



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CAMPUS FLORIANÓPOLIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

André Victória Matias

**Uso de Redes Neurais Convolucionais para Detecção Precoce de Câncer de  
Cavidade Oral em Exames de Citologia Convencional com Coloração de  
Papanicolau**

Florianópolis/SC  
2022

André Victória Matias

**Uso de Redes Neurais Convolucionais para Detecção Precoce de Câncer de  
Cavidade Oral em Exames de Citologia Convencional com Coloração de  
Papanicolau**

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Santa Catarina como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Aldo von Wangenheim, Dr. rer.nat.

Florianópolis/SC

2022

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

Matias, André

Uso de Redes Neurais Convolucionais para Detecção  
Precoce de Câncer de Cavidade Oral em Exames de Citologia  
Convencional com Coloração de Papanicolau / André Matias ;  
orientador, Aldo von Wangenheim, 2022.

60 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.  
Citologia. 4. Visão Computacional. I. von Wangenheim,  
Aldo. II. Universidade Federal de Santa Catarina. Programa  
de Pós-Graduação em Ciência da Computação. III. Título.

André Victória Matias

**Uso de Redes Neurais Convolucionais para Detecção Precoce de Câncer de  
Cavidade Oral em Exames de Citologia Convencional com Coloração de  
Papanicolau**

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

Prof. Mateus Grellert, Dr.

Instituição UFSC

Prof. Felipe Perozzo Daltoé, Dr.

Instituição UFSC

Prof. Antonio Carlos Sobieranski, Dr.

Instituição UFSC

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. rer.nat.  
Orientador

Florianópolis/SC, 2022.

## **AGRADECIMENTOS**

Agradeço ao meu noivo, Fabio Martins Vieira, por todo o apoio e carinho para me manter motivado. Agradeço à minha família pela base para poder concluir meus objetivos. Agradeço ao meu orientador pelo suporte e pelos ensinamentos para possibilitar este trabalho. Agradeço ao Professor Dr. Felipe Daltoé por ter contribuído tanto com este projeto. Agradeço aos meus colegas de projeto e a todos que contribuíram com a criação do *dataset* por toda a parceria e colaboração: João Amorim, Luiz Antonio Macarini, Allan Cerentini, Bruna Zanicoski, Filipe Landa e Ricardo Pereira.

*“Robots are not going to replace humans, they are going to make their jobs much more humane.  
Difficult, demeaning, demanding, dangerous, dull – these are the jobs robots will be taking.”  
(HAUERT, 2017)*

## RESUMO

O câncer de cavidade oral é um problema de saúde global, com mais de 350.000 novos casos a cada ano. Essa patologia, se descoberta em fases iniciais, possui grandes chances de cura. A citologia usando o corante de Papanicolau é um método de baixo custo e não invasivo que pode ser usado como exame preventivo e a análise computacional tem a capacidade de diminuir a subjetividade inerente à análise manual desse processo. Neste trabalho, procura-se identificar qual a melhor abordagem para essa automatização usando Redes Neurais Convolucionais. Para isso, foi realizado um levantamento do estado da arte dessa aplicação e elaborado um conjunto de dados com 8 pacientes e 6.753 imagens para treinamento das redes. A partir disso, foram testados e comparados diferentes modelos de segmentação (U-Net e HRNet), detecção (Faster R-CNN e Mask R-CNN) e classificação de imagens (HRNet e ResNet) com abordagens binárias e multiclasse para localizar e classificar núcleos celulares em imagens de citologia convencional de cavidade oral usando a coloração de Papanicolau. Os experimentos mostraram que a abordagem de detecção de objetos usando o modelo Mask R-CNN com *backbone* ResNet 101 é a melhor para detecção de núcleos anormais no conjunto de dados (0,74 de  $F_1$  Score). Além disso, o modelo de detecção Faster R-CNN com *backbone* HRNet 48 mostrou os melhores resultados para detecção de núcleos saudáveis (0,78 de  $F_1$  Score). Os resultados reforçam ainda que a análise assistida por computador de amostras de citologia pode ajudar a aumentar a adesão a esse método como exame preventivo.

**Palavras-chave:** Citologia; Deep Learning; Detecção; Segmentação; Visão Computacional.

## ABSTRACT

Oral cancer is a global health problem with over 350,000 people affected each year. This pathology, if early detected, has a high cure rate. The Papanicolaou stained cytology is a low-cost, non-invasive method that can be used for cancer screening and the computer-assisted analysis can help to reduce the subjectivity of the human analysis of this process. This work aims to identify the best approach to this automatization using Convolutional Neural Networks. To do that, the state-of-the-art of this application was identified and a dataset comprising 8 patients and 6,753 images was created to train the networks. After that, different models of segmentation (U-Net e HRNet), detection (Faster R-CNN e Mask R-CNN), and image classification (HRNet e ResNet) were tested and compared using multiclass and binary approaches to localize and classify cell nuclei in Papanicolaou stained oral cavity conventional cytology images. The experiments showed that the object detection approach with Mask R-CNN with ResNet 101 backbone was the best to detect abnormal nuclei in the dataset (0.74  $F_1$  Score). Additionally, the Faster R-CNN detection model with HRNet 48 showed the best results for healthy nuclei detection (0.78  $F_1$  Score). The results reinforce that a computer-assisted analysis of cytology samples can help to spread the use of this method as a screening exam.

**Keywords:** Cytology; Deep Learning; Detection; Segmentation; Computer Vision.

## LISTA DE FIGURAS

|   |    |
|---|----|
| Figura 1 – Exemplo de imagem de citologia oral usando Papanicolau. . . . .  | 13 |
| Figura 2 – Apresentação do processo de análise de citologias de pacientes de cidades pequenas como feito hoje e como se propõe ser realizado através da digitalização de lâminas e análise assistida por computador. . . . .  | 14 |
| Figura 3 – Visão geral da metodologia. . . . .  | 23 |
| Figura 4 – Exemplos de núcleos para cada classe presente no <i>dataset</i> . . . . .  | 25 |
| Figura 5 – Processamento dos dados para cada abordagem. . . . .   | 26 |
| Figura 6 – Distribuição de classes na quarta versão do UFSC-OCPap. . . . .  | 26 |
| Figura 7 – <i>Ground Truth</i> (GT) e predições dos melhores modelos de cada abordagem para núcleos ( <i>Nucleus</i> ), núcleos saudáveis ( <i>Healthy</i> ), núcleos anormais ( <i>Abnormal</i> ), núcleos fora de foco ( <i>Out of Focus</i> ) e núcleos sanguíneos ( <i>Blood</i> ). . . . . | 33 |
| Figura 8 – Resultados em nível de pixel de modelos binários de segmentação semântica. Os maiores valores de cada métrica estão destacados em negrito. . . . .   | 34 |
| Figura 9 – Resultados em nível de instância de modelos binários de detecção de objetos e da combinação de segmentação semântica e Componentes Conexos (CC). Os maiores valores de cada métrica estão destacados em negrito. . . . .   | 35 |
| Figura 10 – Resultados dos modelos de classificação de imagem. Os maiores valores de cada métrica estão destacados em negrito. . . . .  | 36 |
| Figura 11 – Resultados em nível de instância da abordagem de detecção de objetos para a classe “Anormal”. Os maiores valores de cada métrica por classe estão destacados em negrito. . . . .  | 37 |
| Figura 12 – Resultados em nível de instância da abordagem de detecção de objetos para a classe “Saudável”. Os maiores valores de cada métrica por classe estão destacados em negrito. . . . .   | 38 |
| Figura 13 – Resultados em nível de instância da abordagem de detecção de objetos para a classe “Fora de Foco”. Os maiores valores de cada métrica por classe estão destacados em negrito. . . . .   | 39 |

## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 1 – Comparação com trabalhos relacionados em relação ao tipo de citologia e ao uso de <i>Deep Learning</i> (DL) e métodos de Segmentação, Localização e Classificação dos núcleos celulares. . . . .             | 21 |
| Tabela 2 – Resultados em nível de pixel de modelos de segmentação semântica multiclasse (%) por classe e tamanho da imagem. Os melhores resultados por métrica e tamanho de imagem estão destacados em negrito. . . . . | 36 |
| Tabela 3 – Comparação dos melhores resultados dos principais trabalhos relacionados. . . . .  | 41 |

## LISTA DE ABREVIATURAS E SIGLAS

|       |  |
|-------|--|
| AgNOR | <i>Argyrophilic Nucleolar Organizer Region</i> |
| ANN   | <i>Artificial Neural Network</i>               |
| CNN   | <i>Convolutional Neural Network</i>            |
| DL    | <i>Deep Learning</i>                           |
| FCN   | <i>Fully Connected Network</i>                 |
| FPN   | <i>Feature Pyramid Network</i>                 |
| FRCN  | <i>Fully Convolutional Regression Networks</i> |
| GT    | <i>Ground Truth</i>                            |
| IoU   | <i>Intersection over Union</i>                 |
| NMS   | <i>Non Maximum Suppression</i>                 |
| RSL   | Revisão Sistemática de Literatura              |
| SVM   | <i>Support-vector Machine</i>                  |

## SUMÁRIO

|              |   |           |
|--------------|---|-----------|
| <b>1</b>     | <b>INTRODUÇÃO</b>   | <b>12</b> |
| 1.1          | OBJETIVOS   | 15        |
| <b>1.1.1</b> | <b>Objetivos Específicos</b>  | <b>15</b> |
| <b>2</b>     | <b>REVISÃO BIBLIOGRÁFICA</b>  | <b>16</b> |
| 2.1          | REDES NEURAIS CONVOLUCIONAIS  | 16        |
| <b>2.1.1</b> | <b>Citologia de Cavidade Oral e o Corante de Papanicolau</b>                        | <b>17</b> |
| 2.2          | TRABALHOS RELACIONADOS - REDES NEURAIS CONVOLUCIONAIS NA CITOLOGIA DE CAVIDADE ORAL | 18        |
| <b>3</b>     | <b>DESENVOLVIMENTO</b>  | <b>22</b> |
| 3.1          | METODOLOGIA   | 22        |
| <b>3.1.1</b> | <b>Conjunto de Dados</b>  | <b>22</b> |
| <b>3.1.2</b> | <b>Treinamento e Avaliação dos Modelos</b>  | <b>27</b> |
| 3.1.2.1      | Modelos de Classificação de Imagens   | 27        |
| 3.1.2.2      | Modelos de Segmentação Semântica  | 28        |
| 3.1.2.3      | Modelos de Detecção de Objetos  | 29        |
| 3.1.2.4      | Combinações de Detecção de Objetos  | 29        |
| 3.1.2.5      | Métricas de Avaliação   | 30        |
| 3.1.2.6      | Aumento de Dados  | 31        |
| 3.2          | RESULTADOS  | 31        |
| <b>3.2.1</b> | <b>Localização</b>  | <b>32</b> |
| <b>3.2.2</b> | <b>Classificação</b>  | <b>35</b> |
| <b>3.2.3</b> | <b>Detecção</b>   | <b>35</b> |
| 3.3          | DISCUSSÃO   | 37        |
| <b>4</b>     | <b>CONCLUSÃO</b>  | <b>42</b> |
| 4.1          | SE EU FOSSE DESENVOLVER UM SISTEMA HOJE, O QUE EU USARIA?                           | 43        |
|              | <b>REFERÊNCIAS</b>  | <b>44</b> |
|              | <b>APÊNDICE A – PRIMEIRAS PÁGINAS DE ARTIGOS PUBLICADOS</b>                         | <b>54</b> |
|              | <b>APÊNDICE B – TABELA DETALHANDO METODOLOGIA DE TODOS OS EXPERIMENTOS</b>          | <b>59</b> |

## 1 INTRODUÇÃO

As chances de cura de um câncer são proporcionais à quão precoce é o seu diagnóstico. A detecção de anomalias em células de forma precoce pode, portanto, aumentar as chances de sucesso do tratamento (BELL *et al.*, 2006). Para os casos de câncer de cavidade oral, as chances de sobrevivência são de 80% quando diagnosticado nos primeiros estágios e 19% quando em estágios avançados (PEKTAŞ *et al.*, 2006). Dessa forma, o uso de métodos não invasivos para exames preventivos, como a citologia esfoliativa, é fortemente recomendado para o diagnóstico precoce. A indicação de exames preventivos é reforçada pelo fato de que os casos desse tipo de tumor vêm aumentando nos últimos anos, sendo responsáveis por cerca de 370 mil dos novos casos de câncer em todo mundo e causando cerca de 175 mil mortes em 2020 (SUNG *et al.*, 2021).

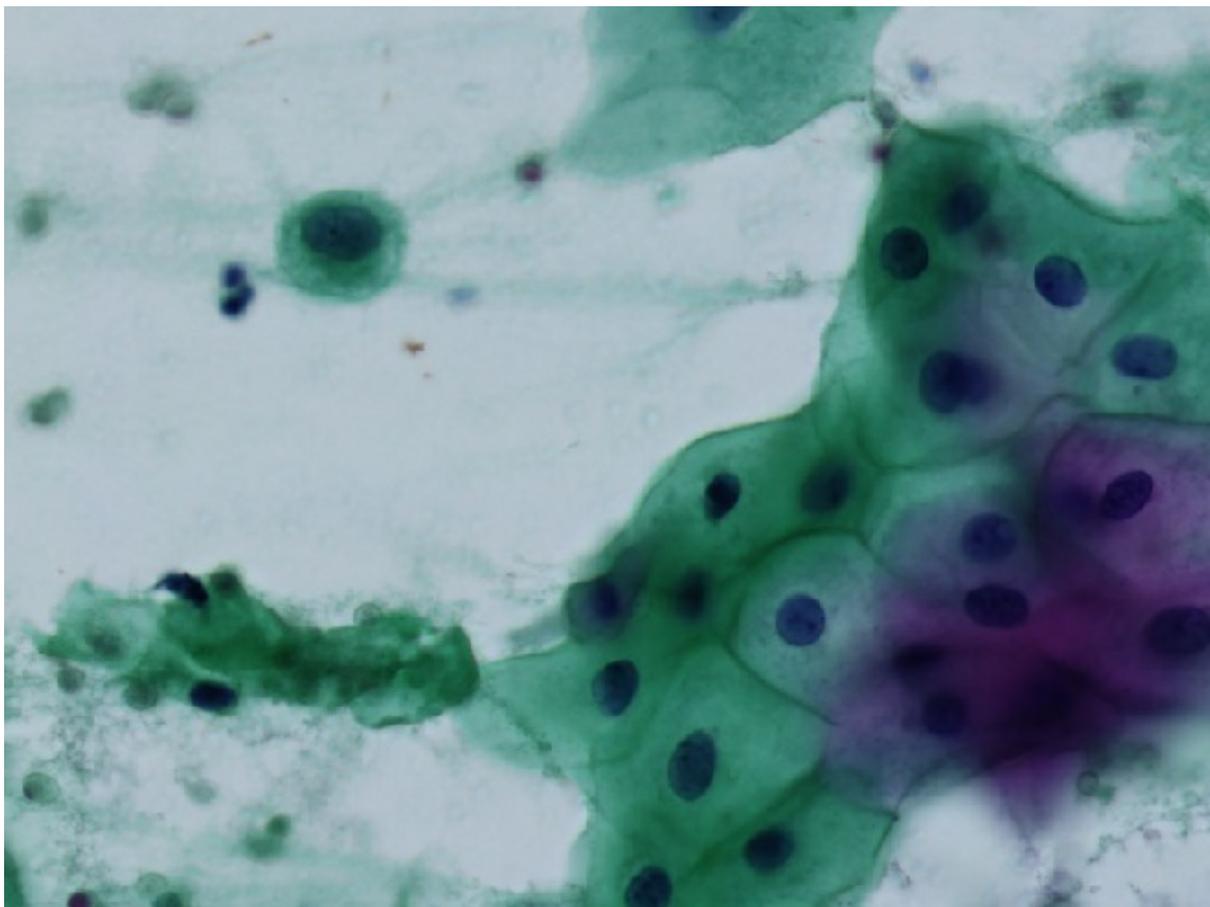
Hoje, o principal método para detecção deste tipo de tumor é a histologia feita através de biópsia. Essa técnica necessita da extração cirúrgica de uma quantidade significativa de tecido do paciente para sua realização e, por isso, é considerada muito invasiva para exames preventivos (LUCENA *et al.*, 2011). Os exames de citologia, por sua vez, são realizados através da avaliação microscópica de lâminas preparadas com pequenos conjuntos de células. Na citologia esfoliativa, as células analisadas são obtidas por meio de raspagem superficial dos tecidos, permitindo a detecção de alterações celulares através de uma coleta de amostras minimamente invasiva e de baixo custo (LUCENA *et al.*, 2011).

Para diferenciar e destacar características de interesse das células analisadas, é necessária a utilização de uma técnica de coloração da amostra. Na citologia esfoliativa uma das mais utilizadas é a coloração de Papanicolau (CHATTERJEE *et al.*, 2018; LUCENA *et al.*, 2011), que gera imagens como a mostrada na Figura 1.

As motivações para uso da coloração de Papanicolau nesse exame vêm de seu baixo custo, de sua boa difusão na área médica e dos bons resultados que possui na prevenção e diagnóstico de tumores malignos de colo de útero. Como os tecidos da mucosa do colo de útero e da mucosa oral são semelhantes, pode-se usar a mesma coloração para análises de citologias de cavidade oral. Apesar disso, essa ainda é uma técnica relativamente pouco usada na odontologia (LUCENA *et al.*, 2011). Considerando que o foco da pesquisa é sua aplicação na rotina clínica do Brasil, os experimentos foram limitados à amostras de citologia convencional, que gera imagens de menos qualidade que a citologia em base líquida, mas é a técnica mais viável no sistema de saúde público (LUCENA *et al.*, 2011).

Uma das principais desvantagens do exame de citologia esfoliativa é a sua sensibilidade à experiência do profissional responsável pela análise das lâminas e seu alto grau de resultados não conclusivos (LUCENA *et al.*, 2011). A análise computacional

Figura 1 – Exemplo de imagem de citologia oral usando Papanicolau.



Fonte – Elaborada pelo autor.

