

EXPLORANDO O USO DE VISION TRANSFORMER NA CLASSIFICAÇÃO DE LESÕES ANAIS E CERVICAIAS

Ingrid Bromerschenckel
Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Biológicas
Departamento de Computação
PCC146 - Reconhecimento de Padrões - Profa. Andrea Gomes
Programa de Pós-Graduação em Ciência da Computação
Maio/2025



Introdução

- O câncer é a **segunda principal causa de morte global**, sendo responsável por aproximadamente 9,6 milhões de óbitos por ano – e a tendência é que, nos próximos 25 anos, se torne a principal
- O câncer de canal anal (CCA) representa cerca de 4% das neoplasias malignas do trato digestivo inferior
- Embora o CCA apresente baixa incidência entre os tumores do trato digestivo, observa-se um **aumento progressivo de casos**, especialmente entre populações vulneráveis



Introdução - Diagnóstico

- Não existe consenso internacional para a triagem do CCA
- Em população de risco → citologia convencional
 - Pessoas vivendo com HIV;
 - Imunossuprimidos e receptores de transplantes;
 - Indivíduos com histórico de neoplasias ou displasias cervicais e/ou vulvares de alto grau;
 - Indivíduos que praticam sexo anal receptivo com penetração por pênis – prática associada ao maior risco de microlesões na mucosa anal, facilitando a infecção por agentes como o HPV e o HIV.



Imagen 1

Introdução - Triagem recomendada

Grupos-alvo para triagem	Institutos Nacionais de Saúde dos EUA Escritório de Pesquisa em AIDS	Sociedade Brasileira de Oncologia Cirúrgica	Sociedade Portuguesa de Coloproctologia	Sociedade Nacional Francesa de Coloproctologia	Sociedade Internacional de Neoplasia Anal
HSH vivendo com HIV	X	X	X	X	X
HSM vivendo com HIV	X				X
Homens trans vivendo com HIV	X				X
Mulheres trans vivendo com HIV	X				X
Mulheres cis vivendo com HIV	X				X
Mulheres com transplante de órgão sólido > 10 anos			X	X	X
Homens com transplante de órgão sólido > 10 anos					X
Mulheres com histórico de câncer vulvar			X	X	X

HSH: Homens que fazem sexo com homens

HSM: Homens que fazem sexo com mulheres

Fonte: Adaptado de ALBUQUERQUE; FONTES, 2025



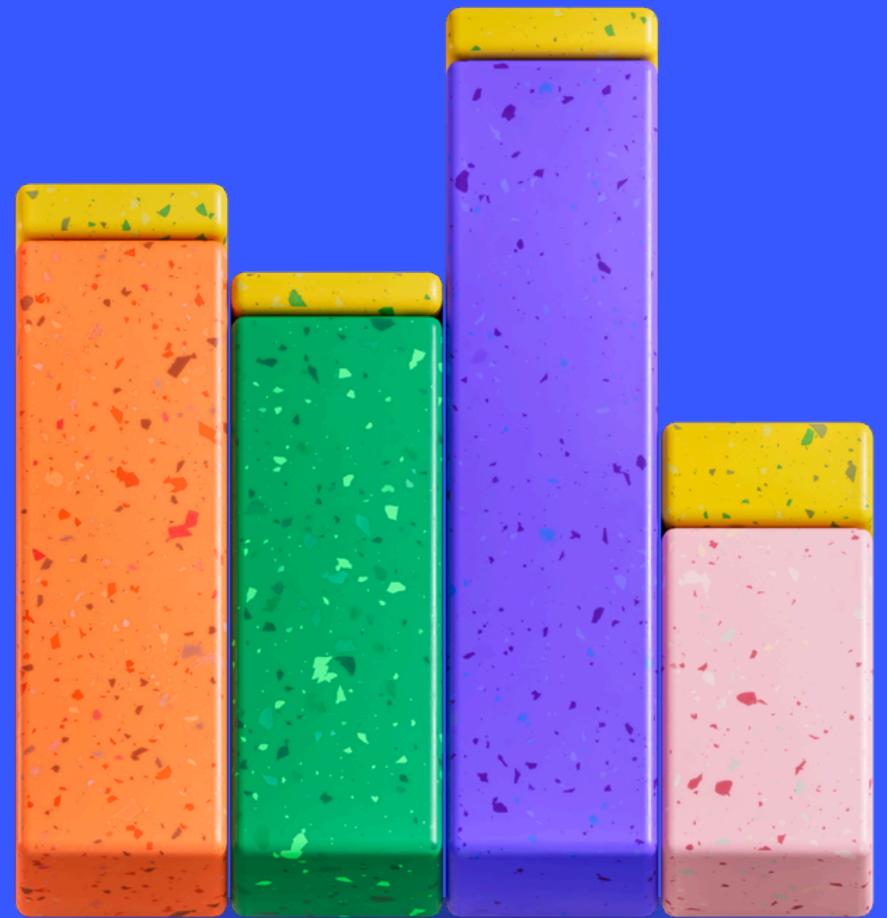
Introdução

- A descoberta de que o HPV, particularmente os sorotipos 16 e 18, é a principal causa desse tipo de câncer, consolidou a compreensão do CCA como uma possível consequência de uma **infecção sexualmente transmissível**
- Esse entendimento reforça seu potencial de prevenção por meio de estratégias de rastreamento baseadas em evidências, vacinação e **diagnóstico precoce**
- Como há uma **relação inversa entre o estágio da doença e a sobrevida**, a detecção precoce surge como uma ferramenta essencial para reduzir a mortalidade e posiciona o CCA como uma preocupação relevante e tratável de saúde pública

