

# Segmentação de núcleos de células cervicais em exame de Papanicolau

Paulo Henrique Calaes Oliveira

Orientadora: Andrea G. C. Bianchi



# Agenda

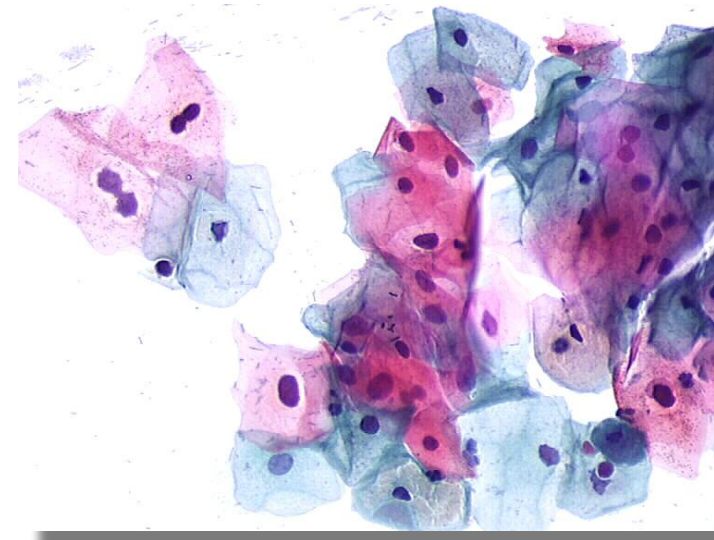
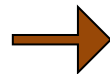
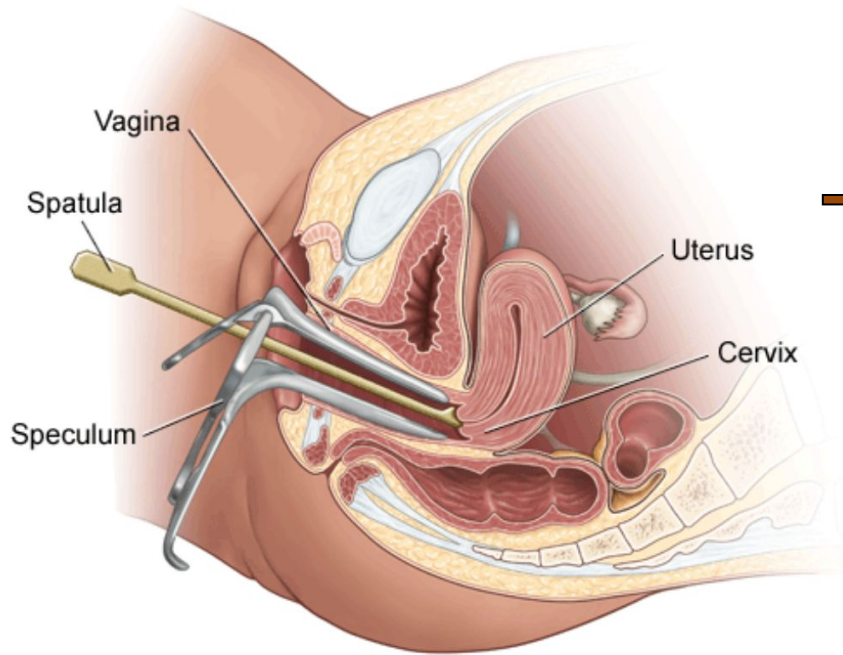
1. Introdução
2. Objetivos
3. Bases de Dados
4. Método Proposto
5. Testes
6. Resultados
7. Conclusão

# 1. Câncer Cervical

- Terceira maior causa de mortes relacionadas à câncer em mulheres;
  - Acima de 527,600 casos de câncer anualmente;
  - 265,700 mortes em 2012
- 
- 100% chance de cura se diagnosticado prematuramente

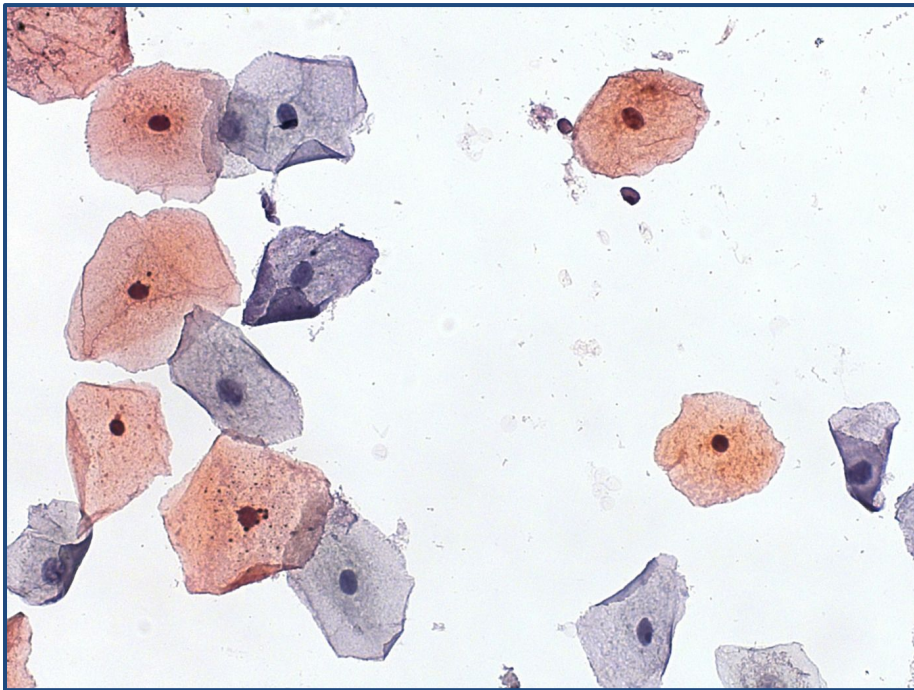
# 1. Câncer Cervical: Diagnóstico

Pap Test

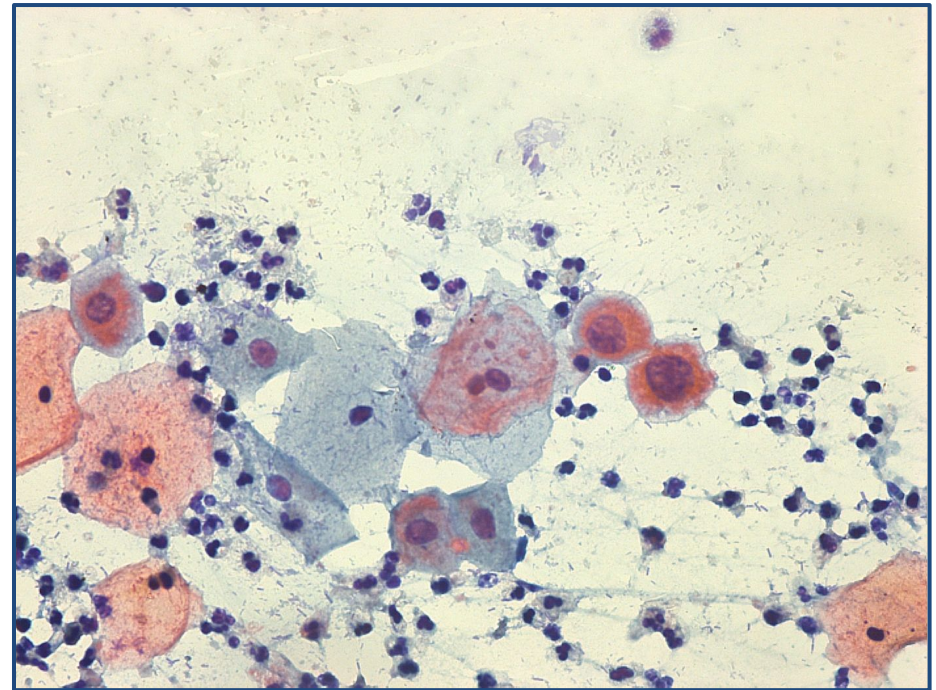


# 1. Câncer Cervical: Diagnóstico

Exame em meio líquido



X Exame Convencional



# 1. Câncer Cervical: Diagnóstico

- Milhões de testes são realizados anualmente
- Os exames são feitos de forma manual e tem uma taxa de erro de 5% a 55%

Por que não criar uma ferramenta de auxílio ao diagnóstico?

## 2. Objetivos

- Proposta de um método de segmentação de núcleos de células cervicais

## 2. Objetivos

- Proposta de um método de segmentação de núcleos de células cervicais
- Comparação do método com o estado da arte



## 2. Objetivos

- Proposta de um método de segmentação de núcleos de células cervicais
- Comparação do método com o estado da arte
- Base de dados de imagens reais

## 2. Objetivos

Importância da segmentação dos núcleos:

