



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE

CENTRO DE TECNOLOGIA



PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

# **Técnica para Segmentação Automática de Imagens Microscópicas de Componentes Sanguíneos e Classificação Diferencial de Leucócitos Baseada em Lógica *Fuzzy***

**Alessandra Mendes Pacheco Guerra Vale**

Orientador: Profa. Dra. Ana Maria Guimarães Guerreiro

**Tese de Doutorado** apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação da UFRN (área de concentração: Engenharia de Computação) como parte dos requisitos para obtenção do título de Doutor em Ciências.

**Natal, RN, dezembro de 2014.**

UFRN / Biblioteca Central Zila Mamede  
Catalogação da Publicação na Fonte

Vale, Alessandra Mendes Pacheco Guerra.

Técnica para segmentação automática de imagens microscópicas de componentes sanguíneos e classificação diferencial de leucócitos baseada em lógica fuzzy / Alessandra Mendes Pacheco Guerra Vale. – Natal, RN, 2014.

143 f. : il.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Ana Maria Guimarães Guerreiro.

Tese (Doutorado) – Universidade Federal do Rio Grande do Norte. Centro de Tecnologia. Programa de Pós-Graduação em Engenharia Elétrica e da Computação.

1. Processamento digital de imagens – Tese. 2. Lógica fuzzy – Tese.  
3. Segmentação de imagens – Tese. 4. Classificação diferencial de leucócitos – Tese. 5. Componentes sanguíneos – Tese. I. Guerreiro, Ana Maria Guimarães. II. Universidade Federal do Rio Grande do Norte. III. Título.

RN/UF/BCZM

CDU 621.397

# **Técnica para Segmentação Automática de Imagens Microscópicas de Componentes Sanguíneos e Classificação Diferencial de Leucócitos Baseada em Lógica *Fuzzy***

**Alessandra Mendes Pacheco Guerra Vale**

Tese de doutorado aprovada em 26 de dezembro de 2014 pela banca examinadora composta pelos seguintes membros:

---

Profa. Dra. Ana Maria Guimarães Guerreiro (orientadora) .....	DEB/UFRN
Prof. Dr. Adrião Duarte Dória Neto .....	DCA/UFRN
Prof. Dr. Marco Antônio Garcia de Carvalho .....	FT/UNICAMP
Prof. Dr. Allan de Medeiros Martins .....	DEE/UFRN
Profa. Dra. Cicília Raquel Maia Leite .....	DI/UERN



Universidade Federal do Rio Grande do Norte  
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

ATA N° 133

Aos vinte e seis dias do mês de dezembro do ano de dois mil e quatorze, foi realizada a 133ª sessão de defesa de tese de doutorado do Programa de Pós-Graduação em Engenharia Elétrica e de Computação da UFRN, na qual a doutoranda Alessandra Mendes Pacheco Guerra Vale apresentou o trabalho que tem como título: Técnica para Auto-Segmentação de Componentes Sanguíneos e Classificação Diferencial de Leucócitos Baseada em Lógica Fuzzy. A sessão teve inicio às 15h00min, tendo a banca examinadora sido constituída pelos participantes: Ana Maria Guimarães Guerreiro (Dra. UFRN, Orientadora), Adrião Duarte Doria Neto (Dr. UFRN, Examinador Interno ao Programa), Allan de Medeiros Martins (Dr. UFRN, Examinador Externo ao Programa), Cicilia Raquel Maia Leite (Dra. UERN, Examinadora Externa à Instituição) e Marco Antonio Garcia de Carvalho (Dr. UNICAMP, Examinador Externo à Instituição). Após a apresentação do trabalho e o exame pela banca, a doutoranda foi considerada APROVADA, tendo sido lavrada a presente ata, que vai assinada pelos examinadores e pela doutoranda. A versão final da tese deverá ser entregue ao programa, no prazo máximo de 60 dias, contendo as modificações sugeridas pela banca examinadora e constante na folha de correção anexa. Conforme o Artigo 49 da Resolução 197/2013 - CONSEPE, o candidato não terá o título se não cumprir as exigências acima.

Dr. CICILIA RAQUEL MAIA LEITE, UERN

Examinador Externo à Instituição

Marco Antônio C. de Carvalho

Dr. MARCO ANTONIO GARCIA DE CARVALHO, UNICAMP

Examinador Externo à Instituição

Dr. ALLAN DE MEDEIROS MARTINS, UFRN

Examinador Externo ao Programa

Dr. ADRIÃO DÚARTE DORIA NETO, UFRN

Examinador Interno

Dr. ANA MARIA GUIMARAES GUERREIRO, UFRN

Presidente

Doutorando

*Às minhas filhas, minhas flores,  
porque toda a minha vida é com elas, por elas e para elas.*

*Ao meu marido,  
porque no seu coração encontro a minha casa.*

*À minha casa,  
porque se existe um “melhor lugar do mundo”, este lugar é aqui.*

*Aos meus pais,  
porque sem eles, todas as dedicatórias acima desta seriam vãs.*

---

# Agradecimentos

---

À minha orientadora, Profa. Dra. Ana Maria Guimarães Guerreiro, por acreditar no meu trabalho e nesta caminhada.

Ao prof. Dr. Adrião Duarte Dória Neto, sempre presente no meu crescimento acadêmico e no meu coração, pela atenção amiga, palavra certeira, disponibilidade inabalável e fé constante na minha capacidade de concluir os meus propósitos. Obrigada por ter acreditado e lutado pelo meu sonho, que aqui se torna realidade.

Ao prof. Allan de Medeiros Martins por ter contribuído valorosamente com o seu conhecimento nos momentos mais duros desta estrada.

Aos professores Dr. Marco Antônio Garcia de Carvalho, Dra. Heliana Bezerra Soares e Dra. Cicília Raquel Maia Leite por adicionarem ao trabalho críticas construtivas importantes que o fizeram alcançar um patamar superior de qualidade.

Ao prof. Dr. Geraldo Barroso Cavalvanti Junior, ao amigo Victor Cesar Lucena Tavares de Sá Leitão e a todos os profissionais do Hemocentro do Rio Grande do Norte Dalton Cunha (Hemonorte) pelos ensinamentos e disponibilidades inestimáveis para a realização deste trabalho.

Ao amigo e prof. Dr. Fellipe Araújo Aleixo pelas contribuições sempre presentes.

Aos meus queridos colegas da Escola Agrícola de Jundiaí/UFRN, especialmente aos amigos inestimáveis que fazem parte do grupo de informática e aos professores Gerbson Azevedo de Mendonça e Júlio César de Andrade, pelo apoio, confiança e conhecimento que permitiram que este trabalho pudesse ser executado e concluído da melhor forma possível.

Aos demais professores da UFRN e amigos do Programa de Pós-Graduação em Engenharia Elétrica e de Computação/PPgEEC por todos os esclarecimentos e suporte necessários à execução deste trabalho.

Aos meus familiares, tios, tias, primos, primas, irmãos, cunhadas, sobrinhos e amigos por existirem e por serem sempre o oásis no meu deserto. Amo vocês.

Ao meu avô Eider e à minha avó Ruth, que vocês possam sentir, dos diferentes lugares nos quais se encontram, que me sinto grata por tê-los e sabê-los meus.

Ao meu tio Gilson pela sua disponibilidade constante em ajudar sempre que foi preciso, mesmo sem saber como. Obrigada pelo apoio e pela torcida. Amo você.

Aos meus pais, Robson e Juçara, por seus exemplos, apoios, ombros, colos, palavras, abraços, torcidas, olhares e corações. Obrigado por me ensinarem que não se desiste de um caminho e que a cada passo, por mais doloroso que seja, nos tornamos mais fortes, resistentes e valiosos. Pai, sinta-se orgulhoso de si mesmo. A sua história fez com que eu tomasse este seu caminho para mim mesma; foi o “ser professor” que vi em você que me trouxe até aqui. Mãe, sinta-se orgulhosa de si mesma. Você me fez gente, me fez luta, me fez resiliência, me fez resistência, me fez amor, me fez fé, me fez paz. Amo vocês.

Às minhas filhas, Julia e Marcelle. Queridas, vocês são os meus motivos, as minhas alegrias, a minha insistência em ser mais e melhor a cada dia. Em vocês encontro o que há de mais doce e perfeito do amor de Deus em minha vida. Por vocês enfrento qualquer desafio. Com vocês me sinto plena. Para vocês dedico a minha história. Amo vocês. Ah! Obrigado também pelas células e mais células que preencheram seus olhares e desejos de ajudar. ;)

Ao meu marido, Alicsson Roberto Guerra Vale. Você me deu todas as ferramentas que eu precisava para viver essa escolha: seu conhecimento, seu tempo, sua companhia, seus ensinamentos, seu colo, seu abraço, sua fé, seu amor. Em você encontrei o meu porto e nele me refiz, me recompus, me reconstruí. Com você escrevi essa conquista. Para você entrego esse mérito, minha alma e o meu coração. Te Amo.

À Deus que, silenciosamente, me sustenta. Silêncio e Paz.

---

# Resumo

---

A detecção automática de componentes sanguíneos em imagens microscópicas é um importante tópico da área hematológica. A segmentação permite que os componentes sanguíneos sejam agrupados em áreas comuns e a classificação diferencial dos leucócitos possibilita que os mesmos sejam analisados separadamente. Com a segmentação automática e classificação diferencial, contribui-se no processo de análise dos componentes sanguíneos, fornecendo ferramentas que propiciem a diminuição do trabalho manual e o aumento da sua precisão e eficiência. Utilizando técnicas de processamento digital de imagens associadas a uma abordagem *fuzzy* genérica e automática, este trabalho apresenta dois *Sistemas de Inferência Fuzzy*, definidos como *I* e *II*, para a segmentação automática de componentes sanguíneos e classificação diferencial de leucócitos, respectivamente, em imagens microscópicas de esfregaços. Utilizando o *Sistema de Inferência Fuzzy I*, a técnica desenvolvida realiza a segmentação da imagem em quatro regiões: núcleo e citoplasma leucocitários, eritrócitos e área de plasma e utilizando o *Sistema de Inferência Fuzzy II* e os leucócitos segmentados (núcleo e citoplasma leucocitários), os classifica diferencialmente em cinco tipos: basófilos, eosinófilos, linfócitos, monócitos e neutrófilos. Foram utilizadas para testes 530 imagens contendo amostras microscópicas de esfregaços sanguíneos corados com métodos variados. As imagens foram processadas e seus índices de Acurácia e *Gold Standards* foram calculados e comparados com os resultados manuais e com outros resultados encontrados na literatura para os mesmos problemas. Quanto à segmentação, a técnica desenvolvida demonstrou percentuais de acurácia de 97,31% para leucócitos, 95,39% para eritrócitos e 95,06% para plasma sanguíneo. Quanto à classificação diferencial, os percentuais variaram entre 92,98% e 98,39% para os diferentes tipos leucocitários. Além de promover a segmentação automática e classificação diferencial, a técnica desenvolvida contribui ainda com definição de novos descritores e a construção de um banco de imagens utilizando diversos processos de coloração hematológicos.

**Palavras-chave:** Processamento digital de imagens, lógica *fuzzy*, segmentação de imagens, classificação diferencial de leucócitos, componentes sanguíneos.