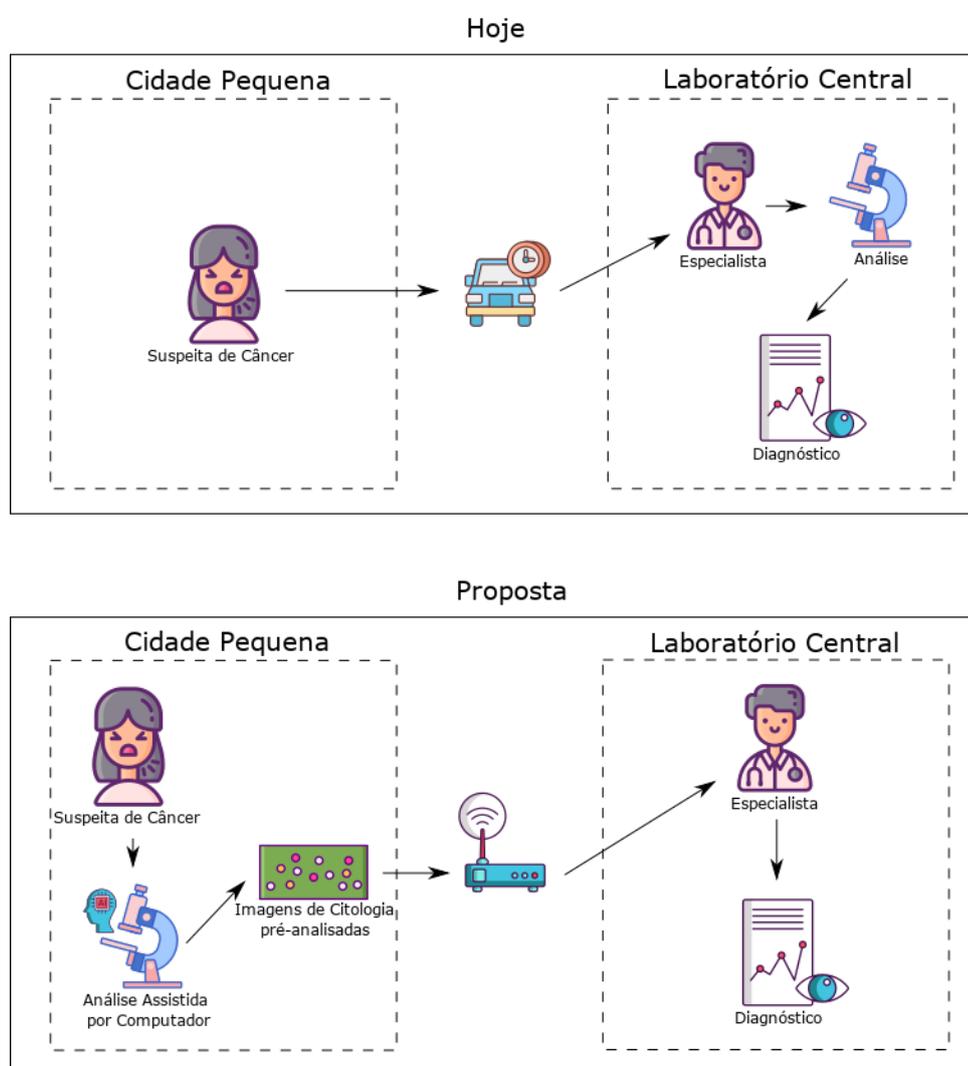
tem a capacidade de diminuir a subjetividade da análise manual e proporcionar resultados mais precisos e menos variáveis (LI *et al.*, 2017). Por sua vez, Redes Neurais Convolucionais (do inglês, *Convolutional Neural Network* (CNN)) tem se mostrado uma das melhores técnicas de extração de informações a partir de imagens, particularmente em imagens médicas (KHAN *et al.*, 2019).

CNNs são modelos de redes neurais que buscam reconhecer padrões, principalmente em imagens, e têm sua estrutura baseada nos processos da visão humana. Os resultados apresentados por estes modelos são o estado da arte em muitos trabalhos envolvendo reconhecimento de imagens, segmentação, detecção de características e outras tarefas relacionadas (LIU *et al.*, 2019). Isso indica que CNNs têm potencial para detectar anomalias que podem passar despercebidas por uma análise humana qualitativa, sendo assim um método candidato para aprimoramento do diagnóstico precoce de câncer através da análise de citologias.

Além disso, hoje, como apresentado na Figura 2, quando um paciente apresenta suspeita de câncer oral no sistema público de saúde brasileiro, ele precisa ir até um laboratório central para realizar a coleta da amostra. Essa amostra, posteriormente, é

diagnosticada por um especialista que analisa a lâmina em um microscópio óptico. Os laboratórios centrais geralmente estão localizados apenas em cidades grandes, o que faz com que pacientes de cidades pequenas precisem percorrer grandes distâncias e demorem mais para receber um resultado. A análise de lâminas através de imagens de microscópio digitalizadas pode simplificar e facilitar esse processo, já que estas podem ser enviadas através da rede para o especialista. Nesse contexto, a análise assistida por computador pode ser usada para auxiliar o patologista no diagnóstico e, assim, incentivar a adoção do modelo de análise através de citologias digitalizadas.

Figura 2 – Apresentação do processo de análise de citologias de pacientes de cidades pequenas como feito hoje e como se propõe ser realizado através da digitalização de lâminas e análise assistida por computador.



Fonte – Elaborada pelo autor.

Na pesquisa apresentada neste trabalho, a análise de imagens deste tipo de exame tem por objetivo a detecção precoce de células que possuam alterações que caracterizem fases iniciais de um câncer de cavidade oral. Além disso, é pertinente

para a ciência da computação demonstrar o desempenho desse tipo de rede neural em diferentes problemas. Elencar e testar múltiplos modelos de CNNs em diferentes domínios de problemas é importante para mostrar pontos fortes e fracos dessa técnica e para quais tipos de dados ela é mais adequada.

## 1.1 OBJETIVOS

O principal objetivo deste projeto é avaliar a viabilidade da utilização de CNNs em imagens de exames de citologia esfoliativa convencional com a coloração de Papanicolau para detecção precoce de risco de câncer de cavidade oral.

### 1.1.1 Objetivos Específicos

Os objetivos específicos são os seguintes:

- Levantar o estado da arte em análise de imagens de exames de citologia;
- Elencar técnicas computacionais já usadas para identificar células alteradas nesse tipo de exame;
- Gerar um conjunto de dados de imagens de citologia de cavidade oral coradas com o corante de Papanicolau, usando a técnica convencional, marcadas por especialistas para treinamento e avaliação das CNNs;
- Enumerar diferentes abordagens e estruturas de CNNs para localização e classificação de núcleos celulares em exames de citologia;
- Avaliar o desempenho de CNNs na detecção de núcleos em imagens de exames de citologia esfoliativa de cavidade oral;
- Propor uma abordagem e arquitetura de CNN para a detecção de núcleos celulares que indiquem, de forma precoce, o desenvolvimento de um câncer de cavidade oral;
- Publicar os resultados obtidos através desta pesquisa.

## 2 REVISÃO BIBLIOGRÁFICA

Neste capítulo são apresentados os temas relacionados ao trabalho através de uma revisão de literatura.

### 2.1 REDES NEURAIS CONVOLUCIONAIS

Redes Neurais Artificiais (do inglês, *Artificial Neural Network (ANN)*) são dispositivos da área de aprendizado de máquina que buscam reconhecer padrões através de mecanismos baseados no funcionamento dos neurônios biológicos. A arquitetura desse tipo de rede é composta por uma camada de neurônios de entrada, que recebem os dados a serem processados; uma ou mais camadas ocultas, que geram abstrações e reconhecem padrões nos dados; e uma camada de saída, que devolve os resultados processados pela rede. O processo de aprendizagem supervisionado desse tipo de modelo é realizado através da exposição do modelo à uma quantidade suficientemente grande de dados marcados, de forma que o modelo possa aprender padrões nos dados de entrada e generalizar o aprendizado para dados ainda não vistos (JAIN *et al.*, 1996). No contexto deste trabalho, os dados usados para treinamento são campos de lâminas de citologia marcados por especialistas e é necessária uma grande quantidade desses dados para que as redes possam reconhecer os padrões das células de forma a localizar e classificar regiões de interesse nessas imagens.

Nos últimos anos, os métodos de aprendizagem profunda (do inglês, DL) permitiram a construção de redes maiores e mais complexas, revolucionando a pesquisa com ANNs (KRIZHEVSKY *et al.*, 2012). Através das CNNs, esses avanços foram percebidos principalmente no processamento de imagens, onde mostraram robustez e eficácia na solução de problemas de diferentes domínios (LECUN *et al.*, 2015). As CNNs são capazes de aprender através de operações de convolução que representam múltiplos níveis de abstração das características de uma imagem, partindo de informações dos pixels da imagem e gerando uma representação semântica do seu conteúdo em um único passo (LECUN *et al.*, 2015). Além disso, trabalhos recentes mostraram que modelos de CNN são capazes de solucionar problemas complexos nas áreas da genética e da biomedicina (LI *et al.*, 2017).

A principal desvantagem desse dispositivo é que, assim como qualquer ANN, ele é um processo caixa preta: as operações de convolução são aprendidas e codificadas como dados na estrutura da rede e na maioria dos casos não é possível interpretar o que exatamente foi aprendido pela rede, o que pode ser um impasse para a adoção destes modelos na área médica (HAYASHI, 2020). Apesar disso, já existem esforços para melhorar a interpretabilidade e explanabilidade de ANNs (HAYASHI, 2020) e contornar esse problema.

### 2.1.1 Citologia de Cavidade Oral e o Corante de Papanicolau

O exame de citologia é uma ferramenta diagnóstica que analisa células retiradas de tecidos através de aspiração ou esfoliação, preparadas com uma técnica de coloração, em um microscópio. Desde que foi descrita pela primeira vez pelo Dr. Alfred François Donné em 1837, quando descobriu o *Trichomonas vaginalis*, essa técnica vem sendo usada como suporte ao diagnóstico de uma grande variedade de patologias. A citologia se tornou popular em 1941, quando o Dr. Georgios Papanikolaou provou que esse exame era capaz de auxiliar no diagnóstico de células cancerígenas em esfregaços de colo de útero (DIAMANTIS; MAGIORKINIS, 2013).

Hoje, a citologia é uma ferramenta importante no diagnóstico dos principais cânceres do corpo humano, incluindo os de cavidade oral (IVANOVIC, 2013). Quando comparado a outros métodos de diagnóstico que exigem a extração de tecidos, como a biópsia, a citologia apresenta as vantagens de ser um exame mais simples e barato, além de ser menos invasivo para o paciente e mais fácil de ser realizado pelo profissional (BEDROSSIAN, 2007).

Em relação à preparação das amostras em citologias esfoliativas, é importante destacar a existência de dois principais métodos: a citologia convencional e a citologia em base líquida. Na técnica convencional as células são fixadas diretamente sobre a lâmina, enquanto na citologia em base líquida a amostra é diluída, centrifugada e filtrada antes de ser fixada sobre a lâmina. No contexto desta pesquisa, a citologia em base líquida gera imagens mais simples de serem computacionalmente analisadas, porém essa preparação possui um alto custo de equipamentos e manutenção, o tempo de preparação é maior e existe a necessidade de transporte e disposição do meio líquido (LUCENA *et al.*, 2011). Por esses motivos, e considerando que uma das motivações desta pesquisa é ampliar o uso de citologias como exames preventivos, não foram consideradas amostras preparadas em base líquida para o treinamento das redes.

No exame de citologia, a condição das células é analisada através de diferentes abordagens, como a quantificação de DNA, identificação de proteínas chave, presença de patógenos e, mais comumente, a morfologia celular. Cada abordagem necessita de uma técnica de coloração e processamento diferente, para destacar organelas específicas ou outras partes da célula. Independente disso, a análise final geralmente é feita manualmente utilizando um microscópio de luz (IVANOVIC, 2013). Apesar de este procedimento ser realizado por um médico especializado e treinado para esta tarefa, muitos dos critérios diagnósticos são vulneráveis a interpretações humanas e vieses (CHAPMAN; OTIS, 2011).

Uma das principais técnicas de coloração usadas nos exames de citologia é a técnica de Papanicolau. Esta é uma técnica de baixo custo e que geralmente é aplicada para detecção de câncer cervical, sendo composta por cinco corantes: Hematoxilina,

Alaranjado G, Pardo Bismark Y, Verde luz amarelado e Eosina Y. Estes corantes são usados para definir detalhes das células, destacar seus limites e para diferenciar entre diferentes tipos de células. Apesar de ser largamente usada, a técnica de Papanicolaou não é padronizada, o que causa grandes variações nos seus resultados (GILL, 2013).

## 2.2 TRABALHOS RELACIONADOS - REDES NEURAIAS CONVOLUCIONAIS NA CITOLOGIA DE CAVIDADE ORAL

Para identificar o estado da arte na aplicação de CNNs na citologia de cavidade oral, foram realizadas duas Revisão Sistemática de Literaturas (RSLs): a primeira (SERGIO *et al.*, 2019) foi publicada em dezembro de 2019 e abordou a análise computacional de exames de citologia oral, enquanto a segunda (VICTÓRIA MATIAS, André *et al.*, 2021) foi publicada em julho de 2021 e focou em métodos computacionais para detectar anomalias em todos os tipos de exames de citologia. O objetivo do primeiro trabalho, intitulado “Computer-assisted technologies for diagnosis of oral cancer on cytology samples – A Systematic Literature Review” (SERGIO *et al.*, 2019), foi identificar quais métodos de visão computacional já foram testados no diagnóstico de câncer de cavidade oral usando amostras de citologia. Foram analisados 404 artigos, sem restrição de data de publicação, e 6 artigos foram selecionados a partir dos critérios de inclusão e exclusão. Já a segunda revisão, intitulada “What is the state of the art of computer vision-assisted cytology? A Systematic Literature Review” (VICTÓRIA MATIAS, André *et al.*, 2021), buscou identificar qual o estado da arte dos métodos computacionais usados para suportar a análise de imagens de citologia. Foram analisados 431 artigos, publicados entre janeiro de 2016 e setembro de 2020, dos quais 157 foram selecionados a partir dos critérios de inclusão e exclusão. Além disso, para identificar trabalhos realizados após essa última RSL, foi realizada uma outra busca, usando os mesmos critérios desta, por trabalhos publicados entre setembro de 2020 e fevereiro de 2022. Nessa última pesquisa, foi encontrado mais um trabalho relevante (BERNANDER *et al.*, 2021) que usa métodos computacionais na análise de citologias orais. Todos os trabalhos encontrados nas RSLs que usam visão computacional na análise de imagens de citologia oral serão detalhados nos próximos parágrafos em ordem de publicação.

Foram encontrados 2 trabalhos que não utilizam o corante de Papanicolaou. Em Bell *et al.* (2006), as Regiões Argirofílicas Organizadoras Nucleolares (do inglês *Argyrophilic Nucleolar Organizer Region (AgNOR)*) são identificados a partir de uma imagem de Alta Faixa Dinâmica. O conjunto de imagens usado no trabalho inclui 40 amostras de mucosa oral, além de 145 amostras de outras partes do corpo. Ao comparar a contagem manual de AgNORs com a segmentação automática em imagens de alta e baixa faixas dinâmicas, o método proposto foi o que mais se aproximou da contagem realizada por especialistas. Já os autores de Pektaş *et al.* (2006) avaliaram

a eficácia da análise citomorfométrica e do estado da ploidia de ADN para a detecção de alterações malignas em células orais usando o corante de Feulgen. Esse estudo usa esfregaços de 22 mucosas orais saudáveis e 22 lesões orais. As características citomorfométricas foram extraídas e analisadas usando o *software* de análise de imagens Zeiss Vision KS 400. O trabalho conclui que a análise citomorfométrica é útil na identificação de lesões pré-malignas e lesões cancerígenas em estágios iniciais, com bons resultados de especificidade e sensibilidade.

No trabalho de Banerjee *et al.* (2016), é apresentado um classificador para imagens de amostras de citologia exfoliativa coradas com o corante de Papanicolau. Foram usados como características de interesse a morfologia e a textura das células para treinar uma *Support-vector Machine* (SVM) linear para classificação entre carcinoma e leucoplasia. O *dataset* usado é composto por 198 amostras de leucoplasia e 140 amostras de carcinoma. O modelo proposto alcançou 89,9% de sensibilidade e 100% de especificidade no subconjunto de validação usando validação cruzada (k-fold).

O trabalho de Dey *et al.* (2017) realiza a análise em imagens de microscopia de contraste de interferência diferencial com coloração de Papanicolau, extraindo características morfológicas e texturais das células. Essas informações são então passadas para um classificador SVM que deve indicar se o paciente está ou não em fase de pré-câncer. No conjunto de teste, com 14 indivíduos, o classificador obteve sensibilidade de 80%, especificidade de 88,89% e acurácia de 85,71%.

A pesquisa realizada em Lu *et al.* (2020) é a que mais se aproxima da proposta deste projeto. Os autores treinaram múltiplos modelos para a localização e classificação de núcleos corados com Papanicolau. Para isso, primeiramente, um modelo *Fully Convolutional Regression Networks* (FRCN) foi treinado em imagens de amostras em base líquida para localização de núcleos. Esse modelo obteve, no subconjunto de validação (119 núcleos), 95,8% de acurácia, 92% de  $F_1$  Score, 90% de precisão e 94% de sensibilidade em nível de instância. Numa segunda etapa, os autores treinaram um modelo U-Net para segmentação dos núcleos e, posteriormente, treinar os modelos ResNet-50 e DenseNet-201 para classificar esses núcleos entre normal e anormal. Os melhores resultados para classificação de núcleos anormais foram obtidos usando o modelo DenseNet-201 sem pré-treino, que obteve, em nível de instância, uma acurácia de 75,4%, um  $F_1$  Score de 61,4%, precisão de 84,3% e 48,3% de sensibilidade. Esses valores foram obtidos a partir de testes realizados com 10 lâminas com um total de 68.509 núcleos. Para os treinos e testes, os autores consideraram que todas os núcleos de lâminas diagnosticadas como casos de câncer são anormais e todos os núcleos de lâminas saudáveis são saudáveis.

Na pesquisa de Bernander *et al.* (2021), o mais recente trabalho relacionado encontrado, os autores avaliaram o impacto da equivariância de rotação na classificação de núcleos em imagens de citologia. Para isso, os autores desenvolveram uma

variação do modelo VGG16 chamada GCNN. Esse novo modelo rotaciona a imagem de entrada entre as camadas da CNN usando múltiplos de 90 graus para tornar a arquitetura equivariante em relação à rotação da entrada. O método foi treinado e testado em um *dataset* de amostras preparadas em base líquida contendo 14.196 imagens de núcleos isolados de 12 pacientes (6 saudáveis e 6 com diagnóstico de câncer). Assim como no trabalho anterior, os autores marcaram todas os núcleos dos casos de câncer como anormais e todos os núcleos dos casos saudáveis como saudáveis. Os resultados mostram que o modelo GCNN teve um melhor desempenho que o modelo VGG16 padrão em termos de especificidade e sensibilidade, além de reduzir *overfitting* e a quantidade de dados necessária para alcançar a maior acurácia obtida.

Através das RSLs (VICTÓRIA MATIAS, André *et al.*, 2021; SERGIO *et al.*, 2019), pode-se perceber uma tendência de crescimento de pesquisas envolvendo DL para a automatização da análise de citologias, apesar de métodos clássicos de visão computacional ainda serem amplamente usados. Foram encontrados apenas dois *datasets* públicos de imagens de citologias coradas com Papanicolau usando a técnica tradicional, sendo apenas um deles contendo imagens de campos de lâminas com múltiplas células e nenhum deles envolvendo exames de cavidade oral. Além disso, foi destacado o fato de que os trabalhos analisados estão em fase experimental, sem testes na rotina clínica de profissionais patologistas. Considerando as aplicações de CNNs em imagens médicas, estas são mais comuns em exames de histologia (KHAN *et al.*, 2019). As medidas de desempenho e a definição dos modelos de CNN mais apropriados são específicos de cada domínio de aplicação e, portanto, é importante que o uso destas seja explorado em diferentes casos, sendo a citologia esfoliativa com coloração Papanicolau um domínio de imagens ainda pouco explorado.

De forma geral, apesar de a técnica de Papanicolau ser a mais comum dentre os trabalhos analisados nas RSLs, é evidente que a análise de imagens de citologias orais não é um problema bem explorado na pesquisa de visão computacional: foram encontrados apenas seis trabalhos que executaram experimentos com este tipo de imagem. Com exceção de Lu *et al.* (2020), que usa amostras de citologia convencional apenas para testar a etapa de classificação de imagens do modelo proposto, todos os trabalhos relacionados citados usam apenas amostras preparadas em meio líquido. Apesar de esta técnica de preparação gerar imagens com menos células sobrepostas e menos artefatos, tornando sua análise por técnicas de visão computacional menos complexa (ARAÚJO *et al.*, 2019), ela possui um custo mais alto quando comparada à citologia convencional, tornando seu uso inviável para sistemas de saúde públicos de países em desenvolvimento (LUCENA *et al.*, 2011). Além disso, este trabalho compara múltiplas abordagens de DL, o que não é abordado por nenhum dos trabalhos relacionados, e é o primeiro a tornar público um *dataset* de imagens de citologias de cavidade oral, que também é o primeiro conjunto de imagens de lâminas de Papanicolau

convencional marcadas a ser tornado público.