- Relação Núcleo / Citoplasma
- Tamanho
- Granulometria
- Pigmentação
- Forma

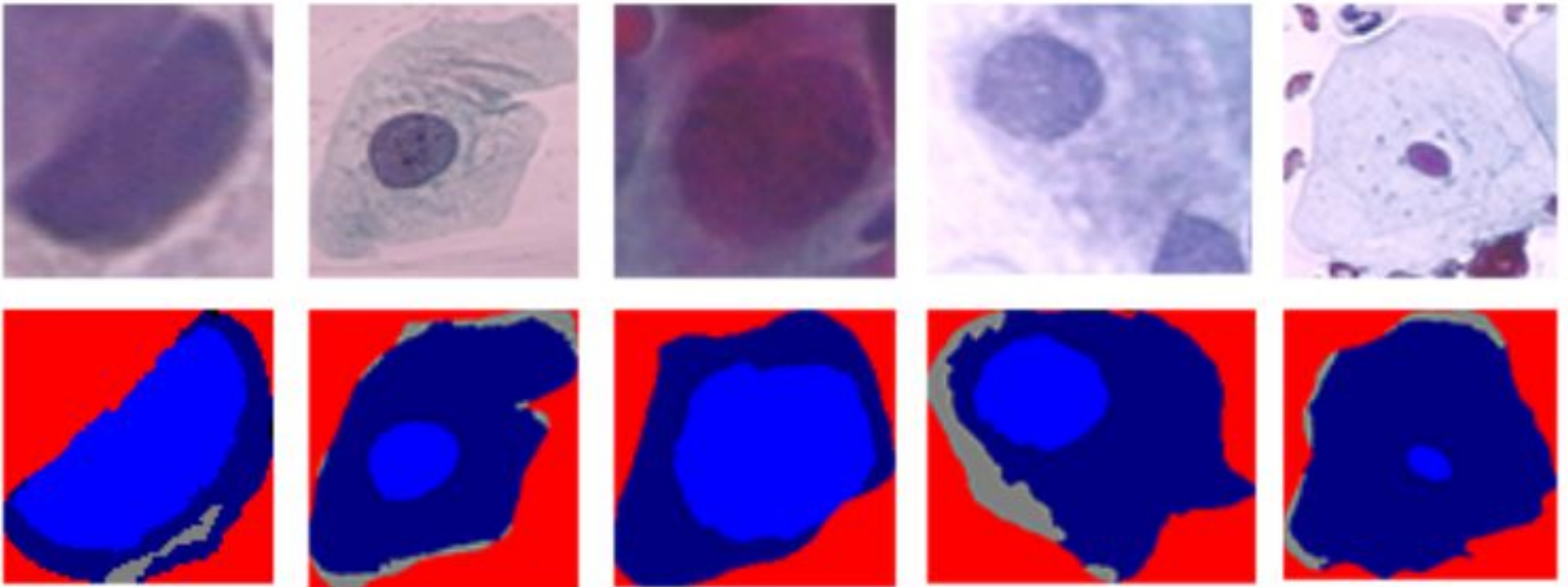
# 3. Bases de Datos

- Herlev
- Hacettepe
- ISBI 2014

### 3. Bases de Dados: Herlev

- 917 imagens
- 7 classes

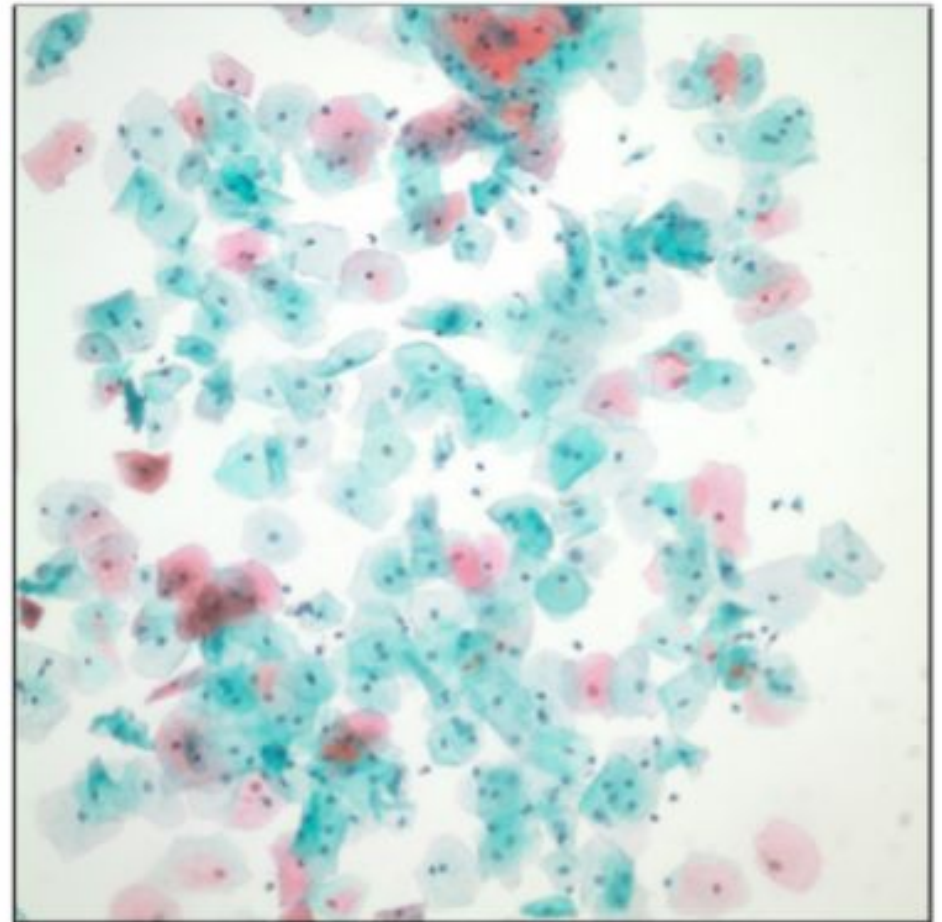
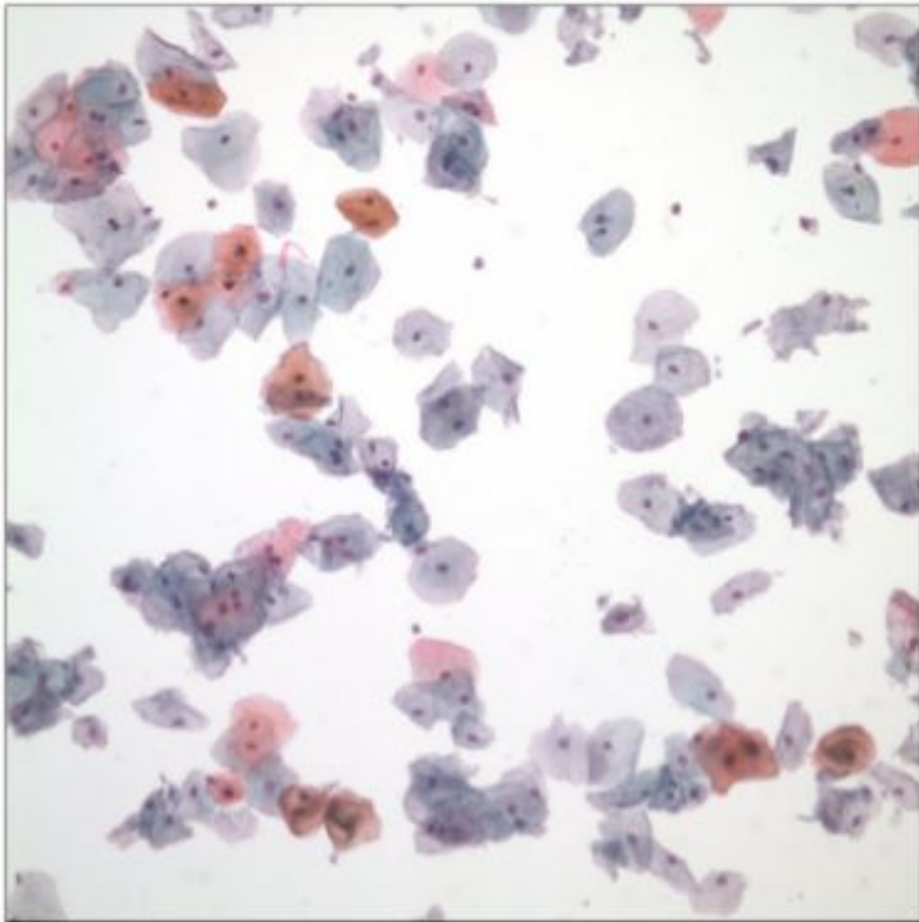
Desvantagem: Uma célula por imagem.



### 3. Bases de Dados: Hacettepe

- 82 imagens

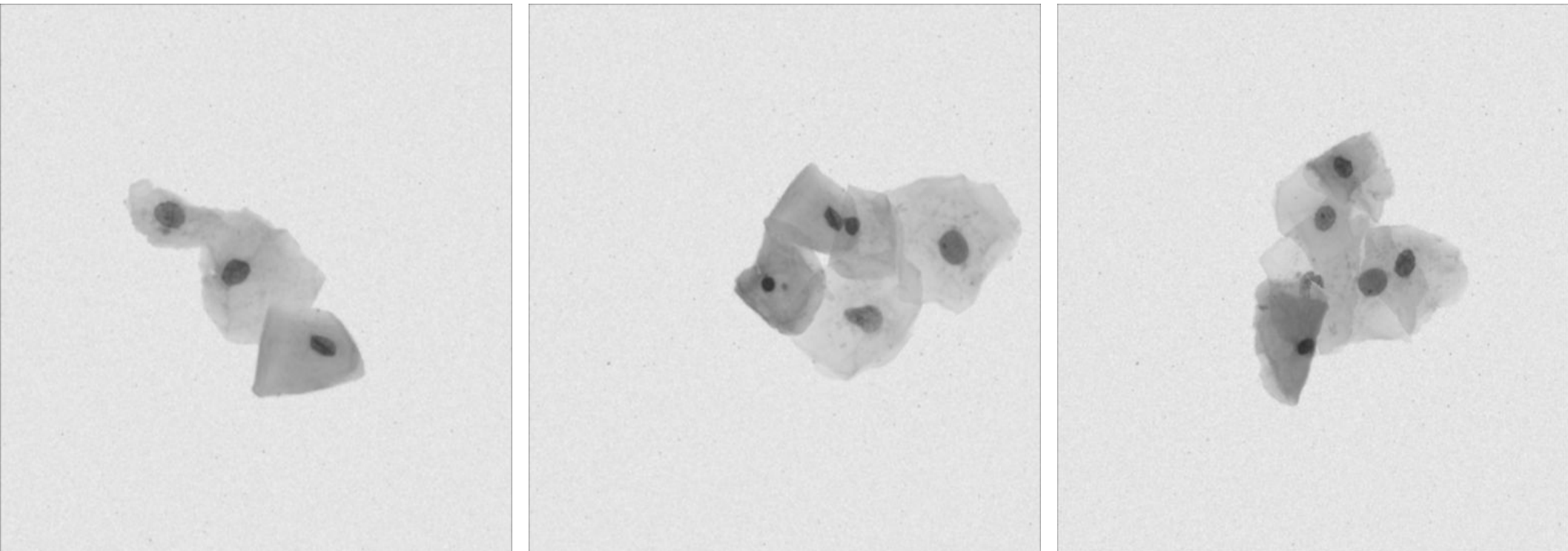
Desvantagem: Não é pública



### 3. Bases de Dados: ISBI 2014

- First Challenge of Overlapping Cervical Cell Segmentation at International Symposium on Biomedical Imaging(ISBI)

Desvantagens: Imagens simuladas  
Não trata classificação



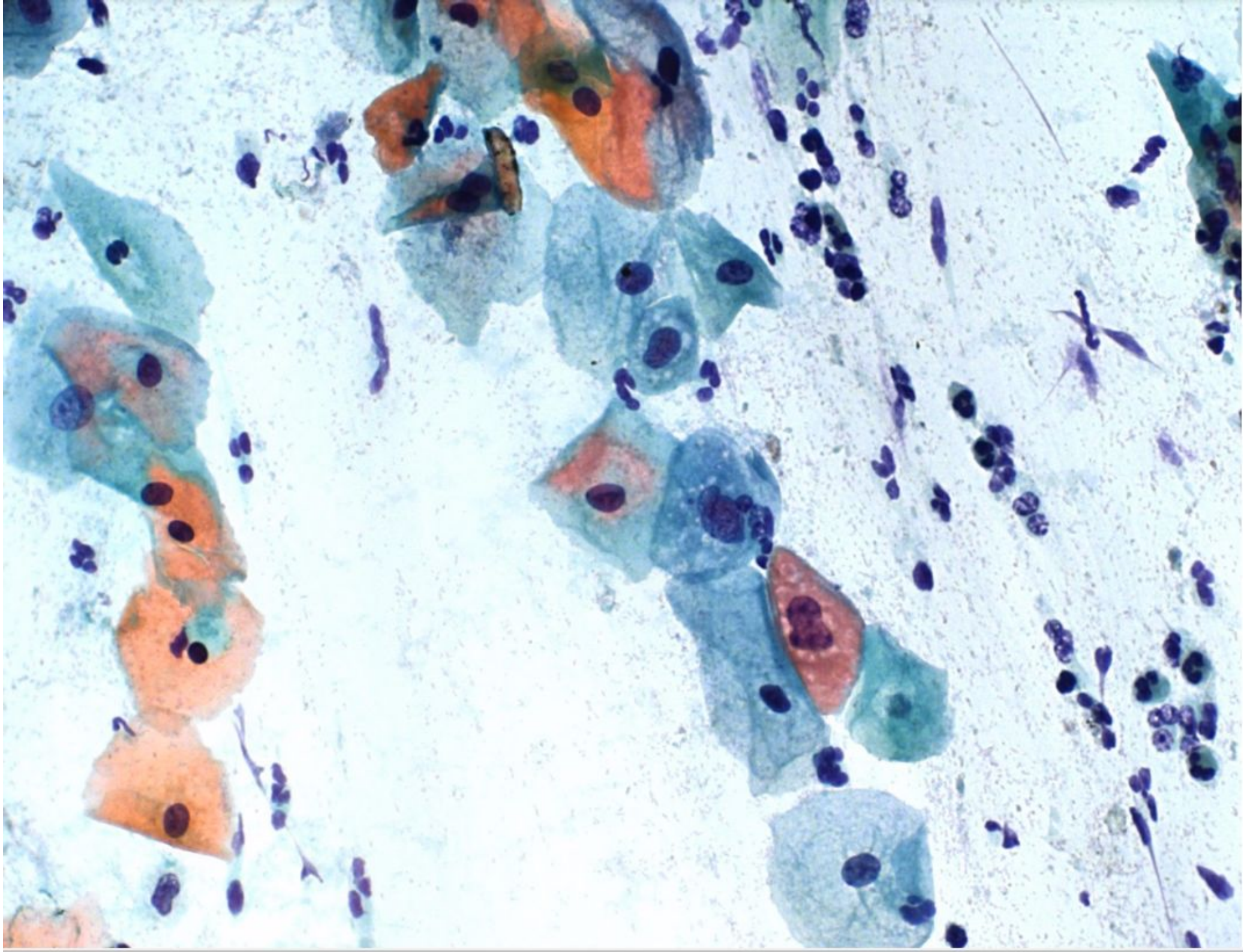
# Trabalhos Relacionados

Trabalho	Base de dados	Precision %	Recall	F-value
Kale & Aksoy 2010	Herlev/Hacettepe	86,36	92,61	89,37
Gençtav, Aksoy & Onder 2012	Herlev/Hacettepe	74,71	93,53	83,06
Nosrati 2014	ISBI 2014	90,30	89,30	89,79
Ushizima et al 2014	ISBI 2014	95,90	89,50	92,58
Nosrati 2015	ISBI 2014	-	-	99,00



