

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

FREDERICO MAY DE LIZ

**CONTAGEM E VERIFICAÇÃO DE  
ANOMALIAS EM CÉLULAS DA MUCOSA  
BUCAL UTILIZANDO PROCESSAMENTO  
DE IMAGENS**

Porto Alegre

2020

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

FREDERICO MAY DE LIZ

**CONTAGEM E VERIFICAÇÃO DE ANOMALIAS EM  
CÉLULAS DA MUCOSA BUCAL UTILIZANDO  
PROCESSAMENTO DE IMAGENS**

Projeto de Diplomação apresentado ao Departamento de Engenharia Elétrica da Escola de Engenharia da Universidade Federal do Rio Grande do Sul, como requisito parcial para Graduação em Engenharia Elétrica

Orientador: Prof. Dr. Ronaldo Husemann

Porto Alegre

2020

FREDERICO MAY DE LIZ

**CONTAGEM E VERIFICAÇÃO DE ANOMALIAS EM  
CÉLULAS DA MUCOSA BUCAL UTILIZANDO  
PROCESSAMENTO DE IMAGENS**

Projeto de Diplomação apresentado ao Departamento de Engenharia Elétrica da Escola de Engenharia da Universidade Federal do Rio Grande do Sul, como requisito parcial para Graduação em Engenharia Elétrica

---

**Prof. Dr. Ronaldo Husemann**  
Orientador - UFRGS

BANCA EXAMINADORA

---

**Prof. Dr. Altamiro Amadeu Susin**  
UFRGS

---

**Prof. Dr. Luiz Fernando Ferreira**  
UFRGS

*Dedico este trabalho à minha família,  
pois sempre estive comigo apoiando incondicionalmente em todos os  
momentos difíceis da minha trajetória acadêmica,  
permitindo que eu chegasse até aqui.*

# Agradecimentos

Agradeço primeiramente aos meus pais pela sua presença e amor incondicional na minha vida sempre. Este trabalho é a prova de que os esforços deles pela minha educação não foram em vão e valeram a pena.

Sou grato pela confiança depositada na minha proposta de projeto pelo meu professor Ronaldo Husemann, orientador do meu trabalho. Obrigado por me manter motivado durante todo o processo.

Agradeço também ao meu colega e amigo Lucas Royes por sempre me apoiar com novas ideias, contribuindo para a conclusão deste trabalho.

*A coisa boa que a ciência tem é que é verdade,  
quer você acredite ou não.*

Neil deGrasse Tyson

# Resumo

O processo de formação de um câncer é determinado por uma lesão em nível do DNA celular, formando um processo denominado carcinogênese. O Teste de Micronúcleos pode ser utilizado para analisar esses processos em qualquer tipo de célula que sofra divisão. Esse teste traduz-se num importante marcador para avaliar o grau de exposição da mucosa bucal a carcinógenos. A princípio, o método de análise desse exame, feito por contagem manual de células anômalas de forma visual, demanda muito tempo aos profissionais da área. Este trabalho propõe uma ferramenta de software para realizar a contagem de células normais ou anômalas de forma automática, percorrendo por toda a lâmina capturada pelo microscópio. Os dados foram obtidos através de amostras de células do tecido esfoliativo da mucosa bucal de agricultores. Para as inspeções das imagens foram utilizadas funções de processamento morfológico, eliminando ruído e aspectos indesejados da imagem capturada, e funções de segmentação, salientando as características de interesse da célula e seu interior. Foi também criada uma interface gráfica para facilitar o usuário, sendo possível observar as ocorrências de irregularidades durante a análise das imagens capturadas. Os resultados mostraram que com a utilização do software foi possível localizar as células de cada imagem analisada e classificar conforme suas características, tais como células normais, binucleadas, micronucleadas e kariolíticas. Foi realizada a comparação dos resultados obtidos de diferentes amostras, analisadas tanto de forma digital quanto manual, demonstrando uma taxa de acerto de aproximadamente 90% na tarefa de identificação de células normais e de 80% para identificação das anomalias celulares.

**Palavras-chave:** anomalia; células; processamento de imagem.

# Abstract

The formation of a cancer is determined by an injury at the level of cellular DNA, forming a process called carcinogenesis. The Micronucleus Test can be used to analyze these processes in any type of cell that undergoes division. This test is an important marker to assess the degree of exposure of oral mucosa to carcinogens. The method to analyze this exam, made by manually counting anomalous cells in a visual way, demands a lot of time from professionals. This project proposes a software tool to count normal or anomalous cells automatically, going through the entire slide captured by the microscope. The data were obtained through cell samples from the exfoliative tissue of the buccal mucosa of farmers. Morphological processing functions were used to inspect the images, eliminating noise and unwanted aspects, and segmentation functions, highlighting the cell and its interior. An interface was also created to facilitate users, making it possible to observe the occurrences of cell anomalies during the analysis of the captured images. The results showed that the tool was able to locate the cells of each image analyzed and classify according to their characteristics, such as normal, binucleated, micronucleated and karyolytic cells. The results were obtained from different samples, analyzed both digitally and manually, demonstrating a success rate of approximately 90 % for identifying normal cells and 80 % for cellular anomalies.

**Keywords:** anomaly; cells; image processing.



# Lista de Figuras

Figura 1 – Diagrama de representação dos diversos tipos de anomalias avaliadas no Teste de Micronúcleos. . . . .	17
Figura 2 – Exemplos de células micronucleadas coradas pelo método de Feulgen. . . . .	18
Figura 3 – Exemplos de células com brotamento nuclear coradas pelo método de Feulgen. . . . .	18
Figura 4 – Exemplos de células com ponte núcleo-plasmática ( <i>broken-egg</i> ) coradas pelo método de Feulgen. . . . .	18
Figura 5 – Exemplos de células binucleadas pelo método de Feulgen. . . . .	19
Figura 6 – Exemplos de células com cromatina condensada coradas pelo método de Feulgen. . . . .	19
Figura 7 – Exemplos de células em cariorrexe coradas pelo método de Feulgen. . . . .	19
Figura 8 – Exemplos de células picnóticas coradas pelo método de Feulgen. . . . .	20
Figura 9 – Exemplos de células em cariólise coradas pelo método de Feulgen. . . . .	20
Figura 10 – Exemplo de processo de aquisição de uma imagem digital. . . . .	21
Figura 11 – Representação de uma imagem digital colorida no sistema RGB. . . . .	22
Figura 12 – Imagem RGB (a) e os componentes de seus correspondentes espaços em HSI: matiz (b), saturação (c) e intensidade (d). . . . .	23
Figura 13 – Exemplo de erosão progressiva de uma imagem. . . . .	24
Figura 14 – Exemplo de dilatação em uma imagem. . . . .	25
Figura 15 – Representação de duas abordagens de segmentação. . . . .	26
Figura 16 – Processo de segmentação utilizando algoritmo <i>k-mean clustering</i> com diversos valores de k. . . . .	27
Figura 17 – Processo de segmentação <i>watershed</i> . . . . .	28
Figura 18 – Imagem apresentando os marcadores internos e externos (esquerda) e o resultado da segmentação (direita). . . . .	29
Figura 19 – Operações lógicas em imagens binárias. . . . .	30
Figura 20 – Principais atributos dos objetos independentes na imagem. . . . .	31
Figura 21 – Imagens capturadas para análise. . . . .	32
Figura 22 – Bordas das células aplicadas à imagem capturada. . . . .	33
Figura 23 – Microscópio automatizado controlado pelo programa SAIMOScan. . . . .	35
Figura 24 – Interface gráfica do programa SAIMOScan. . . . .	36
Figura 25 – Máscara H da imagem de entrada. . . . .	38
Figura 26 – Imagem filtrada com reconstrução de abertura e fechamento. . . . .	39
Figura 27 – Imagem limiarizada pelo algoritmo de Otsu. . . . .	40
Figura 28 – Transformada de distância com base na imagem binarizada. . . . .	40
Figura 29 – Bordas encontradas pelo algoritmo de <i>watershed</i> . . . . .	41

Figura 30 – Imagem limiarizada e aplicada ao algoritmo de <i>watershed</i> . . . . .	42
Figura 31 – Resultado da operação da máscara da célula com a imagem de entrada. . . . .	42
Figura 32 – Binarização do núcleo dentro da célula. . . . .	43
Figura 33 – Resultado da imagem processada. . . . .	44
Figura 34 – Interface gráfica do aplicativo. . . . .	45
Figura 35 – Imagem Resultado em células de tonalidade clara. . . . .	48
Figura 36 – Taxa de acerto em células claras considerando os casos de sobreposição celular. . . . .	49
Figura 37 – Taxa de acerto em células claras desconsiderando os casos de sobreposição celular. . . . .	50
Figura 38 – Imagem Resultado em células de tonalidade escura. . . . .	50
Figura 39 – Taxa de acerto em células escuras considerando os casos de sobreposição celular. . . . .	51
Figura 40 – Taxa de acerto em células escuras desconsiderando os casos de sobreposição celular. . . . .	52
Figura 41 – Taxa de acerto total considerando os casos de sobreposição celular. . . . .	52
Figura 42 – Taxa de acerto total desconsiderando os casos de sobreposição celular. . . . .	53
Figura 43 – Exemplos de agrupamentos celulares. . . . .	54
Figura 44 – Exemplos de células muito escuras. . . . .	55
Figura 45 – Exemplos de células pouco pigmentadas. . . . .	55
Figura 46 – Exemplos de células sobrepostas. . . . .	56

# Lista de Tabelas

Tabela 1 – Comparação de técnicas de segmentação e captura de imagens em cada ferramenta. . . . .	36
Tabela 2 – Comparação de tipos de células analisadas em cada ferramenta. . . . .	37
Tabela 3 – Resultados para células claras considerando os casos de sobreposição celular. . . . .	49
Tabela 4 – Resultados para células claras desconsiderando os casos de sobreposição celular. . . . .	49
Tabela 5 – Resultados para células escuras considerando os casos de sobreposição celular. . . . .	51
Tabela 6 – Resultados para células escuras desconsiderando os casos de sobreposição celular. . . . .	51
Tabela 7 – Resultado total considerando os casos de sobreposição celular. . . . .	52
Tabela 8 – Resultado total desconsiderando os casos de sobreposição celular. . . . .	53

# Lista de Abreviaturas e Siglas

BMP	<i>Bitmap</i>
CCD	<i>Charged Coupled Device</i>
CSV	<i>Comma-separated values</i>
DELET	<i>Departamento de Engenharia Elétrica</i>
DNA	<i>Ácido Desoxirribonucleico</i>
HSI	<i>Hue, Saturation and Intensity</i>
HSL	<i>Hue, Saturation and Lightness</i>
HSV	<i>Hue, Saturation and Value</i>
INCA	<i>Instituto Nacional do Câncer</i>
JPEG	<i>Joint Photographics Experts Group</i>
LAPSI	<i>Laboratório de Processamento de Sinais e Imagens</i>
PNG	<i>Portable Network Graphics</i>
RGB	<i>Red, Green and Blue</i>
RNA	<i>Ácido Ribonucleico</i>
UERGS	<i>Universidade Estadual do Rio Grande do Sul</i>
UFRGS	<i>Universidade Federal do Rio Grande do Sul</i>

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>15</b>
<b>2.1</b>	<b>Princípios básicos sobre Exames de Citologia</b>	<b>15</b>
2.1.1	Teste de Micronúcleos como Biomonitoramento	16
<b>2.2</b>	<b>Alterações celulares</b>	<b>16</b>
<b>2.3</b>	<b>Processamento Digital de Imagens</b>	<b>20</b>
2.3.1	Aquisição de imagens	21
2.3.2	Morfologia de imagens	23
2.3.3	Segmentação de imagens	25
2.3.3.1	Método de Otsu	26
2.3.3.2	Algoritmo <i>k-mean clustering</i>	26
2.3.3.3	Algoritmo de <i>Watershed</i>	27
2.3.3.4	Operações Lógicas em Imagens	29
2.3.4	Reconhecimento de imagens	30
<b>2.4</b>	<b>Detecção Automática de Células Micronucleadas</b>	<b>31</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>34</b>
<b>3.1</b>	<b>Organização do Banco de Dados</b>	<b>34</b>
<b>3.2</b>	<b>Solução Proposta</b>	<b>36</b>
3.2.1	Segmentação das células	37
3.2.2	Segmentação do núcleo	42
<b>3.3</b>	<b>Desenvolvimento da Interface</b>	<b>44</b>
<b>3.4</b>	<b>Avaliação do Algoritmo</b>	<b>46</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b>	<b>48</b>
<b>4.1</b>	<b>Resultados</b>	<b>48</b>
4.1.1	Análise com Células Claras	48
4.1.2	Análise com Células Escuras	50
4.1.3	Análise Total	52
<b>4.2</b>	<b>Discussões</b>	<b>53</b>
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>58</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>60</b>

# 1 Introdução

Segundo o Instituto Nacional de Câncer-INCA (2018), a estimativa mundial mostra que em 2012 ocorreram 14,1 milhões de novos casos de câncer e 8,2 milhões de óbitos causados pela doença. Estima-se, para o Brasil, biênio 2018-2019, a ocorrência de 600 mil novos casos de câncer, para cada ano. O processo de carcinogênese, ou seja, a formação do câncer, em geral, não ocorre de forma rápida, podendo levar vários anos para que uma célula cancerosa prolifere e dê origem a um tumor visível. Segundo Andrade *et al.* (2005), o fumo e o uso excessivo de álcool, comuns a inúmeros processos patológicos, têm tido especial interesse na etiopatogenia do câncer de boca e do trato aerodigestivo. O mecanismo que associa esses agentes a um aumento do risco, fundamenta-se em suas composições químicas e na interação com outros fatores. A exposição da mucosa bucal a agentes cancerígenos resulta, inicialmente, em lesões inflamatórias inespecíficas que são reversíveis, caso a agressão seja suspensa. Porém, se os agentes continuam agredindo a mucosa bucal, eles provocam reações que podem levar ao desenvolvimento de alterações celulares chamadas de displasias.

Segundo Barbon *et al.* (2013), o Teste de Micronúcleos se traduz em um importante biomarcador para diversas patologias, tais como os efeitos genotóxicos do tabaco e do álcool, câncer de colo uterino e câncer oral. Através de um método não invasivo, simples e de baixo custo, pode-se monitorar o efeito de carcinógenos a células que sofrem divisão celular, podendo atuar como um exame preventivo. Uma das aplicações desse teste é analisar células esfoliadas da mucosa bucal exposta, sendo capaz de detectar quebra cromossômica e interferências na mitose, fenômenos importantes na carcinogênese.

O micronúcleo é uma estrutura composta por DNA, RNA e proteínas, contida sob uma envoltória nuclear, de tamanho menor do que o núcleo principal e que se forma após o processo de divisão celular. A alta frequência de aparecimento dos micronúcleos em uma amostra de células pode indicar a ocorrência de patologias ainda em seu estágio inicial, possibilitando uma atuação preventiva. O método de determinação da quantidade de ocorrência de micronúcleos em uma amostra é feito através da contagem manual do número total de células e de quantas delas são micronucleadas.

O trabalho de Testa (2019) consistiu na apresentação de uma ferramenta de software para auxiliar no processo de contagem de células micronucleadas através do tratamento computadorizado de imagens microscópicas com lente de aumento de 10x. Com essa técnica foi possível identificar e classificar células normais e micronucleadas. O resultado permitiu uma eficiência de aproximadamente 95% na tarefa de identificação de células e 66% para identificação de micronúcleos.

Entretanto, as poucas imagens tratadas, restringiram a validação do método. Ademais, segundo Tolbert *et al.* (1992), junto ao Teste de Micronúcleos, são também constatados outros tipos de anomalias celulares na detecção que não foram categorizadas pela ferramenta desenvolvida.

A partir desse trabalho, foi dada continuidade com o desenvolvimento de um novo software que permitiu o aperfeiçoamento da detecção de células e suas anomalias. A interface gráfica da ferramenta foi aprimorada para facilitar a visualização da imagem fornecida pelo algoritmo. A interface possui também o recurso de analisar a contagem das células com varredura por todo o diretório selecionado, com a possibilidade de salvar um relatório no formato CSV.

Para melhor identificar os detalhes da célula na varredura de lâminas, foi utilizado o microscópio óptico com lente de aumento de 20x. Com o objetivo do algoritmo se ajustar às diferentes tonalidades das células, foram aplicadas segmentações adaptativas na etapa de processamento. A utilização do algoritmo de reconhecimento de contorno combinado com a segmentação iterativa permitiu a identificação de outras anomalias tais como binucleação, micronucleação e cariólise.

As imagens avaliadas são provenientes de lâminas de amostras de células do tecido esfoliativo da mucosa bucal fornecidas pela Universidade Estadual do Rio Grande do Sul (UERGS). A avaliação consistiu na análise de 122 imagens incluindo 460 células. A análise das células foi realizada manualmente com base em referencial teórico, separando-as em distintas categorias tais como: células claras, escuras, sobrepostas e anomalias celulares.

Os resultados obtidos permitiram uma taxa de acerto de aproximadamente 90% na detecção de células normais, 85% para células binucleadas, 56% para células micronucleadas e 72% para células com cariólise. Aproximadamente 10% das células detectadas pelo algoritmo representam células sobrepostas, classificando de forma errônea essas células identificadas. Ao remover os casos de sobreposições celulares, as taxas de acerto para células binucleadas sobem para 94% e para as micronucleadas sobem para 88%, representando um aumento significativo de 8% e 33% de acerto, respectivamente.

O documento a seguir apresenta o conteúdo teórico o qual servirá como base para o desenvolvimento da pesquisa, bem como os métodos os quais são propostos para realizar o projeto.

## 2 Revisão Bibliográfica

### 2.1 Princípios básicos sobre Exames de Citologia

Segundo Montanari *et al.* (2016), a célula é a unidade estrutural e funcional básica dos seres vivos. As células organizam-se em tecidos, e estes, em órgãos e sistemas. A Biologia celular preocupa-se com a compreensão da célula quanto aos seus constituintes e à sua atividade funcional. A citopatologia é o estudo que analisa as células individualizadas, descamadas, expelidas ou retiradas da superfície de órgãos de diferentes partes do organismo e posteriormente suas alterações morfológicas em casos patológicos. (CAPUTO *et al.*, 2013).

De acordo com Caputo *et al.* (2013), para não haver a perda de conteúdo, a amostra deve passar após a coleta por um procedimento de acondicionamento do material. A natureza da amostra irá definir posteriormente a forma de coleta e preparo do material segundo as etapas da técnica citológica escolhida.

Após a etapa de acondicionamento se sucede para a etapa de coloração, facilitando a visualização da amostra durante e após o processo. Esse procedimento inicialmente requer uma etapa de hidratação, quando as células passam por uma reposição gradual de água através de banhos alcoólicos ou de água destilada. As células hidratadas recebem um corante aquoso para corar núcleos e são novamente desidratadas para depois receber corantes alcoólicos citoplasmáticos, que têm como objetivo distinguir o citoplasma das células, que é a porção gelatinosa no interior da célula que carrega o núcleo e as demais organelas.

Por fim, as células são novamente desidratadas, clarificadas e seladas com fixadores hidrofóbicos, que protegem a amostra de interações com outras moléculas do ambiente de clausura. A qualidade da coloração citológica está diretamente relacionada às características tintoriais dos corantes, à espessura dos esfregaços (camada de matéria orgânica) e à fixação. Esses cuidados devem ser observados para se evitar artefatos e dificuldade de análise do material.