A Tabela 1 mostra uma comparação da metodologia deste trabalho com a metodologia dos trabalhos relacionados ordenados por data de publicação. Pode-se notar uma maior adoção de técnicas de DL nos últimos trabalhos publicados, apesar de a maioria ainda usar técnicas clássicas de visão computacional. Percebe-se também que, em relação aos trabalhos selecionados, este trabalho é o único a aplicar DL em imagens de Papanicolau convencional e também o único a comparar técnicas de segmentação, localização e classificação de núcleos celulares.

Tabela 1 – Comparação com trabalhos relacionados em relação ao tipo de citologia e ao uso de DL e métodos de Segmentação, Localização e Classificação dos núcleos celulares.

| Trabalho                       | Corante            | Preparação          | DL         | Seg.       | Loc.       | Clas.      |
|--------------------------------|--------------------|---------------------|------------|------------|------------|------------|
| Bell <i>et al.</i> (2006)      | AgNOR              | Convencional        | Não        | Sim        | Não        | Não        |
| Pektaş <i>et al.</i> (2006)    | Feulgen            | Convencional        | Não        | Sim        | Não        | Não        |
| Banerjee <i>et al.</i> (2016)  | Papanicolau        | Em base líquida     | Não        | Não        | Não        | Sim        |
| Dey <i>et al.</i> (2017)       | Papanicolau        | Convencional        | Não        | Sim        | Não        | Sim        |
| Lu <i>et al.</i> (2020)        | Papanicolau        | Em base líquida     | Sim        | Não        | Sim        | Sim        |
| Bernander <i>et al.</i> (2021) | Papanicolau        | Em base líquida     | Sim        | Não        | Não        | Sim        |
| <b>Este</b>                    | <b>Papanicolau</b> | <b>Convencional</b> | <b>Sim</b> | <b>Sim</b> | <b>Sim</b> | <b>Sim</b> |

Fonte – Elaborada pelo autor.

### 3 DESENVOLVIMENTO

Neste capítulo serão apresentados a metodologia e os resultados do projeto de pesquisa.

#### 3.1 METODOLOGIA

Para avaliar possíveis métodos de análise de lâminas coradas com Papanicolau, foram comparadas três abordagens usando CNNs para localizar e classificar núcleos celulares: segmentação semântica, detecção de objetos e classificação de imagens. A segmentação semântica é baseada em atribuir uma classe a cada pixel da imagem, tornando possível quantificar e analisar a morfologia dos núcleos. A detecção de objetos é menos precisa que a segmentação semântica, atribuindo a cada objeto de interesse uma caixa de delimitação (*bounding box*) e uma classe, tornando possível a localização e classificação dos núcleos, mas não a extração de características morfológicas por não identificar o contorno dos núcleos. A classificação de imagens busca apenas atribuir uma classe à imagem como um todo, sem identificar regiões de interesse.

O levantamento do estado da arte em relação à análise de exames de citologia esfoliativa para câncer de cavidade oral foi realizado através da RSL realizada pelo grupo de pesquisa e publicada como relatório técnico (SERGIO *et al.*, 2019) e de uma segunda RSL focada em na análise computacional de exames de citologia e publicada como artigo (VICTÓRIA MATIAS, André *et al.*, 2021). Ambas estão detalhadas na Seção 2.2.

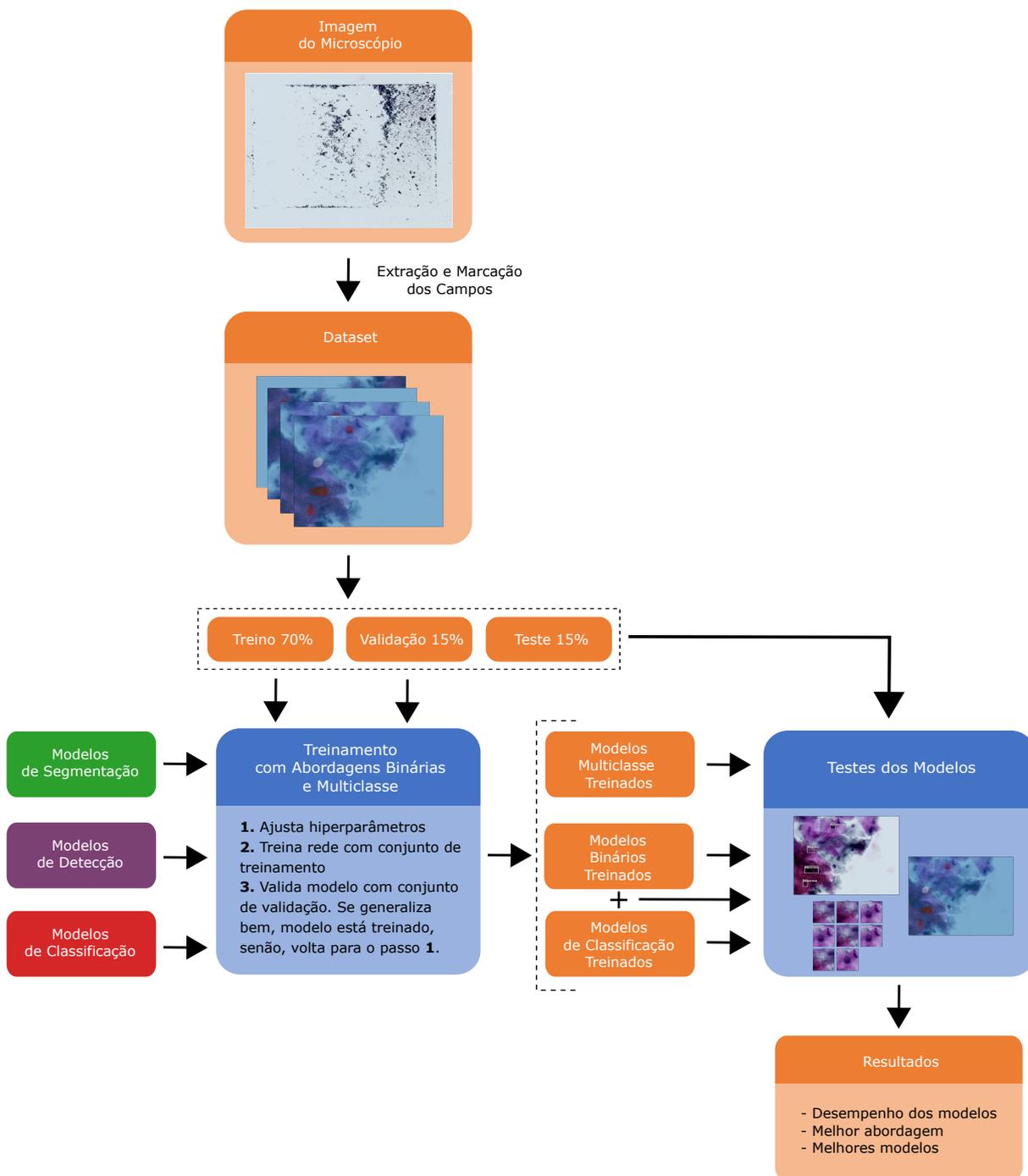
Como não foram identificados *datasets* de citologia oral públicos, um conjunto de dados foi construído para realizar avaliação de diferentes modelos de CNNs, identificando vantagens e desvantagens de cada um, para identificar os mais adequados na análise de imagens de exames de citologia esfoliativa de cavidade oral.

Depois disso, foram implementados experimentos com múltiplos modelos de CNNs e combinações de segmentação, detecção e classificação, avaliando-se o desempenho de cada abordagem para o problema da pesquisa. Na Figura 3 é mostrada uma visão geral dos métodos usados neste trabalho.

##### 3.1.1 Conjunto de Dados

Para treinar e validar os métodos propostos, foi elaborado o *dataset* UFSC-OCPap v4 (VICTÓRIA MATIAS, André, 2022). Esse conjunto de imagens é composto por 6.753 campos de 1200x1600 pixels obtidos de 8 lâminas de citologia convencional de cavidade oral de 8 diferentes pacientes coradas com o corante de Papanicolau. Destas amostras, 5 foram previamente diagnosticadas como casos de câncer e 3 diagnosticadas como casos saudáveis. As imagens das lâminas foram obtidas usando

Figura 3 – Visão geral da metodologia.



Fonte – Elaborada pelo autor.

um microscópio Axio Scan.Z1 e uma câmera Hitachi HV-F202SCL no Laboratório Multiusuário de Estudos em Biologia (LAMEB/UFSC). As lâminas foram fornecidas pelo Centro de Odontologia do Hospital Universitário Professor Polydoro Ernani de São Thiago da Universidade Federal de Santa Catarina (HU/UFSC)<sup>1</sup>. De cada lâmina foi originalmente obtida uma imagem de 214.000x161.000 pixels (0,111 $\mu$ m x 0,111 $\mu$ m por pixel) que foi dividida para gerar os campos que compõem o *dataset*.

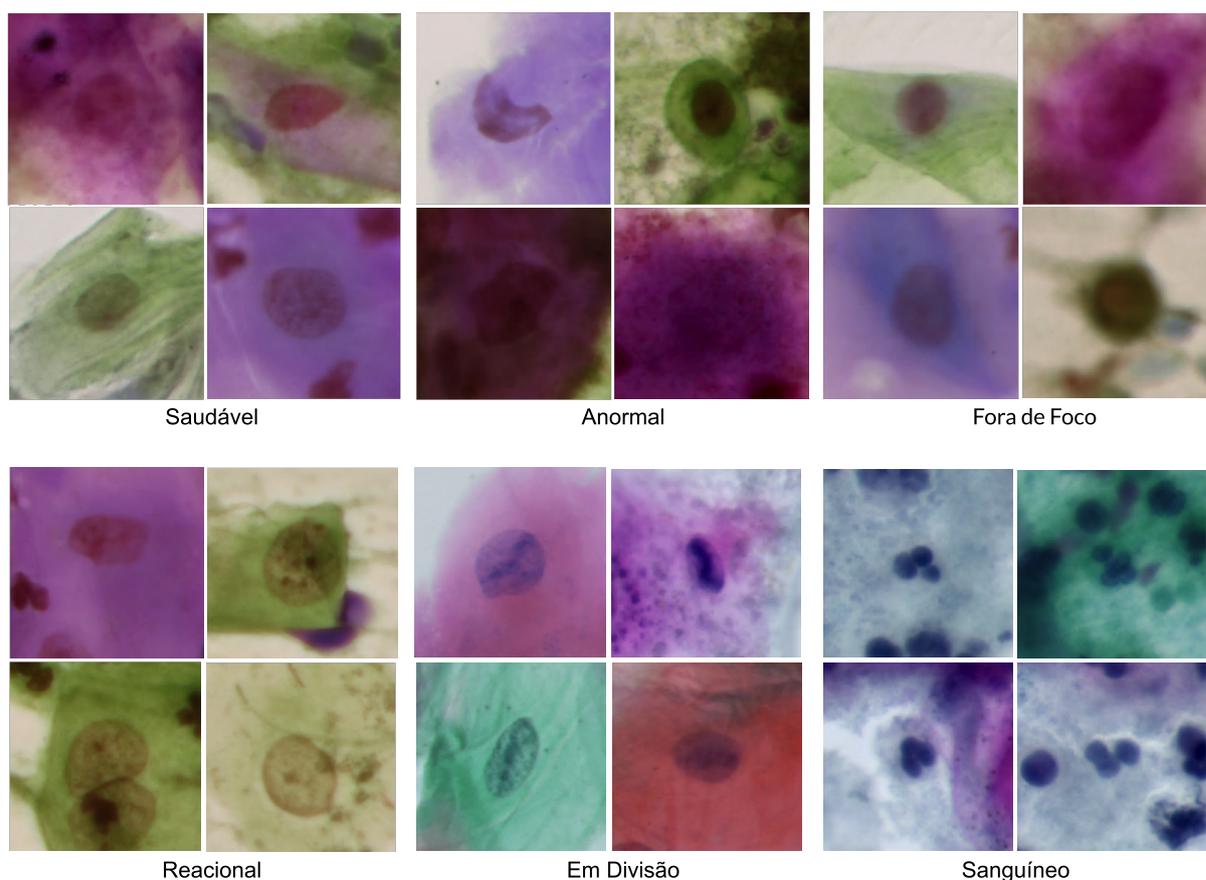
Um grupo de especialistas marcou estas imagens e, para garantir a qualidade das marcações, um terceiro especialista com experiência na análise de lâminas de citologia de cavidade oral revisou cerca de 35% das marcações usando as ferramentas Labelbox (LABELBOX, 2021) e LabelMe (RUSSELL *et al.*, 2008). O processo de marcação e revisão gerou uma máscara para cada imagem com os pixels classificados nas classes “background”, “Núcleo Anormal”, “Núcleo Saudável”, “Núcleo em Divisão”, “Núcleo Fora de Foco”, “Núcleo de Célula Sanguínea” e “Núcleo Reacional”. Essas classes são definidas da seguinte forma:

- **background**: Pixels que não pertencem a nenhum núcleo;
- **Núcleo Saudável**: Núcleos epiteliais que não apresentam anormalidades;
- **Núcleo Anormal**: Núcleos epiteliais que apresentam alguma anormalidade que possa indicar o desenvolvimento de um carcinoma;
- **Núcleo em Divisão**: Núcleos epiteliais que estão em processo de divisão celular saudável;
- **Núcleo Reacional**: Núcleos epiteliais que apresentam anomalias causadas por processos não ligados à carcinomas;
- **Núcleo de Célula Sanguínea**: Núcleos que pertencem a células sanguíneas;
- **Núcleo Fora de Foco**: Núcleos com qualidade de imagem insuficiente para definir a sua classe.

A Figura 4 mostra exemplos de núcleos de cada classe presentes nas imagens que compõem o *dataset*.

O conjunto de imagens foi dividido em três subconjuntos: ~70% para treinamento (4.745 imagens), ~15% (976 imagens) para validação e ~15% (1.032 imagens) para teste. O processo de geração de dados para cada modelo a partir das máscaras está ilustrado na Figura 5: A partir de um campo da lâmina, um dos especialistas marca manualmente os núcleos celulares indicando suas classes. Essas marcações

<sup>1</sup> Esta pesquisa foi aprovada pelo Comitê de Ética em Pesquisa da UFSC (CEPSH), protocolo 23193719.5.0000.0121. Todos os pacientes foram apresentados sobre os objetivos do estudo e, aqueles que concordaram em participar, assinaram um termo de consentimento livre e esclarecido.

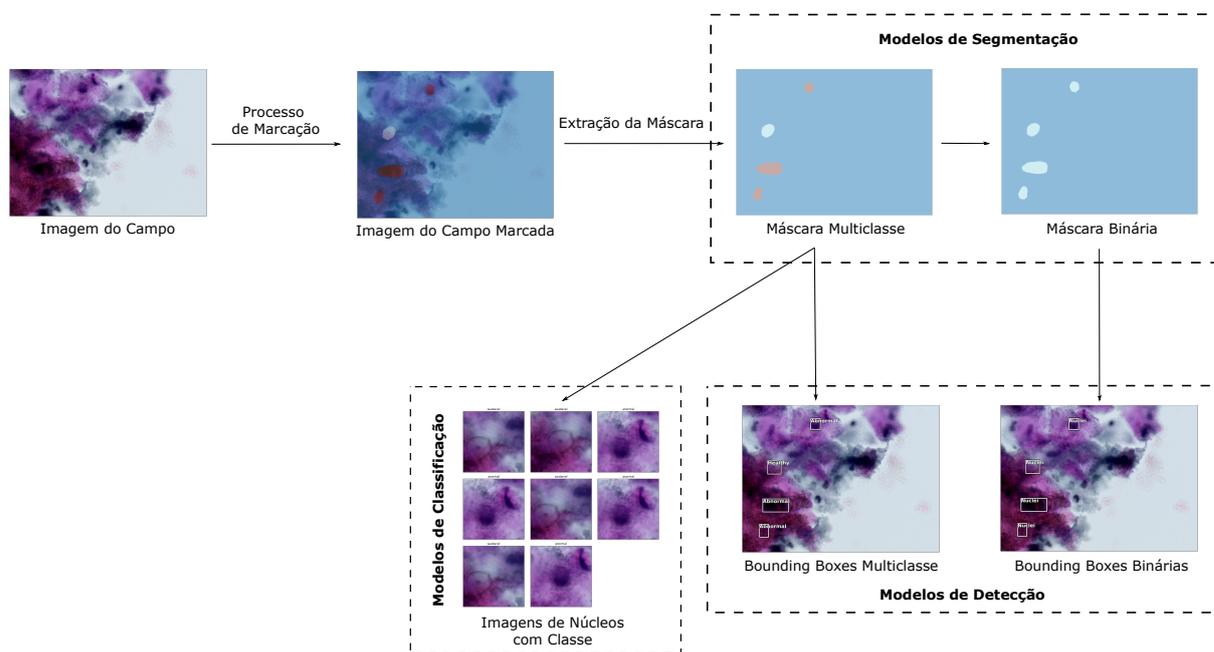
Figura 4 – Exemplos de núcleos para cada classe presente no *dataset*.

Fonte – Elaborada pelo autor.

são então usadas para gerar máscaras multiclasse, usadas para treinar os modelos multiclasse, e binárias (classes “Núcleo” e “background”) que são então usadas nos modelos binários de segmentação e, também, para gerar as *bounding boxes* para cada núcleo usadas nos modelos de detecção de objetos binários.

Foram publicadas 4 versões do UFSC-OCPap. Na primeira versão, ele possui 1.934 campos de 2 lâminas de diferentes pacientes diagnosticados como casos de câncer. Na segunda versão, todas as marcações da primeira versão foram revisadas por um patologista com experiência na área e foi adicionada mais uma lâmina de um caso de câncer, totalizando 7.449 imagens marcadas. Na terceira versão do *dataset*, foram marcados campos de mais 5 lâminas de diferentes pacientes, sendo 3 deles de casos saudáveis para equilibrar o conjunto de dados. Na quarta versão, buscou-se aumentar a qualidade do *dataset* através de revisão das imagens e equilíbrio das classes. A distribuição de classes da última versão do *dataset* (v4) pode ser vista na Figura 6. Essa última versão está publicamente disponível em <https://data.mendeley.com/datasets/dr7ydy9xbk/2>.

