



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

João Gustavo Atkinson Amorim

Aprendizagem profunda em imagens citológicas: identificação e caracterização de objetos corados com a técnica AgNOR

Florianópolis
2022

João Gustavo Atkinson Amorim

Aprendizagem profunda em imagens citológicas: identificação e caracterização de objetos corados com a técnica AgNOR

Dissertação submetida ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Santa Catarina para a obtenção do título de mestre em Ciência da Computação.

Orientador: Prof. Aldo von Wangenheim, Dr.

Florianópolis

2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Amorim, João Gustavo Atkinson

Aprendizagem profunda em imagens citológicas :
identificação e caracterização de objetos corados com a
técnica AgNOR / João Gustavo Atkinson Amorim ; orientador,
Aldo von Wangenheim, 2022.

62 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico, Programa de Pós-Graduação em
Ciência da Computação, Florianópolis, 2022.

Inclui referências.

1. Ciência da Computação. 2. Deep Learning. 3. Visão
Computacional. 4. AgNOR. 5. Citologia. I. Wangenheim, Aldo
von. II. Universidade Federal de Santa Catarina. Programa
de Pós-Graduação em Ciência da Computação. III. Título.

João Gustavo Atkinson Amorim

Aprendizagem profunda em imagens citológicas: identificação e caracterização de objetos corados com a técnica AgNOR

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Mauro Roisenberg, Dr.
Universidade Federal de Santa Catarina

Prof.(a) Fabiana Botelho de Miranda Onofre, Dr(a).
Universidade Federal de Santa Catarina

Prof. Jonata Tyska Carvalho, Dr.
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Ciência da Computação.

Coordenação do Programa de
Pós-Graduação

Prof. Aldo von Wangenheim, Dr.
Orientador

Florianópolis, 2022.

AGRADECIMENTOS

Gostaria de agradecer a minha família pelo apoio. Assim como para todo meu grupo de pesquisa pelas discussões, insights e apoio ao longo desta jornada.

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (Fapesc).

RESUMO

O câncer do colo do útero é o quarto tipo de câncer mais comum e mortal nas mulheres nos últimos anos. Uma maneira de diagnosticar este tipo de câncer é através da citologia, que se mostra uma técnica viável e capaz de detectar precocemente o câncer. Porém, a análise e/ou diagnóstico deste tipo de exame é feito de maneira manual com profissionais treinados, e isto é propenso a erros humanos por diversos fatores. Na citologia é necessário o uso de um corante para destacar as células e suas características, sendo o Papanicolau a técnica padrão para o câncer do colo do útero, porém em alguns casos métodos coadjuvantes são necessários para um diagnóstico completo. Neste trabalho, amostras cervicais coradas com a técnica coadjuvante AgNOR foi utilizada. A partir destas amostras, imagens são capturadas para realização da identificação e caracterização das células utilizando modelos de *Deep Learning* do estado da arte. O diagnóstico destas células ocorre através da contagem/detecção dos núcleos e NORs — Regiões Organizadoras Nucleolares. Os NORs são estruturas celulares muito pequenas presentes nos núcleos, para detectá-los foi utilizado o método de segmentação semântica. Foi comparado modelos do estado da arte (HRNet e SegFormer) com a rede U-net tradicional usando ResNets como backbone. Os resultados apresentaram o modelo *transformer*, SegFormer, como o melhor modelo para segmentar núcleos. Sendo a HRNet o melhor modelo para segmentar os NORs. Foi utilizado como pós-processamento um algoritmo de componentes conexas com uma filtragem por tamanho, para extrair as instâncias dos objetos a partir da saída destes modelos. A HRNet mostrou-se mais adequado para segmentação de imagens AgNOR na totalidade, com 0,85 de precisão, 0,85 de recall e 0,85 F-score para detecção dos núcleos, e 0,7 de precisão, 0,55 de recall e 0,62 F-score na detecção de NORs.

Palavras chaves: Deep Learning; Visão Computacional; AgNOR; Citologia.

ABSTRACT

Cervical cancer is the fourth most common and deadly cancer in women in recent years. One way to diagnose this type of cancer is through cytology, which is a viable technique capable of detecting cancer in its early stages. However, the analysis and/or diagnosis of this type of exam is done manually with trained professionals, and this is prone to human errors due to several factors. In cytology, it is necessary to use a stain to highlight the cells and their characteristics, the Pap-smear being the standard technique for cervical cancer, but in some cases supporting methods are necessary for a complete diagnosis. In this work, has been utilized cervical samples stained with the AgNOR, an adjuvant technique. From these samples, images are captured to perform cell identification and characterization using state-of-the-art *Deep Learning* models. The diagnosis of these cells occurs through the counting/detection of nuclei and NORs - Nucleolar Organizing Regions. The NORs are very small objects present in the nuclei, to detect them has been employed semantic segmentation methods. Comparing state-of-the-art models (HRNet and SegFormer) with the traditional U-net network using ResNets as a backbone. The results showed the transformer model, SegFormer, as the best model to segment the nucleus. HRNet is the best model to segment NORs. An algorithm of connected components with filtering by area was used as post-processing to extract the object instances from the output of these models. HRNet showed to be more suitable for segmenting AgNOR images in their entirety, with 0.85 precision, 0.85 recall, and 0.85 F-score for nuclei detection, and 0.7 precision, 0.55 recall, and 0.62 F-score in the detection of NORs.

Keywords: Deep Learning; Computer vision; AgNOR; Cytology.

LISTA DE FIGURAS

Figura 1 – Informações sobre a distribuição de novos casos e mortes por país para o ano de 2020.	16
Figura 2 – Exemplos de diferentes campos de uma lâmina capturadas com uma ampliação x40 (a) sem coloração; e (b) Argyrophilic Nucleolar Organizer Regions (AgNOR)	19
Figura 3 – Etapas macros para aquisição das imagens de amostras cervicais .	24
Figura 4 – Etapas macros e itens para produção do conjunto de dados	25
Figura 5 – Etapas macros de preparo do conjunto de dados, treinamento de modelos e avaliação de resultados	25
Figura 6 – Categorias do CCAgT dataset	27
Figura 7 – Organização geral dos experimentos	31
Figura 8 – Exemplo de separação de uma imagem em 16 partes.	32
Figura 9 – Diagrama sobre do estágio de treino.	35
Figura 10 – Diagrama sobre o pós-processamento.	36
Figura 11 – Diagrama da separação dos dados	37
Figura 12 – Exemplo de resultado obtido pela HRNet-32 após pós-processamento ao nível de instâncias e de pixels	41

LISTA DE TABELAS

Tabela 1 – Estatísticas gerais da primeira versão do CCAgT dataset	28
Tabela 2 – Estatísticas gerais da segunda versão do CCAgT dataset	29
Tabela 3 – Estatísticas dos subdatasets	33
Tabela 4 – Resultado ao nível de instâncias de cada modelo treinado com todas as categorias (SA)	39
Tabela 5 – Média dos resultados ao nível de instâncias de cada modelo treinado com todas as categorias (SA)	40
Tabela 6 – Resultado ao nível de instâncias de cada modelo treinado com categorias principais e núcleos fora de foco (SB)	42
Tabela 7 – Resultado ao nível de pixels do modelo HRNet-32 no SA e SB . . .	42
Tabela 8 – Resultado ao nível de instâncias do modelo HRNet-32 no SA e SB na WSI de teste.	43

LISTA DE ABREVIATURAS E SIGLAS

AgNOR	Argyrophilic Nucleolar Organizer Regions
CCAgT	Cervical Cells with AgNOR Stain Technique
CNN	Convolutional Neural Network
DL	Deep Learning
DSC	Dice Similarity Coefficient
FPN	Feature Pyramid Networks
HPV	Human Papillomavirus
IoU	Intersection Over Union
ML	Machine Learning
NORs	Nucleolar Organizer Regions
OMS	Organização Mundial da Saúde
RSL	revisão sistemática de literatura
SOTA	state-of-the-art
VC	Visão Computacional
WSI	Whole Slide Images

SUMÁRIO

1	INTRODUÇÃO	12
1.1	CONTEXTUALIZAÇÃO DO PROBLEMA	12
1.2	OBJETIVOS	13
1.2.1	Objetivos Específicos	14
1.3	JUSTIFICATIVA	14
1.4	ESTRUTURA DA DISSERTAÇÃO	15
2	REVISÃO DA LITERATURA E FUNDAMENTAÇÃO TEÓRICA	18
2.1	CITOLOGIA	18
2.1.1	AgNOR	18
2.2	VISÃO COMPUTACIONAL E DEEP LEARNING	20
2.3	VISÃO COMPUTACIONAL APLICADA A CITOLOGIA	21
3	MATERIAIS E MÉTODOS	24
3.1	ELABORAÇÃO DO CONJUNTO DE DADOS	26
3.1.1	Versão 1	28
3.1.2	Versão 2	28
3.2	MODELOS	29
3.3	CONFIGURAÇÃO DOS EXPERIMENTOS	30
3.3.1	Sub-datasets	30
3.3.2	Treinamento	33
3.3.3	Pós-processamento	35
3.4	AVALIAÇÃO DE RESULTADOS	36
4	RESULTADOS E DISCUSSÃO	39
4.1	DISCUSSÃO	42
5	CONCLUSÃO	45
5.1	TRABALHOS FUTUROS	46
	REFERÊNCIAS	47
	APÊNDICE A – ARTIGOS PUBLICADOS	59
A.1	ARTIGOS COMO AUTOR PRINCIPAL	59
A.1.1	Relatório técnico — INCoD - UFSC	59
A.1.2	IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS 2020)	59
A.1.3	XVII Congresso Brasileiro de Informática em Saúde (CBIS 2020)	59
A.1.4	Computerized Medical Imaging and Graphics	60
A.1.5	XIII Computer on the Beach (COTB 2022)	60
A.1.6	Social Science Research Network - SSRN	60
A.2	CONTRIBUIÇÃO EM OUTRAS PUBLICAÇÕES	61

A.2.1	IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS 2020)	61
A.2.2	XVII Congresso Brasileiro de Informática em Saúde (CBIS 2020)	61
A.2.3	SN Computer Science	61
A.2.4	medRxiv - Cold Spring Harbor Laboratory	62
A.2.5	Social Science Research Network - SSRN	62
A.3	PREMIAÇÕES	62
A.3.1	Terceiro colocado no <i>Outdoor Semantic Segmentation Challenge</i> da <i>43rd DAGM German Conference on Pattern Recognition</i> . . .	62

1 INTRODUÇÃO

Visão Computacional (VC) tem aplicações interdisciplinares, buscando através dos computadores processar imagens, identificar e entender objetos, conceitos, ações, entre outros itens. Algoritmos e/ou métodos de VC tendem, por muitas vezes, solucionar problemas de maneiras similares a visão humana. Estes algoritmos e/ou métodos são para as áreas de sensoriamento remoto, segurança, robótica, biomedicina, entre outras (GONZALEZ; WOODS, 2008).

Entre as aplicações encontradas na biomedicina, há a análise automatizada de imagens: de raios-X para as mais diversas partes do corpo, por exemplo, em exames de angiografia e tomografia axial computadorizada; geradas por raios gama, para escaneamento ósseo; de ressonância magnética (do inglês, magnetic resonance imaging (MRI)) que permite verificar a anatomia e os processos fisiológicos do corpo; de microscópio eletrônicos, permitindo analisar diferentes tecidos, por exemplo, provenientes de exames de histologia ou de citologia (GONZALEZ; WOODS, 2008).

A análise automatizada de células é um campo explorado desde os anos 60, mas sem um método genérico que possa ser aplicado a diferentes categorias de células (MEIJERING, 2012). Até os anos 90, a detecção de células em imagens era realizada principalmente por abordagens baseadas em limiares, e nas últimas décadas têm sido mais frequentemente explorados o uso de outras características, como a morfologia, e também métodos baseados em regiões. Na última década, com os avanços dos computadores e algoritmos, as abordagens de aprendizagem profunda (do inglês, Deep Learning (DL)) na análise de imagens de células começaram a mostrar o protagonismo (MEIJERING, 2012; MATIAS *et al.*, 2021a).

1.1 CONTEXTUALIZAÇÃO DO PROBLEMA

Um grande problema de saúde pública, que pode ser melhorado com a análise automatizada de imagens celulares, é o câncer do colo do útero. O câncer do colo do útero (também chamado de câncer cervical) é o quarto tipo de câncer mais comum e mortal nas mulheres no mundo nos últimos anos (SUNG *et al.*, 2021). Isto também é verdade no Brasil, em 2020, ocorreram 17743 novos casos que representa 6,1% referente a novos casos de câncer em mulheres de todas as idades (SUNG *et al.*, 2021). A causa mais comum é o Papilomavírus Humano (do inglês, Human Papillomavirus (HPV)), onde outros cofatores de risco são o vírus da imunodeficiência humana (HIV), tabagismo e aumento da paridade, mas outros fatores são necessários para o desenvolvimento do câncer (LEI *et al.*, 2020; STELZLE *et al.*, 2021).

As Regiões Organizadoras de Nucléolo (do inglês, *Nucleolar Organizer Regions (NORs)*) são segmentos de DNA (ácido desoxirribonucleico) que contém todos os elementos necessários e essenciais para sintetizar o RNAr (ácido ribonucleico ri-

bossômico), e podem ser destacados quando corados com prata, chamado Regiões organizadoras nucleolares argirofílicas (do inglês, (*AgNOR*)). Como o RNA é o principal local para a síntese de proteínas, o número e o tamanho dos NORs em um núcleo celular podem representar a atividade do núcleo. A transformação de células saudáveis para malignas é caracterizada pelo aumento da síntese proteica, conseqüentemente, a quantidade de NORs pode ser considerada um marcador para a quantificação da proliferação celular, diferenciação e transformação maligna (EROZ, 2015).

Em alguns estudos, como em Miranda Onofre *et al.* (2008) e em Pedrini *et al.* (2020), primeiramente foram contadas as agregações de pontos corados com prata ou nucléolos parcialmente desagregados (clusters) tratados como uma estrutura. Segundo, pontos individuais (satélites) fora dos aglomerados de estruturas coradas com prata foram contados. Terceiro, os clusters e satélites foram contados juntos para obter a contagem total de AgNORs. A avaliação das lâminas coradas com AgNOR geralmente tem que ser feita manualmente, acarretando alguns problemas como, variação nas percepções visuais, demora para o diagnóstico, entre outros, tornando o processo propenso a erros humanos (LI *et al.*, 2017).

A citologia com a técnica de AgNOR, pode oferecer opções seguras de auxílio ao diagnóstico em casos duvidosos ou suspeitos para células malignas. Para tornar o processo menos propenso a erros humanos, o uso de análises automatizadas com métodos DL pode ser explorado. Como mostrado em Matias *et al.* (2021a), modelos para segmentação e detecção de objetos estão sendo amplamente explorados nos últimos anos para análise de imagens citológicas.

Este trabalho visa analisar métodos de DL para a segmentação e/ou detecção de AgNORs para fins de auxílio ao diagnóstico e prognóstico de tumores em imagens citológicas. As imagens utilizadas são campos extraídos de lâminas escaneadas em laboratório com microscópio. Estas imagens foram manualmente marcadas e revisadas, e fazem parte do conjunto de dados criado. Utilizando este conjunto de dados, diferentes modelos de segmentação semântica foram testadas para verificar qual demonstra o melhor resultado em relação à identificação correta dos objetos para o auxílio ao diagnóstico de câncer cervical. Os dados são provenientes de pacientes do Hospital Universitário Professor Polidoro Ernani de São Thiago da Universidade Federal de Santa Catarina (HU-UFSC).

1.2 OBJETIVOS

Analisar a possibilidade da automação do processo de caracterização de lâminas coradas com prata (AgNOR), utilizando diferentes modelos e métodos de DL, para identificar características dos núcleos em células provenientes do colo de útero, realizando um auxílio ao diagnóstico de maneira precoce de possíveis casos de câncer.

1.2.1 Objetivos Específicos

Considerando o desenvolvimento do trabalho e o objetivo apresentado, destacam-se os seguintes objetivos específicos:

- Analisar o estado da arte da computação aplicado a imagens citológicas;
- Analisar o estado da arte em relação aos métodos de aprendizagem profunda para segmentação semântica e detecção de objetos;
- Criar/anotar um conjunto de imagens para treinamento dos modelos de DL;
- Comparar diferentes modelos de DL visando encontrar o melhor modelo para a segmentação semântica e/ou detecção de objetos para aplicação em imagens citológicas;
- Apresentar um método de DL para a caracterização de imagens citológicas coradas com prata (AgNOR);
- Divulgar/publicar os resultados da pesquisa.

1.3 JUSTIFICATIVA

Segundo os dados de Sung *et al.* (2021), entre o período de janeiro de 2016 e dezembro de 2020, estima-se cerca de 1,5 milhões de casos de câncer cervical. Foram cerca de 604 mil novos casos e 342 mil mortes durante o ano de 2020. No Brasil, foram cerca de 18 mil novos casos e mais de 9 mil mortes ao longo de 2020, apenas Índia, China e Indonésia possuem valores (em quantidade) superiores ao Brasil. Estas distribuições podem ser observadas nas Figuras 1a e 1b respectivamente. Nas figuras 1c e 1d, os gráficos no mundo (por país) para a taxa bruta¹ de novos casos e mortes. Observa-se uma alta taxa de novos casos e mortes, principalmente em países subdesenvolvidos. Em 2020, o Brasil apresentou cerca de 16,4 novos casos e 8,5 mortes a cada 100 mil mulheres. No continente Africano estão as maiores taxas, sendo as três maiores taxas de mortes do mundo apresentando valores acima de 30 e onde diversos outros países ao longo do continente com mais de 20 mortes a cada 100 mil mulheres.

Conforme o levantamento de dados em WHO (2020), cerca de 90% das 311 mil mortes de câncer cervical em 2018 ocorreram em países de baixa e média renda (do inglês, *low-income and middle-income countries (LMIC)*). Mais que 95% dos casos de câncer ou pré-câncer ocorrem devido ao HPV, cerca de 70% dos casos são devido a

¹ As taxas brutas são calculadas dividindo o número de novos casos ou mortes por câncer pelo número correspondente de indivíduos na população em risco. Para o câncer, o resultado é geralmente expresso como uma taxa anual por 100.000 indivíduos em risco.

dois tipos específicos do HPV (16 e 18). Nos últimos anos uma vacina contra estas variantes do HPV começou a ser utilizada em diversos países, porém ainda é necessário um diagnóstico e tratamento para aquelas que desenvolverem a doença.

Com um diagnóstico precoce, capaz de detectar lesões pré-cancerosas, estas podem ser tratadas e o câncer pode ser evitado. De maneira similar, caso detectado o câncer em uma fase precoce, o tratamento tem um alto potencial de cura. Porém, a detecção precoce ainda é um grande desafio (WHO, 2020; COHEN *et al.*, 2019).

