

# Detecção de Pólipos em Imagens de Colonoscopia utilizando Aprendizado Profundo

Diene Xie<sup>1</sup>, Larissa S. X. Rosa<sup>2</sup>, Vilmar C. P. Filho<sup>2</sup>, Alisson A. Cardoso<sup>2</sup>, Ricardo A. P. Franco<sup>3</sup>,

<sup>1</sup>Centro de Excelência em Inteligência Artificial (CEIA) - Universidade Federal de Goiás (UFG) - Goiânia - GO - Brasil

<sup>2</sup>Escola de Engenharia Elétrica, Mecânica e de Computação – Universidade Federal de Goiás (UFG) – Goiânia – GO – Brasil

<sup>3</sup>Instituto de Informática – Universidade Federal de Goiás (UFG) – Goiânia – GO – Brasil

diene12374@gmail.com, {larissarosa, vilmarfilho}@discente.ufg.br, {alsnac, ricardofranco}@ufg.br

**Abstract.** *This paper proposes to carry out a polyp detection study in colonoscopy images using deep neural object detection models. For this, a review of works in the literature is carried out, where more recent object detection models are used. Results in terms of precision, recall, mean average precision (mAP) and F1 score are presented. The results obtained indicate that the models trained for the detection of polyps in the Kvasir dataset presented superior results in relation to those obtained in the literature, showing that these models can be a powerful ally in medicine to aid in the detection of polyps.*

**Resumo.** *Este artigo propõe realizar um estudo de detecção de pólipos em imagens de colonoscopia utilizando modelos neurais profundos de detecção de objetos. Para isto, uma revisão de trabalhos na literatura é realizada, onde modelos mais recentes de detecção de objetos são utilizados. Resultados em termos de precisão, recall, mean average precision (mAP) e score F1 são apresentados. Os resultados obtidos indicam que os modelos treinados para a detecção de pólipos no dataset Kvasir apresentaram resultados superiores em relação aos obtidos na literatura, mostrando que esses modelos podem ser um poderoso aliado na medicina no auxílio da detecção de pólipos.*

## 1. Introdução

O câncer ainda é uma doença complexa e que representa uma ameaça significativa à saúde global. De acordo com a Agência Internacional de Pesquisa sobre o Câncer (IARC), em 2020, foram estimados cerca de 19,3 milhões de novos casos de câncer e 10 milhões de mortes relacionadas ao câncer em todo o mundo [IARC 2020]. São vários os fatores de risco associados à essa doença, tais como genética, estilo de vida e exposição ambiental [ACS 2020].

O câncer é uma doença que pode se manifestar de diferentes formas nas diversas partes do corpo humano. Uma dessas formas de câncer que merece destaque é o câncer de intestino, também conhecido como câncer de cólon e reto ou câncer colorretal. Ele é

o terceiro câncer mais comum em homens e o segundo câncer mais comum em mulheres. De acordo com o Instituto Nacional de Câncer, no Brasil, o número estimado de novos casos para cada ano entre 2023 e 2025 é de 45.630 novos casos [INCA 2022]. Já em termos de mortalidade, de acordo com o Instituto Nacional de Câncer José Alencar Gomes da Silva, em 2020, ocorreram 20.245 óbitos por câncer colorretal, sendo 9.889 óbitos entre os homens e 10.356 entre mulheres [INCA 2022].

O câncer colorretal é um crescimento desordenado de células no cólon (intestino grosso) ou no reto. A primeira etapa observada geralmente é a formação de pólipos no revestimento interno do cólon ou reto. Os pólipos são pequenos crescimentos de tecido anormal que se projetam para dentro do lúmen intestinal. Somente certos tipos de pólipos, chamados de adenomas, se transformam em câncer ao longo do tempo. Com o crescimento, os pólipos adenomatosos podem eventualmente invadir camadas mais profundas e se entrarem na corrente sanguínea ou no sistema linfático, podem se espalhar em outras partes do corpo, podendo gerar tumores secundários [Fearon & Vogelstein 1990].

