



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Instituto de Ciências Exatas e de Informática

Métodos para classificação e segmentação de imagens do teste de Papanicolau

Bacharel Ciências da Computação

Bruno Rodrigues Faria
Guilherme Dantas Caldeira Fagundes
Lucas de Paula Vieira Santos

Resumo

O teste de Papanicolaou é um teste de prevenção do câncer cervical que rastreia alterações pré-neoplásicas no epitélio cervical células. A classificação de imagens do exame Papanicolau são de extrema importância para área de tecnologia, porém é um desafio muito grande classificar as células. Para tentar contornar esse desafio esse trabalho considera segmentar as imagens para deixar so o núcleo e após isso utilizar classificadores de Mahalanobis e Redes Convolucionais (ResNet50), respectivamente. O dataset foi utilizado pela UFOP e disponibilizado no site do CRIC.

Palavras-chave: Papanicolau. Visão Computacional. Rede neural. Distância de Mahalanobis.

Sumário

1	INTRODUÇÃO	3
2	Trabalhos relatados	3
2.1	Base de Dados	3
2.2	Pré-processamento	4
2.3	Divisão do banco de dados	5
2.4	Arquiteturas	6
3	Resultados	7
3.1	Caracterização das Células	7
3.1.1	Área	7
3.1.2	Compacidade	7
3.1.3	Excentricidade	8
3.2	Cálculo do Centro da Célula	8
3.3	Precisão do Modelo	8
3.4	Mahalanobis	9
4	Como Executar	11

1 INTRODUÇÃO

O câncer cervical é uma das principais causas de morbidade e mortalidade entre as mulheres em todo o mundo. O exame de Papanicolaou, também conhecido como citologia cervical, tem sido uma ferramenta essencial na detecção precoce de alterações nas células cervicais, possibilitando uma intervenção oportuna e eficaz na prevenção do câncer cervical. No entanto, a interpretação precisa das imagens do exame de Papanicolaou representa um desafio significativo devido à complexidade das informações visuais contidas nas lâminas.

Ao longo deste artigo, vamos explorar a metodologia empregada, analisar os resultados alcançados e destacar a importância dessas técnicas inovadoras na melhoria da precisão do diagnóstico de câncer cervical. Essas técnicas representam um avanço significativo no campo da tecnologia aplicada à saúde, contribuindo para a melhoria dos cuidados de saúde e a prevenção do câncer cervical.

2 TRABALHOS RELATADOS

Ao examinar os trabalhos publicados pelo CRIC, observamos uma tendência crescente no emprego de técnicas de segmentação para isolar áreas específicas de interesse nas imagens do exame de Papanicolaou. A concentração no núcleo celular, conforme sugerido neste estudo, surge como uma estratégia inovadora para identificar regiões próximas ao centro do núcleo, conforme indicado pelo CSV fornecido pelo criador do conjunto de dados (DINIZ et al., 2021). Esta abordagem permite a utilização de formas descritivas para classificar as imagens.

Entre os métodos de classificação empregados em estudos anteriores, o classificador de Mahalanobis se destaca por sua habilidade em modelar distribuições estatísticas complexas. Além disso, as Redes Convolucionais, especialmente a arquitetura ResNet50, têm se mostrado notavelmente eficazes em tarefas de visão computacional, demonstrando sua capacidade de extrair padrões sutis em imagens médicas.

2.1 Base de Dados

Este trabalho utiliza um conjunto de dados derivado da classificação de células cervicais, obtido a partir do banco de imagens pesquisáveis CRIC. As imagens foram geradas pelo Laboratório de Citologia da Faculdade de Farmácia na Unidade de Microscopia do Núcleo de Pesquisas em Ciências Biológicas (NUPEB) da Universidade Federal de Ouro Preto (DINIZ et al., 2021).

