

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

Virgínia Xavier Nunes

**Uma abordagem de aprendizagem profunda baseada em Internet das
Coisas para segmentação de imagens médicas em exames de
Tomografia Computadorizada**

Fortaleza – CE

15 de julho de 2021

Virgínia Xavier Nunes

**Uma abordagem de aprendizagem profunda baseada em
Internet das Coisas para segmentação de imagens
médicas em exames de Tomografia Computadorizada**

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: Visão Computacional

Orientador: Prof. Dr. Pedro Pedrosa Rebouças Filho

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

15 de julho de 2021

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)

N972a Nunes, Virginia Xavier.

Uma abordagem de aprendizagem profunda baseada em Internet das Coisas para segmentação de imagens médicas em exames de Tomografia Computadorizada / Virginia Xavier Nunes. - 2021.
74 f. : il.

Dissertação (Mestrado) - Instituto Federal do Ceará, Mestrado em Ciência da Computação, Campus Fortaleza, 2021.

Orientação: Prof. Dr. Pedro Pedrosa Rebouças Filho.

1. IoT. 2. Classificação. 3. Segmentação. 4. Transferência de Aprendizagem. 5. Ajuste Fino. 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)

Virgínia Xavier Nunes

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)

Prof. Dr. Francisco Nivando Bezerra

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

Prof. Dr. Edson Cavalcanti Neto

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

Fortaleza – CE
15 de julho de 2021

Agradecimentos

Como tudo em minha vida, agradeço primeiramente a Deus, que me auxiliou e guiou até aqui.

Agradeço também aos meus pais, Solange Maria Xavier Nunes e Fernando Antônio Travassos Nunes, por me fornecer a estrutura básica para que eu pudesse estudar.

Ao meu orientador, Prof. Dr. Pedro Pedrosa, por me acolher como orientanda e por todo o suporte oferecido para que eu pudesse desenvolver minha pesquisa.

Aos professores do Programa de Pós-Graduação em Ciência da Computação do Instituto Federal do Ceará (PPGCC-IFCE) por contribuírem com meu conhecimento.

À Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUN-CAP) pelo suporte financeiro.

Ao Aldísio Medeiros por todos os ensinamentos e orientações prestadas desde o início do meu trabalho e pela grande colaboração em ideias e pesquisas. Aproveito para agradecer também ao Luís Fabrício pelo trabalho em equipe e disposição em ajudar e a todos os integrantes do Laboratório de Processamento de Imagens, Sinais e Computação Aplicada (LAPISCO) que contribuíram de forma direta e indireta neste trabalho.

Também agradeço, em especial, ao meu esposo Raylson Silva de Lima, pelo amor, cuidado e dedicação depositados ao longo da minha jornada desde o primeiro semestre da graduação. Hoje compreendo que o propósito da nossa união não é apenas construir uma bela relação, mas vai muito além disso.

*“Deus não escolhe os capacitados, capacita os escolhidos.
Fazer ou não fazer algo só depende da nossa vontade e perseverança.”*
Albert Einstein

Resumo

A classificação e segmentação de patologias através de sistemas inteligentes é um desafio significativo para a análise de imagens médicas em sistemas de visão computacional. Doenças como problemas pulmonares e acidentes vasculares cerebrais têm um efeito sério na saúde humana em todo o mundo. As doenças pulmonares estão entre as principais causas de morte ao redor do mundo ficando atrás dos acidentes vasculares cerebrais (AVC), que em 2016 se tornaram a segunda principal causa de morte por doenças. A tomografia computadorizada (TC) é um dos principais exames de diagnóstico clínico vinculado aos sistemas de Diagnóstico Auxiliado por Computador (CAD) que estão se tornando soluções para tecnologias de saúde. Neste trabalho, propomos um método baseado na Internet das Coisas (IoT) para a classificação e segmentação do pulmão e AVC hemorrágico em imagens de tomografia. O sistema denominado *HTSCS - Medical Images: Health-of-Things System for the Classification and Segmentation of Medical Images* usa transferência de aprendizagem entre modelos baseados em aprendizagem profunda (Detectron2) combinados com métodos clássicos para ajuste fino. O método proposto obteve resultados promissores para a classificação de presença ou ausência de AVC hemorrágico e regiões pulmonares em imagens, com valores próximos a 100% de acurácia. Os modelos também alcançaram desempenhos excelentes na tarefa de segmentação, com acurácia acima de 99% e coeficiente de Dice acima de 97% nos melhores casos com um tempo médio de segmentação entre 0,095 e 1,7 segundos. Para validar essa abordagem comparamos os melhores modelos para a segmentação do pulmão e AVC hemorrágico em tomografias com trabalhos relacionados encontrados no estado da arte. Esse método traz uma abordagem inovadora para classificação e segmentação por meio do uso de IoT para diferentes tipos de imagens médicas com resultados promissores para análise de imagens médicas.

Palavras-chave: IoT, Classificação e Segmentação, Transferência de Aprendizagem, Ajuste Fino.

Abstract

The classification and segmentation of pathologies through intelligent systems is a significant challenge for medical image analysis and computer vision systems. Diseases, such as lung problems and strokes, have a serious effect on human health worldwide. Lung diseases are among the leading causes of death worldwide, lagging behind strokes that in 2016 became the second leading cause of death from illnesses. Computed tomography (CT) is one of the main clinical diagnostic exams, linked to Computerized Diagnostic Assistance Systems (CAD), which are becoming solutions for health technologies. In this work, we propose a method based on the health of things for the classification and segmentation of computed tomography (CT) images of the lung and hemorrhagic stroke. The system called HTSCS - Medical Images: Health-of-Things System for the Classification and Segmentation of Medical Images, uses transfer learning between models based on deep learning (Detec-tron2) combined with classical methods for fine-tuning. The proposed method obtained excellent results for the classification of hemorrhagic stroke and pulmonary regions, with values of up to 100% accuracy. The models also achieved outstanding performances for segmentation, with Accuracy above 99% and Dice coefficient (DICE) above 97% in the best cases with an average segmentation time between 0.095 and 1.7 seconds. To validate our approach, we compared our best models for the segmentation of lung and hemorrhagic stroke in CTs, with related works found in state of the art. Our method brings an innovative approach to classification and segmentation through the use of the Health of Things for different types of medical images with promising results for medical image analysis and computer vision fields.

Keywords: Health of Things, Classification and Segmentation, Transfer Learning, Fine-tuning.

Lista de ilustrações

Figura 1 – MLP com uma camada escondida.	37
Figura 2 – Pseudocódigo do algoritmo <i>K-means</i> para a segmentação de imagens. .	39
Figura 3 – Apresentação do <i>HTSCS - Medical Images: Health-of-Things System for the Classification and Segmentation of Medical Images</i> . A figura ilustra as etapas do método baseado na Saúde das Coisas. Na imagem, o símbolo * representa a etapa inicial de treinamento do método para classificação e segmentação de imagens tomográficas de pulmão e AVC hemorrágico. A etapa 1 consiste na classificação da imagem de entrada (com a detecção ou não de pulmão ou AVC hemorrágico). Na etapa 2 a rede Detectron2 inicia o processo de segmentação do pulmão ou AVC hemorrágico gerando mapas característicos através das caixas delimitadoras. A etapa 3 continua do processo de segmentação usando o ajuste fino e, finalmente, apresenta o melhor resultado.	44
Figura 4 – LINDA framework. 1- Tela de login e senha. 2- Criação do projeto, que inclui definir nome, número de classes e que ação (extração de atributos, classificação ou predição do modelo treinado) o usuário deseja realizar. 3- Caso o usuário decida treinar um novo modelo, as imagens separadas por classes serão carregadas. 4- Ao fazer o upload das imagens, o usuário escolhe quais descritores e classificadores serão utilizados no treinamento. 5- Análise dos resultados após o treinamento com gráficos de barras e matrizes de confusão.	45
Figura 5 – Resultado da instância de segmentação da rede Detectron2 em imagens de TC do tórax e imagens de TC do crânio, respectivamente. As instâncias detectadas foram sobrepostas com várias cores.	47
Figura 6 – Ilustração gráfica da Tabela 3. Diferentes modelos propostos com base na Saúde das Coisas com o uso de aprendizagem profunda aliado à técnica de ajuste fino.	51
Figura 7 – Resultado da segmentação com os diferentes modelos baseados na rede Detectron2 combinados com métodos de ajuste fino.	53
Figura 8 – Ilustração do tempo de segmentação dos modelos Detectron2, Detectron- $\mathbf{f}\lambda$, Detectron- $\mathbf{f}\delta$ e Detectron- $\mathbf{f}\mu$ para a segmentação do pulmão e AVC hemorrágico. As colunas de creme representam o tempo médio para segmentação de imagens de AVC hemorrágico por TC. As colunas de cor azul representam o tempo médio de segmentação das imagens de TC do pulmão.	54

Figura 9 – Ilustração gráfica das métricas obtidas pelos métodos Detectron2, Detectron- $f\lambda$, Detectron- $f\delta$ e Detectron- $f\mu$ na segmentação do AVC hemorrágico.	56
Figura 10 – Ilustração do resultado da segmentação da rede Detectron2 para o conjunto de dados de imagens de TC de AVC Hemorrágico.	57
Figura 11 – Ilustração gráfica da Tabela Comparativa 9.	59
Figura 12 – Gráfico de barras ilustrando o coeficiente de Dice de cada método. . . .	60

Lista de tabelas

Tabela 1	– Tabela com os dois melhores resultados da combinação extrator-classificador para a etapa de detecção, quando há ou não pulmão na imagem. Como pode ser visto, a combinação do extrator VGG16 e do extrator Xception, ambos com o classificador SVM-RBF, obteve 97% em todas as métricas de avaliação.	50
Tabela 2	– Tabela com os dois melhores resultados da combinação extrator-classificador para detecção de presença ou não de AVC hemorrágico em imagens de crânio. Como pode ser visto, a melhor combinação é do extrator VGG16 com o classificador SVM-RBF, que pontuou 100% em todas as métricas de avaliação com o menor tempo de extração.	50
Tabela 3	– Resultados gerados pelo método proposto usando aprendizagem profunda e ajuste fino para segmentar o pulmão em imagens de TC. As mesmas imagens foram utilizadas nos trabalhos de Filho et al. (2019) e Hu et al. (2020).	51
Tabela 4	– Comparação do tempo de segmentação entre o modelo Detectron2 e os modelos propostos baseados em técnicas de ajuste fino para a segmentação do pulmão em imagens de TC.	54
Tabela 5	– Tabela de Teste Estatístico Kolmogorov-Smirnov. Resultados das métricas na segmentação das imagens de TC do pulmão, utilizando as técnicas: A) Detectron2, B) Detectron- $f\lambda$ C) Detectron- $f\delta$, e D) Detectron- $f\mu$	55
Tabela 6	– Resultados gerados pelo método proposto, baseado na rede Detectron2 em conjunto com técnicas de ajuste fino no conjunto de dados de TC de AVC hemorrágico.	56
Tabela 7	– Comparação dos tempos de segmentação entre modelos Detectron2 e modelos usando técnicas de ajuste fino. Os resultados foram gerados a partir de 20% das imagens de AVC hemorrágico.	58
Tabela 8	– Tabela de Teste Estatístico Kolmogorov-Smirnov. Resultados das métricas utilizadas na segmentação de imagens de TC do cérebro, utilizando as seguintes técnicas: A) Detectron2, B) Detectron- $f\lambda$, C) Detectron- $f\delta$ e D) Detectron- $f\mu$	58
Tabela 9	– Tabela de validação para o método proposto na etapa de segmentação do sistema HTSCS com o conjunto de dados de imagens pulmonares. Os métodos usados para validação podem ser encontrados no artigo de Hu et al. (2020).	58

Tabela 10 – Tabela de Validação do método proposto na etapa de segmentação do sistema HTSCS para o conjunto de dados de imagem pulmonar, os métodos utilizados na validação podem ser encontrados nos artigos (HU, GVF, VFC, SISDEP, OPS, CRAD) e Rebouças et al. (2019).	60
Tabela 11 – Comparação entre métodos relatados em Rebouças et al. (2019), com nossa melhor abordagem (Detectron2) para segmentação de AVC hemorrágico.	60
Tabela 12 – Comparação do tempo de segmentação para imagens de AVC hemorrágico.	60
Tabela 13 – Tabela de <i>Highlights</i> : Comparação de trabalhos utilizados como motivação para o método proposto.	62

Lista de abreviaturas e siglas

AVC	Acidente Vascular Cerebral
TC	Tomografia Computadorizada
CAD	Diagnóstico Auxiliado por Computador (<i>Computer-Aided Diagnosis</i>)
HTSCS Medical Images	Sistema de Saúde das Coisas para a Classificação e Segmentação de Imagens Médicas (<i>Health-of-Things System for the Classification and Segmentation of Medical Images</i>)
DPOC	Doença Pulmonar Obstrutiva Crônica
OMS	Organização Mundial da Saúde
RoI	Região de Interesse (<i>Region of Interest</i>)
CNN	Rede Neural Convolutacional (<i>Convolutional Neural Network</i>)
Mask R-CNN	Rede Neural Convolutacional Regional de Máscara (<i>Mask Region-based Convolutional Neural Network</i>)
IoT	Internet das Coisas (<i>Internet of Things</i>)
PDI	Processamento Digital de Imagens
MRI	Imagem por Ressonância Magnética (<i>Magnetic Resonance Imaging</i>)
DICOM	Comunicação de Imagens Digitais em Medicina (<i>Digital Imaging and Communications in Medicine</i>)
SVM	Máquinas de Vetores de Suporte (<i>Support Vector Machine</i>)
MLP	Perceptron Multicamadas (<i>Multilayer Perceptron</i>)
RF	Floresta Aleatória (<i>Random Forest</i>)
KNN	k-Vizinhos Mais Próximos (<i>k-Nearest Neighbor</i>)
ACC	Acurácia
DICE	Coefficiente de Dice
SEN	Sensibilidade
ESP	Especificidade

API	Interface de Programação de Aplicação (<i>Application Programming Interface</i>)
LINDA	Lapisco Image Interface for Application Development

Sumário

1	INTRODUÇÃO	21
1.1	Objetivos Geral e Específicos	24
1.2	Produção Científica	24
1.3	Estrutura	26
2	TRABALHOS RELACIONADOS	27
2.1	Processamento de Imagens Médicas	27
2.2	Sistemas IoT em imagens médicas	28
2.3	Visão computacional aplicada a imagens de regiões pulmonares	28
2.4	Visão computacional aplicada à imagens de regiões cerebrais	29
3	MATERIAIS E MÉTODOS	31
3.1	Conjuntos de Dados	31
3.2	Redes Neurais Convolucionais	31
3.2.1	Extração Profunda de Atributos	32
3.2.2	Detectron2	33
3.3	Métodos de Aprendizagem de Máquina	34
3.3.1	Naive Bayes	34
3.3.2	k-Vizinhos Mais Próximos	35
3.3.3	Máquinas de Vetores de Suporte	36
3.3.4	Floresta Aleatória	37
3.3.5	Perceptron Multicamadas	37
3.4	Técnicas de PDI como Ajuste Fino para a Segmentação de Objetos	38
3.4.1	Crescimento de Regiões	38
3.4.2	K-means	38
3.4.3	Janela de Parzen	39
3.5	Métricas de Avaliação	40
3.6	Teste de Kolmogorov-Smirnov	41
4	METODOLOGIA	43
4.1	Saúde das Coisas	43
4.2	Classificação	43
4.3	Segmentação	46
4.4	Ajuste Fino	47
5	RESULTADOS	49

5.1	Primeira Etapa do Experimento - Classificação	49
5.2	Segunda Etapa do Experimento - Segmentação	50
5.2.1	Segunda Etapa do Experimento - Segmentação do Pulmão	50
5.2.2	Segunda Etapa do Experimento - Segmentação de AVC Hemorrágico	55
5.2.3	Terceira Etapa - Comparação entre métodos da literatura	58
6	CONCLUSÃO E TRABALHOS FUTUROS	63
	REFERÊNCIAS	65

1 Introdução

Diversas patologias afetam gravemente a saúde humana em todo o mundo, sendo que as principais estão relacionadas aos pulmões, cérebro e coração. A Doença Pulmonar Obstrutiva Crônica (DPOC) é a principal causa de mortalidade respiratória no mundo (SORIANO et al., 2017) e foi a terceira causa de morte no mundo, segundo a Organização Mundial da Saúde (OMS), em 2016 (ORGANIZATION, 2017). Além disso, de acordo com a OMS, cerca de 3,2 milhões de mortes foram causadas por DPOC em 2015. Isso representa um vigésimo de todas as mortes, a nível global, naquele ano e mais de 90% dessas mortes ocorreram em países de baixa e média renda. Em 2020 um total de 200 milhões de pessoas em todo o mundo foram diagnosticadas com DPOC sendo que muitas estão vivendo com doenças não diagnosticadas (FORD et al., 2020).

A DPOC envolve bronquiolite obstrutiva, enfisema ou características de ambas. A primeira causa um estado permanente de inflamação das vias aéreas causando inchaço dentro dos tubos respiratórios interferindo na capacidade de fluxo de ar e eficiência dos pulmões, enquanto a segunda destrói os alvéolos, que são estruturas que promovem as trocas gasosas no órgão (LAREAU et al., 2019). Essas condições são perigosas pois além do potencial de inibir totalmente a respiração, diminui a circulação de oxigênio no sangue e desencadeia outras respostas inflamatórias por todo o corpo fazendo com que o risco de ataque cardíaco e derrame sejam duplicados. Os pacientes também podem sofrer de fraqueza muscular, raciocínio prejudicado e até mesmo depressão.

Embora a maioria da DPOC seja causada pela fumaça do tabaco (cerca de 85%) (FOREY; THORNTON; LEE, 2011), outros compostos prejudiciais, como a forte exposição à poeiras no trabalho, produtos químicos e poluição do ar interno ou externo (incluindo fumaça de lenha ou combustíveis de biomassa) e fatores genéticos (herdados) podem contribuir para desenvolver DPOC. O primeiro desafio para identificar e classificar a DPOC é detectar o contorno da região do pulmão. Lembrando que esta doença altera visualmente a aparência do pulmão no exame de tomografia computadorizada (TC) e altera parcialmente a nitidez dos limites pulmonares (HU; HOFFMAN; REINHARDT, 2001).

Assim como a DPOC, o Acidente Vascular Cerebral (AVC) é responsável por taxas significativas de mortalidade e incapacidade humana com custos econômicos consideráveis para o atendimento pós-AVC, em escala global (RAJSIC et al., 2019). Os acidentes vasculares cerebrais são um grupo de doenças que envolvem a interrupção repentina do fluxo sanguíneo no cérebro. A obstrução das artérias cerebrais pode causar deficiências neurológicas. Os derrames hemorrágicos e isquêmicos são os dois tipos mais comuns

de derrame. Um derrame hemorrágico ocorre quando uma artéria no cérebro se rompe causando um sangramento amplo e devastador no cérebro, enquanto um acidente vascular cerebral isquêmico ocorre quando um vaso sanguíneo, que leva sangue para o cérebro, é bloqueado (SARMENTO et al., 2019).

