UNIVERSIDADE FEDERAL DE OURO PRETO INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS DEPARTAMENTO DE COMPUTAÇÃO

ANA CRISTINA LOPES DE OLIVEIRA Orientadora: Prof<sup>a</sup>. Dra. Andrea Gomes Campos Bianchi Coorientador: Prof<sup>a</sup>. Msc. Daniela Costa Terra

## SEGMENTAÇÃO DE NÚCLEO E CITOPLASMA EM IMAGENS DE CÉLULAS CERVICAIS COM APRENDIZADO PROFUNDO USANDO UMA REDE NEURAL U-NET

Ouro Preto, MG 2023

### UNIVERSIDADE FEDERAL DE OURO PRETO INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS DEPARTAMENTO DE COMPUTAÇÃO

ANA CRISTINA LOPES DE OLIVEIRA

## SEGMENTAÇÃO DE NÚCLEO E CITOPLASMA EM IMAGENS DE CÉLULAS CERVICAIS COM APRENDIZADO PROFUNDO USANDO UMA REDE NEURAL U-NET

Monografia apresentada ao Curso de Ciência da Computação da Universidade Federal de Ouro Preto como parte dos requisitos necessários para a obtenção do grau de Bacharel em Ciência da Computação.

**Orientadora:** Prof<sup>a</sup>. Dra. Andrea Gomes Campos Bianchi **Coorientador:** Prof<sup>a</sup>. Msc. Daniela Costa Terra

> Ouro Preto, MG 2023



MINISTÉRIO DA EDUCAÇÃO UNIVERSIDADE FEDERAL DE OURO PRETO REITORIA INSTITUTO DE CIENCIAS EXATAS E BIOLOGICAS DEPARTAMENTO DE COMPUTACAO



### FOLHA DE APROVAÇÃO

Ana Cristina Lopes de Oliveira

### Segmentação de Núcleo e Citoplasma em Imagens de Células Cervicais com Aprendizado Profundo usando uma Rede Neural U-Net

Monografia apresentada ao Curso de Ciência da Computação da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Ciência da Computação

Aprovada em 30 de Agosto de 2023.

Membros da banca

Andrea Gomes Campos Bianchi (Orientadora) - Doutora - Universidade Federal de Ouro Preto Daniela Terra Costa (Coorientadora) - Mestre - Instituto Federal de Minas Gerais - Campus Ouro Branco Guillermo Cámara Chávez (Examinador) - Doutor - Universidade Federal de Ouro Preto Pedro Henrique Lopes Silva (Examinador) - Doutor - Universidade Federal de Ouro Preto

Andrea Gomes Campos Bianchi, Orientadora do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 30/08/2023.



Documento assinado eletronicamente por Andrea Gomes Campos Bianchi, PROFESSOR DE MAGISTERIO SUPERIOR, em 05/09/2023, às 17:01, conforme horário oficial de Brasília, com fundamento no art. 6°, § 1°, do Decreto n° 8.539, de 8 de outubro de 2015.



A autenticidade deste documento pode ser conferida no site <u>http://sei.ufop.br</u> /sei/controlador\_externo.php?acao=documento\_conferir&id\_orgao\_acesso\_externo=0, informando o código verificador **0586115** e o código CRC **4B68F766**.

Referência: Caso responda este documento, indicar expressamente o Processo nº 23109.010000/2023-18

## Agradecimentos

Este trabalho é resultado da colaboração e esforços de muitos que fizeram parte da minha jornada durante a graduação. Por isso, gostaria aqui de oferecer meus agradecimentos, primeiramente à Universidade Federal de Ouro Preto, por me proporcionar as ferramentas e oportunidades necessárias para a minha profissionalização.

Agradeço especialmente ao apoio da Fapemig, CNPq e Capes que possibilitou o desenvolvimento desse trabalho.

Agradeço também aos professores do DECOM que tive a oportunidade de conhecer. Em especial, Andrea Bianchi, Daniela Terra, Dayanne Coelho, Guilherme de Assis e Rodrigo Ribeiro, que desempenharam um exímio trabalho em me guiar, ensinar e mostrar os caminhos que eu poderia percorrer em minha carreira. Obrigada por todas as oportunidades.

À Voluta, obrigada pela oportunidade, por todos os ensinamentos, experiências vividas e projetos desenvolvidos.

Aos amigos e colegas que conheci nessa graduação, obrigada por todo suporte e carinho oferecidos durante esses anos. Em especial, agradeço à Fernanda, Maycon, Valmir e Júlia. Obrigada por terem tornado os momentos mais leves e por estarem ao meu lado.

Por fim, agradeço à minha família por sempre ter acreditado em mim e fornecido tudo ao seu alcance para que eu fosse bem-sucedida nessa trajetória. Obrigada, mãe, avó, José, e aos meus irmãos, por todo apoio e carinho.

A todos que não foram citados aqui, mas que fizeram parte dessa jornada também, meus sinceros agradecimentos por terem contribuído com essa trajetória.

## Resumo

O câncer de colo de útero é um problema grave de saúde pública, afetando milhares de mulheres anualmente e sendo a quarta causa de morte por câncer mais frequente no sexo feminino, segundo o Instituto Nacional do Câncer. Essa doença se caracteriza por um desenvolvimento lento e é tratável quando diagnosticada em seus estágios iniciais. Dentre as principais formas de prevenção e detecção, destaca-se o exame de Papanicolau, que reduziu significativamente a incidência e mortalidade da doença desde sua criação. Contudo, um dos principais desafios do exame reside na alta proporção de falsos positivos e diagnósticos imprecisos, problema muitas vezes causado pela qualidade no processo de coleta e análise do material. Na etapa de análise, os profissionais em citopatologia precisam buscar manualmente por anomalias nas imagens microscópicas das células, um procedimento sujeito à exaustão e falhas de percepção humana. Como alternativa, diversos estudos foram desenvolvidos com o objetivo de proporcionar sistemas que auxiliam no diagnóstico, ao automatizar a etapa de segmentação e posterior classificação de células cervicais quanto à sua malignidade. Dentre esses estudos, as técnicas de aprendizagem profunda têm ganhado ênfase por serem capazes de segmentar regiões celulares de forma precisa e eficiente. Diante disso, este trabalho propõe um método baseado em aprendizagem profunda para segmentar as regiões do núcleo e citoplasma de células cervicais, utilizando a rede neural U-Net e a base de dados ISBI 2014. Com essa metodologia, foi possível alcançar resultados de grande relevância ao serem contrastados com o estado da arte da literatura existente, obtendo a pontuação de 0,9824 para o coeficiente de Dice na segmentação do citoplasma.

**Palavras-chave**: 1. Segmentação de células. 2. Redes Neurais Convolucionais. 3. Células Cervicais.

## Abstract

Cervical cancer is a serious public health problem, affecting thousands of women annually and being the fourth most frequent cause of cancer death in females, according to the National Cancer Institute. This disease is characterized by slow development and is treatable when diagnosed in its early stages. Among the main forms of prevention and detection, the Papanicolaou test stands out, which has significantly reduced the incidence and mortality of the disease since its creation. However, one of the main challenges of the test lies in the high proportion of false positives and inaccurate diagnoses, a problem often caused by the quality of the collected material and analysis process. During the inspection, cytology professionals must manually search for anomalies in microscopic images of cells, a procedure subject to exhaustion and human perception interpretation. As an alternative, several studies have been developed to provide systems that assist the professional in diagnosis by automating the segmentation and subsequent classification of cervical cells regarding their malignancy. Among these studies, deep learning techniques have gained emphasis for being able to accurately and efficiently segment cellular regions. This work proposes a deep learning-based method for segmenting the nucleus and cytoplasm regions of cervical cells, using the U-Net neural network and the ISBI 2014 database. With this methodology, it was possible to achieve highly relevant results compared to the state of the art in the existing literature, obtaining a Dice coefficient score of 0.9824 for cytoplasm segmentation.

Keywords: Citology Segmentation, Convolutional Neural Networks, Cervical Cells.

# Lista de Ilustrações

Figura 1.1 –	Incidência de novos casos de câncer no mundo em 2020	1
Figura 1.2 –	- Imagem microscópica extraída da base de dados CRIC Cervix, obtida a partir	
	de uma lâmina gerada pela amostra coletada no exame papanicolau	2
Figura 1.3 –	- Aplicação da técnica de segmentação, onde cada cor é representada por uma	
	classe de objeto.	3
Figura 2.1 –	Exemplo de segmentação semântica com duas classes cat e background (à	
	esquerda), e segmentação por instância (à direita) onde cada gato é identificado	
	como um indivíduo pertencente à classe <i>cat</i>	6
Figura 2.2 –	Arquitetura comum entre redes neurais convolucionais.	7
Figura 2.3 –	Exemplo da aplicação de um Smoothing Kernel e um Sharpening Kernel em	
	uma imagem de citologia digital.	8
Figura 2.4 –	Filtros gerados nas primeiras camadas da AlexNet ao ser treinada com a base	
	de dados ImageNet.	9
Figura 2.5 –	- Aplicação da função Max Pool com um deslocamento de 2 pixels (stride) em	
	um mapa de características 4*4	9
Figura 2.6 –	Exemplo de aplicação da técnica de <i>Dropout</i> em uma RNA	11
Figura 2.7 –	- Aplicação de aumento de dados na coloração de uma imagem utilizada para	
	classificação de melanomas.	12
Figura 2.8 –	Arquitetura da rede U-Net. Cada retângulo em azul representa um mapa de	
	características multi-canal.	12
Figura 2.9 –	Levantamento da frequência das principais métricas usadas em trabalhos de	
	citologia assistida por visão computacional.	14
Figura 2.10-	-Representação visual do coeficiente de Dice aplicado á segmentação de imagens.	14
Figura 2.11-	-Aglomerado de células sobrepostas extraído da base de dados CRIC.	17
Figura 3.1 –	Imagem de profundidade estendida (EDF) coletada da base ISBI 2014	21
Figura 3.2 –	Exemplo de imagem rotacionada que gerou bordas preenchidas automatica-	
	mente utilizando o algoritmo nearest neighbor	21
Figura 3.3 –	Exemplo de transformações aplicadas á uma mesma imagem (a) encontrada	
	na base ISBI 2014	22
Figura 3.4 –	Fluxograma proposto para o desenvolvimento.	23
Figura 3.5 –	Fluxograma da divisão da base de dados ISBI 2014 em <i>k-folds</i>	24
Figura 3.6 –	Procedimento de remoção da área do núcleo na região citoplasmática. A	
	imagem (c) corresponde á máscara do núcleo que foi combinada á imagem	
	(b) para resultar a imagem (d) que contém apenas a região do citoplasma (em	
	branco).	24

Figura 3.7 - Fluxograma de treinamento da rede U-Net utilizado por Teixeira (2021)	e
aplicado também neste trabalho.	25
Figura 4.1 – Exemplos de predição da U-Net para segmentação do núcleo	28
Figura 4.2 – Exemplos de predição da U-Net para segmentação do citoplasma	28
Figura 4.3 – Evolução do treinamento da U-Net para segmentação do núcleo	29
Figura 4.4 – Evolução do treinamento da U-Net para segmentação do citoplasma	29
Figura 4.5 – Imagens originais e as suas respectivas predições com as células individua	ıl-
mente delimitadas.	30

