

Segmentação do Pâncreas em Tomografia Computadorizada com *Ensemble* Guiado por Atlas e Backbone Swin Transformer

Felipe R. S. Teles¹, Neilson P. Ribeiro^{1,2}, Celso L. S. Soares Filho¹,
Luana B. da Cruz³, João O. B. Diniz^{1,2}, Geraldo B. Júnior¹, Anselmo C. de Paiva¹

¹Núcleo de Computação Aplicada
Programa de Pós Graduação em Ciência da Computação
Universidade Federal do Maranhão (UFMA)

²Fábrica de Inovação Grajáu
Instituto Federal do Maranhão (IFMA)

³Laboratório de Inteligência Computacional Aplicada (LICA)
Universidade Federal do Cariri (UFCA)

felipe.teles@nca.ufma.br

Abstract. *Automatic segmentation of the pancreas in abdominal computed tomography scans is a challenging task due to the organ's anatomical variability and low contrast. This work proposes a deep learning-based method structured in three stages: (1) filtering irrelevant slices with a Convolutional Neural Network, (2) detection of the region of interest via probabilistic atlas, and (3) final segmentation using an ensemble of convolutional and transformer-based networks. The experiments achieve an average Dice score of 78.55%. The results demonstrate competitive performance compared to recent pancreas segmentation methods, combining preprocessing and robust deep learning techniques.*

Resumo. *A segmentação automática do pâncreas em Tomografias Computadorizadas abdominais é uma tarefa desafiadora devido à variabilidade anatômica e ao baixo contraste do órgão. Este trabalho propõe um método baseado em aprendizado profundo estruturado em três etapas: (1) filtragem de fatias irrelevantes com uma Rede Neural Convolutiva, (2) detecção da região de interesse via atlas probabilístico e (3) segmentação final utilizando um ensemble de redes convolucionais e baseadas em Transformer. Os experimentos alcançam uma pontuação média de Dice de 78,55%. Os resultados demonstram desempenho competitivo em relação a métodos recentes de segmentação do pâncreas, combinando pré-processamento e técnicas robustas de aprendizado profundo.*

1. Introdução

Dados da Agência Internacional de Pesquisa sobre o Câncer apontam que o câncer de pâncreas é o sexto mais letal entre todos os casos levantados em sua pesquisa de nível global em 2022, respondendo por mais de 450 mil óbitos [Bray et al. 2024]. Neste cenário, é essencial desenvolver soluções que auxiliem os especialistas no diagnóstico, a fim de aumentar as chances de sucesso no tratamento e reduzir a mortalidade.

A identificação do pâncreas em imagens médicas é uma etapa crucial no tratamento, no planejamento da radioterapia e no monitoramento da evolução estrutural [Silva et al. 2021]. A Tomografia Computadorizada (TC) abdominal é o exame mais utilizado nesse processo, pois fornece imagens detalhadas que permitem a diferenciação do pâncreas dos tecidos circundantes, facilitando sua delimitação para uma análise clínica mais precisa [Daniel et al. 2024].

Modelos baseados em aprendizado profundo têm sido amplamente adotados para diversas tarefas na área médica [Diniz et al. 2021, Neto et al. 2024, Junior et al. 2024, Teles et al. 2025]. Neste contexto, as Redes Neurais Convolucionais (CNNs) [LeCun et al. 1998] e os Transformers [Vaswani 2017] têm obtido avanços contínuos, possibilitando o desenvolvimento de soluções para desafios como a segmentação automática do pâncreas em TCs.

Desta forma, o objetivo deste trabalho é propor um método computacional combinando técnicas de pré-processamento e aprendizado profundo para segmentação do pâncreas. A hipótese deste estudo é que, ao utilizar técnicas de redução de escopo, como filtrar fatias sem o pâncreas e empregar um atlas probabilístico para a localização inicial da região do pâncreas, o desempenho do *ensemble* das redes de segmentação pode ser aprimorado. Portanto, o método proposto oferece as seguintes contribuições:

- Redução de informações desnecessárias por meio da filtragem de fatias sem o pâncreas e do uso de um Atlas Probabilístico para refinar a região de interesse.
- Um método de segmentação do pâncreas totalmente automático baseado em *ensemble* de redes neurais que supera a maioria dos trabalhos relacionados.

Além disso, métodos de segmentação eficazes podem aprimorar o Diagnóstico Auxiliado por Computador (CAD) para médicos, reduzir custos operacionais para hospitais e diminuir o tempo de espera para pacientes, melhorando, conseqüentemente, sua qualidade de vida.