O banco de dados inclui tanto arquivos CSV quanto as próprias imagens. O CSV contém os seguintes atributos: **image_id**, **image_filename**, **image_doi**, **cell_id**, **bethesda_system**, **nu-**

cleus_x, nucleus_y. Cada atributo é explicado detalhadamente abaixo:

- **image_id:** Número inteiro que identifica a imagem em <<http://database.cric.com.br/>>.
- **image_filename:** Nome que identifica a imagem no arquivo ZIP que você possui.
- **image_doi:** DOI que identifica a imagem.
- **cell_id:** Número inteiro que identifica a célula em <<http://database.cric.com.br/>>.
- **bethesda_system:** Classificação da célula utilizando o sistema Bethesda, podendo ser:
 - Negativo para lesão intraepitelial.
 - ASC-US: Células escamosas atípicas de significado indeterminado.
 - ASC-H: Células escamosas atípicas não podem excluir HSIL.
 - LSIL: Lesão intraepitelial escamosa de baixo grau.
 - HSIL: Lesão intraepitelial escamosa de alto grau.
 - SCC: Carcinoma espinocelular.
- **núcleo_x:** Número inteiro entre 1 e 1384, igual à coordenada x do pixel que representa a célula.
- **núcleo_y:** Número inteiro entre 1 e 1384, igual à coordenada y do pixel que representa a célula.

Essa abordagem metodológica permite uma análise abrangente das características das células cervicais, utilizando tanto informações textuais quanto visuais disponíveis no conjunto de dados. As imagens do dataset tem as dimensões **1376 x 1020** onde as células das imagens são identificadas pelo CSV, como uma forma de facilitar. A identificação das células na lâmina.

2.2 Pré-processamento

Para otimizar a segmentação das células cervicais, inicialmente, realizou-se um recorte em uma escala de 100x100 ao redor dos núcleos fornecidos. Esse procedimento visa facilitar a segmentação subsequente. Em seguida, para isolar o núcleo do fundo, empregou-se a técnica de crescimento de regiões no pixel fornecido.

O crescimento de regiões envolve a busca das cores ao redor do pixel inicial, permitindo assim a expansão para áreas com cores semelhantes ao núcleo. Após a conclusão do crescimento, a borda da imagem é identificada para criar uma máscara que será aplicada à imagem original.

Essa máscara visa minimizar erros dentro da célula que podem surgir durante o processo de crescimento de regiões.

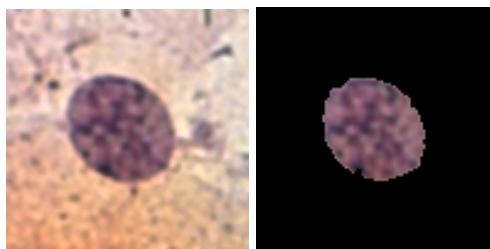


Figura 1 – (a) Imagem original (b) Imagem segmentada

Essas etapas de pré-processamento são cruciais para preparar as imagens e garantir uma segmentação mais precisa, contribuindo para uma análise mais eficaz das características das células cervicais no contexto do estudo.

2.3 Divisão do banco de dados

A divisão do banco de dados foi realizada em proporções de 80% para treinamento e 20% para teste, respeitando essa distribuição para cada classe. No entanto, observou-se que essa abordagem não resultou em um balanceamento adequado entre as classes. A distribuição final ficou da seguinte maneira:

<i>Divisão</i>	ASC-H	ASC-US	HSIL	LSIL	Negativa	SCC
<i>Treino</i>	67	312	229	394	3.395	65
<i>Teste</i>	17	79	58	99	849	17

Tabela 1 – Divisão padrão do DataSet

Para promover o balanceamento e facilitar o treinamento do modelo, adotamos uma estratégia semelhante à proposta no artigo (DINIZ et al., 2021). Para a classe negativa, nenhuma transformação foi aplicada, enquanto para as demais classes realizamos diversas transformações nas imagens, como rotações de 90, 180 e 270 graus, além de espelhamentos das imagens originais e rotacionadas. Essas transformações resultaram em um aumento significativo na quantidade de arquivos disponíveis para treinamento.

Após a aplicação dessas transformações, para a tarefa binária entre a classe negativa e as demais classes, obtivemos um conjunto de treinamento composto por 2400 arquivos negativos e 2400 arquivos positivos. Para o conjunto de teste, utilizamos 600 arquivos para cada classe.

Para o treinamento do modelo com as seis classes, a divisão ficou com 520 arquivos para treino e 130 arquivos para teste em cada classe. Essa abordagem assegurou um balanceamento adequado, fornecendo dados suficientes para o treinamento eficaz do modelo ResNet50.