HOG Features + Random Forest



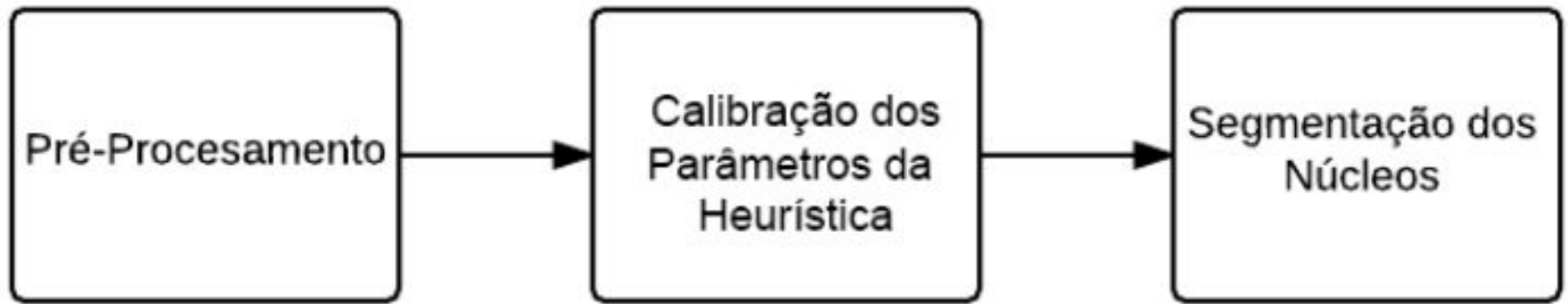




# 3. Bases de Dados

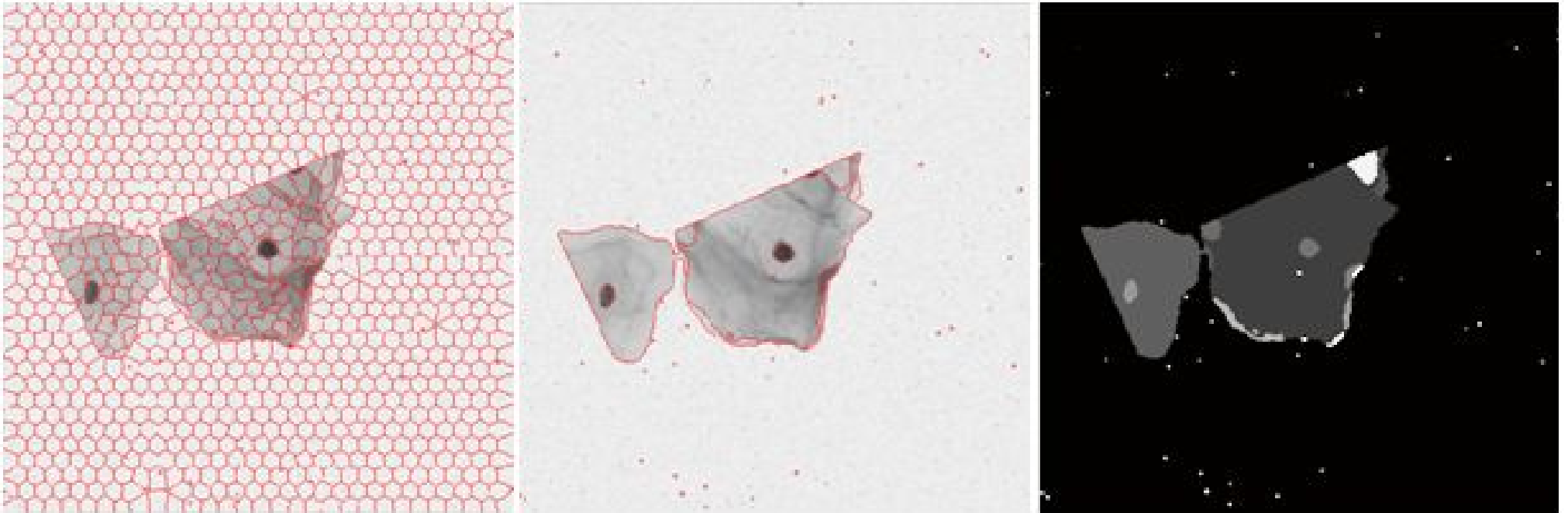
- ISBI 2014
- Herlev
- Hacettepe
- Base de Imagens Reais

## 4. Método Proposto



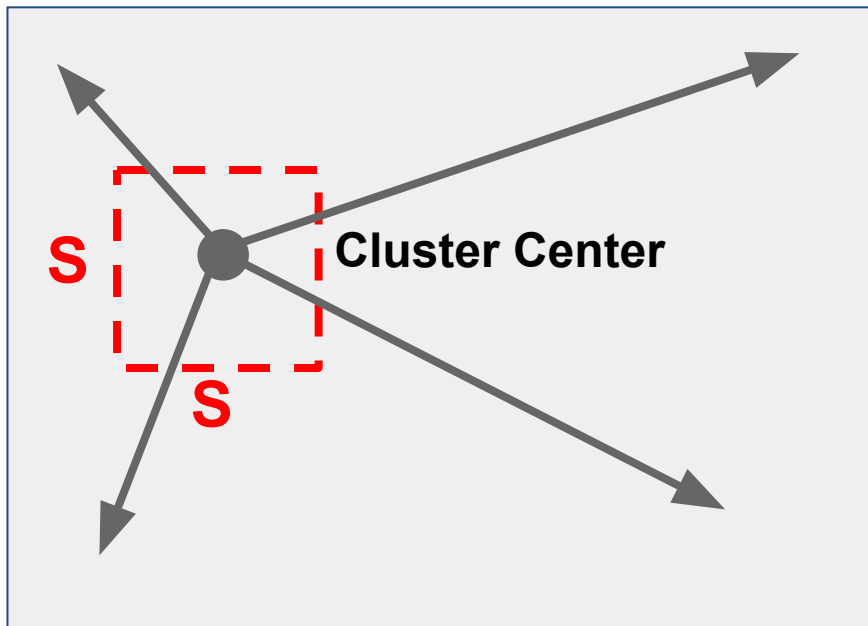
# Pré-Processamento

- *Simple Linear Iterative Clustering (SLIC)*
- *Density Based Spatial Clustering of Applications with Noise (DBSCAN)*

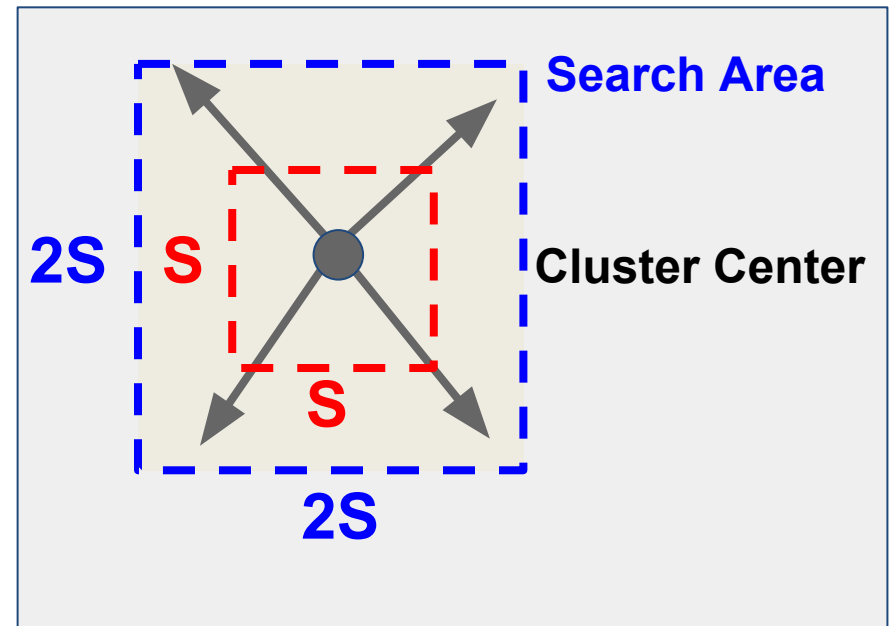


# SLIC

- Uma adaptação do K-means para a geração de superpixels
- Distância de Cor + Distância Espacial



