



UFBA

UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS GRADUAÇÃO EM
ENGENHARIA INDUSTRIAL - PEI

MESTRADO EM ENGENHARIA INDUSTRIAL

ALINE SILVA RAMOS

Contagem de células somáticas em leite de búfalas
usando um classificador *fuzzy* e técnicas de
processamento de imagens



SALVADOR
2019



UFBA

UNIVERSIDADE FEDERAL DA BAHIA

ESCOLA POLITÉCNICA

PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA INDUSTRIAL – PEI

MESTRADO EM ENGENHARIA INDUSTRIAL

ALINE SILVA RAMOS

CONTAGEM DE CÉLULAS SOMÁTICAS EM
LEITE DE BÚFALAS USANDO UM
CLASSIFICADOR *FUZZY* E TÉCNICAS DE
PROCESSAMENTO DE IMAGENS



Salvador, 2019

ALINE SILVA RAMOS

**CONTAGEM DE CÉLULAS SOMÁTICAS EM LEITE DE
BÚFALAS USANDO UM CLASSIFICADOR FUZZY E TÉCNICAS
DE PROCESSAMENTO DE IMAGENS**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Industrial, da Universidade Federal da Bahia, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Industrial.

Orientador: Prof. Dr. Cristiano Fontes
Coorientadora: Prof.^a Dr.^a Viviani Gomes

Salvador
2019

Ficha catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA),
com os dados fornecidos pelo(a) autor(a).

Ramos, Aline Silva

Contagem de Células Somáticas em Leite de Búfalas
usando um Classificador Fuzzy e Técnicas de
Processamento de Imagens / Aline Silva Ramos. --
Salvador, 2019.
104 f. : il

Orientador: Cristiano Hora Fontes.

Coorientadora: Viviani Gomes.

Dissertação (Mestrado - Engenharia Industrial) --
Universidade Federal da Bahia, Escola Politécnica,
2019.

1. Leite de búfalas. 2. Células somáticas. 3.
Fuzzy C-Means. 4. Processamento de imagens. 5.
Transformada de Watershed. I. Hora Fontes, Cristiano.
II. Gomes, Viviani. III. Título.


CONTAGEM DE CÉLULAS SOMÁTICAS EM LEITE DE BÚFALAS USANDO UM CLASSIFICADOR FUZZY E TÉCNICAS DE PROCESSAMENTO DE IMAGENS


ALINE SILVA RAMOS


Dissertação submetida ao corpo docente do programa de pós-graduação em Engenharia Industrial da Universidade Federal da Bahia como parte dos requisitos necessários para a obtenção do grau de mestre em Engenharia Industrial.

Examinada por:

Prof. Dr. Cristiano Hora de Oliveira Fontes 
Doutor em Eng. Química, pela Universidade Estadual de Campinas, BRASIL, 2001.

Prof^a. Dr^a. Viviani Gomes 
Doutora em Clínica Veterinária, pela FMVZ da Universidade de São Paulo - USP, BRASIL, 2008.

Prof^a. Dr^a. Karina Medici Madureira 
Doutora em Clínica Veterinária, pela FMVZ da Universidade de São Paulo - USP, BRASIL, 2012.

Prof. Dr. Ricardo Araújo Rios 
Doutor em Ciência da Computação, pelo Instituto de Ciências Matemáticas e Computação da Universidade de São Paulo, BRASIL, 2013.

Salvador, BA - BRASIL
Maio/2019

Dedico este trabalho a Deus, família, amigos, colegas de trabalho e orientadores pelo apoio, força, incentivo, companheirismo e amizade. Sem eles nada seria possível.

Agradecimentos

- A Deus por me dar força interior, sabedoria e confiança para superar todas as dificuldades.
- À minha mãe Jânia e minhas tias Jussara e Elza, pelo amor e carinho incondicional e pelos ensinamentos obtidos, estes estarão comigo por toda vida!
- À minha amiga Karine e meu amigo Gabriel, por me apoiar em um dos momentos mais difíceis desta etapa da minha vida, pela cumplicidade e afeto, por me escutar e me ajudar sempre que possível.
- À minha amiga Denise, por toda ajuda, companheirismo e alegria.
- Às minhas irmãs Rojânia, Jamilla e Gabriela por confiarem e acreditarem em mim em todo o tempo.
- A meu orientador, professor Cristiano Hora Fontes, por todo empenho, sabedoria e compreensão. E a minha orientadora, professora Viviani Gomes por toda dedicação.
- Às minhas amigas Ana Célia, Francielle e Nina pelos ensinamentos, conselhos e apoio.
- Aos professores, funcionários e colegas do Programa de Pós-graduação em Engenharia Industrial da UFBA, em especial a Tatiane, que sempre me acolheu tão bem e com tanto carinho e ao professor Adonias Magdiel por toda ajuda dispensada.
- Aos meus amigos, pela amizade e pelos bons momentos de descontração.

"Não fui eu que ordenei a você? Seja forte e corajoso! Não se apavore nem desanime, pois o Senhor, o seu Deus, estará com você por onde você andar". (JOSUÉ 1:9)

Resumo da Dissertação apresentada ao PEI/UFBA como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

CONTAGEM DE CÉLULAS SOMÁTICAS EM LEITE DE BÚFALAS UTILIZANDO UM CLASSIFICADOR *FUZZY* E TÉCNICAS DE PROCESSAMENTO DE IMAGENS

Aline Silva Ramos
Maio/2019

Orientador: Prof. Dr. Cristiano Hora Fontes
Coorientadora: Prof.^a Dr.^a Viviani Gomes

A produção do leite de búfalas e seus derivados vem aumentando no Brasil e no mundo, juntamente com o aumento das exigências em relação ao seu padrão de qualidade. A mastite, doença inflamatória da glândula mamária (GM), é responsável por perdas qualitativas e quantitativas em relação ao leite produzido. A contagem de células somáticas (CCS) no leite é o principal biomarcador tanto para a detecção quanto para a avaliação da qualidade do leite. A CCS é determinada tradicionalmente por métodos trabalhosos que consistem na observação visual das células em esfregaços lácteos através do microscópio. Esta técnica tradicional é exaustiva e possui um inerente grau de subjetividade na medida em que está sujeita à interpretação humana em relação à identificação e reconhecimento das células. Este trabalho propõe e apresenta um método automático para a contagem das células somáticas (CCS) no leite de búfalas que compreende, entre outros, a aplicação de um método de agrupamento *Fuzzy* e técnicas de processamento de imagens. Ao contrário de outros trabalhos similares da literatura, o *Fuzzy C-Means* foi utilizado na etapa de pré-processamento das imagens e não na etapa de segmentação das mesmas. Esta abordagem viabilizou a separação das imagens (objetos) das células somáticas em leite de búfalas em grupos que apresentassem similaridades em relação à intensidade de cor, possibilitando uma melhor aplicação posterior das técnicas de processamento como a limiarização, segmentação e reconhecimento/interpretação das imagens de células somáticas. Três métodos de limiarização foram avaliados e comparados e a Transformada de *Watershed* foi utilizada para a separação de células bastante próximas, o que contribuiu para

a correta identificação e contagem das mesmas. Por fim, realizou-se uma comparação entre os resultados obtidos pela contagem manual (técnica microscópica direta) e pelo método proposto neste trabalho. Foi utilizada uma prova estatística não-paramétrica (*Kruskal Wallis*) que comprovou a obtenção de resultados de contagens consistentes. Em relação ao método padrão empregado para a contagem das células somáticas do leite, a utilização do *Fuzzy C-means* no pré-processamento das imagens revelou ser uma alternativa potencial e eficiente para o agrupamento das imagens em grupos que apresentam similaridade na intensidade de cor, o que proporciona um melhor desempenho do processo de limiarização e conseqüentemente da contagem das células somáticas nas imagens.

Abstract of Dissertation presented to PEI/UFBA as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

SOMATIC CELL COUNT IN BUFFALOS MILK USING A FUZZY CLASSIFIER AND IMAGE PROCESSING TECHNIQUES

Aline Silva Ramos
May/2019

Advisors: Prof. Dr. Cristiano Hora Fontes
Prof.^a Dr.^a Viviani Gomes

Milk production of buffaloes and their derivatives has been increasing in Brazil and in the world, together with the increasing demands on its quality standard. Mastitis, inflammatory disease of the mammary gland (GM), is responsible for qualitative and quantitative losses in relation to the milk produced. The somatic cell count (CCS) in milk is the main biomarker for both detection and evaluation of milk quality. CCS is traditionally determined by laborious methods consisting of the visual observation of cells in milk smears through the microscope. This traditional technique is exhaustive and has an inherent degree of subjectivity in that it is subject to human interpretation in relation to the identification and recognition of cells. For this reason, this research proposes and presents an automatic method for counting somatic cells in buffalo milk which includes, among others, the application of a Fuzzy clustering method and image processing techniques. Unlike other similar works, the Fuzzy C-Means was used in the preprocessing stage of the images and not in the segmentation stage of the images. This approach enabled the separation of the somatic cell images (objects) present in buffalo milk in clusters that showed similarities in relation to the color intensity, allowing a better posterior application of processing techniques such as thresholding, segmentation and image recognition (interpretation of somatic cells). Three methods of thresholding were evaluated and compared, and the Watershed Transform was used to separate cells closely together, which contributed to the correct identification and counting of the same. Finally, a comparison was made between the results obtained by manual counting by the direct microscopic technique and by the method proposed in this work. A non-parametric statistical test (Kruskal Wallis) was used, which proved

to obtain consistent counts results. The use of a Fuzzy C-means in the preprocessing of the images was a potential and efficient alternative for the clustering of images in clusters that show similarity in color intensity, which provides a better performance of the thresholding process and consequently the somatic cell count in the images.

Lista de Figuras

Figura 1: Distribuição de bubalinos em diferentes estados do Brasil (Adaptado do MAPA, 2017).....	18
Figura 2: Produção e venda de leite bubalino em 2006 (IBGE - CENSO AGROPECUÁRIO, 2006).....	19
Figura 3: Produção Mundial de leite de búfalas (Adaptado de FAO. Elaboração: Intelactus/Embrapa Gado de Leite, 2018).....	19
Figura 4: Migração das células de defesa para o tecido mamário (HULSEN & LAM, 2007).....	21
Figura 5: Células somáticas em leite de búfala (VIVIANI GOMES, 2016).....	26
Figura 6: Úbere inchado e vermelho, devido ao processo inflamatório da glândula mamária (VIVIANI GOMES, 2019).....	27
Figura 7: Campo microscópico observado em uma lâmina de leite de búfala em diferentes aumentos, (Imagem da esquerda em aumento de 400x e imagem da direita em aumento de 1000x).....	30
Figura 8: Citometria de fluxo (Adaptado de DANIELLE LIMA, 2010).....	31
Figura 9: Representação de uma imagem digital bidimensional (DE QUEIROZ E GOMES, 2006).....	33
Figura 10: Sensibilidade do olho humano para as cores primárias (HUNT, 1995).....	34
Figura 11: Criando cores com pixels RGB: (a) Pixels branco, preto, azul claro e amarelo ouro; (b) Componentes RGB (Intensidade de cada cor primária).....	35
Figura 12: Posicionamento dos eixos para representação de imagens (GONZALEZ AND WOODS, 2000).....	35
Figura 13: Os componentes iluminância (I) e refletância (R) de uma imagem (VIEIRA E MARQUES, 1999).....	36
Figura 14: Uma imagem digital monocromática: (a) Representação gráfica usual; (b) Representação matricial (BATISTA, 2005).....	37
Figura 15: Representação do pixel (CONCI, 2015).....	37
Figura 16: Imagem em RGB: (a) Representação gráfica usual; (b) Matriz da cor vermelha; (c) Matriz da cor verde; (d) Matriz da cor azul.....	38
Figura 17: Imagem original e seu histograma (SOUZA E CORREIA, 2007).....	39
Figura 18: Histograma: (a) Original; (b) Modificado (VIEIRA E MARQUES, 1999).....	40
Figura 19: Método de Otsu (OTSU, 1979).....	42
Figura 20: Método de Limiarização de Rosin (ROSIN, 2001).....	42
Figura 21: Método Rosin: (a) Histograma unimodal; (b) Histograma Bimodal (Jessie NG et al., 2009).....	43
Figura 22: Limiarização: Método de Melo (DE MELO, 2015).....	44
Figura 23: Segmentação: (a) Imagem original; (b) Imagem monocromática; (c) Imagem binária segmentada (VIVIANI GOMES, 2016).....	45
Figura 24: Princípio de imersão de uma imagem: (a) Imagem original e seus mínimos; (b) Linhas divisórias e regiões; (c) Imagem imersa (PINTO, 2014).....	46
Figura 25: Transformada de Watershed aplicada em imagem de células somáticas (DE MELO, 2015).....	47
Figura 26: Montagem de lâminas (esfregaços) de leite.....	51
Figura 27: (a) Observação das lâminas e (b) Sistema de aquisição de imagens.....	51
Figura 28: Resumo do método utilizado para contagem de células somáticas.....	52
Figura 29: Representação da imagem RGB para conversão: (a) Imagem original; (b) Área menor da imagem RGB; (c) Parte das matrizes que representam a área menor da imagem RGB.....	53
Figura 30: Imagem convertida em monocromática: (a) Imagem monocromática; (b) Área menor da imagem monocromática; (c) Parte da matriz que representa a área menor da imagem monocromática.....	54

Figura 31: Representação da intensidade da imagem: (a) Imagem monocromática; (b) Matriz da imagem; (c) Histograma da imagem.	55
Figura 32: Matrizes: (a) Parte da matriz com os valores dos pixels; (b) Parte da matriz com as frequências dos pixels; (c) Parte da matriz com a distribuição da intensidade de cinza da imagem.	55
Figura 33: Exemplos de imagens das amostras (VIVIANI GOMES, 2016).	56
Figura 34: Matrizes: (a) Parte da matriz com a distribuição da intensidade de cinza da imagem; (b) Parte da matriz com os valores dos pixels; (c) Parte da matriz com as frequências dos pixels.	57
Figura 35: Exemplos de imagens tratadas com o mesmo limiar.....	57
Figura 36: Procedimento para a determinação do limiar de intensidade em um histograma unimodal (DE MELO, 2015).....	58
Figura 37: Índice de correlação C_n em função do parâmetro de ponderação L (DE MELO, 2015).....	59
Figura 38: Imagem de Células Somáticas: (a) Duas células; (b) União dos pixels das células formando apenas uma (VIVIANI GOMES, 2016).....	60
Figura 39: Etapa de contagem das células: (a) Imagem original; (b) Reconhecimento dos objetos na imagem; (c) Reconhecimento do resíduo; (d) Retirada do resíduo, deixando as células.	61
Figura 40: Células somáticas: (a) Imagem RGB original; (b) Imagem transformada em monocromática; (c) Parte da matriz da imagem monocromática; (d) Histograma da imagem monocromática; (e) Parte da matriz com a distribuição da intensidade de cinza da imagem.	62
Figura 41: Histograma gerado a partir de informações da matriz do centro do grupo 1.	63
Figura 42: Histograma gerado a partir de informações da matriz do centro do grupo 2.	64
Figura 43: Histograma gerado a partir de informações da matriz do centro do grupo 3.	64
Figura 44: Histograma gerado a partir de informações da matriz do centro do grupo 4.	64
Figura 45: Histograma gerado a partir de informações da matriz do centro do grupo 5.	65
Figura 46: Gráfico das pertinências das imagens nos grupos.....	65
Figura 47: Uso de Watershed para separar células em grupos.....	69
Figura 48: Exemplos de histogramas unimodais das imagens tratadas.	70
Figura 49: Exemplo de uma imagem original: (a) Tratada com Limiar de Rosin; (b) Tratada com Limiar de Melo.	71
Figura 50: Boxplot comparativo entre CCSFPI (com limiar de Melo) e especialistas.....	72
Figura 51: Teste de normalidade com os dados da contagem do método CCSFPI.	73
Figura 52: Teste de normalidade com os dados da contagem do especialista 1.	73
Figura 53: Teste de normalidade com os dados da contagem do especialista 2.	74
Figura 54: Teste de normalidade com os dados da contagem do especialista 3.	74
Figura 55: Imagem tirada com iluminação inadequada: (a) Imagem escura demais; (b) Contagem das células com erro.	77
Figura 56: Amostra com excesso de gordura: (a) Células com núcleos desbotados;(b) Contagem das células com erro.	77

Lista de Tabelas

<i>Tabela 1: Quantidade de imagens capturadas por amostra de leite.....</i>	<i>51</i>
<i>Tabela 2: Número de imagens por grupo separadas pelo FCM.</i>	<i>63</i>
<i>Tabela 3: Índices de Xie Beni.....</i>	<i>66</i>
<i>Tabela 4: Valores de L (Limiar Melo).....</i>	<i>67</i>
<i>Tabela 5: Médias dos limiares de Melo, Otsu e Rosin de cada grupo.</i>	<i>67</i>
<i>Tabela 6: Limiares obtidos sem o uso do Classificador FCM</i>	<i>68</i>
<i>Tabela 7: Contagem final das células somáticas - Limiar Melo.....</i>	<i>69</i>
<i>Tabela 8: Contagem final das células somáticas – Limiar Otsu.....</i>	<i>69</i>
<i>Tabela 9: Contagem final das células somáticas – Limiar Rosin</i>	<i>70</i>
<i>Tabela 10: Contagem das células somáticas sem uso do FCM.....</i>	<i>71</i>
<i>Tabela 11: Contagens células (Limiar Melo).....</i>	<i>76</i>
<i>Tabela 12: Falsos positivos e falsos negativos (Limiar Melo).....</i>	<i>76</i>

Lista de Abreviaturas

ABCB	Associação Brasileira de Criação de Búfalas
CCS	Contagem de Células Somáticas
CCSFPI	Contagem de Células Somáticas utilizando <i>Fuzzy</i> e Processamento de Imagem
CMT	<i>California Mastitis Test</i>
DMSCC	<i>Direct Microscopy Somatic Cell Counter</i>
DNA	<i>Deoxyribonucleic Acid</i>
FCM	<i>Fuzzy C-Means</i>
FMVZ	Faculdade de Medicina Veterinária e Zootecnia
GIF	<i>Graphics Interchange Format</i>
HE	Histograma Equalizado
IBGE	Instituto Brasileiro de Geografia e Estatística
JPEG	<i>Joint Photographics Experts Group</i>
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
PH	Potencial Hidrogeniônico
RGB	<i>Red – Green - Blue</i>
TW	Transformada <i>de Watershed</i>
USP	Universidade de São Paulo
WMT	<i>Wiscosin Mastitis Test</i>

Sumário

CAPÍTULO 1. INTRODUÇÃO	17
1.1 CONSIDERAÇÕES INICIAIS	17
1.2 OBJETIVOS	25
CAPÍTULO 2. FUNDAMENTAÇÃO TEÓRICA E REVISÃO BIBLIOGRÁFICA	26
2.1 CÉLULAS SOMÁTICAS	26
2.1.1 Contagem de células somáticas	29
2.1.2 Métodos de detecção de mastite	29
2.2 TÉCNICAS DE PROCESSAMENTO DE IMAGENS.....	32
2.2.1 Visão geral.....	32
2.2.2 Sistema RGB e imagens monocromáticas	34
2.2.3 Histograma.....	39
2.2.4 Limiarização	41
2.2.5 Segmentação e extração de características/ reconhecimento/interpretação	44
2.2.6 Transformada de <i>Watershed</i> (TW)	45
2.3 FUZZY C-MEANS	47
CAPÍTULO 3. METODOLOGIA E RESULTADOS.....	50
3.1 RESULTADOS	61
CAPÍTULO 4. CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS	78
4.1 CONCLUSÕES	78
4.2 SUGESTÕES PARA TRABALHOS FUTUROS.....	79
REFERÊNCIAS BIBLIOGRÁFICAS	80
APÊNDICE A. IMAGENS DOS GRUPOS	92

CAPÍTULO 1

INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

Segundo Ricci e Domingues (2012), as búfalas representam uma importante contribuição para a produção de leite e possuem destaque na atividade agrícola como fornecedoras de alimentos, tais como a carne. A Índia é um dos principais países produtores de leite bubalino (SINDHU e ARORA, 2011). A Itália é o país mais desenvolvido em relação à produção de leite de búfala na Europa tendo como principal produto a muçarela (PITACAS & RODRIGUES, 2013). No continente Americano, os países que mais se destacam em relação à quantidade desses animais são a Colômbia, Brasil, Venezuela e Argentina, cuja produção contempla a venda de leite, alguns subprodutos (muçarela, manteiga, coalhada, iogurte, parmesão, ricota) e carne (JORGE *et al.*, 2011). No Brasil, o rebanho de búfalos é estimado em até 1,8 milhões de animais espalhados por todas as regiões (EMBRAPA, 2018).

De acordo com a Organização das Nações Unidas para Agricultura e Alimentação, a produção mundial de leite em 2016 foi de 798 milhões de toneladas, onde 83% foi de leite de vaca, 14% de leite de búfalas, 2% de cabra, 1% de ovelha e menos e 1% de camela. Já o IFCN, uma rede de pesquisadores que estudam a produção de leite no mundo, avaliou que o volume oriundo da vaca e das búfalas somam 97% do total, e outras espécies, como cabra, ovelha e camela, produziram 3% (EMBRAPA, 2018).

Segundo a Organização das Nações Unidas para Agricultura e Alimentação - FAO, em 2017 o rebanho mundial de bubalinos era de 200.967.747 cabeças. Dentre os países com maiores rebanhos, os que mais se destacam são: Índia (110.000.000 cabeças), Paquistão (34.553.000 cabeças), China (23.345.000 cabeças), Bangladesh

(14.570.000 cabeças), Egito (3.949.262 animais) (SEAB - SECRETARIA DE ESTADO DA AGRICULTURA E DO ABASTECIMENTO, 2017).

A cultura de criar búfalos está se desenvolvendo no Brasil como uma alternativa rentável e saudável do ponto de vista nutricional. Apesar do Brasil ter um rebanho de búfalos pequeno em comparação ao rebanho bovino (ANUALPEC, 2009), a bubalinocultura representa uma importante atividade econômica devido à facilidade de adaptação dos animais às condições climáticas do país; potencial de rentabilidade devido à expansão comercial do mercado da carne de búfalo; características do animal (dócil e rústico); elevada qualidade do leite e derivados (ANDRIGHETTO, 2011). Segundo a Associação Brasileira de Criadores de Búfalos (ABCB), o custo da criação de búfalos chega a ser 20% menor que a de bovinos e para o produtor os ganhos são maiores com o leite de búfala, devido ao alto teor de sólidos, representando um rendimento industrial de 50% a 80% superior ao do leite bovino (SOCIEDADE NACIONAL DA AGRICULTURA, 2013).

O Brasil tem o maior rebanho de bubalinos do Ocidente (BERNARDES, 2007) e desponta na condição de 10º no rebanho mundial, representando aproximadamente 0,68% desse rebanho. O Estado do Amazonas teve destaque no período de 2008 a 2013 com um crescimento da produção de 50,5% (região Norte com crescimento da ordem de 22,3%) e o Brasil como um todo teve o incremento da ordem de 16,2% (SEAB, 2017). A Figura 1 mostra um gráfico com a quantidade de cabeças de búfalo no Distrito Federal e nos estados brasileiros.

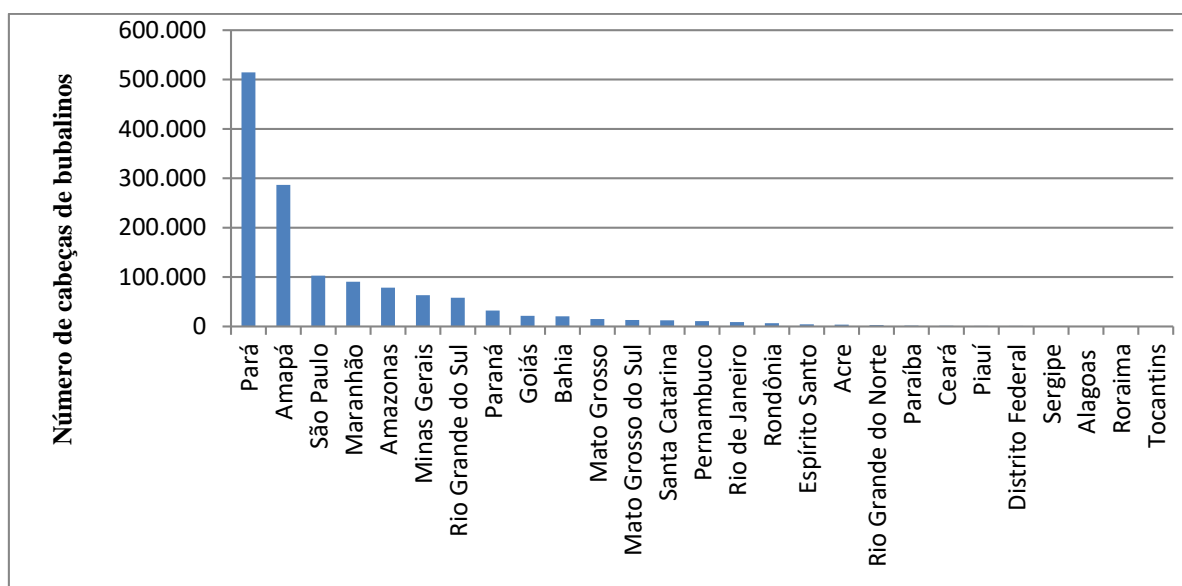


Figura 1: Distribuição de bubalinos em diferentes estados do Brasil (Adaptado do MAPA, 2017).

De acordo com o último Censo Agropecuário do Instituto Brasileiro de Geografia e Estatística (IBGE, 2006), relacionado à produção do leite de búfala, os produtores brasileiros venderam mais de 35 mil litros de leite bubalino. A Figura 2 mostra o desempenho das vendas em todas as regiões do Brasil.

Brasil e Grande Região	Número de estabelecimentos agropecuários que produziram leite de búfala no ano (unidades)	Búfalas ordenhadas (cabeças)	Quantidade produzida de leite de búfala no ano (mil litros)	Quantidade vendida de leite de búfala no ano (mil litros)	Valor da venda de leite de búfala no ano (mil reais)	Valor do litro de leite de búfala (RS)
Norte	1.298	23.140	18.929	13.963	10.862	1.28
Nordeste	420	4.855	3.491	2.874	2.018	1.42
Sudeste	632	15.623	18.032	16.347	12.233	1.33
Sul	161	1.799	1.405	1.233	782	1.57
Centro-Oeste	285	3.447	4.098	3.672	1.937	1.89
Brasil	2.796	48.864	45.955	38.087	27.832	1.36

Figura 2: Produção e venda de leite bubalino em 2006 (IBGE - CENSO AGROPECUÁRIO, 2006).

Na Figura 3, pode-se observar a evolução da produção mundial de leite de búfalas entre os anos de 1990 e 2016.

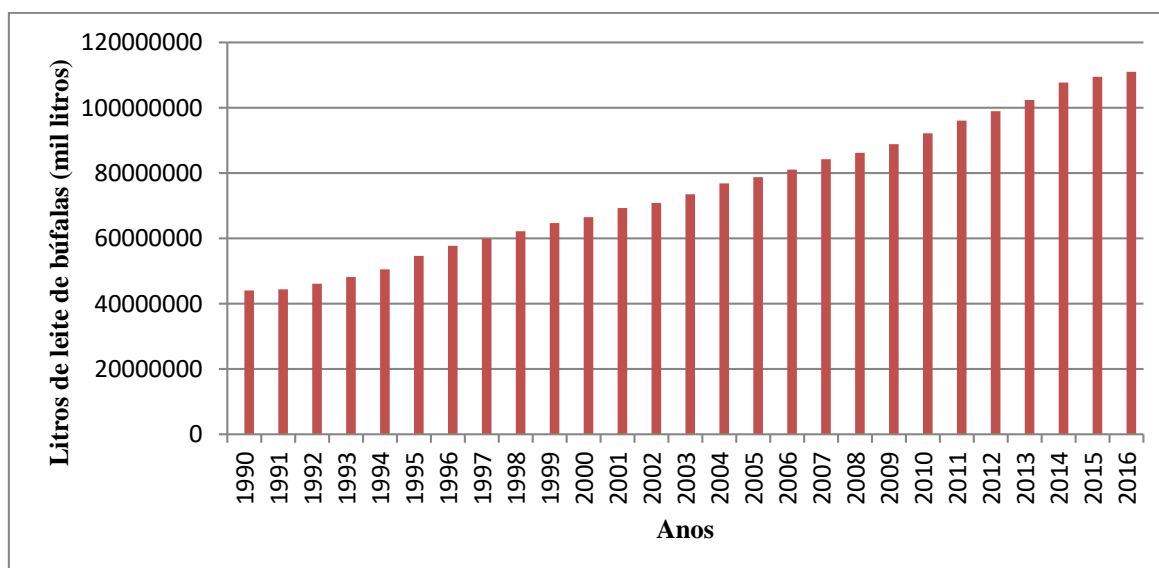


Figura 3: Produção Mundial de leite de búfalas (Adaptado de FAO. Elaboração: Intelactus/Embrapa Gado de Leite, 2018).

O leite de búfala possui alta qualidade nutricional com propriedades específicas que o tornam diferente de qualquer outro tipo de leite oriundo de outras espécies. Isso se

deve aos percentuais de suas proteínas, lipídeos, lactose, sólidos totais e resíduos minerais fixos, sendo o leite de búfala mais concentrado e com menor teor de água em relação ao leite bovino. Além disso, possui uma tonalidade branca devido à falta de betacaroteno, sem que haja alteração em seu poder nutritivo, e apresenta um sabor adocicado. Sua alta quantidade de cálcio o torna especial para o consumo humano, sendo indicado para prevenção da osteoporose (DUBEY, 1997; MACEDO *et al.*, 2001).

Devido ao maior teor de gordura, o leite bubalino é avaliado como sendo de 40% a 50% mais eficiente na fabricação de derivados, como manteiga, queijos, iogurte, doce de leite e outros. Se compararmos a quantidade de leite para a fabricação de 1 Kg de manteiga, são necessários 14 litros de leite de búfala, enquanto que seriam necessários 20 litros de leite de vaca (SILVA *et al.* 2003; MEENA *et al.* 2014).

A análise e o monitoramento da qualidade do leite de búfala são exigidos pelos órgãos de controle no Brasil (Ministério da Agricultura, Pecuária e Abastecimento MAPA), garantindo a segurança no consumo de leite e seus derivados cuja instrução normativa seguida é a nº 76, de 26 de novembro de 2018.

Uma doença comum que pode afetar a quantidade e qualidade do leite é a mastite. Estudos realizados demonstraram que as búfalas são menos suscetíveis à mastite se comparado com as vacas (DELLA LIBERA *et al.*, 2005; KEE, 2012); JINGAR *et al.*, 2014). A resistência das búfalas à mastite tem sido associada a algumas peculiaridades em relação às características do seu úbere e dos tetos. As búfalas possuem uma maior concentração de melanina, o canal do teto composto por epitélio estratificado queratinoso é mais espesso se comparado com a vaca, além de possuírem menor diâmetro do lúmen do canal do teto (UPPAL *et al.*, 1994). Outra característica marcante é o fato do canal dos tetos ser mais musculoso, onde se observa uma quantidade maior de vasos sanguíneos e de fibras, que funciona como uma barreira de proteção mais eficiente contra infecções (KAPRONEZAI *et al.*, 2005). Apesar destas vantagens, o incremento na produção de leite de búfalas trouxe problemas que não eram observados. Assim como a vaca, a búfala também está sujeita à mastite, e por falta de conhecimento, muitas vezes as medidas sanitárias que são adotadas para as vacas adotam-se para as búfalas, o que é inadequado, pois são animais com características diferenciadas (MESQUITA *et al.*, 2016).

A mastite é um processo inflamatório da glândula mamária, normalmente originária de um processo de infecção bacteriana, sendo classificada em dois tipos:

clínica e subclínica. Na forma clínica, ocorrem alterações inflamatórias do úbere (calor, rubor e dor) e pode-se ter a presença de grumos, pus ou sangue no leite. Em um estágio avançado, pode apresentar além das alterações no leite, alterações sistêmicas, tais como febre, depressão, fraqueza, pulsação fraca. Na forma subclínica, apesar do úbere se mostrar normal, o leite possui alterações em sua composição caracterizada principalmente pelo aumento da quantidade de células somáticas (ACOSTA *et al.*, 2016), em sua maioria leucócitos, que migram da circulação sanguínea para o tecido mamário (Figura 4). Desta forma, a Contagem de Células Somáticas (CCS) se constitui em uma análise biomarcadora para a detecção da mastite subclínica.

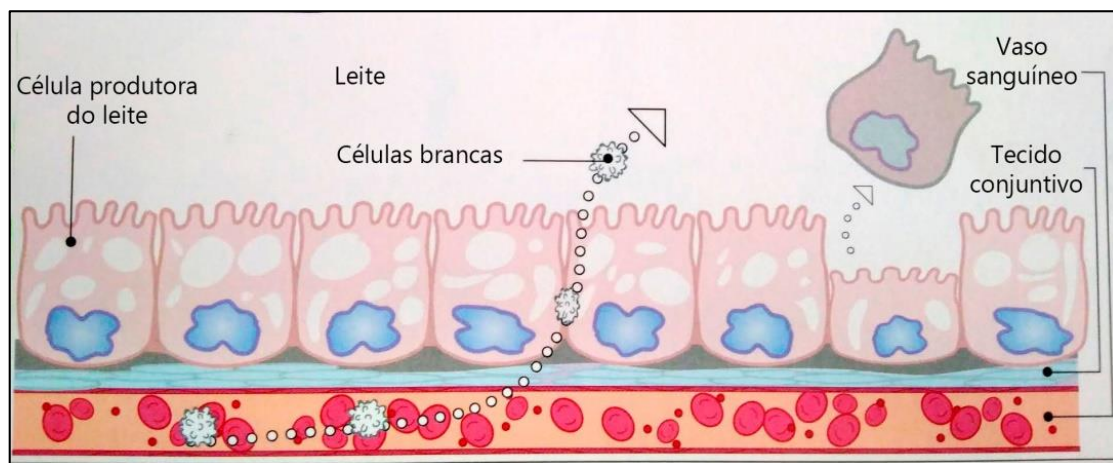


Figura 4: Migração das células de defesa para o tecido mamário (HULSEN & LAM, 2007).