---

# Abstract

---

*Automatic detection of blood components is an important topic in the field of hematology. The segmentation is an important stage because it allows components to be grouped into common areas and processed separately and leukocyte differential classification enables them to be analyzed separately. With the auto-segmentation and differential classification, this work is contributing to the analysis process of blood components by providing tools that reduce the manual labor and increasing its accuracy and efficiency. Using techniques of digital image processing associated with a generic and automatic fuzzy approach, this work proposes two Fuzzy Inference Systems, defined as I and II, for auto-segmentation of blood components and leukocyte differential classification, respectively, in microscopic images smears. Using the Fuzzy Inference System I, the proposed technique performs the segmentation of the image in four regions: the leukocyte's nucleus and cytoplasm, erythrocyte and plasma area and using the Fuzzy Inference System II and the segmented leukocyte (nucleus and cytoplasm) classify them differentially in five types: basophils, eosinophils, lymphocytes, monocytes and neutrophils. Were used for testing 530 images containing microscopic samples of blood smears with different methods. The images were processed and its accuracy indices and Gold Standards were calculated and compared with the manual results and other results found at literature for the same problems. Regarding segmentation, a technique developed showed percentages of accuracy of 97.31% for leukocytes, 95.39% to erythrocytes and 95.06% for blood plasma. As for the differential classification, the percentage varied between 92.98% and 98.39% for the different leukocyte types. In addition to promoting auto-segmentation and differential classification, the proposed technique also contributes to the definition of new descriptors and the construction of an image database using various processes hematological staining.*

*Keywords:* *digital image processing, fuzzy logic, image segmentation, leukocytes differential classification, blood components.*

---

# Sumário

---

<b>Sumário</b>	<b>x</b>
<b>Lista de Figuras</b>	<b>xii</b>
<b>Lista de Tabelas</b>	<b>xv</b>
<b>Introdução</b>	<b>16</b>
<b>1.1. Motivação</b>	<b>16</b>
<b>1.2. Justificativas</b>	<b>18</b>
<b>1.3. Objetivos</b>	<b>19</b>
<b>1.4. Organização do Trabalho</b>	<b>20</b>
<b>Estado da Arte</b>	<b>22</b>
<b>2.1 Introdução</b>	<b>22</b>
<b>2.2 Conceitos Fundamentais Especialistas</b>	<b>22</b>
2.2.1 Microscopia	23
2.2.2 Técnicas de Coloração	26
2.2.3 Hematologia	27
2.2.4 Leucócitos	32
<b>2.3 Trabalhos Relacionados</b>	<b>35</b>
2.3.1 Segmentação de Componentes Sanguíneos e Classificação Diferencial de Leucócitos Utilizando Técnicas Variadas	35
2.3.2 Lógica Fuzzy Aplicada a Análises de Imagens	41
2.3.3 Segmentação de Componentes Sanguíneos e Classificação Diferencial de Leucócitos utilizando Lógica Fuzzy	45
<b>Processamento Digital de Imagens e Lógica Fuzzy</b>	<b>50</b>
<b>3.1 Introdução</b>	<b>50</b>
<b>3.2 Imagens Digitais</b>	<b>50</b>
<b>3.3 Processamento Digital de Imagens</b>	<b>54</b>
3.3.1 Aquisição de imagens	55
3.3.2 Pré-processamento	55
3.3.3 Segmentação	56
3.3.4 Representação e Descrição	56
3.3.5 Reconhecimento e interpretação	58
3.3.6 Base de Conhecimento	58
<b>3.4 Lógica Fuzzy</b>	<b>58</b>
3.4.1 Os Conjuntos Fuzzy	60
3.4.2 Conceitos Básicos em Conjuntos Fuzzy	60
3.4.3 Variáveis Linguísticas	62
3.4.4 Funções de Pertinência	64
3.4.5 Modificadores	66
3.4.6 Operações	66
3.4.7 Regras Linguísticas	67
3.4.8 Sistema de Inferência Fuzzy	69

<b>Técnica Desenvolvida</b>	<b>72</b>
<b>4.1 Introdução</b>	<b>72</b>
<b>4.2 Contextualização</b>	<b>72</b>
<b>4.3 Apresentação da Técnica e Etapas Desenvolvidas</b>	<b>73</b>
4.3.1 Etapa: Aquisição de Imagens	74
4.3.2 Base de Conhecimento	75
4.3.3 Etapa: Pré-Processamento	78
4.3.4 Etapa: Segmentação	83
4.3.5 Etapa: Representação e Descrição	91
4.3.6 Etapa: Reconhecimento e Interpretação	98
<b>Análises de Resultados</b>	<b>103</b>
<b>5.1 Introdução</b>	<b>103</b>
<b>5.2 Experimentos</b>	<b>103</b>
<b>5.3 Resultados</b>	<b>106</b>
5.3.1 Segmentação	106
5.3.2 Classificação	107
<b>5.4 Discussões</b>	<b>107</b>
5.4.1 Segmentação	108
5.4.2 Classificação	112
<b>Conclusões</b>	<b>114</b>
<b>6.1 Introdução</b>	<b>114</b>
<b>6.2 Conclusões</b>	<b>114</b>
<b>6.3 Contribuições</b>	<b>115</b>
<b>6.4 Trabalhos Futuros</b>	<b>116</b>
<b>6.5 Publicações</b>	<b>117</b>
<b>Referências</b>	<b>118</b>
<b>Apêndice A</b>	<b>129</b>

---

# **Lista de Figuras**

---

Figura 1. Microscópio óptico e suas divisões (GENOMA, 2012).	24
Figura 2. Imagem de fungo, feita em microscópio óptico (LVP, 2012).	24
Figura 3. Esquema de representação de elementos sanguíneos (CEB, 2012).	28
Figura 4. Esquema de representação de esfregaços sanguíneos, sendo (a) e (b) confecção do esfregaço e (c) esfregaço com coloração. (Biomedicina Brasil, 2011)	30
Figura 5. Esquema de leitura de células utilizado em exames manuais (Moura, 1997).	31
Figura 6. Classificação dos leucócitos (CEB, 2012).	32
Figura 7. Lâminas de esfregaços sanguíneos apresentando diversos leucócitos.	34
Figura 8. Convenção dos eixos para representação de imagens digitais (Gonzalez & Woods, 2000).	51
Figura 9. Representações do Modelo RGB (Gonzalez & Woods, 2009).	52
Figura 10. Máscaras espaciais para imagens RGB (Gonzalez & Woods, 2009).	53
Figura 11. Representações de uma imagem: a) cores (modelo RGB), b) escalas de cinza, c) preto e branco, d) componente R, e) componente G, f) componente B. Imagem: Lena.jpg (Gonzalez & Woods, 2000).	53
Figura 12. Etapas do Processamento Digital de Imagens (Gonzalez & Woods, 2000)	55
Figura 13. Variável linguística Velocidade (Speed) (Lee, 1990b).	62
Figura 14. Função de pertinência para a variável linguística temperatura (Tanscheit, 2012).	64
Figura 15. Formatos de funções de pertinência mais utilizados: triangular (trimf), trapezoidal (trapmf) e gaussiana (gaussmf). (MathWorks, 2014).	65
Figura 16. Sistema de Inferência Fuzzy (adaptado de Leite, 2009).	69
Figura 17. Representação gráfica da técnica desenvolvida neste trabalho.	73
Figura 18. Exemplo de imagem microscópica de esfregaço sanguíneo contendo um leucócito (púrpura), vários eritrócitos (róseos) e plasma leucocitário (área mais clara da amostra) (Heckner & Freund, 2000).	76
Figura 19. a) Imagens originais RGBmxnx3 (amostras coradas com os métodos Panótico Rápido, Leishman, Rosenfeld e com coloração não especificada, respectivamente); b) Canais G; c) Histogramas dos canais G com aumento (escalas de cinza x frequência); d) Picos dos histogramas: repetição das frequências mais altas; e) Picos dos histogramas: PicoEscuro – azul, PicoMédio– verde, PicoClaro– vermelho; f) Limiarização dos núcleos leucocitários (branco), centróides (1-vermelho), AltaProx (2-azul), BaixaProx (3-amarelo) e máximo valor da matriz D (4-verde); g) Distâncias Euclidianas de cada pixel em relação ao centróide do núcleo leucocitário mais próximo (mesmos pontos 1, 2, 3 e 4).	82
Figura 20. Sistema de Inferência Fuzzy I – Variáveis linguísticas fuzzy de entrada e saída.	84

Figura 21. Sistema de Inferência Fuzzy I – Superfície do Sistema.	85
Figura 22. Sistema de Inferência Fuzzy I – funcionamento do sistema para um pixel considerando $g_{ij} = 200$ e $d_{ij} = 50$ .	86
Figura 23. Sistema de Inferência Fuzzy I – Exemplos de Funções de Pertinência para valores PicoEscuro, PicoMédio, PicoClaro, AltaProx e BaixaProx previamente definidos na etapa de pré-processamento para uma imagem específica.	86
Figura 24. Processo de fuzificação (amostras coradas com os corantes Panótico Rápido, Leishman, Rosenfeld e sem coloração especificada, respectivamente): a) fuzificação para tonalidade escura; b) fuzificação para tonalidade média; c) fuzificação para tonalidade clara; d) fuzificação para proximidade alta; e) fuzificação para proximidade baixa.	88
Figura 25. Regras fuzzy (amostras coradas com os corantes Panótico Rápido, Leishman, Rosenfeld e sem coloração especificada, respectivamente): a) Regra fuzzy 1 – classe Núcleo; b) Regra fuzzy 2 – classe Plasma; c) Regra fuzzy 3 – classe Eritrócito; d) Regra fuzzy 4 – classe Citoplasma; e) Variável de saída Classe.	89
Figura 26. Pós-processamento (amostras coradas com os corantes Panótico Rápido, Leishman, Rosenfeld e sem coloração especificada, respectivamente): a) Descarte de falsos positivos para citoplasma; b) Classificação final.	91
Figura 27. Segmentação e descritores leucocitários: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.	93
Figura 28. Basófilos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.	94
Figura 29. Eosinófilos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.	95
Figura 30. Linfócitos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.	96
Figura 31. Monócitos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.	97
Figura 32. Neutrófilos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro	

equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.	98
<b>Figura 33. Sistema de Inferência Fuzzy II.</b>	<b>99</b>
Figura 34. Função de Pertinência para as variáveis de entrada fuzzy (neste exemplo, RazãoNL) do Sistema de Inferência Fuzzy II.	100
<b>Figura 35. Classificação final da imagem mostrada na Figura 27 utilizando o Sistema de Inferência Fuzzy II: Neutrófilo.</b>	<b>102</b>
Figura 36. Exemplos de imagens originais (RGB) e segmentadas pelo Sistema de Inferência Fuzzy I proposto, cujas colorações hematológicas são, respectivamente: a) G1: Panótico Rápido; b) G2: Leishman; c) G3: Rosenfeld; d) G4: coloração não especificada.	104
<b>Figura 37. Interfaces desenvolvidas para o processamento individual das imagens: a) Interface de Segmentação; b) Interface de Classificação.</b>	<b>105</b>
Figura 38. Exemplos de imagens com resultados negativos: a) FP para núcleos leucocitários e FN para citoplasmas leucocitários - Semelhança entre as suas cores; b) FP para núcleos leucocitários - presença de artefatos; c) FN para núcleos leucocitários - borda da segmentação manual; d) FP para citoplasmas leucocitários - eritrócitos adjacentes; e) FN para citoplasmas leucocitários - não foram corados de forma significativa; f) FN para eritrócitos - localizados nas bordas; g) FP para eritrócitos - presença de plaquetas; h) FP para eritrócitos - perda de foco do microscópio.	109
Figura 39. Percentuais médios obtidos para as métricas Gold Standards e Acurácia da técnica desenvolvida para segmentação. a) Verdadeiro Positivo (TP) e Falso Positivo (FP); b) Verdadeiro Negativo (TN) e falso negativo (FN); c) Acurácia (AC).	110
<b>Figura 40. Análise comparativa para a segmentação de leucócitos entre os percentuais de acurácia (AC) utilizando a técnica desenvolvida e os percentuais disponíveis na literatura, sendo [1] Técnica Fuzzy Proposta, [2] Zheng et. al. (2014), [3] Jati et. al. (2014) – com ruído, [4] Putzu &amp; Ruberto (2013), [5] Ramesh et. al. (2012), [6] Faticahah et. al. (2012) - núcleo leucocitário, [7] Faticahah et. al. (2012) – citoplasma leucocitário, [8] Rezatofighi &amp; Soltanian-Zadeh (2011), [9] Ko et. al. (2011) – núcleo leucocitário, [10] Ko et. al. (2011) - citoplasma leucocitário, [11] Hamghalam &amp; Aytollahi (2009) e [12] Ramoser et. al. (2005).</b>	<b>111</b>