## Lista de Tabelas

Tabela 2.1 – Trabalhos relacionados que serão usados para fins de comparação.	19
Tabela 3.1 – Transformações e suas variações de valores.	22
Tabela 3.2 – Valores atribuídos os parâmetros de treinamento do modelo.       .       .	26
Tabela 3.3 – Configuração de máquina disponibilizada pela plataforma Google Colabora-	
tory para a execução dos experimentos.	26
Tabela 4.1 – Resultado final para segmentação do núcleo	27
Tabela 4.2 – Resultado final para segmentação do citoplasma	27
Tabela 4.3 – Comparação de resultados da nossa proposta com outros trabalhos voltados	
que realizam a segmentação do citoplasma	30

## Lista de Abreviaturas e Siglas

- ABNT Associação Brasileira de Normas Técnicas
- CAD Computer-aided Diagnosis
- CNN Convolutional Neural Network
- CRIC Center for Recognition and Inspection of Cells
- DECOM Departamento de Computação
- DPA Digital Pathology Association
- DSC Dice Similarity Coefficient
- EDF Extended Depth of Field
- FCM Fuzzy C-Means
- FCN Fully Convolutional Network
- FN False Negative
- FP False Positive
- GT Ground Truth
- MPFW Multi-Pass Fast Watershed
- OMS Organização Mundial da Saúde
- P Precisão
- R Revocação
- ReLU Rectified linear unit
- RGB Red Green and Blue
- RNA Redes Neurais Artificiais
- SZU Shenzhen University
- TN True Negative
- TP True Positive
- UFOP Universidade Federal de Ouro Preto

# Lista de Símbolos

- \* Convolução
- ∫ Integral
- au Letra grega tau

## Sumário

1	Introdução				
	1.1	Justific	cativa	3	
	1.2	Objeti	VOS	5	
	1.3	Organi	ização do Trabalho	5	
2	Rev	isão Bib	oliográfica	6	
	2.1	.1 Fundamentação Teórica			
		2.1.1	A segmentação de imagens de células cervicais	6	
		2.1.2	Redes neurais convolucionais para a segmentação de imagens	7	
		2.1.3	U-Net	12	
		2.1.4	Métricas para a análise de resultados da segmentação	13	
	2.2	Trabal	hos Relacionados	15	
		2.2.1	Metodologias baseadas em <i>deep learning</i>	15	
		2.2.2	Outras metodologias	18	
		2.2.3	Discussão	18	
3	Des	envolvir	nento	20	
	3.1	Base d	le dados	20	
		3.1.1	ISBI 2014	20	
		3.1.2	Aumento de dados	20	
	3.2	Metod	ologia	22	
		3.2.1	Preparação dos dados	22	
		3.2.2	Treinamento	24	
4	Res	ultados		27	
	4.1	Discus	ssões	28	
5	Con	siderac	ões finais	31	
	5.1	Trabal	hos futuros	31	
Re	eferên	icias .		33	

## 1 Introdução

Em 2020, o câncer do colo do útero, ou câncer cervical, apresentou mais de 600.000 casos no mundo todo e resultou em uma taxa de mortalidade de 4,60 óbitos/100 mil mulheres no Brasil, sendo o terceiro tipo de câncer mais comum entre o sexo feminino sem considerar aqueles advindos de tumores de pele não melanoma (Instituto Nacional do Câncer, 2022). Na FIG. 1.1 é possível observar o número estimado de novos casos de câncer em 2020, com destaque para o câncer cervical que compôs cerca de 6,5% do total de casos em mulheres dentre todas as idades.

Figura 1.1 – Incidência de novos casos de câncer no mundo em 2020.



Estimated number of new cases in 2020, World, females, all ages

Data source: Globocan 2020 Graph production: Global Cancer Observatory (http://gco.iarc.fr)

Fonte: International Agency for Research on Cancer (2023).

O exame citopatológico Papanicolau, criado na década de 1950, tornou-se um dos métodos mais relevantes para o rastreamento do câncer cervical, sendo um dos principais responsáveis pela redução em mais de 70% da taxa de incidência e mortalidade nos Estados Unidos (SAFAEIAN; SOLOMON; CASTLE, 2007). O procedimento envolve a coleta de amostras de esfregaços no colo do útero e uma posterior análise manual feita por um profissional de citopatologia, que procura por anomalias celulares nas imagens microscópicas geradas pelas lâminas obtidas. Nessa segunda fase, o procedimento se torna exaustivo e suscetível á falhas, visto que são geradas milhares de imagens contendo diversos grupos celulares, acompanhados de interferências que prejudicam a visualização, como sangue, inflamações e detritos. Um exemplo de uma imagem obtida por esse exame pode ser observado na FIG. 1.2.

Com o objetivo de melhorar a performance desse procedimento, trazendo uma maior precisão no diagnóstico, diversos profissionais, como Tareef et al. (2018), Phoulady et al. (2017) e

Figura 1.2 – Imagem microscópica extraída da base de dados CRIC Cervix, obtida a partir de uma lâmina gerada pela amostra coletada no exame papanicolau.



Fonte: Rezende, Bianchi e Carneiro (2021).

Song et al. (2015), têm direcionado seus estudos á automatização desse processo através do uso de algoritmos e técnicas de aprendizagem de máquina aplicados no campo de visão computacional. Esses trabalhos se propõem a extrair as regiões do núcleo e citoplasma das imagens para em um segundo passo conceder dados dessas regiões para algoritmos de classificação, que por sua vez entregarão predições quanto à presença de neoplasias malignas. O resultado desse processo fornece ao profissional da saúde uma estimativa da presença do câncer, podendo auxiliar significantemente o diagnóstico.

Para extrair as regiões celulares, é possível usar a técnica de segmentação, que sendo amplamente estudada no campo de visão computacional, têm por objetivo final separar regiões de uma imagem cujos pixels pertencentes àquela região possuem características em comum e formam um objeto, ou instância. O resultado desse processo possibilita separar os objetos de estudo de outras regiões da imagem, sendo então possível extrair os dados necessários das regiões de interesse para a solução do problema. Um exemplo dessa técnica é demonstrado na FIG. 1.3

No contexto de citologia cervical, existem duas regiões principais que são objetos de estudo na classificação da célula quanto à sua natureza: o citoplasma e seu núcleo correspondente. Para que seja possível extrair os dados dessas regiões, é necessário, especialmente, segmentar a região do núcleo, pois anormalidades em seu formato, cor e outras características entregam indícios de uma célula potencialmente cancerígena (PANGARKAR, 2022). Em conjunto, as informações contidas no citoplasma também desempenham um papel fundamental para o estudo de anomalias da célula, e obter uma segmentação precisa dessa região torna-se um fator crucial para uma análise completa. Contudo, muitos trabalhos já realizados nesse campo se concentraram exclusivamente na segmentação da região do núcleo (DINIZ et al., 2021; ZOU et al., 2020), ou então não segmentam o citoplasma de cada célula de forma individual. Dado essa problemática,



Figura 1.3 – Aplicação da técnica de segmentação, onde cada cor é representada por uma classe de objeto.

Fonte: Jeong, Yoon e Park (2018).

nesse trabalho é proposto o treinamento da rede neural U-Net (RONNEBERGER; FISCHER; BROX, 2015) para segmentar ambas as regiões do núcleo e citoplasma de células cervicais, utilizando imagens da bases de dados ISBI 2014 (LU; CARNEIRO; BRADLEY, 2015).

### 1.1 Justificativa

O câncer do colo de útero é uma doença que se caracteriza pela replicação desordenada das células que revestem o órgão (INCA, 2022). Esse processo é lento e, em sua fase inicial, geralmente é assintomático. A evolução da doença é dividida em quatro estágios, sendo que, no estágio IV, o tumor pode se infiltrar em estruturas extra uterinas, comprometendo órgãos como a bexiga e o reto (INCA, 2001). Por essa razão, a detecção precoce é fundamental para o sucesso do tratamento, que pode eliminar efetivamente todas as células cancerígenas, resultando em altas chances de cura.

O exame citopatológico Papanicolau é amplamente utilizado para diagnosticar o câncer do colo do útero em duas etapas. Primeiramente, durante o processo de coleta, é introduzido um espéculo na vagina, causando uma leve descamação na superfície externa e interna do colo do útero. Em seguida, o material coletado é depositado em lâminas para posterior análise microscópica por um profissional especializado em citopatologia (BVS, 2022). No processo de análise do material coletado, é essencial que o profissional responsável seja capaz de examinar minuciosamente as anomalias celulares frequentemente encontradas no câncer do colo do útero. Essas anomalias podem ser identificadas por meio de alterações morfológicas, como um núcleo significativamente maior em relação ao citoplasma (MOSHAVEGH et al., 2012) ou mudanças na forma, cor e outros aspectos do núcleo e citoplasma celulares (PANGARKAR, 2022). No entanto, esse procedimento se torna altamente escalável e suscetível a falhas humanas, uma vez que uma única lâmina de coleta pode gerar até 15.000 imagens (DINIZ, 2019) contendo vários agrupamentos celulares.

Embora a rapidez no diagnóstico não seja considerada crítica para o sucesso no tratamento do câncer do colo do útero, a precisão das informações obtidas por meio do exame Papanicolau é crucial para garantir um diagnóstico correto. A falha em detectar anomalias celulares pode resultar em consequências graves para o paciente (CHAPMAN; OTIS, 2011), sendo assim, a rastreabilidade de neoplasias malignas em imagens coletadas pelo exame é um desafio importante em todo o processo, como evidenciado pela alta taxa de falsos negativos relatada na literatura, que pode variar de 15% a 65% (CASTILLO et al., 2016). Esses falsos negativos podem ser causados por erros nas fases pré-analítica e analítica do exame, onde na fase analítica, a imprecisão nas análises manuais das lâminas pode resultar em equívocos no diagnóstico final, como a classificação incorreta de uma célula maligna como benigna (COSTA CARLA MAYLA SILVA DE MELO, 2021). Os fatores que contribuem para essa falha humana na análise estão ligados à fadiga e à repetição de uma atividade que requer uma observação crítica e detalhada constante de um grande conjunto de informações.

Com o objetivo de tornar essa fase analítica do exame mais eficiente e menos suscetível à falhas humanas, um sistema *Computer-aided Diagnosis* (CAD) torna-se uma alternativa. De acordo com a Associação de Patologia Digital (DPA) (LAAK, 2017), um sistema CAD pode aumentar a eficiência em trabalhos rotineiros, assumindo tarefas exaustivas e que dependem de alta acurácia. Além disso, podem fornecer informações que normalmente passam despercebidas pelo olho humano e seriam decisivas em um diagnóstico. Contudo, para que isso ocorra no contexto de detecção de células cervicais cancerígenas, é necessário que o software seja capaz de identificar e segmentar as regiões do núcleo e citoplasma da célula, com o intuito de extrair informações importantes a respeito da morfologia celular dessas regiões, tornando assim possível utilizá-las para classificar essas células quanto a sua naturalidade, a fim de obter um diagnóstico mais preciso. Esse tipo de sistema pode auxiliar o profissional de citopatologia no diagnóstico de forma parcial ou integral, fornecendo essas informações que podem também vir acompanhadas de uma estimativa prévia da classificação da amostra.