Em 2016, o AVC foi a segunda maior causa de morte no mundo (5,5 milhões). O número de mulheres que morreram de AVC foi ligeiramente inferior ao número de homens (2,6 milhões e 2,9 milhões, respectivamente). O número de mortes por AVC isquêmico foi de 2,7 milhões, um pouco menos do que o número de mortes devido ao AVC hemorrágico que foi de 2,88 milhões (JOHNSON et al., 2019).

Em 2015, segundo a American Heart Association e a American Stroke Association, ocorreram aproximadamente 800.000 AVCs nos Estados Unidos, sendo responsáveis por uma a cada 20 mortes (MOZAFFARIAN et al., 2016). Além do alto índice de mortalidade, grande parte das pessoas que sobrevivem ao AVC acaba apresentando algum tipo de deficiência em relação às atividades básicas comprometendo a qualidade e a expectativa de vida (TEASELL; HUSSEIN, 2016). A deficiência varia de acordo com o grau de recuperação neurológica, a localização da lesão, o estado pré-mórbido do paciente e os sistemas de suporte disponíveis (SARMENTO et al., 2019).

Atualmente, várias áreas médicas realizam diagnósticos por meio de imagens (VANNIER, 2009). O equipamento de TC se destaca como o instrumento mais importante para a aquisição dessas imagens (HERMAN, 2009) devido à sua disponibilidade em quase todas as unidades de emergência e à rápida obtenção dos resultados. Além disso, a TC tem ganhado importância cada vez maior, pois o diagnóstico é menos invasivo do que alguns outros sistemas e dá resultados precisos (DIAZ; CELLI; CELEDÓN, 2018). Além disso, pode ser usado para adquirir imagens de pulmão, coração, cérebro, artérias, ossos, entre outros (MUZAMMAL et al., 2020).

O diagnóstico de DPOC por meio de TC ajuda a avaliar a extensão e distribuição da doença (MADANI; KEYZER; GEVENOIS, 2001), que é estimada por quantificação visual ou pela análise da distribuição da densidade pulmonar (LYNCH; AL-QAISI, 2013) proporcionando uma avaliação mais precisa e objetiva da doença (MADANI; KEYZER; GEVENOIS, 2001). Ao contrário da DPOC, o AVC é considerado uma emergência médica e precisa ser diagnosticado e tratado prontamente para minimizar as implicações que podem ocorrer (FILHO et al., 2017). Desta forma, a TC se apresenta como a técnica mais adequada e financeiramente viável, devido ao seu baixo custo e agilidade (FERREIRA; RUANO, 2011).

Os diagnósticos por exames de TC podem ser melhorados por meio de sistemas de Diagnóstico Auxiliado por Computador (CAD). Graças à atuação dos sistemas CAD a eficiência e precisão dos diagnósticos clínicos por meio da detecção e/ou classificação automática de anormalidades e/ou doenças em exames médicos radiológicos podem ser

aprimorados. Vale ressaltar que diversos sistemas comerciais foram desenvolvidos para propósitos específicos. Por exemplo, diversos sistemas CAD têm como objetivo detectar câncer de mama, pulmão ou cólon usando raios X, TC ou ressonância magnética.

Para que o sistema CAD auxilie no diagnóstico, o primeiro desafio consiste em localizar e segmentar a região de interesse (RoI) da imagem. Sendo assim, diversas técnicas de segmentação são aplicadas para localizar a RoI em uma imagem. No caso das imagens médicas, é comum o uso de métodos de segmentação para demarcar órgãos e associá-los ao estudo de patologias. Diante disso, um número significativo de estudos é produzido enfatizando a utilização de sistemas CAD no diagnóstico de doenças.

Dentre as técnicas desenvolvidas para automatizar a tarefa de segmentação de regiões pulmonares, temos as técnicas de segmentação tradicionais baseadas no crescimento da região (NÓBREGA; RODRIGUES; FILHO, 2017), crescimento da região com limiares locais e abordagens de bacias hidrográficas (AVINASH; MANJUNATH; KUMAR, 2016). Porém, essas técnicas tradicionais não obtêm segmentação consistente para regiões com parênquima de baixo contraste levando a resultados inadequados quando aplicadas em exames de TC (FILHO et al., 2014).

Trabalhos recentes usando técnicas de Rede Neural Convolucional (CNN) para a detecção (YU et al., 2019), segmentação (DURASAMY; DURASAMY, 2012) e classificação (DING et al., 2019) de objetos de interesse em imagens foram propostos com resultados significativos. Porém, quando se trata de imagens médicas, essa técnica torna-se um desafio (SHIN et al., 2016), tendo em vista que a realização de um treinamento eficiente desses modelos profundos necessita de um grande conjunto de imagens médicas (JR et al., 2019). Para superar isso, uma estratégia chamada transferência de aprendizagem (WEISS; KHOSHGOFTAAR; WANG, 2016) sugere que os recursos aprendidos para resolver um problema específico podem ser usados para resolver problemas em outros domínios (NÓBREGA et al., 2018).

Apesar disso, vários estudos usando métodos de aprendizado profundo para segmentação em imagens médicas foram desenvolvidos (KER et al., 2017). Hu et al. (2020) propôs a Rede Neural Convolucional Regional de Máscara (Mask R-CNN) combinada com métodos de aprendizado de máquina supervisionados e não supervisionados para segmentação automática dos pulmões em imagens de TC. Medeiros et al. (2019) propôs uma nova abordagem usando a Mask R-CNN para segmentar o ventrículo esquerdo com sucesso. Já Zhang et al. (2018) propôs uma segmentação automática de AVC isquêmico agudo usando DenseNets totalmente convolucionais.

As técnicas de ajuste fino surgiram com novas possibilidades no campo da visão computacional (ZHOU et al., 2017) e ciência de dados (MIKOLOV et al., 2013). Paralelamente ao uso da transferência de aprendizagem, diferentes sistemas de monitoramento na área médica baseados nos princípios de Internet das Coisas (IoT) (RACHEDI et al., 2016)

têm sido propostos. Esses sistemas, denominados de Saúde das Coisas, trabalham com a troca rápida de informações médicas sobre as condições dos pacientes, imagens médicas e diagnósticos otimizando o trabalho dos profissionais de saúde que, além disso, promove melhorias significativas na qualidade dos tratamentos médicos para diminuir os custos médicos por consulta, acompanhamento e diagnóstico (SCARPATO et al., 2017).

1.1 Objetivos Geral e Específicos

Motivados pelo sucesso dos sistemas de Saúde das Coisas propomos um sistema de auxílio ao diagnóstico médico denominado *HTSCS - Medical Images: Health-of-Things System for the Classification and Segmentation of Medical Images* utilizando princípios de IoT, transferência de aprendizagem, aprendizagem profunda e ajuste fino.

Especificamente, este trabalho visa:

1. Extrair atributos profundos de bases de dados de imagens do pulmão e de AVC hemorrágico usando dois modelos de CNN pré-treinados no Desafio de Reconhecimento Visual de Grande Escala ImageNet (ILSVRC), como: Xception (CHOLLET, 2017) e VGG16 (SZEGEDY et al., 2015).
2. Realizar o treinamento do classificador Máquinas de Vetores de Suporte (SVM) para cada conjunto de atributos extraídos.
3. Classificar as imagens de TC do pulmão e AVC hemorrágico e separar as imagens que contenham ou não RoI para a etapa de segmentação.
4. Realizar o treinamento da CNN Detectron2 através da técnica de ajuste fino para cada bases de dados.
5. Segmentar a área de interesse usando a CNN Detectron2 já ajustada.
6. Aplicar três técnicas de ajuste fino baseadas no Crescimento da Região (ADAMS; BISHOF, 1994), Agrupamento K-means (DHANACHANDRA; MANGLEM; CHANU, 2015) e Janela de Parzen (REBOUÇAS et al., 2019) para melhorar a qualidade da segmentação.
7. Comparar os resultados entre os métodos propostos e outros trabalhos da literatura por meio de cinco métricas de avaliação: Acurácia (ACC), Coeficiente de Dice (DICE), Sensibilidade (SEN), Especificidade (ESP) e Tempo.

1.2 Produção Científica

Durante este projeto de mestrado, os seguintes artigos científicos foram publicados:

- **NUNES, VIRGINIA XAVIER**; MEDEIROS, A. G.; LIMA, R. S.; DE F. SOUZA, LUÍS FABRÍCIO; ALBUQUERQUE, V. H. C.; REBOUÇAS FILHO, PEDRO P. A Novel Web Platform for COVID-19 diagnosis using X-Ray exams and Deep Learning Techniques. In: International Joint Conference on Neural Networks (IJCNN), 2021, On line. IEEE Computational Intelligence Society and the International Neural Network Society, 2021. (Qualis A1 C.C.);
- LIMA, R. S.; **NUNES, VIRGINIA XAVIER**; MARQUES, ADRIELL GOMES; SOUZA, L. F. F.; ROCHA NETO, P. D.; BRAYNER, A. R. A.; ALBUQUERQUE, V. H. C.; REBOUÇAS FILHO, PEDRO P. New AIoT-Based Approach for Automatic Detection of COVID-19 Using Deep Learning Techniques in CT Images. In: International Conference on Systems, Signals and Image Processing (IWSSIP), 2021, Bratislava, Eslováquia. 28th International Conference on Systems, Signals and Image Processing, IWSSIP 2021. v. 28. (Qualis B1 C.C.);
- HERCULES SILVA, FRANCISCO DOS SANTOS; GOMES, ADRIELL; DE F. SOUZA, LUIS FABRICIO; **NUNES, VIRGINIA X.**; NASCIMENTO, J. J.; ROCHA NETO, P. D.; REBOUÇAS FILHO, PEDRO P. New Approach Detection of COVID-19 by Cough Spectrogram Using IoT and Learning Transfer Technique for Classification. In: International Conference on Systems, Signals and Image Processing (IWSSIP), 2021, Bratislava, Eslováquia. 28th International Conference on Systems, Signals and Image Processing, IWSSIP 2021, 2021. v. 28. (Qualis B1 C.C.);
- HAN, TAO; **NUNES, VIRGINIA XAVIER**; SOUZA, LUIS FABRICIO DE F.; MARQUES, ADRIELL GOMES; SILVA, IAGSON CARLOS L.; F., MARCOS AURELIO A.; SUN, JINGHUA; FILHO, PEDRO P. REBOUCAS. Internet of Medical Things-based Deep Learning Techniques for Segmentation of Lung and Stroke Regions in CT Scans. IEEE Access, v.8, p.71117-71135, 2020. (Qualis A2 C.C.);
- **NUNES, VIRGINIA X.**; MEDEIROS, ALDISIO G.; SILVA, FRANCISCO H. S.; BEZERRA, GABRIEL M.; FILHO, PEDRO P. R. Adaptive Level Set with region analysis via Mask R-CNN: A comparison against classical methods. In: 2020 International Joint Conference on Neural Networks (IJCNN), 2020, Glasgow. 2020 International Joint Conference on Neural Networks (IJCNN), 2020. p.1. (Qualis A1 C.C.);
- DE SOUZA, JOAO W. M.; Holanda, Gabriel B.; IVO, ROBERTO F.; ALVES, SHARA S. A.; DA SILVA, SUANE P. P.; **NUNES, VIRGINIA X.**; LOUREIRO, LUIZ LANNES; DIAS-SILVA, C. H.; REBOUCAS FILHO, PEDRO P. Predicting body measures from 2D images using Convolutional Neural Networks. In: 2020 International Joint Conference on Neural Networks (IJCNN), 2020,

- Glasgow. 2020 International Joint Conference on Neural Networks (IJCNN), 2020. p.1. (Qualis A1 C.C.);
- SOUZA, LUÍS FABRÍCIO DE FREITAS; SILVA, IÁGSON CARLOS LIMA; MARQUES, ADRIELL GOMES; SILVA, FRANCISCO HÉRCULES DOS S.; **NUNES, VIRGÍNIA XAVIER**; HASSAN, MOHAMMAD MEHEDI; AL-BUQUERQUE, VICTOR HUGO C. DE; FILHO, PEDRO P. REBOUÇAS. Internet of Medical Things: An Effective and Fully Automatic IoT Approach Using Deep Learning and Fine-Tuning to Lung CT Segmentation. SENSORS, v.20, p.6711, 2020. (Qualis A2 C.C.);

1.3 Estrutura

Este trabalho apresenta no Capítulo 1 uma contextualiza do problema, abordando os principais trabalhos literatura recente. No Capítulo 2 são apresentados os trabalhos relacionados que serão utilizados para validar a proposta desta dissertação, além dos objetivos gerais e específicos. No Capítulo 3 os materiais e fundamentos teóricos utilizados neste trabalho serão apresentados. O Capítulo 4 descreve a metodologia proposta nesta pesquisa. O Capítulo 5 apresenta os resultados alcançados, bem como um estudo comparativo com técnicas recentes aplicadas aos mesmos conjuntos de dados utilizados neste trabalho. Por fim, o Capítulo 6 traz as conclusões do estudo e propostas para trabalhos futuros.

2 Trabalhos Relacionados

A seguir, vários trabalhos nas áreas de pesquisa referenciadas por este trabalho serão apresentados, bem como trabalhos relacionados com o tema em diferentes contextos.

2.1 Processamento de Imagens Médicas

As técnicas de processamento de imagens têm se tornado cada vez mais comuns na área médica, especialmente métodos que usam filtros em visão computacional com sistemas CAD. Estas técnicas são muito procuradas devido aos resultados rápidos e aos processos de segmentação e classificação das imagens médicas (DOUGHERTY, 2009) serem eficazes.

Compreendendo a importância da aplicabilidade médica dos sistemas CAD, Bouchet, Pastore e Ballarín (2007) propuseram o uso da morfologia matemática *fuzzy* (LI et al., 2016) na segmentação de ramos em imagens angiográficas. Os resultados foram visualmente superiores quando comparados às técnicas convencionais de processamento digital de imagens (PDI) para os mesmos casos. No entanto, a análise foi apenas visual e sem comparação com os resultados fornecidos por um médico especialista contendo o padrão ouro.

Ciente do poder da técnica de crescimento da região, Duan et al. (2019) propôs um método de segmentação dos vasos pulmonares em imagens de TC usando operações computacionais para filtrar a difusão anisotrópica e o crescimento da região. Seus resultados obtiveram uma sensibilidade de 92,9% e especificidade de 91,6%. Porém, o algoritmo desenvolvido só consegue segmentar com sucesso os vasos de um pulmão saudável, além do fato de o método apresentar dificuldades em encontrar os parâmetros de filtragem ideais, o que consequentemente pode ter reduzido os valores das métricas de avaliação.

Raja et al. (2018) propôs um método de segmentação de imagens T1 de ressonância magnética do cérebro e mama doentes, usando a integração do algoritmo Chaotic Bat, limiar baseado em Tsallis e a técnica de crescimento da região. Depois de testar o algoritmo no *Brain Tumor Segmentation Challenge*, o método obteve 97,5% de acurácia e 90,36% de DICE. Esse método encontrou dificuldades apenas em segmentar algumas fatias do cérebro. No entanto, o método não é automático, o que exige intervenção humana para ajustar os parâmetros de cada imagem.

Pei et al. (2017) desenvolveu uma técnica baseada na densidade de imagens de TC para pré-definir o número de grupos visando um agrupamento automático que auxilie na segmentação de imagens médicas. Os autores usaram um limite simples para separar a região agrupada de interesse. Os resultados usando diferentes conjuntos de dados, como

Ecoli e Iris, obtiveram uma acurácia média de 83,15%. Resta saber se o método proposto é capaz de agrupar efetivamente imagens de TC do cérebro.

2.2 Sistemas IoT em imagens médicas

Os sistemas de monitoramento médico, denominado Saúde das Coisas (MA et al., 2016), realizam uma rápida troca de informações sobre as condições dos pacientes, imagens médicas e diagnósticos para otimizar o trabalho dos profissionais da área clínica e gerar uma melhoria significativa na qualidade de tratamentos (SCARPATO et al., 2017).

Santos, Perkusich e Almeida (2014) propuseram em seu trabalho uma arquitetura que engloba a utilização de diferentes tipos de gerenciamento de saúde e gateways. Além da interoperabilidade, por meio da utilização de padrões adotados, a arquitetura proposta torna a troca de dados entre máquinas muito mais rápida fazendo com que o gerenciamento seja mais eficaz. Porém, por ser uma técnica recente, esse método ainda apresenta algumas falhas de segurança e problemas de interface.

Seguindo a mesma linha de pensamento, Al-Hamadi e Chen (2017) propôs um protocolo de comunicação entre os sistemas IoT de saúde com base na tomada de decisão. Este protocolo permite a criação de uma base de dados de conhecimento coletivo entre os dispositivos, que posteriormente possibilitará a tomada de decisões pelos dispositivos agendados.

Ciente da eficácia dos sistemas baseados na Saúde das Coisas no monitoramento da saúde, Ray (2014) desenvolveu o H3IoT, um framework arquitetônico baseado no IoT Home Health Hub. A arquitetura dessa solução tem como foco o acompanhamento da saúde dos idosos em suas residências. Os autores concluíram que o sistema é leve e fácil de instalar, mas que necessita de alguns ajustes para funcionar em ambiente clínico.

2.3 Visão computacional aplicada a imagens de regiões pulmonares

Sabendo da importância da segmentação dos pulmões para a detecção de doenças pulmonares, Hu et al. (2020) propôs um método automático usando a rede Mask R-CNN e Aprendizagem de Máquina junto com o auxílio de transferência de aprendizagem para segmentar imagens de TC de pulmão. O método obteve excelentes resultados (acurácia de 97,68%) em comparação com a segmentação de um médico especialista. A técnica realizou a segmentação em um tempo médio de 11,2 segundos, que é considerado um tempo razoável para a segmentação de tomografias de pulmão quando se trata de redes de aprendizado profundo e processos automáticos.

Com o mesmo propósito, Wang et al. (2017) usou CNN com Aprendizado de Máquina para a segmentação de nódulos pulmonares. O método obteve acurácia de apenas

80%, devido à dificuldade do método em segmentar pequenos nódulos e aos diferentes tamanhos de pulmão nas diversas imagens tomográficas. Essas diferenças de tamanho se devem ao processo do sistema respiratório durante a realização do exame.

Duraisamy e Duraisamy (2012) desenvolveram uma abordagem para segmentar imagens de ressonância magnética de AVC e pulmão. Essa técnica envolvia o uso de CNN, lógica difusa C e azul K vizinhos mais próximos e os resultados foram visualmente superiores ao uso apenas da CNN na segmentação. O método obteve 95% em acurácia. No entanto, o fluxo do conjunto de dados de imagens pulmonares pode influenciar os valores métricos.