---

# **Lista de Tabelas**

---

Tabela 1. Limiares das funções de pertinência utilizadas na fuzificação do sistema de inferência fuzzy utilizado na etapa de classificação.	101
Tabela 2. Resultados percentuais (Gold Standards) encontrados para a segmentação dos componentes sanguíneos para os grupos de imagens G1, G2, G3, G4 e média dos resultados para todos os grupos: Verdadeiro Positivo (TP), Verdadeiro Negativo (TN), Falso Positivos (FP), Falso Negativo (FN) e Acurácia (AC).	107
Tabela 3. Matriz de confusão, precisão e acurácia geral para as classificações dos cinco tipos de leucócitos.	107
Tabela 4. Análise comparativa entre índices de acurácia (AC) para segmentação de leucócitos utilizando a técnica fuzzy proposta e os descritos na literatura.	112
Tabela 5. Análise comparativa entre taxas de acurácia (AC) para classificação diferencial de leucócitos (Basófilos – B, Eosinófilos – E, Linfócitos – L, Monócitos – M, Neutrófilos – N) encontradas nos trabalhos descritos na literatura e resultantes da técnica desenvolvida neste trabalho para os grupos G1, G2 e G3.	113

---

# Capítulo 1

## Introdução

---

### 1.1. Motivação

O olho humano tem um poder de resolução de, aproximadamente, 0,1 mm. Pode-se afirmar então que dois objetos separados por uma distância menor que 0,1mm serão vistos, a olho nu, como um objeto único. Portanto, para possibilitar a distinção entre tais objetos faz-se necessária a utilização de instrumentos ópticos que tenham poder de resolução aumentada.

Entretanto, é importante salientar a diferença entre poder de aumento e poder de resolução. Se a uma imagem qualquer for inferido sucessivas vezes um limiar de aumento, os pontos originalmente separados por distâncias menores que 0,1mm continuarão sendo percebidos como um único ponto, muito embora a nova imagem esteja perceptivelmente maior. Tem-se assim um aumento do seu tamanho, mas não uma melhora na sua resolução. Entretanto, se à resolução de tal imagem também for aplicado um limiar de aumento, os pontos tenderão a afastar-se.

Solucionando tais limitações, os microscópios propiciaram ao homem a observação de estruturas com poder muito maior de ampliação e de resolução. O limite de resolução dos microscópios ópticos atuais é de até 2.000 vezes. Sem o surgimento da microscopia não seria possível a observação de inúmeros elementos que têm fundamental importância na qualidade e manutenção da vida.

Imprescindível para a evolução de diversas áreas, a análise microscópica colaborou especial e consideravelmente com a hematologia, área médica que estuda o sangue, seus órgãos originários, funções, distúrbios e doenças, possibilitando a observação das células e seus componentes invisíveis a olho nu e essenciais para o processo de diagnóstico. Nesta área, a análise dos diversos componentes sanguíneos, especialmente leucócitos e eritrócitos, foco deste trabalho, pode ser feita a partir das observações microscópicas de esfregaços, que consistem em amostras sanguíneas coradas e fixadas em lâminas de extensão. As análises diagnósticas quantitativas e qualitativas baseadas nos conteúdos celulares do sangue são

importantes, visto que as células sanguíneas são indicadoras de perturbações ou degradações nos seus órgãos de origem, de mais difícil acesso.

Os componentes sanguíneos consistem em três tipos básicos de células produzidos na medula óssea: as hemácias ou eritrócitos (glóbulos vermelhos), os leucócitos (glóbulos brancos) e as plaquetas. Existem várias doenças relacionadas ao sangue, como anemia, hemofilia e leucemia, diagnosticáveis a partir de análise sanguínea.

O hemograma é considerado o principal exame no estudo da função hematológica e é solicitado na prática médica em diversas situações para detecção e diagnóstico de alergias, inflamações, infecções e outras moléstias. Com o auxílio de um microscópio, uma análise atenta e cuidadosa dos elementos celulares implica, além da quantificação de cada um dos tipos de células sanguíneas, no estudo da sua morfologia, que pode contribuir para o diagnóstico diferencial de diversas doenças. De acordo com o foco na série vermelha, branca e plaquetária, o hemograma pode ser subdividido em eritrograma, leucograma e plaquetograma, respectivamente. O leucograma consiste na contagem total e diferencial dos leucócitos e na avaliação do esfregaço sanguíneo ao microscópio, podendo evidenciar ou não a presença de infecção ou inflamação e acompanhar a evolução de doenças infecciosas ou inflamatórias e no controle de tratamentos que usam quimioterapia ou radiação.

Mesmo com a evolução da microscopia óptica para a eletrônica, a qual possibilita a quantificação automática de aspectos da amostra em análise, a observação do analista clínico ao microscópio óptico ainda é indispensável. Os elementos chegam a ser registrados e/ou quantificados pelo aparelho eletrônico, porém não são identificados sob o ponto de vista qualitativo (Heckner & Freund, 2000). A análise diferencial de leucócitos é feita através do reconhecimento visual das células com o auxílio de um microscópio, identificando, selecionando e contando cada tipo, a fim de emitir o resultado em proporção ao número total de elementos, o que permitirá ao médico diagnosticar o tipo de doença, se existir, através dos dados obtidos pelo exame. A contagem diferencial de leucócitos consiste em um dos mais valiosos métodos entre os exames do sangue.

Com o avanço tecnológico, alternativas têm sido desenvolvidas objetivando automatizar os processos quantitativo e qualitativo da microscopia na hematologia. Tem-se buscado equipamentos e sistemas computacionais capazes de realizar estudos microscópios visando à uniformização dos resultados e maior velocidade nas análises. Fornecendo soluções que capturam imagens microscópicas e as processam, utilizando técnicas de processamento digital de imagens, lógica fuzzy, sistemas especialistas e redes neurais, entre outras, o uso de sistemas inteligentes têm alcançado resultados satisfatórios.

## 1.2. Justificativas

Na hematologia, o estudo microscópico quantitativo e qualitativo de amostras sanguíneas é vital para o diagnóstico de diversas moléstias que comprometem o sistema hematológico. Minuciosos e demorados, tais estudos ocasionam muitas vezes para o observador um custo razoável na obtenção dos resultados ou só podem ser feitos de forma manual.

A fim de possibilitar a análise microscópica, as amostras passam previamente por processos de elaboração e coloração. Utilizando-se técnicas específicas, como *Giemsa*, *Wrigth*, *May-Grünwald*, *Rosenfeld* e *Leishman*, as amostras são distribuídas em esfregaços sanguíneos e coradas. Os processos de coloração são os responsáveis por atribuir aos componentes sanguíneos cores que os diferenciam. Assim, a coloração exerce fundamental importância no processo, pois permite a distinção por tons e cores dos diversos elementos amostrados. Quando corados, normalmente são encontrados os seguintes padrões: a área de plasma se mantém com os tons mais claros da amostra, os eritrócitos e os citoplasmas leucocitários apresentam tons intermediários e os núcleos leucocitários apresentam os tons mais escuros. Porém, como existe uma variação de coloração para cada componente, o processo de identificação possui características inexatas e subjetivas.

Ao microscópio óptico, as análises manuais feitas por especialistas para identificação, caracterização e contagem dos diferentes componentes sanguíneos são minuciosas e demoradas, possuindo um caráter intrínseco de imprecisão, difícil reprodução e subjetividade, ocasionando, muitas vezes, um custo razoável na obtenção dos resultados. Conforme descrito em Guo *et. al.* (2007), a análise automática visa identificar diferentes classes de células e é um importante tópico no diagnóstico de doenças como o câncer e a anemia. Neste contexto, as técnicas de processamento digital de imagens, que propiciam melhorias na qualidade e análise dos seus elementos e lidam constantemente com dados imprecisos em diferentes níveis, têm alcançado bons resultados.

Segundo Zadeh (1973), a lógica *fuzzy* propõe-se a expressar matematicamente as formulações do pensamento humano em linguagem natural sem, contudo, diminuir a potência expressiva das mesmas. Segundo Cox (1994), combinando a imprecisão associada aos eventos naturais e o poder computacional das máquinas, podem ser produzidos sistemas de resposta inteligentes, robustos e flexíveis. Assim, segundo Boaventura (2010), o conceito de precisão ou imprecisão, dependendo do contexto aplicado, expressar-se-á numericamente, indicando a possibilidade, e não a probabilidade, de uma inferência estar correta.

Nos conjuntos tradicionais, o conceito de pertinência de um elemento a um conjunto é bem definido. Nos conjuntos *fuzzy*, a ideia está associada a um grau de pertinência, que varia de acordo com o elemento em questão. A pertinência de um elemento em relação a determinado conjunto deve ser entendida como a intensidade com que este elemento está relacionado a tal conjunto. Desta forma, evidenciada em Klir & Folger (1988), segundo a variação do seu grau de pertinência – da completa exclusão até a total pertinência – um elemento pode pertencer muito ou pouco e até não pertencer ao conceito representado pelo conjunto *fuzzy*, podendo assumir qualquer um dos valores intermediários. É esta característica de subjetividade que permite que a lógica *fuzzy* atue fortemente nos problemas que modelam conceitos imprecisos como os de segmentação e classificação, que visam detectar e classificar regiões distintas da imagem.

Visando auxiliar nos processos de análise microscópica dos componentes sanguíneos, este trabalho apresenta uma técnica utilizando sistemas de inferência *fuzzy* para segmentar a imagem em quatro regiões distintas: leucócito (e suas sub-regiões: núcleo e citoplasma leucocitários), eritrócito e área de plasma e, posteriormente, classificar cada leucócito em um dos cinco tipos principais: basófilo, eosinófilo, neutrófilo, linfócito e monócito.

Dante do contexto apresentado, faz-se necessária a utilização de técnicas inteligentes na análise destas imagens, permitindo que os elementos sejam processados de forma automática e inferindo autonomamente em respostas como análises quantitativas, qualitativas e diagnósticas. Como ferramenta inteligente, optou-se pela utilização da lógica *fuzzy* por possuir, como característica intrínseca, o processamento baseado em conceitos não exatos tal qual aqueles inerentes ao problema pesquisado, como cor, forma, textura, granulosidade e grau de similaridade.

### 1.3. Objetivos

O objetivo principal deste trabalho é apresentar um conjunto de técnicas baseadas em lógica *fuzzy* para segmentar automaticamente os núcleos e citoplasmas leucocitários, eritrócitos e áreas de plasma sanguíneo em imagens microscópicas de esfregaços e, após a segmentação, classificar diferencialmente os leucócitos nos cinco tipos principais: basófilos, eosinófilos, linfócitos, monócitos e neutrófilos.

Como objetivos específicos, podem ser citados:

- Detectar e acompanhar o estado da arte relacionado ao objetivo principal deste trabalho;

- Definir, a partir dos conhecimentos de especialistas da área hematológica, bases de regras que possibilitem a identificação automática dos diferentes componentes sanguíneos, como leucócitos, eritrócitos e áreas de plasma, independente da coloração hematológica utilizada, e dos principais tipos leucocitários;
- Definir técnicas de segmentação automáticas e adaptativas com conceitos bem definidos e valores que variam para cada imagem de esfregaço sanguíneo;
- Calcular automaticamente os domínios das variáveis linguísticas *fuzzy* para a etapa de segmentação utilizando o histograma da imagem e as distâncias Euclidianas entre os centróides dos núcleos leucocitários e demais *pixels*;
- Identificar quatro regiões na imagem para histogramas que apresentam apenas três regiões bem definidas;
- Reduzir as taxas de falsos negativos para citoplasmas leucocitários que apresentarem eritrócitos adjacentes;
- Definir novos descritores que, associados a outros já existentes, possibilitem uma classificação diferencial mais precisa entre os cinco principais tipos leucocitários;
- Aplicar as técnicas desenvolvidas em uma base de imagens reais construída utilizando-se diversos processos de coloração hematológicos;
- Comparar os resultados da técnica desenvolvida com a segmentação e classificação manuais feitas por especialistas;
- Comparar os índices de acurácia encontrados com os relatados na literatura para o mesmo problema.