2.4 Arquiteturas

A arquitetura da Rede Neural empregada neste estudo é fundamentada na ResNet50, uma robusta rede neural convolucional (CNN) que se destacou em diversas tarefas de visão computacional. A ResNet50 é uma variante da arquitetura ResNet, que utiliza blocos residuais para facilitar o treinamento de redes profundas, mitigando desafios como o desvanecimento do gradiente.

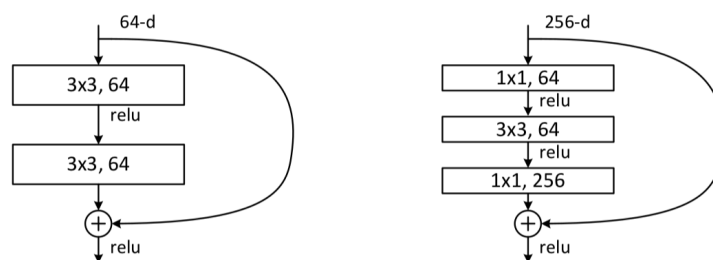


Figura 2 – Arquitetura ResNet50

Os blocos residuais permitem que a rede aprenda identidades, o que é crucial para o treinamento eficaz de redes profundas. A ResNet50 é composta por 50 camadas, incluindo camadas convolucionais, de normalização e de ativação. Ela é reconhecida por sua capacidade de extrair características complexas e aprender representações hierárquicas em imagens.

Além da ResNet50, exploramos o uso de um classificador de Mahalanobis para a tarefa de classificação de células cervicais. O classificador de Mahalanobis é uma técnica estatística que calcula a distância entre um ponto e uma distribuição de pontos. Ele se destaca na modelagem de distribuições estatísticas complexas, sendo útil quando os dados apresentam correlações e covariâncias. A seguinte imagem exemplifica o processo de classificação de Mahalanobis, considerando a distância entre classes:

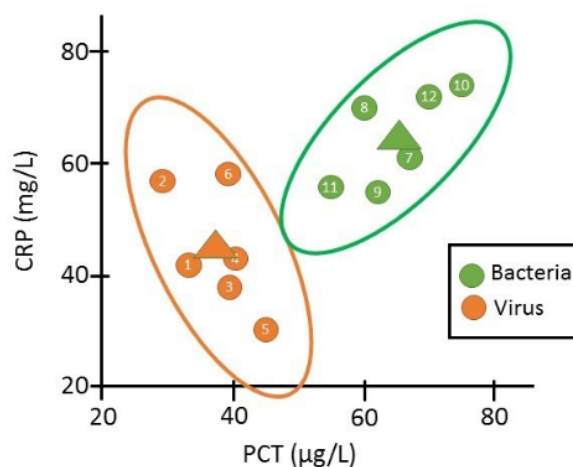


Figura 3 – Exemplo classificando Mahalanobis com a distância entre classes

Sua função pode ser definida como:

$$D_M(\mathbf{x}) = \sqrt{(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}$$

- $(\mathbf{x} - \boldsymbol{\mu})$: Diferença entre o vetor de características \mathbf{x} e o vetor de médias $\boldsymbol{\mu}$;
- $(\mathbf{x} - \boldsymbol{\mu})^T$: Transposto da diferença (para a multiplicação de matrizes);
- Σ^{-1} : Matriz inversa da matriz de covariância Σ .

A proposta é utilizar o classificador de Mahalanobis para avaliar a similaridade entre as características extraídas das imagens e as distribuições de cada classe. Essa abordagem complementar à ResNet50 visa aprimorar a precisão do diagnóstico, considerando aspectos estatísticos, na nossa classificação foi utilizada a compacidade, área e excentricidade da imagem segmentada para melhorar a forma de descobrir a distância entre as classes, para assim classificar de forma mais correta.

3 RESULTADOS

3.1 Caracterização das Células

A caracterização das células passa por definir diversos valores sobre a célula sendo elas: Area, compacidade e excentricidade.

3.1.1 Área

O cálculo da área se baseia em utilizar o fundo preto da célula já segmentada. O algoritmo contabiliza a quantidade de pixels não pretos presentes na imagem e utiliza a soma desse valor como área.

3.1.2 Compacidade

O algoritmo utiliza de uma função para calcular o perímetro P do contorno. A partir disso a compacidade é calculada como $\frac{4\pi \times A}{P^2}$. Os valores mais próximos de 1 indicam formas mais compactas, enquanto valores menores indicam formas menos compactas.

