

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Adailton dos Santos Holanda

**Segmentação de Acidente Vascular Cerebral Hemorrágico com
Abordagem Híbrida Baseada em Redes Grafo Convolucionais**

Fortaleza – CE

18 de novembro de 2025

Adailton dos Santos Holanda

Segmentação de Acidente Vascular Cerebral Hemorrágico com Abordagem Híbrida Baseada em Redes Grafo Convolucionais

Dissertação submetida ao Programa de Pós-Graduação em Ciência da Computação do Instituto Federal de Educação, Ciência e Tecnologia do Ceará, como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

Linha de Pesquisa: Computação Aplicada

Orientador: Prof. Dr. Pedro Pedrosa Rebouças Filho

Coorientador: Prof. Dr. Roger Moura Sarmiento

Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PRPI)

Programa de Pós-Graduação em Ciência da Computação (PPGCC)

Fortaleza – CE

18 de novembro de 2025

Dados Internacionais de Catalogação na Publicação
Instituto Federal do Ceará - IFCE
Sistema de Bibliotecas - SIBI
Ficha catalográfica elaborada pelo SIBI/IFCE, com os dados fornecidos pelo(a) autor(a)

- H722s Holanda, Adailton dos Santos.
Segmentação de Acidente Vascular Cerebral Hemorrágico com Abordagem Híbrida Baseada em Redes Grafo Convolucionais / Adailton dos Santos Holanda. - 2025.
101 f. : il. color.
- Dissertação (Mestrado) - Instituto Federal do Ceará, Mestrado em Ciência da Computação, Campus Fortaleza, 2025.
Orientação: Prof. Dr. Pedro Pedrosa Rebouças Filho.
Coorientação: Prof. Dr. Roger Moura Sarmento.
1. Visão Computacional. 2. Machine Learning. 3. Redes Grafo Convolucionais . 4. Acidente Vascular Cerebral. I. Título.

CDD 004



Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)
Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PRPI)
Programa de Pós-Graduação em Ciência da Computação (PPGCC)

Adailton dos Santos Holanda

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, sendo aprovada pela Coordenação do Programa de Pós-Graduação em Ciência da Computação do Instituto Federal de Educação, Ciência e Tecnologia do Ceará e pela banca examinadora:

**Orientador: Prof. Dr. Pedro Pedrosa Rebouças
Filho**

Instituto Federal de Educação, Ciência e Tecnologia do
Ceará (IFCE)

Coorientador: Prof. Dr. Roger Moura Sarmiento
Instituto Federal de Educação, Ciência e Tecnologia do
Ceará (IFCE)

Prof^a. Dr^a. Suane Pires Pinheiro da Silva
Instituto Federal de Educação, Ciência e Tecnologia do
Ceará (IFCE)

Prof. Dr. Pedro Henrique Feijó de Sousa
Universidade Federal do Ceará (UFC)

Fortaleza – CE
18 de novembro de 2025

*Este trabalho é dedicado aos meus pais José Sergio e Antonia,
que sempre me apoiaram com amor, valores e ensinamentos.
Sem vocês, nada disso seria possível.*

...

*À minha esposa Thalita,
pela paciência, carinho e por caminhar ao meu lado
em todos os momentos dessa jornada.
Seu apoio foi essencial para esta conquista.*

Agradecimentos

Realizar este mestrado foi uma etapa marcante na minha trajetória acadêmica e pessoal. Ao longo do caminho, enfrentei desafios que exigiram dedicação e perseverança. Felizmente, pude contar com o apoio de pessoas importantes, que contribuíram de forma significativa para que esse percurso fosse possível.

Ao meu orientador, Prof. Dr. Pedro Pedrosa, agradeço profundamente pela orientação, pelas sugestões precisas e pelo apoio constante. Sua dedicação, mesmo diante de tantos compromissos, foi fundamental para que eu pudesse crescer academicamente.

Ao meu coorientador, Prof. Dr. Róger Sarmiento, agradeço pelas contribuições e pelo constante apoio. Sua disposição em colaborar foi essencial para o desenvolvimento deste trabalho.

Aos meus pais, José Sergio e Antonia, minha eterna gratidão. Obrigado por nunca medirem esforços para me apoiar, mesmo quando as dificuldades pareciam maiores que os sonhos. A educação e os valores que vocês me transmitiram são a base de tudo que conquistei até aqui.

À minha esposa Thalita, minha companheira de todos os dias, meu alicerce e inspiração. Obrigado por estar ao meu lado em cada momento, por compreender minhas ausências, incentivar minhas escolhas e compartilhar comigo os dias bons e ruins. Sem você, nada disso teria o mesmo sentido.

Agradeço também aos professores, colegas e amigos que contribuíram, direta ou indiretamente, para a construção deste trabalho e para o meu amadurecimento pessoal e profissional.

E, acima de tudo, agradeço a Deus, por ter me sustentado nos momentos de fraqueza, por me conceder sabedoria quando eu não sabia o caminho, e por me dar forças para seguir mesmo diante das incertezas.

Muito obrigado a todos!

“Porque dEle, e por Ele, e para Ele, são todas as coisas; glória, pois, a Ele eternamente.

Amém”

(Romanos 11:36)

Resumo

O Acidente Vascular Cerebral (AVC) representa um grave problema de saúde pública global, sendo responsável por elevados índices de morbidade e mortalidade. Segundo a Organização Mundial da Saúde (OMS), aproximadamente 15 milhões de pessoas sofrem AVC anualmente, desse total, 5 milhões morrem e outros 5 milhões ficam permanentemente incapacitados. Diante desse cenário, a reabilitação precoce, iniciada ainda no ambiente hospitalar, desempenha um papel fundamental na recuperação funcional, contribuindo para que os sobreviventes alcancem o maior nível possível de independência ao longo do tempo. A Tomografia Computadorizada (TC) é o principal exame de imagem utilizado na rotina clínica para identificação e avaliação da gravidade do AVC, principalmente em sua fase aguda. Neste contexto, sistemas baseados em inteligência artificial têm se mostrado promissores para auxiliar na detecção e caracterização das lesões cerebrais associadas ao AVC. Este trabalho propõe uma abordagem híbrida, que combina redes convolucionais (CNN) com Redes Grafo Convolucionais (GCN), visando realizar a segmentação automática de regiões afetadas por AVC hemorrágico em imagens de TC do crânio. Inicialmente, um modelo CNN é empregado para extrair atributos representativos de cada imagem, os quais são utilizados para a construção de um grafo que modela a similaridade entre os exames. Em seguida, uma GCN é utilizada para realizar a propagação de informações entre os nós do grafo, permitindo a classificação dos exames nas classes normal, isquêmico e hemorrágico. Nos casos classificados como hemorrágicos, é aplicada uma segunda etapa de segmentação, baseada em GCN, com foco na delimitação precisa da região lesionada. Para avaliação da proposta, foi utilizada uma base de dados composta por exames de TC rotulados manualmente por especialistas. Os resultados obtidos demonstraram desempenho superior em relação aos trabalhos da literatura, com destaque para a acurácia de 99,50% na classificação e coeficiente de Dice de 95,98% na segmentação. As demais métricas de segmentação obtidas foram: Jaccard de 92,54%, sensibilidade de 95,94%, especificidade de 95,94% e precisão de 95,88%, evidenciando a robustez da abordagem proposta. Esses resultados indicam que a abordagem híbrida proposta é eficaz para a identificação e segmentação de AVC hemorrágico, apresentando grande potencial para integrar sistemas de apoio ao diagnóstico clínico.

Palavras-chave: Acidente vascular cerebral, Tomografia computadorizada, Redes grafo convolucionais, Classificação de AVC, Segmentação de AVC.

Abstract

Stroke represents a serious global public health issue, being responsible for high rates of morbidity and mortality. According to the World Health Organization (WHO), approximately 15 million people suffer a stroke each year. Of these, 5 million die, and another 5 million are left permanently disabled. In this scenario, early rehabilitation—initiated still in the hospital setting—plays a crucial role in functional recovery, helping survivors achieve the highest possible level of independence over time. Computed Tomography (CT) is the main imaging exam used in clinical practice for identifying and assessing the severity of stroke, especially in its acute phase. In this context, artificial intelligence-based systems have shown promise in assisting with the detection and characterization of brain lesions associated with stroke. This work proposes a hybrid approach that combines Convolutional Neural Networks (CNNs) with Graph Convolutional Networks (GCNs) to automatically segment regions affected by hemorrhagic stroke in cranial CT images. Initially, a CNN model is employed to extract representative features from each image, which are then used to construct a graph modeling the similarity between exams. Subsequently, a GCN is applied to propagate information across the graph nodes, enabling the classification of exams into three clinical categories: normal, ischemic, and hemorrhagic. For cases classified as hemorrhagic, a second segmentation stage based on GCNs is performed to delineate the affected region precisely. To evaluate the proposed method, a dataset consisting of manually annotated CT scans by specialists was used. The results demonstrated superior performance compared to existing methods in the literature, achieving an accuracy of 99.50% in classification and a Dice coefficient of 95.98% in segmentation. Other segmentation metrics included: Jaccard index of 92.54%, sensitivity of 95.94%, specificity of 95.94% and precision of 95.88%, highlighting the robustness of the proposed approach. These results indicate that the proposed hybrid method is effective for identifying and segmenting hemorrhagic stroke and shows strong potential for integration into clinical decision support systems.

Keywords: Stroke, Computed Tomography, Graph Convolutional Networks, Stroke Classification, Stroke Segmentation.

Lista de ilustrações

Figura 1 – Representação esquemática do AVC isquêmico. Fonte: Caplan (2016). . .	31
Figura 2 – Representação esquemática do AVC hemorrágico. Fonte: Caplan (2016). . .	33
Figura 3 – Modelo inicial do tomógrafo criado por Hounsfield, desenvolvido nos Laboratórios Centrais de Pesquisa da EMI, em Londres. Fonte: Buzug (2009).	35
Figura 4 – Ilustração da escala de Hounsfield. Fonte: Freitas (2013).	36
Figura 5 – Estrutura de um neurônio artificial utilizada no modelo Perceptron. Fonte: Haykin (2009).	49
Figura 6 – Arquitetura de uma rede MLP. Fonte: Haykin (2009).	51
Figura 7 – Diagrama ilustrativo do processo de retropropagação. Fonte: Silva, Spatti e Flauzino (2010).	52
Figura 8 – Comparação entre bloco tradicional e bloco residual com conexão de atalho na arquitetura ResNet. Fonte: Zhang et al. (2023).	57
Figura 9 – Representação de grafo não direcionado com suas respectivas arestas e vértices. Fonte: Ma e Tang (2021).	59
Figura 10 – Processo de filtragem em grafos para atualização das representações nodais. Fonte: Ma e Tang (2021).	67
Figura 11 – Operação de pooling em grafos para redução da dimensionalidade estrutural e nodal. Fonte: Ma e Tang (2021).	68
Figura 12 – Processo de decomposição, modulação espectral e reconstrução de sinais em grafos. Fonte: Ma e Tang (2021).	70
Figura 13 – Fluxo da Metodologia Proposta para Classificação e Segmentação de AVCs em Imagens de TC. Fonte: Autor.	74
Figura 14 – Exemplos de exames de TC pertencentes a cada classe do dataset. Fonte: Autor.	75
Figura 15 – Etapas do pré-processamento aplicadas às imagens de TC. Fonte: Autor. . .	76
Figura 16 – Modificações na ResNet para extração de atributos. Fonte: Autor. . . .	77
Figura 17 – Extração de vetores de atributos para cada imagem do batch utilizando a ResNet modificada. Fonte: Autor.	77
Figura 18 – Geração do mapa de atributos a partir da ResNet para segmentação. Fonte: Autor.	79
Figura 19 – Criação do grafo para classificação a partir das características extraídas das imagens. Fonte: Autor.	80
Figura 20 – Criação do grafo para segmentação com base nos mapas de características extraídos. Fonte: Autor.	81

Figura 21 – Fluxo de Classificação com GCN a partir dos atributos extraídos pela ResNet. Fonte: Autor.	82
Figura 22 – Fluxo do processo de segmentação de AVCs hemorrágicos com uso de GCNs. Fonte: Autor.	83
Figura 23 – Fluxo da segmentação utilizando GCN. Fonte: Autor.	84
Figura 24 – Arquitetura do modelo GCN para segmentação de lesões hemorrágicas. Fonte: Autor.	84
Figura 25 – Matriz de confusão para a tarefa de classificação em três categorias. Fonte: Autor.	87
Figura 26 – Evolução da função de perda e do coeficiente de Dice ao longo das épocas para $k = 10$, distância euclidiana. Fonte: Autor.	91
Figura 27 – Curvas de evolução da função de perda e do Dice score durante o treinamento. Fonte: Autor.	92

Lista de tabelas

Tabela 1	–	Resumo comparativo dos estudos sobre classificação de AVC em imagens médicas.	39
Tabela 2	–	Resumo comparativo dos estudos sobre segmentação de AVC hemorrágico em imagens médicas.	42
Tabela 3	–	Estudos que aplicam GCNs em tarefas de classificação e segmentação em imagens médicas.	47
Tabela 4	–	Comparação dos Resultados de Classificação com Trabalhos Relacionados.	88
Tabela 5	–	Comparação das métricas por número de vizinhos e tipo de distância. .	90
Tabela 6	–	Comparação entre Dice, Jaccard e Tempo por número de vizinhos e tipo de distância.	91
Tabela 7	–	Comparação de Métricas de Segmentação: Sensibilidade, Especificidade e Acurácia.	93
Tabela 8	–	Comparação de Métricas de Sobreposição e Precisão: Dice, Jaccard e Precisão.	93

Lista de abreviaturas e siglas

ABTD	Analysis of Brain Tissue Densities
ACR	American College of Radiology
ASIR	Reconstrução Estatística Iterativa Adaptativa
AVC	Acidente Vascular Cerebral
AVCh	Acidente Vascular Cerebral Hemorrágico
AVCi	Acidente Vascular Cerebral Isquêmico
CAD	Diagnóstico Auxiliado por Computador
CNN	Redes Neurais Convolucionais
DGC	Grafo com Convoluções Dinâmicas
DICOM	Digital Imaging and Communications in Medicine
DL	Deep Learning
FN	Falsos Negativos
FP	Falsos Positivos
GACM	Método de Contorno Ativo Geodésico
GCN	Redes Grafo Convolucionais
GNN	Redes Grafo Neurais
HD	Distância de Hausdorff
HIMS	Himalayan Institute of Medical Sciences
IA	Inteligência Artificial
ICH	Hemorragia Intracerebral
IRIS	Reconstrução Iterativa no Espaço da Imagem
KNN	k-Nearest Neighbors
LSBRD	Level Set Based on Radiological Densities

MBIR	Reconstrução Iterativa Baseada em Modelo
MC	Matriz de Confusão
ML	Machine Learning
MLP	Rede Perceptron Multicamadas
NAM	Academia Nacional de Medicina
NEMA	National Electrical Manufacturers Association
OMS	Organização Mundial da Saúde
PET	Tomografia por Emissão de Pósitrons
PDI	Processamento Digital de Imagens
RBB	Residual Building Block
ResNet	Redes Neurais Residuais
RM	Ressonância Magnética
ROI	Região de Interesse
SAH	Hemorragia Subaracnoidea
SCM	Matriz de Coocorrência Estrutural
SVD	Valores Singulares
TC	Tomografia Computadorizada
VC	Visão Computacional
VN	Verdadeiros Negativos
VP	Verdadeiros Positivos

Sumário

1	INTRODUÇÃO	25
1.1	Descrição do Problema	26
1.2	Motivação	28
1.3	Objetivos	29
1.3.1	Objetivo Geral	29
1.3.2	Objetivos Específicos	29
1.4	Desafios	29
1.5	Produção Científica	30
1.6	Organização da Dissertação	30
2	ACIDENTE VASCULAR CEREBRAL	31
2.1	AVC Isquêmico	31
2.2	AVC Hemorrágico	32
2.3	Tratamento e Recomendações	33
2.4	Tomografia Computadorizada	35
2.5	Estado da Arte	37
2.5.1	Classificação	37
2.5.2	Segmentação	40
2.5.3	Redes Grafo Convolucionais	43
3	REDES NEURAIS	49
3.1	Perceptron Simples	49
3.2	Perceptron Multicamadas	51
3.3	Redes Neurais Convolucionais	54
3.4	Redes Neurais Residuais	56
4	REDES GRAFO NEURAIS	59
4.1	Conceitos Básicos	59
4.1.1	Propriedades e Medidas	60
4.2	Teoria Espectral de Grafos	63
4.3	Grafos Complexos	64
4.4	Fundamentos das Redes Grafo Neurais	66
4.5	Filtros em Grafos	68
4.5.1	Filtragem Espectral em Grafos	68
4.5.2	Filtragem Espacial em Grafos	70
4.6	Convolução de Chebyshev	71

4.7	Tarefas Computacionais em Grafos	72
5	METODOLOGIA	74
5.1	Base de Dados	74
5.2	Pré-Processamento	75
5.3	Extração de Atributos	76
5.3.1	Extração de Atributos Para a Etapa de Classificação	76
5.3.2	Extração de Atributos Para a Etapa de Segmentação	78
5.4	Criação do Grafo	79
5.4.1	Criação do Grafo Para a Etapa de Classificação	80
5.4.2	Criação do Grafo Para a Etapa de Segmentação	80
5.5	Classificação	81
5.6	Segmentação	83
5.7	Avaliação	84
6	RESULTADOS E DISCUSSÕES	87
6.1	Classificação de AVC's	87
6.2	Segmentação de AVC Hemorrágico	89
6.3	Comparação dos Resultados com Trabalhos Relacionados	92
7	CONCLUSÕES E TRABALHOS FUTUROS	95
7.1	Trabalhos Futuros	96
	REFERÊNCIAS	97

1 Introdução

A Inteligência Artificial (IA) tem evoluído rapidamente nas últimas décadas (BARUA; AHMED; BEGUM, 2023), e suas aplicações podem ser encontradas em uma ampla variedade de áreas, como saúde, robótica, varejo e mercado financeiro (SHEHAB et al., 2022). Na área da saúde, destaca-se por contribuir com a melhoria de sistemas clínicos, gerenciamento de informações, registros de pacientes e apoio à detecção e tratamento de doenças (KUMAR et al., 2023). Nesse contexto, muitas pesquisas têm concentrado esforços no desenvolvimento de soluções para o diagnóstico de doenças por meio de sistemas inteligentes, os quais têm se mostrado eficazes na identificação de diversas enfermidades, como câncer de pele, doenças hepáticas, cardiovasculares, Alzheimer, entre outras, cuja detecção precoce é essencial para um tratamento mais eficiente (KUMAR et al., 2023).

Na área da IA, os modelos de *Deep Learning* (DL) têm obtido excelentes resultados quando aplicados à análise de imagens médicas, demonstrando alta capacidade de aprendizado de padrões em dados de alta dimensionalidade, como os obtidos por Ressonância Magnética (RM) e Tomografia Computadorizada (TC) (YAN et al., 2022). Além disso, a combinação desses modelos com técnicas de Processamento Digital de Imagens (PDI) tem contribuído significativamente para o avanço da Visão Computacional (VC) aplicada à área da saúde.

Os exames por imagem contém uma grande quantidade de informações que são fundamentais para a identificação de patologias e, conseqüentemente, para o tratamento de diversas doenças. Estima-se que 90% dos dados gerados na área da saúde sejam compostos por imagens, sendo, portanto, consideradas a principal fonte de dados para intervenções e análises médicas. Diferentes modalidades de imagem, como TC, ultrassonografia, radiografia por raios X, RM e imagens patológicas, são comumente utilizadas no diagnóstico por imagem (PARVAIZ et al., 2023).

A RM e a TC são métodos bastante utilizados para avaliação clínica por imagem, que permitem a visualização de estruturas internas do corpo humano. A RM é não invasiva, livre de radiação e apresenta excelente contraste entre tecidos, sendo preferencialmente utilizada na análise de diferentes sistemas orgânicos (FLORKOW et al., 2022). A TC, por sua vez, é mais indicada para a visualização de estruturas ósseas, por oferecer imagens de alta resolução com contraste baseado na radiodensidade, o que favorece a identificação de tecidos ósseos. Sua alta resolução temporal também a torna eficaz na obtenção de imagens de múltiplos órgãos, apesar de sua menor sensibilidade na diferenciação de tecidos moles (RAHMAN, 2023). Diversas ferramentas automáticas ou semiautomáticas têm sido desenvolvidas especificamente para imagens de TC, possibilitando, por exemplo,

reconstruções tridimensionais valiosas (FLORKOW et al., 2022).

Esses exames podem ser analisados com o auxílio de algoritmos de *Machine Learning* (ML), aplicados em sistemas de Diagnóstico Auxiliado por Computador (Computer-Aided Diagnosis – CAD). Esses algoritmos são treinados a partir de um grande conjunto de exames médicos rotulados por especialistas, com o objetivo de apoiar os profissionais da saúde na identificação de patologias (SHEHAB et al., 2022). Desde seu surgimento, os modelos de DL têm se destacado na análise automática de diferentes modalidades de imagem médica, com avanços significativos sendo relatados ao longo do tempo (PARVAIZ et al., 2023). Além disso, esses exames muitas vezes contêm informações sutis que não são perceptíveis ao olho humano, o que pode levar à omissão de detalhes relevantes durante a avaliação clínica.

Exames de TC e RM são compostos por centenas de imagens bidimensionais, que segundo o protocolo clínico padrão, devem ser analisadas individualmente por médicos especialistas. No entanto, o diagnóstico humano está sujeito a variáveis como cansaço, fadiga visual e viés cognitivo, o que pode comprometer a precisão da avaliação. Além disso, esses exames contêm informações que não são perceptíveis a olho nu, uma vez que o sistema visual humano é limitado na distinção de níveis de tons de cinza. Diante desse cenário, algoritmos de ML têm sido utilizados em aplicações de CAD, aprendendo com grandes volumes de exames médicos rotulados por especialistas para auxiliar na identificação de patologias (SHEHAB et al., 2022).

Neste contexto, este trabalho tem como foco a segmentação do Acidente Vascular Cerebral hemorrágico (AVCh) em imagens de TC do crânio, por meio de uma abordagem que combina Redes Neurais Convolucionais (Convolutional Neural Networks - CNN) e Redes Grafo Convolucionais (Convolutional Graph Networks - GCN). A principal contribuição consiste na proposta de uma metodologia integrada que utiliza CNNs para extrair características visuais e GCNs para capturar relações estruturais nos dados, uma abordagem ainda pouco explorada na segmentação de AVC.

1.1 Descrição do Problema

Apesar dos avanços em ferramentas computacionais e sistemas de apoio à decisão clínica, erros diagnósticos continuam a ocorrer e têm consequências graves para os pacientes e suas famílias. Em 2015, a Academia Nacional de Medicina (National Academy of Medicine - NAM) dos Estados Unidos alertou que a maioria das pessoas vivenciará um erro diagnóstico ao longo da vida (WALLACE; LIBERMAN, 2021).

Na área da radiologia, essas falhas estão frequentemente associadas a limitações perceptivas. A má interpretação de estruturas anatômicas, por exemplo, pode levar a diagnósticos equivocados com potencial risco à saúde do paciente. Entre os equívocos mais

comuns na análise de imagens de TC, destacam-se a pareidolia e o efeito de contraste simultâneo. A pareidolia ocorre quando o profissional percebe padrões familiares, como rostos, objetos ou formas, em imagens médicas que não representam essas figuras. Radiologistas experientes, por sua vez, podem utilizar essas ilusões como ferramentas diagnósticas, uma vez que certos padrões pareidólicos aparecem de forma consistente em determinadas condições clínicas. Já o efeito de contraste simultâneo refere-se à alteração na percepção da densidade de objetos adjacentes, causada por diferenças na densidade do fundo, o que também pode levar a interpretações equivocadas da imagem (ALEXANDER et al., 2021).

Erros diagnósticos são uma realidade na prática radiológica e podem comprometer a qualidade do atendimento ao paciente. O estudo de Kasalak et al. (2023) revela que, nos dias em que cometeram erros, os radiologistas apresentaram um aumento médio de 21% na carga de trabalho, em comparação com sua própria produtividade habitual, medida pelo número de exames relatados. Esse dado sugere que a sobrecarga de trabalho pode ser um fator contribuinte importante para a ocorrência de erros diagnósticos, especialmente em contextos de alta demanda, pressão clínica ou equipes reduzidas. Além disso, como a produtividade diária de um profissional pode variar ao longo do tempo, é fundamental considerar essas oscilações ao analisar a frequência de erros diagnósticos.

A falha em diagnosticar o AVC está associada a elevada morbidade e mortalidade para o paciente. Em ambientes ambulatoriais e de consultório, evidências indicam que pacientes com manifestações leves, transitórias ou atípicas de doenças cerebrovasculares sejam os mais propensos a receber um diagnóstico incorreto ou tardio (WALLACE; LIBERMAN, 2021).

O AVC é uma das doenças com maior impacto em termos de danos graves decorrentes de erros de diagnóstico. Quase 1 em cada 5 casos de AVC não é diagnosticado corretamente a tempo, o que pode levar a desfechos severos, como morte ou invalidez permanente. Isso reforça a importância de melhorias na precisão e agilidade no diagnóstico, especialmente em contextos de atendimento inicial (NEWMAN-TOKER et al., 2024).

Alguns problemas relacionados ao reconhecimento dos padrões radiológicos cerebrais referem-se a:

- Os erros diagnósticos afetam diretamente a qualidade do atendimento em saúde, podendo resultar em morte, invalidez permanente e sofrimento prolongado. No caso do AVC, uma condição tempo-dependente, o diagnóstico incorreto ou tardio compromete a eficácia do tratamento. Pacientes com manifestações clínicas leves ou atípicas estão mais suscetíveis a erros, principalmente durante atendimentos ambulatoriais. Esse cenário evidencia a necessidade de maior precisão e agilidade nos processos diagnósticos.
- A interpretação de imagens de TC está sujeita a limitações perceptivas e cognitivas dos profissionais, como a percepção equivocada de padrões ou de densidade

entre estruturas adjacentes. Esses fatores podem levar à má interpretação de estruturas anatômicas. Além disso, a sobrecarga de trabalho, a pressão clínica e a redução das equipes contribuem para o aumento da ocorrência de falhas humanas no diagnóstico por imagem.

Esses problemas constituem uma área propícia para a aplicação de análises computadorizadas e estudos mais aprofundados. Sendo assim, com o propósito de amenizar os problemas citados, torna-se necessário desenvolver um sistema capaz de auxiliar os médicos na avaliação de exames de TC, possibilitando uma identificação correta e precisa da região onde o AVC ocorreu.

1.2 Motivação

O Acidente Vascular Cerebral (AVC) representa um importante desafio de saúde pública global, sendo uma das principais causas de morbidade, mortalidade e incapacidade (RUSSELL et al., 2020). Entre os anos de 2000 e 2019, foi responsável por 11% de todas as mortes no mundo (HUANG et al., 2023), afetando anualmente cerca de 800.000 pessoas apenas nos Estados Unidos (ANANTH et al., 2023), e mantendo taxas elevadas de mortalidade em países como o Brasil (MOREIRA et al., 2021).

Na prática clínica, a TC é o exame mais amplamente utilizado para o diagnóstico inicial do AVC, especialmente em sua fase aguda, por ser rápida, acessível e eficaz na exclusão de hemorragias (ABRAMOVA et al., 2021; LEE et al., 2021). Apesar da maior sensibilidade da RM, suas limitações de tempo e disponibilidade tornam a TC o exame preferencial em emergências (PATIL et al., 2022; LEE et al., 2021).

Entretanto, a análise das imagens de TC depende da experiência do radiologista e está sujeita a vieses humanos, como fadiga, sobrecarga de trabalho e variação interobservadores. Esse cenário é agravado pela escassez de profissionais especializados no diagnóstico de AVC, sobretudo em contextos de alta demanda e em regiões menos assistidas. Além disso, o tempo entre o início dos sintomas e a definição do tratamento é um fator crítico para o desfecho clínico, o que exige agilidade e precisão na identificação do tipo de lesão (LEE et al., 2021).

Nesse contexto, sistemas de apoio à decisão clínica baseados em IA têm se mostrado promissores para auxiliar na detecção, classificação e segmentação de lesões cerebrais (HUANG et al., 2023; PATIL et al., 2022). Ferramentas computacionais podem contribuir para reduzir o tempo de análise, minimizar erros humanos e ampliar o acesso a diagnósticos especializados, tornando o processo mais eficiente e seguro.

Diante desses desafios, este trabalho propõe o desenvolvimento de uma ferramenta computacional voltada à análise automatizada de imagens de TC do crânio, com o objetivo de auxiliar profissionais da saúde na identificação de regiões afetadas por AVC. A proposta

visa melhorar a acurácia e a rapidez do diagnóstico, contribuindo para intervenções mais eficazes e para a redução dos impactos do AVC sobre os pacientes e o sistema de saúde.