A evolução para o câncer colorretal é um processo lento, e a detecção precoce e a intervenção médica podem salvar vidas, pois aumentam significativamente as chances de tratamento bem-sucedido [Krishnendu, S. *et al.* 2020]. Por esta razão, é importante fazer o rastreamento dos pólipos através da realização constante de consultas médicas e exames, permitindo a remoção deles antes que se tornem malignos. Um dos exames é o de colonoscopia, que ajuda no diagnóstico e na detecção precoce do câncer colorretal, bem como na avaliação de pólipos e outras condições do cólon e reto. Durante o exame, um tubo flexível, chamado colonoscópio, com uma câmera na sua extremidade é inserido no reto e avançado lentamente pelo cólon. A câmera transmite imagens em tempo real para um monitor, permitindo que o médico examine o revestimento interno do cólon em busca de anomalias, como os pólipos [Levin B. *et al.* 2008].

A análise dos resultados de exames de colonoscopia é demorada e pode apresentar falhas de detecção de pólipos ou identificação de falsos pólipos, levando a procedimentos errados de tratamento. A detecção automática dos pólipos surge como uma das abordagens para auxiliar nesta tarefa [Krishnendu, S. *et al.* 2020].

A utilização da inteligência artificial (IA) tem se destacado na área da saúde como uma ferramenta de apoio à tomada de decisões. Ela permite analisar grandes quantidades de dados médicos e identificar padrões complexos, fornecendo insights valiosos aos profissionais de saúde [Obermeyer *et al.* 2016]. A aprendizagem profunda (*deep learning*) é um ramo da IA que se baseia em redes neurais profundas e tem a capacidade de extrair automaticamente características relevantes em imagens médicas. Dessa forma, estudos recentes têm surgido sobre o uso de modelos de redes neurais convolucionais para detecção de pólipos em imagens de exames de colonoscopia [ELKarazle, K. *et al.* 2023].

Detecção de pólipo colorretal no contexto da inteligência artificial significa o processo de treinar um modelo de aprendizado de máquina (*machine learning*) para aprender um conjunto de características específicas que representam os pólipos para então detectar eles por meio de uma caixa delimitadora (*bounding box*) dado uma imagem ou vídeo de entrada [ELKarazle, K. *et al.* 2023].

São várias as etapas envolvidas na construção de um sistema de detecção de pólipos usando inteligência artificial. Primeiramente a seleção de um conjunto de dados

para treinamento e validação, em seguida um pré-processamento dos dados, depois aplicação de técnicas de aumento de dados para aumentar a variedade de imagens, seguido do treinamento do modelo e finalizando com a validação do modelo. Existem diversas formas de diversificar cada etapa, desde a escolha de parâmetros até a escolha do modelo a ser treinado [LeCun, Y. *et al.* 2015].

Existem diversos modelos de detecção de objetos, de arquiteturas diferentes com parâmetros diferentes e aplicados em objetivos diferentes, ou seja, conjunto de dados diferentes. Um modelo pode não ter o mesmo nível de resultados ao mudar o tipo de objeto a ser detectado. O objetivo deste artigo é apresentar uma revisão de modelos já existentes, escolher os mais atuais e realizar a implementação prática deles para validação de suas eficácias na detecção de pólipos em imagens de colonoscopia.

O artigo está organizado da seguinte forma: a Seção 2 apresenta a revisão bibliográfica de trabalhos relacionados; a metodologia incluindo coleta de dados, modelos e treinamento e inferência é apresentada na Seção 3, os resultados obtidos na Seção 4 e finalizando com a conclusão na Seção 5.

## 2. Trabalhos Relacionados

Grande parte dos estudos recentes sobre o uso de técnicas de inteligência artificial na área da saúde baseiam-se em métodos de aprendizagem profunda. Algoritmos de aprendizado profundo têm sido aplicados com sucesso em muitos campos da saúde, especialmente na análise de imagens médicas [Krishnendu 2020]. Dessa forma, esta seção apresenta alguns trabalhos desenvolvidos na literatura para realizar a tarefa de detecção de pólipos utilizando imagens de colonoscopia e algoritmos de detecção de objetos.

*Region-based Convolutional Neural Network* (R-CNN) é um algoritmo de detecção de objetos desenvolvido em 2014 [Aston Zhang *et al.* 2023]. R-CNN usa redes neurais convolucionais para localizar objetos de interesse e realizar a extração de características de forma independente de cada região de interesse para processamento posterior. Alguns trabalhos apresentaram bons resultados na detecção de pólipos utilizando modelos baseados no modelo R-CNN, detectando pólipos de formas e tamanhos distintos [ELKarazle *et al.* 2023].