Diversos algoritmos têm sido utilizados ao longo dos anos no campo da visão computacional para segmentar imagens, tais como o algoritmo de *k-means clustering* (LI et al., 2012), o método *region growing* (LIN; CHEN, 2010), os métodos *watershed* (TAREEF et al., 2018) e cortes em grafos (ZHANG et al., 2014). No entanto, na última década, as técnicas de aprendizagem profunda têm se destacado significativamente, apresentando resultados com uma acurácia e performance consideravelmente superiores. Conforme evidenciado em Minaee et al. (2021), após a comparação de mais de cem métodos de aprendizagem profunda para segmentação de imagens, pode-se concluir que essa abordagem tem obtido melhorias significativas de desempenho nos últimos 5-6 anos, atraindo a atenção de pesquisadores.

Em 2021, resultados relevantes foram obtidos por Teixeira (2021) na segmentação do núcleo através do treinamento obtido com a Rede Neural Convolucional U-Net (RONNEBERGER;

FISCHER; BROX, 2015) combinada à base de dados CRIC Cervix (REZENDE; BIANCHI; CARNEIRO, 2021), que contém imagens reais de células cervicais. A partir dessa abordagem, este trabalho propõe avançar na pesquisa e realizar a segmentação das regiões do núcleo e citoplasma de células cervicais, treinando a arquitetura de Rede Neural Convolucional U-Net com imagens contidas na base de dados ISBI 2014 (LU; CARNEIRO; BRADLEY, 2015).

### 1.2 Objetivos

O trabalho aqui proposto tem por objetivo principal obter a segmentação das regiões do núcleo e do citoplasma de células cervicais em imagens citológicas microscópicas sintéticas, bem como posteriormente comparar a metodologia proposta com outras abordagens encontradas na literatura. Os objetivos específicos são:

- Treinar a arquitetura U-Net desenvolvida por Ronneberger, Fischer e Brox (2015) para segmentação do núcleo, como proposto por Teixeira (2021), para a base de dados ISBI 2015 (LU; CARNEIRO; BRADLEY, 2015).
- Treinar a arquitetura U-Net para segmentação do citoplasma, utilizando a mesma base de dados.

### 1.3 Organização do Trabalho

No Capítulo 2 deste documento é possível encontrar a base teórica fundamental para o entendimento da temática juntamente à menção dos trabalhos relacionados. Posteriormente, encontram-se nos Capítulos 3 e 4, a metodologia usada e os resultados obtidos, respectivamente. Por fim, o Capítulo 5 dispõe das considerações finais sobre esse trabalho.

## 2 Revisão Bibliográfica

### 2.1 Fundamentação Teórica

Para uma melhor contextualização acerca do desenvolvimento do trabalho, essa Seção descreve o embasamento teórico a respeito dos conceitos e métodos utilizados, sendo eles organizados nas seguintes Subseções: 2.1.1 A segmentação de imagens de células cervicais 2.1.2, Redes neurais convolucionais para a segmentação de imagens, e **??** Métricas para a análise de resultados da segmentação.

#### 2.1.1 A segmentação de imagens de células cervicais

A segmentação de imagens é um problema altamente relevante para a área de visão computacional. O processo consiste na separação de pixels que apresentam características similares em relação aos aspectos de uma imagem, como intensidade, cor e textura. Ao agrupar esses pixels, é possível formar regiões distintas dentro da imagem e assim extrair as informações relevantes para o contexto do problema.

Existem dois tipos principais de segmentação: a segmentação semântica e a segmentação de instâncias. No primeiro caso, as regiões delimitadas correspondem aos pixels pertencentes as regiões de interesse para cada classe definida (HARALICK; SHAPIRO, 1985). Já a segmentação de instâncias vai além ao identificar os pixels para cada objeto de uma classe alvo, individualmente. Na FIG. 2.1 é possível diferenciar ambos os tipos de segmentação aplicados em uma mesma imagem.

Ja a segmentacao de instancias

Figura 2.1 – Exemplo de segmentação semântica com duas classes *cat* e *background* (à esquerda), e segmentação por instância (à direita) onde cada gato é identificado como um indivíduo pertencente à classe *cat*.



Fonte: Chollet (2021).

Com o intuito de garantir a precisão da segmentação das células cervicais e, consequentemente, contribuir para o processo de classificação, a técnica de segmentação semântica se destaca por obter regiões bem definidas sem a necessidade de separar individualmente cada agrupamento de células em instâncias.

#### 2.1.2 Redes neurais convolucionais para a segmentação de imagens

As redes neurais artificias (RNA) são algoritmos inspirados no sistema nervoso humano e têm por finalidade o aprendizado de máquina para a solução de problemas. Elas foram primeiramente apresentadas por Warren McCullough e Walter Pitts em 1944 (HARDESTY, 2017) e desde então, diversos modelos e novas arquiteturas de redes surgiram; em destaque as redes neurais convolucionais (do inglês, *Convolutional Neural Networks* - CNN), que já demonstraram importantes resultados quanto a sua performance, principalmente no campo de visão computacional (ALBAWI; MOHAMMED; ALZAWI, 2017). Além das camadas de entrada e saída, uma CNN comumente é composta por três tipos diferentes de camadas como demonstra a FIG. 2.2 sendo elas: camada de convolução, camada de *pooling* e um conjunto de camadas completamente conectadas.



Figura 2.2 – Arquitetura comum entre redes neurais convolucionais.

#### Camada de convolução

O termo que dá origem ao nome CNN é proveniente da operação matemática linear de convolução, representada por \* na Equação 2.1, que pode ser brevemente definida como uma integral que expressa a sobreposição de uma função g, medida em relação a outra função f, enquanto esta é deslocada, operando com dados de entrada cujos valores são reais (WEISSTEIN, 2023).

$$[f * g](t) = \int_0^t f(\tau)g(t - \tau) \, d\tau$$
 (2.1)

A operação de convolução aplicada à arquitetura de uma CNN usualmente representa a função *f* como a camada de dados de entrada, enquanto *g* é a matriz de convolução, também conhecida como *kernel* ou máscara. O *kernel* consiste em uma matriz de valores, cujo seu produto operado sobre o valor de cada pixel da matriz de entrada resulta nos mapas de características, ou *feature maps*. Quando aplicado a imagens, um *kernel* pode produzir efeitos como *sharpening* ou *blurring*, como pode ser observado na FIG. 2.3.





**Smoothing and Sharpening Convolution Kernels** 

Fonte: Spring John C. Russ (2022).

Em uma rede convolucional, os neurônios são representados pelos filtros e se comportam como detectores de características. Eles são inicializados aleatoriamente e têm seus pesos atualizados à medida que o aprendizado é propagado pela rede. Um exemplo de alguns filtros gerados por uma CNN pode ser observado na FIG. 2.4.

#### Camada de pooling

A camada de *pooling*, também conhecida como subamostragem, tem por objetivo extrair as informações mais importantes do mapa de características obtido através da camada de convolução. Esse processo reduz o número de parâmetros entregues para a próxima camada, é a etapa de aprendizado, tornando-o mais eficiente. Para realizar essa redução, é aplicada uma função de agregação no *feature map*, deslocando-a com base no valor definido pelo parâmetro *stride* (*S*), que indica o número de pixels á serem deslocados pela função. A escolha da função de agregação e do valor do *stride* são feitas com base na problemática em questão: para CNNs, a função *Max Pool* é interessante por extrair o maior valor dentre a matriz de pixels, entregando assim parâmetros relevantes como bordas, formas e texturas. A FIG. 2.5 demonstra esse processo.

#### Função de ativação

A fim de evitar a propagação de resultados lineares e garantir que a rede neural não se torne um modelo de regressão linear, as funções de ativação são aplicadas nas arquiteturas das



Figura 2.4 – Filtros gerados nas primeiras camadas da AlexNet ao ser treinada com a base de dados ImageNet.

Fonte: Garg, Panda e Roy (2019).

Figura 2.5 – Aplicação da função *Max Pool* com um deslocamento de 2 pixels (*stride*) em um mapa de características 4\*4



Fonte: CS231n Convolutional Neural Networks for Visual Recognition (2023).

RNAs. O papel de uma função de ativação consiste em determinar se um neurônio será ou não ativado. Esse passo é crucial, visto que os problemas do mundo real frequentemente apresentam padrões não lineares e, portanto, a capacidade da CNN de aprender esses padrões complexos é fundamental para a obtenção de resultados assertivos.

Ao longo da evolução das redes neurais, a função *Sigmoid* desempenhou um papel importante na ativação de neurônios, sendo amplamente utilizada. Contudo, percebeu-se que ela apresentava um problema em sua derivada, conhecido como *vanishing gradient*, que pode prejudicar significativamente o processo de aprendizado (RASAMOELINA; ADJAILIA; SINčáK, 2020). Atualmente, outras funções de ativação têm sido usadas, como tangente hiperbólica e a *Rectified Linear Units* (ReLU), que por sua vez pode ser facilmente calculada como dado na

Equação (2.2). Em uma CNN como a U-Net, essa função sucede os cálculos feitos na camada de convolução, atribuindo zero para os pixels de valores negativos e mantendo o valor original caso contrário.

$$ReLU(x) = \begin{cases} x & \text{se } x > 0\\ 0 & \text{senão} \end{cases}$$
(2.2)

A função ReLU tem se destacado em aplicações de CNNs devido à sua simplicidade computacional. Ela é capaz de tornar as conexões entre os neurônios mais esparsas, o que acelera o processo de aprendizado e torna a rede mais eficiente (RASAMOELINA; ADJAILIA; SINčáK, 2020).

#### Função de perda

Com o propósito de avaliar a diferença entre os resultados obtidos e os desejados, a função de perda (*loss function*) é aplicada para mensurar o erro cometido pelo modelo. Baseado nesse indicador é feito o ajuste necessário nos parâmetros da rede durante o treinamento para melhorar sua assertividade. De acordo com Sultana, Sufian e Dutta (2020), algumas das funções de perda utilizadas em modelos de segmentação semântica no estado da arte são: *Dice coefficient loss, Cross entropy loss, Sum of cross-entropy loss, Standard binary cross entropy loss for boundary refinement, Per-pixel multinomial logistic loss, Semantic Encoding loss* e a *Binary Cross Entropy*, que foi a escolhida neste trabalho e se encontra representada pela Equação (2.3).

$$L(y, \hat{y}) = -(y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y}))$$
(2.3)

#### Estratégias para overfitting

O *overfitting* é um problema recorrente em RNAs onde a rede se especializa demasiadamente aos padrões encontrados nos dados de treinamento, sendo capaz de trazer estimativas corretas com alta acurácia para esses dados, porém ao mesmo tempo se tornando incapaz de generalizar para novos dados de entrada. Para minimizar esse problema, algumas técnicas são aplicadas á arquitetura da rede, como o *Dropout*, a Normalização em lote e o *Fine Tuning*.