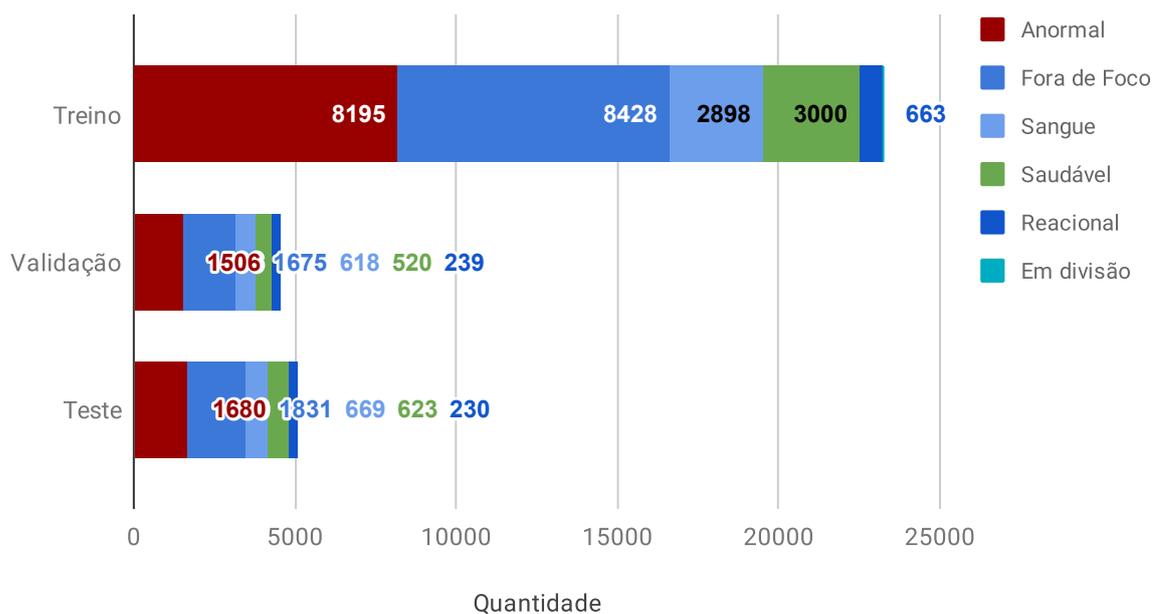
Figura 5 – Processamento dos dados para cada abordagem.



Fonte – Elaborada pelo autor.

Figura 6 – Distribuição de classes na quarta versão do UFSC-OCPap.

Distribuição de Instâncias



Fonte – Elaborada pelo autor.

### 3.1.2 Treinamento e Avaliação dos Modelos

Diferentes modelos foram avaliados para cada abordagem, considerando o que foi encontrado no levantamento do estado da arte e nos resultados parciais, detalhados na Seção 3.2, publicados em A. Victória Matias *et al.* (2020) e Matias *et al.* (2020). Para otimizar o desempenho das redes, diferentes técnicas de *data augmentation* e *transfer learning* foram testadas e serão detalhadas nas próximas seções. Os hiperparâmetros usados foram escolhidos baseando-se nos resultados obtidos por cada rede no subconjunto de validação. As particularidades metodológicas de cada experimento estão detalhadas no Apêndice B.

Em relação ao uso dos subconjuntos, apenas o de treinamento foi usado para treinamento da rede, o conjunto de validação foi usado para ajustes de hiperparâmetros e o conjunto de teste foi reservado apenas para avaliação das métricas após a conclusão de todas as iterações de treinamento (RIPLEY, 2007).

Os modelos e estratégias avaliados neste trabalho serão detalhados nas próximas subseções.

#### 3.1.2.1 Modelos de Classificação de Imagens

Para a abordagem de classificação de imagens, foi usada a biblioteca Fastai (FASTAI, 2019), foram selecionados os modelos ResNet 18, ResNet 34, ResNet 50, HRNet 18 e HRNet 48 e foram selecionados a técnica ADAM (KINGMA; BA, 2014) como otimizador e a entropia cruzada como função de perda.

As *Residual Convolutional Neural Networks* (ou ResNets) são modelos estado da arte para classificação de imagens que possuem uma arquitetura interna composta por uma sequência de Blocos Residuais (do inglês *Residual Blocks*). Esta característica possibilita que as ResNets contornem o problema dos gradientes evanescentes, treinando mais rápido e com melhores resultados. Isto torna o modelo ResNet uma boa alternativa quando se busca explorar novas aplicações de classificação de imagens. Já o modelo *High-Resolution Neural Network* (ou HRNet) (WANG, J. *et al.*, 2020a) é um modelo de CNN que usa uma abordagem de aprendizado de multi-resoluções simultâneas e permite que sua arquitetura seja facilmente modificada, permitindo que seja adaptada para as tarefas de classificação de imagens, detecção de objetos e segmentação semântica (WANG, J. *et al.*, 2020b; YUAN *et al.*, 2020; SUN *et al.*, 2019). Além disso, esta rede apresenta uma estrutura matematicamente elegante e homogênea.

No processo de treino, foi usada a técnica *One Cycle Policy* (SMITH, 2018) para otimizar as taxas de aprendizado durante o treino e também foi usado o método discutido em (SMITH, 2015) para encontrar a melhor limite superior para taxa de aprendizagem. Além disso, os modelos de classificação foram pré-treinados no *ImageNet Dataset* (DENG *et al.*, 2009).

### 3.1.2.2 Modelos de Segmentação Semântica

Para a abordagem de segmentação semântica, foram usadas as bibliotecas Fastai (FASTAI, 2019) e Sementorch (CASTILLO, 2020) e foi selecionada arquitetura U-Net considerando que ela já demonstrou desempenhos estado da arte nesta tarefa em várias modalidades de imagens médicas (LITJENS *et al.*, 2017; WANG, P. *et al.*, 2019). Adicionalmente, também foram realizados experimentos com a arquitetura HRNet, já que esta também pode ser adaptada para a tarefa de segmentação.

O modelo U-Net é composto por dois componentes principais:

- a) Um *framework* de segmentação, que é uma estrutura de DL em formato de “U” composta de duas partes: um braço de codificação, que recebe os valores dos pixels e aprende a codificá-los em uma representação abstrata da distribuição de pixels para cada classe; e um braço decodificador que gera uma imagem onde cada pixel recebe o valor de uma classe;
- b) um modelo de classificação de imagens como *backbone* para fazer o papel de módulo de processamento e aprendizado para ambos os braços.

Neste trabalho, os modelos de classificação selecionados como *backbone* foram ResNet 34, ResNet 50, e ResNet 101 (HE, Kaiming *et al.*, 2015) com um otimizador ADAM (KINGMA; BA, 2014) e a função de perda de entropia cruzada. Além disso, os modelos U-Net usam conexões horizontais adicionais entre as camadas correspondentes do *backbone* de cada braço que aprendem com diferentes níveis de abstração do processo de segmentação.

Nessa abordagem, também foi usada a técnica *One Cycle Policy* para otimizar as taxas de aprendizado durante o treino e também foi usado o método discutido em (SMITH, 2015) para encontrar o melhor limite superior para taxa de aprendizagem. Além disso, cada modelo de segmentação foi treinado iterativamente com versões de 300x400 pixels (1/4), 600x800 pixels (1/2) e 1200x1600 pixels (1) das imagens dos campos, nessa ordem. Para cada uma dessas resoluções, o treino foi executado em duas etapas: uma onde os parâmetros do codificador são congelados com os valores de pré-treinamento e uma outra de ajuste fino onde todos os parâmetros são descongelados. Nesse ciclo, os parâmetros da resolução anterior são usados para inicializar os modelos para a próxima resolução após a conclusão de cada treino. Este método iterativo com resolução progressiva é uma adaptação da estratégia apresentada em (HOWARD, 2019) e já obteve sucesso em conjuntos de dados de difícil treinamento como volumes micro-tomográficos na área de Paleontologia Marinha (CARVALHO *et al.*, 2020). Os modelos de segmentação foram pré-treinados no *ImageNet Dataset* (DENG *et al.*, 2009).

### 3.1.2.3 Modelos de Detecção de Objetos

Para a detecção de objetos, foram usadas as ferramentas Detectron2 (WU *et al.*, 2019) e MMDetection (MMDTECTION CONTRIBUTORS, 2018) e foram selecionadas as arquiteturas Faster R-CNN (REN *et al.*, 2017) e Mask R-CNN (HE, K. *et al.*, 2017), considerando que elas já mostraram desempenho estado da arte nessa tarefa em vários domínios de aplicação. Apesar de o modelo Mask R-CNN ser considerado um modelo de segmentação de instâncias, neste trabalho ele é avaliado apenas em relação à sua capacidade de detecção, sendo então considerado um modelo de detecção de objetos neste contexto.

A rede Faster R-CNN é um *framework* de detecção de objetos composto de:

- a) Uma rede de sugestão de regiões (do inglês, *region-proposal network*);
- b) uma rede de sugestão de máscaras (do inglês, *mask-proposal network*);
- c) um mecanismo de *pooling* de regiões de interesse;
- d) uma rede de classificação de regiões.

A rede de sugestão de regiões é uma CNN simples e a rede de sugestão de máscaras é uma *Fully Connected Network* (FCN). O classificador de regiões é composto de dois *backbones* que são as redes de classificação de imagens que atuam efetivamente como seletor das regiões candidatas e como mecanismo de classificação de regiões.

O modelo Faster R-CNN usa uma rede de classificação que é capaz de operar em múltiplas escalas de imagem. Para os experimentos deste trabalho, foi usado um par de *backbones* composto por uma *Feature Pyramid Network* (FPN) (LIN *et al.*, 2017) como seletor de regiões combinada com as redes ResNet 34, ResNet 50, HRNet 18 e HRNet 48 como módulos de classificação. A arquitetura Mask R-CNN é uma extensão do modelo Faster R-CNN que é capaz de segmentar instâncias com um módulo adicional para predição das máscaras de segmentação para cada região selecionada. Para essa abordagem, também foi aplicado o algoritmo *Non Maximum Suppression* (NMS) com 0,7 de limite de *Intersection over Union* (IoU) e 0,5 de limite de escore de predição. Os modelos de detecção foram pré-treinados no *COCO Dataset* (LIN *et al.*, 2014).

### 3.1.2.4 Combinações de Detecção de Objetos

Para avaliar o desempenho dos modelos de segmentação em nível de instância, foi usado um algoritmo de Componentes Conexos com conectividade de 8 pixels (BOLELLI *et al.*, 2020) para extrair instâncias das máscaras preditas pelos modelos. A classe de cada instância foi definida pela classe mais comum sob a máscara dessa instância. O escore de predição foi definido pela média de escores para a classe da instância sob a mesma máscara. Também, para otimizar esse *pipeline*, instâncias muito

pequenas para serem núcleos (menos de  $2.22\mu m$  de altura ou comprimento) ou com escore de predição menor que 0,5 foram descartadas. Também foi aplicado o algoritmo NMS com 0,7 de limite de IoU.

Também foi avaliada a combinação de localização e classificação usando os modelos que mostraram os melhores valores de  $F_1$  Score na localização de núcleos e na classificação da classe “Anormal”. Para esse experimento, as caixas de delimitação identificadas pelo modelo de localização foram usadas para extrair imagens de  $256 \times 256$  pixels de núcleos isolados que foram então passados como entrada para o modelo de classificação de imagens. A classe predita pelo modelo de classificação foi então associada à caixa de delimitação para calcular as métricas de avaliação multiclasse.

### 3.1.2.5 Métricas de Avaliação

Através das RSLs, foram levantadas as métricas precisão, sensibilidade e  $F_1$  Score para avaliação dos resultados. A métrica de acurácia não foi considerada pois o *dataset* é consideravelmente desbalanceado.

Quando expressas em termos de Verdadeiros Positivos (VP), Falsos Positivos (FP) e Falsos Negativos (FN) em relação à referência (ou *Ground truth* (GT)), as métricas de avaliação usadas podem ser descritas da seguinte forma:

- **Precisão:** Indica a probabilidade de um predição positiva ser verdadeira. Essa métrica é calculada conforme a Equação (1).

$$Precisão = \frac{VP}{VP + FP} \quad (1)$$

- **Sensibilidade ou Revocação:** Indica a capacidade do modelo de identificar instâncias positivas. Essa métrica é calculada conforme a Equação (2).

$$Sensibilidade = \frac{VP}{VP + FN} \quad (2)$$

- **$F_1$  Score:** É a média harmônica entre a precisão e a sensibilidade do modelo. Essa métrica é calculada conforme a Equação (3).

$$F_1 \text{ Score} = \frac{2VP}{2VP + FP + FN} \quad (3)$$

Foi usado um limite de 0.5 de IoU na abordagem de detecção de objetos para considerar se uma instância da referência foi ou não detectada. Para calcular essas métricas de forma padronizada para todos os modelos e abordagens, independentemente da ferramenta utilizada, foi usada a biblioteca de código, desenvolvida por nosso grupo de pesquisa, LapixDL (VICTÓRIA MATIAS; ATKINSON AMORIM, 2022). Este trabalho e a biblioteca usam as definições descritas por André Victória Matias *et al.* (2021) para implementar o cálculo das métricas.

### 3.1.2.6 Aumento de Dados

O Aumento de Dados (do inglês *Data Augmentation*) é uma técnica que busca aumentar a generalização de modelos de aprendizado de máquina através da adição de modificações nos dados de entrada que simulem variações que podem ocorrer em amostras reais, como oclusão de objetos, distorções na imagem e variações de iluminação. Nos experimentos realizados nesta pesquisa foi utilizada a biblioteca *Albumentations* (BUSLAEV *et al.*, 2020) e as modificações foram incluídas como etapa apenas do *pipeline* de treino.

As *Data Augmentations* aplicadas neste trabalho foram escolhidas levando em consideração a natureza dos dados e os resultados de múltiplos experimentos com modelos de detecção de objetos. Dessa forma, o conjunto de modificações aplicadas foi:

- Espelhamento horizontal e vertical aleatório com probabilidade de 0,5;
- Rotação aleatória entre -90 e 90 graus com probabilidade de 0,75 (interpolação por vizinho mais próximo);
- *Zoom* aleatório entre 1 e 1,1 com probabilidade de 0,75;
- Alterações aleatórias de iluminação e contraste com probabilidade de 0,75;
- *Shear* de magnitude entre -0,2 e 0,2 com probabilidade de 0,75.

## 3.2 RESULTADOS

Usando a primeira versão do UFSC-OCPap, o artigo “*Segmentation, Detection and Classification of Cell Nuclei on Oral Cytology Samples Stained with Papanicolaou*” (VICTÓRIA MATIAS, A. *et al.*, 2020) foi apresentado no *33rd IEEE International Symposium on Computer-Based Medical Systems (IEEE CBMS 2020)* e publicado como versão estendida na *SN Computer Science* (MATIAS *et al.*, 2021). Nesse trabalho foram testadas abordagens multiclasse e binárias com múltiplos modelos de segmentação (U-Net com ResNet 34 e ResNet 50), detecção (Fast R-CNN e RetinaNet) e classificação (ResNet 34 e ResNet 50), além de mecanismos de atenção. De acordo com os resultados desse artigo, nessa versão das marcações a melhor abordagem seria a combinação de localização dos núcleos usando Faster R-CNN, que obteve 0,76 de IoU, e classificação desses usando ResNet 34, que obteve 0,86 de  $F_1$  score. Os mecanismos de atenção não mostraram melhorias significativas nos modelos.

Experimentos com a segunda versão do *dataset* foram apresentados em formato de artigo no XVII Congresso Brasileiro de Informática em Saúde (CBIS 2020). Em “Localização de Núcleos Celulares em Citologia Oral Usando Métodos de Deep Learning” (MATIAS *et al.*, 2020), foram testados métodos de localização de núcleos

usando os modelos U-Net com ResNet 34, 50 e 101 além de Faster R-CNN e RetinaNet. O modelo Faster R-CNN foi o que apresentou o melhor desempenho entre os testados, com IoU de 0,81. A primeira página dos trabalhos citados nessa seção estão expostas no Apêndice A.

A partir dos resultados desses experimentos, percebeu-se que melhores resultados podem ser obtidos reduzindo-se o número de classes. Assim, as classes “Em divisão” e “Reacional” foram mapeadas para a classe “Saudável”. A classe “Sangue”, apesar de ter sido usada no treino das redes, não foi considerada na avaliação dos resultados por não ser relevante para o objetivo deste trabalho.

Por terem sido realizados experimentos com modelos de segmentação semântica e detecção de objetos, os resultados desses modelos serão apresentados em dois níveis de detalhes: em nível de pixel e em nível de instância. Métricas em nível de pixel são calculadas considerando a referência e os valores preditos para cada pixel da imagem e são válidas apenas para tarefas de segmentação semântica. Por sua vez, métricas em nível de instância são calculadas considerando os valores de referência e os valores preditos das *bounding boxes*. As métricas usadas neste trabalho e suas fórmulas estão descritas na Seção 3.1.2.5. Todas as métricas foram calculadas no subconjunto de teste. Um exemplo das predições dos melhores modelos pode ser visto na Figura 7.

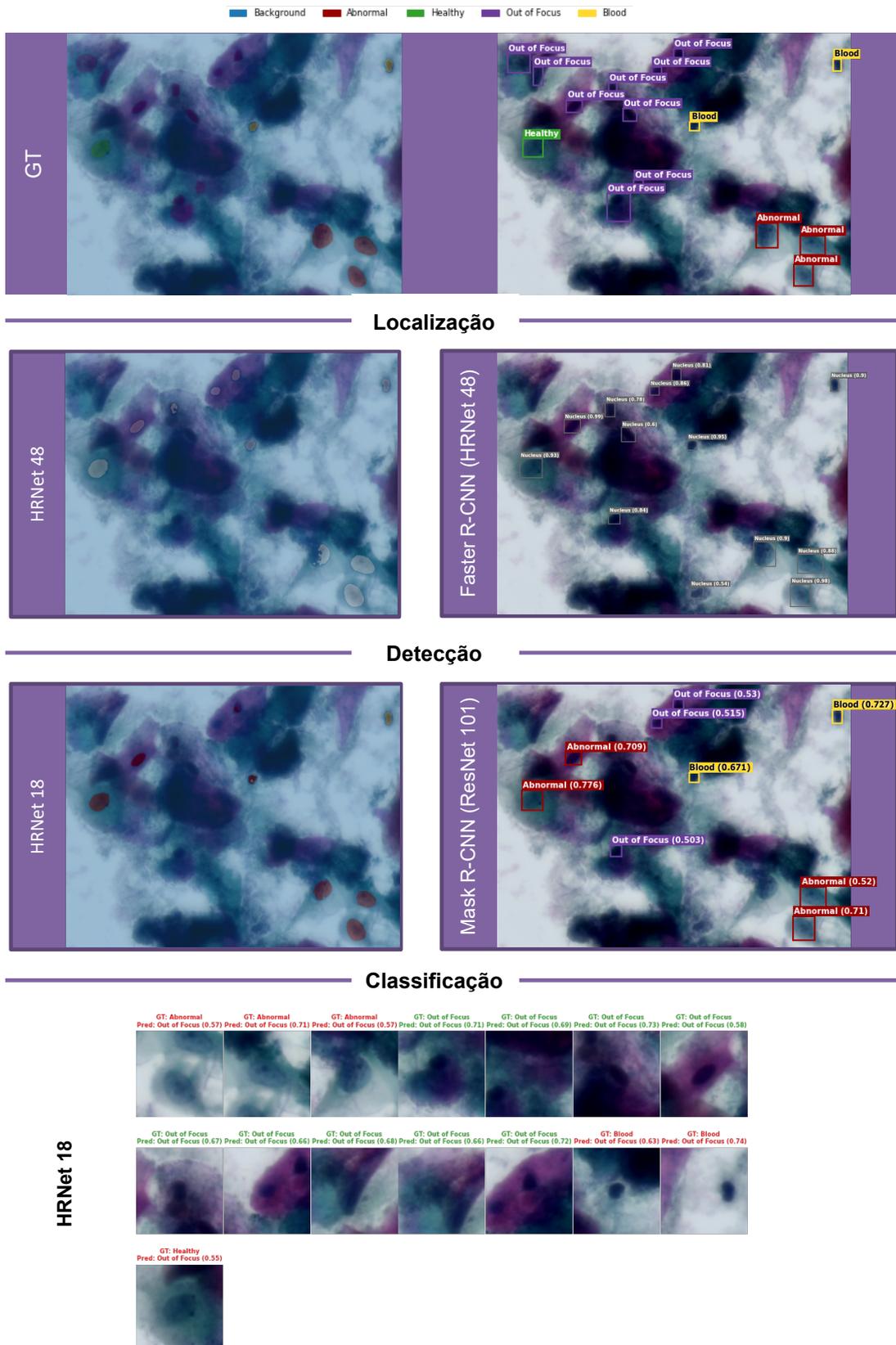
Para melhor avaliar as abordagens e modelos, a tarefa de detecção de núcleos foi dividida em duas sub-tarefas: localização e classificação. Neste trabalho, a tarefa “Localização” é responsável por determinar a localização de cada núcleo em uma imagem de campo da lâmina e é realizada pelos modelos binários de detecção e segmentação. Já a tarefa “Classificação” é responsável por determinar a classe de cada núcleo e é realizada pelos modelos multiclases de detecção, segmentação e classificação de imagens. Dessa forma, a tarefa “Detecção”, sendo a combinação dessas duas sub-tarefas, pode ser realizada por um modelo de detecção ou pela combinação de um modelo de localização e um modelo de classificação. Os resultados de localização, detecção e classificação são detalhados nas próximas subseções.