A técnica de esfregaço cervical descrito na década de 1940 por George Papanicolaou (PAPANICOLAOU, 1942), é recomendada pela Organização Mundial da Saúde (OMS) para o rastreamento do câncer cervical, pois detecta células atípicas em amostras cervicais e tem se mostrado eficaz em reduzir a incidência e mortalidade (WHO, 2020). O esfregaço cervical é realizado como um exame de rotina, e dependendo dos resultados encontrados no exame a mulher deve ser encaminhada para a colposcopia. Em casos de achados anormais maiores na colposcopia a biópsia ainda se mantém necessária para avaliação histopatológica.

Anormalidades da morfologia nucleolar são frequentemente encontradas nas células cancerígenas, portanto sua análise pode auxiliar no diagnóstico. Os NORs são coradas com métodos de coloração utilizando prata, chamadas de AgNOR. Estes são identificados como pontos castanhos escuros a pretos dentro de um núcleo acastanhado dentro de um citoplasma amarelo (CROCKER *et al.*, 1989a; AN *et al.*, 2013). A contagem da quantidade de AgNOR em um núcleo pode refletir a atividade de proliferação das células e o grau de transformação maligna de certos tecidos (AN *et al.*, 2013; KHOT *et al.*, 2015). A quantidade de NORs coradas com prata tem sido utilizada para o diagnóstico de malignidade de numerosas lesões neoplásicas (PEDRINI *et al.*, 2020; DARKWAH *et al.*, 2018; KHUSHBU *et al.*, 2017).

Como este processo de análise e contagem das lâminas coradas com prata é feito manualmente, ele é propenso a erros humanos (LI *et al.*, 2017). Com o avanço dos métodos de DL, principalmente utilizando modelos de rede neural convolucional (do inglês, Convolutional Neural Network (CNN)), para visão é apresentado na literatura recente a vantagem dos seus usos, inclusive em imagens citológicas (MATIAS *et al.*, 2021a). Os modelos de CNN realizam a extração de características em sua estrutura, permitindo processar as imagens em sua forma bruta, sem a necessidade de pré-processamentos. Portanto, problemas da medicina, como a automatização do diagnóstico de lâminas citológicas, têm sido melhorados ou resolvidos para os mais diversos tipos de cânceres e corantes (MATIAS *et al.*, 2021a).

1.4 ESTRUTURA DA DISSERTAÇÃO

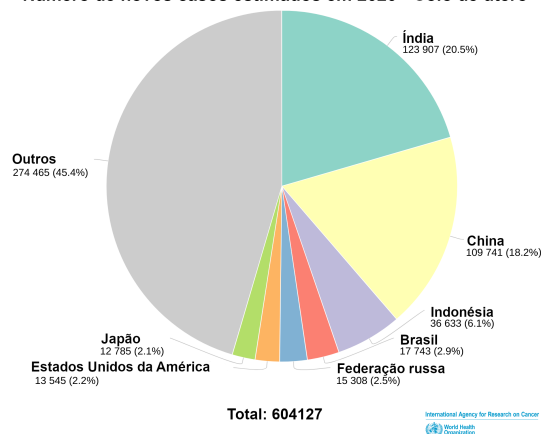
Essa dissertação está dividida em 5 capítulos. Inicialmente, com a introdução (capítulo 1), apresentando a contextualização do problema, os objetivos, a justificativa,

Figura 1 – Informações sobre a distribuição de novos casos e mortes por país para o ano de 2020.

Fonte: Ferlay *et al.* (2020)

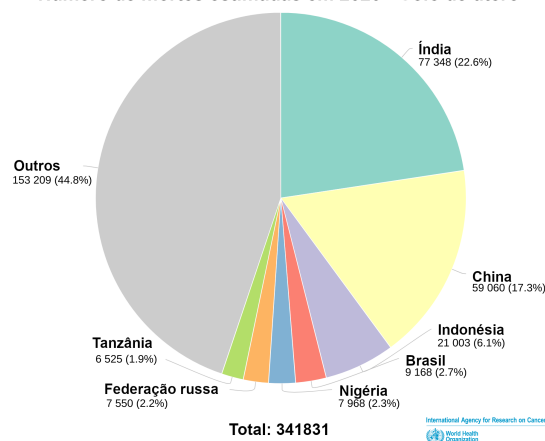
(a) Distribuição de novos casos por país

Número de novos casos estimados em 2020 - Colo de útero



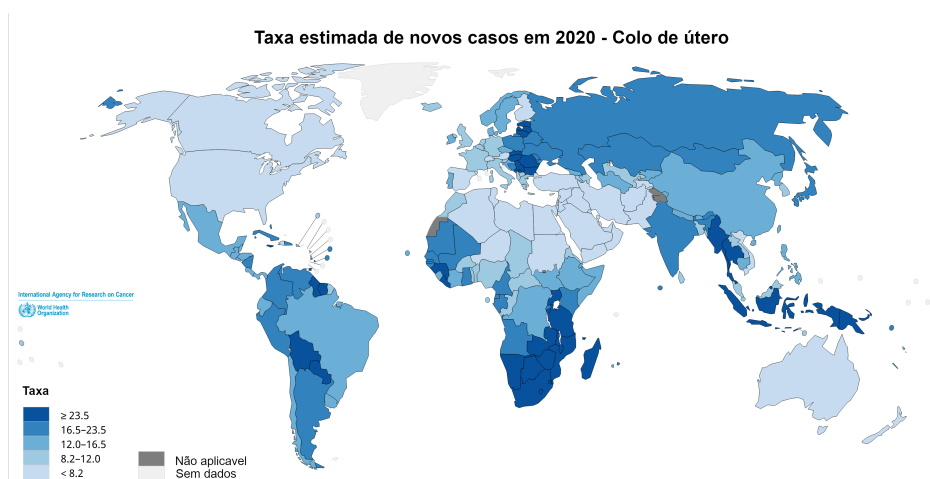
(b) Distribuição de mortes por país

Número de mortes estimadas em 2020 - Colo de útero



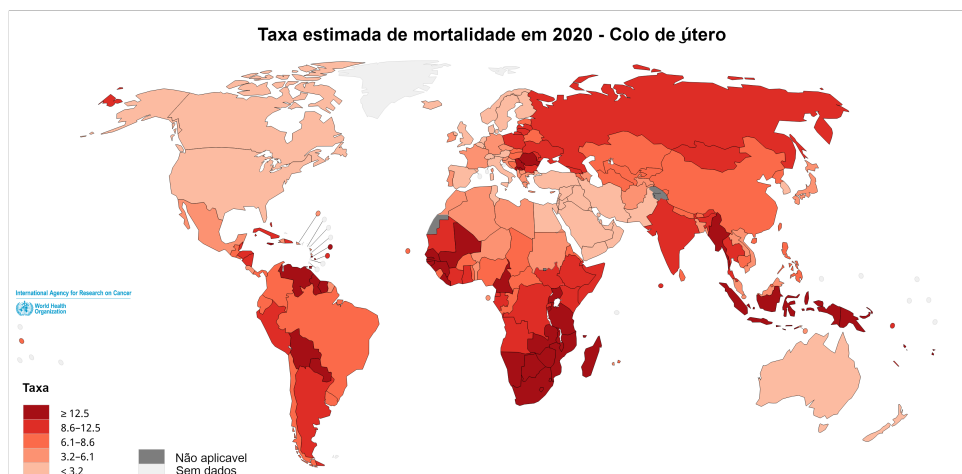
(c) Gráfico da taxa bruta de novos casos no mundo

Taxa estimada de novos casos em 2020 - Colo de útero



(d) Gráfico da taxa bruta de mortes no mundo

Taxa estimada de mortalidade em 2020 - Colo de útero



e por fim esta estrutura do trabalho. No capítulo 2 é apresentado uma revisão da literatura, primeiramente sobre citologia e focado ao método AgNOR, segundo sobre visão computacional e DL, e por fim sobre trabalhos de VC aplicado a citologia obtidos principalmente de duas revisões sistemáticas de literatura realizadas. O capítulo 3 apresenta os materiais e métodos, explicando os modelos e processos utilizados nesta dissertação. Os resultados obtidos são apresentados e discutidos no capítulo 4, seguido pela conclusão da dissertação no capítulo 5.

2 REVISÃO DA LITERATURA E FUNDAMENTAÇÃO TEÓRICA

2.1 CITOLOGIA

A citologia ou citopatologia é um procedimento diagnóstico de baixo custo e não invasivo empregado para auxiliar no diagnóstico de uma ampla gama de patologias. Estas células são coletadas por aspiração ou raspagem, sendo a técnica de raspagem cervical descrita na década de 1940 por George Papanicolaou (PAPANICOLAOU, 1942), que tornou a técnica de citologia mais popular (DIAMANTIS; MAGIORKINIS, 2013). Apesar da citologia existir desde 1837, a técnica só foi amplamente utilizada após os estudos de George Papanicolaou. Nas décadas seguintes, foram realizados diversos estudos para diagnosticar diversos tipos de câncer, e também outras doenças (DEMAY, 1996). A citologia apresenta diversas vantagens, como: baixo custo, e ser menos invasiva e mais simples de ser realizada, quando comparado com outras maneiras de coletar amostras, como a biópsia (BEDROSSIAN, 2007).

Há duas maneiras de se realizar exames citológicos, através da aspiração de agulha fina (do inglês, *Fine Needle Aspiration (FNA)*) e por citologia esfoliativa (do inglês, *Exfoliative Cytology*). Em cada uma há diferentes técnicas dedicadas a um órgão e/ou região a ser examinada. Exames cervicais (exame de papanicolau) podem ser obtidos utilizando uma espátula Ayre (ou dispositivo semelhante) em conjunto com uma escova endocervical. A amostra é colocada na lâmina sendo fixada imediatamente ou ser por meios de citologia de base líquida (do inglês, *liquid-based approach*). Apesar de ser necessário a coloração das lâminas, cada método utiliza diferentes corantes em diferentes processos sendo normalmente realizado o diagnóstico através de um microscópico (AL-ABBADI, 2011; IVANOVIC, 2013).

O método padrão em amostras cervicais é o exame e corante papanicolau, porém em alguns casos corantes coadjuvantes podem ser utilizados para auxiliar no diagnóstico. Entretanto, ainda é um procedimento predominantemente realizado manualmente por citologistas amplamente treinados para esse fim. É um processo demorado e repetitivo onde muitos critérios diagnósticos são subjetivos e vulneráveis à interpretação humana (LI *et al.*, 2017).

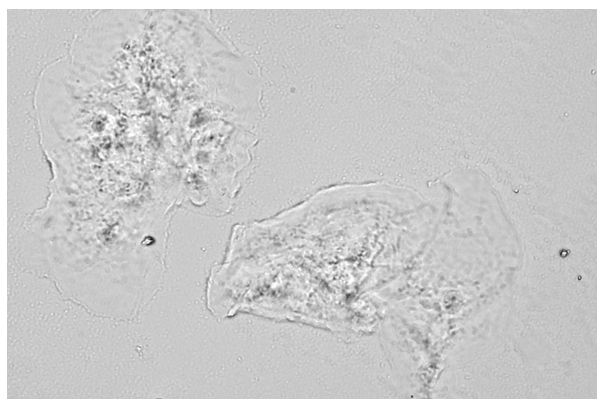
2.1.1 AgNOR

O AgNOR pode ser utilizado como uma técnica coadjuvante ao Papanicolau no diagnóstico ou prognóstico de exames cervicais, assim como para outras partes do corpo. Onde a identificação das proteínas ácidas associadas ao local de transcrição do rRNA é realizado através das AgNORs, este nome é dado ao método de coloração das NORs utilizando prata. A coloração pode ser observada na figura 2, onde na figura 2a está representado um campo de uma lâmina sem coloração, não sendo possível

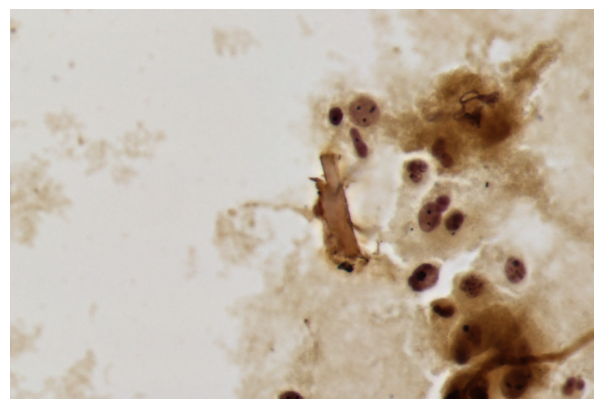
observar grandes diferenciações entre núcleos, citoplasma, etc. A figura 2b representa um campo de uma lâmina corada com prata (AgNOR), onde é possível identificar os NORs como pontos castanhos escuros a pretos dentro de um núcleo acastanhado dentro de um citoplasma amarelo (CROCKER *et al.*, 1989a; AN *et al.*, 2013).

Figura 2 – Exemplos de diferentes campos de uma lâmina capturadas com uma ampliação x40 (a) sem coloração; e (b) AgNOR

(a) Lâmina sem coloração



(b) Lâmina corada com prata (AgNOR)



Fonte: Autor.

A análise das lâminas coradas com prata, é feita manualmente, onde é contado a quantidade de NORs, os pontos pretos nos núcleos como na figura 2b. Portanto, estes NORs corados com prata, descritas em (CROCKER *et al.*, 1989b), podem ser classificados em 1) normais, uma região argirofílica totalmente agregada e arredondada (esférica) denominadas satélites, ou pontos isolados; 2) atípica, uma região de nucléolos quase agregada ou parcialmente desagregada, tratada como uma única região (grande, em forma de rim ou aglomerada) chamada *cluster*. Os *clusters* são um agrupamento de NORs observável na microscopia óptica, ou em imagens utilizando objetivas de maior aumento, também é possível identificá-los pela medida de sua área que é relativamente superior à dos satélites (AN *et al.*, 2013).

Embora não haja padronização para o diagnóstico utilizando AgNOR, o tamanho e a quantidade de AgNORs (também chamadas por somente NORs) estão relacionados com a capacidade proliferativa da célula (TRERÈ, 2000). Alguns trabalhos consideram a contagem dos NORs separadamente (PEDRINI *et al.*, 2020; MIRANDA ONOFRE *et al.*, 2008; CARMEN ALARCÓN-ROMERO *et al.*, 2009), ou apenas *clusters* (AN *et al.*, 2013) ou a quantidade total (DOGENSKI *et al.*, 2020; SHARMA, H. D. *et al.*, 2021), e em outros casos, consideram também a morfologia e área dos NORs e núcleos (SANCER *et al.*, 2020; GARCÍA-VIELMA *et al.*, 2015).

2.2 VISÃO COMPUTACIONAL E DEEP LEARNING

O princípio geral por trás da visão computacional clássica é que a realização da interpretação da imagem começa a partir de píxeis sem significado e move-se, passo a passo, em direção a uma representação significativa do conteúdo dessa imagem. Portanto, a visão computacional pode ser definida como o campo de estudos que desenvolve técnicas responsáveis por dar sentido para o computador através uma imagem digital (GONZALEZ; WOODS, 2008).

Aprendizado profundo (do inglês, *Deep Learning* (DL)) é uma sub-área do aprendizado de máquina (do inglês, *Machine Learning* (ML)) composta por diversos modelos para diferentes propostas, como processamento de linguagem natural, processamento de fala, entre outros. *Convolutional Neural Network* (CNN) (ou Rede Neural Convolucional) é o modelo mais adequado para trabalhar com visão computacional, onde originalmente era utilizado apenas para reconhecimento (classificação) de imagens. Atualmente, CNN não é utilizada exclusivamente para visão computacional, mas ainda é a mais utilizada para a área. Diversas abordagens — como segmentação semântica e detecção de objetos — são realizadas na visão computacional utilizando CNN (POUYANFAR *et al.*, 2019).

Detecção de objetos (do inglês, *Object detection*), consiste em detectar objetos pertencentes a determinada categoria a partir de uma imagem, determinando a localização absoluta e também atribuir a cada instancia detectada uma categoria predefinida. Nos últimos anos, diversos modelos surgiram baseadas em DL, sendo os modelos no estado da arte (do inglês, *state-of-the-art* (SOTA) models) baseados em dois estágios. Onde primeiramente é localizado regiões com maiores valores de *recall*, de maneira que todos os objetos na imagem se adéquem as regiões propostas. O segundo estágio consiste em modelos de classificação, geralmente CNNs, utilizados para determinar a categoria de cada região proposta (instancias). Os modelos SOTA para detecção de objetos são *Faster R-CNN* e *R-FCN*, e suas variações adaptando em seu uso a *Feature Pyramid Networks* (FPN) (SHARMA, V.; MIR, 2020). Apesar de a detecção de objetos ter evoluído nos últimos anos, a detecção de objetos pequenos ou minúsculos ainda continua sendo um desafio. Alguns autores tentam resolver criando modelos específicos para resolver problemas específicos, porém, não há modelo que tenha obtido destaque neste quesito (TONG; WU, 2022).

Segmentação semântica (do inglês, *Semantic Segmentation*), é a técnica que realiza a classificação de cada píxel a partir de uma imagem. Desde a publicação dos modelos *SegNet* e *U-net*, diversas arquiteturas de redes e adaptações foram realizadas. Ambos modelos são arquiteturas baseada em uma estrutura de *encoder/decoder*. Onde o *encoder* é o estágio responsável por extrair as características de uma determinada entrada utilizando camadas de convolução, e o *decoder* é responsável por a partir destes mapas de características, construir uma camada de mesmo tamanho

como saída, que tenha uma categoria associada a cada píxel. A U-net é tratada como linha base para problemas de segmentação semântica, onde sua principal característica quando comparado a SegNet foi a adição de conexões “diretas” entre o *encoder* e *decoder*. A arquitetura da U-net para o *encoder* e *decoder*, é popularmente chamada *backbone* da rede, e pode ser adaptada a partir de modelos de CNN para classificação como, por exemplo, os modelos VGGNet, ResNet, EfficientNet, etc (TAGHANAKI *et al.*, 2020).

Desde os relatos de (VASWANI *et al.*, 2017) sobre modelos *transformers* (baseados em mecanismos de atenção), vários trabalhos adaptaram os métodos de atenção para modelos VC, por exemplo, a ResNet com atenção (WANG, F. *et al.*, 2017) e depois a *U-net* com atenção (OKTAY *et al.*, 2018). A revisão sistemática de literatura (RSL) sobre *transformers* em visão computacional descrito em (KHAN *et al.*, 2022) relata o aumento maciço em publicações de novos modelos e aplicações de modelos *transformers*. Sendo segmentação semântica uma abordagem em destaque, diversos modelos estão surgindo diariamente, porém para este trabalho destaque os modelos SOTA HRNet baseado em CNN e SegFormer baseado em *transformers* (WANG, J. *et al.*, 2021; XIE *et al.*, 2021).