A análise posterior ao processo de preparação das amostras ocorre em um aparelho apropriado, permitindo visualizar as células e seus devidos componentes microscópicos. Ao realizar o processo de preparação da lâmina de forma bem-sucedida, é possível garantir uma ótima visualização e diferenciar cada componente celular, auxiliando no processo de análise de imagens.



### 2.1.1 Teste de Micronúcleos como Biomonitoramento

Segundo Fenech (2002), o Teste de Micronúcleos em linfócitos humanos é um dos métodos mais comumente usados para medir as taxas de dano ao DNA em populações humanas, pois é relativamente mais fácil classificar micronúcleos do que aberrações cromossômicas. Também avalia a capacidade de recuperação do tecido através da eliminação de células alteradas por necrose e apoptose. O teste é um método não invasivo de simples execução, de baixo custo e de fácil interpretação no biomonitoramento.

Para analisar os efeitos da genotoxicidade induzida pelos agrotóxicos, em indivíduos expostos de formas direta e indireta, Godoy (2018) avaliou células epiteliais esfoliadas da mucosa bucal de agricultores, consumidores de alimentos expostos à agrotóxicos e de consumidores de alimentos exclusivamente orgânicos através do Teste de Micronúcleos. Ao realizar essa análise, foi identificado maior frequência de micronúcleos e de células alteradas no grupo dos agricultores e também no grupo dos consumidores de alimentos produzidos de forma convencional, ambos comparados ao o grupo controle.

## 2.2 Alterações celulares

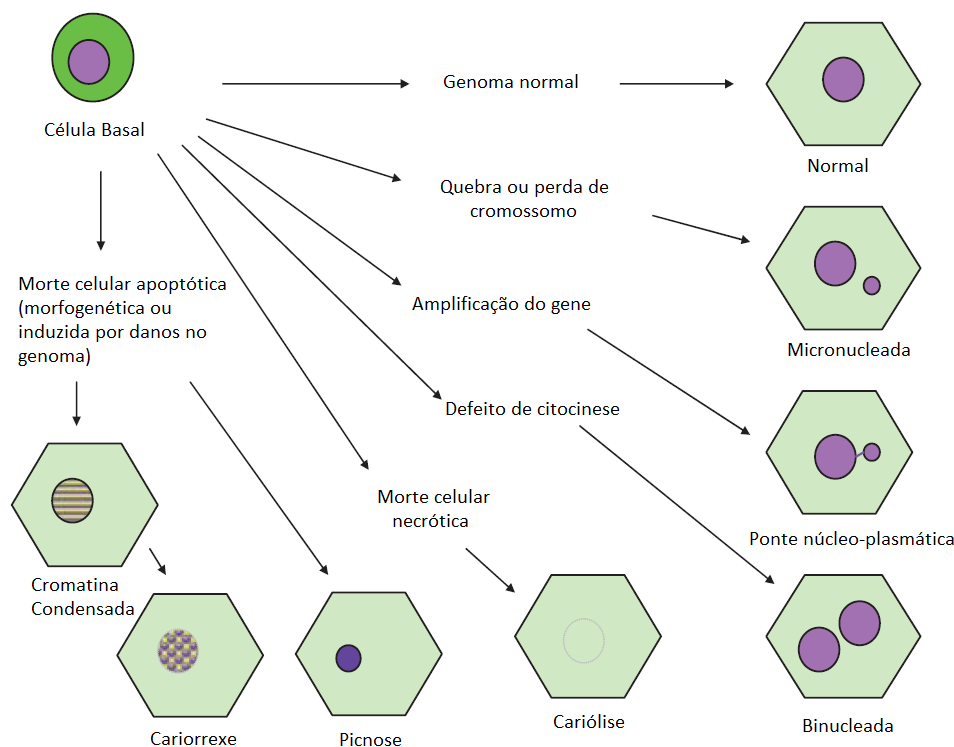
Segundo Thomas *et al.* (2009), a capacidade regenerativa dos tecidos e órgãos do corpo é fundamental para um envelhecimento saudável. A regeneração depende do número e da taxa de divisão das células em proliferação, chamadas de células basais, de sua estabilidade genômica e de sua propensão para a morte celular. Esses eventos podem ser estudados na mucosa bucal, que é um tecido de fácil acesso para coleta de células de forma minimamente invasiva e não causa estresse indevido aos sujeitos do estudo. Este método está sendo cada vez mais usado em estudos epidemiológicos moleculares para investigar o impacto da nutrição, fatores de estilo de vida, exposição à genotoxina e genótipo no dano ao DNA e morte celular.

De acordo com Bonassi *et al.* (2011), abordagens baseadas em biomarcadores têm sido aplicadas na avaliação da exposição de agentes genotóxicos e os aumentos desses biomarcadores são considerados eventos precoces, associados a alterações relacionadas a doenças. Assumindo que os mecanismos para a indução de dano cromossômico são semelhantes em diferentes tecidos, a extensão do dano cromossômico avaliado em linfócitos e outros tecidos substitutos provavelmente reflete o nível de dano em tecidos propensos ao câncer e, por sua vez, o risco de câncer.

A detecção de micronúcleos em citologia esfoliativa deve ser interpretada como resultante da exposição recente a carcinógenos ou o reparo frente a erros espontâneos durante a duplicação do DNA (BLOCHING *et al.*, 2000). Outras alterações nucleares, como podem ser observadas pela Figura 1, podem ser identificadas quando se estuda micronúcleos. O Teste de Micronúcleos atualmente tem sido usado para medir biomarcadores de danos

ao DNA (micronúcleos e/ou botões nucleares), defeitos citocinéticos (células binucleadas), potencial proliferativo (frequência celular basal) e/ou morte celular (cromatina condensada, cariorrexe, células picnóticas e cariólíticas) (THOMAS E FENECH, 2011).

Figura 1 – Diagrama de representação dos diversos tipos de anomalias avaliadas no Teste de Micronúcleos.

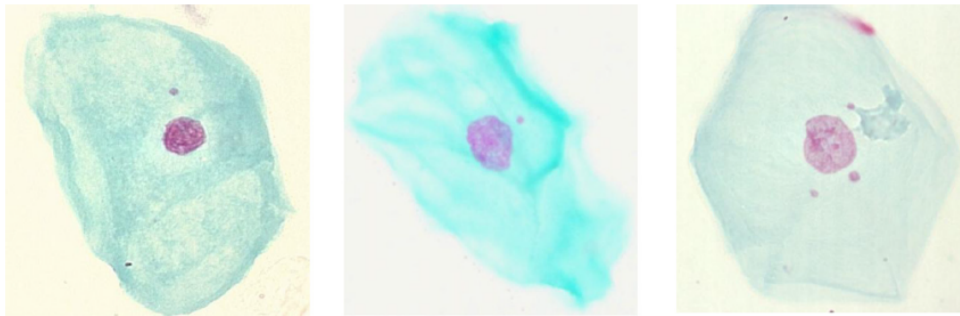


Fonte: THOMAS E FENECH, 2011.

Tolbert *et al.* (1992) propuseram que esses biomarcadores também fossem levados em conta na medida em que representam diferentes fenômenos celulares degenerativos e/ou adaptativos próprios do tecido epitelial. Esses fenômenos sofrem influência de processos inflamatórios, do processo de reparo e até mesmo da intensidade da raspagem. Os critérios de amostragem e padrões morfológicos para cada anomalia celular foram embasados nas recomendações descritas e exemplificadas por Bolognesi *et al.* (2013) conforme os itens a seguir:

- Célula micronucleada: contém o núcleo e fragmentos de núcleo que podem variar de  $\frac{1}{3}$  a  $\frac{1}{6}$  do tamanho original do núcleo (micronúcleos); possuem formato arredondado e coram-se com intensidade semelhante ao núcleo principal (Figura 2);

Figura 2 – Exemplos de células micronucleadas coradas pelo método de Feulgen.



Fonte: BOLOGNESI *et al.*, 2013.

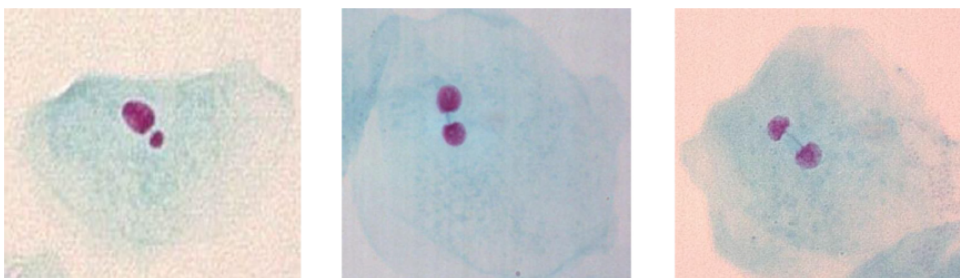
- Célula com brotamento nuclear: o núcleo está em redução, formando um broto; esse broto possui mesma morfologia e coloração do núcleo e seu diâmetro varia de  $\frac{1}{2}$  a  $\frac{1}{4}$  do núcleo original (Figura 3);

Figura 3 – Exemplos de células com brotamento nuclear coradas pelo método de Feulgen.



Fonte: BOLOGNESI *et al.*, 2013.

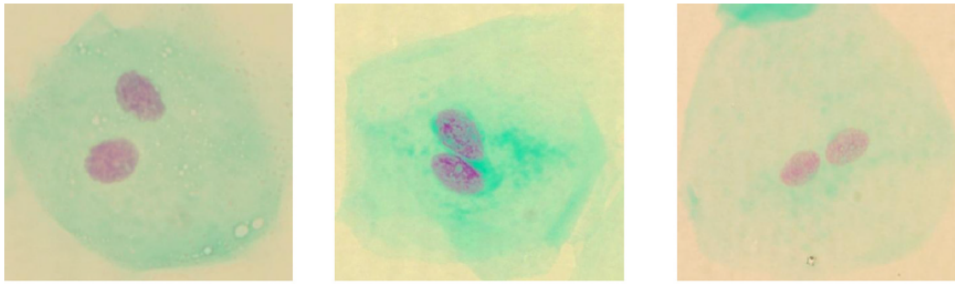
- Célula com ponte núcleo-plasmática (*broken-egg*): varia entre o micronúcleo e o brotamento celular. É caracterizada por um fragmento do núcleo com uma conexão ao núcleo principal (Figura 4);

Figura 4 – Exemplos de células com ponte núcleo-plasmática (*broken-egg*) coradas pelo método de Feulgen.

Fonte: BOLOGNESI *et al.*, 2013.

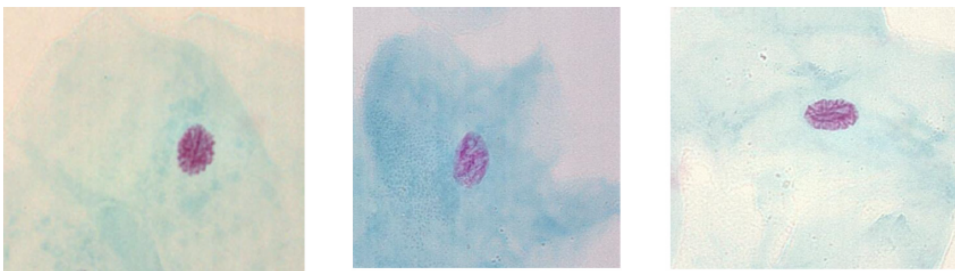
- Célula binucleada: células com dois núcleos principais com semelhança de intensidade e tamanho. (Figura 5);

Figura 5 – Exemplos de células binucleadas pelo método de Feulgen.

Fonte: BOLOGNESI *et al.*, 2013.

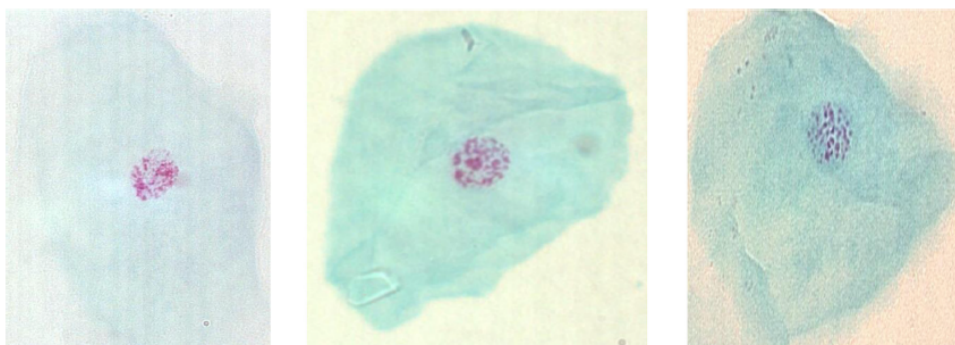
- Célula com cromatina condensada: caracterizado por um núcleo com um padrão estriado devido a áreas paralelas de cromatina condensada que são intensamente coradas. (Figura 6);

Figura 6 – Exemplos de células com cromatina condensada coradas pelo método de Feulgen.

Fonte: BOLOGNESI *et al.*, 2013.

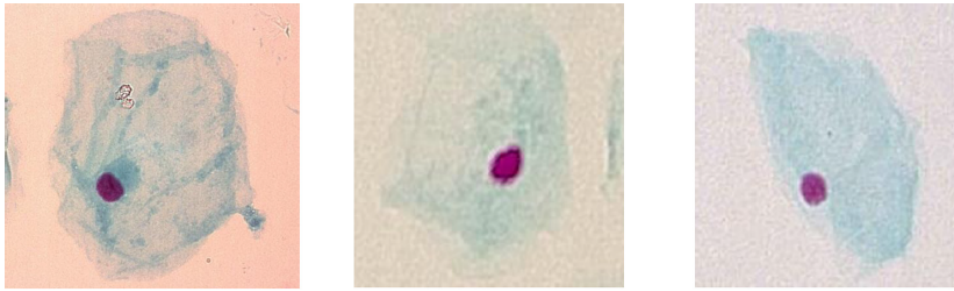
- Célula em cariorrexe: apresenta padrão nuclear heterogêneo em sua coloração, indicando uma espécie de fragmentação nuclear (Figura 7);

Figura 7 – Exemplos de células em cariorrexe coradas pelo método de Feulgen.

Fonte: BOLOGNESI *et al.*, 2013.

- Célula picnótica: Núcleo se aparenta encolhido, com uma alta densidade de material nuclear, manifestando-se em coloração intensa e uniforme com diâmetro nuclear de  $\frac{1}{2}$  a  $\frac{1}{3}$  do núcleo original (Figura 8);

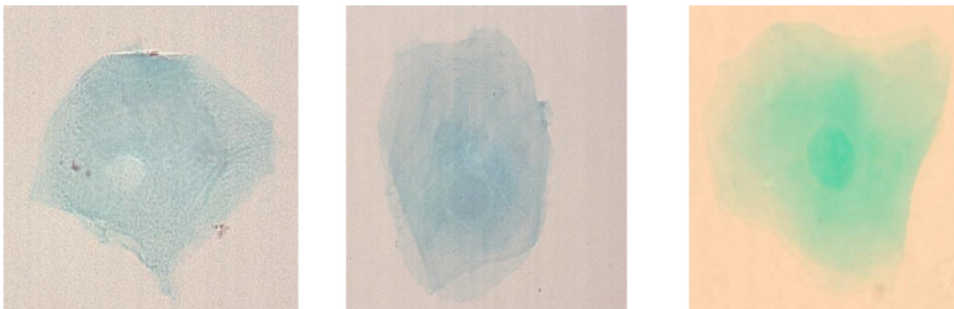
Figura 8 – Exemplos de células picnóticas coradas pelo método de Feulgen.



Fonte: BOLOGNESI *et al.*, 2013.

- Células em cariólise: apresenta-se anucleada, não havendo coloração nuclear (Figura 9);

Figura 9 – Exemplos de células em cariólise coradas pelo método de Feulgen.



Fonte: BOLOGNESI *et al.*, 2013.

## 2.3 Processamento Digital de Imagens

Gonzalez (2010) afirma que uma imagem pode ser definida como uma função bidimensional,  $f(x,y)$ , em que  $x$  e  $y$  são coordenadas espaciais, e a amplitude de  $f$  em qualquer par de coordenadas  $(x,y)$  é chamada de intensidade ou nível de cinza da imagem nesse ponto. Segundo Queiroz e Gomes (2001), o processamento digital de imagens não é uma tarefa simples, na realidade envolve um conjunto de tarefas interconectadas. Tudo se inicia com a captura de uma imagem, a qual, normalmente, corresponde à iluminação que é refletida na superfície dos objetos, realizada através de um sistema de aquisição. Após a captura por um processo de digitalização, uma imagem precisa ser representada de forma apropriada para tratamento computacional.

De acordo com Gonzalez (2010), o processamento de imagens pode ser dividido em diferentes aplicações, tais como aquisição, realce, restauração, compressão, morfologia, segmentação, representação e reconhecimento. Essas são categorias amplas, cada uma engloba um conjunto de métodos específicos para cumprir com um determinado objetivo. A seguir são apresentados as áreas de interesse deste trabalho.

### 2.3.1 Aquisição de imagens

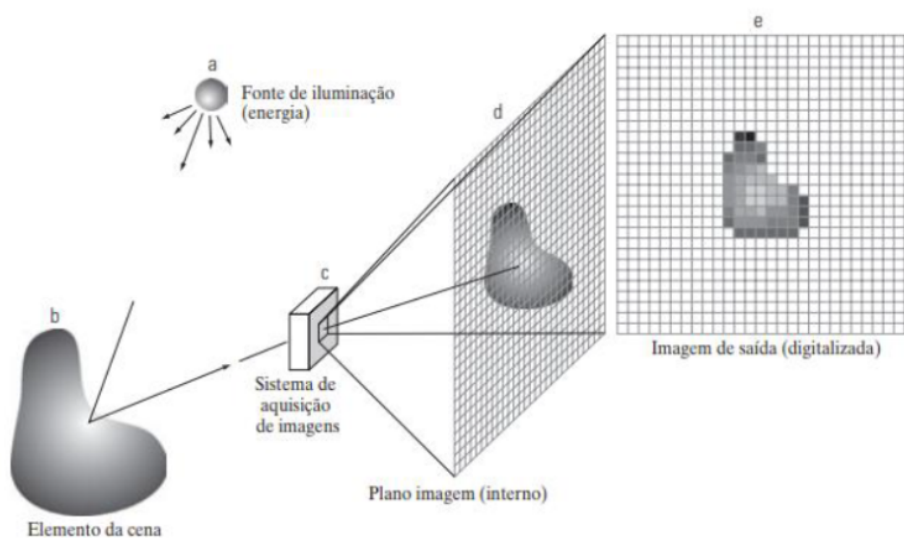
De acordo com Battaiola (2003), as projeções mais consideradas em computação gráfica são as que projetam um sistema de coordenadas 3D em um 2D, realizando a projeção em um plano ao utilizar raios projetores lineares. Esta classe de projeções é conhecida como projeções geométricas planares e podem ser subclassificadas em relação à distância do plano de projeção ao centro de projeção. Se a distância é finita, a projeção é perspectiva, caso for infinita, a projeção é paralela.