Em relação à importância de cada métrica para a comparação dos modelos, a métrica  $F_1$  Score é considerada a mais importante, seguida da sensibilidade e da precisão. Essa prioridade foi definida pois, para a detecção de células anormais, falsos negativos são mais críticos que falsos positivos, porém ainda é importante que o modelo tenha uma precisão alta, o que é refletido pela métrica  $F_1$  Score.

### 3.2.1 Localização

A Figura 8 mostra as métricas em nível de pixel, por tamanho de imagem, para os modelos de segmentação semântica. Os resultados que não estão presentes não foram computados por indisponibilidade de capacidade de VRAM de GPU. Apesar de

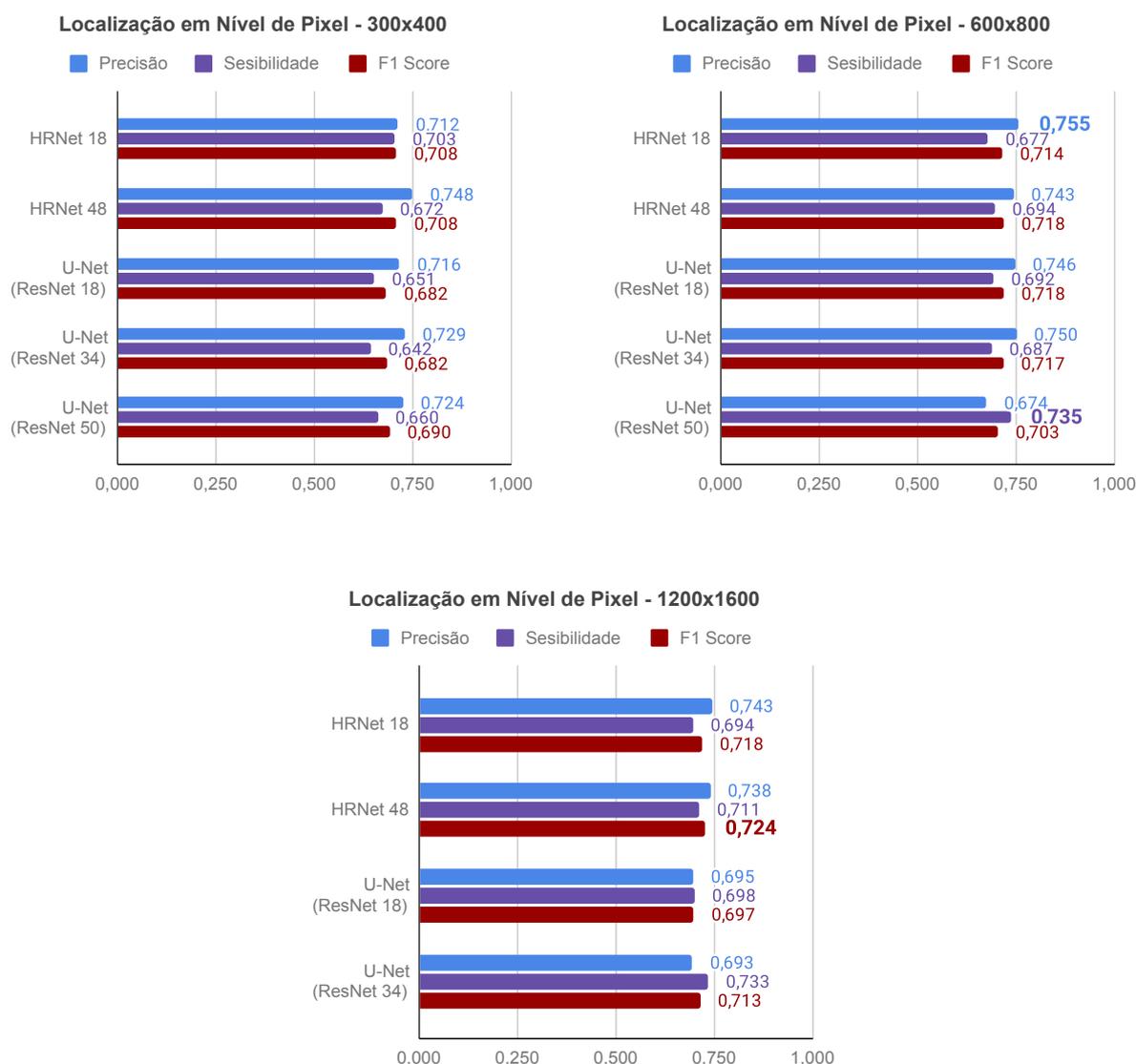
Figura 7 – GT e predições dos melhores modelos de cada abordagem para núcleos (*Nucleus*), núcleos saudáveis (*Healthy*), núcleos anormais (*Abnormal*), núcleos fora de foco (*Out of Focus*) e núcleos sanguíneos (*Blood*).



Fonte – Elaborada pelo autor.

o modelo U-Net com ResNet 50 ter mostrado a melhor sensibilidade em imagens de 600x800 pixels, o modelo HRNet 48 foi considerado o melhor modelo para segmentação semântica binária por ter obtido o maior  $F_1$  Score.

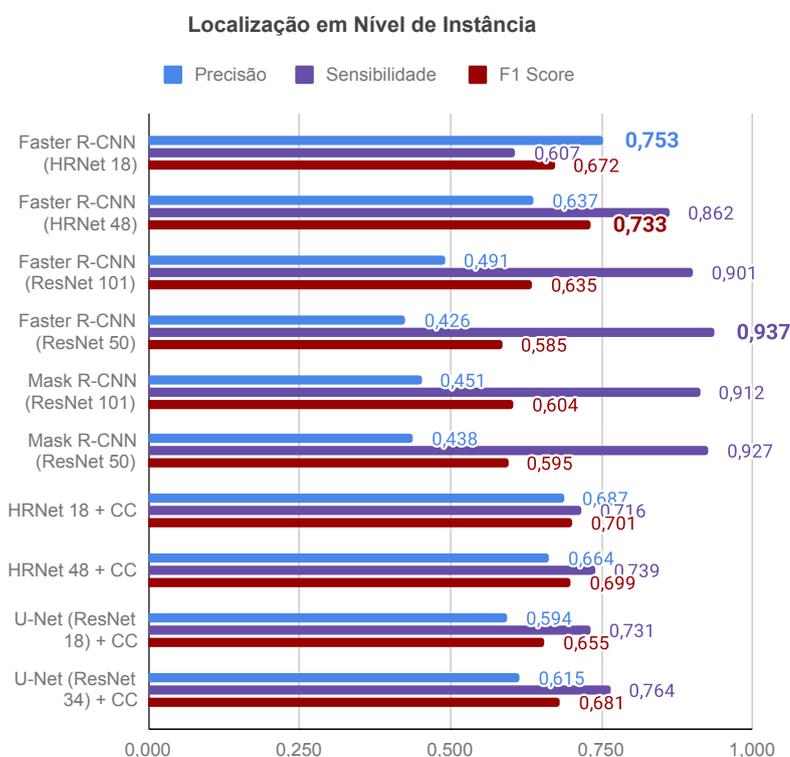
Figura 8 – Resultados em nível de pixel de modelos binários de segmentação semântica. Os maiores valores de cada métrica estão destacados em negro.



Fonte – Elaborada pelo autor.

Em relação à abordagem de detecção de objetos binária, como mostrado na Figura 9, os modelos de detecção isolados obtiveram melhores resultados que as combinações de modelos de segmentação semântica e Componentes Conexos. Ainda, entre os modelos de detecção, o modelo Faster R-CNN com *backbone* HRNet 48 obteve o maior valor de  $F_1$  Score.

Figura 9 – Resultados em nível de instância de modelos binários de detecção de objetos e da combinação de segmentação semântica e Componentes Conexos (CC). Os maiores valores de cada métrica estão destacados em negrito.



Fonte – Elaborada pelo autor.

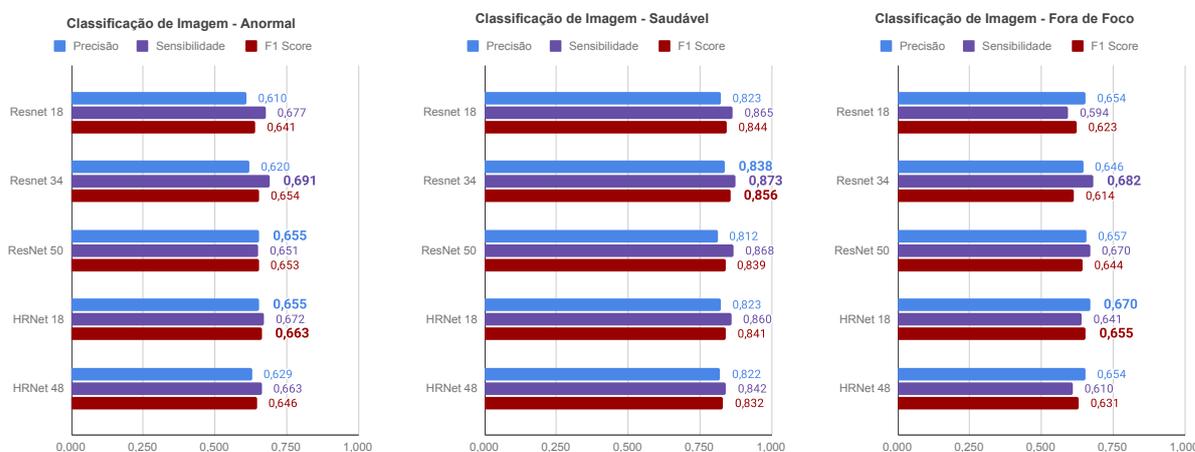
### 3.2.2 Classificação

Os valores de precisão, sensibilidade e  $F_1$  Score obtidos para cada classe pelos modelos de classificação de imagem são mostrados na Figura 10. Com base nesses resultados, o modelo ResNet 34 obteve os maiores valores para todas as métricas para a classe “Saudável”. No entanto, considerando que o modelo HRNet 18 mostrou os melhores valores de  $F_1$  Score para ambas as classes “Anormal” e “Fora de Foco”, este foi considerado o melhor modelo de classificação de imagem para este trabalho.

### 3.2.3 Detecção

A Tabela 2 mostra os resultados de segmentação semântica multiclasse para todos os tamanhos de imagem e classes avaliados. Considerando os melhores valores e que a classe “Anormal” é a mais relevante neste trabalho, o modelo HRNet 18 foi avaliado como o melhor modelo de segmentação semântica multiclasse. Apesar disso, o modelo HRNet 48 mostrou melhores métricas para as classes “Saudável” e “Fora de Foco” e também um valor razoável de  $F_1$  Score para a classe “Anormal”.

Figura 10 – Resultados dos modelos de classificação de imagem. Os maiores valores de cada métrica estão destacados em negrito.



Fonte – Elaborada pelo autor.

Tabela 2 – Resultados em nível de pixel de modelos de segmentação semântica multi-classe (%) por classe e tamanho da imagem. Os melhores resultados por métrica e tamanho de imagem estão destacados em negrito.

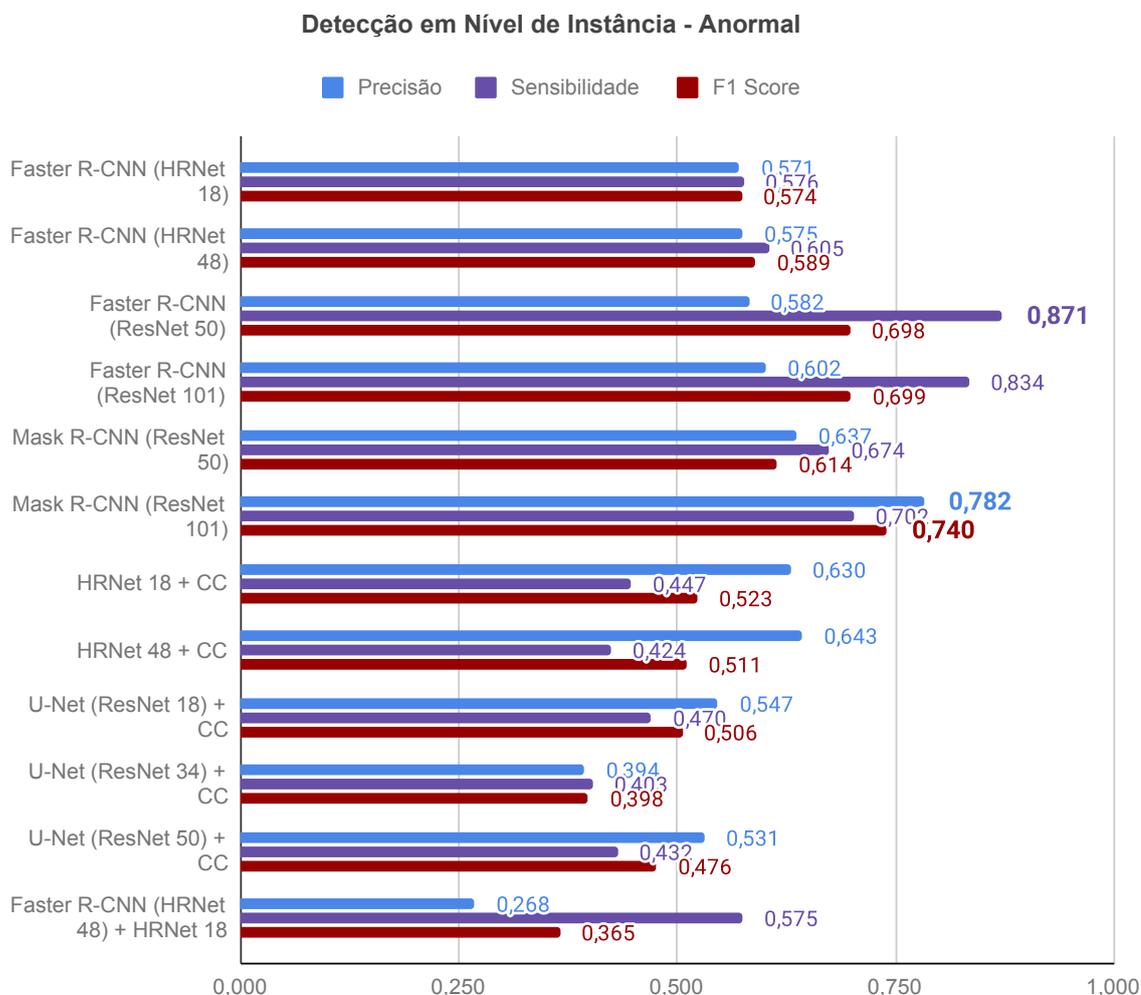
| Modelo               | Tamanho | Anormal     |             |                | Saudável    |             |                | Fora de Foco |             |                |
|----------------------|---------|-------------|-------------|----------------|-------------|-------------|----------------|--------------|-------------|----------------|
|                      |         | P           | S           | F <sub>1</sub> | P           | S           | F <sub>1</sub> | P            | S           | F <sub>1</sub> |
| U-Net<br>(ResNet 18) | 1/4     | 52,3        | 51,1        | 51,7           | 50,4        | 67,1        | 57,6           | <b>53,0</b>  | 07,4        | 12,9           |
|                      | 1/2     | <b>60,1</b> | 55,9        | 57,9           | 53,9        | 75,1        | 62,8           | 49,8         | 14,2        | 22,1           |
|                      | 1       | 50,1        | 60,0        | 54,6           | 63,2        | 49,9        | 55,8           | 39,4         | 15,8        | 22,5           |
| U-Net<br>(ResNet 34) | 1/4     | 54,7        | 46,7        | 50,4           | 51,7        | 66,3        | 58,1           | 51,2         | 04,3        | 07,9           |
|                      | 1/2     | 57,0        | 54,3        | 55,6           | 52,1        | 73,7        | 61,0           | 47,9         | 08,3        | 14,2           |
|                      | 1       | 39,2        | 58,8        | 47,1           | 53,7        | 11,4        | 18,8           | 28,7         | 20,3        | 23,8           |
| U-Net<br>(ResNet 50) | 1/4     | 52,5        | 52,5        | 52,5           | 52,0        | 66,5        | 58,4           | 48,2         | 07,9        | 13,6           |
|                      | 1/2     | 48,0        | 54,5        | 51,1           | <b>65,3</b> | 44,3        | 52,8           | 30,6         | <b>22,1</b> | 25,7           |
| HRNet 18             | 1/4     | 53,9        | 53,4        | 53,6           | 51,1        | 69,2        | 58,8           | 47,5         | 12,4        | 19,6           |
|                      | 1/2     | 57,4        | 60,4        | 58,9           | 51,3        | 78,4        | 62,0           | 51,7         | 12,8        | 20,5           |
|                      | 1       | 58,1        | <b>61,3</b> | <b>59,6</b>    | 52,7        | 84,2        | 64,8           | 48,7         | 09,6        | 16,1           |
| HRNet 48             | 1/4     | 59,4        | 48,6        | 53,4           | 51,8        | 69,5        | 59,4           | 45,8         | 19,4        | 27,3           |
|                      | 1/2     | 58,2        | 58,4        | 58,3           | 54,2        | 77,4        | 63,7           | 44,7         | 18,3        | 26,0           |
|                      | 1       | 57,0        | 60,7        | 58,8           | 53,7        | <b>86,0</b> | <b>66,1</b>    | 42,7         | <b>22,1</b> | <b>29,1</b>    |

Fonte – Elaborada pelo autor.

Por fim, as Figuras 11, 12 e 13 mostram os resultados de todos os modelos e combinações propostas para detecção de núcleos celulares neste trabalho. Analisando esses valores, é possível observar que as combinações de modelos de localização e classificação não obtiveram resultados melhores do que os modelos de detecção de objetos isolados. O modelo Mask R-CNN com *backbone* ResNet 101 obteve o melhor F<sub>1</sub> Score para a classe “Anormal” e, portanto, é considerado a melhor abordagem para a detecção de células anormais para o *dataset*. Esse modelo, no entanto, não obteve os melhores resultados para as outras duas classes, onde a arquitetura Faster R-CNN

com HRNet 48 mostrou os maiores valores de  $F_1$  Score.

Figura 11 – Resultados em nível de instância da abordagem de detecção de objetos para a classe “Anormal”. Os maiores valores de cada métrica por classe estão destacados em negrito.