2.3 VISÃO COMPUTACIONAL APLICADA A CITOLOGIA

Em nosso grupo de pesquisa, foi primeiramente realizado uma RSL publicada como relatório técnico junto ao INCoD - UFSC (AMORIM *et al.*, 2020a), onde foi posteriormente publicada uma versão estendida desta RSL em um periódico (MATIAS *et al.*, 2021a). Nesta seção serão analisado os resultados, principalmente do trabalho Matias *et al.* (2021a), por se tratar de um trabalho mais completo. Esta RSL é focada em identificar o estado da arte das técnicas de visão computacional atualmente aplicadas à citologia, buscando abordagens para segmentação, detecção, quantificação e classificação de células e organelas usando visão computacional em lâminas citológicas. Sendo analisado artigos publicados entre 2016 e 2020, sendo que a busca inicial foi realizada em setembro de 2020 e resultou em 431 artigos. Após a aplicação dos critérios de inclusão/exclusão, restaram 157 artigos, analisados para construir um retrato das tendências e problemas presentes nesta área de pesquisa, destacando os métodos de visão computacional, técnicas de coloração, métricas de avaliação e a disponibilidade dos conjuntos de dados utilizados e de códigos para avaliação de modelos.

Nesta RSL, foi identificado que majoritariamente (101 artigos) em alguma das etapas — pré-processamento, processamento, e/ou pós-processamento — ainda é utilizado algoritmos de visão computacional clássica. Os demais trabalhos (56 artigos) empregaram processos puramente baseados em DL, porém entre os 101 artigos que utilizam VC clássica 14 deles também utilizam processos de DL. Em relação às técnicas de coloração, o Papanicolaou foi o mais utilizado (130 artigos), seguido

por H&E (20 artigos) e Feulgen (5 artigos). Foram identificados que doze conjuntos de dados estão disponíveis publicamente, sendo o conjunto de dados DTU/Herlev o mais utilizado. Contudo, foi detectado que ainda há falta de conjuntos de dados de alta qualidade para muitos tipos de corantes e a maioria dos trabalhos não está testada o suficiente para ser aplicada na rotina diária de diagnóstico clínico.

Para os anos de 2019 e 2020, majoritariamente os trabalhos publicados utilizaram DL, apontando uma tendência crescente para a adoção deste tipo de abordagem. Detectou-se que também não havia padronização na utilização das métricas e uma ferramenta/biblioteca para cálculo das mesmas, dificultando a comparação de resultados entre trabalhos. A métrica *Dice Similarity Coefficient* (também chamada *F-score*) a mais recorrente (38 artigos) para a segmentação semântica e a acurácia (5 artigos) para a detecção de objetos.

Para este trabalho, foi ampliado a busca para artigos publicados entre 2020 e abril de 2022, e não foram encontrados artigos relacionados a visão computacional com DL para imagens coradas com a técnica de AgNOR de outros autores. Os únicos trabalhos entre 2016 e abril de 2022, relacionado a AgNOR, foram os resultados iniciais ou intermediários relacionados a esta dissertação. Todos os artigos publicados ao decorrer da dissertação encontram-se listados no apêndice A.

Os resultados iniciais de segmentação semântica utilizando um modelo de U-net com ResNet, assim como a primeira versão do conjunto de dados criado (*Cervical Cells with AgNOR Stain Technique — CCAgT dataset*) foi relatado em (AMORIM *et al.*, 2020b). No trabalho (AMORIM *et al.*, 2020c), utilizando a primeira versão do conjunto de dados CCAgT, foi reportado resultados utilizando o modelo de detecção de objetos Faster R-CNN, onde foi testado a possibilidade desta abordagem para a detecção de apenas núcleos. No trabalho (AMORIM *et al.*, 2022b), utilizando uma versão intermediária (versão 1.5) do conjunto de dados CCAgT, e o modelo Fast R-CNN de detecção de objetos com diferentes backbones, para a detecção de diferentes categorias de núcleos.

Embora não tenha sido possível localizar trabalhos relacionados a métodos de VC aplicado a imagens de AgNOR, pode-se citar alguns trabalhos, pois trabalham com amostras cervicais e/ou métodos DL. Em (ZHANG, L. *et al.*, 2017), os autores combinaram uma CNN com uma abordagem baseada em grafos para segmentação de núcleos corados com papanicolau, apresentando resultados melhores que os métodos de VC clássica testados. Os autores em (HUSSAIN *et al.*, 2020) aplicaram um processo onde é primeiramente produzida a segmentação de instância utilizando um modelo baseado na U-net, e posteriormente classificaram com um modelo de representação de formas, para classificação de núcleos normais e anormais corados com Papanicolau.

Diversos trabalhos podem ser encontrados utilizando o conjunto de dados *Overlapping Cervical Cytology Image Segmentation Challenge - ISBI 2014*, que consiste em

imagens do colo do útero corado com papanicolau em base líquida. Em (UMADI *et al.*, 2020) utilizam agrupamento k-means para diferenciar a região do fundo (branco) das células, e posteriormente utilizaram uma U-net para a segmentação do núcleo da célula. Em (WAN *et al.*, 2019), são utilizados um método baseado em janelamento duplo para localizar as células individualmente e então são utilizados o modelo baseado no DeepLab V2 para a segmentação semântica. Em (HUANG *et al.*, 2020) apresentam um modelo que utiliza um novo bloco residual que tira as vantagens dos blocos residuais (da ResNet), dos blocos da Inception e assim como dos blocos densos (*Dense blocks*) mostrando melhora quando comparado a U-net original.

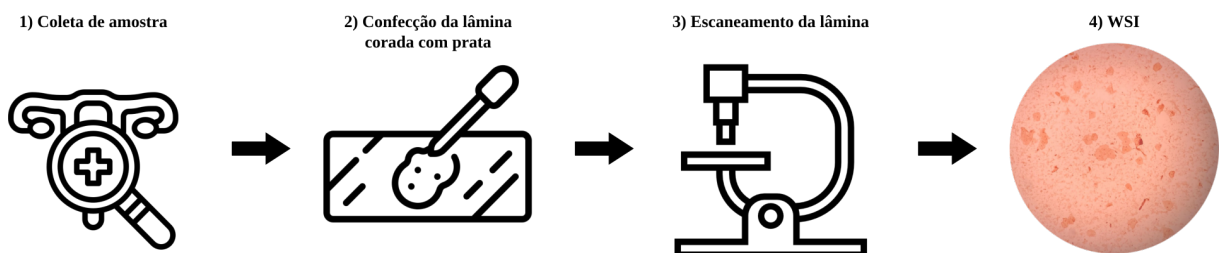
Em (DU *et al.*, 2019) foi reportado resultados utilizando o modelo de detecção de objetos Faster R-CNN para a detecção de células coradas com Papanicolau. Já em (MATIAS *et al.*, 2021b) é realizado uma comparação dos métodos de segmentação semântica, detecção de objetos e classificação para amostras orais coradas com papanicolau, para segmentação foi utilizado o modelo U-net com ResNet, e para detecção os modelos Faster R-CNN e RetinaNet. Para detectar células em imagens de microscópio, em (PRANGEMEIER *et al.*, 2020) adaptaram o modelo *transformer* DETR, para realizar a segmentação de instâncias não sobrepostas. Em (LIN *et al.*, 2021) utilizaram diferentes modelos para a classificação e segmentação semântica para imagens citológicas de pulmão, sendo o melhor resultado para segmentação obtido pela HRNet.

3 MATERIAIS E MÉTODOS

Nesta seção será explicado as etapas, processos e modelos adotados para construção desta dissertação, sendo estas: a elaboração do conjunto de dados, explicação geral sobre os modelos testados, a configuração dos experimentos e tecnologias utilizadas, e por fim a metodologia e métricas adotadas para a avaliação de resultados. A seguir, de maneira breve, será abordado sobre as etapas macros do trabalho e subsequentemente cada etapa essencial será abordada ao longo das próximas seções.

O projeto em que esta dissertação está inserida, é multidisciplinar, onde a coleta das amostras e análise das mesmas é feita por membros do departamento de análises clínicas da UFSC. As primeiras etapas destacadas na figura 3 são essenciais para o desenvolvimento desta dissertação. Portanto, as etapas realizadas ou supervisionadas por membros do departamento de análises clínicas, consiste em: (1) aquisição da amostra do colo uterino; (2) confecção da lâmina a partir da amostra e coloração da lâmina; (3) realizar o escaneamento da lâmina, para gerar uma (4) imagem da lâmina por completo (do inglês, Whole Slide Images (WSI)). Esta dissertação é responsável pelas etapas subsequentes, apresentadas nas figuras 4 e 5.

Figura 3 – Etapas macros para aquisição das imagens de amostras cervicais



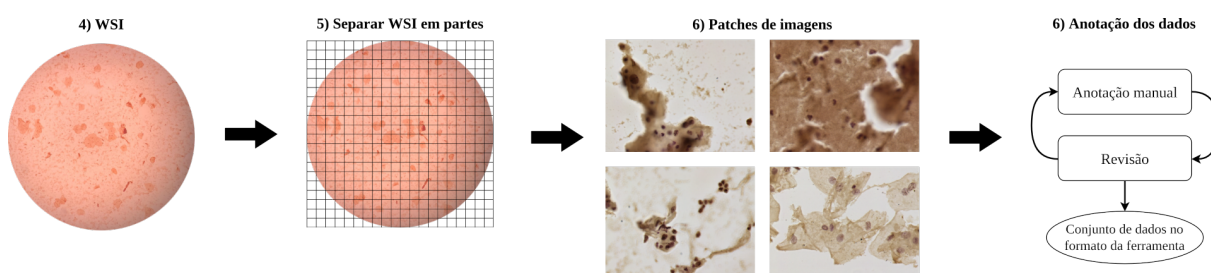
Fonte: Autor.

Da mesma maneira que não foi possível localizar trabalhos relacionados a métodos de VC aplicado a imagens de AgNOR, não foi possível identificar nenhum conjunto de dados (do inglês, *dataset*) público de imagens de células coradas com AgNOR. Portanto, foi construído o *Cervical Cells with AgNOR Stain Technique (CCAgT)* dataset conforme as etapas e itens apresentados na figura 4 exemplificam. A criação do dataset CCAgT, consiste em extrair de uma WSI suas sub-imagens (ou “patches” de imagens) como ilustrado na etapa 5. Estas imagens (item 6), foram anotados e revisados em um software para a criação do conjunto de dados ¹.

Após o dataset ser revisado (etapa 6), utilizando a ferramenta criada *CCAgT-utils*, o dataset é convertido e padronizado a partir das anotações do software para um formato de tabelas (etapa 7). Este dataset gerado na etapa 7, pode ser exportado

¹ CCAgT-utils: ferramenta de utilidades para ser utilizada junto ao dataset. Disponível em <https://doi.org/10.5281/zenodo.6350442>.

Figura 4 – Etapas macros e itens para produção do conjunto de dados



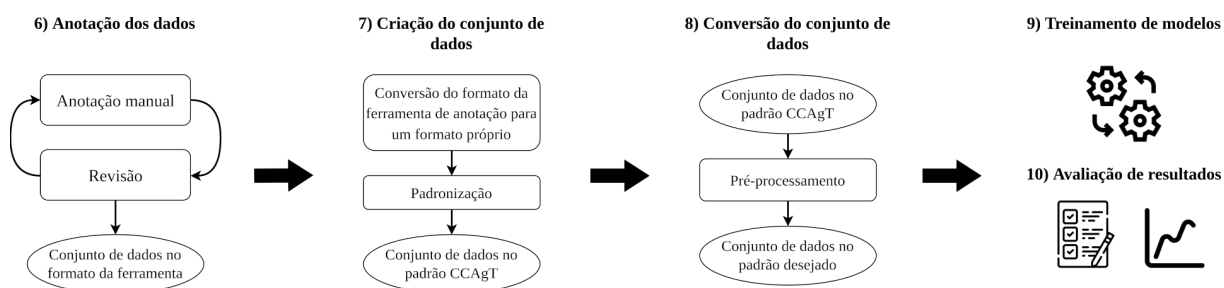
Fonte: Autor.

para diferentes formatos assim como realizar pré-processamentos para a criação de sub-datasets (etapa 8). Por fim, o foco principal deste trabalho, treinamento dos modelos (etapa 9) e avaliação dos resultados dos modelos (etapa 10). Neste trabalho, é demonstrado resultados de três diferentes modelos em dois sub-datasets.

As etapas essenciais explicadas nas seções subsequentes, estão organizadas da seguinte maneira:

- **Elaboração do conjunto de dados:** abordando as configurações de aquisição de imagem, a metodologia utilizada para anotação das categorias nas imagens, e por fim informações gerais sobre o dataset;
- **Modelos:** uma breve explicação sobre os modelos adotados;
- **Configuração dos experimentos:** abordando quais as tecnologias e bibliotecas utilizadas, assim como quais as configurações e estratégias foram utilizadas ao longo do treinamento dos modelos.
- **Avaliação de resultados:** explicando quais as metodologias adotada para a separação dos dados, assim como sobre as métricas utilizadas.

Figura 5 – Etapas macros de preparo do conjunto de dados, treinamento de modelos e avaliação de resultados



Fonte: Autor.

3.1 ELABORAÇÃO DO CONJUNTO DE DADOS

O conjunto de dados (*dataset*) utilizado nesta dissertação foi elaborado pelo nosso grupo de pesquisa, um trabalho em conjunto entre os laboratórios da UFSC: *Image Processing and Computer Graphics Lab.* — LAPIX e Laboratório de Citologia Clínica do Departamento de Análises Clínicas. As amostras do colo do útero (etapa 1 da figura 3) foram obtidas de pacientes do Ambulatório de Ginecologia e Colonoscopia do *Hospital Universitário Polidoro Ernani de São Thiago da Universidade Federal de Santa Catarina* (HU-UFSC). Todas as amostras são de mulheres que tiveram previamente alterações em seus exames de citologia cervical².

As etapas 2 e 3, referentes a confecção das lâminas, coloração e escaneamento da lâmina foram realizadas ou supervisionadas por membros do departamento de análises clínicas, e parte deste processo foi relatado em (PEDRINI *et al.*, 2020). A aquisição das imagens foram feitas através de um *scanner*, onde a partir de uma configuração de aquisição, a câmera realiza a captura de milhares de imagens utilizadas para compor a imagem da lâmina inteira (WSI). Estas imagens podem ser chamadas subimagens ou *patches*, podendo apresentar objetos fora de foco e há imagens que não possuem nenhum objeto de interesse.

As WSI foram obtidas usando um *ZEISS Axio Scanner.Z1* com uma câmera *Hitachi HV-F202SCL*. Para compor uma WSI são necessários mais de 20 mil subimagens com uma resolução de 1200×1600 pixels, onde cada pixel corresponde a $0,111\mu m \times 0,111\mu m$. Para a criação do dataset, primeiramente (etapa 5 apresentada na figura 4) é necessário separar os patches do arquivo *RAW* da WSI mantendo toda a resolução da aquisição. Em seguida, os patches (apresentados como item 6) são selecionados para irem para a anotação manual dos dados (etapa 7). Esta seleção é realizada de maneira automatizada baseado na porcentagem de pixels não branco que a mesma possui, utilizando uma análise de histograma destas imagens. Desta maneira é possível diminuir de mais de 20 mil patches para cerca de 5 a 10 mil patches (a depender da lâmina) para serem analisados e anotados manualmente.

Na etapa 6, para realizar a marcação manual dos objetos foram utilizados dois softwares. Para a primeira versão do dataset foi utilizado o LabelMe (RUSSELL *et al.*, 2007) e na segunda versão foi utilizado o LabelBox (LABELBOX, 2021). Todas as anotações foram realizadas manualmente, sendo que na primeira versão do dataset foram anotadas 3 categorias e na segunda versão foram anotadas 7 categorias. Dentre estas anotações, os núcleos, clusters, núcleos fora de foco, núcleos sobrepostos, núcleos não viáveis e leucócitos foram anotadas como polígonos. As anotações dos satélites

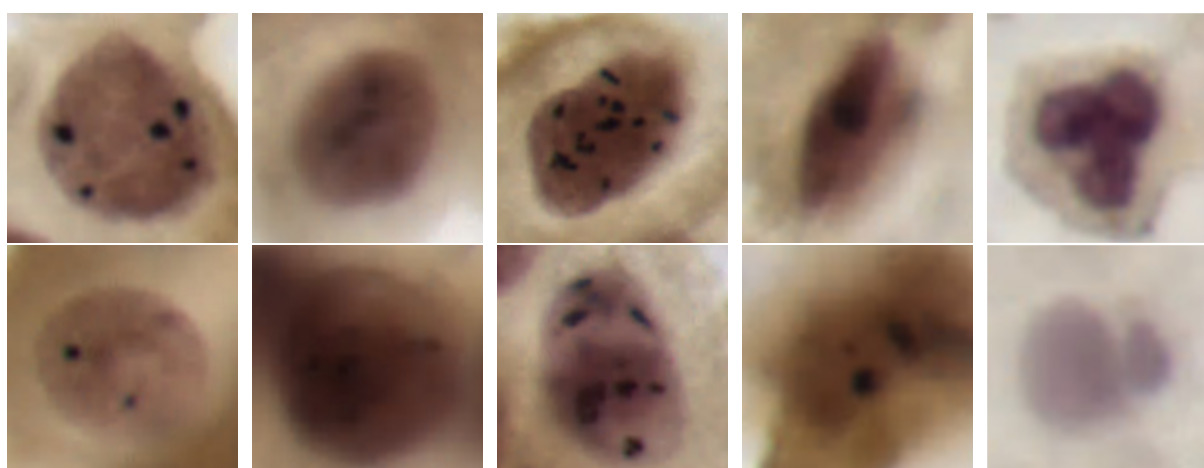
² Pesquisa aprovada pelo Comitê de Ética em Pesquisa com Seres Humanos (CEPSH) da UFSC, protocolo número 57423616.3.0000.0121. Primeiramente, todas as pacientes envolvidas foram informadas sobre os objetivos do estudo, e aqueles que concordaram em participar assinaram o termo de consentimento.

foram anotados como pontos, e posteriormente transformados em polígonos redondos com área igual a 88 píxeis.

Sendo as categorias do conjunto de dados:

1. *Núcleos* (figura 6a): o núcleo celular, região arredondada em tons de marrom;
2. *Cluster*: NORs atípicos, itens não arredondados (em forma de rim ou aglomerado de pontos) ou grandes pontos pretos sempre dentro de um núcleo;
3. *Satélites*: NORs típicos, pontos pretos arredondados e não grandes sempre dentro de um núcleo;
4. *Núcleos fora de foco* (figura 6b): núcleos que a câmera não conseguiu focar e consequentemente não é possível identificar seus NORs;
5. *Núcleos sobrepostos* (figura 6c): onde não é possível determinar a qual núcleo o NOR pertence;
6. *Núcleos não viáveis* (figura 6d): objeto que não são núcleos identificáveis, porém nas imagens tem visualmente os mesmos tons, texturas ou forma de um núcleo, mas é um artefato que queremos ignorar na avaliação do WSI;
7. *Leucócitos* (figura 6e): qualquer categoria de leucócito;

Figura 6 – Categorias do CCAgT dataset



(a) Núcleo

(b) Núcleo fora de
foco

(c) Núcleos sobre-
postos

(d) Núcleo não-
viável

(e) Leucócitos

Fonte: Autor.