#### Dropout

O *Dropout* é uma técnica de regularização cuja ideia está em desativar aleatoriamente alguns neurônios e suas respectivas conexões durante a fase de treinamento da RNA. Seu princípio de funcionamento reside na ideia de eliminar a interdependência entre os neurônios e forçar a rede á buscar por rotas alternativas de propagação da informação pelas camadas. Apesar de aumentar o tempo de treinamento por conta da constante atualização de parâmetros, essa técnica demonstrou resultados significativos quanto á performance e melhora do *overfitting* em metodologias que

alcançaram o estado da arte utilizando notórias bases de dados como a ImageNet e a MNIST (SRIVASTAVA et al., 2014a). A FIG. 2.6 ilustra essa técnica.



Figura 2.6 – Exemplo de aplicação da técnica de Dropout em uma RNA.

Fonte: Srivastava et al. (2014b).

#### Normalização em lote

Um fenômeno comum em redes neurais profundas que gera o *overfitting* é o *Internal Covariate Shift*. Este é causado pelas recorrentes variações nos parâmetros de entrada de cada camada que são propagadas pela rede e resultam em saídas inconsistentes, dificultando o aprendizado por estar associado á uma baixa taxa de aprendizado, o *learning rate* (IOFFE; SZEGEDY, 2015).

De acordo com Bjorck et al. (2018), a normalização em lote é um procedimento capaz de melhorar a generalização de um modelo e está associada á uma alta taxa de aprendizado. Se a rede não está normalizada, as atualizações de parâmetros que ocorrem após cada *mini-batch* acabam por serem bruscas, causando o *Internal Covariate Shift*. A utilização dessa técnica também dispensa a preocupação quanto à forma de inicialização de parâmetros e em alguns casos até mesmo elimina a necessidade de aplicar a técnica de *Dropout* em conjunto (IOFFE; SZEGEDY, 2015).

#### Aumento de dados

Uma outra técnica frequentemente utilizada com o objetivo de atenuar o fenômeno de *overfitting* é conhecida como aumento de dados. Ela tem apresentado resultados satisfatórios em modelos de aprendizagem profunda, tornando as bases de dados mais robustas e é particularmente vantajosa em situações onde há pouca disponibilidade de dados (SHORTEN; KHOSHGOFTAAR, 2019a). Essencialmente, sua premissa consiste em expandir o conjunto de dados disponível mediante a geração de réplicas das imagens, às quais são aplicadas transformações variadas. No

contexto de imagens, tais alterações podem ocorrer por meio de transformações geométricas, como rotações, ou pela aplicação de filtros, mudanças nos canais RGB e remoção de determinadas partes da imagem. Um exemplo ilustrativo pode ser visualizado na FIG. 2.7.

Figura 2.7 – Aplicação de aumento de dados na coloração de uma imagem utilizada para classificação de melanomas.



Fonte: Mikołajczyk e Grochowski (2018).

#### 2.1.3 U-Net

Criada em 2015 por Ronneberger, Fischer e Brox (2015), a U-Net é uma arquitetura de rede neural convolucional que propõe uma abordagem eficaz para a segmentação semântica de imagens. Ao contrário de outras CNNs convencionais, ela possibilita um treinamento eficiente para bases de dados limitadas, tornando-a uma metodologia interessante para a solução de problemas em imagens médicas, que por muitas vezes dispõe de um pequeno número de amostras.





Fonte: Ronneberger, Fischer e Brox (2015).

Originalmente projetada contendo 23 camadas de convolução seguidas pela função de ativação ReLU, a arquitetura da U-Net, mostrada na FIG. 2.8, é dividida em duas etapas principais: a codificação (*encoder*) e a decodificação (*decoder*). Diferente de outras RNAs, essa rede não é formada por camadas totalmente conectadas e foi chamada pelos autores (RONNEBERGER; FISCHER; BROX, 2015) de uma "rede completamente convolucional".

Durante a fase de codificação, a imagem de entrada passa por um processo de *downsampling*, onde é aplicada a função *max pool* e sua dimensão é reduzida pela metade á cada nível que percorre. Em cada um desses níveis, são aplicadas convoluções com um determinado número de filtros de tamanho  $3 \times 3$  a fim de formar os mapas de características, extraindo os atributos relevantes que serão posteriormente usados na criação da máscara de segmentação.

Após a codificação, são aplicadas convoluções com filtros de dimensão  $2 \times 2$  nos próximos níveis presentes na camada de decodificação, reduzindo pela metade o número de filtros a cada nível, enquanto o mapa de características obtido passa pelo processo de *upsampling*. Para que não ocorra perca de dados, esse mapa é redimensionado para a dimensão original da imagem utilizando a concatenação através de *skip connections*, que por sua vez irá reintroduzir características da fase de codificação que poderiam se perder ao longo da rede. Ao fim é aplicada uma convolução  $1 \times 1$ , produzindo a saída que será então um mapa de segmentação.

#### 2.1.4 Métricas para a análise de resultados da segmentação

Com o objetivo de avaliar a precisão da segmentação obtida, serão aplicados alguns métodos de avaliação supervisionados que utilizam imagens manualmente segmentadas como referência, conhecidas como *ground truth* (GT). Tais metodologias fazem uso de conceitos como o verdadeiro positivo (*true positive* - TP) e verdadeiro negativo (*true negative* - TN) que indicam a classificação correta dos pixels pertencentes e não pertencentes a região segmentada, respectivamente. De forma análoga, o falso positivo (*false positive* - FP) e o falso negativo (*false negative* - FN) se propõe á representar os pixels que foram incorretamente dados como presentes (FP) e não presentes (FN) na máscara de segmentação. Dentre diversas métricas existentes para avaliar o resultado de uma segmentação, um dos critérios usados na escolha foi a recorrência: de acordo com Amorim et al. (2020), alguns dos índices mais empregados na solução de problemas de segmentação de imagens citológicas foram Acurácia, Dice, *Recall* e Precisão, como pode ser observado no gráfico proposto por Amorim et al. (2020) e apresentado na FIG. 2.9.

#### **Coeficiente de Dice**

O coeficiente de Dice (DSC) é uma métrica de avaliação de desempenho amplamente utilizada na área de processamento de imagens. O cálculo é obtido a partir de uma comparação entre a máscara de segmentação resultante e a *ground truth*, uma razão entre as regiões em comum de ambos, que foram corretamente segmentadas, e o total de pixels presentes nos dois grupos. O resultado assume maior relevância à medida que se aproxima de 1, indicando que a





Fonte: Amorim et al. (2020).

máscara está próxima do valor esperado pela *ground truth*, como ilustra a FIG. 2.10, obtendo assim a sobreposição entre as duas áreas segmentadas. Sua Equação é dada em (2.4).

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{2.4}$$

Figura 2.10 – Representação visual do coeficiente de Dice aplicado á segmentação de imagens.



Fonte: Zou et al. (2004).

#### Acurácia

A acurácia é uma métrica que analisa a proporção entre o número de acertos obtidos (TP + TN) para o total de testes realizados, indicando a acertabilidade do modelo, conforme Equação (2.5).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$
(2.5)

#### Precisão

Com uma ideia similar á métrica de acurácia, a precisão entrega a parcela de acertos positivos TP obtidos pelo modelo dentre o total de acertos e falsos positivos (TP + TN), e é definida na Equação (2.6). Quando aplicada a problemas de segmentação de imagens, é um indicativo potencialmente útil para conhecer a capacidade da rede em segmentar corretamente a região de interesse, contudo, a precisão deve ser combinada á outras métricas para uma análise completa.

$$P = \frac{TP}{TP + FP} \tag{2.6}$$

#### Revocação

A revocação (ou *recall*) mede a proporção dos resultados positivos obtidos em relação à *ground truth*. Ou seja, o objetivo dessa métrica nesse contexto é extrair o percentual de dados corretamente inseridos na região de interesse da segmentação em comparação com aqueles, que quanto á mesma classe, foram incorretamente classificados. A Equação para o *recall* (R) é definida em (2.7).

(2.7).

$$R = \frac{TP}{TP + FN} \tag{2.7}$$

### 2.2 Trabalhos Relacionados

O estudo em segmentação de imagens para citologia cervical têm atraído a atenção de pesquisadores, tendo como objetivo final diagnósticos mais precisos e menos suscetíveis à falhas humanas. Diversas abordagens foram apresentadas nos últimos anos, evoluindo à medida que tecnologias e algoritmos mais eficientes eram introduzidos. Com os avanços no campo da inteligência artificial, trabalhos introduzindo a aprendizagem de máquina para a segmentação e classificação de imagens têm se tornado mais frequentes desde 2007 até 2018 (SARWAR et al., 2020). Nas próximas Seções 2.2.1 e 2.2.2, serão apresentados os trabalhos mais relevantes já realizados com metodologias envolvendo redes neurais e outros utilizando outros métodos, respectivamente. Logo após, na Seção 2.2.3 é feita uma discussão acerca desses trabalhos.

#### 2.2.1 Metodologias baseadas em *deep learning*

Apesar do surgimento das CNNs ter iniciado em 1998 com a criação da LeNet-5 desenvolvida por Lecun et al. (1998), sua popularização no ramo de visão computacional começou com o êxito da AlexNet, proposta por Krizhevsky, Sutskever e Hinton (2012), onde pela primeira vez uma CNN ganhou o desafio ImageNet 2012 e revolucionou o campo de aprendizagem profunda ao demonstrar sua discrepante performance em comparação aos outros competidores. Posteriormente, em 2014 surgiu a VGGNet (SIMONYAN; ZISSERMAN, 2014), que com uma arquitetura mais profunda, ganhou o desafio ImageNet em 2014 e se provou eficiente na classificação de imagens.

Em 2015 surgem as Redes Totalmente Convolucionais (FCN) (LONG; SHELHAMER; DARRELL, 2014), um marco que possibilitou o surgimento de novas arquiteturas de CNN para realizar a segmentação semântica, visto que a maioria das aplicações anteriores eram voltadas para classificação. Sua ideia principal reside em substituir as camadas totalmente conectadas por camadas de convolução transposta, sendo assim capazes de identificar e localizar as regiões de interesse na imagem ao atribuir uma saída para cada pixel de entrada. Ela também utiliza a técnica de *skip connections* para que informações de uma camadas sejam diretamente repassadas para outros níveis, evitando a perca de dados.

Arquiteturas baseadas nessa ideia, como a ResNet (HE et al., 2015), SegNet (BADRINA-RAYANAN; KENDALL; CIPOLLA, 2015), e U-Net (RONNEBERGER; FISCHER; BROX, 2015) foram desenvolvidas e começaram a ser amplamente aplicadas em tarefas de segmentação. A ResNet foi uma das primeiras a aplicar a ideia de normalização em lote, e também á introduzir a ideia de conexões residuais, proporcionando um aprendizado mais eficiente com resultados significativos quando comparados com a VGGNet (HE et al., 2015). Já a arquitetura da SegNet inovou ao propor um novo passo, o *unpooling*, que reconstrói os mapas de características com o objetivo de recuperar a resolução perdida nos processos de codificação, preservando assim os pixels mais significativos da imagem. Contudo, quando comparada á SegNet, a U-net já demonstrou resultados significativos e uma melhor performance na predição para a segmentação de núcleo de células cervicais, quando usada a base de dados Herlev (MARINAKIS; DOUNIAS; JANTZEN, 2009), como apresentado por Teixeira (2021).