3.1.3 Excentricidade

O cálculo de excentricidade é feito a partir de uma elipse e do contorno da imagem. A excentricidade é calculada como $\sqrt{1 - \frac{a^2}{b^2}}$ em que a é o eixo menor da elipse e b é o eixo maior da elipse.

3.2 Cálculo do Centro da Célula

O cálculo do centro da célula se baseia na imagem já segmentada com o fundo preto. A distância utiliza da média das posições dos pixels não pretos em relação ao eixo vertical.

A partir disso é feito o cálculo da distância estimada em relação a distância real utilizando a distância euclidiana que é dada por $\sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2}$ em que os valores de índice 0 representam o centro estimado e os valores de índice 1 representam o centro real.

Em termos gerais, a precisão do cálculo do centro da célula se mostrou eficiente. Existe uma correlação direta entre a excentricidade e a precisão do cálculo. Quanto mais próximo de um círculo a célula tiver, maior a precisão do cálculo entre levando em conta o centro correto e o previsto.

3.3 Precisão do Modelo

O modelo treinado obteve resultados diferentes dependendo do escopo do teste. O treinamento feito com a ResNet50 para classificação binária desempenhou melhor com a acurácia na casa dos 76% tanto para treino quanto para teste treinando por 100 épocas, com 200 épocas a acurácia subiu para a casa de 80%. Tal valor pode ser considerado como interessante, entretanto, foram desempenhadas alterações no input do modelo com o intuito de aumentar a acurácia resultante. Pode-se citar, por exemplo, o aumento do número de épocas.

O modelo para multiclases apresentou um desempenho consideravelmente inferior. Para o treinamento com 100 épocas a acurácia do treinamento por volta dos 45% e de teste em 40%. Para o treinamento com 200 épocas, o valor da acurácia se manteve o mesmo. Tal resultado pode estar vinculado com uma maior complexidade do escopo do problema. Diversas ações foram tomadas para aumentar a acurácia: foi aumentado o número de camadas no modelo, foi aplicado o descongelamento das camadas do imagenet e foram executadas mudanças no pré-processamento, entretanto, nenhuma delas aumentou consideravelmente a acurácia do modelo. Aplicamos técnicas apresentadas no (DINIZ et al., 2021) com o intuito de aumentar a acurácia mas sem maior sucesso.

Tal problema com a acurácia pode estar relacionada com a dificuldade de fazer ajustes finos na

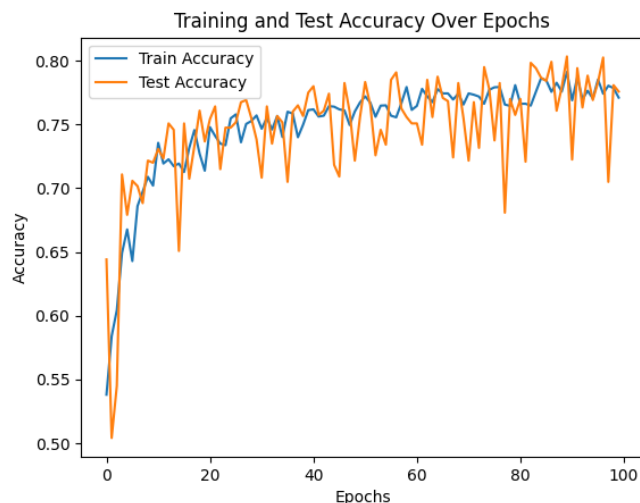


Figura 4 – Treinamento ResNet50 - Binário

parametrização do modelo, além de problemas com o balanceamento que foi feito no dataset.



Figura 5 – Treinamento ResNet50 - Multiclasse

3.4 Mahalanobis

Para análise da matrizes de confusão do Mahalanobis, as seguintes métricas foram utilizadas.