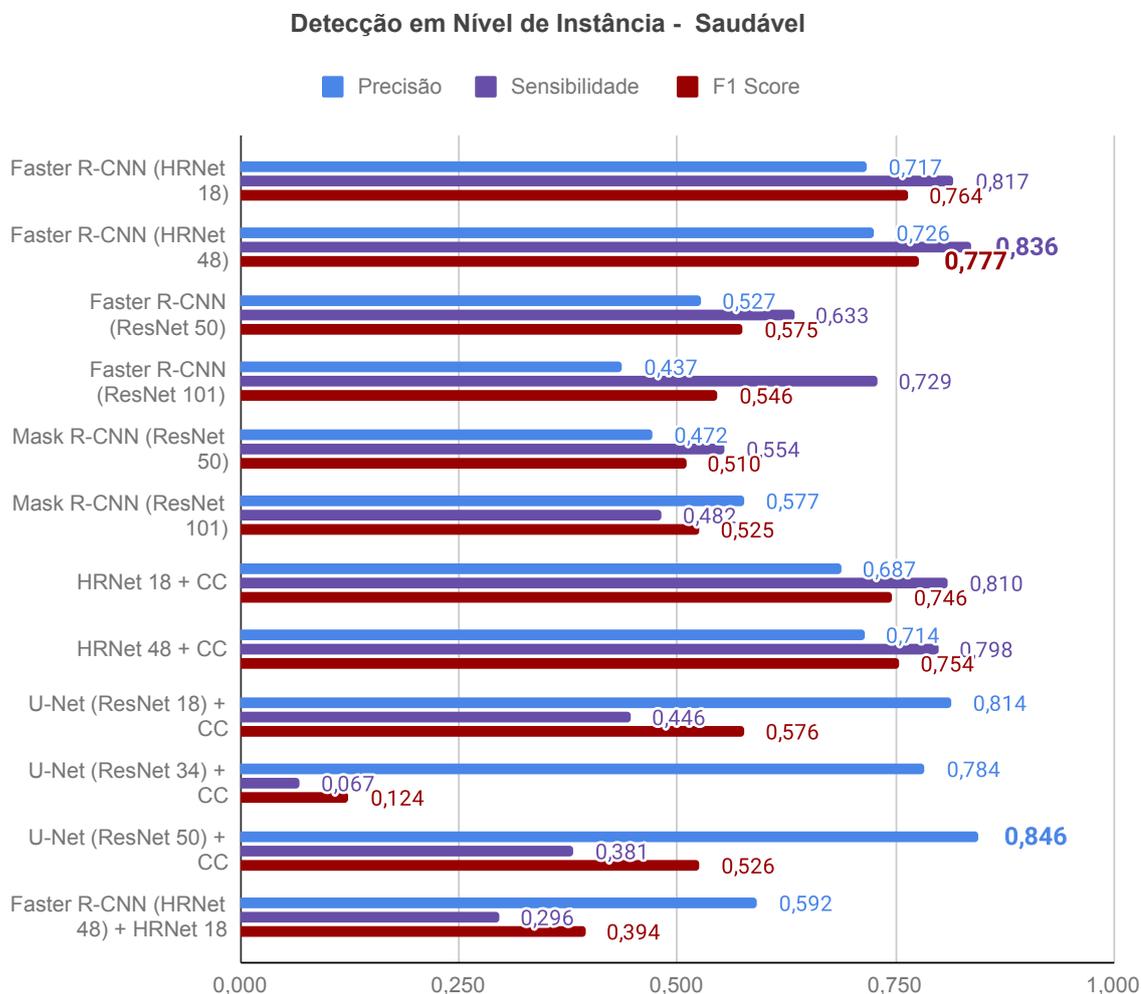


Fonte – Elaborada pelo autor.

### 3.3 DISCUSSÃO

Neste trabalho, foram apresentados três diferentes métodos de processamento de imagens para imagens de citologia oral corada com o corante de Papanicolau. Todos esses métodos foram desenvolvidos usando técnicas de DL e testados com diferentes modelos de DL. Os dois primeiros métodos usaram apenas uma arquitetura de DL enquanto o terceiro usou combinações de duas abordagens sequencialmente em forma de cascata.

Figura 12 – Resultados em nível de instância da abordagem de detecção de objetos para a classe “Saudável”. Os maiores valores de cada métrica por classe estão destacados em negrito.

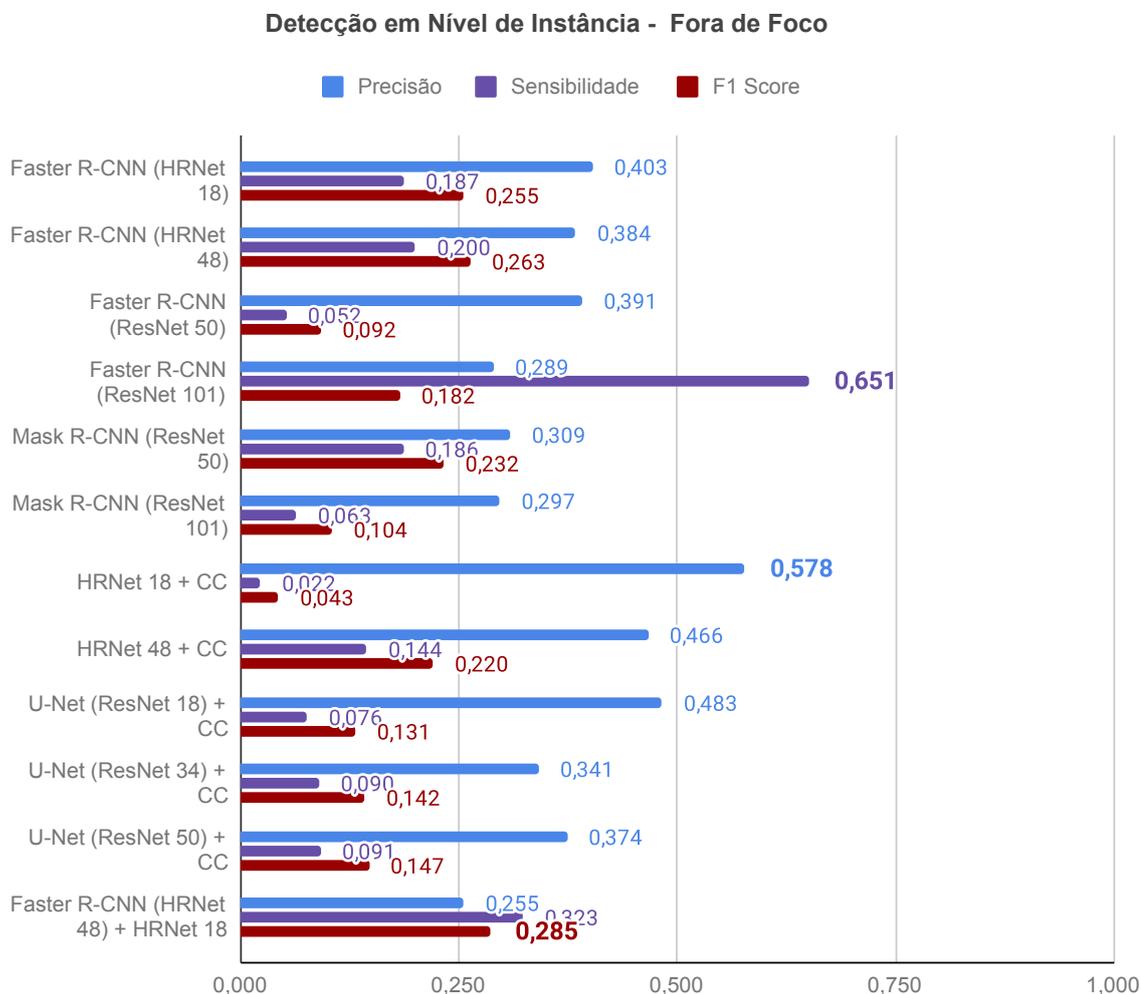


Fonte – Elaborada pelo autor.

De todos os métodos propostos, o resultado que mais se destaca é o de que os *pipelines* compostos não mostraram métricas melhores que os modelos isolados. Esses experimentos foram realizados porque os modelos de localização e classificação mostraram resultados promissores nos trabalhos publicados com versões anteriores do *dataset*. O resultado talvez possa ser explicado em virtude de os modelos de detecção terem atingido um desempenho melhor com um conjunto de dados maior e mais diverso. Os modelos ainda apresentam um desempenho consideravelmente melhor na tarefa de localização do que na tarefa de classificação de núcleos anormais, porém o melhor modelo de detecção (0.74 de  $F_1$  Score para a classe “Anormal”) superou o melhor modelo de localização (0.73 de  $F_1$  Score).

Nenhum dos modelos analisados foi capaz de detectar núcleos fora de foco com

Figura 13 – Resultados em nível de instância da abordagem de detecção de objetos para a classe “Fora de Foco”. Os maiores valores de cada métrica por classe estão destacados em negrito.



Fonte – Elaborada pelo autor.

valores consideráveis de métricas. Esse tipo de artefato é um problema comum nesse tipo de análise e geralmente não está presente em conjuntos de dados de imagens de citologia. Como este trabalho busca identificar um método que possa ser aplicado na rotina clínica, é importante incluir esses resultados para mostrar que, possivelmente, CNNs não são capazes de diferenciar entre células em foco e fora de foco em imagens de citologia e, portanto, este problema exige pesquisas mais aprofundadas de métodos de pré-processamento que possam auxiliar nessa tarefa.

Outra limitação deste trabalho é o número de pacientes incluídos no conjunto de dados: apesar de ser composto por mais de 6.500 imagens, elas foram obtidas de apenas 8 lâminas preparadas no mesmo centro de coletas. Isso pode limitar a capacidade de generalização dos modelos, mas foi considerado que essa amostra é

suficiente para o objetivo do trabalho, que, principalmente, é comparar modelos para a detecção de núcleos anormais. Para trabalhos futuros em que o objetivo é prever o diagnóstico de lâminas inteiras, serão coletadas amostras de mais pacientes para aumentar o tamanho e a qualidade do *dataset*.

Levando em conta as abordagens avaliadas, os resultados mostram que a detecção de objetos é a melhor para o domínio de imagens de citologia oral, já que os modelos de segmentação semântica e classificação de imagens não alcançaram resultados satisfatórios de  $F_1$  Score para a classe “Anormal”, nem em nível de instância, nem em nível de pixel. Outro resultado que vale ser destacado é que o melhor modelo de segmentação binária obteve 0,724 de  $F_1$  Score para localização em nível de pixel, mas não foi o melhor modelo em nível de instância quando combinado com o algoritmo de Componentes Conexos.

Os resultados também mostraram que o modelo HRNet é uma arquitetura promissora para ser aplicada na análise de imagens de citologia. De acordo com as RSLs usadas neste trabalho, esse modelo nunca havia sido testado nesse domínio antes e, apesar de não ser o melhor modelo dentre os avaliados, ele mostrou os melhores resultados de localização e classificação do trabalho e também o melhor valor de  $F_1$  Score na detecção das classes “Saudável” e “Fora de Foco”. Considerando que o modelo HRNet foi comparado com outros bem estabelecidos, como U-Net e ResNet, ele mostrou resultados promissores e deve ser mais bem explorado em trabalhos futuros.

Em relação aos trabalhos relacionados, a Tabela 3 mostra uma comparação dos resultados dos experimentos que mais se aproximam da proposta desta pesquisa. É importante ressaltar que essa comparação mostra apenas uma visão geral dos resultados obtidos em experimentos com DL em imagens de citologia oral e não tem o objetivo de definir qual o melhor modelo para esta aplicação, já que todos os trabalhos usam diferentes conjuntos de dados para calcular as métricas dos resultados e isso impossibilita uma comparação justa. Levando isso em consideração, esta pesquisa é a que usa o maior *dataset* em relação ao número de imagens e é a primeira a usar um conjunto de imagens de citologia oral público, apesar de ser a com menor número de pacientes.

Em relação aos resultados de  $F_1$  Score para núcleos anormais, ocorreu uma piora em relação ao trabalho anterior (MATIAS *et al.*, 2021). Porém, os experimentos anteriores usam uma versão menor do UFSC OCPap com apenas 2 lâminas, o que pode ter enviesado os resultados. Quando comparado aos resultados dos trabalhos que usam amostras em meio líquido para testar os modelos, o resultado de  $F_1$  Score deste trabalho ficou entre os experimentos de Lu *et al.* (2020) e Banerjee *et al.* (2016). O trabalho de Banerjee *et al.* (2016) obteve o melhor resultado entre todos os trabalhos relacionados, mas também usa o *dataset* com menor número de imagens. Considerando que imagens de citologia convencional são mais complexas de serem

analisadas do que as preparadas em meio líquido, o fato de que os resultados deste trabalho tenham se revelado melhores que os obtidos por Lu *et al.* (2020) mostra que o melhor modelo testado nesta pesquisa é capaz de lidar relativamente bem com os ruídos produzidos em citologias convencionais.

Tabela 3 – Comparação dos melhores resultados dos principais trabalhos relacionados.

| Artigo                                  | Meio Líquido                    |                           | Convencional                  |               |
|---|---------------------------------|---------------------------|-------------------------------|---------------|
|   | (BANERJEE <i>et al.</i> , 2016) | (LU <i>et al.</i> , 2020) | (MATIAS <i>et al.</i> , 2021) | Este trabalho |
| <b>Pacientes</b>                        | 39                              | 10                        | 2                             | 8             |
| <b>Imagens</b>                          | 338                             | 1.014                     | 1.934                         | 6,753         |
| <b><math>F_1</math> Score (Anormal)</b> | 0,95                            | 0,61                      | 0,86                          | 0,74          |

Fonte – Elaborada pelo autor.

## 4 CONCLUSÃO

Neste trabalho, foram avaliadas abordagens de aprendizagem profunda para localizar e classificar núcleos celulares em imagens de citologia convencional coradas com o corante de Papanicolau para detecção de anomalias. Para isso, foi levantado o estado da arte na aplicação de técnicas computacionais em imagens de citologia oral através de duas RSLs (VICTÓRIA MATIAS, André *et al.*, 2021; SERGIO *et al.*, 2019), onde foi identificada uma tendência de crescimento de pesquisas envolvendo DL para a automatização da análise de citologias, apesar de métodos clássicos de visão computacional ainda serem amplamente usados. Além disso, foi destacado o fato de que os trabalhos analisados estão em fase experimental, sem testes na rotina clínica.

Para avaliar a aplicação de DL na detecção de anomalias em núcleos celulares de citologias orais, foram elencados os modelos ResNet, HRNet, U-Net, Faster R-CNN e Mask R-CNN. Comparando os resultados de precisão, sensibilidade e  $F_1$  Score, o melhor modelo encontrado para localização e classificação de núcleos anormais foi o Mask R-CNN com *backbone* ResNet 101 que alcançou 0,74 de  $F_1$  Score para essa classe. Esse resultado vai contra o que foi assumido nos experimentos parciais de que uma combinação de modelos de DL teria desempenho melhor do que um modelo isolado. Outro destaque dos resultados é o desempenho das redes HRNet no domínio das imagens de citologia. Essa estrutura de CNN não havia sido testada nessa área antes e mostrou resultados promissores, com desempenho melhor que arquitetura bem estabelecida ResNet nas avaliações de localização e classificação e também os melhores resultados na detecção de núcleos saudáveis e fora de foco. Isso mostra que as HRNets devem ser mais exploradas na pesquisa de automação da análise de citologias. Os resultados parciais dessa pesquisa foram publicados na forma de artigos (VICTÓRIA MATIAS, A. *et al.*, 2020; MATIAS *et al.*, 2020, 2021).

A principal contribuição deste trabalho para a ciência da computação é mostrar a possibilidade de uso de CNNs na citologia e outros tipos de imagens semelhantes e comparar o desempenho de possíveis abordagens e modelos de DL para a automação da análise de citologias convencionais coradas com a técnica de Papanicolau. Outra importante contribuição é a publicação do *dataset* UFSC OCPap v4, que é o primeiro conjunto de dados público de imagens de campos de amostras de citologia oral com marcações para treinamento de modelos de segmentação semântica, detecção de objetos e classificação de imagens. Para as ciências da saúde, este trabalho mostra um possível método mais eficiente de detecção de células tumorais em exames de citologia esfoliativa, auxiliando o diagnóstico precoce de tumores malignos de cavidade oral e diminuindo a subjetividade da análise humana no resultado do exame. Com os resultados dessa pesquisa, pretende-se também incentivar uma maior difusão do

método de Papanicolau para exames preventivos para detecção precoce de câncer oral, aumentando as chances de sucesso dos tratamentos.

Dessa forma, como trabalhos futuros, se propõe testar os melhores modelos identificados nesta pesquisa em nível de lâmina inteira para validar se esses modelos são capazes de distinguir entre amostras saudáveis e de câncer quando aplicados em todos os campos de uma lâmina. Ainda, se pretende aumentar a qualidade e o tamanho do conjunto de imagens e generalizar a análise para outros tipos de tecidos e núcleos celulares para dar suporte ao diagnóstico de outros tipos de câncer. O Objetivo, nos próximos anos, é integrar esse processo automatizado no módulo de Telepatologia (ANDREÓLI PETROLINI *et al.*, 2019) e no módulo de Telemedicina Oral (MEURER *et al.*, 2020) de um sistema público de Telemedicina (NOBRE; WANGENHEIM, 2012; WANGENHEIM; NUNES, 2019) e realizar um estudo de validação em larga escala da abordagem proposta neste trabalho que também irá considerar a viabilidade de integrar a Tele-Citologia nos procedimentos de Telemedicina.

#### 4.1 SE EU FOSSE DESENVOLVER UM SISTEMA HOJE, O QUE EU USARIA?

A partir dos experimentos realizados, pôde-se perceber que diferentes modelos foram particularmente melhores para diferentes classes de núcleos e nenhum modelo se mostrou capaz de lidar com núcleos fora de foco. Além disso, de forma geral, modelos de detecção isolados obtiveram melhores resultados do que *pipelines* mais complexos ou compostos. Considerando isso, acredito que uma solução para filtrar núcleos fora de foco pode melhorar consideravelmente o desempenho dos modelos, já que estes acabam gerando falsos positivos para as outras classes. Os núcleos filtrados poderiam então ser entrada em um sistema de *Ensemble Learning* composto pelos modelos Mask R-CNN com *backbone* ResNet 101, o melhor para detecção de núcleos anormais, e Faster R-CNN com *backbone* HRNet 48, o melhor para detecção de núcleos saudáveis.

## REFERÊNCIAS

ANDREÓLI PETROLINI, V.; BECKHAUSER, E.; SAVARIS, A.; INES MEURER, M.; VON WANGENHEIM, A.; KRECHEL, D. Collaborative Telepathology in a Statewide Telemedicine Environment - First Tests in the Context of the Brazilian Public Healthcare System. *In: 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*. [S.l.: s.n.], 2019. P. 684–689. DOI: 10.1109/CBMS.2019.00139.

ARAÚJO, Flávio H.D.; SILVA, Romuere R.V.; USHIZIMA, Daniela M.; REZENDE, Mariana T.; CARNEIRO, Cláudia M.; CAMPOS BIANCHI, Andrea G.; MEDEIROS, Fátima N.S. Deep learning for cell image segmentation and ranking. **Computerized Medical Imaging and Graphics**, v. 72, p. 13–21, 2019. ISSN 0895-6111. DOI: <https://doi.org/10.1016/j.compmedimag.2019.01.003>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S089561111830048X>.

BANERJEE, S.; KARRI, S. P. K.; CHATTERJEE, S.; PAL, M.; PAUL, R. R.; CHATTERJEE, J. Multimodal diagnostic segregation of oral leukoplakia and cancer. *In: 2016 International Conference on Systems in Medicine and Biology (ICSMB)*. [S.l.: s.n.], 2016. P. 67–70.