Três estudos realizados entre 2016 e 2019 obtiveram resultados consideráveis ao utilizar redes neurais convolucionais para segmentar núcleos e citoplasmas em células cervicais. Esses trabalhos se destacam por terem introduzido metodologias para contornar o problema de citoplasmas sobrepostos em agrupamentos de células cervicais, uma situação comum nesse contexto onde o citoplasma de uma célula muitas vezes se sobrepõe ao de outra, o que dificulta a tarefa de segmentação pelos algoritmos, como pode ser observado na FIG. 2.11.

Em 2016, Song et al. (2016) propôs uma metodologia baseada na ideia de que cada célula é representada por seu núcleo correspondente, e que consistia em três etapas principais: segmentação de componentes celulares, rotulagem de várias células e refinamento de borda celular e inferência. Na primeira, foi utilizada uma CNN *multi-scale* para aprender os atributos das imagens de entrada. A saída dessa CNN é um vetor de características que foi usado em uma outra rede neural de duas camadas para assim classificar cada pixel em uma das três classes: núcleo, citoplasma e fundo da imagem. Logo após ocorre uma rotulagem em cada pixel presente



Figura 2.11 – Aglomerado de células sobrepostas extraído da base de dados CRIC.

Fonte: Rezende, Bianchi e Carneiro (2021).

em uma região citoplasmática, onde ele é associado ao seu núcleo correspondente através de um cálculo de probabilidade feito usando um filtro Gaussiano. Por último, é feito um refinamento de borda e inferência através de um modelo dinâmico de deformação *multi-template*, para tratar situações onde um pixel de um citoplasma é associado á duas diferentes regiões celulares. Os resultados do coeficiente de Dice para as bases de dados ISBI 2015 e SZU<sup>1</sup> foram 0,89 e 0,84, respectivamente.

Já em 2017, Tareef et al. (2017b) apresentou um *framework* que opera em duas etapas: classificação dos componentes celulares com base nas características aprendidas através de uma CNN baseada em *superpixel*, e a segmentação individual do citoplasma a partir do aglomerado de células utilizando Segmentação Voronoi e evolução de nível baseada em prioridade de forma dinâmica (*dynamic shape prior-based level set evolution*). Na primeira fase, a imagem é dividida em 3 classes: núcleo, citoplasma e fundo da imagem para que na fase posterior cada célula contenha seu devido contorno nas regiões do núcleo e citoplasma. Com essa técnica aplicada na base de dados ISBI 2014 foi possível obter resultados significativos com 0,94 de precisão e 0,95 de revocação.

Propondo um método de aprendizagem profunda, Wan et al. (2019) entregou uma metodologia que combina diversas técnicas capazes de segmentar células sobrepostas com interferências, como o ruído do fundo da imagem e um baixo contraste. Inicialmente, é usada a técnica de *double-window* para localizar individualmente cada célula através da TernausNet, que por sua vez irá classificar cada pixel da imagem em núcleo, citoplasma e fundo da imagem. Logo após é utilizado um modelo de *DeepLab V2* para operar a segmentação do citoplasma. Para aumentar o número de amostras de treinamento, foi utilizado um método de sínteses para gerar agrupamentos celulares com células sobrepostas. Os resultados dessa abordagem se mostraram promissores ao atingir no coeficiente de Dice 0,93 para a base de dados ISBI 2014 e 0,92 para a ISBI 2015.

#### 2.2.2 Outras metodologias

Embora sejam muito aplicadas em problemas de classificação de imagens, as técnicas de aprendizagem profunda ainda não foram amplamente exploradas para a segmentação de imagens de células cervicais. Antes disso, os principais trabalhos empregavam algoritmos comuns no campo de visão computacional como *k-means*, *fuzzy C-means*, e *Extended Depth of Field (EDF)*.

Em 2013, Chankong, Theera-Umpon e Auephanwiriyakul (2014) propôs a segmentação e classificação de células cervicais, onde no primeiro passo uma única célula foi separada em núcleo, citoplasma e fundo da imagem através da técnica de clusterização *fuzzy c-means* (FCM); e logo após cada uma foi classificada em quatro classes quanto à sua natureza. Os resultados demostraram boa performance para a base de dados Herlev: 0,95 de precisão na segmentação completa da célula e 0,80 de revocação.

Propondo a melhoria de um método anteriormente apresentado pelos próprios autores Phoulady et al. (2016), esse trabalho parte de imagens fornecidas através de múltiplos planos focais por meio de amostras do exame de Papanicolaou e localiza os núcleos através do algoritmo iterativo EDF. A partir disso, os aglomerados de células são segmentados usando um modelo já aprendido de mistura gaussiana através de um algoritmo de maximização de expectativa (EM) partindo das intensidades dos pixel. Por fim, técnicas de pós-processamento são aplicadas para refinar o contorno do citoplasma. Quanto aos resultados, foi possível obter 0,861 de DSC para a base de dados ISBI 2015.

No ano seguinte, os mesmos autores Phoulady et al. (2017) proporam um framework com uma metodologia similar á anterior, porém aplicada á uma segunda base de dados ISBI 2014, obtendo resultados melhores: 0,901 de DSC, ultrapassando o estado da arte naquele período.

Logo em 2018, Tareef et al. (2018) apresentou uma metodologia distinta das anteriores ao utilizar um método *watershed-based* (MPFW), para segmentar ambos núcleo e citoplasma de grandes massas de células cervicais sobrepostas em três Passes de Watershed. Essa técnica tem por objetivo separar as regiões de uma imagem em mínimos locais e máximos locais por meio da transformada de Watershed. Em seguida, é realizada a segmentação adaptada para o formato e a localização das células que se encontram isoladas, que estão em contato umas com as outras ou que estão parcialmente sobrepostas. O método foi aplicado na base de dados ISBI 2014, apresentando um DSC de 0,898, e também na base ISBI 2015, obtendo um DSC de 0,851.

#### 2.2.3 Discussão

Foram encontrados alguns trabalhos relacionados que apresentaram resultados interessantes, porém quando segmentaram o citoplasma, fizeram de forma á extrair o agrupamento celular como um todo do fundo da imagem, sem delimitar cada região citoplasmática. Dentre

Listagem de trabalhos relacionados			
Artigo	Metodologia	Base de dados	Resultados para cito- plasma
Tareef et al. (2017b)	Superpixelwise Convolutional Neutral Network	ISBI 2014	TP=0,95
Tareef et al. (2017a)	Support Vector Machine (SVM)	ISBI 2014	DSC=0,84
Tareef et al. (2018)	Multi- pass Fast Fatershed-based Method (MPFW)	ISBI 2014 e ISBI 2015	DSC=0,851
Phoulady et al. (2016)	Extended Depth of Field (EDF)	ISBI 2015 e SZU	DSC=0,861
Phoulady et al. (2017)	Extended Depth of Field (EDF)	ISBI 2014 e ISBI 2015	DSC=0,869
Song et al. (2016)	Multi-scale CNN + Dynamic multi- template deformation	ISBI 2015 e uma base de dados privada com 21 ima- gens, cada uma contendo de 30 á 80 células.	ISBI 2015: DSC=0,89; Base de da- dos privada: DSC=0,84
Wan et al. (2019)	TernausNet, Adaboost classifier, double- window method, DeepLab V2, fully con- nected CRFs e regularized level set evo- lution (DRLSE).	ISBI 2014 e ISBI 2015	ISBI 2014: DSC=0,93 e ISBI 2015: DSC=0,92
Song et al. (2015)	CNN multiescalar e particionamento de grafos.	Base de dados privada com 53 amostras	DSC=0,950

Tabela 2.1 – Trabalhos relacionados que serão usados para fins de comparação.

essas metodologias, destacam-se os trabalhos de Song et al. (2015), Zhang et al. (2017), Song et al. (2014) e Li et al. (2012). Consequentemente, o problema da sobreposição de citoplasma, que é comum em imagens obtidas pelo exame papanicolau, não é tratado, o que interfere na extração das informações necessárias para o exame. Portanto, foram incluídos como referência apenas os trabalhos que propuseram soluções para esse problema.

Uma limitação comum dentre os trabalhos encontrados que possuem os melhores resultados está nas bases de dados usadas para treinamento: Song et al. (2014) obteve um DSC igual á 0,95 para segmentação do citoplasma, contudo fez uso de apenas uma base de dados privada com um baixo número de amostras, sendo 53 ao todo.

Outros métodos que visam exclusivamente a extração do núcleo também não foram considerados para fins de comparação, visto que o objetivo desse trabalho é obter a segmentação completa da célula cervical. Os trabalhos usados como referência se encontram listados na TAB. 2.1.

## **3** Desenvolvimento

A metodologia, assim como a base de dados á ser usada na execução desse trabalho se encontram descritas nas próximas Seções 3.1 e 3.2.

### 3.1 Base de dados

A escolha da base de dados em metodologias de aprendizagem profunda é um passo crucial que interfere diretamente na eficiência e qualidade do treinamento da rede neural. Uma CNN treinada com imagens que contêm apenas uma célula cada, como é o caso da base de dados Herlev (MARINAKIS; DOUNIAS; JANTZEN, 2009), não proporcionará à rede a oportunidade de aprender as características de um conjunto de células, ou de sobreposição. Diante disso, foi escolhida uma base de dados que visa lidar com essas limitações, ao propor conjuntos de imagens de células cervicais sobrepostas. A seguir, essa base é descrita, assim como a aplicação do método de aumento de dados que foi feita sob ela.

#### 3.1.1 ISBI 2014

A base de dados sintética ISBI 2014 (LU; CARNEIRO; BRADLEY, 2015) proposta no *Overlapping Cervical Cytology Image Segmentation Challenge - ISBI 2014* no *IEEE International Symposium on Biomedical Imaging (ISBI 2014)*, dispõe de 45 imagens sintéticas para o treinamento e 90 imagens sintéticas para fins de teste. Elas foram geradas a partir de uma base de dados de citologia cervical reais, a partir de 5 imagens *Extended Depth Field* (EDF). As imagens sintéticas foram geradas de modo a conterem agrupamentos de 2 a 10 células, onde o coeficiente de sobreposição varia de 0,1 a 0,5. Posteriormente ao desafio, foram disponibilizadas 810 imagens sintéticas que seriam usadas no teste para fins de avaliação do modelo no contexto do desafio. Contudo, nesse trabalho, essas 810 imagens foram estrategicamente realocadas para agregarem o conjunto das 45 imagens de treino, totalizando 855 imagens para treino. Essa estratégia foi adotada como o intuito de melhorar o aprendizado do modelo ao fornecer a ele um grupo maior de amostras. Uma imagem real EDF usada para gerar a base sintética é ilustrada na FIG. 3.1.