Este conjunto de dados criado (etapas 6 e 7 da figura 5) recebeu o nome de *Cervical Cells with AgNOR Stain Technique (CCAgT) dataset*, tendo sido publicada em duas diferentes versões. As estatísticas e informações sobre cada uma das versões publicada são explicadas nas subseções seguintes.

3.1.1 Versão 1

Esta versão do CCAgT dataset foi construída com anotações apenas das categorias principais (núcleo, cluster e satélite). A mesma foi disponibilizada na versão COCO de detecção de objetos, e também como máscara para segmentação semântica.

Possui dados de três lâminas de diferentes pacientes ($n = 3$), com 2540 imagens e 16711 anotações, onde cada imagem possui pelo menos um núcleo anotado. Esta versão foi publicada em anexo junto ao trabalho descrito em (AMORIM *et al.*, 2020b), e encontra-se publicamente acessível³. A tabela 1 apresenta estatísticas gerais sobre esta versão do dataset.

Categoria	Qtd. de instâncias	Área média	Dev. pad.
Núcleos	4515	4815	3108
Cluster	10025	120	79
Satélites	2171	100	3

Tabela 1 – Estatísticas gerais da primeira versão do CCAgT dataset

3.1.2 Versão 2

Esta versão do CCAgT dataset foi construída com anotações de setes categorias (descritas anteriormente) (AMORIM *et al.*, 2022a). A mesma foi disponibilizada em uma versão própria (tabela com geometrias), na versão como COCO para detecção de objetos, e também máscara para segmentação semântica. Para esta versão foi criado e utilizado a ferramenta *CCAgT-utils*, que auxilia na conversão do formato dos dados, assim como nas tarefas de divisão dos dados, pré ou pós processamento (AMORIM, 2022). Esta ferramenta foi utilizada também para garantir a qualidade das anotações, excluindo aquelas com áreas muito pequenas, unindo anotações de núcleos sobrepostos, garantindo que todos os núcleos anotados possuem pelo menos um NOR em seu interior, assim como garantindo que todos os NORs estavam dentro de algum núcleo ou núcleo sobreposto.

As novas categorias foram adicionadas em todas as lâminas, com intuito de ter a maior quantidade de categorias de objetos anotados que aparecem recorrentemente nas imagens. Ou seja, algumas destas categorias podem não ser essenciais para o diagnóstico do paciente, porém foram adicionadas com o intuito de auxiliar os modelos de DL no momento da aprendizagem.

Possui dados de quinze lâminas de diferentes pacientes ($n = 15$) com sete diferentes diagnósticos, 9339 imagens e mais de 63 mil anotações, onde cada ima-

³ Disponível em <https://arquivos.ufsc.br/d/373be2177a33426a9e6c>.

gem possui pelo menos uma anotação. Encontra-se publicamente acessível na base de dados *mendelay*⁴. A tabela 2 apresenta estatísticas gerais sobre esta versão do dataset.

Categoria	Qtd. de instâncias	Área média	Dev. pad.
Núcleos	14013	4372	1827
Cluster	25710	104	63
Satélite	6478	88	1
Núcleo fora de foco	9270	3566	1607
Núcleos sobrepostos	1373	7891	4632
Núcleo não-viável	1412	3745	1723
Leucócitos	4934	1826	825

Tabela 2 – Estatísticas gerais da segunda versão do CCAgT dataset

3.2 MODELOS

Neste trabalho, a segmentação semântica foi utilizada como a única etapa de visão computacional com DL. É chamado “semântica” porque associa um píxel individual diretamente ao seu significado: cada píxel é classificado conforme a classe de objeto a que pertence e não apenas agregado em regiões sem sentido com base em critérios sintáticos, como homogeneidade de cor ou bordas e variação de cor. A segmentação semântica elimina a necessidade de reconhecimento de padrões adicionais para a classificação das regiões resultantes de uma segmentação. Os três diferentes modelos para a segmentação utilizados são os seguintes:

- *U-net*: é o modelo que revolucionou as abordagens de segmentação do VC é um modelo com duas partes: codificador (do inglês, encoder), responsável por extrair características da imagem original, e o decodificador (do inglês, decoder), responsável por reconstruir as características extraídas em uma máscara de saída com o mesmo tamanho de entrada (RONNEBERGER *et al.*, 2015). Para este trabalho, as arquiteturas de backbone testadas foi a ResNet em diferentes tamanhos.
- *HRNet*: este modelo introduz uma nova abordagem para a segmentação semântica que mantém representações de alta resolução e trabalha com blocos de

⁴ Disponível em <https://doi.org/10.17632/wg4bpm33hj.2>.

convolução com resolução 1/4, 1/8, 1/16 e 1/32 da imagem de entrada (WANG, J. *et al.*, 2021). Assim, ao contrário da U-net, o pipeline desta arquitetura funciona com blocos de convoluções paralelas com troca de informações entre resoluções. Esta abordagem multi-resolução melhora a representação espacial, e a informação semântica, conseqüentemente, pode realizar a segmentação de pequenos objetos.

- *SegFormer*: é um modelo Transformer com um decoder leve feito por uma rede neural de multicamadas (XIE *et al.*, 2021). O encoder é um bloco extrator de características sequenciais de multi-resolução (1/4, 1/8, 1/16 e 1/32) que passa as características de todas as resoluções para o decoder. A fusão dessas características no decoder permite combinar a atenção local e atenção global de diferentes escalas, tornando este modelo uma opção para problemas de segmentação com pequenos objetos.

3.3 CONFIGURAÇÃO DOS EXPERIMENTOS

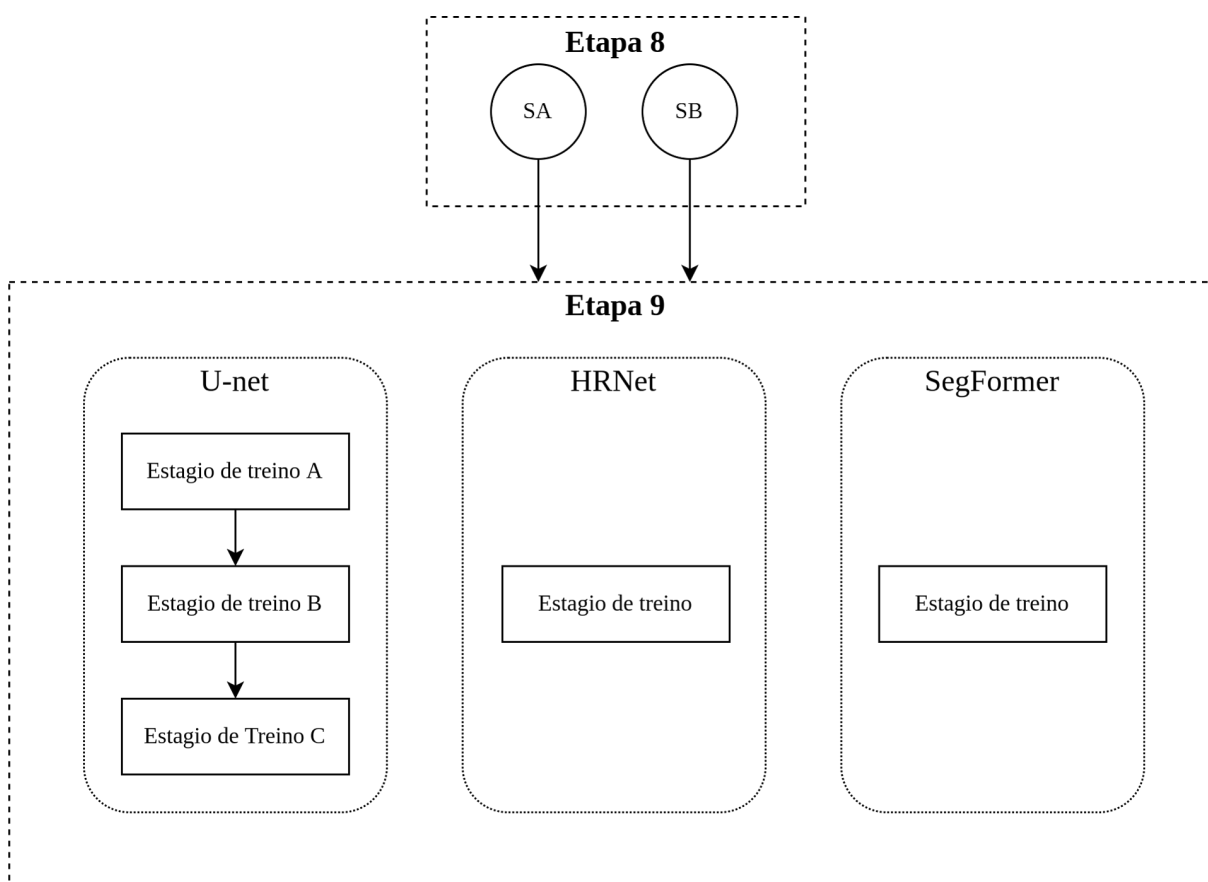
Para a realização dos experimentos, diferentes subetapas foram realizadas para as etapas 8 e 9 da figura 5. Estas subetapas podem ser observadas na figura 7, na etapa 8 diferentes sub-datasets são criados, e posteriormente treinados na etapa 9 para cada um dos modelos avaliados. Além disto, é exibido os estágios de treino de cada modelo. Nesta seção será detalhada cada uma destas etapas caracterizando a configuração dos experimentos.

Todos os experimentos foram realizados utilizando Python 3.7.12 (VAN ROSSUM; DRAKE, 2009) e Pytorch 1.11.0 (PASZKE *et al.*, 2019). As arquiteturas dos modelos utilizadas disponibilizadas através das bibliotecas *fast.ai* (HOWARD; GUGGER, 2020) e *Hugging Face transformers* (WOLF *et al.*, 2020), com a ferramenta *Weight and Biases* para realizar o acompanhamento dos modelos (BIEWALD, 2020). Para a manipulação do dataset a ferramenta *CCAgT-utils* foi criada (AMORIM, 2022). O ambiente de desenvolvimento utilizado foi o Google Colaboratory (BISONG, 2019), que dispõe de máquinas com os seguintes hardwares: CPU Intel(R) Xeon(R) @ 2.20GHz × 4; com 16 a 26 GiB de memória RAM; placa de video Nvidia Tesla K80 (12 GiB VRAM), T4 (16 GiB VRAM), P100 (16 GiB VRAM), ou A100 (40 GiB VRAM); com o Ubuntu 18.04.3 LTS 64-bit como sistema operacional.

3.3.1 Sub-datasets

Conforme mostra a figura 7, dois sub-datasets foram utilizados. Estes sub-datasets foram criados para se adequar aos modelos, e criar diferentes experimentos utilizando o mesmo dataset como base. Para realizar esta tarefa, é utilizado a ferramenta *CCAgT-utils* (AMORIM, 2022).

Figura 7 – Organização geral dos experimentos

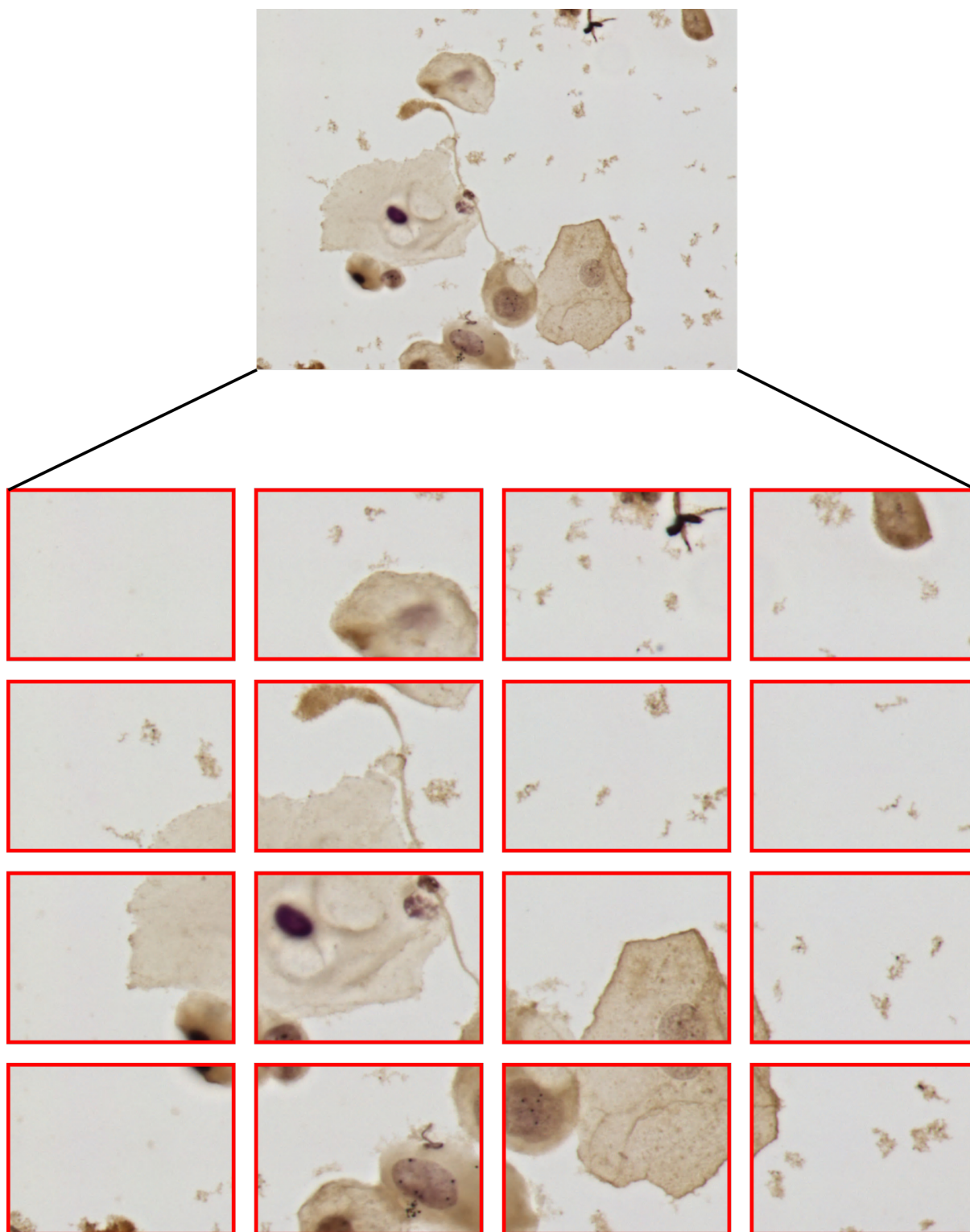


Fonte: Autor.

Estes sub-datasets criados, utilizaram como base os dados da segunda versão do CCAgT dataset. Para criação do sub-dataset A (SA) para segmentação semântica, o seguinte pipeline foi adotado: manter todas as categorias de anotações; separar as imagens utilizando uma grade 4×4 , criando imagens de 400×300 píxeis de resolução conforme ilustrado na figura 8; eliminar anotações inválidas (cada categoria possui uma área mínima aceita), manter apenas imagens que possuem pelo menos uma anotação de NORs.

Para a criação do sub-dataset B (SB), o mesmo pipeline foi adotado, porém, foram mantidas apenas anotações das categorias principais (núcleos e NORs) e núcleos fora de foco. Esta resolução foi selecionada para se adequar com as arquiteturas utilizadas, que apesar de poderem ser adaptadas para qualquer resolução, em geral, foram construídas ou pré-treinadas para imagens de resolução de 256×256 , 512×512 , etc. O SA possui 11954 imagens com um total de 54413 anotações, sendo 47321 anotações apenas das categorias principais. O SB possui 11119 imagens com um total de 45740 anotações, sendo 42942 anotações das categorias principais. A tabela 3 possui as distribuições de ambos sub-datasets.

Figura 8 – Exemplo de separação de uma imagem em 16 partes.



Fonte: Autor.

Categoria	Sub-dataset A			Sub-dataset B		
	Qtd. de instâncias	Área média	Dev. pad.	Qtd. de instâncias	Área média	Dev. pad.
Núcleos Cluster	14978	3716	1883	14978	3706	1886
Satélite	25819	105	67	22057	105	68
Núcleo fora de foco	6524	86	6	5907	86	6
Núcleos sobrepostos	2949	2901	1535	2798	2929	1497
Núcleo não-viável	1737	6739	7347	-	-	-
Leucócitos	673	2305	1319	-	-	-
	1733	1137	602	-	-	-

Tabela 3 – Estatísticas dos subdatasets

3.3.2 Treinamento

O treinamento de um modelo consiste na iteração por cada imagem, por N vezes, o número de vezes que o modelo percorre todo o conjunto de dados é um hiperparâmetro chamado *época*. Para auxiliar no processo de iterações, lotes (do inglês, batches) são criados sendo cada lote um subconjunto de imagens. O tamanho destes lotes (considera também um hiperparâmetro chamado *batch size*) foram definidos conforme a disponibilidade do hardware (quantidade de memória de vídeo disponível na placa gráfica) (HAYKIN, 2009).

Para auxiliar os modelos a aprenderem as características mais importantes das imagens e evitar *overfitting*, foi utilizada a técnica de *data augmentation*. Esta técnica consiste um conjunto de técnicas que aprimoram o tamanho e a qualidade dos conjuntos de dados de treinamento (SHORTEN; KHOSHGOFTAAR, 2019). Este conjunto de técnicas consiste em técnicas que consigam realizar transformações nas imagens, portanto sendo i uma imagem e T uma função de transformação, portanto $T(i) = i_t$, onde i_t consiste em uma imagem transformada. Estas funções de transformações podem ser técnicas clássicas de visão computacional como, por exemplo, rotacionar, espelhar e/ou distorcer uma imagem, assim como podem ser imagens sintéticas geradas a partir de modelos de DL (SANDFORT *et al.*, 2019).

Para melhor proveito desta técnica, aplicam-se estas funções de transformações de maneira encadeada e aleatória. Elas são aplicadas para cada imagem a cada iteração, fornecendo para rede a cada época imagens transformadas de maneira diferente. Neste trabalho para realizar o *pipeline* de *data augmentation*, foi utilizado a ferramenta *Albumentations* (BUSLAEV *et al.*, 2020). As funções de transformações adotadas para compor o *pipeline* foram: jitter de cor que aplica até 20% de modificação

no brilho, contraste, saturação e matriz de cor; distorção em grade com limite de $\pm 30\%$; deslocamento, escalonamento e rotação com limites são iguais a $\pm 25\%$, $\pm 50\%$ e $\pm 90^\circ$ respectivamente; espelhamento vertical e horizontal. Todas as transformações são aplicadas aleatoriamente com probabilidades iguais a 50% de ocorrerem aleatoriamente entre os limites determinados de cada uma.

A função de perda (do inglês, loss function) é a função que calcula a distância entre a saída atual do modelo e a saída esperada, utilizada para otimizar o modelo. A função de perda utilizada foi a de entropia cruzada (do inglês, cross-entropy loss) com pesos para cada categoria, esta abordagem auxiliara na aprendizagem mesmo sendo o dataset desbalanceado (ZHANG, Z.; SABUNCU, 2018). Os pesos escolhidos foram calculados a partir das distribuições das instâncias de cada categoria, portanto o peso ω de cada categoria C , sendo a_C o número total de anotações com a categoria C , e a_t o número total de anotações, os pesos podem ser determinado pela equação $\omega_C = 1 - a_C/a_t$.