Os métodos utilizados na contagem de células somáticas classificam-se em duas categorias, quais sejam, diretos e indiretos. Nos métodos indiretos, mistura-se o leite a um detergente provocando uma coagulação que torna a mistura viscosa. O *California Mastitis Test (CMT)* e o *Wiscosin Mastitis Test (WMT)* são os métodos indiretos mais utilizados. No CMT, o grau de viscosidade é analisado e estimado visualmente, ao passo que no WMT a verificação do grau da viscosidade é feita através da passagem dessa mistura (leite e detergente) por um viscosímetro constituído de um suporte metálico com tubos individuais de acrílico ou plástico no qual consta um escala de leitura (SCHALM & NOORLANDER, 1957).

Os métodos diretos são baseados no uso de aparelhos eletrônicos tais como o *Fotosomatic* e o *Somacount* ou por meio de microscópio óptico do tipo comum (DMSCC - Direct Microscopy Somatic Cell Counter). A contagem por meio dos aparelhos eletrônicos é a mais precisa e mais rápida e seu uso está restrito aos

laboratórios aprovados pela Rede Brasileira de Controle da Qualidade do Leite devido ao alto custo desses instrumentos de medição. De acordo com a *Dairy* equipamentos, que é a representante da *Bentley* no Brasil, o custo desses equipamentos varia de R\$580.000,00 até R\$780.000,00, para equipamentos já nacionalizados (aplicados os impostos), e de USD85.000,00 à USD105.000,00 nos EUA. A contagem com o uso do microscópio óptico comum permite que sejam identificadas estruturas de observação como as células através da incidência de luz e de lentes objetivas que promovem um aumento de até 1000x. O microscópio deve possuir objetiva com visão panorâmica de aumento aproximado de 40x, 100x, 400x e 1000x, respectivamente, sendo necessário o uso de óleo de imersão para a visualização de estruturas com aumento de 1000x sem prejuízo às lâminas. Nas lentes objetivas a imagem pode ser ajustada pelo seu limiar de resolução. Uma marca compatível com essas características e que pode ser utilizada na contagem de células somáticas é o Eclipse E800, Nikon®, Japan. A DMSCC é realizada diretamente através da observação de lâminas com esfregaços de leite corados, cujo objetivo é a identificação e reconhecimento das células somáticas (DE MELO *et al.*, 2015).

A mastite e CCS em búfalas são temas de vários trabalhos que analisam o efeito dessa doença sobre a quantidade do leite, além das alterações em suas propriedades (teores de proteínas, gorduras, açúcares, minerais e vitaminas) no estágio de gravidez, fase lútea e de lactação do animal (Singh *et al.* (2000), Shen *et al.* (2002), Tripaldi *et al.* (2003), Bansal *et al.* (2007), Kavitha *et al.* (2009), Kaur *et al.* (2015), Mansour *et al.* (2016)). Nestes trabalhos, a técnica mais adotada para a CCS é a DMSCC pelo seu baixo custo e aceitação ou confiança do mercado.

Os métodos diretos de detecção de mastite são considerados mais precisos em relação aos indiretos, sendo recomendados pela FDA (Food and Drug Administration). Dentre os diretos, se destaca a citometria de fluxo por ser preciso e apresentar alta produtividade na contagem das células, sendo padronizado por organizações internacionais de produtos lácteos (GARCIA-CORDERO *et al.*, 2010). Por outro lado, a citometria de fluxo requer um equipamento para a contagem eletrônica cujos custos de aquisição e manutenção são elevados, principalmente para o pequeno/médio produtor. Uma opção mais barata de método direto é o de contagem microscópica de células somáticas (DMSCC), considerado como uma técnica de referência para calibração de outros equipamentos e controle do número de células. O DMSCC é um método de

referência nas normas oficiais e de suma importância para suprir a demanda de produtores que não dispõem de recursos para a contagem eletrônica das células. Trata-se também de um método de caráter subjetivo, exaustivo e que dispende tempo para a preparação das lâminas e contagem visual por parte dos especialistas (GAO *et al.*, 2016; GRILO *et al.*, 2001).

Na tentativa de realizar o controle da mastite algumas já foram apresentadas (muitas delas utilizando como parâmetro a CCS) por meio de processamento de imagens, uso de classificadores, redes neurais e outras técnicas computacionais (*K-means*, o *Fuzzy C-means* (FCM), a Rede Neural SOM, fusão de imagens, PCA, entre outros).

Wang e Xue (2010) e Xue *et al.* (2009) apresentam técnicas de processamento de imagens e redes neurais SOM (Redes Neurais Auto-Organizáveis, *Self-Organizing Maps*) para identificação/detecção das células. Wang e Xue (2010) empregam a RNA SOM na etapa de segmentação. As imagens em RGB são lidas, as componentes R, G, B são extraídas e o treinamento não-supervisionado pela RNA é realizado visando o reconhecimento de 3 classes de *pixels* (núcleo, citoplasma e fundo). Xue *et al.* (2009) combinam técnicas de processamento de imagens, redes neurais e PCA também aplicados especificamente na etapa de segmentação. O PCA é aplicado em 4 espaços de cores (RGB, YIQ, XYZ e Ycbr) e através da análise de componentes principais, três classes (núcleo, citoplasma, fundo) são obtidas e usadas como vetor de entrada da rede neural SOM. A rede é treinada, e os *pixels* das classes são classificados para o processo de segmentação da imagem. A técnica de PCA também é aplicada no trabalho de GAO *et al.* (2017) para a redução de dimensionalidade do espaço dos dados juntamente com a Máquina de Vetores de Suporte (Support Vector Machine, SVM) que consiste em um método de classificação não-supervisionada.

Bai *et al.* (2015) aplicam o método de agrupamento não-hierárquico *k-means* também na etapa de segmentação. As imagens RGB são segmentadas com base nas informações dos histogramas que apresentou uma diferença considerável entre as componentes R e G no núcleo, citoplasma e fundo, de forma que usando o *k-means* a segmentação foi realizada nos canais RG e RB da imagem. De Melo *et al.* (2015) exploram as funcionalidades do *k-means* também na etapa de segmentação para a contagem de células somáticas em leite de vaca usando um sistema de Microscopia de

Vídeo (VM) automatizado. A metodologia combina o uso do algoritmo *k-means* e Transformada *Watershed* (Roerdink e Meijster, 2000).

Feng *et al.* (2009) aplicaram o método FCM na etapa de segmentação das imagens em RGB das células somáticas do leite, mostrando que o mesmo teve um bom resultado e também melhorou a velocidade da segmentação.

Outros trabalhos apresentam outras técnicas aplicadas também na etapa de segmentação, a exemplo de Xue *et al.* (2007) (classificação morfológica no espaço de cores 2D e técnica de fusão) e Gao *et al.* (2016) (leite bovino, limiarização com modelo de nuvem, LI *et al.*, 2009). Cavero *et al.* (2008) aplicaram a Rede Neural *Feedforward* multicamadas para a predição do número de células somáticas usando como entradas da Rede a taxa de produção de leite, condutividade elétrica e fluxo de leite. Um modelo baseado em sistemas de inferência *fuzzy* para a detecção direta de mastite em vacas é proposto por Cavero *et al.* (2006) usando as mesmas entradas (condutividade elétrica, taxa de produção e vazão/fluxo de leite).

Xu e Lu (2014) propõem um método para segmentar núcleos celulares em imagens histopatológicas da pele através de uma técnica de limiarização adaptativa. Os autores utilizam identificador elíptico para detectar os núcleos isolados com formas elípticas. O descritor tem o papel de classificar as regiões dos núcleos com base em dois parâmetros de elipticidade.

Em trabalhos relacionados à identificação e contagem de células sanguíneas, o FCM e/ou *k-means* são utilizados na etapa de segmentação (em Li *et al.* (2014), Chaira (2014), Theera-Umpon (2005), Mondal *et al.* (2014), Sarrafzadeh e Dehnavi (2015), Savkare e Narote (2015) e Mandal *et al.*, (2016)). Outros trabalhos abordam a contagem de células no sangue humano e se baseiam em técnicas de processamento de imagens, reconhecimento de padrões e uso de classificadores (Selinummi *et al.* (2005), Sinha e Ramakrishnan (2003), Ongun *et al.* (2001), Kareem *et al.* (2011), Nazlibilek *et al.* (2014)).

Considerando as evidências apresentadas que ressaltam a importância do setor de produção do leite bubalino no mercado brasileiro e mundial e a inexistência de uma metodologia de CCS rápida, automática, viável tecnicamente e financeiramente e menos suscetível à interpretação humana, este trabalho propõe um método de contagem de células somáticas em leite bubalino, baseado em agrupamento *Fuzzy* e em técnicas de

processamento de imagens com uma metodologia diferente das apresentadas pelos trabalhos encontrados na área.

Na seção seguinte é apresentado o objetivo geral e os objetivos específicos deste trabalho. O segundo capítulo traz a fundamentação teórica e revisão acerca do conteúdo que foi utilizado como base para o estudo. O terceiro capítulo apresenta a metodologia aplicada e os resultados. O quarto capítulo apresenta as principais conclusões e sugere futuros trabalhos.

1.2 OBJETIVOS

O objetivo geral deste trabalho é desenvolver um método para a CCS (contagem de células somáticas) em amostras de leite de búfalas de forma automática, utilizando agrupamento *Fuzzy* e técnicas de processamento de imagens, tendo os seguintes objetivos específicos:

- Desenvolver um método de agrupamento das imagens com base na similaridade entre suas respectivas intensidades e que seja baseado na técnica de agrupamento *Fuzzy C-Means*;
- Avaliar quais métodos de limiarização são mais adequados para imagens em tons de cinza nas células somáticas do leite bubalino;
- Desenvolver um método para a contagem de células somáticas em leite de búfala em imagens capturadas pelo microscópio óptico.

CAPÍTULO 2

FUNDAMENTAÇÃO TEÓRICA E REVISÃO BIBLIOGRÁFICA

2.1 CÉLULAS SOMÁTICAS

As células somáticas (Figura 5) compreendem as células de defesa, que têm origem no sangue e migram para o úbere, além das células de descamação da glândula mamária (células epiteliais). As células provenientes da corrente sanguínea são os leucócitos e estes constituem cerca de 75% a 98% do total das células somáticas do leite (PHILPOT E NICKERSON, 1991; RIBAS, 2001).

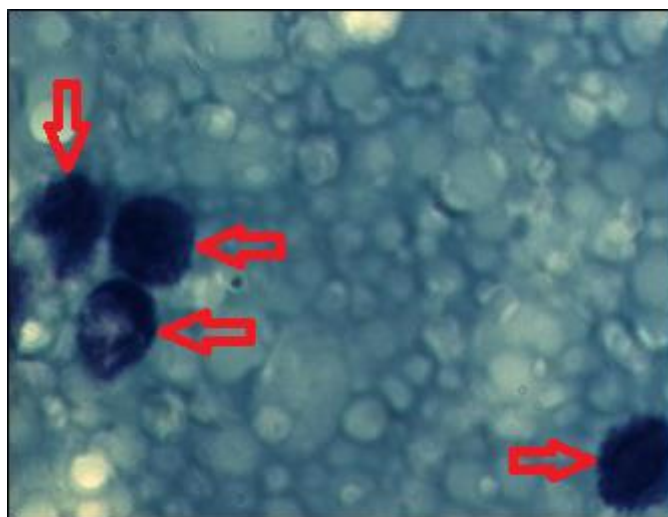


Figura 5: Células somáticas em leite de búfala (VIVIANI GOMES, 2016).

Quando acontece a invasão de bactérias no úbere ocorre imediatamente uma resposta do sistema de defesa com o desencadeamento de resposta inflamatória. Na tentativa de destruir essas bactérias, as células de defesa do sangue migram para a

glândula mamária, com subsequente aumento destas células no leite, o que justifica o uso deste biomarcador como indicativo de infecção e/ou inflamação da glândula mamária (SHARMA *et al.*, 2011).

Segundo Schalm (1971), a mastite é uma inflamação da glândula mamária proveniente de causas diversas, com variações na duração e diferentes graus de intensidade. A forma mais comum se dá pela penetração de bactérias no interior da glândula mamária produzindo irritações e inflamação nos tecidos (Figura 6). Existem dois tipos de mastites: clínica e subclínica. Na forma clínica, a inflamação se caracteriza grosseiramente pelo inchaço, calor, dor e vermelhidão. A forma subclínica não provoca tais efeitos e é detectável através de análises do leite que indicam a presença de produtos gerados pela inflamação, tais como, leucócitos, coágulos de fibrina e soro. Ocorrem também mudanças na composição química do leite por alterações na permeabilidade vascular local com a transferência de cloreto de sódio e bicarbonato do sangue para o leite, alterando o PH da secreção láctea.



Figura 6: Úbere inchado e vermelho, devido ao processo inflamatório da glândula mamária (VIVIANI GOMES, 2019).

A mastite é uma doença de grande impacto em relação à quantidade e qualidade do leite produzido pelo animal (DURR *et al.*, 2008), sendo a CCS utilizada como um biomarcador não apenas para o diagnóstico da doença, mas para o monitoramento da qualidade do leite.

Animais infectados com mastite apresentam alteração dos componentes do leite, com redução dos teores de cálcio, lactose, caseína e gordura; aumento dos níveis de íons

sódio, cloro e de proteínas séricas, associados ao aumento no número de células somáticas. Apesar de controles realizados, a doença causa perdas econômicas significativas, principalmente para as indústrias de laticínios, considerada como uma das doenças mais dispendiosas na indústria pecuária de leite no mundo (DAMM *et al.*, 2017; OLIVEIRA *et al.*, 1999).

Keefe (2012) afirma que a mastite é responsável por perdas econômicas significativas nos países desenvolvidos, mesmo com investimento em estudos e medidas de prevenção. A doença causa uma diminuição na produção do leite (SOMMERHAUSER *et al.*, 2003), e a alta contagem de células somáticas está relacionada a quedas de rendimento e na produção dos seus derivados, acarretando modificações organolépticas no leite e derivados, conseqüentemente promovendo a redução na vida de prateleira dos produtos. Portanto, pode-se dizer que a CCS está diretamente relacionada com o rendimento de derivados lácteos, sendo muito importante a realização de um monitoramento adequado e preciso (LOPES *et al.*, 2012).

Um dos derivados mais consumidos e que apresenta problemas durante a produção é o queijo. Durante o processo de coagulação pode ocorrer o aumento da quantidade de água e no tempo para a formação do coágulo; o rendimento industrial é reduzido; as propriedades sensoriais sofrem alterações negativas; a textura é comprometida; e há perda significativa de sólidos no soro. As alterações relacionadas de forma direta com a alta contagem de células somáticas estão ligadas, por exemplo, às alterações na composição da proteína e balanço de minerais. A alta CCS provoca a redução de vida de prateleira de alimentos como a manteiga, leite em pó e outros derivados, pois as células somáticas fazem a síntese de enzimas que são resistentes à pasteurização que deterioram a gordura, tendo como consequência um sabor desagradável e alteração do valor nutricional (FONSECA & SANTOS, 2001; PETROVA E TZANKOVA, 1999; TRIPALDI *et al.*, 2003; SINGH E SINGH, 1981).

Outros fatores que também podem influenciar na produção e composição do leite de búfala são os estádios de lactação e os efeitos da estação do ano. Estudos têm sido realizados em relação aos meios que influenciam os constituintes do leite de búfala, cujo objetivo é controlá-los. Algumas variáveis que compõem esta composição estão em função da raça, idade do animal, fatores relacionados ao meio ambiente tais como as estações do ano, tipo de alimentação, manejo durante a ordenha e a evolução do período de lactação (AMARAL *et al.*, 2005; MATTOS, 2007). Os parâmetros utilizados para se

avaliar a qualidade do leite são cada vez mais importantes e devem ser utilizados para proporcionar a detecção de falhas durante as práticas de manejo, pois a matéria-prima deve ser valorizada e esses parâmetros servem como referência nesse processo (DÜRR, 2004).

2.1.1 Contagem de células somáticas

A contagem de células somáticas (CCS) é um biomarcador utilizado para a detecção da mastite e monitoramento da qualidade do leite em bovinos (SCHALM, 1971). Em bubalinos, a menor susceptibilidade da espécie tem sido associada a um sistema de proteção mamário diferenciado dos bovinos, destacando-se diferenças quantitativas e qualitativas da celularidade do leite desses dois animais (DELLA LIBERA *et al.*, 2005; KEE, 2012); JINGAR *et al.*, 2014).

A legislação da Comunidade Europeia, considerando leite de conjunto, em 1994, determinou que o leite cujas origens sejam bovina, bubalina, caprina ou ovina, quando destinado ao humano, deve apresentar uma contagem máxima de células somáticas de 400.000 células/mL (BIERENS, 1993). Em relação às búfalas, não existe uma legislação no Brasil que regulamenta um padrão de CCS e os limiares utilizados para os bovinos não são adequados, visto que os valores das contagens em bubalinos são significativamente menores do que nos bovinos (ARAÚJO & GHELLER, 2005; AMARAL, 2005).

No trabalho de Piccinini *et al.* (2006) foi sugerido a contagem de 400.000 células/mL como ponto de separação para o leite bubalino. Porém, outros autores (Vianni *et al.*, 1990, Tuteja *et al.*, 2001, Singh *et al.*, 2002, Vivek *et al.*, 2002, Dhakal 2004, Singh *et al.*, 2004) adotaram o valor de 500.000 células/mL como limiar da forma subclínica.

2.1.2 Métodos de detecção de mastite

Para se obter um leite com um alto padrão de qualidade deve-se definir boas estratégias para a prevenção e controle da mastite dentro do rebanho. Para tanto, deve-se identificar o problema, a dinâmica da infecção e os agentes que estão causando a infecção através do uso de ferramentas de diagnóstico.

Dentre os métodos para detecção têm-se os diretos, baseados na CCS, que compreendem o DMSCC e a contagem eletrônica. E os indiretos, dentre eles pode-se citar: o *California Mastitis Test (CMT)* e o *Wisconsin Mastitis Test (WMT)*.

O DMSCC é o método de referência atual para a CCS no leite e requer uma equipe bem treinada, com precisão e garantia de repetibilidade de resultados. É considerada uma estratégia lenta devido aos requisitos de trabalho exigidos tais como montagem das lâminas, aquisição das imagens no microscópio e contagem visual das células (MOON *et al.*, 2007). O método consiste em distribuir uma quantidade de leite de 0,01 mL em uma área de 1 cm² na superfície de uma lâmina de microscopia (PRESCOTT & BREED, 1910). Em seguida, a lâmina passa por um processo de secagem durante 24 h, em temperatura ambiente, e então o esfregaço lácteo é corado com a possibilidade do uso de diferentes colorações. Por fim, as células que foram coradas são identificadas e contadas por 100 diferentes campos microscópicos (MARSHALL, 1992).

O campo microscópico é a área de preparação da lâmina que se está observando no microscópio. A Figura 7 mostra um exemplo de um campo microscópico observado em uma lâmina de leite de búfala. A soma do número de células contadas em 100 campos microscópicos deve ser multiplicada pelo fator do microscópio e expresso em número de células por mililitro. O fator do microscópio, de acordo com o método utilizado, equivale à área do círculo cujo raio corresponde ao aumento verificado.

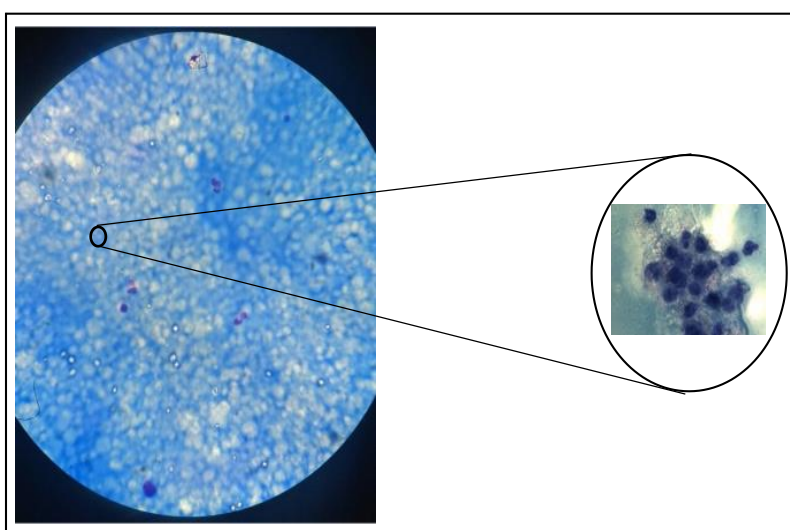


Figura 7: Campo microscópico observado em uma lâmina de leite de búfala em diferentes aumentos, (Imagem da esquerda em aumento de 400x e imagem da direita em aumento de 1000x).

Os contadores eletrônicos trabalham de uma forma diferenciada, utilizando técnicas de análises baseadas na contagem de impulsos elétricos a partir da passagem de partículas entre dois eletrodos, o uso de citometria de fluxo e a fluorescência óptica (BABAK E RYSANEK, 1999). Na citometria de fluxo (Figura 8), as células são coradas e colocadas em um líquido para seguir um fluxo para excitação das mesmas por um feixe de luz (laser), por fluorescência. Os núcleos corados emitem impulsos luminosos e esses são ampliados por um foto multiplicador, contabilizados e convertidos em células somáticas (CECALAIT, 1993).

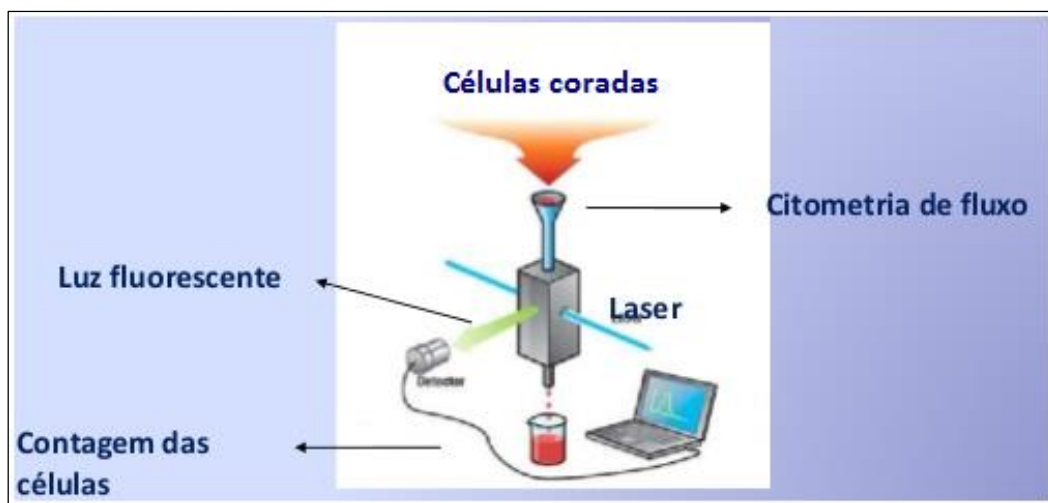


Figura 8: Citometria de fluxo (Adaptado de DANIELLE LIMA, 2010).

Na técnica baseada em fluorescência óptica, o DNA das células passa por um processo de coloração e quando sofrem estímulos através de uma fonte de luz, ocorre a emissão de fluorescência e esta é medida na forma de pulsos elétricos e convertida em quantidade de células somáticas que estão na amostra (CHRISTEN, 1993).

Dentre os equipamentos eletrônicos, o *Fotosomatic* e o *Somacount* são os mais utilizados no Brasil. Em um *Somacount 500 – Bentley*, por exemplo, a capacidade de análise é de 500 amostras por hora. Cada amostra equivale a uma porção de leite com volume menor que 5 mL, contidas em frascos. No processo da avaliação um percentual do material é recolhido pelo aparelho, exposto a uma temperatura de 67°C, e colocado em uma seringa contendo o que se chama de corante tampão. Logo após, 50 µL do material são conduzidos até o *cell* que é uma das partes do equipamento que armazena o material recebido através de um fluído carreador RBS. Fluídos carreadores são utilizados para transportar substâncias com menor quantidade. No *cell* ocorre a

incidência de raio laser havendo uma pequena explosão. Essa luz emitida passa por vários filtros ópticos e lentes que estão com foco adequado para tal comprimento de onda e é captada como pulso elétrico, sendo este amplificado, filtrado e convertido em contagem de células (BENTLEY INSTRUMENTS, 1994).

Outra forma de detectar a mastite consiste nos chamados testes “no local”, ou seja, na sala de ordenha, realizados pela equipe de ordenha e médicos veterinários. O *California Mastitis Test (CMT)* é um dos mais conhecidos, antigos e mundialmente utilizados para o diagnóstico da mastite subclínica. Como vantagens, apresenta um custo efetivo, é rápido e fácil de usar, mas apresenta baixa sensibilidade e o resultado pode ser de difícil interpretação, pois o operador estima de forma indireta a quantidade de células de acordo com a viscosidade da amostra de leite que foi avaliada (VIGUIER *et al.*, 2009). Este método consiste na adição de reagente ao leite em uma raquete apropriada e como resultado ocorre a quebra das membranas das células somáticas e a liberação de ácidos nucleicos, formando uma matriz semelhante a um gel cuja viscosidade é proporcional ao número de células somáticas. Quanto maior a viscosidade do leite, maior é a quantidade de células na amostra (BRITO *et al.*, 1997).

O *Wisconsin Mastitis Test (WMT)* é um teste considerado como um aperfeiçoamento do CMT, devido à precisão dos componentes (leite e reagente) e do procedimento realizado. Quantidades precisas de reagentes e leite são colocados em um tubo graduado. O reagente deve ser diluído em água destilada e para que fique homogêneo deve-se realizar movimento de rotação no tubo, invertendo-o e deixando escoar por 15 segundos e logo em seguida voltando à posição original. O resultado é expresso em milímetros e está correlacionado com a contagem de células somáticas (LANGONI, 2000).

2.2 TÉCNICAS DE PROCESSAMENTO DE IMAGENS

2.2.1 Visão geral

A representação de uma imagem é feita através de uma matriz cujos pontos são chamados de *pixels* (Figura 9). Em uma imagem típica de uma câmera digital existe de 4 a 6 milhões de *pixels*, cada um representando a intensidade de uma cor em cada ponto da imagem. Essa grande quantidade de dados é armazenada como um arquivo, em um

formato específico, por exemplo, GIF ou JPEG, viabilizando-se a manipulação desta informação através de softwares (PARKER, 2010). A quantidade de *pixels* define a resolução da imagem e, quanto maior, melhor a resolução e o tamanho da imagem (SOUZA e CORREIA, 2007).

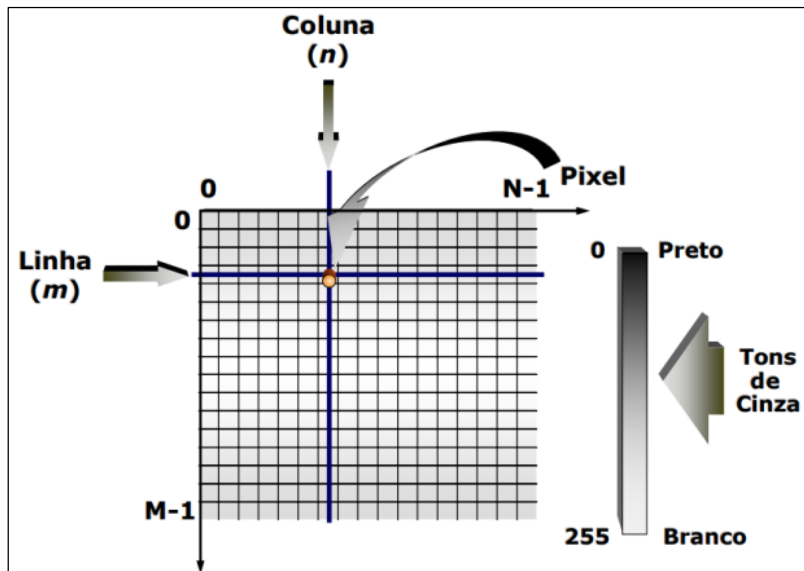


Figura 9: Representação de uma imagem digital bidimensional (DE QUEIROZ E GOMES, 2006).

O processamento de imagens consiste em técnicas de análise de dados multidimensionais que permitem a manipulação e o reconhecimento de características. As etapas que constituem um sistema de processamento de imagens são: pré-processamento, segmentação, extração de características e reconhecimento/interpretação (COELHO *et al.*, 2012).

O processamento de imagens evoluiu bastante, sendo aplicado em quase todas as áreas das atividades humanas. Isso pode ser observado na biologia, quando se utiliza o processamento de imagens microscópicas; na medicina, aperfeiçoando o diagnóstico médico através de equipamentos que sejam capazes de interpretar imagens; na arqueologia, através de técnicas que possibilitam a restauração de imagens; na geografia, sensoriamento remoto, geoprocessamento e meteorologia, através da interpretação das imagens de satélites; e tantas outras áreas como direito, publicidade, segurança, etc. (VIEIRA E MARQUES, 1999).

2.2.2 Sistema RGB e imagens monocromáticas

O RGB (*Red-Green-Blue*, Vermelho-Verde-Azul) é um dos espaços de cores mais conhecidos. Um espaço de cores consiste em um modelo matemático para se descrever uma cor através de uma formatação padronizada (ROCHA, 2011).

Ao trabalhar com imagens coloridas se faz necessário entender de que forma as cores são originadas. Os sistemas que trabalham com esse espaço de cores têm as cores primárias (vermelho, verde e azul) como base para obtenção das demais cores, em diferentes proporções. Quando se fala de cor, associa-se ao conceito de luz visível que consiste no conjunto de ondas eletromagnéticas na faixa de comprimento cujo sistema visual humano é capaz de detectar (Figura 10). O comprimento de onda determina qual a tonalidade percebida da luz. A teoria sobre a visão das cores pelo olho humano retrata que a retina é constituída por células fotossensíveis, divididas em três classes com diferentes picos de sensibilidade cujos comprimentos de ondas estão situados entre o vermelho (*Red*), verde (*Green*) e o azul (*Blue*) (GEVERS e SMEULDERS, 1999).



Figura 10: Sensibilidade do olho humano para as cores primárias (HUNT, 1995).

Em uma imagem em RGB cada pixel é composto por três valores e cada um destes refere-se às intensidades de vermelho, verde e azul, respectivamente. As cores são representadas por matrizes e os elementos de cada matriz variam de 0 a 255, pois cada uma dessas três cores possui 256 tonalidades representando as intensidades dessas cores. Por exemplo, para formar a cor branca tem-se RGB (255, 255, 255) (KUMAR & VERMA, 2010). A Figura 11 ilustra as criações de *pixels* com o RGB.

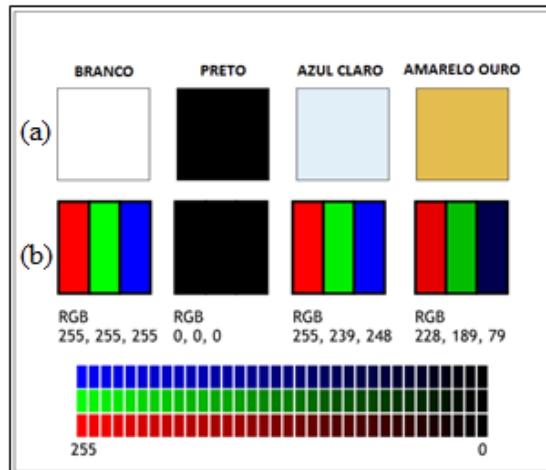


Figura 11: Criando cores com *pixels* RGB: (a) *Pixels* branco, preto, azul claro e amarelo ouro; (b) Componentes RGB (Intensidade de cada cor primária).

As imagens compostas por componentes distintas (vermelho, verde e azul), como é o caso das que estão no espaço de cores RGB, possuem uma função bidimensional $f(x,y)$ (x e y são coordenadas de posição) para cada matriz referente as cores (*Red*-Vermelho, *Green*-Verde, *Blue*-Azul). O entendimento sobre as imagens em RGB pressupõe a compreensão sobre a representação matemática de imagens monocromáticas.

Uma imagem monocromática pode ser descrita por uma função $f(x,y)$ de intensidade luminosa cujo valor em qualquer ponto do espaço bidimensional (coordenadas (x,y)) é proporcional ao brilho ou nível de cinza da imagem naquele ponto (Figura 12) (VIEIRA E MARQUES, 1999).



Figura 12: Posicionamento dos eixos para representação de imagens (GONZALEZ AND WOODS, 2000).

A definição pontual de uma imagem consiste na intensidade de iluminação, $i(x,y)$, referente à quantidade de luz incidente sobre o ponto (x,y) , e a refletância, $r(x,y)$, referente à quantidade de luz refletida sobre o mesmo ponto (x,y) (Figura 13). Tem-se, portanto:

$$f(x,y) = i(x,y).r(x,y) \quad (1)$$

Onde:

$$0 < i(x,y) < \infty \quad (2)$$

$$0 < r(x,y) < 1 \quad (3)$$

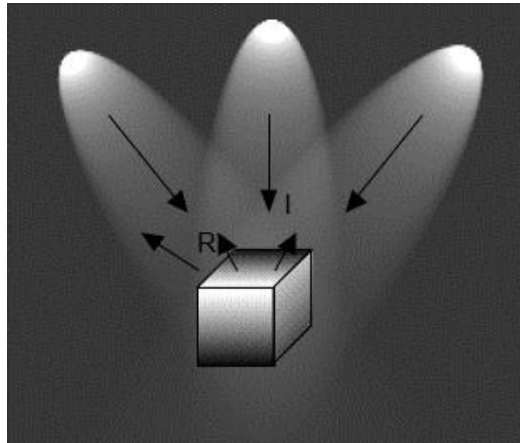


Figura 13: Os componentes iluminância (I) e refletância (R) de uma imagem (VIEIRA E MARQUES, 1999).

De acordo com Gonzalez e Woods (2000) a refletância se situa entre 0 (absorção total da luz) e 1 (refletância total). A iluminância ($i(x,y)$) é definida pela fonte de iluminação e $r(x,y)$ é definida pelas características dos objetos que compõem a cena, como por exemplo sua cor e textura. Alguns exemplos de valores de reflectância em lúmen/m² são: 900 para dia ensolarado, 100 para dia nublado, 10 para iluminação média de escritório e 0,001 para lua cheia. Exemplos de reflectância incluem 0,93 para a neve; 0,80 para parede branco-fosco; 0,65 para aço inoxidável; 0,01 para veludo preto (VIEIRA e MARQUES, 1999).