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>.

BELL, A. A.; KAFTAN, J. N.; AACH, T.; MEYER-EBRECHT, D.; BÖCKING, A. High dynamic range images as a basis for detection of argyrophilic nucleolar organizer regions under varying stain intensities. **Proceedings - International Conference on Image Processing, ICIP**, p. 2541–2544, 2006. ISSN 15224880. DOI: 10.1109/ICIP.2006.312959.

BERNANDER, Karl Bengtsson; LINDBLAD, Joakim; STRAND, Robin; NYSTRÖM, Ingela. Replacing Data Augmentation with Rotation-Equivariant CNNs in Image-Based Classification of Oral Cancer. *In: TAVARES, João Manuel R. S.; PAPA, João Paulo; GONZÁLEZ HIDALGO, Manuel (Ed.). Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Cham: Springer International Publishing, 2021. P. 24–33.

BOLELLI, Federico; ALLEGRETTI, Stefano; BARALDI, Lorenzo; GRANA, Costantino. Spaghetti Labeling: Directed Acyclic Graphs for Block-Based Connected Components

Labeling. **IEEE Transactions on Image Processing**, v. 29, p. 1999–2012, 2020. DOI: 10.1109/TIP.2019.2946979.

BUSLAEV, Alexander; IGLOVIKOV, Vladimir I.; KHVEDCHENYA, Eugene; PARINOV, Alex; DRUZHININ, Mikhail; KALININ, Alexandr A. Albumentations: Fast and Flexible Image Augmentations. **Information**, v. 11, n. 2, 2020. ISSN 2078-2489. DOI: 10.3390/info11020125. Disponível em: <https://www.mdpi.com/2078-2489/11/2/125>.

CARVALHO, L.E.; FAUTH, G.; BAECKER FAUTH, S.; KRAHL, G.; MOREIRA, A.C.; FERNANDES, C.P.; VON WANGENHEIM, A. Automated Microfossil Identification and Segmentation using a Deep Learning Approach. **Marine Micropaleontology**, v. 158, p. 101890, 2020. ISSN 0377-8398. DOI: <https://doi.org/10.1016/j.marmicro.2020.101890>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0377839819300830>.

CASTILLO, David Lacalle. **SemTorch**. [S.l.: s.n.], ago. 2020. Disponível em: <https://github.com/WaterKnight1998/SemTorch>.

CHAPMAN, Christopher N.; OTIS, Christopher N. From critical values to critical diagnoses. **Cancer Cytopathology**, Wiley, v. 119, n. 3, p. 148–157, abr. 2011. DOI: 10.1002/cncy.20158. Disponível em: <https://doi.org/10.1002/cncy.20158>.

CHATTERJEE, Saunak; NAWN, Debaleena; MANDAL, Mousumi; CHATTERJEE, Jyotirmoy; MITRA, Somdutta; PAL, Mousumi; PAUL, Ranjan Rashmi. Augmentation of Statistical Features in Cytopathology Towards Computer Aided Diagnosis of Oral Precancer/Cancer. *In: March. 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII)*. [S.l.]: IEEE, mar. 2018. P. 206–212. DOI: 10.1109/ICBSII.2018.8524706. Disponível em: <https://ieeexplore.ieee.org/document/8524706/>.

DENG, Jia; DONG, Wei; SOCHER, Richard; LI, Li-Jia; LI, Kai; FEI-FEI, Li. ImageNet: A large-scale hierarchical image database. *In: 2009 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2009. P. 248–255. DOI: 10.1109/CVPR.2009.5206848.

DEY, Susmita; SARKAR, Ripon; CHATTERJEE, Kabita; DATTA, Pallab; BARUI, Ananya; MAITY, Santi P. Pre-cancer risk assessment in habitual smokers from DIC images of oral exfoliative cells using active contour and SVM analysis. **Tissue**

- and Cell**, Elsevier BV, v. 49, n. 2, p. 296–306, abr. 2017. DOI: 10.1016/j.tice.2017.01.009. Disponível em: <https://doi.org/10.1016/j.tice.2017.01.009>.
- 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>.
- FASTAI. **Fastai Python Library v1.0.57**. [S.l.: s.n.], 2019. <http://docs.fast.ai/3>. [Online; accessed 3-March-2020].
- GILL, Gary W. **Cytopreparation**. [S.l.]: Springer New York, 2013. DOI: 10.1007/978-1-4614-4933-1. Disponível em: <https://doi.org/10.1007/978-1-4614-4933-1>.
- HAYASHI, Yoichi. New unified insights on deep learning in radiological and pathological images: Beyond quantitative performances to qualitative interpretation. **Informatics in Medicine Unlocked**, v. 19, p. 100329, 2020. ISSN 2352-9148. DOI: <https://doi.org/10.1016/j.imu.2020.100329>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S235291482030023X>.
- HE, K.; GKIOXARI, G.; DOLLÁR, P.; GIRSHICK, R. Mask R-CNN. *In*: 2017 IEEE International Conference on Computer Vision (ICCV). [S.l.: s.n.], 2017. P. 2980–2988. DOI: 10.1109/ICCV.2017.322.
- HE, Kaiming; ZHANG, Xiangyu; REN, Shaoqing; SUN, Jian. Deep Residual Learning for Image Recognition. **CoRR**, abs/1512.03385, 2015. arXiv: 1512.03385. Disponível em: <http://arxiv.org/abs/1512.03385>.
- HOWARD, J. **Deep Learning 2019 - Fastai Course**. [S.l.], 2019. YouTube, Jan. 25 [Video file], Available: [https://youtu.be/XfoYk\\_Z5AkI](https://youtu.be/XfoYk_Z5AkI) . [Accessed: May. 14, 2020].
- 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](https://doi.org/10.1007/978-3-642-38850-7_1).

JAIN, Anil K.; MAO, Jianchang; MOHIUDDIN, K. M. Artificial Neural Networks: A Tutorial. **Computer**, IEEE Computer Society Press, Washington, DC, USA, v. 29, n. 3, p. 31–44, mar. 1996. ISSN 0018-9162. DOI: 10.1109/2.485891. Disponível em: <https://doi.org/10.1109/2.485891>.

KHAN, SanaUllah; ISLAM, Naveed; JAN, Zahoor; DIN, Ikram Ud; RODRIGUES, Joel J. P. C. A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. **Pattern Recognition Letters**, Elsevier BV, v. 125, p. 1–6, jul. 2019. DOI: 10.1016/j.patrec.2019.03.022. Disponível em: <https://doi.org/10.1016/j.patrec.2019.03.022>.

KINGMA, Diederik P.; BA, Jimmy. **Adam: A Method for Stochastic Optimization**. [S.l.: s.n.], 2014. cite arxiv:1412.6980Comment: Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015. Disponível em: <http://arxiv.org/abs/1412.6980>.

KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. Imagenet classification with deep convolutional neural networks. *In*: ADVANCES in neural information processing systems. [S.l.: s.n.], 2012. P. 1097–1105.

LABELBOX. **Labelbox**. [S.l.: s.n.], 2021. <https://labelbox.com>. Acesso em: 26 fev. 2021.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. **Nature**, Springer Science e Business Media LLC, v. 521, n. 7553, p. 436–444, mai. 2015. DOI: 10.1038/nature14539. Disponível em: <https://doi.org/10.1038/nature14539>.

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, Tsung-Yi; DOLLAR, Piotr; GIRSHICK, Ross; HE, Kaiming; HARIHARAN, Bharath; BELONGIE, Serge. Feature Pyramid Networks for Object Detection. **2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, IEEE, jul. 2017. DOI: 10.1109/cvpr.2017.106. Disponível em: <http://dx.doi.org/10.1109/CVPR.2017.106>.

LIN, Tsung-Yi *et al.* Microsoft COCO: Common Objects in Context. **CoRR**, abs/1405.0312, 2014. arXiv: 1405.0312. Disponível em: <http://arxiv.org/abs/1405.0312>.

LITJENS, Geert; KOOI, Thijs; BEJNORDI, Babak Ehteshami; SETIO, Arnaud Arindra Adiyoso; CIOMPI, Francesco; GHAFORIAN, Mohsen; LAAK, Jeroen A.W.M. van der; GINNEKEN, Bram van; SÁNCHEZ, Clara I. A survey on deep learning in medical image analysis. **Medical Image Analysis**, Elsevier BV, v. 42, p. 60–88, dez. 2017. DOI: 10.1016/j.media.2017.07.005. Disponível em: <https://doi.org/10.1016/j.media.2017.07.005>.

LIU, Xiaolong; DENG, Zhidong; YANG, Yuhan. Recent progress in semantic image segmentation. **Artificial Intelligence Review**, 2019. ISSN 15737462. DOI: 10.1007/s10462-018-9641-3.

LU, Jiahao; SLADOJE, Nataša; STARK, Christina Runow; RAMQVIST, Eva Darai; HIRSCH, Jan-Michaél; LINDBLAD, Joakim. A Deep Learning Based Pipeline for Efficient Oral Cancer Screening on Whole Slide Images. *In*: LECTURE Notes in Computer Science. [S.l.]: Springer International Publishing, 2020. P. 249–261. DOI: 10.1007/978-3-030-50516-5\_22. Disponível em: [https://doi.org/10.1007/978-3-030-50516-5\\_22](https://doi.org/10.1007/978-3-030-50516-5_22).

LUCENA, Eudes; MIRANDA, Alessandra; ARAÚJO, Fábio; GALVÃO, Carlos; MEDEIROS, Ana. Collection Method and the Quality of the Smears from Oral Mucosa. **Revista de Cirurgia e Traumatologia Buco-maxilo-facial**, v. 11, n. 2, p. 55–62, 2011. ISSN 1808-5210.

MATIAS, André Victória; CERENTINI, Allan; MACARINI, Luiz Antonio Buschetto; AMORIM, João Gustavo Atkinson; DALTOÉ, Felipe Perozzo; WANGENHEIM, Aldo von. Localização de Núcleos Celulares em Citologia Oral Usando Métodos de Deep Learning. *Journal of Health Informatics*, 2020. ISSN 2175-4411. Disponível em: <http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/view/810>.

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. 2021. DOI: 10.1007/s42979-021-00676-8. Disponível em: <https://doi.org/10.1007/s42979-021-00676-8>.

MEURER, Maria InÃ¡s; VON WANGENHEIM, Aldo; ZIMMERMANN, Caroline; SAVARIS, Alexandre; PETROLINI, VinÃ¡cius AndreÃ¡li; WAGNER, Harley Miguel. Launching a public statewide tele(oral)medicine service in Brazil during COVID-19 pandemic. **Oral Diseases**, n/a, n/a, 2020. DOI: 10.1111/odi.13528. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/odi.13528>. DisponÃ­vel em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/odi.13528>.

MMDTECTION CONTRIBUTORS. **OpenMMLab Detection Toolbox and Benchmark**. [S.l.: s.n.], ago. 2018. DisponÃ­vel em: <https://github.com/open-mmlab/mmdetection>.

NOBRE, Luiz Felipe; WANGENHEIM, Aldo von. Development and Implementation of a Statewide Telemedicine/Telehealth System in the State of Santa Catarina, Brazil. *In: TECHNOLOGY Enabled Knowledge Translation for eHealth: Principles and Practice*. New York, NY: Springer New York, 2012. P. 379–400. ISBN 978-1-4614-3495-5. DOI: 10.1007/978-1-4614-3495-5\_22. DisponÃ­vel em: [https://doi.org/10.1007/978-1-4614-3495-5\\_22](https://doi.org/10.1007/978-1-4614-3495-5_22).

PEKTAŞ, Zafer Özgür; KESKIN, Ahmet; GÜNHAN, Ömer; KARSLIOĞLU, Yildirim. Evaluation of nuclear morphometry and DNA ploidy status for detection of malignant and premalignant oral lesions: Quantitative cytologic assessment and review of methods for cytomorphometric measurements. **Journal of Oral and Maxillofacial Surgery**, v. 64, n. 4, p. 628–635, 2006. ISSN 02782391. DOI: 10.1016/j.joms.2005.12.010.

PEKTAŞ, Zafer Özgür; KESKIN, Ahmet; GÜNHAN, Ömer; KARSLIOĞLU, Yildirim. Evaluation of Nuclear Morphometry and DNA Ploidy Status for Detection of Malignant and Premalignant Oral Lesions: Quantitative Cytologic Assessment and Review of Methods for Cytomorphometric Measurements. **Journal of Oral and Maxillofacial Surgery**, Elsevier BV, v. 64, n. 4, p. 628–635, abr. 2006. DOI: 10.1016/j.joms.2005.12.010. DisponÃ­vel em: <https://doi.org/10.1016/j.joms.2005.12.010>.

REN, Shaoqing; HE, Kaiming; GIRSHICK, Ross; SUN, Jian. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Institute of Electrical e Electronics Engineers (IEEE), v. 39, n. 6, p. 1137–1149, jun. 2017. ISSN 2160-9292. DOI: 10.1109/tpami.2016.2577031. DisponÃ­vel em: <http://dx.doi.org/10.1109/TPAMI.2016.2577031>.

RIPLEY, Brian D. **Pattern recognition and neural networks**. [S.l.]: Cambridge university press, 2007.

RUSSELL, BryanC.; TORRALBA, Antonio; MURPHY, KevinP.; FREEMAN, WilliamT. LabelMe: A Database and Web-Based Tool for Image Annotation. English. **International Journal of Computer Vision**, Springer US, v. 77, n. 1-3, p. 157–173, 2008. ISSN 0920-5691. DOI: 10.1007/s11263-007-0090-8. Disponível em: <http://dx.doi.org/10.1007/s11263-007-0090-8>.

SERGIO, Bruna Zanicoski; LUIZ ANTONIO BUSCHETTO MACARINI; TOÉ, Felipe Perozzo Dal; WANGENHEIM, Aldo Von. Computer-assisted technologies for diagnosis of oral cancer on cytology samples - A Systematic Literature Review. en. Federal University of Santa Catarina, 2019. DOI: 10.13140/RG.2.2.14207.76964. Disponível em: <http://rgdoi.net/10.13140/RG.2.2.14207.76964>.

SMITH, Leslie N. A disciplined approach to neural network hyper-parameters: Part 1 - learning rate, batch size, momentum, and weight decay. **CoRR**, abs/1803.09820, 2018. arXiv: 1803.09820. Disponível em: <http://arxiv.org/abs/1803.09820>.

SMITH, Leslie N. No More Pesky Learning Rate Guessing Games. **CoRR**, abs/1506.01186, 2015. arXiv: 1506.01186. Disponível em: <http://arxiv.org/abs/1506.01186>.

SUN, Ke *et al.* High-resolution representations for labeling pixels and regions. **arXiv preprint arXiv:1904.04514**, 2019.

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**, v. 71, n. 3, p. 209–249, 2021. DOI: <https://doi.org/10.3322/caac.21660>. Disponível em: <https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.3322/caac.21660>.

VICTÓRIA MATIAS, A.; CERENTINI, A.; BUSCHETTO MACARINI, L. A.; ATKINSON AMORIM, J. G.; PEROZZO DALTOÉ, F.; VON WANGENHEIM, A. Segmentation, Detection and Classification of Cell Nuclei on Oral Cytology Samples Stained with Papanicolaou. In: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS). [S.l.: s.n.], jul. 2020. P. 53–58. DOI: 10.1109/CBMS49503.2020.00018.

VICTÓRIA MATIAS, André. **UFSC OCPap: Papanicolaou Stained Oral Cytology Dataset (v4)**. [S.l.]: Mendeley, 2022. DOI: 10.17632/dr7ydy9xbk.2. Disponível em: <https://data.mendeley.com/datasets/dr7ydy9xbk/2>.

VICTÓRIA MATIAS, André; ATKINSON AMORIM, João Gustavo. **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>.

VICTÓRIA MATIAS, André; ATKINSON AMORIM, João Gustavo; BUSCHETTO MACARINI, Luiz Antonio; CERENTINI, Allan; CASIMIRO ONOFRE, Alexandre Sherlley; DE MIRANDA ONOFRE, Fabiana Botelho; DALTOÉ, Felipe Perozzo; STEMMER, Marcelo Ricardo; VON WANGENHEIM, Aldo. What is the state of the art of computer vision-assisted cytology? A Systematic Literature Review. **Computerized Medical Imaging and Graphics**, v. 91, p. 101934, 2021. ISSN 0895-6111. DOI: <https://doi.org/10.1016/j.compmedimag.2021.101934>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0895611121000835>.

WANG, Jingdong *et al.* Deep High-Resolution Representation Learning for Visual Recognition. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, p. 1–1, 2020a. DOI: 10.1109/TPAMI.2020.2983686.

WANG, Jingdong *et al.* Deep High-Resolution Representation Learning for Visual Recognition. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, p. 1–1, 2020b. DOI: 10.1109/TPAMI.2020.2983686.

WANG, Pin; WANG, Lirui; LI, Yongming; SONG, Qi; LV, Shanshan; HU, Xianling. Automatic cell nuclei segmentation and classification of cervical Pap smear images. **Biomedical Signal Processing and Control**, Elsevier BV, v. 48, p. 93–103, fev. 2019. DOI: 10.1016/j.bspc.2018.09.008. Disponível em: <https://doi.org/10.1016/j.bspc.2018.09.008>.

WANGENHEIM, Aldo von; NUNES, Daniel Holthausen. Creating a Web Infrastructure for the Support of Clinical Protocols and Clinical Management: An Example in Teledermatology. **Telemedicine and e-Health**, v. 25, n. 9, p. 781–790, 2019. PMID: 30499753. DOI: 10.1089/tmj.2018.0197. eprint: <https://doi.org/10.1089/tmj.2018.0197>. Disponível em: <https://doi.org/10.1089/tmj.2018.0197>.

WU, Yuxin; KIRILLOV, Alexander; MASSA, Francisco; LO, Wan-Yen; GIRSHICK, Ross. **Detectron2**. [S.l.: s.n.], 2019. <https://github.com/facebookresearch/detectron2>.

YUAN, Yuhui; CHEN, Xilin; WANG, Jingdong. Object-contextual representations for semantic segmentation. *In*: SPRINGER. COMPUTER Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI 16. [S.l.: s.n.], 2020. P. 173–190.

# **Apêndices**

**APÊNDICE A – PRIMEIRAS PÁGINAS DE ARTIGOS PUBLICADOS**

# Segmentation, Detection and Classification of Cell Nuclei on Oral Cytology Samples Stained with Papanicolaou

André Victória Matias

*Department of Informatics and Statistics  
Federal University of Santa Catarina  
Florianópolis, Brazil  
andre.v.matias@posgrad.ufsc.br*

Allan Cerentini

*Department of Informatics and Statistics  
Federal University of Santa Catarina  
Florianópolis, Brazil  
allan.c@posgrad.ufsc.br*

Luiz Antonio Buschetto Macarini

*Automation and Systems Department  
Federal University of Santa Catarina  
Florianópolis, Brazil  
luiz.buschetto@posgrad.ufsc.br*

João Gustavo Atkinson Amorim

*Department of Informatics and Statistics  
Federal University of Santa Catarina  
Florianópolis, Brazil  
joao.atkinson@posgrad.ufsc.br*

Felipe Perozzo Daltoé

*Department of Pathology  
Federal University of Santa Catarina  
Florianópolis, Brazil  
felipe.daltoa@ufsc.br*

Aldo von Wangenheim

*Brazilian Institute for Digital Convergence  
Federal University of Santa Catarina  
Florianópolis, Brazil  
aldo.vw@ufsc.br*

**Abstract**—Although oral cancer is considered a global health issue with 350.000 people diagnosed over a year it can successfully be treated if diagnosed at early stages. Papanicolaou is an unexpensive and non-invasive method, generally applied to detect cervical cancer, but it can also be useful to detect cancer on oral cavities. The manual process of analyzing cells to detect abnormalities is a time-consuming cell analysis and is subject to variations in perceptions from different professionals. This paper compares models for three different deep learning approaches: segmentation, object detection and image classification. Our results show that the binary object detection with Faster R-CNN is the best approach for nuclei detection and localization (0.76 IoU). Since ResNet 34 had a good performance on abnormal nuclei classification (0.88 accuracy, 0.86  $F_1$  score) we concluded that these two models can be used in combination to make a localization and classification pipeline. This work reinforces that the automated analysis of oral cytology to make a pipeline for nuclei classification and localization using deep learning can help to minimize the subjectivity of the human analysis and also to detect cancer at early stages.