1.3 Objetivos

1.3.1 Objetivo Geral

Este trabalho propõe uma metodologia integrada para a segmentação de regiões afetadas por AVCh em imagens de TC do crânio, por meio da combinação de CNN e GCN. A abordagem proposta busca oferecer uma ferramenta eficaz para a extração e análise de características em imagens de TC, com o intuito de auxiliar profissionais da saúde na delimitação de lesões causadas por AVCh.

1.3.2 Objetivos Específicos

Além do foco geral, este trabalho possui os seguintes objetivos específicos:

- classificar três padrões cerebrais, a saber: normal, AVC Hemorrágico ou Isquêmico;
- avaliar o impacto dos métodos de criação de arestas baseados em distância euclidiana e similaridade do cosseno na tarefa de segmentação;
- avaliar o impacto do número de vizinhos conectados na qualidade da segmentação;
- realizar um estudo comparativo entre diferentes métodos de criação de grafo e os seus impactos no tempo de processamento;

1.4 Desafios

Com base nos problemas apresentados na seção 1.1, os seguintes desafios foram levantados:

- Definir uma arquitetura de extração de atributos eficiente, utilizando CNNs modificadas;
- Reduzir o custo computacional da criação do grafo para as tarefas de classificação e segmentação;
- Construir grafos adequados às tarefas de classificação e segmentação, de forma que a conectividade entre regiões ou amostras represente suas relações.

Em síntese, os principais desafios envolvem o desenvolvimento de um pipeline robusto, eficiente e com baixo custo computacional, capaz de extrair atributos relevantes para as tarefas de classificação e segmentação. Isso inclui a definição de uma estratégia

eficaz para a construção dos grafos, a modelagem adequada das arestas e o projeto de uma arquitetura de GCN compatível com as exigências de cada etapa.

1.5 Produção Científica

Como resultado da pesquisa desenvolvida, o seguinte artigo foi aceito e publicado:

- **HOLANDA, Adailton Dos Santos; SARMENTO, Roger Moura; REBOUÇAS FILHO, Pedro P. Segmentation of Hemorrhagic Stroke Using Graph Convolutional Network.** In: 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). IEEE, 2025. p. 1-4. *Qualis A1* em Ciência da Computação.

Além disto, o seguinte artigo foi aceito mas não publicado:

- **HOLANDA, Adailton Dos Santos; SARMENTO, Roger Moura; REBOUÇAS FILHO, Pedro P. Stroke Classification Using Graph Convolutional Networks.** Submetido para World Conference on Information Systems and Technologies. *Sem Qualis*.

1.6 Organização da Dissertação

Esta dissertação está organizada em sete capítulos. No Capítulo 1 é feita a introdução geral do trabalho, contextualizando o problema de pesquisa. Em seguida, o Capítulo 2 apresenta os principais aspectos relacionados ao AVC, incluindo definições, tipos, métodos de diagnóstico por imagem e uma revisão do estado da arte em técnicas de classificação e segmentação do AVC em imagens de TC do crânio, além dos avanços recentes no uso de GCNs aplicadas a imagens médicas.

Na sequência, o Capítulo 3 aborda os fundamentos teóricos das Redes Neurais Artificiais, com ênfase nas CNNs, amplamente utilizadas em tarefas de visão computacional aplicadas a imagens médicas, enquanto o Capítulo 4 explora os princípios e fundamentos das GCNs, destacando sua aplicabilidade no processamento de dados estruturados em forma de grafos.

O Capítulo 5 descreve a metodologia proposta para a extração de características, classificação e segmentação do AVC a partir de imagens de TC. Já o Capítulo 6 apresenta os experimentos realizados, a definição dos parâmetros empregados na aplicação das GCNs, bem como a análise e discussão dos resultados obtidos. Por fim, o Capítulo 7 reúne as conclusões do trabalho e aponta sugestões de pesquisas futuras.

2 Acidente Vascular Cerebral

O Acidente Vascular Cerebral (AVC) é uma condição neurológica resultante da interrupção súbita do fluxo sanguíneo cerebral, podendo ocorrer por obstrução ou ruptura de vasos sanguíneos. Trata-se de uma das principais causas de mortalidade e incapacidade no mundo, exigindo diagnóstico rápido e intervenções eficazes. Este capítulo apresenta informações sobre os dois principais tipos de AVC, isquêmico e hemorrágico. Além disso, são discutidos seus mecanismos fisiopatológicos, fatores de risco, estratégias de tratamento e reabilitação, além do papel central da TC no diagnóstico e no suporte à tomada de decisão clínica.

2.1 AVC Isquêmico

O AVCi consiste na interrupção ou redução do fluxo sanguíneo para determinada região do cérebro, privando as células nervosas do fornecimento de oxigênio e nutrientes essenciais à sua sobrevivência. Essa condição leva à morte celular e pode ser confirmada por evidências patológicas, de imagem ou outros dados objetivos (SACCO et al., 2013). As causas mais comuns incluem fatores como a formação de coágulos no próprio vaso sanguíneo, a migração de coágulos vindos de outras partes do corpo ou o acúmulo de placas de gordura que dificultam a passagem do sangue (CAMPBELL; KHATRI, 2020).

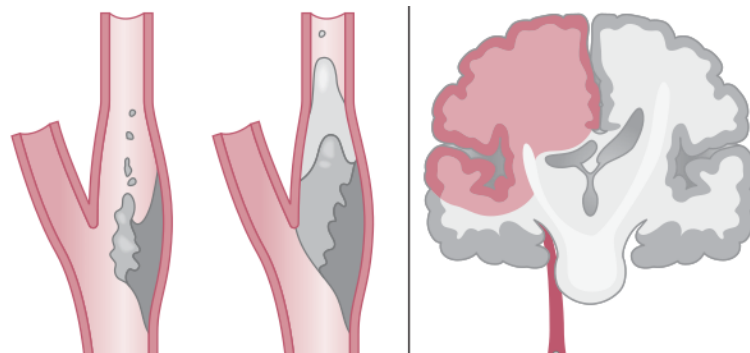


Figura 1 – Representação esquemática do AVC isquêmico. Fonte: Caplan (2016).

Entre as principais artérias acometidas, destaca-se a artéria cerebral média, frequentemente observada em exames de imagem de pacientes com AVC isquêmico (MANIVA; FREITAS, 2012). A obstrução dessas artérias pode ter origem em placas ateroscleróticas, que progridem com a idade e aumentam o risco de estreitamento significativo dos vasos (OVBIAGELE; NGUYEN-HUYNH, 2011). A redução no suprimento sanguíneo cerebral compromete o funcionamento normal das células, levando à falha no controle de íons como sódio e potássio, ao acúmulo extracelular de glutamato e à ativação de mecanismos lesivos,

como o estresse oxidativo e a apoptose (RAMOS-CABRER et al., 2011).

As consequências da isquemia são devastadoras, além de representar uma das principais causas de mortalidade e morbidade no mundo (FEIGIN et al., 2022), é responsável por déficits motores, sensoriais, de memória, linguagem e cognição (BAUMANN et al., 2014). Estima-se que 22,2% dos pacientes com AVC isquêmico venham a óbito em até um ano após o evento (KHAN et al., 2013). Além disso, infartos cerebrais silenciosos aumentam significativamente o risco de infarto clínico e estão relacionados a disfunções cognitivas e físicas (SACCO et al., 2013).

A chamada zona de penumbra isquêmica corresponde a uma área de tecido cerebral ainda viável, porém com irrigação sanguínea insuficiente, situada ao redor do núcleo infartado. Embora essa região ainda não esteja irreversivelmente comprometida, sua preservação depende diretamente do tempo decorrido e da restauração adequada da perfusão (RAMOS-CABRER et al., 2011). A proteção desse tecido representa um dos principais objetivos da terapia trombolítica (MANIVA; FREITAS, 2012).

Nesse cenário, a neuroimagem desempenha papel fundamental no diagnóstico e na tomada de decisão clínica. Exames como a TC e a RM permitem identificar áreas infartadas e distinguir tecido comprometido da penumbra, sendo fundamentais para confirmar o diagnóstico e excluir outras condições que podem simular o AVC. Técnicas como RM ponderada por perfusão e difusão, além da Tomografia por Emissão de Pósitrons (PET), são amplamente utilizadas na avaliação da penumbra (RAMOS-CABRER et al., 2011). No entanto, em muitos hospitais brasileiros, ainda persistem dificuldades relacionadas à disponibilidade e à agilidade na realização desses exames (CARVALHO et al., 2011).

2.2 AVC Hemorrágico

O AVCh representa aproximadamente 10% a 20% de todos os casos de AVC e é considerado uma das formas mais graves da doença, com altas taxas de morbidade e mortalidade associadas (GONZÁLEZ-PÉREZ et al., 2013). Segundo definição atualizada da *American Heart Association*, o AVCh refere-se a um acúmulo focal de sangue no parênquima cerebral ou no sistema ventricular, sem evidência de trauma, resultando em sinais neurológicos rapidamente desenvolvidos.

Essa condição ocorre devido a hemorragia para o interior do tecido cerebral ou para os espaços extravasculares intracranianos. Esse sangramento interrompe conexões neurais e provoca lesões por compressão do tecido cerebral. Além disso, substâncias bioquímicas liberadas durante o sangramento podem agravar o dano ao parênquima e aos vasos adjacentes (CAPLAN, 2016). A fisiopatologia varia conforme o local e a extensão do sangramento, sendo os principais subtipos a hemorragia intracerebral (ICH), caracterizada pelo sangramento dentro do parênquima cerebral, e a hemorragia subaracnoidea (SAH),

que ocorre no espaço entre a membrana aracnóide e a pia-máter (SACCO et al., 2013).

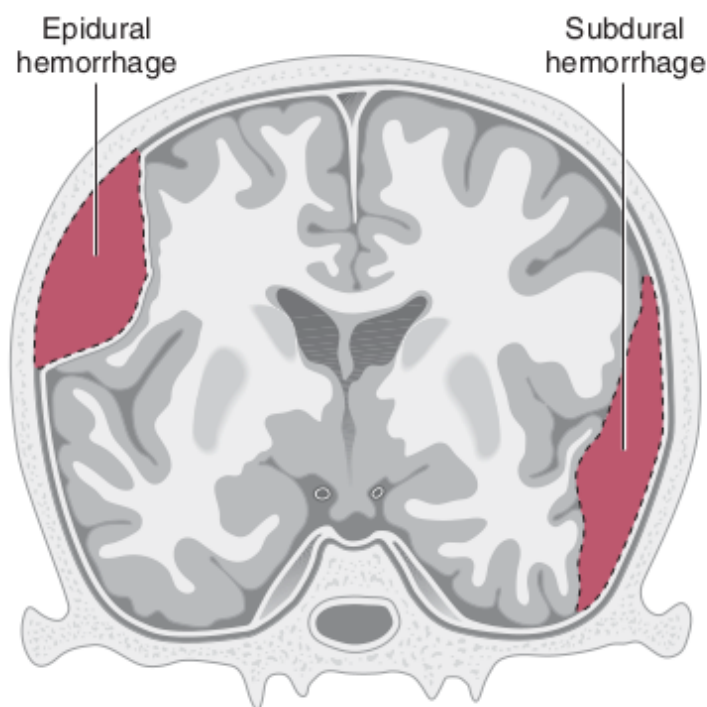


Figura 2 – Representação esquemática do AVC hemorrágico. Fonte: Caplan (2016).

As principais causas do AVCh incluem hipertensão, angiopatia amilóide cerebral, malformações vasculares, distúrbios de coagulação e consumo excessivo de álcool (SACCO et al., 2013; OVBIAGELE; NGUYEN-HUYNH, 2011). Estudos demonstram que a hipertensão está presente em até 91% dos pacientes acometidos por hemorragia cerebral, sendo considerada o fator de risco modificável mais frequentemente associado à hemorragia intracerebral profunda (CARVALHO et al., 2011). Em análises populacionais, observou-se que pacientes com hipertensão crônica apresentam maior risco de ruptura de pequenas artérias penetrantes, mecanismo associado a hemorragias profundas espontâneas (OVBIAGELE; NGUYEN-HUYNH, 2011).

A ruptura de vasos cerebrais pode ser atribuída a diferentes mecanismos, dependendo do tipo de lesão vascular. A pressão arterial elevada é o principal fator associado às hemorragias profundas, enquanto nas hemorragias lobares, especialmente em idosos, a angiopatia amilóide cerebral destaca-se como uma causa comum devido ao acúmulo de proteínas nos vasos corticais (GONZÁLEZ-PÉREZ et al., 2013). Ambas as condições envolvem fragilidade vascular e aumento do risco de sangramento.

2.3 Tratamento e Recomendações

Devido à sua alta letalidade e impacto funcional, o AVC ocupa a posição de segunda principal causa de morte no mundo e é reconhecido como a principal causa de incapacidade

adquirida em adultos (VOS et al., 2020). A literatura internacional reforça essa perspectiva ao equiparar a síndrome cerebrovascular aguda à síndrome coronariana aguda, destacando a urgência no diagnóstico e na intervenção terapêutica, inclusive com a utilização de recursos de telemedicina em áreas geograficamente remotas (SACCO et al., 2013). Destaca-se também a importância do manejo adequado na fase aguda, recomendando-se evitar quadros de hipotensão e hipovolemia que possam agravar o infarto cerebral (MEAD et al., 2023).

A estruturação de unidades especializadas para o atendimento a pacientes com AVC tem demonstrado impacto positivo na redução da mortalidade e da incapacidade funcional, promovendo benefícios sociais e econômicos ao organizar o cuidado com equipe multidisciplinar especializada (RINGELSTEIN et al., 2013). Apesar dos avanços, o acesso rápido ao diagnóstico por imagem e terapias adequadas ainda representa um desafio relevante em países em desenvolvimento, como o Brasil, onde são observadas longas demoras na admissão hospitalar e na realização de exames de neuroimagem (CARVALHO et al., 2011).

O tratamento do AVC evoluiu significativamente com a introdução das terapias de reperfusão, como a trombólise endovenosa e a trombectomia mecânica. Ambas visam restaurar o fluxo sanguíneo cerebral, sendo a trombectomia indicada mesmo até 24 horas após o início dos sintomas em casos selecionados (SCHWAMM et al., 2010). A reperfusão arterial continua sendo o principal alvo terapêutico, e o uso precoce de agentes trombolíticos tem se mostrado seguro e eficaz em pacientes com AVCi agudo (FAN; YANG; JI, 2025). No Brasil, entretanto, estudos demonstram baixa taxa de aplicação de trombólise, inferior a 2%, evidenciando lacunas na organização do cuidado (CARVALHO et al., 2011).

O atendimento ao paciente deve ser realizado por uma equipe multidisciplinar composta por médicos, enfermeiros, fisioterapeutas, terapeutas ocupacionais, fonoaudiólogos, nutricionistas, farmacêuticos e assistentes sociais, cuja atuação integrada tem se mostrado fundamental para a qualidade do cuidado (RINGELSTEIN et al., 2013). A complexidade do diagnóstico e tratamento do AVC exige conhecimento especializado, especialmente para distinguir entre os diferentes mecanismos de lesão cerebral e selecionar intervenções adequadas (CAPLAN, 2016). Além disso, evidências demonstram que o cuidado conduzido por neurologistas e equipes treinadas está associado à redução da mortalidade, institucionalização e dependência funcional (OVBIAGELE; NGUYEN-HUYNH, 2011).

A prevenção primária e secundária do AVC envolve o controle rigoroso de fatores de risco como hipertensão arterial, diabetes, dislipidemia e fibrilação atrial. A linha de cuidado do AVC deve contemplar medidas preventivas, incluindo o refinamento da investigação etiológica e reavaliações periódicas (ADEOYE et al., 2019). Estudos apontam que até 90% dos AVCs poderiam ser evitados com controle eficaz desses fatores (OVBIAGELE; NGUYEN-HUYNH, 2011). Em populações com menor nível socioeconômico, a conscienti-

zação sobre medidas preventivas ainda é limitada, impactando negativamente na qualidade de vida (BAUMANN et al., 2014).

A recuperação funcional após um episódio de AVC depende de múltiplos fatores, incluindo a gravidade da lesão inicial, o tempo até o atendimento, e a qualidade da reabilitação. A reabilitação deve ser iniciada o quanto antes, envolvendo todas as especialidades da equipe multidisciplinar. Contudo, a recuperação dos pacientes ainda é incerta em muitos casos, cerca de 30% dos pacientes apresentam dependência funcional ou sintomas depressivos seis meses após o evento (OVBIAGELE; NGUYEN-HUYNH, 2011; BAUMANN et al., 2014).

2.4 Tomografia Computadorizada

A TC revolucionou a prática médica desde sua introdução clínica em 1971, com a obtenção das primeiras imagens do cérebro humano. Seu desenvolvimento é creditado a Godfrey Hounsfield, que, ao lado de Allan Cormack, recebeu o Prêmio Nobel em 1979 (MAZONAKIS; DAMILAKIS, 2016; RUBIN, 2014). A tecnologia evoluiu rapidamente, passando de sistemas com feixe em lápis e único detector para configurações com múltiplos detectores e reconstruções helicoidais, capazes de adquirir milhares de cortes em poucos segundos (HSIEH; FLOHR, 2021). A Figura 3 exibe o primeiro protótipo de tomógrafo idealizado por Hounsfield.

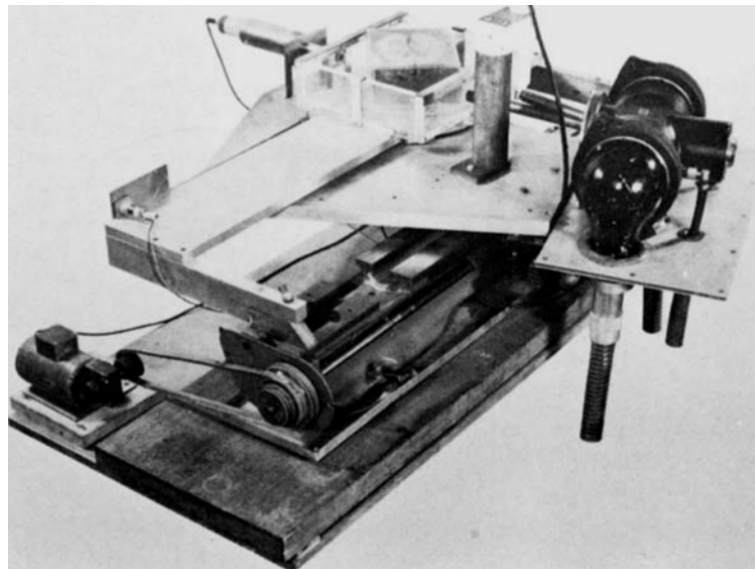


Figura 3 – Modelo inicial do tomógrafo criado por Hounsfield, desenvolvido nos Laboratórios Centrais de Pesquisa da EMI, em Londres. Fonte: Buzug (2009).

O funcionamento da TC baseia-se na emissão de raios-X que atravessam o corpo e são atenuados de acordo com as características dos tecidos. Os feixes atenuados são captados por detectores, convertidos em sinais elétricos e processados por algoritmos que reconstroem

imagens seccionais (HATHCOCK; STICKLE, 1993; MAZONAKIS; DAMILAKIS, 2016). Com o avanço da TC helicoidal e da TC *multislice*, a resolução espacial e o tempo de aquisição melhoraram significativamente, permitindo a aquisição volumétrica e imagens com voxels isotrópicos (HSIEH; FLOHR, 2021; SAMEI; PELC, 2020).

Os tipos de reconstrução evoluíram de técnicas bidimensionais baseadas em retroprojeção filtrada para reconstruções tridimensionais com algoritmos iterativos, como Reconstrução Estatística Iterativa Adaptativa (ASIR), Reconstrução Iterativa Baseada em Modelo (MBIR) e Reconstrução Iterativa no Espaço da Imagem (IRIS), que reduzem ruídos e artefatos (FLEISCHMANN; BOAS, 2011). A reconstrução secundária permite gerar imagens multiplanares, renderizações volumétricas e endoscopia virtual a partir de cortes axiais (BUZUG, 2011).

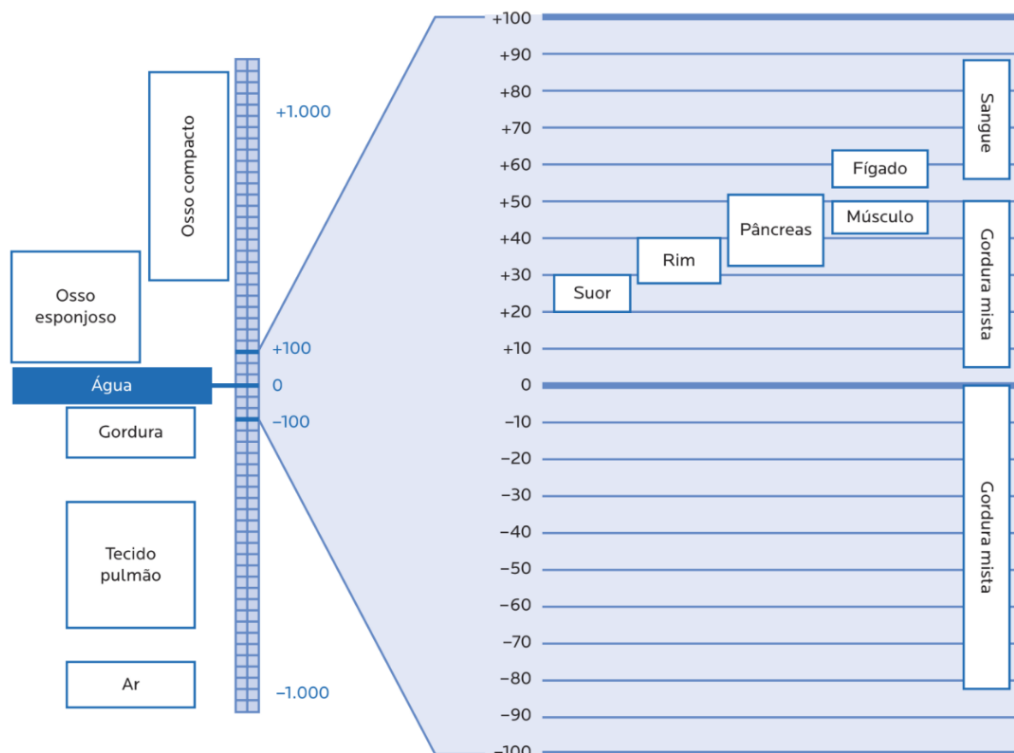


Figura 4 – Ilustração da escala de Hounsfield. Fonte: Freitas (2013).

As imagens produzidas pela TC são compostas por matrizes de pixels que representam voxels, blocos tridimensionais de tecido. A qualidade da imagem está relacionada à dimensão dos voxels, ao número de projeções e à eficiência dos detectores (KAK; SLANEY, 2001; SAMEI; PELC, 2020). Quanto menor o voxel, maior a resolução espacial, embora isso também aumente o ruído e a dose de radiação.

A escala de Hounsfield é um parâmetro fundamental na TC, pois representa a atenuação dos tecidos em unidades padronizadas. Nessa escala, a água é atribuída a 0 HU, o ar a -1000 HU e o osso cortical a +1000 HU (MAZONAKIS; DAMILAKIS, 2016; HATHCOCK; STICKLE, 1993). A manipulação da janela de visualização permite otimizar

o contraste entre tecidos de diferentes densidades, adaptando a imagem às necessidades de diagnóstico (HATHCOCK; STICKLE, 1993). Na Figura 4 é ilustrado a escala de Hounsfield.

No diagnóstico do AVC, a TC é essencial tanto para identificar isquemias quanto hemorragias. Desde a publicação de Ambrose em 1973, a TC tem sido usada para visualizar diretamente infartos cerebrais e hemorragias intracranianas (RUBIN, 2014). Técnicas como TC de perfusão e angiografia por TC têm ampliado a capacidade de avaliação da perfusão cerebral e detecção precoce de AVCi (RUBIN, 2014).

2.5 Estado da Arte

A classificação e a segmentação de doenças por meio de imagens médicas constituem um tema amplamente explorado na literatura científica especializada em IA, especialmente no contexto do AVC. Diversos estudos têm abordado essa problemática sob diferentes perspectivas e metodologias. Considerando o escopo deste trabalho e os avanços relatados na literatura científica sobre o AVC, destacam-se os seguintes enfoques:

- A classificação de exames de TC em três categorias: normais, isquêmicos e hemorrágicos;
- A segmentação das regiões afetadas nos casos de AVCh;
- A aplicação de Redes Grafo Convolucionais (GCNs).

2.5.1 Classificação

A seguir, são apresentados trabalhos relevantes da literatura que tratam da classificação de exames de TC em três categorias: normais, isquêmicos e hemorrágicos.

A pesquisa de Gautam e Raman (2021), utilizou 900 imagens de TC cerebral, distribuídas igualmente entre casos de AVCh, AVCi e normais, obtidas no *Himalayan Institute of Medical Sciences* (HIMS), na Índia. As imagens passaram por um processo de pré-processamento que inclui ajuste de contraste, filtragem, e fusão de imagens. Para a extração de características e classificação, foi proposta uma CNN denominada P_CNN, composta por 13 camadas. No melhor cenário, com validação cruzada 10-fold, a metodologia proposta alcançou 98,77% de acurácia, 98,78% de precisão, 98,77% de sensibilidade, 1,23% de FPR e 98,77% de F-measure.

No trabalho de Chen et al. (2022), foram utilizadas 24.769 imagens de TC cerebral sem contraste, obtidas de 1.715 pacientes, distribuídas entre os seguintes grupos: AVCh, AVCi, normal e outros. As imagens passaram por um processo de pré-processamento que incluiu redimensionamento, normalização dos pixels e *data augmentation*. Utilizou-se o modelo ResNet-50, pré-treinado e ajustado por transferência de aprendizado e otimização de hiperparâmetros. O desempenho alcançado foi de 98,72% de acurácia, 98,75% de

precisão, 98,75% de sensibilidade, 99,00% de F1-score e 99,00% de AUC.

O estudo de Neethi et al. (2022), empregou exames de TC sem contraste de 234 pacientes, distribuídos entre casos de AVCi, AVCh e normais. As imagens passaram por um processo de pré-processamento denominado *P-Proc Tomogram*, que incluiu correção de inclinação, reamostragem isotrópica dos voxels, conversão para unidades Hounsfield com janela [0, 80], remoção de estruturas irrelevantes e redimensionamento. Adicionalmente, foi aplicado *data augmentation* para equilibrar as classes, com normalização por caso. Para extração de características e classificação, foi desenvolvida uma arquitetura de CNN tridimensional. A melhor configuração foi obtida com profundidade de amostra de 8 fatias, alcançando 88,00% de F1-score, 94,00% de precisão, 84,00% de sensibilidade e 92,00% de acurácia.

No estudo realizado por Peixoto e Filho (2018), foram utilizados 300 exames de TC, sendo 100 normais, 100 com AVCi e 100 com AVCh. O pré-processamento consistiu na aplicação da transformada de Fourier para extração das principais componentes estruturais, seguida da reconstrução da imagem no domínio espacial por transformada inversa. A extração de atributos foi realizada por meio da Matriz de Coocorrência Estrutural (SCM) configurada no domínio da frequência, sem necessidade de parametrizações manuais. Para a etapa de classificação, o modelo com melhor desempenho foi o LS SVM com *kernel* RBF. A metodologia proposta alcançou 99,11% de especificidade, 97,00% de sensibilidade, 98,16% de precisão, 98,00% de F1-score, 98,50% de média harmônica e 98,00% de acurácia.

No trabalho Souza et al. (2019), foram utilizadas 420 imagens de TC do crânio, distribuídas igualmente entre casos normais, de AVCi e de AVCh. A etapa de pré-processamento incluiu a aplicação de filtro mediana, limiarização e extração da região de interesse (ROI) cerebral por meio da operação morfológica *Convex Hull* e divisão da ROI em hemisférios esquerdo e direito. A extração de características foi realizada com o método *Analysis of Brain Tissue Densities* (ABTD). A melhor configuração, denominada ABTD-5, considerou cinco classes de densidade: líquido cefalorraquidiano, substância branca, substância cinzenta, sangue e ossos. O classificador utilizado foi o *Optimum-Path Forest* (OPF) com distância euclidiana. A metodologia alcançou 99,30% de acurácia, 98,96% de sensibilidade, 98,98% de precisão, 99,40% de especificidade e 98,95% de F1-score.

A pesquisa de Xu et al. (2020), utilizou 356 imagens de TC do crânio, sendo 142 de pacientes com AVCh e 174 de indivíduos sem lesões. As imagens, originalmente no formato DICOM com resolução de 512×512 pixels e profundidade de 16 bits, foram convertidas para o formato PNG como etapa inicial de pré-processamento. A extração de características foi realizada por meio de redes profundas, sendo a MobileNet a que apresentou melhor desempenho. Para a etapa de classificação, utilizou-se a Rede Perceptron Multicamadas (MLP), combinado à MobileNet. Essa configuração obteve 99,98% de acurácia, 99,98% de F1-score, 99,89% de sensibilidade e 99,81% de precisão.

Tabela 1 – Resumo comparativo dos estudos sobre classificação de AVC em imagens médicas.