Uma imagem monocromática é representada por uma matriz (Figura 14) cujos valores dos componentes estão limitados entre o zero (que corresponde ao preto) e 255

(que corresponde ao branco). Valores intermediários representam diferentes níveis de cinza (SONKA *et. al*, 2014).

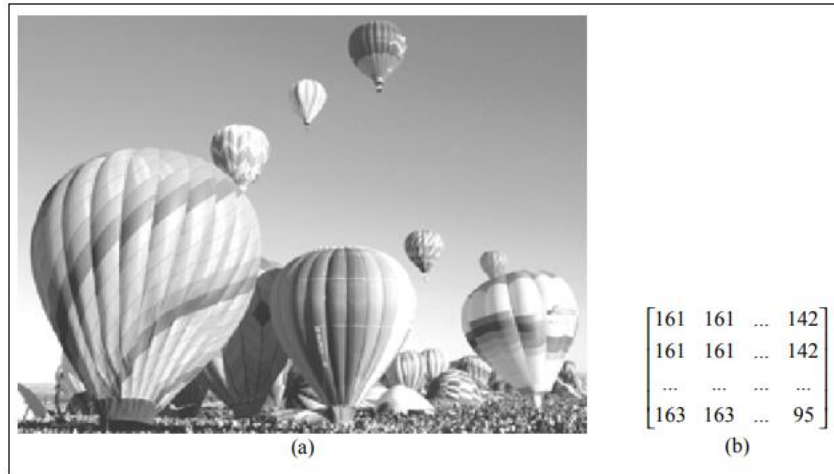


Figura 14: Uma imagem digital monocromática: (a) Representação gráfica usual; (b) Representação matricial (BATISTA, 2005).

Quando se trata de imagens coloridas, a representação segue a mesma lógica. Levando em consideração o espaço de cores RGB, cada pixel (Figura 15) é representado por um vetor de três componentes $f(x,y) = (f_R(x,y), f_G(x,y), f_B(x,y))$ cujos valores representam as intensidades de vermelho, verde e azul no ponto (x,y) . Portanto, uma imagem colorida é a composição de três imagens monocromáticas, $f_R(x,y), f_G(x,y), f_B(x,y)$, denominadas, respectivamente, de banda vermelha, banda verde, e banda azul, no ponto (x,y) (QUEIROZ e GOMES, 2006).

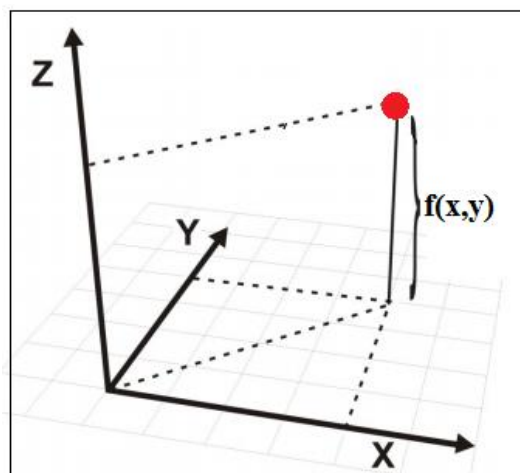


Figura 15: Representação do pixel (CONCI, 2015).

Uma imagem em RGB é representada por 3 matrizes (Figura 16) cujos valores dos componentes dessa matriz estão limitados entre o zero (que corresponde ao preto) e 255 (que corresponde ao branco).

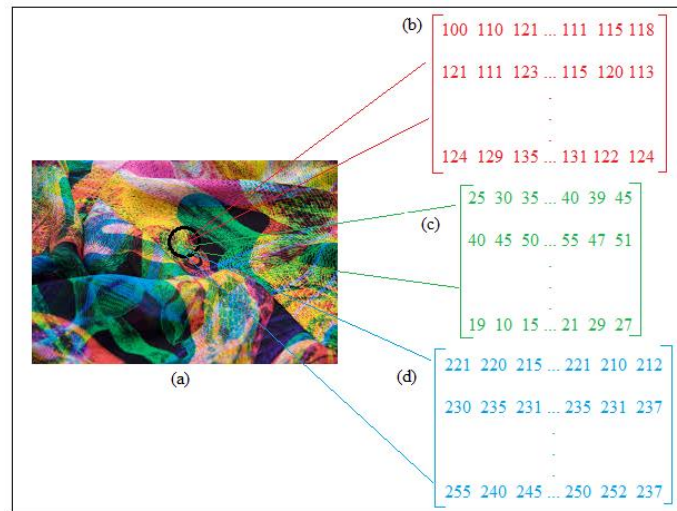


Figura 16: Imagem em RGB: (a) Representação gráfica usual; (b) Matriz da cor vermelha; (c) Matriz da cor verde; (d) Matriz da cor azul.

O método clássico para conversão de uma imagem RGB em seu nível aproximado de tons de cinza baseia-se em se obter as suas cores primárias vermelho, verde e azul e adicionar 30% do vermelho mais 59% do verde e 11% do azul. Estas porcentagens estão relacionadas à sensibilidade visual do olho humano para as cores primárias. A equação de conversão é:

$$\text{EscalaCinza} = 0.299 \times \text{Vermelho} + 0.587 \times \text{Verde} + 0.114 \times \text{Azul}$$

O cálculo deve ser feito para cada *pixel* da imagem e em seguida os valores obtidos em EscalaCinza devem ser aplicados para as componentes RGB do *pixel*, considerando Vermelho=EscalaCinza, Verde=EscalaCinza e Azul=EscalaCinza (CAVINATO, 2009).

Utilizando como exemplo a imagem da Figura 16, o valor dos *pixels* vermelho, verde e azul, cujos elementos de suas respectivas matrizes (linha 1, coluna 1), Vermelho= 100, Verde=25 e Azul=221, tem-se:

$$\text{EscalaCinza} = 0.299 \times 100 + 0.587 \times 25 + 0.114 \times 221 = 69,76970.$$

O valor do elemento da linha 1 e coluna 1 na matriz da respectiva imagem monocromática será 70. Todos os elementos da matriz da imagem monocromática serão obtidos através deste procedimento.

2.2.3 Histograma

O histograma de uma imagem indica a distribuição de *pixels* associados a diferentes níveis de cinza (Figura 17). Normalmente são representados por um gráfico de barras onde para cada nível de cinza é fornecido o número ou o percentual de *pixels* que tem na imagem. Ao observar o histograma de uma determinada imagem pode-se obter a indicação de sua qualidade em relação ao contraste e ao brilho, indicando se a imagem é predominantemente clara ou escura (VIEIRA E MARQUES, 1999). No eixo das abcissas do histograma tem-se o intervalo de 0 a 255 valores de luminosidade e o eixo das ordenadas apresenta a quantidade de *pixels* associados a cada valor de intensidade (GERLACH, 2013). Na forma discreta, o histograma é representado por uma função discreta (VIEIRA E MARQUES, 1999):

$$p_r(r_k) = \frac{n_k}{n} \quad (4)$$

Onde, $0 \leq r_k \leq 1$, $K = 0, 1, \dots, L-1$, sendo L o número de níveis de cinza na imagem digitalizada, n é o número total de *pixels* na imagem, $p_k(r_k)$ é a probabilidade de ocorrência do k -ésimo nível de cinza e n_k é número de *pixels* cujo nível de cinza corresponde à intensidade k .

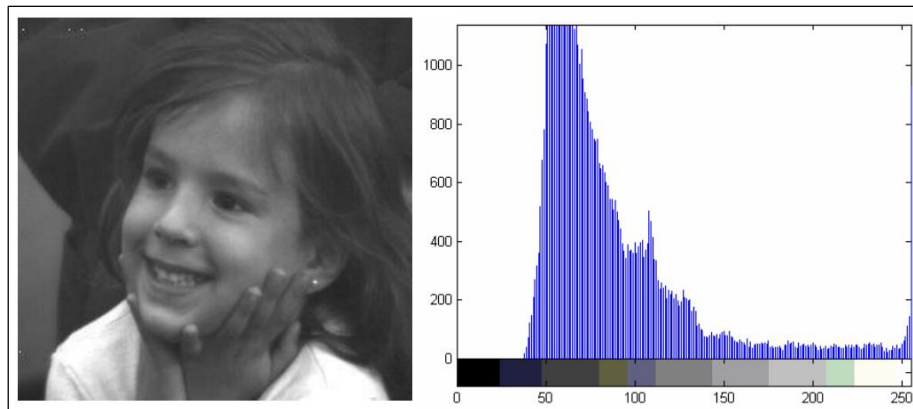


Figura 17: Imagem original e seu histograma (SOUZA E CORREIA, 2007).

Existem situações que são necessárias modificações na imagem original para a melhoria de sua qualidade. Esta melhoria consiste no aprimoramento da interpretação das informações captadas pelos seres humanos e possibilidade de fornecimento de uma

entrada (imagem aprimorada) com melhor qualidade para outros sistemas de processamento de imagens.

Uma das técnicas mais populares para o melhoramento de imagens é a equalização de histogramas (HE), pois é considerada uma técnica simples e que apresenta um desempenho satisfatório. Consiste em alterar a intensidade dos *pixels* na imagem original de modo a obter uma imagem com maior diferenciação entre as tonalidades e com níveis de detalhes perceptíveis amplificados (GONZALEZ e WOODS, 2002).

Nessa técnica procura-se redistribuir os valores de tons de cinza dos *pixels* em uma nova imagem (modificada) cujo histograma apresente uma uniformidade (percentual de *pixels* de qualquer nível de cinza seja similar) (VIEIRA E MARQUES, 1999). Ocorre a diminuição e o aumento da faixa dinâmica (faixa que contém o número de *pixels*) do histograma tendo-se um aperfeiçoamento do nível de contraste da imagem, conforme mostra a Figura 18 (KIM, 1997).

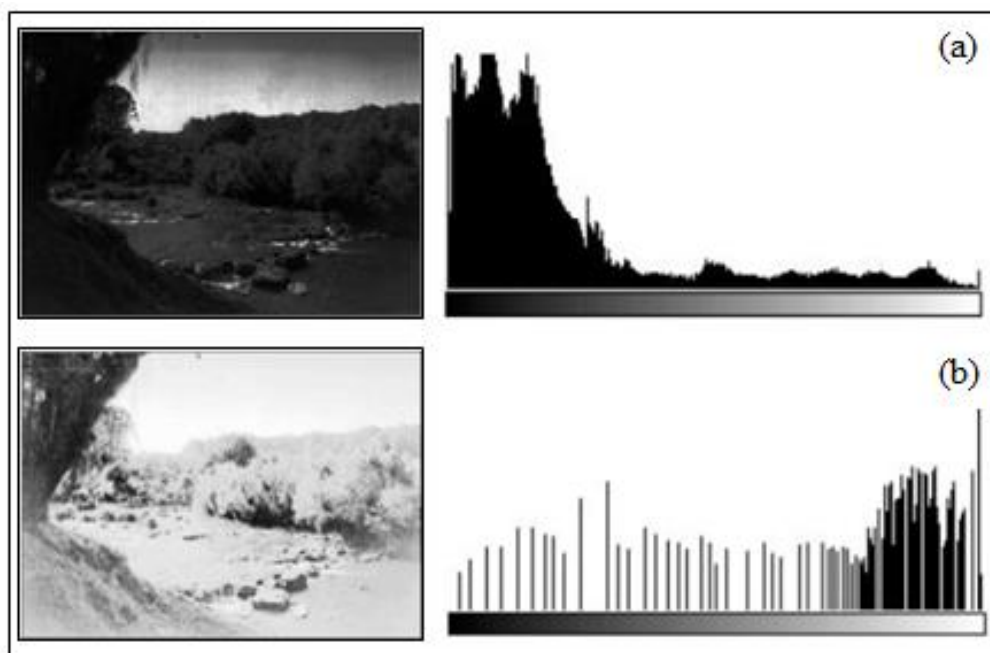


Figura 18: Histograma: (a) Original; (b) Modificado (VIEIRA E MARQUES, 1999).

Após a técnica de equalização aplicada na imagem original o histograma equalizado, mesmo estando longe de se representar uma distribuição uniforme, exhibe melhor distribuição dos *pixels* ao longo da escala de cinza em relação ao original. Isto

contribui para destacar detalhes sutis de pequenas porções da imagem. A imagem mais nítida, com maior riqueza de detalhes facilita a sua análise durante o seu processamento.

O histograma apresenta tipos diferentes, de acordo com as imagens. Dentre alguns exemplos tem-se o simétrico e unimodal, que centraliza os dados na média, contendo no centro da distribuição a maior frequência, apresentando apenas um pico. O bimodal, cuja característica principal é a presença de dois picos. E o multimodal, cuja característica é a presença de vários picos. Quando se trata de imagens, esses picos representam as regiões do espaço com maior quantidade de *pixels* (ARTES, 2014).

2.2.4 Limiarização

O processo de limiarização consiste na conversão das imagens que estão em tons de cinza para imagens binárias (zero ou um), preta ou branca (SOUZA E CORREIA, 2007). O objetivo é criar uma representação binária da imagem, classificando cada *pixel* em uma das duas categorias, como "escuro" ou "claro". A forma de limiarização mais simples consiste em definir um valor de limiar fixo e comparar cada *pixel* da imagem com esse valor. A imagem é segmentada em dois grupos: o grupo de *pixels* com níveis de cinza acima do limiar e o grupo de *pixels* com níveis de cinza abaixo do limiar. Em seguida, atribui-se um valor binário (0 ou 1) para todos os *pixels* que estejam em um mesmo grupo. Por exemplo, se o limiar for igual a 50 *pixels*, os *pixels* que forem menores ou iguais a esse limiar terão seus valores alterados para 0 (zero) e os que estiverem acima desse valor de limiar terão seus valores alterados para 1 (um) (BRADLEY ROTH, 2007). O principal desafio desta estratégia é a definição do melhor limiar que dependerá das condições da imagem no que diz respeito ao contraste e iluminação. O melhor limiar é aquele que funciona como um filtro, ou seja, permite que a imagem de um objeto seja capaz de eliminar os *pixels* que não fazem parte desse objeto.

Dentre os métodos de limiarização tradicionais, temos o método de Otsu (1979) e o de Rosin (2001). O método de Otsu considera a limiarização em uma imagem como sendo a divisão da mesma com L tons de cinza em duas classes, quais sejam, Classe 0 e Classe 1 que representam o objeto e o fundo. A partição se dará no tom de cinza T. A Classe 0 contém os *pixels* da imagem com tons de 0 a T e a Classe 1 contém os *pixels*

com tons de T+1 a L. O limiar de Otsu (Figura 19) é obtido pela média entre os tons médios das duas classes particionadas por esse limiar T (OTSU, 1979; XU *et al.*, 2011).

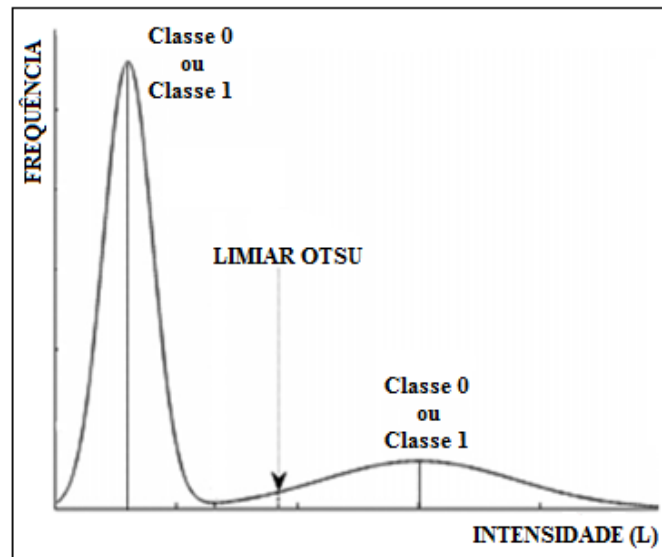


Figura 19: Método de Otsu (OTSU, 1979).

O método de limiarização de Rosin foi proposto para encontrar o limiar em imagens que apresentassem a distribuição de frequência unimodal, ou seja, na imagem existe uma porção maior de apenas uma classe de *pixels* e que domina o histograma. A Figura 20 ilustra todos os parâmetros necessários que estão associados à limiarização de Rosin.

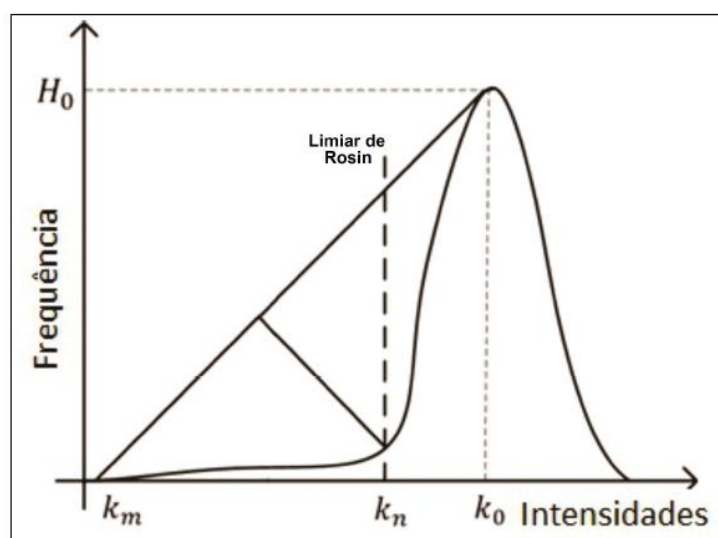


Figura 20: Método de Limiarização de Rosin (ROSIN, 2001).

Rosin (2001) parte do princípio que existe um único pico principal de frequência H_0 para o tom de cinza k_0 . O k_m é o primeiro valor de tom cinza para frequência nula. O limiar de Rosin é obtido traçando uma reta entre o ponto (k_0, H_0) até o ponto $(k_m, 0)$. O limiar escolhido é o valor de tom k_n dado pela frequência do histograma que maximiza a distância d perpendicular à reta. Ou seja, é o valor do histograma que maximiza a distância perpendicular entre a curva que representa a distribuição de frequência dos níveis de cinza e a reta que passa pelo valor máximo (A) e mínimo (B) do histograma, de acordo com a Figura 21 (ROSIN, 2001).

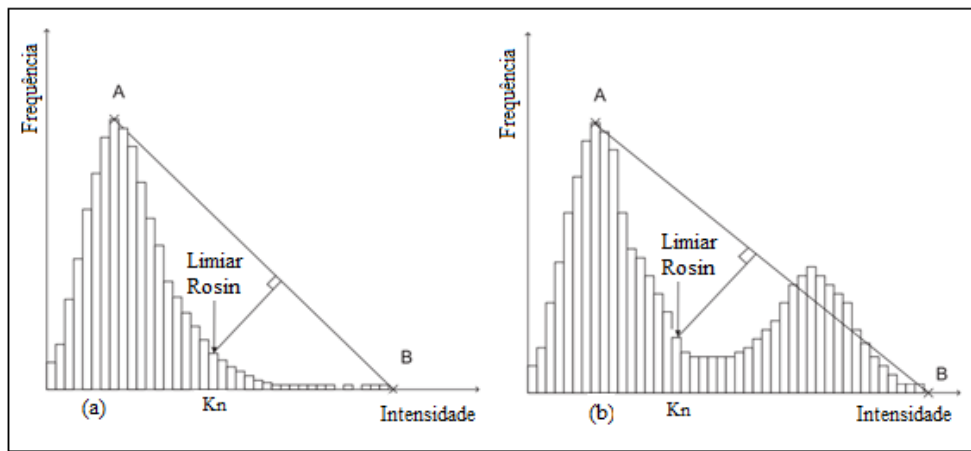


Figura 21: Método Rosin: (a) Histograma unimodal; (b) Histograma Bimodal (Jessie NG *et al.*, 2009).

Um método de limiarização proposto por De Melo (2015) assume que o histograma é unimodal com pico principal único e tem as premissas semelhantes às de Rosin. O nível de limiar previsto deve ser o melhor possível para atender a separação entre objetos e fundo, cujo objetivo é obter o menor erro de classificação.

Esse método parte do princípio que o valor do limiar ideal pode não ser o valor k_n que maximiza a distância entre a reta e a curva, conforme Rosin (2001), mas assume que existe uma classe dominante em relação à classe secundária. O limiar nesse método é definido por:

$$k = \frac{(k_m + k_0)}{L} \quad (5)$$

onde $k \in [k_m, k_0]$ e $L \in [b, a]$ é o parâmetro de ponderação limitado por:

$$a = \frac{(k_m + k_0)}{k_m} \quad (6)$$

$$b = \frac{(k_m + k_0)}{k_0} \quad (7)$$

A Figura 22 ilustra o método de Melo.

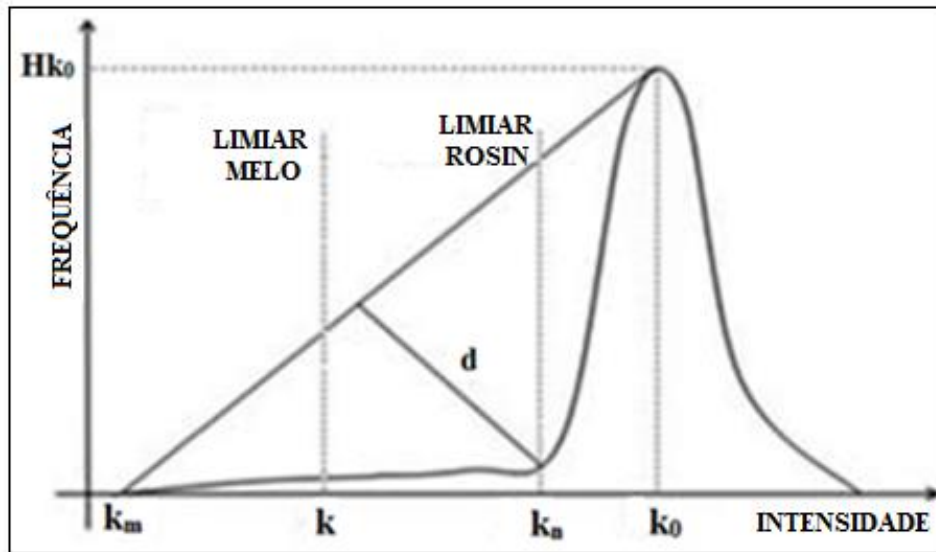


Figura 22: Limiarização: Método de Melo (DE MELO, 2015).

2.2.5 Segmentação e extração de características/reconhecimento/interpretação

A etapa de segmentação é uma das mais importantes visto que se uma imagem for segmentada de forma incorreta, a extração das características ficará prejudicada. A segmentação consiste em separar os elementos (*pixels*) em regiões diversas, a partir de características comuns, ou seja, extraindo-se objetos ou áreas de interesse na imagem (VIEIRA E MARQUES, 1999).

É comum chamar de “objetos” da imagem os grupos de *pixels* que se tem interesse em se separar e chamar de “fundo” da imagem os grupos de *pixels* que devem ser descartados. Vale ressaltar que não se tem um modelo exato e preciso para o processo de segmentação, visto que ele é considerado empírico e adaptativo, pois procura sempre se adequar as particularidades de cada imagem e ao que se deseja alcançar como resultado. A técnica de segmentação normalmente mais utilizada é a que trabalha com a similaridade entre os *pixels*. Esta consiste em realizar a binarização das imagens, onde um nível de cinza é considerado um limiar de separação entre os *pixels* que compõem os objetos e o fundo. Como resultado, teremos uma imagem binária, com apenas dois níveis (preto e branco), conforme ilustra a Figura 23 (GONZALEZ E WOODS, 1993).

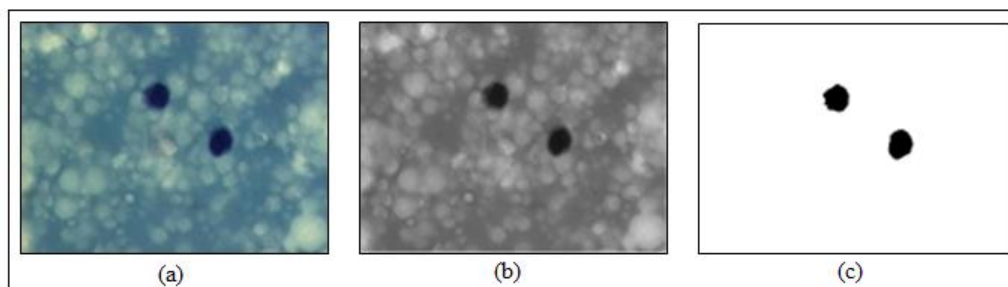


Figura 23: Segmentação: (a) Imagem original; (b) Imagem monocromática; (c) Imagem binária segmentada (VIVIANI GOMES, 2016).

Após a etapa de segmentação vem a etapa de reconhecimento dos objetos ou das regiões resultantes. O processo de reconhecimento de padrões identifica os objetos (os grupos de *pixels* de interesse) na imagem baseado em um conjunto de medições ou atributos (cor, textura, forma, etc.). Cada um dos objetos consiste em um padrão e os valores que são medidos são as características desse padrão. Os objetos são agrupados/classificados conforme a similaridade entre suas características ou atributos (QUEIROZ E GOMES, 2006).

Ao se extrair as características o reconhecimento/interpretação atribui significado ao conjunto de objetos que foram reconhecidos com base nas características/atributos desses objetos (COELHO *et al.*, 2012; VIEIRA E MARQUES, 1999).

2.2.6 Transformada de *Watershed* (TW)

A transformada *Watershed* sugere uma abordagem morfológica para o problema de segmentação de imagens. Cada imagem é interpretada como uma superfície, cada *pixel* corresponde a uma posição e os níveis de cinza determinam as altitudes (KORBES, *et al.* 2010).

A TW pode ser entendida através das noções da inundação de um relevo topográfico. Para isso, considera-se que a altura da superfície será representada pelos tons de cinza que formam uma imagem (BEUCHER, 1991). Essa técnica de *Watershed* complementa o processo de segmentação de uma imagem e se baseia na divisão de águas de uma superfície. Considera-se a imagem de entrada (imagem que está sendo processada) em níveis de cinza como uma superfície topográfica e o objetivo é produzir linhas de divisão de águas nesta superfície. Esta análise mimetiza uma situação na qual

o relevo é inundado de forma progressiva por correntes de água. O acúmulo de água começaria a princípio nas partes mais baixas deste relevo e várias bacias hidrográficas seriam formadas em lugares diversos até chegarem a níveis mais altos e a água dessas bacias se juntarem. Cada bacia formada é chamada de região e quando a inundação acontece, toda vez que uma região chegar à proximidade de se juntar com outra, uma linha divisória é criada, cujo objetivo é evitar que águas de diferentes regiões se misturem. Ao final, cada região formará um segmento na imagem e as linhas formarão o seu contorno tendo como consequência a separação dos mesmos (PINTO, 2014). A Figura 24 exemplifica esse processo.

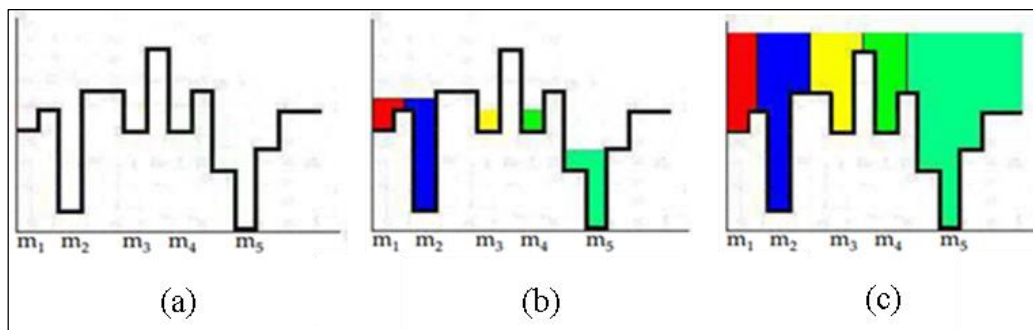


Figura 24: Princípio de imersão de uma imagem: (a) Imagem original e seus mínimos; (b) Linhas divisórias e regiões; (c) Imagem imersa (PINTO, 2014).

A transformada *Watershed* é aplicada sobre uma imagem g^t , obtendo uma nova imagem f apenas com as bordas dos objetos segmentados. As linhas de separação criadas neste momento são denominadas de linhas de partição de águas (GRACIANO, 2002). Em imagens binárias, para separar objetos em contato, a transformada de *Watershed* utiliza-se da ferramenta da transformada da distância, cuja métrica utilizada é a distância euclidiana e a principal função é calcular o afastamento de cada *pixel* para o *pixel* com valor não nulo mais próximo. Ao definir valores semelhantes de distanciamento para uma determinada interação entre pontos de mínimo (nesse caso, centro de cada célula), o método identifica a fronteira de contato entre as duas células. Análogo a bacias hidrográficas, o limite encontrado pelo método serve como um divisor de águas entre duas bacias, permitindo a individualização destas e consequentemente de cada célula.

A Figura 25 ilustra a importância do uso de *Watershed* em imagens de células somáticas muito próximas ou em grupos. Na imagem original observa-se 8 células, na

imagem binarizada apenas 6 células devido a junção dos *pixels*. Com o uso de *Watershed*, ocorre a separação e a contagem será feita de forma correta.

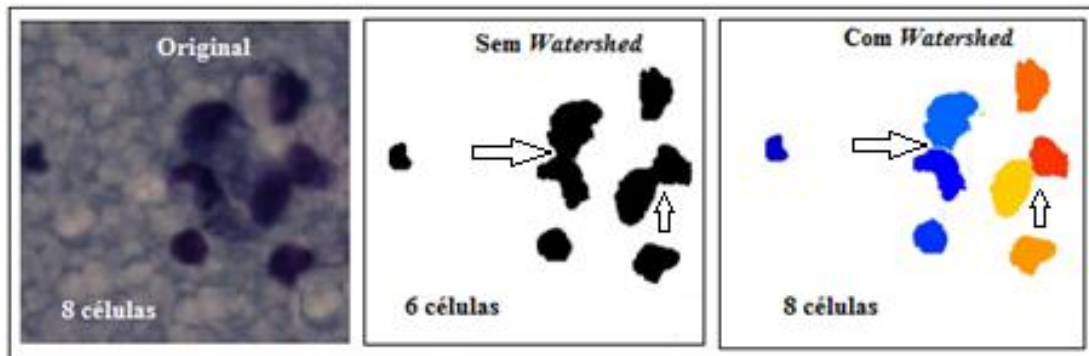


Figura 25: Transformada de Watershed aplicada em imagem de células somáticas (DE MELO, 2015).

2.3 FUZZY C-MEANS

Agrupamento de dados consiste na divisão de um conjunto de N objetos em c grupos de tal forma que haja uma elevada similaridade entre objetos do mesmo grupo e uma elevada dissimilaridade entre objetos de grupos distintos (DUDA e HART, 1973; DE CARVALHO, 2010).

O agrupamento *fuzzy* tem sido amplamente estudado, e o algoritmo de agrupamento *Fuzzy C-Means* (FCM) proposto por Dunn (1973) e estendido por Bezdek (1984) é um dos métodos não hierárquicos mais conhecidos e utilizados. O FCM é capaz de particionar um conjunto de dados/objetos em um determinado número de grupos homogêneos com base em uma métrica de similaridade. Considerando-se objetos descritos por vetores de atributos no espaço multidimensional, a distância euclidiana é uma métrica de similaridade amplamente adotada (CHUANG, 2006; ROSA, 2017).

Na lógica *Fuzzy* um objeto pode pertencer a mais de um grupo com diferentes graus de pertinência, podendo variar entre 0 (pertinência nula) e 1 (pertinência total) (BEZDEK *et al.*, 1984). O FCM compreende essencialmente um problema de otimização cujas variáveis de decisão são as pertinências de cada objeto a todos os grupos e os centros (padrões) de cada grupo.

Sendo um método não hierárquico, o FCM requer a pré-especificação do número de grupos (BEZDEK, 1981; LIAO, 2005; BEZDEK, KELLER E KRISHNAPURAM, 1993) e consiste no seguinte problema de otimização:

$$\min_{(U,C)} J_m(U,C) = \sum_{i=1}^p \sum_{j=1}^n u_{ij} d(x_j; c_i)^2 \quad (8)$$

onde p é o número de grupos, n é o número de objetos, u_{ij} é o grau de pertinência do j^{th} objeto ao i^{th} grupo e U é a matriz de partição ($p \times n$). O parâmetro m ($m > 1$) é o coeficiente fuzzificador (recomendado na literatura $m=2$) e está relacionado ao nível de incerteza do problema de partição. C é o conjunto de padrões ou vetores protótipos $\{c_1, c_2, \dots, c_p\}$ (Vargas *et al.* (2011); Liao, 2005). $d(x_j; c_i)$ é a distância entre o objeto x_i e o centro do grupo j (c_j) (geralmente calculada através da norma Euclidiana).

Duas restrições adicionais devem ser consideradas:

$$u_{ij} \in [0,1] \quad e \quad \sum_{i=1}^n u_{ij} > 0 \quad \forall i \in \{1, \dots, c\} \quad (9)$$

$$\sum_{i=1}^c u_{ij} = 1 \quad \forall j \in \{1, \dots, n\} \quad (10)$$

A restrição 9 garante que nenhum grupo ficará vazio e a restrição 10 estabelece que a soma das pertinências de um objeto, a todos os grupos, deve ser igual a unidade (abordagem probabilística).

A aplicação das condições de primeira ordem ao problema descrito pelas equações 8-10 produz a seguinte solução analítica para o método FCM:

$$\mathbf{v}_i = \frac{\sum_{k=1}^n (u_{ik})^m \cdot x_k}{\sum_{k=1}^n (u_{ik})^m} \quad i = 1, \dots, p \quad (11)$$

e

$$u_{ik} = \frac{\left(\frac{1}{\|x_k - \mathbf{v}_i\|^2} \right)^{\frac{1}{m-1}}}{\sum_{j=1}^c \left(\frac{1}{\|x_k - \mathbf{v}_j\|^2} \right)^{\frac{1}{m-1}}} \quad k = 1, \dots, n \quad i = 1, \dots, p \quad (12)$$

O algoritmo FCM clássico compreende um procedimento iterativo (*Alternating Optimization*) envolvendo as equações 11 e 12, a partir de um chute inicial para a matriz de pertinência.