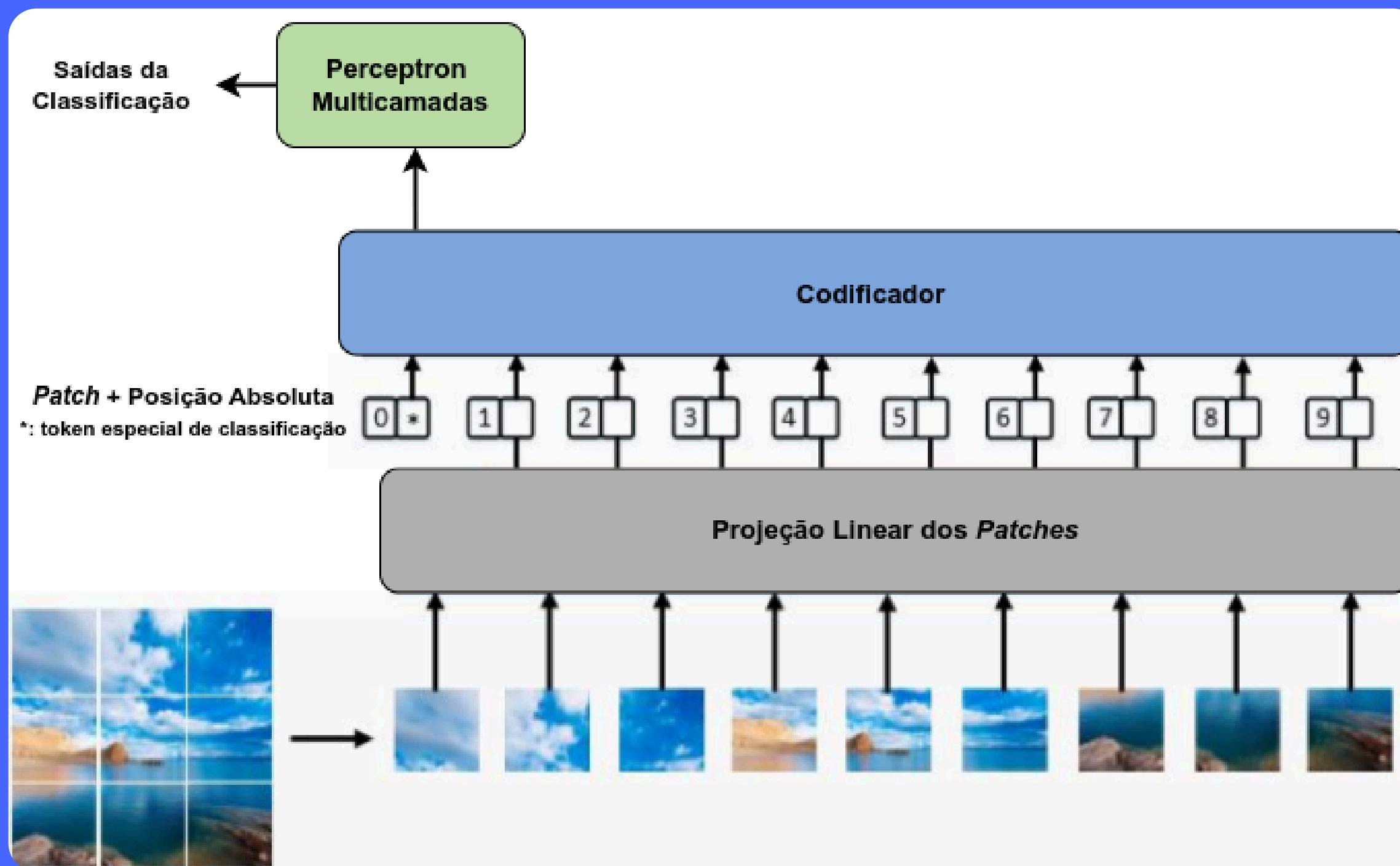


Hipótese

- Sucesso previamente obtido para classificação de lesões celulares de colo de útero em exames de Papanicolaou (citologia convencional), a partir de CNNs
- Publicação de conjunto de dados representativo de classificação e segmentação celular seguindo o protocolo internacional Bethesda
 - CRIC Database (Centro de Reconhecimento e Inspeção de Células)
 - 400 imagens de lâminas multicelulares
 - 11.534 células classificadas e rotuladas
- Histologia semelhante entre células cervicovaginais e anais
- Sensibilidade da citologia convencional equiparada em ambos tipos celulares
- Mesma interpretação de resultados (Bethesda)
- Vision Transformer possui capacidade de processamento de imagens de maneira global (mecanismo de *self-attention*), o que obtém melhor desempenho na classificação celular e evita vieses em estruturas locais, além do grande potencial para transferência de aprendizado



Vision Transformer



O token* agrega as informações globais ao longo das camadas do codificador ("resumo da imagem")

Cada patch é transformado em um vetor de características

Revisão Bibliográfica - Metodologia



Perguntas de Pesquisa

- Como o ViT tem sido empregado nos modelos de classificação de lesões e câncer de colo do útero e anal em lâminas de citologia convencional?
- É possível extender o modelo de classificação de lesões e câncer de colo do útero para lesões e câncer anal?