Os modelos utilizados foram pré-treinados no já estabelecido dataset, *imagenet* (RUSSAKOVSKY *et al.*, 2015). O *imagenet* consiste em mais de 14 milhões de imagens anotadas manualmente com mais de 20 mil categorias. Após os modelos serem exaustivamente treinados no *imagenet*, é possível utilizar esse “conhecimento” adquirido como um ponto de partida para novos treinamentos em outros conjuntos de dados (KORNBLITH *et al.*, 2019). Esta técnica de transferir o aprendizado (do inglês, transfer learning), permite que os modelos convirjam realizando treinamento por menos épocas e utilizando datasets menores.

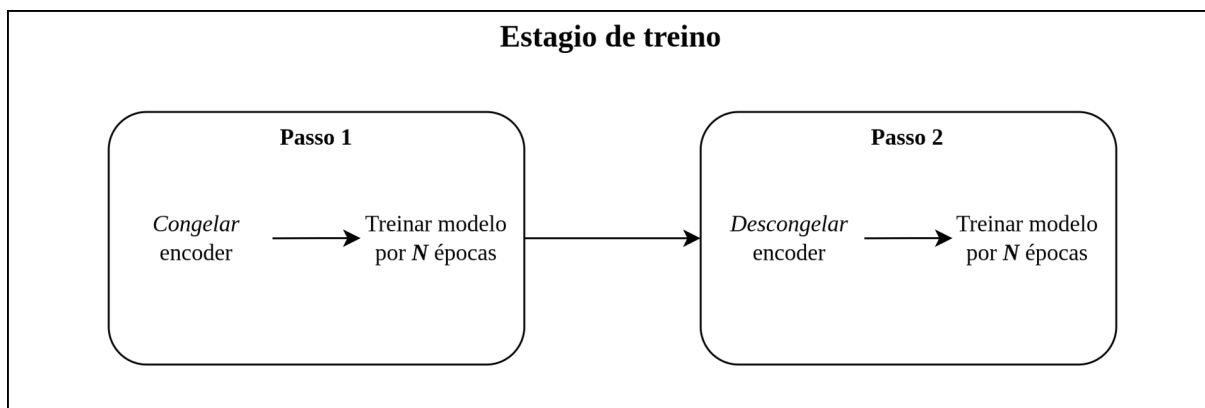
Outro método utilizado para não demandar longos períodos de treinamento foi a utilização de um “agendador”/função para determinar o hiperparâmetro de taxa de aprendizado (do inglês, learning rate). O “agendador” (do inglês, learning rate scheduler) utilizado foi o *one cycle scheduler*. Esta função de um ciclo, combina a taxa de aprendizado de uma taxa de aprendizado inicial para uma taxa de aprendizado máxima e, em seguida da taxa de aprendizado máxima para uma taxa de aprendizado mínima muito inferior à taxa de aprendizado inicial (SMITH; TOPIN, 2017).

Por fim, conforme descrito na figura 7, para a U-net foi utilizado três estágios de treinamento e para a HRNet e SegFormer apenas um estágio. Uma visão geral do que é referido como um estágio de treino pode ser observado na figura 9. Para a U-net foi utilizado três estágios para utilizar uma estratégia de treinamento com multi-resoluções (25%, 50% e 100% do tamanho original das imagens respectivamente) que permite o modelo aprender diferentes características das imagens em cada estágio (HOWARD; GUGGER, 2017). A HRNet e o SegFormer, não necessitam desta abordagem por apresentarem mecanismos de multi-resoluções dentro de suas arquiteturas.

Cada estágio de treino, conforme mostra a figura 9, é separado em duas etapas. Onde primeiro é treinado apenas o decoder do modelo por N épocas, e posteriormente

a arquitetura por inteiro (encoder e decoder) é treinada. A U-net foi treinada por 50 épocas para cada etapa ($N = 50$), totalizando 300 épocas ao final dos três estágios. A HRNet e o SegFormer, foram treinados por 150 épocas para cada etapa ($N = 150$), também totalizando 300 épocas ao final do processo de treinamento.

Figura 9 – Diagrama sobre do estágio de treino.



Fonte: Autor.

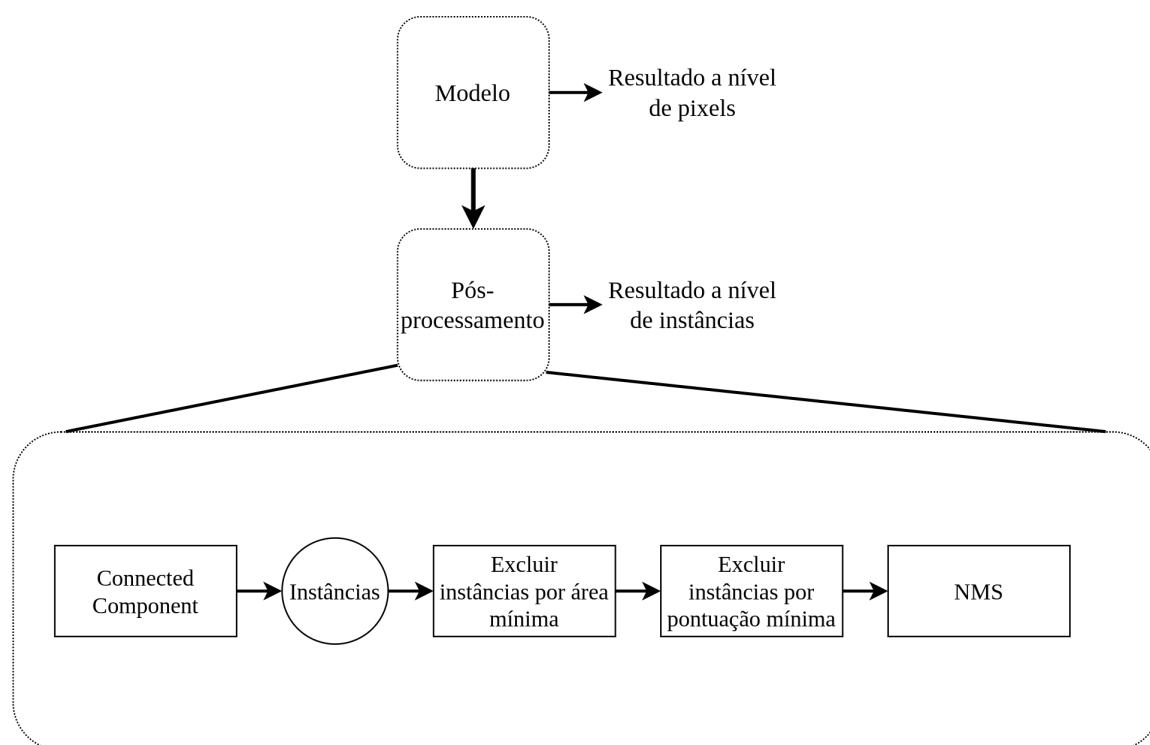
3.3.3 Pós-processamento

As saídas dos modelos de segmentação semântica são ao nível de píxel. Para transformar estas saídas para resultados de instâncias, onde há apenas uma categoria associada a uma determinada região, é necessário realizar um pós-processamento. Isto foi realizado utilizando um simples pipeline de algoritmos clássicos de VC, como mostra a figura 10.

Primeiramente, utiliza-se o algoritmo *connected components Spaghetti Labeling* (BOLELLI *et al.*, 2020) para encontrar as instâncias de cada categoria. Em seguida, para filtrar estas instâncias, todos os casos que tinham uma área menor que um valor de limiar em píxeis é excluído: 500 píxeis para Núcleo, Núcleo fora de foco e Núcleos sobrepostos; 40 píxeis para NORs (Clusters e Satélites); 200 píxeis para núcleo não viável e leucócitos.

Estes limiares de áreas de cada categorias foram adotados baseado nas estatísticas de todas as anotações do dataset. Além disso, as anotações são filtradas com base na pontuação (do inglês, score) do modelo, utilizando como limiar para instâncias pelo menos 50% de score, onde o score de cada instância é a média entre as pontuações de cada píxel daquela instância. Por fim, como último processo do pipeline, usamos o algoritmo de "supressão não-máxima" (do inglês, non-maximum suppression (NMS)), com limiar de 70%.

Figura 10 – Diagrama sobre o pós-processamento.



Fonte: Autor.

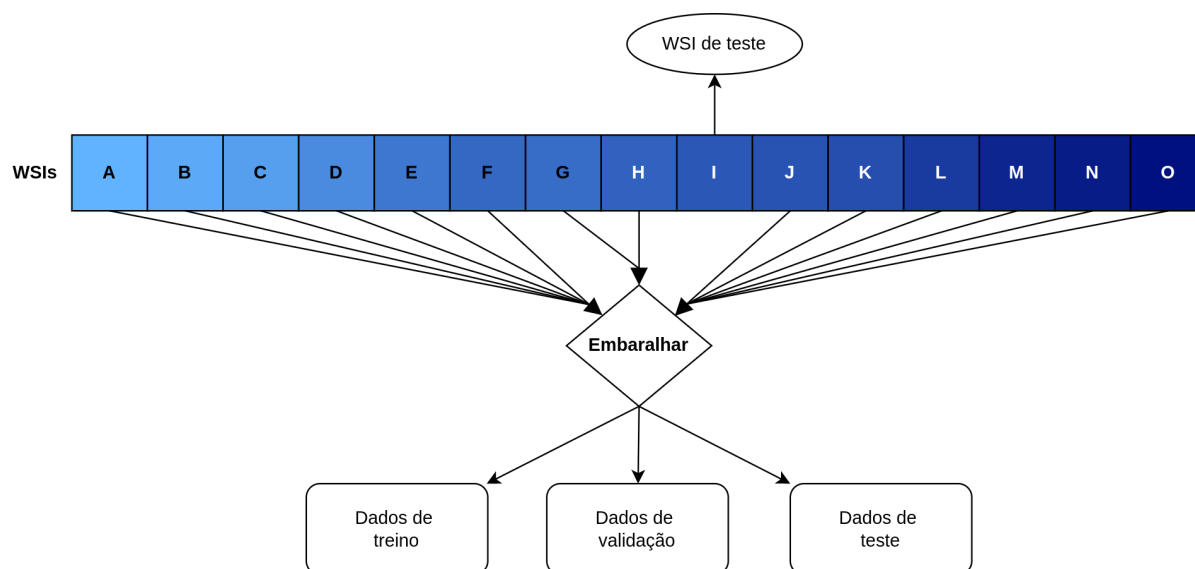
3.4 AVALIAÇÃO DE RESULTADOS

Para avaliar os modelos de maneira correta, utilizou-se a técnica de subamostragem aleatória. Para isto, os dados foram separados em quatro conjuntos, seguindo os passos ilustrados na figura 11. Primeiramente, foi separado todas as imagens pertencentes a uma mesma lâmina para utilizar-se como um segundo conjunto de teste, chamada de WSI de teste. Posteriormente, as imagens das lâminas restantes foram embaralhadas e separadas em três conjuntos de dados de maneira que todos possuam uma distribuição similar das instâncias de categorias. Conforme ilustrado na figura 11, estes conjuntos de dados são chamados treino, validação e teste que possuem 70%, 15% e 15% respectivamente dos dados das quatorze lâminas.

O conjunto de treino e validação foram utilizados durante o treinamento, sendo o conjunto de treino apenas para treinar os modelos, e o conjunto de validação para calcular as métricas e verificar o desempenho da rede ao longo do treinamento. Após ser finalizado o processo de treinamento dos modelos, eles são avaliados por métricas no conjunto de teste e nas imagens da lâmina de teste.

A avaliação dos modelos é realizada em ambos os níveis de resultado, primeiramente ao nível de píxel e posteriormente por nível de instâncias. As métricas utilizadas foram *Intersection Over Union (IoU)* e *F-score* (também conhecido como *Dice Similarity Coefficient (DSC)*) para os resultados em píxeis. Precisão (P), recall (R) e *F-score* para resultados ao nível de instâncias, todas as métricas são calculadas separadamente

Figura 11 – Diagrama da separação dos dados



Fonte: Autor.

para cada categoria (MATIAS *et al.*, 2021a).

Para descrever cada métrica, primeiro é necessário descrever os quatro parâmetros básicos usados para calcular essas métricas. Esses parâmetros estão listados abaixo:

- **Verdadeiro Positivo:** também conhecido como “hit”; do inglês True Positive (TP); quando o valor é verdadeiro e a previsão é verdadeira.
- **Verdadeiro Negativo:** também conhecido como “rejeição correta”; do inglês True Negative (TN); quando o valor é falso e a previsão é falsa.
- **Falso Positivo:** do inglês False Positive (FP); quando o valor é falso, porém a previsão é positiva.
- **Falso Negativo:** do inglês False Negative (FN); quando o valor é positivo, porém a previsão é falsa.

Para o cálculo das métricas de instâncias, é necessário utilizar um limiar para definir o que é uma ocorrência, o limiar utilizado foi de 0,5 de IoU. As equações 1 e 2, respectivamente, das métricas de precisão e recall, utilizam esse valor de limiar.

$$P = \frac{TP}{FP + TP} \quad (1)$$

$$R = \frac{TP}{FN + TP} \quad (2)$$

As equações 3 e 4, respectivamente, para as métricas F-score e IoU, o mesmo limiar para a avaliação ao nível de instância é utilizado. Para avaliação em píxeis: n_{jj} é a ocorrência TP, onde o valor e a previsão do modelo são j ; n_{ij} é uma ocorrência FP, onde o valor é i , mas a previsão do modelo foi j ; n_{ji} é uma ocorrência FN, onde a previsão do modelo é i , mas o valor é j . Para ambos os níveis de resultados, a mesma ferramenta de avaliação (*LAPiX DL*) para comparar os resultados foi utilizada (MATIAS; AMORIM, 2022).

$$\begin{aligned} DSC &= 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2 \times TP}{FP + FN + 2TP} \\ &= 2 \times \sum_{j=1}^k \frac{n_{jj}}{n_{ij} + n_{ji} + 2n_{jj}}, \quad i \neq j \end{aligned} \quad (3)$$

$$IoU = \frac{TP}{FP + FN + TP} = \sum_{j=1}^k \frac{n_{jj}}{n_{ij} + n_{ji} + n_{jj}}, \quad i \neq j \quad (4)$$

4 RESULTADOS E DISCUSSÃO

Neste estudo, foram testados modelos de segmentação SOTA de DL para identificar e caracterizar NORs em amostras de citologia cervical corados com prata. Nesta seção, primeiro será analisado o desempenho dos modelos treinados no conjunto de teste para o subdataset A (SA), mostrado nas tabelas 4 e 5, e para o subdataset B (SB), mostrado na tabela 6. Em seguida, os melhores resultados de cada modelo, e também as métricas ao nível de píxel para os melhores modelos. Então é apresentado os resultados avaliados no WSI de teste para averiguar se o modelo conseguiu generalizar seu desempenho. Por fim a discussão geral dos resultados.

A tabela 4 apresenta a comparação de U-net com ResNet34, HRNet-18, HRNet-32, SegFormer-B1 e SegFormer-B3 em SA para as métricas de precisão (P), recall (R) e F-score (DSC) no nível de instâncias. Na tabela 5 é apresentado a média entre todas as categorias e as categorias principais (núcleos e NORs). Nesta tabela, os modelos SOTA (HRNet e SegFormer) apresentam melhor desempenho quando comparados ao modelo U-net por uma grande margem (cerca de 0,2 em valores de DSC).

Modelo	Métrica	Núcleos	Cluster	Satélite	Núcleo fora de foco	Núcleos sobrepostos	Núcleo não-viável	Leucócitos
U-net+	P	0,64	0,57	0,35	0,32	0,27	0,02	0,00
	R	0,56	0,78	0,70	0,21	0,01	0,03	0,05
ResNet34	DSC	0,70	0,63	0,27	0,03	0,05	0,03	0,00
	P	0,82	0,82	0,57	0,45	0,41	0,32	0,45
HRNet-18	R	0,83	0,63	0,43	0,45	0,36	0,22	0,44
	DSC	0,83	0,71	0,49	0,45	0,38	0,26	0,45
HRNet-32	P	0,83	0,83	0,55	0,46	0,35	0,27	0,40
	R	0,65	0,83	0,63	0,48	0,39	0,37	0,28
SegFormer B1	DSC	0,83	0,72	0,51	0,42	0,36	0,27	0,41
	P	0,82	0,81	0,58	0,44	0,41	0,23	0,45
SegFormer B3	R	0,83	0,62	0,32	0,44	0,30	0,21	0,35
	DSC	0,83	0,70	0,41	0,44	0,34	0,22	0,39
SegFormer B3	P	0,83	0,84	0,54	0,47	0,47	0,30	0,45
	R	0,57	0,83	0,59	0,30	0,47	0,31	0,26
	DSC	0,83	0,69	0,38	0,47	0,37	0,28	0,42

Tabela 4 – Resultado ao nível de instâncias de cada modelo treinado com todas as categorias (SA)

No SA o SegFormer e HRNet apresentaram desempenhos aproximados para núcleos (diferença inferior a 0,01), núcleos sobrepostos, núcleos não viáveis e leucócitos. O SegFormer foi cerca de 0,04 melhor na classificação de núcleos fora de foco, considerando a métrica DSC. Além disso, a HRNet teve melhor desempenho na detecção de NORs, apresentando também uma diferença razoável para a categoria

satélite (cerca de 0,13), sendo o melhor modelo com DSC médio entre as categorias principais de 0,69 e no DSC médio entre todas as categorias foi de 0,5.