CAPÍTULO 3

METODOLOGIA E RESULTADOS

As amostras de leite utilizadas foram provenientes de cinco quartos mamários de cinco búfalas diferentes que apresentaram escores 2+ e 3+ no *California Mastitis Test* (CMT) precedidos da higienização dos tetos com solução clorada a 2% e secagem dos mesmos com papel toalha individual. A coleta do leite foi realizada por ordenha manual em frascos plásticos de 15mL, transportados ao laboratório sob refrigeração em caixa de material isotérmico. O CMT consiste na adição de detergente à amostra de leite, causando o rompimento das membranas celulares e liberação do material nucléico, formando uma matriz semelhante a um gel cuja viscosidade é proporcional ao número de células somáticas. Utiliza-se de uma escala para a interpretação que varia de ausente a três cruzes (+++). De acordo com a intensidade da reação classifica-se em: negativa (0), reação leve (+), moderada (++) e intensa (+++) (FONSECA & SANTOS, 2000).

A confecção dos esfregaços lácteos foi baseada no método de PRESCOTT & BREED (1910), no qual uma quantidade de leite equivalente a 0,01 mL foi distribuída de forma homogênea em uma área demarcada por 1 cm², com o auxílio de uma pipeta calibrada (Figura 26). Posteriormente, as lâminas passaram por um processo de secagem durante 24h, em temperatura ambiente, e em seguida foram coradas pelo método de ROSENFELD (1947).



Figura 26: Montagem de lâminas (esfregaços) de leite.

Cada lâmina foi colocada no microscópio, percorrida por um mínimo de 100 diferentes campos (Figura 27-a) e em seguida foram feitas as aquisições das respectivas imagens pelo sistema da Figura 27-b.

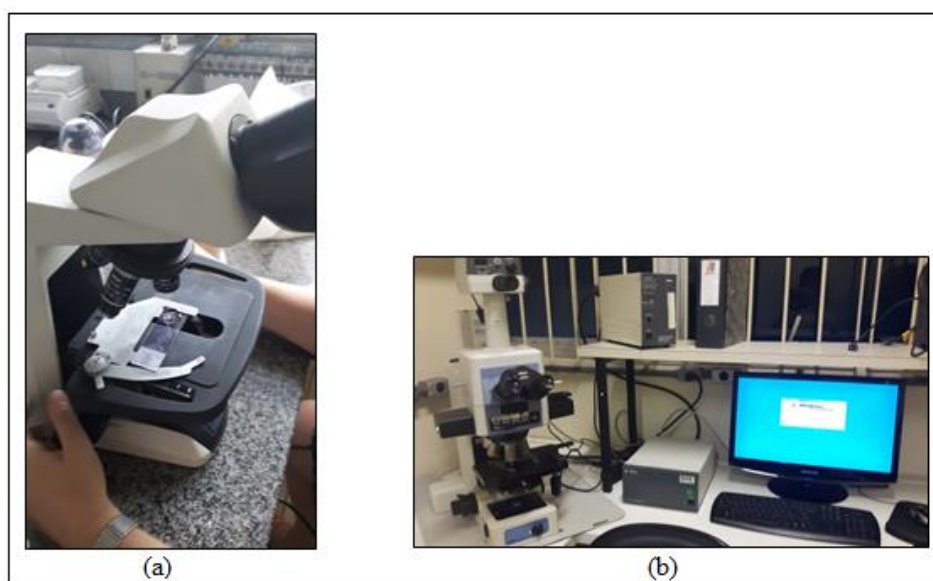


Figura 27: (a) Observação das lâminas e (b) Sistema de aquisição de imagens.

A Tabela 1 mostra o número de imagens obtidas por cada lâmina de cada animal, contabilizando-se um total de 511 imagens capturadas.

Tabela 1: Quantidade de imagens capturadas por amostra de leite

Animal 1	Animal 2	Animal 3	Animal 4	Animal 5
100	103	100	100	108

As imagens foram cedidas pela Faculdade de Medicina Veterinária e Zootecnia – FMVZ/USP. A captura das imagens foi feita por uma câmera fotográfica digital (CoolSnap Color, Media Cybernetics™, USA) acoplada ao microscópio óptico (Eclipse E800, Nikon®, Japan) com aumento de 1000 vezes.

Após a aquisição das imagens no espaço de cores RGB, o procedimento experimental compreendeu as seguintes etapas:

- Transformação das imagens em RGB para o padrão monocromático;
- Uso do *Fuzzy C-Means* cujo objetivo foi separar em diferentes grupos as imagens da amostra total que apresentassem maior similaridade de intensidade de cinza;
- Cálculo dos limiares dos grupos utilizando os métodos de Otsu (1979), Rosin (2001) e De Melo (2015);
- Segmentação das células somáticas, uso da Transformada de *Watershed* e contagem das células.

A Figura 28 apresenta um resumo da metodologia desenvolvida.

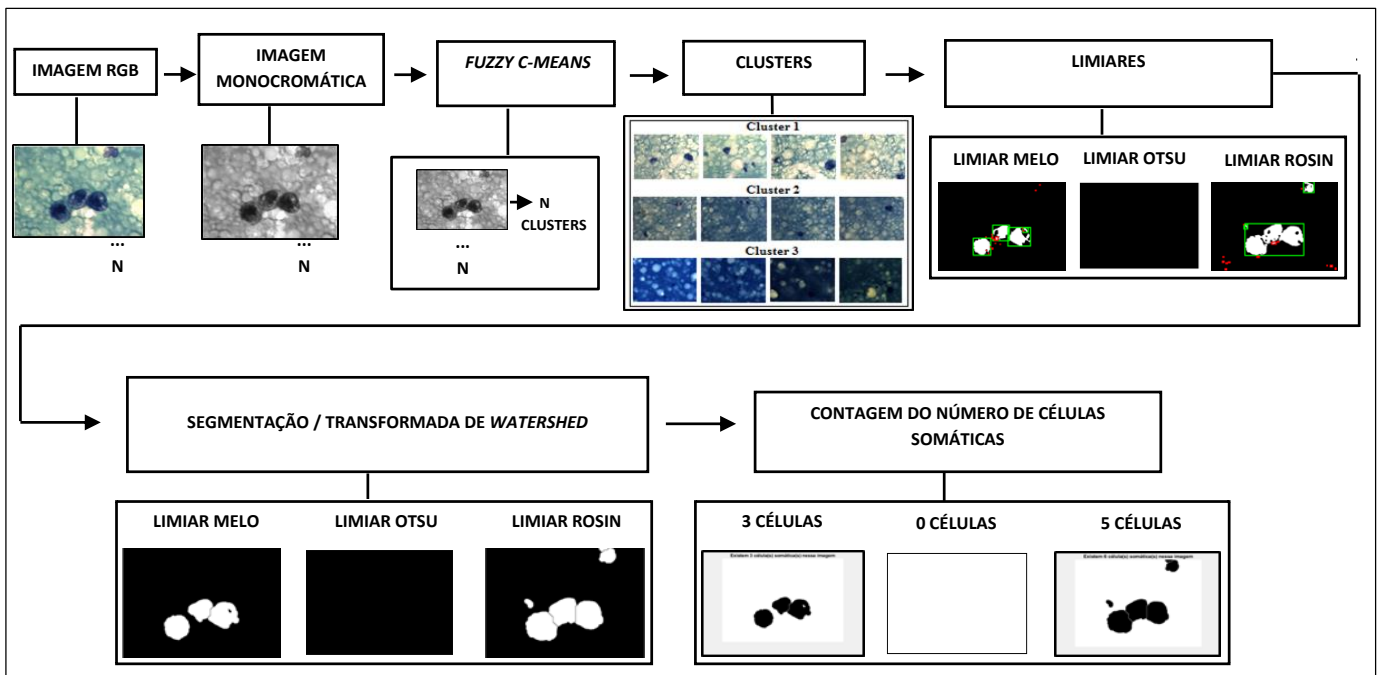


Figura 28: Resumo do método utilizado para contagem de células somáticas

Para a realização das etapas que envolviam as técnicas de pré-processamento e de processamento de imagens foi desenvolvido um algoritmo no *Matlab* R2012b.

As 511 imagens feitas das cinco lâminas foram obtidas no espaço de cores RGB e tinham um tamanho original de 1392x1040 *pixels*. Para viabilizar a manipulação destas imagens e a redução da dimensão das matrizes de intensidade de cores que as representam, elas tiveram o tamanho modificado para a dimensão de 535x400 *pixels* através de um programa de edição, sem prejuízo da qualidade da visualização (sem perdas para identificação das células).

O método proposto, CCSFPI (Contagem de Células Somáticas através de agrupamento *Fuzzy* e Processamento de Imagens), recebe como entrada as imagens em RGB já modificadas (de dimensão 535x400 *pixels*) e as converte para monocromática de acordo com o que foi descrito no capítulo anterior (seção 2.2.2).

A conversão de uma imagem colorida para tons de cinza é requisito para muitas aplicações em processamento de imagens e permite que a imagem seja representada por apenas uma matriz de tons de cinza (valores variam de 0 a 255) ao invés de três matrizes como na imagem em RGB. Por se tratar de uma redução de dimensão no espaço de cores, haverá perda de informação de cor, mas preservando-se o máximo possível do contraste original. A Figura 29 ilustra uma imagem em RGB e, para efeito de compreensão e representação das matrizes, foi retirada uma área menor dessa imagem visando ilustrar o processo de conversão para o monocromático.

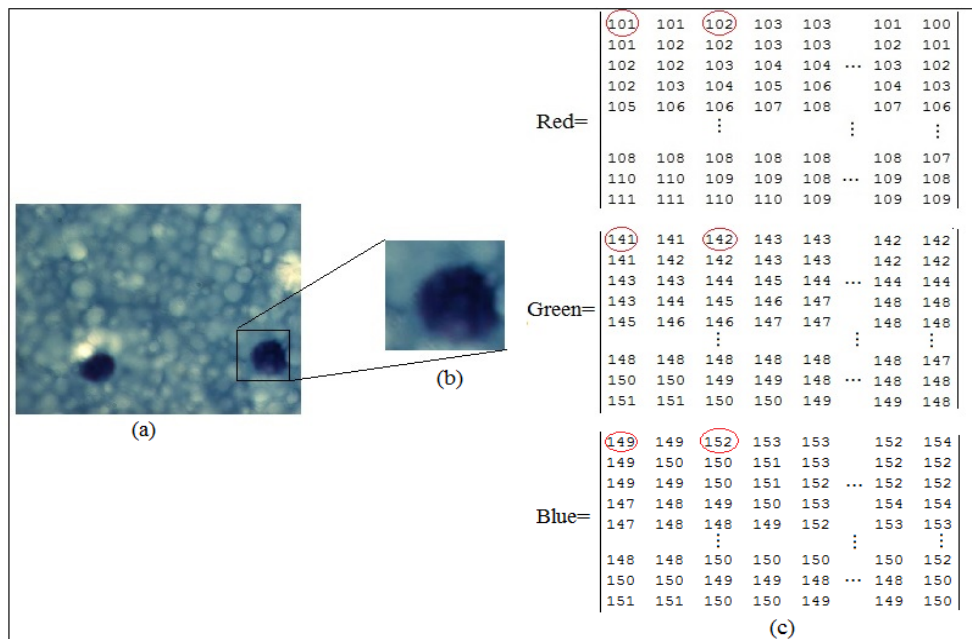


Figura 29: Representação da imagem RGB para conversão: (a) Imagem original; (b) Área menor da imagem RGB; (c) Parte das matrizes que representam a área menor da imagem RGB.

Tomando como base os elementos da linha 1, coluna 1 e da linha 1, coluna 3 (em destaque na Figura 29-c), tem-se:

Elemento linha 1, coluna 1:

$$\text{EscalaCinza} = 0.299 \times \text{Red} + 0.587 \times \text{Green} + 0.114 \times \text{Blue}$$

$$\text{EscalaCinza} = 0.299 \times 101 + 0.587 \times 141 + 0.114 \times 149$$

$$\text{EscalaCinza} = 129,952130.$$

Elemento linha 1, coluna 3:

$$\text{EscalaCinza} = 0.299 \times \text{Red} + 0.587 \times \text{Green} + 0.114 \times \text{Blue}$$

$$\text{EscalaCinza} = 0.299 \times 102 + 0.587 \times 142 + 0.114 \times 152$$

$$\text{EscalaCinza} = 131,18131$$

A Figura 30 mostra o resultado da conversão da imagem em RGB para monocromática. A matriz 30-c apresenta a comprovação dos valores encontrados para os elementos ilustrados (linha 1, coluna 1 e linha 1, coluna 3).

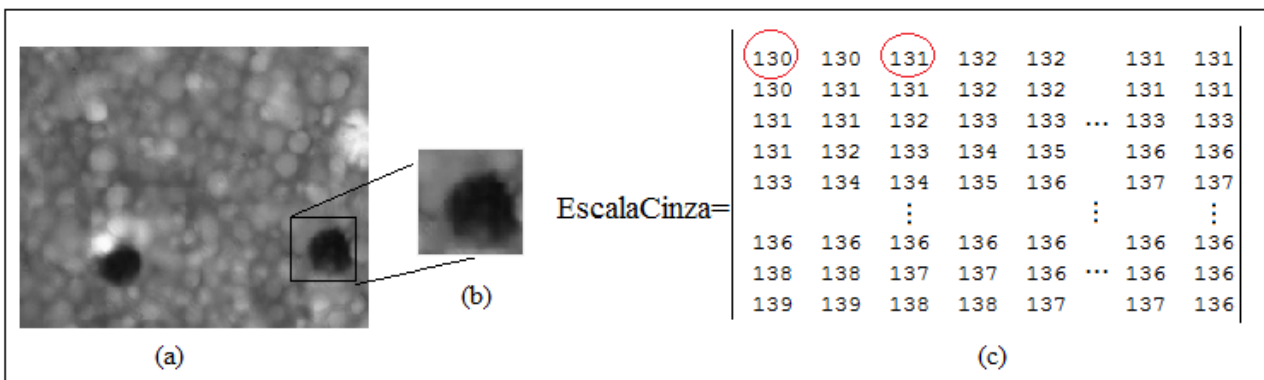


Figura 30: Imagem convertida em monocromática: (a) Imagem monocromática; (b) Área menor da imagem monocromática; (c) Parte da matriz que representa a área menor da imagem monocromática.

A intensidade dos níveis de cinza é a única informação disponível em imagens monocromáticas (CHUANG *et al.*, 2006). A Figura 31 ilustra o procedimento de aquisição de informações tomando como exemplo uma área menor de uma das imagens da amostra. A partir da imagem em tons de cinza se obtém a matriz de valores (variação de 0 a 255) que a representa. Em seguida é gerado o histograma desta imagem com base nos valores da matriz. Conforme foi visto, o histograma de uma imagem indica a distribuição de *pixels* associados a diferentes níveis de cinza, sendo representado por um gráfico de barras onde para cada nível de cinza é fornecido o número ou o percentual de *pixels* presentes na imagem.

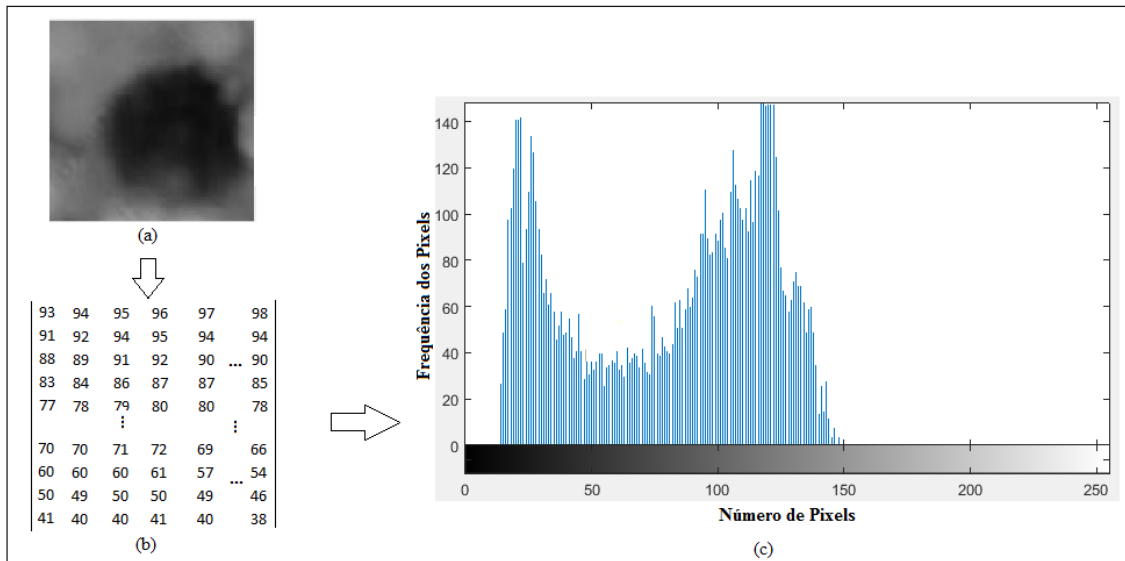


Figura 31: Representação da intensidade da imagem: (a) Imagem monocromática; (b) Matriz da imagem; (c) Histograma da imagem.

Para cada histograma referente às 511 imagens, os produtos entre o valor de cada *pixel* e sua respectiva frequência fornecem uma matriz com a distribuição da intensidade de cinza de cada imagem. A Figura 32 ilustra parte das matrizes de *pixels* e das frequências referentes à imagem da Figura 31-a e a respectiva matriz com a distribuição de intensidade de cinza. Para cada imagem foi gerada uma única matriz na qual cada linha contém a distribuição de intensidade de cinza.

Matriz com Valores dos Pixels																																						
Colunas	1	2	3	4	5	6	7	8	9	10	...	50	51	52	53	54	55	56	...	100	101	102	103	104	105	...	245	246	247	248	249	250	251	252	253	254	255	256
Linha 1	0	1	2	3	4	5	6	7	8	9	...	49	50	51	52	53	54	55	...	99	100	101	102	103	104	...	244	245	246	247	248	249	250	251	252	253	254	255

(a)

Matriz com Valores das Frequências dos Pixels																																										
Colunas	1	2	3	4	5	6	7	8	9	10	...	50	51	52	53	54	55	56	...	100	101	102	103	104	105	...	245	246	247	248	249	250	251	252	253	254	255	256				
Linha 1	0	0	0	0	0	0	0	0	0	0	...	31	39	33	41	40	40	26	...	92	89	98	101	86	81	...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(b)

Matriz com Valores da Distribuição da Intensidade de Cinza da Imagem																																											
Colunas	1	2	3	4	5	6	7	8	9	10	...	50	51	52	53	54	55	56	...	100	101	102	103	104	105	...	245	246	247	248	249	250	251	252	253	254	255	256					
Linha 1	0	0	0	0	0	0	0	0	0	0	...	1519	1950	1683	2132	2120	2160	1430	...	9108	8900	9898	10302	8858	8424	...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(c)

Figura 32: Matrizes: (a) Parte da matriz com os valores dos *pixels*; (b) Parte da matriz com as frequências dos *pixels*; (c) Parte da matriz com a distribuição da intensidade de cinza da imagem.

A matriz final obtida é um dos parâmetros que serão fornecidos para o *Fuzzy C-means*, juntamente com o número de grupos que se deseja identificar/reconhecer. Cada linha dessa matriz que contém a distribuição de intensidade de cinza em relação à posição, referente a cada imagem, foi avaliada de modo que imagens com similaridade

na distribuição espacial de intensidade de tons de cinza fossem colocadas no mesmo grupo e as que apresentassem dissimilaridade fossem colocadas em grupos diferentes. O FCM reconheceu um padrão que representa a distribuição da intensidade de cinza para cada grupo. A métrica de similaridade adotada pelo *Fuzzy C-means* foi a distância Euclidiana. Desta forma, foi possível abordar o problema de agrupamento e reconhecimento em imagens bidimensionais usando uma metodologia de agrupamento não hierárquico onde cada imagem (objeto) é representada por um vetor de características que compreende a distribuição de intensidade de cinza.

A aplicação de um método de agrupamento de imagens se justifica pela heterogeneidade da amostra o que, por sua vez, inviabiliza o uso de um único limiar para todas as imagens. Dessa forma, foi adotado um limiar para cada grupo identificado/reconhecido. A Figura 33 ilustra diferentes imagens e mostra o nível de heterogeneidade entre elas.

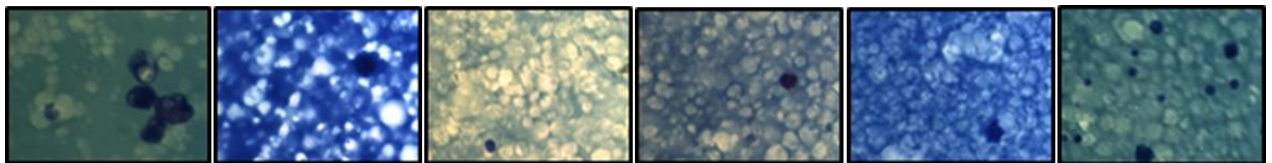


Figura 33: Exemplos de imagens das amostras (VIVIANI GOMES, 2016).

A definição do número de grupos a ser considerado no método FCM foi realizada através de observação visual com base na variação da tonalidade de fundo das imagens e na contagem do número de células.

A partir do procedimento de agrupamento das imagens pelo FCM, duas matrizes foram obtidas: a matriz com o centro de cada grupo e a matriz com os valores das pertinências de cada imagem aos respectivos grupos. A matriz do centro contém os valores de distribuição de cinza da imagem padrão (representativa) do grupo e, portanto, é possível obter o histograma da imagem padrão de cada grupo. Neste caso, uma vez que os valores da distribuição de intensidade de cinza são resultantes do produto entre o valor do *pixel* e de sua frequência, os valores de frequência foram obtidos considerando-se que os valores de *pixels* são fixos e conhecidos (0 a 255) (Figura 34).

Matriz com Valores da Distribuição da Intensidade de Cinza da Imagem																												
Coluna de 1 a 256	1	2	3	4	...	40	41	42	43	44	45	46	47	48	49	50	51	...	100	101	102	103	...	252	253	254	255	256
Linha 1	0	0	0	0	...	1131	1680	1681	1764	1376	1716	1564	1692	1440	1715	1250	...	280764	315800	340875	435846	...	0	0	0	0	0	0

(a)

Matriz com Valores dos Pixels																													
Coluna de 1 a 256	1	2	3	4	...	40	41	42	43	44	45	46	47	48	49	50	51	...	100	101	102	103	...	252	253	254	255	256	
Linha 1	0	1	2	3	...	39	40	41	42	43	44	45	46	47	48	49	50	51	...	99	100	101	102	...	251	252	253	254	255

(b)

Matriz com Valores das Frequências dos Pixels																												
Coluna de 1 a 256	1	2	3	4	...	40	41	42	43	44	45	46	47	48	49	50	51	...	100	101	102	103	...	252	253	254	255	256
Linha 1	0	0	0	0	...	29	42	41	42	32	39	38	34	36	30	35	25	...	2836	3158	3375	4273	...	0	0	0	0	0

(c)

Figura 34: Matrizes: (a) Parte da matriz com a distribuição da intensidade de cinza da imagem; (b) Parte da matriz com os valores dos *pixels*; (c) Parte da matriz com as frequências dos *pixels*.

Conhecidos os valores das frequências dos *pixels* e dos *pixels* foi possível construir o histograma de cada imagem padrão do respectivo grupo.

Através da limiarização, cada imagem é segmentada em duas partes (fundo e objeto), permitindo que apenas os objetos (a(s) célula(s) propriamente dita(s)) sejam identificados e contabilizados. A aplicação de um único limiar para toda a amostra (todas as imagens) gera a ocorrência de resíduos, ou seja, pontos que deveriam pertencer ao fundo da imagem e erroneamente são identificados como célula, acarretando erro de contagem. A Figura 35 apresenta quatro imagens que foram tratadas da mesma forma (com o mesmo limiar) e no processo de segmentação pode-se observar que o resultado não é aceitável para duas delas (imagens 2 e 3) uma vez que partes do fundo da imagem foram identificadas como objeto (células). Observa-se, porém, que as imagens 1 e 4 que apresentam similaridade no tom de fundo foram segmentadas de forma correta. Neste caso, o mesmo limiar foi adequado para ambas pela semelhança na principal característica (distribuição da intensidade de cinza).

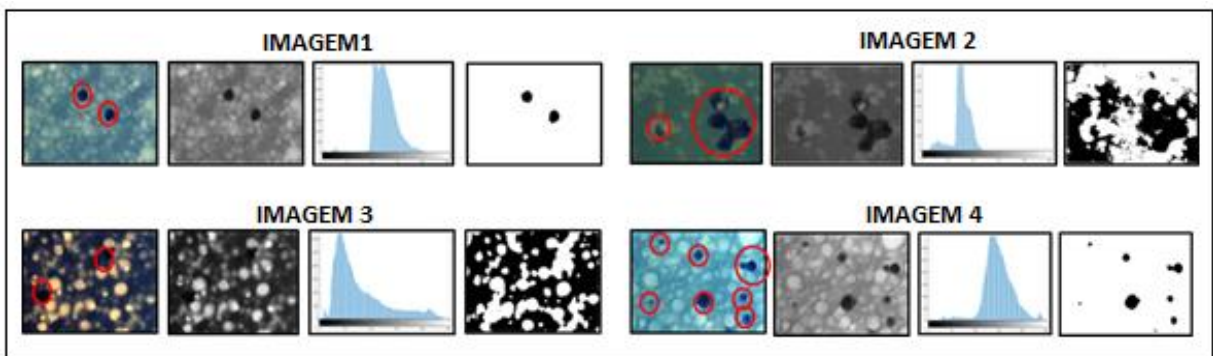


Figura 35: Exemplos de imagens tratadas com o mesmo limiar.

A etapa de segmentação compreendeu a aplicação dos métodos de limiarização de De Melo (2015) e os tradicionais Métodos de Otsu (1979) e Rosin (2001).

Nos métodos de Otsu e Rosin os limiares foram encontrados de forma automática. O algoritmo de Otsu (1979) consiste em um método iterativo que busca valores de limiares associados à mínima soma da variância entre classes da imagem. Neste sentido, o histograma da imagem é calculado em tons de cinza e, para cada valor de um possível limiar, são obtidos os pesos (probabilidade de um *pixel* pertencer à classe do fundo ou da frente) e as médias para as classes que representam a frente e o fundo da imagem. Os valores obtidos são usados no cálculo da variância entre classes. O limiar escolhido é o que fornece o maior valor para a variância entre classes. Ocorrera a iteração para cada *pixel* da imagem e a verificação do seu valor, se esse valor for maior ou igual ao limiar, a cor desse *pixel* será redefinida como branca, caso seja menor, a cor será redefinida como preta, tendo-se como resultado uma imagem binária (TOROK & CONCI, 2016).

O algoritmo de Rosin considera que existe uma classe dominante na imagem que produzirá um pico principal e uma classe secundária que pode ou não apresentar um pico que possa ser identificado. Uma reta é traçada no histograma da imagem ligando o pico (intensidade com maior número de *pixels*) à primeira intensidade sem *pixel* (vazia). O limiar escolhido é a intensidade que maximiza o comprimento do segmento perpendicular a esta reta que finaliza no nível de cinza do histograma (Figura 36) (ROSIN, 2001).

O método de Melo (Figura 36) parte da mesma premissa do método de Rosin (2001), sendo, porém necessário encontrar os valores dos parâmetros de ponderação L (seção 2.2.4).

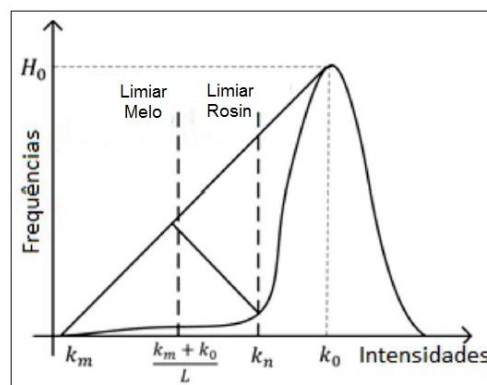


Figura 36: Procedimento para a determinação do limiar de intensidade em um histograma unimodal (DE MELO, 2015).

De acordo com o método de Melo, para se obter o valor do parâmetro L , o mesmo deve ser considerado igual a L_{max} . Para isso, são determinados valores de k (limiar) para N valores discretos de, (L_1, L_2, \dots, L_N) , considerando intervalos uniformes com espaçamento de 0,1 gerando-se uma sequência de 100 pontos. Para cada valor de L nesta sequência, um único valor de k é obtido para todo o conjunto de imagens. O coeficiente de correlação entre os dados obtidos da contagem da metodologia CCSFPI e os dados obtidos através das respectivas contagens dos especialistas foram utilizados para verificar qual o melhor valor de L para as imagens analisadas (Figura 37). O melhor valor de k é dado por:

$$k = \frac{k_m + k_0}{L_{max}}$$

Os valores de L_{max} foram computados para cada grupo de imagens.

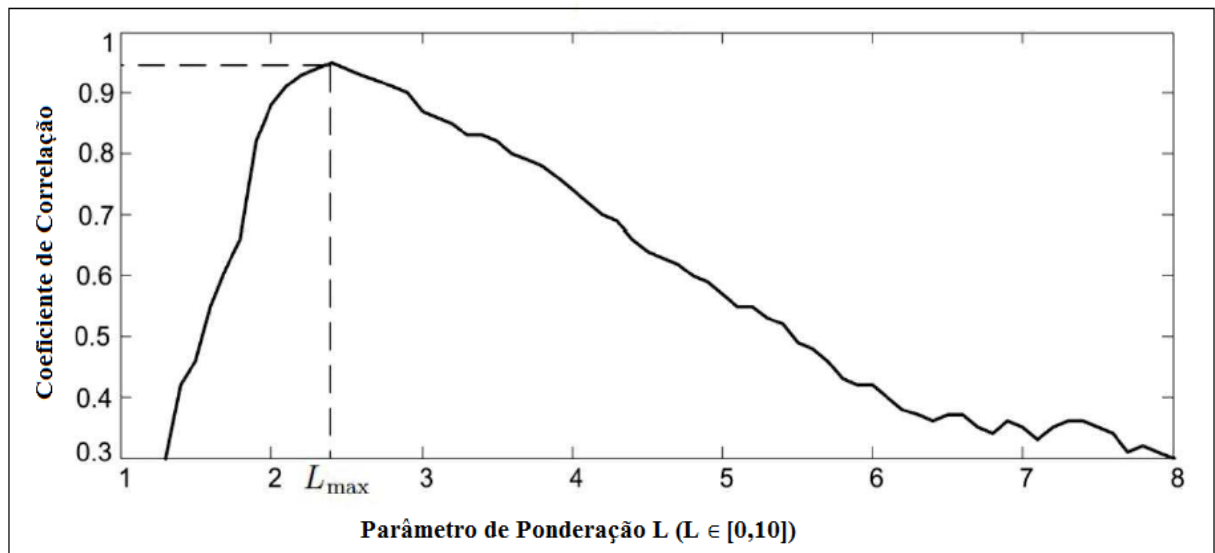


Figura 37: Índice de correlação C_n em função do parâmetro de ponderação L (DE MELO, 2015).

Para cada imagem pertencente a um dado grupo se obteve um valor de limiar específico e, como as imagens apresentam similaridade, a média aritmética dos limiares de todas as imagens foi utilizada como limiar do grupo. Após a limiarização com os três métodos, cada imagem foi transformada em preto e branco, através do processo de segmentação.

Algumas imagens em preto e branco apresentaram células em grupos ou muito próximas (Figura 38) causando a união dos *pixels* e transformando o que seriam duas

células em apenas uma. A Transformada de *Watershed* foi aplicada para inserir linhas divisórias entre os objetos, separando-os.

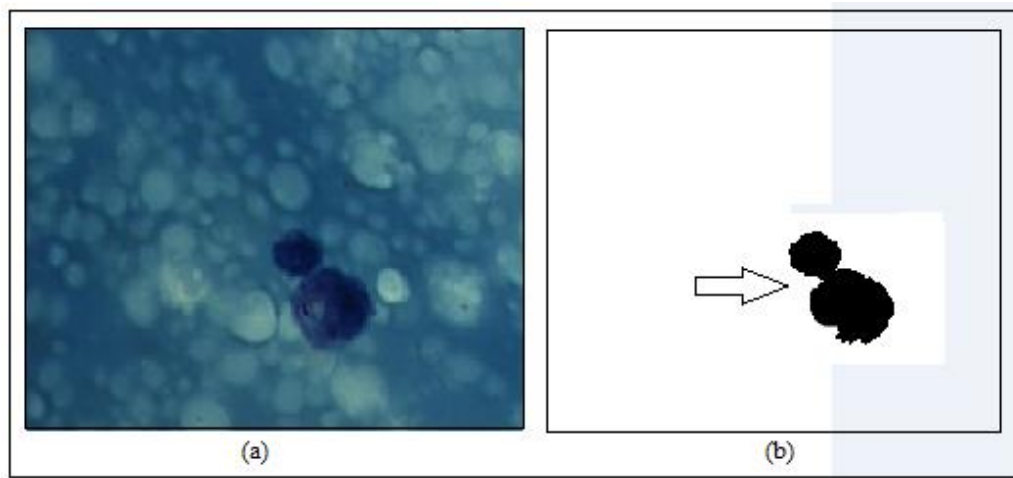


Figura 38: Imagem de Células Somáticas: (a) Duas células; (b) União dos *pixels* das células formando apenas uma (VIVIANI GOMES, 2016).

No processo de contagem, as células somáticas foram identificadas através de um procedimento que traça os contornos dos objetos que estão presentes em uma imagem preto e branco (binária) utilizando como referência uma matriz que contém a localização de *pixels* nos limites de uma área. Os *pixels* 1 (Branco) são do objeto (célula) e 0 (Preto) referem-se ao fundo da imagem.

Após a identificação dos objetos (células), em algumas imagens foi necessário a retirada de resíduos que estavam presentes (considerados como ruídos). Para este procedimento foi calculada a área de todos os objetos presentes em cada imagem e foram excluídos os objetos com áreas muito pequenas (inferior a 110 *pixels*). Durante o processo de contagem, os objetos eram identificados, contados e somados para o cálculo do total de células na amostra. A Figura 39 ilustra o processo de identificação/contagem de células em uma imagem.