O modelo *Single-Shot Detector* (SSD) depende de uma única rede neural profunda para detectar um objeto [Aston Zhang *et al.* 2023]. O modelo SSD discretiza o espaço de saída das caixas delimitadoras em várias caixas com diferentes proporções. Depois que as caixas delimitadoras são discretizadas, a rede é dimensionada por localização do mapa de características. A localização de um objeto de interesse é prevista usando vários mapas de características com diferentes resoluções. Os trabalhos [Aston Zhang *et al.* 2023] propuseram utilizar o modelo SSD para realizar a detecção de pólipos colorretais.

A primeira versão do *You Only Look Once* (YOLO) [Aston Zhang *et al.* 2023] foi apresentada em 2016 e é um dos algoritmos mais utilizados recentemente para a tarefa de detecção de objetos. O algoritmo YOLO foi modificado diversas vezes, gerando novas versões do algoritmo no decorrer dos últimos anos. O YOLO também é um dos modelos de detecção de objetos que vem sendo utilizado com frequência na detecção de pólipos [Joseph Redmon 2016].

O modelo *Real-Time Detection Transformer* (RT-DETR) [Wenyu Lv *et al.* 2023]

foi desenvolvido pela Baidu e foi apresentado em julho de 2023. Uma de suas características é o uso de *transformers* e consiste em usar um backbone, um codificador híbrido e um decodificador transformador com cabeçotes auxiliares de predição. Os autores apresentaram ainda os resultados que superaram os resultados com outros modelos, como a YOLO e suas diversas versões, utilizando o conjunto de dados da COCO 2017.

Os autores em [Aston Zhang *et al.* 2023] projetaram e treinaram uma Convolutional Neural Network (CNN) profunda para realizar a detecção de pólipos usando um conjunto diversificado e representativo de imagens de colonoscopias de triagem coletadas de mais de 2.000 pacientes. Foram treinadas diferentes arquiteturas CNN neste estudo.

### **3. Metodologia**

Este trabalho tem dois principais focos: 1 - estudo dos métodos, técnicas e conjuntos de dados usados recentemente para detecção de pólipos; e 2 - implementação prática dos detectores mais atuais, RT-DETR e YOLO, para validação de suas eficácias no problema em questão. O estudo dos métodos e técnicas foi abordado na seção anterior, então será tratada a busca por conjuntos de dados na subseção 3.1, sobre os modelos escolhidos na subseção 3.2 e sobre as métricas escolhidas para comparações na subseção 3.3.

#### **3.1. Conjunto de dados**

A coleta de dados desempenha um papel fundamental em qualquer pesquisa baseada em aprendizado profundo, pois influencia diretamente na qualidade de resultados, mas sabe-se que há uma certa dificuldade para disponibilização de imagens médicas, por questões éticas. Diante disso, foi realizado um levantamento de conjuntos de dados públicos disponíveis na internet, contendo imagens de colonoscopia e anotações no formato correto para o modelo YOLO. Entre os materiais levantados, destaca-se: *Deep Neural Networks approaches for detecting and classifying colorectal polyps* [Alba Nogueira-Rodríguez *et al.* 2021], pois disponibiliza uma série de tabelas e comparações de conjuntos de dados, incluindo públicos.

Após o levantamento, foi selecionado o conjunto de dados Kvasir-SEG [Debesh Jha *et al.* 2020], contendo 1000 imagens de pólipos e suas anotações. A resolução das imagens varia de 332x487 a 1920x1072 pixels, e a divisão para treinamento e teste foi na proporção de 95/5.

#### **3.2. Modelos**

Já em sua primeira versão, a YOLO revolucionou o que se conhecia como detecção de objetos, trazendo uma abordagem inovadora e maior velocidade em relação a outros modelos da época, como a Fast-RCNN, Faster-RCNN e DPM. Além disso, a velocidade de inferência era consideravelmente mais rápida, sendo possível a inferência em tempo real. Todas essas vantagens foram possíveis devido à sua principal característica: o algoritmo de regressão. Ele é o fator determinante para o próprio nome da rede: *You Only Look Once*, que é a inovação de fato - o algoritmo “olha” apenas uma

vez para a imagem, dividindo-a em diversos retângulos e inferindo onde estão os objetos através de probabilidade.