Autor	Dataset	Pré-processamento	Extrator	Classificador	Métricas
Gautam e Raman (2021)	900 imagens de TC (HIMS)	Ajuste de contraste, filtro 3×3, fusão de imagens	P_CNN	P_CNN	Acurácia: 98,77% Precisão: 98,78% Sensibilidade: 98,77% FPR: 1,23% F-measure: 98,77%
Chen et al. (2022)	24.769 imagens de TC	Redimensionamento, normalização, data augmentation	ResNet-50	ResNet-50	Acurácia: 98,72% Precisão: 98,75% Sensibilidade: 98,75% F1-score: 99,00% AUC: 99,00%
Neethi et al. (2022)	Exames NCCT de 234 pacientes	Correção de inclinação, voxel isotrópico, janela HU, strided slicing	3D CNN	3D CNN	F1-score: 88,00% Precisão: 94,00% Sensibilidade: 84,00% Acurácia: 92,00%
Peixoto e Filho (2018)	300 exames de TC	Fourier, reconstrução espacial	SCM	LS SVM com RBF	Especificidade: 99,11% Sensibilidade: 97,00% Precisão: 98,16% F1: 98,00% Acurácia: 98,00%
Souza et al. (2019)	420 imagens TC	Conversão para 8 bits, filtro 3×3, limiarização 300 HU, Convex Hull	ABTD-5	OPF	Acurácia: 99,30% Sensibilidade: 98,96% Precisão: 98,98% Especificidade: 99,40% F1: 98,95%
Xu et al. (2020)	356 imagens de TC	Conversão para PNG	MobileNet	MobileNet + MLP	Acurácia: 99,98% F1-score: 99,98% Sensibilidade: 99,89% Precisão: 99,81%

A análise dos estudos voltados à classificação de exames do AVC mostra a predominância de abordagens baseadas em CNNs que demonstraram alto desempenho. A variedade de técnicas de pré-processamento, extração de atributos e classificadores utilizados reflete a busca dos pesquisadores em automatizar a distinção entre exames em casos normais, isquêmicos e hemorrágicos. Tais resultados reforçam o potencial das metodologias computacionais como ferramentas de suporte à decisão clínica.

2.5.2 Segmentação

A seguir, são apresentados trabalhos relevantes da literatura que abordam a segmentação do AVCh em exames de TC.

A pesquisa de Yuan et al. (2023), utilizou imagens de TC de 40 pacientes com lesões hemorrágicas cerebrais irregulares. Para o processo de segmentação, os autores propuseram uma versão aprimorada da arquitetura U-Net, incorporando os módulos ResOctConv e MAM. O módulo ResOctConv combina aprendizado residual com convolução em oitavas, permitindo a extração de características multiescalares e reduzindo o gap semântico entre camadas profundas e rasas. Já o módulo MAM agrega atenção espacial e por canal para priorizar regiões informativas e reduzir interferências. A metodologia proposta alcançou 86% no coeficiente de Dice, 78% no coeficiente de Jaccard.

No trabalho de Xu et al. (2024), foram utilizados 401 exames de TC de pacientes com diferentes tipos de hemorragia intracraniana. As imagens foram divididas em 280 amostras para treinamento e 121 para teste. O pré-processamento incluiu aplicação de janela com bases no valores HU, normalização e *data augmentation*. A extração de características foi realizada por uma arquitetura em cascata baseada em CNN do tipo *encoder-decoder* com duas etapas. O classificador integrado ao modelo realizava a segmentação voxel a voxel. A metodologia proposta alcançou 84,54% no coeficiente de Dice, 86,03% de sensibilidade e erro médio absoluto de 3,149 ml na estimativa do volume da hemorragia.

O estudo de Peng et al. (2022), utilizou 1027 exames de TC de pacientes com hemorragia intracerebral. As imagens foram redimensionadas e truncadas para 40 cortes por paciente, submetidas à janela de intensidade e posteriormente normalizadas. Para extração de características e segmentação, foi proposto o modelo AttFocusNet, que combina a arquitetura U-Net com mecanismos de atenção e uma estrutura Focus inspirada no YOLOv5, permitindo preservar detalhes relevantes durante a codificação. No conjunto de teste interno, a metodologia alcançou 90,8% no coeficiente de Dice, 87,4% de IoU, 91,3% de sensibilidade, 95,7% de precisão.

No estudo realizado por Souza et al. (2023), foram utilizadas 100 imagens axiais de TC de pacientes com AVCh. O pré-processamento envolveu a detecção inicial da região de interesse utilizando a rede Detectron2. A partir do centróide dessas regiões,

foi aplicado o método proposto denominado *Divisible Cell-Segmentation*, que simula o movimento e a divisão de células computacionais em diferentes direções, replicando um processo de crescimento morfológico até a delimitação completa da lesão. A segmentação é finalizada com aplicação do operador *Convex Hull*. A metodologia atingiu 99,80% de acurácia, 98,82% de sensibilidade, 99,81% de especificidade, 93,25% no coeficiente de Dice, 98,55% de precisão e 99,98% de valor preditivo negativo.

O estudo de Santos, Medeiros e Filho (2022), propôs o método *Level Set Based on Radiological Densities with μ -law* (LSBRD- μ) para segmentação totalmente automática de lesões hemorrágicas em imagens de TC do crânio. O pré-processamento incluiu a segmentação do crânio por limiarização, aplicação de filtro de suavização, equalização de intensidades com o algoritmo μ -law, operações morfológicas para isolamento da lesão e redimensionamento da imagem. A extração e evolução da lesão foram conduzidas por uma curva de nível inicializada a partir da região destacada com μ -law e guiada por uma função de densidade estimada por janela de Parzen. A segmentação foi realizada de forma não supervisionada, com classificação probabilística dos pixels baseada nas intensidades radiológicas. A metodologia atingiu 93,39% de MCC, 93,35% no coeficiente de Dice, 93,35% de sensibilidade, 99,94% de especificidade, 93,78% de precisão e 99,85% de acurácia.

No trabalho de Barroso et al. (2022), foram utilizadas 100 imagens de TC com diagnóstico de AVCh. As imagens foram convertidas para o formato JPG e divididas em 80% para treinamento e 20% para teste. A metodologia proposta baseia-se na combinação de CNNs e técnicas de processamento de imagens. Inicialmente, a rede Detectron2 foi treinada para identificar a região de interesse correspondente à lesão, sendo posteriormente aplicada a técnica Método de Contorno Ativo Geodésico (GACM) para refinar a segmentação por meio da evolução de curvas geodésicas. A abordagem proposta obteve 99,85% de acurácia, 94,07% de sensibilidade, 99,92% de especificidade e 94,28% de coeficiente de Dice.

Tabela 2 – Resumo comparativo dos estudos sobre segmentação de AVC hemorrágico em imagens médicas.

Autor	Dataset	Pré-processamento	Modelo	Métricas
Yuan et al. (2023)	40 pacientes	Seleção de camada com maior lesão, contorno manual por especialistas	U-Net com ResOctConv + MAM	Dice: 86% Jaccard: 78% RVD: 11%
Xu et al. (2024)	401 exames	Janela (40/90 HU), normalização, data augmentation	CNN encoder-decoder	Dice: 84,54% Sensibilidade: 86,03% Erro vol.: 3,149 ml
Peng et al. (2022)	1027 exames	Reamostragem, truncagem, normalização, janela [-60, 140]	AttFocusNet	Dice: 90,8% IoU: 87,4% Sensibilidade.: 91,3% Precisão: 95,7% HD95: 5,96 mm ICC: 0,997
Souza et al. (2023)	100 imagens	Detectron2, centroide, Divisible Cell-Segmentation, Convex Hull	Divisible Cell-Segmentation	Acurácia: 99,80% Sensibilidade.: 98,82% Especificidade: 99,81% Dice: 93,25% Precisão: 98,55% NPV: 99,98%
Santos et al. (2022)	CTs do crânio	Limiar 200 HU, μ -law, Parzen, operações morfológicas, redimensionamento	LSBRD- μ	MCC: 93,39% Dice: 93,35% Jaccard: 87,66% Sensibilidade.: 93,35% Especificidade: 99,94% Precisão: 93,78% Acurácia: 99,85%
Barroso et al. (2022)	100 imagens DICOM	Conversão JPG	Detectron2 + Geodesic Active Contour Method	Acurácia: 99,85% Sensibilidade.: 94,07% Especificidade: 99,92% Dice: 94,28%

A análise dos estudos voltados à segmentação do AVCh em imagens de TC mostra o uso crescente de arquiteturas avançadas de CNNs, combinadas a técnicas de pré-processamento e extração de características. Modelos como U-Net aprimorada, AttFocusNet e arquiteturas encoder-decoder têm se mostrado eficazes na delimitação precisa das regiões afetadas, com resultados expressivos em métricas como Dice e Jaccard. Além disso, abordagens baseadas em processamento morfológico, segmentação por curvas de nível e aprendizado não supervisionado também demonstraram alto desempenho, evidenciando a diversidade de soluções aplicáveis a essa tarefa. De modo geral, os estudos destacam o potencial das técnicas de segmentação como componentes fundamentais em sistemas automatizados de apoio ao diagnóstico de AVCh.

2.5.3 Redes Grafo Convolucionais

A seguir, são apresentados e discutidos trabalhos relevantes da literatura que abordam a aplicação de GCNs aplicadas em exames de TC.

A pesquisa de Khoshkhabar et al. (2023), utilizou o conjunto de dados *Liver Tumor Segmentation Challenge*, composto por 4158 imagens. O pré-processamento incluiu a exclusão de pixels com intensidade fora do intervalo $[0, 150 \text{ HU}]$, redimensionamento, além de *data augmentation*. A segmentação foi conduzida por meio do modelo SLIC-DGN, que combina a geração de superpixels com a técnica SLIC e sua representação em grafos, sendo cada nó definido com base na média de intensidade dos superpixels. A arquitetura da rede é baseada em GCN, com convoluções de Chebyshev. A metodologia alcançou 98,7% de acurácia, 90,00% no coeficiente de Dice, 89,2% de IoU médio, 97,9% de sensibilidade e 88,7% de precisão. Em comparação com modelos tradicionais, como o Modified U-Net e o Shortcut CNN, o SLIC-DGN demonstrou desempenho superior, atingido acurácia superior a 90,00% na segmentação de tumores hepáticos.

No trabalho de Xuan et al. (2022), foi utilizado o conjunto de dados *Kidney Tumor Segmentation Challenge*, composto por 210 volumes de TC. O pré-processamento incluiu *data augmentation*, além da adição de ruído de brilho e gama. As imagens foram divididas em blocos tridimensionais e processadas pelo U-Net para extração de atributos. A metodologia proposta, denominada DGC-Seg, integra um autoencoder baseado em grafos com convoluções dinâmicas (DGC) e um mecanismo de atenção por atributo de nó. A estrutura do grafo é formada com base na similaridade entre regiões e evolui dinamicamente durante o treinamento. Os resultados obtidos para a segmentação de tumor foram de 86,5% de Dice e 77,2% de IoU, enquanto para o rim os valores foram 96,1% de Dice e 92,6% de IoU. A arquitetura DGC-Seg superou métodos tradicionais como 3D U-Net, 3D nnU-Net e 3D ResNet, bem como abordagens baseadas em GCNs estáticas e convoluções convencionais.

O estudo de Hu et al. (2024), utilizou um conjunto de dados públicos de TC de

pacientes com câncer de cabeça e pescoço, com anotações manuais dos órgãos de risco, incluindo o músculo constritor naris, olhos, lentes, nervos ópticos, lobos temporais e tireoides. O pré-processamento incluiu a fusão de estruturas simétricas em uma única classe para evitar viés posicional e a segmentação tridimensional dos volumes. A arquitetura proposta é baseada na 3D U-Net, estendida com um módulo denominado *Adaptive Sparse Graph Convolutional Network*, que trata cada canal da rede como um nó em um grafo. A matriz de adjacência é construída a partir de convoluções 3D e regularizada com *dropout* adaptativo para formar uma estrutura esparsa e adaptativa. A saída segmentada é obtida por meio da multiplicação da matriz de adjacência com os vetores de características dos canais. A metodologia proposta alcançou, dentre os melhores resultados, 89,51% de Dice e 1,1163 mm de distância de Hausdorff (HD) na segmentação dos olhos; 64,95% de Dice e 2,4956 mm nas lentes; e 86,90% de Dice e 2,1383 mm nos lobos temporais. Em comparação com a U-Net tradicional e variantes como Swin UNETR, DCGCN e Attention Gate, o modelo ASGCN demonstrou desempenho superior em quatro dos seis órgãos avaliados.

No estudo conduzido por Yu et al. (2021), foram utilizadas 296 imagens de TC do tórax, sendo 148 de pacientes diagnosticados com COVID-19 e 148 de indivíduos saudáveis. As imagens foram redimensionadas e processadas pela arquitetura ResNet101-C, uma versão modificada da ResNet101. A classificação foi realizada pelo modelo ResGNet-C, que combina os vetores extraídos com uma GCN de uma camada, construída com base na vizinhança k-mais próxima, definida por distância euclidiana. A segmentação final é obtida pela multiplicação da matriz de adjacência normalizada pelos vetores de características e aplicação da GCN. Na melhor configuração, o modelo alcançou 96,62% de acurácia, 96,21% de precisão, 96,65% de F1-score, 97,33% de sensibilidade e 95,91% de especificidade. Quando comparado com métodos tradicionais, como a ResNet101-C e a arquitetura simples NNet-C, a ResGNet-C demonstrou desempenho superior.

O estudo de Liu et al. (2023), utilizou o conjunto de dados TCGA-LGG, composto por 402 exames de RM da modalidade FLAIR, classificados em três grupos: livre de tumor, glioma IDH-mutante e glioma IDH wild-type. O pré-processamento envolveu a seleção de 64 fatias centrais por exame, redimensionamento e *data augmentation*. A extração de características foi realizada por meio da *Convolutional Deep Belief Network*, cuja camada *bottleneck* produziu vetores de características de 128 dimensões. A classificação foi conduzida por uma GCN de duas camadas, denominada SASG-GCN. A metodologia proposta alcançou 93,62% de acurácia e 92,30% de AUC. Comparado a abordagens convencionais baseadas em redes convolucionais como ResNet, DenseNet, Pyramid Dilated CNN, GAN e MMD-VAE, o modelo SASG-GCN apresentou desempenho superior, mesmo utilizando uma arquitetura mais rasa e com menor número de amostras.

No trabalho de Wang et al. (2025), foram utilizados dois conjuntos de dados de imagens de ultrassonografia mamária automatizada: o HDH-ABUS, contendo 547 casos, e

o TDSC-ABUS, com 88 volumes. O pré-processamento incluiu normalização das imagens, recorte central e seleção das 11 fatias centrais de cada volume tumoral. A extração de características foi realizada com uma ResNet50 pré-treinada, utilizando as saídas das quatro primeiras etapas da rede. A classificação foi conduzida por uma GCN de duas camadas, na qual cada fatia é representada como um vértice em um grafo construído com base na correlação de Pearson/Spearman entre as fatias. A metodologia alcançou, no teste com os dados do dataset HDH-ABUS, 93,7% de AUC, 87,3% de acurácia, 91,7% de precisão, 81,5% de sensibilidade, 92,9% de especificidade e 86,3% de F1-score. No teste com os dados do dataset TDSC-ABUS, os valores foram 82,6% de AUC, 86,4% de acurácia, 92,6% de precisão, 71,4% de sensibilidade, 96,2% de especificidade, 67,7% de Youden's Index e 80,6% de F1-score. Em comparação com redes convolucionais tradicionais, como ResNet, DenseNet, ResNeXt, Vision Transformer e Swin Transformer, o modelo ABUS-Net demonstrou desempenho superior.

A pesquisa de Fu et al. (2023), utilizou 483 nódulos pulmonares extraídos de exames de TC. As imagens passaram por conversão para unidades Hounsfield com corte de intensidade para o intervalo $[-1400, 400]$, normalização e extração de regiões de 60×60 pixels centradas nos nódulos. Os canais das imagens foram replicados para adequação ao formato RGB, e técnicas de *data augmentation* foram aplicadas. A extração de características foi realizada com uma rede VGG16 pré-treinada aprimorada com o módulo de atenção *Convolutional Block Attention Module*, que incorpora atenção espacial e de canal, resultando em vetores de 512 dimensões por fatia. A classificação foi conduzida por uma GCN de duas camadas utilizando grafos completamente conectados entre todas as fatias de um mesmo nódulo. A melhor configuração da metodologia, denominada AE-GCN, alcançou 95,24% de AUC, 93,55% de acurácia, 84,62% de sensibilidade, 91,04% de especificidade, 78,57% de precisão e 81,48% de F1-score. Em comparação com modelos baseados apenas em redes convolucionais (ResNet, DenseNet, 3D ResNet) e Transformers (ViT, ViT-GCN), a AE-GCN demonstrou desempenho superior.

O trabalho de Xiao et al. (2023), utilizou um conjunto de dados privado fornecido pelo *Tianjin Haihe Hospital*, composto por 373 imagens de TC de 59 pacientes com tuberculose pulmonar, todos tratados como uma única classe para fins de segmentação. O pré-processamento incluiu a conversão das imagens para unidades Hounsfield, com recorte da faixa de intensidade para o intervalo $[-1400, 400]$, normalização e transformação dos mapas de características em vetores para redução da dimensionalidade e construção dos grafos. A extração de características foi realizada por redes convolucionais VGG-16 ou ResNet-50, cujos vetores de saída foram utilizados como nós na arquitetura proposta, denominada R2GCN. Essa arquitetura combina uma GCN, responsável por capturar as relações topológicas entre os nós, com uma Rede Neural Recorrente Bidirecional, que modela as dependências contextuais horizontais e verticais. Na melhor configuração, utilizando VGG-16 como *backbone*, o modelo atingiu 96,93% de Dice e 94,05% de IoU.

Em comparação com abordagens baseadas exclusivamente em CNNs, como FCN, U-Net, PSPNet, FPN e TransUNet com ResNet-50, o modelo R2GCN demonstrou desempenho superior.

No estudo conduzido por Choudhry et al. (2025), foi utilizado o conjunto de dados do desafio *Brain Tumor Segmentation Challenge*, composto por 2000 exames de RM. Os volumes foram fatiados ao longo do eixo Z, totalizando 193.905 imagens bidimensionais. As máscaras de segmentação foram utilizadas apenas na fase de avaliação. A extração de características foi realizada por uma arquitetura baseada em ResNet com blocos residuais e *pooling* piramidal, enquanto a segmentação foi conduzida por um modelo híbrido que integra uma CNN com uma GCN. A integração entre as duas arquiteturas foi realizada por um mecanismo de atenção adaptativa e mapas de ativação. A metodologia proposta atingiu 97,43% de Dice, sendo os valores individuais de Dice: 96,17% para tumor realçado, 95,47% para tumor total e 96,76% para núcleo tumoral. Em comparação com modelos tradicionais baseados exclusivamente em CNN, como UNet, VNet e UNet++, o modelo híbrido demonstrou desempenho significativamente superior.

No estudo conduzido por Cui et al. (2024), foram utilizados dois conjuntos de dados para avaliação da metodologia proposta: o KiTS19, composto por 210 exames de TC, e o SDCH, com 245 exames torácicos de pacientes com câncer de pulmão. Os volumes foram divididos em blocos tridimensionais e foram aplicadas técnicas de *data augmentation*. A extração de características foi realizada por meio do *backbone* 3D nnU-Net, normalização e ativação ReLU. A segmentação foi conduzida por uma arquitetura híbrida denominada *Hybrid Graph Convolution Transformer*, composta por três módulos: CTrans, responsável por aprender relações contextuais entre canais por meio de transformers; convG-Trans, que integra GCN e transformer com evolução dinâmica da topologia do grafo; e uma camada de fusão com atenção orientada à categoria. No conjunto KiTS19, o modelo atingiu 96,78% de Dice, 93,82% de IoU e 15,75 mm de distância de Hausdorff para o rim, e 87,64% de Dice, 78,96% de IoU e 14,53 mm de HD para o tumor. No conjunto SDCH, os valores foram de 99,21% de Dice para pulmão e 60,85% para tumor, com HDs de 92,13 mm e 34,72 mm, respectivamente. Em comparação com arquiteturas tradicionais como 3D U-Net, nnU-Net, TransUNet, TransBTS e diversas variantes de GCN, o modelo HCGT apresentou desempenho superior em todos os cenários, com destaque para o ganho de 1,75% no Dice e redução de 57,44% no HD na segmentação de tumores renais em relação ao nnU-Net.

Tabela 3 – Estudos que aplicam GCNs em tarefas de classificação e segmentação em imagens médicas.

Autor	Dataset	Objetivo	Rede GCN	Resultados (comparação com CNNs)
Khoshkhabar et al. (2023)	Liver Tumor Segmentation Challenge (4158 imagens)	Segmentação de tumores hepáticos	SLIC-DGN (GCN com convoluções de Chebyshev)	Acurácia: 98,7%, Dice: 90%, IoU: 89,2%, supera U-Net e Shortcut CNN
Xuan et al. (2022)	Kidney Tumor Segmentation Challenge (210 volumes)	Segmentação de tumores renais	DGC-Seg (GCN com convoluções dinâmicas e atenção)	Dice: 86,5%, IoU: 77,2%, supera 3D U-Net, nnU-Net e ResNet
Hu et al. (2024)	TC cabeça/pescoço (público)	Segmentação de órgãos de risco	ASGCN (GCN com estrutura esparsa adaptativa)	Dice: 89,5% (olhos), supera Swin UNETR, DCGCN, SE
Yu et al. (2021)	TC tórax (296 imagens)	Classificação COVID-19	ResGNet-C (GCN + ResNet101-C)	Acurácia: 96,62%, supera ResNet101-C isolada (91,59%)
Liu et al. (2023)	TCGA-LGG (402 exames)	Classificação de gliomas	SASG-GCN (GCN com autoatenção)	Acurácia: 93,62%, AUC: 92,30%, supera ResNet, DenseNet, GAN
Wang et al. (2025)	HDH-ABUS e TDSC-ABUS (ultrassom)	Classificação de tumores mamários	GCN com ResNet50 + correlação inter-fatia	Acurácia: até 87,3%, supera DenseNet, ViT, Swin Transformer
Fu et al. (2023)	483 nódulos pulmonares (TC)	Classificação de nódulos pulmonares	AE-GCN (GCN + atenção espacial e de canal)	Acurácia: 93,55%, supera VGG16, ResNet, ViT
Xiao et al. (2023)	59 pacientes com tuberculose (TC)	Segmentação pulmonar	R2GCN (GCN + Bi-RNN)	Dice: 96,93%, supera U-Net, FCN, PSPNet, TransUNet
Arshad et al. (2025)	Brain Tumor Segmentation Challenge (193.905 imagens)	Segmentação de tumores cerebrais	CNN + GCNN com atenção adaptativa	Dice médio: 97,43%, supera UNet, VNet, UNet++
Cui et al. (2024)	KiTS19 e SDCH (455 exames)	Segmentação renal e pulmonar	HCGT (GCN + Transformer com fusão dinâmica)	Dice: até 99,21%, supera TransUNet, GCNs tradicionais

A revisão dos estudos que empregam GCNs em imagens médicas evidencia a crescente relevância dessa abordagem para tarefas de classificação e segmentação em diversos contextos clínicos. Os trabalhos analisados demonstram que as GCNs, frequentemente combinadas com CNNs, apresentam desempenho superior às arquiteturas tradicionais em métricas como Dice, acurácia e AUC, sobretudo em cenários onde as relações espaciais ou estruturais entre regiões da imagem desempenham papel crucial. As estratégias adotadas variam desde a modelagem de superpixels como nós do grafo até a integração de mecanismos de atenção espacial e temporal, reforçando a versatilidade das GCNs na representação de dados médicos complexos. Esse panorama fundamenta a proposta deste trabalho, que visa aplicar uma abordagem híbrida baseada em CNNs e GCNs para classificação e segmentação do AVCh em exames de TC do crânio, com o objetivo de aprimorar a identificação automática de regiões afetadas e contribuir com o desenvolvimento de ferramentas de suporte à decisão clínica.

3 Redes Neurais

Este capítulo apresenta os principais conceitos sobre redes neurais artificiais, fundamentais para o desenvolvimento de sistemas de aprendizado profundo. São abordadas desde estruturas mais simples até modelos mais avançados, evidenciando a evolução das redes em termos de profundidade e capacidade de generalização. A discussão inicia-se com o Perceptron simples, considerado a base histórica das redes neurais. Essa arquitetura elementar permite compreender os princípios fundamentais do funcionamento de um neurônio artificial, servindo de ponto de partida para as demais arquiteturas exploradas ao longo do capítulo.

3.1 Perceptron Simples

O Perceptron, idealizado por Rosenblatt em 1958, representa a forma mais simples de configuração de uma rede neural artificial. Seu objetivo era implementar um modelo computacional inspirado na retina, atuando como um elemento de percepção eletrônica de sinais (SILVA; SPATTI; FLAUZINO, 2010). Essa estrutura é composta por um combinador linear e uma função de limiar.

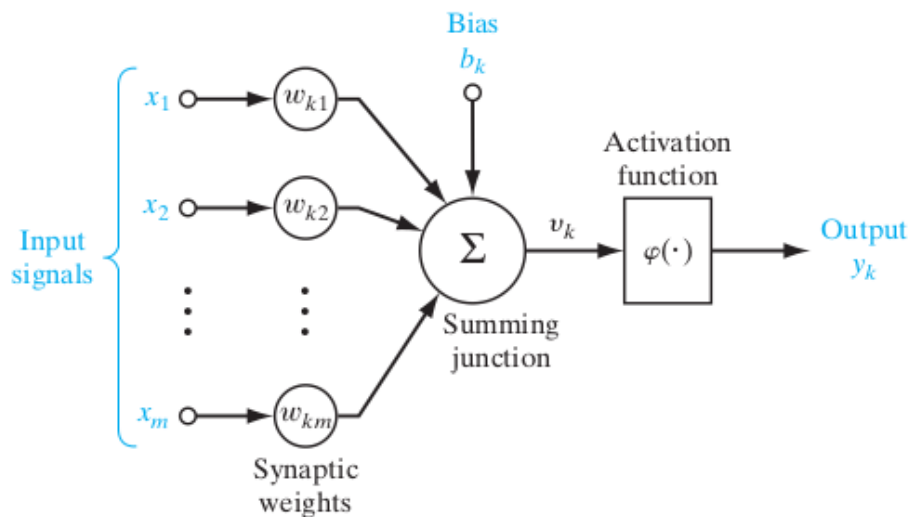


Figura 5 – Estrutura de um neurônio artificial utilizada no modelo Perceptron. Fonte: Haykin (2009).

A simplicidade do Perceptron está no fato de ser constituído por apenas uma camada neural e, frequentemente, conter um único neurônio de saída. Na Figura 5 é exibida essa configuração, composta por n sinais de entrada, representando o problema a ser mapeado, e uma única saída, visto que a rede contém apenas um neurônio. Apesar de sua simplicidade, o Perceptron atraiu à época de sua proposição, o interesse de diversos

pesquisadores da área de inteligência artificial, sendo amplamente estudado e valorizado pela comunidade científica (SILVA; SPATTI; FLAUZINO, 2010).

A partir da análise da Figura 5, observa-se que cada entrada x_i , representando informações do processo a ser modelado, é inicialmente ponderada por um peso sináptico w_{ki} , atribuindo-se importância relativa a cada entrada em relação ao comportamento desejado do sistema. O valor resultante da combinação dessas entradas, somado ao limiar de ativação θ , é então processado por uma função de ativação, resultando na saída y do Perceptron.

Em seguida, o valor resultante da composição de todas as entradas já devidamente ponderadas pelos respectivos pesos, adicionado ainda do limiar de ativação θ , é repassada como argumento da função de ativação, cujo resultado de retorno será a saída y produzida pelo Perceptron.

O funcionamento matemático do Perceptron é descrito pelas seguintes equações:

$$\begin{cases} u = \sum_{i=1}^n w_i \cdot x_i - \theta & (1) \\ y = g(u) & (2) \end{cases} \quad (3.1)$$

Nessas expressões, x_i representa a i -ésima entrada da rede, w_i é o peso sináptico correspondente, θ é o limiar de ativação, $g(\cdot)$ é a função de ativação e u é o potencial de ativação.

As funções de ativação utilizadas no Perceptron são, tipicamente, a função degrau ou degrau bipolar. Assim, a saída da rede pode assumir apenas dois valores possíveis: 0 ou 1, no caso da função degrau padrão, ou -1 e 1, quando se utiliza a função degrau bipolar. As entradas x_i podem assumir quaisquer valores numéricos, a depender do tipo de problema modelado. Na prática, técnicas de normalização são frequentemente aplicadas, de modo a ajustar os valores de entrada aos limites de saída definidos pelas funções de ativação, contribuindo para a eficiência do processo de treinamento.

O ajuste dos pesos sinápticos e do limiar é realizado por meio de treinamento supervisionado. Isso significa que, para cada amostra de entrada, conhece-se a respectiva saída desejada. Como o Perceptron é comumente utilizado para problemas de reconhecimento de padrões com duas classes, sua saída binária é associada diretamente às categorias que o modelo deve distinguir.