Fontes

IEEE, PUBMED, SCOPUS, WEB OF SCIENCE

Revisão Bibliográfica - Metodologia

Critérios de Inclusão

- Artigos que estivessem dentro do escopo da proposta, ou seja, que abordassem sobre as células-alvo (cervicais e anais) e empregassem a citologia convencional como método de triagem

Critérios de Exclusão

- Documento Duplicado
- Literatura Cinza
- Menciona superficialmente ViT em células-alvo, não sendo essas células empregadas no desenvolvimento do modelo
- Imagens de outros órgãos (que não cólon, reto, útero)
- Imagens obtidas por outros métodos (que não a citologia convencional)



Revisão Bibliográfica - Metodologia

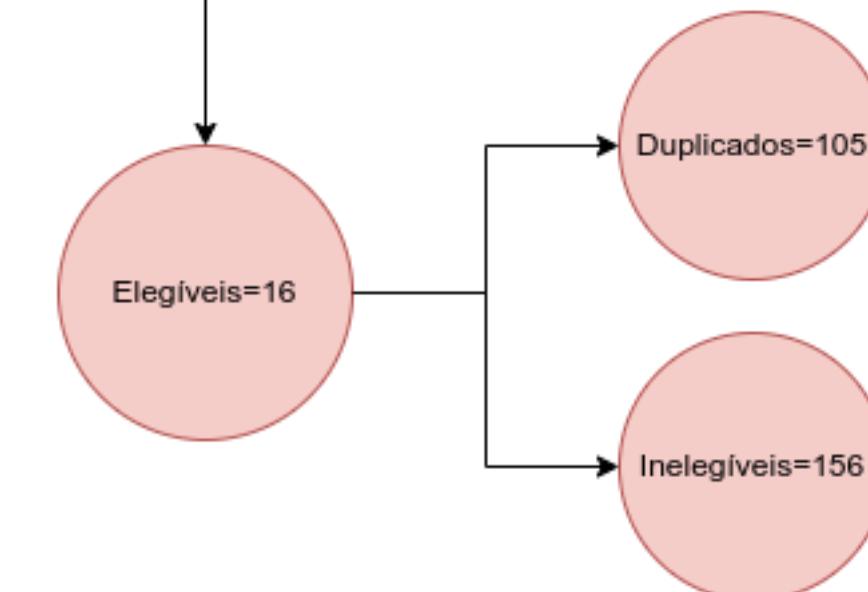
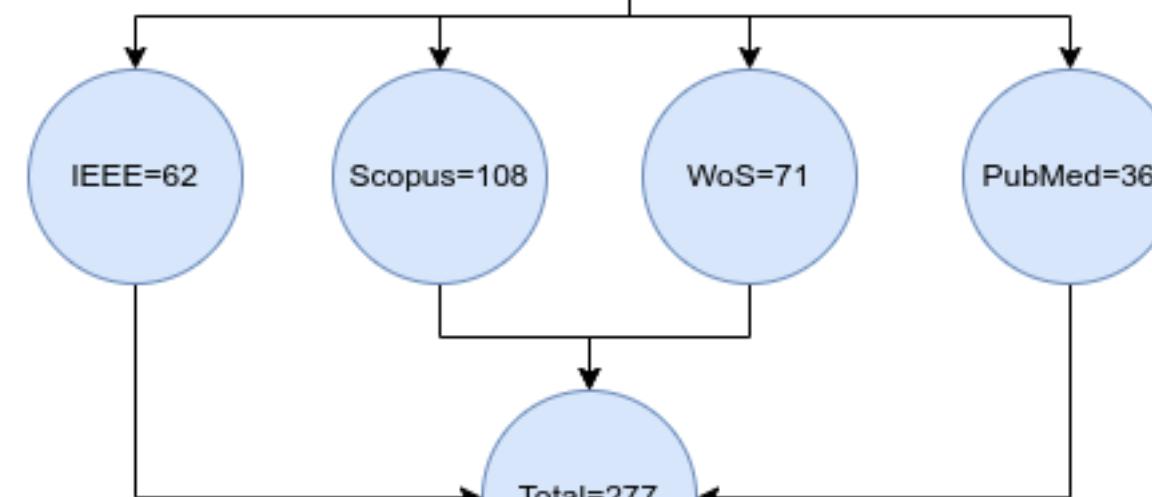