2. Trabalhos Relacionados

A tarefa de segmentação do pâncreas utilizando aprendizado profundo segue duas abordagens principais. Sendo elas a abordagem 3D, processando dados volumétricos onde os *voxels* são as unidades fundamentais para análise, e, a segunda abordagem, por sua vez, opera em fatias bidimensionais individuais, utilizando pixels como elemento base para extração de características. Esta seção apresenta uma revisão sistemática dos principais trabalhos relacionados a essas metodologias, destacando seus avanços, limitações e contribuições para o campo.

[Li et al. 2021] propôs um método híbrido que integra abordagens baseadas em atlas e em rede neural. Seu estudo usou o conjunto de dados *National Cancer Institute* (NIH), com binarização aplicada entre $-100HU$ e $240HU$. Além disso, eles avaliaram o método no conjunto de dados *Medical Segmentation Decathlon* (MSD). A entrada consistiu em imagens 2.5D, processadas por uma rede FCN. A pontuação média de *Dice* alcançada foi 71,36%.

[Zhu et al. 2023] introduziu um *framework* de controle de qualidade de imagem combinada com um algoritmo de segmentação para pâncreas. Eles usaram uma U-Net 3D para segmentar o pâncreas, avaliando sua abordagem nos conjuntos de dados MSD e

NIH. A o melhor resultado foi a pontuação média de *Dice* de 75,43% no conjunto de dados NIH.

[Ferrara et al. 2024] propôs o uso de Redes Neurais Convolucionais Profundas (DCNNs) como uma etapa de pré-processamento para segmentar a região do pâncreas antes de identificar tumores malignos. Eles treinaram seu modelo usando o conjunto de dados MSD, adotando uma abordagem 2.5D com uma arquitetura SRSNet. A melhor pontuação média de *Dice* foi 66,8%.

[Kurnaz et al. 2024] propôs a Pascal U-Net para segmentação do pâncreas. Ao contrário da U-Net tradicional, a Pascal U-Net estabelece conexões por meio de camadas convolucionais em vez de links diretos entre o codificador e o decodificador. Seu modelo foi avaliado no conjunto de dados *The Cancer Imaging Archive* (TCIA), atingindo 71,35% para a pontuação média de *Dice*.

[Juwita et al. 2025] propôs a MOMUNet (MobileNetV2 + U-Net), uma rede leve para segmentação do pâncreas em TC, combinada com uma estratégia de *External Contour Cropping* (ECC) para mitigar o desbalanceamento entre fundo e pâncreas. Avaliada nas bases de dados NIH e MSD, a MOMUNet com ECC obteve 75,29% como pontuação média de *Dice* na base da dados MSD, fazendo uso da MOMUNet e do ECC.

Por fim, [Tsewalo Tondji et al. 2025] propuseram um modelo de segmentação do pâncreas baseado em U-Net 3D com mecanismo de atenção, projetado para aprimorar a capacidade da rede em destacar regiões anatômicas relevantes em cenários de baixo contraste e alta complexidade estrutural. A abordagem foi avaliada em dois conjuntos de dados de referência amplamente utilizados, o NIH e o MSD. No conjunto MSD, o método alcançou uma pontuação média de *Dice* de 78,1%.

Embora modelos 3D, como o proposto por Zhu et al.(2023) e Tsewalo Tondji et al.(2025), apresentem bom desempenho na segmentação do pâncreas, eles demandam elevado custo computacional, o que limita sua aplicação em cenários práticos. Trabalhos como o de Ferrara et al.(2024) exploram abordagens 2.5D, geralmente associadas a técnicas de pré-processamento para enriquecimento do contexto espacial. Já métodos puramente 2D, como os de Kurnaz et al.(2024), Li et al.(2021) e Juwita et al.(2025), oferecem menor complexidade computacional, porém tendem a apresentar desempenho inferior.

Neste contexto, nossa abordagem combina técnicas de pré-processamento e aprendizado profundo para segmentação automática do pâncreas em imagens de TC. O método utiliza um atlas probabilístico para auxiliar na localização do órgão e aplica filtragem de fatias irrelevantes para reduzir o custo computacional. Além disso, emprega um *ensemble* híbrido de CNNs e Transformers para melhorar a qualidade da segmentação.

3. Método Proposto

Esta seção apresenta o método proposto, descrevendo a base de dados utilizada e as principais etapas do processamento das imagens. A Figura 1 ilustra de forma geral o fluxo do método, incluindo o pré-processamento, o treinamento das redes neurais de segmentação, a configuração do *ensemble* e a etapa de avaliação dos resultados por meio de métricas consolidadas na literatura.