**Index Terms**—Papanicolaou, cytology, deep learning, detection, segmentation, classification

## I. INTRODUCTION

Oral cancer is a global health problem, with over 350,000 people being affected each year [1]. The high mortality rate is mostly related to the late diagnosis due the fact this disease doesn't present symptoms at early stages of development. Nonetheless, signs are present at the oral mucosa since the early beginning of the disease development and the use of non-invasive methods such as brush biopsy is strongly recommended. The early diagnosis is so important that it can increase survival rates from 19% to 80% [2].

We would like to thank Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (Fapescc) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for funding this work.

Chromosome instability can lead to an abnormal amount of DNA in the nuclei of each cell, which is called aneuploidy. Aneuploidy seems to play a major role on cancer development. Although this quantitative genetic material content analysis has been proven to be a promising and cost effective solution for early cancer detection, this approach is not widely used in clinical routine, because there is still no fully automated method for data analysis [3]. Currently, biopsy is still recognized as the gold-standard method for oral cancer diagnosis but it is an invasive surgical procedure that require specific room, staff and does not allow the cell DNA analysis. On the other hand, brush biopsy is cheaper, don't require special infrastructure because it is a non invasive method for cell collection and it is useful for DNA analysis.

Cytology is widely used as diagnostic tool in medicine and dentistry and always requires three things to be fully analysed: a stain for cells visualisation (the most popular is the Papanicolaou Staining method); 2) a light microscope for image amplification and; 3) a Pathology consultant. Although, the professionals have years of training and the other processes are easy to be standardize, diagnose on cytology samples is still considered a manual and time-consuming process, subject to variations in perceptions and human error. In order to avoid this problems, computer-assisted analysis such as deep learning methods can be used to detect alterations of cellular status, helping on early cancer diagnosis [4]. In this context, Machine Learning and Neural Networks have been employed to develop methods for the automated support of quantitative analysis of cytological samples for more than 25 years [5]–[9]. Recent advances in image processing, however, especially in the field of Deep Learning, have opened new possibilities.

In this work, we compare three different approaches using state-of-the-art methods to localize and classify the cells nu-



# Segmentation, Detection, and Classification of Cell Nuclei on Oral Cytology Samples Stained with Papanicolaou

André Victória Matias<sup>1</sup> · Allan Cerentini<sup>1</sup> · Luiz Antonio Buschetto Macarini<sup>2</sup> · João Gustavo Atkinson Amorim<sup>1</sup> · Felipe Perozzo Daltoé<sup>3</sup> · Aldo von Wangenheim<sup>4</sup>

Received: 23 November 2020 / Accepted: 3 May 2021  
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd 2021

## Abstract

Although oral cancer is considered a global health issue with 350,000 people diagnosed over a year, it can successfully be treated if diagnosed at early stages. Papanicolaou is an inexpensive and non-invasive method, generally applied to detect cervical cancer, but it can also be useful to detect cancer on oral cavities. The manual process of analyzing cells to detect abnormalities is a time-consuming cell analysis and is subject to variations in perceptions from different professionals. This paper compares three different deep learning (DL) approaches: segmentation, object detection, and image classification. Our results show that the binary object detection with Faster R-CNN is the best approach for nuclei detection and localization (0.76 IoU). Since ResNet 34 had a good performance on abnormal nuclei classification (0.86  $F_1$  score), we concluded that these two models can be used in combination to perform a reliable localization and classification pipeline. This work reinforces that the automated analysis of oral cytology to build a pipeline for nuclei classification and localization using DL can contribute to minimize the subjectivity of the human analysis and also support the detection of cancer at early stages.

**Keywords** Cytology · Deep learning · Segmentation · Classification

---

We would like to thank Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (Fapesc) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for funding this work.

---

This article is part of the topical collection “AI and Deep Learning Trends in Healthcare” guest edited by KC Santosh, Paolo Soda and Zalelam Temesgen.

---

✉ André Victória Matias  
andre.v.matias@posgrad.ufsc.br

Allan Cerentini  
allan.c@posgrad.ufsc.br

Luiz Antonio Buschetto Macarini  
luiz.buschetto@posgrad.ufsc.br

João Gustavo Atkinson Amorim  
joao.atkinson@posgrad.ufsc.br

Felipe Perozzo Daltoé  
felipe.daltoe@ufsc.br

Aldo von Wangenheim  
aldo.vw@ufsc.br

## Introduction

Oral cancer is a global health problem, with over 350,000 people being affected each year [5]. The high mortality rate is mostly related to the late diagnosis due to the fact this disease does not present symptoms at early stages of development. Nonetheless, signs are present at the oral mucosa since the early beginning of the disease development and the use of non-invasive methods such as brush biopsy is strongly recommended. The early diagnosis is so important

<sup>1</sup> Department of Informatics and Statistics, Federal University of Santa Catarina, Florianópolis, Brazil

<sup>2</sup> Automation and Systems Department, Federal University of Santa Catarina, Florianópolis, Brazil

<sup>3</sup> Department of Pathology, Federal University of Santa Catarina, Florianópolis, Brazil

<sup>4</sup> Brazilian Institute for Digital Convergence, Federal University of Santa Catarina, Florianópolis, Brazil



## Localização de Núcleos Celulares em Citologia Oral usando Métodos de Deep Learning

Cell Nuclei Localization on Oral Cytology Samples using Deep Learning Methods

Localización de Núcleos Celulares en Citología Oral usando Métodos de Deep Learning

André Victória Matias<sup>1</sup>, Allan Cerentini<sup>1</sup>, Luiz Antonio Buschetto Macarini<sup>2</sup>, João Gustavo Atkinson Amorim<sup>1</sup>, Ricardo Pereira<sup>3</sup>, Felipe Perozzo Daltoé<sup>4</sup>, Aldo Von Wangenheim<sup>5</sup>

### RESUMO

**Descritores:**

Citopatologia;  
Aprendizado Profundo;  
Neoplasias Bucais

**Objetivo:** Verificar a melhor abordagem para a detecção precoce de câncer na análise de imagens provenientes de lâminas de cavidade oral coradas com Papanicolaou. **Método:** São comparados os métodos de *Deep Learning* segmentação e detecção de objetos para localização de núcleos em imagens provenientes de lâminas de três diferentes pacientes. **Resultado:** O melhor resultado indica um IoU de 0,59 para a segmentação e 0,81 para a detecção de objetos. **Conclusão:** Os resultados mostram que a detecção de objetos usando o modelo Faster R-CNN possui potencial de ser usado juntamente com um modelo de classificação de imagens para auxiliar na detecção precoce de câncer. A base de imagens e os códigos usados foram disponibilizados publicamente.

### ABSTRACT

**Keywords:**

Cytopathology; Deep Learning; Mouth Neoplasms

**Objective:** Verify the best approach for the early detection of cancer in the analysis of images from oral cavity slides stained with Papanicolaou. **Method:** Deep Learning approaches of segmentation and object detection are compared to detect nuclei in images from three different patients' slides. **Result:** The best result indicates an IoU of 0.59 for segmentation and 0.81 for object detection. **Conclusion:** The results show that object detection using the Faster R-CNN model has the potential to be used in conjunction with an image classification model to aid in early cancer detection. The image base and codes used have been made publicly available.

### RESUMEN

**Descriptores:**

Citopatología;  
Aprendizaje Profundo;  
Neoplasias de la Boca

**Objetivo:** Verificar el mejor enfoque para la detección temprana del cáncer en el análisis de imágenes de diapositivas de la cavidad oral teñidas con el colorante de Papanicolaou. **Método:** Los enfoques de aprendizaje profundo de segmentación y detección de objetos se comparan para detectar núcleos en las imágenes de las diapositivas de tres pacientes diferentes. **Resultado:** El mejor resultado indica un IOU de 0,59 para la segmentación y 0,81 para la detección de objetos. **Conclusión:** Los resultados muestran que la detección de objetos mediante el modelo faster R-CNN tiene el potencial de ser usado en conjunto con un modelo de clasificación de imágenes para ayudar a la detección temprana de cáncer. La base de imágenes y los códigos utilizados se han puesto a disposición del público.

<sup>1</sup> Mestrando do Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de Santa Catarina – UFSC, Santa Catarina (SC), Brasil.

<sup>2</sup> Mestrando do Programa de Pós-Graduação em Engenharia de Automação e Sistemas, Universidade Federal de Santa Catarina – UFSC, Santa Catarina (SC), Brasil.

<sup>3</sup> Graduando em Engenharia Elétrica, Universidade Federal de Santa Catarina – UFSC, Santa Catarina (SC), Brasil.

<sup>4</sup> Professora do Departamento de Patologia, Universidade Federal de Santa Catarina – UFSC, Santa Catarina (SC), Brasil.

<sup>5</sup> Coordenador do Instituto Nacional para Convergência Digital, Universidade Federal de Santa Catarina – UFSC, Santa Catarina (SC), Brasil.



## What is the state of the art of computer vision-assisted cytology? A Systematic Literature Review

André Victória Matias<sup>a,\*</sup>, João Gustavo Atkinson Amorim<sup>a</sup>, Luiz Antonio Buschetto Macarini<sup>b</sup>, Allan Cerentini<sup>a</sup>, Alexandre Sherlley Casimiro Onofre<sup>d</sup>, Fabiana Botelho De Miranda Onofre<sup>d</sup>, Felipe Perozzo Daltoé<sup>c</sup>, Marcelo Ricardo Stemmer<sup>b</sup>, Aldo von Wangenheim<sup>e</sup>

<sup>a</sup> Department of Informatics and Statistics, Federal University of Santa Catarina, Florianópolis, Brazil

<sup>b</sup> Automation and Systems Department, Federal University of Santa Catarina, Florianópolis, Brazil

<sup>c</sup> Department of Pathology, Federal University of Santa Catarina, Florianópolis, Brazil

<sup>d</sup> Clinical Analyses Department, Federal University of Santa Catarina, Florianópolis, Brazil

<sup>e</sup> Brazilian Institute for Digital Convergence, Federal University of Santa Catarina, Florianópolis, Brazil

### ARTICLE INFO

#### Keywords:

Cytology  
Segmentation  
Classification  
Deep learning  
Computer vision

### ABSTRACT

Cytology is a low-cost and non-invasive diagnostic procedure employed to support the diagnosis of a broad range of pathologies. Cells are harvested from tissues by aspiration or scraping, and it is still predominantly performed manually by medical or laboratory professionals extensively trained for this purpose. It is a time-consuming and repetitive process where many diagnostic criteria are subjective and vulnerable to human interpretation. Computer Vision technologies, by automatically generating quantitative and objective descriptions of examinations' contents, can help minimize the chances of misdiagnoses and shorten the time required for analysis. To identify the state-of-art of computer vision techniques currently applied to cytology, we conducted a Systematic Literature Review, searching for approaches for the segmentation, detection, quantification, and classification of cells and organelles using computer vision on cytology slides. We analyzed papers published in the last 4 years. The initial search was executed in September 2020 and resulted in 431 articles. After applying the inclusion/exclusion criteria, 157 papers remained, which we analyzed to build a picture of the tendencies and problems present in this research area, highlighting the computer vision methods, staining techniques, evaluation metrics, and the availability of the used datasets and computer code. As a result, we identified that the most used methods in the analyzed works are deep learning-based (70 papers), while fewer works employ classic computer vision only (101 papers). The most recurrent metric used for classification and object detection was the accuracy (33 papers and 5 papers), while for segmentation it was the Dice Similarity Coefficient (38 papers). Regarding staining techniques, Papanicolaou was the most employed one (130 papers), followed by H&E (20 papers) and Feulgen (5 papers). Twelve of the datasets used in the papers are publicly available, with the DTU/Herlev dataset being the most used one. We conclude that there still is a lack of high-quality datasets for many types of stains and most of the works are not mature enough to be applied in a daily clinical diagnostic routine. We also identified a growing tendency towards adopting deep learning-based approaches as the methods of choice.

### 1. Introduction

Cytology is a diagnostic technique where cells are harvested from tissues by aspiration or scraping, prepared with different staining techniques and examined on the microscope. Since first described by Dr. Alfred François Donné in 1837 when he discovered *Trichomonas*

*vaginalis*, this technique has been used to support the diagnosis of a broad range of pathologies. It became popular in 1941 when Dr. Georgios Papanicolaou proved cytology was useful in the diagnostic of malignant cells of vaginal smears of the uterus (Diamantis and Magiorkinis, 2013).

Cytology has been successfully applied to the diagnostic of some of

\* Corresponding author.

E-mail addresses: [andre.v.matias@posgrad.ufsc.br](mailto:andre.v.matias@posgrad.ufsc.br) (A. Victória Matias), [joao.atkinson@posgrad.ufsc.br](mailto:joao.atkinson@posgrad.ufsc.br) (J.G. Atkinson Amorim), [luiz.buschetto@posgrad.ufsc.br](mailto:luiz.buschetto@posgrad.ufsc.br) (L.A. Buschetto Macarini), [allan.c@posgrad.ufsc.br](mailto:allan.c@posgrad.ufsc.br) (A. Cerentini), [alexandre.onofre@ufsc.br](mailto:alexandre.onofre@ufsc.br) (A.S. Casimiro Onofre), [fabiana.onofre@ufsc.br](mailto:fabiana.onofre@ufsc.br) (F.B. De Miranda Onofre), [felipe.daltoe@ufsc.br](mailto:felipe.daltoe@ufsc.br) (F.P. Daltoé), [marcelo.stemmer@ufsc.br](mailto:marcelo.stemmer@ufsc.br) (M.R. Stemmer), [aldo.vw@ufsc.br](mailto:aldo.vw@ufsc.br) (A. von Wangenheim).

<https://doi.org/10.1016/j.compmedimag.2021.101934>

Received 1 October 2020; Received in revised form 16 April 2021; Accepted 4 May 2021

Available online 15 May 2021

0895-6111/© 2021 Elsevier Ltd. All rights reserved.

## APÊNDICE B – TABELA DETALHANDO METODOLOGIA DE TODOS OS EXPERIMENTOS

A tabela abaixo detalha a metodologia de todos os experimentos realizados neste trabalho. Os experimentos estão ordenados conforme suas abordagens: Segmentação Semântica (SS), Detecção de Objetos (DO) e classificação de Imagens (CI). O método para treinamento *One Cycle policy* (OCP) está descrito em (SMITH, 2018).

| Abor. | Classes     | Otimiz. | Perda | WD     | LR Máximo     | OCP | Modelo                             | Tamanho da Imagem |
|-------|-------------|---------|-------|--------|---------------|-----|------------------------------------|-------------------|
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 18)                  | 1200x1600         |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 18)                  | 600x800           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 18)                  | 300x400           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 50)                  | 600x800           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 50)                  | 300x400           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 34)                  | 1200x1600         |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 34)                  | 600x800           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 34)                  | 300x400           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 18                           | 1200x1600         |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 18                           | 600x800           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 18                           | 300x400           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 48                           | 1200x1600         |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 48                           | 600x800           |
| SS    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 48                           | 300x400           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 34)                  | 1200x1600         |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 34)                  | 600x800           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 34)                  | 300x400           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 18)                  | 1200x1600         |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 18)                  | 600x800           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 18)                  | 300x400           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 50)                  | 600x800           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | U-Net (ResNet 50)                  | 300x400           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 18                           | 1200x1600         |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 18                           | 600x800           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 18                           | 300x400           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 48                           | 1200x1600         |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 48                           | 600x800           |
| SS    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | HRNet 48                           | 300x400           |
| DO    | Binário     | SGD     | L1    | 0,001  | 0,005         | Não | Faster R-CNN (ResNet 50)           | 1200x1600         |
| DO    | Binário     | SGD     | L1    | 0,001  | 0,005         | Não | Faster R-CNN (ResNet 101)          | 1200x1600         |
| DO    | Binário     | SGD     | L1    | 0,0001 | 0,005         | Não | Mask R-CNN (ResNet 50)             | 1200x1600         |
| DO    | Binário     | SGD     | L1    | 0,0001 | 0,005         | Não | Mask R-CNN (ResNet 101)            | 1200x1600         |
| DO    | Binário     | ADAM    | CE    | 0,001  | 0,0001        | Não | Faster R-CNN (HRNet 18)            | 1200x1600         |
| DO    | Binário     | ADAM    | CE    | 0,001  | 0,0001        | Não | Faster R-CNN (HRNet 48)            | 1200x1600         |
| DO    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | HRNet 18 + CC                      | 1200x1600         |
| DO    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | U-Net (ResNet 50) + CC             | 600x800           |
| DO    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | U-Net (ResNet 34) + CC             | 1200x1600         |
| DO    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | U-Net (ResNet 18) + CC             | 1200x1600         |
| DO    | Binário     | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | HRNet 48 + CC                      | 1200x1600         |
| DO    | Multiclasse | SGD     | L1    | 0,001  | 0,005         | Não | Faster R-CNN (ResNet 50)           | 1200x1600         |
| DO    | Multiclasse | SGD     | L1    | 0,001  | 0,005         | Não | Faster R-CNN (ResNet 101)          | 1200x1600         |
| DO    | Multiclasse | SGD     | L1    | 0,0001 | 0,005         | Não | Mask R-CNN (ResNet 50)             | 1200x1600         |
| DO    | Multiclasse | SGD     | L1    | 0,0001 | 0,005         | Não | Mask R-CNN (ResNet 101)            | 1200x1600         |
| DO    | Multiclasse | ADAM    | L1    | 0,001  | 0,0001        | Não | Faster R-CNN (HRNet 18)            | 1200x1600         |
| DO    | Multiclasse | ADAM    | L1    | 0,001  | 0,0001        | Não | Faster R-CNN (HRNet 48)            | 1200x1600         |
| DO    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | U-Net (ResNet 34) + CC             | 1200x1600         |
| DO    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | U-Net (ResNet 18) + CC             | 1200x1600         |
| DO    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | U-Net (ResNet 50) + CC             | 600x800           |
| DO    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | HRNet 18 + CC                      | 1200x1600         |
| DO    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | HRNet 48 + CC                      | 1200x1600         |
| DO    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Não | Faster R-CNN (HRNet 48) + HRNet 18 | 1200x1600         |
| CI    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | ResNet 18                          | 256x256           |
| CI    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | ResNet 34                          | 256x256           |
| CI    | Multiclasse | ADAM    | CE    | 0,001  | (SMITH, 2015) | Sim | ResNet 50                          | 256x256           |
| CI    | Multiclasse | ADAM    | CE    | 0,0001 | 0.05          | Sim | HRNet 18                           | 256x256           |
| CI    | Multiclasse | ADAM    | CE    | 0,0001 | 0.05          | Sim | HRNet 48                           | 256x256           |