Pulando para a versão 3 [Joseph Redmon, Ali Farhadi 2018], suas principais melhorias são: nova espinha dorsal (DarkNet53), com mais camadas convolucionais; adição de conexões residuais; previsões em múltiplas escalas (três grades em escalas diferentes), melhorando a detecção de objetos pequenos; e a possibilidade de múltiplas classes em uma mesma caixa delimitadora, devido à troca da função softmax para entropia binária cruzada.

A YOLOv5 pode ser considerada um ponto de revolução na própria história da YOLO, visto que foi o primeiro modelo Pytorch da série, disponibilizado pela Ultralytics em 2020. Além disso, a versão 5 possui vasta contribuição da comunidade, através de seu repositório no GitHub. Suas melhorias mais relevantes foram: um algoritmo de âncora automática, que aplica um algoritmo k-means nos dados e evolui as posições das caixas predefinidas para serem usadas nas previsões; a espinha dorsal é a CSPDarkNet, uma versão modificada que começa com uma camada de convolução ampliada, reduzindo custos computacionais; uma série de transformações para aumento de dados, que melhora o aprendizado; há também uma camada SPPF (*Spatial Pyramid Pooling Fast*), que acelera o cálculo agrupando características de diferentes escalas em um mapa de características de tamanho fixo.

A versão YOLOv6 foi publicada em 2022 pelo Meituan Vision AI Department, e superou os modelos anteriores em métricas de precisão e velocidade. As novidades destacadas são: uma nova espinha dorsal e baseada em RepVGG [Xiaohan Ding *et al.* 2021], chamada EfficientRep, que usa maior paralelismo; e uma estratégia de auto-destilação (*self-distillation*), que permite transferir o aprendizado de camadas profundas para camadas superficiais através de módulos de atenção.

Assim como a versão 6, a YOLOv7 superou os detectores da época em precisão e velocidade, na faixa de 5 FPS a 160 FPS. Sua precisão foi incrementada sem afetar a velocidade de inferência, em especial devido à soma de suas inovações: o uso da estratégia E-ELAN (*Extended Efficient Layer Aggregation Network*), que permite que um modelo profundo aprenda e convirja de forma mais eficiente, controlando o caminho de gradiente; normalização dos lotes (*batch normalization*) durante a convolução, que normaliza as ativações intermediárias, melhorando a estabilidade e o desempenho geral da rede.

A versão YOLOv8 foi lançada em janeiro de 2023, é considerada um modelo de última geração para tarefas de detecção de objetos. Ela deriva diretamente da YOLOv5 e também foi disponibilizada pela *Ultralytics*. A respeito de suas melhorias, foram as seguintes: uma alteração da espinha dorsal (CSPDarkNet), permitindo combinar recursos na convolução e conseqüentemente melhorando a precisão; são utilizadas funções de ativação separadas para objetividade - probabilidade de a caixa delimitadora conter um objeto - e classificação - probabilidade de o objeto ser de determinada classe -, sendo sigmóide para a primeira e softmax para a segunda; as funções de perda de caixa delimitadora e de classificação também são separadas, o que auxilia na detecção de objetos pequenos.

O modelo RT-DETR foi lançado em julho de 2023, e apresenta como principal característica o uso de transformadores. Mais detalhadamente, o modelo utiliza um codificador híbrido eficiente para processar recursos em várias escalas e propõe uma

seleção de consultas com reconhecimento de *IoU*, fornecendo consultas de objetos iniciais de maior qualidade ao decodificador [Wenyu Lv *et al.* 2023].

Para fins de comparação de resultados, utilizou-se os seguintes modelos pré-treinados, realizando a transferência de aprendizado (*transfer learn*) no conjunto de dados Kvasir-SEG: YOLOv5: modelos Nano; Small; Medium; Large e Extra-large. YOLOv6: modelos Nano; Small; Medium e Large. YOLOv7: modelo Tiny e X. YOLOv8: modelos Nano; Small; Medium; Large e Extra-large. RT-DETR: modelo Large e Extra-large. Todos os modelos foram treinados no Google Colab, com GPU T4, ao longo de 300 épocas, utilizando um batch size variável de 8 a 32. Os resultados obtidos serão discutidos na seção 4.

### 3.3. Métricas

As métricas utilizadas para comparar os resultados entre os modelos de treinamento YOLO nas versões 5 a 8 e RT-DETR foram: Precisão, Recall, mAP@50 e mAP@50-95. Antes de explicar o significado de cada métrica é preciso entender o significado de *IoU* (Intersection over Union) e os possíveis resultados de classificação obtidos de um modelo de detecção de objetos dado um limite de *IoU*.