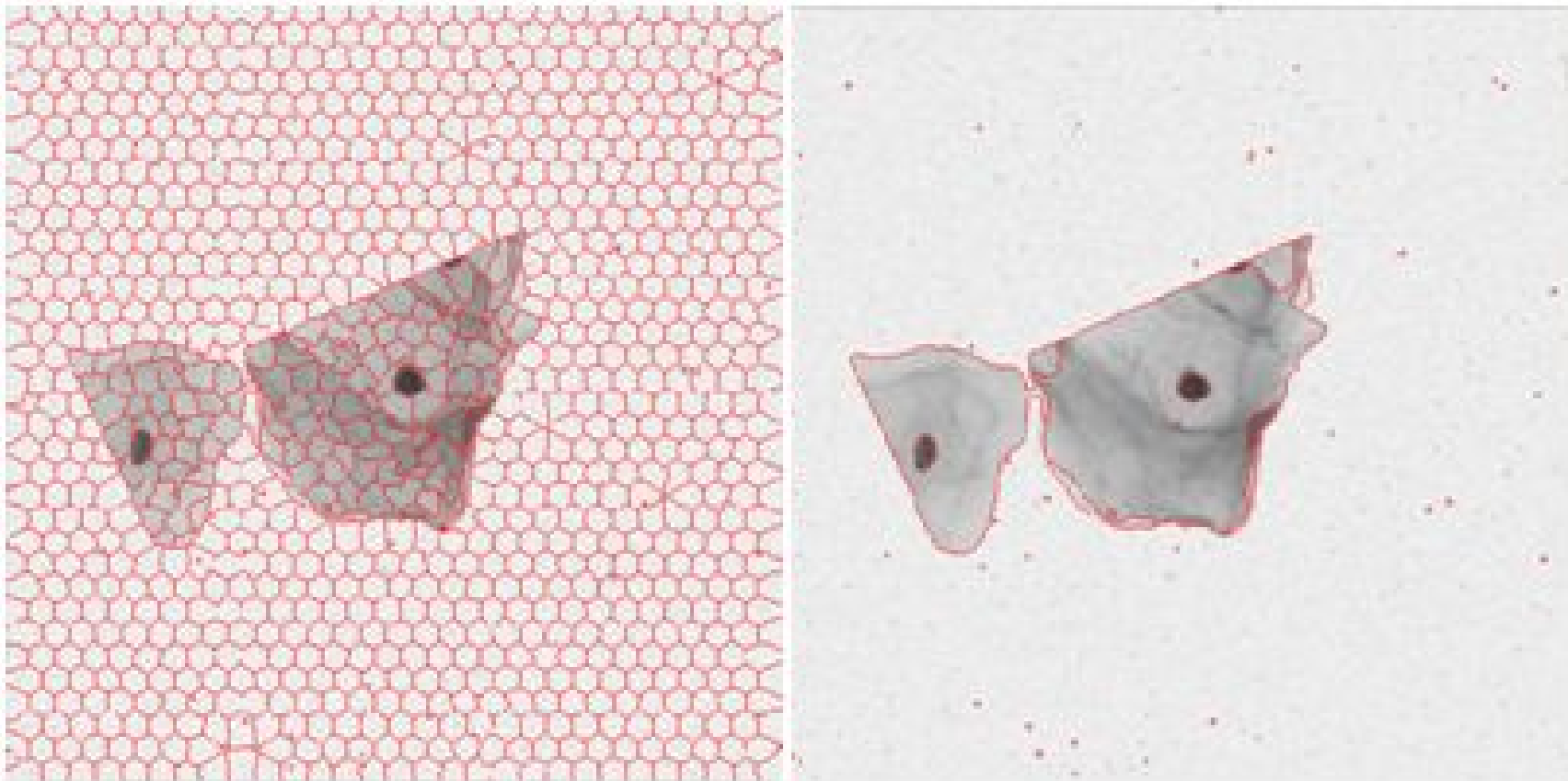
K-means



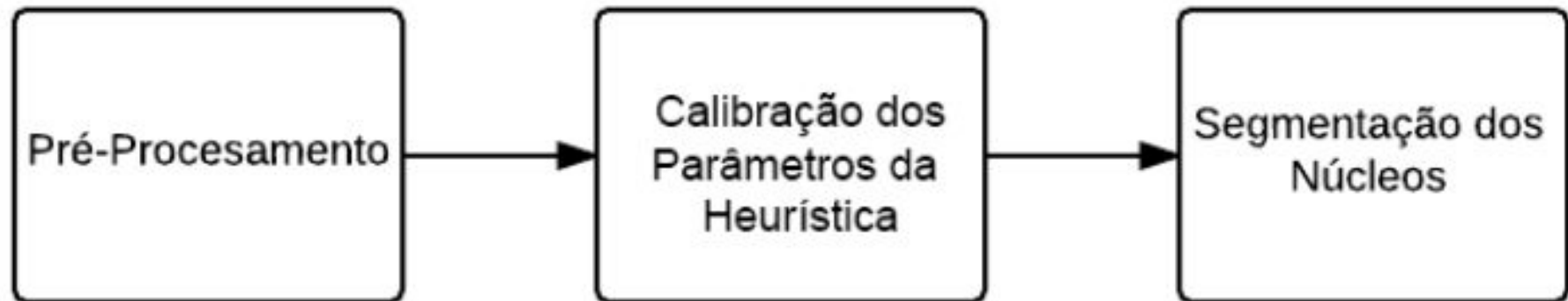
SLIC

# DBSCAN

- Usado para agrupar regiões similares



## 4. Método Proposto

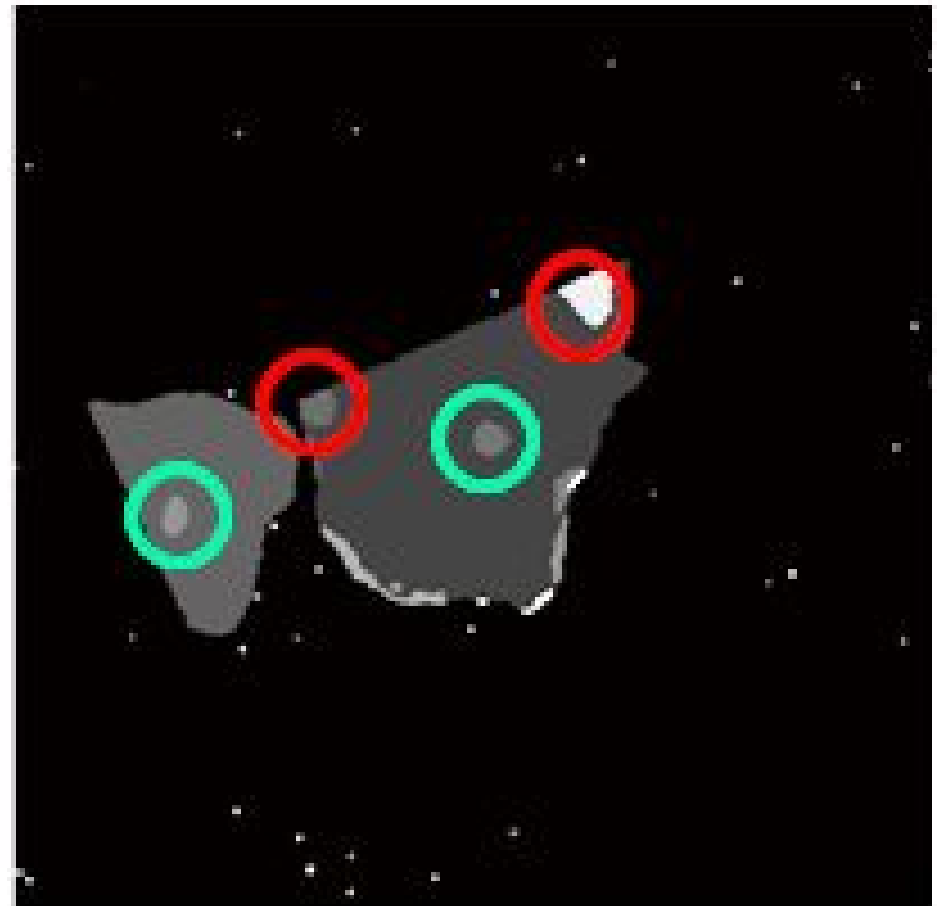
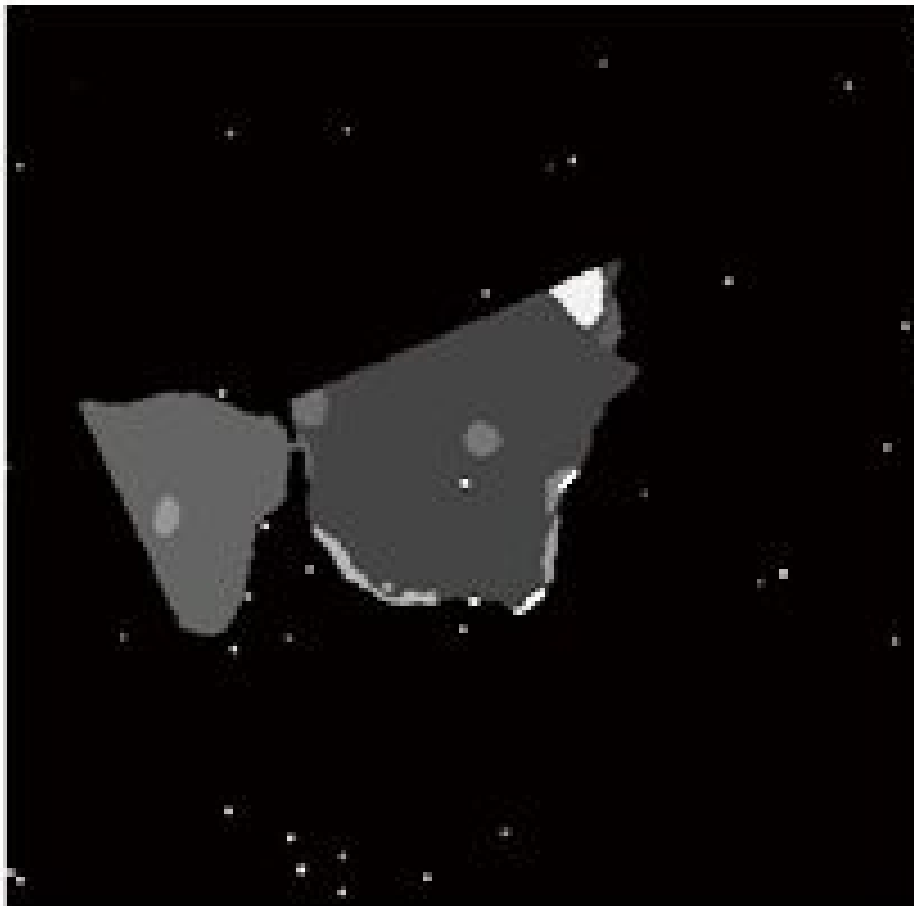


# Avaliação da Solução

## Heurística CIA

(Circularidade, Intensidade e Área)

$$circularidade = \frac{4\pi Area}{Perimetro^2}$$



# Heurística CIA

1. **input:** *superpixel* as *s*
2. **if** *s.Intensity* > *maxIntensity* **then**
3.     **return** **false**;
4. **else if** *s.Area* < *minArea* **or** *s.Area* > *maxArea* **then**
5.     **return** **false**;
6. **else if** *s.Circularity* > *minCircularity* **or** *s.Circularity* > *maxCircularity* **then**
7.     **return** **false**;
8. **else**
9.     **return** **true**;

*Se o superpixel satisfaz todos os critérios ele é classificado como núcleo, caso contrário, é descartado.*



# Calibração dos Parâmetros da Heurística

- Algoritmo baseado em NSGA-II
- Cada indivíduo tem os seguintes 5 parâmetros:
  - Intensidade máxima
  - Área mínima
  - Área máxima
  - Circularidade mínima
  - Circularidade máxima

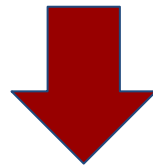
# Objetivos



Maximizar taxa de TP



Minimizar taxa de FP



# Calibração dos Parâmetros da Heurística

1. Gera população inicial ( $n$  indivíduos)
  - a. Avalia indivíduos
  - b. Cruzamento
  - c. Mutação
  - d. Avalia População
  - e. Seleciona nova população
2. Repete a-e

# Cruzamento

A

15	10	600	0.9	1.2
----	----	-----	-----	-----



B

10	11	400	0.8	1.1
----	----	-----	-----	-----



A + B

15	11	400	0.9	1.1
----	----	-----	-----	-----

# Mutação

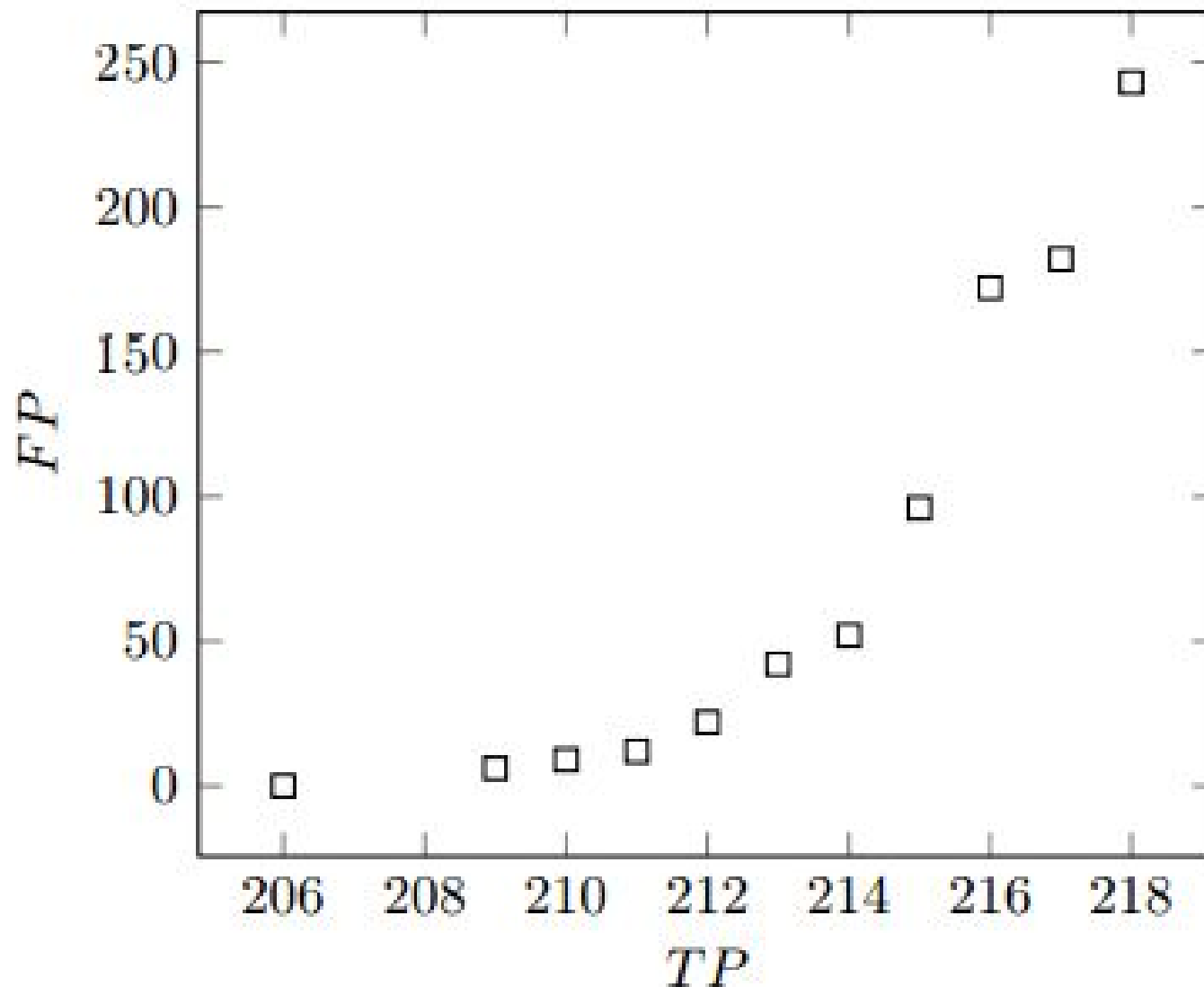
Max Intensity	Min Area	Max Area	Min Circularity	Max Circularity
10	20	500	0.8	1.2

Probabilidade de mutação: X %,  
faz mutação?