Para a aquisição digital de imagens, dois elementos são necessários. O primeiro é um dispositivo físico que deve ser sensível ao espectro de energia eletromagnético, como por exemplo ao espectro de raio-x, luz ultravioleta, visível, ou infravermelha. Este dispositivo transdutor deve produzir em sua saída um sinal elétrico proporcional ao nível de energia percebido. O segundo, chamado digitalizador, é um dispositivo que converte o sinal elétrico analógico produzido na saída do sensor em um sinal digital. (VEGA, 2018)

Gonzalez (2010) afirma que o sensor típico para a utilização em câmeras digitais e outros instrumentos de luz é formado por uma matriz CCD (*charge-coupled device*), que é construída com um arranjo de 4000 ou mais sensores, cada qual possuindo suas propriedades sensoras. A resposta fornecida por cada sensor está relacionada a integral da energia luminosa que a sua superfície recebe.

Pela Figura 10 é possível verificar um exemplo do processo de aquisição. A fonte de iluminação (a) aplica energia luminosa sob uma cena. O elemento da cena (b) reflete a energia luminosa em direção ao sistema de aquisição de imagens (c). A cena é então projetada no plano de imagem (d), fornecendo uma imagem de saída digitalizada (e).

Figura 10 – Exemplo de processo de aquisição de uma imagem digital.



Fonte: GONZALEZ, 2010.

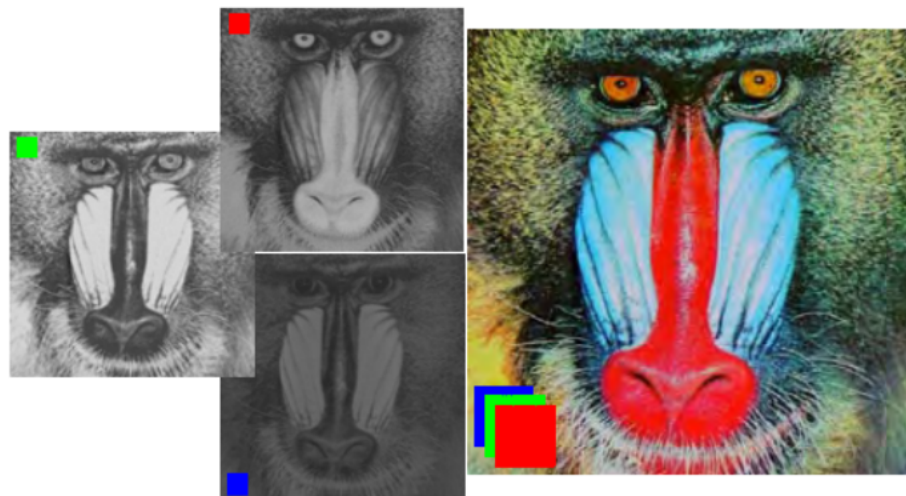
Devido ao fato do processamento ser feito de forma digital, a imagem é discretizada em pequenos elementos, denominados pixels, e seus valores são ordenados em uma matriz bidimensional de coordenadas  $m$  e  $n$ , onde  $m$  indica a linha e  $n$  a coluna onde o pixel se encontra (QUEIROZ e GOMES, 2001).

Segundo Queiroz e Gomes (2001), para imagens digitais coloridas no sistema RGB, um pixel pode ser visto como um vetor cujas componentes representam as intensidades de vermelho, verde e azul de sua cor. A imagem colorida pode ser vista como a composição de três imagens monocromáticas, conforme a Equação 1, na qual  $f_R(x, y)$ ,  $f_G(x, y)$ ,  $f_B(x, y)$  representam, respectivamente, as intensidades luminosas das componentes vermelha, verde e azul da imagem, no ponto  $(x, y)$ .

$$f(x, y) = f_R(x, y) + f_G(x, y) + f_B(x, y) \quad (1)$$

Na Figura 11 são apresentados os planos monocromáticos de uma imagem e o resultado da composição dos três planos. Os mesmos conceitos formulados para uma imagem digital monocromática aplicam-se a cada plano de uma imagem colorida.

Figura 11 – Representação de uma imagem digital colorida no sistema RGB.



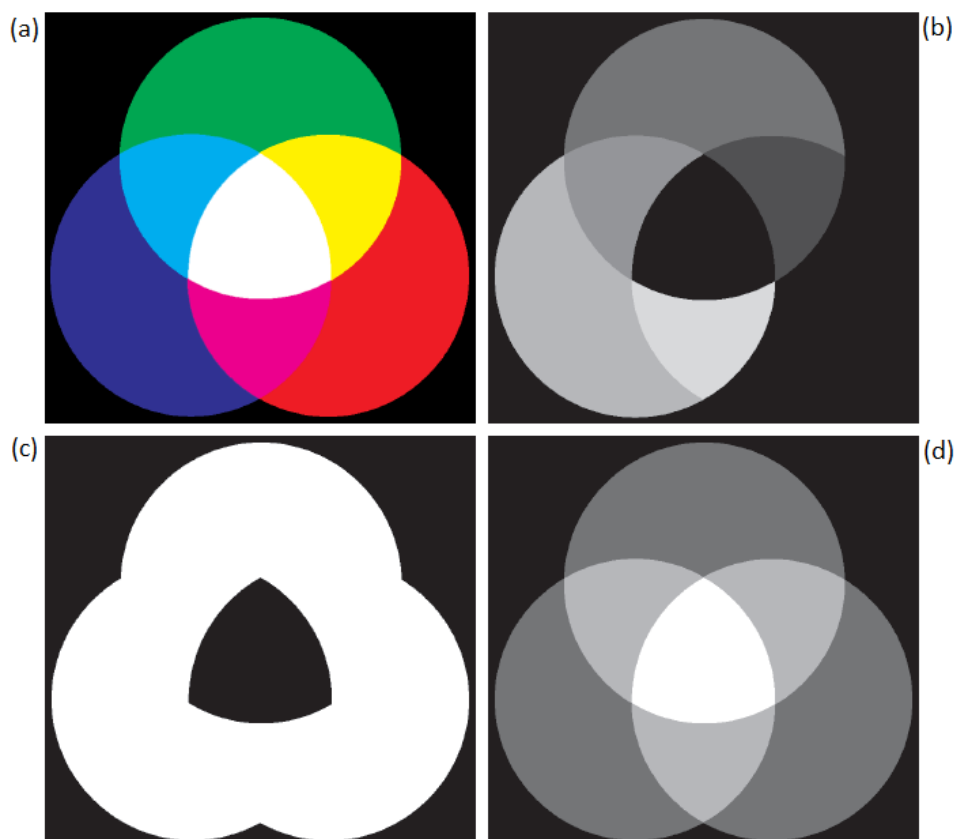
Fonte: QUEIROZ E GOMES, 2001.

Imagens coloridas também podem ser apresentadas em outros espaços de cor como matiz, saturação e brilho, podendo ser observado na Figura 12. Esse método de representação de imagens coloridas é denominado HSI (*Hue, Saturation and Intensity*). Outros sistemas de cores semelhantes são o HSL (L de *luminosity* ou luminosidade) e o HSV (V de *value* ou valor).

O matiz é um atributo de cor que descreve a cor pura, como amarelo, laranja ou vermelho. Já a saturação, como pode ser observada pela mesma figura no campo inferior esquerdo, fornece uma medida do grau em que uma cor pura é diluída pela luz branca. Por

fim, o brilho é um descritor que incorpora a noção de intensidade luminosa, sendo assim um dos principais elementos para análise da imagem (GONZALEZ, 2010). Pela Figura 12 é possível observar as representações de cada canal HSI em relação à uma imagem em RGB.

Figura 12 – Imagem RGB (a) e os componentes de seus correspondentes espaços em HSI: matiz (b), saturação (c) e intensidade (d).



Fonte: GONZALEZ, 2010.

### 2.3.2 Morfologia de imagens

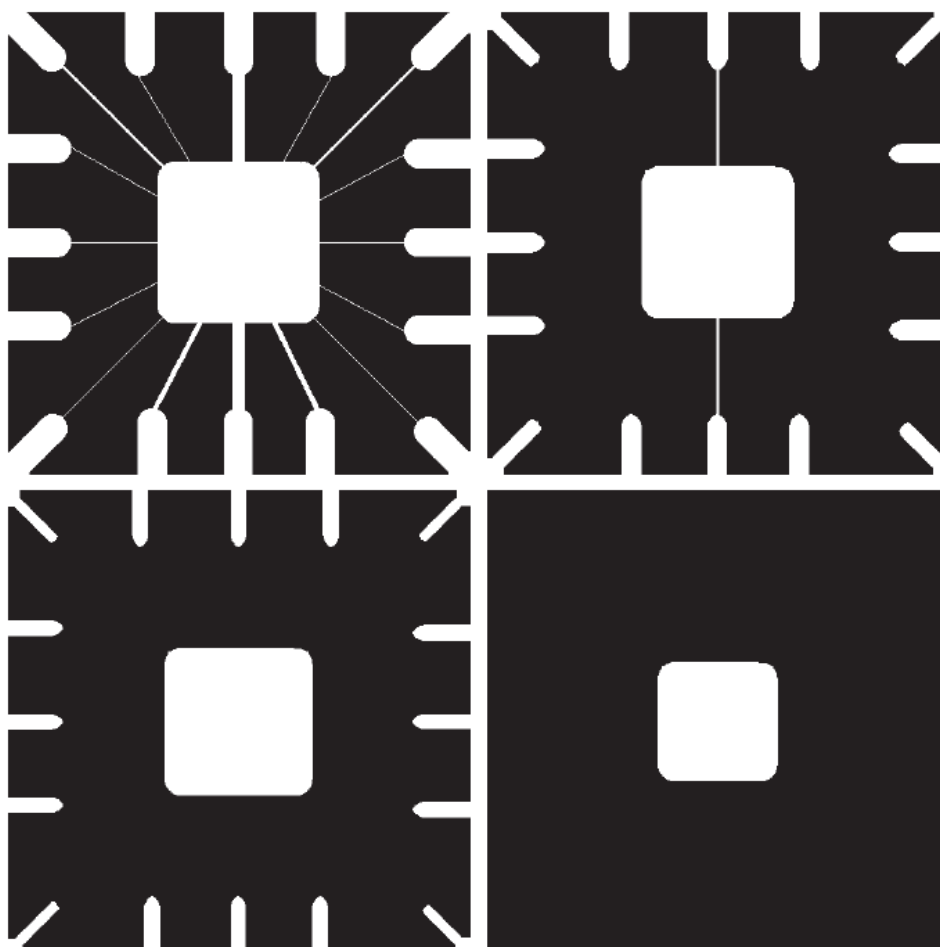
O princípio de morfologia digital se embasa no fato de que a imagem é um conjunto de pixels que formam subconjuntos elementares bi ou tridimensionais. Os subconjuntos e a inter-relação entre eles formam estruturalmente a morfologia da imagem. (QUEIROZ, 2001)

De acordo com Neto (2010), o processo da morfologia matemática se baseia na geometria. A ideia básica é percorrer uma imagem com um elemento estruturante e quantificar a maneira com que este se encaixa ou não na imagem. No caso afirmativo, marca-se o local ou nível de cinza onde o elemento estruturante coube na imagem. Dessa forma, pode-se extrair informações relevantes sobre tamanho e forma de estruturas.



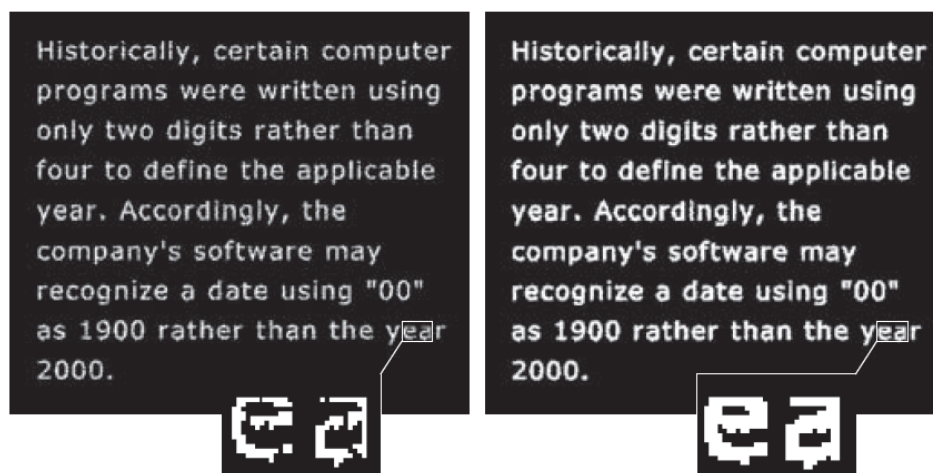
As operações básicas da morfologia digital são a erosão e a dilatação. Na erosão, conforme a Figura 13, são removidos da imagem os pixels que não atendem a um dado padrão. Já para a dilatação, uma pequena área relacionada a um pixel é alterada para um dado padrão, como pode ser observado pela Figura 14. Segundo Fisher *et al.* (2003), o efeito básico do operador em uma imagem binária é aumentar gradualmente os limites das regiões dos pixels do primeiro plano, normalmente os pixels brancos. Portanto, as áreas de pixels do primeiro plano aumentam de tamanho, enquanto os buracos nessas regiões se tornam menores.

Figura 13 – Exemplo de erosão progressiva de uma imagem.



Fonte: GONZALEZ, 2010.

Figura 14 – Exemplo de dilatação em uma imagem.



Fonte: GONZALEZ, 2010.

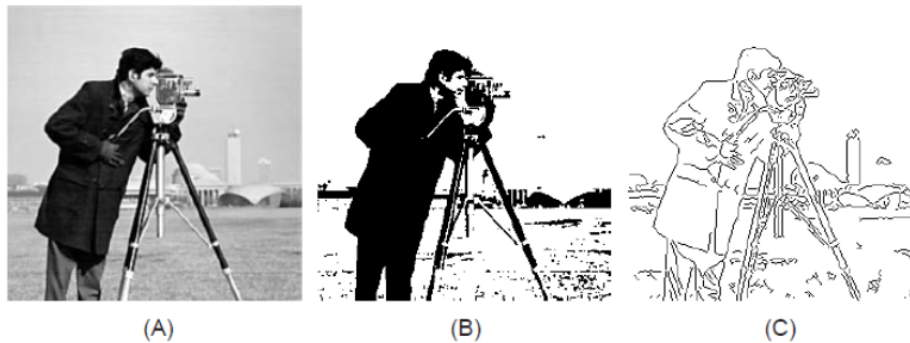
Ainda de acordo com Fisher *et al.* (2003), todos os outros operadores morfológicos básicos são uma combinação das operações básicas de erosão e dilatação. Esses operadores são frequentemente usados para selecionar ou suprimir características de uma determinada forma, como por exemplo remover ruído de imagens ou selecionar objetos com uma direção particular.

### 2.3.3 Segmentação de imagens

Segundo Esquef *et al.* (2003), a segmentação é considerada, dentre todas as etapas do processamento de imagens, a etapa mais crítica do tratamento da informação. É na etapa de segmentação que são definidas as regiões de interesse para processamento e análise posteriores. A segmentação é um processo empírico e adaptativo, procurando sempre se adequar às características particulares de cada tipo de imagem e aos objetivos que se pretende alcançar.

De acordo com Gonzalez (2010), os algoritmos de segmentação para imagens monocromáticas geralmente estão baseados em uma de duas categorias básicas relacionadas às propriedades dos valores de intensidade: descontinuidade e similaridade. Na primeira categoria, o pressuposto é que as fronteiras das regiões são suficientemente diferentes entre si e em relação ao fundo da imagem para permitir a detecção de limite com base nas descontinuidades locais em intensidade. A segmentação baseada nas bordas é a principal abordagem usada nesta categoria. As abordagens de segmentação baseada na região na segunda categoria estão baseadas na divisão de uma imagem em regiões semelhantes de acordo com um conjunto de critérios predefinidos. Na Figura 15 é possível verificar as 2 abordagens para segmentação definidas anteriormente, sendo (a) a imagem original, (b) a imagem segmentada baseada na região e (c) a imagem segmentada baseada nas bordas.

Figura 15 – Representação de duas abordagens de segmentação.



Fonte: ESQUEF *et al.*, 2003.

### 2.3.3.1 Método de Otsu

Para abordagem de segmentação baseada na região pode-se utilizar o método de Otsu. Segundo Torok (2016), a partir de uma imagem em tons de cinza, o objetivo do método de Otsu é determinar o valor ideal de um *threshold* que separe os elementos do fundo e da frente da imagem em dois *clusters* ou agrupamentos, atribuindo a cor branca ou preta para cada um deles. O conceito proposto é de iterar por todos os valores possíveis para o *threshold* em uma imagem, buscando aquele que minimiza a soma da variância entre as classes da imagem. Esse valor corresponderá ao melhor *threshold* para o caso, separando frente e fundo, atribuindo uma cor para cada classe.

### 2.3.3.2 Algoritmo *k-mean clustering*

Em relação à separação de elementos da imagem em *clusters*, uma técnica utilizada para essa abordagem é a *k-mean clustering*. Segundo Gautam e Bhadauria (2014), *k-mean clustering* é o método de algoritmo de aprendizagem não supervisionado para resolver os problemas de agrupamento. O objetivo principal deste algoritmo é minimizar a distância quadrada do objeto ao centro do *cluster*. O algoritmo primeiramente seleciona  $k$  pontos como as centroides iniciais em uma imagem. Formam-se  $k$  grupos, atribuindo cada ponto ao seu centroide mais próximo, que tem distância quadrada mínima. Ao final, se recalcula o centróide de cada agrupamento até que não haja mudança no seu valor. Uma das utilidades desse algoritmo é realizar a clusterização para segmentar em mais de um grupo as cores de uma imagem, como pode ser visto na Figura 16.

Figura 16 – Processo de segmentação utilizando algoritmo *k-mean clustering* com diversos valores de  $k$ .

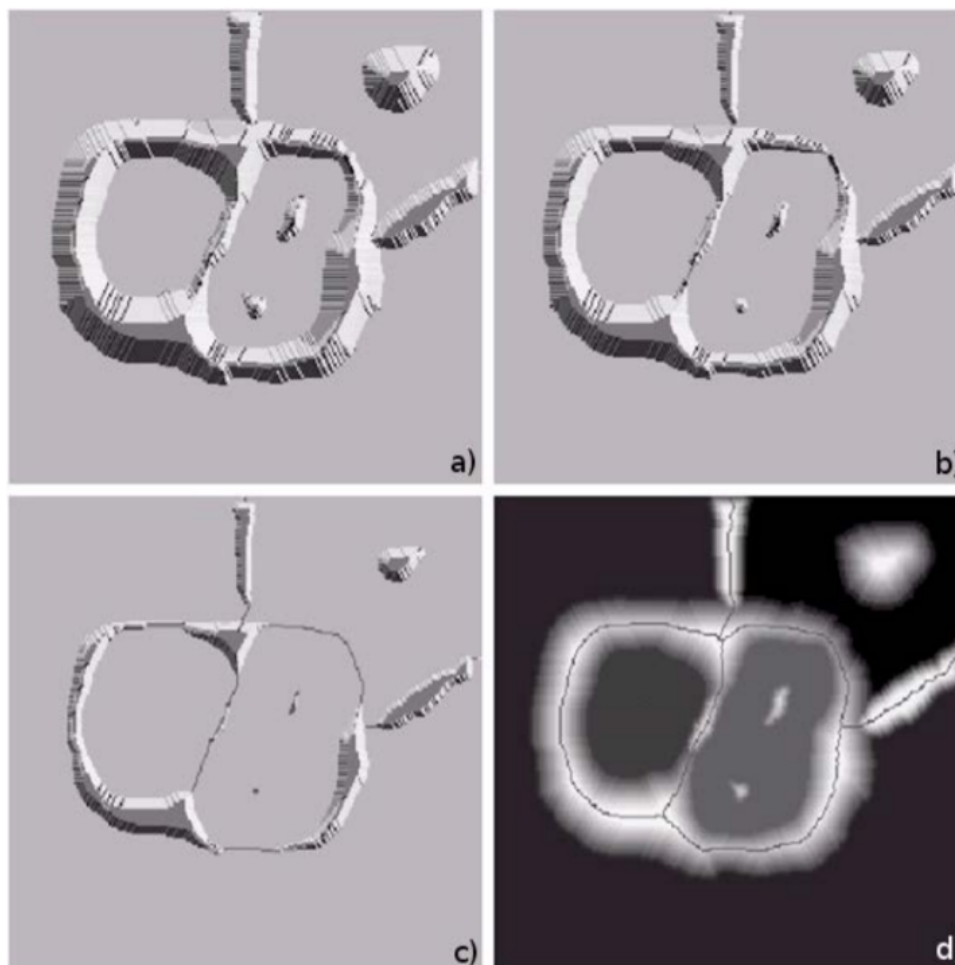


Fonte: OPENCV, 2020.