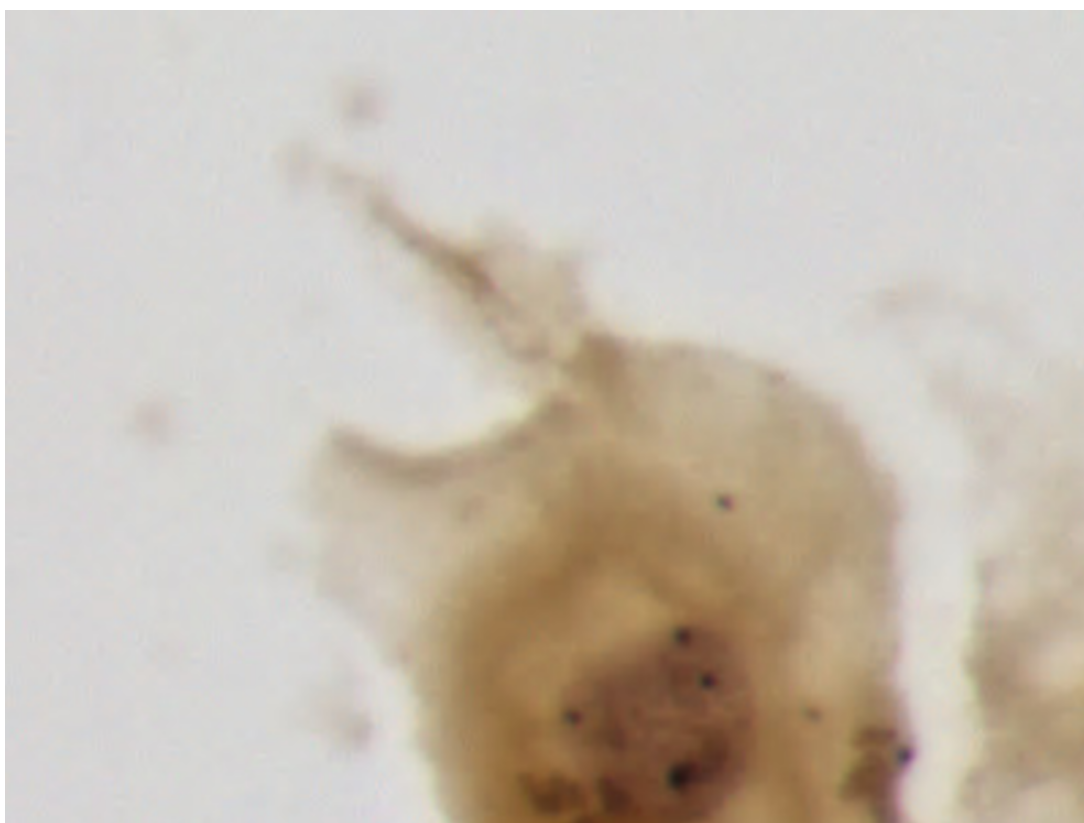
Modelo	Métrica	Médias entre as categorias principais	Média
U-net+	P	0,52	0,31
ResNet34	R	0,68	0,34
	DSC	0,53	0,24
HRNet-18	P	0,74	0,55
	R	0,63	0,48
	DSC	0,68	0,51
HRNet-32	P	0,74	0,53
	R	0,71	0,52
	DSC	0,69	0,50
SegFormer B1	P	0,74	0,53
	R	0,59	0,44
	DSC	0,65	0,48
SegFormer B3	P	0,74	0,56
	R	0,67	0,48
	DSC	0,63	0,49

Tabela 5 – Média dos resultados ao nível de instâncias de cada modelo treinado com todas as categorias (SA)

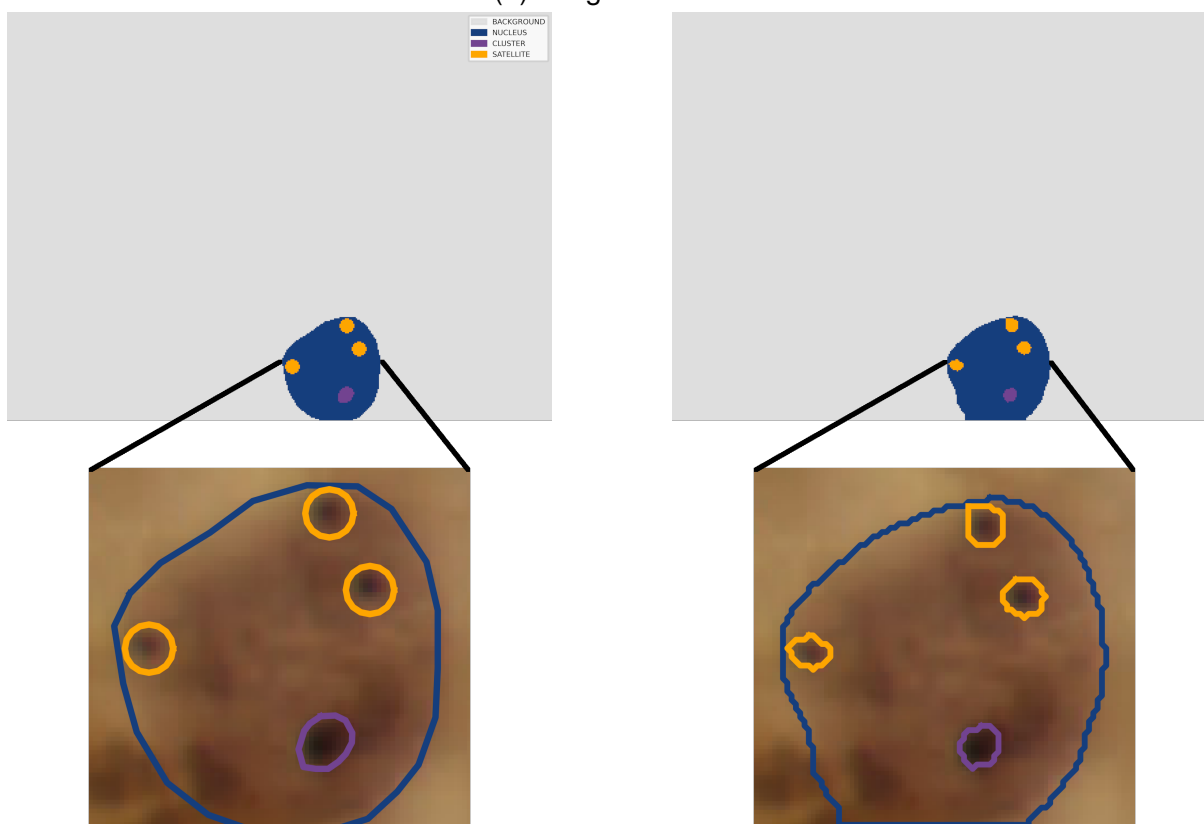
A tabela 6 apresenta a comparação entre os mesmos modelos e métricas no nível de instância, que na tabela 4, porém são resultados com o SB. Esses resultados mostram que o U-net teve um desempenho cerca de 0,10 pior que os outros modelos para todas as métricas. Novamente, o SegFormer foi melhor para detectar núcleos e núcleos fora de foco, mas teve um desempenho pior para detecção de NORs. Os modelos em SB apresentam valores mais baixos para recall e valores mais altos para precisão e DSC, quando comparados aos mesmos modelos aplicados em SA. Os resultados entre as categorias principais mostram que o HRNet teve um desempenho melhor em cerca de 0,01 no SB comparado ao SA.

Os resultados da HRNet-32, que apresentou melhores resultados em SA e manteve-os em SB, ao nível de píxel é apresentado na tabela 7 para as métricas IoU e DSC de cada categoria. O IoU médio é de 0,54 e 0,55 para SA e SB, respectivamente, considerando as principais categorias. O DSC médio entre as principais categorias é 0,69 e 0,70 para SA e SB. O DSC para satélites foi de 0,49 e 0,53, e para clusters 0,74 e 0,72, respectivamente para SA e SB. Um exemplo de resultado é apresentado na figura 12, com a figura 12b mostrando o que foi manualmente anotado e a figura 12c mostrando a previsão da HRNet após o pós-processamento.

O WSI de teste foi de uma paciente diagnosticada sem lesão. Este WSI tem 242, 310, 113 e 29 anotações, respectivamente, de núcleos, clusters, satélites e núcleos fora de foco. Os resultados na tabela 8 apresentam um desempenho com queda



(a) Imagem do SB



(b) Manualmente anotada

(c) Predição

Figura 12 – Exemplo de resultado obtido pela HRNet-32 após pós-processamento ao nível de instâncias e de pixels

Modelo	Métrica	Núcleos	Cluster	Satélite	Núcleo fora de foco	Médias entre as categorias principais
Unet+	P	0,79	0,76	0,36	0,02	0,64
	R	0,79	0,76	0,36	0,02	0,64
	DSC	0,72	0,61	0,40	0,04	0,58
ResNet34	P	0,85	0,79	0,62	0,43	0,76
	R	0,84	0,61	0,50	0,41	0,65
	DSC	0,85	0,69	0,55	0,42	0,70
HRNet-18	P	0,85	0,79	0,61	0,43	0,75
	R	0,85	0,59	0,52	0,46	0,65
	DSC	0,85	0,68	0,56	0,44	0,69
HRNet-32	P	0,84	0,78	0,57	0,45	0,73
	R	0,84	0,56	0,37	0,47	0,59
	DSC	0,84	0,65	0,45	0,46	0,65
SegFormer B1	P	0,85	0,79	0,57	0,45	0,74
	R	0,85	0,58	0,42	0,44	0,62
	DSC	0,85	0,67	0,48	0,45	0,67

Tabela 6 – Resultado ao nível de instâncias de cada modelo treinado com categorias principais e núcleos fora de foco (SB)

		Núcleos	Cluster	Satélite	Núcleo fora de foco	Núcleos sobrepostos	Núcleo não-viável	Leucócitos
SA	IoU	0,71	0,58	0,32	0,33	0,39	0,18	0,31
	DSC	0,83	0,74	0,49	0,50	0,56	0,30	0,47
SB	IoU	0,74	0,56	0,36	0,33	-	-	-
	DSC	0,85	0,72	0,53	0,50	-	-	-

Tabela 7 – Resultado ao nível de pixels do modelo HRNet-32 no SA e SB

aproximada de 0,05 para o DSC médio entre as categorias principais, em comparação ao conjunto de teste. Além disso, os resultados obtidos com este WSI indicam uma queda de cerca de 0,12 para DSC na detecção de instâncias de NORs.

4.1 DISCUSSÃO

Para esta dissertação foi analisado o estado da arte de modelos de DL para visão computacional, assim como para quais modelos são utilizados em diversos trabalhos para imagens citológicas. Utilizando o conjunto de dados criado, realizando os experimentos descritos para identificação e caracterização de imagens coradas com prata (técnica AgNOR), foi identificado o modelo mais adequado para esta categoria de imagens. Contudo, visto que um dos principais itens deste trabalho era identificar e caracterizar NORs, poderia, em teoria, ter também sido explorado a abordagem de detecção de objetos. Todavia, a detecção de objetos pequenos e/ou minúsculos

	SA			SB		
	P	R	DSC	P	R	DSC
Núcleos	0,94	0,91	0,93	0,96	0,89	0,93
Cluster	0,84	0,40	0,55	0,83	0,41	0,55
Satélite	0,58	0,30	0,40	0,49	0,39	0,44
Núcleo fora de foco	0,58	0,66	0,61	0,51	0,76	0,61
Médias entre as categorias principais	0,79	0,54	0,62	0,76	0,56	0,64

Tabela 8 – Resultado ao nível de instâncias do modelo HRNet-32 no SA e SB na WSI de teste.

em imagens utilizando DL, é uma tarefa extremamente desafiadora, especialmente ao usar modelos de detecção de objetos, que não tem um bom desempenho nesta tarefa (TONG; WU, 2022). Quando elaborado os trabalhos (AMORIM *et al.*, 2020c, 2022b) também foi experimentado a tentativas sem sucesso da utilização de detecção de objetos para a detecção de NORs. Por esse motivo, foi selecionado a segmentação semântica como o método para a identificação de satélites e clusters individuais, pois tem melhores resultados ao se trabalhar com objetos muito pequenos.

O conjunto de dados criado ao longo do desenvolvimento desta dissertação, entre todos os trabalhos averiguados, mostrou-se ser o único disponível publicamente de imagens citológicas coradas com prata. Apesar de ter sido identificado ao longo do processo outras categorias para ser anotada (evolução da versão 1 para a versão 2 do dataset CCAgT) para cobrir todos os objetos existentes em imagens coradas por AgNOR e, conseqüentemente, minimizar detecções erradas. Entretanto, os resultados mostraram que usar apenas as categorias principais e a categoria de núcleo fora de foco permite que os modelos tenham resultado similares na detecção dos NORs.

Apesar de o dataset ter sido ampliado e a sua segunda versão apresentar dados de quinze pacientes com sete diferentes diagnósticos, o dataset apresenta um desbalanceamento entre a quantidade das instâncias de satélites e clusters. Este desbalanceamento entre as categorias de NORs pode ter causado a diferença dos resultados obtidos entre satélites e clusters. Estas diferenças nos resultados também podem estar associadas a maneira que os satélites foram anotados, porque um ponto único não têm a precisão da representação ao nível de píxel. Essas anotações menos precisas causam valores de erro enganosos para os píxeis anotados como satélite, mas, na verdade, são píxeis de núcleo. Isso pode ser visto comparando os resultados entre as tabelas 4 e 7, e também na figura 12b em comparação com 12c.

A qualidade do scanner, da câmera e, conseqüentemente, das imagens também podem ser consideradas como fatores que afetam o desempenho dos modelos.

Em conjuntos de dados públicos de citologia amplamente utilizados, as imagens são geralmente capturadas manualmente para cada célula com alta resolução. O conjunto de dados CCAgT tem apenas WSI digitalizadas usando foco automático e zoom fixo. A configuração de aquisição de imagem que adotamos para imagens coradas por AgNOR dificulta a análise da imagem porque os ajustes de foco e resolução de objetos muito pequenos, principalmente satélites, são difíceis de automatizar. A falta de foco correto e de resolução adequada para esses objetos muito pequenos resulta em objetos sem uma definição detalhada de textura e formato. No entanto, é considerado importante utilizar tais imagens nesse estudo, pois, amostras com tais níveis de qualidade de aquisição refletem melhor a qualidade que se espera em amostras que seriam adquiridas na rotina clínica, tornando nossos resultados mais realistas para uma possível aplicação clínica futura.

Outro aspecto que foi possível observar a partir dos resultados, é que os modelos que empregam mecanismos de *attention* (HRNet e SegFormer) mostraram desempenho consideravelmente melhor, quando comparados com o modelo U-net mais tradicional. Isso pode indicar que o mecanismo de atenção em geral pode melhorar os modelos DL ao trabalhar com objetos muito pequenos. É importante observar aqui, entretanto, que os mecanismos de atenção empregados pela HRNet e SegFormer diferem consideravelmente, ainda que tenham raízes conceituais semelhantes.

Embora a tentativa de compensar a falta de uma abordagem multi-resolução da U-net usando a estratégia de várias etapas no treinamento (que funciona melhor do que apenas treinar ela com uma resolução fixa), os modelos SOTA superaram a U-net com ResNet por uma grande margem. O modelo SegFormer-B3 apresentou resultados inesperados, cerca de 0,04 pior que o HRNet no nível de instância, mostrando que os modelos *transformers* ainda não são melhores que CNN em alguns casos. Considerando os modelos testados, os resultados indicam que o HRNet é o modelo mais adequado para a segmentação de objetos muito pequenos. Sendo os NORs, objetos muito pequenos e também os itens mais importantes para realizar a caracterização de amostras coradas com prata, a HRNet mostrou ser o modelo mais ideal para trabalhar com imagens de amostras cervicais coradas com a técnica AgNOR. Os artigos publicados ao decorrer da dissertação encontram-se listado no apêndice A.

5 CONCLUSÃO

Neste trabalho foi apresentado um estudo comparativo entre diferentes modelos de *Deep Learning* para a identificação e caracterização de objetos corados com a técnica AgNOR. Os três modelos de segmentação semântica utilizados neste trabalho foram U-net, HRNet e SegFormer. Sendo que estes modelos utilizam diferentes mecanismos em suas arquiteturas, considerado a U-net como modelo base, a HRNet modelo convolucional que utiliza mecanismos de atenção, e por fim o SegFormer, um modelo totalmente baseado em mecanismo de atenção (conhecido como modelo *Transformer*).

Após a análise do estado da arte dos modelos de visão computacional, e também destes quais são amplamente aplicados a imagens citológicas. Foi adotado a utilização apenas da abordagem de segmentação semântica devido ao tamanho muito pequeno dos principais objetos (NORs). Observou-se também que não existem dados similares aos utilizados neste trabalho que estejam disponíveis de maneira pública. Desta maneira, uma das contribuições deste trabalho é a publicação do dataset criado.

O conjunto de dados *Cervical Cells with AgNOR Stain Technique - CCAgT* (AMORIM *et al.*, 2022a), conta com dados de quinze pacientes e 63190 anotações. O dataset conta com as principais categorias para a realização do diagnóstico, sendo elas núcleos e NORs, onde os NORs são separadas em atípicos (clusters) e típicos (satélites). Outras categorias anotadas para abranger os objetos mais recorrentes nas lâminas coradas com a técnica de AgNOR são os núcleos fora de foco, núcleos sobrepostos, núcleos não viáveis e leucócitos.

Foram apresentados os resultados e delimitações de cada modelo com dois subconjuntos de dados. O modelo HRNet obteve melhores resultados para as categorias muito pequenas, e o modelo SegFormer obteve melhores resultados para as demais categorias. Ambos tiveram um desempenho melhor do que o U-net com ResNet como backbone.

Utilizando um simples pipeline de técnicas clássicas como pós-processamento, foi possível extrair instâncias a partir dos resultados obtidos pelos modelos de segmentação semântica com sucesso. A HRNet-32 mostrou-se o melhor modelo para trabalhar com imagens coradas com a técnica AgNOR com resolução de $0,111\mu m \times 0,111\mu m$ por píxel. Este modelo teve um desempenho de DSC (F-score) de 0,83 para núcleo, 0,72 para NORs atípicos - clusters, 0,51 para NORs típicos — satélite, 0,42 para núcleo fora de foco, 0,36 para núcleos sobrepostos, 0,27 para núcleos inviáveis e 0,41 para leucócitos em nível de instância. Os experimentos mostraram que treinar os modelos apenas com as categorias principais (núcleo e NORs) e o núcleos fora de foco pode aumentar o desempenho da detecção de satélites em cerca de 0,05. Usando esta configuração, o modelo também obteve um DSC médio de 0,69 para as principais

categorias, uma melhora inferior a 0,01.

Os melhores resultados foram 0,85 de precisão, 0,85 de recall e 0,85 de F-score na detecção de núcleos corado com prata (AgNOR). Para a detecção de NORs, os resultados médios entre as amostras típicas e atípicas foram 0,70 de precisão, 0,55 de recall e 0,62 de F-score. Ao nível de pixels, os resultados da segmentação em DSC foram 0,85 para pixels de núcleos e 0,62 para os pixels de NORs.

Concluindo, este trabalho apresentou a possibilidade de utilização de modelos de visão computacional de última geração para a segmentação de estruturas celulares muito pequenas, assim como um novo dataset. Espera-se que estes resultados ajudem a difundir o AgNOR como método coadjuvante ao diagnóstico de câncer, visto que ainda não é um método padronizado entre os patologistas. Além disso, a automação da detecção de NORs por meio de métodos de deep learning possibilita sua utilização em sistemas de telemedicina, diminuindo o tempo necessário para analisar uma WSI e aumentando a precisão dos diagnósticos.

5.1 TRABALHOS FUTUROS

Apesar deste trabalho utilizar apenas amostras cervicais, a técnica AgNOR pode ser utilizada com células de diferentes partes do corpo e, em alguns casos, esses outros tipos de células apresentam semelhanças com as do colo do útero. Portanto, este trabalho também pode ser usado como base para identificar núcleos e NORs nestas amostras.

Para melhores resultados considerando os mesmos modelos, a otimização do conjunto de dados mostra-se uma opção. Principalmente na maneira das anotações dos satélites, para garantir que haja apenas pixels de satélites (tons escuros e/ou pretos) em suas anotações.