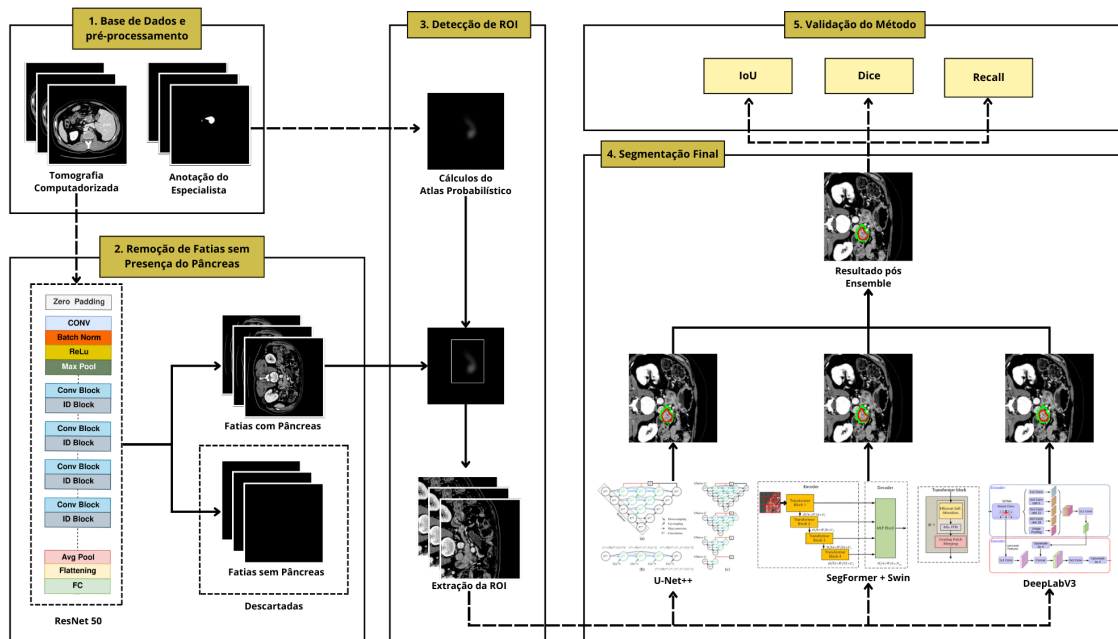


Figura 1. Ilustração do Método.

3.1. Base de Dados e Pré-Processamento

A base de dados utilizado por este estudo é o *Medical Segmentation Decathlon* (MSD) [Antonelli et al. 2022], que contém 420 volumes de TC abdominais coletados pelo *Memorial Sloan Kettering Cancer Center*. Cada volume tem aproximadamente 100 fatias (imagens 2D) com resolução de 512×512 . O pâncreas foi anotado manualmente em cada fatia por um radiologista usando o aplicativo *Scout*. Exemplos de fatias do conjunto de dados podem ser vistos na Figura 2, onde as anotações do pâncreas variam entre uma única região ou dois objetos distintos, dependendo da fatia.

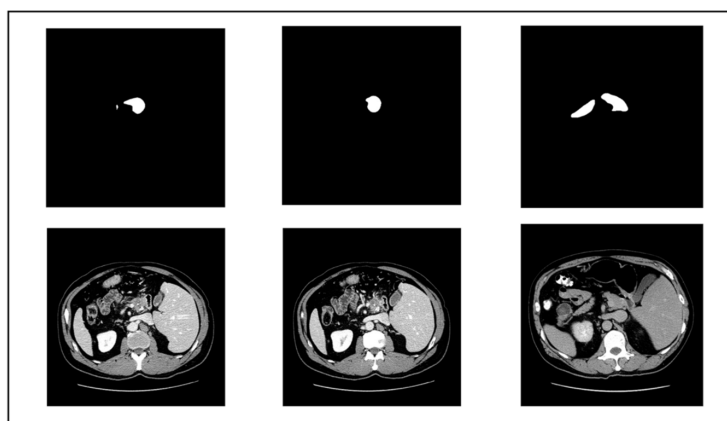


Figura 2. Exemplo da Base de Dados.

Antes de enviar as imagens para a CNN de classificação, os volumes passaram por um processo de janelamento, onde os valores de intensidade dos volumes foram limitados ao intervalo de $[-150, 250]$ na escala da Unidade de Hounsfield (HU), melhorando

a visualização das estruturas abdominais ao excluir regiões de ar e estruturas altamente densas, como ossos [Fernandes et al. 2023].