#### 3.1.2 Aumento de dados

Treinar um modelo de rede neural com uma base de dados limitada exige a aplicação de técnicas para contornar essa limitação e melhorar o aprendizado. Nesse contexto, a técnica de aumento de dados emerge como uma abordagem relevante, permitindo a criação de múltiplas instâncias de dados com base nas amostras existentes e se mostrando eficaz em diversos trabalhos, como demonstra Shorten e Khoshgoftaar (2019b). Nesse trabalho, a base de dados de treino da



Figura 3.1 – Imagem de profundidade estendida (EDF) coletada da base ISBI 2014.

Fonte: Lu, Carneiro e Bradley (2015).

ISBI 2014 foi submetida ao aumento através de transformações que fossem capazes de alterar significativamente aspectos da imagem sem que a morfologia das células fossem deformadas, e ao mesmo tempo demonstrassem ser eficazes para esse contexto. Por isso, as operações foram escolhidas com base no sucesso do trabalho de Teixeira (2021), sendo elas: inversão horizontal, inversão vertical, rotação, variação das faixas de brilho e contraste. Um exemplo de aplicação dessas transformações pode ser encontrado na FIG. 3.3. Para a rotação, foram definidos os ângulos de 90°, 180° e 270° para serem escolhidos aleatoriamente. Essa seleção foi feita para que não ocorressem rotações com ângulos que gerassem bordas preenchidas por algoritmos como o *nearest neighbor*, que nesse caso estava corrompendo o contorno da célula e com isso prejudicando a integridade da imagem, como pode ser observado na FIG. 3.2. Além da rotação, outras transformações também foram aplicadas aleatoriamente, com os detalhes dessas transformações e os valores associados a elas listados na TAB. 3.1. Esse procedimento foi repetido três vezes para cada amostra, com o objetivo de introduzir variações significativas na aparência da imagem, sem gerar dados redundantes.

Figura 3.2 – Exemplo de imagem rotacionada que gerou bordas preenchidas automaticamente utilizando o algoritmo *nearest neighbor*.



Fonte: elaboração própria a partir de uma imagem coletada da base ISBI 2014.

Transformações	Intervalo de variação
Inversão horizontal,	Verdadeiro
Inversão vertical,	Verdadeiro
Rotação,	[90°, 180°, 270°]
Faixa de brilho,	[0,2 - 0,9]
Faixa de contraste	[0,2 - 0,9]

Tabela 3.1 – Transformações e suas variações de valores.





(e) Alteração na faixa de contraste

(f) Alteração na faixa de brilho

Fonte: elaboração própria a partir de uma imagem coletada da base ISBI 2014.

### 3.2 Metodologia

A metodologia adotada neste estudo foi estruturada em três etapas fundamentais: a primeira consistiu na preparação dos dados (Fase 1); a segunda envolveu o treinamento da U-Net para a segmentação tanto do núcleo quanto do citoplasma (Fase 2); por fim, procedeu-se à análise dos resultados obtidos (Fase 3). O diagrama desse processo é apresentado na FIG. 3.4, e a descrição das etapas 1 e 2 segue nas Seções 3.2.1 e 3.2.2, respectivamente, enquanto a terceira etapa, relacionada aos resultados, será abordada na Seção 4.

#### 3.2.1 Preparação dos dados

Na fase 1, a base de dados foi organizada de forma a compor os grupos de treino, teste e validação. Para que os experimentos fossem executados com uma avaliação robusta através de múltiplas divisões da base, reduzindo o viés na seleção de dados, a base de dados ISBI foi



Figura 3.4 – Fluxograma proposto para o desenvolvimento.

Fonte: elaboração própria.

dividida seguido a metodologia *k-fold cross validation*, técnica essa que já se provou muito eficaz em métodos de *deep learning*, como no trabalho de Lyu et al. (2022). Dessa forma, as 855 imagens disponíveis para treino foram subdividas em 5 grupos, sendo 4 direcionados ao grupo treino e um ao grupo validação. Tal procedimento foi executado cinco vezes de forma aleatória, e posteriormente foi aplicada a técnica de aumento de dados descrita na Seção 3.1.2. Com isso, foram geradas 3 imagens á partir de cada amostra, resultando em 684 imagens em cada *fold* de treino. Esse processo entregou cinco agrupamentos diferentes da mesma base, como demonstra a FIG. 3.5. Ainda na fase 1, realizou-se o pré-processamento dos dados com o objetivo de prepará-los para inserção na rede neural. Isso incluiu o redimensionamento das imagens para se adequarem às dimensões apropriadas para a camada de entrada. Além disso, com o intuito de aprimorar os resultados do citoplasma e entregar á rede imagens *ground truth* apenas com a região citoplasmática definida, as máscaras do núcleo e do citoplasma foram combinadas, excluindo o núcleo. Um exemplo desse procedimento pode ser observado na FIG. 3.6.



Figura 3.5 – Fluxograma da divisão da base de dados ISBI 2014 em k-folds.

Fonte: elaboração própria.

Figura 3.6 – Procedimento de remoção da área do núcleo na região citoplasmática. A imagem (c) corresponde á máscara do núcleo que foi combinada á imagem (b) para resultar a imagem (d) que contém apenas a região do citoplasma (em branco).



(a) Imagem original



(b) Máscara do citoplasma



(c) Máscara do núcleo



(d) Resultado final

Fonte: elaboração própria a partir de imagens coletadas da base ISBI 2014.

#### 3.2.2 Treinamento

Na segunda etapa do processo, procedeu-se ao treinamento da rede U-Net para a segmentação tanto do núcleo quanto do citoplasma, de maneira separada e independente. No que diz respeito à implementação da arquitetura da rede, foi empregada uma versão adaptada da arquitetura U-Net, disponibilizada em Karol Żak (2023), que permite a customização de parâmetros da rede, como os valores de *dropout* e resolução das imagens de entrada, a configuração do uso de

Figura 3.7 – Fluxograma de treinamento da rede U-Net utilizado por Teixeira (2021) e aplicado também neste trabalho.



Fonte: elaboração própria.

normalização em lote e da função de saída. Os valores escolhidos para esses parâmetros nesse trabalho se encontram na TAB. 3.2.

O ciclo de treinamento para cada região celular, ilustrado na FIG. 3.7, foi configurado para abranger no máximo 300 épocas, com um tamanho de lote fixo de 64. Cada ciclo se inicia com o treino do modelo, onde ao final de cada época, a rede conduz validações internas para ajustar seus parâmetros e avaliar se houve ou não melhoria em relação à iteração anterior. Se sim, os pesos são salvos, mas caso contrário, os resultados são agora comparados aos das últimas 25 épocas, e se não houver melhoria, é realizado o processo de parada antecipada (*early stop*). Quando ocorre uma parada antecipada, o valor da taxa de aprendizado é reduzido, a menos que já tenha atingido o seu limite inferior, e o treinamento da rede continua. No entanto, se essa época resultou em melhorias nos resultados, é avaliado se a época final foi atingida, e se for o caso, encerra-se o ciclo. Quanto à variação da taxa de aprendizado, foram realizadas quatro iterações dessa variação, com valores decrescentes de  $10^{-3}$ ,  $10^{-4}$ ,  $10^{-5}$  e  $10^{-6}$ . Todo esse processo foi executado uma vez para cada *k-fold gerado*, resultando em 5 treinos para núcleo e 5 treinos para citoplasma.

Para que o treinamento do modelo fosse eficaz e bem sucedido, foi utilizado como referência os parâmetros que obtiveram os melhores resultados no treinamento da rede U-Net para segmentação da base de dados ISBI 2014, conforme demonstrado no trabalho de Teixeira (2021). Dessa forma, a TAB. 3.2 dispõe dos valores atribuídos aos parâmetros.

Parâmetros	Valores atribuídos	
Tamanho de lote	64	
Função de ativação	Relu	
Normalização em lote	Verdadeiro	
Dropout	0,5	
Função de ativação de saída	sigmoid	
Resolução das imagens de entrada	128*128*1	

Tabela 3.2 – Valores atribuídos os parâmetros de treinamento do modelo.

#### Ferramentas

As principais ferramentas utilizadas na realização desse trabalho incluíram a linguagem Python (PYTHON, 2023) para o desenvolvimento do código, as plataformas Google Colaboratory<sup>1</sup> para a execução do código de treinamento do modelo, e o Google Scripts<sup>2</sup> para a manipulação das imagens armazenadas no Google Drive <sup>3</sup>. Foram usadas bibliotecas como NumPy (NUMPY, 2023), TensorFlow (TENSORFLOW, 2023) e xlwt (XLWT, 2023). Os experimentos executados no Google Colaboratory contaram com a seguinte configuração de máquina, disposta na TAB. 3.3.

Tabela 3.3 – Configuração de máquina disponibilizada pela plataforma Google Colaboratory para a execução dos experimentos.

Configuração	Capacidade
RAM	12,7 GB
DISK	78,2 GB
GPU	T4 GPU

<sup>&</sup>lt;sup>1</sup> https://colab.google/

<sup>&</sup>lt;sup>2</sup> www.google.com/script/start/

<sup>&</sup>lt;sup>3</sup> www.google.com/intl/pt-br/drive/about.html

## **4 Resultados**

Este capítulo apresenta os resultados do treinamento da rede neural U-Net aplicada à segmentação de núcleo e citoplasma, realizada como parte da metodologia deste estudo. Durante este capítulo, serão discutidos os resultados obtidos em ambas as tarefas de segmentação, juntamente com a análise das métricas de avaliação, incluindo *Loss*, Precisão, Revocação, Dice e Acurácia. O objetivo é fornecer uma compreensão detalhada do desempenho da rede neural em relação à segmentação das estruturas que compõem a célula. Quanto aos limiares de precisão, revocação e Dice, eles foram estabelecidos para monitorar a qualidade da segmentação, seguindo o estado da arte, que requer que pelo menos 70% da região seja segmentada. A função de perda escolhida para a métrica *Loss* foi a *Binary Cross Entropy* (RUBY; YENDAPALLI, 2020).

A fim de avaliar os resultados alcançados em cada métrica para os 5 modelos gerados, foi calculada a média e o desvio padrão á partir das métricas encontradas ao fim de cada treino. Esses cálculos foram feitos separadamente para os resultados do núcleo e citoplasma, como pode ser observado nas Tabelas 4.1 e 4.2, respectivamente.

Métrica	Média	Desvio Padrão
Perda	0,0062	0,0002
Precisão	0,979	0,002
Revocação	0,80	0,01
Dice	0,867	0,008
Acurácia	0,99813	0,00009

Tabela 4.1 – Resultado final para segmentação do núcleo

Tabela 4.2 – Resultado final para segmentação do citoplasma

Métrica	Média	Desvio padrão
Perda	0,0149	0,0003
Precisão	0,979	0,003
Revocação	0,986	0,003
Dice	0,9824	0,0006
Acurácia	0,9954	0,0002

Após o treinamento e obtenção dos resultados, foram realizadas predições para a segmentação tanto do núcleo, conforme ilustrado em dois exemplos na FIG. 4.1, quanto do citoplasma, como demonstrado em dois exemplos na FIG. 4.2. Também foram gerados gráficos para observar a evolução do treinamento de rede para as duas regiões celulares, dispostos nas FIG. 4.3 e 4.4.