Com base no exposto, Shin et al. (2016) propôs um estudo com as CNNs mais comumente usadas, como CifarNet, AlexNet, GoogleNet e outras. Embora todas essas redes sejam capazes de classificar os mais variados tipos de objetos, elas foram treinadas para classificar doenças pulmonares, como os linfonodo toracoabdominais e doença pulmonar intersticial. O estudo proposto pelos autores consistiu na utilização de transferência de aprendizagem entre essas redes para que o processo de segmentação fosse otimizado. No entanto, o método obteve apenas 85% de sensibilidade. Além disso, os modelos mostraram-se limitados à medida que os pesos foram reaproveitados e sua sensibilidade foi superada por métodos mais modernos.

Wang et al. (2018), propôs uma abordagem para a segmentação interativa de imagens médicas usando aprendizado profundo com ajuste fino no final do processo. Embora os resultados não sejam robustos e precisos o suficiente para uso clínico devido à variação insuficiente das amostras. O método proposto obteve resultados satisfatórios na segmentação automática de imagens médicas.

2.4 Visão computacional aplicada à imagens de regiões cerebrais

Ciente da importância de segmentar o cérebro de forma rápida e eficaz, Havaei et al. (2015) propôs um método para segmentação de tumor cerebral em imagens de ressonância magnética com a ajuda de uma arquitetura em cascata CNN. Os resultados obtiveram um DICE de 87% em comparação com a segmentação do médico especialista. O tempo médio de segmentação por imagem foi de 25 segundos. No entanto, o método teve dificuldade em detectar pequenos tumores no cérebro.

Chen, Bentley e Rueckert (2017) desenvolveram uma arquitetura composta por CNN EDD Net e MUSCULE Net para a segmentação automática de um banco de dados de 741 imagens de MRI-DWI de AVC isquêmico de diferentes pacientes. Os resultados alcançaram taxa de detecção de 94% e DICE de 67% quando comparados ao padrão ouro. No entanto, os autores admitem que o método apresenta limitações na segmentação de pequenas lesões.

Haan et al. (2015) elaborou um método de segmentação semiautomática de AVC isquêmico em imagens de TC e imagens por ressonância magnética (MRI), baseado em técnicas de agrupamento. Os resultados foram encorajadores com um Jaccard de 87%. No entanto, por ser semiautomático, o método requer a intervenção do operador para configurar os parâmetros iniciais. Sun et al. (2015) também propôs um método de segmentação 3D para Hemorragia Intra-Craniana em imagens de TC com o auxílio de um algoritmo Supervoxel, que por sua vez é baseado em Agrupamento Linear Iterativo Simples e refinado com o algoritmo *Graph*. O método proposto obteve uma taxa de verdadeiros positivos de 97,94% e taxa de falsos positivos de 92,26%.

No estudo de Rebouças et al. (2019), os autores propuseram um método de segmentação semiautomática de imagens de TC de crânio através do uso de uma janela de Parzen. O método obteve uma acurácia de 99,84% em comparação com o padrão ouro e superou os resultados baseados na nebulosa C, *watershed* e crescimento da região que foram usados para validar este método. Porém, como o método não é totalmente automático, o usuário precisa selecionar o ponto inicial de crescimento da região lesada até sua borda.

Com o objetivo de contribuir com o estado da arte na literatura, este estudo propõe um método baseado na Saúde das Coisas para classificar a presença ou ausência de pulmão e AVC hemorrágico e segmentar a região de interesse em imagens de TC. O sistema é denominado *HTSCS - Medical Images: Health-of-Things System for the Classification and Segmentation of Medical Images*. A classificação faz uso da rede VGG16 para a extração dos atributos e do classificador SVM-RBF. A segmentação é baseada no uso de aprendizagem profunda usando a CNN Detectron2 combinada com os métodos clássicos para o ajuste fino da segmentação.

3 Materiais e Métodos

Este capítulo reúne os materiais e métodos da literatura utilizados como fundamentação teórica no desenvolvimento deste trabalho. A Seção 3.1 apresenta os conjuntos de dados usados nos experimentos deste estudo. A Seção 3.2 apresenta as Redes Neurais Convolucionais voltadas para extração profunda de atributos e ajuste fino para a segmentação de região de interesse. Na Seção 3.3, os métodos de aprendizagem de máquina aplicados nos experimentos deste trabalho são apresentados. Na Seção 3.4 são apresentadas as técnicas de ajuste fino baseadas em PDI propostas neste trabalho. As métricas de avaliação aplicadas para validar a metodologia proposta são apresentadas na Seção 3.5. Por fim, o teste de Kolmogorov-Smirnov, também utilizado para validar os resultados obtidos, é descrito na Seção 3.6.

3.1 Conjuntos de Dados

O conjunto de dados de imagens de TC do pulmão contém 1.265 imagens, no formato Digital Imaging and Communications in Medicine (DICOM), com dimensões de 512×512 *pixels* e profundidade de 16 bits, juntamente com o padrão ouro. As imagens foram adquiridas em parceria com o Hospital Walter Cantídio da Universidade Federal do Ceará (UFC), Brasil, e aprovadas pelo Comitê de Ética em Pesquisa. Comitê - COMEPE (Protocolo nº 35/06).

O conjunto de dados de TC de AVC hemorrágico consiste em 100 imagens axiais de diferentes pacientes. O Hospital do Coração de Fortaleza, Brasil, forneceu as imagens, que foram geradas por um GE MEDICAL SYSTEM CT modelo Optima CT660. A configuração das imagens (parâmetros) foi uma espessura de corte de 0,7 mm, campo de visão de 230 mm, tensão do tubo de 120 kV e corrente elétrica de 80 mA. As dimensões das imagens são 512 *pixels* de altura por 512 *pixels* de largura, com um voxel de $0,488 \times 0,488 \times 1,5$ mm. A quantificação das imagens está no padrão 16 bits, salvas no formato DICOM.

3.2 Redes Neurais Convolucionais

As amplas possibilidades de aplicação de CNNs no campo médico ainda são, de maneira geral, desconhecidas. Isso se deve à escassez de dados com grandes quantidades de amostras necessárias para treinar esses modelos a partir do início (SHIN et al., 2016).

Para lidar com essa problemática e fazer uso das ferramentas de aprendizado profundo, tendo em vista seus excelentes resultados, a abordagem de Transferência de

Aprendizagem foi amplamente aplicada para o diagnóstico assistido por computador (NÓBREGA et al., 2018; HUYNH; LI; GIGER, 2016; SOUZA et al.,). Extração de atributos profundos é uma técnica de transferência de aprendizagem responsável por transformar o problema em um espaço de domínio diferente para aumentar o poder de discriminação de um conjunto de dados genérico. No caso específico de CNNs, um modelo pré-treinado em um grande conjunto de dados é usado para realizar a extração em um conjunto de dados não visto. A última camada do modelo é removida e a saída do modelo terá um tamanho equivalente ao tamanho da última camada remanescente remodelada em um vetor unidimensional (MOLCHANOV et al., 2016; KARPATHY et al., 2014).

Com o novo conjunto de dados de atributos extraídos via transferência de aprendizagem, podemos usar os algoritmos clássicos de aprendizagem de máquina para realizar o treinamento e generalização do conhecimento.

Além da transferência de aprendizagem por extração de atributos profundos, o ajuste fino também é um processo de transferência de aprendizagem que reaproveita um modelo de CNN pré-treinado em uma tarefa de contexto diferente visando solucionar uma segunda tarefa. Frequentemente, a técnica de ajuste fino em CNNs é aplicada em redes que são treinadas em datasets como o ImageNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) (com 1.2M de imagens com suas classes). No processo de ajuste fino, a rede continua a ser treinada com um dataset menor.

Na Subseção 3.2.1, uma apresentação sobre as CNNs utilizadas nesse trabalho, como os extratores de atributos profundos, é feita de forma breve. A subseção 3.2.2 apresenta a Rede Neural Profunda Detectron2, que é uma CNN de última geração proposta pelo *Facebook Artificial Intelligence Research* (WU et al., 2019).

3.2.1 Extração Profunda de Atributos

A CNN Extreme Inception (Xception) é um modelo CNN que contém a camada de convolução separável em profundidade, conforme proposto por Chollet (2017), no qual é apresentada a camada de convolução separável em profundidade. O modelo obteve bons resultados usando o conjunto de dados JFT (HINTON; VINYALS; DEAN, 2015), que contém 350 milhões de imagens. Embora o Xception tenha o mesmo número de parâmetros que o Inception V3, esta CNN usa os mesmos parâmetros do modelo com mais eficiência.

Visual Geometry Group (VGG) (Oxford University) (SZEGEDY et al., 2015) foi vice-campeão no desafio ILSVRC 2014 (RUSSAKOVSKY et al., 2015). A arquitetura da rede consiste em 11, 13, 16 ou 19 camadas convolucionais uniformes. O uso de convoluções fatoradas foi a estratégia utilizada para aumentar a profundidade, sem causar *overfitting* do modelo.

A MobileNet possui 28 camadas e é construída fundamentalmente, com exceção

da primeira camada, em convoluções separáveis por profundidade. Essas camadas foram introduzidas por (HOWARD et al., 2017a) em 2016. Os resultados Alcançaram 99.1% de acurácia de detecção do InceptionV3 (SZEGEDY et al., 2015) no banco de dados ImageNet com 87,7% menos de parâmetros e 88.6% menos operações (HOWARD et al., 2017b).

A Inception-ResNet-V2 (SZEGEDY et al., 2017) é uma CNN composta pela combinação das ideias de Conexões Residuais (WU; ZHONG; LIU, 2018) com o modelo Inception (SZEGEDY et al., 2015). Esse modelo de rede atingiu o erro de classificação top 1 de 16,5% e um top 5 de 3,1%, de forma a superar os resultados da ResNet e Inception para o mesmo problema proposto.

A Rede Convolutacional Densamente Conectada (DenseNet) amplia o número de conexões entre as camadas adjacentes. Essa CNN conecta cada respectiva camada a todas as outras com a mesma resolução de entrada, tanto em largura como em altura. Para esse conjunto, de camadas interconectadas, é denominado o termo Bloco Denso (HUANG et al., 2017).

A Rede Residual (ResNet) introduzida pela Microsoft Research, cuja profundidade é de até 152 camadas e possui o tamanho da imagem de entrada padrão de 224x224x3, obteve 92,1% de acurácia no conjunto de validação do ImageNet (HE et al., 2016).

3.2.2 Detectron2

A CNN Detectron2 vem do framework Mask R-CNN (HE et al., 2017), proposto pelo *Facebook Artificial Intelligence Research*. A estrutura é implementada em PyTorch e inclui implementações de alta qualidade dos algoritmos de detecção de objetos. A escolha deste framework para a etapa de detecção e segmentação se deve a sua modularidade, facilidade, rapidez para realizar o treinamento e ao alto desempenho no trabalho com imagens médicas.

Para realizar a detecção e segmentação de objetos, a Detectron2 requer, além do conjunto de imagens e seus respectivos padrões ouro, que o conjunto de dados seja especificado em uma lista de anotações. O formato das anotações para a Detectron2 segue o formato adotado pelo conjunto de dados COCO (LIN et al., 2014). As anotações devem conter todos os objetos individuais de todas as imagens no conjunto de dados. Em geral, essas anotações contêm uma lista de vértices de polígonos ao redor de cada objeto de interesse, sua respectiva caixa delimitadora, categoria e área. Assim, para realizar o ajuste fino da Detectron2 com os conjuntos de imagens médicas estudadas neste trabalho, foi criado um *script* para analisar e preparar as imagens para o formato padrão da Detectron2. O *script* recebe a categoria, uma imagem e seu respectivo padrão ouro e gera caixas delimitadoras e máscaras dos objetos presentes na imagem.

O modelo pré-treinado utilizado neste trabalho foi o COCO Instance Segmentation

Baseline with Mask R-CNN R50-FPN 3x, disponibilizado por Wu et al. (2019) em Model Zoo and Baselines, em que todo o modelo foi reaproveitado com todos os parâmetros já treinados, exceto a última camada. De acordo com o conceito de ajuste fino, assumiu-se que esses parâmetros continham algum conhecimento prévio aprendido do dataset inicial e esse conhecimento foi aproveitado no processo de treinamento com os conjuntos de imagens médicas. A última camada, a camada de classificação do modelo original, foi removida por ser diretamente relacionada com as classes do conjunto de dados iniciais, de forma que, para o problema atual, as categorias dessa camada não condiz com as novas classes. Dessa forma, uma nova camada de saída com tamanho igual ao número de classes atual foi criada com pesos inicializados aleatoriamente. Por fim, o treinamento da CNN Detectron2 foi realizado com os conjuntos de imagens médicas. A camada de saída foi treinada do zero (com os parâmetros iniciados aleatoriamente), enquanto que os parâmetros das camadas restantes foram ajustados (ajuste-fino) com base nos parâmetros do modelo inicial.

Após o ajuste fino, a rede neural recebe uma nova imagem e realiza a etapa de detecção generalizada, demarcando as possíveis regiões de interesse por meio das caixas delimitadoras. Essas caixas demarcam as regiões com maior número de *pixels* cujas características são quantitativamente semelhantes às características das regiões de interesse aprendidas durante a fase de treinamento. A etapa de detecção especializada consiste em escanear as regiões dentro das caixas delimitadoras *pixel* a *pixel* para que os *pixels* da imagem sejam classificados como pertencentes ou não à região de interesse. O resultado da classificação é uma máscara binária composta por *pixels* brancos cujas coordenadas coincidem com os *pixels* classificados como pertencentes à região de interesse (HE et al., 2017).

3.3 Métodos de Aprendizagem de Máquina

3.3.1 Naive Bayes

Naive Bayes é um conjunto de classificadores probabilísticos. Com base no teorema de Bayes, esses classificadores assumem independência condicional de classe entre os atributos. Em outras palavras, um classificador Naive Bayes presume que, para uma dada classe, a probabilidade de presença (ou ausência) de uma determinada característica não está relacionada à presença de qualquer outra característica (THEODORIDIS; KOUTROUMBAS, 2009).

O algoritmo de classificação Bayesiana supõe que as probabilidades $P(\omega_i)$ e as densidades de probabilidade condicionais $p(\mathbf{x}|\omega_i)$ para cada classe ω_i , $i \in 1, 2, \dots, c$, são conhecidas. A partir dessas informações, podemos utilizar o teorema de Bayes (equação 3.8) e calcular a probabilidade a posteriori $P(\omega_i|\mathbf{x})$ para cada classe dado um padrão $\mathbf{x}$,

ou seja,

$$P(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i)P(\omega_i)}{p(\mathbf{x})} \quad (3.1)$$

em que $P(\omega_i)$ é probabilidade a priori, $p(\mathbf{x}|\omega_i)$ é a função de densidade de probabilidade condicional, às vezes também referida como função de verossimilhança ou *likelihood* de ω_i com o respectivo $\mathbf{x}$, $p(\mathbf{x})$ é a evidência dada por $p(\mathbf{x}) = \sum_{i=1}^c p(\mathbf{x}|\omega_i)P(\omega_i)$ e $P(\omega_i|\mathbf{x})$ é a probabilidade a posteriori. Dessa forma, a classificação do padrão $\mathbf{x}$ com a classe ω_i é feita baseado na maior probabilidade a posteriori.

Uma das funções de densidade de probabilidade mais comumente encontrada na prática é a função de densidade de probabilidade Gaussiana ou normal. A principal razão para essa popularidade é o fato de que ela modela adequadamente um grande número de casos evidenciados por um dos teoremas mais célebres na estatística conhecido como o Teorema do limite central.

A função de densidade de probabilidade condicional Gaussiana univariada com média μ e variância σ^2 é dada pela equação 3.2:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in R \quad (3.2)$$

em que x é um escalar.

Em problemas reais, os padrões costumam estar em R^d . Dessa forma, a função de densidade de probabilidade condicional Gaussiana multivariada com média μ e covariância Σ é dada pela equação 3.3:

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)\right\}, \quad \mathbf{x} \in R^d \quad (3.3)$$

em que $\mathbf{x}$ é um vetor d dimensional, $\mu \in R^d$ é o vetor de médias, Σ é uma matriz $d \times d$, Σ^{-1} é a inversa de Σ e $|\Sigma|$ é o determinante de Σ .

Neste trabalho, foi utilizado o algoritmo Naive Bayes Gaussiano Multivariado como classificador probabilístico para avaliar o desempenho da metodologia.

3.3.2 k-Vizinhos Mais Próximos

Os k-Vizinhos Mais Próximos (KNN) é um algoritmo de aprendizado não paramétrico e baseado em instâncias (COVER; HART, 1967). O objetivo é descobrir os k vizinhos mais próximos de um padrão de entrada e, dessa forma, classificá-los com a classe dos padrões mais similares. O KNN classifica um novo padrão baseado no cálculo das distâncias entre este e cada padrão pertencente a uma base de treinamento dada. O rótulo

do novo padrão é atribuído de acordo com a classe que mais se repete entre os K vizinhos mais próximos (DUDA; HART; STORK, 2012).

As medidas de dissimilaridade mais comuns são as distâncias Euclidiana, Mahalanobis, Minkowski, Hamming e outras. A distância Euclidiana, por exemplo, pode ser descrita como a Equação 3.4:

$$d(\mathbf{X}, \mathbf{X}^i) = \sqrt{(x_1 - x_1^i)^2 + (x_2 - x_2^i)^2 + \dots + (x_p - x_p^i)^2} \quad (3.4)$$

onde $\mathbf{X}$ e $\mathbf{X}^i$ são dois vetores em um espaço dimensional p e d é uma função de medida de dissimilaridade. Neste trabalho foi utilizado o KNN com a distância Euclidiana.

3.3.3 Máquinas de Vetores de Suporte

As Máquinas de Vetores de Suporte (SVM) (VAPNIK; VAPNIK, 1998) é um classificador baseado em um hiperplano, margem máxima e margem suave. A ideia principal do classificador SVM é transformar o espaço dos atributos de entrada de uma forma específica para que um limite de decisão linear (neste novo espaço) possa separar corretamente as amostras.

A principal desvantagem do classificador de hiperplano é que, se um conjunto de amostras for linearmente separável (perfeitamente separado usando um hiperplano), um número infinito de hiperplanos podem existir.

Para selecionar os melhores hiperplanos dentre esses o classificador de margem máxima foi proposto. Esse classificador consiste em selecionar o hiperplano mais distante de todas amostras do conjunto de treinamento de forma que este hiperplano é a solução ideal para o problema de classificação. A expressão “mais distante” significa o somatório com maior valor de distâncias perpendiculares para cada amostra do conjunto de treinamento em relação a um determinado hiperplano. As amostras com menor distância perpendicular ao hiperplano são conhecidas como margem. Contudo, o classificador de margem máxima é limitado pela sua sensibilidade à amostras individuais.