Identificação

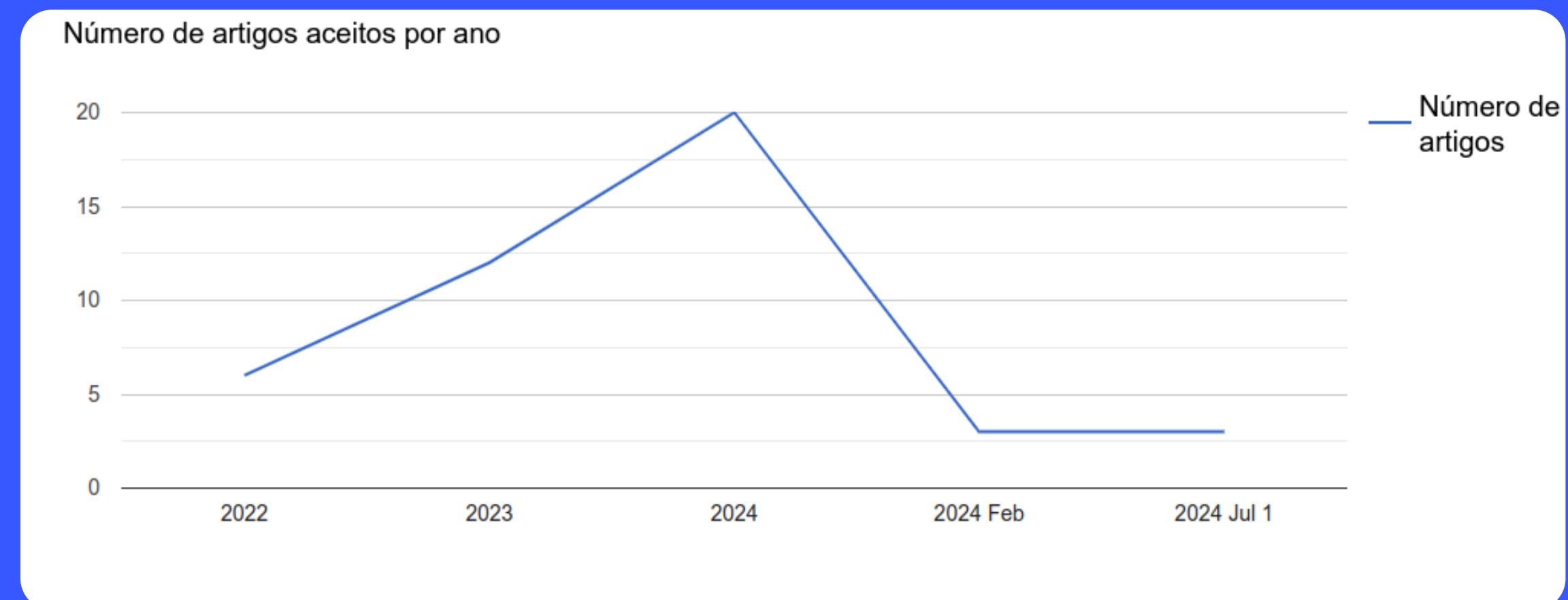
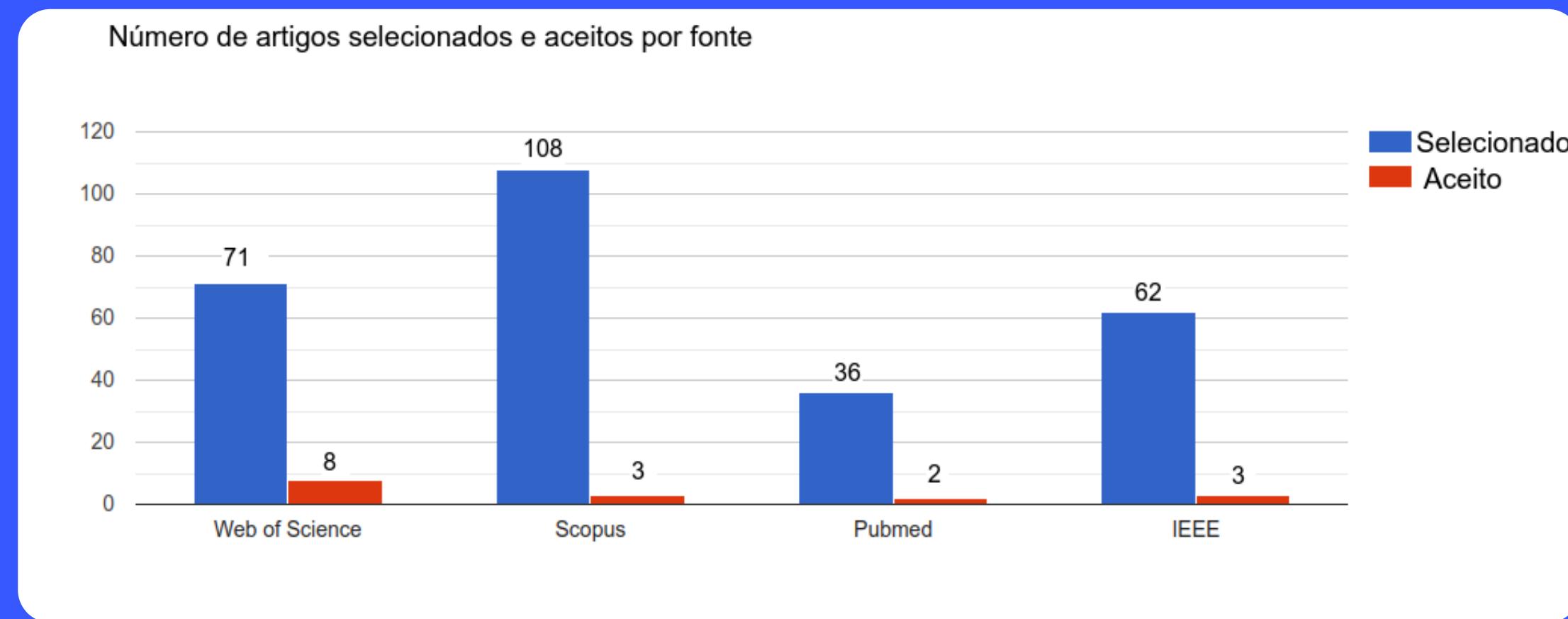
Seleção

Palavras-chave: ("cervical cancer" OR "colorectal cancer"
OR "anal cancer" OR "colon cancer" OR "Papanicolaou" OR
"Pap Smear" OR "PapSmear" OR "conventional cytology")
AND ("vision transformer" OR "visual transformer")

Artigos identificados de acordo com a base de dados



Revisão Bibliográfica - Metodologia



Revisão Bibliográfica - Resultados

- Dos 16 artigos analisados, 9 utilizaram a combinação entre CNN e ViT: a CNN é responsável por extrair características locais detalhadas, como bordas e texturas, enquanto o ViT capta relações globais, contribuindo para uma detecção mais precisa
- 6 artigos utilizaram ViT isolado, ou com outras abordagens que não envolveram CNN
- Não foram identificados estudos que explorassem a transferência de aprendizado para adaptação do modelo a outros tipos celulares. A única abordagem observada envolveu a aplicação de um modelo treinado em células cervicais para a classificação de células sanguíneas



Revisão Bibliográfica - Resultados

- Foram identificados dois estudos que empregaram ViT sobre a base de dados CRIC:
CVM-Cervix: combinação de CNN, ViT e um Perceptron Multicamadas para a classificação final. O método foi treinado em uma base de dados composta pela união dos datasets CRIC e SIPaKMeD, totalizando 8838 imagens divididas em 11 classes. Acurácia de 91,72% no conjunto de teste, com precisão média de 91,80%, revocação de 91,60% e F1-Score de 91,70%, superando modelos tradicionais como DenseNet, ResNet e o ViT sem agregação a outro modelo
Deep Integrated Feature Fusion: combinou interativamente os mapas de características locais (CNN) e globais (ViT) por concatenação, convoluções 1x1 e 3x3, e conexões residuais, promovendo uma fusão mais eficaz entre os dois tipos de informação. O método atingiu acurácia de 85,06%, com precisão de 85,47%, revocação de 85,23% e F1-Score de 85,12% na CRIC para as seis classes