Outra via para evoluções futuras, é abordagens de outros modelos, que principalmente possuem capacidade de lidar com detalhes finos ou objetos minúsculos. Assim como testar abordagens em cascata entre modelos de detecção e segmentação, com um modelo especialista em detectar núcleos e outro em caracterizá-lo.

REFERÊNCIAS

AL-ABBADI, MousaA. Basics of cytology. **Avicenna Journal of Medicine**, Medknow, v. 1, n. 1, p. 18, 2011. DOI: 10.4103/2231-0770.83719. Disponível em: <https://doi.org/10.4103/2231-0770.83719>.

AMORIM, João Gustavo Atkinson. **CCAgT-utils**. [S.l.: s.n.], fev. 2022. DOI: 10.5281/zenodo.6350442. Disponível em: <https://github.com/johnnv1/CCAgT-utils>.

AMORIM, João Gustavo Atkinson; CERENTINI, Allan; MACARINI, Luiz Antonio Buschetto; MATIAS, André Victória; WANGENHEIM, Aldo von. **Systematic Literature Review of Computer Vision-Aided Cytology - A Review of Classic Computer Vision and Deep Learning-Based Approaches published between January/2016 - March/2020**. en. [S.l.], 2020a. DOI: 10.13140/RG.2.2.13304.67840. Disponível em: <http://rgdoi.net/10.13140/RG.2.2.13304.67840>.

AMORIM, João Gustavo Atkinson; MACARINI, Luiz Antonio Buschetto; MATIAS, André Victória; CERENTINI, Allan; ONOFRE, Fabiana Botelho De Miranda; ONOFRE, Alexandre Sherlley Casimiro; WANGENHEIM, Aldo von. A Novel Approach on Segmentation of AgNOR-Stained Cytology Images Using Deep Learning. *In*: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS). [S.l.: s.n.], 2020b. P. 552–557. DOI: 10.1109/CBMS49503.2020.00110. Disponível em: <https://doi.org/10.1109/CBMS49503.2020.00110>.

AMORIM, João Gustavo Atkinson; MACARINI, Luiz Antonio Buschetto; MATIAS, André Victória; CERENTINI, Allan; ONOFRE, Fabiana Botelho De Miranda; ONOFRE, Alexandre Sherlley Casimiro; WANGENHEIM, Aldo von. Detecção de núcleos em imagens citológicas de AgNOR utilizando Aprendizado Profundo. *In*. Disponível em: <http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/view/811>.

AMORIM, João Gustavo Atkinson; MATIAS, André Victória; BOTTAMEDDI, Tainee; SANCHES, Vinícius; COSTA, Ane Francyne; ONOFRE, Fabiana Botelho De Miranda; ONOFRE, Alexandre Sherlley Casimiro; WANGENHEIM, Aldo von. **CCAgT: Images of Cervical Cells with AgNOR Stain Technique**. [S.l.: s.n.], 2022a. Mendelay Data. DOI: 10.17632/wg4bpm33hj.2. Disponível em: <https://doi.org/10.17632/wg4bpm33hj.2>.

AMORIM, João Gustavo Atkinson; SANCHES, Viníus Moreno; BOTTAMEDDI, Tainee; MATIAS, André Victória; CAVACO, Marco; ONOFRE, Alexandre Sherlley Casimiro; ONOFRE, Fabiana Botelho De Miranda; WANGENHEIM, Aldo von. Nucleus detection in cervical samples stained with AgNOR. *In: ANAIS do XIII Computer on the Beach - COTB'22*. [S.l.: s.n.], jul. 2022b.

AN, Srivastava; S, Srivastava; C, Bansal; JS, Misra. Diagnostic importance of AgNOR pleomorphism in cervical carcinogenesis. **ecancer**, Ecancer Global Foundation, p. 287–293, jan. 2013. DOI: 10.3332/ecancer.2013.287. Disponível em: <https://doi.org/10.3332/ecancer.2013.287>.

BEDROSSIAN, Carlos W. M. Cytopathology: A global perspective. **Diagnostic Cytopathology**, Wiley, v. 35, n. 12, p. 745–748, 2007. DOI: 10.1002/dc.20773. Disponível em: <https://doi.org/10.1002/dc.20773>.

BIEWALD, Lukas. **Experiment Tracking with Weights and Biases**. [S.l.: s.n.], 2020. Software available from wandb.com. Disponível em: <https://www.wandb.com/>.

BISONG, Ekaba. Google Colaboratory. *In: BUILDING Machine Learning and Deep Learning Models on Google Cloud Platform*. [S.l.]: Apress, 2019. P. 59–64. DOI: 10.1007/978-1-4842-4470-8_7. Disponível em: https://doi.org/10.1007/978-1-4842-4470-8_7.

BOLELLI, Federico; ALLEGRETTI, Stefano; BARALDI, Lorenzo; GRANA, Costantino. Spaghetti Labeling: Directed Acyclic Graphs for Block-Based Connected Components Labeling. **IEEE Transactions on Image Processing**, Institute of Electrical e Electronics Engineers (IEEE), v. 29, p. 1999–2012, 2020. DOI: 10.1109/tip.2019.2946979. Disponível em: <https://doi.org/10.1109/tip.2019.2946979>.

BUSLAEV, Alexander; IGLOVIKOV, Vladimir I.; KHVEDCHENYA, Eugene; PARINOV, Alex; DRUZHININ, Mikhail; KALININ, Alexandr A. Albuementations: Fast and Flexible Image Augmentations. **Information**, MDPI AG, v. 11, n. 2, p. 125, fev. 2020. DOI: 10.3390/info11020125. Disponível em: <https://doi.org/10.3390/info11020125>.

CARMEN ALARCÓN-ROMERO, Luz del; ILLADES-AGUIAR, Berenice; FLORES-ALFARO, Eugenia; TERÁN-PORCAYO, Marco Antonio; ANTONIO-VÉJAR, Verónica; REYES-MALDONADO, Elba. AgNOR polymorphism

association with squamous intraepithelial lesions and invasive carcinoma with HPV infection. **Salud Pública de México**, FapUNIFESP (SciELO), v. 51, n. 2, p. 134–140, abr. 2009. DOI: 10.1590/s0036-36342009000200009. Disponível em: <https://doi.org/10.1590/s0036-36342009000200009>.

COHEN, Paul A; JHINGRAN, Anjua; OAKNIN, Ana; DENNY, Lynette. Cervical cancer. **The Lancet**, Elsevier BV, v. 393, n. 10167, p. 169–182, jan. 2019. DOI: 10.1016/s0140-6736(18)32470-x. Disponível em: [https://doi.org/10.1016/s0140-6736\(18\)32470-x](https://doi.org/10.1016/s0140-6736(18)32470-x).

CROCKER, John; BOLDY, David A. R.; EGAN, Mark J. How should we count AgNORS? Proposals for a standardized approach. **The Journal of Pathology**, v. 158, n. 3, p. 185–188, 1989a. DOI: <https://doi.org/10.1002/path.1711580303>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/path.1711580303>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/path.1711580303>.

CROCKER, John; BOLDY, David A. R.; EGAN, Mark J. How should we count AgNORS? Proposals for a standardized approach. **The Journal of Pathology**, Wiley, v. 158, n. 3, p. 185–188, jul. 1989b. DOI: 10.1002/path.1711580303. Disponível em: <https://doi.org/10.1002/path.1711580303>.

DARKWAH, Williams Kweku; AIDOO, Gideon; AO, Yanhui; DANQUAH, Kwabena Owusu; ADJEI, Ernest; ADANKWAH, Ernest; KADRI, Alhassan. Assessment of proliferative index in different grades of breast cancers using AgNOR (Agyrophilic Nuclear Organizer Region) expression. **Beni-Suef University Journal of Basic and Applied Sciences**, Springer Science e Business Media LLC, v. 7, n. 4, p. 587–592, dez. 2018. DOI: 10.1016/j.bjbas.2018.06.011. Disponível em: <https://doi.org/10.1016/j.bjbas.2018.06.011>.

DEMAY, R.M. **The Art and Science of Cytopathology**. [S.l.]: American Society for Clinical Pathology, 1996. ISBN 9780891894087.

DIAMANTIS, A.; MAGIORKINIS, E. Pioneers of exfoliative cytology in the 19th century: the predecessors of George Papanicolaou. **Cytopathology**, Wiley, v. 25, n. 4, p. 215–224, jun. 2013. DOI: 10.1111/cyt.12074. Disponível em: <https://doi.org/10.1111/cyt.12074>.

DOGENSKI, Letícia Copatti; FIGUEIREDO RIBEIRO, Sara de; GAMBIN, Diego José; MASO, Patrícia Canova; LINDEN, Maria Salete Sandini; TRENTIN, Micheline Sandini;

ZOEHLER, Bernardo; CRIVELINI, Marcelo Macedo; CARLI, João Paulo De. Oral leukoplakia—epidemiological survey and histochemical analysis of 107 cases in Brazil. **Clinical Oral Investigations**, Springer Science e Business Media LLC, v. 25, n. 4, p. 1859–1867, ago. 2020. DOI: 10.1007/s00784-020-03488-x. Disponível em: <https://doi.org/10.1007/s00784-020-03488-x>.

DU; LI, Xueyu; LI, Qinghua. Detection and Classification of Cervical Exfoliated Cells Based on Faster R-CNN. *In*: 2019 IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT). [S.l.]: IEEE, out. 2019. DOI: 10.1109/icaait.2019.8935931. Disponível em: <https://doi.org/10.1109/icaait.2019.8935931>.

EROZ, Recep. **Argyrophilic nucleolar organizing regions associated proteins in oncocytopology**. v. 5. [S.l.]: Edorium Journals Pvt. Ltd., 2015. P. 21–24. DOI: 10.5348/ocy-2015-10-ed-4. Disponível em: <https://doi.org/10.5348/ocy-2015-10-ed-4>.

FERLAY, J.; ERVIK, M.; LAM, F.; COLOMBET, M.; MERY, L.; PIÑEROS, M.; ZNAOR, A.; SOERJOMATARAM, I.; BRAY, F. **Global Cancer Observatory: Cancer Today**. [S.l.: s.n.], dez. 2020. Disponível em: <https://gco.iarc.fr/today>. Acesso em: 19 jan. 2021.

GARCÍA-VIELMA, C; DÁVILA-RODRÍGUEZ, MI; HERNÁNDEZ-GARZA, F; CERDA-FLORES, RM; CORTÉS-GUTIÉRREZ, EI. Digital image analysis of AgNORs in cervical smears of women with premalignant and malignant lesions of the uterine cervix. **Biotechnic & Histochemistry**, Informa UK Limited, v. 91, n. 2, p. 102–107, nov. 2015. DOI: 10.3109/10520295.2015.1078909. Disponível em: <https://doi.org/10.3109/10520295.2015.1078909>.

GONZALEZ, Rafael C.; WOODS, Richard E. **Digital image processing**. Upper Saddle River, N.J.: Prentice Hall, 2008. ISBN 9780131687288.

HAYKIN, Simon S. **Neural networks and learning machines**. Third. Upper Saddle River, NJ: Pearson Education, 2009.

HOWARD, Jeremy; GUGGER, Sylvain. Fastai: A Layered API for Deep Learning. **Information**, MDPI AG, v. 11, n. 2, p. 108, fev. 2020. DOI: 10.3390/info11020108. Disponível em: <https://doi.org/10.3390/info11020108>.

HOWARD, Jeremy; GUGGER, Sylvain. **Practical Deep Learning for Coders (fast.ai courses) - V1**. v. 1. [S.l.]: fast.ai courses, 2017. Disponível em:

<https://github.com/fastai/courses/>.

HUANG, Yongfeng; LI, Xueyang; YAN, Cairong; LIU, Lihao; DAI, Hao. MIRD-Net for Medical Image Segmentation. *In: ADVANCES in Knowledge Discovery and Data Mining*. [S.l.]: Springer International Publishing, 2020. P. 207–219. DOI:

10.1007/978-3-030-47436-2_16. Disponível em:

https://doi.org/10.1007/978-3-030-47436-2_16.

HUSSAIN, Elima; MAHANTA, Lipi B.; DAS, Chandana Ray; CHOUDHURY, Manjula; CHOWDHURY, Manish. A shape context fully convolutional neural network for segmentation and classification of cervical nuclei in Pap smear images. **Artificial Intelligence in Medicine**, Elsevier BV, v. 107, p. 101897, jul. 2020. DOI:

10.1016/j.artmed.2020.101897. Disponível em:

<https://doi.org/10.1016/j.artmed.2020.101897>.

IVANOVIC, M. Overview of Cytopathology Procedures and Techniques. *In: CANCER Treatment and Research*. [S.l.]: Springer Berlin Heidelberg, out. 2013. P. 1–12. DOI:

10.1007/978-3-642-38850-7_1. Disponível em:

https://doi.org/10.1007/978-3-642-38850-7_1.

KHAN, Salman; NASEER, Muzammal; HAYAT, Munawar; ZAMIR, Syed Waqas; KHAN, Fahad Shahbaz; SHAH, Mubarak. Transformers in Vision: A Survey. **ACM Computing Surveys**, Association for Computing Machinery (ACM), jan. 2022. DOI:

10.1145/3505244. Disponível em: <https://doi.org/10.1145/3505244>.

KHOT, K. P.; KOTAK, K. J.; BAGRI-MANJREKAR, K.; ALEX, S.; CHOUDHARI, S. S. Pleomorphism of argyrophilic nucleolar organizer regions in oral submucous fibrosis and oral squamous cell carcinoma. **J Nat Sci Biol Med**, v. 6, Suppl 1, s35–39, ago. 2015. DOI: 10.4103/0976-9668.166064. Disponível em:

10.4103/0976-9668.166064. Disponível em:

<https://doi.org/10.4103/0976-9668.166064>.

KHUSHBU, BP.; CHALISHAZAR, M.; KALE, H.O; BARANWAL, M.; MODI, T.

Quantitative and qualitative assessment of argyrophilic nucleolar organizer regions in normal, premalignant and malignant oral lesions. **J Oral Maxillofac Pathol**, v. 21, p. 360–366, 2017. DOI: 10.4103/jomfp.JOMFP_52_15. Disponível em:

https://dx.doi.org/10.4103/jomfp.JOMFP_52_15.

KORNBLITH, Simon; SHLENS, Jonathon; LE, Quoc V. Do Better ImageNet Models Transfer Better? *In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.]: IEEE, jun. 2019. DOI: 10.1109/cvpr.2019.00277. Disponível em: <https://doi.org/10.1109/cvpr.2019.00277>.

LABELBOX. **Labelbox**. [S.l.: s.n.], 2021. Disponível em: <https://labelbox.com>. Acesso em: 28 jan. 2021.

LEI, Jiayao; PLONER, Alexander; ELFSTRÖM, K. Miriam; WANG, Jiangrong; ROTH, Adam; FANG, Fang; SUNDSTRÖM, Karin; DILLNER, Joakim; SPARÉN, Pär. HPV Vaccination and the Risk of Invasive Cervical Cancer. **New England Journal of Medicine**, Massachusetts Medical Society, v. 383, n. 14, p. 1340–1348, out. 2020. DOI: 10.1056/nejmoa1917338. Disponível em: <https://doi.org/10.1056/nejmoa1917338>.

LI, Huamin; SHAHAM, Uri; STANTON, Kelly P; YAO, Yi; MONTGOMERY, Ruth R; KLUGER, Yuval. Gating mass cytometry data by deep learning. Edição: Janet Kelso. **Bioinformatics**, Oxford University Press (OUP), v. 33, n. 21, p. 3423–3430, jul. 2017. DOI: 10.1093/bioinformatics/btx448. Disponível em: <https://doi.org/10.1093/bioinformatics/btx448>.

LIN, Ching-Kai; CHANG, Jerry; HUANG, Ching-Chun; WEN, Yueh-Feng; HO, Chao-Chi; CHENG, Yun-Chien. Effectiveness of convolutional neural networks in the interpretation of pulmonary cytologic images in endobronchial ultrasound procedures. **Cancer Medicine**, Wiley, v. 10, n. 24, p. 9047–9057, nov. 2021. DOI: 10.1002/cam4.4383. Disponível em: <https://doi.org/10.1002/cam4.4383>.

MATIAS, André Victória; AMORIM, João Gustavo Atkinson. **LAPiX DL – Utils for Computer Vision Deep Learning research**. [S.l.: s.n.], fev. 2022. DOI: 10.5281/zenodo.5963342. Disponível em: <https://github.com/lapix-ufsc/lapixdl>.

MATIAS, André Victória; AMORIM, João Gustavo Atkinson; MACARINI, Luiz Antonio Buschetto; CERENTINI, Allan; ONOFRE, Alexandre Sherlley Casimiro; ONOFRE, Fabiana Botelho De Miranda; DALTOÉ, Felipe Perozzo; STEMMER, Marcelo Ricardo; WANGENHEIM, Aldo von. What is the state of the art of computer vision-assisted cytology? A Systematic Literature Review. **Computerized Medical Imaging and Graphics**, Elsevier BV, v. 91,

p. 101934, jul. 2021a. DOI: 10.1016/j.compmedimag.2021.101934. Disponível em: <https://doi.org/10.1016/j.compmedimag.2021.101934>.

MATIAS, André Victória; CERENTINI, Allan; MACARINI, Luiz Antonio Buschetto; AMORIM, João Gustavo Atkinson; DALTOÉ, Felipe Perozzo; WANGENHEIM, Aldo von. Segmentation, Detection, and Classification of Cell Nuclei on Oral Cytology Samples Stained with Papanicolaou. **SN Computer Science**, Springer Science e Business Media LLC, v. 2, n. 4, mai. 2021b. DOI: 10.1007/s42979-021-00676-8. Disponível em: <https://doi.org/10.1007/s42979-021-00676-8>.

MEIJERING, Erik. Cell Segmentation: 50 Years Down the Road [Life Sciences]. **IEEE Signal Processing Magazine**, Institute of Electrical e Electronics Engineers (IEEE), v. 29, n. 5, p. 140–145, set. 2012. DOI: 10.1109/msp.2012.2204190. Disponível em: <https://doi.org/10.1109/msp.2012.2204190>.

MIRANDA ONOFRE, Fabiana Botelho de; ONOFRE, Alexandre Sherlley Casimiro; POMJANSKI, Natalia; BUCKSTEGGE, Birgit; GROTE, Hans Juergen; BOCKING, Alfred. 9p21 Deletion in the diagnosis of malignant mesothelioma in serous effusions additional to immunocytochemistry, DNA-ICM, and AgNOR analysis. **Cancer**, Wiley, v. 114, n. 3, p. 204–215, fev. 2008. DOI: 10.1002/cncr.23413. Disponível em: <https://doi.org/10.1002/cncr.23413>.

OKTAY, Ozan *et al.* **Attention U-Net: Learning Where to Look for the Pancreas**. [S.l.]: arXiv, 2018. DOI: 10.48550/ARXIV.1804.03999. Disponível em: <https://arxiv.org/abs/1804.03999>.