O Perceptron simples, embora pioneiro, possui limitações importantes. Uma das mais conhecidas é sua incapacidade de resolver problemas que não são linearmente separáveis, como o caso clássico da porta lógica XOR. Essa limitação evidenciou a necessidade de arquiteturas mais complexas, capazes de representar funções não lineares. Foi nesse contexto que surgiram as redes MLP, apresentando uma estrutura com camadas ocultas

que possibilitam a modelagem de relações mais complexas entre entradas e saídas.

3.2 Perceptron Multicamadas

As redes MLP representam uma evolução do modelo Perceptron simples, sendo compostas por uma ou mais camadas ocultas entre as camadas de entrada e saída (SILVA; SPATTI; FLAUZINO, 2010). Cada camada é formada por neurônios artificiais que recebem entradas ponderadas, aplicam uma função de ativação e propagam o sinal para a próxima camada. A principal característica da MLP é sua capacidade de modelar relações não lineares e de resolver problemas de classificação que não são linearmente separáveis.

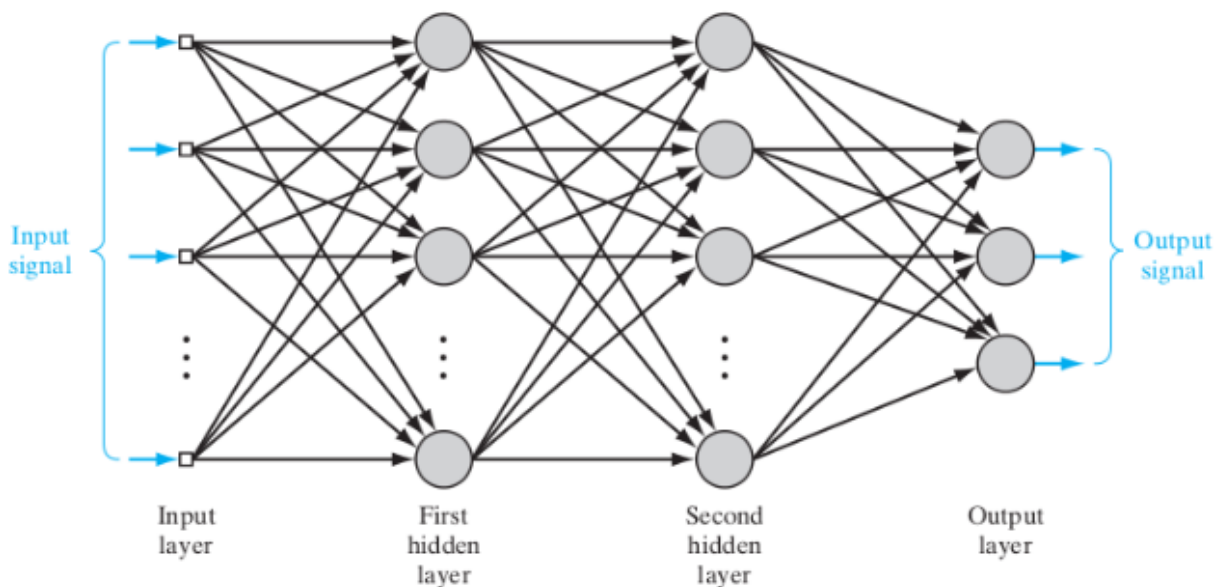


Figura 6 – Arquitetura de uma rede MLP. Fonte: Haykin (2009).

Na figura 6 é ilustrada a propagação dos sinais desde a camada de entrada até a de saída, seguindo um único sentido, o que caracteriza uma arquitetura do tipo feedforward. As saídas dos neurônios da primeira camada oculta tornam-se as entradas da camada seguinte, e assim sucessivamente.

Diferentemente do Perceptron simples, a camada de saída pode conter múltiplos neurônios, permitindo a classificação em diversas categorias. O mapeamento entre entradas e saídas é distribuído entre os neurônios da rede, o que confere robustez e capacidade de generalização ao modelo. A topologia da rede, definida pelo número de camadas ocultas e pela quantidade de neurônios em cada uma delas, depende de fatores como a complexidade do problema, a distribuição dos dados de treinamento e os valores iniciais dos parâmetros.

O algoritmo *backpropagation* é o método mais amplamente utilizado para o treinamento supervisionado de uma MLP. Esse algoritmo utiliza o gradiente descendente para minimizar o erro entre a saída obtida pela rede e a saída desejada.

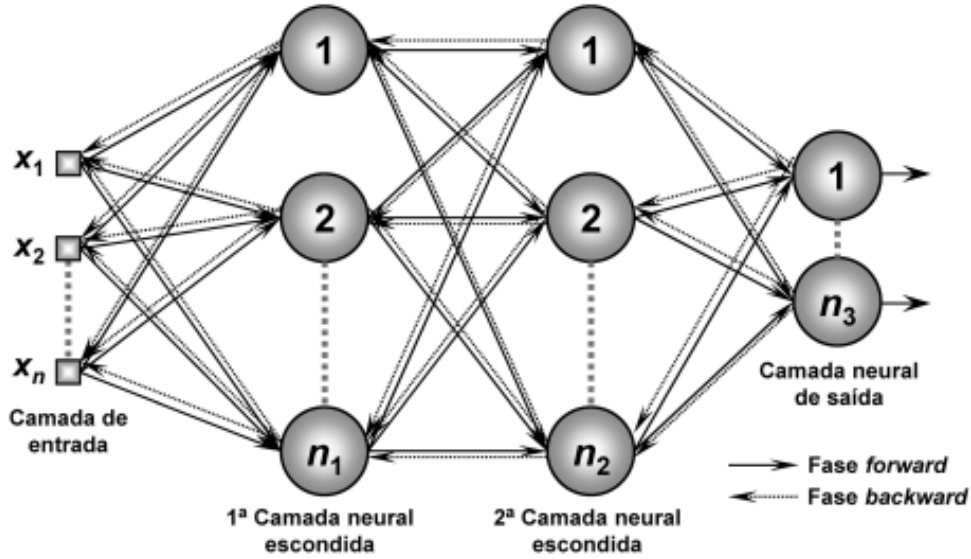


Figura 7 – Diagrama ilustrativo do processo de retropropagação. Fonte: Silva, Spatti e Flauzino (2010).

O processo de aprendizado ocorre em duas fases principais:

1. Propagação direta (*forward*): os sinais de entrada são transmitidos camada por camada até a saída, passando por pesos sinápticos e funções de ativação. Nessa etapa, os pesos são utilizados apenas para calcular a saída da rede.
2. Propagação reversa (*backward*): o erro entre a saída obtida e a saída desejada é calculado e propagado no sentido inverso, da saída para as camadas anteriores. Os pesos são então atualizados com base no gradiente do erro, utilizando o método do gradiente descendente.

O campo local induzido no neurônio j , na iteração n , é dado por:

$$v_j(n) = \sum_{i=1}^m w_{ji}(n)y_i(n) + w_{j0}(n) \quad (3.2)$$

A saída do neurônio é obtida por meio de uma função de ativação não linear:

$$y_j(n) = \phi(v_j(n)) \quad (3.3)$$

O erro entre a saída desejada $d_j(n)$ e a saída obtida $o_j(n)$ é calculado como:

$$e_j(n) = d_j(n) - o_j(n) \quad (3.4)$$

A atualização do peso sináptico w_{ji} segue a regra do gradiente descendente:

$$\Delta w_{ji}(n) = -\eta \cdot \frac{\partial e(n)}{\partial w_{ji}(n)} \quad (3.5)$$

onde η é a taxa de aprendizagem. O gradiente local (ou sinal de erro) do neurônio j é dado por:

$$\delta_j(n) = e_j(n) \cdot \phi'(v_j(n)) \quad (3.6)$$

Nas camadas ocultas, os gradientes locais são calculados com base nos erros propagados da saída, ponderados pelos respectivos pesos. Esse processo é repetido iterativamente até que os pesos da rede converjam para valores que minimizem a função de erro.

Para melhorar a eficiência e a estabilidade do treinamento, podem ser adotadas estratégias como o uso de *momentum* e esquemas de redução progressiva da taxa de aprendizagem. O *momentum* suaviza as atualizações e ajuda a evitar mínimos locais, enquanto a redução progressiva da taxa de aprendizagem reduz gradualmente o valor de η , promovendo estabilização nas fases finais do treinamento.

A função de ativação introduz não linearidade ao modelo e determina como o campo local $v_j(n)$ se transforma na saída do neurônio $y_j(n)$. A derivada da função de ativação é essencial para o cálculo do gradiente local durante a retropropagação do erro.

$$y_j(n) = \phi(v_j(n)) \quad (3.7)$$

Durante o treinamento, o erro local é calculado por:

$$e_j(n) = d_j(n) - y_j(n) \quad (3.8)$$

E o gradiente local do neurônio j , utilizado para o ajuste dos pesos sinápticos, é definido como:

$$\delta_j(n) = e_j(n) \cdot \phi'(v_j(n)) \quad (3.9)$$

A escolha da função ϕ influencia diretamente a suavidade da curva de erro e a taxa de convergência do treinamento.

Apesar da expressiva capacidade das redes MLP em modelar relações complexas, sua estrutura totalmente conectada pode se tornar limitante ao lidar com dados de alta dimensão, como imagens, nos quais a preservação da estrutura espacial é essencial. Essa limitação impulsionou o desenvolvimento das CNNs, que utilizam operações específicas de convolução e pooling para explorar características espaciais locais de maneira mais adequada a esse tipo de dado.

3.3 Redes Neurais Convolucionais

As CNNs se destacam pela capacidade de extrair características locais a partir de dados estruturados espacialmente, como imagens. A operação de convolução é central neste tipo de rede, sendo responsável por aplicar filtros que identificam padrões como bordas, texturas e regiões de interesse nos dados de entrada, conforme discutido por LeCun et al. (1998) e Dumoulin e Visin (2018). Para compreender seu funcionamento, é fundamental apresentar a base matemática dessa operação, que sustenta a extração eficiente de características em redes convolucionais.

Matematicamente, a convolução é uma operação que combina duas funções reais para produzir uma terceira função (MA; TANG, 2021). No domínio contínuo, a convolução entre duas funções $f(t)$ e $g(t)$ é definida por:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g(t - \tau) d\tau \quad (3.10)$$

Essa operação pode ser interpretada como uma média ponderada da função f , onde a função g , chamada de função de peso ou *kernel*, determina a influência relativa dos valores próximos a t . Quando aplicada a sinais ruidosos, por exemplo, essa técnica permite suavizar flutuações locais, atribuindo pesos maiores aos valores mais próximos no tempo.

Na prática, entretanto, a maioria dos dados é representada de forma discreta. Assim, para funções definidas em pontos inteiros, a convolução discreta assume a forma:

$$(f * g)[n] = \sum_{k=-\infty}^{\infty} f[k] \cdot g[n - k] \quad (3.11)$$

No caso em que $g[k]$ é não nula apenas em uma janela de tamanho finito $2n + 1$, a convolução é restrita a um intervalo local, tornando-se computacionalmente mais eficiente:

$$(f * g)[n] = \sum_{k=-n}^n f[n - k] \cdot g[k] \quad (3.12)$$

No contexto do processamento de imagens, a convolução bidimensional é amplamente utilizada para realçar características visuais, como bordas ou regiões homogêneas. A operação bidimensional entre uma imagem de entrada S e um kernel K é dada por:

$$S'(i, j) = (S * K)(i, j) = \sum_m \sum_n S(m, n) \cdot K(i - m, j - n) \quad (3.13)$$

em que S' representa a imagem resultante da operação (DUMOULIN; VISIN, 2018). Filtros com pesos alternados positivos e negativos são geralmente utilizados para detecção de bordas, enquanto filtros suavizantes, que se comportam como médias ponderadas, são eficazes na redução de ruídos e realce de áreas uniformes.

Nas CNNs, o conceito de convolução é implementado por meio da aplicação de filtros sobre regiões locais da entrada. Essa estrutura resulta em conexões esparsas entre os neurônios das camadas, em contraste com as camadas densamente conectadas, onde todos os neurônios de entrada afetam todos os de saída.

Uma das principais vantagens dessa arquitetura é a eficiência computacional. Enquanto camadas densas exigem $N \times M$ parâmetros para N entradas e M saídas, as camadas convolucionais necessitam de apenas $K \times M$ parâmetros, onde K é o tamanho do kernel e $K \ll N$. Assim, a complexidade computacional é reduzida de $O(N \times M)$ para $O(K \times M)$, proporcionando um desempenho significativamente melhor em tarefas de visão computacional (MA; TANG, 2021).

Além da redução no número de parâmetros, outros aspectos estruturais influenciam diretamente o comportamento e o desempenho das CNNs. Um desses aspectos é o tipo de *padding* utilizado, que determina como as bordas da entrada são tratadas durante a aplicação da convolução e afeta diretamente as dimensões da saída, com implicações na profundidade da rede e na preservação de informações espaciais. Abaixo são descritos os principais tipos de *padding* utilizados em redes convolucionais:

- Sem *padding*: não há adição de zeros nas bordas da imagem. A saída tem dimensão menor que a entrada.
- Half *padding*: adiciona-se *padding* suficiente para manter as dimensões da saída iguais às da entrada.
- Full *padding*: adiciona-se zeros de modo que todos os elementos do kernel possam ser aplicados, resultando em uma saída maior que a entrada.

Esses conceitos são fundamentais para controlar o tamanho da saída em redes convolucionais profundas.

Outro parâmetro estrutural importante nas CNNs é o *stride*, que define o deslocamento do filtro sobre a entrada. Esse fator influencia diretamente a resolução da saída e pode ser ajustado conforme os objetivos de compactação ou detalhamento da representação espacial.

O *stride* representa o passo com que o filtro é movido sobre a entrada. Um *stride* de 1 aplica o filtro em cada posição possível, enquanto valores maiores resultam em uma saída com resolução reduzida. *Strides* maiores podem ser utilizados para reduzir dimensionalidade ou implementar *downsampling* sem *pooling* adicional (DUMOULIN; VISIN, 2018). No entanto, quando se deseja realizar essa redução de forma mais controlada ou com objetivos estatísticos específicos, utilizam-se as camadas de *pooling*, frequentemente aplicadas após as convoluções.

As camadas de *pooling* geralmente são aplicadas após as camadas de convolução e detecção, com o objetivo de resumir estatisticamente as informações de regiões locais da

entrada. Essas camadas reduzem a largura e a altura dos mapas de ativação, mantendo inalterada a profundidade (número de canais). Dentre as operações mais comuns estão o *max pooling* e o *average pooling*, que atuam sobre vizinhanças locais gerando um único valor representativo por região. O *max pooling* seleciona o maior valor da vizinhança, destacando as ativações mais intensas, enquanto o *average pooling* calcula a média dos valores, proporcionando um resumo mais suave da região. Ambas as operações reduzem a dimensionalidade espacial da entrada, diminuem o custo computacional e ajudam no controle do *overfitting*, ao mesmo tempo que preservam as informações mais relevantes das ativações (DUMOULIN; VISIN, 2018).

Essas operações de *pooling* contribuem para a compactação dos dados ao longo da rede. Por outro lado, em algumas aplicações, como na reconstrução de imagens ou na segmentação semântica, é necessário realizar o caminho inverso, aumentando novamente a dimensão espacial das representações. Nesses casos, utiliza-se a convolução transposta, que permite a recuperação da resolução original dos dados de entrada.

A convolução transposta é usada para aumentar a dimensão espacial dos dados (*upsampling*). É frequentemente empregada em arquiteturas de segmentação semântica e *autoencoders* para reconstruir imagens a partir de representações compactas. Essa operação é o inverso da convolução no sentido de mapeamento de espaço, mas é implementada como uma convolução com espaçamento entre os elementos (DUMOULIN; VISIN, 2018).

Paralelamente ao avanço das CNNs, foram desenvolvidas arquiteturas alternativas projetadas para enfrentar os desafios impostos pelo aumento da profundidade das redes. Nesse contexto, as Redes Neurais Residuais (ResNet) surgem como uma ramificação das CNNs, com o objetivo de mitigar o problema da degradação do desempenho em redes muito profundas.

3.4 Redes Neurais Residuais

As Redes Neurais Residuais (ResNet) foram introduzidas por He et al. (2016) como uma solução para mitigar o problema de degradação de desempenho observado em redes neurais muito profundas. Esse problema ocorre, em grande parte, devido ao fenômeno do *vanishing gradient*, no qual os gradientes se tornam extremamente pequenos durante a retropropagação, dificultando o aprendizado efetivo nas camadas iniciais da rede (HE et al., 2016; LIU; ZHANG, 2020).

Diferentemente das redes tradicionais, onde cada camada tenta aprender diretamente um mapeamento $H(x)$, a ResNet reformula essa tarefa para que cada bloco aprenda uma função residual:

$$F(x) = H(x) - x, \quad (3.14)$$

de forma que a saída do bloco seja obtida como:

$$y = F(x) + x. \quad (3.15)$$

Essa abordagem parte da hipótese de que é mais fácil otimizar a diferença entre a entrada e a saída esperada do que aprender a transformação completa. Para isso, são utilizadas conexões de atalho que realizam mapeamentos de identidade e transferem diretamente a entrada para a saída do bloco.

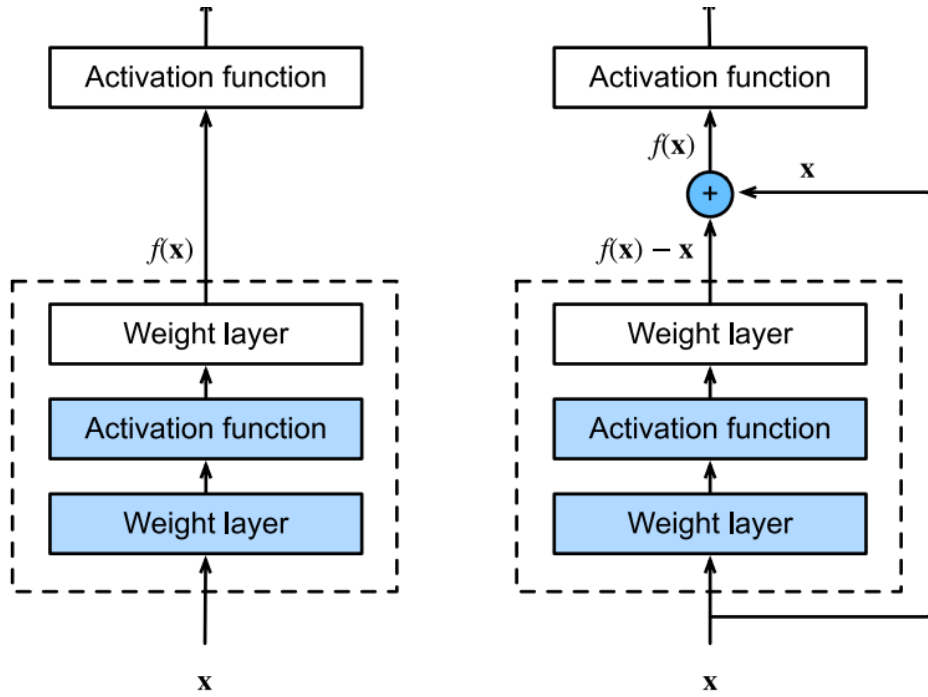


Figura 8 – Comparação entre bloco tradicional e bloco residual com conexão de atalho na arquitetura ResNet. Fonte: Zhang et al. (2023).

A Figura 8 ilustra a estrutura de um bloco residual. À esquerda, observa-se o fluxo tradicional de uma rede sem atalho, no qual a saída é resultado apenas da transformação interna do bloco. Já à direita, é apresentada a estrutura característica da ResNet, em que a entrada x é somada à saída da função residual $F(x)$, permitindo a preservação da identidade do sinal original ao longo da rede. Essa arquitetura favorece a propagação dos gradientes, facilita o treinamento e contribui para o aprendizado mais estável em redes profundas.

Nos casos em que a dimensão da entrada e da saída não coincidem, a ResNet utiliza projeções lineares com convoluções 1×1 , representadas por $W_s x$, para ajustar as dimensões:

$$y = F(x) + W_s x \quad (3.16)$$

Esses blocos são conhecidos como Residual *Building Blocks* (RBBs) e são compostos por camadas convolucionais seguidas de normalização e ativação ReLU. A versão ResNet-50, por exemplo, utiliza um padrão de bottleneck com três camadas: convoluções 1×1 para redução e expansão de dimensões e uma convolução 3×3 intermediária (HE et al., 2016).

O uso de atalhos diretos facilita a propagação do gradiente durante o treinamento, resolvendo de forma eficaz o problema de degradação observado em redes muito profundas. Conforme observado por He et al. (2016), redes com até 152 camadas podem ser treinadas com sucesso utilizando esse princípio, superando modelos anteriores como o VGG em desempenho e eficiência.

4 Redes Grafo Neurais

As Redes Grafo Neurais (Neural Graph Networks — GNNs) surgiram como uma classe de modelos para tratar dados com estrutura relacional, cada vez mais comuns em aplicações do mundo real. Diferentemente de estruturas tradicionais como vetores, sequências ou imagens, os grafos permitem representar entidades e suas interações de maneira explícita e flexível. Neste capítulo, são abordados os fundamentos matemáticos, os tipos de representações, as principais arquiteturas e filtros utilizados.

4.1 Conceitos Básicos

Grafos são estruturas fundamentais para representar relações par-a-par entre entidades em diversos domínios do mundo real. Diferentemente de dados estruturados convencionais, os nós em um grafo estão interligados, o que implica em uma dependência natural entre eles. No entanto, muitas técnicas tradicionais de aprendizado de máquina assumem que os dados são independentes e identicamente distribuídos (MA; TANG, 2021).

Um grafo pode ser denotado como $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, onde $\mathcal{V} = \{v_1, \dots, v_N\}$ é um conjunto de $N = |\mathcal{V}|$ nós e $\mathcal{E} = \{e_1, \dots, e_M\}$ é um conjunto de M arestas.

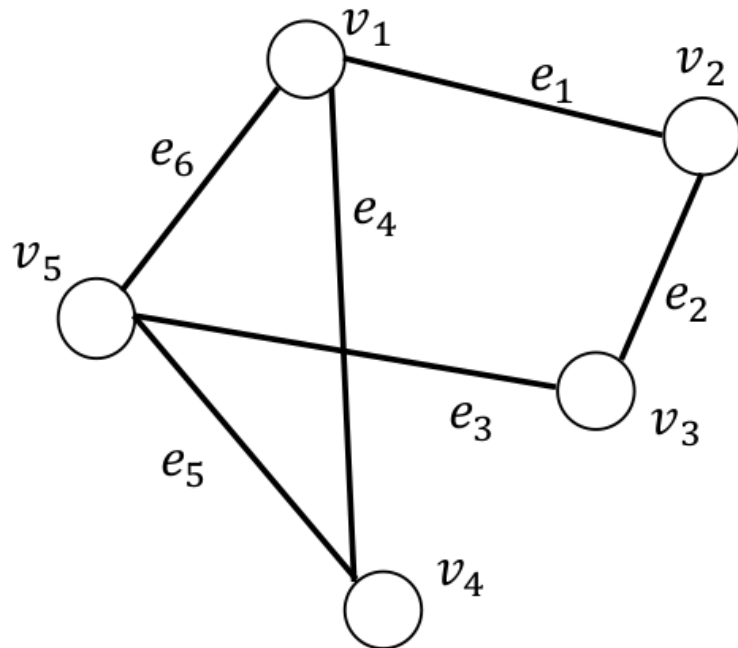


Figura 9 – Representação de grafo não direcionado com suas respectivas arestas e vértices.
Fonte: Ma e Tang (2021).

As arestas $e \in \mathcal{E}$ representam conexões entre pares de nós. Cada aresta e conecta

dois nós v_{1e} e v_{2e} , podendo ser representada como (v_{1e}, v_{2e}) . Em grafos direcionados, a aresta tem uma direção de v_{1e} para v_{2e} . Já em grafos não direcionados, a ordem dos nós não altera a definição da aresta, ou seja, $(v_{1e}, v_{2e}) = (v_{2e}, v_{1e})$. Neste capítulo, consideramos apenas grafos não direcionados.

Dado um grafo $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, a matriz de adjacência correspondente é denotada como $\mathbf{A} \in \{0, 1\}^{N \times N}$. A entrada $A_{i,j}$ representa a conectividade entre os nós v_i e v_j , sendo:

$$A_{i,j} = \begin{cases} 1, & \text{se } v_i \text{ é adjacente a } v_j; \\ 0, & \text{caso contrário.} \end{cases} \quad (4.1)$$

Em um grafo não direcionado, se v_i é adjacente a v_j , então v_j também é adjacente a v_i , o que implica que a matriz de adjacência $\mathbf{A}$ é simétrica, ou seja, $A_{i,j} = A_{j,i}$.

No contexto do aprendizado de máquina, a estrutura de grafos apresenta novos desafios e oportunidades. Técnicas tradicionais de classificação geralmente ignoram as relações entre os nós, assumindo amostras independentes. Em contrapartida, abordagens como a classificação coletiva utilizam informações dos vizinhos de um nó para inferir seu rótulo.

Uma direção alternativa é representar os nós por um conjunto de atributos que capturem sua posição e contexto no grafo, permitindo o uso de classificadores tradicionais. A construção de representações de nós eficazes é o ponto chave para o sucesso dessa abordagem, que tem se tornado dominante na área devido à sua capacidade de incorporar conhecimento estrutural utilizando modelos clássicos de aprendizado.

Essa estratégia de aprendizado baseada em representações reforça a importância de compreender as propriedades estruturais dos grafos. Para isso, é essencial conhecer medidas e características fundamentais dessas estruturas, como grau, conectividade, caminhos e centralidade, que fornecem informações cruciais sobre o papel e a influência de cada nó dentro do grafo.

4.1.1 Propriedades e Medidas

Grafos podem apresentar estruturas e propriedades variadas que revelam aspectos importantes sobre os dados representados. Nesta seção, são discutidos conceitos fundamentais como grau, conectividade, caminhos e medidas de centralidade.

O grau de um nó $v_i \in \mathcal{V}$ em um grafo $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ representa o número de conexões que ele possui com outros nós. Esse valor indica a quantidade de arestas incidentes em v_i , e pode ser formalmente definido como:

$$d(v_i) = \sum_{v_j \in \mathcal{V}} 1_{\mathcal{E}}(\{v_i, v_j\}), \quad (4.2)$$

onde a função indicadora $1_{\mathcal{E}}(\{v_i, v_j\})$ retorna 1 se existir uma aresta entre v_i e v_j , e 0 caso contrário:

$$1_{\mathcal{E}}(\{v_i, v_j\}) = \begin{cases} 1, & \text{se } (v_i, v_j) \in \mathcal{E}, \\ 0, & \text{caso contrário.} \end{cases} \quad (4.3)$$

Essa formulação considera todos os pares possíveis de vértices, contabilizando apenas aqueles que formam arestas com o nó v_i .

Além disso, o grau de um nó também pode ser obtido diretamente da matriz de adjacência $\mathbf{A} \in R^{N \times N}$, onde $N = |\mathcal{V}|$. Cada elemento $A_{i,j}$ indica se há uma aresta entre os nós v_i e v_j . Assim:

$$d(v_i) = \sum_{j=1}^N A_{i,j}. \quad (4.4)$$

Nesse caso, a soma dos elementos da linha i da matriz de adjacência fornece o número total de conexões do nó v_i , isto é, seu grau.

O conjunto de vizinhos de um nó, denotado por $\mathcal{N}(v_i)$, é formado por todos os nós diretamente conectados a v_i . Portanto, o número de elementos desse conjunto é igual ao grau do nó:

$$d(v_i) = |\mathcal{N}(v_i)|. \quad (4.5)$$

O grau total de um grafo, definido como a soma dos graus de todos os seus nós, é sempre igual ao dobro do número de arestas da estrutura:

$$\sum_{v_i \in \mathcal{V}} d(v_i) = 2 \cdot |\mathcal{E}|. \quad (4.6)$$

Essa relação ocorre porque cada aresta contribui com uma unidade para o grau de dois vértices, um em cada extremidade da aresta.

A análise do grau e da vizinhança de um nó fornece informações locais sobre sua importância e centralidade na rede. No entanto, para compreender de forma mais abrangente o papel dos vértices na estrutura global do grafo, é necessário considerar a conectividade entre as regiões, os caminhos existentes e a distribuição das conexões. Tais aspectos são fundamentais para avaliar a coesão estrutural da rede e a acessibilidade entre diferentes partes do grafo.

Antes de discutir conectividade, é importante compreender os seguintes conceitos:

- Caminhada: uma sequência alternada de nós e arestas, onde cada aresta conecta os nós imediatamente anterior e posterior.
- Trilha: uma caminhada onde as arestas não se repetem.
- Caminho: uma caminhada em que os nós são distintos.

Seja $\mathbf{A}$ a matriz de adjacência de um grafo. O elemento (i, j) da matriz $\mathbf{A}^n$ indica o número de caminhadas de comprimento n entre os nós v_i e v_j .

