclusão de curso (licenciatura e bacharelado - Física) - Universidade Estadual Paulista, Instituto de Geociências e Ciências Exatas, 2011. Disponível em: <<http://hdl.handle.net/11449/121195>>.

SILVA, Fabrício.M. D.; LENZ, Maikon. L.; FREITAS, Pedro.H. C.; SANTOS, Sidney.C.Bispo. D. **Inteligência artificial:** Grupo A, 2019. 9788595029392. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595029392/>. Acesso em: 12 out. 2021.

SIMPSON, A. L., ANTONELLI, M., BAKAS, S. et al. **A large annotated medical image dataset for the development and evaluation of segmentation algorithms.** (2019). <https://doi.org/10.48550/arXiv.1902.09063>

SUPERDATASCIENCE. **Deep Learning A-Z™: Download Practice Datasets.** 2018. Disponível em: <https://www.superdatascience.com/deep-learning>. Acesso em: 07 nov. 2021.

SWARBRICK, S.; WRAGG, N.; GHOSH, S.; STOLZING, A. **Systematic review of miRNA as biomarkers in alzheimer's disease.** *Molecular Neurobiology*, v. 56, p.6156-6167, 2019.

VARMA, D. (2012). Managing DICOM images: **Tips and tricks for the radiologist.** *Indian Journal of Radiology and Imaging*, 22(1), 4. doi:10.4103/0971-3026.95396.

WU, J. **Introduction to convolutional neural networks.** National Key Lab for Novel Software Technology Nanjing University, China, 2017.

ZHANG, J. **U-Net - Line by line explanation.** (2019). Disponível em <https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5>. Acesso em: 26 jun. 2022.