**IoU** é a métrica que representa a interseção sobre a união, comparando as caixas delimitadoras da verdade e prevista. Seja A a caixa delimitadora verdadeira e B a caixa delimitadora prevista, a fórmula do *IoU* é:

$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Dado um valor limite de *IoU*, um resultado de previsão de modelo pode ser considerado como: Verdadeiro Positivo (VP) - classificação certa da classe positivo; Falso Negativo (FN) - modelo previu como classe negativo sendo o valor real a classe positivo; Falso Positivo (FP) - modelo previu como classe positivo sendo o valor real a classe negativo; ou Verdadeiro Negativo (VN) - classificação certa da classe negativo.

**Precisão** é a métrica que irá medir dentre todas as classificações de classe positivo que o modelo fez, quantas de fato possuem esse valor real, ou seja:

$$precisão = \frac{VP}{VP+FP}$$

**Recall** é a métrica que irá medir quantas classificações corretas foram feitas da classe positivo em relação ao total esperado, ou seja, em relação a todas as situações de classe positivo como valor esperado:

$$recall = \frac{VP}{VP+FN}$$

**mAP@50 e mAP@50-95:** Para entender o *Mean Average Precision* é necessário entender primeiramente a *Average Precision*, ou seja, a precisão média (AP). Ela representa a área debaixo da curva de um gráfico das métricas *precision X recall*, conhecida também como curva AP. Entretanto este gráfico possui um aspecto de zigue-zague, já que a precisão diminui com o aumento de falsos positivos (FP) e aumenta com o aumento de verdadeiros positivos (VP), tornando difícil a comparação

entre modelos diferentes. Dessa forma, a forma da curva AP é resumida calculando a precisão em um classes de valores de recall igualmente espaçados.

O mAP@50 é então a média da AP dado um limiar de sobreposição de pelo menos 50% de IoU e o mAP@50-95 é a média dos valores de AP para cada limiar de sobreposição no intervalo de 50% a 95%, com incrementos de 5%.

#### 4. Resultados

No processo de treinamento no conjunto de dados Kvasir-SEG [Debesh Jha *et al.* 2020], foram registradas as métricas exibidas na tabela 1 para os modelos YOLO nas versões 5 a 8 e RT-DETR.

**Tabela 1** - Métricas de desempenho dos modelos treinados para o Kvasir.

Modelo	Métricas			
	Precisão	Recall	mAP@50	mAP@50-95
YOLOv5-N	0,884	0,830	0,910	0,726
YOLOv5-S	0,923	0,871	0,913	0,743
YOLOv5-M	0,889	0,872	0,916	0,744
YOLOv5-L	0,889	0,927	0,932	0,755
YOLOv5-X	0,918	0,815	0,906	0,725
YOLOv6-N	0,872	0,866	0,917	0,707
YOLOv6-S	0,840	0,873	0,897	0,718
YOLOv6-M	0,854	0,850	0,887	0,673
YOLOv6-L	0,815	0,855	0,877	0,650
YOLOv7-Tiny	0,896	0,782	0,877	0,604
YOLOv7	0,877	0,909	0,917	0,720
YOLOv7-X	0,938	0,836	0,915	0,692
YOLOv8-N	0,873	0,875	0,923	0,723
YOLOv8-S	0,876	0,903	0,931	0,751
YOLOv8-M	0,922	0,855	0,931	0,757
YOLOv8-L	0,920	0,873	0,930	<b>0,774</b>
YOLOv8-X	0,887	0,891	0,915	0,734
RT-DETR L	0,882	0,855	0,885	0,736
RT-DETR X	0,905	0,855	0,872	0,737

Destaca-se o modelo YOLOv8L, considerando o map@50-95, e também ao ser avaliado com a matriz de confusão normalizada, evidenciando seu desempenho na detecção de pólipos. Quando se observa valores favoráveis nessa matriz, como proporções elevadas de verdadeiros positivos e verdadeiros negativos, isso sinaliza que o modelo está acertando com precisão as classificações, minimizando os erros de classificação, como pode ser visto na tabela 2.