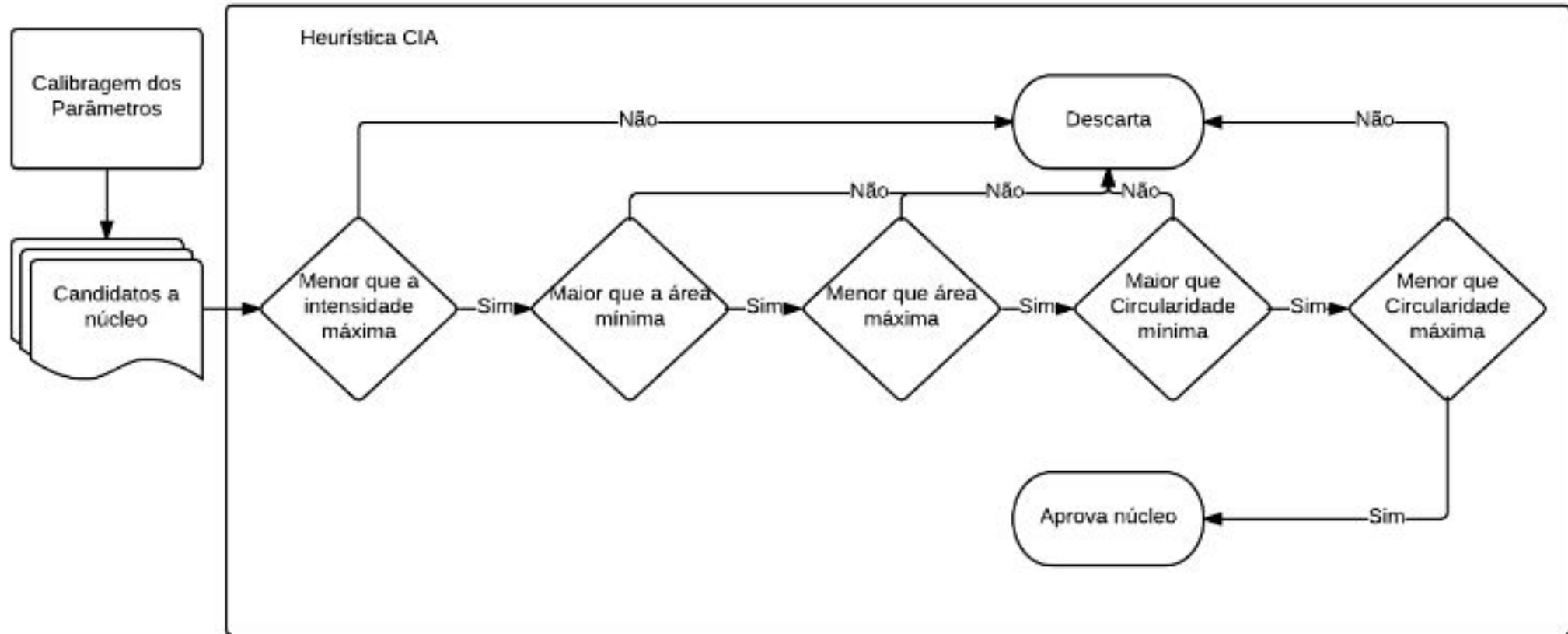
Sim	Não	Sim	Sim	Não
12	20	530	0.85	1.2

- Intensidade máxima,  $\delta \in [-5, 5]$ ;
- Área mínima,  $\delta \in [-5, 5]$ ;
- Área máxima,  $\delta \in [-5, 5]$ ;
- Circularidade mínima,  $\delta \in [-0.1, 0.1]$ ;
- Circularidade máxima,  $\delta \in [-0.1, 0.1]$ .

# Avalia Soluções: Fronteira Pareto



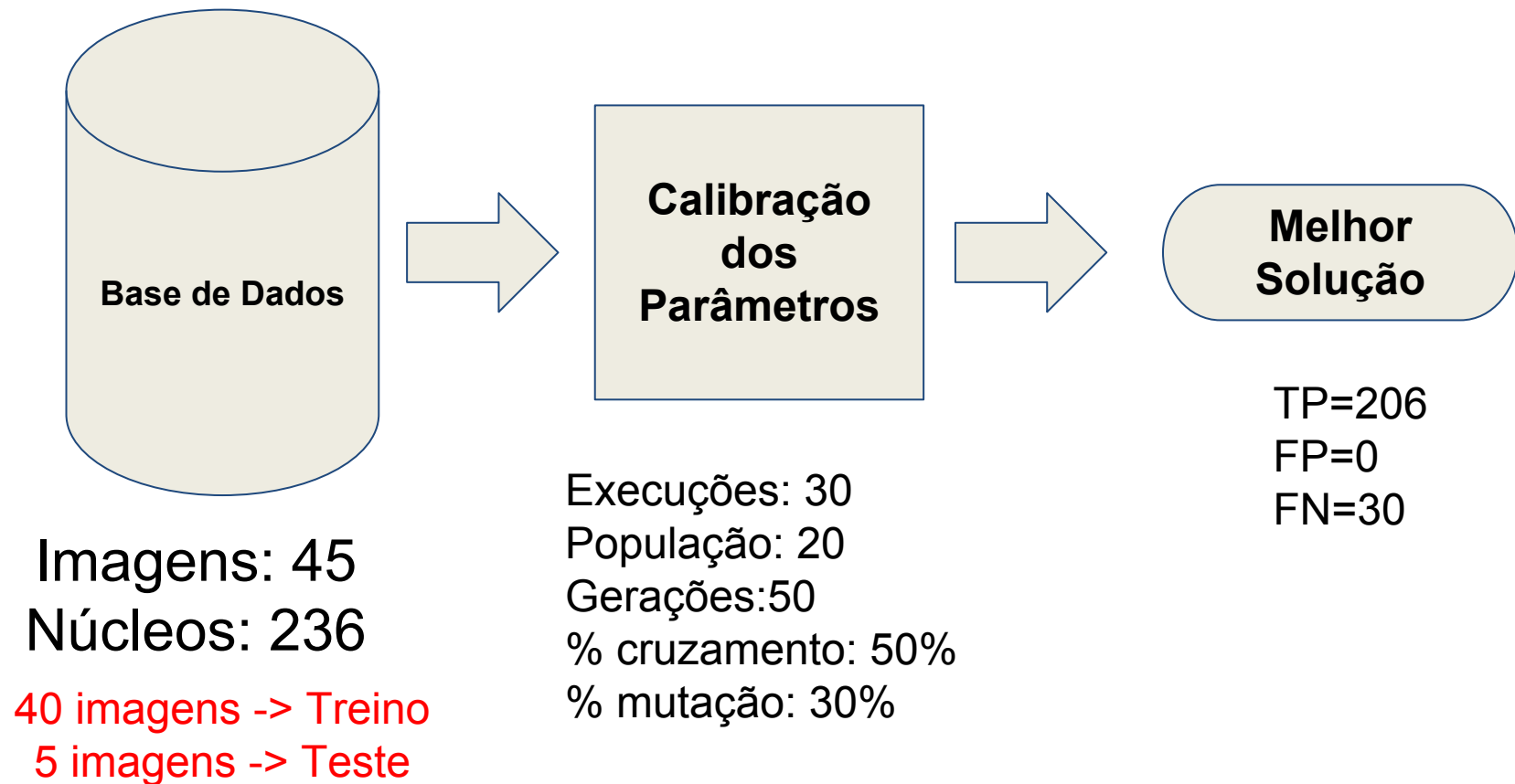
# Segmentação dos Núcleos



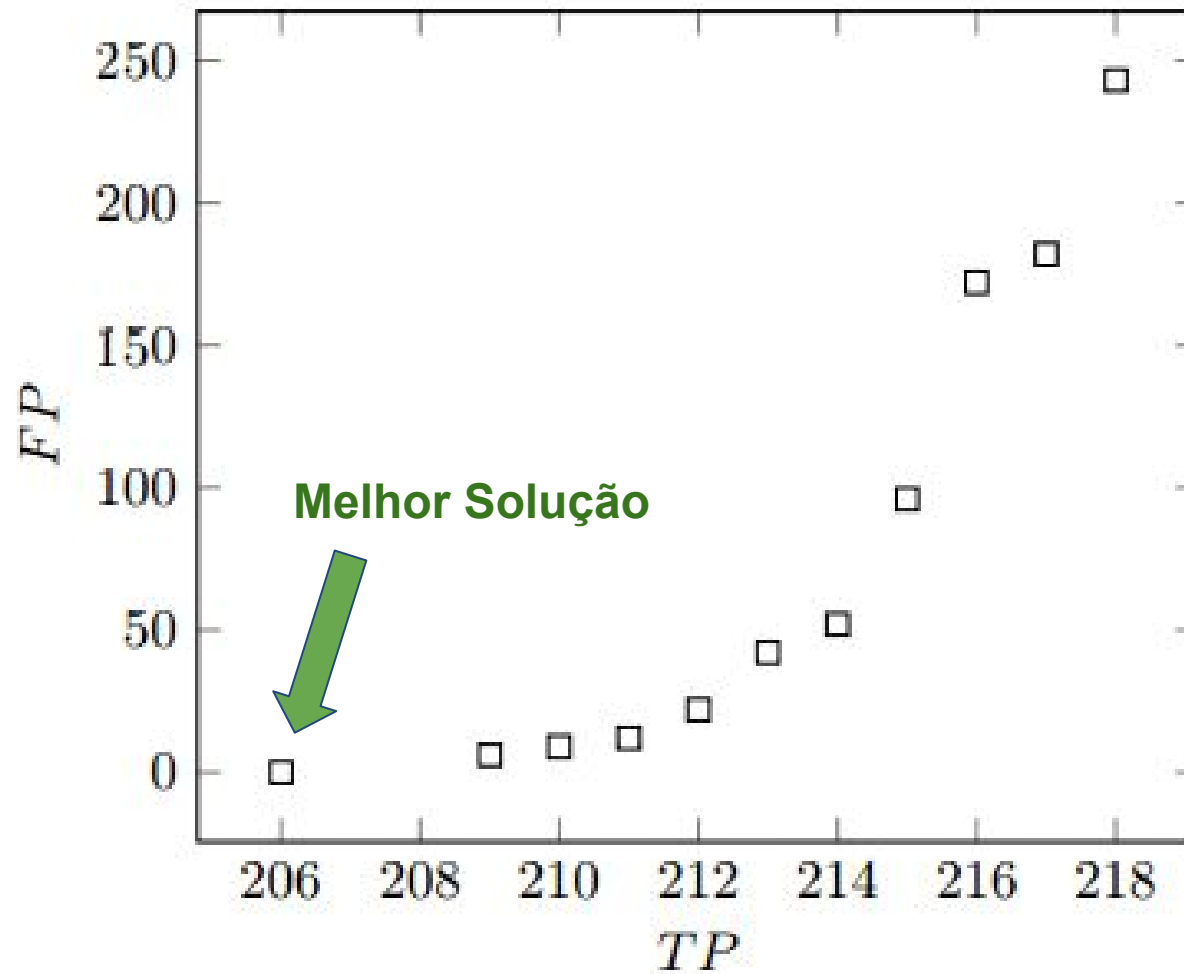
# Experimentos e Resultados



# Testes de Calibração(*isbi\_train\_45*)



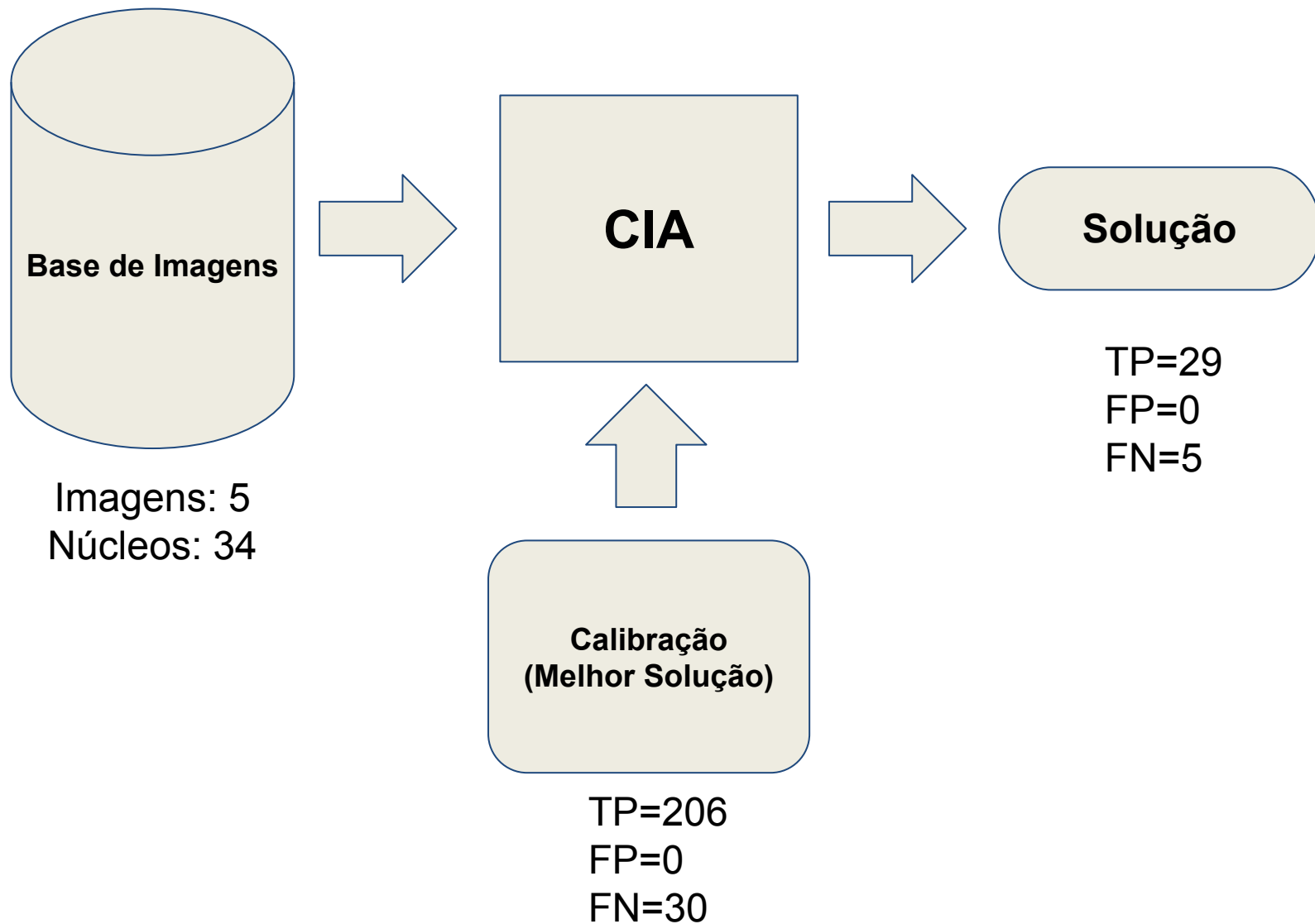
# Fronteira Pareto(*isbi\_train\_45*)