Dessa forma, o classificador de margem suave adiciona uma variável de folga denominada ϵ para cada padrão do conjunto de treinamento. Essa variável permite que amostras sejam classificadas erroneamente de forma a aumentar a margem do classificador.

Apesar disso, uma das principais desvantagens do classificador de margem suave é que sua superfície de decisão é sempre linear. Para resolver esse problema e estender o classificador a problemas não-lineares, o SVM foi proposto. A principal ideia do classificador SVM é transformar o espaço dos atributos das amostras de entrada para que um limite de decisão linear (no novo espaço) possa separar as amostras corretamente.

3.3.4 Floresta Aleatória

A Floresta Aleatória (RF) é uma rede não supervisionada com base no método de conjunto de dados escolhido com base no conjunto inicial, criando muitas árvores de decisão. Em seu treinamento, o algoritmo melhora a classificação de cada árvore de decisão calculando a margem das classificações corretas para as classificações incorretas (LIAW; WIENER et al., 2002).

3.3.5 Perceptron Multicamadas

O Perceptron Multicamadas (MLP) (HAYKIN, 2010) é um método de aprendizado de máquina supervisionado baseado no classificador Perceptron introduzido por Rosemblatt. O Perceptron calcula uma soma ponderada das entradas e aplica o resultado em uma função de ativação. Cada peso indica a relevância de um determinado recurso para prever uma determinada classe.

A Figura 1 mostra a arquitetura de uma MLP com uma camada escondida. A camada de entrada é composta apenas de unidades receptoras dos sinais externos, não havendo cálculos nela. A camada intermediária é composta por neurônios somadores com função de ativação não-linear, sendo que a camada de saída também é composta por neurônios somadores com funções de ativação que podem ou não ser lineares.

A camada de entrada possui P neurônios mais um bias, em que P é a dimensão do vetor de entrada. O número de neurônios Q da camada oculta é um hiperparâmetro e deve ser dado ao algoritmo. A camada de saída possui C neurônios, onde C é o número de classes do problema.

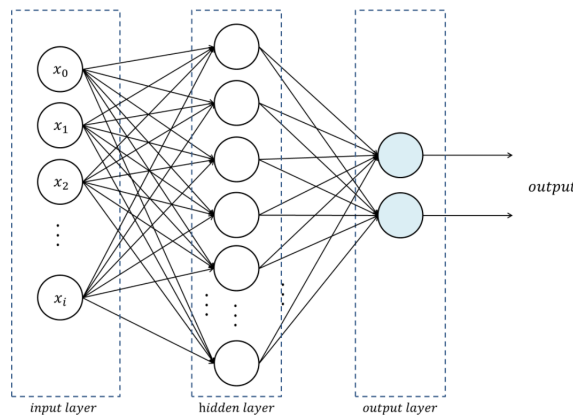


Figura 1 – MLP com uma camada escondida.

3.4 Técnicas de PDI como Ajuste Fino para a Segmentação de Objetos

As técnicas de ajuste fino utilizadas neste artigo são baseadas em PDI, onde as técnicas recebem uma região inicial ou semente e, partir desta segmentação inicial, age nas bordas da segmentação pela combinação das técnicas Crescimento de Região (ADAMS; BISCHOF, 1994), agrupamento K-means (DHANACHANDRA; MANGLEM; CHANU, 2015) e Janela de Parzen (REBOUÇAS et al., 2019), a fim de melhorar os contornos da região considerada de interesse.

3.4.1 Crescimento de Regiões

O Crescimento de Região é um procedimento que agrupa os *pixels* ou sub-regiões com base em regras predefinidas para o crescimento. A abordagem básica é começar com um conjunto de pontos semente e, a partir deles, fazer as regiões crescerem anexando a cada semente aqueles *pixels* vizinhos que têm propriedades predefinidas semelhantes às das sementes (GONZALEZ; WOODS, 2009).

O algoritmo básico para segmentação baseado em crescimento da região pode ser escrito da seguinte maneira: 1- Iniciar um conjunto de pontos sementes (manual ou automaticamente definidos). 2- Agrupar *pixels* candidatos aos *pixels* sementes por meio de um critério de similaridade ou propriedade comum. Um *pixel* é adicionado a uma região se ele não tiver sido designado para outra região, ou ele for vizinho desta região (i.e., conectado a um *pixel* dela) ou a nova região criada pela adição do *pixel* continuar conexa. 3- Definir um critério de parada. Um critério de parada pode ser considerado quando o perímetro da região parar evoluir ou atingir um número de iterações.

3.4.2 K-means

O algoritmo *K-means* é um método para agrupar automaticamente dados semelhantes. No contexto de imagens, recebemos uma imagem $M \times N \times C$ e desejamos agrupar os *pixels* em alguns grupos coesos baseados em níveis de intensidade. A intuição por trás do *K-means* é um procedimento iterativo que começa supondo os centróides iniciais e refina esse palpite atribuindo repetidamente exemplos para seus centróides mais próximos e, em seguida, recalculando os centróides com base nas atribuições.

Um pseudocódigo (Algoritmo 1) mostra os principais passos do algoritmo *K-means* para segmentação de imagens:

O laço do algoritmo executa repetidamente duas etapas: (i) Atribuir cada *pixel* (x, y) ao centróide mais próximo e (ii) re-computar a média de cada centróide usando os

Algoritmo 1 K-means(imagem, centroidesIniciais, maxIters)

```

1: K = len(centroidesIniciais)
2: centroides = centroidesIniciais
3: para i in range(maxIters)
4:     idx = encontraCentroidesMaisProximos(imagem, centroides)
5:     centroides = computaCentroides(imagem, idx, K)
6: fim para
7: retorna centroides

```

Figura 2 – Pseudocódigo do algoritmo *K-means* para a segmentação de imagens.

pixels atribuídos a ele. O algoritmo *K-means* sempre convergirá para algum conjunto final de médias para os centróides.

O método “encontraCentroidesMaisProximos” na linha 4 atribui todos os *pixels* ao centróide mais próximo, dadas as coordenadas (x, y) atuais dos centróides. Especificamente, para cada *pixel*, estabelecemos:

$$idx^{(i)} := j \text{ que minimiza } ||x^{(i)} - \mu_j||^2 \quad (3.5)$$

em que $x^{(i)}$ é o i -ésimo *pixel* da imagem, μ_j é valor do centróide j (média dos níveis de intensidade dos *pixels* atribuído ao centróide) e $idx^{(i)}$ é o índice do j -ésimo centróide mais próximo do *pixel* $x^{(i)}$.

Dadas as atribuições de cada *pixel* a um centróide, o método “computaCentroides” recalcula, para cada centróide, a média dos *pixels* que foram atribuídos a ele (linha 5). Especificamente, para cada centróide j temos:

$$\mu_j := \frac{1}{|C_j|} \sum_{i \in C_j} x^{(i)} \quad (3.6)$$

em que C_j é o conjunto de *pixels* atribuídos ao centróide j .

3.4.3 Janela de Parzen

A Janela de Parzen é uma técnica avançada para ajustar as bordas da região de interesse com base na densidade de probabilidade baseado em um Elemento Estruturante Gaussiano (REBOUÇAS et al., 2019). As bordas da região de interesse passam por um processo de expansão ou contração de acordo com a probabilidade dos *pixels* vizinhos pertencerem ou não a ela como parte da região a ser segmentada.

As Janelas de Parzen não supõe uma distribuição a priori. Estas técnicas são variações da aproximação por histograma de uma função de densidade de probabilidade

desconhecida (THEODORIDIS; KOUTROUMBAS, 2009). Uma importante vantagem dessa abordagem é que o modelo de densidade pode estimar com maior precisão a distribuição dos dados. Por exemplo, se a distribuição dos dados é multimodal, então este aspecto da distribuição nunca poderá ser capturado por uma Gaussiana, que é necessariamente unimodal.

A Janela de Parzen consiste num histograma contínuo, cujas janelas são centralizadas em cada um dos pontos de dados de onde se quer estimar a densidade, em que um hiperparâmetro h auxilia na definição da janela.

Caso se utilize como janela a função de densidade Gaussiana, temos a verossimilhança na Equação 3.7:

$$p(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi)^{d/2} h^d} \exp\left(-\frac{(\mathbf{x} - \mathbf{x}_i)^T (\mathbf{x} - \mathbf{x}_i)}{2h^2}\right) \quad (3.7)$$

em que N é o número de *pixels*, d é a dimensão dos dados e h representa o desvio padrão das componentes Gaussianas e auxilia na definição da janela. Assim, a densidade de probabilidade condicional $p(\mathbf{x})$ do modelo é obtida colocando uma Gaussiana em cada ponto dos dados somando as contribuições em todo o conjunto de dados e, em seguida, dividindo por N para que a densidade seja corretamente normalizada.

A fase de treinamento trata-se de armazenar as probabilidades a priori de cada classe. Após isso, o algoritmo pode ser empregado para realizar a classificação de novos *pixels*. A classificação de um padrão $\mathbf{x}$ com a classe ω_i é feita baseado na maior probabilidade a posteriori computada pela regra de Bayes (Equação 3.8):

$$P(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i)P(\omega_i)}{p(\mathbf{x})} \quad (3.8)$$

em que $p(\mathbf{x}|\omega_i)$ é a função de densidade de probabilidade condicional da classe ω_i dada pelo cômputo da Equação 3.7 e $P(\omega_i)$ é a probabilidade a priori da classe ω_i .

3.5 Métricas de Avaliação

As métricas de avaliação são ferramentas importantes para analisar a eficácia da segmentação. Dentre os principais itens a serem avaliados, a semelhança do resultado com o do médico especialista e o número de acertos em relação ao total (TAHA; HANBURY, 2015) devem ser avaliados. As métricas de avaliação adotadas neste trabalho foram baseadas na matriz de confusão, fornecendo os índices Verdadeiro Positivo (TP), Verdadeiro Negativo (TN), Falso Positivo (FP) e Falso Negativo (FN).

Acurácia: é uma métrica de classificação e está diretamente relacionada ao número de *pixels* que foram segmentados corretamente sobre o número total de *pixels* segmentados em uma imagem (ALJABAR et al., 2009). Sua fórmula é dada pela Equação 3.9:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.9)$$

Coeficiente de Dice: é uma métrica de sobreposição que analisa a similaridade entre a região segmentada pelo algoritmo e a região segmentada pelo médico especialista, sendo uma boa segmentação a que mais se aproxima da segmentação do especialista (MILLETARI; NAVAB; AHMADI, 2016). Sua fórmula é dada pela Equação 3.10:

$$DICE = \frac{2TP}{2TP + FP + FN} \quad (3.10)$$

Sensibilidade: é uma métrica de classificação responsável por avaliar o número de *pixels* corretamente segmentados como pertencentes à região de interesse entre todos os *pixels* realmente pertencentes a esse grupo (BHADAURIA; DEWAL, 2014). Sua fórmula é dada pela Equação 3.11:

$$SEN = \frac{TP}{TP + FN} \quad (3.11)$$

Especificidade: é uma métrica de classificação responsável por indicar o número de *pixels* corretamente segmentados como não-pertencentes à região de interesse, ou *background*, entre todos os *pixels* realmente pertencentes ao background (BHADAURIA; DEWAL, 2014). Sua fórmula é dada pela Equação 3.12:

$$ESP = \frac{TN}{TN + FP} \quad (3.12)$$

3.6 Teste de Kolmogorov-Smirnov

Os testes estatísticos são ferramentas importantes para comparar valores métricos, visto que em algumas situações os valores podem diferir visualmente, mas do ponto de vista estatístico são equivalentes. Usando testes estatísticos, duas quantidades de amostras podem ser comparadas como sendo maiores, menores ou iguais, mas esta conclusão só pode ser tirada após o teste estatístico apropriado ter sido realizado (PATERNOSTER et al., 1998).

O Teste Estatístico de Kolmogorov-Smirnov é um teste paramétrico cuja funcionalidade é testar a igualdade entre duas distribuições de amostras contínuas do mesmo tamanho. As hipóteses deste teste são: H_0 - as duas amostras são estatisticamente idênticas,

H1 - as duas amostras são diferentes, uma das quais pode ser maior ou menor que a outra (PRÆSTGAARD, 1995).

O coeficiente de confiabilidade (α) é um fator determinante na decisão de qual hipótese estará em vigor sobre os dados. O produto final será um coeficiente denominado P-Value. Este valor P será comparado com o valor α . Esta comparação define a hipótese estatística atual de acordo com a Equação 3.13 (CORTINA, 1993).

$$K \quad \text{Teste} \begin{cases} P > \alpha & = & H_0 & - & \text{Igualdade Estatística} \\ P < \alpha & = & H_1 & - & \text{Diferença Estatística} \end{cases} \quad (3.13)$$

O valor do coeficiente de confiabilidade (α) comumente adotado é 5%, e para testes mais rigorosos, um valor de 3% (CORTINA, 1993). Para realizar o teste com os valores da métrica, adotamos α com valor de 3%.

4 Metodologia

Neste capítulo, apresentamos a metodologia que consiste em três etapas: Passo 1 - Classificação das imagens de entrada com a presença ou ausência de RoI. Passo 2 - Início do processo de segmentação das imagens que possuem RoI (Pulmão ou AVC hemorrágico). Passo 3: Uso de ajuste fino para aprimorar a segmentação e o resultado final.

4.1 Saúde das Coisas

Os sistemas baseados em Saúde das Coisas estão sendo aplicados na área da saúde para conectar sistemas de visão computacional com o usuário final, especialistas em saúde e até mesmo pacientes que podem usar a aplicação. Nesse estudo será proposto um protótipo denominado *Health-of-Things System for the Classification and Segmentation of Medical Images* que foi desenvolvido tendo como objetivo a comunicação através do protocolo Representational State Transfer (REST), via JavaScript Object Notation (JSON). Desta forma, aplicações com acesso à Internet podem se comunicar com o sistema, pois a *Application Programming Interface* (API) foi desenvolvida de forma a interagir através da IoT.

A parte central da aplicação, que é responsável pelo processamento de visão computacional foi implementada em Python utilizando principalmente as bibliotecas *Scikit-Learn* e *Tensorflow*. O acesso à API foi desenvolvido utilizando as linguagens Python versão 3.6, Java versão 8.0 e PHP versão 7.0.0.

4.2 Classificação

A etapa 1 corresponde à detecção de dois problemas de forma binária sendo um relacionado ao pulmão (presença ou ausência de pulmão na imagem) e o outro ao cérebro (presença ou ausência de AVC hemorrágico na imagem).

Para o desenvolvimento da ferramenta, que corresponde ao passo 1, foi utilizada a ferramenta de visão computacional Lapisco Image Interface for Application Development (LINDA) (JR et al., 2019). A etapa 1 cobre o início do modelo, conforme mostrado na Figura 3. LINDA é um aplicação hospedada na nuvem e qualquer dispositivo com acesso à internet pode utilizá-la.

Para encontrar a melhor combinação extrator-classificador seguimos o fluxo indicado na Figura 4. Inicialmente as imagens estavam no formato *Digital Imaging and Communications in Medicine* (DICOM), que é o padrão para imagens médicas. Todas as

HTSCS - Medical Images:
Health-of-Things System for the Classification and Segmentation of Medical Images

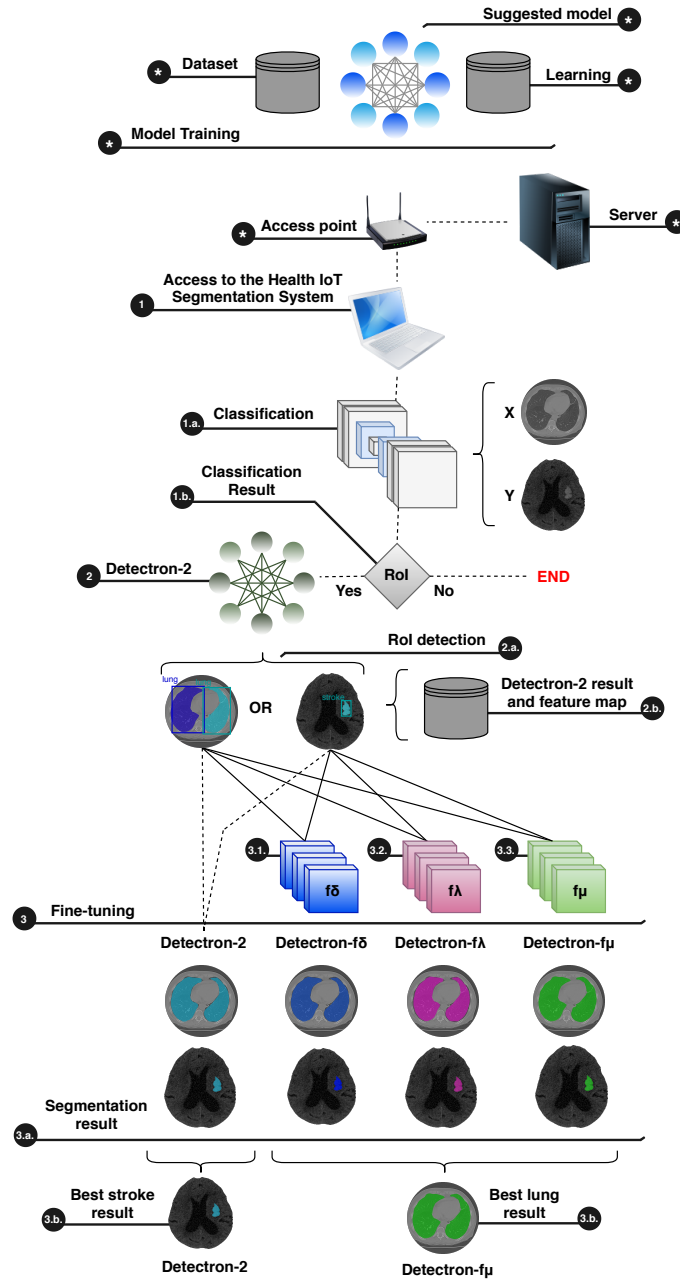


Figura 3 – Apresentação do *HTSCS - Medical Images: Health-of-Things System for the Classification and Segmentation of Medical Images*. A figura ilustra as etapas do método baseado na Saúde das Coisas. Na imagem, o símbolo * representa a etapa inicial de treinamento do método para classificação e segmentação de imagens tomográficas de pulmão e AVC hemorrágico. A etapa 1 consiste na classificação da imagem de entrada (com a detecção ou não de pulmão ou AVC hemorrágico). Na etapa 2 a rede Detectron2 inicia o processo de segmentação do pulmão ou AVC hemorrágico gerando mapas característicos através das caixas delimitadoras. A etapa 3 continua do processo de segmentação usando o ajuste fino e, finalmente, apresenta o melhor resultado.

imagens DICOM foram convertidas para *Portable Network Graphics* (PNG). A partir do momento que as imagens são carregadas, a extração de dados é iniciada e, posteriormente, classificada. Por fim, a melhor combinação extrator-classificador é analisada usando gráficos e matrizes de confusão.