### 2.3.3.3 Algoritmo de *Watershed*

Para segmentação de bordas em imagens, uma abordagem é o algoritmo de bacias hidrográficas ou *watershed*. De acordo com Shen e Wang (2018), o algoritmo *watershed* é um algoritmo clássico de segmentação de imagens, baseado na teoria topológica em morfologia matemática, tratando a imagem como uma superfície topográfica 3D, distinguindo entre alta e baixa altitude. Diferentes pontos de pixel são classificados em ordem crescente, sendo os pontos de pixel com menor valor de cinza representando a baixa altitude enquanto aqueles com maior valor de cinza corresponderiam à maior altitude.

Segundo Gonzalez (2010), a analogia de imersão é o modelo mais clássico do algoritmo de bacias hidrográficas: a hipótese é perfurar o menor pixel da superfície, sendo perfurado de baixo para cima, a água será derramada na superfície por esses pequenos orifícios de pixel de forma uniforme. Quando a água acumulada em uma *watershed* está prestes a se juntar a outra, uma barragem é criada a fim de impedir a fusão. Esses limites das barragens correspondem às linhas de divisão das *watersheds*. Pela Figura 17 é possível observar o processo do algoritmo, onde (a) demonstra o início da inundação na imagem, (b) é o começo da fusão entre duas *watersheds*, criando uma barragem entre elas que se fortalece em (c), resultando na segmentação em linhas da *watershed* em (d).

Figura 17 – Processo de segmentação *watershed*.

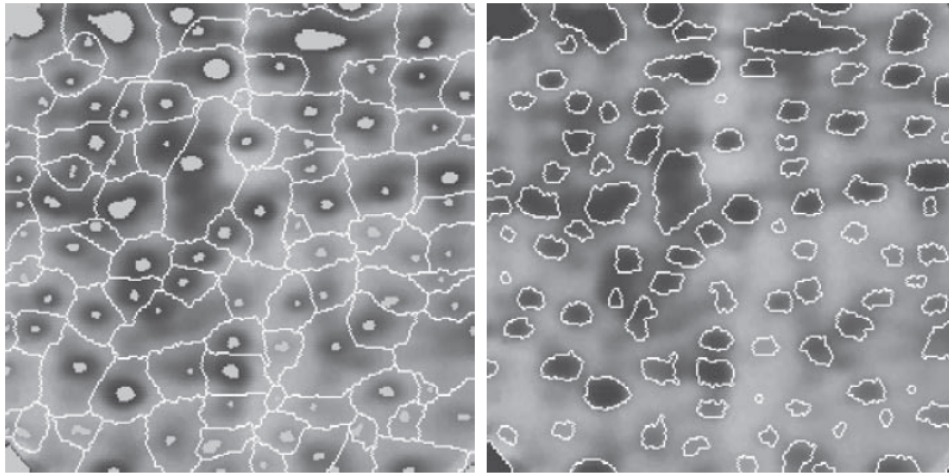
Fonte: GONZALEZ, 2010.

A aplicação direta do algoritmo de *watershed* é sensível a ruído e textura, as desvantagens da segmentação excessiva não conduzem ao processamento e reconhecimento de imagem subsequente (SHEN E WANG, 2018). De acordo com Parvati *et al.* (2008), a segmentação de *watersheds* controlada por marcadores tem se mostrado um método robusto e flexível para segmentação de objetos com contornos fechados e complexos. A imagem do marcador usada para segmentação de bacias hidrográficas é uma imagem binária que consiste em pontos de marcadores únicos ou regiões de marcadores maiores, onde cada marcador conectado é colocado dentro de um objeto de interesse.

Ainda de acordo com Parvati *et al.* (2008), cada marcador inicial tem uma relação de um para um com uma região de *watershed* específica, portanto, o número de marcadores será igual ao número final de regiões de *watershed*. Após a segmentação, os limites das regiões de *watershed* são dispostos nas cristas desejadas, separando cada objeto de seus vizinhos. Os marcadores podem ser selecionados manual ou de forma automática, porém para garantir rápido processamento computacional geralmente empregam marcadores gerados automaticamente para economizar tempo e recursos humanos. Na Figura 18 é

possível observar os marcadores internos e externos gerados para garantir a segmentação correta.

Figura 18 – Imagem apresentando os marcadores internos e externos (esquerda) e o resultado da segmentação (direita).



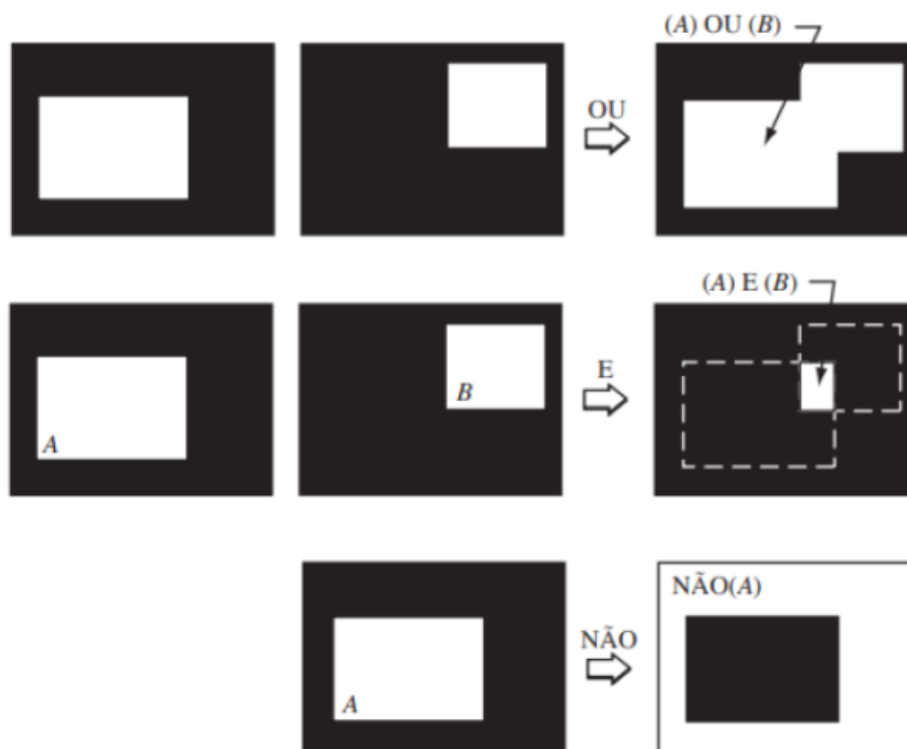
Fonte: GONZALEZ, 2010.

#### 2.3.3.4 Operações Lógicas em Imagens

Queiroz e Gomes (2001) destacam que existem operações básicas que podem ser realizadas como forma auxiliar e secundária. Segundo Gonzalez (2010), trabalhando com imagens binárias, podemos pensar em um grupo de pixels de frente, com valor 1, e outro grupo de pixels de fundo, com valor 0. É possível realizar a interseção, união ou complemento de imagens binárias, ao utilizar operações E (*AND*), OU (*OR*) e NÃO (*NOT*), provenientes da lógica matemática, na qual 1 e 0 equivalem a verdadeiro e falso, respectivamente.

Gonzalez (2010) exemplifica as operações pela Figura 19 considerando duas regiões A e B, ambas compostas por pixels de frente (de valor 1), a operação OU entre essas duas regiões é o conjunto de elementos que pertencem a A ou a B, ou a ambos. A operação E resulta no conjunto de elementos que são comuns às duas regiões A e B. A operação NÃO, aplicada em A, representa todos os elementos que não pertencem ao conjunto de A, sendo esses pixels os elementos de fundo de A.

Figura 19 – Operações lógicas em imagens binárias.



Fonte: GONZALEZ, 2010.

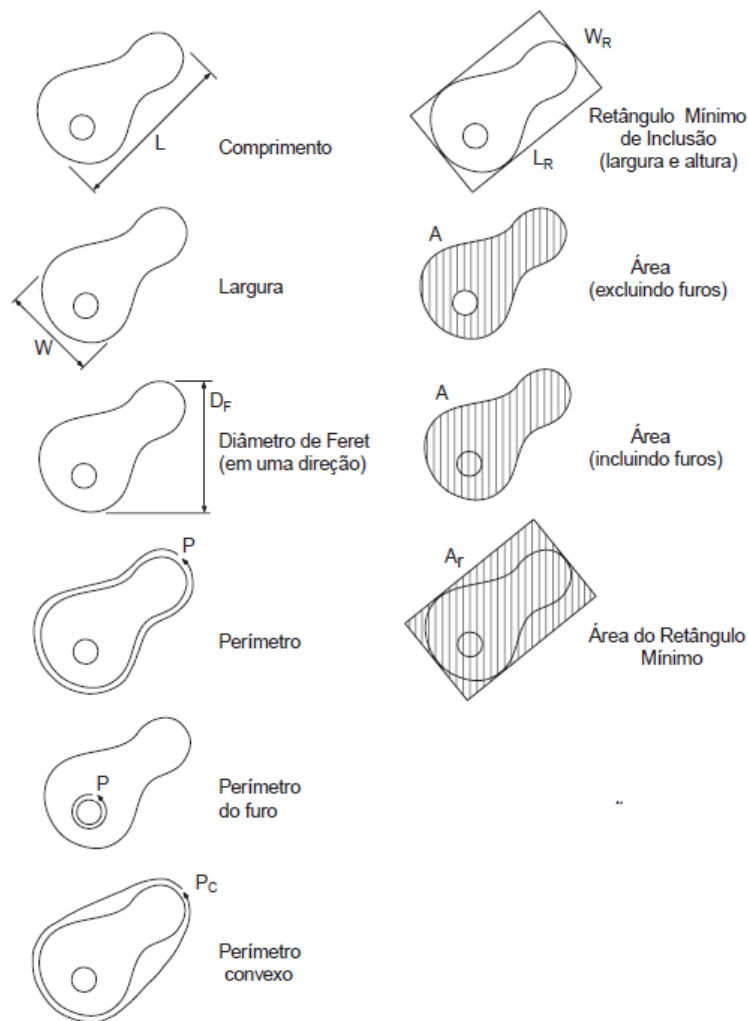
### 2.3.4 Reconhecimento de imagens

A etapa final de um sistema de processamento de imagens é aquela em que se extrai as informações úteis da imagem processada. Segundo Esquef *et al.* (2003), o objetivo do reconhecimento é realizar, de forma automática, a “identificação” dos objetos segmentados na imagem.

Uma das formas de extração útil das informações é o reconhecimento de contornos. Segundo Gong *et al.* (2018), um contorno é definido como uma representação que delimita a forma de um objeto. A detecção de contorno tenta extrair curvas que representam formas de objetos de imagens. O contorno está intimamente relacionado a borda e limite, que correspondem às discontinuidades nas características geométricas e físicas dos objetos nas imagens.

De acordo com Esquef *et al.* (2003), ao definir corretamente o objeto encontrado na imagem, pode-se extrair diferentes atributos de acordo com sua região no espaço, como pode ser observado na Figura 20. A escolha e a qualidade dos parâmetros para construir este espaço é essencial para a realização de um bom processo de reconhecimento de forma.

Figura 20 – Principais atributos dos objetos independentes na imagem.



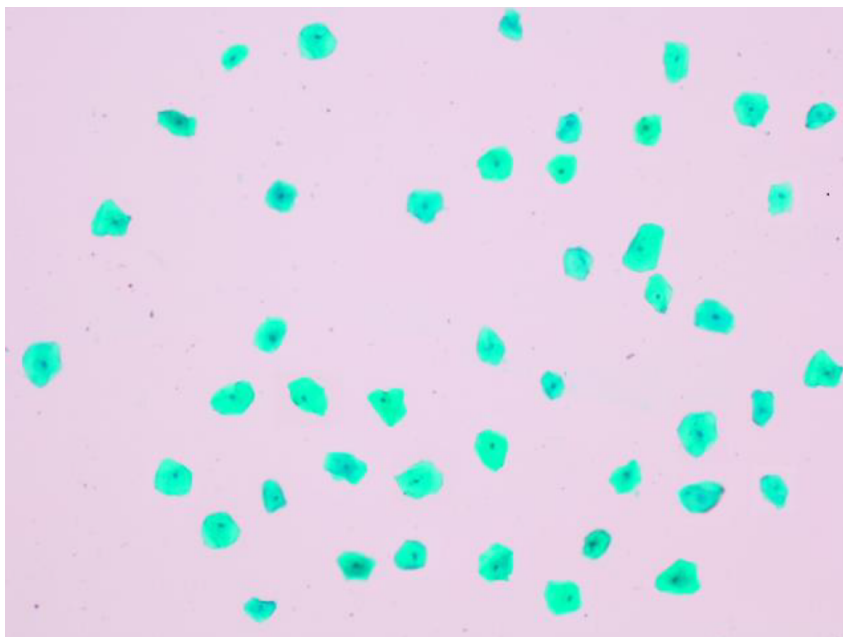
Fonte: Adaptado de ESQUEF *et al.*, 2003.

## 2.4 Detecção Automática de Células Micronucleadas

Em seu trabalho de conclusão de curso, Pedro Testa (2019) apresenta uma ferramenta de detecção de células para automatizar o procedimento de contagem no Teste de Micronúcleos. Para isso, foi utilizado cerca de 12 imagens diferentes para a experimentação do seu algoritmo, cada imagem contendo cerca de 40 células em seu conteúdo. As imagens foram capturadas a partir do microscópio óptico com lente de aumento de 10x, como pode ser observado na Figura 21.



Figura 21 – Imagens capturadas para análise.



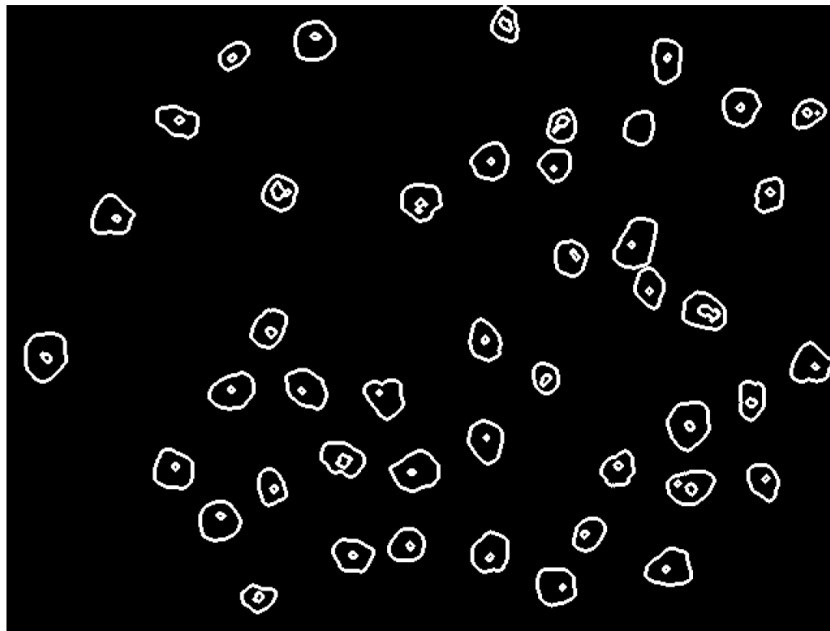
Fonte: TESTA, 2019.

A primeira estratégia de seu trabalho para as amostras capturadas foi desenvolvida usando as matrizes vermelha, verde e azul da imagem (RGB). Porém, devido aos problemas encontrados, optou-se por trabalhar com as matrizes de matiz, saturação e intensidade (HSI), encontrando assim melhores resultados.

Sua solução consistiu em primeiramente aplicar na imagem um filtro morfológico de dilatação e erosão a fim de remover impurezas sem alterar de tamanho. É aplicado uma limiarização fixa, salientando o formato exato das células inspecionadas. Após é realizada a operação de inversão na figura binarizada, tornando pretas as regiões onde antes representava a célula e brancas as regiões consideradas fundo da imagem. Em seguida foi feita uma dilatação, para que a região branca avance exatamente um pixel em tamanho. Por fim, a operação lógica E foi aplicada entre a imagem original do corpo da célula e a imagem invertida e dilatada, resultando na borda do objeto encontrado (TESTA, 2019).

O mesmo método foi aplicado para encontrar os núcleos das células. Testa (2019) considera o fato dos núcleos apresentarem tons mais escuros para definir o novo valor fixo a ser utilizado pela limiarização. Em seguida a operação deve ser aplicada, tornando brancas as regiões mais escuras que o valor escolhido e pretas as regiões mais claras. O resultado final dessa estratégia pode ser observada pela Figura 22.

Figura 22 – Bordas das células aplicadas à imagem capturada.



Fonte: TESTA, 2019.

Testa (2019) desenvolveu também o algoritmo de contorno que recebe a imagem que passou pela rotina de tratamento de imagem e percorre por toda a extensão da imagem, analisando pixel a pixel a procura de objetos e seus núcleos. Para os casos de células micronucleadas, foi parametrizado que o diâmetro do segundo núcleo da célula deve respeitar a proporção de  $\frac{1}{3}$  para ser considerado um micronúcleo. Além disso, para qualquer caso em que o algoritmo encontrasse mais de dois elementos no interior da célula, a mesma seria descartada e não entraria para a contagem geral.

Segundo Testa (2019), foram encontradas limitações no algoritmo, devido ao fato de não ter sido desenvolvido para receber como entrada imagens contendo qualquer informação. Foi realizada a captura de uma área grande de imagem, reduzindo a qualidade das células e impede a observação minuciosa. Ao utilizar imagens contendo células agrupadas ou que apresentem mutações genéticas que vão além do micronúcleo, o programa teve problemas para interpretar essas informações. Além disso, células que apresentavam cor fraca ou forte não puderam ser identificadas pelo programa, pois o tratamento da imagem não foi capaz de manter essas células sem acarretar em problemas para as demais.

Mesmo com essas limitações, o algoritmo desenvolvido por Testa (2019) foi capaz de identificar as características do objetos que compõem a imagem, resultando em uma taxa de acerto de aproximadamente 95% no total de células. Já para a contagem de micronúcleos, a taxa de acerto encontrada foi de 66%. Os resultados de seu trabalho mostram que a aplicação consegue diferenciar casos que se tratam de células normais de casos de células micronucleadas.