PAPANICOLAOU, G. N. A NEW PROCEDURE FOR STAINING VAGINAL SMEARS. **Science**, American Association for the Advancement of Science (AAAS), v. 95, n. 2469, p. 438–439, abr. 1942. DOI: 10.1126/science.95.2469.438. Disponível em: <https://doi.org/10.1126/science.95.2469.438>.

PASZKE, Adam *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: WALLACH, H.; LAROCHELLE, H.; BEYGELZIMER, A.; D'ALCHÉ-BUC, F.; FOX, E.; GARNETT, R. (Ed.). **Advances in Neural Information Processing Systems 32**. [S.l.]: Curran Associates, Inc., 2019. P. 8024–8035. Disponível em: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.

PEDRINI, Carolina Marian *et al.* Análise das regiões nucleolares com prata (AgNOR) em lesões cervicais. **USFC/Farmácia**, Universidade Federal de Santa Catarina - UFSC, 2020. Disponível em:

<https://repositorio.ufsc.br/handle/123456789/216317>.

POUYANFAR, Samira; SADIQ, Saad; YAN, Yilin; TIAN, Haiman; TAO, Yudong; REYES, Maria Presa; SHYU, Mei-Ling; CHEN, Shu-Ching; IYENGAR, S. S. A Survey on Deep Learning. **ACM Computing Surveys**, Association for Computing Machinery (ACM), v. 51, n. 5, p. 1–36, jan. 2019. DOI: 10.1145/3234150. Disponível em:

<https://doi.org/10.1145/3234150>.

PRANGEMEIER, Tim; REICH, Christoph; KOEPPL, Heinz. Attention-Based Transformers for Instance Segmentation of Cells in Microstructures. *In*: 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). [S.l.]: IEEE, dez. 2020. P. 700–707. DOI: 10.1109/bibm49941.2020.9313305. Disponível em:

<https://doi.org/10.1109/bibm49941.2020.9313305>.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation. *In*: LECTURE Notes in Computer Science. [S.l.]: Springer International Publishing, 2015. P. 234–241. DOI:

10.1007/978-3-319-24574-4_28. Disponível em:

https://doi.org/10.1007/978-3-319-24574-4_28.

RUSSAKOVSKY, Olga *et al.* ImageNet Large Scale Visual Recognition Challenge. **International Journal of Computer Vision (IJCV)**, v. 115, n. 3, p. 211–252, 2015. DOI: 10.1007/s11263-015-0816-y.

RUSSELL, Bryan C.; TORRALBA, Antonio; MURPHY, Kevin P.; FREEMAN, William T. LabelMe: A Database and Web-Based Tool for Image Annotation. **International Journal of Computer Vision**, Springer Science e Business Media LLC, v. 77, n. 1-3, p. 157–173, out. 2007. DOI: 10.1007/s11263-007-0090-8. Disponível em:

<https://doi.org/10.1007/s11263-007-0090-8>.

SANCER, Okan; ŞIRIN, Mümtaz Cem; ARIDOĞAN, Buket; ÇETIN, Emel Sesli; YIĞIT, Ahmet; TEPEBAŞI, Muhammet Yusuf; KOŞAR, Pınar Aslan. Comparison of DNA damage and proliferative capacities in smear samples of HPV positive and negative patients by micronucleus counting and AgNOR staining. **Biotechnic & Histochemistry**, Informa UK Limited, v. 96, n. 5, p. 354–358, ago. 2020. DOI:

10.1080/10520295.2020.1800091. Disponível em:
<https://doi.org/10.1080/10520295.2020.1800091>.

SANDFORT, Veit; YAN, Ke; PICKHARDT, Perry J.; SUMMERS, Ronald M. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. **Scientific Reports**, Springer Science e Business Media LLC, v. 9, n. 1, nov. 2019. DOI: 10.1038/s41598-019-52737-x. Disponível em: <https://doi.org/10.1038/s41598-019-52737-x>.

SHARMA, H. D.; MAHADESH, J.; MONALISA, W.; GOPINATHAN, P. A.; LAXMIDEVI, B. L.; SANJENBAM, N. Quantitative assessment of tumor-associated tissue eosinophilia and nuclear organizing region activity to validate the significance of the pattern of invasion in oral squamous cell carcinoma: A retrospective study. **Journal of Oral and Maxillofacial Pathology**, v. 25, n. 2, p. 258–265, 2021. DOI: 10.4103/0973-029X.325124. Disponível em:
<https://doi.org/10.4103/0973-029x.325124>.

SHARMA, Vipul; MIR, Roohie Naaz. A comprehensive and systematic look up into deep learning based object detection techniques: A review. **Computer Science Review**, Elsevier BV, v. 38, p. 100301, nov. 2020. DOI: 10.1016/j.cosrev.2020.100301. Disponível em:
<https://doi.org/10.1016/j.cosrev.2020.100301>.

SHORTEN, Connor; KHOSHGOFTAAR, Taghi M. A survey on Image Data Augmentation for Deep Learning. **Journal of Big Data**, Springer Science e Business Media LLC, v. 6, n. 1, jul. 2019. DOI: 10.1186/s40537-019-0197-0. Disponível em:
<https://doi.org/10.1186/s40537-019-0197-0>.

SMITH, Leslie N.; TOPIN, Nicholay. **Super-Convergence: Very Fast Training of Neural Networks Using Large Learning Rates**. [S.l.]: arXiv, 2017. DOI: 10.48550/ARXIV.1708.07120. Disponível em: <https://arxiv.org/abs/1708.07120>.

STELZLE, Dominik *et al.* Estimates of the global burden of cervical cancer associated with HIV. **The Lancet Global Health**, Elsevier BV, v. 9, n. 2, e161–e169, fev. 2021. DOI: 10.1016/s2214-109x(20)30459-9. Disponível em:
[https://doi.org/10.1016/s2214-109x\(20\)30459-9](https://doi.org/10.1016/s2214-109x(20)30459-9).

SUNG, Hyuna; FERLAY, Jacques; SIEGEL, Rebecca L.; LAVERSANNE, Mathieu; SOERJOMATARAM, Isabelle; JEMAL, Ahmedin; BRAY, Freddie. Global Cancer

Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. **CA: A Cancer Journal for Clinicians**, Wiley, v. 71, n. 3, p. 209–249, fev. 2021. DOI: 10.3322/caac.21660. Disponível em: <https://doi.org/10.3322/caac.21660>.

TAGHANAKI, Saeid Asgari; ABHISHEK, Kumar; COHEN, Joseph Paul; COHEN-ADAD, Julien; HAMARNEH, Ghassan. Deep semantic segmentation of natural and medical images: a review. **Artificial Intelligence Review**, Springer Science e Business Media LLC, v. 54, n. 1, p. 137–178, jun. 2020. DOI: 10.1007/s10462-020-09854-1. Disponível em: <https://doi.org/10.1007/s10462-020-09854-1>.

TONG, Kang; WU, Yiquan. Deep learning-based detection from the perspective of small or tiny objects: A survey. **Image and Vision Computing**, Elsevier BV, v. 123, p. 104471, jul. 2022. DOI: 10.1016/j.imavis.2022.104471. Disponível em: <https://doi.org/10.1016/j.imavis.2022.104471>.

TRERÈ, D. AgNOR staining and quantification. **Micron**, Elsevier BV, v. 31, n. 2, p. 127–131, abr. 2000. DOI: 10.1016/s0968-4328(99)00069-4. Disponível em: [https://doi.org/10.1016/s0968-4328\(99\)00069-4](https://doi.org/10.1016/s0968-4328(99)00069-4).

UMADI, Ashwin; NAGARAJAN, Kushal; VENKATESHA, Juthik Bangalore; GANESH, Aashutosh; GEORGE, Koshy. Automated Segmentation of Overlapping Cells in Cervical Cytology Images Using Deep Learning. *In*: 2020 IEEE 17th India Council International Conference (INDICON). [S.l.]: IEEE, dez. 2020. DOI: 10.1109/indicon49873.2020.9342328. Disponível em: <https://doi.org/10.1109/indicon49873.2020.9342328>.

VAN ROSSUM, Guido; DRAKE, Fred L. **Python 3 Reference Manual**. Scotts Valley, CA: CreateSpace, 2009. ISBN 1441412697.

VASWANI, Ashish; SHAZEER, Noam; PARMAR, Niki; USZKOREIT, Jakob; JONES, Llion; GOMEZ, Aidan N.; KAISER, Lukasz; POLOSUKHIN, Illia. **Attention Is All You Need**. [S.l.]: arXiv, 2017. DOI: 10.48550/ARXIV.1706.03762. Disponível em: <https://arxiv.org/abs/1706.03762>.

WAN, Tao; XU, Shusong; SANG, Chen; JIN, Yulan; QIN, Zengchang. Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. **Neurocomputing**, Elsevier BV, v. 365, p. 157–170, nov. 2019. DOI:

10.1016/j.neucom.2019.06.086. Disponível em:
<https://doi.org/10.1016/j.neucom.2019.06.086>.

WANG, Fei; JIANG, Mengqing; QIAN, Chen; YANG, Shuo; LI, Cheng;
ZHANG, Honggang; WANG, Xiaogang; TANG, Xiaoou. **Residual Attention Network for Image Classification**. [S.l.]: arXiv, 2017. DOI: 10.48550/ARXIV.1704.06904.
Disponível em: <https://arxiv.org/abs/1704.06904>.

WANG, Jingdong *et al.* Deep High-Resolution Representation Learning for Visual Recognition. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Institute of Electrical e Electronics Engineers (IEEE), v. 43, n. 10, p. 3349–3364, out. 2021. DOI: 10.1109/tpami.2020.2983686. Disponível em:
<https://doi.org/10.1109/tpami.2020.2983686>.

WHO. **Human papillomavirus (HPV) and cervical cancer**. [S.l.: s.n.], nov. 2020.
Disponível em: [https://www.who.int/news-room/fact-sheets/detail/human-papillomavirus-\(hpv\)-and-cervical-cancer](https://www.who.int/news-room/fact-sheets/detail/human-papillomavirus-(hpv)-and-cervical-cancer). Acesso em: 20 jan. 2021.

WOLF, Thomas *et al.* Transformers: State-of-the-Art Natural Language Processing. *In: PROCEEDINGS of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Online: Association for Computational Linguistics, out. 2020. P. 38–45. DOI: 10.18653/v1/2020.emnlp-demos.6. Disponível em: <https://aclanthology.org/2020.emnlp-demos.6>.

XIE, Enze; WANG, Wenhai; YU, Zhiding; ANANDKUMAR, Anima; ALVAREZ, Jose M.; LUO, Ping. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. *In: RANZATO, M.; BEYGELZIMER, A.; DAUPHIN, Y.; LIANG, P.S.; VAUGHAN, J. Wortman (Ed.). Advances in Neural Information Processing Systems*. [S.l.]: Curran Associates, Inc., 2021. P. 12077–12090. Disponível em: <https://proceedings.neurips.cc/paper/2021/file/64f1f27bf1b4ec22924fd0acb550c235-Paper.pdf>.

ZHANG, Ling; SONKA, Milan; LU, Le; SUMMERS, Ronald M.; YAO, Jianhua. Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei. *In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. [S.l.]: IEEE, abr. 2017. P. 406–409. DOI: 10.1109/isbi.2017.7950548. Disponível em:
<https://doi.org/10.1109/isbi.2017.7950548>.

ZHANG, Zhilu; SABUNCU, Mert R. Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels. **CoRR**, abs/1805.07836, 2018. arXiv: 1805.07836. Disponível em: <http://arxiv.org/abs/1805.07836>.

APÊNDICE A – ARTIGOS PUBLICADOS

A.1 ARTIGOS COMO AUTOR PRINCIPAL

A maioria dos resultados apresentados neste trabalho foram publicados em trabalhos publicados como artigos de periódicos ou conferencias. Assim como uma prévia da RSL publicada como relatório técnico.

A.1.1 Relatório técnico — INCoD - UFSC

- **Título:** Systematic Literature Review of Computer Vision-Aided Cytology - A Review of Classic Computer Vision and Deep Learning-Based Approaches published between January/2016 - March/2020
- **Autores:** AMORIM, João Gustavo Atkinson; CERENTINI, Allan; MACARINI, Luiz Antonio Buschetto; MATIAS, André Victória; WANGENHEIM, Aldo von.
- **Data:** maio de 2020.
- **DOI:** <http://rgdoi.net/10.13140/RG.2.2.13304.67840>

A.1.2 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS 2020)

- **Título:** A Novel Approach on Segmentation of AgNOR-Stained Cytology Images Using Deep Learning
- **Autores:** AMORIM, João Gustavo Atkinson; MACARINI, Luiz Antonio Buschetto; MATIAS, André Victória; CERENTINI, Allan; ONOFRE, Fabiana Botelho De Miranda; ONOFRE, Alexandre Sherlley Casimiro; WANGENHEIM, Aldo von.
- **Data:** julho de 2020.
- **DOI:** <https://doi.org/10.1109/CBMS49503.2020.00110>

A.1.3 XVII Congresso Brasileiro de Informática em Saúde (CBIS 2020)

- **Título:** Detecção de núcleos em imagens citológicas de AgNOR utilizando Aprendizado Profundo (publicado como suplemento da revista *Journal of Health Informatics*)
- **Autores:** AMORIM, João Gustavo Atkinson; MACARINI, Luiz Antonio Buschetto; MATIAS, André Victória; CERENTINI, Allan; ONOFRE, Fabiana Botelho De Miranda; ONOFRE, Alexandre Sherlley Casimiro; WANGENHEIM, Aldo von.
- **Data:** dezembro de 2020.

- **Url:** <http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/view/811>

A.1.4 Computerized Medical Imaging and Graphics

- **Título:** What is the state of the art of computer vision-assisted cytology? A Systematic Literature Review
- **Autores:** MATIAS, André Victória; AMORIM, João Gustavo Atkinson; MACARINI, Luiz Antonio Buschetto; CERENTINI, Allan; ONOFRE, Alexandre Sherlley Casimiro; ONOFRE, Fabiana Botelho De Miranda; DALTOÉ, Felipe Perozzo; STEMMER, Marcelo Ricardo; WANGENHEIM, Aldo von.
- **Data:** junho de 2021.
- **DOI:** <https://doi.org/10.1016/j.compmedimag.2021.101934>

A.1.5 XIII Computer on the Beach (COTB 2022)

- **Título:** Nucleus detection in cervical samples stained with AgNOR
- **Autores:** AMORIM, João Gustavo Atkinson; SANCHES, Viníus Moreno; BOTAMEDI, Tainee; MATIAS, André Victória; CAVACO, Marco; ONOFRE, Alexandre Sherlley Casimiro; ONOFRE, Fabiana Botelho De Miranda; WANGENHEIM, Aldo von.
- **Data:** maio de 2022.
- **DOI:** <http://dx.doi.org/10.14210/cotb.v13.p045-050>

A.1.6 Social Science Research Network - SSRN

- **Título:** Semantic Segmentation for the Detection of Very Small Objects on Cervical Cell Samples Stained with the AgNOR Technique
- **Autores:** AMORIM, João Gustavo Atkinson; MATIAS, André Victória; CERENTINI, Allan; MIRANDA ONOFRE, Fabiana Botelho de; ONOFRE, Alexandre Sherlley Casimiro; WANGENHEIM, Aldo von.
- **Data:** junho de 2022.
- **DOI:** <https://dx.doi.org/10.2139/ssrn.4126881>

A.2 CONTRIBUIÇÃO EM OUTRAS PUBLICAÇÕES

Durante o mestrado, eu também colaborei em outras publicações, que não necessariamente estão diretamente relacionados a este trabalho. O grupo e projeto de pesquisa apesar de estar focado em trabalhar com diferentes amostras citológicas com visão computacional, também atuou em algumas outras frentes. Apesar disto, estes trabalhos forneceram ferramentas uteis para o desenvolvimento deste trabalho.

A.2.1 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS 2020)

- **Título:** Segmentation, Detection and Classification of Cell Nuclei on Oral Cytology Samples Stained with Papanicolaou
- **Autores:** MATIAS, André Victória; CERENTINI, Allan; MACARINI, Luiz Antonio Buschetto; AMORIM, João Gustavo Atkinson; DALTOÉ, Felipe Perozzo; WANGENHEIM, Aldo von.
- **Data:** julho de 2020.
- **DOI:** <https://doi.org/10.1109/cbms49503.2020.00018>

A.2.2 XVII Congresso Brasileiro de Informática em Saúde (CBIS 2020)

- **Título:** Localização de Núcleos Celulares em Citologia Oral Usando Métodos de Deep Learning (publicado como suplemento da revista *Journal of Health Informatics*)
- **Autores:** MATIAS, André Victória; CERENTINI, Allan; MACARINI, Luiz Antonio Buschetto; AMORIM, João Gustavo Atkinson; DALTOÉ, Felipe Perozzo; WANGENHEIM, Aldo von.
- **Data:** dezembro de 2020.
- **URL:** <http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/view/810>

A.2.3 SN Computer Science

- **Título:** Segmentation, Detection, and Classification of Cell Nuclei on Oral Cytology Samples Stained with Papanicolaou
- **Autores:** MATIAS, André Victória; CERENTINI, Allan; MACARINI, Luiz Antonio Buschetto; AMORIM, João Gustavo Atkinson; DALTOÉ, Felipe Perozzo; WANGENHEIM, Aldo von.

- **Data:** maio de 2021.
- **DOI:** <https://doi.org/10.1007/s42979-021-00676-8>

A.2.4 medRxiv - Cold Spring Harbor Laboratory

- **Título:** Comparison of Object Detection Approaches Applied to Field Images of Papanicolaou Stained Cytology Slides
- **Autores:** MATIAS, André Victória; CERENTINI, Allan; MACARINI, Luiz Antonio Buschetto; AMORIM, João Gustavo Atkinson; DALTOÉ, Felipe Perozzo; WANGENHEIM, Aldo von.
- **Data:** Agosto de 2021.
- **DOI:** <https://doi.org/10.1101/2021.08.25.21262605>

A.2.5 Social Science Research Network - SSRN

- **Título:** Computer-Aided Analysis of Oral Conventional Papanicolaou Cytology Samples
- **Autores:** MATIAS, André Victória; AMORIM, João Gustavo Atkinson; MACARINI, Luiz Antonio Buschetto; CERENTINI, Allan; DALTOÉ, Felipe Perozzo; WANGENHEIM, Aldo von.
- **Data:** junho de 2022.
- **DOI:** <https://dx.doi.org/10.2139/ssrn.4119212>

A.3 PREMIAÇÕES

A.3.1 Terceiro colocado no *Outdoor Semantic Segmentation Challenge da 43rd DAGM German Conference on Pattern Recognition*

- **Título:** Learning to Identify the Navigable Path in Unstructured and Off-Road Scenes from a Single Dataset
- **Autores:** MATIAS, André Victória; MARIAN, Juliana; CERENTINI, Allan; MARTINS, Juncklaus; AMORIM, João Gustavo Atkinson; WANGENHEIM, Aldo von.
- **Data:** dezembro de 2020.
- **URL:** https://unstructured-scene-understanding.com/assets/pdf/DeepUFSC_DAGM_Paper.pdf