3.2. Remoção de Fatias sem Presença do Pâncreas

Nesta etapa, as fatias por paciente são reduzidas retirando do conjunto aquelas que não incluem o pâncreas. Utiliza-se uma CNN treinada para classificação binária, onde a classe positiva corresponde a uma fatia com o pâncreas e a classe negativa a uma sem o pâncreas.

Para esta tarefa, um modelo ResNet50 [He et al. 2016]. Para validar a qualidade do treinamento foi utilizado um método estatístico de validação chamado de *Cross-Validation* para verificar se o modelo estaria apto para ser utilizado no estudo, isto é, com uma boa capacidade de generalização.

Após o treinamento, o modelo foi capaz de filtrar fatias que não contêm o pâncreas. A perda de fatias com pâncreas foi corrigida com um pós-processamento, ao considerar como positivas as imagens que estavam entre duas fatias com a presença do pâncreas.

3.3. Detecção de ROI

Com fatias contendo o pâncreas selecionadas, a Região de Interesse (ROI) será obtida reduzindo a imagem para conter apenas a área onde o órgão está localizado, otimizando a análise. Para isso, foi-se usado um Atlas Probabilístico, um método estatístico avançado que quantifica e modela a variabilidade funcional em uma determinada população, gerando uma representação média de características específicas.

Ao calcular a posição média das máscaras no conjunto de dados de treinamento, o atlas define uma área probabilística para o pâncreas, permitindo a remoção de recursos irrelevantes por meio do corte da imagem. O atlas é obtido usando a Equação 1:

$$A = \frac{1}{N} \sum_{i=1}^N \frac{I_i}{255} \quad (1)$$

onde N é o número total de pixels na área, e I_i denota a intensidade do i -ésimo pixel, normalizado para o intervalo $[0, 1]$.

Uma *bounding box* é então gerada para definir a ROI. Finalmente, todas as fatias filtradas são cortadas de acordo com o ROI calculado, preservando apenas a região do pâncreas para a etapa final de segmentação. A Figura 3 ilustra o atlas com a *bounding box* sendo aplicado para obter a ROI.

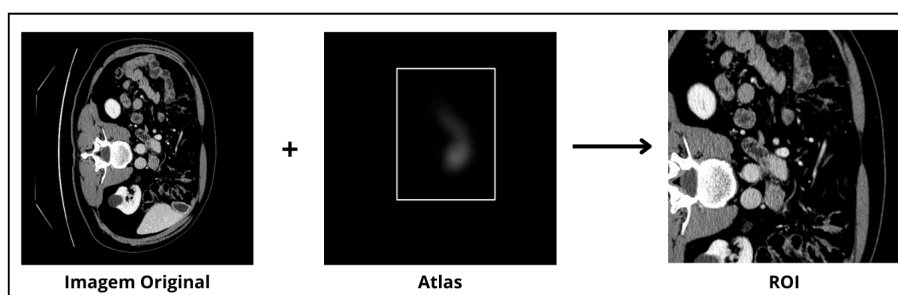


Figura 3. Extração da ROI.

3.4. Segmentação Final

Após filtrar as fatias sem o pâncreas e detectar o ROI usando o atlas probabilístico, o próximo passo é a segmentação do pâncreas. Foram treinadas três redes neurais, sendo elas, a U-Net++ [Zhou et al. 2018], a DeepLabV3 [Chen et al. 2017] e a SegFormer [Xie et al. 2021] utilizando como encoder a Swin Transformer [Carion et al. 2020]. Todas as redes foram treinadas de forma independente sobre as regiões de interesse previamente delimitadas, permitindo uma comparação justa entre os modelos e fornecendo as previsões necessárias para a etapa posterior de combinação via *ensemble* [Naderalvojoud and Hernandez-Boussard 2024].

Para que o *ensemble* apresente um desempenho superior ao dos modelos individuais, é fundamental definir adequadamente o peso atribuído a cada rede no processo de combinação [Naderalvojoud and Hernandez-Boussard 2024]. Dessa forma, foi realizada a estratégia de busca *Grid Search*, que consiste em levantar diferentes configurações de pesos priorizando a configuração que agrega o melhor desempenho para a segmentação. Os pesos atribuídos a cada modelo foram inicializados de forma aleatória.