## 1.4. Organização do Trabalho

Este trabalho está organizado conforme apresentação abaixo relacionada:

- O Capítulo 2 apresenta a fundamentação teórica especialista necessária contendo os conceitos básicos pertinentes ao contexto do problema, como microscopia, técnicas de coloração, hematologia, componentes sanguíneos e tipos leucocitários e o estado da arte, mostrando a utilização da lógica *fuzzy* em sistemas computacionais inteligentes de análises de imagens e outras abordagens para segmentação de componentes sanguíneos e classificação diferencial de leucócitos utilizando lógica *fuzzy* e técnicas variadas;

- O Capítulo 3 apresenta as imagens digitais e seus conceitos, o processamento digital de imagens e suas etapas e a lógica *fuzzy* e seus conceitos principais e pertinentes para este trabalho;
- O Capítulo 4 apresenta as técnicas propostas para a segmentação automática de componentes sanguíneos e classificação diferencial de leucócitos utilizando lógica *fuzzy*, suas definições, etapas, sistemas de inferência, variáveis linguísticas e bases de regras.
- O Capítulo 5 apresenta os experimentos realizados, resultados encontrados e as discussões;
- O Capítulo 6 apresenta as conclusões, as contribuições, os trabalhos futuros e as publicações.

---

# Capítulo 2

## Estado da Arte

---

### 2.1 Introdução

Neste capítulo são apresentados o estado da arte e a fundamentação teórica especialista relacionados ao foco principal deste trabalho. Para tal, subdivide-se em:

- **Conceitos Fundamentais Especialistas:** apresenta um breve estudo sobre a microscopia e seus conceitos, as técnicas de coloração mais comumente utilizadas em amostras de esfregaços sanguíneos para fins de análise microscópica direta, uma descrição da área de hematologia, seus componentes e tipos de análises e o detalhamento dos leucócitos.
- **Segmentação de Componentes Sanguíneos e Classificação Diferencial de Leucócitos Utilizando Técnicas Variadas:** apresenta trabalhos relacionados às áreas de segmentação de componentes sanguíneos e classificação diferencial de leucócitos em tipos leucocitários distintos, como eosinófilos, basófilos, linfócitos, monócitos e neutrófilos, utilizando diversas técnicas computacionais.
- **Lógica Fuzzy Aplicada a Análises de Imagens:** apresenta trabalhos diversos que mostram a relevância da utilização da lógica *fuzzy* associada a técnicas de processamento digital de imagens, inclusive na área médica.
- **Segmentação de Componentes Sanguíneos e Classificação Diferencial de Leucócitos Utilizando Lógica Fuzzy:** apresenta trabalhos relacionados às áreas de segmentação de componentes sanguíneos e classificação diferencial de leucócitos utilizando lógica *fuzzy*.

### 2.2 Conceitos Fundamentais Especialistas

Para o completo entendimento do problema foco deste trabalho, a segmentação de componentes sanguíneos e classificação diferencial de leucócitos em imagens de esfregaços

sanguíneos, alguns conceitos especializados, à parte dos sistemas computacionais e inteligentes, mostram-se relevantes. Segue-se uma explanação destes conceitos fundamentais, chamados aqui de *especialistas*, suficiente para o entendimento do contexto no qual se encontra inserido o problema. Não se tem, entretanto, a intenção de exaurir estes assuntos além do relevante para o entendimento deste trabalho, visto que são demasiadamente extensos e fogem do escopo definido.

### 2.2.1 Microscopia

Não são perceptíveis para a visão humana objetos com diâmetros inferiores a um décimo de milímetro (0,1mm ou 100 $\mu$ m). Para possibilitar a visualização desses objetos, o microscópio é a ferramenta mais amplamente utilizada. A sua principal função é tornar visível ao olho humano o que for muito pequeno para tal (Goodhew *et. al.*, 2001).

A partir da microscopia, foram possíveis a observação e descoberta de inúmeras estruturas e seres vivos microscópicos até aqui desconhecidos, como bactérias, protozoários e leveduras. Primordialmente importante, em 1835, Schleiden & Schwann propuseram as bases da Teoria Celular, primeiro grande princípio unificador da biologia, o qual define como pressuposto que todos os organismos vivos são constituídos por células, sendo estas as unidades estruturais e funcionais dos mesmos (Amabis & Martho 2008).

A observação de estruturas diversas utilizando microscopia como uma extensão natural da observação a olho nu representou papel importante no surgimento das ciências da natureza. São exemplos dessas estruturas as células vivas e mortas (após fixação e coloração), as bactérias, os ovos de vermes, o sangue humano e seus componentes, os micro-organismos em alimentos, as partículas formadoras do solo e muitas outras. A forma mais antiga e usual da microscopia é a lupa, seguida pelo microscópio óptico e pelos microscópios eletrônicos.

#### Microscopia Óptica

A microscopia óptica permite a visualização de estruturas diversas, cujas observações seriam impossíveis a olho nu, através da incidência de luz e de lentes objetivas que promovem grandes aumentos, como o de até 2000 vezes.

O microscópio óptico (Figura 1) possui lentes objetivas variadas que proporcionam visões panorâmicas de aumentos diversos. O aumento final é resultado da multiplicação do aumento dado pela lente objetiva pelo aumento da lente ocular. Como existem várias lentes objetivas num mesmo microscópio, uma grande variedade de aumentos pode ser facilmente

atingida, bastando girar o revólver. Assim, se for utilizada uma objetiva de 20 vezes e uma ocular de 10 vezes o aumento final será de 200 vezes.

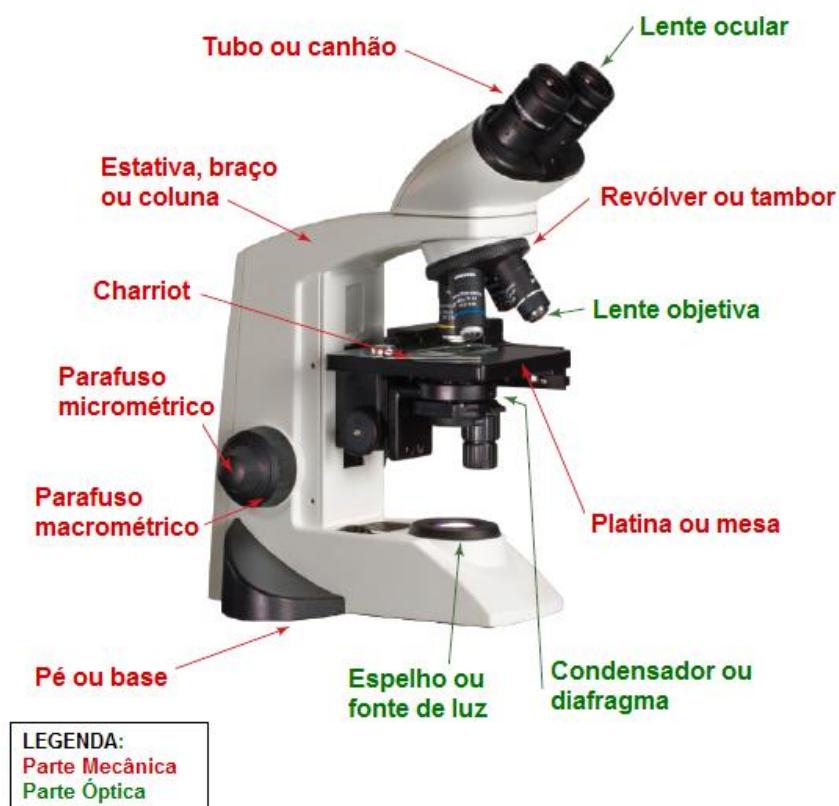


Figura 1. Microscópio óptico e suas divisões (GENOMA, 2012).

Para a visualização de estruturas com aumento de mais de 1000 vezes (Figura 2) é necessária utilização de óleo de imersão para que a lâmina não seja danificada (LFF, 2012).

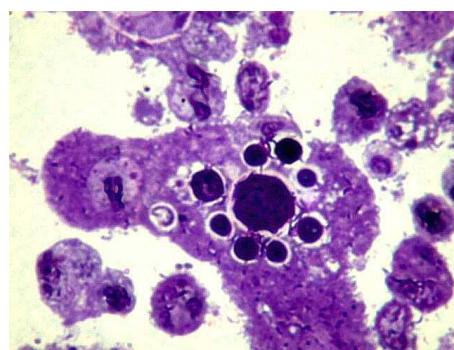


Figura 2. Imagem de fungo, feita em microscópio óptico (LVP, 2012).

Para a melhor utilização do microscópio, diversos materiais e técnicas foram desenvolvidos, como: corantes, fixadores, micrótomo, esfregaço e esmagamento. As diferentes técnicas utilizadas em microscopia dependem das finalidades laboratoriais. Por

exemplo, se as lâminas forem para fins educacionais, devem-se utilizar técnicas que propiciem montar uma lâmina permanente. No entanto, se a lâmina for preparada para testes laboratoriais na área de saúde, como contagem de células, devem-se utilizar técnicas que propiciem o descarte seguindo as normas de biossegurança que forem necessárias (UNIP, 2012).

Durante a evolução histórica da microscopia, o microscópio óptico foi aperfeiçoado até o ponto em que a única limitação era o grande comprimento de onda da radiação (luz visível) utilizada para iluminação, obstáculo que impedia a obtenção de um maior poder de resolução.

### **Microscopia Eletrônica**

Em 1924, o físico Louis de Broglie constatou que um feixe de elétrons apresentava um comportamento idêntico aos raios luminosos, mas com um comprimento de onda 10.000 vezes menor. Estavam assim elaboradas as bases teóricas do microscópio eletrônico que, em 1933, já ultrapassava o limite de resolução do microscópio óptico. A microscopia eletrônica progrediu rapidamente a partir de grandes aperfeiçoamentos técnicos que permitiram não apenas maiores valores de ampliação como também aumentos sucessivos da capacidade de resolução e da qualidade das imagens obtidas. Estes progressos foram também possíveis graças ao aperfeiçoamento dos métodos de preparação do material para observação, sendo desenvolvidas várias técnicas, como a de obtenção de cortes ultrafinos e a de fixação de estruturas através do uso de resinas sintéticas.

A diferença básica entre o microscópio óptico e o eletrônico é que neste último não é utilizada a luz, mas sim feixes de elétrons. Além disso, no microscópio eletrônico não há lentes de cristal e sim bobinas, chamadas lentes eletromagnéticas.

O microscópio eletrônico possui potencial de aumento superior ao microscópio óptico, permitindo aumentos de 300.000 vezes ou até superiores a 1 milhão de vezes, dependendo do material (Maliska, 2007).

Não é possível observar material vivo neste tipo de microscópio. O material a ser estudado passa por um complexo processo de desidratação, fixação e inclusão em resinas especiais, muito duras, que permitem cortes ultrafinos obtidos através das navalhas de vidro do instrumento conhecido como ultramicrótomo (EMBRAPA, 2012).

Atualmente as técnicas de microscopia, tanto óptica quanto eletrônica, continuam sendo de vital importância em inúmeras áreas. São exemplos de áreas que utilizam a microscopia:

histologia, anatomia, hematologia, bacteriologia, microbiologia, radiologia, alimentos, morfologia, análise geográfica de solos e sólidos, mineralogia e petrografia (Bradbury, 1989).

O foco deste trabalho é a análise de imagens provenientes da aplicabilidade da microscopia óptica na área de hematologia. Com técnicas especificamente desenvolvidas para coloração dessas amostras, as imagens são analisadas quantitativa e qualitativamente para os mais diversos fins diagnósticos.