**Tabela 2** - Matriz de Confusão Normalizada. Fonte: próprios autores, 2023.

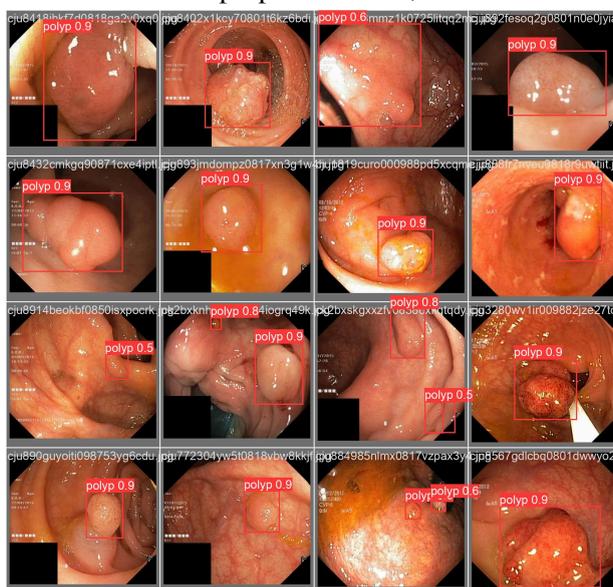
Matriz de Confusão		Classe Original	
		pólipo	background
Classe Preditá	pólipo	0.89	1.00
	background	0.11	0.00

A figura 1 exibe alguns pólipos detectados pelo modelo, bem como os níveis de precisão associados a cada detecção. Ela permite uma compreensão visual e quantitativa do quão preciso e confiável é o modelo na tarefa de detecção de pólipos em imagens médicas.

No contexto deste estudo, após a apresentação dos resultados obtidos, também foi feita uma análise comparativa, investigando na literatura os resultados de outros modelos já empregados no mesmo conjunto de dados. Essa análise comparativa, em conformidade com a metodologia de [Jha et al. 2020] no artigo "Real-Time Polyp Detection, Localisation and Segmentation in Colonoscopy Using Deep Learning," contribuiu significativamente para uma compreensão mais sólida das estratégias adotadas no contexto deste estudo. A comparação está ilustrada na tabela 3, e a métrica considerada foi o Mean Average Precision 50 (mAP@50).

**Figura 1** - Pólipos detectados na validação do modelo YOLOv8L.

Fonte: próprios autores, 2023.



**Tabela 3** - Métricas de desempenho dos modelos obtidos na literatura e do melhor modelo treinado (YOLOv8-L) para o conjunto de dados Kvasir.

Método	Rede Principal	mAP@50
EfficientDet-D0	EfficientNet-b0, biFPN	0.5047
Faster R-CNN	ResNet50	0.8418

RetinaNet	ResNet50	<b><u>0.9095</u></b>
RetinaNet	ResNet101	<b><u>0.9095</u></b>
YOLOv3+spp	Darknet53	0.8532
YOLOv4	Darknet53, CSP	0.8234
ColonSegNet	-	0.8166
YOLOv8-L ( <b>proposto</b> )	-	<b><u>0.930</u></b>

## 5. Conclusão

Este estudo abordou a detecção precoce de pólipos colorretais por meio do uso de inteligência artificial, com um foco significativo na metodologia adotada. A seleção criteriosa de um conjunto de dados público, composto por uma variedade de imagens de colonoscopia, foi essencial para garantir a representatividade da amostra.

Os treinamentos dos modelos YOLO (versões 5 até a 8) e RT-DETR (*Large e Extra Large*) foram conduzidos de maneira sistemática e controlada, com ajustes de hiperparâmetros para otimizar o desempenho. A análise comparativa com outros modelos, como o RetinaNet [T.-Y. Lin *et al.*, 2023], destaca a superioridade dos modelos YOLO em termos de detecção de pólipos.

Em síntese, este artigo enfatizou que a metodologia é a base fundamental para o sucesso de pesquisas na área de detecção de pólipos colorretais por meio de inteligência artificial. A qualidade da metodologia adotada não apenas garantiu resultados sólidos, mas também estabeleceu um padrão para avaliações rigorosas. A detecção precoce de pólipos é crucial para a prevenção do câncer colorretal e, ao enfatizar a importância da metodologia sólida, este estudo contribui para avanços contínuos nessa área, promovendo a saúde e a qualidade de vida dos pacientes.