3.5. Validação do Método

Para validar o método, métricas padrão de imagens médicas são adotadas. Na etapa de filtragem de fatias, Acurácia, Especificidade, Precisão, Sensibilidade e *F1-Score* avaliaram o desempenho do modelo na detecção de fatias com o pâncreas, minimizando a quantidade de fatias da classe negativa. Para segmentação, a pontuação *Dice* mediu a sobreposição entre a máscara prevista e a marcação do especialista, enquanto o *Recall* garante que a maior parte do pâncreas seja segmentada corretamente, reduzindo o risco de regiões críticas ausentes. Além disso, a métrica *Intersection over Union (IoU)* foi utilizada para avaliar a similaridade entre a segmentação predita e a anotação de referência, mensurando a proporção entre a interseção e a união das regiões segmentadas.

4. Resultados e Discussão

Esta seção apresenta os resultados dos experimentos para avaliação do método proposto. Primeiro, define-se o ambiente de treinamento usado para o desenvolvimento. Em seguida, os experimentos projetados para validar cada etapa são apresentados. Finalmente, nossos resultados são comparados com os trabalhos relacionados discutidos na Seção 2, e, em seguida, os estudos de caso com uma análise qualitativa são apresentados.

4.1. Ambiente de Treinamento

Os experimentos foram conduzidos usando o ambiente de computação em nuvem Google Colab Pro, que fornece acesso a uma GPU Tesla T4 com 12 GB de RAM. A implementação foi escrita em Python, utilizando como bibliotecas principais *TensorFlow*, *PyTorch*, *NumPy*, *OpenCV* e *Matplotlib*.

4.2. Experimentos

Esta seção descreve os diferentes cenários experimentais projetados para validar cada etapa do método proposto.

Para a validação do método, utiliza-se o conjunto de dados descrito na Seção 3.1, onde 420 volumes foram divididos em 80% para treinamento e 20% para teste, garantindo

que fatias do mesmo paciente não fossem compartilhadas entre os conjuntos, evitando assim vazamento de dados. Todas as técnicas foram aplicadas apenas ao conjunto de treinamento, enquanto o conjunto de teste foi usado exclusivamente para avaliação do método proposto.

4.2.1. Experimento 1: Validação da Etapa de Filtragem de Fatias

Nesta etapa, a *ResNet50* [He et al. 2016] foi avaliada para a tarefa de classificar se as fatias continham o pâncreas ou não (Seção 3.2). O modelo de classificação foi inicializado com os pesos pré-treinados na ImageNet [Deng et al. 2009] para, em seguida, ocorrer o treinamento em 30 épocas usando o otimizador *Adam* com uma taxa de aprendizado de 0,0001 e um tamanho de *batch* de 16. Para encontrar os hiperparâmetros ótimos, aplica-se um método de otimização baseado em *Grid Search*, que explora sistematicamente um conjunto predefinido de valores no otimizador (*Adam* e *AdamW*), a taxa de aprendizado (0,01, 0,001 e 0,0001) e o tamanho do lote (8, 12 e 16).

Antes de realizar o pós processamento, foi avaliado o desempenho completo do modelo, foi realizado um método de validação cruzada, *K-Fold*. O resultado individual de cada *Fold* pode ser analisado na Tabela 1, que apresenta as métricas de desempenho e o desvios-padrão.

Tabela 1. Resultados da validação cruzada *K-Fold* para a etapa de filtragem de fatias.

Fold	Sensibilidade(%)	Precisão(%)	<i>F1-Score</i> (%)	Especificidade(%)
0	93,54	93,71	93,40	96,62
1	94,53	92,63	93,46	96,02
2	93,83	93,17	93,31	96,18
3	94,23	95,51	94,68	97,32
4	94,87	96,57	95,64	97,71
Média	94,20% ± 0,53	94,32% ± 1,66	94,10% ± 1,03	96,77% ± 0,73

Com um *F1-Score* médio de 94,10% ± 1,03, percebe-se um desempenho efetivo do modelo na distinção de fatias contendo o pâncreas. Além disso, o baixo desvio padrão indica a estabilidade e confiabilidade do modelo.

Tabela 2. Impacto do pós-processamento no desempenho da filtragem de fatias.

Configuração	Sensibilidade(%)	Precisão(%)	<i>F1-Score</i> (%)	Especificidade(%)	Acurácia(%)
Sem pós-processamento	94,20	94,32	94,10	94,10	96,77
Com pós-processamento	98,07	94,26	95,99	96,36	96,96

Os resultados da 2 destacam a importância da etapa de pós processamento para a filtragem, que contribui para reduzir o número de falsos negativos nas imagens que serão processadas na segmentação, otimizando a eficiência do método.

4.2.2. Experimento 2: Treinamento das Redes de Segmentação

No segundo experimento, três arquiteturas de segmentação semântica foram avaliadas individualmente: *DeepLabV3*, *UNet++* e *SegFormer* com backbone *Swin Transformer*. Todas as redes foram treinadas utilizando apenas as fatias previamente filtradas pela etapa de classificação, assegurando que o processo de segmentação fosse aplicado exclusivamente a imagens contendo o pâncreas.