### 2.2.2 Técnicas de Coloração

A maioria das observações microscópicas iniciais é feita com preparações coradas. Tanto microrganismos quanto materiais celulares são frequentemente incolores e podem ser distinguidos entre si através da utilização de técnicas de coloração. (Martinez *et. al.*, 2005).

Segundo Pelczar *et. al.* (2005),

*“Existem dois métodos gerais utilizados para preparar espécimes microbiológicos para observações por meio de microscópio luminoso. Um utiliza uma suspensão de microrganismos vivos em uma gota ou uma camada líquida. No outro uma camada fina do espécime é seca e corada, assim os microrganismos ficam fixados à superfície e apresentam-se corados para facilitar a visualização.”*

As técnicas de coloração são de importância fundamental para este trabalho visto que as amostras não tratadas têm pouca ou nenhuma diferenciação óptica. Nestes casos, são utilizados corantes que tingem as amostras, montadas em lâminas ou lamínulas, aumentando o contraste de seus componentes e possibilitando a observação microscópica. Esta visualização direta permite a distinção da composição celular e morfologia dos microrganismos expostos (Tortora *et. al.*, 2003).

Cada corante reage apenas com certos elementos, que ficam contrastados em relação aos outros, o que facilita a observação. Além disso, certos cuidados precisam ser tomados. Para a análise microscópica de células vivas, por exemplo, deve-se ter o cuidado de usar corantes que não alterem nem destruam o material biológico. Tais corantes, chamados de corantes vitais, como o azul-de-metileno e a eosina, são utilizados normalmente nestas preparações, feitas para exames de ocasião. Entretanto, quando se pretende fazer um grande

número de observações ou exames demorados, são geralmente utilizadas preparações definitivas. Neste caso, além da coloração, há necessidade de preservar, da maneira mais perfeita possível, a estrutura original das amostras através de fixadores, como o álcool, o formol, o éter e o ácido acético.

As colorações de um modo geral se efetuam por processos físico-químicos ou puramente físicos e podem ser consideradas, segundo a modalidade, a ação, o caráter, o grau de ação, o tempo, o número de corantes e a cromatização. Quanto à cromatização, ou seja, de acordo com o número de cores conferidas às estruturas pelas colorações simples ou combinadas, estas tomam a denominação de colorações monocrônicas (uma cor), bicrônicas (duas cores), tricrônicas (três cores) e policrônicas (mais de três cores).

Em análises microscópicas sanguíneas, os principais corantes utilizados são o azul-de-metileno e a eosina. Tais corantes são preparados segundo metodologias propostas por vários autores: *Leishman*, *May-Grünwald*, *Giemsa*, *Wright*, *Rosenfeld* e outros (que dão os respectivos nomes ao corantes). Estes corantes são dissolvidos em álcool e, na solução envelhecida, o azul de metileno se oxida em graduações diferentes, originando diversos “azuis” de metileno. Teremos então uma solução alcoólica de um complexo eosinato de azul e “azuis” de metileno. Os componentes celulares sanguíneos coram-se diferentemente, em tons vermelho e roxo, permitindo assim uma fácil identificação (FCF, 2012).

### 2.2.3 Hematologia

O sangue é o meio líquido que flui pelo sistema circulatório, transportando oxigênio e outros nutrientes, hormônios, eletrólitos, água e resíduos do metabolismo celular. O sistema circulatório provê uma ligação entre os diversos órgãos e células do organismo e o sangue mantém o equilíbrio do meio ambiente celular ao circular através dos tecidos.

Segundo Souza & Elias (2005),

*“Pode-se definir, em linhas gerais, a hematologia como o estudo do sangue e dos tecidos formadores das células sanguíneas; abrange o estudo dos elementos celulares sanguíneos: hemácias, leucócitos e plaquetas, suspensos no seu meio líquido, o plasma sanguíneo; estuda ainda as funções do sangue no organismo e as doenças primárias do sangue e dos tecidos hematopoiéticos (formadores dos elementos sanguíneos, como medula óssea). A*

*hemoterapia estuda o emprego do sangue e dos seus principais componentes, isolados mediante diversos processos de separação, como recursos necessários à reposição de eventuais perdas e ao tratamento de diversas condições patológicas”.*

O sangue se divide em duas fases: líquida e sólida. A fase líquida é representada pelo plasma, enquanto que a fase sólida é constituída por três diferentes tipos celulares: eritrócitos (glóbulos vermelhos ou hemácias), leucócitos (glóbulos brancos) e plaquetas (Figura 3). Estes elementos celulares têm forma, tamanho e funções distintas. Os eritrócitos são as células que existem em maior quantidade no sangue e são responsáveis pela coloração avermelhada deste. Os leucócitos distinguem-se em cinco tipos, neutrófilos, eosinófilos, basófilos, linfócitos e monócitos, e constituem a parte celular do sistema imunológico ou de defesa do organismo. De um modo bastante simplificado, podemos afirmar que os glóbulos vermelhos sustentam a vida do organismo, enquanto os glóbulos brancos a defendem. As plaquetas, por sua vez, possuem importância fundamental nos mecanismos da hemostase (coibição de hemorragia) e coagulação sanguíneos (Theml *et. al.*, 2004).

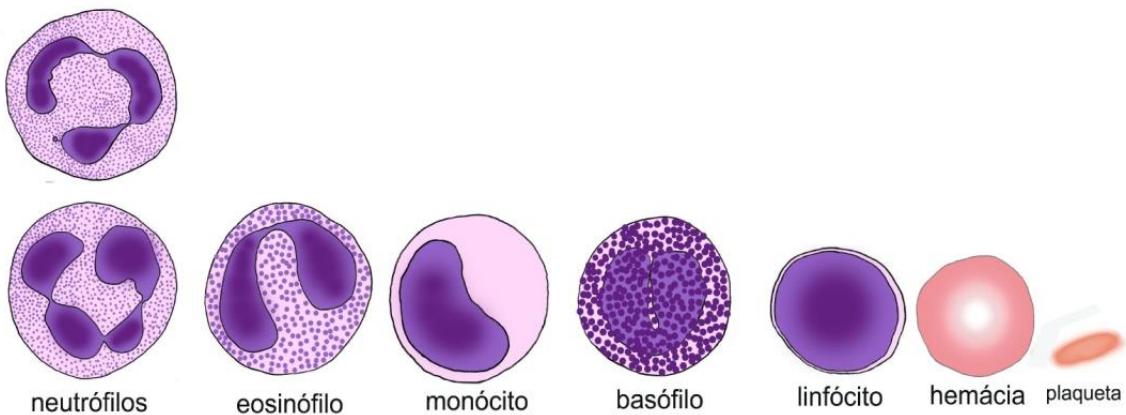


Figura 3. Esquema de representação de elementos sanguíneos (CEB, 2012).

O exame hematológico ou hemograma constitui a análise quantitativa e qualitativa das células e estruturas que compõe o tecido sanguíneo, incluindo em sua abrangência o estudo da concentração, estrutura e função das células. O hemograma normalmente subdivide-se em eritrograma, leucograma e plaquetograma. O eritrograma estuda a contagem e alterações nos eritrócitos, na hemoglobina, no hematocrito, nos índices globulares e na morfologia eritrocitária. O leucograma estuda a contagem total e diferencial de leucócitos assim como as

fórmulas percentual e absoluta e o estudo da sua morfologia. O plaquetograma faz uma estimativa do número de plaquetas e estuda sua morfologia.

As análises das amostras sanguíneas podem ser realizadas utilizando-se métodos manuais, como observações microscópicas diretas ou automáticas, baseadas no uso de técnicas que avaliam as variações de impedância do fluxo elétrico ou da dispersão de luz produzida pelas diferentes células. Os métodos automáticos, apesar de propiciarem análises mais precisas e rápidas, se baseiam em características médias, podendo assim conduzir a resultados errôneos. Além disso, uma análise por métodos manuais reveste-se de grande importância nos casos em que é necessária uma avaliação qualitativa.

Segundo Santos *et. al.* (2009),

*“Mesmo que a contagem eletrônica de células seja mais precisa por avaliar um maior número de células, esta pode induzir a um elevado número de falso-positivos que devem ser confirmados pela microscopia, para garantia de um resultado reproduzível. Assim, mesmo sendo a automação indispensável no laboratório de análises clínicas, fornecendo resultados sensíveis, reproduzíveis e precisos para alterações quantitativas, as técnicas manuais, particularmente de concentrado de leucócitos, mostram-se insubstituíveis para alterações qualitativas e continuam úteis na rotina laboratorial, pelo fato de que algumas alterações hematológicas só podem ser diferenciadas pela microscopia.”*

No eritrograma, os métodos automáticos são mais utilizados pela precisão dos resultados. No leucograma, entretanto, a contagem é feita através do reconhecimento visual das células com o auxílio de um microscópio óptico. O processamento de imagem médica desempenha um papel importante na hematologia. A segmentação de uma imagem de esfregaço de sangue periférico em suas regiões constituintes ajuda o hematologista a avaliar o paciente com maior precisão (Jati *et. al.*, 2014). A identificação, seleção e análise diferencial permite ao médico diagnosticar o tipo de doença, se existir (Richetto, 2007).

Para a análise manual dos leucócitos, é confeccionado um esfregaço sanguíneo, posteriormente corado, contendo a amostra de sangue a ser analisada.

Segundo Richetto (2007),

*“O sangue, objeto do exame, é coletado do paciente através de tubos comerciais com pressões negativas. Após a coleta, o sangue, anticoagulado por agentes especiais adicionados ao tubo de coleta, é colocado em uma lâmina de vidro, onde é formado o ‘esfregaço’, nome dado à região da lâmina de sangue, onde duas gotas do material coletado são distribuídas uniformemente e coradas com reagente adequado para o exame que se deseja realizar.”*

A avaliação de um esfregaço sanguíneo é uma parte importante na avaliação da doença hematológica. Embora um diagnóstico específico possa ser sugerido com base em resultados obtidos por métodos automáticos, muitas doenças têm uma contagem celular normal com morfologia celular anormal. A confecção do esfregaço sanguíneo é padronizada e consiste em um ponto crucial para a realização de um hemograma confiável (Figura 4). O esfregaço ideal deve ser livre de falhas e paradas, não muito espesso, nem fino demais, e sem falhas na cauda. Na observação ao microscópio, as bordas onde são realizadas as contagens devem apresentar os eritrócitos mais separados e os leucócitos bem distribuídos.

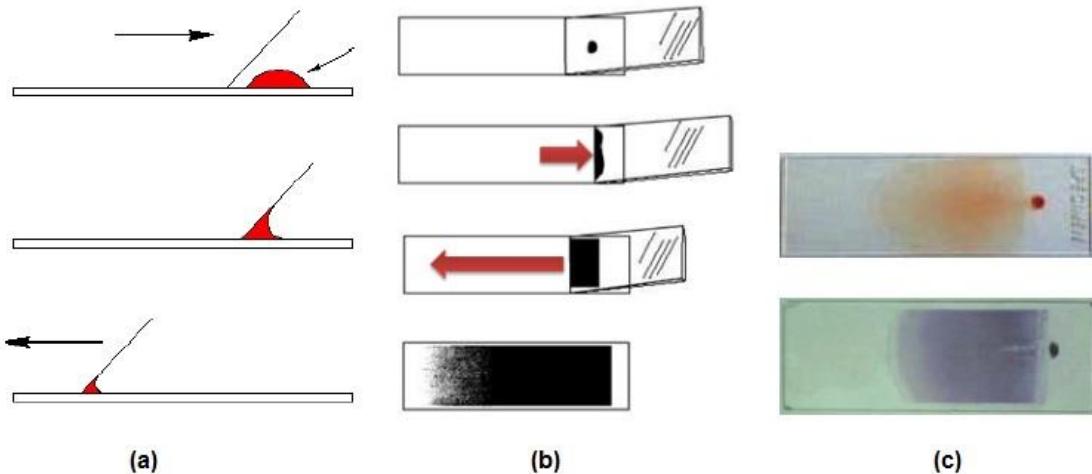


Figura 4. Esquema de representação de esfregaços sanguíneos, sendo (a) e (b) confecção do esfregaço e (c) esfregaço com coloração. (Biomedicina Brasil, 2011)

Os componentes sanguíneos não têm cor, mas apresentam coloração quando as lâminas são coradas com produtos químicos para torná-los visíveis ao microscópio (Kumar *et. al.*, 2010). Para destacar os componentes incolores disponíveis em esfregaço, alguns tipos

especiais de corantes são utilizados e este processo é conhecido como coloração. Diferentes tipos ou métodos são utilizados para coloração de esfregaços: *Wright* (Pan *et. al.*, 2012), *May-Grünwald Giemsa* (Hamghalam *et. al.*, 2009), *Leishman* (Ghosh *et. al.*, 2010) *Rosenfeld* (Vendruscolo *et. al.*, 2006), entre outros.