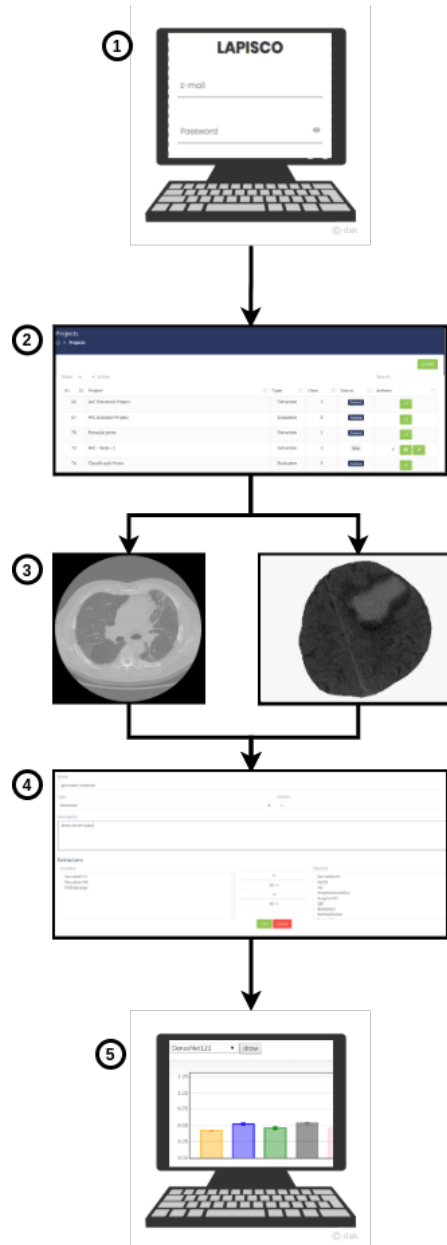


Figura 4 – LINDA framework. 1- Tela de login e senha. 2- Criação do projeto, que inclui definir nome, número de classes e que ação (extração de atributos, classificação ou predição do modelo treinado) o usuário deseja realizar. 3- Caso o usuário decida treinar um novo modelo, as imagens separadas por classes serão carregadas. 4- Ao fazer o upload das imagens, o usuário escolhe quais descritores e classificadores serão utilizados no treinamento. 5- Análise dos resultados após o treinamento com gráficos de barras e matrizes de confusão.

O primeiro passo no pré-processamento da imagem é normalizar a imagem para ajustar a profundidade de cor das imagens, principalmente nas imagens do tipo DICOM.

Uma vez que a melhor combinação tenha sido definida podemos encapsular o modelo treinado para prever novos dados de imagem na nuvem e prosseguir com o modelo de Saúde das Coisas conforme mostrado na Figura 3.

Na Figura 3, Passo 1, ilustra quando o usuário carrega a imagem para ser classificada e segmentada após treinar o modelo. No item 1.a a imagem é classificada em ausência ou presença de RoI. Caso a imagem não contenha RoI, o processamento computacional é encerrado para aquela imagem, conforme mostrado no item 1.b. Se o objeto de interesse for detectado pelo modelo, o método avança para a Etapa 2 (segmentação da RoI) e, no caso deste estudo, segue para o processo de segmentação do pulmão ou AVC hemorrágico.

4.3 Segmentação

O Passo 2 da Figura 3, que corresponde ao primeiro estágio da segmentação, é representado pela CNN Detectron2 conforme explicado na Seção 3.2.2. A rede Detectron2 identifica a região real pertencente às caixas delimitadoras do objeto de interesse (Figura 3 item 2.a) e gera um mapa de características do pulmão ou AVC hemorrágico, conforme mostrado na Figura 3 no item 2.b.

No item 2.a, a rede detecta a região ou regiões de interesse de forma generalizada por meio de caixas delimitadoras. Essas caixas detectam as regiões que são mais semelhantes à região de interesse aprendida no estágio de treinamento. No conjunto de dados do pulmão as caixas delimitadoras se formam ao redor dos pulmões e no conjunto de dados de AVC hemorrágico as caixas delimitadoras se formam ao redor da região mais clara da região do cérebro considerando que esta é possivelmente a região afetada pelo vazamento. O resultado do processo de detecção nesta etapa através da rede Detectron pode ser visto na Figura 5, que apresenta resultados diferentes para as imagens de TC do pulmão e as imagens de TC de AVC hemorrágico.

A Figura 5 mostra a instância de segmentação da Detectron2 em diferentes imagens. A rede detecta e demarca cada objeto de interesse distinto que aparece em uma imagem.

Nesse estágio de segmentação, os *pixels* da região demarcada pela caixa delimitadora são classificados de acordo com seus atributos e a rede realiza a construção da região de interesse. Esses *pixels* classificados são usados para criar uma máscara binária, que é o resultado da segmentação, representada no item 2.b da Figura 3. Posteriormente, ela será utilizada como uma segmentação inicial da região de interesse no estágio 3 da figura.

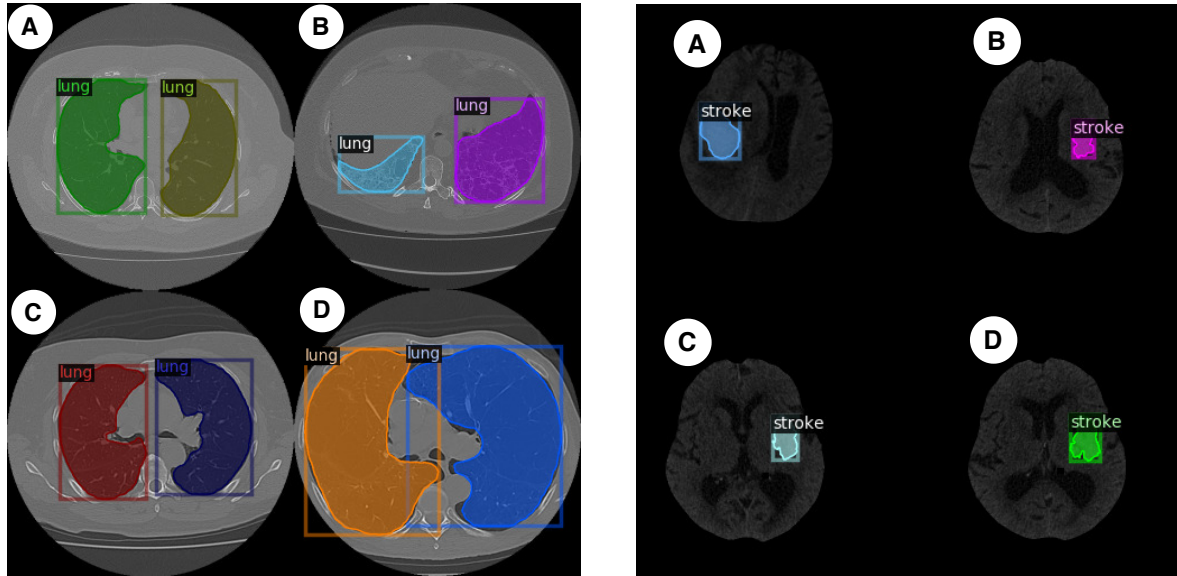


Figura 5 – Resultado da instância de segmentação da rede Detectron2 em imagens de TC do tórax e imagens de TC do crânio, respectivamente. As instâncias detectadas foram sobrepostas com várias cores.

4.4 Ajuste Fino

O ajuste fino é a última etapa do sistema HTSCS sendo utilizado como um ajuste de segmentação a fim de melhorar a eficiência da segmentação da região de interesse.

Três técnicas de ajuste fino foram propostas juntamente com a rede Detectron2 conforme mostrado no item 3 da Figura 3. As técnicas são baseadas em PDI que são Crescimento de Região, Agrupamento K-Means e Janela de Parzen. A técnica Detectron2 + Janela de Parzen (Detectron- $f\lambda$) consiste na aplicação direta da técnica de Janela de Parzen no resultado gerado pela rede a fim de melhorar o contornos da região segmentados suavemente pelo elemento estruturante gaussiano. Os cálculos de densidade de probabilidade realizados para contrair ou expandir as bordas da região são obtidos a partir dos *pixels* da imagem do exame de TC original.

A segunda técnica, Detectron2 + Janela de Parzen + Agrupamento K-means (Detectron- $f\delta$) é a aplicação direta da técnica da Janela de Parzen sobre a região segmentada. Desta forma, o agrupamento que atua normalizando a imagem visa melhorar os limites da região de interesse em relação ao fundo da imagem, o que torna a Janela de Parzen mais eficaz.

A terceira técnica, Detectron2 + Janela de Parzen + Crescimento da região + Agrupamento K-Means (Detectron- $f\mu$) consiste na reconstrução da região de interesse inicialmente pela técnica de crescimento de região na imagem agrupada com um $K = 3$ do exame de TC em que o próprio resultado da rede serve como uma semente para o crescimento, a fim de ajustar as bordas.

No item 3.a da Figura 3 os resultados da segmentação são apresentados com os

diferentes modelos baseados no ajuste fino incluindo os resultados diretos (sem o processo de ajuste fino). Nos itens 3.b e 3.c são apresentados apenas os resultados dos modelos que obtiveram o melhor desempenho na segmentação. Ambos os modelos funcionam em paralelo para cada imagem de entrada no sistema proposto.

5 Resultados

Este capítulo apresenta os resultados e discussões do método proposto, de acordo com a Metodologia apresentada no Capítulo 4.

Os resultados estão em três etapas. Na primeira etapa será apresentado os experimentos e resultados do estágio de classificação das imagens. A classificação é feita utilizando o método SVM e extratores baseados em deep learning para detectar a presença do pulmão ou AVC hemorrágico nas imagens de TC. A segunda etapa é subdividida em dois experimentos: segmentação pulmonar e segmentação de acidente vascular cerebral hemorrágico. No primeiro experimento temos os resultados da segmentação pulmonar em imagens de TC. No segundo experimento temos os resultados da segmentação do AVC hemorrágico. Por fim, a Terceira Etapa valida o método proposto neste estudo. Os melhores resultados alcançados são apresentados e comparados com os resultados dos métodos relatados na literatura.

5.1 Primeira Etapa do Experimento - Classificação

Todos os conjuntos de dados usados foram pré-processados da mesma forma. Foram aplicadas dez interações nas quais os padrões foram divididos arbitrariamente em dois grupos, 80% para treinamento e o restante para teste. Os conjuntos de treinamento e teste foram normalizados (média de zero e variação unitária) utilizando as mesmas regras de normalização. A fim de encontrar o melhor conjunto de hiperparâmetros para o classificador foi aplicada a técnica de busca em grade com validação cruzada com *k-fold*. Os hiperparâmetros que alcançaram a maior acurácia no conjunto de validação foram armazenados e os valores mais repetidos ao final das dez iterações foram escolhidos como os melhores hiperparâmetros.

A técnica de busca em grade com validação cruzada com 10-*folds* foi adotada para escolher os melhores hiperparâmetros C e γ do classificador SVM com kernel RBF. O intervalo de C e γ variou entre $[2^{-5}, 2^{15}]$ e $[2^{-15}, 2^3]$, respectivamente.

As Tabelas 1 e 2 apresentam os resultados detalhados das duas melhores combinações extratores-classificadores para a Etapa 1 do modelo. Várias combinações extratores-classificadores foram testadas para encontrar a combinação que resolve o problema proposto. Porém, para a classificação de presença ou ausência de pulmão, dois extratores de atributos foram destacados: Xception e VGG 16 com o classificador SVM configurado com o kernel RBF. O extrator Xception teve uma acurácia de 96,63% com um tempo médio de 14,7ms enquanto que o extrator VGG16 teve 97,04% de acurácia com um tempo médio de 9,44ms.

O conjunto de dados foi organizado da seguinte forma: classe 0 sem pulmão e classe 1 com pulmão. Para detectar a presença ou ausência de AVC, os extratores Xception e VGG16 foram aplicados com o classificador SVM-RBF. Os dois extratores combinados com o SVM-RBF alcançaram acurácia de 100% apenas com diferenças no tempo médio de extração por imagem com 16,10 ms para Xception e 10,88 ms para VGG16.

Tabela 1 – Tabela com os dois melhores resultados da combinação extrator-classificador para a etapa de detecção, quando há ou não pulmão na imagem. Como pode ser visto, a combinação do extrator VGG16 e do extrator Xception, ambos com o classificador SVM-RBF, obteve 97% em todas as métricas de avaliação.

Classificador	Extrator	ACC	F1	SEN	ESP	Tempo de Extração
SVM-RBF	Xception	96.63±1.48	96.62±1.58	96.63±1.48	96.79±1.53	14.70 ms
SVM-RBF	VGG16	97.04±1.23	97.11±1.19	97.04±1.23	97.35±1.09	0.44 ms

Tabela 2 – Tabela com os dois melhores resultados da combinação extrator-classificador para detecção de presença ou não de AVC hemorrágico em imagens de crânio. Como pode ser visto, a melhor combinação é do extrator VGG16 com o classificador SVM-RBF, que pontuou 100% em todas as métricas de avaliação com o menor tempo de extração.

Classificador	Extrator	ACC	F1	SEN	ESP	Tempo de Extração
SVM-RBF	Xception	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	16.10 ms
SVM-RBF	VGG16	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	10.88 ms

5.2 Segunda Etapa do Experimento - Segmentação

Todos os experimentos de segmentação foram conduzidos em um sistema operacional Ubuntu 18.04 com 16 GB de RAM, processador Intel Core i7 e GPU NVIDIA GeForce GTX 1660 TI. Esses recursos foram utilizados para o treinamento e inferência da rede neural. Inicialmente, o modelo utilizado foi o Mask-RCNN R-50-FPN-3x, como pode ser visto na visão geral destes modelos ¹. Como havia pouco mais de 1000 imagens de tomografia computadorizada de pulmão e 80 de AVC hemorrágico para treinamento, optamos por iniciar o treinamento com pesos pré-treinados. As redes neurais foram treinadas por 2.000 épocas com taxa de aprendizagem de 0,00025. O tempo total de treinamento para as imagens de AVC hemorrágico foi de aproximadamente 11 minutos, enquanto nas imagens de pulmão foi de 12 minutos.

5.2.1 Segunda Etapa do Experimento - Segmentação do Pulmão

Os resultados gerados nesta seção são baseados no conjunto de dados apresentado na Seção 3.1, que representa o mesmo conjunto de dados utilizado nos trabalhos de Hu et al. (2020) e Filho et al. (2019).

¹ <https://github.com/facebookresearch/detectron2/blob/master/MODEL_ZOO.md>

A Tabela 3 apresenta os resultados gerados por este estudo para segmentar o pulmão de imagens de tomografia. A primeira coluna apresenta os modelos Detectron2, Detectron- $f\lambda$, Detectron- $f\delta$ e Detectron- $f\mu$.

Tabela 3 – Resultados gerados pelo método proposto usando aprendizagem profunda e ajuste fino para segmentar o pulmão em imagens de TC. As mesmas imagens foram utilizadas nos trabalhos de Filho et al. (2019) e Hu et al. (2020).

MODELOS	ACC	DICE	SEN	ESP
Detectron2	99.00 ± 0.60	96.98 ± 1.21	96.74 ± 2.69	99.41 ± 0.50
Detectron- $f\lambda$	98.95 ± 0.61	96.74 ± 1.44	96.52 ± 3.16	99.29 ± 0.76
Detectron- $f\delta$	99.02 ± 0.59	97.03 ± 1.24	96.54 ± 2.73	99.45 ± 0.47
Detectron- $f\mu$	99.02 ± 0.60	97.01 ± 1.25	96.74 ± 2.78	99.45 ± 0.49

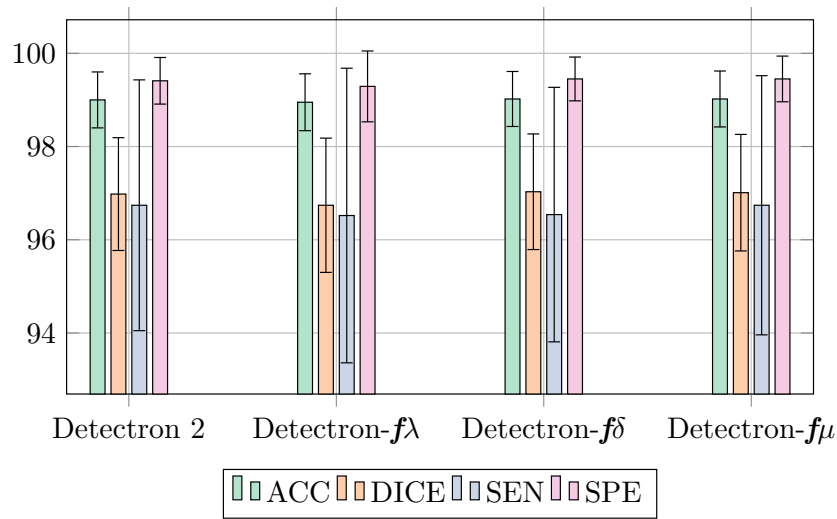


Figura 6 – Ilustração gráfica da Tabela 3. Diferentes modelos propostos com base na Saúde das Coisas com o uso de aprendizagem profunda aliado à técnica de ajuste fino.

O Modelo Detectron- $f\lambda$ é a combinação da Detectron2 com a técnica da Janela de Parzen para o ajuste fino da segmentação.

O Modelo Detectron- $f\delta$ é a combinação da Detectron2 com as técnicas de Janela de Parzen e Agrupamento.

O Modelo Detectron- $f\mu$ é baseado na combinação da Detectron2 com a Janela de Parzen, Crescimento de Região e Agrupamento.

O método foi treinado com 80% das 1.265 imagens de TC de pulmão sendo que todas as imagens continham o padrão. Isso valida ainda mais o método proposto considerando que o modelo foi treinado com base nas anotações de especialistas.

De acordo com a coluna ACC, que representa a métrica de Acurácia, pode-se ver o sucesso da rede Detectron2 em localizar e segmentar a região pulmonar de forma eficaz e

precisa. A rede Detectron2 obteve um excelente resultado com $99,00 \pm 0,60$. Com exceção da rede Detectron- $f\lambda$, os resultados obtiveram uma ligeira melhora a partir da aplicação do ajuste fino com variação mínima de 0,02 permitindo que ambos atingissem 99,02% de acurácia. Isso se deve à eficiência da rede Detectron2 em detectar verdadeiramente a localização do pulmão na imagem e segmentar o contorno que pertence ao objeto. Essa ligeira melhora com variações positivas refletiu nos dois modelos: Detectron- $f\delta$ e Detectron- $f\mu$ que conseguiram melhorar o ajustes de imagem com mais precisão. Isso porque ambos os métodos utilizaram técnicas de Agrupamento, que ajudaram a identificar com maior profundidade a relação de variações entre os *pixels* pertencentes ao pulmão na imagem da TC.