Os valores obtidos foram comparados com a contagem feita pelos especialistas e uma análise estatística foi utilizada para comprovar a eficiência desta metodologia.

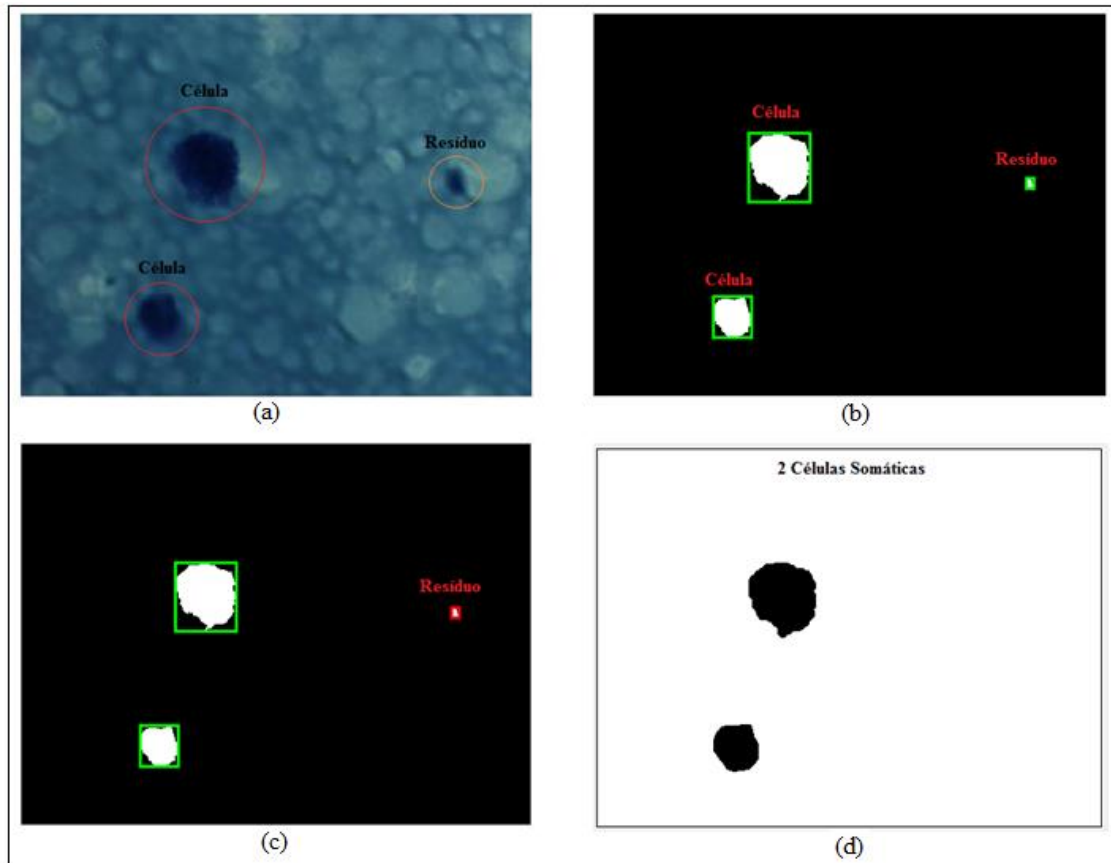


Figura 39: Etapa de contagem das células: (a) Imagem original; (b) Reconhecimento dos objetos na imagem; (c) Reconhecimento do resíduo; (d) Retirada do resíduo, deixando as células.

3.1 RESULTADOS

A Figura 40 ilustra a transformação da configuração RGB em monocromática para uma das imagens da amostra juntamente com o histograma e a matriz que a representa. Verifica-se através da Figura 40-b o desempenho satisfatório do processo de monocromatização realizado através do método descrito na seção 2.2.2.

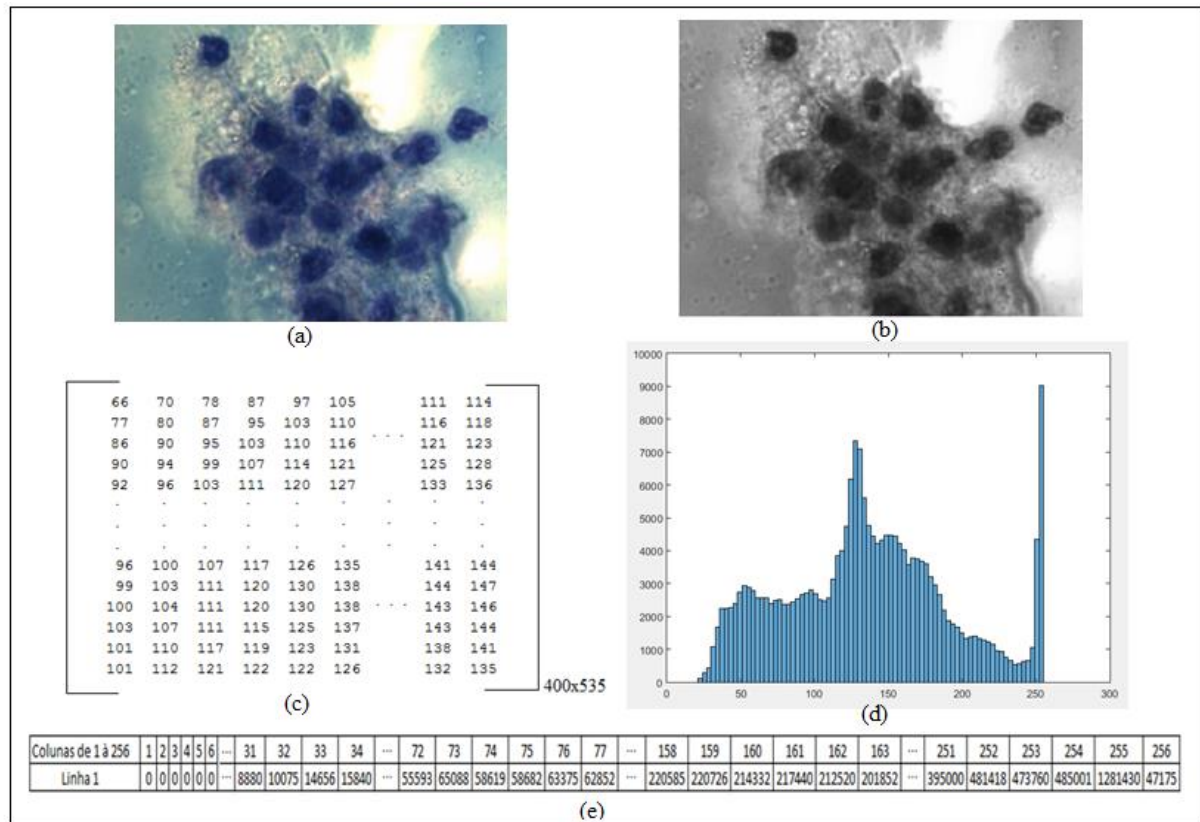


Figura 40: Células somáticas: (a) Imagem RGB original; (b) Imagem transformada em monocromática; (c) Parte da matriz da imagem monocromática; (d) Histograma da imagem monocromática; (e) Parte da matriz com a distribuição da intensidade de cinza da imagem.

O procedimento de monocromatização foi realizado para todas as imagens da amostra global e posteriormente foram obtidas as distribuições das intensidades de cinza, gerando uma matriz de 511x256. Cada linha da matriz corresponde a um objeto (uma imagem) e as colunas correspondem aos atributos de cada objeto (intensidade de cinza em uma dada posição ou *pixel*). A Figura 40-e mostra um vetor de atributos da imagem ilustrada.

Para a escolha do número de grupos foram feitos testes com 2, 3, 4, 5 e 6 grupos. Verificou-se que um número de grupos superior a 5 não implicou em alteração significativa dos resultados visuais e de contagem e abaixo deste valor os resultados não foram satisfatórios.

A tabela 2 apresenta o número de imagens em cada um dos 5 grupos identificados/reconhecidos pelo método FCM.

Tabela 2: Número de imagens por grupo separadas pelo FCM.

Número de Imagens por Grupo				
Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
84	108	130	127	62

A validação da qualidade do agrupamento foi realizada através de verificação da similaridade visual entre imagens de cada grupo e da contagem do número de células. O apêndice A apresenta todas as imagens de cada um dos grupos identificados.

A aplicação de um método de agrupamento *Fuzzy* (FCM) foi capaz de efetivamente reconhecer padrões de similaridade entre as diversas imagens disponíveis da amostra. Isto sugere, portanto uma abordagem sistemática de agrupamento inicial de imagens importante para a qualidade das etapas seguintes do processo de contagem de células somáticas.

A partir dos valores das matrizes de cada centro foi possível gerar o histograma da imagem padrão criada para cada grupo. Estes histogramas estão representados nas Figuras 41 a 45.

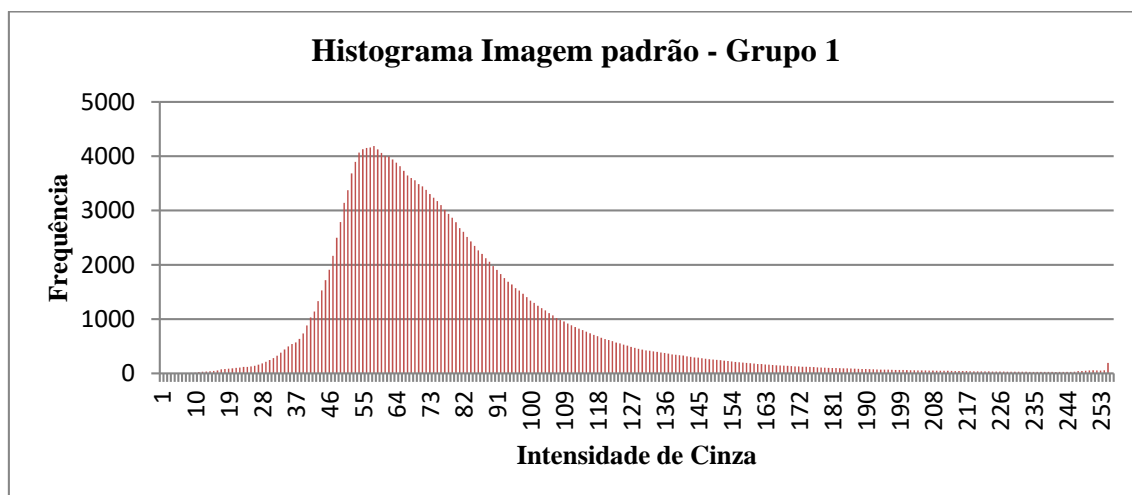


Figura 41: Histograma gerado a partir de informações da matriz do centro do grupo 1.

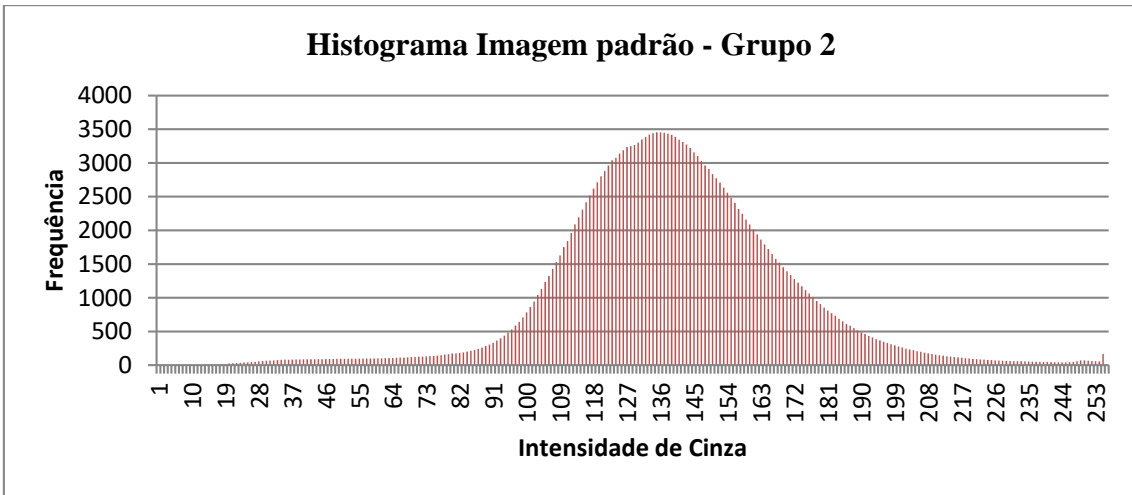


Figura 42: Histograma gerado a partir de informações da matriz do centro do grupo 2.

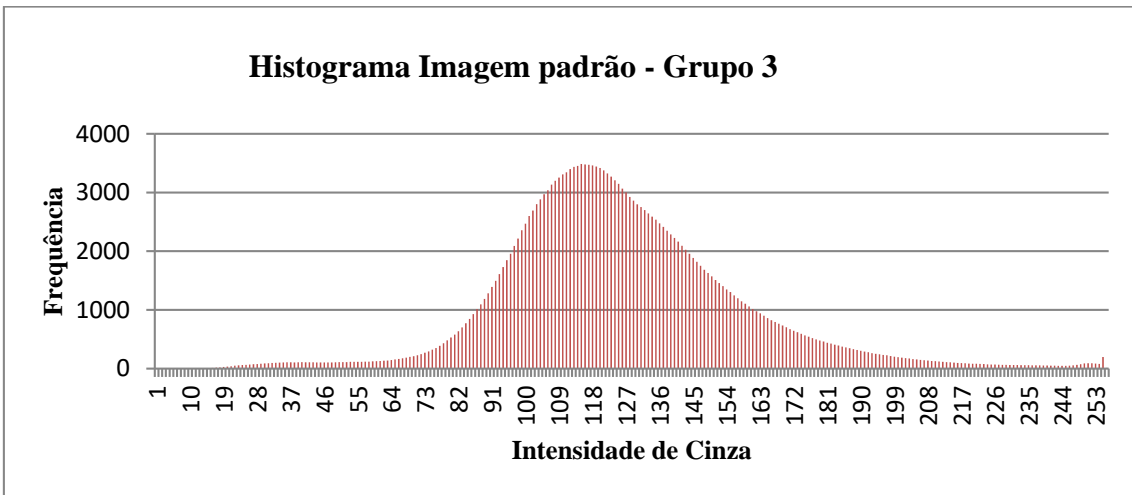


Figura 43: Histograma gerado a partir de informações da matriz do centro do grupo 3.

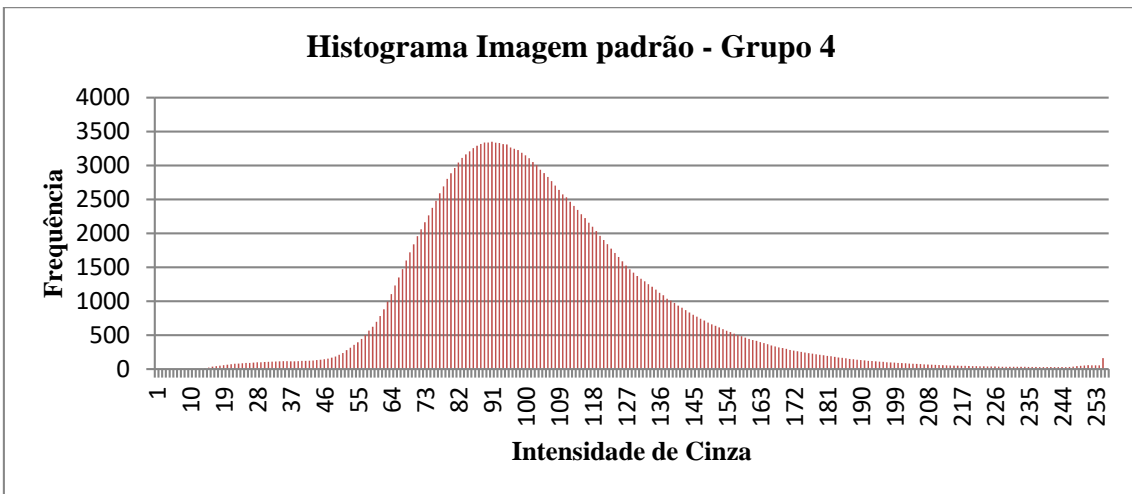


Figura 44: Histograma gerado a partir de informações da matriz do centro do grupo 4.

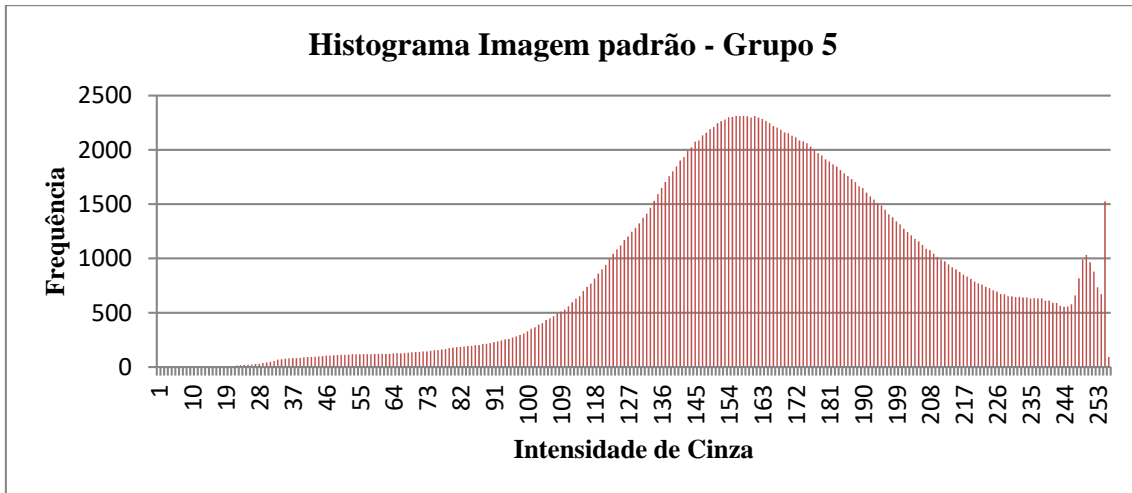


Figura 45: Histograma gerado a partir de informações da matriz do centro do grupo 5.

Verifica-se que os grupos 1 e 4 apresentam características de imagens com tons mais escuros, sendo o grupo 1 o que representa a imagem mais escura. Os grupos 2, 3 e 5 apresentam características de imagens com tons intermediários, tendendo para mais claros, sendo o do grupo 5 o que representa a imagem mais clara em relação a todas as outras. Através das imagens agrupadas é possível observar e comprovar essa análise (Ver Apêndice A).

A Figura 46 apresenta o grau de pertinência de cada objeto ao seu respectivo grupo. O intervalo de valores verificados revela um bom nível de polarização do processo de agrupamento que está associado à qualidade do método em reconhecer a incerteza inerente às fronteiras entre os grupos.

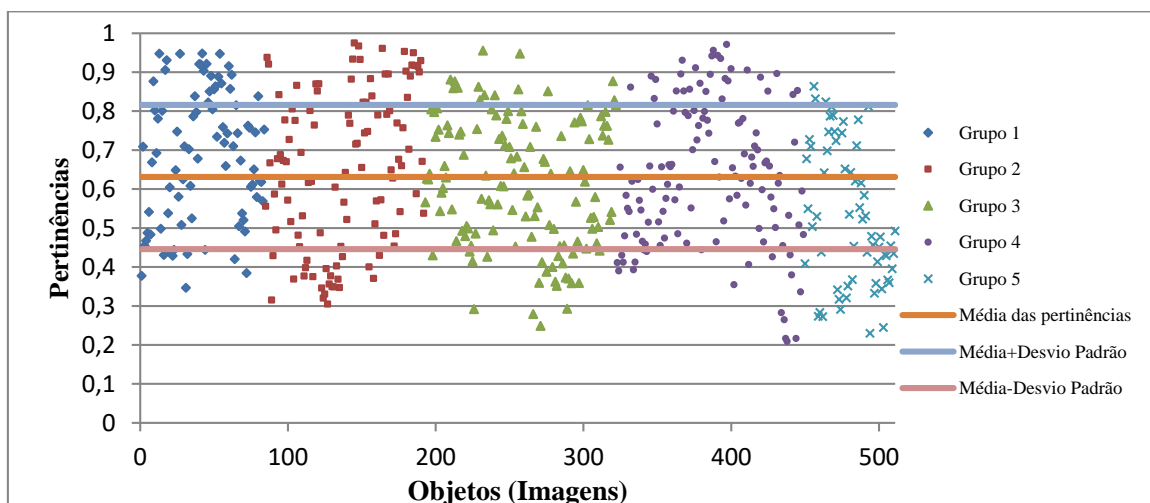


Figura 46: Gráfico das pertinências das imagens nos grupos.

A aplicação do método FCM compreende essencialmente um problema de agrupamento no qual o objetivo é a separação das imagens em grupos de acordo com a tonalidade de cinza. Embora o resultado final da contagem (apresentado posteriormente) seja por si só, um indicador (indireto) da qualidade do agrupamento obtido pelo método FCM, é importante comparar os resultados obtidos (Figuras 41-45) com diferentes quantidades de grupos (2 a 6), segundo um critério interno de validação capaz de avaliar a qualidade do agrupamento em cada caso.

Dentre os índices internos de validação de agrupamento propostos na literatura (XIONG & LI, 2013), foi adotado nesta etapa o índice de Xie-Beni que avalia a qualidade do resultado considerando a matriz de partição (pertinências) juntamente com os centros/padrões reconhecidos para cada grupo.

$$XB = \frac{\sum_{i=1}^p \sum_{k=1}^n \mu_{ik}^2 d(x_k; c_i)}{n \cdot \left(\min_{i \neq j} \|c_i - c_j\|^2 \right)}, \text{ onde } c_i \text{ (} i = 1, \dots, p \text{) é o centro do grupo } i.$$

O numerador é o valor da função objetivo (FO) final obtida pelo FCM. O denominador calcula a distância mínima entre centros de grupos diferentes (XIONG & LI, 2013).

O índice de XB relaciona compacidade (homogeneidade) dos grupos com a distância entre os seus respectivos centros (grau de separação entre os grupos). Portanto, quanto menor a compactação e maior a distância entre os centros (menor XB), melhor a qualidade do agrupamento obtido.

A Tabela 3 apresenta os índices de XB referentes aos diferentes números de grupos (2 a 6), a partir da aplicação do método FCM. Os valores mostram que os melhores resultados de agrupamento estão associados ao reconhecimento de 4 ou 5 grupos. A opção pela escolha de 5 grupos foi reforçada pelos melhores resultados de contagem obtidos posteriormente.

Tabela 3: Índices de Xie Beni

Índice de Xie Beni				
2 grupos	3 grupos	4 grupos	5 grupos	6 grupos
0,413	0,359	0,292	0,301	0,348

Definidos os grupos com suas respectivas imagens, a próxima etapa compreendeu a definição dos limiares utilizando os métodos de Melo, Otsu e Rosin.

Para o método de limiarização de Melo foram necessários os valores de parâmetros de ponderação L. A Tabela 4 apresenta os valores de L obtidos para cada grupo.

Tabela 4: Valores de L (Limiar Melo)

Parâmetro de ponderação L (L ∈ R)				
Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
2,4	2,2	2,2	2,2	2,2

O maior valor de L (grupo 1 - menor limiar) está de acordo com o esperado, pois à medida que a imagem é mais escura, o limiar necessário para segmentá-la é menor. Em relação aos demais grupos, a Tabela 5 mostra que os valores dos limiares são maiores na medida em que as tonalidades (Apêndice A) decrescem.

As médias dos limiares de Melo, Otsu e Rosin para cada grupo estão apresentadas na Tabela 5.

Tabela 5: Médias dos limiares de Melo, Otsu e Rosin de cada grupo.

Limiares (em Pixels)					
	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
Melo	27,8522	62,9756	54,1189	42,4911	74,0836
Otsu	0,3544	0,5521	0,5147	0,4419	0,4435
Rosin	41,1190	93,7222	78,2769	58,8462	59,2231

Estes resultados confirmam a necessidade de se utilizar um limiar específico para cada grupo e, portanto, a importância da técnica de um agrupamento prévio das imagens, pois é possível observar que cada grupo obteve valores distintos para os 3 métodos (exceção apenas para os grupos 4 e 5 utilizando o método de Otsu). A proximidade entre os limiares neste caso demonstra que as imagens destes grupos (4 e 5) não apresentaram diferenças em suas tonalidades. Vale ressaltar, inclusive, que foi uma avaliação errônea do método, visto que o mesmo não apresentou resultados satisfatórios no processo de contagem das células.

As médias dos limiares dos grupos obtidas pelo método de Melo foram as mais adequadas uma vez que o resultado na contagem geral das células, neste caso, foi muito próximo das contagens feitas pelos especialistas.

Os limiares obtidos utilizando o método de Rosin foram maiores do que os limiares do método de Melo em 4 dos 5 grupos. Considerando-se que o método de Melo apresentou melhor desempenho no resultado final de contagem das células, isto justifica o desempenho insatisfatório do método de Rosin uma vez que valores mais elevados de limiares neste caso conduziriam a um maior número de resíduos (Falso Positivo) nas imagens.

A Tabela 6 apresenta as médias dos limiares obtidos pelos métodos de Melo com $L=2,4$ e Rosin considerando a amostra original sem agrupamento prévio.

Tabela 6: Limiares obtidos sem o uso do Classificador FCM

Limiar Melo (com $L= 2.4$) (em <i>pixels</i>)	Limiar Rosin (em <i>pixels</i>)
47,3198	73,7065

Comparando os valores das tabelas 5 e 6 é possível observar que utilizar um único limiar obtido tanto com o método de Melo quanto com o de Rosin não seria suficiente para “representar” todas as imagens diante dos diferentes valores de limiares que foram encontrados para os respectivos grupos. Valores de limiares diferentes confirmam que a amostra é bastante heterogênea.

A imagem da Figura 47 foi classificada pelo FCM com maior pertinência ao grupo 2 cujo limiar utilizado para segmentá-la se encontra na Tabela 5. Nesta imagem a proximidade entre as células dificultou a correta identificação, mantendo-se a junção de algumas mesmo depois da segmentação. Neste caso, os *pixels* que estão ao redor das células estão abaixo do valor de limiar utilizado para segmentação. O uso da transformada de *Watershed* foi essencial em situações similares à Figura 47, pois viabiliza a identificação da fronteira de contato entre as células, aprimorando o processo de separação e identificação.

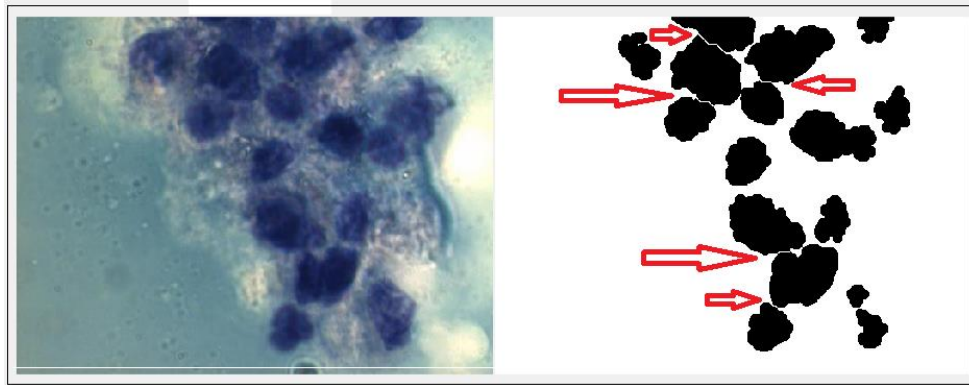


Figura 47: Uso de *Watershed* para separar células em grumos.

A contagem automática das células de todos os animais foi realizada através do método CCSFPI e comparada com a contagem das células realizada por três especialistas diferentes, sem contato uns com os outros para evitar qualquer tipo de interferência. Os 3 especialistas aplicaram o método tradicional DMSCC para efetuar a contagem das células. As Tabelas 7, 8 e 9 apresentam os resultados de contagem.

Tabela 7: Contagem final das células somáticas - Limiar Melo

Contagem Final das Células Somáticas com Limiar Melo			
CCSFPI	Especialista 1	Especialista 2	Especialista 3
1476	1475	1507	1510

Pode-se observar pelos valores apresentados na Tabela 7 a proximidade dos valores encontrados pelo método CCSFPI utilizando o método de limiarização de Melo com a contagem dos especialistas. O agrupamento prévio das imagens, bem como o uso do parâmetro de ponderação L do método proporcionou a adequação dos limiares para os diferentes grupos de imagens com similaridade, tendo como consequência um bom resultado na identificação/contagem das células.

Tabela 8: Contagem final das células somáticas – Limiar Otsu

Contagem Final das Células Somáticas com Limiar Otsu			
CCSFPI	Especialista 1	Especialista 2	Especialista 3
0	1475	1507	1510

De acordo com a Tabela 8, o método de limiarização Otsu foi ineficiente para esta aplicação. Segundo NG (2006), a técnica de Otsu é bastante utilizada e fornece

resultados satisfatórios no processo de limiarização de imagens que apresentem histograma de distribuição bimodal, podendo falhar se o histograma for unimodal ou próximo do unimodal. Verifica-se em alguns histogramas de imagens (Figura 48) que há um comportamento tipicamente unimodal (uma classe maior de *pixels* representada pelo fundo da imagem), o que pode justificar a limitação do método de Otsu para a amostra analisada.

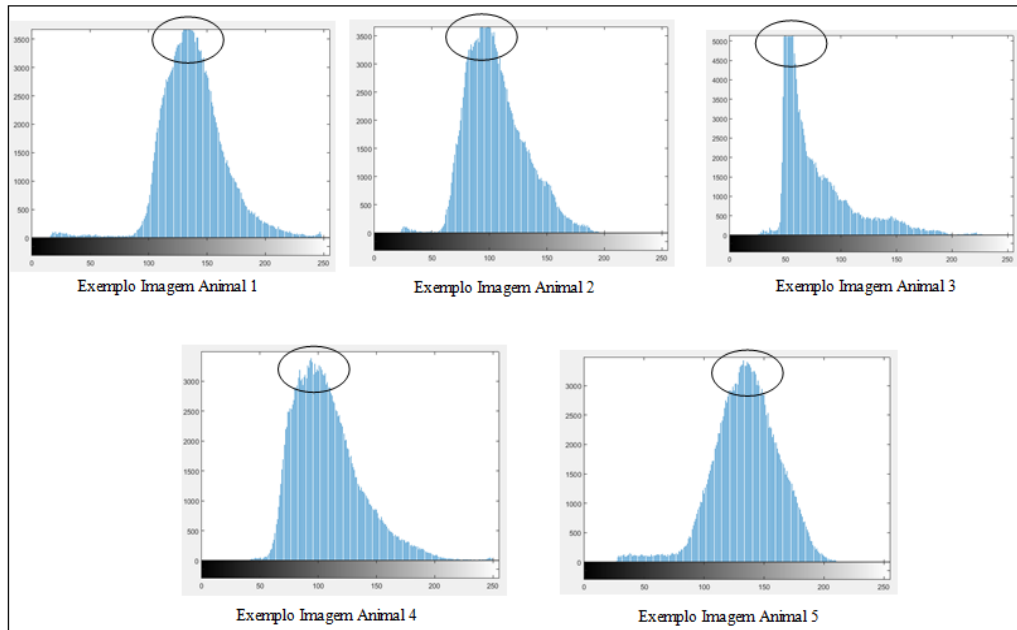


Figura 48: Exemplos de histogramas unimodais das imagens tratadas.

O método de limiarização de Rosin proporcionou a contagem, porém com resultados inconsistentes (Tabela 9).

Tabela 9: Contagem final das células somáticas – Limiar Rosin

Contagem Final das Células Somáticas com Limiar Rosin			
CCSFPI	Especialista 1	Especialista 2	Especialista 3
3838	1475	1507	1510

Apesar do método de Rosin assumir que existe uma classe dominante na imagem (fundo), em relação à classe secundária (células), os limiares encontrados não foram os adequados, gerando resíduos que foram contabilizados como células. O exemplo na Figura 49 mostra que a mesma imagem original tratada com o limiar de Rosin apresenta erro de contagem das células, enquanto que com o limiar de Melo a

contagem é feita de forma correta na imagem analisada. Isso aconteceu com muita frequência, o que sugere que o método de Rosin tende a superestimar a quantidade de células.

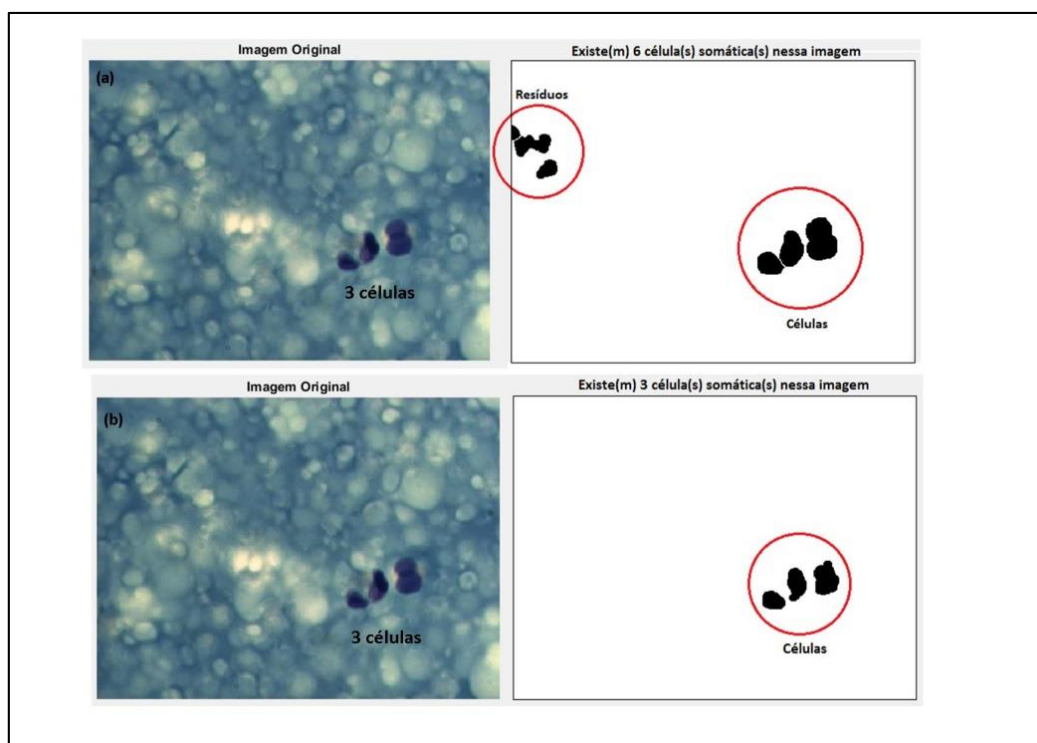


Figura 49: Exemplo de uma imagem original: (a) Tratada com Limiar de Rosin; (b) Tratada com Limiar de Melo.