As arquiteturas foram treinadas sob o mesmo protocolo experimental, incluindo divisão dos dados, função de perda e critérios de avaliação, de modo a garantir uma comparação justa entre os modelos. A avaliação foi realizada no conjunto de teste utilizando as métricas *Dice*, *IoU* e *Recall*, apresentadas na seção 3.5. Os resultados quantitativos são apresentados na Tabela 3.

Tabela 3. Desempenho dos modelos individuais e do *ensemble* no conjunto de teste.

Modelo	Dice(%)	IoU(%)	Recall(%)
DeepLabV3	75,70	64,73	81,61
UNet++	75,94	65,19	82,46
SegFormer + SwinTR	77,46	66,63	84,11
Ensemble	78,55	68,17	84,41

Observa-se que o modelo *SegFormer + SwinTR* obteve o melhor desempenho individual, alcançando coeficiente *Dice* de 76,18% e *Recall* de 83,33%. Esse resultado evidencia a contribuição do modelo baseado em Transformer, capaz de capturar características de longo alcance, além de modelar melhor o contexto global da imagem. As arquiteturas *UNet++* e *DeepLabV3* também apresentaram resultados competitivos, porém com desempenho um pouco inferior.

4.2.3. Experimento 3: Aplicação do *Ensemble*

No terceiro experimento, foi aplicada uma estratégia de *ensemble* combinando as predições das três arquiteturas de segmentação treinadas individualmente. O objetivo dessa etapa foi explorar a complementaridade entre modelos convolucionais e baseados em Transformer, reduzindo erros individuais e aumentando a robustez do método.

A estratégia de *ensemble* resultou em melhora consistente em todas as métricas avaliadas, atingindo, respectivamente, *Dice* e *Recall* médios de 78.55% e 86.24%, superando os resultados obtidos por qualquer modelo individualmente. Este resultado foi encontrado após a otimização da rede após a aplicação de um *Grid Search* com os possíveis pesos, onde os três melhores resultados estão presentes na Tabela 4. Esses ganhos confirmam que a combinação das redes permite integrar informações locais, capturadas principalmente pelas arquiteturas convolucionais, com o contexto global fornecido pelo modelo baseado em Transformer.

A utilização do *ensemble* permite explorar a complementaridade entre arquiteturas convolucionais, como a *DeepLabV3* e a *UNet++*, e modelos baseados em *Transformer*, como o *SegFormer*, resultando em ganhos consistentes tanto em desempenho quanto em

Tabela 4. Melhores distribuições de pesos do *ensemble* e seus respectivos desempenhos.

Rank	DeepLabV3	UNet++	SwinTR	Dice (%)	Recall (%)
1°	0.20	0.30	0.50	78.55	84.41
2°	0.15	0.25	0.60	78.08	86.24
3°	0.10	0.30	0.60	78.08	86.22

robustez. Na Tabela 4, observa-se que as melhores distribuições de pesos priorizam o SwinTR, evidenciando a importância da modelagem de dependências de longo alcance, enquanto as CNNs contribuem para a preservação de detalhes locais e estabilidade na segmentação.

4.3. Comparação com Trabalhos Relacionados

Nesta seção, o desempenho do nosso método com abordagens existentes na literatura é comparado e exibido na Tabela 5. A avaliação é baseada na abordagem (3D, 2.5D e 2D) e no Dice, destacando as vantagens e limitações da nossa abordagem.

Tabela 5. Comparação do método proposto com trabalhos relacionados em segmentação do pâncreas.

Trabalho	Abordagem	Arquitetura	Dataset	Dice (%)
Li et al. (2021)	2.5D	FCN + Atlas	NIH/MSD	71.36
Zhu et al. (2023)	3D	U-Net 3D	NIH	75.43
Ferrara et al. (2024)	2.5D	DCNN	MSD	66.80
Kurnaz et al. (2024)	2D	Pascal U-Net	TCIA	71.35
Juwita et al. (2025)	2D	MOMUNet + ECC	MSD	75.29
Tondji et al. (2025)	3D	U-Net 3D com atenção	MSD	78.10
Método proposto	2D	Ensemble (CNN + Transformer)	MSD	78.55

Em relação aos trabalhos descritos na Tabela 5, nosso método 2D proposto se destaca como a abordagem de melhor desempenho, alcançando uma pontuação *Dice* de 78.55%. Este resultado supera os métodos 2.5D e 3D, que são normalmente considerados arquiteturas mais robustas devido à sua capacidade de capturar informações espaciais mais completas.