A coluna de DICE, que representa o Coeficiente de Dice, apresentou algumas diferenças positivas em comparação ao modelo Detectron2 (Modelo sem uso de ajuste fino), exceto a Detectron- $f\lambda$, que apresentou valores inferiores à Detectron2. A Detectron- $f\lambda$ não foi capaz de superar a Detectron2, considerando que o uso somente da Janela de Parzen como um ajuste fino teve dificuldade em encontrar as bordas dos ângulos mais fechados, como mostrado na Figura 7. Os modelos Detectron- $f\delta$ e Detectron- $f\mu$ tiveram resultados ligeiramente melhores que a Detectron2, mas se comportaram de forma semelhante. Com a avaliação de DICE que analisa o desempenho calculado da sobreposição os resultados foram relativamente semelhantes, ambos com 97,03% e 97,01% com um desvio padrão aproximado de 1,25%, contra $96,98\% \pm 1,21$ da rede Detectron2. Tais resultados mostram a eficiência dos modelos em fazer pequenos ajustes na segmentação relacionada à borda do pulmão.

Em relação à métrica de Sensibilidade (SEN) da Tabela 3, podemos encontrar similaridade entre os modelos Detectron2 e Detectron- $f\mu$, com valores relativamente equivalentes. Em outras palavras, ambas as redes sem ajuste fino e com ajuste fino foram capazes de prever corretamente os *pixels* pertencentes à área do pulmão e classificá-los como uma região pulmonar. Embora a diferença seja pequena entre os modelos Detectron- $f\lambda$ e Detectron- $f\delta$, esta variação demonstra a eficácia do método para diferentes tipos de pulmões em uma imagem de TC. Lembrando que a variação do tamanho do pulmão muda a cada imagem adquirida pelo tomógrafo, visto que o processo respiratório responsável por inflar e desinsuflar os pulmões produz diferentes tamanhos de pulmões na tomografia, tornando essa variação de sensibilidade um tanto complexa. Sem falar na variedade de diferentes patologias registradas pela TC, que aparecem como pontos específicos que podem ser caracterizados como anormalidades em uma TC, conforme mostrado na Figura 7.

A métrica de Especificidade na Tabela 3 obteve um excelente desempenho acima de 99% com variação máxima de 0,76% para todos os modelos. Os modelos obtiveram um excelente desempenho na classificação da região não pulmonar. Isso significa que os

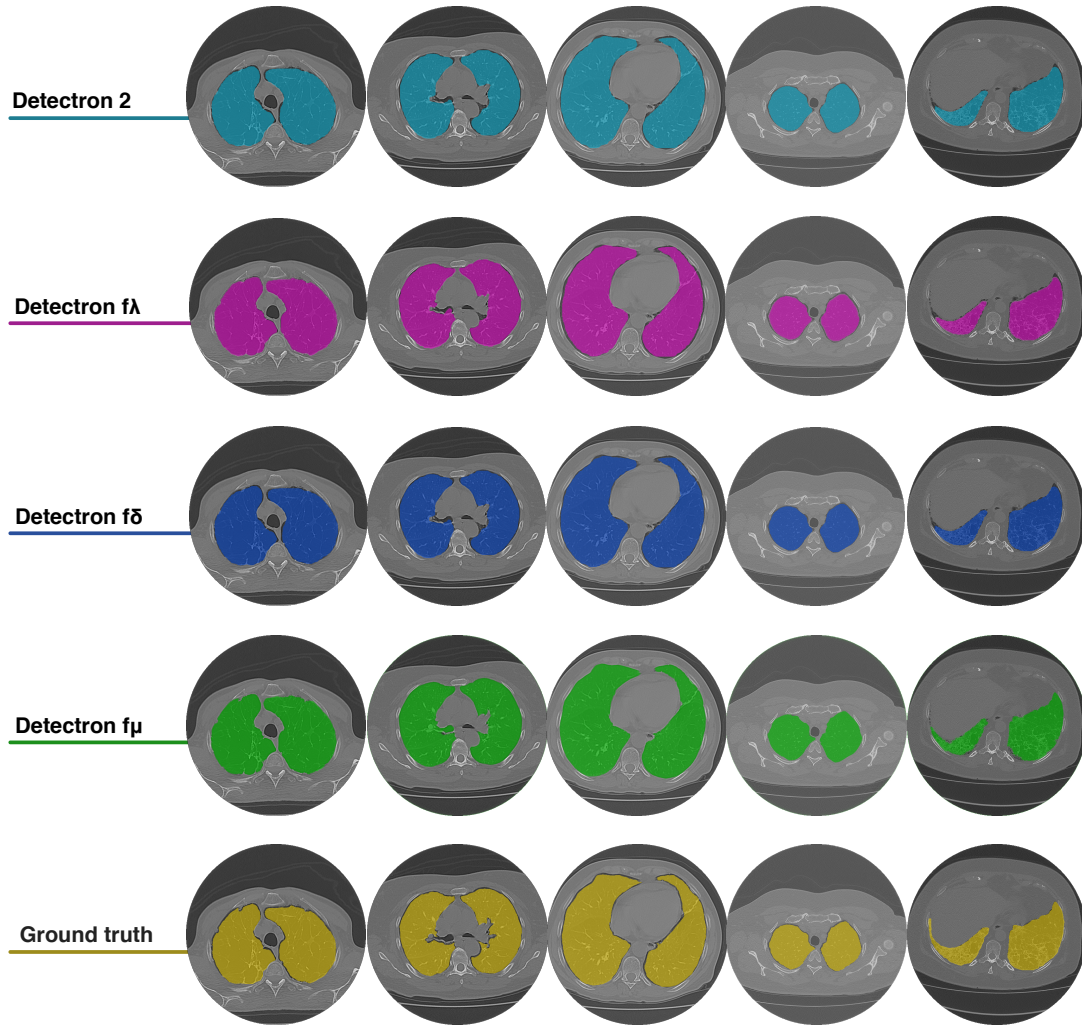


Figura 7 – Resultado da segmentação com os diferentes modelos baseados na rede Detectron2 combinados com métodos de ajuste fino.

modelos são capazes de apontar áreas que não fazem parte da RoI. No gráfico da Figura 6, que representa a Tabela 3, podemos analisar visualmente as semelhanças entre os modelos.

A Figura 7 apresenta os resultados da segmentação de cada modelo em diferentes formatos de pulmão. Houve uma ligeira melhora feita pelos modelos com o uso do ajuste fino na busca de contornar o objeto de interesse a partir do resultado inicial de segmentação da Detectron2.

A Tabela 4 refere-se ao tempo de segmentação e também destaca o sucesso do método com base nos resultados obtidos de forma rápida e precisa na segmentação do pulmão. A estimativa média de 1,7 segundos para os modelos demonstra um excelente desempenho e torna o modelo eficaz e robusto. Visualmente podemos ver a semelhança entre os resultados da Figura 7, também confirmada pelo teste estatístico apresentado na Tabela 5. A Figura 7 apresenta um gráfico de caixa que ilustra o tempo médio de segmentação dos modelos para imagens de AVC hemorrágico por TC e imagens de TC de pulmão.

Tabela 4 – Comparação do tempo de segmentação entre o modelo Detectron2 e os modelos propostos baseados em técnicas de ajuste fino para a segmentação do pulmão em imagens de TC.

Técnica	Tempo
Detectron2	0.096 ± 0.005
Detectron- $f\lambda$	2.231 ± 1.730
Detectron- $f\delta$	2.186 ± 1.890
Detectron- $f\mu$	2.295 ± 2.004

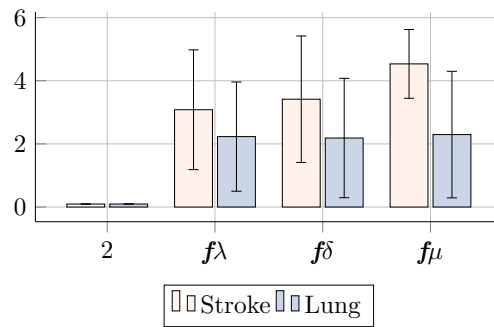


Figura 8 – Ilustração do tempo de segmentação dos modelos Detectron2, Detectron- $f\lambda$, Detectron- $f\delta$ e Detectron- $f\mu$ para a segmentação do pulmão e AVC hemorrágico. As colunas de creme representam o tempo médio para segmentação de imagens de AVC hemorrágico por TC. As colunas de cor azul representam o tempo médio de segmentação das imagens de TC do pulmão.

A Figura 8 ilustra o tempo médio de segmentação por imagem do pulmão, referenciado no gráfico com uma barra azul. Analisando também a Tabela 4, que apresenta o tempo de segmentação em conjunto com a Figura 7 dos resultados de segmentação, podemos concluir que após o processo de ajuste fino a Detectron- $f\mu$ obteve um desempenho relativamente melhor no reajuste da forma do objeto, a partir do resultado já gerado pelo Detectron2 dado como um início de crescimento para alcançar as bordas do pulmão, principalmente, no caso de pontos de curvatura onde as pontas são bastante acentuadas. Os modelos são estatisticamente equivalentes e o contorno do modelo Detectron- $f\mu$ mostrou-se relativamente flexível em alguns pontos de curvatura pulmonar fazendo com que o modelo tivesse pequenas diferenças no ajuste para as regiões pulmonares e não pulmonares.

Para apresentar a equivalência entre os modelos foi realizado o teste de Kolmogorov-Smirnov no conjunto de amostras em que cada amostra foi testada entre si e juntas geraram as métricas médias da Tabela 5. A igualdade estatística entre as amostras foi testada sendo que a Equação 5.1 explica as figuras utilizadas para indicar a situação estatística de

comparação do conjunto de métricas.

$$K \quad teste \begin{cases} \square = & \text{Igualdade Estatística} \\ \boxtimes = & \text{Diferença Estatística} \end{cases} \quad (5.1)$$

Tabela 5 – Tabela de Teste Estatístico Kolmogorov-Smirnov. Resultados das métricas na segmentação das imagens de TC do pulmão, utilizando as técnicas: A) Detectron2, B) Detectron- $\mathbf{f}\lambda$ C) Detectron- $\mathbf{f}\delta$, e D) Detectron- $\mathbf{f}\mu$.

Métodos	ACC	DICE	SEN	SPE
A-B	$\square$	$\square$	$\square$	$\square$
A-C	$\square$	$\square$	$\square$	$\square$
A-D	$\square$	$\square$	$\square$	$\square$
B-C	$\square$	$\square$	$\square$	$\square$
B-D	$\square$	$\square$	$\square$	$\square$
C-D	$\square$	$\square$	$\square$	$\square$

A Tabela 5 mostra que as técnicas de ajuste fino desenvolvidas e aplicadas ao resultado da Detectron2, mesmo que tenha havido um ligeiro aumento como houve nos valores das métricas, do ponto de vista estatístico, essas métricas permaneceram semelhantes.

5.2.2 Segunda Etapa do Experimento - Segmentação de AVC Hemorrágico

Nesta subseção, é apresentado o resultado da segmentação do AVC hemorrágico. Também utilizado no estudo de Rebouças et al. (2019), o conjunto de dados contém o padrão, e, com isso, também foi possível comparar com a segmentação do especialista validando o estudo assim como foi feito no primeiro experimento com a segmentação do pulmão.

A Tabela 6 apresenta os resultados da segmentação no conjunto de teste contendo 20 imagens de AVC hemorrágico. A rede Detectron2 conseguiu superar os modelos em quase todas as métricas, com uma ACC de $99,89 \pm 0,05$, DICE de $94,81 \pm 2,11$ e SEN de $92,79 \pm 3,87$. Isso significa que a rede Detectron2 foi capaz de detectar e segmentar um AVC hemorrágico de forma relativamente mais precisa do que os modelos com ajuste fino. A rede conseguiu ser equivalente ao modelo Detectron- $\mathbf{f}\mu$ em termos de especificidade. No entanto, ambos foram bem-sucedidos na detecção das regiões sem AVC. Isso é extremamente importante considerando que os valores de acurácia que representam a localização real do AVC são semelhantes com todos os modelos próximos a 100% de acurácia.

A principal diferença entre os modelos foi encontrada na métrica de sensibilidade (SEN), variando 7% entre o melhor e o pior caso. Esta variação ocorreu apenas entre os modelos Detectron2 e Detectron- $\mathbf{f}\lambda$. Eles tiveram uma pequena dificuldade para contornar

Tabela 6 – Resultados gerados pelo método proposto, baseado na rede Detectron2 em conjunto com técnicas de ajuste fino no conjunto de dados de TC de AVC hemorrágico.

Técnica	ACC	DICE	SEN	ESP
Detectron2	99.89 ± 0.05	94.81 ± 2.11	92.79 ± 3.87	99.97 ± 0.03
Detectron- $f\lambda$	99.83 ± 0.06	91.37 ± 3.70	85.21 ± 6.29	99.99 ± 0.02
Detectron- $f\delta$	99.88 ± 0.05	94.09 ± 2.40	90.89 ± 4.57	99.97 ± 0.03
Detectron- $f\mu$	99.88 ± 0.05	94.04 ± 2.42	90.61 ± 4.54	99.98 ± 0.03

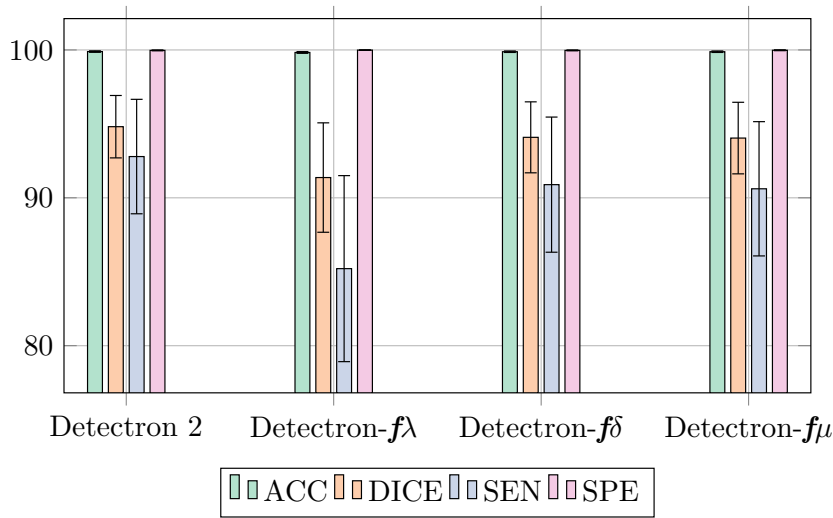


Figura 9 – Ilustração gráfica das métricas obtidas pelos métodos Detectron2, Detectron- $f\lambda$, Detectron- $f\delta$ e Detectron- $f\mu$ na segmentação do AVC hemorrágico.

a borda do traço que já estava perto do limite. Somente a técnica da Janela de Parzen para reajustar às bordas do traço torna o método mais difícil de reajustar. Isso também pode ser analisado visualmente pela observação do contorno com leves deformidades na imagem segmentada. Na Figura 10 podemos analisar visualmente os resultados gerados pela Detectron2 em diferentes imagens de TC do cérebro.

As diferenças também podem ser vistas no gráfico da Figura 9 que apresentam os valores da segmentação em forma de barra. Este gráfico de barras mostra o grande potencial da rede em segmentar diferentes objetos. Isso se deve ao fato da rede Detectron2 possuir em sua estrutura pesos pré-treinados em diferentes formas de aprendizagem, o que é uma espécie de generalização entre modelos de aprendizagem profunda.

A Tabela 7 mostra o tempo de segmentação do AVC hemorrágico em imagens de TC alcançadas pelas técnicas Detectron2, Detectron- $f\lambda$, Detectron- $f\delta$ e Detectron- $f\mu$. Claramente, a rede Detectron2 sem ajuste fino superou as outras técnicas com o menor tempo médio de segmentação, com aproximadamente 0.095 ± 0.006 segundos. A Figura 8 mostra o resultado em formato de barras.

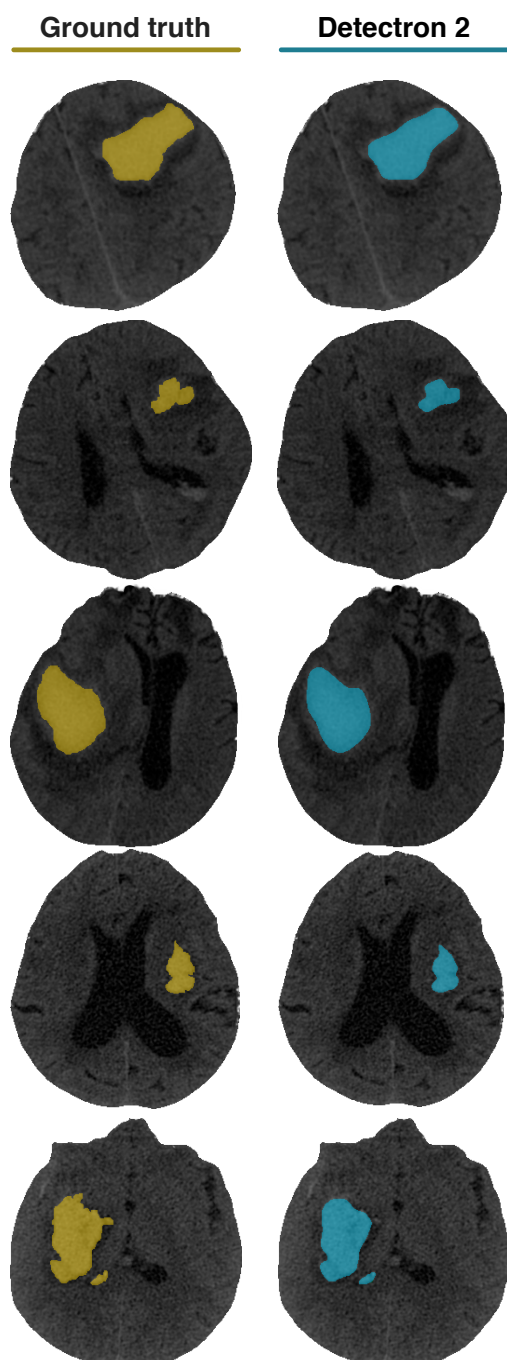


Figura 10 – Ilustração do resultado da segmentação da rede Detectron2 para o conjunto de dados de imagens de TC de AVC Hemorrágico.

A Tabela 8 demonstra que os valores métricos da técnica de ajuste fino Detectron- $f\lambda$ são estatisticamente diferentes dos valores de Detectron2, então pode-se dizer que este método foi menos eficaz na segmentação do AVC Hemorrágico, conforme mostrado na Tabela 6 e ilustrado no gráfico da Figura 9.