A Tabela 10 apresenta os resultados obtidos utilizando apenas um limiar para o processo de limiarização/segmentação das células, ou seja, sem a aplicação de um agrupamento prévio das imagens, o que confirma mais uma vez a necessidade deste procedimento.

Tabela 10: Contagem das células somáticas sem uso do FCM.

CCSFPI com Limiar Melo (L=2.4)	CCSFPI com Limiar Rosin	Especialista 1	Especialista 2	Especialista 3
2235	7473	1475	1507	1510

A Figura 50 apresenta uma análise comparativa dos resultados obtidos da contagem das células de todas as imagens da amostra, utilizando o método de limiarização de Melo.

De acordo com os “boxplots” apresentados, verifica-se que as medianas das distribuições são semelhantes, o que mostra que a contagem automática realizada pelo método CCSFPI foi eficiente se comparada à contagem manual realizada pelos três especialistas.

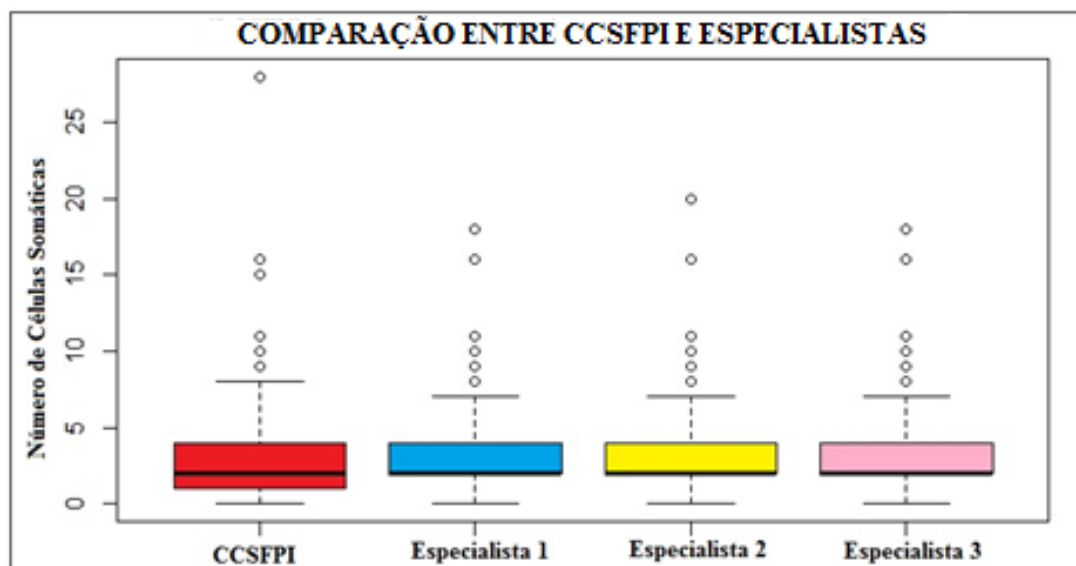


Figura 50: Boxplot comparativo entre CCSFPI (com limiar de Melo) e especialistas.

Em relação à dispersão, verifica-se que todas as medianas (método CCSFPI e especialistas) foram iguais a 2 células somáticas e o terceiro quartil (75% de valores acima da mediana) também foi o mesmo em todas as contagens. O primeiro quartil (25% de valores abaixo da mediana) obtido pelo método CCSFPI foi inferior aos resultados dos especialistas.

Para identificar qual o teste estatístico mais apropriado para a validação dos resultados foi necessário verificar inicialmente um critério importante, qual seja, o comportamento normal da amostra. Os dados são considerados normais ou homogêneos quando os seus valores apresentam uma distribuição em torno de um valor supostamente central.

Foram avaliados todos os dados da contagem (toda a amostra) do método CCSFPI e dos especialistas com base no teste gráfico quantil-quantil que é um procedimento adotado para verificação da normalidade (WILK & GNANADESIKAN, 1968). Este teste compara duas distribuições traçando seus quantis amostrais. Os quantis amostrais são os resultados da contagem real versus os quantis teóricos esperados para uma distribuição normal. Cada quantil é obtido em intervalos regulares que dividem os

dados em subconjuntos de mesma dimensão. Se os pontos dos quantis amostrais se situarem em torno de uma reta (quantis teóricos), então a distribuição se aproxima da distribuição normal. A reta que indica a normalidade é obtida a partir dos pontos $[-1, \bar{x} - s]$, $[1, \bar{x} + s]$, onde \bar{x} é a média e s é o desvio padrão da amostra. Conforme pode ser visto das Figuras 51 a 54 o afastamento da distribuição normal é evidente em todos os casos, mostrando que não há normalidade nos resultados de contagens obtidos.

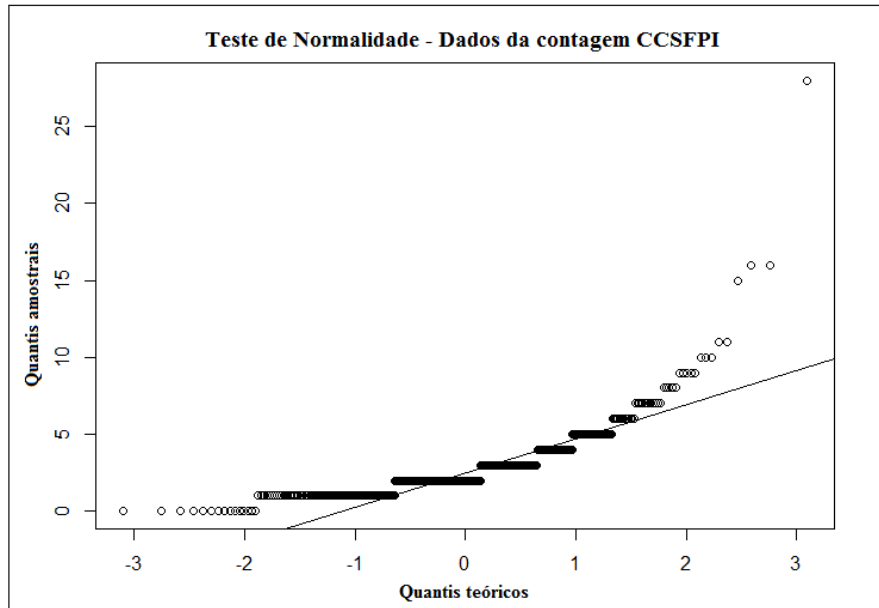


Figura 51: Teste de normalidade com os dados da contagem do método CCSFPI.

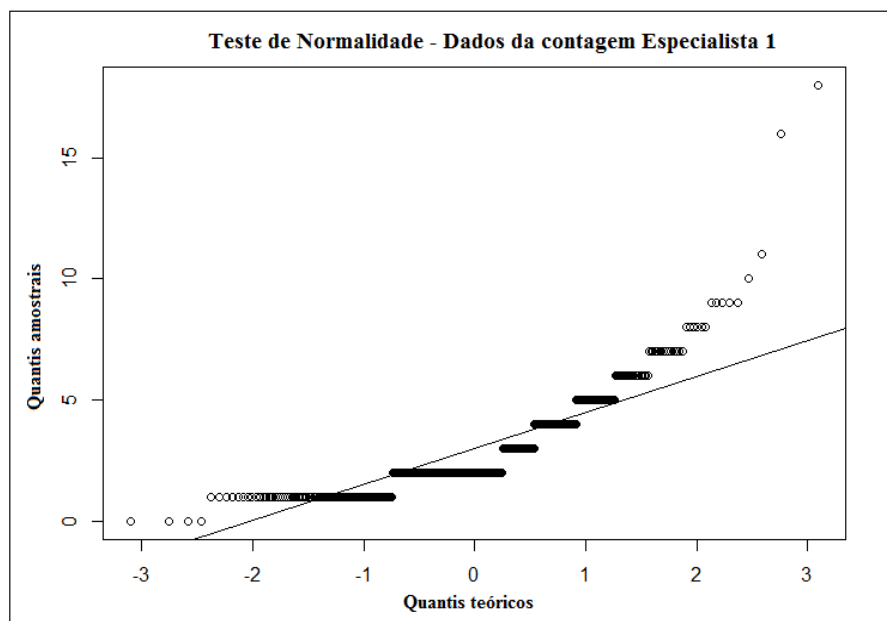


Figura 52: Teste de normalidade com os dados da contagem do especialista 1.

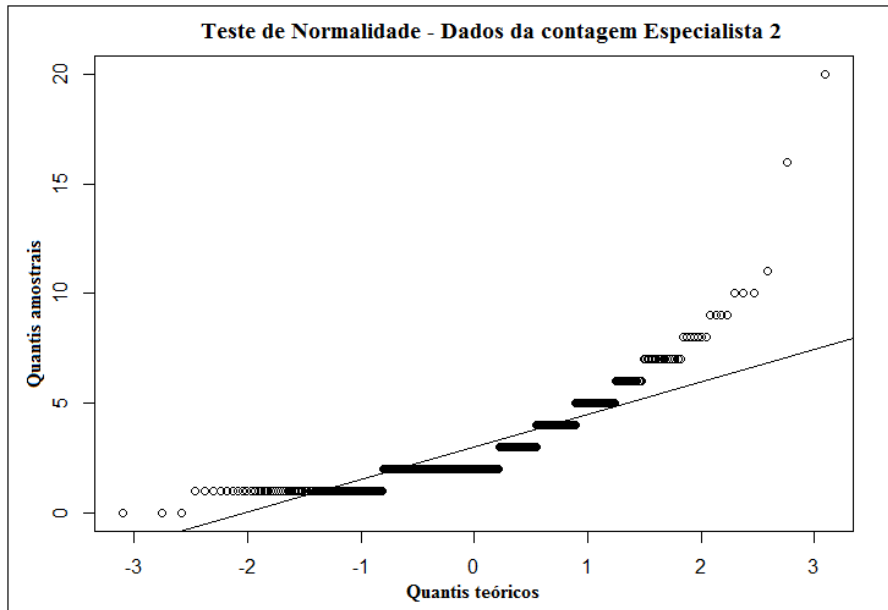


Figura 53: Teste de normalidade com os dados da contagem do especialista 2.

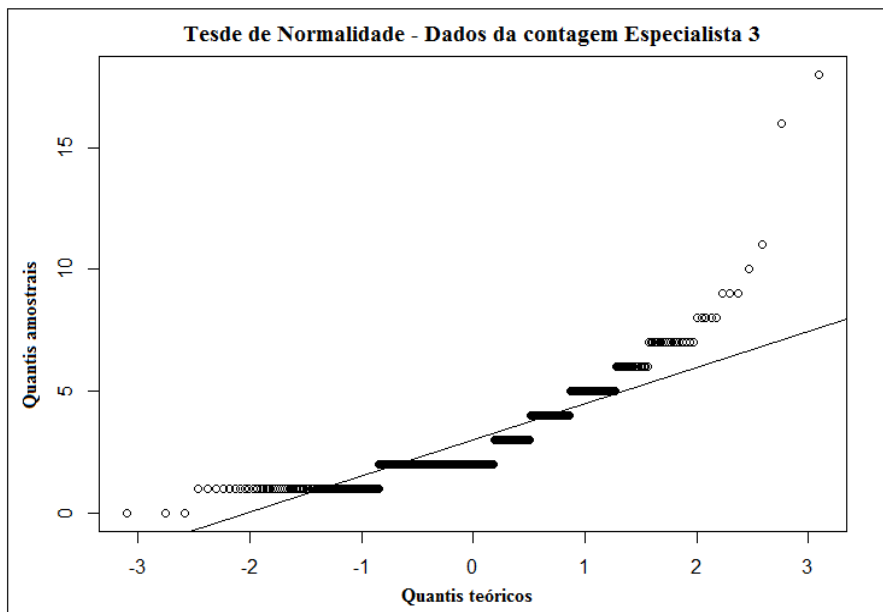


Figura 54: Teste de normalidade com os dados da contagem do especialista 3.

Para uma distribuição que não tenha comportamento normal, testes estatísticos não paramétricos de validação podem ser aplicados (*“Um teste não paramétrico é baseado em um modelo que especifica somente condições gerais e nenhuma a respeito da forma específica da distribuição da qual a amostra foi extraída”*, SIEGEL e CASTELLAN, 1975).

Foi utilizada como prova estatística não-paramétrica a *Kruskal Wallis*, pois esta apresenta duas condições favoráveis que permitem a sua aplicação. Primeiro, a comparação entre três ou mais grupos e, segundo, a existência de amostras independentes. O teste consiste em determinar se as medianas de dois ou mais grupos diferem. Um fator importante é que os dados para todas as contagens (CCSFPI, Especialistas 1, 2 e 3) devem ter distribuições similares, o que se verifica neste trabalho, conforme mostrado nas Figuras 51-54. O teste *Kruskal Wallis* é utilizado em outros trabalhos (Nyame *et. al*, 2018, Chaloupková *et. al*, 2018, Khan & Khan, 2018, Rajaraman *et. al*, 2018).

Utiliza-se um nível de significância α de 5% (tradicionalmente usado) que se refere à probabilidade de se concluir a existência de diferença entre as medianas. Juntamente com o nível de significância se adota o *valor-p* que é uma probabilidade cuja função é medir a evidência contra a hipótese. Deve-se, portanto comparar o *valor-p* com o nível de significância a fim de avaliar a hipótese nula para determinar se alguma das diferenças entre as medianas (CCSFPI, Especialista 1, Especialista 2, Especialista 3) é estatisticamente significativa.

Se $\text{valor-p} \leq \alpha$ as diferenças entre algumas das medianas são estatisticamente significativas. Caso contrário, as medianas não são estatisticamente diferentes. Feito o teste de *Kruskal Wallis* com os dados das quatro contagens, o *valor-p* obtido foi de 0,5481, comprovando que as medianas de todas as contagens obtidas (pelo método proposto e pelos 3 especialistas) não são estatisticamente diferentes. Esta análise estatística demonstra que o método proposto foi capaz de gerar resultados consistentes com as contagens realizadas pelos especialistas.

Outra análise feita compreendeu a verificação entre as diferenças da contagem das células de cada imagem individualmente, e não da contagem geral.

Para cada imagem, um Falso Positivo (FP) refere-se à identificação de pontos, pelo método, que não são células e um Falso Negativo (FN) refere-se à células que não foram efetivamente contadas pelo método proposto. A tabela 11 apresenta os valores das respectivas contagens e a Tabela 12 apresenta os resultados de FP e FN (Esp1, Esp2, Esp3 referem-se aos especialistas 1, 2 e 3, respectivamente). O percentual de FP foi calculado com base na quantidade de células contabilizadas pelo método, ou seja, qual o percentual de células contadas pelo método que realmente não são células. Por sua vez, o percentual de FN foi calculado com base no total de células contadas pelo

especialista, ou seja, qual percentual de células contadas pelo especialista que não foram contabilizadas pelo método.

Tabela 11: Contagens células (Limiar Melo)

Contagens	Amostra total
Contagem células Esp1	1475
Contagem células Esp2	1507
Contagem células Esp3	1510
Contagem células método CCSFPI	1476

Tabela 12: Falsos positivos e falsos negativos (Limiar Melo)

Análise dos falsos positivos e falsos negativos			
FP Esp1	FN Esp1	Percentual FP Esp1	Percentual FN Esp1
207	206	$(207/1476)*100= 14,02\%$	$(206/1475)*100= 13,97\%$
FP Esp2	FN Esp2	Percentual FP Esp2	Percentual FN Esp2
197	228	$(197/1476)*100= 13,35\%$	$(228/1507)*100= 15,13\%$
FP Esp3	FN Esp3	Percentual FP Esp3	Percentual FN Esp3
204	238	$(204/1476)*100= 13,82\%$	$(238/1510)*100= 15,76\%$
FP- Falso Positivo; FN- Falso Negativo			

Pode-se observar pelos resultados obtidos, que os percentuais dos falsos positivos e os percentuais dos falsos negativos foram muitos próximos. Isso indica que a metodologia CCSFPI utilizada com o limiar de Melo mantém um equilíbrio entre os valores encontrados e mostra que a mesma não é tendenciosa.

Duas situações podem justificar a presença dos falsos positivos e falsos negativos e estão relacionadas à presença de imperfeições nas imagens. O primeiro caso se refere ao processo de geração da imagem (Figura 55) no qual a iluminação inadequada pode contribuir para a ocorrência de erros. O segundo, o fato do leite de búfalas apresentar elevada quantidade de gordura pode gerar uma dificuldade no processo de confecção de alguns esfregaços, fazendo com que o corante utilizado não tenha a aderência necessária ao núcleo da célula, deixando-o muito claro na imagem (Figura 56). Ambas imperfeições dificultam o processo de limiarização/segmentação causando problemas no processamento das imagens.

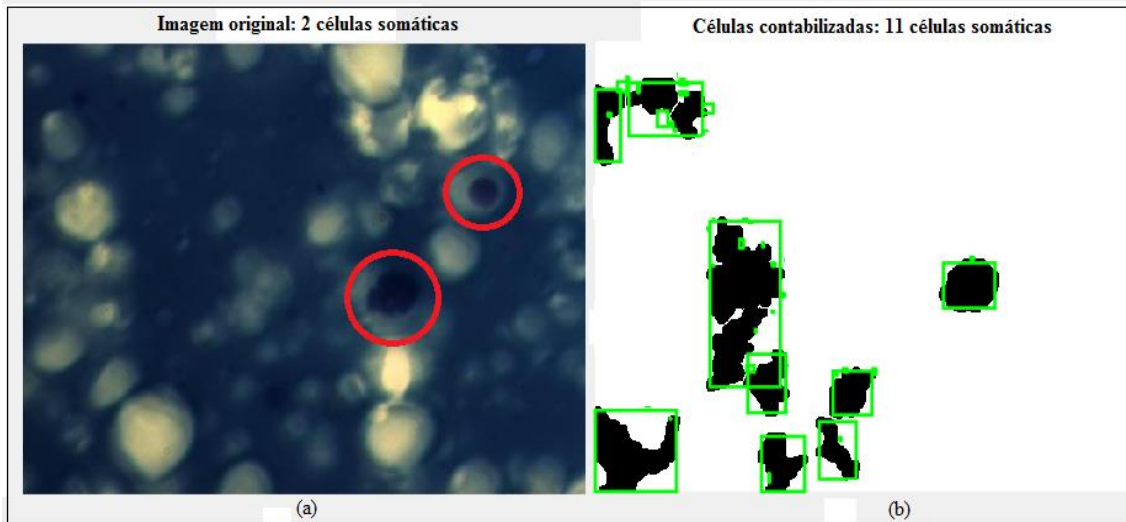


Figura 55: Imagem tirada com iluminação inadequada: (a) Imagem escura demais; (b) Contagem das células com erro.

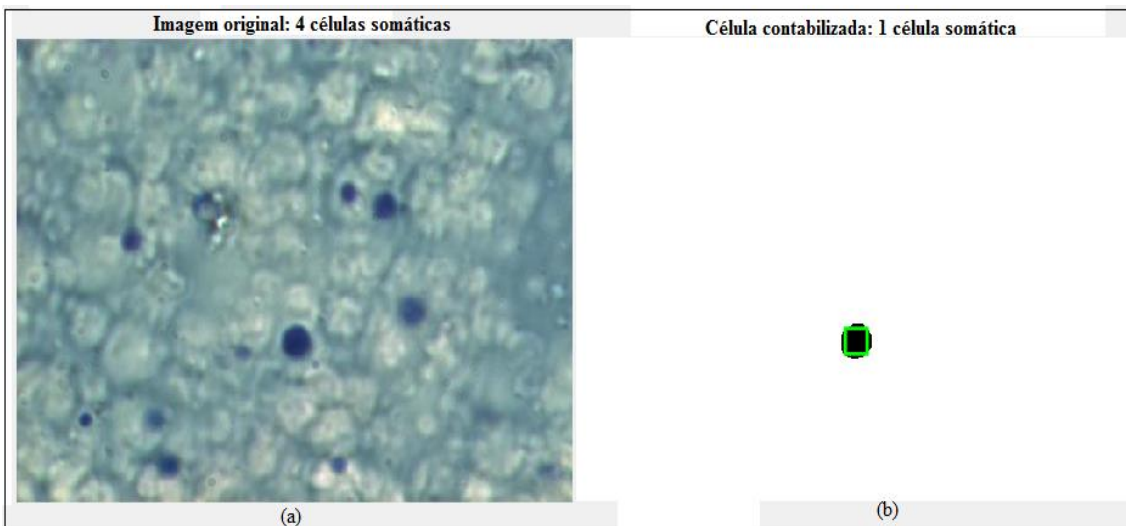


Figura 56: Amostra com excesso de gordura: (a) Células com núcleos desbotados; (b) Contagem das células com erro.

CAPÍTULO 4

CONCLUSÕES E SUGESTÕES

PARA TRABALHOS FUTUROS

4.1 CONCLUSÕES

O uso do *Fuzzy C-means* foi de grande relevância na etapa de pré-processamento, mostrando a sua importância para o sucesso na separação de imagens com tonalidades de fundos variadas e que necessitem de tratamentos diferenciados.

Os trabalhos envolvendo a aplicação de um método de agrupamento não hierárquico (tal como o FCM) no reconhecimento de imagens contemplam o uso desta técnica na etapa de processamento, mais especificamente na segmentação. O método proposto neste trabalho (CCSFPI) propõe o uso do FCM na etapa de pré-processamento que consiste na preparação das imagens coletadas (amostra disponível) cuja qualidade no resultado final da contagem possui um efeito muito significativo. Além do agrupamento, a aplicação do FCM viabiliza o reconhecimento de padrões de imagens referentes a cada um dos grupos, o que oferece um referencial de comparação e conhecimento da própria amostra.

Na contagem geral, a metodologia proposta com o método de limiarização de Melo obteve um excelente resultado, conforme foi mostrado no teste de *Kruskal Wallis*. O método de Melo permite a alteração do parâmetro de ponderação (L), podendo ajustar os valores dos limiares para atender aos diferentes grupos de acordo com a característica das imagens em cada um destes grupos.

Um fator que deve ser considerado é a forma da aquisição das imagens, com relação à intensidade de iluminação, evitando que se tenha muita variação entre elas. A iluminação incorreta pode levar a resultados não satisfatórios, pois dificulta a aplicação das etapas de processamento na imagem. Isso foi perceptível em algumas imagens aqui tratadas. Outro fator a ser observado é a elaboração da lâmina, pois o leite a ser examinado possui alto teor de gordura, podendo provocar dificuldade na confecção da lâmina, acarretando dificuldade de detecção das células.

A contagem geral do método CCSFPI utilizando a limiarização de Melo, comparado com a contagem dos especialistas mostrou ser eficiente, sendo promissor para substituir o método de contagem DMSCC.

4.2 SUGESTÕES PARA TRABALHOS FUTUROS

- Implementar um método de limiarização ainda mais eficaz ou o aperfeiçoamento de um já existente para proporcionar um melhor resultado. Considerando que a criação de um novo parâmetro (método de Melo) aperfeiçoou um método já existente (Rosin), a aplicação de um outro parâmetro associado ao histograma da imagem pode aumentar o ajuste do limiar e viabilizar a obtenção de várias medidas estatísticas a partir de um histograma tais como valores de máximo e mínimo, valor médio, variância e desvio padrão dos níveis de cinza da imagem, proporcionando melhores ajustes na imagem através destes dados.
- Realizar a coleta e fazer testes com amostras de tanques de expansão, pois esta é a alternativa mais frequente para identificação de problemas de mastite em rebanhos.
- Desenvolver um produto tecnológico com o CCSFPI embarcado, com custo e condições de utilização e transporte acessíveis para produtores de leite.

REFERÊNCIAS

BIBLIOGRÁFICAS

ACOSTA, A. C., SILVA, L. B. G. D., MEDEIROS, E. S., PINHEIRO-JÚNIOR, J. W., & MOTA, R. A. Mastitis in ruminants in Brazil. **Pesquisa Veterinária Brasileira**, v. 36, n. 7, p. 565-573, 2016.

AMARAL, F. R., CARVALHO, L. D., SILVA, N. D., & Brito, J. R. F. Qualidade do leite de búfalas: composição. **Revista Brasileira de Reprodução Animal**, v. 29, n. 2, p. 106-110, 2005.

AMARAL, Fabricio Rodrigues. **Fatores que interferem na contagem de células somáticas e constituintes do leite de búfalas**, 2005.

ANDRIGHETTO, C. Cadeia produtiva do leite de búfala: visão da universidade. **Simpósio da cadeia produtiva da bubalinocultura**, v. 2, 2011.

ANUALPEC. Anuário da pecuária brasileira. **FNP Consultoria/Agros Comunicação**, São Paulo, SP. 400p, 2009.

ARAÚJO, D. K. G.; GHELLER, V. A. Aspectos morfológicos, celulares e moleculares da imunidade da glândula mamária de búfalas (*Bubalus bubalis*): revisão de literatura. **Revta Bras. Reprod. Anim**, v. 29, p. 77-83, 2005.

ARTES, Rinaldo. **Coefficiente de Assimetria**. Insper. p. 1. Consultado em 26 de janeiro de 2014.

BABAK, V.; RYSANEK, D. Inter-laboratory trials of milk somatic cell counters: a comparison of the Fossomatic and Somacount systems. **Milchwissenschaft**, v. 54, n. 3, p. 126-128, 1999.

BAI, Jie; XUE, Heru; ZHOU, Yanqing. The milk somatic cell image segmentation method based on dimension reduction and fusion. In: International Conference on Computer and Computing Technologies in Agriculture. **Springer**, Cham, p. 580-586, 2015.

BANSAL, B. K., HAMANN, J., Lind, O., SINGH, S. T., & DHALIWAL, P. S. Somatic cell count and biochemical components of milk related to udder health in buffaloes. **Italian Journal of Animal Science**, v. 6, n. sup2, p. 1035-1038, 2007.

BARO, J. A., ROLDÁN, P., CARLEOS, C. E., GRILLO, G. J., & PÉREZ, M. A. Video microscopy as an alternative method for somatic cell count in milk. **Journal of dairy research**, v. 72, n. 1, p. 93-100, 2005.

BATISTA, Leonardo Vidal. **Introdução ao Processamento Digital de Imagens**. lucianowjr. googlepages. com/PDI2006. ppt, Acessado em: 23 de novembro de 2018, v. 8, 2005.

BENTLEY Instruments, Inc. Bentley 2000. **Operator's manual**. Minesota, 1994.

BERNARDES, O. Bubalinocultura no Brasil: situação e importância econômica. **Revista Brasileira de Reprodução Animal**, v. 31, n. 3, p. 293-298, 2007.

BEUCHER, S. The watershed transformation applied to image segmentation. **In The Pfefferkorn Conference on Signal Image Processing in Microscopy and Microanalysis**, pp. 299–314, 1991.

BEZDEK, J., KELLER, J. and KRISHNAPURAM, R. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. 1st Edn., **Kluwer Academy Publishers**, Norwell, MA., USA., ISBN: 0792385217, pp: 792, 1981.

BEZDEK, J.C., Pattern Recognition with Fuzzy Objective Function Algorithms. 1st Edn., **Plenum Press**, New York, USA, 1981.

BEZDEK, James C.; EHRLICH, Robert; FULL, William. FCM: The fuzzy c-means clustering algorithm. **Computers & Geosciences**, v. 10, n. 2-3, p. 191-203, 1984.

BIERENS, M. Stricter hygiene regulations for milk and milk products from 1994. **Lait et Nous**, v. 3, p. 22-23, 1993.

BRADLEY, Derek; ROTH, Gerhard. Adaptive thresholding using the integral image. **Journal of graphics tools**, v. 12, n. 2, p. 13-21, 2007.

BRITO, J. R. F., CALDEIRA, G. A. V., VERNEQUE, R. D. S., & BRITO, M. A. V. P. Sensibilidade e especificidade do “California Mastitis Test” como recurso diagnóstico da mastite subclínica em relação à contagem de células somáticas. **Pesquisa Veterinária Brasileira**, v. 17, n. 2, p. 49-53, 1997.

CAVERO, D., TÖLLE, K. H., BUXADÉ, C., & KRIETER, J. Mastitis detection in dairy cows by application of fuzzy logic. **Livestock Science**, v. 105, n. 1-3, p. 207-213, 2006.

CAVERO, D., TÖLLE, K. H., HENZE, C., BUXADÉ, C., & KRIETER, J. (2008). Mastitis detection in dairy cows by application of neural networks. **Livestock Science**, v. 114, n. 2-3, p. 280-286, 2008.

CAVINATO, M.V. Fundamentos de Processamento de Imagens - Projeto 01-A: Realce de imagens utilizando transformações de intensidade. **Universidade Federal do Rio Grande do Sul. Instituto de Informática**, 2009.

CECALAIT (Centre d'études et de contrôle des analyses en industrie laitière). **La lettre de CECALAIT**, n.7, 1993.

CHAIRA, Tamalika. Intuitionistic fuzzy color clustering of human cell images on different color models. **Journal of Intelligent & Fuzzy Systems**, v. 23, n. 2, 3, p. 43-51, 2012.

CHALOUPKOVÁ, V., IVANOVA, T., EKRT, O., KABUTEY, A., & HERÁK, D. Determination of particle size and distribution through image-based macroscopic analysis of the structure of biomass briquettes. **Energies**, v. 11, n. 2, p. 331, 2018.

CHRISTEN, G.L. Analysis. In: HUA, Y.E. **Dairy science and technology handbook**. New York: VCH Publishers, V.1, p.83-156,1993.

CHUANG, K. S., TZENG, H. L., CHEN, S., WU, J., & CHEN, T. J. Fuzzy c-means clustering with spatial information for image segmentation. **computerized medical imaging and graphics**, v. 30, n. 1, p. 9-15, 2006.

COELHO, T. S., FERNANDES, M. A. R., MIOT, H. A., & YORIYAZ, H. Uso do método fuzzy c-means para segmentação de imagens dermatoscópicas de lesões de pele. **Revista Brasileira de Física Médica**, 99-102, 2012.

CONCI, Aura. **Importância do histograma em análise de imagens**. Universidade Federal Fluminense - Instituto de Computação, 2015.

DAMM, M., HOLM, C., BLAABJERG, M., BRO, M. N., & SCHWARZ, D. Differential somatic cell count—A novel method for routine mastitis screening in the frame of Dairy Herd Improvement testing programs. **Journal of dairy science**, v. 100, n. 6, p. 4926-4940, 2017.

DE CARVALHO, Francisco de AT; TENÓRIO, Camilo P. Fuzzy K-means clustering algorithms for interval-valued data based on adaptive quadratic distances. **Fuzzy Sets and Systems**, v. 161, n. 23, p. 2978-2999, 2010.

DE MELO, G. J. A. *et al.* A robust segmentation method for counting bovine milk somatic cells in microscope slide images. **Computers and Electronics in Agriculture**, v. 115, p. 142-149, 2015.

DE QUEIROZ, José Eustáquio Rangel; GOMES, Herman Martins. **Introdução ao processamento digital de imagens**. RITA, v. 13, n. 2, p. 11-42, 2006.

DE VARGAS, Rogerio R.; BEDREGAL, Benjamin Rene Callejas; PALMEIRA, Eduardo Silva. A comparison between K-Means, FCM and ckMeans Algorithms.

DELLA LIBERA, A. M. M. P., ARAUJO, W. P. D., KITAMURA, S. S., ROSENFELD, A. M. F., & BIRGEL, E. H. Citologia do leite de búfalas (*Bubalus bubalis*) híginas criadas no Estado de São Paulo, Brasil. **Ciência Rural**, v. 34, n. 4, p. 1087-1092, 2004.

DHAKAL I.P. Normal somatic cell count and subclinical mastitis in Murrah buffaloes. **Buffalo J.** 20:261-270, 2004.

DUBEY, P. C., SUMAN, C. L., SANYAL, M. K., PANDEY, H. S., SAXENA, M. M., & YADAV, P. L. Factors affecting composition of milk of buffaloes. **The Indian Journal of Animal Sciences**, v. 67, n. 9, 1997.

DUDA, Richard O.; HART, Peter E. **Pattern recognition and scene analysis**. 1973.

DUNN, Joseph C. **A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters**, 1973.

DÜRR, J. W., CUE, R. I., MONARDES, H. G., MORO-MÉNDEZ, J., & WADE, K. M. (2008). Milk losses associated with somatic cell counts per breed, parity and stage of lactation in Canadian dairy cattle. **Livestock Science**, v. 117, n. 2-3, p. 225-232, 2008.

DÜRR, J. W.; CARVALHO, M. P.; SANTOS, M. V. O Compromisso com a Qualidade do Leite no Brasil. **Passo Fundo: EdiUPF**, p. 38-55, 2004.

EMBRAPA. **Anuário Leite 2018**: Indicadores, tendências e oportunidades para quem vive no setor leiteiro, 2018.

FAILACE, Rafael; PRANKE, Patricia Helena Lucas. Avaliação dos critérios de liberação direta dos resultados de hemogramas através de contadores eletrônicos. **Revista brasileira de hematologia e hemoterapia= Brazilian journal of hematology and hemotherapy**. São Paulo. Vol. 26, n. 3 (2004), p. 159-166, 2004.

FENG, Mei-fang; XUE, He-ru; YAN, Ke-yu. Application of Fuzzy c-Means Algorithm in the Color Images Segmentation of Milk Somatic Cells. **Journal of Inner Mongolia University**, v. 3, 2009.

FONSECA, L. F. L.; SANTOS, M. V. Qualidade do leite e controle de mastite. São Paulo, Lemos Editorial. 175p. (p. 27 - 50), 2000.