Um subgrafo $\mathcal{G}' = \{\mathcal{V}', \mathcal{E}'\}$ é uma componente conexa de $\mathcal{G}$ se existe pelo menos um caminho entre quaisquer pares de nós de $\mathcal{V}'$, e os nós em $\mathcal{V}'$ não são adjacentes a nós fora de $\mathcal{V}'$. Um grafo é dito conexo se possui exatamente uma componente conexa.

Entre dois nós quaisquer podem existir múltiplos caminhos com comprimentos diferentes. Definimos o menor desses caminhos como:

$$p_{st}^{sp} = \arg \min_{p \in \mathcal{P}_{st}} |p|, \quad (4.7)$$

onde $\mathcal{P}_{st}$ representa o conjunto de caminhos entre v_s e v_t .

Com base nisso, o diâmetro de um grafo é definido como o maior comprimento dentre todos os caminhos mínimos:

$$\text{diâmetro}(\mathcal{G}) = \max_{v_s, v_t \in \mathcal{V}} \min_{p \in \mathcal{P}_{st}} |p|. \quad (4.8)$$

A centralidade de um nó representa sua importância dentro da estrutura do grafo. Uma forma simples de mensurá-la é através do número de conexões diretas que o nó possui, o que é chamado de centralidade por grau:

$$c_d(v_i) = d(v_i) = \sum_{j=1}^N A_{i,j}. \quad (4.9)$$

No entanto, essa medida considera todos os vizinhos com o mesmo peso. Uma abordagem mais sofisticada é a centralidade por autovetor, que pondera a importância dos vizinhos. A centralidade $c_e(v_i)$ de um nó é definida como:

$$c_e(v_i) = \frac{1}{\lambda} \sum_{j=1}^N A_{i,j} \cdot c_e(v_j), \quad (4.10)$$

ou em forma matricial:

$$\mathbf{c}_e = \frac{1}{\lambda} \mathbf{A} \cdot \mathbf{c}_e, \quad (4.11)$$

onde $\mathbf{c}_e \in R^N$ é o vetor de centralidades e λ é o maior autovalor da matriz $\mathbf{A}$. Pelo teorema de Perron–Frobenius, a matriz de adjacência garante um único autovalor dominante e um autovetor com elementos positivos, o que permite interpretar $\mathbf{c}_e$ como um vetor de relevância nodal.

4.2 Teoria Espectral de Grafos

A teoria espectral de grafos estuda as propriedades estruturais de um grafo por meio da análise dos autovalores e autovetores de sua matriz Laplaciana. Essa abordagem fornece informações fundamentais sobre conectividade, fluxo de informação e agrupamentos dentro do grafo.

Além da matriz de adjacência $\mathbf{A}$, uma outra forma de representar um grafo é por meio de sua matriz Laplaciana. Para um grafo $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, com matriz de adjacência $\mathbf{A}$ e matriz de grau diagonal $\mathbf{D} = \text{diag}(d(v_1), \dots, d(v_N))$, a matriz Laplaciana é definida como:

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \quad (4.12)$$

Note que $\mathbf{L}$ é uma matriz simétrica, já que $\mathbf{D}$ e $\mathbf{A}$ são simétricas em grafos não direcionados. Além disso, a matriz Laplaciana admite uma versão normalizada que é amplamente utilizada em algoritmos espectrais:

$$\mathbf{L} = \mathbf{D}^{-1/2}(\mathbf{D} - \mathbf{A})\mathbf{D}^{-1/2} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} \quad (4.13)$$

Seja $\mathbf{f} \in R^N$ um vetor cujos elementos $f[i]$ estão associados a cada nó v_i . O produto da matriz Laplaciana pela função $\mathbf{f}$ gera um novo vetor $\mathbf{h}$, definido por:

$$\mathbf{h} = \mathbf{L}\mathbf{f} = (\mathbf{D} - \mathbf{A})\mathbf{f} = \mathbf{D}\mathbf{f} - \mathbf{A}\mathbf{f} \quad (4.14)$$

O i -ésimo elemento do vetor $\mathbf{h}$ pode ser escrito como:

$$\begin{aligned} \mathbf{h}[i] &= d(v_i) \cdot f[i] - \sum_{j=1}^N A_{i,j} \cdot f[j] \\ &= d(v_i) \cdot f[i] - \sum_{v_j \in \mathcal{N}(v_i)} A_{i,j} \cdot f[j] \\ &= \sum_{v_j \in \mathcal{N}(v_i)} (f[i] - f[j]) \end{aligned} \quad (4.15)$$

Ou seja, $\mathbf{h}[i]$ representa a soma das diferenças entre o valor de $f[i]$ e os valores de seus vizinhos. A expressão quadrática $\mathbf{f}^\top \mathbf{L}\mathbf{f}$ fornece uma métrica global de variação entre os valores associados a nós adjacentes:

$$\begin{aligned}
\mathbf{f}^\top \mathbf{L} \mathbf{f} &= \sum_{v_i \in \mathcal{V}} f[i] \sum_{v_j \in \mathcal{N}(v_i)} (f[i] - f[j]) \\
&= \sum_{v_i \in \mathcal{V}} \sum_{v_j \in \mathcal{N}(v_i)} f[i]^2 - f[i]f[j] \\
&= \sum_{v_i \in \mathcal{V}} \sum_{v_j \in \mathcal{N}(v_i)} \left(\frac{1}{2} f[i]^2 - f[i]f[j] + \frac{1}{2} f[j]^2 \right) \\
&= \frac{1}{2} \sum_{v_i \in \mathcal{V}} \sum_{v_j \in \mathcal{N}(v_i)} (f[i] - f[j])^2
\end{aligned} \tag{4.16}$$

A equação acima revela que $\mathbf{f}^\top \mathbf{L} \mathbf{f}$ mede a diferença total entre os valores de nós conectados. Como é uma soma de quadrados, essa expressão é sempre não negativa para qualquer vetor real $\mathbf{f}$, indicando que a matriz Laplaciana é semi-definida positiva.

Esse resultado fundamenta várias técnicas espectrais, como o particionamento de grafos, onde os autovalores e autovetores de $\mathbf{L}$ são utilizados para identificar grupos de nós densamente conectados.

4.3 Grafos Complexos

Nas seções anteriores, discutimos os grafos simples e suas propriedades fundamentais. No entanto, nas aplicações do mundo real, os grafos podem ser muito mais complexos. Nesta subseção, descrevemos formalmente os principais tipos de grafos complexos que aparecem com frequência em contextos práticos.

Os grafos simples considerados anteriormente são homogêneos, contendo apenas um tipo de nó e um tipo de aresta. Entretanto, em muitos cenários reais, deseja-se modelar múltiplos tipos de relações entre diferentes tipos de entidades. Isso nos leva à definição de grafos heterogêneos.

Um grafo heterogêneo $\mathcal{G}$ consiste em um conjunto de nós $\mathcal{V} = \{v_1, \dots, v_N\}$ e um conjunto de arestas $\mathcal{E} = \{e_1, \dots, e_M\}$, onde cada nó e cada aresta estão associados a um tipo. Seja $\mathcal{T}_n$ o conjunto de tipos de nós e $\mathcal{T}_e$ o conjunto de tipos de arestas. Existem duas funções de mapeamento:

$$\phi_n : \mathcal{V} \rightarrow \mathcal{T}_n, \quad \phi_e : \mathcal{E} \rightarrow \mathcal{T}_e \tag{4.17}$$

que associam, respectivamente, cada nó e cada aresta ao seu tipo.

Um caso particular de grafo heterogêneo é o grafo bipartido. Nesse tipo de grafo, os nós podem ser divididos em dois subconjuntos disjuntos $\mathcal{V}_1$ e $\mathcal{V}_2$, de forma que todas as arestas conectam apenas elementos entre esses dois conjuntos.

Dado um grafo $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, ele é bipartido se, e somente se, $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, com $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$, e toda aresta $e = (v_e^1, v_e^2) \in \mathcal{E}$ conecta um nó $v_e^1 \in \mathcal{V}_1$ a um nó $v_e^2 \in \mathcal{V}_2$.

Outra classe de grafos complexos é composta pelos grafos multidimensionais, nos quais diferentes tipos de relações são modelados como dimensões distintas.

Um grafo multidimensional consiste em um conjunto de N nós $\mathcal{V} = \{v_1, \dots, v_N\}$ e D conjuntos de arestas $\{\mathcal{E}_1, \dots, \mathcal{E}_D\}$, onde cada conjunto $\mathcal{E}_d$ representa um tipo específico de relação. Cada tipo pode ser representado por uma matriz de adjacência:

$$\mathbf{A}^{(d)} \in R^{N \times N}, \quad d = 1, \dots, D \quad (4.18)$$

A entrada $\mathbf{A}_{i,j}^{(d)} = 1$ indica a existência de uma aresta entre v_i e v_j na dimensão d ; caso contrário, $\mathbf{A}_{i,j}^{(d)} = 0$.

Com o crescimento das redes sociais, tornou-se comum o uso de grafos com relações positivas e negativas, chamados de grafos assinados. Seja $\mathcal{G} = \{\mathcal{V}, \mathcal{E}^+, \mathcal{E}^-\}$ um grafo assinado, onde $\mathcal{V} = \{v_1, \dots, v_N\}$ é o conjunto de nós, e $\mathcal{E}^+, \mathcal{E}^- \subset \mathcal{V} \times \mathcal{V}$ são os conjuntos de arestas positivas e negativas, com $\mathcal{E}^+ \cap \mathcal{E}^- = \emptyset$. A matriz de adjacência assinada $\mathbf{A}$ é definida por:

$$\mathbf{A}_{i,j} = \begin{cases} 1, & \text{se } (v_i, v_j) \in \mathcal{E}^+ \\ -1, & \text{se } (v_i, v_j) \in \mathcal{E}^- \\ 0, & \text{caso contrário} \end{cases} \quad (4.19)$$

Outra estrutura que permite capturar relações de ordem superior são os hipergrafos, que generalizam arestas para conectar múltiplos nós simultaneamente. Seja $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{W}\}$ um hipergrafo, onde $\mathcal{V}$ é o conjunto de nós, $\mathcal{E}$ é o conjunto de hiperarestas, e $\mathbf{W} \in R^{|\mathcal{E}| \times |\mathcal{E}|}$ é uma matriz diagonal de pesos. Um hipergrafo pode ser representado por uma matriz de incidência $\mathbf{H} \in R^{|\mathcal{V}| \times |\mathcal{E}|}$, com:

$$\mathbf{H}_{i,j} = \begin{cases} 1, & \text{se } v_i \in e_j \\ 0, & \text{caso contrário} \end{cases} \quad (4.20)$$

O grau de um nó é dado por:

$$d(v_i) = \sum_{j=1}^{|\mathcal{E}|} \mathbf{H}_{i,j}, \quad d(e_j) = \sum_{i=1}^{|\mathcal{V}|} \mathbf{H}_{i,j} \quad (4.21)$$

Utilizamos $\mathbf{D}_v$ e $\mathbf{D}_e$ para representar as matrizes diagonais de grau dos nós e das hiperarestas, respectivamente.

Por fim, os grafos dinâmicos representam estruturas em constante evolução, onde nós e arestas podem aparecer ou desaparecer ao longo do tempo. Um grafo dinâmico $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ consiste em um conjunto de nós $\mathcal{V} = \{v_1, \dots, v_N\}$ e um conjunto de arestas $\mathcal{E} = \{e_1, \dots, e_M\}$, onde cada nó e cada aresta está associado a uma referência temporal. Funções de mapeamento ϕ_v e ϕ_e indicam os instantes em que os nós e arestas emergem:

$$\phi_v : \mathcal{V} \rightarrow T, \quad \phi_e : \mathcal{E} \rightarrow T \quad (4.22)$$

Em cenários práticos, pode-se registrar apenas observações discretas de tempo t , resultando em múltiplas “fotos” do grafo ao longo do tempo: $\mathcal{G}_t$. Esse tipo de estrutura é conhecido como grafo dinâmico discreto.

4.4 Fundamentos das Redes Grafo Neurais

As redes neurais profundas tradicionais, como as CNNs, foram projetadas para lidar com dados estruturados em dados regulares, como imagens. No entanto, muitos dados do mundo real estão organizados como grafos, cuja estrutura é irregular e não pode ser diretamente processada por essas arquiteturas. Nesse contexto, surgem as GNNs, uma classe de métodos que permite aplicar modelos de aprendizado profundo a dados estruturados em grafos.

As GNNs podem ser vistas como um processo de aprendizado de representações em grafos. Em tarefas centradas nos nós, o objetivo das GNNs é aprender boas representações para cada nó individual, facilitando tarefas como classificação ou predição de conexões. Em tarefas centradas no grafo, busca-se extrair uma representação global do grafo inteiro, sendo o aprendizado das características dos nós apenas uma etapa intermediária.

Esse processo de aprendizado explora tanto as características dos nós quanto a estrutura do grafo, definida por sua matriz de adjacência $\mathbf{A} \in \mathbb{R}^{N \times N}$. As representações dos nós antes e depois do processo de filtragem são dadas por $\mathbf{F}^{(\text{if})} \in \mathbb{R}^{N \times d_{\text{if}}}$ e $\mathbf{F}^{(\text{of})} \in \mathbb{R}^{N \times d_{\text{of}}}$, respectivamente, onde os índices denotam entrada e saída da filtragem. A operação de filtragem pode ser expressa por:

$$\mathbf{F}^{(\text{of})} = h(\mathbf{A}, \mathbf{F}^{(\text{if})}) \quad (4.23)$$

A função $h(\cdot, \cdot)$ representa um filtro em grafos, responsável por combinar as informações estruturais e atributos dos nós para atualizar suas representações. A Figura 10 ilustra esse processo, onde apenas as características dos nós são modificadas, enquanto a estrutura do grafo é mantida.

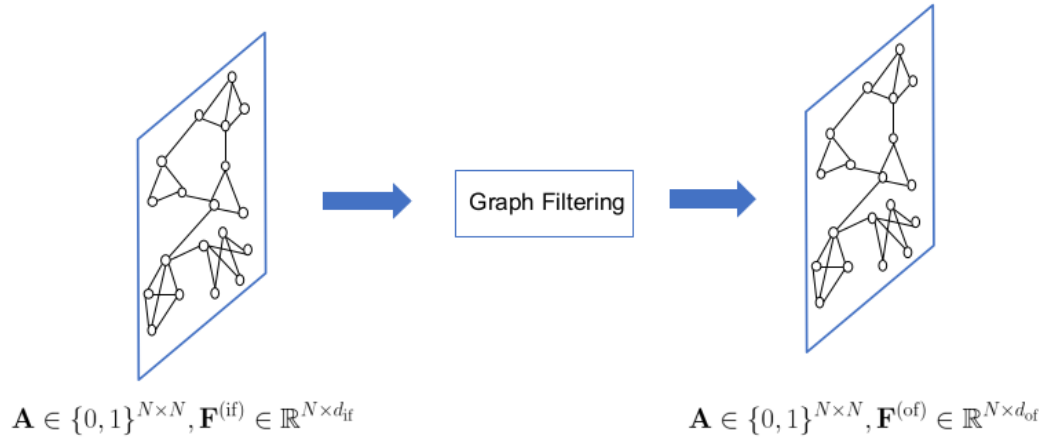


Figura 10 – Processo de filtragem em grafos para atualização das representações nodais.
Fonte: Ma e Tang (2021).

Para tarefas centradas nos nós, as operações de filtragem são suficientes. É comum empilhar múltiplas camadas de filtragem consecutivamente, sendo a saída de uma camada a entrada da próxima. Esse empilhamento progressivo permite capturar informações de vizinhança de ordens maiores a cada camada.

Por outro lado, para tarefas centradas no grafo, como classificação de grafos, são necessárias operações adicionais que sintetizam informações dos nós para produzir uma representação do grafo como um todo. Inspirando-se nas CNNs, operações de *pooling* em grafos foram propostas para reduzir a dimensionalidade do grafo e agregar informações dos nós.

Como a estrutura dos grafos é irregular, o processo de *pooling* deve considerar a topologia do grafo durante a agregação. A operação de *pooling* reduz o número de nós no grafo, gerando uma nova matriz de adjacência e novas representações dos nós. O processo pode ser descrito como:

$$\mathbf{A}^{(op)}, \mathbf{F}^{(op)} = \text{pool}(\mathbf{A}^{(ip)}, \mathbf{F}^{(ip)}) \quad (4.24)$$

onde $\mathbf{A}^{(ip)} \in \{0, 1\}^{N_{ip} \times N_{ip}}$ e $\mathbf{F}^{(ip)} \in \mathbb{R}^{N_{ip} \times d_{ip}}$ são a matriz de adjacência e a matriz de características antes do pooling, e $\mathbf{A}^{(op)} \in \{0, 1\}^{N_{op} \times N_{op}}$, $\mathbf{F}^{(op)} \in \mathbb{R}^{N_{op} \times d_{op}}$ representam os resultados após o pooling, com $N_{op} < N_{ip}$. A Figura 11 ilustra esse processo, no qual tanto as características nodais quanto a estrutura do grafo são reduzidas. Esse tipo de operação é utilizado para condensar a informação presente no grafo original, produzindo uma versão mais compacta e representativa, com menor número de nós e conexões.

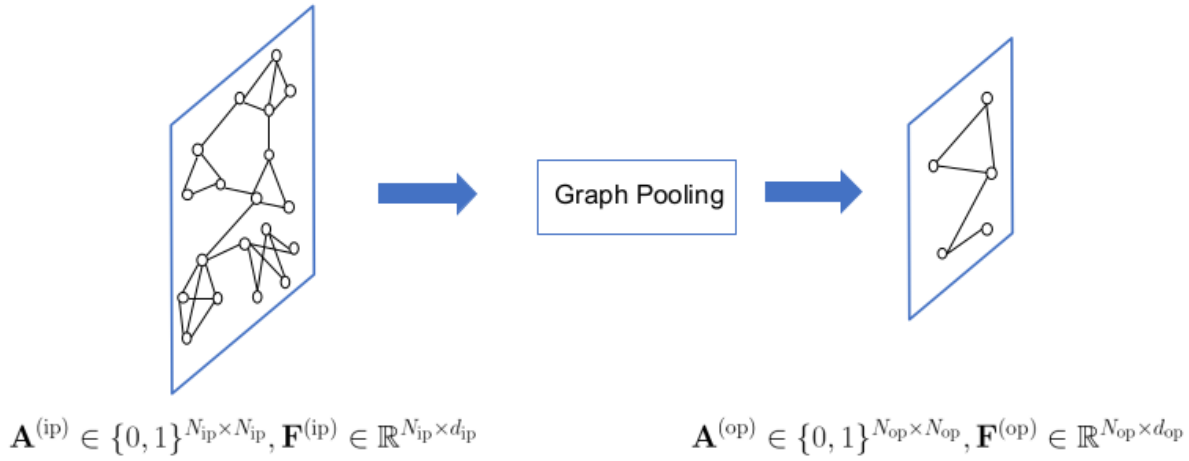


Figura 11 – Operação de pooling em grafos para redução da dimensionalidade estrutural e nodal. Fonte: Ma e Tang (2021).

A arquitetura de uma GNN típica consiste na combinação de camadas de filtragem e pooling. Para tarefas centradas nos nós, o modelo é composto apenas por camadas de filtragem empilhadas, enquanto para tarefas centradas no grafo, as camadas de pooling são intercaladas com blocos de camadas de filtragem, permitindo a extração hierárquica de características globais.

4.5 Filtros em Grafos

A tarefa de projetar filtros sobre grafos pode ser abordada por diferentes perspectivas, que podem ser divididas, de forma geral, em duas categorias principais: filtros baseados no domínio espacial e filtros baseados no domínio espectral. Filtros espaciais aproveitam diretamente a estrutura do grafo, isto é, as conexões entre os nós, para refinar suas características no próprio domínio do grafo. Em contraste, os filtros espectrais utilizam a teoria espectral dos grafos para realizar a filtragem no domínio das frequências.

Essas abordagens não são mutuamente exclusivas. Em muitos casos, é possível interpretar certos filtros espectrais como filtros espaciais, especialmente quando o operador espectral é aproximado por polinômios locais.

4.5.1 Filtragem Espectral em Grafos

Como ilustrado na Figura 12, a ideia central da filtragem espectral em grafos é modular as componentes de frequência de um sinal definido sobre o grafo, amplificando ou suprimindo determinadas faixas espectrais.

Dado um sinal $\mathbf{f} \in \mathbb{R}^N$ definido sobre um grafo $\mathcal{G}$, o primeiro passo consiste em aplicar a Transformada de Fourier para grafos, definida por:

$$\hat{\mathbf{f}} = \mathbf{U}^\top \mathbf{f} \quad (4.25)$$

onde $\mathbf{U} \in R^{N \times N}$ é a matriz de autovetores da matriz Laplaciana $\mathbf{L}$ do grafo, ou seja, $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$, com $\mathbf{\Lambda}$ sendo a matriz diagonal de autovalores.

As componentes $\hat{\mathbf{f}}[i]$ representam os coeficientes da base espectral associada ao autovalor λ_i . Para aplicar um filtro espectral $\gamma(\cdot)$, modulamos esses coeficientes da seguinte forma:

$$\hat{\mathbf{f}}'[i] = \hat{\mathbf{f}}[i] \cdot \gamma(\lambda_i), \quad \text{para } i = 1, \dots, N \quad (4.26)$$

Esse processo pode ser escrito de maneira matricial como:

$$\hat{\mathbf{f}}' = \gamma(\mathbf{\Lambda}) \cdot \hat{\mathbf{f}} = \gamma(\mathbf{\Lambda}) \cdot \mathbf{U}^\top \mathbf{f} \quad (4.27)$$

onde $\gamma(\mathbf{\Lambda})$ é a aplicação da função filtro sobre a diagonal da matriz de autovalores:

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}, \quad \gamma(\mathbf{\Lambda}) = \begin{pmatrix} \gamma(\lambda_1) & & \\ & \ddots & \\ & & \gamma(\lambda_N) \end{pmatrix} \quad (4.28)$$

A reconstrução do sinal filtrado no domínio dos grafos é realizada através da Transformada de Fourier Inversa para grafos:

$$\mathbf{f}' = \mathbf{U}\hat{\mathbf{f}}' = \mathbf{U} \cdot \gamma(\mathbf{\Lambda}) \cdot \mathbf{U}^\top \mathbf{f} \quad (4.29)$$

Portanto, a operação de filtragem espectral pode ser interpretada como a aplicação do operador:

$$\mathbf{H}_\gamma = \mathbf{U} \cdot \gamma(\mathbf{\Lambda}) \cdot \mathbf{U}^\top$$

à entrada $\mathbf{f}$. A função $\gamma(\cdot)$ é frequentemente chamada de filtro espectral, pois determina como cada frequência do sinal será tratada. Em casos extremos, se $\gamma(\lambda_i) = 0$, então a componente espectral $\hat{\mathbf{f}}'[i]$ é completamente removida da reconstrução.

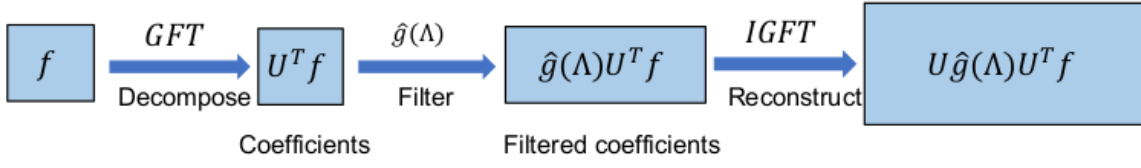


Figura 12 – Processo de decomposição, modulação espectral e reconstrução de sinais em grafos. Fonte: Ma e Tang (2021).

O processo de filtragem espectral é útil não apenas pela interpretação em termos de frequência, mas também pela capacidade de manipular as propriedades de suavidade ou oscilação dos sinais no grafo, permitindo, por exemplo, enfatizar regiões homogêneas ou bordas estruturais da topologia.

4.5.2 Filtragem Espacial em Grafos

Os filtros espaciais representam uma abordagem fundamental no projeto de redes neurais em grafos, realizando a agregação de informações com base diretamente na estrutura topológica do grafo. Diferente dos filtros espectrais, que atuam no domínio de frequências dos autovalores da matriz Laplaciana, os filtros espaciais operam no domínio local dos nós, explorando suas vizinhanças imediatas.

Historicamente, filtros baseados no domínio espacial foram propostos ainda antes da popularização do aprendizado profundo. O conceito de redes neurais em grafos foi introduzido por Scarselli, propondo um mecanismo iterativo de atualização das características de um nó com base nas informações dos seus vizinhos.

No modelo original, cada nó v_i do grafo é associado a um rótulo l_i , que serve como informação inicial fixa no processo de filtragem. As características do grafo são representadas por uma matriz de atributos $\mathbf{F}$, onde $\mathbf{F}_i$ é a linha correspondente ao nó v_i . A saída do filtro é uma nova matriz de características $\mathbf{F}'$, definida como:

$$\mathbf{F}'_i = \sum_{v_j \in \mathcal{N}(v_i)} g(l_i, \mathbf{F}_j, l_j) \quad (4.30)$$

A função $g(\cdot)$ é chamada de função de transição local e é tipicamente modelada por uma rede neural do tipo *feedforward*. Essa função é compartilhada entre todos os nós do grafo, o que garante a invariância da operação de filtragem em relação à ordem dos nós.

O processo descrito considera apenas os nós que estão diretamente conectados ao nó central v_i , o que garante uma operação local e eficiente. Esse tipo de arquitetura é especialmente útil para tarefas de classificação de nós, onde a influência da vizinhança direta é crucial para o aprendizado de representações discriminativas.

A ideia original foi posteriormente generalizada, originando diversas variantes de filtros espaciais que compõem as arquiteturas modernas de GNNs. Em geral, esses filtros operam por meio de uma agregação de características da vizinhança seguida de uma transformação linear paramétrica. Um exemplo comum é o filtro da *Graph Convolutional Network*, que pode ser descrito como:

$$\mathbf{H}^{(l+1)} = \sigma(\widetilde{\mathbf{D}}^{-1/2} \widetilde{\mathbf{A}} \widetilde{\mathbf{D}}^{-1/2} \mathbf{H}^{(l)} \boldsymbol{\Theta}^{(l)}) \quad (4.31)$$

onde $\widetilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$ é a matriz de adjacência com auto-laços, $\widetilde{\mathbf{D}}$ é a matriz de graus correspondente, $\mathbf{H}^{(l)}$ representa as características dos nós na camada l , $\boldsymbol{\Theta}^{(l)}$ é a matriz de pesos aprendíveis, e $\sigma(\cdot)$ é uma função de ativação não-linear (como ReLU). Esse filtro realiza uma normalização simétrica da matriz de adjacência e aplica uma transformação linear às características dos nós.

O sucesso das GNNs modernas reside na capacidade desses filtros espaciais de capturar dependências locais em grafos arbitrários, utilizando mecanismos de agregação flexíveis, eficientes e escaláveis.

4.6 Convolução de Chebyshev

A convolução de Chebyshev representa uma abordagem eficiente de convolução espectral para grafos, possibilitando a generalização das CNNs tradicionais para domínios não euclidianos. Sua principal inovação está na utilização de uma expansão de polinômios de Chebyshev, o que evita a necessidade de decomposição espectral explícita (DEFFERRARD; BRESSION; VANDERGHEYNST, 2016).

Considere um grafo não direcionado $G = (V, E, W)$, onde V é o conjunto de vértices, E o conjunto de arestas e $W \in \mathbb{R}^{n \times n}$ a matriz de adjacência ponderada. O Laplaciano normalizado do grafo é dado por:

$$\mathbf{L} = \mathbf{I}_n - \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}, \quad (4.32)$$

onde $\mathbf{D}$ é a matriz de graus e $\mathbf{I}_n$ a matriz identidade.

Como $\mathbf{L}$ é simétrica e semidefinida positiva, ela admite uma decomposição espectral:

$$\mathbf{L} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^T, \quad (4.33)$$

sendo $\mathbf{U}$ os autovetores e $\boldsymbol{\Lambda}$ a matriz diagonal de autovalores.

A convolução espectral de um sinal $x \in \mathbb{R}^n$ com um filtro espectral g_θ é definida por:

$$y = g_\theta(\mathbf{L})x = \mathbf{U}g_\theta(\mathbf{\Lambda})\mathbf{U}^T x. \quad (4.34)$$

Para evitar a decomposição espectral computacionalmente custosa, $g_\theta(\mathbf{L})$ é aproximado por uma expansão de polinômios de Chebyshev:

$$g_\theta(\mathbf{L}) = \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\mathbf{L}}), \quad (4.35)$$

onde θ_k são coeficientes aprendidos, T_k é o polinômio de Chebyshev de ordem k , e $\tilde{\mathbf{L}} = \frac{2\mathbf{L}}{\lambda_{\max}} - \mathbf{I}$ é o Laplaciano reescalado com $\lambda_{\max}$ sendo o maior autovalor de $\mathbf{L}$.