A Figura 10 mostra que a segmentação realizada pela Detectron2 é a que mais se aproxima da segmentação realizada pelo especialista. Isso porque, neste caso específico de AVC hemorrágico, a Detectron2 obteve o melhor desempenho entre todos os modelos.

Tabela 7 – Comparação dos tempos de segmentação entre modelos Detectron2 e modelos usando técnicas de ajuste fino. Os resultados foram gerados a partir de 20% das imagens de AVC hemorrágico.

Técnica	Tempo
Detectron2	0.095 ± 0.006
Detectron- $f\lambda$	3.082 ± 1.897
Detectron- $f\delta$	3.415 ± 2.003
Detectron- $f\mu$	4.534 ± 1.089

Tabela 8 – Tabela de Teste Estatístico Kolmogorov-Smirnov. Resultados das métricas utilizadas na segmentação de imagens de TC do cérebro, utilizando as seguintes técnicas: A) Detectron2, B) Detectron- $f\lambda$, C) Detectron- $f\delta$ e D) Detectron- $f\mu$.

Métodos	ACC	DICE	SEN	SPE
A-B	☒	☒	☒	☒
A-C	☐	☐	☒	☐
A-D	☐	☐	☒	☒
B-C	☒	☒	☒	☒
B-D	☒	☒	☒	☒
C-D	☐	☐	☐	☐

5.2.3 Terceira Etapa - Comparação entre métodos da literatura

Nesta subseção, a fim de validar nosso método, comparamos nossa melhor abordagem com outros métodos relatados na literatura que usaram os mesmos bancos de dados apresentados na Seção 3.1.

A Tabela 9 apresenta nosso melhor modelo baseado neste estudo em comparação com o trabalho de Hu et al. (2020). O experimento usou o mesmo banco de dados de TC de pulmão com diferentes tipos de métodos.

Tabela 9 – Tabela de validação para o método proposto na etapa de segmentação do sistema HTSCS com o conjunto de dados de imagens pulmonares. Os métodos usados para validação podem ser encontrados no artigo de Hu et al. (2020).

Métrica	Detectron2	Mask R-CNN	Detectron- $f\mu$	Mask + Bayes	Mask + SVM	Mask + K-means	Mask + EM
Dice	99.00 ± 0.60	76.81 ± 16.90	99.02 ± 0.60	76.10 ± 16.49	86.05 ± 11.21	97.33 ± 3.24	87.63 ± 9.39
Acurácia	96.98 ± 1.21	89.96 ± 4.38	97.01 ± 1.25	86.42 ± 11.11	95.78 ± 2.62	97.68 ± 3.42	97.28 ± 3.85
Sensibilidade	96.74 ± 2.69	87.72 ± 16.82	96.74 ± 2.78	91.06 ± 14.66	96.69 ± 10.24	96.58 ± 8.58	95.86 ± 8.67
Especificidade	99.41 ± 0.50	86.70 ± 6.12	99.45 ± 0.49	78.05 ± 17.25	92.18 ± 4.87	97.11 ± 3.65	96.42 ± 4.49

A Tabela 9 mostra os resultados gerados por nossa abordagem usando a rede Detectron2 contra a rede Mask R-CNN. Os valores são muito mais altos em comparação com a Mask R-CNN. A rede Detectron2 usada em nossa abordagem conseguiu ter um resultado melhor do que Mask R-CNN por uma diferença de mais de 23% na métrica de DICE. Detectron2 com $99,00 \pm 0,60$ contra $76,81 \pm 16,90$ em DICE. ACC de $96,98 \pm 1,21$ contra $89,96 \pm 4,38$. SEN de $96,74 \pm 2,69$ contra $87,72 \pm 16,82$ e ESP de $99,41 \pm 0,50$ contra

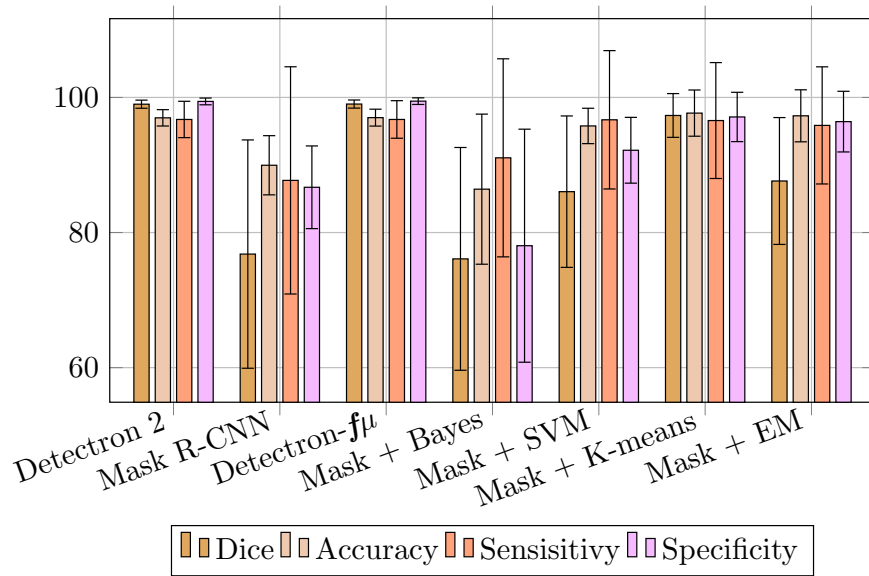


Figura 11 – Ilustração gráfica da Tabela Comparativa 9.

$86,70 \pm 6,12$. Isso mostra o alto desempenho da rede Detectron2 na detecção e segmentação da região do pulmão nas camadas mais profundas da rede. Avaliando Tabela 9, nosso modelo Detectron- $f\mu$ teve um desempenho melhor do que todos os modelos estudados por Hu et al. (2020), como os modelos Máscara + bayes, Máscara + SVM, Máscara + K-médias e Máscara + EM em todas as métricas. Todos os modelos usaram aprendizado profundo com métodos combinados com técnicas de ajuste fino. O gráfico representado pela Figura 11 distingue claramente a variação entre as métricas de segmentação pulmonar realizadas em Hu et al. (2020).

A Tabela 10 compara alguns resultados de Hu et al. (2020) e Rebouças et al. (2019), com diferentes métodos automáticos. O método de Hu et al. (2020) também usou métodos de aprendizado profundo combinados com o ajuste fino para segmentar imagens de TC de pulmão. Nosso modelo Detectron- $f\delta$ apresentou os melhores resultados entre os da Tabela 10 e foi equivalente ao método HU para a métrica DICE com 97%. No entanto, nosso modelo foi mais rápido com uma diferença de 9,54 segundos em comparação com o modelo baseado na Mask R-CNN combinado com ajuste fino por meio de aprendizado de máquina. Nossa abordagem foi melhor do que os outros métodos, inclusive no tempo de segmentação, variando de 240 segundos no pior caso a 2 segundos no melhor caso de tempo médio de segmentação pulmonar. Nossa abordagem foi melhor do que o modelo CRAD na métrica DICE que pontuou 94%, enquanto nosso melhor modelo teve 97%, superando assim a maioria dos trabalhos renomados no estado da arte.

Nas Tabelas 11 e 12 os resultados obtidos pelo trabalho de Rebouças et al. (2019) são comparados com nossa melhor abordagem (Detectron2) para segmentação de derrame

Tabela 10 – Tabela de Validação do método proposto na etapa de segmentação do sistema HTSCS para o conjunto de dados de imagem pulmonar, os métodos utilizados na validação podem ser encontrados nos artigos (HU, GVF, VFC, SISDEP, OPS, CRAD) e Rebouças et al. (2019).

Método	DICE	Time (s)
Detectron-$f\mu$	0.97 ± 0.12	2.295 ± 2.004
Method Hu (HU et al., 2020)	0.97 ± 0.03	11.24 ± 2.57
GVF	0.82 ± 0.06	240.00 ± 3.05
VFC	0.84 ± 0.04	30.00 ± 2.67
SISDEP	0.93 ± 0.02	4.00 ± 0.68
OPS	0.93 ± 0.02	6.73 ± 2.38
CRAD (REBOUÇAS et al., 2019)	0.94 ± 0.11	2.00 ± 0.16

Tabela 11 – Comparação entre métodos relatados em Rebouças et al. (2019), com nossa melhor abordagem (Detectron2) para segmentação de AVC hemorrágico.

	Método de Segmentação						
	Detectron2	LSCPM	CR	WS	OPS Euclidean	OPS Gaussian	OPS Manhattan
Dice	0.94 ± 0.02	0.81 ± 0.10	0.79 ± 0.13	0.77 ± 0.11	0.83 ± 0.07	0.83 ± 0.07	0.84 ± 0.08

Tabela 12 – Comparação do tempo de segmentação para imagens de AVC hemorrágico.

Método	Tempo
Detectron2	0.09 ± 0.01
Ws	4.80 ± 0.62
CR	3.10 ± 1.69
OPS Euclidean	2.88 ± 0.51
OPS Gaussian	2.88 ± 0.58
OPS Manhattan	2.86 ± 0.49
LSCPM	1.76 ± 0.28

cerebral.

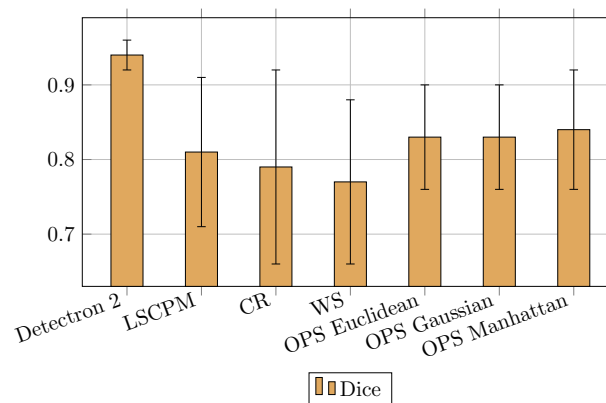


Figura 12 – Gráfico de barras ilustrando o coeficiente de Dice de cada método.

A Tabela 11 mostra que Detectron2 teve um melhor desempenho, de acordo com os

trabalhos publicados no estudo de Rebouças et al. (2019). A Detectron2 alcançou 94% de DICE contra o melhor caso OPS Manhattan com 84%, uma diferença de 10%. A diferença é ainda maior em comparação com o pior caso WS com 17%. Em outras palavras, o modelo Detectron2 provou ser mais eficaz na segmentação do AVC hemorrágico. Vale ressaltar também que essa abordagem é um método totalmente automático, com deep learning, sem interferência humana para o processo de segmentação, ao contrário dos outros métodos apresentados na Tabela. A Figura 12 ilustra os resultados para a métrica de DICE da Tabela 11.

A Rede Detectron2 conseguiu resultados melhores do que todos os trabalhos referidos em Rebouças et al. (2019) incluindo o tempo médio de segmentação das imagens. Embora esta abordagem use deep learning, que requer um custo computacional maior em comparação com os métodos clássicos (sem o uso de deep learning), nossa abordagem por meio de deep learning conseguiu ser mais rápida do que o melhor estudo de caso em Rebouças et al. (2019) para segmentação de AVC hemorrágico. A Detectron2 levou 0,09 segundos contra 1,76 para o método LSCPM. Houve uma diferença de 4,71 segundos em comparação com o pior caso denominado método Ws, conforme mostrado na Tabela 12 comparando os modelos apresentados no estudo de Rebouças et al. (2019).

O método proposto mostrou-se superior e totalmente automático para a classificação e segmentação de imagens de tomografia computadorizada de pulmão e acidente vascular cerebral hemorrágico. As comparações na Primeira Etapa de Resultados mostraram que o método obteve excelentes resultados em diferentes modelos sendo a melhor e mais eficaz abordagem para segmentação pulmonar. A comparação na Segunda Etapa de Resultados e Discussões mostrou que a rede Detectron2 foi capaz de detectar e segmentar a região do AVC sem a necessidade de ajuste fino. Isso mostra a capacidade excepcional da rede de aprendizagem profunda para se aproximar do padrão ouro realizado por um especialista podendo ser utilizado como uma espécie de pré-diagnóstico para AVC hemorrágico com segmentação automática em menos de 1 segundo. A terceira etapa proporcionou uma comparação atualizada de métodos consagrados relatados na literatura, bem como a validação de nossa abordagem. A Tabela 13 apresenta um resumo comparando as vantagens e desvantagens do trabalho proposto e os trabalhos usados para comparação.

Tabela 13 – Tabela de *Highlights*: Comparação de trabalhos utilizados como motivação para o método proposto.

Trabalhos	Técnica	Objetivo	Vantagens	Desvantagens
Método Proposto	Internet of Medical Things-based Deep Learning Techniques for Segmentation	Segmentação das regiões do pulmão e AVC em tomografias computadorizadas	Um método robusto e eficaz, com resultados alcançado o estado da arte	Ainda requer uma quantidade significativa de dados para o treinamento do modelo
(BOUCHET; PASTORE; BALLARÍN, 2007)	Morfologia Matemática	Segmentação de ramos em imagem angiográfica	Representação de imprecisões durante o processo de segmentação	Não validado
(DUAN et al., 2019)	Filtro de difusão anisotrópica e crescimento de região pulmonares em imagens de TC	Segmentação de vasos pulmonares em imagens de TC	Redução de ruído	Apenas capaz de segmentar com sucesso os vasos de um pulmão saudável e dificuldades em encontrar os parâmetros ideais
(RAJA et al., 2018)	Integração do algoritmo Chaotic Bat limite baseado em Tsallis e crescimento de região	Segmentação de imagens de ressonância magnética T1 de cérebro e seios dentais	Melhor contraste da imagem de varredura da ressonância magnética	Dificuldades em segmentar alguns slices do cérebro
(PEI et al., 2017)	Limiar	Segmentação do cérebro em imagens de ressonância magnética	Número de clusters indicados automaticamente	Dificuldade em segmentar imagens com certos tons de cinza
(SANTOS; PERKUSICH; ALMEIDA, 2014)	Diferentes tipos de gestores de saúde e porta de entrada	Gerenciamento de sistema de saúde	Distribuição dos gestores de saúde em diferentes localidades, permitindo o transporte da informação do PHD através de diferentes tecnologias e protocolos	Falhas de segurança e problemas de interface
(AL-HAMADI; CHEN, 2017)	Protocolo de comunicação baseado em tomada de decisão	Criação de um conhecimento coletivo banco de dados entre os dispositivos	O protocolo IoT baseado em confiança considera classificação de risco, confiabilidade e probabilidade de perda de saúde como três dimensões do projeto para a tomada de decisão	Não validado
(RAY, 2014)	Estrutura arquitetônica baseada na Internet das Things Home Health Hub	Acompanhar a saúde dos idosos nas suas casas	A estrutura é promissora em termos de design e visão de futuro do uso da implementação H3IoT da vida real	Não ajustado para funcionar em um ambiente clínico
(HU et al., 2020)	Um método automático composto pela utilização da rede Mask R-CNN e Aprendizado de Máquina com o auxílio de Transferência de Aprendizagem	Segmentação do pulmão em imagens de TC	Segmentação Totalmente Automática	Tempo de resposta muito longo em comparação com os métodos mais atuais
(WANG et al., 2017)	CNN	Segmentação de nódulos pulmonares	Uma técnica inovadora de classificação de voxel	Dificuldade em segmentar nódulos menores e com diferentes tamanhos de pulmões nas diferentes imagens tomográficas devido ao processo do sistema respiratório durante o exame ao processo de redução de ruído
(DURASAMY; DURASAMY, 2012)	CNN, lógica fuzzy C e K-vizinhos mais próximos	Segmentação de imagens de MRI de acidente vascular cerebral e pulmão	Suavização de ruído	Perda de informações em algumas imagens devido ao processo de redução de ruído
(SHIN et al., 2016)	CifarNet, AlexNet, GoogleNet, e outros	Classificação de doenças pulmonares, como linfonodos toracocelomais	CNNs aprendem recuses de imagem de uma maneira sofisticada	Obtenção de conjuntos de dados com anotações abrangentes como ImageNet
(WANG et al., 2018)	Aprendizagem profunda com ajuste fino	Segmentação de imagens médicas	Segmentação robusta e mais eficaz	Ajuste de parâmetro não automático
(HAVAET et al., 2015)	CNN com arquitetura em cascata	Segmentação de tumor cerebral em imagens de MRI	Arquitetura robusta e segmentação automática	Alto tempo de resposta
(CHEN; BENTLEY; RUECKERT, 2017)	Uma arquitetura composta por CNN EDD Net e MUSCULE Net	Segmentação de imagens MRI-DWI de AVC isquêmico	Arquitetura robusta e eficiente	Alta demanda por imagens para validação
(HAAN et al., 2015)	Técnicas de agrupamento	Segmentação de AVC isquêmico em imagens de TC e MRI	Rápida segmentação	Técnica não automática
(SUN et al., 2015)	Agrupamento linear iterativo combinado com o algoritmo Graph	Método de segmentação para hemorragia intracraniana em imagens de TC	Generalização da segmentação 2D tradicional de hemorragia intracraniana para uma abordagem 3D	Dificuldade em segmentar regiões com menos contraste
(REBOUÇAS et al., 2019)	Janela de Parzen	Segmentação craniana em imagens de TC	Crterios de início e parada automáticos	Método de demarcação inicial manual

6 Conclusão e Trabalhos Futuros

Este trabalho teve como objetivo desenvolver um método inovador de pré-diagnóstico médico baseado em Saúde das Coisas por meio de aprendizagem profunda e aplicações da técnica de ajuste fino. O método propõe que a aplicação de técnicas de aprendizagem profunda em conjunto com métodos de ajuste fino é capaz de generalizar o aprendizado e solucionar diferentes problemas na médica. O método foi dividido em dois estágios sendo eles a classificação e segmentação de pulmão e AVC hemorrágico em imagens de TC. No primeiro estágio do nosso método, um modelo (extrator de atributos profundos + classificador) foi desenvolvido para classificar a existência de regiões de interesses em imagens de TC.

Nesse processo, se a rede identifica o objeto de interesse na TC ela segue o Processo de Segmentação, que é o segundo estágio do método. O segundo estágio é composto por quatro modelos para segmentar a região pulmonar e a lesão por AVC hemorrágico. Os modelos apresentados neste estudo utilizaram deep learning combinado com ajuste fino, e todos os modelos obtiveram excelentes resultados.