FONSECA, Luís Fernando Laranja da; SANTOS, Marcos Veiga dos. Qualidade do leite e controle de mastite. 2001.

GAO, X., XUE, H., PAN, X., JIANG, X., BO, Y., & WANG, Y. Segmentation of somatic cells based on cloud model. **Rev. Téc. Ing. Univ. Zulia**, v. 39, n. 2, p. 93-101, 2016.

GAO, X., XUE, H., PAN, X., JIANG, X., ZHOU, Y., & LUO, X. Somatic Cells Recognition by Application of Gabor Feature-Based (2D) 2PCA. **International Journal of Pattern Recognition and Artificial Intelligence**, v. 31, n. 12, p. 1757009, 2017.

GARCIA-CORDERO, J. L., BARRETT, L. M., O'KENNEDY, R., & RICCO, A. J. Microfluidic sedimentation cytometer for milk quality and bovine mastitis monitoring. **Biomedical microdevices**, v. 12, n. 6, p. 1051-1059, 2010.

GERLACH, John. **Fotografia Digital de Paisagens**. Elsevier Brasil, 2013.

GEVERS, Theo; SMEULDERS, Arnold WM. Color-based object recognition. **Pattern recognition**, v. 32, n. 3, p. 453-464, 1999.

GONZALEZ, R. AND WOODS, R. **Processamento de Imagens Digitais**. Editora Edgard Blucher, São Paulo, 2000.

GONZALEZ, R. C. and WOODS, R. E., **Digital Image Processing**, Addison Wesley, 1993.

GONZALEZ, R. C.; WOODS, R. E. Digital Image Processing, for. **Computer Science and Engineering Series**. Prentice Hall, v. 14, n. 3, 2002.

GRACIANO, Ana Beatriz V., **Estudo de técnicas para suavização de bordas detectadas na segmentação de imagens em sequências de vídeo**. Relatório final do projeto de iniciação científica - PIBIC-CNPq. Departamento de Ciência da Computação –IME-USP, 2002.

GRILLO, G. J., ANTON, J. C., FERRERO, F., & PÉREZ, M. A. Low-cost video-microscopy system for evaluation of fresh cow milk somatic cell concentration. In: **Instrumentation and Measurement Technology Conference, 2001. IMTC 2001. Proceedings of the 18th IEEE**. IEEE, p. 906-909, 2001.

HARTIGAN, John A.; WONG, Manchek A. Algorithm AS 136: A k-means clustering algorithm. **Journal of the Royal Statistical Society. Series C (Applied Statistics)**, v. 28, n. 1, p. 100-108, 1979.

HULSEN, Jan; LAM, Theo. **Udder health: a practical guide to first-rate udder health**. Roodbont, 2007.

IBGE; Diretoria de Pesquisas, Coordenação e Agropecuária, Pesquisa da Pecuária Municipal 2006. Acesso em: 30 de set. 2018.

INSTRUMENTS, BENTLEY. Inc. Bentley 2000. **Operator's manual**. 1994.

INTELACTUS. Indicadores: Leite e Derivados. – Ano 9, n. 78 – Juiz de Fora: **Embrapa Gado de Leite**, 2018.

- JINGAR, S. C., MEHLA, R. K., SINGH, M., KUMAR, A., KANTWA, S. C., & SINGH, N. Comparative study on the incidence of mastitis during different parities in cows and buffaloes. **Indian J. Anim. Res.**, v. 48, n. 2, p. 194-197, 2014.
- JORGE, A. M., de GUSMAO Couto, A., CRUDELI, G. A., & PATIÑO, E. M. **Produção de búfalas de leite**. Fepaf (Fundação de Estudos e Pesquisas Agrícolas e Florestais), 2011.
- KAPRONEZAI, J.; MELVILLE, P.; BENITES, N. R. Análise microbiológica, teste de Tamis e California Mastitis Test realizados em amostras de leite de fêmeas bubalinas pertencentes a rebanhos do Estado de São Paulo. **Arq. Inst. Biol.**, v. 72, p. 179-83, 2005.
- KAREEM, Saumya; MORLING, Richard CS; KALE, Izzet. A novel method to count the red blood cells in thin blood films. In: **Circuits and Systems (ISCAS), 2011 IEEE International Symposium on**. IEEE. p. 1021-1024, 2011.
- KAUR, M., VERMA, R., BANSAL, B. K., MUKHOPADHYAY, C. S., & ARORA, J. S. Status of sub-clinical mastitis and associated risk factors in Indian water buffalo in Doaba region of Punjab, India. **Indian J. Dairy Sci**, 68, 483-487, 2015.
- KAVITHA, K. L., RAJESH, K., SURESH, K., SATHEESH, K., & SUNDAR, N. S. Buffalo mastitis-risk factors. **Buffalo Bull**, v. 28, n. 3, p. 134-137, 2009.
- KEE, C. N. R. D. Bovine mastitis: an Asian perspective. **Asian Journal of Animal and Veterinary Advances**, v. 7, n. 6, p. 454-476, 2012.
- KEEFE, Greg. Update on control of Staphylococcus aureus and Streptococcus agalactiae for management of mastitis. **Veterinary Clinics: Food Animal Practice**, v. 28, n. 2, p. 203-216, 2012.
- KHAN, Mohammad Farhan; KHAN, Mohammad Arsalan. Information preserving histogram segmentation of low contrast images using fuzzy measures. **Optik**, v. 157, p. 1397-1404, 2018.
- KIM, Yeong-Taeg. Contrast enhancement using brightness preserving bi-histogram equalization. **IEEE transactions on Consumer Electronics**, v. 43, n. 1, p. 1-8, 1997.
- KORBES, André. Análise de algoritmos da transformada watershed. 2010.
- KUMAR, Tarun; VERMA, Karun. A Theory Based on Conversion of RGB image to Gray image. **International Journal of Computer Applications**, v. 7, n. 2, p. 7-10, 2010.
- LANGONI, Hélio. Tendências de modernização do setor lácteo: monitoramento da qualidade do leite pela contagem de células somáticas. **Revista de Educação Continuada em Medicina Veterinária e Zootecnia do CRMV-SP**, v. 3, n. 3, p. 57-64, 2000.

- LI, Deyi; LIU, Changyu; GAN, Wenyan. A new cognitive model: Cloud model. **International Journal of Intelligent Systems**, v. 24, n. 3, p. 357-375, 2009.
- LI, Q., WANG, Y., LIU, H., HE, X., XU, D., WANG, J., & GUO, F. Leukocyte cells identification and quantitative morphometry based on molecular hyperspectral imaging technology. **Computerized Medical Imaging and Graphics**, v. 38, n. 3, p. 171-178, 2014.
- LIAO, T. Warren. Clustering of time series data—a survey. **Pattern recognition**, v. 38, n. 11, p. 1857-1874, 2005.
- LIMA, Danielle Dantas. Citometria de fluxo. Universidade Federal da Bahia – Instituto de Ciências da Saúde (ICS), 2010.
- LOPES, M. A., DEMEU, F. A., DA ROCHA, C. M. B. M., DA COSTA, G. M., FRANCO NETO, A., & SANTOS, G. D. Avaliação do impacto econômico da mastite em rebanhos bovinos leiteiros. **Arquivos do Instituto Biológico**, v. 79, n. 4, p. 477-483, 2012.
- MACEDO, M. P., WECHSLER, F. S., RAMOS, A. D. A., AMARAL, J. B. D., SOUZA, J. C. D., RESENDE, F. D. D., & OLIVEIRA, J. V. D. Composição físico-química e produção do leite de búfalas da raça Mediterrâneo no oeste do Estado de São Paulo. **Revista Brasileira de Zootecnia**, 1084-1088, 2001.
- MANDAL, Rupesh; CHOUDHURY, Nupur; GOSWAMI, Baishali. Histological Image Segmentation using Fuzzy C-Means. **International Journal of Computer Applications**, v. 975, p. 8887, 2016.
- MANSOUR, Mohamed Mohsen; HENDAWY, Amin O.; ZEITOUN, Moustafa M. Effect of mastitis on luteal function and pregnancy rates in buffaloes. **Theriogenology**, v. 86, n. 5, p. 1189-1194, 2016.
- MARSHALL, Robert T. Standard methods for the examination of dairy products. 1992.
- MATTOS, B. C. 2007. Aspectos qualitativos do leite bubalino. **PUBVET**, v. 1, n. 9, 2007.
- MEENA, Sunita; RAJPUT, Y. S.; SHARMA, Rajan. Comparative fat digestibility of goat, camel, cow and buffalo milk. **International Dairy Journal**, v. 35, n. 2, p. 153-156, 2014.
- MESQUITA, A. A., da COSTA, G. M., de ALMEIDA, F. V. M., DEMEU, F. A., & REIS, E. M. B. Mastite em rebanhos bubalinos e sua suscetibilidade a antimicrobianos. **PUBVET**, 11, 1-102, 2016.
- MONDAL, P. K., PRODHAN, U. K., Al MAMUN, M. S., RAHIM, M. A., & HOSSAIN, K. K. Segmentation of white blood cells using fuzzy C means segmentation algorithm. **IOSR Journal of Computer Engineering**, v. 1, n. 16, p. 1-5, 2014.

MOON, J. S., KOO, H. C., JOO, Y. S., JEON, S. H., HUR, D. S., CHUNG, C. I., ... & PARK, Y. H. Application of a new portable microscopic somatic cell counter with disposable plastic chip for milk analysis. **Journal of dairy science**, v. 90, n. 5, p 2253-2259, 2007.

MORONI, P., Rossi, C. S., PISONI, G., BRONZO, V., CASTIGLIONI, B., & BOETTCHER, P. J. Relationships between somatic cell count and intramammary infection in buffaloes. **Journal of dairy science**, v. 89, n. 3, p. 998-1003, 2006.

NAZLIBILEK, S., KARACOR, D., ERCAN, T., SAZLI, M. H., KALENDER, O., & EGE, Y. Automatic segmentation, counting, size determination and classification of white blood cells. **Measurement**, 55, 58-65, 2014.

NG, Hui-Fuang. Automatic thresholding for defect detection. **Pattern recognition letters**, v. 27, n. 14, p. 1644-1649, 2006.

NG, Jessie; ROHLING, Robert; LAWRENCE, Peter D. Automatic measurement of human subcutaneous fat with ultrasound. **IEEE transactions on ultrasonics, ferroelectrics, and frequency control**, v. 56, n. 8, p. 1642-1653, 2009.

NYAME, Y. A., GRIMBERG, D. C., GREENE, D. J., GUPTA, K., KARTHA, G. K., BERGLUND, R., ... & KLEIN, E. A. Genomic scores are independent of disease volume in men with favorable risk prostate cancer: implications for choosing men for active surveillance. **The Journal of urology**, v. 199, n. 2, p. 438-444, 2018.

OLIVEIRA, Carlos Augusto Fernandes de; LARANJA-DA-FONSECA, Luís Fernando; GERMANO, Pedro Manuel Leal. Aspectos relacionados à produção, que influenciam a qualidade do leite. **Higiene Alimentar**, v. 13, n. 62, p. 10-16, 1999.

ONGUN, G., HALICI, U., LEBLEBICIOGLU, K., ATALAY, V., BEKSAÇ, M., & BEKSAÇ, S. Feature extraction and classification of blood cells for an automated differential blood count system. In: **Neural Networks, 2001. Proceedings. IJCNN'01. International Joint Conference on. IEEE**. p. 2461-2466, 2001.

OTSU, Nobuyuki. A threshold selection method from gray-level histograms. **IEEE transactions on systems, man, and cybernetics**, v. 9, n. 1, p. 62-66, 1979.

PARKER, Jim R. **Algorithms for image processing and computer vision**. John Wiley & Sons, 2010.

PETROVA, N.; TZANKOVA, M. SCC of milk from three breeds of buffaloes in Shoumen region. **Bulgarian Journal of Agricultural Science (Bulgaria)**, 1999.

PHILPOT, W. Nelson; NICKERSON, Stephen C. Mastitis: counter attack. 1991.

PICCININI, R., MIARELLI, M., FERRI, B., TRIPALDI, C., BELOTTI, M., DAPRA, V., ... & ZECCONI, A. Relationship between cellular and whey components in buffalo milk. **Journal of Dairy Research**, v. 73, n. 2, p. 129-133, 2006.

PINTO, Tiago Willian *et al.* Segmentação de imagens digitais combinando watershed e corte normalizado em grafos. 2014.

PITACAS, F.I; RODRIGUES, A.M. Agrioforum: Produção de búfalos de leite na Europa Buffalo milk raising in Europe, n.º 31, Ano 21, 2013.

PRESCOTT, S. C.; BREED, R. S. The determination of the number of the body cells in milk by a direct method. **Journal of Infectious Diseases**, v. 7, 632-640, 1910.

RAJARAMAN, S., ANTANI, S. K., POOSTCHI, M., SILAMUT, K., HOSSAIN, M. A., MAUDE, R. J., ... & THOMA, G. R. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. **PeerJ**, 6, e4568, 2018.

RAMARAJ, M., NIRAIMATHI, S., Application of color based image segmentation paradigm on rgb color pixels using fuzzy c-means and k means algorithms. **International Journal of Computer Science and Mobile Computing**, Vol.6 Issue.6, pg. 430-440, 2017.

RIBAS, N. P. Valor e uso da contagem de células somáticas no manejo de rebanhos leiteiros. In: **Congresso Holstein de las Americas**. p. 47-61, 2001.

RICCI, Gisele Dela; DOMINGUES, Paulo Francisco. O leite de búfala. **Revista de Educação Continuada em Medicina Veterinária e Zootecnia do CRMV-SP**, v. 10, n. 1, p. 14-19, 2012.

ROCHA, João Carlos. Cor luz, cor pigmento e os sistemas RGB e CMY. **Revista Belas Artes**, 2011.

ROERDINK, Jos BTM; MEIJSTER, Arnold. The watershed transform: Definitions, algorithms and parallelization strategies. **Fundamenta informaticae**, v. 41, n. 1, 2, p. 187-228, 2000.

ROSA, Ekuikui Vanilson dos Anjos. Agrupamento não supervisionado fuzzy C-Means. Trabalho de Conclusão de Curso. Universidade Tecnológica Federal do Paraná, 2017.

ROSENFELD, G. Corante pancrômico para hematologia e citologia clínica. Nova combinação dos componentes de May-Grunwald e do Giemsa num só corante de emprego rápido. **Memórias do Instituto Butantan**, v. 20, p. 329-335, 1947.

ROSIN, Paul L. Unimodal thresholding. **Pattern recognition**, v. 34, n. 11, p. 2083-2096, 2001.

SARRAFZADEH, Omid; DEHNAVI, Alireza Mehri. Nucleus and cytoplasm segmentation in microscopic images using K-means clustering and region growing. **Advanced biomedical research**, v. 4, 2015.

SAVKARE, S. S.; NAROTE, S. P. Blood cell segmentation from microscopic blood images. In: **2015 International Conference on Information Processing (ICIP)**. IEEE, p. 502-505, 2015.

SCHALM O. W. & NOORLANDER D. D.. Experiments and observations leading to development of the California Mastitis Test. **J. Am. Vet. Med. Assoc.** 130: 199-204, 1957.

SCHALM, Oscar William *et al.* **Bovine mastitis**. Philadelphia, USA: Lea & Febiger., 1971.

SEAB – Secretaria de Estado da Agricultura e do Abastecimento. DERAL - Departamento de Economia Rural. Bubalinocultura, 2017.

SELINUMMI, J., SEPPÄLÄ, J., YLI-HARJA, O., & PUHAKKA, J. A. Software for quantification of labeled bacteria from digital microscope images by automated image analysis. **Biotechniques**, v. 39, n. 6, p. 859-863, 2005.

SERFLING, Robert J. Approximation theorems of mathematical statistics. John Wiley & Sons, 2009.

SHARMA, N.; SINGH, N. K.; BHADWAL, M. S. Relationship of somatic cell count and mastitis: An overview. **Asian-Australasian Journal of Animal Sciences**, v. 24, n. 3, p. 429-438, 2011.

SIEGEL, Sidney; CASTELLAN JR, N. John. **Estatística não-paramétrica para ciências do comportamento**. Artmed Editora, 1975.

SILVA, M. S. T., JÚNIOR, J. B. L., & GONÇALVES, I. A.. Programa de incentivo a criação de búfalos por pequenos produtores–PRONAF. **Belém, PA: CPATU**, 2003.

SINDHU, J. S., & ARORA, S. Buffalo milk. Encyclopedia of Dairy Sciences (2nd ed.), p. 503–511, 2011.

SINGH, M; LUDRI, R. S. Somatic Cell Counts in Marrah buffaloes (*Bubalus bubalis*) during different stages of lactation, parity and season. **Asian Australasian Journal of Animal Sciences**, v. 14, n. 2, p. 189-192, 2000.

SINGH, R. S., BANSAL, B. K., RANDHAWA, S. S., & MAVI, P. S. Effect of lactation therapy on quarter infection and milk composition in specific mastitis of buffaloes. **Indian Journal of Veterinary Medicine**, v. 24, n. 1, p. 16-18, 2004.

SINGH, S. P.; SINGH, R. S. Influence of somatic cell count on the physicochemical properties of buffalo milk. 1981.

SINHA, Neelam; RAMAKRISHNAN, A. G. Automation of differential blood count. In: **TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region**. IEEE, p. 547-551, 2003.

SOCIEDADE NACIONAL DE AGRICULTURA; Mercado de carne, leite e queijo de bubalinos pode atrair novos criadores, 2013. Disponível em: <https://www.sna.agr.br/mercado-diferenciado-pode-atrair-interesse-para-criacao-de-bufalos/>. Acesso dia: 30 de set de 2018.

SOMMERHAUSER, J.; KLOPPERT, B.; WOLTER, W.; ZSCHOCK, M.; SOBIRAJ, A.; FAILING K. The epidemiology of Staphylococcus aureus infections from subclinical mastitis in dairy cows during a control programme. **Veterinary Microbiology**, v. 96, p. 91-102, 2003.

SONKA, Milan; HLAVAC, Vaclav; BOYLE, Roger. **Image processing, analysis, and machine vision**. Cengage Learning, 2014.

SOUZA, Taciana; CORREIA, Suzete. Estudo de técnicas de realce de imagens digitais e suas aplicações. **João Pessoa. Paraíba**, 2007.

THEERA-UMPON, Nipon. White blood cell segmentation and classification in microscopic bone marrow images. In: **International Conference on Fuzzy Systems and Knowledge Discovery**. Springer, Berlin, Heidelberg, p. 787-796, 2005.

TOROK, Leonardo; CONCI, Aura. Método de Otsu. **Instituto de Computação (UFF)**, 2016.

TRIPALDI, C., TERRAMOCCIA, S., BARTOCCI, S., ANGELUCCI, M., & DANESE, V. The effects of the somatic cell count on yield, composition and coagulating properties of Mediterranean buffalo milk. **Asian-australasian journal of animal sciences**, v. 16, n. 5, p. 738-742, 2003.

TUTEJA, F. C., KAPUR, M. P., SHARMA, A., & MANUJA, B. Mastitis pathogens from apparently healthy buffaloes and their relationship to somatic cell count of milk. **Indian Journal of Comparative Microbiology Immunology and Infectious diseases**, v. 22, n. 2, p. 162-163, 2001.

UPPAL, S. K., SINGH, K. B., ROY, K. S., NURIYAL, D. S., & BANSAL, B. K. Natural defence mechanism against mastitis: A comparative histomorphology of buffalo teat canal. **Buffalo Journal**, v. 2, p. 125-131, 1994.

VIANNI, M. C. E. Eficiência do California Mastitis Test (CMT) na estimativa do número de células somáticas do leite bovino. *Ciência Veterinária Jaboticabal (Brasil)* v. 3, n. 2, p. 5-6, 1989.

VIEIRA, N. H.; MARQUES, F. O. Processamento digital de imagens, 1999.

VIGUIER, C., ARORA, S., GILMARTIN, N., WELBECK, K., & O'KENNEDY, R. Mastitis detection: current trends and future perspectives. **Trends in biotechnology**, v. 27, n. 8, p. 486-493, 2009.

VIVEK S., Anshu S., RAVINDER S. & ASHOK K. Comparison of various indirect tests for the detection of subclinical mastitis. **Buffalo J.** v. 18, p. 267-271, 2002.

WANG, Yuedong; XUE, Heru. The color image processing technology of the milk somatic cells. **2nd International Conference on Computer Engineering and Technology.** IEEE, p. V4-280-V4-283, 2010.

WILK, Martin B.; GNANADESIKAN, Ram. Probability plotting methods for the analysis for the analysis of data. **Biometrika**, v. 55, n. 1, p. 1-17, 1968.

XIONG H, LiZ. Clustering validation measures. In: Agarrawal CK, Reddy CC, editors. Data clustering – algorithms and applications. New York: CRC Press, Taylor & Francis Group; p.571–605, 2013.

XU, Hongming; LU, Cheng; MANDAL, Mrinal. An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm. **IEEE journal of biomedical and health informatics**, v. 18, n. 5, p. 1729-1741, 2014.

XU, X., XU, S., JIN, L., & SONG, E. Characteristic analysis of Otsu threshold and its applications. **Pattern recognition letters**, v. 32, n. 7, p. 956-961, 2011.

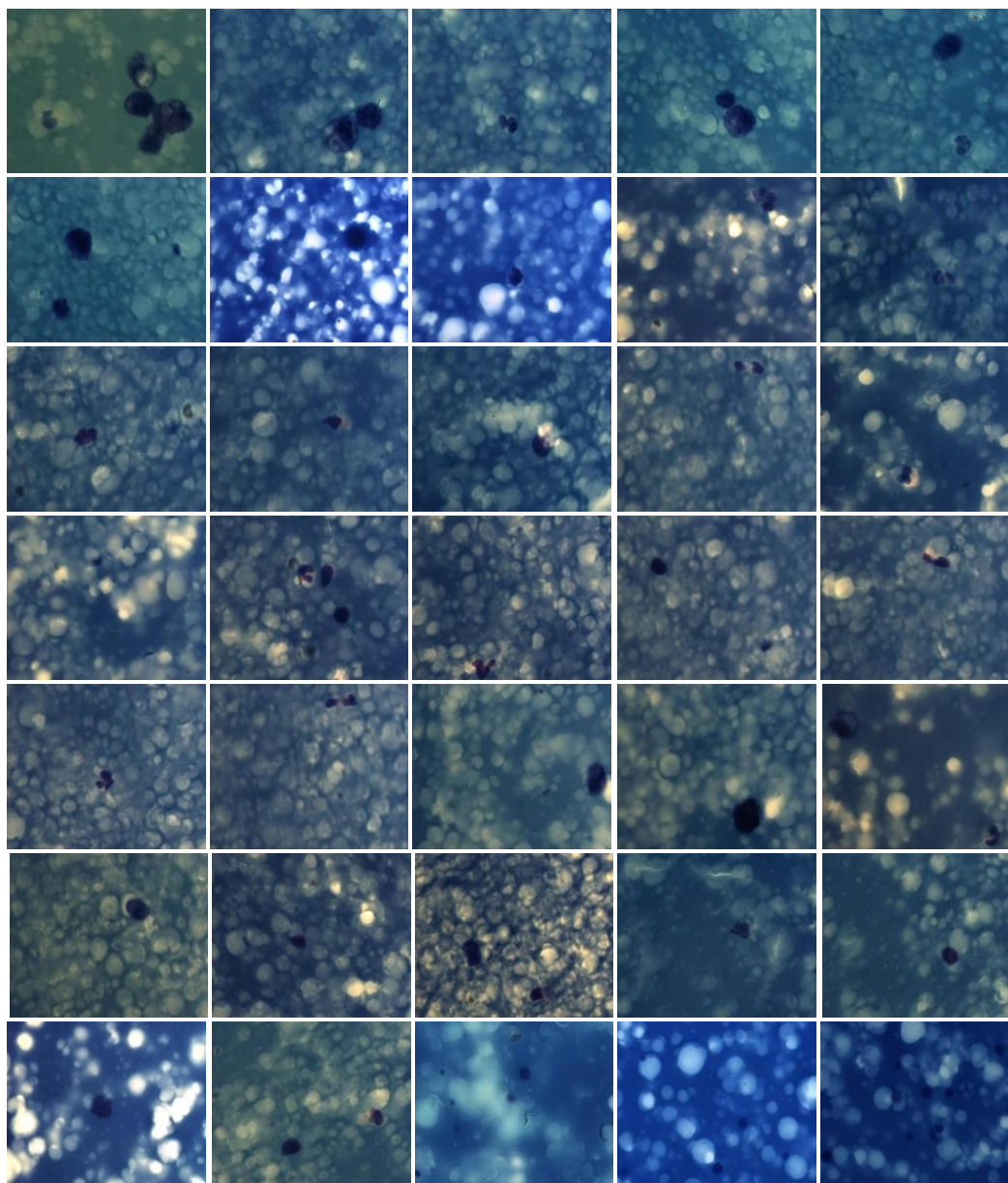
XUE, H., LI, H., WANG, Y., & ZHAO, T. The segmentation of the color milk somatic cells images. In 2009 2nd **International Congress on Image and Signal Processing** (pp. 1-4). IEEE, 2009.

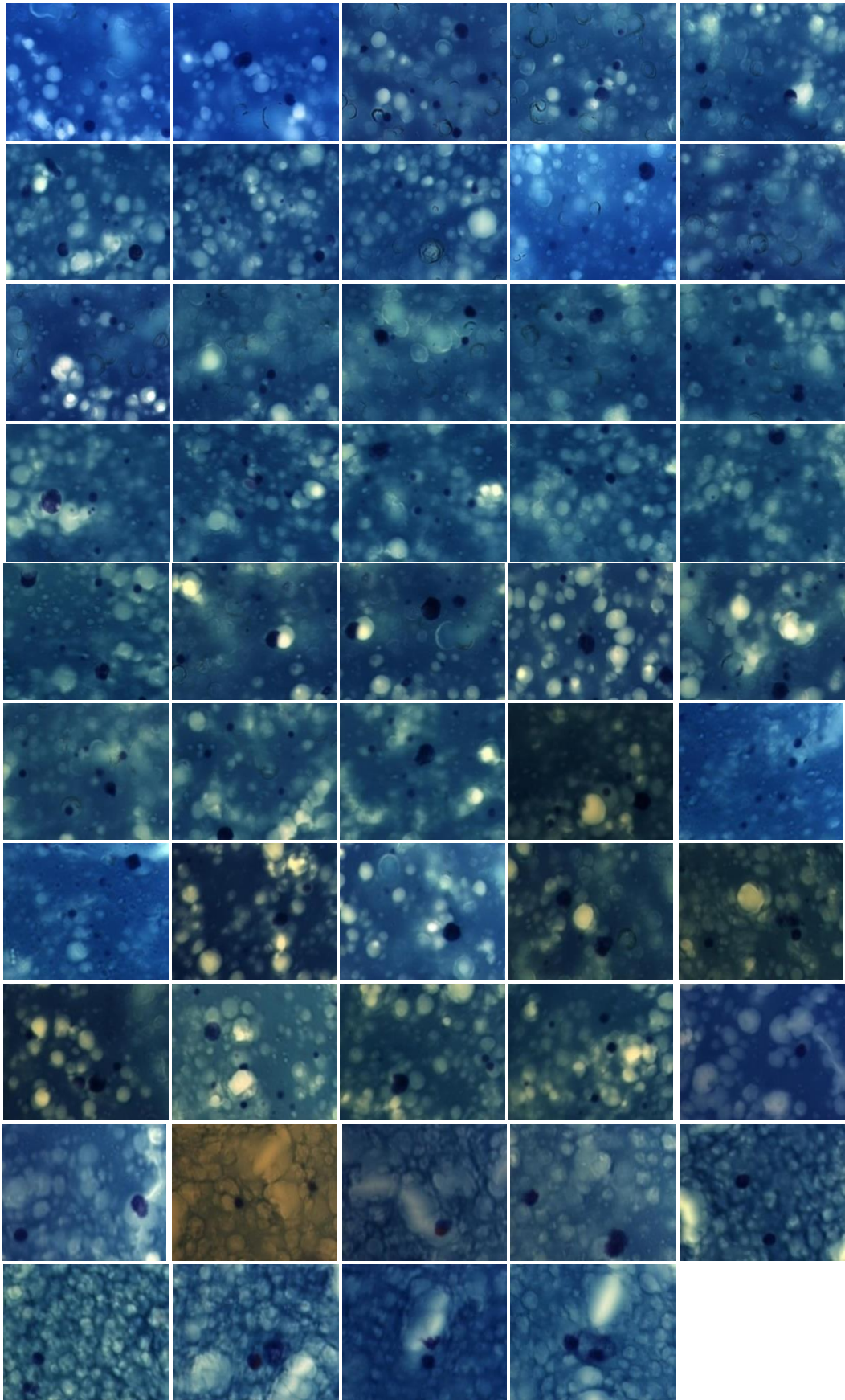
XUE, Heru; MA, Shuoshi; PEI, Xichun. An efficient segmentation method for milk somatic cell images. In: **International Conference on Computer and Computing Technologies in Agriculture.** Springer, Boston, MA, p. 1253-1258, 2007.