## 3 Materiais e Métodos

Como sugerido por Testa (2019), o trabalho propõe a aplicação dos métodos de processamento de imagens para possibilitar a automatização do Teste de Micronúcleos em exames citológicos laboratoriais. Com base nessa estratégia, foi proposto um algoritmo aprimorado para conseguir com maior precisão a leitura das células em uma imagem. Esse software fornecerá auxílio ao profissional responsável por fazer o diagnóstico das amostras.

As imagens do banco de dados foram obtidas a partir de lâminas cedidas pela professora Dra. Jane Marlei Boeira, da Universidade Estadual do Rio Grande do Sul (UERGS). Essas lâminas contêm amostras de células esfoliativas da mucosa bucal de uma pesquisa para avaliar a relação do efeito dos agrotóxicos com o surgimento de alterações celulares.

Essa pesquisa possui três grupos participantes. O primeiro grupo é composto por agricultores que tiveram contato direto com agrotóxicos. O segundo grupo são de pessoas que consomem alimentos que receberam agrotóxicos. Por fim, o último grupo é composto por pessoas que consomem apenas alimentos orgânicos. Cada grupo apresenta uma taxa diferenciada de alterações celulares pelas amostras coletadas.

As amostras foram devidamente tratadas e tingidas de acordo com os processos descritos no capítulo de referencial teórico. Com essa etapa concluída, é possível distinguir o citoplasma e o núcleo de cada célula separadamente. Estes processos de separação das amostras e destaque das células foram realizados pelos alunos da UERGS.

Nas próximas seções serão descritos os métodos de tratamento de imagem considerados para este trabalho juntamente com o desenvolvimento do algoritmo de detecção de células. Ao final, será abordado a forma como os resultados foram testados para a avaliação do software.

### 3.1 Organização do Banco de Dados

Ao todo foram separadas 20 lâminas de cada grupo para serem escaneadas, digitalizadas e, por último, analisadas computacionalmente. A aquisição das imagens foi realizada utilizando o microscópio óptico Nikon Eclipse LV100 (2005) do Laboratório de Processamento de Sinais de Imagens (LaPSI) do Departamento de Engenharia Elétrica (DELET) da UFRGS, como pode ser observado na Figura 23. A câmera utilizada, acoplada ao microscópio, é do modelo Digital Sight DS-Fi1, da fabricante Nikon.

Figura 23 – Microscópio automatizado controlado pelo programa SAIMOScan.

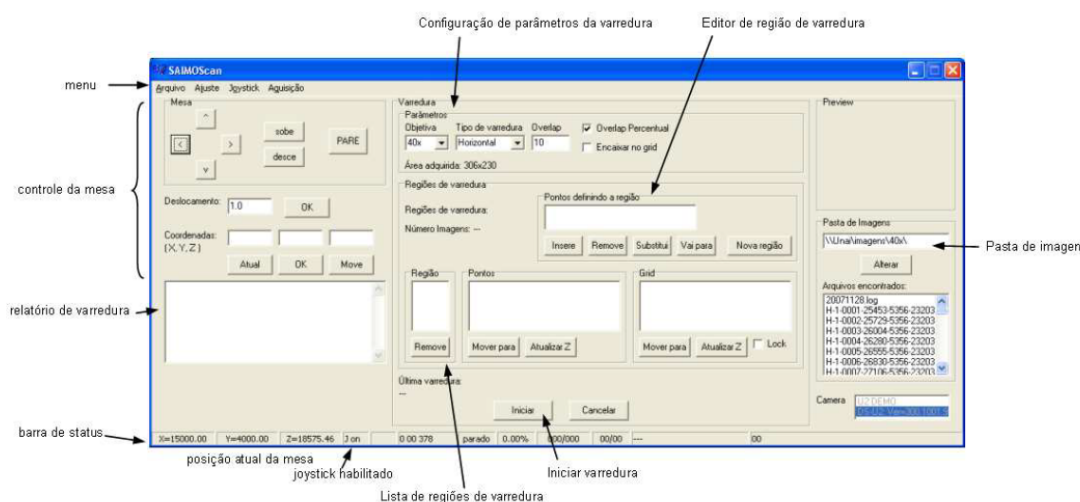


Fonte: NEGREIROS *et al.*, 2008.

O microscópio realiza a varredura da lâmina, delimitada por uma coordenada espacial inicial e uma coordenada final pelo seu software, gerando um mapeamento da mesma. A mesa digitalizadora utilizada para movimentar a lâmina é a OptiScan II, fabricada pela PRIOR. Para digitalizar as lâminas foi utilizado a lente de aumento de 20x.

De acordo com Negreiros *et al.* (2008), o SAIMOScan é o programa criado pelo LaPSI responsável pela digitalização automática de uma lâmina. O programa efetua o posicionamento automático da mesa e captura as imagens em sequência, permitindo sua visualização posteriormente. Ademais, o programa possui uma interface gráfica onde é possível alterar diversos parâmetros da câmera e calibrar a posição e luz da mesa digitalizadora para que a captura seja a melhor possível. A interface gráfica do SAIMOScan pode ser observada na Figura 24.

Figura 24 – Interface gráfica do programa SAIMOScan.



Fonte: NEGREIROS *et al.*, 2008.

Foi feita a calibração manual da câmera acoplada ao microscópio para o programa efetuar a varredura das amostras sem diferenças de coordenadas na sua captura. A fim de evitar a presença de substâncias indesejadas nas imagens, cada lâmina também foi cuidadosamente limpada e conservada.

### 3.2 Solução Proposta

A principal tarefa do software é realizar uma análise por toda a extensão da imagem de entrada e verificar as células existentes. Para cada objeto, o programa deve identificar o seu formato para a sua validação. Caso o formato esteja de acordo com o esperado de uma célula, deve-se então percorrer o interior dela e analisar a presença de núcleos.

A partir do trabalho de Testa (2019), foi dado continuidade com o desenvolvimento de um novo software que permitiu o aperfeiçoamento da detecção de células e suas anomalias. Para aumentar o índice de leitura de células, foi utilizado para a captura lente de aumento de 20x e implementado na imagem o algoritmo de *watershed* para separar cada objeto corretamente. A fim do algoritmo se adequar às diferentes pigmentações de células capturadas pelas amostras, foram aplicadas segmentações adaptativas na etapa de processamento. A Tabela 1 apresenta um comparativo das principais estratégias de segmentação aplicadas e captura das imagens para análise.

Tabela 1 – Comparação de técnicas de segmentação e captura de imagens em cada ferramenta.

	Estratégia de Segmentação		Lente de Aumento Utilizada na Captura de Imagens
	Segmentação Celular	Segmentação Nuclear	
Testa (2019)	Limiarização fixa.	Limiarização fixa.	10x
Este Trabalho	Método de Otsu aplicado ao algoritmo de <i>watershed</i> .	Limiarização iterativa combinado com algoritmo de reconhecimento de contorno.	20x

A utilização do algoritmo de reconhecimento de contorno combinado com a segmentação iterativa permitiu a identificação de outras anomalias tais como binucleação, micronucleação e cariólise. Um comparativo com as detecções de diferentes células pode ser observado na Tabela 2

Tabela 2 – Comparação de tipos de células analisadas em cada ferramenta.

	<b>Tipos de Células Analisadas</b>			
	<b>Binucleadas</b>	<b>Cariólise</b>	<b>Micronucleadas</b>	<b>Normais</b>
<b>Testa (2019)</b>	Não	Não	Sim	Sim
<b>Este Trabalho</b>	Sim	Sim	Sim	Sim

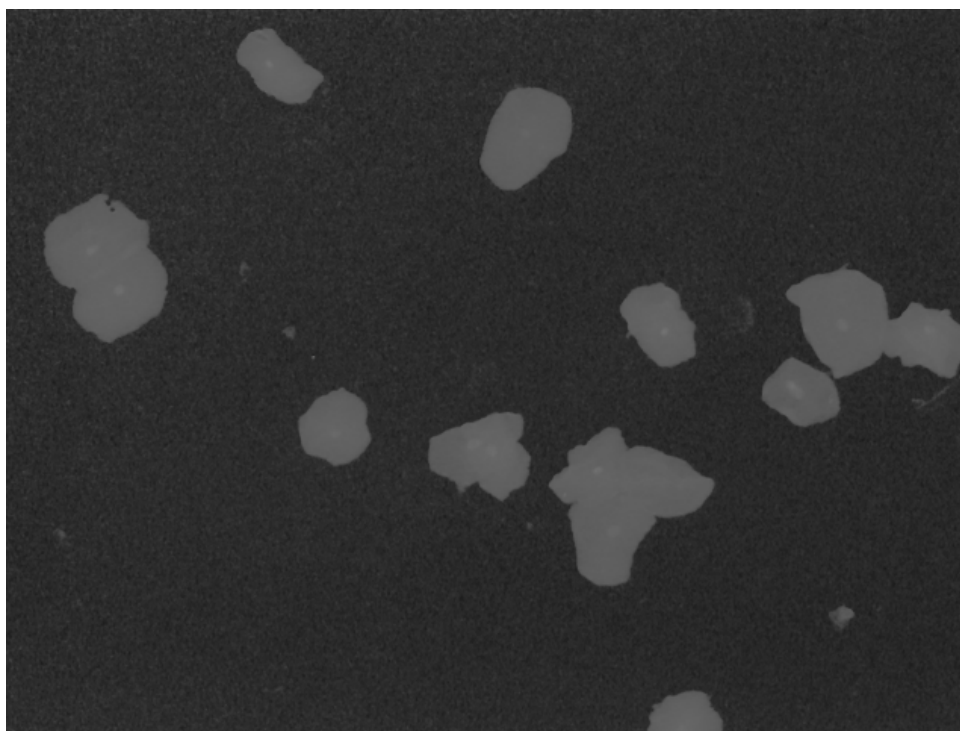
As funções utilizadas para a solução proposta são da biblioteca OpenCV. Essa é uma biblioteca de código aberto que fornece uma infraestrutura de visão computacional para o auxílio no desenvolvimento de aplicações na área de processamento de imagens. A seguir são apresentados as etapas de desenvolvimento da solução deste projeto.

### 3.2.1 Segmentação das células

Para eliminar uma parcela de agrupamento celulares, foi utilizado o algoritmo de *watershed* para separar cada célula de forma mais precisa. Segundo Zhang e Xu (2012), a segmentação de imagens é uma das tarefas fundamentais e importantes na análise de imagens médicas, uma vez que somente com base em uma segmentação bem-sucedida como o *watershed* pode ser extraída uma característica significativa da célula.

A imagem de entrada lida é colorida, representada em formato RGB. Como observado no capítulo de revisão teórica, em um espaço de cor HSV, a matiz é um atributo de cor que descreve a cor pura. Como o canal da matiz possui maior contraste, foi necessário primeiramente converter a imagem de entrada do formato RGB para HSV para melhorar a qualidade do processo de segmentação. A imagem de entrada representada pela sua matiz pode ser observada na Figura 25.

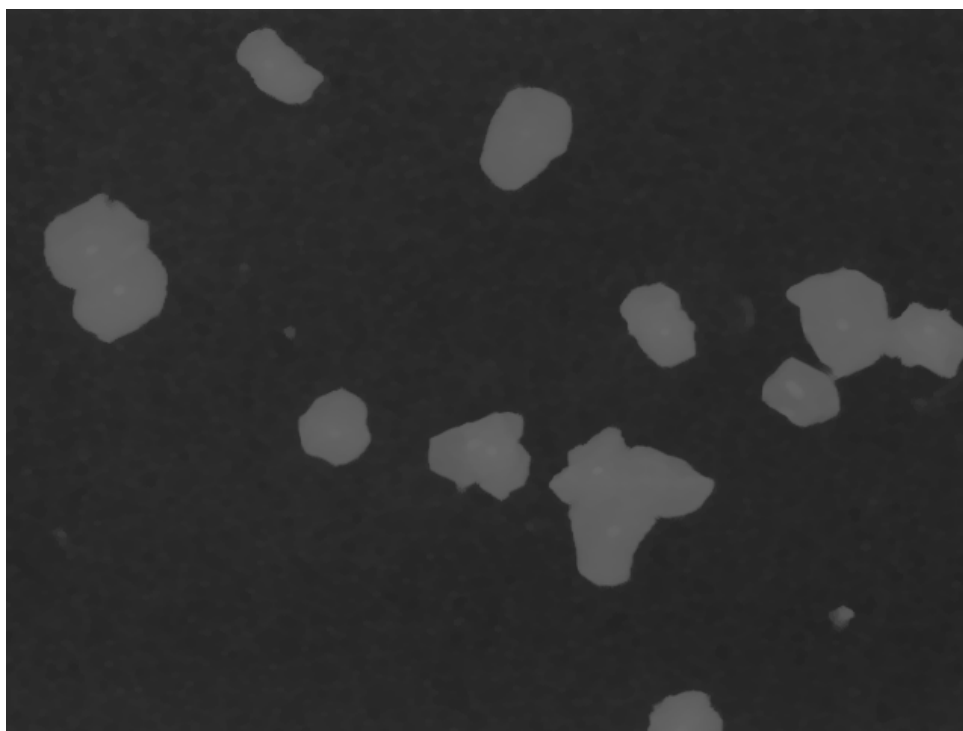
Figura 25 – Máscara H da imagem de entrada.



Foi utilizado o algoritmo de reconstrução morfológica de abertura para remover certas impurezas e objetos irrelevantes para o processo de análise. Essa morfologia é uma erosão seguida de uma dilatação, conforme observado no capítulo de referencial teórico. Em seguida, foi utilizado o algoritmo de fechamento, sendo essa morfologia uma dilatação seguida de uma erosão.

Através da aplicação desses filtros à imagem da célula, eliminamos as rebarbas na célula e suavizamos seus limites, enquanto isso, não alteramos o tamanho e formato das imagens das células (ZHANG E XU, 2012). Com essa abordagem se elimina também alguns artefatos pequenos que ficam na lâmina, tornando mais fácil segmentar a imagem da célula. A Figura 26 ilustra o que foi descrito.

Figura 26 – Imagem filtrada com reconstrução de abertura e fechamento.



Com a imagem filtrada, deve ser implementada a operação de limiarização para separar as células do fundo da imagem. Para contornar os diferentes tipos de pigmentações celulares, foi definido o tipo de limiarização Otsu por ser mais adaptativo para diferentes tons de célula. O resultado obtido deve ser uma imagem binária, conforme é apresentado na Figura 27.

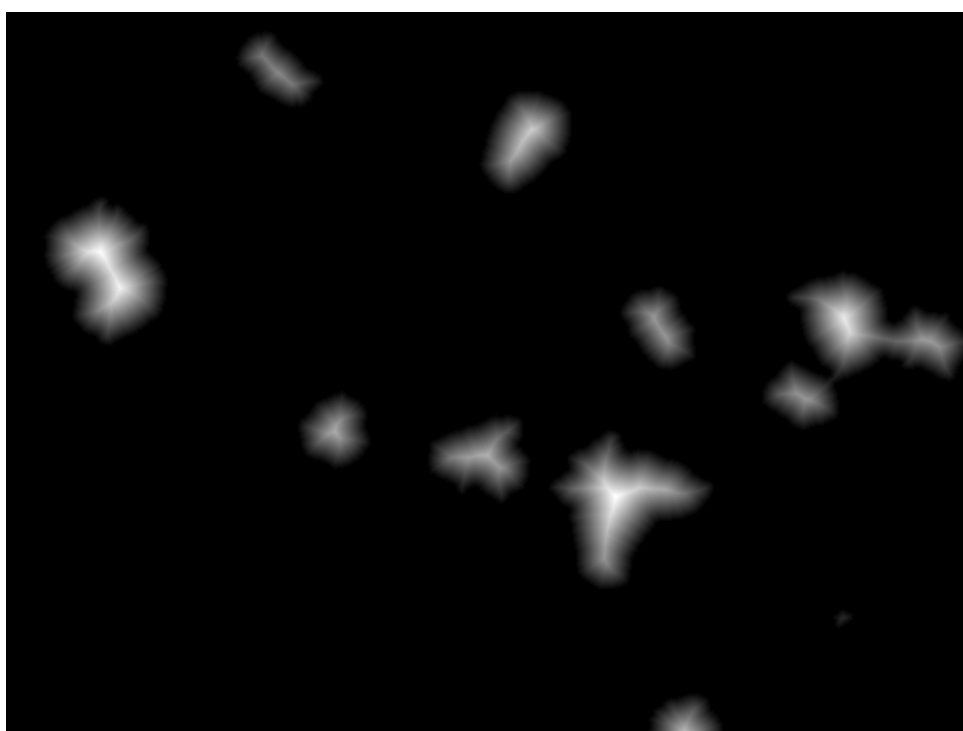


Figura 27 – Imagem limiarizada pelo algoritmo de Otsu.



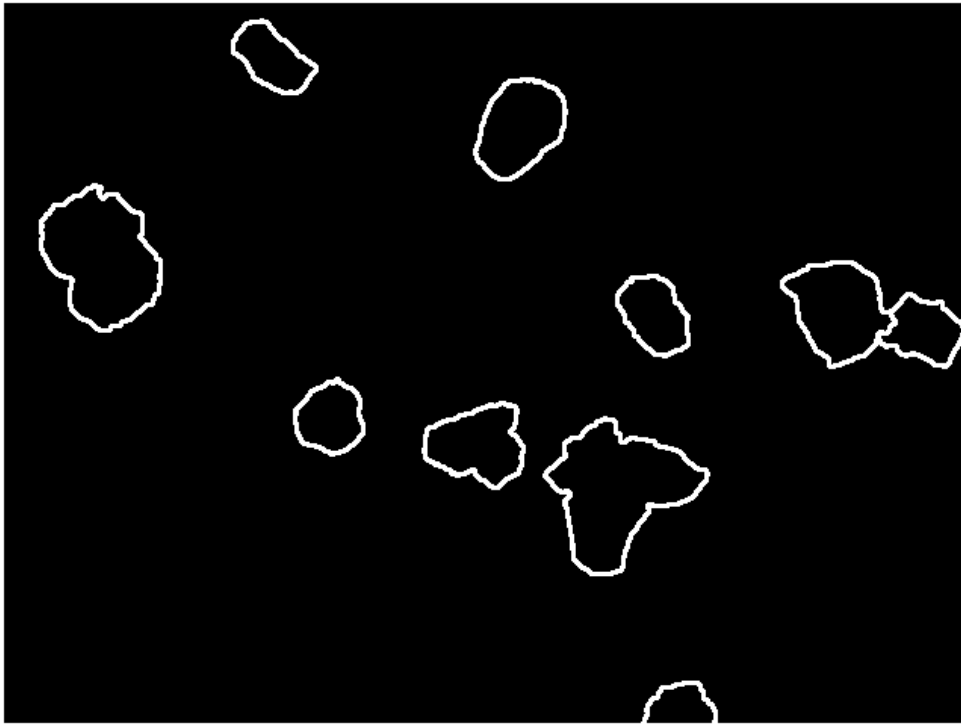
Em seguida é aplicada uma transformada de distância na imagem binarizada, conforme se observa na Figura 28. Nesta transformação, as intensidades do nível de cinza dos pontos dentro das regiões de primeiro plano são alteradas para distanciar suas respectivas distâncias do valor 0 mais próximo, que nesse caso está representando o fundo.

Figura 28 – Transformada de distância com base na imagem binarizada.



A imagem gerada pela transformada de distância é utilizada como marcadores para o algoritmo de *watershed*. Sendo assim, o resultado das bordas geradas pelo *watershed* é observado na Figura 29. Foi aplicada juntamente à essa imagem um algoritmo de dilatação para expandir um pouco mais as linhas de segmentação das células.