4.4. Estudo de Caso

Para uma melhor avaliação do método proposto, serão apresentados estudos de caso que fornecem uma análise qualitativa dos resultados da segmentação. Esses estudos de caso ilustram como o método funciona em diferentes cenários e destacam sua capacidade de segmentar com precisão o pâncreas em imagens de TC médicas.

A Figura 4 mostra os melhores casos do conjunto de dados de teste, onde, nas fatias, a anotação do especialista está em verde e a segmentação está em vermelho. Observa-se uma boa sobreposição entre as regiões segmentadas automaticamente e as anotações manuais, indicando que o método proposto é capaz de capturar adequadamente a forma e a localização do pâncreas mesmo em diferentes variações anatômicas e níveis de contraste presentes nas imagens de TC.

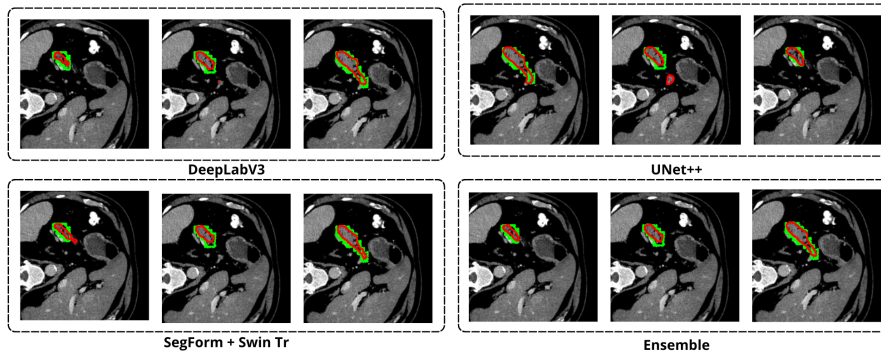


Figura 4. Cenário Favorável.

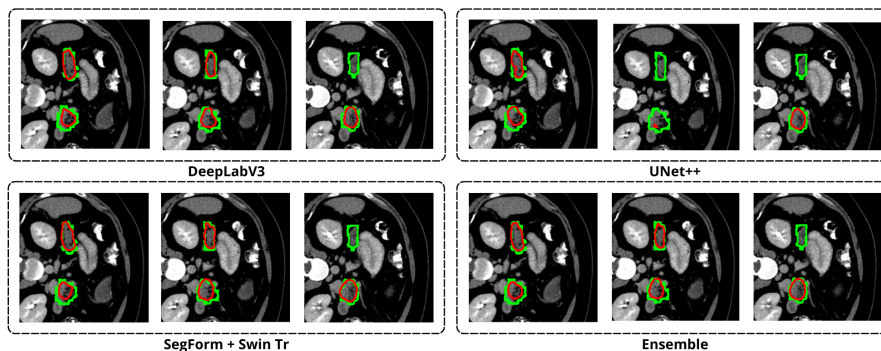


Figura 5. Cenário Desafiador.

Percebe-se que no melhor mesmo as redes tendo um bom desempenho individual, o ensemble foi capaz de melhorar ainda mais o resultado, comprovando sua eficácia. Esse comportamento evidencia que a combinação das previsões permite reduzir erros individuais e aumentar a robustez do método frente a variações nos dados. Entretanto também existe uma limitação no pior caso, presente na Figura 5, pois ela não é capaz de aprender padrões, apenas priorizar os melhores resultados.

5. Conclusão e Trabalhos Futuros

Este trabalho apresentou um método automático para segmentação do pâncreas em TC abdominal, combinando filtragem de fatias irrelevantes, localização de ROI via atlas probabilístico e segmentação por meio de um *ensemble* de redes profundas. A abordagem reduziu o custo computacional ao focar apenas em regiões relevantes, mantendo desempenho competitivo. Os experimentos alcançaram uma pontuação média de *Dice* médio de 78,55% no conjunto MSD, superando métodos 2D, 2.5D e 3D recentes. Como trabalhos futuros, pretende-se explorar ponderação adaptativa no *ensemble*, incorporar contexto volumétrico parcial e avaliar o método em outras bases de dados e cenários clínicos.

6. Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, da Fundação de Amparo à Pesquisa do Maranhão (FAPEMA) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Além disso, reconhece-se o uso de *Large*

Language Models (LLM) para verificação ortográfica, correção gramatical e auxílio na tradução de termos específicos.