A segmentação celular é um problema desafiador devido à natureza complexa das células e à incerteza presente na análise microscópica. Os métodos manuais são onerosas, imprecisos e altamente subjetivos, exigindo, portanto, métodos automatizados que executem essa tarefa de forma objetiva e eficiente (Dorini *et. al.*, 2007). As imagens de esfregaço sanguíneo para análise clínica e pré-clínica são amplamente adquirido por meio de microscopia óptica de imersão. Portanto, um exame de microscopia de um esfregaço de sangue devidamente preparado e bem marcado é necessário e clinicamente útil em várias circunstâncias (Gulati *et. al.*, 2013).

A contagem manual dos leucócitos pode, através de comparação de valores relativos, mostrar excesso ou falta de determinadas células. Durante o exame, a lâmina contendo o esfregaço sanguíneo é disposta no microscópio óptico, que tem o seu foco ajustado. No processo de identificação e contagem das células, a lâmina é movimentada, conforme Figura 5, e cada tipo de célula é contado até que se tenha um total de 100 células. Na sequência, é feito um cálculo diferencial de valores relativos de cada tipo de leucócito encontrado.

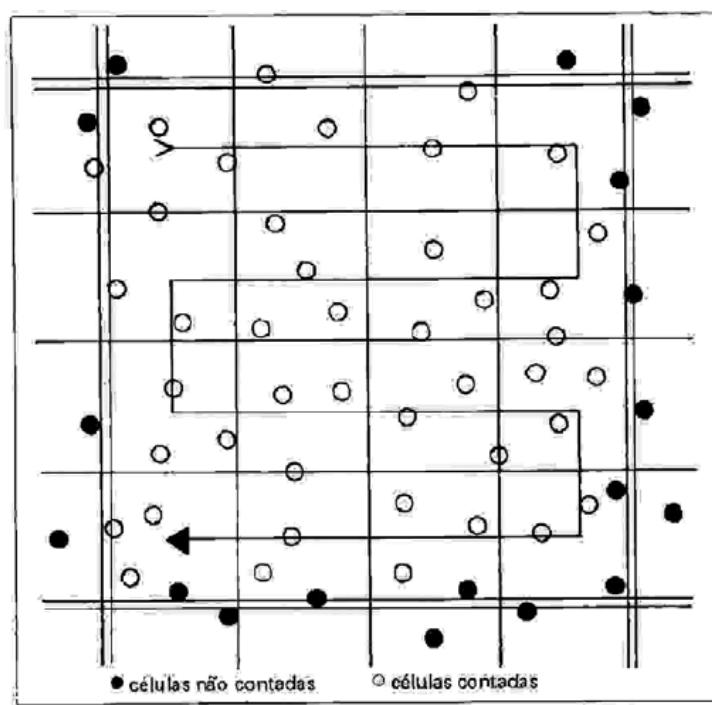


Figura 5. Esquema de leitura de células utilizado em exames manuais (Moura, 1997).

A identificação de células de leucócitos é um dos exames de sangue mais realizados. A observação quantitativa das células de leucócitos é muitas vezes complementada por análise morfológica (Li *et. al.*, 2014). A quantificação de leucócitos como uma parte da definição de seu papel no processo da doença ou durante o rastreio de drogas anti-inflamatórias (Kuse *et. al.*, 2010) é de grande importância. Além disso, a classificação de células faz-se essencial para saber o tipo e a duração do processo de inflamação e também pode ser útil em aspectos clínicos terapêuticos.

#### 2.2.4 Leucócitos

Os leucócitos são, na verdade, um grupo de diferentes células, com diferentes funções no sistema imunológico. Alguns leucócitos atacam diretamente o invasor, outros produzem anticorpos, outros apenas fazem a identificação. O valor normal dos leucócitos não depende do sexo do paciente e varia entre 4000 e 11000 células por ml de sangue.

Os leucócitos são reunidos de início em dois grupos: granulócitos e agranulócitos (Figura 6). Esta denominação se prende à presença ou ausência de granulação no citoplasma dos mesmos.

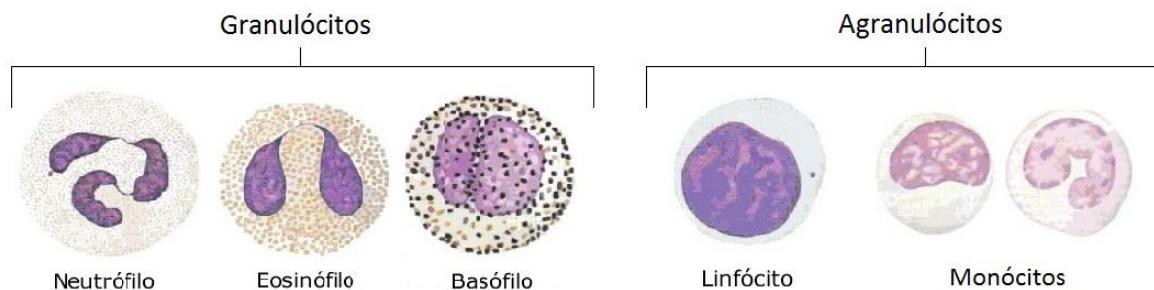


Figura 6. Classificação dos leucócitos (CEB, 2012).

A família de leucócitos é composta de eosinófilos, basófilos, linfócitos, monócitos e neutrófilos. Os cinco tipos de leucócitos podem ser distinguidos pelos seus grânulos citoplasmáticos, manchando propriedades dos grânulos, o tamanho da célula, a proporção do núcleo para o material citoplasmático, e do tipo de lóbulos de nucléolos (Chaira, 2014).

Segundo Saraswat & Arya (2014), são leucócitos granulócitos:

- **Neutrófilos:** primeira linha de defesa celular contra invasão de microrganismos, o neutrófilo é o tipo de leucócito mais comum. Representando em média de 50% a 70% dos leucócitos circulantes, os neutrófilos são especializados no combate a bactérias. Quando há uma infecção bacteriana, a medula óssea aumenta a sua

produção, fazendo com que sua concentração sanguínea se eleve. Portanto, quando temos um aumento do número de leucócitos totais, causado basicamente pela elevação dos neutrófilos, estamos provavelmente diante de um quadro infeccioso bacteriano. Os neutrófilos têm um tempo de vida de aproximadamente 24-48 horas. Por isso, assim que o processo infeccioso é controlado, a medula óssea reduz a produção de novas células e seus níveis sanguíneos retornam rapidamente aos valores normais. Os neutrófilos possuem o núcleo segmentado em lobos, em número que varia de dois a cinco, sendo denominados “neutrófilos segmentados”.

- **Eosinófilos:** Os eosinófilos são os leucócitos responsáveis pelo combate de parasitas e pelo mecanismo da alergia. Apenas 1 a 5% dos leucócitos circulantes são eosinófilos. Estes possuem grânulos que aparecem em tons de rosa ou vermelho em lâminas coradas, cujos núcleos possuem frequentemente dois lobos conectados. O aumento de eosinófilos ocorre em pessoas alérgicas, asmáticas ou em casos de infecção intestinal por parasitas. Os eosinófilos são um pouco maiores que os neutrófilos.
- **Basófilos:** tipo menos comum de leucócitos no sangue, representam menos de 1% de todos os leucócitos e apresentam grandes grânulos, corados do azul marinho ao roxo, muitas vezes tão numerosos que mascaram o núcleo. Sua elevação normalmente ocorre em processos alérgicos e estados de inflamação crônica. Normalmente em pequeno número, possuem um núcleo segmentado e granulações maiores do que aquelas existentes nos demais granulócitos.

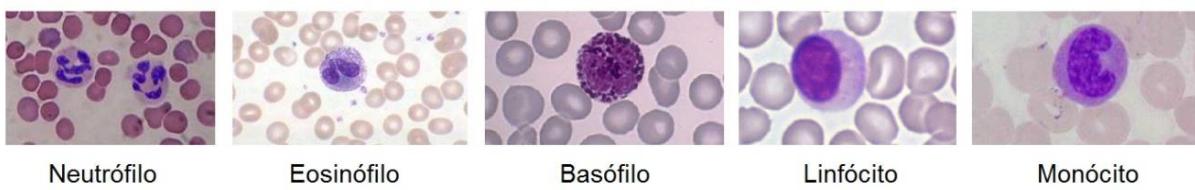
#### E agranulócitos:

- **Linfócitos:** segundo tipo mais comum, representam de 25 a 35% dos leucócitos no sangue. Os linfócitos são as células que fazem o reconhecimento de organismos estranhos, iniciando o processo de ativação do sistema imunológico. Principais linhas de defesa contra infecções por vírus e surgimento de tumores, são eles também os responsáveis pela produção dos anticorpos. Quando temos um processo viral em curso, é comum que o número de linfócitos aumente, às vezes, ultrapassando o número de neutrófilos e tornando-se o tipo de leucócito mais presente na circulação. Possuem um núcleo regular grande e arredondado que ocupa praticamente toda a célula, deixando uma borda muito fina de

citoplasma. Em lâminas coradas, apresenta coloração azul escura ou púrpura. Esta célula é menor do que os basófilos, eosinófilos e neutrófilos.

- **Monócitos:** correspondem aos maiores leucócitos do sangue circulante e representam em média de 3 a 9% do total de leucócitos. Apresentam núcleo grande em forma de rim ou farradura e sem segmentação com a cromatina delicada disposta em forma de rede. Quando corados, apresentam núcleo azul marinho ou púrpura e citoplasma azul claro. São ativados tanto em processos virais quanto bacterianos. Quando um tecido está sendo invadido por algum germe, o sistema imunológico encaminha os monócitos para o local infectado. Os monócitos tipicamente se elevam nos casos de infecções, principalmente naquelas mais crônicas como a tuberculose.

Alguns exemplos de imagens microscópicas de leucócitos podem ser vistos na Figura 7.



*Figura 7. Lâminas de esfregaços sanguíneos apresentando diversos leucócitos.*

As alterações quantitativas dos leucócitos na corrente sanguínea são chamadas de leucocitose e leucopenia. Enquanto a leucocitose é caracterizada pelo aumento do número de leucócitos, a leucopenia é caracterizada pela sua diminuição. A leucocitose pode ser causada por uma linfocitose, aumento do número de linfócitos, ou por uma neutrofilia, aumento do número de neutrófilos, por exemplo. Já a leucopenia pode surgir devido a uma linfopenia, diminuição do número de linfócitos, ou neutropenia, diminuição do número de neutrófilos. Quando existe aumento ou redução dos valores dos leucócitos é importante ver qual das linhagens descritas anteriormente é a responsável por essa alteração. Como neutrófilos e linfócitos são os tipos mais comuns, estes geralmente são os responsáveis pelo aumento ou diminuição da concentração dos leucócitos.

Grandes leucocitoses podem ocorrer nas leucemias, o câncer nos leucócitos. Enquanto processos infecciosos podem elevar os leucócitos até 20.000 – 30.000 células/ml, na leucemia estes valores ultrapassam facilmente os 50.000 células /ml. As leucopenias normalmente

ocorrem por lesões na medula óssea. Podem ser por quimioterapia, por drogas, por invasão de células cancerígenas ou por invasão por micro-organismos.