- **Acurácia:** Proporção de predições corretas em relação ao total de predições por classe;

$$Acuracia = \frac{VerdadeirosPositivos + VerdadeirosNegativos}{Total}$$

- **Precisão:** Proporção de instâncias positivas previstas corretamente em relação ao total de

instâncias *previstas* como positivas;

$$Precisao = \frac{VerdadeirosPositivos}{VerdadeirosPositivos + FalsosPositivos}$$

- **Recall:** Proporção de instâncias positivas previstas corretamente em relação ao total de instâncias que *realmente* são positivas;

$$Recall = \frac{VerdadeirosPositivos}{VerdadeirosPositivos + FalsosNegativos}$$

- **F1-Score:** Média entre precisão e recall, fornecendo uma métrica balanceada entre as duas.

$$F1 - Score = \frac{2 * Precisao * Recall}{Precisao + Recall}$$

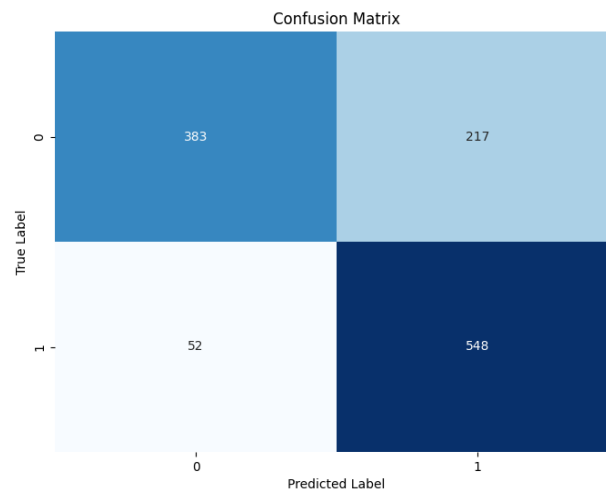


Figura 6 – Matriz de confusão - Binários

A matriz de confusão gerada pela aplicação de Mahalanobis Binário (Figura 6) apresentou acurácia de 73.5%, um resultado acima do aceitável, porém com possibilidade de melhorias. Foi possível observar que o algoritmo se destacou ao prever corretamente quando um caso era de fato "Other"(Verdadeiro Negativo), todavia os números de casos que consideraram imagens de "Other"como "Negative for intraepithelial lesion"(Falso Positivo) foi acima do esperado, enquanto o número de Falsos Negativos apresentou-se consideravelmente baixo.

A precisão foi de aproximadamente 63.8%, um resultado não muito decente para acertar quais imagens eram de fato "Other", enquanto o recall foi de aproximadamente 88%, um resultado muito bom, o que sugere que alta capacidade de capturar casos negativos. Essas métricas resultam em um F1-Score de 74,1%, um equilíbrio acima do esperado.

No caso da matriz de confusão gerada pela aplicação de Mahalanobis Multiclasse (Figura 7) os calculos das matrizes permanece com as mesmas funções, porém para cada classe individualmente. A tabela a seguir apresenta os valores encontrados:

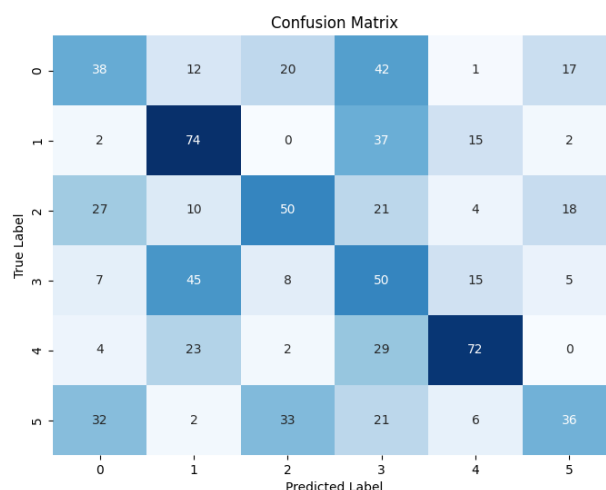


Figura 7 – Matriz de confusão - Multiclasse

Classe	Precisão	Recall	F1-Score
ASC-H	0.375	0.388	0.381
ASC-US	0.542	0.493	0.517
HSIL	0.455	0.260	0.331
LSIL	0.269	0.297	0.282
Negative for intraepithelial lesion	0.720	0.510	0.598
SCC	0.387	0.300	0.338

Foi possível observar uma drástica redução na qualidade das métricas do modelo, apresentando valores muito reduzidos e abaixo do esperado. Tal queda deve-se, potencialmente, aos seguintes fatores:

1. Problemas de balanceamento do Dataset;
2. Insuficientes dados no Dataset;
3. Falha no Pré-processamento de dados;
4. Parâmetros mal ajustados;
5. Insuficiente caracterização das classes.