Referências

- Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B. A., Litjens, G., Menze, B., Ronneberger, O., Summers, R. M., et al. (2022). The medical segmentation decathlon. *Nature communications*, 13(1):4128.
- Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., and Jemal, A. (2024). Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 74(3):229–263.
- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., and Zagoruyko, S. (2020). Hierarchical vision transformer using shifted windows. In *European Conference on Computer Vision (ECCV)*, volume 2.
- Chen, L.-C., Papandreou, G., Schroff, F., and Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*.
- Daniel, N., Farinella, R., Chatziioannou, A. C., Jenab, M., Mayén, A.-L., Rizzato, C., Belluomini, F., Canzian, F., Tavanti, A., Keski-Rahkonen, P., et al. (2024). Genetically predicted gut bacteria, circulating bacteria-associated metabolites and pancreatic ductal adenocarcinoma: a mendelian randomisation study. *Scientific Reports*, 14(1):25144.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255.
- Diniz, J. O., Ferreira, J. L., da Silva, G. L., Quintanilha, D. B., Silva, A. C., and Paiva, A. (2021). Segmentação de coração em tomografias computadorizadas utilizando atlas probabilístico e redes neurais convolucionais. In *Anais do XXI Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 83–94. SBC.
- Fernandes, A. G. S., Braz Junior, G., Diniz, J. O. B., Silva, A. C., and Matos, C. E. F. (2023). Efficientdeeplab for automated trachea segmentation on medical images. In *Brazilian Conference on Intelligent Systems*, pages 154–166. Springer.
- Ferrara, N., Andria, G., Scarpetta, M., Lanzolla, A. M. L., Attivissimo, F., Di Nisio, A., and Ramos, D. (2024). 2d and 2.5 d pancreas and tumor segmentation in heterogeneous ct images of pdac patients. In *2024 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–5. IEEE.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Junior, D. A. D., da Cruz, L. B., and Diniz, J. O. (2024). Classificação da camada lipídica do filme lacrimal usando k-means e deep learning. In *Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*, pages 1–12. SBC.
- Juwita, J., Hassan, G. M., and Datta, A. (2025). Pancreas segmentation in ct scans: A novel momunet based workflow. *Computers in Biology and Medicine*, 193:110346.

- Kurnaz, E., Ceylan, R., Bozkurt, M. A., Cebeci, H., and Koplay, M. (2024). A novel deep learning model for pancreas segmentation: Pascal u-net. *Inteligencia Artificial*, 27(74):22–36.
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- Li, J., Lin, X., Che, H., Li, H., and Qian, X. (2021). Pancreas segmentation with probabilistic map guided bi-directional recurrent unet. *Physics in Medicine & Biology*, 66(11):115010.
- Naderalvojud, B. and Hernandez-Boussard, T. (2024). Improving machine learning with ensemble learning on observational healthcare data. In *AMIA Annual Symposium Proceedings*, volume 2023, page 521.
- Neto, C. M. S., Silva, A. L., Pessoa, A. C., Quintanilha, D. B., de Almeida, J. D., Junior, G. B., and Diniz, J. O. (2024). Diagnóstico de tuberculose em imagens de radiografia utilizando cvt. In *Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*, pages 342–353. SBC.
- Silva, G., Oliveira, F., Diniz, J., Diniz, P., Quintanilha, D., Silva, A., Paiva, A., and Cavalcanti, E. (2021). An automatic method for prostate segmentation on 3d mri scans using local phylogenetic indexes and xgboost. In *Anais do XXI Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 165–176, Porto Alegre, RS, Brasil. SBC.
- Teles, F. R., Ribeiro, N. P., da Cruz, L. B., Júnior, G. B., de Paiva, A. C., Diniz, J. O., and Cortes, O. A. (2025). Atlas-guided u-net++ with efficientnetb5 for automatic pancreas segmentation in abdominal ct scans. In *Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*, pages 104–115. SBC.
- Tsewalo Tondji, I. C., Scapicchio, C., Lizzi, F., Fantacci, M. E., Oliva, P., and Retico, A. (2025). Deep learning model with attention mechanism for a 3d pancreas segmentation in ct scans. *Mathematics*, 13(24):3942.
- Vaswani, A. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.
- Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., and Luo, P. (2021). Segformer: Simple and efficient design for semantic segmentation with transformers. *Advances in neural information processing systems*, 34:12077–12090.
- Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., and Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*, pages 3–11. Springer.
- Zhu, Y., Hu, P., Li, X., Tian, Y., Bai, X., Liang, T., and Li, J. (2023). An end-to-end data-adaptive pancreas segmentation system with an image quality control toolbox. *Journal of Healthcare Engineering*, 2023(1):3617318.