Modelo: Pré-processamento dos dados

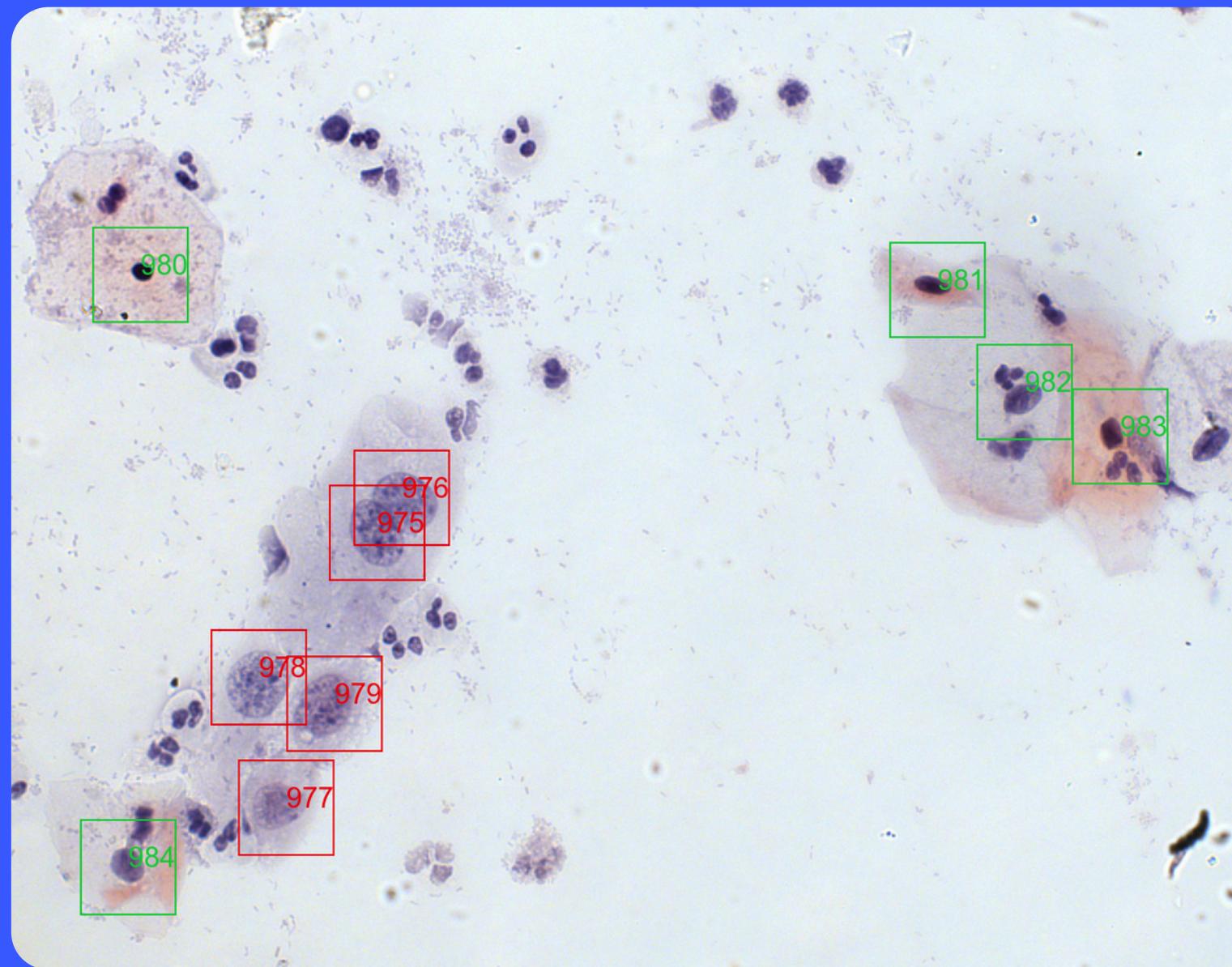
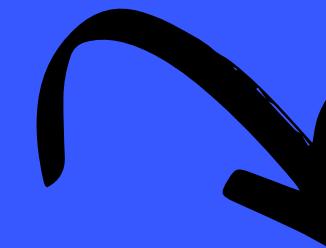
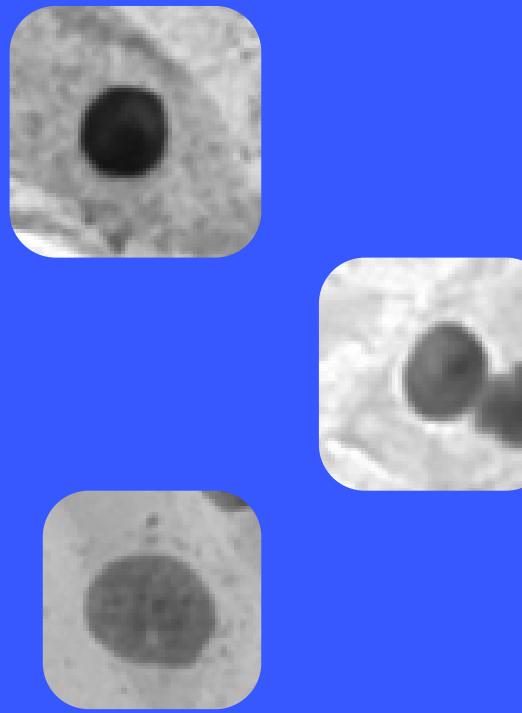