Figura 4.1 – Exemplos de predição da U-Net para segmentação do núcleo.

Fonte: elaboração própria a partir de imagens coletadas da base ISBI 2014.

Figura 4.2 – Exemplos de predição da U-Net para segmentação do citoplasma.



Fonte: elaboração própria a partir de imagens coletadas da base ISBI 2014.

## 4.1 Discussões

Conforme os resultados obtidos por Teixeira (2021), era esperado para esta análise uma pontuação aproximada de 0,9743 para o coeficiente Dice na segmentação do núcleo e uma pontuação relativamente inferior para a segmentação do citoplasma, uma vez que esta última envolve uma região estruturalmente mais complexa para fins de segmentação. No entanto, foi observado que a base de dados ISBI alcançou resultados superiores na segmentação do citoplasma, registrando uma pontuação notável de 0,9825.



Figura 4.3 – Evolução do treinamento da U-Net para segmentação do núcleo.

Fonte: elaboração própria.

Figura 4.4 – Evolução do treinamento da U-Net para segmentação do citoplasma.



Fonte: elaboração própria.

Considerando alguns estudos que alcançaram resultados significativos para o citoplasma, nossa abordagem também se destaca, apresentando um coeficiente Dice superior, conforme evidenciado na TAB. 4.3.

No que diz respeito aos resultados da segmentação do núcleo, a explicação para os valores abaixo do esperado pode residir na falta de recortes aplicados aos dados de treinamento, o que poderia auxiliar a rede na identificação de núcleos com áreas menores, entre outros fatores. Ou seja, os núcleos são estruturas muito pequenas quando comparadas ao tamanho da imagem. Uma alternativa viável seria iniciar o treinamento do núcleo com um pré-treinamento realizado a partir de uma base de dados distinta, ou então, inicializar os pesos da rede com os valores obtidos a partir do treinamento do módulo de citoplasma.

Um aspecto de relevância a ser destacado neste trabalho diz respeito à realocação das 810 imagens de teste para o conjunto de treinamento. Em linhas gerais, as 45 imagens iniciais

não apresentavam citoplasmas segmentados individualmente, o que restringiria a capacidade da rede em aprender a separar os citoplasmas. No entanto, nas 810 imagens subsequentes, foram fornecidas máscaras que continham essa delimitação, permitindo à rede efetuar previsões mais precisas e separar as células dos aglomerados celulares, como ilustrado na FIG. 4.5.

Figura 4.5 – Imagens originais e as suas respectivas predições com as células individualmente delimitadas.



Fonte: elaboração própria a partir de imagens coletadas da base ISBI 2014.

Tabela 4.3 – Comparação de resultados da nossa proposta com outros trabalhos voltados que realizam a segmentação do citoplasma.

Trabalho	Dice
Tareef et al. (2018)	0,851
Phoulady et al. (2017)	0,869
Song et al. (2015)	0,950
Nossa abordagem	0,982

## **5** Considerações finais

A segmentação do citoplasma representa um desafio substancial no âmbito da citologia digital, especialmente devido à ocorrência de sobreposição entre as células. Entretanto, é inegável a contribuição significativa que as técnicas de aprendizado profundo (*deep learning*) oferecem para essa tarefa, como evidenciado não apenas pelos estudos de Song et al. (2015), mas também pelos resultados alcançados por Teixeira (2021), e nesta proposta.

Neste trabalho realizamos a segmentação do núcleo e do citoplasma em imagens citológicas sintéticas utilizando uma rede neural convolucional (U-Net). Realizamos treinamentos independentes para núcleo e citoplasma, alcançando resultados significativos, cuja pontuação índice de Dice de 0, 9824 superou consideravelmente os resultados mais relevantes na literatura que utilizaram metodologias similares, como os de Tareef et al. (2018), Phoulady et al. (2017), e Song et al. (2015). Para a tarefa de segmentação do núcleo, o índice de Dice (DSC) alcançou o valor de 0, 867, indicando a necessidade de implementação de estratégias adicionais visando uma segmentação mais precisa do núcleo e uma aprendizagem mais eficiente pelo modelo.

Assim, ao longo do desenvolvimento deste estudo, foi possível alcançar resultados consideravelmente superiores na segmentação do citoplasma, bem como na geração de previsões com células delimitadas de forma individual, o que se configura como um diferencial em relação a outros trabalhos similares. Esse desfecho ilustra com êxito a eficácia da metodologia adotada aqui, que inclui o aumento de dados, a divisão da base por meio da técnica de *k-fold cross validation*, o tratamento correto das imagens na fase de pré-processamento, o treinamento, e, finalmente a seleção criteriosa dos parâmetros utilizados durante os treinamentos.

### 5.1 Trabalhos futuros

Quanto às perspectivas futuras de pesquisa, é possível elencar algumas áreas de interesse, a saber:

- A aplicação das técnicas empregadas neste estudo para a segmentação de imagens reais, como as contidas na base de dados CRIC *database* (REZENDE; BIANCHI; CARNEIRO, 2021);
- O desenvolvimento de um algoritmo capaz de combinar as máscaras de segmentação do núcleo com as máscaras de segmentação do citoplasma, a fim de proporcionar imagens com uma definição nítida e clara das regiões devidamente separadas.

- Realizar comparações entre a metodologia proposta e a ferramenta Segment Anything Model - SAM<sup>1</sup> da Meta, como forma de comparações entre ferramentas mais genéricas.
- Propor novas metodologias usando *cross dataset* para treinamento, treina com uma informação e testa em outra.

<sup>&</sup>lt;sup>1</sup> https://segment-anything.com/

## Referências

ALBAWI, S.; MOHAMMED, T. A.; ALZAWI, S. Understanding of a convolutional neural network. In: . [S.l.: s.n.], 2017.

AMORIM, J. A.; CERENTINI, A.; MACARINI, L.; MATIAS, A.; WANGENHEIM, A. V. 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. 2020.

BADRINARAYANAN, V.; KENDALL, A.; CIPOLLA, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *CoRR*, abs/1511.00561, 2015. Disponível em: <a href="http://arxiv.org/abs/1511.00561">http://arxiv.org/abs/1511.00561</a>.

BJORCK, N.; GOMES, C. P.; SELMAN, B.; WEINBERGER, K. Q. Understanding batch normalization. In: BENGIO, S.; WALLACH, H.; LAROCHELLE, H.; GRAUMAN, K.; CESA-BIANCHI, N.; GARNETT, R. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2018. v. 31. Disponível em: <a href="https://proceedings.neurips.cc/paper/2018/file/36072923bfc3cf47745d704feb489480-Paper.pdf">https: //proceedings.neurips.cc/paper/2018/file/36072923bfc3cf47745d704feb489480-Paper.pdf</a>>.

BVS. *Exame preventivo do câncer de colo uterino (Papanicolau)*. 2022. <https: //bvsms.saude.gov.br/bvs/dicas/237\_papanicolau.html#:~:text=%C3%89%20um%20teste% 20realizado%20para,m%C3%A9todo%20no%20in%C3%ADcio%20do%20s%C3%A9culo.> Acesso em: 19 mar. 2023.

CARVALHO, M. de. *Construindo o saber: técnicas de metodologia científica*. [S.l.]: Papirus Editora, 1989. ISBN 9788530800710.

CASTILLO, M.; ASTUDILLO, A.; CLAVERO, O.; VELASCO, J.; IBáñEZ, R.; SANJOSÉ, S. de. Poor cervical cancer screening attendance and false negatives. a call for organized screening. *PLOS ONE*, Public Library of Science, v. 11, n. 8, p. 1–9, 08 2016. Disponível em: <a href="https://doi.org/10.1371/journal.pone.0161403">https://doi.org/10.1371/journal.pone.0161403</a>.

CHANKONG, T.; THEERA-UMPON, N.; AUEPHANWIRIYAKUL, S. Automatic cervical cell segmentation and classification in pap smears. *Computer Methods and Programs in Biomedicine*, v. 113, n. 2, p. 539–556, 2014. ISSN 0169-2607. Disponível em: <a href="https://www.sciencedirect.com/science/article/pii/S0169260713004021">https://www.sciencedirect.com/science/article/pii/S0169260713004021</a>.

CHAPMAN, C. N.; OTIS, C. N. From critical values to critical diagnoses. *Cancer Cytopathology*, v. 119, n. 3, p. 148–157, 2011. Disponível em: <a href="https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.1002/cncy.20158">https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.1002/cncy.20158</a>.

CHOLLET, F. Deep Learning with Python, Second Edition. [S.I.]: Manning; 2nd edition, 2021.

COSTA CARLA MAYLA SILVA DE MELO, E. d. S. L. J. C. R. d. C. A. P. M. S. H. d. A. M. M. C. O. Fatores que provocam resultados falso-negativos nos exames de citologia oncótica: uma revisão integrativa. 2021.

CS231n Convolutional Neural Networks for Visual Recognition. *Convolutional Neural Networks* (*CNNs / ConvNets*). 2023. <a href="https://cs231n.github.io/convolutional-networks/#norm">https://cs231n.github.io/convolutional-networks/#norm</a>>. Acesso em: 5 mar. 2023.

DINIZ, D. N. Segmentação de núcleos em células cervicais obtidas em exames de Papanicolaou. Dissertação (Mestrado), Ouro Preto, 2019.

DINIZ, D. N.; VITOR, R. F.; BIANCHI, A. G.; DELABRIDA, S.; CARNEIRO, C. M.; USHIZIMA, D. M.; de Medeiros, F. N.; SOUZA, M. J. An ensemble method for nuclei detection of overlapping cervical cells. *Expert Systems with Applications*, Elsevier, v. 185, p. 115642, 2021.

GARG, I.; PANDA, P.; ROY, K. A low effort approach to structured cnn design using pca. *IEEE Access*, PP, p. 1–1, 12 2019.

HARALICK, R.; SHAPIRO, L. Image segmentation techniques. *Computer Vision, Graphics, and Image Processing*, v. 29, n. 1, p. 100–132, 1985. ISSN 0734-189X.

HARDESTY, L. *Explained: Neural networks*. 2017. <https://news.mit.edu/2017/ explained-neural-networks-deep-learning-0414/>.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015. Disponível em: <a href="http://arxiv.org/abs/1512.03385">http://arxiv.org/abs/1512.03385</a>>.

INCA. 1° Seminário em Radioterapia. [S.l.]: INCA, 2001. v. 1.

INCA. Entenda o conceito do câncer do colo do útero e sua magnitude no Brasil. 2022. <https://www.gov.br/inca/pt-br/assuntos/gestor-e-profissional-de-saude/ controle-do-cancer-do-colo-do-utero/conceito-e-magnitude>. Acesso em: 31 jan. 2023.

Instituto Nacional do Câncer. *Incidência — Instituto Nacional de Câncer - INCA*. 2022. <https://www.gov.br/inca/pt-br/assuntos/gestor-e-profissional-de-saude/ controle-do-cancer-do-colo-do-utero/dados-e-numeros/incidencia#:~:text=No%20Brasil% 2C%20exclu%C3%ADdos%20os%20de,mulheres%20(INCA%2C%202022).> Acesso em: 16 mar. 2023.