## 2.3 Trabalhos Relacionados

Diversos estudos envolvendo o processamento digital de imagens, associado à lógica *fuzzy* ou a outras técnicas, e voltado para a segmentação de componentes sanguíneos e classificação diferencial de leucócitos ou outros problemas específicos podem ser encontrados na literatura e alguns deles estão relacionados nos tópicos seguintes.

### 2.3.1 Segmentação de Componentes Sanguíneos e Classificação Diferencial de Leucócitos Utilizando Técnicas Variadas

No trabalho exposto por Kumar *et. al.* (2002), foi proposta uma segmentação do núcleo leucocitário a partir de um novo operador de detecção de bordas que enfatiza a fronteira entre o núcleo e o citoplasma. Neste trabalho, o fundo da imagem foi removido com limiar fixo e o centro do núcleo foi detectado a partir de sucessivas operações de erosão. Uma janela retangular foi calculada a partir do centro para a identificação do citoplasma e falsos positivos (eritrócitos) foram retirados pelo cálculo da área. Este método apresentou bons resultados, desde que os citoplasmas não fossem vizinhos dos eritrócitos. Todas as imagens utilizadas nos testes foram coradas pelo método *May-Grünwald Giemsa*. Este trabalho não apresentou índices de acurácia.

Jiang *et. al.* (2003) propuseram a segmentação de leucócitos combinando filtragem espacial para segmentação do núcleo e agrupamento *watershed* para extração do citoplasma. Este método apresentou resultados melhores para amostras no modelo RGB mas também não foi suficiente para imagens que apresentaram vizinhança entre eritrócitos e leucócitos.

Em Sinha & Ramakrishna (2003) foi proposta uma técnica para contagem diferencial de leucócitos utilizando o algoritmo *k-means* para a detecção de núcleos e os algoritmos *k-means* e Maximização de Expectativas (EM) para a identificação do citoplasma. Para a classificação diferencial, foram extraídas características de forma, cor e textura do núcleo e citoplasma leucocitários e exploradas diferentes combinações de classificadores. Este trabalho apresentou índice de acurácia de 80% para a segmentação em 115 amostras coradas com o método May-Grunwald-Giemsa e de 94% a 97% para a classificação diferencial.

O trabalho apresentado por Sabino *et. al.* (2004) mostrou como a informação de textura pode melhorar a diferenciação entre os leucócitos. Além dos descritores conhecidos, como perímetro, área, circularidade, raio núcleo-citoplasma e convexidade, cinco descritores baseados em textura, inércia, energia, entropia, inércia e homogeneidade, foram calculados a partir das matrizes de co-ocorrência da imagem, aumentando a precisão quanto à identificação dos tipos leucocitários. Os índices de acurácia para a classificação diferencial variaram de 62,5% à 97,73%.

Ramoser *et. al.* (2005) propuseram um método para a segmentação de leucócitos analisando a saturação para localizar o núcleo leucocitário e usando um abordagem de limiarização adaptativa para a segmentação do citoplasma. Este trabalho apresentou índice de acurácia de 94,4% para a segmentação em 534 amostras e de 78,9% a 98% para a classificação diferencial.

O trabalho de Wu *et. al.* (2006) apresentou uma abordagem iterativa utilizando o método de limiarização Otsu (Otsu, 1979) baseado no histograma circular e utilizando os componentes H e S do modelo de cores HSI para segmentação de leucócitos. Este trabalho não apresentou índices de acurácia.

O trabalho de Mircic & Jorgovanovic (2006) propôs a classificação diferencial dos leucócitos utilizando Redes Neurais Artificiais e descritores como área, solidez, circularidade e granularidade. Este trabalho apresentou índice de acurácia de 86%.

Em Zamani & Safabakhsh (2006) foi proposto um método de detecção e segmentação de leucócitos utilizando inicialmente o fecho convexo para determinação do contorno inicial do núcleo e o algoritmo do contorno ativo (*snake*) para segmentação o leucócito considerando a fronteira do núcleo como contorno inicial. Este trabalho apresentou índice de acurácia de 95% para a segmentação de leucócitos em 30 imagens.

No trabalho de Dorini *et. al.* (2007) foi proposta a segmentação do leucócito utilizando operadores morfológicos associados a propriedades de escala e espaço e um operador de alternância para melhorar a precisão da segmentação, porém os eritrócitos vizinhos ainda foram eventualmente classificados como citoplasmas. Este trabalho apresentou testes em 100 imagens, porém sem especificar os índices de acurácia.

Guo *et. al.* (2007) propuseram uma técnica de imagem multiespectral com o método de calibração espectral para adquirir imagens microscópicas de leucócitos na medula óssea independentes do dispositivo. Para a segmentação de imagens, Máquinas de Vetor de Suporte foram aplicadas diretamente sobre espectro de cada *pixel*, usando o algoritmo Sequencial de

Otimização Mínima para a seleção de recurso e redução do tempo de treinamento do classificador. Os índices de acurácia para a segmentação não foram divulgados.

Theera-Umporn & Dhompongsa (2007) investigaram se as informações sobre o núcleo leucocitário por si só eram suficiente para classificar os leucócitos do sangue. As características do núcleo foram analisadas utilizando-se morfologia matemática. Aplicando os classificadores de Bayes e as redes neurais artificiais, classificaram o leucócito usando quatro características do núcleo leucocitário: área do núcleo, a localização do pico de seu espectro padrão e o primeiro e o segundo momentos granulométricos do espectro padrão. Este trabalho apresentou testes em 431 imagens e os índices de acurácia para a classificação variaram de 50,23% à 70,53%.

Hamghalam & Aytollahi (2009) propuseram a segmentação do núcleo leucocitário a partir da binarização da imagem, selecionando como limiar o pico do histograma referente à intensidade dos eritrócitos. O citoplasma leucocitário foi extraído a partir da distância entre o centro do núcleo e os eritrócitos mais próximos. A proposta utilizou somente imagens coradas com o método Giemsa. Em Hamghalam *et. al.* (2009) foi apresentado um método para a segmentação de leucócitos utilizando o método Otsu e o algoritmo do contorno ativo (*snake*) para detectar o limite preciso dos núcleos e do citoplasma, respectivamente. Os trabalhos utilizaram um total de 30 e 20 imagens nos testes, respectivamente, e apresentaram índices de acurácia de 96,70%.

O trabalho proposto por Rezatofighi *et. al.* (2009) abordou um método baseado na teoria da ortogonalidade e processo de *Gram-Schmidt* para segmentar apenas os núcleos dos leucócitos. Os demais componentes não foram considerados. O trabalho não apresentou índices de acurácia.

O trabalho apresentado por Khashman (2009) mostrou uma rede neural baseada no algoritmo de aprendizagem *back propagation* aplicada ao problema de classificação celular. Os desempenhos da rede neural proposta e de uma rede neural convencional, utilizando duas topologias para cada rede, foram comparados. Os resultados experimentais mostraram que os pesos e parâmetros adicionais ofereceram melhorias à taxa de identificação e ao tempo de classificação. O trabalho apresentou um índice de identificação das células de 99,17% em 360 imagens.

Osowski *et. al.* (2009) propuseram um grande conjunto de características para o reconhecimento das células do sangue, utilizando o Algoritmo Genético e uma Máquina de Vetor de Suporte. Tais características foram divididas em quatro grupos principais: texturais, geométricas, estatísticas e morfológicas. As características texturais e geométricas foram

calculadas utilizando os três canais de cores (vermelho, verde e azul) da imagem. As características estatísticas foram baseadas nos histogramas e na matriz de gradiente da imagem para os canais vermelho, verde e azul. As características morfológicas foram calculados após a aplicação de operadores morfológicos, como a erosão e a dilatação. Os índices de acurácia não foram apresentados.

O método proposto por Kuse *et. al.* (2010) utilizou, para a identificação de linfócitos, dezoito características de textura derivadas da matriz de co-ocorrência dos níveis de cinza de imagens histopatológicas. O processo envolveu a segmentação das células no espaço HSV, extração de características, classificação e tratamento de sobreposição. O método utilizou características de textura das células para treinar um classificador utilizando Máquinas de Vetor de Suporte e diferenciar linfócitos de não-linfócitos. O índice de acurácia do classificador foi de 78% em 4 imagens contendo 94 linfócitos e 74 não-linfócitos.

Em Chan *et. al.* (2010) foi proposto um método de segmentação para o leucócito com o objetivo de contar o número de lóbulos nos núcleos das células. Primeiramente os leucócitos foram segmentados da imagem considerando-se as intensidades dos níveis de cinza e os tamanhos dos objetos a partir da detecção de bordas. Após a segmentação inicial, o núcleo do leucócito foi isolado. A forma do núcleo e o número de lóbulos são características importantes para determinar o tipo de leucócito. Um método de contagem de lóbulos foi desenvolvido. Considerando que uma junção entre dois lóbulos no núcleo de um leucócito é geralmente estreita, as operações de erosão e dilatação foram aplicadas para a correta separação e contagem. Porém, o desempenho do método é profundamente afetado pelos valores dos limiares aplicados no algoritmo. Assim, um detector de parâmetro baseado em algoritmos genéticos foi empregado para determinar os valores mais adequados. O método mostrou-se capaz de segmentar o núcleo em lóbulos nos resultados experimentais. Os índices de acurácia da segmentação não foram apresentados.

O trabalho de Hiremath *et. al.* (2010) apresentou uma proposta para segmentação e classificação automática de leucócitos considerando o canal H do modelo de cores HSV e um limiar fixo para a limiarização. Após a segmentação, foram eliminadas regiões com área inferior a 1.000 pixels (valor pré-fixado) para posterior classificação. Os testes em 75 imagens mostraram índices de acurácia entre 96,5% e 98% para a classificação.

Ko *et. al.* (2011) apresentaram uma técnica de segmentação de leucócitos baseada em regras de agrupamento e remoção de fronteiras utilizando o algoritmo *snake*. Regras foram utilizadas para a remoção de bordas e ruídos no limite do citoplasma enquanto o algoritmo *snake* foi forçado a deformá-lo, exigindo, entretanto, um maior tempo computacional. Os

índices de acurácia foram de 69,05% para citoplasma leucocitário e de 88,75% para núcleo leucocitário em 260 imagens classificadas.

O trabalho de Rezatofighi & Soltanian-Zadeh (2011) propôs um algoritmo de processamento de imagens para classificar cinco tipos de células sanguíneas (leucócitos) automaticamente. Primeiramente o núcleo é separado do citoplasma usando o método Gram–Schmidt. Se não há uma fronteira clara entre o núcleo e o citoplasma, são utilizados algoritmos específicos (*snake*). Na sequência são extraídas características de morfologia, cor e textura das áreas do núcleo e citoplasma. Para o processo de classificação, são comparados métodos utilizando máquinas de vetor de suporte e redes neurais artificiais. Neste trabalho foram processadas 150 imagens coradas pelo método Wright-Giemsa e os índices de acurácia foram de 93,09% para segmentação e de 93,10% a 100% para a classificação diferencial.

Ramesh *et. al.* (2012) propuseram a segmentação de leucócitos a partir do canal S da imagem, modelo de cores HSV. Um valor fixo para o limiar foi utilizado objetivando identificar o núcleo leucocitário. O citoplasma leucocitário foi identificado a partir de uma distância fixa e pré-definida do núcleo anteriormente segmentado. Neste trabalho foram processadas 320 imagens coradas pelo método Wright e os índices de acurácia foram de 93,08% para segmentação e de 80,17% a 100% para a classificação diferencial.

O trabalho de Pan *et. al.* (2012) apresentou um método para segmentação de leucócitos baseado em Máquina de Aprendizado Extremo (ELM). Utilizando o crescimento de regiões baseado em entropia para a segmentação de leucócitos, o método localiza, no estágio de amostragem, as regiões de interesse e no estágio de aprendizagem o classificador é treinado e extraí os leucócitos da imagem. Neste trabalho 65 amostras coradas pelo método Wright-Giemsa foram testadas mas os índices de acurácia não foram apresentados.