Figura: Imagem nº 383 da base de dados CRIC -
lâmina de microscópio do colo do útero rotulada



Bounding Box 70x70 pixels
Escala de cinza



70% treino
15% validação
15% teste

4.755 classificadas como positivas para lesão celular e
6.779 como negativas
425 células foram descartadas por estarem fora
dos limites estabelecidos

Modelo: Extração de Atributos Texturais (27)

Método	Descrição	Nº de Atributos
Gray Level Co-occurrence Matrix (GLCM)	Matriz de coocorrência de níveis de cinza	4 (Contraste, Correlação, Energia, Homogeneidade)
Local Binary Pattern (LBP)	Padrão Binário Local – vetor de atributos texturais da imagem	10 bins - frequência de cada padrão de textura
Haralick	Estatísticas de textura calculados pela biblioteca mahotas	13 atributos



Modelo: Escolha dos algoritmos

De acordo com os melhores resultados obtidos na revisão sistemática:

1. ViT adaptado, utilizado no estudo com melhor desempenho - MaxCerVixT
2. ViT combinado a CNNs como EfficientNetBO

Classificação de acordo com o Sistema Bethesda, gradualmente:

- 2 classes: positivo e negativo para lesão celular
- 3 classes: negativo; atípica indefinida (limítrofe): (2) ASC-US, (3) ASC-H; lesão confirmada: (4) LSIL, (5) HSIL, (6) CEC
- 6 classes (completa):
 - (1) Negativo para Lesão Intraepitelial ou Malignidade (NILM)
 - (2) Células Escamosas Atípicas de Significado Indeterminado (ASC-US)
 - (3) Células Escamosas Atípicas, não podendo excluir lesão de alto grau (ASC-H)
 - (4) Lesão Intraepitelial Escamosa de Baixo Grau (LSIL)
 - (5) Lesão Intraepitelial Escamosa de Alto Grau (HSIL)
 - (6) Carcinoma Espinocelular (CEC)

REFERÊNCIAS

- Imagem 1: Hospital Alemão Oswaldo Cruz. Disponível em: <https://www.hospitaloswaldocruz.org.br/centro-especializado/oncologia/cancer-de-anus/>; Acesso em: 01 maio 2025.
- ABINAYA; K.; SIVAKUMAR; B.. A Deep Learning-Based Approach for Cervical Cancer Classification Using 3D CNN and Vision Transformer. *Journal of Imaging Informatics in Medicine*, v. 37, n. , p. 280--296, , 2024.
- AL-HAMADANI; A., M. N.; FADHEL; A., M.; ALZUBAIDI; L.; HARANGI; B.. Reinforcement Learning Algorithms and Applications in Healthcare and Robotics: A Comprehensive and Systematic Review. *Sensors*, v. 24, n. 8, p. 2461, , 2024.
- ALBUQUERQUE, Andreia; FONTES, Filipa. Recent Guidelines on Anal Cancer Screening: A Systematic Review. *Journal of Lower Genital Tract Disease*, v. 29, n. 2, p. 180--185, April, 2025.
- ALMOHIMEED; A.; SHEHATA; M.; AL., et.. ViT-PSO-SVM: Cervical Cancer Prediction Based on Integrating Vision Transformer with Particle Swarm Optimization and Support Vector Machine. *Bioengineering*, v. 11, n. 7, p. 729, , 2024.
- AMBROSIO; R., M.; NICCOLAI; E.; AL., et.. Immune landscape and oncobiota in HPV-Associated Colorectal Cancer: an explorative study. *Clinical and Experimental Medicine*, v. 23, n. 8, p. 5101--5112, , 2023.
- AYANA; G.; BARKI; H.; CHOE; W., S.. Pathological Insights: Enhanced Vision Transformers for the Early Detection of Colorectal Cancer. *Cancers*, v. 16, n. 7, p. 1441, , 2024.
- BERNAL; MONSERRAT; BONONO; CATHERINE-ROSELYNE; EL-HALABI; SILVIA; MARTÍ; JULIETA; MWAKA; D, Amos; VASSILAKOS; PIERRE; JERONIMO; JOSE; G{"U}LMEZOGLU; METIN, A. WHO recommendations and good practice statements on screening for cervical cancer. *International Journal of Gynecology \& Obstetrics*, v. 160, n. S1, p. 5--17, , 2023.
- BILLER; H., Leah; SCHRAG; DEBORAH. Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review. *JAMA*, v. 325, n. 7, p. 669-685, 02, 2021.
- BRAY, F.; FERLAY, J.; SOERJOMATARAM, I.; AL., et. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, v. 68, n. 6, p. 394--424, , 2018.
- CAETANO, Rosângela; VIANNA, Cid Manso de Mello; THULER, Luiz Cláudio Santos; GIRIANELLI, Vania Reis. Custo-efetividade no diagnóstico precoce do câncer de colo uterino no Brasil. *Physis: Revista de Saúde Coletiva*, v. 16, n. 1, p. XX-XX, , 2006.
- CAMPOS; SOUSA, Irving; CARVALHO; BATISTA, Lorena Rocha. Dificuldades para o rastreio do câncer do colo de útero no Brasil: uma revisão de literatura. *Cadernos de Ensino e Pesquisa em Saúde*, v. 4, n. 2, p. 20--32, , 2024.
- CHANDWANI; V.; SARASWATHI; K.; RAO; B., G.; SIVASTAVA; V.. Liquid Based Cytology versus Conventional Cytology for Evaluation of Cervical PAP Smear. *Journal of Pharmaceutical Research International*, v. 33, n. 45B, p. 27--33, , 2021.
- CHEN; RUI; LU; MING; ZHANG; JIAYUAN; ZHAN; XI; ZHAO; YICHENG; WANG; JIAN; ZHU; YUCHENG; WANG; LI. TransPath: Transformer-based self-supervised learning for pathology image analysis. *Medical Image Analysis*, v. 77, n. , p. 102371, , 2022.
- CHEN; S.; LU; S.; WANG; S.; NI; Y.; ZHANG; Y.. Shifted window vision transformer for blood cell classification. *Electronics*, v. 12, n. 11, p. 2442, , 2023.
- CLARO, I. B.; OTHERS. Análise dos motivos de insatisfatoriedade dos exames histopatológicos do colo do útero no Sistema Único de Saúde, Brasil, 2014 a 2017. *Revista Brasileira de Cancerologia*, v. 67, n. 3, p. e-081299, , 2021.
- CLÍNICA, Sociedade Brasileira de Citologia. SBCC defende citologia como método de prevenção do câncer do colo do útero de melhor custo-benefício para o Brasil. **, 2023. Disponível em: <>..
- DAI; Y.; GAO; Y.; LIU; F.. Transmed: Transformers advance multi-modal medical image classification. *Diagnostics*, v. 11, n. 8, p. 1384, , 2021.
- DINIZ; D{\'}E}BORA; REZENDE; MARIANA; CAMPOS; GOMES, Andrea; SOUZA; JF, Marcone. A Deep Learning Ensemble Method to Assist Cytopathologists in Pap Test Image Classification. *Journal of Imaging*, v. 7, n. 7, p. 111, , 2021.
- DOSOVITSKIY; ALEXEY; BEYER; LUCAS; KOLESNIKOV; ALEXANDER; WEISSENBORN; DIRK; ZHAI; XIAOHUA; UNTERTHINER; THOMAS; DEHGHANI; MOSTAFYA; MINDERER; MATTHIAS; HEIGOLD; GEORG; GELLY; SYLVAIN; USZKOREIT; JAKOB; HOULSBY; NEIL. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, v. , n. , p. , , 2020
- Neurocomputing*, v. 188, n. , p. 68--76, , 2020