4 COMO EXECUTAR

1. Executar os arquivos na respectiva ordem:
image_processing.py, *balance_dataset.py* e *segmentation.py*.
 Todos eles estão contidos no diretório *src/train_codes*.
2. Executar o arquivo *main.py* que está no diretório *src*.

3. Ao executar o *main.py* a janela da aplicação irá abrir.
4. Clicar em *Arquivo* no canto superior esquerdo, clicar em *abrir* e selecionar uma imagem do conjunto de células desejado.
5. Digitar o valor de N no campo da parte superior da tela, ele é responsável pelo tamanho da imagem gerado após o corte da célula.
6. Clicar em *Opções* no canto superior esquerdo. A segmentação será exibida no centro da tela e os quadrados serão desenhados.

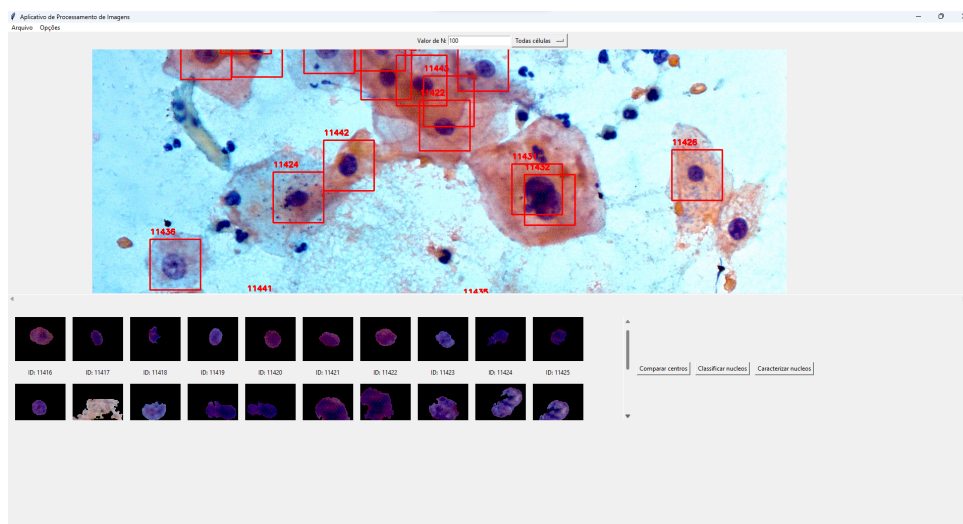


Figura 8 – Imagem da Aplicação 1

7. Ao clicar em *comparar centros*, a aplicação irá exibir uma nova janela exibindo cada célula, seu respectivo centro calculado e a distância total.
8. Ao clicar em "Classificar núcleos" ele testa utilizando o modelo ResNet50 e Mahalanobis, exibindo as métricas resultantes em uma nova janela.

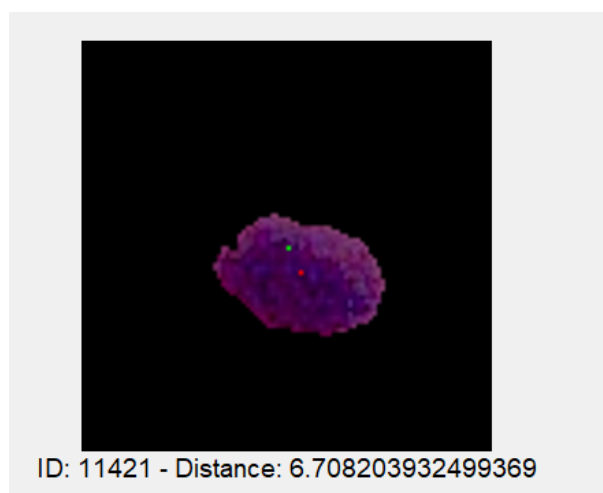


Figura 9 – Imagem da Aplicação 2

9. Ao clicar em "Caracterizar núcleos", o programa gera um CSV contendo os atributo das célulass e, uma nova janela.

Referências

DINIZ, Débora N. et al. A deep learning ensemble method to assist cytopathologists in pap test image classification. **Journal of Imaging**, v. 7, n. 7, 2021. ISSN 2313-433X. Disponível em: <<https://www.mdpi.com/2313-433X/7/7/111>>.