# Fronteira Pareto(*isbi\_train\_45*)


i	Intensity	Minimum Area	Maximum Area	Minimum Circularity	Maximum Circularity	TP	FN	FP
1	140	95	919	0,67	1,31	206	30	0
2	136	92	920	0,6073	1,26	209	27	6
3	153	63	626	0,55	1,49	215	21	96
4	140	95	1359	0,6073	1,49	211	25	12
5	140	95	626	0,5422	1,49	212	24	22
6	153	95	626	0,55	1,49	214	22	52
7	153	95	626	0,6073	1,3	213	23	42
8	136	95	626	0,5422	1,49	210	26	9
9	153	98	625	0,4651	1,25	217	19	182
10	153	92	626	0,476	1,26	216	20	172
11	153	63	626	0,4651	1,3	218	18	243

# Testes CIA(*isbi\_train\_45*)



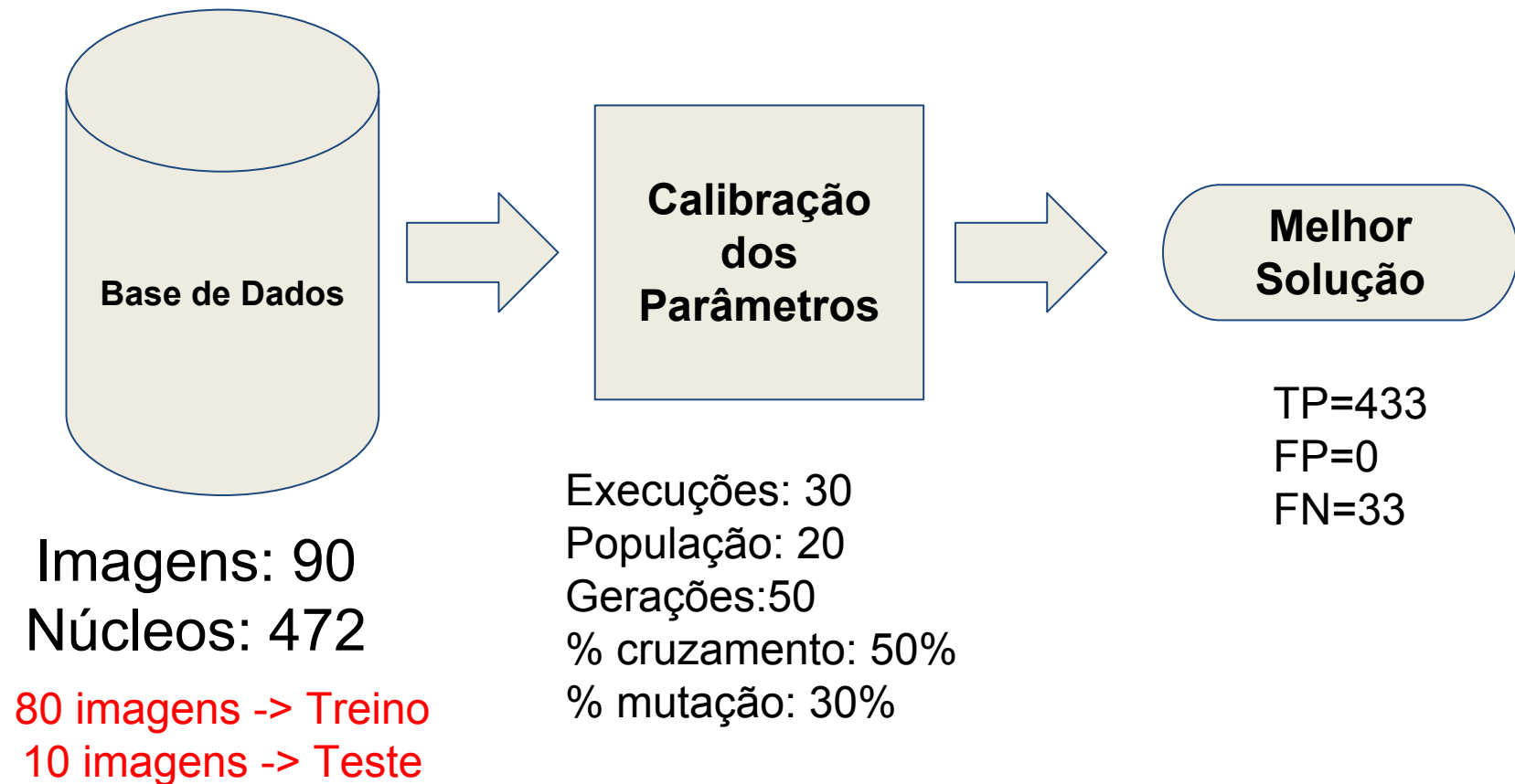
## 6.Resultados (*isbi\_train\_45*)

Metodo	Daniela et al.(2014)*	Método proposto
TP	23	29 
FP	0	0
FN	11	5
% Recall	67,65	85,29
% Precision	100,00	100,00
% F-Value	80,70	92,06

Recall - 17,64% 

F-Value - 11,36% 

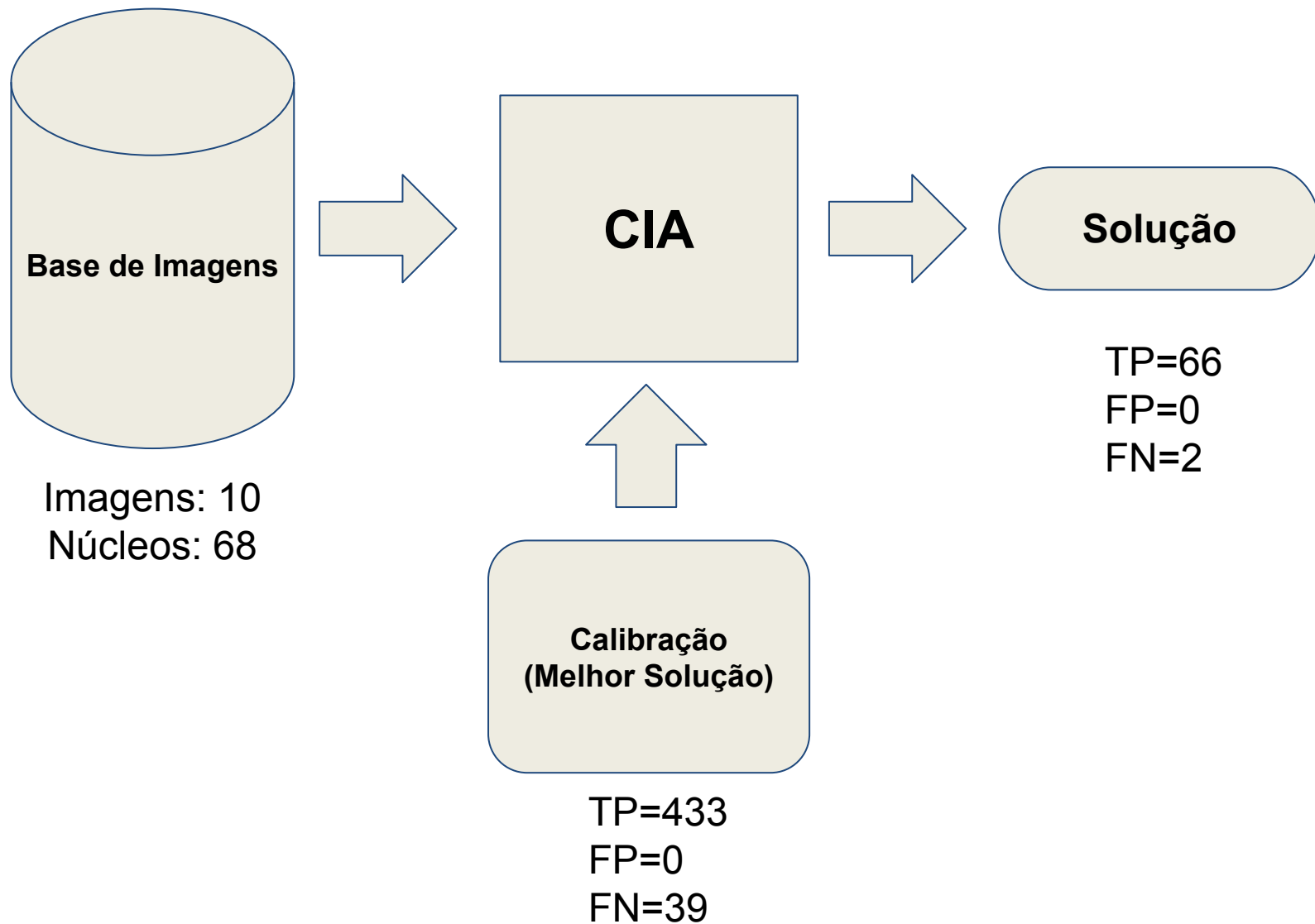
# Testes de Calibração(*isbi\_train\_90*)



# Fronteira Pareto(*isbi\_train\_90*)


i	Intensity	Minimum Area	Maximum Area	Minimum Circularity	Maximum Circularity	TP	FN	FP
1	139	45	914	0,3756	1,2200	433	39	0
2	142	57	914	0,3756	1,4000	439	33	7
3	142	46	906	0,3544	1,2200	440	32	21
4	144	15	913	0,3648	1,3990	442	30	74

# Testes CIA(*isbi\_train\_90*)





## 6.Resultados (*isbi\_train\_90*)

Metodo	Daniela et al.(2014)*	Método proposto
TP	40	66 
FP	0	0
FN	28	2
% Recall	58,82	97,06
% Precision	100,00	100,00
% F-Value	74,07	98,51

Recall - 38,24% 

F-Value - 24,44% 

# CRIC Searchable Database

- Base de imagens foi cedida pelo NUPEB - UFOP
- 400 imagens reais
- 11.480 núcleos
- 6 classes
  - Normais - 6792
  - ASC-US - 627
  - LSIL - 1336
  - ASC-H - 915
  - HSIL - 1687
  - Carcinoma - 123



[www.cricdatabase.com.br](http://www.cricdatabase.com.br)

# CRIC Searchable Database



CELL RECOGNITION  
FOR INSPECTION OF CERVIX

[Home](#) [About](#) [Segmentation Database](#) [Classification Database](#) [Download](#) [Segments Labels](#) [Publications](#) [Contact](#) [Welcome Paulo - Logout](#)

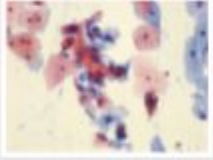

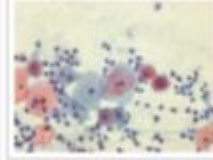

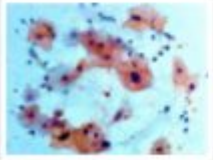
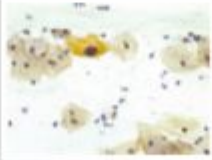



Images(Total: 400)

Click image to zoom and have access to specialist segmentation. Select images and segmentation specialist from list for download (available soon)