Figura 29 – Bordas encontradas pelo algoritmo de *watershed*.



Para detectar cada célula segmentada, deve ser realizada a operação E na imagem de limiarização Otsu juntamente com a imagem das bordas geradas pelo algoritmo de *watershed*, como pode ser observado na Figura 30.

Por fim, foi utilizado o algoritmo de reconhecimento de contorno da biblioteca OpenCV para detectar cada célula. Esse algoritmo retorna uma lista de contornos encontrados dada uma imagem de entrada. Além disso, a biblioteca possui funções específicas que facilitam na extração de características desses contornos. Por conta disso, é realizada uma iteração na lista obtida, se utilizando de funções da biblioteca OpenCV que parametrizam as características da área e perímetro de cada contorno encontrado. Caso o formato verificado não condiz com as condições de uma célula única, esse contorno é ignorado pelo algoritmo.

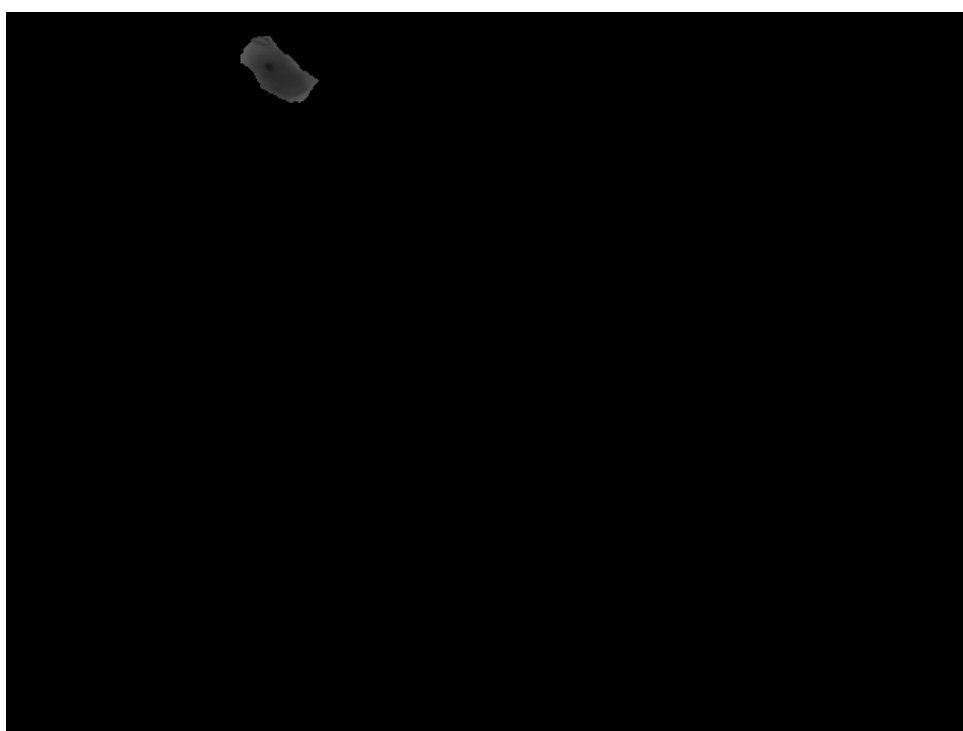
Figura 30 – Imagem limiarizada e aplicada ao algoritmo de *watershed*.



### 3.2.2 Segmentação do núcleo

Com a célula segmentada e encontrada corretamente, é criada uma máscara com seu contorno e preenchido sua parte interna, sendo posteriormente realizada a operação E na imagem da matriz. O resultado da operação pode ser observada na Figura 31.

Figura 31 – Resultado da operação da máscara da célula com a imagem de entrada.

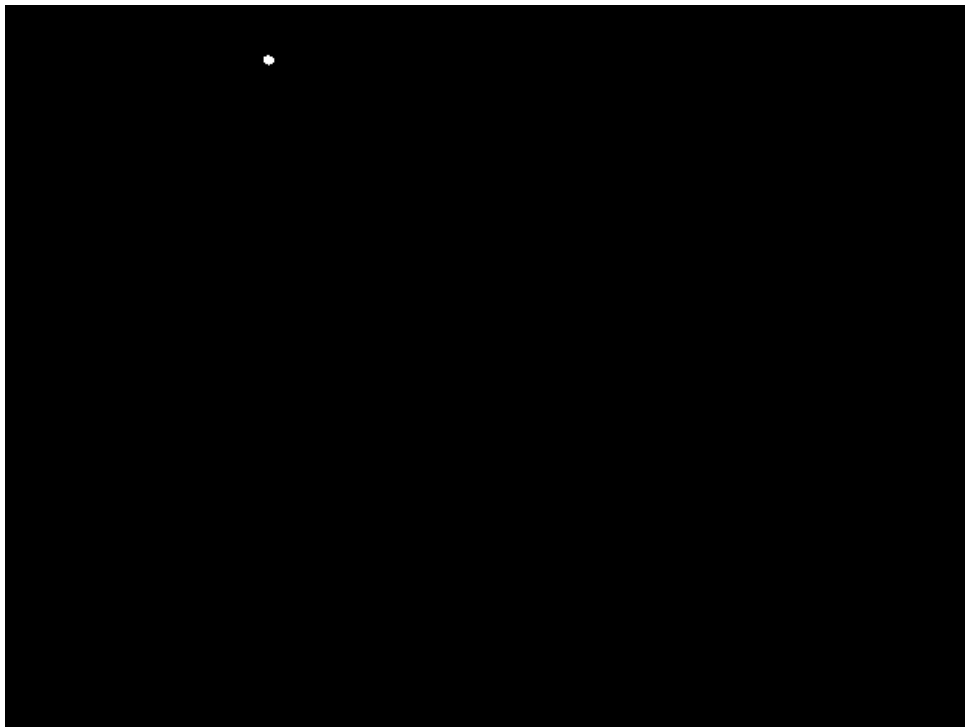


É também realizada a mesma operação da máscara com a imagem no canal azul do RGB. Essa etapa é uma verificação para encontrar a média de intensidade de pixels dentro de cada célula. A intensidade de pixels é importante para as etapas seguintes do processo de segmentação do núcleo. Essa abordagem consegue fazer com que o algoritmo se adapte às diferentes tonalidades de célula que as amostras possuem.

Inicialmente foi pensado em implementar o algoritmo de *k-mean clustering* para agrupar as diferentes tonalidades de cor dentro da célula. Com as regiões estabelecidas, seria possível verificar com o algoritmo de contorno a presença de núcleos e suas possíveis anomalias. Por outro lado, ao implementar o algoritmo, observou-se que o valor de *k* variava muito para células diferentes da mesma imagem. Ademais, ao encontrar o valor de *k* em que era possível verificar a presença do núcleo, esse mesmo valor não encontrava as anomalias dentro da célula necessárias para analisá-la corretamente. Por conta disso, foi necessário alterar a estratégia e utilizar o filtro de limiarização de forma iterativa.

Segundo Phoulady (2016), os núcleos aparecem principalmente como regiões isoladas, redondas e pequenas, com intensidade uniformemente mais escura do que seu citoplasma ou fundo. Por conta disso, foi aplicado uma binarização iterativa com base na intensidade de pixels de cada célula. A binarização da célula da Figura 31 pode ser observada na Figura 32

Figura 32 – Binarização do núcleo dentro da célula.



Para cada binarização é utilizado novamente o algoritmo de contorno da biblioteca OpenCV. O algoritmo retorna uma lista de contornos, sendo parametrizado novamente

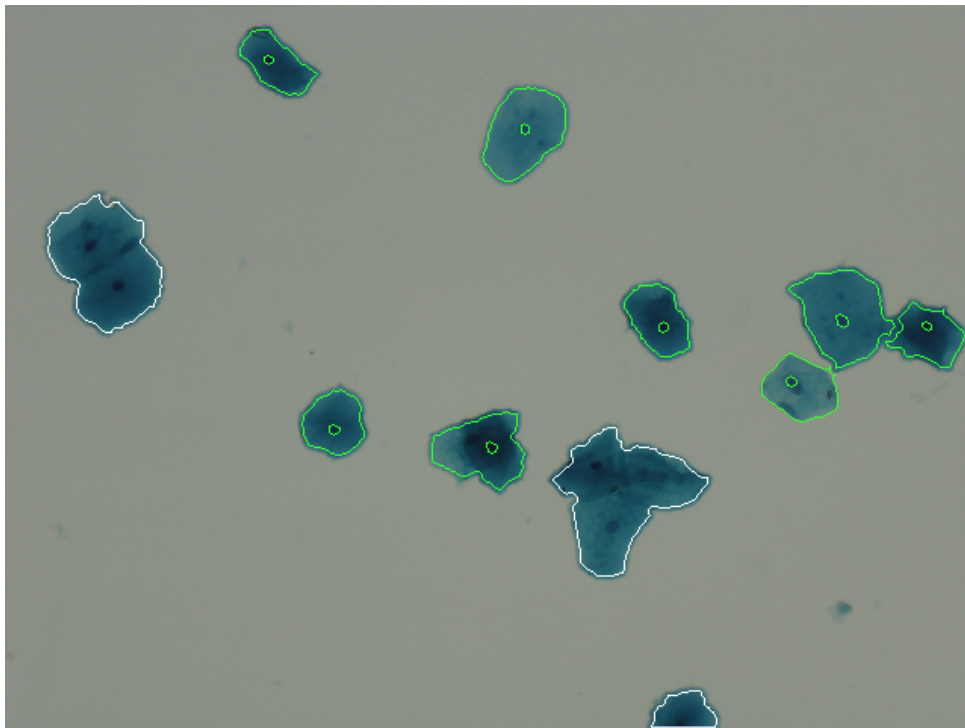
cada contorno utilizando funções de extração de característica de área e perímetro para verificar se dentro da célula existe um possível núcleo dentro dela. Ao verificar a presença desse núcleo, são analisados também os outros contornos caso encontrados.

A célula é então analisada e verificada a sua possível anomalia celular. Caso ocorrer a presença de dois contornos com áreas semelhantes, o algoritmo indicará uma célula binucleada. Se o algoritmo verificar a presença de dois contornos ou mais com diferença de no mínimo  $\frac{1}{3}$  entre eles, o algoritmo apontará como sendo uma célula micronucleada. Por fim, se o algoritmo não encontrar nenhum contorno, a célula é definida como cariólise (célula sem a presença do núcleo).

O processo deve continuar até que o algoritmo percorra por todas as células encontradas na imagem. Ao final da contagem, o software registra as informações obtidas na tela, com a possibilidade de salvar em um arquivo CSV, informando o nome da imagem, a contagem total de células encontradas e suas devidas classificações.

Além disso, deve ser fornecido o resultado visual na interface gráfica, demarcando na imagem original os contornos dos objetos encontrados com cores que denomina cada tipo de anomalia celular correspondente, conforme mostra a Figura 33.

Figura 33 – Resultado da imagem processada.



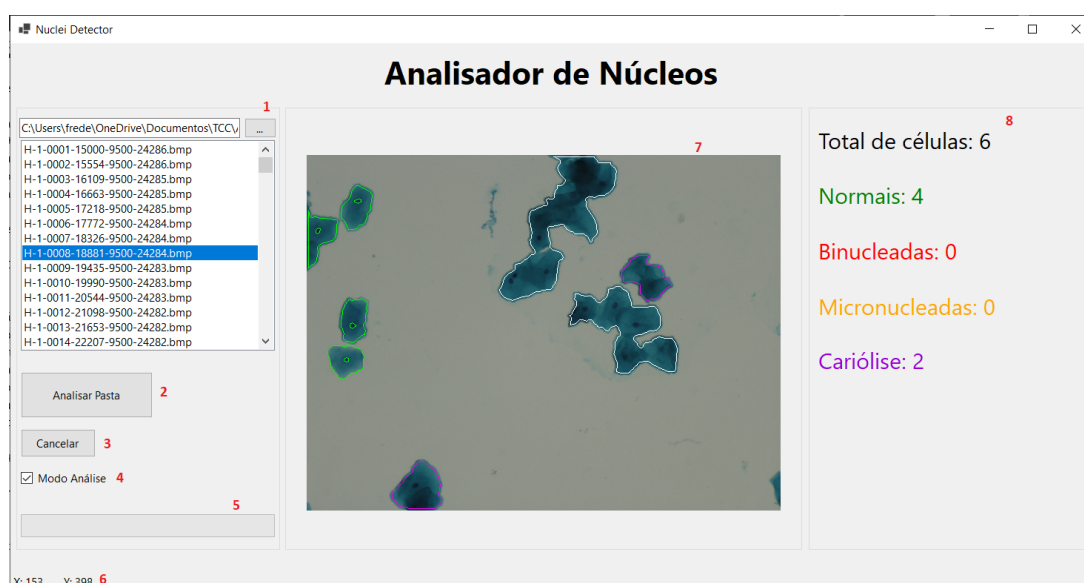
### 3.3 Desenvolvimento da Interface

Com base na interface proposta como parte da solução, foi criado um aplicativo para Windows na plataforma Visual Studio, da Microsoft, utilizando linguagem C#. A

escolha da linguagem de programação permitiu o aprimoramento da aplicação com a possibilidade de utilizar o Windows Forms, plataforma de desenvolvimento de interfaces gráficas, ao mesmo tempo que facilitou a utilização das funções da biblioteca OpenCV.

Para melhorar o manuseio do processo, foram colocados botões na área de exibição do aplicativo que permitem ao usuário selecionar a pasta onde as imagens das amostras se encontram e realizarem tarefas como analisar imagem separadamente, analisar a pasta inteira das imagens, entre outras funções. Na Figura 34 é apresentada a aparência da aplicação desenvolvida, juntamente com sua numeração para ajudar a localizar cada uma de suas funções desempenhadas.

Figura 34 – Interface gráfica do aplicativo.



1. Carregar pasta: A função deste botão é permitir ao usuário escolher a pasta que se encontram as imagens para análise. Ao clicar nele, é possível navegar entre todos os diretórios livremente e escolher qualquer pasta. Ao selecionar uma pasta, aparecerá na caixa abaixo todas as imagens da pasta que respeitam os formatos padrões (bmp, png, jpeg, entre outros).
2. Analisar pasta: Tem a funcionalidade de dar início à execução do algoritmo por toda a pasta selecionada. Ao pressionar o botão, todas as fotos da pasta serão devidamente analisadas. Ao final, o programa salvará um arquivo no formato CSV.
3. Cancelar: A função deste botão é de cancelar a análise em lote que está rodando no momento. Ao apertar esse botão, o sistema salvará aquilo que foi executado até o momento, juntamente com o arquivo CSV criado.
4. Modo Análise: Sua função é ativar a análise da imagem sem realizar uma varredura por toda a pasta selecionada. Assim o usuário poderá analisar o resultado da contagem e sua respectiva imagem resultado separadamente.

5. Barra de progressão: Sua função é indicar ao usuário o tempo restante a aguardar para o sistema realizar a análise completa da pasta.
6. Coordenadas da Imagem: Tem a funcionalidade de divulgar a posição x e y da imagem ao percorrer a imagem com o mouse.
7. Exibir Imagem: Sua função é de apresentar ao usuário tanto a imagem de entrada quanto a imagem resultado quando a função de análise estiver ativada.
8. Resultados: Apresenta ao usuário a contagem das células localizadas e suas devidas análises finais. Cada tipo de anomalia está apresentada com a respectiva cor da análise feita com a imagem resultado, facilitando a verificação do resultado para o usuário.

### 3.4 Avaliação do Algoritmo

A avaliação do algoritmo desenvolvido foi feita através de uma comparação entre a contagem realizada pelo aplicativo e a contagem manual baseada nos exemplos e descrições feitas por Bolognesi *et al.* (2013). Para isso, imagens de três amostras diferentes capturadas pelo microscópio foram analisadas. Visto que a presença de anomalias nucleares se dão mais presentes no grupo de agricultores, todas as imagens contadas manualmente foram feitas para esse grupo.

De acordo com Caputo *et al.* (2013), a qualidade da coloração citológica está diretamente relacionada às características tintoriais dos corantes, à espessura dos esfregaços e à fixação. Esses cuidados devem ser observados para se evitar artefatos e dificuldade de análise do material. Ao capturar as imagens das amostras, foi observado diferentes reações de pigmentação celular em diversos locais da lâmina.

Como a solução implementada tem o objetivo de se adaptar aos diferentes tons de cada célula, foi necessário realizar uma separação de tonalidade celular a fim de validar a estratégia. Para extrair as diferenças de reação pigmentar das células, foram divididas as imagens em dois grupos: células de tonalidade clara e células de tonalidade escura.

Ademais, a análise manual descartou as ocorrências de sobreposição celular, visto que a análise particular de cada célula fica comprometida. Como a análise automática identifica essas células como sendo um único objeto, foi necessário separar essas ocorrências de sobreposição para identificar a influência na taxa de acerto do algoritmo.

Como visto no referencial teórico, a análise de algumas mutações genéticas são mais visíveis se a imagem fotografada estiver com uma lente de maior aumento. Em decorrência disso, para facilitar a contagem e análise correta das anomalias de forma manual, as mutações a serem analisadas foram definidas em: presença de micronúcleos, células binucleadas e cariólise.

Para a presença de micronúcleos, foi parametrizado que o segundo núcleo da célula deve respeitar a proporção de no mínimo  $\frac{1}{3}$  para ser considerado um micronúcleo. Já para células binucleadas, os dois núcleos deveriam possuir aproximadamente o mesmo tamanho. Por fim, a ocorrência de cariólise se dá ao fato da célula não possuir coloração nuclear.



## 4 Resultados e Discussões

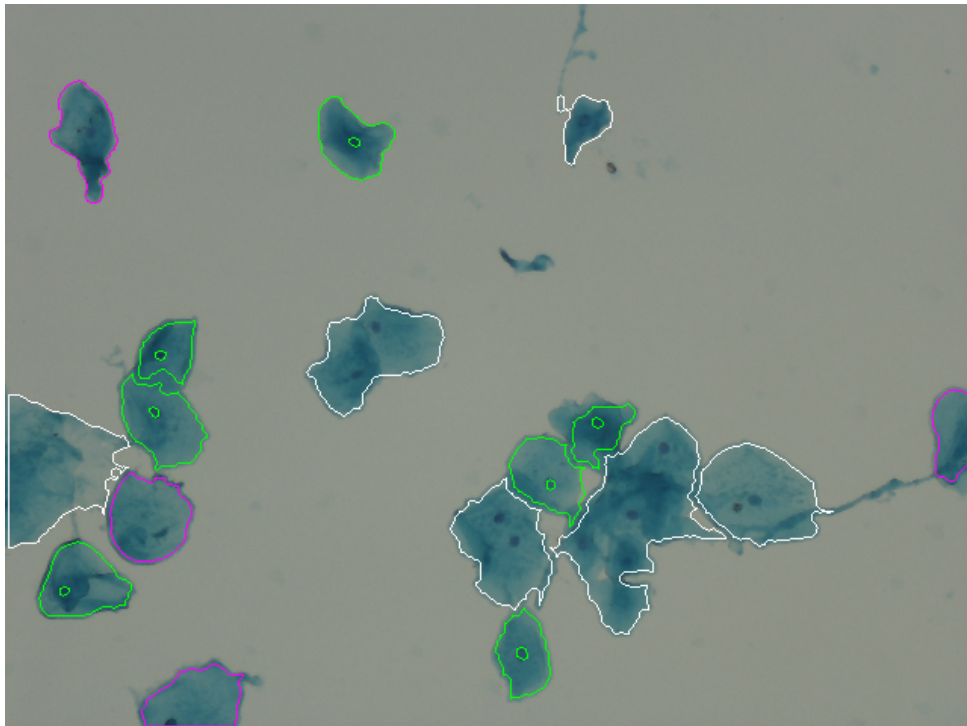
Este capítulo apresenta os resultados obtidos com a aplicação das imagens capturadas das lâminas com amostras de células esfoliativas da mucosa bucal ao algoritmo de análise de células que foi desenvolvido. Para calcular as taxas de acerto de cada reconhecimento, foi determinada a análise manual como sendo a referência dos resultados apresentados. É realizada também a discussão quanto aos resultados alcançados e as limitações da aplicação.