REFERÊNCIAS

- DYER; F., Clare E.; JIN; FENGYI; HILLMAN; J., Richard; NYITRAY; G., Alan; ROBERTS; M., Jennifer; LAW; CARMELLA; GRULICH; E., Andrew; POYNTEN; MARY, I.. Self-collected versus clinician-collected anal swabs for anal cancer screening: A systematic review and meta-analysis. *International Journal of Cancer*, v. 156, n. 1, p. 79--90, , 2025.
- EMARA; M., H.; EL-SHAFAI; W.; AL., et.. Cervical Cancer Detection: A Comprehensive Evaluation of CNN Models, Vision Transformer Approaches, and Fusion Strategies. *IEEE Access*, v. , n. , p. , , 2024.
- FANG; M.; FU; M.; LIAO; B.; AL., et.. Deep integrated fusion of local and global features for cervical cell classification. *Computers in Biology and Medicine*, v. 171, n. , p. 108153, , 2024.
- HEMALATHA; K.; VETRISELVI; V.; DHANDAPANI; M.; GLADYS; A., A.. CervixFuzzyFusion for cervical cancer cell image classification. *Biomedical Signal Processing and Control*, v. 85, n. , p. 104920, , 2023.
- HOPP; M., A.; PUZYRENKO; A.; GIORGADZE; T.. Comprehensive Review of Anal Cytology. *Acta Cytologica*, v. 67, n. 2, p. 185--194, , 2023.
- HOU; XIN; SHEN; GUANGYANG; ZHOU; LIQIANG; LI; YINUO; WANG; TIAN; MA; XIANGYI. Artificial Intelligence in Cervical Cancer Screening and Diagnosis. *Frontiers in Oncology*, v. 12, n. , p. 851367, , 2022.
- JIANG; X.; WANG; S.; ZHANG; Y.. Vision transformer promotes cancer diagnosis: A comprehensive review. *Expert Systems with Applications*, v. 252, n. Part A, p. 124113, , 2024.
- KAVITHA; S., M.; GARGADARAN; P.; AL., et.. Deep Neural Network Models for Colon Cancer Screening. *Cancers (Basel)*, v. 14, n. 15, p. 3707, , 2022.
- KHAN; S.; NASEER; M.; AL., et.. Transformers in vision: A survey. *ACM Computing Surveys (CSUR)*, v. 54, n. 10s, p. Article 200, 1--41, , 2022.
- KHOWAJA; A.; ZOU; B.; KUI; X.. Enhancing cervical cancer diagnosis: Integrated attention-transformer system with weakly supervised learning. *Image and Vision Computing*, v. 149, n. , p. 105193, , 2024.
- KHOWAJA; A.; ZOU; B.; XIAOYAN; K.. Cervix Visionator ELM: A Novel Approach to Early Detection of Cervical Cancer. **, v. , n. , p. , December, 2023.
- KLAS; V., J.; ROTHENBERGER; A., D.; WONG; D., W.; MADOFF; D., R.. Malignant tumors of the anal canal: the spectrum of disease, treatment, and outcomes. *Cancer*, v. 85, n. 8, p. 1686--1693, , 1999.
- LECUN; YANN; BENGIO; YOSHUA; HINTON; GEOFFREY. Deep learning. *Nature*, v. 521, n. 7553, p. 436--444, , 2015.
- LI; M.; QUE; N.; ZHANG; J.; DU; P.; DAI; Y.. VTCNet: A Feature Fusion DL Model Based on CNN and ViT for the Classification of Cervical Cells. *International Journal of Imaging Systems and Technology*, v. 34, n. 5, p. , , 2024.
- LI; T.; ZHANG; Z.; PEI; L.; GAN; Y.. HashFormer: Vision transformer-based deep hashing for image retrieval. *IEEE Signal Processing Letters*, v. 29, n. , p. 827-831, , 2022.
- LI; TAO; ZHANG; XIAODONG; YANG; LEI; WANG; RUI; HUANG; XIAOLEI. Transformer-based architecture for classification of cervical cell images. *Computer Methods and Programs in Biomedicine*, v. 220, n. , p. 106828, , 2022.
- LIU; W.; LI; C.; XU; N.; AL., et.. CVM-Cervix: A hybrid cervical Pap-smear image classification framework using CNN, visual transformer, and multilayer perceptron. *Pattern Recognition*, v. 130, n. , p. 108829, , 2022.
- MANICKADEVI; S., M.; HEMALATHA; V., S.; THANGAMANI; M.. Comparative study of the cytologic diagnosis, specimen adequacy, sensitivity, and cost effectiveness of liquid-based cytology with that of conventional PAP tests. *International Journal of Reproduction, Contraception, Obstetrics and Gynecology*, v. 11, n. 2, p. 474--478, , 2022.
- MAURYA; R.; PANDEY; N., N.; DUTTA; K., M.. VisionCervix: Papanicolaou cervical smears classification using novel CNN-Vision ensemble approach. *Biomedical Signal Processing and Control*, v. 79, n. Part 2, p. 104156, , 2023.
- PACAL; I.. MaxCerVixT: A novel lightweight vision transformer-based Approach for precise cervical cancer detection. *Knowledge-Based Systems*, v. 289, n. , p. 111482, , 2024.
- PHUNG; VAN; RHEE; EUN. A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets. *Applied Sciences*, v. 9, n. , p. 4500, 10, 2019.
- PLISSITI; E., M.; NIKOU; C.; CHARCHANTI; A.. Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering. *IEEE Transactions on Information Technology in Biomedicine*, v. 15, n. 2, p. 233--241, , 2011. *Neurocomputing*, v. 188, n. , p. 68--76, , 2020.