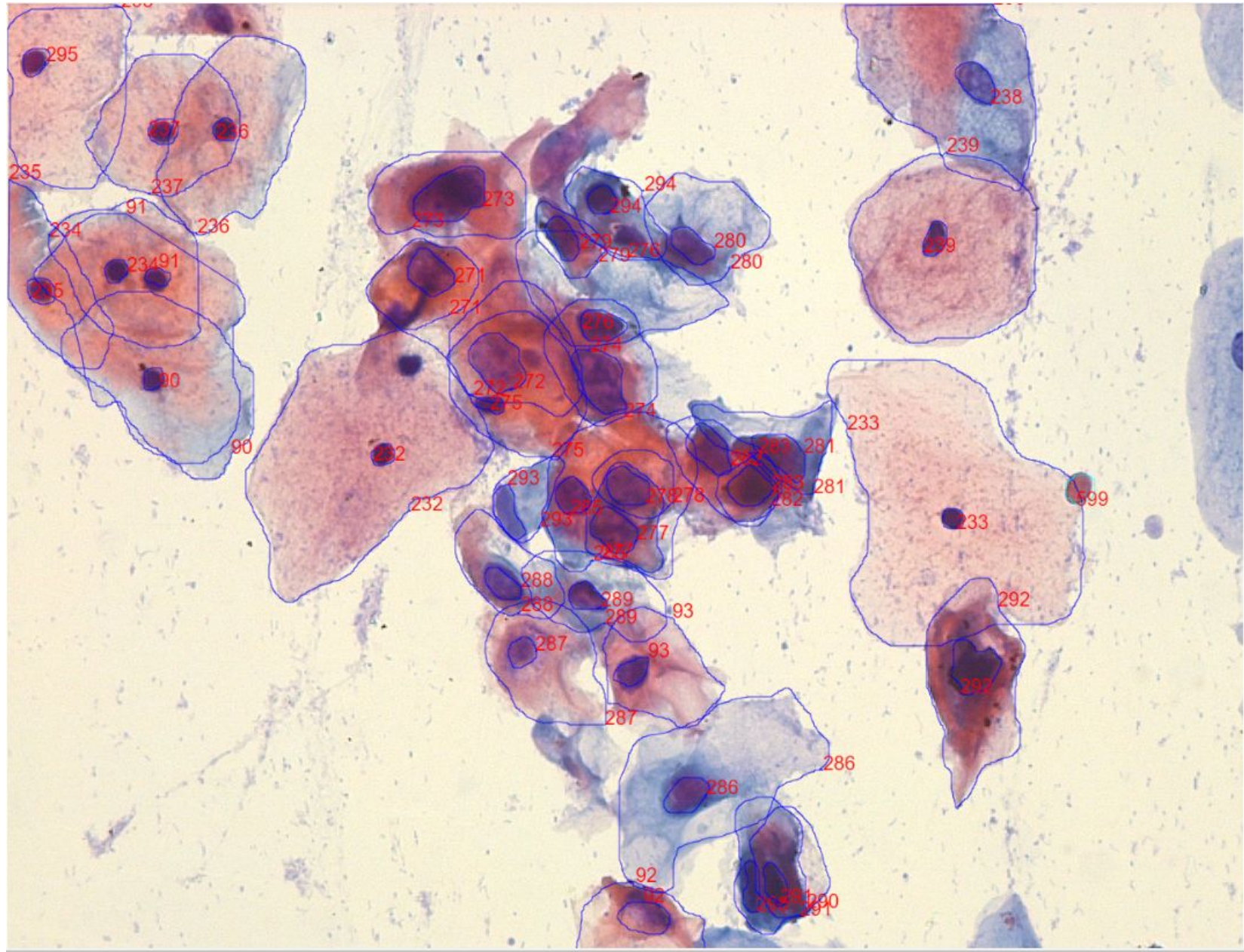
Normal - 6792  
ASC-US - 627  
LSIL - 1336

ASC-H - 915  
HSIL - 1687  
Carcinoma - 123

Label

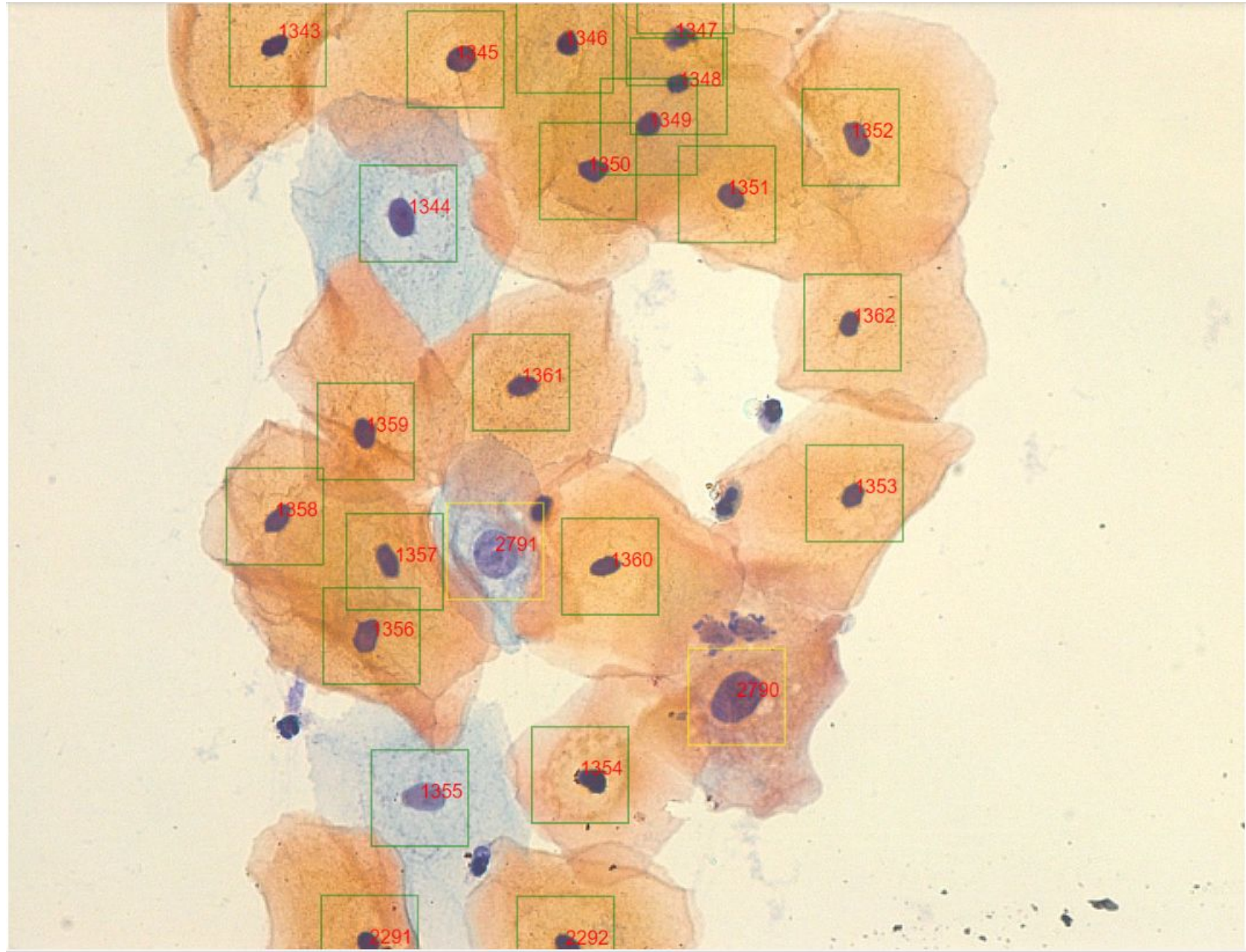
 <p>1 - LSIL 2564 - 14 <input type="button" value="Nucleus(270)"/> <input type="button" value="OK"/></p>	 <p>2 - HSIL 3236 - 14 <input type="button" value="Nucleus(171)"/> <input type="button" value="OK"/></p>	 <p>3 - ASC-US 1430 - 14 <input type="button" value="Nucleus(84)"/> <input type="button" value="OK"/></p>	 <p>4 - ASC-US 1430 - 14 <input type="button" value="Nucleus(68)"/> <input type="button" value="OK"/></p>	 <p>5 - ASC-H 2806 - 14 <input type="button" value="Nucleus(45)"/> <input type="button" value="OK"/></p>	 <p>6 - HSIL 2647 - 14 <input type="button" value="Nucleus(55)"/> <input type="button" value="OK"/></p>
 <p>7 - LSIL 775 - 14 <input type="button" value="Nucleus(103)"/> <input type="button" value="OK"/></p>	 <p>8 - LSIL 270 - 14 <input type="button" value="Nucleus(52)"/> <input type="button" value="OK"/></p>	 <p>9 - LSIL 270 - 14 <input type="button" value="Nucleus(119)"/> <input type="button" value="OK"/></p>	 <p>10 - LSIL 783 - 14 <input type="button" value="Nucleus(103)"/> <input type="button" value="OK"/></p>	 <p>11 - LSIL 783 - 14 <input type="button" value="Nucleus(50)"/> <input type="button" value="OK"/></p>	 <p>12 - LSIL 775 - 14 <input type="button" value="Nucleus(69)"/> <input type="button" value="OK"/></p>