Os resultados foram bastante satisfatórios nas duas etapas do processo. Os melhores modelos obtiveram 97% de acurácia para classificação de imagens em imagens pulmonares e não pulmonares e 100% em imagens com AVC hemorrágico e sem AVC hemorrágico. A segmentação de imagens classificadas contendo regiões de interesse obtidas com nosso o melhor modelo Detectron- μ também alcançou excelentes resultados para segmentação pulmonar com 99% Acurácia e com um tempo médio de 1,7 segundos superando os trabalhos relatados na literatura. Isso demonstra a eficiência e robustez do nosso método. Vale ressaltar que o nosso método também obteve muito sucesso na classificação e segmentação de AVC hemorrágico, atingindo 94% DICE e 99% Acurácia, com tempo médio de segmentação inferior a 1 segundo, superando os trabalhos relatados na literatura apresentados na Subseção 5.2.3.

Para validar nossa abordagem comparamos nossos melhores modelos com as técnicas de segmentação pulmonar e AVC hemorrágico que utilizaram os mesmos bancos de dados encontrados na literatura. Nosso método foi superior a todos os métodos avaliados para segmentação pulmonar superando os melhores modelos Hu et al. (2020) com 97% de DICE e com um tempo de segmentação de 11,24 segundos. O segundo melhor modelo (CRAD), que havia atingido 94% com um tempo de processamento de 2 segundos, também foi superado. Nosso modelo de segmentação de AVC também foi superior ao melhor método (LSCPM) de Rebouças et al. (2019) com 84% DICE, ou seja, uma diferença de 10% em relação ao nosso melhor caso. Além disso, o tempo médio do modelo LSCPM levou 1,76

segundos enquanto que o nosso modelo Detectron2 levou apenas 0,09 segundos. Dessa forma comprovamos a eficácia do método proposto por este estudo para a classificação e segmentação de imagens pulmonares e AVC hemorrágico em imagens de tomografia.

Para estudos futuros propomos testar diferentes bases de imagens para validar nosso método como uma ferramenta poderosa no auxílio ao diagnóstico médico para a classificação e segmentação de diferentes tipos de patologias através do conceito de Saúde das Coisas, como melanomas, mamografias, entre outras. Também propomos como trabalho futuro disponibilizar o sistema *HTSCS - Medical Images* para a Classificação e Segmentação de Imagens Médicas para teste de outros pesquisadores e/ou profissionais de saúde.

Referências

ADAMS, R.; BISCHOF, L. Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 16, n. 6, p. 641–647, 1994. Citado 2 vezes nas páginas 24 e 38.

AL-HAMADI, H.; CHEN, R. Trust-based decision making for health iot systems. *IEEE Internet of Things Journal*, IEEE, v. 4, n. 5, p. 1408–1419, 2017. Citado 2 vezes nas páginas 28 e 62.

ALJABAR, P. et al. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage*, Elsevier, v. 46, n. 3, p. 726–738, 2009. Citado na página 41.

AVINASH, S.; MANJUNATH, K.; KUMAR, S. S. An improved image processing analysis for the detection of lung cancer using gabor filters and watershed segmentation technique. In: IEEE. *2016 International Conference on Inventive Computation Technologies (ICICT)*. [S.l.], 2016. v. 3, p. 1–6. Citado na página 23.

BHADAURIA, H.; DEWAL, M. Intracranial hemorrhage detection using spatial fuzzy c-mean and region-based active contour on brain ct imaging. *Signal, Image and Video Processing*, Springer, v. 8, n. 2, p. 357–364, 2014. Citado na página 41.

BOUCHET, A.; PASTORE, J. I.; BALLARÍN, V. L. Segmentation of medical images using fuzzy mathematical morphology. *Journal of Computer Science & Technology*, v. 7, 2007. Citado 2 vezes nas páginas 27 e 62.

CHEN, L.; BENTLEY, P.; RUECKERT, D. Fully automatic acute ischemic lesion segmentation in dwi using convolutional neural networks. *NeuroImage: Clinical*, Elsevier, v. 15, p. 633–643, 2017. Citado 2 vezes nas páginas 29 e 62.

CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 1251–1258. Citado 2 vezes nas páginas 24 e 32.

CORTINA, J. M. What is coefficient alpha? an examination of theory and applications. *Journal of applied psychology*, American Psychological Association, v. 78, n. 1, p. 98, 1993. Citado na página 42.

COVER, T.; HART, P. Nearest neighbor pattern classification. *IEEE transactions on information theory*, IEEE, v. 13, n. 1, p. 21–27, 1967. Citado na página 35.

DHANACHANDRA, N.; MANGLEM, K.; CHANU, Y. J. Image segmentation using k-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, Elsevier, v. 54, p. 764–771, 2015. Citado 2 vezes nas páginas 24 e 38.

DIAZ, A. A.; CELLI, B.; CELEDÓN, J. C. Chronic obstructive pulmonary disease in hispanics. a 9-year update. *American journal of respiratory and critical care medicine*, American Thoracic Society, v. 197, n. 1, p. 15–21, 2018. Citado na página 22.

DING, G. et al. Feature mask network for person re-identification. *Pattern Recognition Letters*, Elsevier, 2019. Citado na página 23.

DOUGHERTY, G. *Digital image processing for medical applications*. [S.l.]: Cambridge University Press, 2009. Citado na página 27.

DUAN, H.-H. et al. Segmentation of pulmonary vascular tree by incorporating vessel enhancement filter and variational region-growing. *Journal of X-ray Science and Technology*, IOS Press, v. 27, n. 2, p. 343–360, 2019. Citado 2 vezes nas páginas 27 e 62.

DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern classification*. [S.l.]: John Wiley & Sons, 2012. Citado na página 36.

DURASAMY, M.; DURASAMY, S. Cnn-based approach for segmentation of brain and lung mri images. *European Journal of Scientific Research*, v. 81, n. 3, p. 298–313, 2012. Citado 3 vezes nas páginas 23, 29 e 62.

FERREIRA, P. M.; RUANO, A. E. Evolutionary multiobjective neural network models identification: evolving task-optimised models. In: *New Advances in Intelligent Signal Processing*. [S.l.]: Springer, 2011. p. 21–53. Citado na página 22.

FILHO, P. P. R. et al. A new effective and powerful medical image segmentation algorithm based on optimum path snakes. *Applied Soft Computing*, Elsevier, v. 76, p. 649–670, 2019. Citado 3 vezes nas páginas 15, 50 e 51.

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 na página 22.

FILHO, R. et al. Novel adaptive balloon active contour method based on internal force for image segmentation—A systematic evaluation on synthetic and real images. *Expert Systems with Applications*, Elsevier, v. 41, n. 17, p. 7707–7721, 2014. Citado na página 23.

FORD, N. L. et al. Micro-computed tomography imaging of a rodent model of chronic obstructive pulmonary disease (copd). In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging*. [S.l.], 2020. v. 11317, p. 113172F. Citado na página 21.

FOREY, B. A.; THORNTON, A. J.; LEE, P. N. Systematic review with meta-analysis of the epidemiological evidence relating smoking to copd, chronic bronchitis and emphysema. *BMC pulmonary medicine*, Springer, v. 11, n. 1, p. 36, 2011. Citado na página 21.

GONZALEZ, R. C.; WOODS, R. C. *Processamento digital de imagens*. [S.l.]: Pearson Educación, 2009. Citado na página 38.

HAAN, B. de et al. Fast semi-automated lesion demarcation in stroke. *NeuroImage: Clinical*, Elsevier, v. 9, p. 69–74, 2015. Citado 2 vezes nas páginas 30 e 62.

HAVAEI, M. et al. A convolutional neural network approach to brain tumor segmentation. In: SPRINGER. *BrainLes 2015*. [S.l.], 2015. p. 195–208. Citado 2 vezes nas páginas 29 e 62.

- HAYKIN, S. *Neural Networks and Learning Machines, 3/E*. [S.l.]: Pearson Education India, 2010. Citado na página 37.
- HE, K. et al. Mask r-cnn. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2017. p. 2961–2969. Citado 2 vezes nas páginas 33 e 34.
- 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 na página 33.
- HERMAN, G. T. *Fundamentals of computerized tomography: image reconstruction from projections*. [S.l.]: Springer Science & Business Media, 2009. Citado na página 22.
- HINTON, G.; VINYALS, O.; DEAN, J. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*, 2015. Citado na página 32.
- HOWARD, A. G. et al. Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017. Citado na página 33.
- HOWARD, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017. Citado na página 33.
- HU, Q. et al. An effective approach for ct lung segmentation using mask region-based convolutional neural networks. *Artificial Intelligence in Medicine*, Elsevier, p. 101792, 2020. Citado 10 vezes nas páginas 15, 23, 28, 50, 51, 58, 59, 60, 62 e 63.
- HU, S.; HOFFMAN, E. A.; REINHARDT, J. M. Automatic lung segmentation for accurate quantitation of volumetric x-ray ct images. *IEEE transactions on medical imaging*, IEEE, v. 20, n. 6, p. 490–498, 2001. Citado na página 21.
- HUANG, G. et al. Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 4700–4708. Citado na página 33.
- HUYNH, B. Q.; LI, H.; GIGER, M. L. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *Journal of Medical Imaging*, International Society for Optics and Photonics, v. 3, n. 3, p. 034501, 2016. Citado na página 32.
- JOHNSON, C. O. et al. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet Neurology*, Elsevier, v. 18, n. 5, p. 439–458, 2019. Citado na página 22.
- 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 2 vezes nas páginas 23 e 43.
- KARPATHY, A. et al. Large-scale video classification with convolutional neural networks. In: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2014. p. 1725–1732. Citado na página 32.
- KER, J. et al. Deep learning applications in medical image analysis. *Ieee Access*, IEEE, v. 6, p. 9375–9389, 2017. Citado na página 23.

- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 1097–1105. Citado na página 32.
- LAREAU, S. C. et al. Chronic obstructive pulmonary disease (copd). *American journal of respiratory and critical care medicine*, American Thoracic Society, v. 199, n. 1, p. P1–P2, 2019. Citado na página 21.
- LI, B. N. et al. Selective level set segmentation using fuzzy region competition. *IEEE access*, IEEE, v. 4, p. 4777–4788, 2016. Citado na página 27.
- LIAW, A.; WIENER, M. et al. Classification and regression by randomforest. *R news*, v. 2, n. 3, p. 18–22, 2002. Citado na página 37.
- LIN, T.-Y. et al. Microsoft coco: Common objects in context. In: SPRINGER. *European conference on computer vision*. [S.l.], 2014. p. 740–755. Citado na página 33.
- LYNCH, D. A.; AL-QAISI, M. L. Quantitative ct in copd. *Journal of thoracic imaging*, NIH Public Access, v. 28, n. 5, p. 284, 2013. Citado na página 22.
- MA, Y. et al. Big health application system based on health internet of things and big data. *IEEE Access*, IEEE, v. 5, p. 7885–7897, 2016. Citado na página 28.
- MADANI, A.; KEYZER, C.; GEVENOIS, P.-A. Quantitative computed tomography assessment of lung structure and function in pulmonary emphysema. *European Respiratory Journal*, Eur Respiratory Soc, v. 18, n. 4, p. 720–730, 2001. Citado na página 22.
- MEDEIROS, A. G. et al. An automatic left ventricle segmentation on echocardiogram exams via morphological geodesic active contour with adaptive external energy. 2019. Citado na página 23.
- MIKOLOV, T. et al. Distributed representations of words and phrases and their compositionality. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2013. p. 3111–3119. Citado na página 23.
- MILLETARI, F.; NAVAB, N.; AHMADI, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: IEEE. *2016 Fourth International Conference on 3D Vision (3DV)*. [S.l.], 2016. p. 565–571. Citado na página 41.
- MOLCHANOV, P. et al. Pruning convolutional neural networks for resource efficient transfer learning. *arXiv preprint arXiv:1611.06440*, v. 3, 2016. Citado na página 32.
- MOZAFFARIAN, D. et al. Executive summary: heart disease and stroke statistics—2016 update: a report from the american heart association. *Circulation*, Am Heart Assoc, v. 133, n. 4, p. 447–454, 2016. Citado na página 22.
- MUZAMMAL, M. et al. A multi-sensor data fusion enabled ensemble approach for medical data from body sensor networks. *Information Fusion*, Elsevier, v. 53, p. 155–164, 2020. Citado na página 22.
- NÓBREGA, R. V. M. da et al. Lung nodule classification via deep transfer learning in ct lung images. In: IEEE. *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*. [S.l.], 2018. p. 244–249. Citado 2 vezes nas páginas 23 e 32.

- NÓBREGA, R. V. M. da; RODRIGUES, M. B.; FILHO, P. P. R. Segmentation and visualization of the lungs in three dimensions using 3d region growing and visualization toolkit in ct examinations of the chest. In: IEEE. *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*. [S.l.], 2017. p. 397–402. Citado na página 23.
- ORGANIZATION, W. H. *Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth*. [S.l.]: World Health Organization, 2017. Citado na página 21.
- PATERNOSTER, R. et al. Using the correct statistical test for the equality of regression coefficients. *Criminology*, Wiley Online Library, v. 36, n. 4, p. 859–866, 1998. Citado na página 41.
- PEI, H.-X. et al. D-fcm: Density based fuzzy c-means clustering algorithm with application in medical image segmentation. *Procedia Computer Science*, Elsevier, v. 122, p. 407–414, 2017. Citado 2 vezes nas páginas 27 e 62.
- PRÆSTGAARD, J. T. Permutation and bootstrap kolmogorov-smirnov tests for the equality of two distributions. *Scandinavian Journal of Statistics*, JSTOR, p. 305–322, 1995. Citado na página 42.
- RACHEDI, A. et al. Ieee access special section editorial: The plethora of research in internet of things (iot). *IEEE Access*, IEEE, v. 4, p. 9575–9579, 2016. Citado na página 23.
- RAJA, N. S. M. et al. Contrast enhanced medical mri evaluation using tsallis entropy and region growing segmentation. *Journal of Ambient Intelligence and Humanized Computing*, Springer, p. 1–12, 2018. Citado 2 vezes nas páginas 27 e 62.
- RAJSIC, S. et al. Economic burden of stroke: a systematic review on post-stroke care. *The European Journal of Health Economics*, Springer, v. 20, n. 1, p. 107–134, 2019. Citado na página 21.
- RAY, P. P. Home health hub internet of things (h 3 iot): An architectural framework for monitoring health of elderly people. In: IEEE. *2014 International Conference on Science Engineering and Management Research (ICSEMR)*. [S.l.], 2014. p. 1–3. Citado 2 vezes nas páginas 28 e 62.
- 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, n. 19, p. 9265–9286, 2019. Citado 11 vezes nas páginas 16, 24, 30, 38, 39, 55, 59, 60, 61, 62 e 63.
- RUSSAKOVSKY, O. et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, Springer, v. 115, n. 3, p. 211–252, 2015. Citado na página 32.
- SANTOS, D. F.; PERKUSICH, A.; ALMEIDA, H. O. Standard-based and distributed health information sharing for mhealth iot systems. In: IEEE. *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom)*. [S.l.], 2014. p. 94–98. Citado 2 vezes nas páginas 28 e 62.
- SARMENTO, R. M. et al. Automatic neuroimage processing and analysis in stroke-a systematic review. *IEEE reviews in biomedical engineering*, IEEE, 2019. Citado na página 22.

- SCARPATO, N. et al. E-health-iot universe: A review. *management*, v. 21, n. 44, p. 46, 2017. Citado 2 vezes nas páginas 24 e 28.
- SHIN, H.-C. et al. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, IEEE, v. 35, n. 5, p. 1285–1298, 2016. Citado 4 vezes nas páginas 23, 29, 31 e 62.
- SORIANO, J. B. et al. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the global burden of disease study 2015. *The Lancet Respiratory Medicine*, Elsevier, v. 5, n. 9, p. 691–706, 2017. Citado na página 21.
- SOUZA, L. F. de F. et al. Automatic lung segmentation in ct images using mask r-cnn for mapping the feature extraction in supervised methods of machine learning using transfer learning. *International Journal of Hybrid Intelligent Systems*, IOS Press, n. Preprint, p. 1–17. Citado na página 32.
- SUN, M. et al. Intracranial hemorrhage detection by 3d voxel segmentation on brain ct images. In: IEEE. *2015 International Conference on Wireless Communications & Signal Processing (WCSP)*. [S.l.], 2015. p. 1–5. Citado 2 vezes nas páginas 30 e 62.
- SZEGEDY, C. et al. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Thirty-first AAAI conference on artificial intelligence*. [S.l.: s.n.], 2017. Citado na página 33.
- SZEGEDY, C. et al. Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 1–9. Citado 3 vezes nas páginas 24, 32 e 33.
- TAHA, A. A.; HANBURY, A. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, BioMed Central, v. 15, n. 1, p. 29, 2015. Citado na página 40.
- TEASELL, M. R.; HUSSEIN, N. Clinical consequences of stroke. *Evidence-Based Review of Stroke Rehabilitation. Ontario: Heart and Stroke Foundation and Canadian Stroke Network*, p. 1–30, 2016. Citado na página 22.
- THEODORIDIS, S.; KOUTROUMBAS, K. Pattern recognition. 2003. *Google Scholar Google Scholar Digital Library Digital Library*, 2009. Citado 2 vezes nas páginas 34 e 40.
- VANNIER, M. W. Ct clinical perspective: challenges and the impact of future technology developments. In: IEEE. *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. [S.l.], 2009. p. 1909–1912. Citado na página 22.
- VAPNIK, V.; VAPNIK, V. Statistical learning theory wiley. *New York*, v. 1, p. 624, 1998. Citado na página 36.
- WANG, G. et al. Interactive medical image segmentation using deep learning with image-specific fine tuning. *IEEE transactions on medical imaging*, IEEE, v. 37, n. 7, p. 1562–1573, 2018. Citado 2 vezes nas páginas 29 e 62.
- WANG, S. et al. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. *Medical image analysis*, Elsevier, v. 40, p. 172–183, 2017. Citado 2 vezes nas páginas 28 e 62.

- WEISS, K.; KHOSHGOFTAAR, T. M.; WANG, D. A survey of transfer learning. *Journal of Big data*, SpringerOpen, v. 3, n. 1, p. 9, 2016. Citado na página 23.
- WU, S.; ZHONG, S.; LIU, Y. Deep residual learning for image steganalysis. *Multimedia tools and applications*, Springer, v. 77, n. 9, p. 10437–10453, 2018. Citado na página 33.
- WU, Y. et al. *Detectron2*. 2019. <<https://github.com/facebookresearch/detectron2>>. Citado 2 vezes nas páginas 32 e 34.
- YU, Y. et al. Fruit detection for strawberry harvesting robot in non-structural environment based on mask-rcnn. *Computers and Electronics in Agriculture*, Elsevier, v. 163, p. 104846, 2019. Citado na página 23.
- ZHANG, R. et al. Automatic segmentation of acute ischemic stroke from dwi using 3-d fully convolutional densenets. *IEEE transactions on medical imaging*, IEEE, v. 37, n. 9, p. 2149–2160, 2018. Citado na página 23.
- ZHOU, Z. et al. Fine-tuning convolutional neural networks for biomedical image analysis: actively and incrementally. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 7340–7351. Citado na página 23.