### 4.1 Resultados

#### 4.1.1 Análise com Células Claras

Analizando primeiramente as imagens em que a maior parte das células possuem a tonalidade mais clara, como pode ser observado na Figura 35. As imagens que possuem as células com essa tonalidade conseguem obter um maior contraste entre o citoplasma e o núcleo, tornando a inspeção mais precisa.

Figura 35 – Imagem Resultado em células de tonalidade clara.



Nesta análise, foram utilizadas 77 imagens contendo 279 células no total. Após realizar a análise manual de cada célula, foi comparada ao resultado do algoritmo para

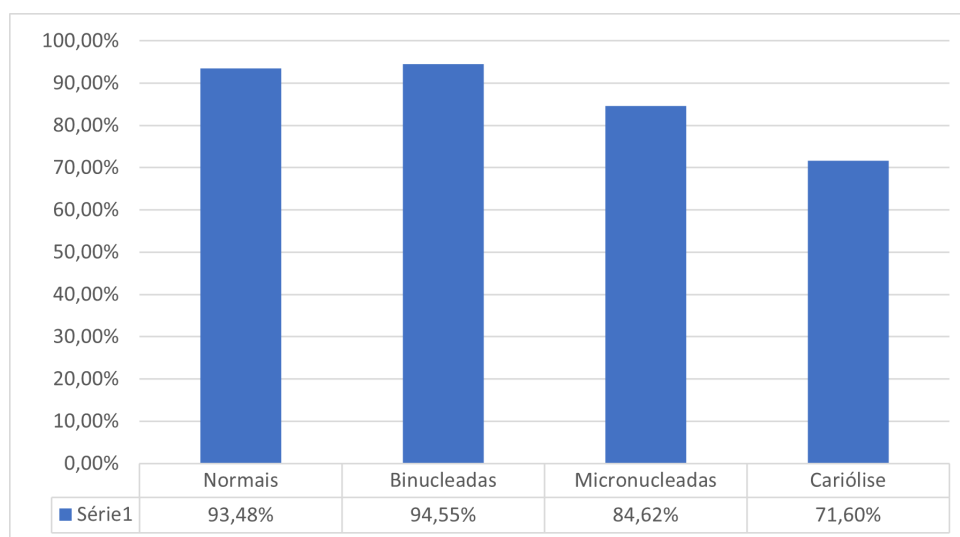
essas mesmas imagens. O resultado foi dividido entre os casos considerando as células sobrepostas e sem considerar a sobreposição, visto que isso afeta o aumento de falsos positivos para anomalias do tipo binucleada e micronucleada.

Os resultados da análise manual e do algoritmo considerando as detecções de sobreposição celular podem ser observados na Tabela 3 e pela Figura 36 é possível verificar a taxa de acerto para cada tipo de célula analisada.

Tabela 3 – Resultados para células claras considerando os casos de sobreposição celular.

	Total de Células	Normais	Binucleadas	Micronucleadas	Cariólise
Contagem Manual	279	147	52	22	58
Contagem Automática	300	138	55	26	81

Figura 36 – Taxa de acerto em células claras considerando os casos de sobreposição celular.

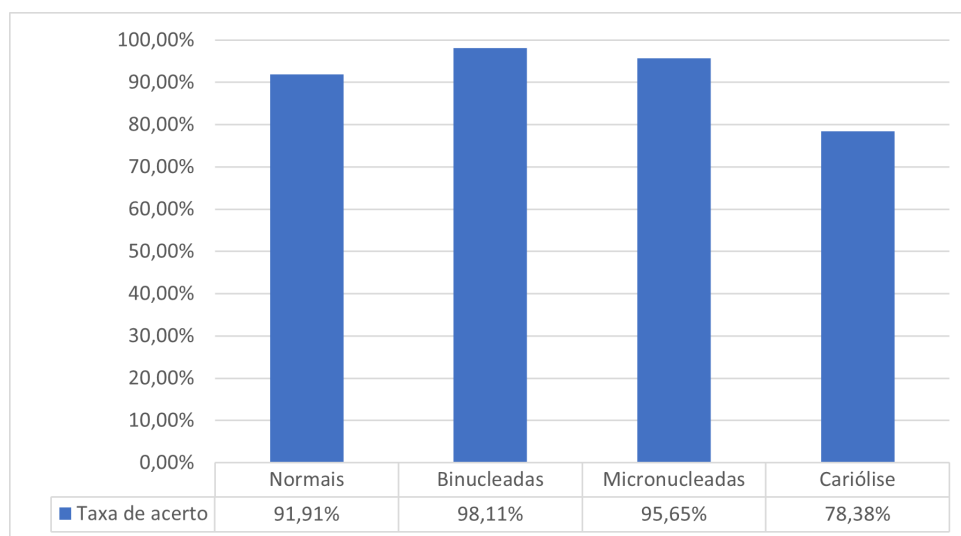


Ao remover as ocorrências de sobreposição celular, foi possível observar um aumento na taxa de acerto em células binucleadas e micronucleadas, visto que esses casos são os mais visíveis nas ocorrências de falsas aprovações para essas anomalias. Os resultados da análise manual e do algoritmo desconsiderando as detecções de sobreposição celular podem ser observados na Tabela 4 e as devidas taxas de acerto pela Figura 37.

Tabela 4 – Resultados para células claras desconsiderando os casos de sobreposição celular.

	Total de Células	Normais	Binucleadas	Micronucleadas	Cariólise
Contagem Manual	279	147	52	22	58
Contagem Automática	286	136	53	23	74

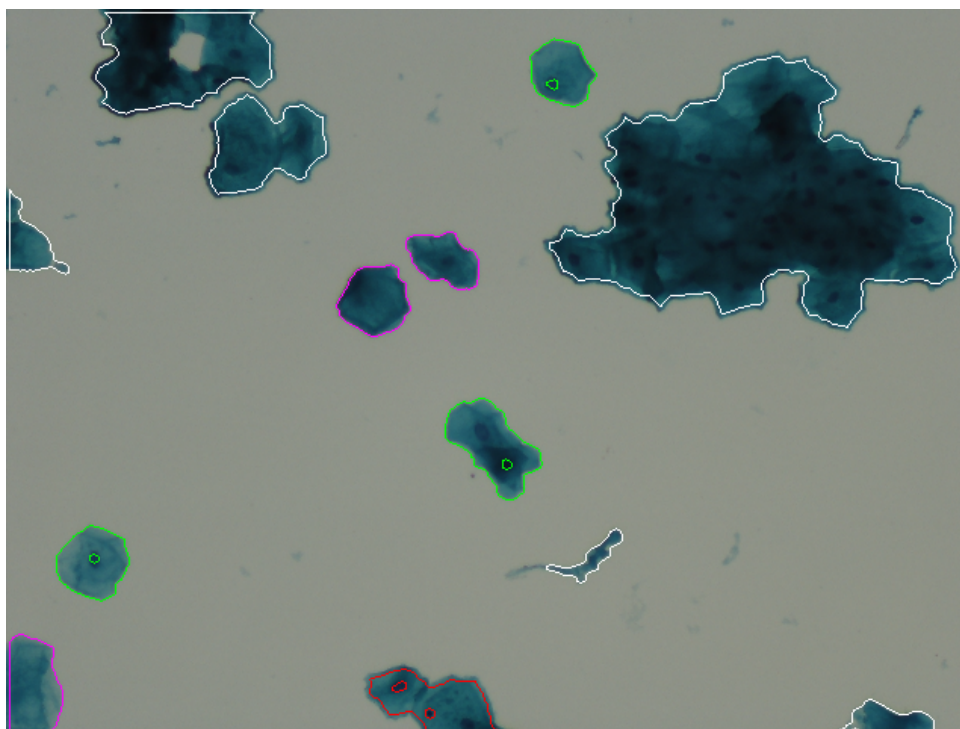
Figura 37 – Taxa de acerto em células claras desconsiderando os casos de sobreposição celular.



#### 4.1.2 Análise com Células Escuras

Analisa-se nessa etapa as imagens em que as células em sua totalidade possuem a tonalidade escura, como pode ser observado na Figura 38. As imagens que possuem as células com essa tonalidade possuem um contraste menor entre o citoplasma e seu núcleo, fazendo com que ocorra mais falsas aprovações.

Figura 38 – Imagem Resultado em células de tonalidade escura.



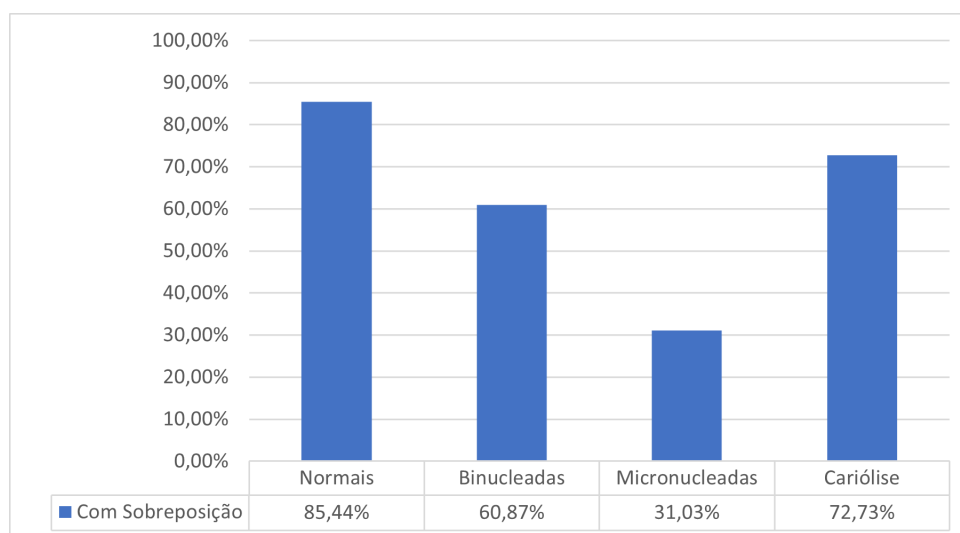
Para esta análise, foram utilizadas 45 imagens contendo 181 células no total. Novamente após realizar a análise manual de cada célula, foi comparada ao resultado do algoritmo para essas mesmas imagens.

Os resultados da análise manual e do algoritmo considerando as detecções de sobreposição celular podem ser observados na Tabela 5 e pela Figura 39 é possível verificar a taxa de acerto para cada tipo de célula analisada.

Tabela 5 – Resultados para células escuras considerando os casos de sobreposição celular.

	Total de Células	Normais	Binucleadas	Micronucleadas	Cariólise
Contagem Manual	181	118	14	9	40
Contagem Automática	210	103	23	29	55

Figura 39 – Taxa de acerto em células escuras considerando os casos de sobreposição celular.

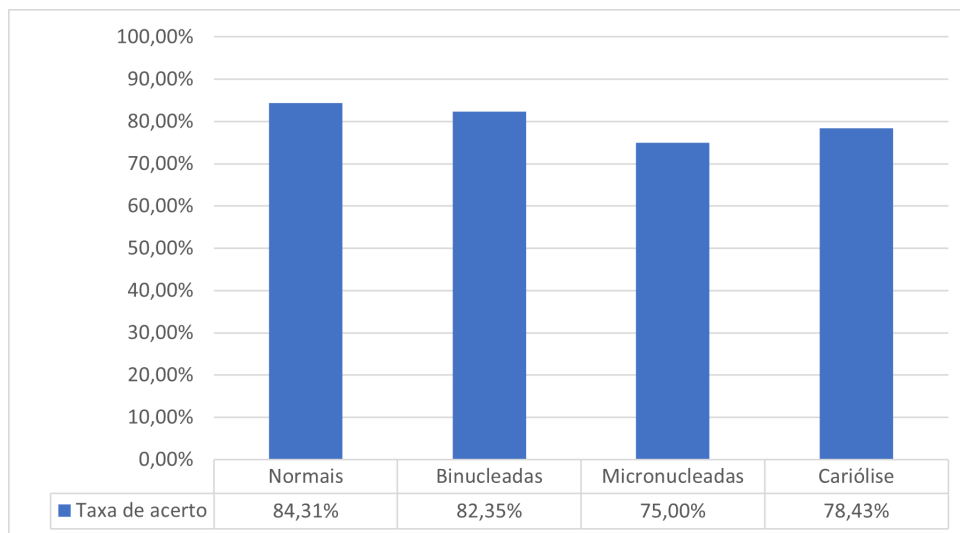


Descartando as ocorrências de sobreposição celular, é possível observar agora um aumento considerável na taxa de acerto em células binucleadas e micronucleadas, visto que esses casos são os mais visíveis nas ocorrências de falsas aprovações para essas anomalias. Os resultados da análise manual e do algoritmo desconsiderando as detecções de sobreposição celular podem ser observados na Tabela 6 e as devidas taxas de acerto pela Figura 40.

Tabela 6 – Resultados para células escuras desconsiderando os casos de sobreposição celular.

	Total de Células	Normais	Binucleadas	Micronucleadas	Cariólise
Contagem Manual	181	118	14	9	40
Contagem Automática	182	102	17	12	51

Figura 40 – Taxa de acerto em células escuras desconsiderando os casos de sobreposição celular.



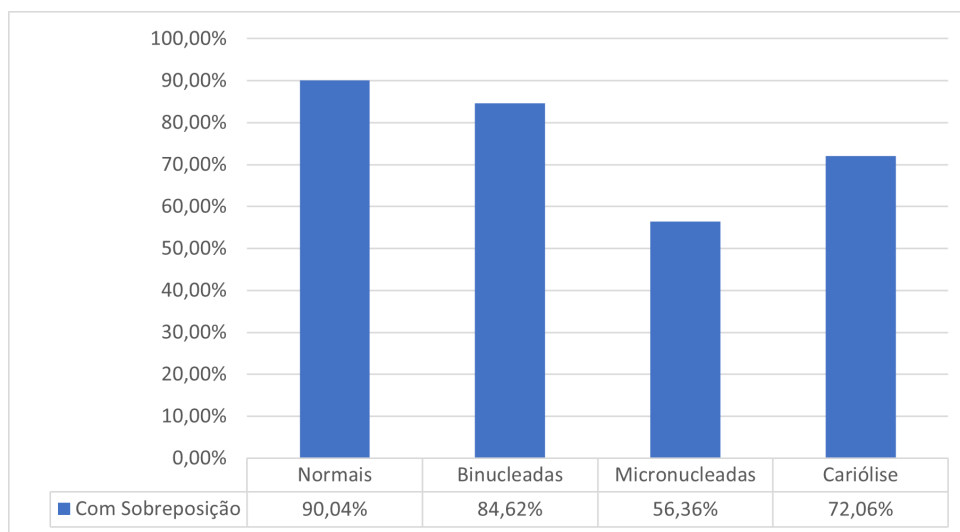
### 4.1.3 Análise Total

Ao considerar ambos os casos citados acima, foi analisado um total de 122 imagens contendo 460 células. O resultado total da análise manual e do algoritmo considerando as detecções de sobreposição celular podem ser observados na Tabela 7, sendo possível verificar a taxa de acerto para cada tipo de célula analisada pela Figura 41.

Tabela 7 – Resultado total considerando os casos de sobreposição celular.

	Total de Células	Normais	Binucleadas	Micronucleadas	Cariólise
Contagem Manual	460	265	66	31	98
Contagem Automática	510	241	78	55	136

Figura 41 – Taxa de acerto total considerando os casos de sobreposição celular.

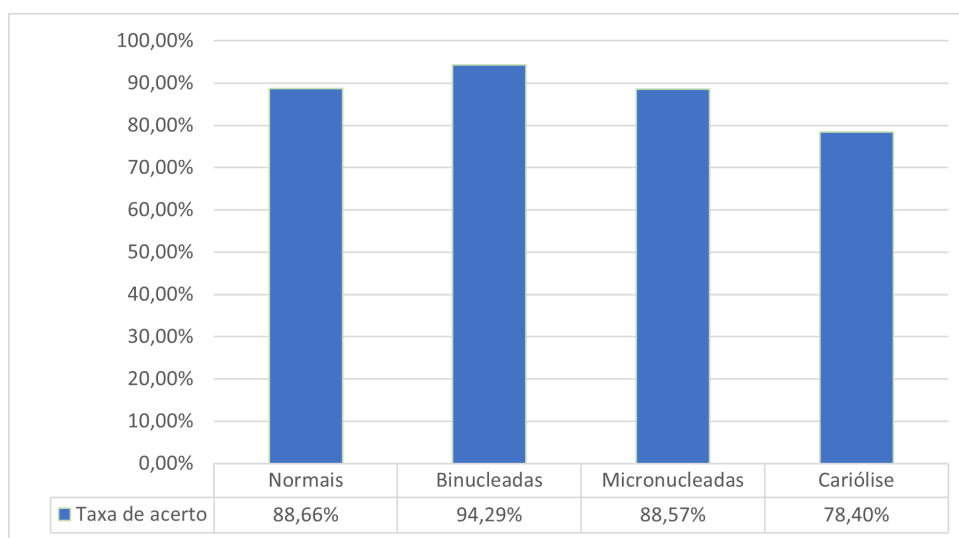


Realizando a mesma análise com o descarte de ocorrências de sobreposição celular, é possível observar os resultados encontrados na Tabela 8. A taxa de acerto total da análise das 122 imagens pode ser encontrada na Figura 42.

Tabela 8 – Resultado total desconsiderando os casos de sobreposição celular.

	Total de Células	Normais	Binucleadas	Micronucleadas	Cariólise
Contagem Manual	460	265	66	31	98
Contagem Automática	468	238	70	35	125

Figura 42 – Taxa de acerto total desconsiderando os casos de sobreposição celular.



## 4.2 Discussões

Utilizando as imagens capturadas a partir da amostra com células esfoliativas da mucosa bucal, o algoritmo desenvolvido é capaz de identificar as características dos objetos que compõem a imagem, bem como as características internas dentro de cada célula encontrada. Os resultados mostram que a aplicação consegue além de diferenciar casos que se tratam de células normais de casos de células micronucleadas, como consegue identificar outras anomalias também.