REFERÊNCIAS

- REZENDE; T.; Mariana; SILVA; RANIERE; BERNARDO; O., Fagner de; TOBIAS; G., Alessandra H.; OLIVEIRA; C., Paulo H.; MACHADO; M., Tales; COSTA; S., Caio; MEDEIROS; S., Fatima N.; USHIZIMA; M., Daniela; CARNEIRO; M., Claudia; BIANCHI; C., Andrea G.. Cric searchable image database as a public platform for conventional pap smear cytology data. *Scientific Data*, v. 8, n. 1, p. 151, , 2021.
- ROBB, Bruce W.; MUTCHE, Matthew G.. Epidermoid Carcinoma of the Anal Canal. *Clinics in Colon and Rectal Surgery*, v. 19, n. 2, p. 108--115, , 2006.
- RYAN, David P.; COMPTON, Carolyn C.; MAYER, Robert J.. Carcinoma of the Anal Canal. *The New England Journal of Medicine*, v. 342, n. 11, p. 792--800, , 2000.
- SHAHINI; E.; SINAGRA; E.; VITELLO; A.; AL., et.. Factors affecting the quality of bowel preparation for colonoscopy in hard-to-prepare patients: Evidence from the literature. *World Journal of Gastroenterology*, v. 29, n. 11, p. 1685--1707, , 2023.
- SHAMSHAD; F.; KHAN; S.; AL., et.. Transformers in medical imaging: A survey. *Medical Image Analysis*, v. 88, n. , p. 102802, , 2023.
- SKEREDE; J., O.; ET.AL. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *The Lancet*, v. 395, n. 10221, p. 350--360, , 2020.
- SOUZA; PIRES, Elissandra Pereira; MUSTAFA; MOACIR, Marcio; SENA; BARBOSA, Amanda. Vantagens e desvantagens da citologia convencional e da citologia em meio líquido na prática clínica: uma revisão integrativa. *Research, Society and Development*, v. 10, n. 14, p. e245101422350, , 2021.
- STEGMÜLLER; T.; ABBET; C.; AL., et.. Self-supervised learning-based cervical cytology for the triage of HPV-positive women in resource-limited settings and low-data regime. *Computers in Biology and Medicine*, v. 169, n. , p. 107809, , 2024.
- STEWART, David B.; GAERTNER, Wolfgang B.; GLASGOW, Sean C.; HERZIG, Daniel O.; FEINGOLD, Daniel; STEELE, Scott R.. The American Society of Colon and Rectal Surgeons Clinical Practice Guidelines for Anal Squamous Cell Cancers (Revised 2018). *Diseases of the Colon \& Rectum*, v. 61, n. 7, p. 755--774, , 2018.
- TANUM, Gunnar; TVEIT, Kjell M.; KARLSEN, Karl Otto. Diagnosis of Anal Carcinoma – Doctor's Finger Still the Best. *Oncology*, v. 48, n. 5, p. 383--386, , 1991.
- WU; B.; XU; C.; AL., et.. Visual transformers: Token-based image representation and processing for computer vision. *arXiv preprint arXiv:2006.03677*, v. , n. , p. , , 2020.
- XU; YIZHE; MO; TIANMENG; FENG; QIANJIN; ZHONG; PEILIN; LAI; MAODE; CHANG; I-CHAO, Eric. Deep learning of feature representation with multiple instance learning for medical image analysis. *Neurocomputing*, v. 188, n. , p. 68--76, , 2020.
- ZHANG; LINYUAN; LU; LISHENG; NOGUES; IRENE; SUMMERS; M, Ronald; LIU; SHAOXIONG; YAO; JIANHUA. DeepPap: deep convolutional networks for cervical cell classification. *IEEE Journal of Biomedical and Health Informatics*, v. 21, n. 6, p. 1633--1643, , 2017.
- ZHANG; Y.; WANG; J.; GORRIZ; M., J.; WANG; S.. Deep learning and vision transformer for medical image analysis. *Journal of Imaging*, v. 9, n. 7, p. 147, , 2023.