Os polinômios de Chebyshev são definidos recursivamente como:

$$T_0(x) = 1, \quad (4.36)$$

$$T_1(x) = x, \quad (4.37)$$

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), \quad \text{para } k \geq 2. \quad (4.38)$$

A convolução Chebyshev é computacionalmente eficiente, com complexidade $\mathcal{O}(K|E|)$, e permite uma implementação localizada, ou seja, o filtro atua apenas sobre os K -vizinhos de um nó. Isso garante extração de características locais similares às convoluções em imagens e evita overfitting, mesmo com menor quantidade de dados rotulados.

Essa abordagem tornou-se uma base fundamental para o desenvolvimento de arquiteturas mais avançadas como a GCN de Kipf e Welling (KIPF; WELLING, 2017), que representa um caso especial da ChebConv com $K = 1$.

4.7 Tarefas Computacionais em Grafos

Diversas tarefas computacionais têm sido propostas para explorar estruturas de grafos. Essas tarefas podem ser, de modo geral, divididas em duas grandes categorias: tarefas focadas em nós e tarefas focadas em grafos. A distinção entre elas baseia-se na granularidade da amostra de dados, enquanto nas tarefas focadas em nós os dados são representados por um único grafo e cada nó é tratado como uma amostra, nas tarefas focadas em grafos cada amostra de dado é, por si só, um grafo completo.

Tarefas focadas em nós são amplamente estudadas e incluem problemas como classificação de nós, ranqueamento, predição de ligações e detecção de comunidades. Em muitos grafos do mundo real, os nós estão associados a informações adicionais que podem ser representadas como rótulos ou categorias. Esses rótulos ajudam a caracterizar os nós e são valiosos em diversas aplicações práticas, como análise de redes sociais, recomendação e

biologia computacional. No entanto, na prática, nem todos os nós estão rotulados. Assim, surge a tarefa de classificação de nós, cujo objetivo é inferir os rótulos dos nós não rotulados com base na topologia do grafo e nas informações disponíveis dos nós vizinhos.

Tarefas focadas em grafos, por outro lado, assumem que cada dado de entrada é um grafo completo. Exemplos dessas tarefas incluem classificação de grafos, correspondência de grafos e geração de grafos. A classificação de grafos busca atribuir um rótulo a grafos inteiros com base em sua estrutura e atributos. Esse tipo de tarefa surge com frequência em áreas como química (classificação de compostos), bioinformática (análise de proteínas) e análise de documentos (classificação de estruturas linguísticas).

Diferente da classificação de nós, onde cada nó em um grafo representa uma amostra, na classificação de grafos cada grafo $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i)$ representa uma única amostra. A complexidade estrutural dos grafos exige métodos especializados, já que classificadores tradicionais não conseguem manipular diretamente representações com topologias variáveis e interdependentes.

5 Metodologia

Neste capítulo, é apresentada uma abordagem cujo foco principal é a segmentação das áreas afetadas por AVC hemorrágico em imagens de TC do crânio. Para identificar automaticamente os casos a serem segmentados, a metodologia proposta inclui uma etapa preliminar de classificação das imagens em três categorias: normal, isquêmico e hemorrágico. A segmentação é então aplicada exclusivamente às imagens identificadas como hemorrágicas, visando delimitar com precisão as regiões comprometidas.

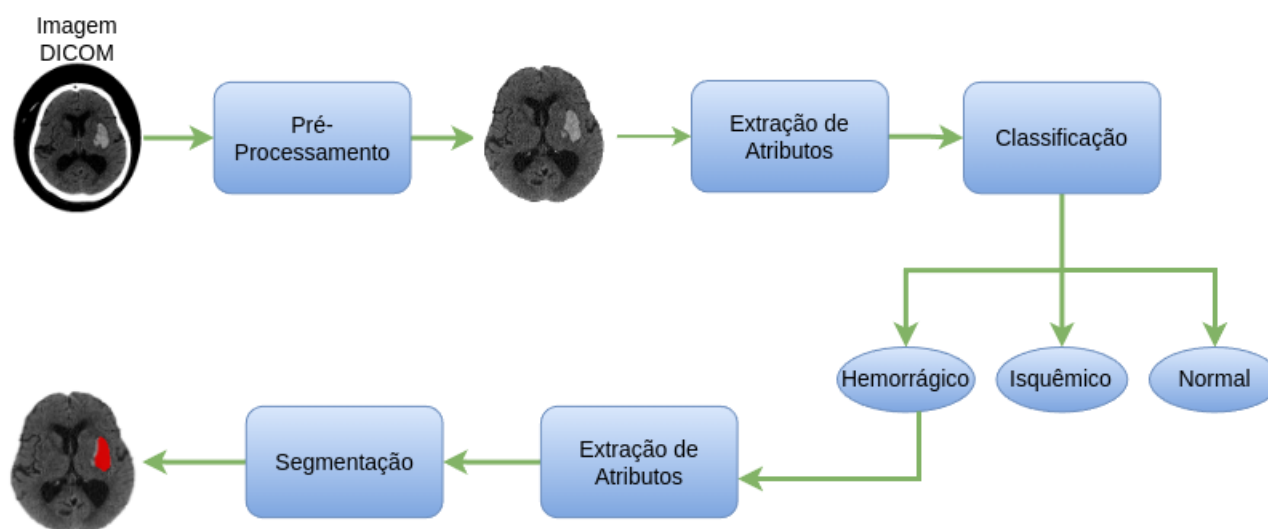


Figura 13 – Fluxo da Metodologia Proposta para Classificação e Segmentação de AVCs em Imagens de TC. Fonte: Autor.

O processo de classificação e segmentação mencionado possui as seguintes etapas gerais: pré-processamento, extração de atributos, classificação e segmentação. Na figura 13 é apresentada a metodologia desta pesquisa. A principal contribuição deste trabalho é a utilização de GCNs para as tarefas de classificação e segmentação.

5.1 Base de Dados

A base de dados utilizada neste trabalho é composto por um total de 473 imagens de TC do cérebro, obtidas com um tomógrafo HiSpeed do fabricante General Electric. Desse total, 174 imagens são de cérebros saudáveis e 299 com AVC, dos quais 157 do tipo isquêmico e 142 do tipo hemorrágico. As imagens de TC foram fornecidas pela Clínica Trajano Almeida e já foram utilizadas em estudos anteriores.

As fatias tomográficas foram feitas a partir das imagens adquiridas com base no

plano axial, sob as seguintes condições: espessura da fatia de 0,7 mm, campo de visão de 230 mm, tensão do tubo de 120 kV, corrente do tubo de 80 mA, dimensões de 512 x 512 pixels e janela de voxel de 0,585 x 0,585 x 1,5 mm, 16 bits, faixa de 40 HU e nível igual a 80 HU. As imagens foram armazenadas utilizando o formato *Digital Imaging and Communications in Medicine* (DICOM), que é o padrão para imagens médicas digitais criado pelo *American College of Radiology* (ACR) e pela *National Electrical Manufacturers Association* (NEMA). Na figura 14 são exibidos alguns exemplos de imagens que estão no conjunto de dados.

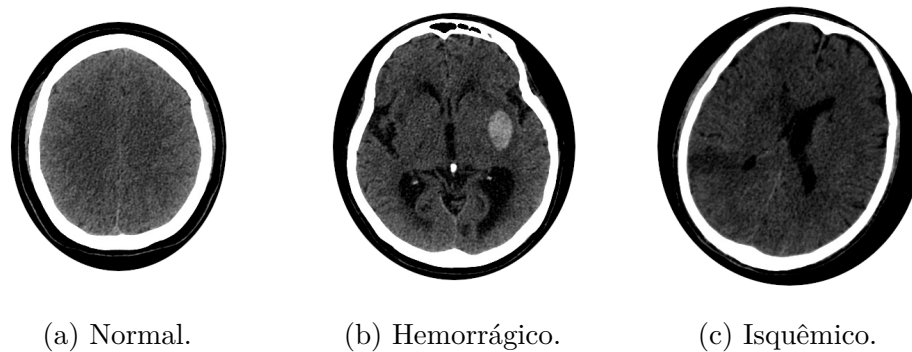


Figura 14 – Exemplos de exames de TC pertencentes a cada classe do dataset. Fonte: Autor.

Neste trabalho, após a leitura dos arquivos, os pixels das imagens, originalmente codificados em 16 bits, com valores variando de -32.768 a 32.768, foram convertidos para 8 bits, cujo intervalo vai de 0 a 255, com o objetivo de simplificar e agilizar o processamento das imagens. Esses valores de intensidade estão diretamente associados à escala de Hounsfield (HU), que representa uma medida quantitativa da radiodensidade dos tecidos.

A fim de preparar as imagens para as etapas posteriores de análise e extração de atributos, faz-se necessário um pré-processamento cuidadoso. Esse processo busca isolar as regiões de interesse mais relevantes para a tarefa de segmentação e eliminar informações irrelevantes ou ruidosas.

5.2 Pré-Processamento

A parte interna do crânio é a região mais relevante da TC para o diagnóstico do AVC. Com o objetivo de segmentar essa ROI, foi adotado o método descrito por Souza et al. (2019). A Figura 15 ilustra visualmente cada uma das etapas envolvidas nesse processo.

Inicialmente, aplica-se um filtro de mediana 3x3 com o intuito de suavizar a imagem e reduzir possíveis ruídos. Em seguida, para isolar a região cerebral, aplica-se um limiar baseado na escala de Hounsfield (HU), que permite distinguir o crânio, com valores superiores a 500 HU, das estruturas cerebrais, tipicamente abaixo de 100 HU. Nessa etapa,

utiliza-se o valor de 300 HU como limiar para destacar a estrutura óssea. Em seguida, por meio do algoritmo de Otsu, é gerada uma máscara binária que realça o crânio, a qual é então subtraída da imagem original.

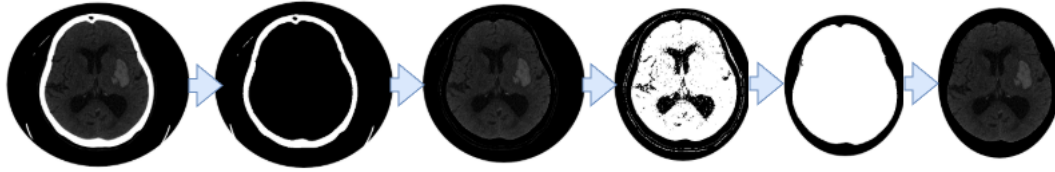


Figura 15 – Etapas do pré-processamento aplicadas às imagens de TC. Fonte: Autor.

Após a remoção do crânio, aplica-se novamente o limiar de Otsu para destacar apenas a região interna. A maior componente conexa é então identificada, representando a massa encefálica. Essa região é utilizada para criar uma nova máscara, que é aplicada à imagem original para obter a ROI final, contendo apenas o cérebro e eliminando ruídos e partes não relevantes da imagem. Essa etapa é essencial para garantir que os métodos subsequentes de extração de características e análise do AVC atuem exclusivamente sobre a região de interesse.

Com a região de interesse devidamente isolada, o próximo passo consiste em extrair representações significativas que possam ser utilizadas nos processos de classificação e segmentação. Essa etapa busca transformar as informações espaciais e de intensidade da imagem em atributos que capturem padrões relevantes para a detecção e análise de lesões.

5.3 Extração de Atributos

A extração de atributos é uma etapa essencial tanto para a classificação quanto para a segmentação de imagens médicas, pois permite a representação mais abstrata e informativa dos dados de entrada. Ao empregar CNNs pré-treinadas, é possível aproveitar o conhecimento adquirido em grandes conjuntos de dados para extrair características relevantes, reduzindo a complexidade do treinamento e aumentando a eficácia das etapas posteriores. Nesta seção, são descritas as abordagens adotadas para a extração de atributos em duas tarefas distintas: classificação de exames em três categorias clínicas e geração de mapas de atributos para segmentação.

5.3.1 Extração de Atributos Para a Etapa de Classificação

Na tarefa de classificação, a extração eficaz de atributos é determinante para o desempenho do modelo. Para essa finalidade, foi utilizada a arquitetura ResNet18, reconhecida por seu desempenho em diversas tarefas de visão computacional. Seguindo abordagens similares às propostas em Yu, Wang e Zhang (2021) e Yu et al. (2021), a

rede foi adaptada para atender a necessidade específica deste trabalho, que envolve a categorização de exames de tomografia em três classes distintas.

As modificações realizadas na arquitetura incluem a remoção das duas últimas camadas da ResNet18, que são voltadas para a tarefa de classificação genérica, e a adição de uma camada totalmente conectada com 256 unidades, precedida por uma camada de dropout, com o objetivo de mitigar o overfitting. Na Figura 16 é apresentada a estrutura da rede após as modificações implementadas.

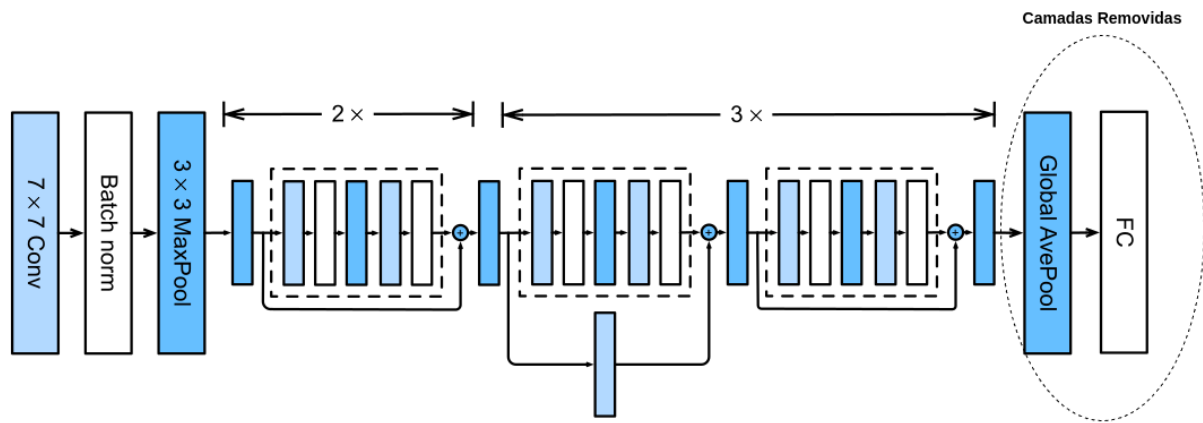


Figura 16 – Modificações na ResNet para extração de atributos. Fonte: Autor.

Inicialmente, a rede é treinada com os dados rotulados, de forma supervisionada, para ajustar os pesos de maneira a maximizar a capacidade discriminativa dos atributos extraídos. Em seguida, o modelo treinado é salvo e recarregado, com a exclusão das camadas finais, de forma que apenas as representações latentes das imagens sejam utilizadas nas etapas subsequentes. Essas representações, oriundas das últimas camadas convolucionais, capturam abstrações de alto nível e são utilizadas como entrada para o classificador final.

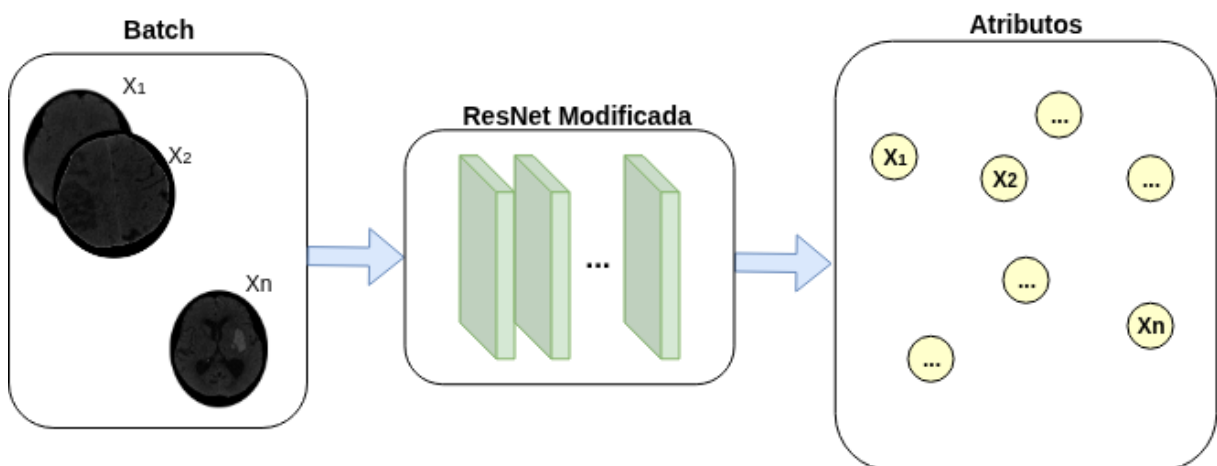


Figura 17 – Extração de vetores de atributos para cada imagem do batch utilizando a ResNet modificada. Fonte: Autor.

Na Figura 17 é exibido o fluxo da etapa de extração de atributos a partir de imagens de TC. À esquerda, observa-se um lote de imagens de entrada (*batch*), representado por $X_1, X_2, \dots, X_n$. O resultado desse processamento, apresentado à direita da figura, é um conjunto de features extraídos para cada imagem do lote. Esses vetores representam abstrações de alto nível da informação visual presente nas imagens originais e são posteriormente utilizados para a criação de grafos.

5.3.2 Extração de Atributos Para a Etapa de Segmentação

Na tarefa de segmentação, uma das principais dificuldades está na elevada resolução das imagens de entrada, que possuem dimensão de 512×512 pixels, totalizando 262.144 pixels por imagem. Caso cada pixel fosse modelado como um nó em um grafo, o número de conexões e o custo computacional associado seriam inviáveis. Dessa forma, torna-se necessário reduzir a dimensionalidade espacial da imagem sem perder informações essenciais para a segmentação das regiões de interesse.

As CNNs se destacam nesse contexto por sua capacidade de realizar *downsampling* por meio de camadas de convolução e *pooling*, permitindo a redução da resolução do mapa de características enquanto ampliam o campo receptivo e preservam informações semânticas. Para a extração do mapa de atributos, utilizou-se a ResNet34, cuja estrutura é análoga à ResNet18 utilizada na tarefa de classificação. Assim como ilustrado na Figura 16, foram removidas as duas últimas camadas da rede, a de *pooling* global e a camada totalmente conectada, de forma a manter apenas os mapas de características espaciais. Essa abordagem permite obter representações condensadas e informativas das imagens originais, compatíveis com a etapa posterior de modelagem com grafos.

A utilização desse *backbone* permite reduzir a imagem original para um mapa de características mais compacto e sem perda significativa de informação, facilitando a posterior construção dos grafos e tornando viável o processamento com a GCN voltada para segmentação.

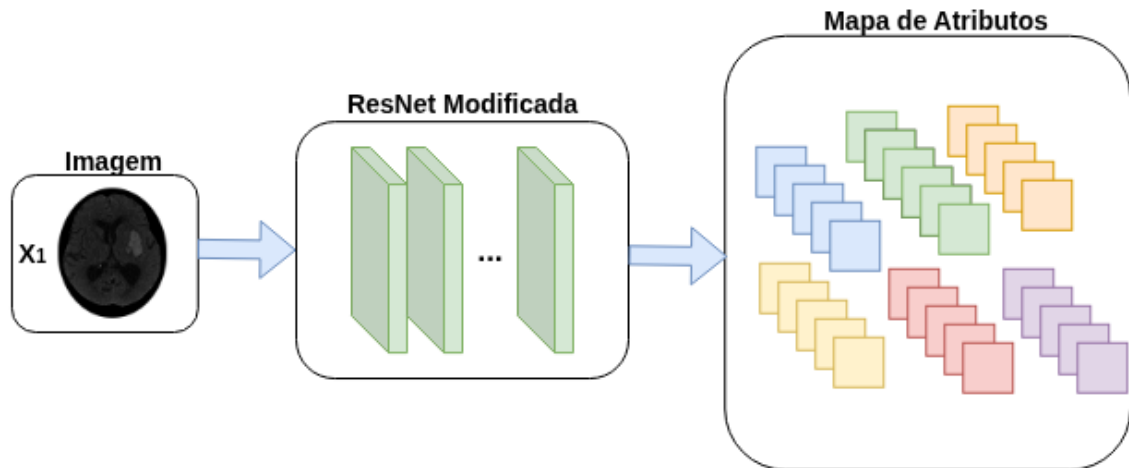


Figura 18 – Geração do mapa de atributos a partir da ResNet para segmentação. Fonte: Autor.

Na Figura 18 é representado o processo de extração de atributos realizado na tarefa de segmentação a partir de uma versão modificada da arquitetura ResNet. No lado esquerdo da figura, visualiza-se a imagem de entrada X_1 . Essas imagens são processadas por uma rede convolucional profunda, que teve suas camadas finais de *pooling* e classificação removidas, de modo semelhante à abordagem utilizada na extração de atributos para a classificação.

O resultado dessa operação está ilustrado à direita da figura, onde é apresentado o mapa de atributos extraído para a imagem. Esse mapa mantém a estrutura espacial da imagem original em menor resolução, contendo representações semânticas mais ricas e compactas. Tais representações são essenciais para a construção dos grafos e para o funcionamento da GCN, que opera sobre os nós e suas relações espaciais extraídas dessas estruturas intermediárias.

5.4 Criação do Grafo

A construção de grafos a partir das características extraídas constitui uma etapa fundamental nas tarefas de classificação e segmentação, pois permite modelar explicitamente as relações estruturais entre as amostras ou regiões da imagem. Através da representação gráfica, é possível capturar tanto a proximidade quanto a semelhança entre diferentes elementos, o que favorece a propagação de informações discriminativas ao longo da topologia do grafo. Nesta seção, são descritas as estratégias empregadas para a construção dos grafos em dois contextos distintos: a representação das imagens como nós para a tarefa de classificação e a estruturação dos mapas de características para a tarefa de segmentação.

5.4.1 Criação do Grafo Para a Etapa de Classificação

Cada conjunto de atributos extraídos de uma única imagem na etapa anterior é representado por um nó com diversas features. Essas características são usadas para construir os grafos, onde as arestas são definidas com base nos k vizinhos mais próximos das características extraídas utilizando o algoritmo *k-Nearest Neighbors* (KNN) para calcular a distância entre os nós. Cada imagem é representada como um nó no grafo. A figura 19 representa o processo de criação do grafo.

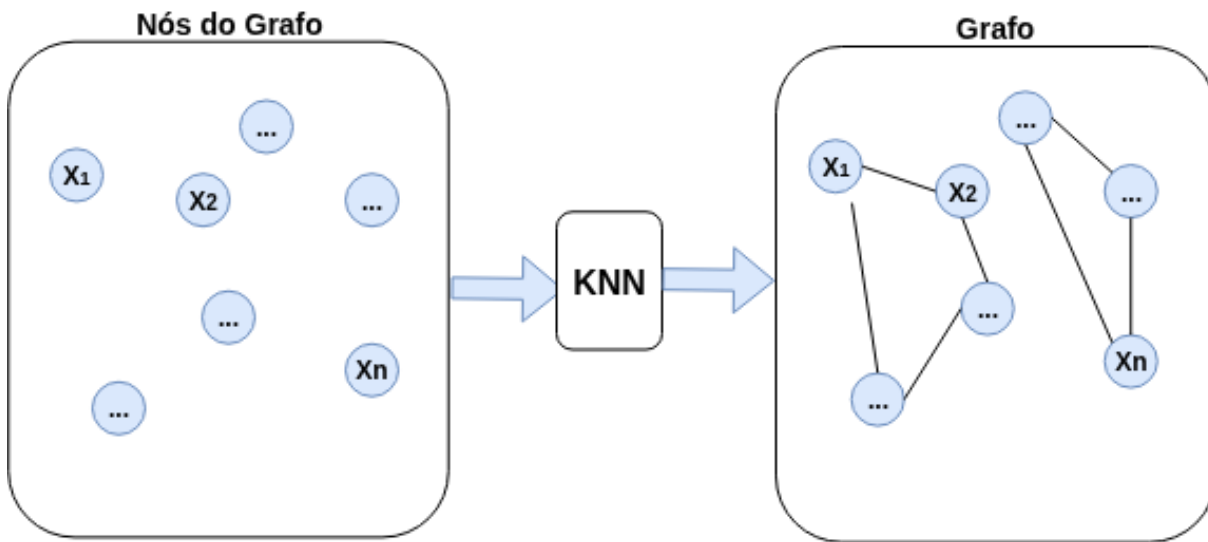


Figura 19 – Criação do grafo para classificação a partir das características extraídas das imagens. Fonte: Autor.

O grafo resultante é definido pelo par (V, E) , onde V representa os nós no grafo e E é o conjunto de arestas. Para a tarefa de classificação de nós, cada nó $v_i \in V$ pertence a R^N , onde N é o número de nós no grafo G . Portanto, $v_i, v_j \in E$ representa a aresta entre os nós v_i e v_j . A matriz de adjacência é usada para descrever as conexões entre os nós, sendo que o elemento na matriz é definido como 1 quando há uma aresta entre os nós, ou seja, quando os nós são identificados como vizinhos mais próximos.

Os grafos construídos são então armazenados para treinamento posterior. Este método de construção de grafos é simples e direto, predefinindo o número de nós conforme o tamanho do lote N e atualizando as arestas com base na matriz de adjacência.

5.4.2 Criação do Grafo Para a Etapa de Segmentação

Os nós do grafo são inicializados com o mapa de características obtido pela ResNet34, assim como mostra a Figura 20. Cada conjunto de atributos é considerado como um nó do grafo, de modo que cada nó do grafo corresponde a uma região da imagem original.

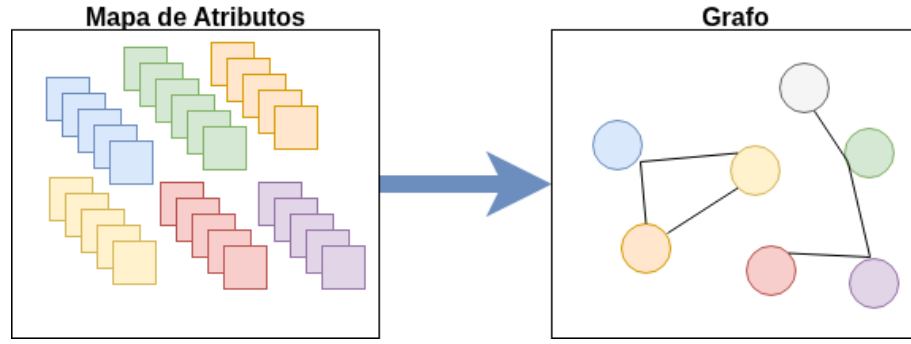


Figura 20 – Criação do grafo para segmentação com base nos mapas de características extraídos. Fonte: Autor.

Os componentes de codificação dessas redes são escolhidos como as redes *backbone* para inicializar o grafo, com o intuito de capturar o menor mapa de características. Com o mapa de atributos, os “super pixels” bidimensionais são convertidos em vetores unidimensionais. O mapa de atributos com dimensões $H \times W \times S$ é reformulado como $\mathbf{x}$ com dimensões $|N| \times S$, onde $|N| = H \times W$ e N é o conjunto de nós. Assim, o conjunto de nós pode ser representado como:

$$\mathbf{x} = [\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{|N|}]^T, \quad \mathbf{x} \in R^{|N| \times S}$$

onde $\mathbf{n}_i \in R^{S \times 1}$ é a anotação do nó, e S é o número de canais na camada de características. Em particular, na estrutura de grafo baseada em distância, há duas dimensões adicionais referentes às posições dos nós no mapa de características. O conjunto de arestas determina a estrutura do grafo, que é a principal diferença em relação ao ponto no espaço euclidiano.

As arestas conectam pares de nós com base nas relações entre eles, enriquecendo a representação estrutural do grafo ao incorporar interações locais. Quando os nós estão predominantemente conectados a outros da mesma classe, torna-se mais provável que recebam informações discriminativas de vizinhos semelhantes, favorecendo a correta propagação de rótulos. Esse comportamento permite que o modelo gráfico capture e aprenda a estrutura subjacente das imagens, uma capacidade que redes convolucionais tradicionais não possuem de forma explícita. Neste trabalho, as arestas são construídas utilizando o algoritmo KNN, que estabelece conexões entre cada nó e seus k vizinhos mais próximos com base em critérios de similaridade.

5.5 Classificação

O processo de classificação dos exames em normal, hemorrágico e isquêmico consiste nas seguintes etapas: divisão em dados de treino e teste, extração de atributos, criação do grafo e treinamento e avaliação da rede grafo convolucional.

A Figura 21 ilustra o fluxo de classificação baseado em GCN. Neste modelo, cada nó do grafo representa uma imagem de TC, enquanto as arestas codificam a similaridade entre essas imagens com base em seus atributos extraídos. O grafo resultante é processado por uma GCN, que realiza a propagação de informações ao longo da vizinhança dos nós, explorando a estrutura relacional dos dados. A saída da rede consiste na atribuição de um rótulo para cada imagem, classificando-a como normal, isquêmica ou hemorrágica.