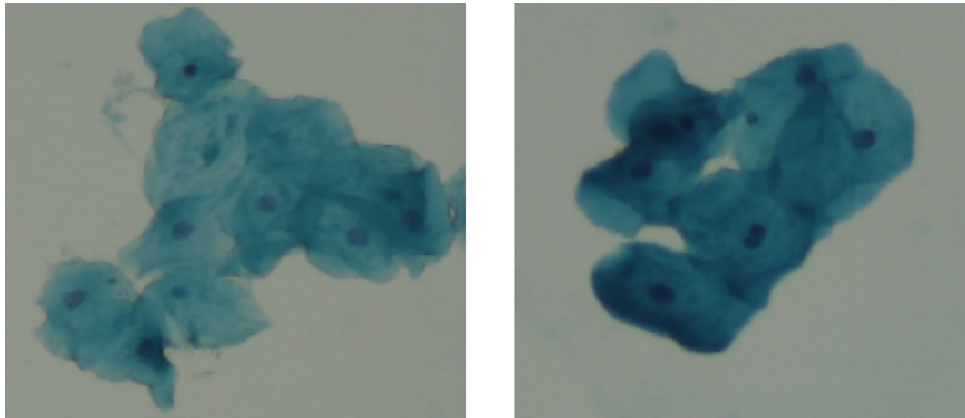
O método utilizado para trabalhar com as imagens foi inicialmente o proposto por Testa (2019). Com a necessidade de identificar outras anomalias e ao mesmo tempo criar um algoritmo adaptável para diferentes reações à coloração na célula, foi preciso alterar a estratégia desde o início da etapa de pré-processamento até o reconhecimento do núcleo.

Testa (2019) afirma que no método realizado pelo seu software foi realizada a captura de uma área grande de imagem, o que conseqüentemente reduz a qualidade das células e impede a observação minuciosa. Para melhorar esse quesito, foram capturadas imagens com maior aumento no microscópio a fim de verificar com precisão o núcleo da célula e suas anomalias.

Foi avaliado primeiramente a possibilidade de utilizar o algoritmo de *k-mean clustering* para encontrar os núcleos e suas irregularidades. O resultado do algoritmo não se mostrou satisfatória, visto que as segmentações não conseguiam localizar certas anomalias do tipo micronúcleo, levando a maiores falsas aprovações.

Muitas imagens apresentaram a ocorrência de agrupamento celular, como pode ser observado na Figura 43. Esse efeito causa a existência de apenas um objeto de reconhecimento no algoritmo.

Figura 43 – Exemplos de agrupamentos celulares.

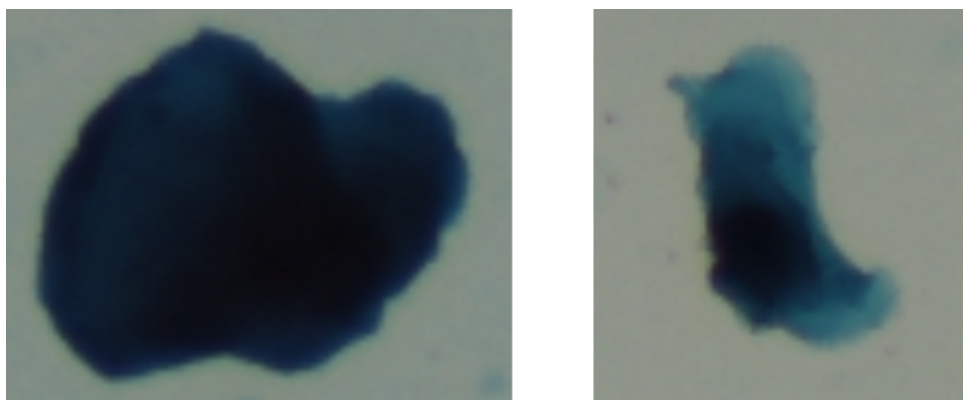


Foi utilizado o algoritmo de segmentação do tipo *watershed* para corrigir alguns casos e separá-las. Por outro lado, os marcadores utilizados no algoritmo se comportam em um determinado tamanho com o objetivo de não supersegmentar objetos a fim de não gerar mais falsos positivos. Por conta disso, o algoritmo implementado não contempla todos os casos de agrupamento celular. Para eliminar essa condição, foi determinado pela detecção de contorno um limite máximo de área e perímetro das células a serem analisadas, caso contrário o algoritmo ignora o contorno encontrado.

Para contornar o problema em relação à diferentes tonalidades de célula, foram utilizadas segmentações iterativas para realizar o reconhecimento dos objetos. A junção de reconhecimento de contorno com a segmentação iterativa levou a resultados satisfatórios tanto para células claras quanto para as escuras.

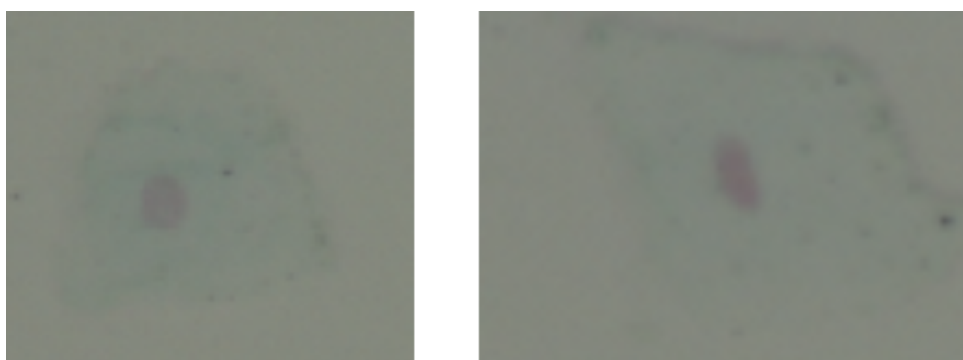
O software encontrou um quantidade relevante de falsos positivos ao se tratar de células com tons mais escuros, visto que essas células não contém contraste relevante entre o citoplasma e seu núcleo. Como resultado, o algoritmo de detecção de núcleo não consegue encontrar o núcleo e o considera como um caso de cariólise. Na Figura 44 podem ser observados alguns desses casos de excesso de coloração dentro da célula.

Figura 44 – Exemplos de células muito escuras.



O mesmo problema pode ser observado em células que a pigmentação foi minimamente absorvida, tornando complexo demais para o algoritmo segmentar os objetos corretamente. Algumas dessas células pouco pigmentadas podem ser observadas na Figura 45.

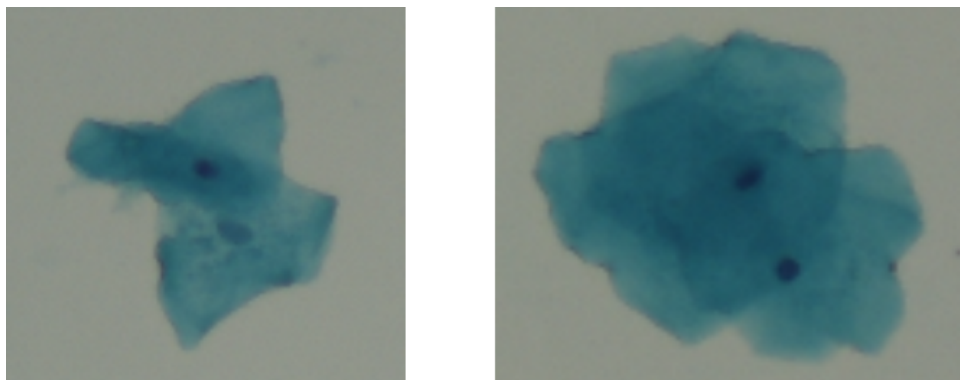
Figura 45 – Exemplos de células pouco pigmentadas.



Outro motivo que levou ao aumento da quantidade de falsos positivos foi a presença de células sobrepostas. Essas células acabam possuindo o mesmo tamanho e área de uma célula comum, como pode ser observado na Figura 46. Por conta disso, acabam gerando falsos positivos para anomalias do tipo binucleadas e micronucleadas. Entretanto, comparando os resultados da análise manual com a análise realizada pelo software, foi possível perceber a eficácia na localização de células e suas anormalidades.



Figura 46 – Exemplos de células sobrepostas.



Testa (2019) com sua ferramenta, levando em comparação a análise feita de forma manual por um profissional da área com a análise do software, conseguiu uma taxa de acerto de aproximadamente 95% no total de células. Porém, para a contagem de micronúcleos, a taxa de acerto encontrada foi de 66%.

Já para o novo algoritmo, levando em comparação a análise feita de forma manual baseada em Bolognesi *et al.* (2013) com a análise do software, houve uma taxa de acerto de aproximadamente 90% na detecção de células normais, 84% para células binucleadas, 56% para células micronucleadas e 72% para células com cariólise. Ao remover os casos de sobreposição celular, que representam 10% da análise automática feita pelo algoritmo, as taxas de acerto para células binucleadas sobem para 94% e para as micronucleadas sobem para 88%, representando uma taxa significativa de acerto, alcançando os requisitos do projeto.

Apesar do objetivo do projeto ter sido alcançado, é possível listar algumas implementações que aumentariam a funcionalidade do software. A capacidade de interpretação pode ser expandida ainda mais para registrar a ocorrência de outras mutações genéticas citadas no referencial teórico, como células *broken-egg*, brotamento nuclear, cariorrexe e picnose. Como não foi possível realizar a validação com o especialista, essas anomalias não foram avaliadas pelo algoritmo, visto que elas precisam de uma análise mais minuciosa dentro do núcleo da célula, cogitando até a necessidade da maior lente de aumento possível do microscópio para analisar o interior da célula.

Além disso, a aplicação não possui uma função para realizar a colagem das imagens capturadas pelo SAIMOScan. Essa funcionalidade poderia ser criada para avaliar a região inteira em que o usuário gostaria de analisar. Esse ajuste tornaria a aplicação integrável com o software de captura do microscópio, facilitando a avaliação completa da amostra.

A aplicação criada para aplicar o algoritmo de reconhecimento retorna um arquivo CSV para o usuário. Esse arquivo pode ser facilmente modificado para salvar a região de cada objeto encontrado e criar um banco de dados de células da mucosa bucal. Esse banco pode ter diversas funcionalidades para o usuário, sendo futuramente possível a

utilização desses dados em uma rede neural para aprimorar ainda mais o reconhecimento das anomalias.

## 5 Considerações Finais

A frequência de aparecimento de anomalias celulares em amostras de tecidos da mucosa bucal é um importante marcador biológico para identificar o efeito de carcinogênese causada pelos agentes genotóxicos. O processo manual de contagem e caracterização dessas células demanda muito tempo por parte do profissional de laboratório, visto que essa atividade é feita de forma manual. Tendo isso em vista, o trabalho desenvolvido como Projeto de Diplomação do curso de Engenharia Elétrica utilizou recursos de captura automática e análise de imagens em lote para criar uma aplicação que auxilie no processo de contagem dessas células em pesquisas laboratoriais.

O algoritmo encontrou barreiras nas tonalidades extremas das imagens, por tratar-se de células com pouco contraste entre o seu núcleo e o citoplasma. Objetos muito escuros ou com falta de pigmentação foram a maior causa de problemas com falsas aprovações para células isoladas.

A presença de sobreposição de células fez com que o algoritmo falhasse, considerando dois objetos distintos em apenas um. Ao analisar, o sistema detectava a presença de binucleação ou micronucleação, gerando falsos positivos. Por outro lado, os problemas encontrados devido à sobreposição das células nas lâminas foram considerados aceitáveis. Como a ocorrência de anomalias celulares é considerada pouco frequente, é possível o usuário da aplicação verificar as células consideradas anômalas pelo relatório gerado, comprovando a sua veracidade sem comprometer a velocidade da análise.

O algoritmo desenvolvido foi capaz de se adaptar às diferentes pigmentações de células capturadas pelas amostras. Mesmo não conseguindo avaliar corretamente todos os casos nos seus extremos, a aplicação analisa com êxito células escuras e claras a partir de uma imagem capturada das lâminas. Ademais, o sistema consegue por meio de análise de contornos ignorar características a serem descartadas, tais como agrupamentos celulares ou algum resíduo que permaneceu na lâmina ao realizar a aquisição da imagem.

Por conta da falta de um profissional da área, foi necessário utilizar o referencial teórico para fazer a detecção manual nas imagens. Como proposta de possível continuidade, seria interessante realizar a análise utilizando o mesmo banco de imagens, as quais tenham sido avaliadas e documentadas pelos profissionais da área.

Como o método de contagem para o Teste de Micronúcleos é realizada atualmente de forma manual, esse programa é de extrema utilidade para o processo, facilitando o trabalho do técnico de laboratório. Ademais, o programa ajuda tanto no tempo de trabalho demandado quanto para estabelecer um critério único para a leitura de toda a lâmina, visto que a interpretação humana é passível de mudanças. Além disso, a aplicação pode ainda

---

melhorar sua performance, identificando outras mutações celulares como broto nuclear, *broken-egg*, cariorrexe e picnose. O sistema pode ainda guardar as informações em um banco de dados para posteriormente utilizar em uma rede neural de identificação de células anômalas, sendo uma solução possível para contornar as limitações descritas no capítulo de resultados.

## Referências Bibliográficas

ANDRADE, M. G. S. *et al.* Micronúcleo: Um importante marcador biológico intermediário na prevenção do câncer bucal. **RevOdonto Ciência**. Porto Alegre, v. 20, n. 48, p. 137-141, abr./jun. 2005.

BARBON, F. *et al.* Micronúcleos e sua relação com o diagnóstico de patologias. **VI Mostra de Pesquisa e Pós-Graduação IMED**, 2013.

BATTAIOLA, A. L., **Apostila do Curso de Computação Gráfica**. Departamento de Computação (UFSCar), São Carlos, 2003.

BLOCHING, M. *et al.* Exfoliative cytology of normal buccal mucosa to predict the relative risk of cancer in the upper aerodigestive tract using the MN-assay. **Oral Oncol.**, Oxford, v. 36, p. 550-555, Nov, 2000.

BOLOGNESI, C. *et al.* The HUMNxl scoring criteria for different cell types and nuclear anomalies in the buccal micronucleus cytome assay - an update and expanded photogallery. **Mutation Research**, v. 753, p. 100-113, Out-Dez, 2013.

BONASSI, S. *et al.* Micronuclei frequency in peripheral blood lymphocytes and cancer risk: evidence from human studies. **Mutagenesis**, v. 26, p. 93-100, Jan, 2011.

CAPUTO, L. F. G. *et al.* **Conceitos e Métodos para a Formação de Profissionais em Laboratórios de Saúde - Volume 2**. Rio de Janeiro: Escola Politécnica de Saúde Joaquim Venâncio, v. 2, p. 190-213, 2013.

CHIN-WEN, C. *et al.* Automatic segmentation of abnormal cell nuclei from microscopic image analysis for cervical cancer screening. **IEEE 3rd International Conference on Nano/Molecular Medicine and Engineering**, Tainan, p. 77-80, 2009.

ESQUEF, I. A. *et al.* **Processamento Digital de Imagens**. Rio de Janeiro, 2003. Disponível em: <<https://www.doccity.com/pt/processamento-digital-de-imagens-1/5073941/>> Acesso em: 23 de set. 2019.

FENECH, M. Biomarkers of genetic damage for cancer epidemiology. **Toxicology**. Adelaide, v. 181-182, p. 411-416, 2002.

FISHER, R. *et al.* Hypermedia Image Processing Reference. **Department of Artificial Intelligence**. Escócia, 2003. Disponível em: <[http://homepages.inf.ed.ac.uk/rbf/HIPR2/hipr\\_top.htm](http://homepages.inf.ed.ac.uk/rbf/HIPR2/hipr_top.htm)> Acesso em: 15 de out. 2020.

GAUTAM, A., BHAUDARIA, H.S. White Blood Nucleus Extraction Using K-Mean Clustering and Mathematical Morphing. **Proceedings of the 5th International**

**Conference on Confluence 2014: The Next Generation Information Technology Summit**. Noida, p. 549-554, Nov, 2014.

GODOY, B. R. B. **Avaliação de Micronúcleos e Outras Alterações Nucleares em Células da Mucosa Bucal de Agricultores e Consumidores de Hortaliças**. 2018. 69 p. Monografia (Graduação em Engenharia de Bioprocessos e Biotecnologia) - Unidade Novo Hamburgo, Universidade Estadual do Rio Grande do Sul, 2018.

GONG, X. *et al.* An Overview of Contour Detection Approaches. **International Journal of Automation and Computing**, v. 15, p. 1-17, 2018.

GONZALEZ, R. C. **Processamento Digital de Imagens**. São Paulo, Pearson Prentice Hall, 2010.

INSTITUTO NACIONAL DE CÂNCER (org.). **Estimativa de Incidência de Câncer no Brasil**, 2018. Disponível em: <[http://www1.inca.gov.br/rbc/n\\_64/v01/pdf/15-resenha-estimativa-2018-incidencia-de-cancer-no-brasil.pdf](http://www1.inca.gov.br/rbc/n_64/v01/pdf/15-resenha-estimativa-2018-incidencia-de-cancer-no-brasil.pdf)> Acesso em: 10 de out. 2019.

MONTANARI, T. **Atlas digital de biologia celular e tecidual**. Porto Alegre, 2016. Disponível em: <<https://www.ufrgs.br/atlasbiocel/>>. Acesso em: 20 de set. 2020.

NEGREIROS, M. *et al.* **PROJETO SAIMOPlus**. Porto Alegre: Laboratório de Processamento de Sinais e Imagens (UFRGS), 2008. pdf.

NETO, R. A. V. **Introdução à Morfologia Matemática Binária e em Tons de Cinza**. Relatório de Iniciação Científica – Departamento de Matemática, Centro de Ciências Exatas, Universidade Estadual de Londrina, Londrina, 2010. Disponível em: <<http://www.ime.unicamp.br/valle/PDFfiles/valente10.pdf>>. Acesso em: 01 de dez. 2020.

NIKON. LV100. Instructions Manual. **Nikon**, 2005. Disponível em: <<https://www.manualslib.com/products/Nikon-Eclipse-Lv100pol-3648708.html>>. Acesso em: 23 de set. 2019.

OPENCV. **OpenCV Reference Manual v4.5**. OpenCV team, 2020. Disponível em: <<https://docs.opencv.org/>>. Acesso em: 20 de set. 2020.

OPTISCAN II. Instructions Manual. **Prior**, 2008. Disponível em: <<https://www.prior.com/wp-content/themes/prior-scientific/download.php?file=243>>. Acesso em: 23 de set. 2019.

PARVATI, K. *et al.* Image Segmentation Using Gray-Scale Morphology and Marker-Controlled Watershed Transformation. **Discrete Dynamics in Nature and Society**, v. 2008, 8 p., 2008.

PHOULADY, H. A. *et al.* A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images. **2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)**, Praga, p. 201-204, 2016.

QUEIROZ, J. E. R. de, GOMES, H. M. Introdução ao Processamento Digital de Imagens. Porto Alegre, **Rev RITA**, v. 13, p. 11-42, 2001.

SHEN, T., WANG, Y. Medical image segmentation based on improved watershed algorithm. **IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference**, Chongqing, p. 1695-1698, 2018.

TESTA, P. **Detecção de Células Micronucleadas através do Processamento de Imagens**. Trabalho de Conclusão de Curso (Graduação em Engenharia Elétrica) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

THOMAS, P., FENECH, M. Buccal micronucleus cytome assay. **Nature**, v. 4, p. 825-837, 2011.

TOLBERT, P. E. *et al.* Micronuclei and other anomalies in buccal smears: methods development. **Mutation Research**, Amsterdã, v. 271, p. 69-77, 1992.

TOROK, L. **Método de Otsu**. Instituto de Computação (UFF). Niterói, 2016. Disponível em: <<http://www2.ic.uff.br/aconci/OtsuTexto.pdf>> Acesso em: 20 out. 2020.

VEGA, A. S. de la. **Apostila de Teoria para Fundamentos de Processamento Digital de Sinais**. Departamento de Engenharia de Telecomunicações (UFF), Niterói, 2019.

ZHANG, Y., XU, D. Improved Watershed Algorithm for Cell Image Segmentation. **Advanced Materials Research**, p. 546-547, 2012.