APÊNDICE A

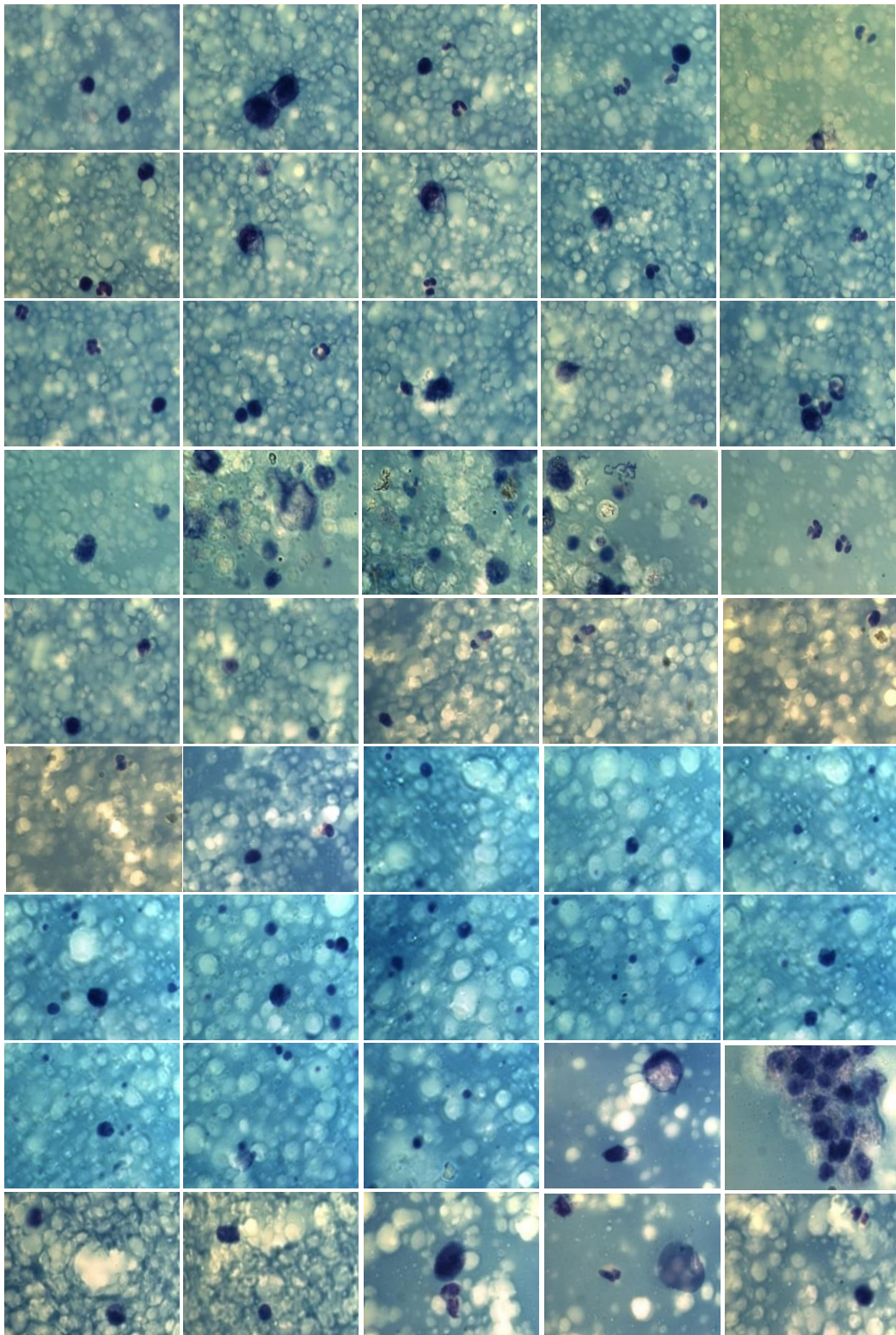
Imagens dos Grupos

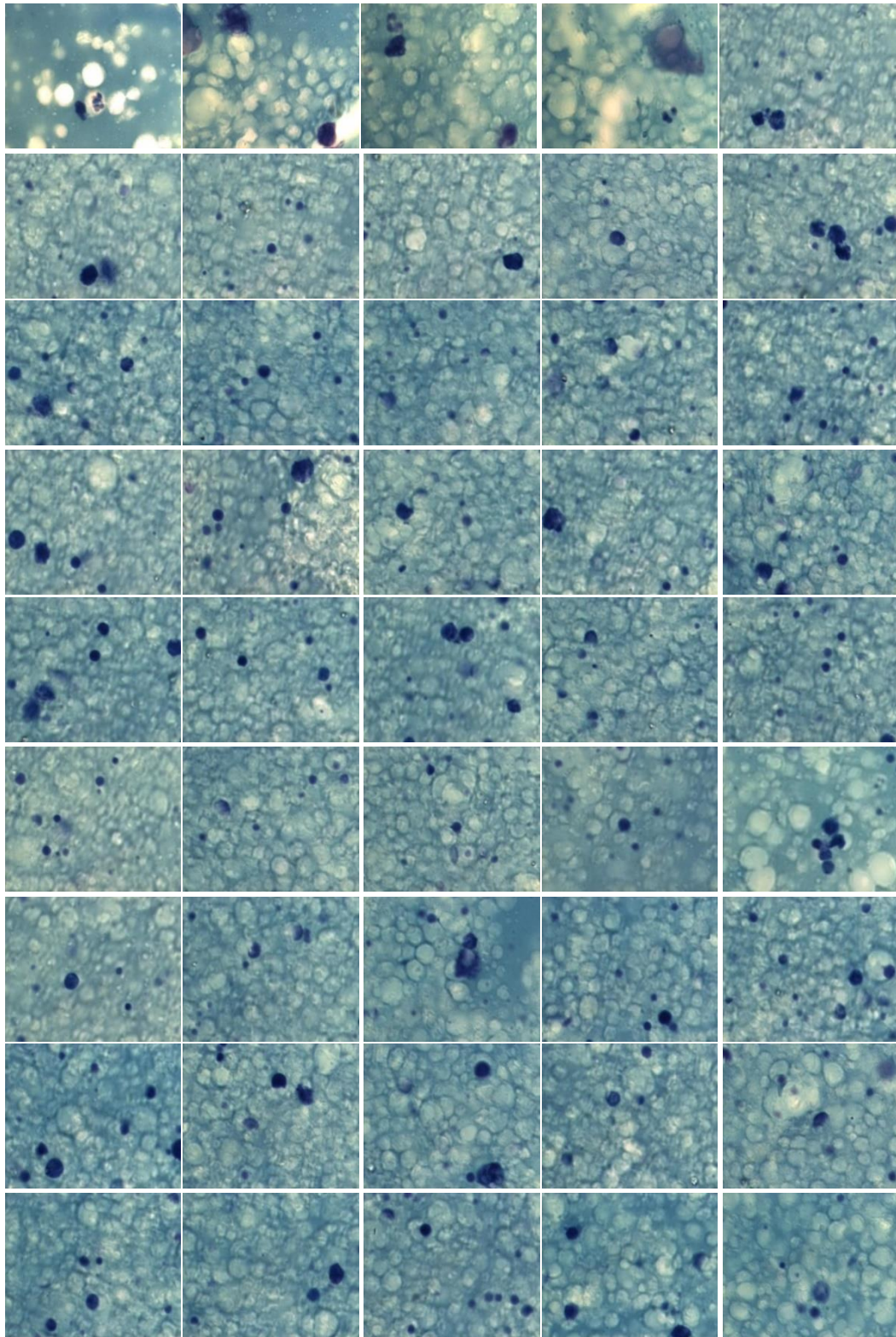
Imagens do Grupo 1

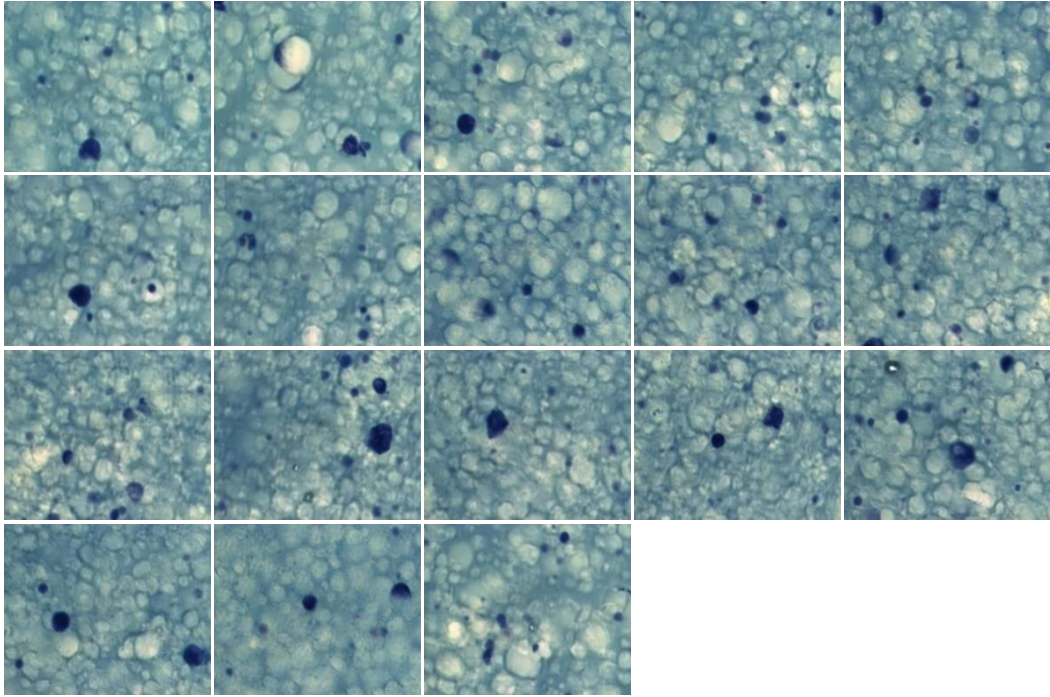




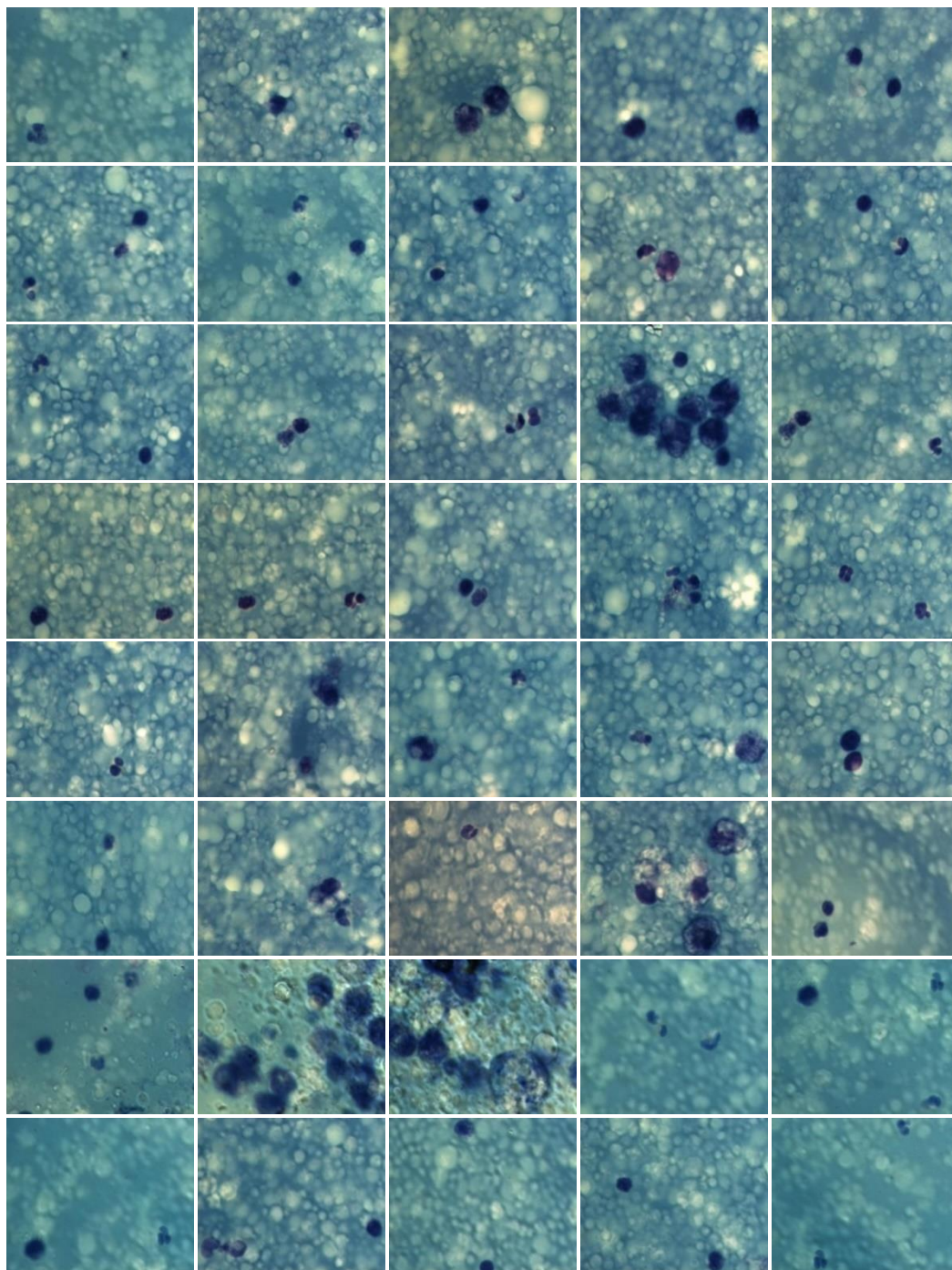
Imagens do Grupo 2

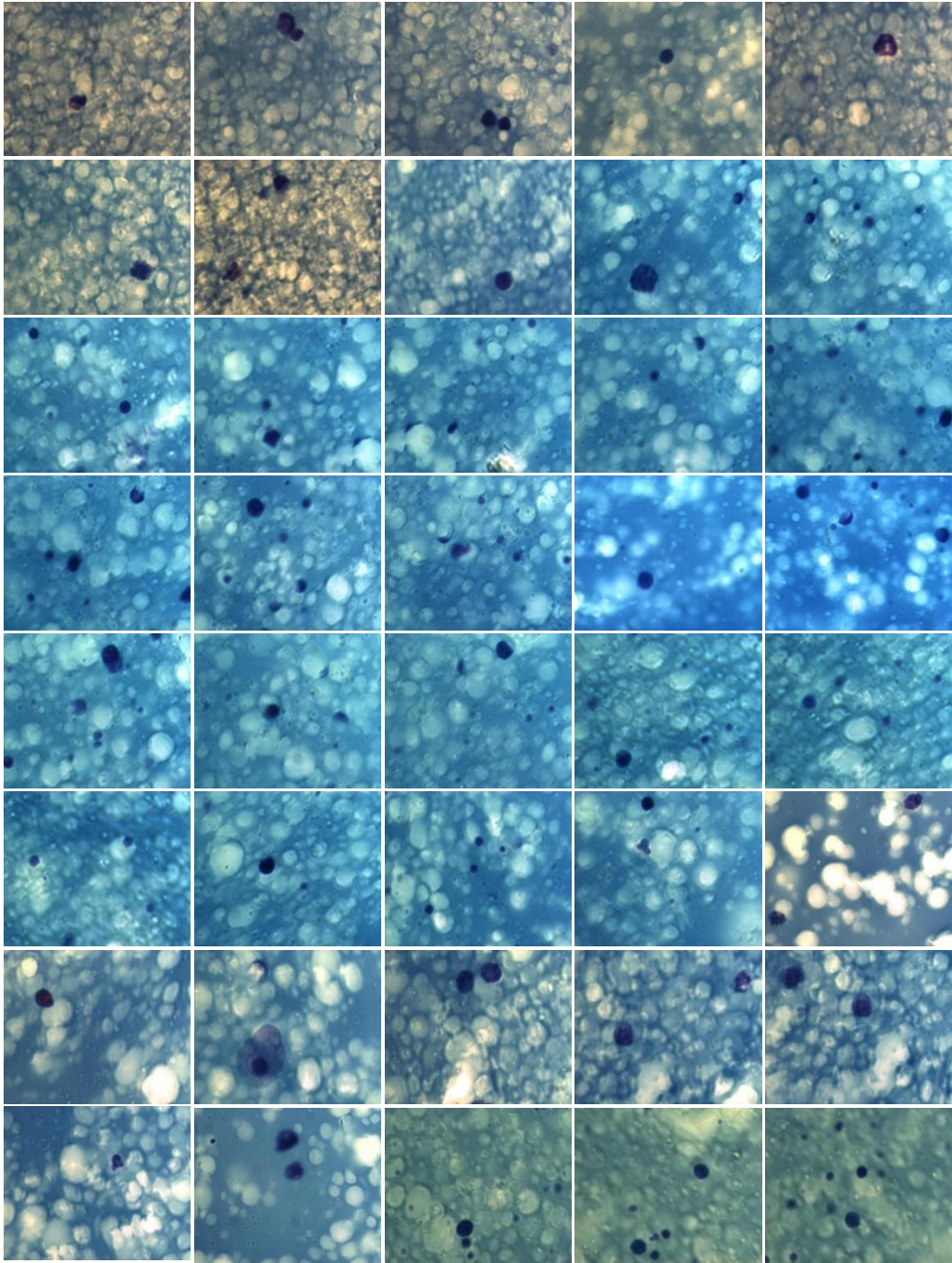


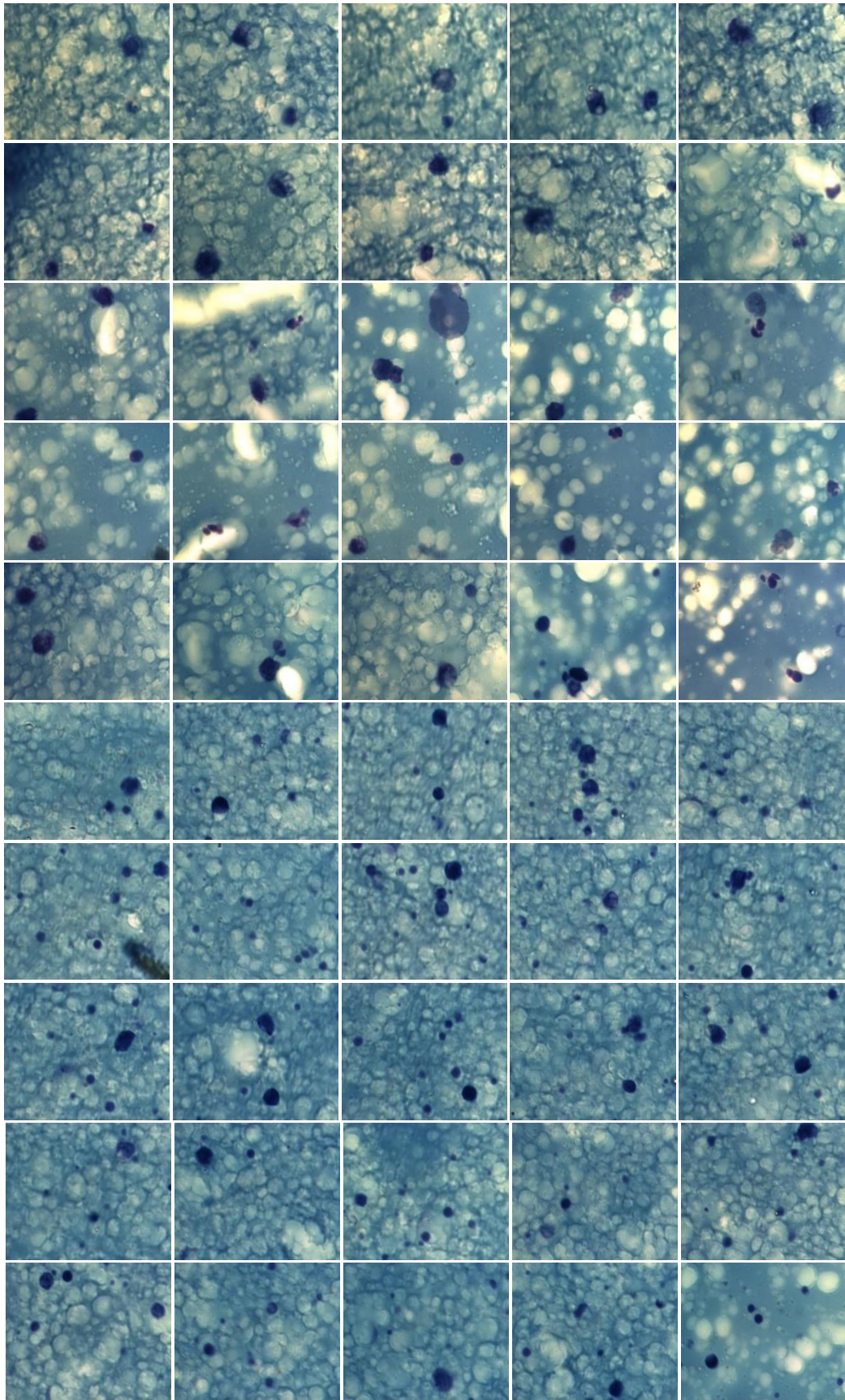




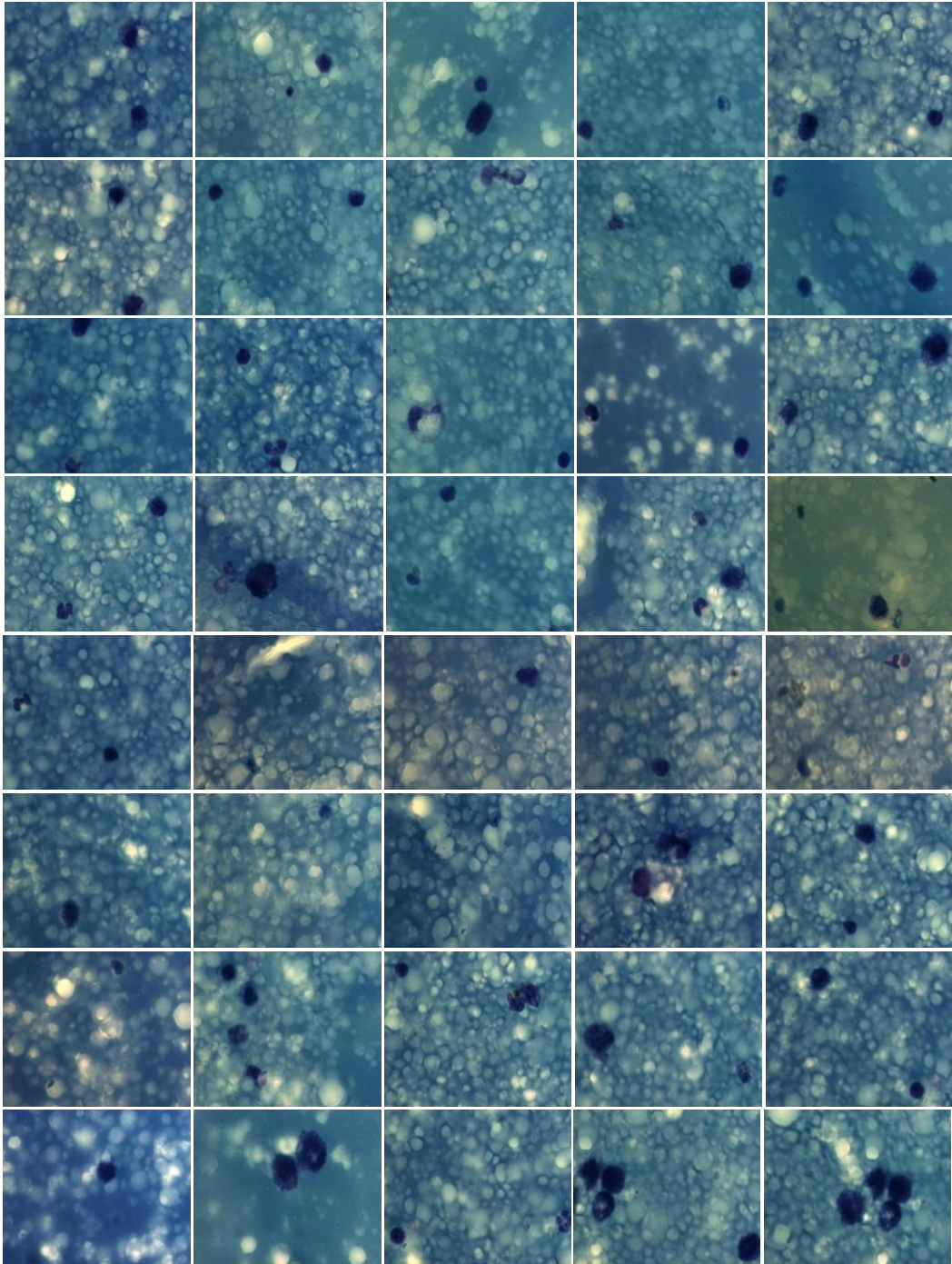
Imagens do Grupo 3

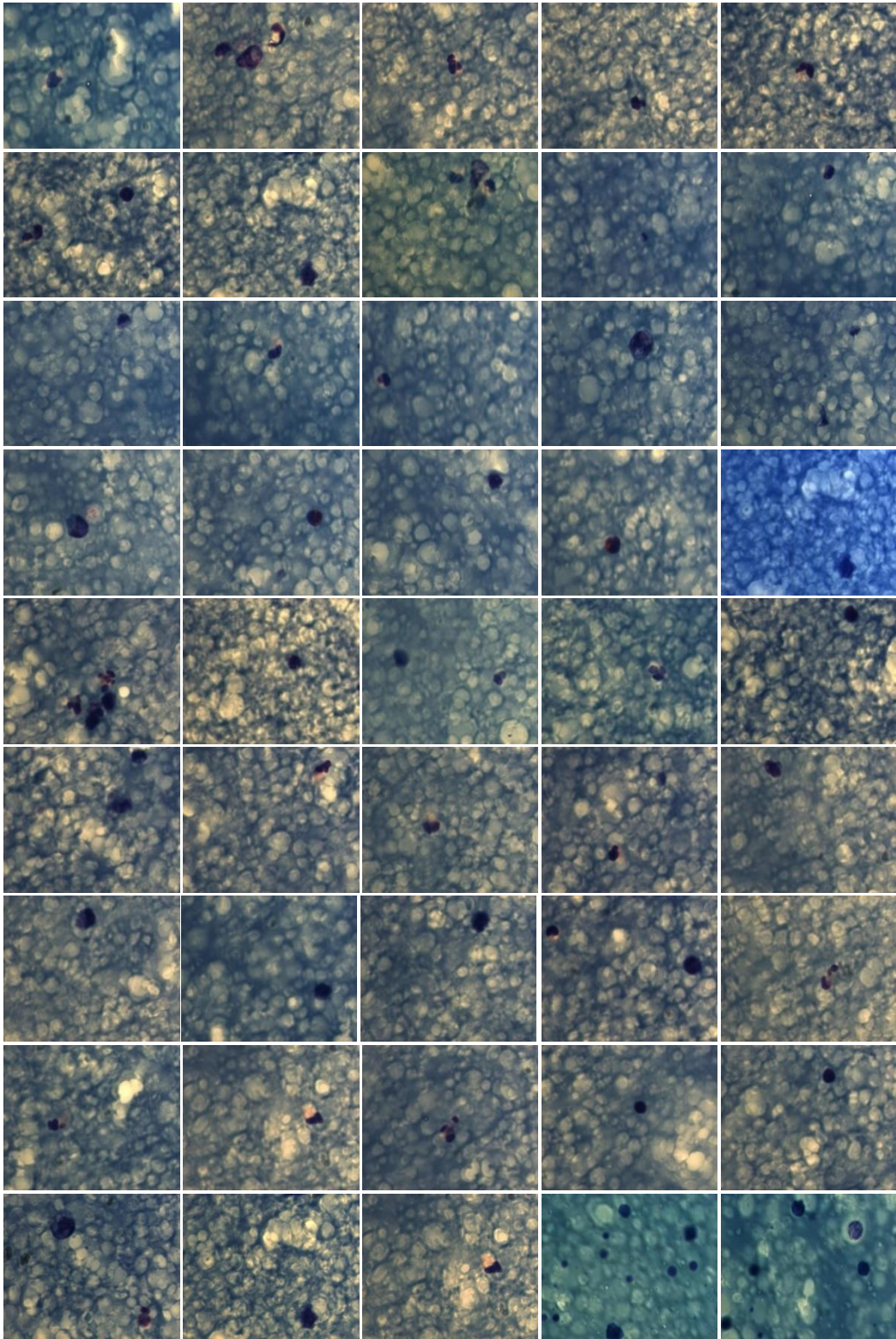


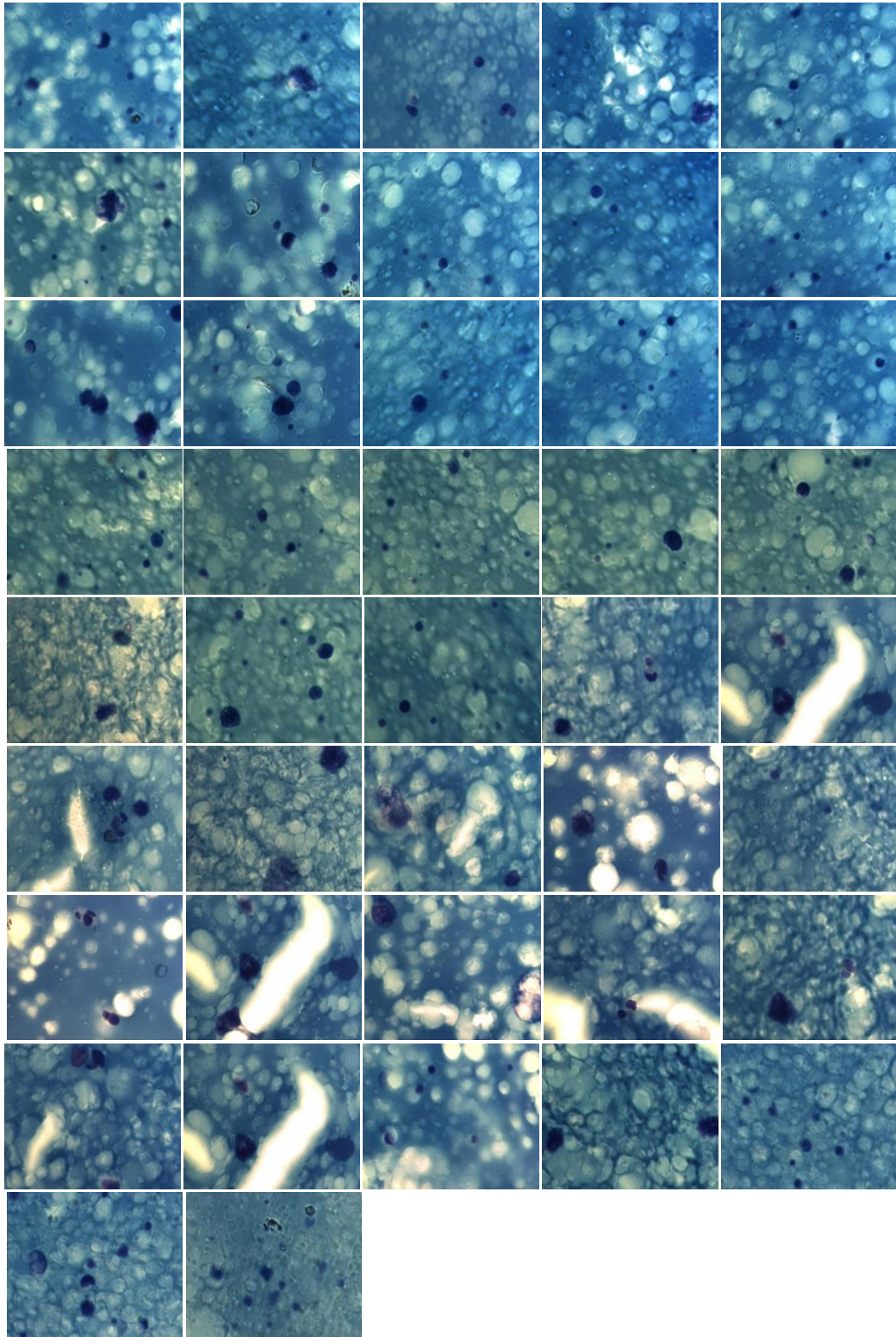




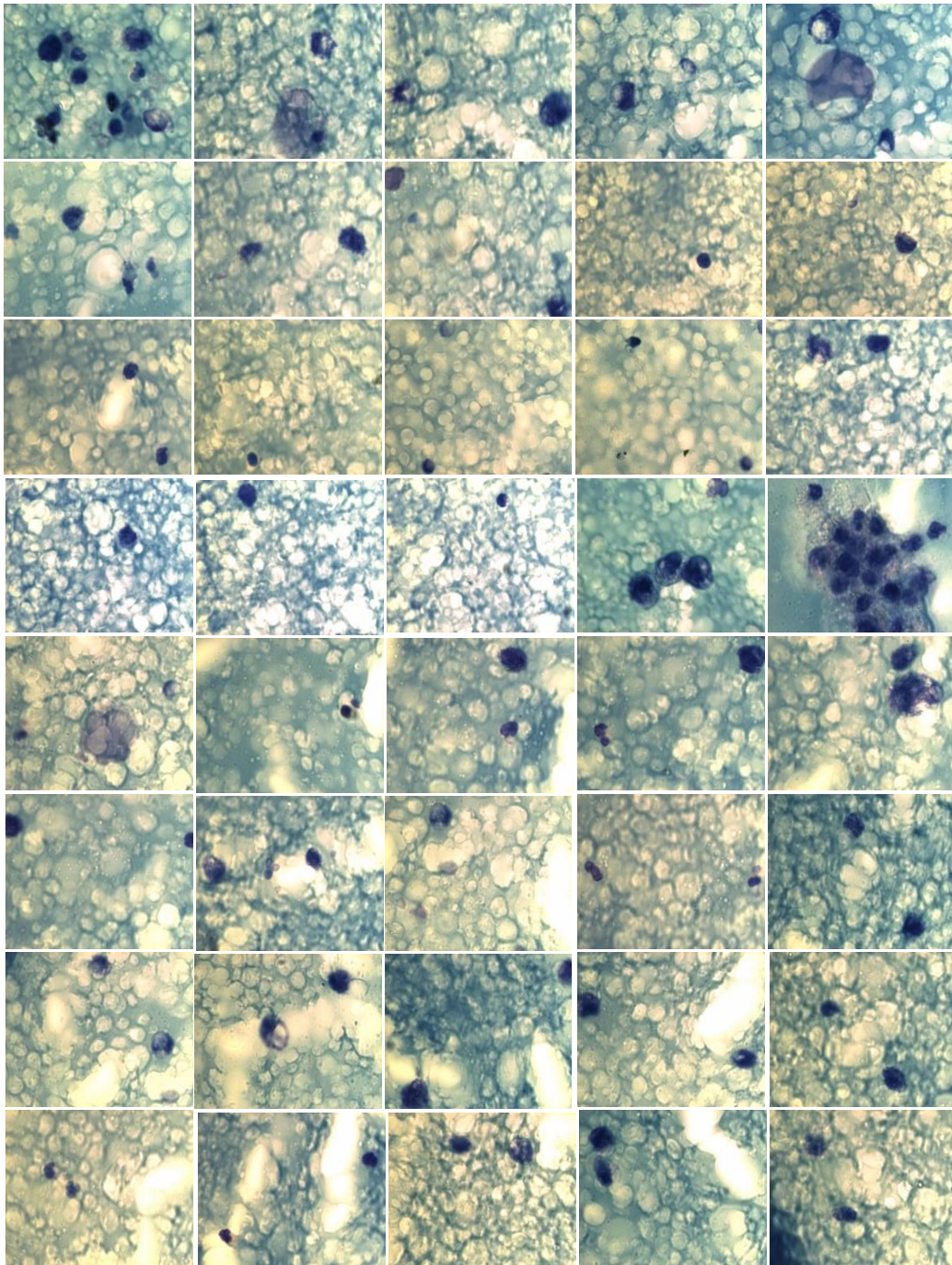
Imagens do Grupo 4

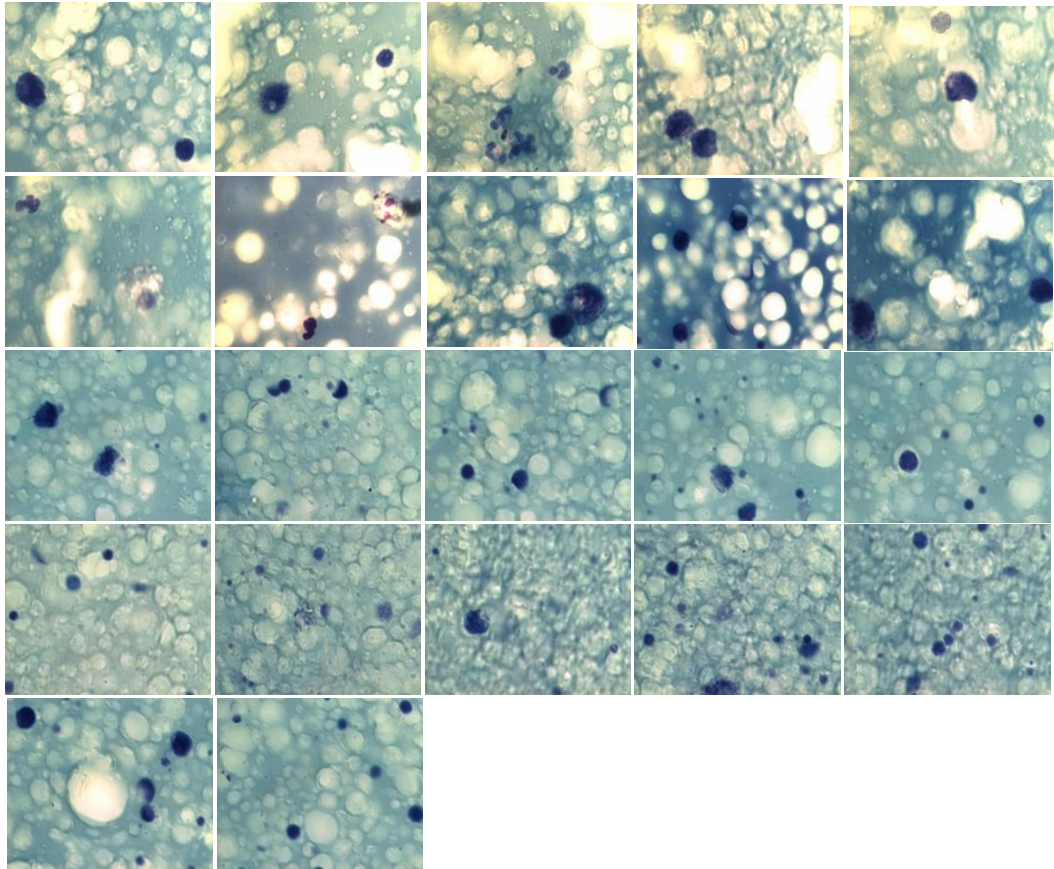






Imagens do Grupo 5





UFBA
UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA

PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA INDUSTRIAL - PEI

Rua Aristides Novis, 02, 6º andar, Federação, Salvador BA
CEP: 40.210-630
Telefone: (71) 3283-9800
E-mail: pei@ufba.br
Home page: <http://www.pei.ufba.br>