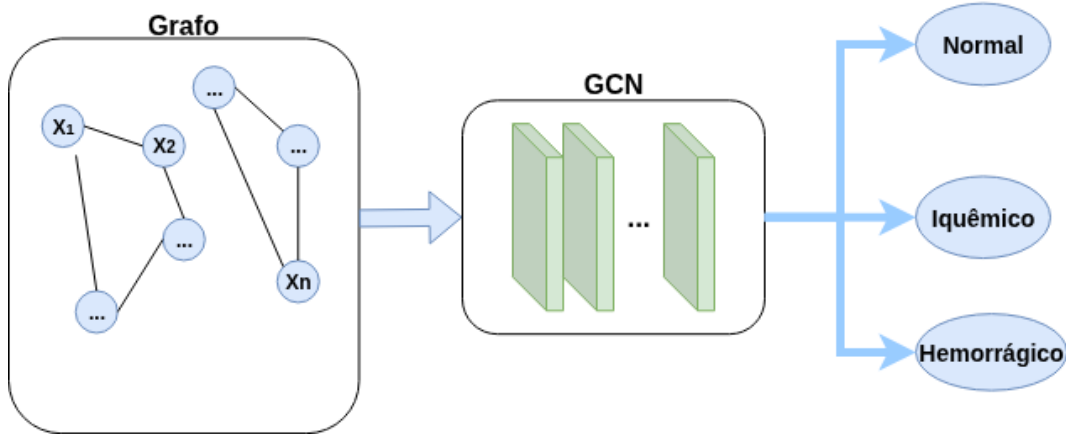


Figura 21 – Fluxo de Classificação com GCN a partir dos atributos extraídos pela ResNet.
Fonte: Autor.

A arquitetura implementada conta com uma camada grafo convolucional, seguida por uma camada de regularização com *dropout* e duas camadas totalmente conectadas, responsáveis por projetar os *embeddings* em um espaço de decisão com três classes de saída: normal, isquêmica e hemorrágica.

A camada convolucional utilizada baseia-se na formulação proposta por Kipf e Welling (2016), expressa por:

$$\mathbf{X}' = \hat{\mathbf{D}}^{-1/2} \hat{\mathbf{A}} \hat{\mathbf{D}}^{-1/2} \mathbf{X} \Theta,$$

em que $\hat{\mathbf{A}} = \mathbf{A} + \mathbf{I}$ é a matriz de adjacência com laços próprios, e $\hat{\mathbf{D}}$ representa a matriz diagonal dos graus correspondentes. A propagação de informações é ponderada por *edge_weight*, permitindo considerar pesos diferenciados nas conexões.

A versão por nó da operação é descrita por:

$$x'_i = \Theta^T \sum_{j \in \mathcal{N}(i) \cup \{i\}} \frac{e_{j,i}}{\sqrt{\hat{d}_i \hat{d}_j}} x_j,$$

em que $e_{j,i}$ é o peso da aresta entre os nós j e i , e $\hat{d}_i = 1 + \sum_{j \in \mathcal{N}(i)} e_{j,i}$ é o grau ajustado do nó i . Essa operação permite que cada nó agregue informações dos seus

vizinhos, respeitando a estrutura do grafo, favorecendo a aprendizagem de representações contextualizadas para cada exame.

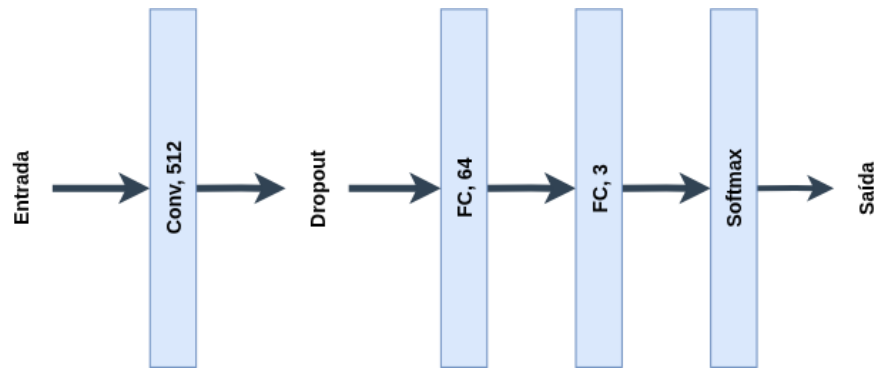


Figura 22 – Fluxo do processo de segmentação de AVCs hemorrágicos com uso de GCNs.
Fonte: Autor.

A Figura 22 apresenta a arquitetura do modelo GCN utilizado para a tarefa de classificação. A entrada do modelo passa inicialmente por uma camada convolucional, seguida de uma camada de dropout para regularização. Em seguida, os dados são processados por duas camadas totalmente conectadas. Por fim, a camada *softmax* gera a distribuição de probabilidades para as três classes de saída: normal, isquêmico e hemorrágico.

5.6 Segmentação

O processo de segmentação do AVCh consiste nas seguintes etapas: divisão em dados de treino e teste, extração do mapa de atributos, criação do grafo e treinamento e avaliação da rede grafo convolucional. Cada uma das etapas é explicada em detalhes nas subseções a seguir.

A Figura 23 apresenta o fluxo geral da segmentação baseado em GCN. A entrada do modelo é um grafo construído a partir dos mapas de características extraídos previamente. Cada nó do grafo representa uma região da imagem e está conectado a outros nós com base em critérios de similaridade. Esse grafo é processado por uma arquitetura GCN, que tem como objetivo propagar informações entre os nós e produzir uma saída segmentada, onde cada região da imagem é classificada quanto à presença ou ausência de lesão.

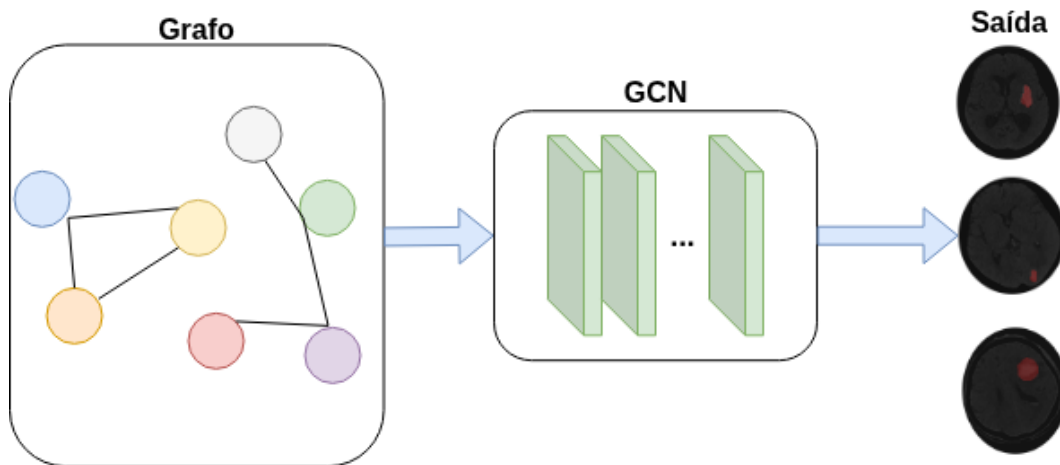


Figura 23 – Fluxo da segmentação utilizando GCN. Fonte: Autor.

A arquitetura interna do modelo é exibida na Figura 24. Inicialmente, são aplicadas duas camadas de convolução espectral do tipo ChebConv, baseadas no polinômio de Chebyshev, com 512 filtros e ordem $K = 2$, conforme proposto por Defferrard, Bresson e Vandergheynst (2016). Essas camadas permitem a propagação eficiente de informações no grafo, considerando a estrutura global das conexões.



Figura 24 – Arquitetura do modelo GCN para segmentação de lesões hemorrágicas. Fonte: Autor.

Após as convoluções, os dados passam por uma camada totalmente conectada, que reduz a dimensionalidade e projeta os embeddings aprendidos. Em seguida, três camadas de convolução transposta são aplicadas com o objetivo de reconstruir a estrutura espacial da imagem segmentada. Por fim, uma camada final de convolução gera a saída com a máscara binária da região lesionada. Essa arquitetura combina aprendizado estrutural com reconstrução espacial eficiente.

5.7 Avaliação

Para avaliar o desempenho dos métodos propostos, foram adotadas métricas derivadas da matriz de confusão (MC). Essa matriz é composta por quatro valores principais: verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN). Com base nesses elementos, é possível calcular diferentes métricas que

auxiliam na análise de modelos de classificação supervisionada, especialmente quando a variável de interesse é categórica.

As métricas selecionadas para validar os resultados obtidos neste trabalho foram: acurácia, precisão, sensibilidade, especificidade, F1-score, índice de Jaccard e coeficiente de Dice.

A acurácia é uma métrica geral que expressa a proporção de classificações corretas em relação ao total de elementos avaliados. No contexto da segmentação, representa a fração de pixels corretamente rotulados como pertencentes ou não à região de interesse:

$$\text{acurácia} = \frac{VP + VN}{VP + VN + FP + FN}. \quad (5.1)$$

A precisão quantifica a proporção de verdadeiros positivos em relação ao total de elementos classificados como positivos pelo modelo. Em segmentação, indica a confiança de que um pixel rotulado como positivo de fato pertence à região de interesse:

$$\text{precisão} = \frac{VP}{VP + FP}. \quad (5.2)$$

A sensibilidade indica a capacidade do modelo de identificar corretamente os elementos que, de fato, pertencem à classe de interesse. No contexto da segmentação de imagens médicas, essa métrica representa a proporção de pixels corretamente classificados como positivos em relação ao total de pixels realmente positivos. Um valor elevado de sensibilidade indica que poucos verdadeiros positivos foram omitidos pelo modelo. Sua formulação é apresentada por:

$$\text{sensibilidade} = \frac{VP}{VP + FN}. \quad (5.3)$$

A especificidade, por sua vez, reflete a habilidade do modelo em reconhecer corretamente os elementos que não pertencem à classe de interesse. Em tarefas de segmentação, refere-se à proporção de verdadeiros negativos corretamente identificados, ou seja, pixels que não pertencem à região de interesse e foram devidamente excluídos:

$$\text{especificidade} = \frac{VN}{VN + FP}. \quad (5.4)$$

O F1-score representa a média harmônica entre precisão e sensibilidade, sendo útil para balancear essas duas métricas, especialmente em cenários com classes desbalanceadas. Um valor elevado de F1-score indica bom desempenho simultâneo em identificar corretamente os positivos (sensibilidade) e evitar falsos positivos (precisão):

$$\text{F1-score} = 2 \cdot \frac{\text{precisão} \cdot \text{sensibilidade}}{\text{precisão} + \text{sensibilidade}}. \quad (5.5)$$

Além dessas métricas, que ajudam a entender a capacidade do modelo em distinguir entre regiões afetadas e não afetadas, também é importante avaliar a qualidade da segmentação propriamente dita. Para isso, é comum utilizar métricas que consideram a sobreposição entre as regiões segmentadas automaticamente e o *ground truth*, como o índice de Jaccard e o coeficiente de Dice.

O índice de Jaccard é uma medida estatística de similaridade entre conjuntos, amplamente utilizada na avaliação da qualidade da segmentação. Ele é definido como a razão entre a interseção e a união das regiões segmentadas pelo algoritmo e aquelas obtidas manualmente. A equação é dada por:

$$\text{jaccard} = \frac{VP}{VP + FP + FN}. \quad (5.6)$$

O coeficiente de Dice é outra métrica amplamente utilizada para mensurar a sobreposição entre regiões segmentadas automaticamente e manualmente. Diferentemente do índice de Jaccard, o coeficiente de Dice calcula uma média ponderada da interseção entre as duas regiões. Valores próximos de 1 indicam uma sobreposição quase perfeita; valores próximos de 0 indicam completa discrepância. Sua fórmula é:

$$\text{dice} = \frac{2 \cdot VP}{2 \cdot VP + FP + FN}. \quad (5.7)$$

Essas métricas complementares, Jaccard e Dice, oferecem uma avaliação mais precisa da qualidade da segmentação, sendo especialmente importantes em contextos médicos, nos quais a delimitação correta das regiões de interesse é fundamental para um diagnóstico confiável.

6 Resultados e Discussões

Esta seção apresenta os resultados referentes à metodologia empregada para classificação e segmentação de AVCh em exames de TC. Foram comparadas as metodologias de classificação utilizando as métricas acurácia, f1-score, sensibilidade, especificidade e precisão, e os modelos de segmentação utilizando as métricas acurácia, sensibilidade, especificidade, precisão, Dice e Jaccard.

6.1 Classificação de AVC's

A primeira etapa da metodologia proposta classificou os exames em hemorrágico, isquêmico e normal. Os resultados obtidos podem ser vistos na Tabela 4 e na Figura 25.

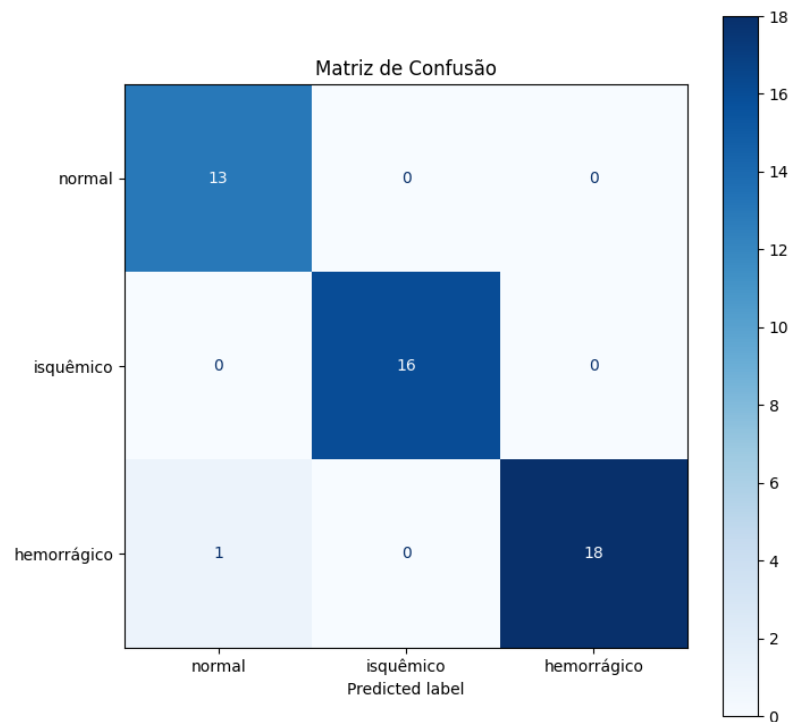


Figura 25 – Matriz de confusão para a tarefa de classificação em três categorias. Fonte: Autor.

A Figura 25 apresenta a matriz de confusão resultante do processo de classificação, evidenciando a capacidade do modelo em distinguir corretamente entre as três classes: normal, isquêmico e hemorrágico. Observa-se que apenas um exemplo da classe

hemorrágica foi incorretamente classificado como normal, enquanto todos os demais foram atribuídos corretamente às suas classes. Esse desempenho indica um bom nível de acurácia, sensibilidade e especificidade da metodologia proposta, aspectos essenciais em cenários clínicos onde decisões diagnósticas devem ser tomadas com elevado grau de segurança e confiabilidade.

A arquitetura proposta apresentou bons resultados em todas as métricas calculadas, obtendo resultados no mesmo nível dos outros trabalhos para todas as métricas. A seguir, detalhamos os resultados obtidos e discutimos sua comparação com os outros trabalhos.

Tabela 4 – Comparação dos Resultados de Classificação com Trabalhos Relacionados.

Trabalho	Acurácia	F1-Score	Sensibilidade	Especificidade	Precisão
Jr et al. (2019)	100% $\pm$ 0,00	100% $\pm$ 0,00	100% $\pm$ 0,00	100% $\pm$ 0,00	100% $\pm$ 0,00
Peixoto e Filho (2018)	95,83% $\pm$ 0,02	95,82% $\pm$ 0,02	94,83% $\pm$ 0,02	98,06% $\pm$ 0,01	96,06% $\pm$ 0,02
Filho et al. (2017)	99,30% $\pm$ 0,01	98,95% $\pm$ 0,01	98,96% $\pm$ 0,01	99,40% $\pm$ 0,01	98,99% $\pm$ 0,01
Souza et al. (2019)	99,40% $\pm$ 0,27	-	99,10% $\pm$ 0,95	98,60% $\pm$ 1,91	98,58% $\pm$ 1,95
Este Trabalho	99,50% $\pm$ 0,47%	99,47% $\pm$ 0,51%	99,34% $\pm$ 0,70%	99,75% $\pm$ 0,70%	99,51% $\pm$ 0,45%

A acurácia da metodologia proposta foi de 99,50%, o que indica que o modelo classificou corretamente a grande maioria dos exemplos. Em comparação com Jr et al. (2019), que alcançou uma acurácia de 100%, a diferença de 0,50% ainda representa alguns casos incorretos. Filho et al. (2017) obteve uma acurácia de 99,30%, o que, comparado ao trabalho atual, mostra uma ligeira melhoria de 0,20%, importante para a precisão diagnóstica em grande escala. Por outro lado, Peixoto e Filho (2018) teve uma acurácia de 95,83%, mostrando uma diferença de aproximadamente 3,67% em relação ao trabalho atual. Essa diferença representa uma redução significativa de erros, com a metodologia proposta apresentando uma melhoria notável na precisão diagnóstica.

O F1-Score de 99,47% indica um bom equilíbrio entre precisão e sensibilidade e que existem alguns falsos positivos ou negativos, embora o modelo seja eficaz, ficando abaixo somente de Jr et al. (2019). Peixoto e Filho (2018) apresentaram um F1-Score de 95,82%, resultando em uma diferença de 3,65% em relação ao trabalho atual, indicando um equilíbrio melhor entre precisão e recall, essencial para a classificação precisa de AVC. Comparado a Filho et al. (2017), que alcançaram um F1-Score de 98,95%, o trabalho atual mostra uma ligeira vantagem de 0,52%, sugerindo uma melhoria no balanceamento entre precisão e recall.

A sensibilidade obtida foi de 99,34%, garantindo que a maioria dos casos positivos de AVC foram detectados corretamente. No entanto, também mostra que houve alguns falsos negativos, mas ainda manteve uma alta eficácia na detecção de casos positivos, ficando abaixo de Jr et al. (2019). Peixoto e Filho (2018) tiveram uma sensibilidade

de 94,83%, mostrando uma diferença de 4,51% em relação ao trabalho atual, indicando uma melhoria significativa na detecção de casos positivos, crucial para evitar a falta de diagnóstico em AVC. Filho et al. (2017) apresentaram uma sensibilidade de 98,96%, com o trabalho atual mostrando uma ligeira vantagem de 0,38%, destacando uma melhoria na identificação correta de casos positivos.

A especificidade do trabalho atual foi de 99,75%, indicando uma alta precisão na identificação de casos negativos, mas sugerindo que ocorreram alguns falsos positivos, embora ainda mantenha uma alta precisão. Peixoto e Filho (2018) tiveram uma especificidade de 98,06%, resultando em uma diferença de 1,69% em relação ao trabalho atual. Isso indica que a metodologia proposta é melhor em identificar corretamente os casos negativos, evitando falsos positivos. Filho et al. (2017) apresentaram uma especificidade de 99,40%, com o trabalho atual mostrando uma vantagem de 0,35%, sugerindo uma melhoria na identificação correta de casos negativos.

A precisão do trabalho atual foi de 99,51%, indicando uma alta confiabilidade nas previsões positivas, embora tenha ocorrido alguns falsos positivos. Peixoto e Filho (2018) tiveram uma precisão de 96,06%, mostrando uma diferença de 3,45% em relação ao trabalho atual. Isso sugere que a metodologia proposta é mais confiável em suas previsões positivas, o que é crucial para garantir diagnósticos precisos. Filho et al. (2017) apresentaram uma precisão de 98,99%, com o trabalho atual mostrando uma ligeira vantagem de 0,52%, destacando uma melhoria na precisão das previsões positivas.

Esses resultados evidenciam que o modelo proposto é altamente eficaz para a tarefa de classificação de exames de TC, sendo competitivo com abordagens existentes na literatura e oferecendo potencial para aplicações clínicas assistidas por inteligência artificial.

6.2 Segmentação de AVC Hemorrágico

Nas Tabelas 5 e 6, são apresentados os resultados das principais métricas de desempenho obtidas a partir da variação dos parâmetros k (número de vizinhos) e do tipo de distância (cosseno ou euclidiana), conforme definidos na função de construção dos grafos. O parâmetro k determina quantos vizinhos cada nó no grafo irá considerar, influenciando diretamente a topologia da estrutura gerada, já que cada nó de destino recebe exatamente k conexões de origem. Por sua vez, o parâmetro de distância define o critério de proximidade entre os nós, podendo ser calculado com base na distância do cosseno ou na distância euclidiana.

Além das métricas de desempenho, a Tabela 6 também apresenta os valores médios e o desvio padrão do tempo total de processamento por execução para cada configuração testada. Essa informação é relevante para a avaliação do custo computacional envolvido,

sobretudo em grafos mais densos ou quando se utilizam métricas de distância mais complexas.

Tabela 5 – Comparação das métricas por número de vizinhos e tipo de distância.

Distância	Vizinhos	Sensibilidade	Especificidade	Acurácia	Precisão
euclidiana	5	95,59% $\pm$ 1,07%	95,59% $\pm$ 1,07%	99,64% $\pm$ 0,04%	95,63% $\pm$ 0,99%
	10	95,94% $\pm$ 1,19%	95,94% $\pm$ 1,19%	99,64% $\pm$ 0,07%	95,88% $\pm$ 0,74%
	20	96,14% $\pm$ 1,27%	96,14% $\pm$ 1,27%	99,65% $\pm$ 0,08%	95,62% $\pm$ 1,38%
	30	95,16% $\pm$ 1,31%	95,16% $\pm$ 1,31%	99,62% $\pm$ 0,07%	95,53% $\pm$ 0,99%
	40	95,30% $\pm$ 1,06%	95,30% $\pm$ 1,06%	99,64% $\pm$ 0,07%	95,60% $\pm$ 1,45%
	50	94,91% $\pm$ 0,70%	94,91% $\pm$ 0,70%	99,66% $\pm$ 0,06%	96,03% $\pm$ 1,40%
cosseno	5	95,07% $\pm$ 0,88%	95,07% $\pm$ 0,88%	99,59% $\pm$ 0,06%	94,65% $\pm$ 1,16%
	10	95,82% $\pm$ 1,02%	95,82% $\pm$ 1,02%	99,66% $\pm$ 0,10%	95,47% $\pm$ 0,76%
	20	94,95% $\pm$ 1,49%	94,95% $\pm$ 1,49%	99,63% $\pm$ 0,07%	95,62% $\pm$ 0,71%
	30	95,18% $\pm$ 1,26%	95,18% $\pm$ 1,26%	99,62% $\pm$ 0,08%	95,28% $\pm$ 1,32%
	40	95,31% $\pm$ 1,61%	95,31% $\pm$ 1,61%	99,62% $\pm$ 0,08%	95,83% $\pm$ 0,86%
	50	95,49% $\pm$ 0,83%	95,49% $\pm$ 0,83%	99,60% $\pm$ 0,07%	94,33% $\pm$ 0,93%

Observa-se nas Tabelas 5 e 6 que o aumento do valor de k tende a melhorar o desempenho do modelo até certo ponto, seguido de uma leve oscilação nas métricas e aumento no tempo de processamento. Entre as configurações avaliadas, destaca-se $k = 10$ com distância euclidiana, que apresentou um desempenho equilibrado e elevado em todas as métricas, com acurácia de 99,64%, Dice de 95,89%, Jaccard de 92,40%, sensibilidade de 95,94%, especificidade de 95,94% e precisão de 95,88%, além de um tempo computacional inferior ao de configurações com maior número de vizinhos. Embora valores mais altos de k também tenham apresentado métricas competitivas, o desempenho de $k = 10$ aliado ao menor custo computacional reforça sua escolha como uma configuração ideal. Além disso, a distância euclidiana mostrou-se consistentemente mais eficaz do que a distância do cosseno, possivelmente por preservar melhor a localidade espacial nas imagens de TC, o que é crucial para capturar estruturas anatômicas contínuas e homogêneas.

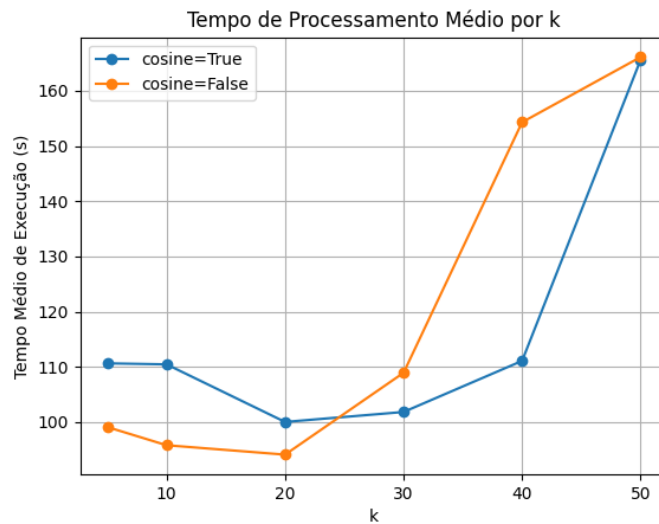
Embora valores mais altos de k gerem grafos mais densamente conectados e, teoricamente, forneçam maior contexto espacial ao modelo, os resultados indicam que a configuração com $k = 10$ e distância euclidiana foi a mais eficiente. Valores superiores de k não proporcionaram melhorias significativas nas métricas de segmentação e ainda elevaram consideravelmente o tempo de processamento, como pode ser observado na Figura 26. Isso sugere que, neste cenário, uma conectividade moderada é suficiente para capturar as relações espaciais relevantes, sem introduzir redundância ou comprometer a eficiência computacional.

Outro aspecto relevante diz respeito à estabilidade dos resultados, evidenciada pelos baixos desvios padrão em praticamente todas as métricas, especialmente nas configurações

Tabela 6 – Comparação entre Dice, Jaccard e Tempo por número de vizinhos e tipo de distância.

Distância	Vizinhos	Dice	Jaccard	Tempo
euclidiana	5	95,59% $\pm$ 0,64%	91,88% $\pm$ 1,08%	99.12s $\pm$ 0.73s
	10	95,89% $\pm$ 0,67%	92,40% $\pm$ 1,14%	95.83s $\pm$ 1.70s
	20	95,87% $\pm$ 1,13%	92,38% $\pm$ 1,90%	94.13s $\pm$ 0.57s
	30	95,32% $\pm$ 0,65%	91,42% $\pm$ 1,08%	108.95s $\pm$ 17.86s
	40	95,44% $\pm$ 0,97%	91,64% $\pm$ 1,63%	154.31s $\pm$ 4.69s
	50	95,45% $\pm$ 0,78%	91,65% $\pm$ 1,31%	166.11s $\pm$ 17.58s
cosseno	5	94,84% $\pm$ 0,52%	90,62% $\pm$ 0,86%	110.66s $\pm$ 12.53s
	10	95,64% $\pm$ 0,74%	91,97% $\pm$ 1,24%	110.47s $\pm$ 12.80s
	20	95,27% $\pm$ 0,92%	91,35% $\pm$ 1,54%	100.04s $\pm$ 1.21s
	30	95,21% $\pm$ 0,93%	91,25% $\pm$ 1,57%	101.85s $\pm$ 0.72s
	40	95,54% $\pm$ 0,80%	91,80% $\pm$ 1,35%	111.10s $\pm$ 10.68s
	50	94,90% $\pm$ 0,75%	90,72% $\pm$ 1,25%	165.61s $\pm$ 5.63s

com $k = 10$. Esse comportamento reforça a consistência do modelo treinado sob essas condições, o que sugere seu potencial de generalização em diferentes subconjuntos de dados.

Figura 26 – Evolução da função de perda e do coeficiente de Dice ao longo das épocas para $k = 10$, distância euclidiana. Fonte: Autor.

A Figura 26 mostra que, para valores baixos de k , o tempo médio de execução se mantém relativamente estável e semelhante entre os dois tipos de distância. No entanto, à medida que k aumenta, o custo computacional cresce de forma mais acentuada, especialmente quando se utiliza a distância do cosseno. Isso pode estar associado ao fato de que essa métrica requer a normalização dos vetores e o cálculo de similaridade angular,

operações computacionalmente mais custosas do que a simples subtração vetorial utilizada na distância euclidiana.

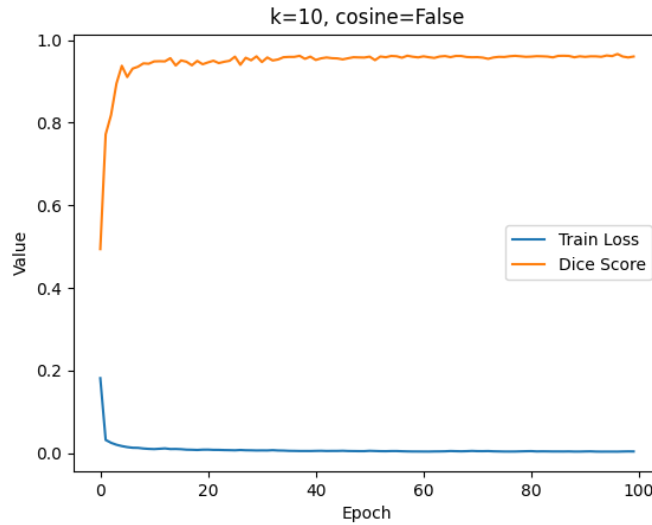


Figura 27 – Curvas de evolução da função de perda e do Dice score durante o treinamento.
Fonte: Autor.