## Referências

- IARC Global Cancer Observatory. (2020). “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries”, <https://pubmed.ncbi.nlm.nih.gov/33538338/>, Outubro.
- American Cancer Society (ACS). (2020). “Cancer Facts & Figures 2020”, em: Atlanta: American Cancer Society; 2020.
- Instituto Nacional de Câncer (INCA). (2022). “Estimativa 2023: incidência de câncer no Brasil”, <https://www.inca.gov.br/publicacoes/livros/estimativa-2023-incidencia-de-cancer-no-brasil>, Outubro.
- Fearon, E. R., & Vogelstein, B. (1990). “A genetic model for colorectal tumorigenesis”, <https://pubmed.ncbi.nlm.nih.gov/2188735/>, Outubro.
- Krishnendu, S. *et al.* (2020). “A Review on Polyp Detection and Segmentation in Colonoscopy Images using Deep Learning”, <https://www.ijert.org/a-review-on-polyp-detection-and-segmentation-in-colonoscopy-images-using-deep-learning>, Outubro.

- ELKarzle, K. *et al.* (2023). "Detection of Colorectal Polyps from Colonoscopy Using Machine Learning: A Survey on Modern Techniques", <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9953705/>, Outubro
- Obermeyer Z, Emanuel EJ. (2016). "Predicting the Future - Big Data, Machine Learning, and Clinical Medicine", <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5070532/>, Outubro.
- Levin B. *et al.* (2008). "Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology", <https://pubmed.ncbi.nlm.nih.gov/18384785/>, Outubro.
- LeCun, Y. *et al.* (2015). "Deep learning", <https://doi.org/10.1038/nature14539>, Outubro.
- Joseph Redmon, *et al.* (2016). "You Only Look Once: Unified, Real-Time Object Detection" arXiv. <https://arxiv.org/abs/1506.02640>, Outubro.
- Wenyu Lv *et al.* (2023). "DETRs Beat YOLOs on Real-time Object Detection", arXiv. <https://arxiv.org/pdf/2304.08069.pdf>, Outubro.
- Alba Nogueira-Rodríguez *et al.* (2021). "Deep Neural Networks approaches for detecting and classifying colorectal polyps". *Neurocomputing*. <https://www.sciencedirect.com/science/article/pii/S0925231220307359>, Outubro.
- Debesh Jha *et al.* (2020). "Kvasir-seg: A segmented polyp dataset". arXiv. <https://arxiv.org/abs/1911.07069>, Outubro.
- Juan Terven, Diana C. (2023). "A Comprehensive Review Of Yolo: From Yolov1 and Beyond". arXiv. <https://arxiv.org/abs/2304.00501>, Outubro.
- Joseph Redmon, Ali Farhadi. (2018). "YOLOv3: An Incremental Improvement". arXiv. <https://arxiv.org/abs/1804.02767>, Outubro.
- Li, C. *et al.* (2023). "YOLOv6 v3.0: A Full-Scale Reloading". arXiv. <https://arxiv.org/abs/2301.05586>, Outubro.
- Xiaohan Ding *et al.* (2021). "RepVGG: Making VGG-style ConvNets Great Again" . *IEEE*. <https://ieeexplore.ieee.org/document/9577516>, Outubro.
- Wang, C.-Y. *et al.* (2022). "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors". arXiv. <https://arxiv.org/pdf/2207.02696.pdf>, Outubro.
- Minh Pham *et al.* (2022). "Revisiting Self-Distillation". arXiv. <https://arxiv.org/abs/2206.08491>, Outubro.
- Cyganek, B. (2013). "Object Detection and Recognition in Digital Images: Theory and Practice" (1<sup>a</sup> ed.). John Wiley & Sons, Ltd, Outubro.
- Aston Zhang, Zachary C. Lipton, Mu Li, e Alexander J. Smola. (2023). "Dive into Deep Learning". <http://d2l.ai/index.html>, Outubro.
- Jha, A., Tushar, A., Gou, D., & Chowdhury, S. A. (2021). "Real-Time Polyp Detection, Localisation and Segmentation in Colonoscopy Using Deep Learning". <https://www.researchgate.net/publication/349824439>, Outubro.
- T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 2980–2988.
- Krishnendu S , Geetha S , Gopakumar G, 2020, A Review on Polyp Detection and Segmentation in Colonoscopy Images using Deep Learning, International Journal of Engineering Research & Technology (IJERT). 2020. volume 09, edição 10.