International Agency for Research on Cancer. *Cancer Today*. 2023. <https://gco.iarc.fr/today/ online-analysis-pie?v=2020&mode=cancer&mode\_population=continents&population= 900&populations=900&key=total&sex=2&cancer=39&type=1&statistic=5&prevalence= 0&population\_group=0&ages\_group%5B%5D=0&ages\_group%5B%5D=17&nb\_items= 7&group\_cancer=1&include\_nmsc=1&include\_nmsc\_other=1&half\_pie=0&donut=0# collapse-group-0-1>. Acesso em: 16 mar. 2023.

IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167, 2015. Disponível em: <a href="http://arxiv.org/abs/1502.03167">http://arxiv.org/abs/1502.03167</a>>.

JEONG, J.; YOON, T.; PARK, J. Towards a meaningful 3d map using a 3d lidar and a camera. *Sensors*, v. 18, p. 2571, 08 2018.

Karol Żak. *keras-unet*. 2023. <https://github.com/karolzak/keras-unet>. Acesso em: 28 ago. 2023.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F.; BURGES, C.; BOTTOU, L.; WEINBERGER, K. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2012. v. 25. Disponível em: <a href="https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b">https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b</a>-Paper.pdf>.

KUMAR, A. *Different Types of CNN Architectures Explained: Examples*). 2022. Disponível em: <a href="https://vitalflux.com/different-types-of-cnn-architectures-explained-examples">https://vitalflux.com/different-types-of-cnn-architectures-explained-examples</a>.

LAAK, J. van der. *Computer-aided Diagnosis: The Tipping Point for Digital Pathology*. 2017. <a href="https://digitalpathologyassociation.org/blog/computer-aided-diagnosis-the-tipping-point-for-digital-pathology/>. Acesso em: 31 jan. 2023.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2324, 1998.

LI, K.; LU, Z.; LIU, W.; YIN, J. Cytoplasm and nucleus segmentation in cervical smear images using radiating gvf snake. *Pattern recognition*, Elsevier, v. 45, n. 4, p. 1255–1264, 2012.

LIN, C.-H.; CHEN, C.-C. Image segmentation based on edge detection and region growing for thinprep-cervical smear. *IJPRAI*, v. 24, p. 1061–1089, 11 2010.

LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. *CoRR*, abs/1411.4038, 2014. Disponível em: <a href="http://arxiv.org/abs/1411.4038">http://arxiv.org/abs/1411.4038</a>>.

LU, Z.; CARNEIRO, G.; BRADLEY, A. P. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *IEEE Transactions on Image Processing*, v. 24, n. 4, p. 1261–1272, 2015.

LYU, Z.; YU, Y.; SAMALI, B.; RASHIDI, M.; MOHAMMADI, M.; NGUYEN, T. N.; NGUYEN, A. Back-propagation neural network optimized by k-fold cross-validation for prediction of torsional strength of reinforced concrete beam. *Materials*, MDPI, v. 15, n. 4, p. 1477, 2022.

MARINAKIS, Y.; DOUNIAS, G.; JANTZEN, J. Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification. *Computers in Biology and Medicine*, Elsevier, v. 39, n. 1, p. 69–78, 2009.

MIKOŁAJCZYK, A.; GROCHOWSKI, M. Data augmentation for improving deep learning in image classification problem. In: IEEE. 2018 international interdisciplinary PhD workshop (*IIPhDW*). [S.1.], 2018. p. 117–122.

MINAEE, S.; BOYKOV, Y. Y.; PORIKLI, F.; PLAZA, A. J.; KEHTARNAVAZ, N.; TERZOPOULOS, D. Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, 2021.

MOSHAVEGH, R.; BEJNORDI, B. E.; MEHNERT, A.; SUJATHAN, K.; MALM, P.; BENGTSSON, E. Automated segmentation of free-lying cell nuclei in pap smears for malignancy-associated change analysis. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, v. 2012, p. 5372–5, 08 2012.

NUMPY. 2023. <https://numpy.org/>.

PANGARKAR, M. A. The bethesda system for reporting cervical cytology. *Cytojournal*, Cytopathology Foundation, Inc., v. 19, 2022.

PHOULADY, H. A.; GOLDGOF, D.; HALL, L. O.; MOUTON, P. R. A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images. *Computerized Medical Imaging and Graphics*, Elsevier, v. 59, p. 38–49, 2017.

PHOULADY, H. A.; GOLDGOF, D. B.; HALL, L. O.; MOUTON, P. R. A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images. In: IEEE. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). [S.1.], 2016. p. 201–204.

PYTHON. 2023. <https://www.python.org/>.

RAMPAZZO, L. Metodologia científica. [S.l.]: Edições Loyola, 2005. ISBN 9788515024988.

RASAMOELINA, A. D.; ADJAILIA, F.; SINčáK, P. A review of activation function for artificial neural network. In: 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI). [S.1.: s.n.], 2020. p. 281–286.

REZENDE, M. T.; BIANCHI, A. G. C.; CARNEIRO, C. M. Cervical cancer: Automation of pap test screening. *Diagnostic Cytopathology*, v. 49, n. 4, p. 559–574, 2021. Disponível em: <a href="https://onlinelibrary.wiley.com/doi/abs/10.1002/dc.24708">https://onlinelibrary.wiley.com/doi/abs/10.1002/dc.24708</a>>.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015. Disponível em: <a href="http://arxiv.org/abs/1505.04597">http://arxiv.org/abs/1505.04597</a>.

RUBY, U.; YENDAPALLI, V. Binary cross entropy with deep learning technique for image classification. *International Journal of Advanced Trends in Computer Science and Engineering*, v. 9, 10 2020.

SAFAEIAN, M.; SOLOMON, D.; CASTLE, P. E. Cervical cancer prevention—cervical screening: Science in evolution. *Obstetrics and Gynecology Clinics of North America*, v. 34, n. 4, p. 739–760, 2007. ISSN 0889-8545. Preventing Malignancies of the Reproductive Tract. Disponível em: <a href="https://www.sciencedirect.com/science/article/pii/S0889854507000861">https://www.sciencedirect.com/science/article/pii/S0889854507000861</a>>.

SARWAR, A.; SHEIKH, A. A.; MANHAS, J.; SHARMA, V. Segmentation of cervical cells for automated screening of cervical cancer: a review. *Artificial Intelligence Review*, Springer, v. 53, p. 2341–2379, 2020.

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of big data*, SpringerOpen, v. 6, n. 1, p. 1–48, 2019.

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of big data*, SpringerOpen, v. 6, n. 1, p. 1–48, 2019.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

SONG, Y.; TAN, E.-L.; JIANG, X.; CHENG, J.-Z.; NI, D.; CHEN, S.; LEI, B.; WANG, T. Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE transactions on medical imaging*, IEEE, v. 36, n. 1, p. 288–300, 2016.

SONG, Y.; ZHANG, L.; CHEN, S.; NI, D.; LI, B.; ZHOU, Y.; LEI, B.; WANG, T. A deep learning based framework for accurate segmentation of cervical cytoplasm and nuclei. In: IEEE. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. [S.1.], 2014. p. 2903–2906.

SONG, Y.; ZHANG, L.; CHEN, S.; NI, D.; LEI, B.; WANG, T. Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 62, n. 10, p. 2421–2433, 2015.

SPRING JOHN C. RUSS, M. W. D. K. R. *Basic Concepts in Digital Image Processing*. 2022. <a href="https://www.olympus-lifescience.com/pt/microscope-resource/primer/digitalimaging/imageprocessingintro/">https://www.olympus-lifescience.com/pt/microscope-resource/primer/digitalimaging/imageprocessingintro/</a>.

SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, JMLR. org, v. 15, n. 1, p. 1929–1958, 2014.

SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, v. 15, n. 56, p. 1929–1958, 2014. Disponível em: <a href="http://jmlr.org/papers/v15/srivastava14a.html">http://jmlr.org/papers/v15/srivastava14a.html</a>.

SULTANA, F.; SUFIAN, A.; DUTTA, P. Evolution of image segmentation using deep convolutional neural network: A survey. *Knowledge-Based Systems*, v. 201-202, p. 106062, 2020. ISSN 0950-7051. Disponível em: <a href="https://www.sciencedirect.com/science/article/pii/S0950705120303464">https://www.sciencedirect.com/science/article/pii/S0950705120303464</a>>.

TAREEF, A.; SONG, Y.; CAI, W.; HUANG, H.; CHANG, H.; WANG, Y.; FULHAM, M.; FENG, D.; CHEN, M. Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation. *Neurocomputing*, Elsevier, v. 221, p. 94–107, 2017.

TAREEF, A.; SONG, Y.; HUANG, H.; WANG, Y.; FENG, D.; CHEN, M.; CAI, W. Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling. *Neurocomputing*, Elsevier, v. 248, p. 28–40, 2017.

TAREEF, A.; SONG, Y.; HUANG, H.; FENG, D.; CHEN, M.; WANG, Y.; CAI, W. Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. *IEEE transactions on medical imaging*, IEEE, v. 37, n. 9, p. 2044–2059, 2018.

TEIXEIRA, J. B. A. Segmentação De Núcleos Reais De Citologia Cervical Com Redes Neurais Convolucionais. 2021. Monografia(Graduação em Ciência da Computação).

TENSORFLOW. 2023. < https://www.tensorflow.org/>.

WAN, T.; XU, S.; SANG, C.; JIN, Y.; QIN, Z. Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. *Neurocomputing*, Elsevier, v. 365, p. 157–170, 2019.

WEISSTEIN, E. "Convolution"From MathWorld–A Wolfram Web Resource. 2023. <a href="https://mathworld.wolfram.com/Convolution.html">https://mathworld.wolfram.com/Convolution.html</a>. Acesso em: 26 fev. 2023.

XLWT. 2023. <https://pypi.org/project/xlwt/>.

ZHANG, L.; KONG, H.; CHIN, C.; LIU, S.; CHEN, Z.; WANG, T.; CHEN, S. Segmentation of cytoplasm and nuclei of abnormal cells in cervical cytology using global and local graph cuts. *Computerized Medical Imaging and Graphics*, v. 38, 07 2014.

ZHANG, L.; KONG, H.; LIU, S.; WANG, T.; CHEN, S.; SONKA, M. Graph-based segmentation of abnormal nuclei in cervical cytology. *Computerized Medical Imaging and Graphics*, Elsevier, v. 56, p. 38–48, 2017.

ZOU, J.; XUE, Z.; BROWN, G.; LONG, R.; ANTANI, S. Deep learning for nuclei segmentation and cell classification in cervical liquid based cytology. In: CHEN, P.-H.; DESERNO, T. M. (Ed.). *Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications*. Houston, USA: SPIE, 2020. v. 11318, p. 1131811.

ZOU, K. H.; WARFIELD, S. K.; BHARATHA, A.; TEMPANY, C. M.; KAUS, M. R.; HAKER, S. J.; WELLS, W. M. r.; JOLESZ, F. A.; KIKINIS, R. Statistical validation of image segmentation quality based on a spatial overlap index. *Academic radiology*, Elsevier, v. 11, n. 2, p. 178–189, Feb 2004.