A Figura 27 apresenta a evolução da função de perda e do coeficiente de Dice ao longo das épocas de treinamento para a configuração $k = 10$ com distância euclidiana. Observa-se uma rápida convergência da função de perda nas primeiras épocas, estabilizando-se em valores baixos, o que indica uma minimização eficaz do erro pelo modelo. Simultaneamente, o Dice Score apresenta crescimento acentuado nas iterações iniciais e mantém-se elevado até o final do treinamento, com baixa variabilidade. Esse comportamento evidencia que o modelo foi capaz de realizar uma segmentação precisa e consistente, mantendo um nível satisfatório de sobreposição entre as predições e o *ground truth*. A estabilidade da curva de Dice também sugere que não houve sinais significativos de *overfitting*.

De forma geral, os resultados obtidos indicam que, em tarefas de segmentação de imagens médicas com GCNs, a definição adequada da vizinhança no grafo é um fator crítico para o desempenho. A combinação entre a distância euclidiana e um número intermediário de vizinhos se mostrou a mais eficiente, conciliando desempenho superior e custo computacional viável. Esses achados estão de acordo com a literatura atual, que destaca a importância da construção da estrutura do grafo como uma etapa determinante no sucesso de arquiteturas baseadas em aprendizado profundo sobre grafos.

6.3 Comparação dos Resultados com Trabalhos Relacionados

Com o objetivo de avaliar a efetividade da abordagem proposta, os resultados obtidos foram comparados com aqueles reportados em estudos prévios que também abordam a segmentação de regiões afetadas por AVCh em exames de TC. As Tabelas 7

e 8 apresentam essa comparação em termos de métricas quantitativas, incluindo acurácia, sensibilidade, especificidade, precisão, Dice e Jaccard.

Em relação à acurácia, o modelo proposto alcançou 99,64% , um valor ligeiramente inferior aos obtidos por Rebouças et al. (2019) e Souza et al. (2023). No entanto, é importante ressaltar que a acurácia pode ser influenciada pela natureza altamente desbalanceada das imagens, nas quais a maior parte do *ground truth* corresponde a pixels de fundo. Isso pode inflar artificialmente os valores de acurácia, tornando essencial uma análise complementar com outras métricas mais sensíveis à qualidade da segmentação, como o Dice e o Jaccard.

Tabela 7 – Comparação de Métricas de Segmentação: Sensibilidade, Especificidade e Acurácia.

Trabalho	Sensibilidade	Especificidade	Acurácia
Souza et al. (2023)	99,82% $\pm$ 1,06%	99,81% $\pm$ 0,09%	99,8% $\pm$ 0,09%
Marques et al. (2022)	99,97% $\pm$ 0,06%	-	99,72% $\pm$ 0,24%
Santos, Medeiros e Filho (2022)	93,35% $\pm$ 4,23%	99,94% $\pm$ 0,05%	99,85% $\pm$ 0,07%
Barroso et al. (2022)	94,07% $\pm$ 3,60%	99,92% $\pm$ 0,70%	99,85% $\pm$ 0,05%
Rebouças et al. (2019)	99,86% $\pm$ 0,09%	97,82% $\pm$ 2,33%	99,84% $\pm$ 0,08%
Este Trabalho	95,94% $\pm$ 1,19%	95,94% $\pm$ 1,19%	99,64% $\pm$ 0,07%

As métricas de sensibilidade e especificidade obtidas foram de 95,94% , com valores iguais para ambas. Esses resultados posicionam o modelo em um nível intermediário em relação à literatura. Trabalhos como o de Souza et al. (2023) e Marques et al. (2022) apresentaram sensibilidades e especificidades superiores a 99%, enquanto Santos, Medeiros e Filho (2022) e Barroso et al. (2022) obtiveram sensibilidades de 93,35% e 94,07%, respectivamente. Tais valores indicam que o modelo proposto apresenta um bom equilíbrio entre a identificação de regiões afetadas e a preservação das regiões normais, sendo competitivo com abordagens mais consolidadas.

Tabela 8 – Comparação de Métricas de Sobreposição e Precisão: Dice, Jaccard e Precisão.

Trabalho	Dice	Jaccard	Precisão
Souza et al. (2023)	93,25% $\pm$ 3,89%	-	98,55% $\pm$ 6,64%
Marques et al. (2022)	-	-	94,5% $\pm$ 10,27%
Santos, Medeiros e Filho (2022)	93,35% $\pm$ 2,80%	87,66% $\pm$ 4,77%	93,78% $\pm$ 5,75%
Barroso et al. (2022)	94,28% $\pm$ 1,85%	-	-
Rebouças et al. (2019)	92,41% $\pm$ 2,80%	92,41% $\pm$ 2,84%	-
Este Trabalho	95,89% $\pm$ 0,67%	92,40% $\pm$ 1,14%	95,88% $\pm$ 0,74%

No que diz respeito às métricas de sobreposição, o coeficiente de Dice e o índice de Jaccard alcançados pelo modelo atual superam os valores obtidos por trabalhos como

Souza et al. (2023) e Santos, Medeiros e Filho (2022). Esses resultados indicam uma excelente capacidade de segmentação das regiões de interesse, com elevada sobreposição entre as máscaras preditas e os rótulos de referência.

Em relação à precisão, o modelo apresentou 95,88%, o que indica que a maioria das regiões segmentadas como positivas realmente corresponde a regiões afetadas. Embora alguns trabalhos, como Souza et al. (2023), tenham obtido valores ligeiramente superiores, a abordagem atual ainda se destaca por apresentar menor variabilidade, o que reforça sua robustez.

De forma geral, os resultados demonstram que o modelo proposto oferece desempenho competitivo em relação aos trabalhos relacionados, com destaque para as métricas de sobreposição (Dice e Jaccard), que são mais representativas da qualidade da segmentação em contextos médicos. Além disso, a estabilidade dos resultados, evidenciada pelos baixos desvios padrão, reforça o potencial da abordagem para aplicações práticas em ambientes clínicos.

7 Conclusões e Trabalhos Futuros

Os resultados obtidos ao longo deste trabalho demonstram a eficácia da metodologia proposta tanto na classificação quanto na segmentação de exames de TC para a identificação de AVCs hemorrágicos. A arquitetura desenvolvida apresentou bom desempenho nas principais métricas de avaliação.

No que se refere à tarefa de classificação, o modelo foi capaz de distinguir com precisão entre três padrões cerebrais distintos: normal, AVCh e AVCi. A abordagem alcançou métricas elevadas, com destaque para a acurácia de 99,50% e F1-score de 99,47%, além de valores igualmente altos de sensibilidade, especificidade e precisão. Os resultados indicam que a metodologia proposta alcança níveis de desempenho comparáveis ou superiores aos das abordagens consolidadas, reforçando sua viabilidade em aplicações clínicas com exigência de precisão diagnóstica.

Com relação à segmentação, foi realizada uma análise do impacto dos métodos de criação de arestas baseados em distância euclidiana e similaridade do cosseno. Os experimentos demonstraram que o uso da distância euclidiana proporcionou melhor desempenho nas métricas Dice, Jaccard e sensibilidade, indicando maior eficácia para representar conexões espaciais relevantes no grafo. A similaridade do cosseno, embora conceitualmente apropriada, não apresentou resultados superiores no cenário avaliado.

Outro fator investigado foi o número de vizinhos conectados em cada grafo. Os testes com diferentes valores de k mostraram que a configuração com $k = 10$ oferece um bom equilíbrio entre conectividade e desempenho computacional. Essa topologia se mostrou eficiente para capturar o contexto espacial relevante, resultando em métricas elevadas de segmentação, como Dice de 95,89% e Jaccard de 92,40%, sem comprometer a eficiência do modelo.

Além disso, foi realizado um estudo comparativo entre diferentes configurações de grafos visando avaliar seus impactos no tempo de processamento. Observou-se que o tempo de execução por configuração se manteve viável para aplicações clínicas, mesmo em cenários com maior conectividade entre os nós. Essa análise reforça a importância do equilíbrio entre complexidade estrutural e desempenho computacional, principalmente em ambientes de uso real.

Outro aspecto relevante está na estabilidade dos resultados. Os baixos desvios padrão registrados ao longo dos experimentos indicam boa capacidade de generalização do modelo, mesmo diante de variações no conjunto de dados. Isso confere robustez à abordagem proposta e sugere aplicabilidade em diferentes contextos clínicos.

Dessa forma, os resultados obtidos confirmam que a metodologia integrada proposta, baseada na combinação de CNNs para extração de atributos e GCNs para modelagem estrutural, é eficaz na segmentação de regiões afetadas por AVCh em imagens de TC. Além de alcançar métricas competitivas nas etapas de classificação e segmentação, a abordagem representa uma contribuição relevante para o desenvolvimento de ferramentas de suporte à decisão clínica, sendo uma alternativa promissora para sistemas inteligentes voltados ao diagnóstico automatizado de AVC.

7.1 Trabalhos Futuros

Com base nos resultados obtidos, diversas possibilidades de aprimoramento e expansão da metodologia proposta podem ser exploradas. Assim, a perspectiva nos trabalhos futuros está voltada para:

- a validação da abordagem em diferentes bases de dados, com maior diversidade de casos;
- a investigação de novas estratégias para construção das arestas do grafo, incluindo técnicas de ponderação baseadas em similaridade contextual ou espacial;
- a utilização de grafos dinâmicos, permitindo a atualização da topologia ao longo do treinamento, visando aprimorar a capacidade adaptativa do modelo;
- a aplicação da metodologia baseada em grafos para segmentação tridimensional, explorando a conectividade entre pixels de diferentes fatias dos exames de TC.

Espera-se que, com essas melhorias, os resultados futuros apresentem maior precisão e menor tempo de processamento em comparação aos obtidos neste trabalho.

Referências

- ABRAMOVA, V. et al. Hemorrhagic stroke lesion segmentation using a 3d u-net with squeeze-and-excitation blocks. *Computerized Medical Imaging and Graphics*, Elsevier, v. 90, p. 101908, 2021. Citado na página 28.
- ADEOYE, O. et al. Recommendations for the establishment of stroke systems of care: a 2019 update: a policy statement from the american stroke association. *Stroke*, Lippincott Williams & Wilkins Hagerstown, MD, v. 50, n. 7, p. e187–e210, 2019. Citado na página 34.
- ALEXANDER, R. G. et al. Visual illusions in radiology: untrue perceptions in medical images and their implications for diagnostic accuracy. *Frontiers in Neuroscience*, Frontiers Media SA, v. 15, p. 629469, 2021. Citado na página 27.
- ANANTH, C. V. et al. Epidemiology and trends in stroke mortality in the usa, 1975–2019. *International journal of epidemiology*, Oxford University Press, v. 52, n. 3, p. 858–866, 2023. Citado na página 28.
- BARROSO, C. B. et al. Stroke detection and segmentation in ct images using convolutional neural networks and active contour geodesic method. In: *Congresso Brasileiro de Automática-CBA*. [S.l.: s.n.], 2022. v. 3, n. 1. Citado 2 vezes nas páginas 41 e 93.
- BARUA, A.; AHMED, M. U.; BEGUM, S. A systematic literature review on multimodal machine learning: Applications, challenges, gaps and future directions. *Ieee access*, IEEE, v. 11, p. 14804–14831, 2023. Citado na página 25.
- BAUMANN, M. et al. Associations between quality of life and socioeconomic factors, functional impairments and dissatisfaction with received information and home-care services among survivors living at home two years after stroke onset. *BMC neurology*, Springer, v. 14, p. 1–12, 2014. Citado 2 vezes nas páginas 32 e 35.
- BUZUG, T. M. Computed tomography: from photon statistics to modern cone-beam ct. *Soc Nuclear Med*, v. 50, n. 7, 2009. Citado 2 vezes nas páginas 17 e 35.
- BUZUG, T. M. Computed tomography. In: *Springer handbook of medical technology*. [S.l.]: Springer, 2011. p. 311–342. Citado na página 36.
- CAMPBELL, B. C. V.; KHATRI, P. Stroke. *The Lancet*, Elsevier, v. 396, n. 10244, p. 129–142, 2020. Disponível em: <[https://doi.org/10.1016/S0140-6736\(20\)31179-X](https://doi.org/10.1016/S0140-6736(20)31179-X)>. Citado na página 31.
- CAPLAN, L. R. *Caplan's Stroke: A Clinical Approach*. 4. ed. [S.l.]: Elsevier Health Sciences, 2016. ISBN 9780323295444. Citado 5 vezes nas páginas 17, 31, 32, 33 e 34.
- CARVALHO, J. J. F. D. et al. Stroke epidemiology, patterns of management, and outcomes in fortaleza, brazil: a hospital-based multicenter prospective study. *Stroke*, Lippincott Williams & Wilkins Hagerstown, MD, v. 42, n. 12, p. 3341–3346, 2011. Citado 3 vezes nas páginas 32, 33 e 34.

CHEN, Y.-T. et al. Deep learning–based brain computed tomography image classification with hyperparameter optimization through transfer learning for stroke. *Diagnostics*, MDPI, v. 12, n. 4, p. 807, 2022. Citado na página 37.

CHOUDHRY, I. A. et al. A novel interpretable graph convolutional neural network for multimodal brain tumor segmentation. *Cognitive Computation*, Springer, v. 17, n. 1, p. 1–25, 2025. Citado na página 46.

CUI, H. et al. Evolving graph convolutional network with transformer for ct segmentation. *Applied Soft Computing*, Elsevier, v. 165, p. 112069, 2024. Citado na página 46.

DEFFERRARD, M.; BRESSON, X.; VANDERGHEYNST, P. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, v. 29, 2016. Citado 2 vezes nas páginas 71 e 84.

DUMOULIN, V.; VISIN, F. A guide to convolution arithmetic for deep learning. *arXiv preprint arXiv:1603.07285*, 2018. Citado 3 vezes nas páginas 54, 55 e 56.

FAN, S.; YANG, L.; JI, X. Reperfusion therapy for acute ischemic stroke: Where we are and where to go. *Journal of Translational Internal Medicine*, De Gruyter, v. 13, n. 1, p. 1–3, 2025. Citado na página 34.

FEIGIN, V. L. et al. World stroke organization (wso): global stroke fact sheet 2022. *International Journal of Stroke*, SAGE Publications Sage UK: London, England, v. 17, n. 1, p. 18–29, 2022. Citado na página 32.

FILHO, P. P. R. et al. New approach to detect and classify stroke in skull ct images via analysis of brain tissue densities. *Computer methods and programs in biomedicine*, Elsevier, v. 148, p. 27–43, 2017. Citado 2 vezes nas páginas 88 e 89.

FLEISCHMANN, D.; BOAS, F. E. Computed tomography—old ideas and new technology. *European radiology*, Springer, v. 21, p. 510–517, 2011. Citado na página 36.

FLORKOW, M. C. et al. Magnetic resonance imaging versus computed tomography for three-dimensional bone imaging of musculoskeletal pathologies: a review. *Journal of Magnetic Resonance Imaging*, Wiley Online Library, v. 56, n. 1, p. 11–34, 2022. Citado 2 vezes nas páginas 25 e 26.

FREITAS, C. F. de. *Imaginologia: série abeno: Odontologia essencial-Parte clínica*. [S.l.]: Artes Medicas Editora, 2013. Citado 2 vezes nas páginas 17 e 36.

FU, X. et al. High-risk factor prediction in lung cancer using thin ct scans: An attention-enhanced graph convolutional network approach. In: IEEE. *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. [S.l.], 2023. p. 1905–1910. Citado na página 45.

GAUTAM, A.; RAMAN, B. Towards effective classification of brain hemorrhagic and ischemic stroke using cnn. *Biomedical Signal Processing and Control*, Elsevier, v. 63, p. 102178, 2021. Citado na página 37.

GONZÁLEZ-PÉREZ, A. et al. Mortality after hemorrhagic stroke: data from general practice (the health improvement network). *Neurology*, Lippincott Williams & Wilkins Hagerstown, MD, v. 81, n. 6, p. 559–565, 2013. Citado 2 vezes nas páginas 32 e 33.

- HATHCOCK, J. T.; STICKLE, R. L. Principles and concepts of computed tomography. *Veterinary Clinics of North America: Small Animal Practice*, Elsevier, v. 23, n. 2, p. 399–415, 1993. Citado 2 vezes nas páginas 36 e 37.
- HAYKIN, S. *Neural networks and learning machines, 3/E*. [S.l.]: Pearson Education India, 2009. Citado 3 vezes nas páginas 17, 49 e 51.
- HE, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 770–778. Citado 2 vezes nas páginas 56 e 58.
- HSIEH, J.; FLOHR, T. Computed tomography recent history and future perspectives. *Journal of Medical Imaging*, Society of Photo-Optical Instrumentation Engineers, v. 8, n. 5, p. 052109–052109, 2021. Citado 2 vezes nas páginas 35 e 36.
- HU, J. et al. Incorporating adaptive sparse graph convolutional neural networks for segmentation of organs at risk in radiotherapy. *International Journal of Intelligent Systems*, Wiley Online Library, v. 2024, n. 1, p. 1728801, 2024. Citado na página 43.
- HUANG, R. et al. Stroke mortality prediction based on ensemble learning and the combination of structured and textual data. *Computers in Biology and Medicine*, Elsevier, v. 155, p. 106176, 2023. Citado na página 28.
- JR, C. M. D. et al. Deep learning iot system for online stroke detection in skull computed tomography images. *Computer Networks*, Elsevier, v. 152, p. 25–39, 2019. Citado na página 88.
- KAK, A. C.; SLANEY, M. *Principles of computerized tomographic imaging*. [S.l.]: SIAM, 2001. Citado na página 36.
- KASALAK, Ö. et al. Work overload and diagnostic errors in radiology. *European Journal of Radiology*, Elsevier, v. 167, p. 111032, 2023. Citado na página 27.
- KHAN, N. A. et al. Risk factors, quality of care and prognosis in south asian, east asian and white patients with stroke. *BMC neurology*, Springer, v. 13, p. 1–8, 2013. Citado na página 32.
- KHOSHKHABAR, M. et al. Automatic liver tumor segmentation from ct images using graph convolutional network. *Sensors*, MDPI, v. 23, n. 17, p. 7561, 2023. Citado na página 43.
- KIPF, T. N.; WELLING, M. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016. Citado na página 82.
- KIPF, T. N.; WELLING, M. Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations (ICLR)*, 2017. Citado na página 72.
- KUMAR, Y. et al. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *Journal of ambient intelligence and humanized computing*, Springer, v. 14, n. 7, p. 8459–8486, 2023. Citado na página 25.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, IEEE, v. 86, n. 11, p. 2278–2324, 1998. Citado na página 54.

LEE, E.-J. et al. Impact of onset-to-door time on outcomes and factors associated with late hospital arrival in patients with acute ischemic stroke. *PLoS One*, Public Library of Science San Francisco, CA USA, v. 16, n. 3, p. e0247829, 2021. Citado na página 28.

LIU, L. et al. Sasg-gcn: self-attention similarity guided graph convolutional network for multi-type lower-grade glioma classification. *IEEE Journal of Biomedical and Health Informatics*, IEEE, v. 27, n. 7, p. 3384–3395, 2023. Citado na página 44.

LIU, Y.; ZHANG, Y. Resnet-50 architecture and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020. Citado na página 56.

MA, Y.; TANG, J. *Deep Learning on Graphs*. [S.l.]: Cambridge University Press, 2021. Citado 7 vezes nas páginas 17, 54, 55, 59, 67, 68 e 70.

MANIVA, S. J. C. d. F.; FREITAS, C. H. A. d. Uso de alteplase no tratamento do acidente vascular encefálico isquêmico agudo: o que sabem os enfermeiros? *Revista Brasileira de Enfermagem*, SciELO Brasil, v. 65, p. 474–481, 2012. Citado 2 vezes nas páginas 31 e 32.

MARQUES, A. G. et al. Automatic segmentation of hemorrhagic stroke on brain ct images using convolutional neural networks through fine-tuning. In: IEEE. *2022 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2022. p. 01–08. Citado na página 93.

MAZONAKIS, M.; DAMILAKIS, J. Computed tomography: What and how does it measure? *European journal of radiology*, Elsevier, v. 85, n. 8, p. 1499–1504, 2016. Citado 2 vezes nas páginas 35 e 36.

MEAD, G. E. et al. A systematic review and synthesis of global stroke guidelines on behalf of the world stroke organization. *International Journal of Stroke*, SAGE Publications Sage UK: London, England, v. 18, n. 5, p. 499–531, 2023. Citado na página 34.

MOREIRA, P. V. L. et al. Coronary heart disease and stroke mortality trends in brazil 2000–2018. *PLoS One*, Public Library of Science San Francisco, CA USA, v. 16, n. 9, p. e0253639, 2021. Citado na página 28.

NEETHI, A. et al. Stroke classification from computed tomography scans using 3d convolutional neural network. *Biomedical Signal Processing and Control*, Elsevier, v. 76, p. 103720, 2022. Citado na página 38.

NEWMAN-TOKER, D. E. et al. Burden of serious harms from diagnostic error in the usa. *BMJ Quality & Safety*, BMJ Publishing Group Ltd, v. 33, n. 2, p. 109–120, 2024. Citado na página 27.

OVBIAGELE, B.; NGUYEN-HUYNH, M. N. Stroke epidemiology: advancing our understanding of disease mechanism and therapy. *Neurotherapeutics*, Elsevier, v. 8, n. 3, p. 319–329, 2011. Citado 4 vezes nas páginas 31, 33, 34 e 35.

PARVAIZ, A. et al. Vision transformers in medical computer vision—a contemplative retrospection. *Engineering Applications of Artificial Intelligence*, Elsevier, v. 122, p. 106126, 2023. Citado 2 vezes nas páginas 25 e 26.

- PATIL, S. et al. Detection, diagnosis and treatment of acute ischemic stroke: current and future perspectives. *Frontiers in medical technology*, Frontiers Media SA, v. 4, p. 748949, 2022. Citado na página 28.
- PEIXOTO, S. A.; FILHO, P. P. R. Neurologist-level classification of stroke using a structural co-occurrence matrix based on the frequency domain. *Computers & Electrical Engineering*, Elsevier, v. 71, p. 398–407, 2018. Citado 3 vezes nas páginas 38, 88 e 89.
- PENG, Q. et al. Deep learning-based computed tomography image segmentation and volume measurement of intracerebral hemorrhage. *Frontiers in Neuroscience*, Frontiers Media SA, v. 16, p. 965680, 2022. Citado na página 40.
- RAHMAN, M. Magnetic resonance imaging and iron-oxide nanoparticles in the era of personalized medicine. *Nanotheranostics*, v. 7, n. 4, p. 424, 2023. Citado na página 25.
- RAMOS-CABRER, P. et al. Targeting the ischemic penumbra. *Stroke*, Lippincott Williams & Wilkins Hagerstown, MD, v. 42, n. 1_suppl_1, p. S7–S11, 2011. Citado na página 32.
- REBOUÇAS, E. d. S. et al. New level set approach based on parzen estimation for stroke segmentation in skull ct images. *Soft Computing*, Springer, v. 23, p. 9265–9286, 2019. Citado na página 93.
- RINGELSTEIN, E. B. et al. European stroke organisation recommendations to establish a stroke unit and stroke center. *Stroke*, Lippincott Williams & Wilkins Hagerstown, MD, v. 44, n. 3, p. 828–840, 2013. Citado na página 34.
- RUBIN, G. D. Computed tomography: revolutionizing the practice of medicine for 40 years. *Radiology*, Radiological Society of North America, v. 273, n. 2S, p. S45–S74, 2014. Citado 2 vezes nas páginas 35 e 37.
- RUSSELL, J. B. et al. Risk factors, clinical outcomes and predictors of stroke mortality in sierra leoneans: a retrospective hospital cohort study. *Annals of medicine and surgery*, Elsevier, v. 60, p. 293–300, 2020. Citado na página 28.
- SACCO, R. L. et al. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the american heart association/american stroke association. *Stroke*, Lippincott Williams & Wilkins Hagerstown, MD, v. 44, n. 7, p. 2064–2089, 2013. Citado 4 vezes nas páginas 31, 32, 33 e 34.
- SAMEI, E.; PELC, B. J. *Computed tomography*. [S.l.]: Springer, 2020. Citado na página 36.
- SANTOS, L. d. O.; MEDEIROS, A. G.; FILHO, P. P. R. Fast fully automatic stroke lesions segmentation based on parzen estimation and μ -law in skull ct images. In: *Congresso Brasileiro de Automática-CBA*. [S.l.: s.n.], 2022. v. 3, n. 1. Citado 3 vezes nas páginas 41, 93 e 94.
- SCHWAMM, L. et al. Translating evidence into practice: a decade of efforts by the american heart association/american stroke association to reduce death and disability due to stroke: a presidential advisory from the american heart association/american stroke association. *Stroke*, Lippincott Williams & Wilkins, v. 41, n. 5, p. 1051–1065, 2010. Citado na página 34.

SHEHAB, M. et al. Machine learning in medical applications: A review of state-of-the-art methods. *Computers in Biology and Medicine*, Elsevier, v. 145, p. 105458, 2022. Citado 2 vezes nas páginas 25 e 26.

SILVA, I. N. d.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. 2010. Citado 5 vezes nas páginas 17, 49, 50, 51 e 52.

SOUZA, J. W. M. de et al. New approach to detect and classify stroke in skull ct images via structural co-occurrence matrix and machine learning. In: SPRINGER. *XXVI Brazilian Congress on Biomedical Engineering: CBEB 2018, Armação de Buzios, RJ, Brazil, 21-25 October 2018 (Vol. 2)*. [S.l.], 2019. p. 341–348. Citado 3 vezes nas páginas 38, 75 e 88.

SOUZA, L. F. D. F. et al. Divisible cell-segmentation: A new approach for stroke detection and segmentation in ct scans using deep learning and fine-tuning. In: IEEE. *2023 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2023. p. 1–9. Citado 3 vezes nas páginas 40, 93 e 94.

VOS, T. et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The lancet*, Elsevier, v. 396, n. 10258, p. 1204–1222, 2020. Citado na página 34.

WALLACE, E. J.; LIBERMAN, A. L. Diagnostic challenges in outpatient stroke: stroke chameleons and atypical stroke syndromes. *Neuropsychiatric Disease and Treatment*, Taylor & Francis, p. 1469–1480, 2021. Citado 2 vezes nas páginas 26 e 27.

WANG, C. et al. Abus-net: Graph convolutional network with multi-scale features for breast cancer diagnosis using automated breast ultrasound. *Expert Systems with Applications*, Elsevier, v. 273, p. 126978, 2025. Citado na página 44.

XIAO, Z. et al. Rnn-combined graph convolutional network with multi-feature fusion for tuberculosis cavity segmentation. *Signal, Image and Video Processing*, Springer, v. 17, n. 5, p. 2297–2303, 2023. Citado na página 45.

XU, W. et al. Automatic segmentation of intracranial hemorrhage in computed tomography scans with convolution neural networks. *Journal of Medical and Biological Engineering*, Springer, p. 1–7, 2024. Citado na página 40.

XU, Y. et al. Deep learning-enhanced internet of medical things to analyze brain ct scans of hemorrhagic stroke patients: a new approach. *IEEE Sensors Journal*, IEEE, v. 21, n. 22, p. 24941–24951, 2020. Citado na página 38.

XUAN, P. et al. Dynamic graph convolutional autoencoder with node-attribute-wise attention for kidney and tumor segmentation from ct volumes. *Knowledge-Based Systems*, Elsevier, v. 236, p. 107360, 2022. Citado na página 43.

YAN, W. et al. Deep learning in neuroimaging: Promises and challenges. *IEEE Signal Processing Magazine*, IEEE, v. 39, n. 2, p. 87–98, 2022. Citado na página 25.

YU, X. et al. Resgnet-c: A graph convolutional neural network for detection of covid-19. *Neurocomputing*, Elsevier, v. 452, p. 592–605, 2021. Citado 2 vezes nas páginas 44 e 76.

YU, X.; WANG, S.-H.; ZHANG, Y.-D. Cgnet: A graph-knowledge embedded convolutional neural network for detection of pneumonia. *Information Processing & Management*, Elsevier, v. 58, n. 1, p. 102411, 2021. Citado na página 76.

YUAN, Y. et al. Computed tomography image segmentation of irregular cerebral hemorrhage lesions based on improved u-net. *Journal of Radiation Research and Applied Sciences*, Elsevier, v. 16, n. 3, p. 100638, 2023. Citado na página 40.

ZHANG, A. et al. *Dive into deep learning*. [S.l.]: Cambridge University Press, 2023. Citado 2 vezes nas páginas 17 e 57.