# CRIC Searchable Database

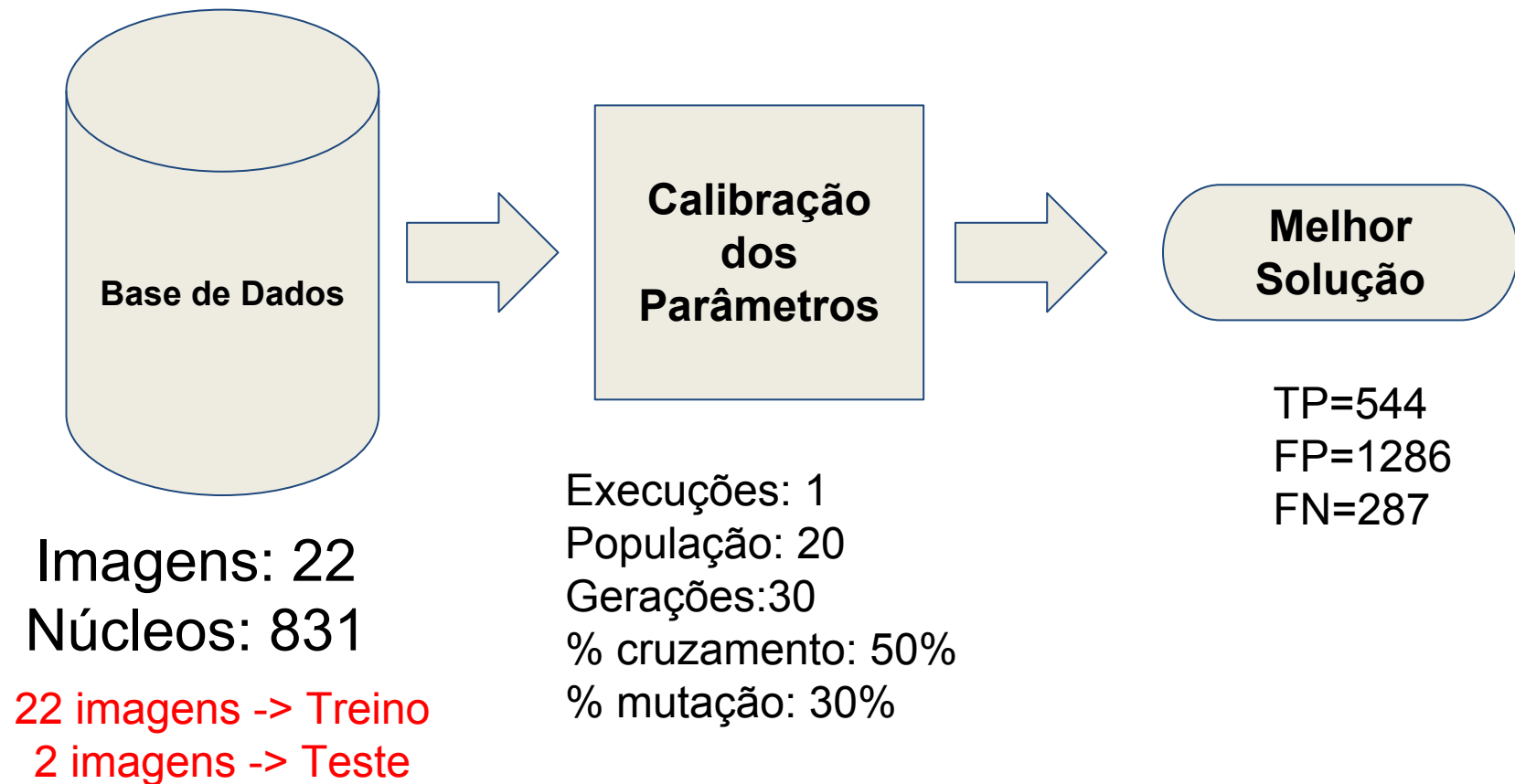




# CRIC Searchable Database



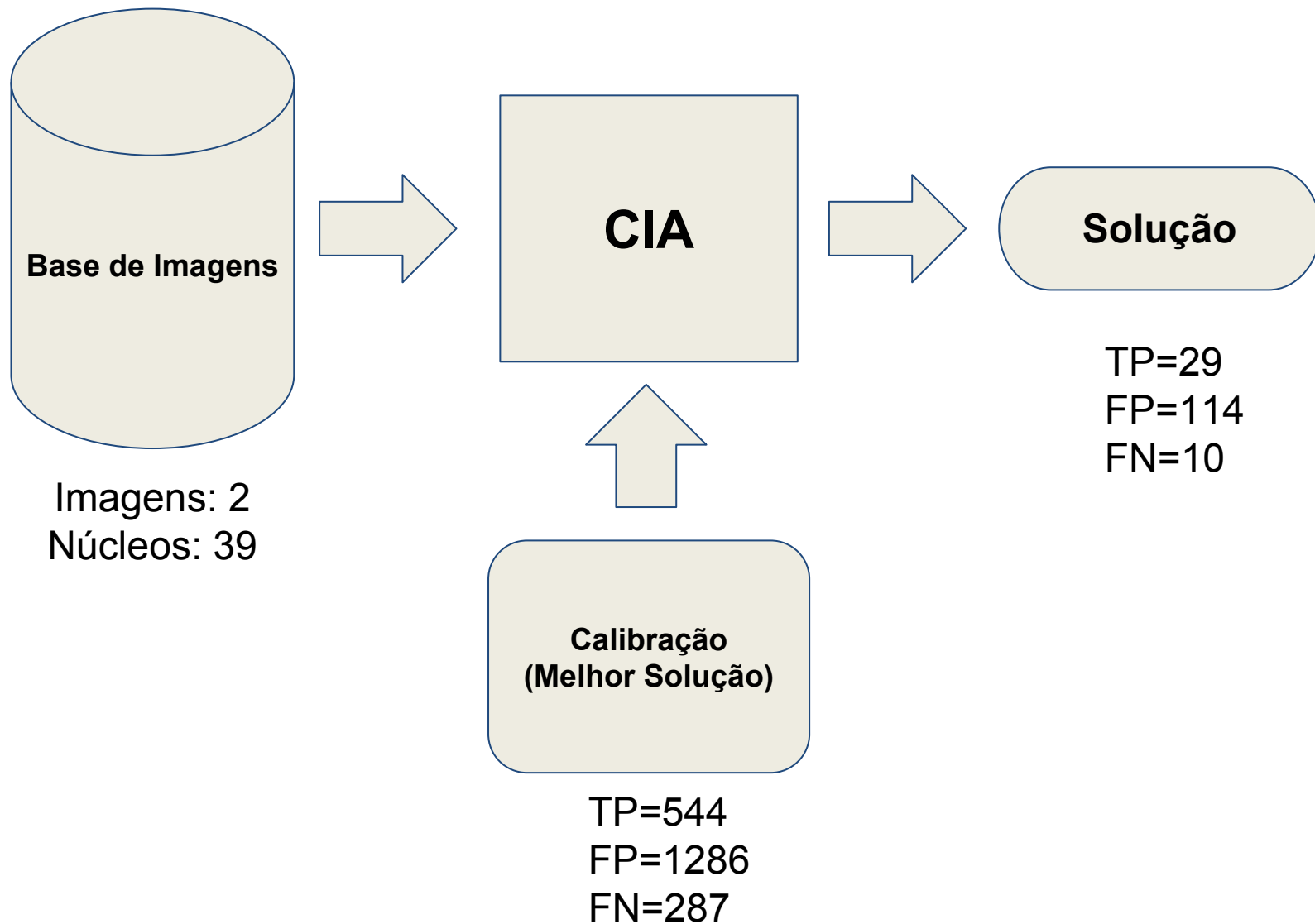
# Testes de Calibração(*CRIC Database*)



# Elementos da Fronteira Pareto(*CRIC Database*)

i	Intensity	Minimum Area	Maximum Area	Minimum Circularity	Maximum Circularity	TP	FN	FP
<b>1</b>	<b>168</b>	<b>253</b>	<b>1451</b>	<b>0,4173</b>	<b>1,2800</b>	<b>544</b>	<b>287</b>	<b>1286</b>
2	184	253	1341	0,4173	1,1455	551	280	1370
3	184	253	1455	0,4173	1,2800	563	268	1393
4	182	217	1450	0,4441	1,1000	566	265	1449
5	185	217	1455	0,4441	1,1000	569	262	1476
6	187	216	1451	0,4441	1,2800	571	260	1503
7	187	214	1451	0,4441	1,2287	573	258	1517
8	189	216	1451	0,4441	1,2800	574	257	1524
9	189	212	1346	0,4441	1,1503	575	256	1533
10	189	214	1451	0,4441	1,2287	576	255	1538
11	189	212	1451	0,4441	1,2800	580	251	1550
<b>12</b>	<b>184</b>	<b>216</b>	<b>1455</b>	<b>0,4173</b>	<b>1,2800</b>	<b>601</b>	<b>230</b>	<b>1627</b>

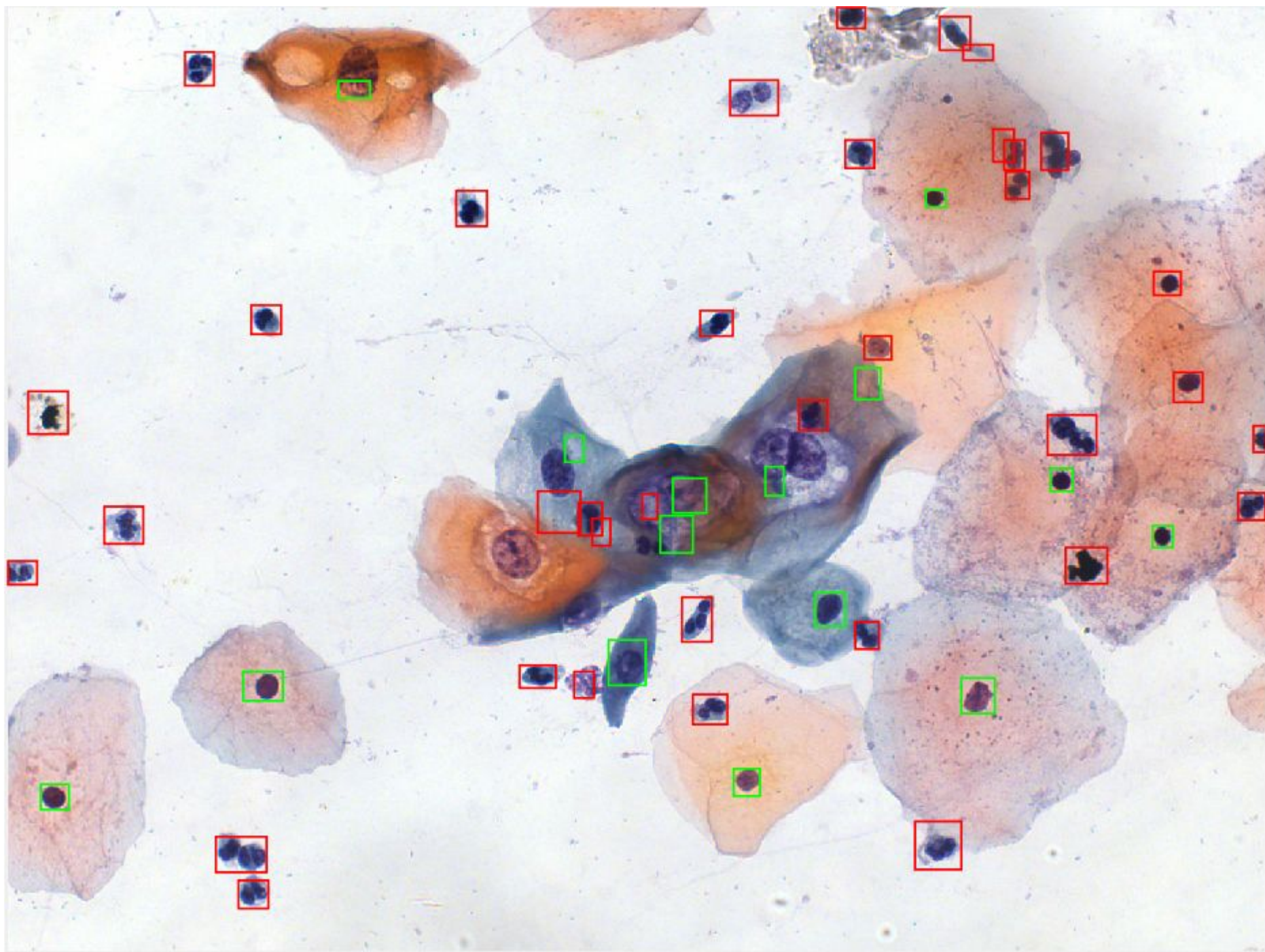
# Testes CIA(*CRIC Database*)



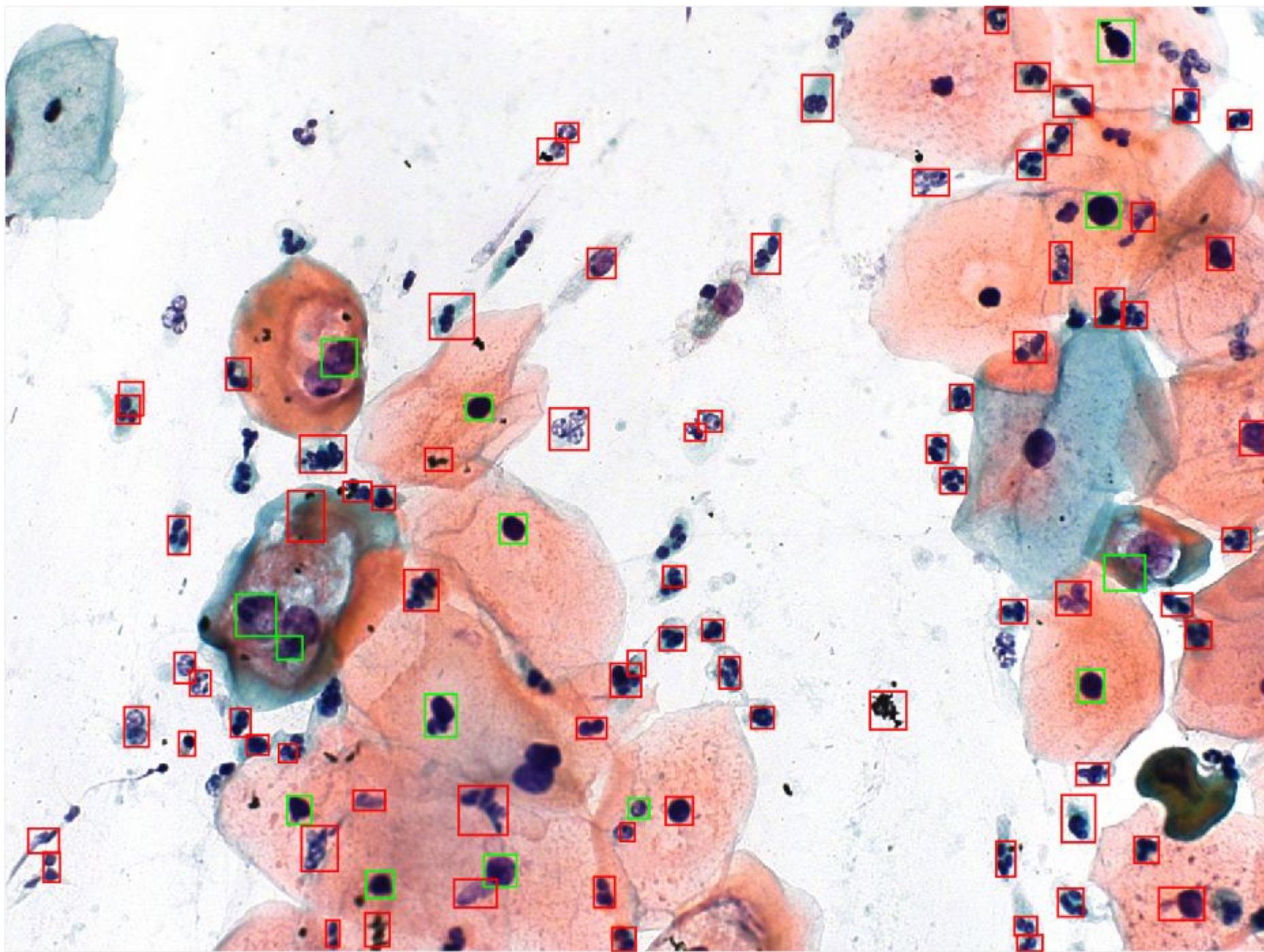


## 6.Resultados (*CRIC Database*)

Medida	Valor
TP	29
FP	114
FN	10
% Recall	74,36
% Precision	20,28
% F-Value	31,87







# 7. Conclusão

- Método eficiente para segmentar núcleos em imagens sintéticas
- Uma nova base de imagens reais e uma ferramenta colaborativa para criação da base
- Uso de algoritmos genéticos

Desvantagens:

- tempo de execução
- imagens reais

# Futuro

- Avaliar outros parâmetros que possam ser relevante e adicionados a heurística
- Investigar o funcionamento do algoritmo na base de imagens reais
- Acrescentar funcionalidades na ferramenta criada para gerar a base de imagens

# Agradecimentos