Putzu & Ruberto (2013) apresentaram um método para a segmentação de leucócitos utilizando um limiar automaticamente calculado pelo algoritmo *Zack* e a segmentação *watershed*. Os índices de acurácia para 245 amostras foram de 92% e 75,1% para segmentação e classificação, respectivamente.

A abordagem proposta por Cuevas *et. al.* (2013) foi baseada no algoritmo de evolução diferencial (ED), transformando a tarefa de detecção em um problema de otimização, cujos indivíduos representaram elipses candidatas. Uma função objetiva verifica se as elipses candidatas estão presentes no mapa de bordas da imagem microscópica sanguínea e, guiadas pelos valores da função, um conjunto de elipses candidatas são evoluídas usando o algoritmo ED para que possam os leucócitos da imagem. Para 517 leucócitos, a abordagem apresentou uma taxa de detecção de 98,26%.

O trabalho apresentado por Osuna *et. al.* (2013) propôs a utilização da distância Hellinger como uma alternativa para a distância Euclidiana, a fim de estimar uma mistura de funções de Gaussianas que melhor se adequasse a um histograma de níveis de cinza de imagens de células sanguíneas. Dois métodos evolucionários (Evolução Diferencial e Colônia Artificial de Abelhas) foram utilizados para realizar a segmentação com base em informações do histograma e um estimador de distância mínima e os resultados foram comparados aos resultados obtidos pelo método clássico de Otsu. Os resultados experimentais mostraram que os três métodos são semelhante, mas os métodos evolucionários são computacionalmente mais rápidos do que o método de Otsu. Os índices de acurácia não foram especificados.

Saraswat *et. al.* (2013) introduziram um método não supervisionado com base na evolução diferencial (DE) para a segmentação de leucócitos e redução de artefatos ou ruídos em imagens de tecido de ratos coradas com a coloração hematoxilina e eosina (H&E). O método realiza um agrupamento multinível não supervisionado em duas fases. Na primeira fase utiliza a intensidade do *pixel* para extrair os leucócitos do fundo da imagem, mas não obtêm êxito na remoção dos artefatos devido a semelhança nas intensidades. Na segunda fase estes artefatos são diferenciados dos leucócitos com base em sua estrutura morfológica através da introdução de níveis de agrupamento sem supervisão utilizando um vetor de características extraído de cada objeto. Em Saraswat & Arya (2014), o método não supervisionado de segmentação baseado em DE da segunda fase foi modificado para o método supervisionado usando o algoritmo genético multiobjetivo NSGA-II, alcançando uma melhor precisão em comparação com outros métodos supervisionadas ou não. Os índices de acurácia não foram especificados.

Zheng *et. al.* (2014) apresentaram um *framework* hierárquico para a localização e segmentação de leucócitos em imagens de amostras rapidamente coradas com fundos complexos e iluminação variável. O *framework* proposto contém duas etapas principais: inicialmente o núcleo é salientado com base na diferença média absoluta, que localiza cada leucócito precisamente enquanto remove fragmentos de eritrócitos ou de coloração, e em seguida o núcleo e o citoplasma são segmentados utilizando o mapa de contraste do histograma, o algoritmo *watershed*, considerando a saliência e a similaridade da cor do núcleo, e a melhora de contraste do citoplasma. Para 54 amostras coradas com corante rápido, o índice de acurácia na segmentação foi de 91,94%.

No trabalho apresentado por Nazlibilek *et. al.* (2014) um novo sistema automático para auxiliar no diagnóstico de algumas doenças sanguíneas foi desenvolvido. Após conversão para níveis de cinza, a imagem da lâmina é limiarizada utilizando o método Otsu. Após

dilatação e preenchimento de regiões, é criada a matriz etiqueta com os componentes conectados. Após a extração de características, os componentes menores são descartados e os leucócitos são extraídos a partir da detecção de bordas e inseridos em matrizes vazias. Após a segmentação, dois classificadores diferentes, ambos utilizando *Multi-Layer Perceptron* (MLP), foram utilizados para a diferenciação leucocitária. Os índices de acurácia dos classificadores foram de 65% e 95%.

### 2.3.2 Lógica Fuzzy Aplicada a Análises de Imagens

A área de processamento de imagens e suas aplicações são valorizadas permanentemente pelo desenvolvimento de novas técnicas e pesquisas científicas que fundamentam o desenvolvimento de novas tecnologias e abordagens. Técnicas de processamento digital de imagens conhecidas que propiciam melhorias na qualidade e análise de seus elementos, como realce, filtragem, segmentação, detecção de bordas, redução de ruídos e restauração, lidam constantemente com dados imprecisos em diferentes níveis. Desde a identificação do nível de cinza de cada *pixel* que compõe a imagem até a delimitação vaga das fronteiras dos objetos componentes, as imagens digitais, que são mapeamentos de cenas naturais, estão sempre acompanhadas por certo grau de incerteza (fuzificidade). A fim de lidar com esse cenário, as técnicas *fuzzy* vêm sendo aplicadas agregando efetivamente valores à área de processamento de imagens.

Segundo Boaventura (2010),

*“As principais razões para a utilização de técnicas fuzzy como uma nova abordagem para processamento de imagem são: a formulação matemática adequada para modelar o conhecimento especialista e o gerenciamento adequado das informações vagas e ambíguas.”*

Estas razões, dentre outras não citadas, justificam a pesquisa e o desenvolvimento de sistemas inteligentes para processamento digital de imagens que utilizem o contexto da lógica *fuzzy* para várias tarefas de análise. Porém, como as técnicas *fuzzy* operam sobre os valores de pertinência, a imagem e seus componentes devem ser previamente fuzificados para que possam ser analisados neste contexto.

Ainda segundo Boaventura (2010),

*“Nas aplicações que envolvem o processamento fuzzy de imagem encontradas na literatura, distingue-se principalmente três maneiras diferentes de fuzificar uma imagem: fuzificação dos níveis de cinza com base nos histogramas das imagens, onde cada nível de cinza da imagem deve ser associado aos valores de pertinência, definidos em relação a uma determinada propriedade, a fuzificação considerando uma determinada vizinhança local dos pixels da imagem, como por exemplo para verificar a homogeneidade, e a fuzificação das características da imagem, a fim de reconhecer os objetos e interpretar as cenas.”*

Vários trabalhos e aplicações de processamento digital de imagens utilizando lógica *fuzzy* associada a outras técnicas de processamento de imagens e sistemas inteligentes são encontrados na literatura. Com focos diversificados, como agrupamento de regiões, classificação, segmentação, filtragem de ruídos e detecção de bordas, as pesquisas foram aplicadas a imagens em escalas de cinza e coloridas. Entre eles, podem ser citados os trabalhos que seguem.

Em Bonventi & Reali Costa (2000) Foi desenvolvido um método para classificar *pixels* em imagens coloridas utilizando as informações cromáticas (matiz) e acromáticas (brilho). As cores dos *pixels* foram representadas no espaço de cores HSI e sua classificação foi feita utilizando lógica *fuzzy*. O espectro de cores obtido da imagem foi dividido em conjuntos *fuzzy* e a atribuição de rótulos foi feita por regras de inferência. Partes de um sistema nebuloso genérico foram empregadas para definir a fuzificação, funções de pertinência, regras e defuzificação. Visando analisar a estabilidade do método proposto, o classificador foi testado mostrou-se bastante robusto, realizando uma classificação satisfatória para imagens com grande variação de cores. O método consistiu de três operações sobre os *pixels* da imagem: a) transformação RGB para HSI, b) atribuição de rótulos linguísticos nebulosos e c) classificação e representação dos *pixels*.

No trabalho de Figueiró (2005), a lógica *fuzzy* foi aplicada nos mapas de regiões dos três canais de cores do modelo RGB a fim de diferenciar as regiões em células, fundo e artefatos. O método conseguiu identificar as regiões, mas não houve a preocupação com a diferenciação celular.

No trabalho de Boaventura *et. al.* (2006) foi apresentada uma abordagem *fuzzy* para a classificação de tons da cor da pele em imagens coloridas. Inicialmente foi feita uma seleção de imagens que continham rostos humanos de cor de pele em diferentes tons. Um subconjunto constituído por estas imagens foi submetido à opinião de um grupo de pessoas com o objetivo de classificá-los em seus respectivos tons de cor da pele: preto, marrom e branco. Estas informações foram associadas às cores RGB e seus respectivos tons formando o conjunto *fuzzy* utilizado e as regras de inferência do sistema. Para determinar a região *fuzzy* a qual pertencia cada *pixel*, um operador *max-min* foi utilizado e três métodos de defuzificação foram testados. Após os testes, o sistema se mostrou eficiente, apresentando uma média de 70% de acerto. Entretanto, o método não foi testado para imagens que não representam os tons de pele, o que deve ser feito para verificar a sua eficiência.

O trabalho de Li *et. al.* (2008) apresenta uma abordagem para filtragem de ruídos em imagens coloridas utilizando um detector por impulso e uma rede neuro-*fuzzy* robusta. O modelo consiste em aplicar um detector de impulso a uma janela da imagem e, considerando seus valores máximos e mínimos, localizar os *pixels* com ruídos. Em seguida, se o *pixel* for detectado como ruidoso, a rede neuro-*fuzzy* o substitui. Os testes mostraram que o filtro proposto no modelo é eficiente para atenuação de ruídos e preservação de detalhes tanto quantitativa quanto qualitativamente quando comparado com técnicas convencionais.

Em Jiji & Ganesan (2011) considerou-se que a textura é uma característica importante e especialmente útil para a identificação de objetos ou regiões de interesse em uma imagem. Considerando que em imagens de textura coloridas pode haver variações de direção, granularidade e outras características, o trabalho apresentou duas abordagens baseadas em lógica *fuzzy* para identificação de texturas em imagens coloridas: unidades de textura *fuzzy*, a partir da decomposição da imagem em um conjunto de pequenas unidades essenciais, e espectro de textura *fuzzy*, ou seja, frequência de distribuição de todas as unidades de textura. O método foi proposto para segmentação de imagens usando descritores de textura colorida. Primeiramente, a presença da textura colorida é representada local e globalmente. A seguir, com o descritor global proposto, a operação de segmentação é realizada baseando-se no agrupamento não supervisionado por meio de um mapa de Kohonen auto-organizável. O algoritmo é implementado por meio de uma rede neural. O método proposto extrai as características da textura na imagem e a operação de segmentação é realizada dividindo a imagem em regiões de cor. O êxito da abordagem descrita baseia-se na existência de texturas contendo segmentos alongados ou naturais como regiões de interesse.

Com uma proposta híbrida, Tan & Isa (2011) apresentaram uma técnica de limiarização do histograma e algoritmo *fuzzy c-means* para melhorar a formação e uniformidade dos agrupamentos de regiões em imagens coloridas. A técnica desenvolvida subdivide-se em dois módulos: o módulo de limiarização do histograma, que é usado para obter a condição de inicialização do centro do agrupamento, e o módulo *fuzzy c-means*, que é utilizado para melhorar a compactação dos agrupamentos. Resultados experimentais demonstraram que a baixa complexidade da abordagem proposta oferece uma melhor qualidade na formação de regiões e segmentação do que outras abordagens conhecidas, podendo ser utilizada em reconhecimento de padrões, visão computacional e segmentação de imagens coloridas.

O trabalho proposto por Lin *et. al.* (2011) apresenta um sistema de inferência *fuzzy* para determinar se uma colônia de células cancerígenas contém ou não mais de 50 células. Um scanner foi utilizado pra capturar a imagem da cultura, analisada posteriormente utilizando-se técnicas de processamento digital, como limiarização e filtro Sobel, e a transformada *Hough*, rotacionando a imagem para o ângulo mais apropriado. Em seguida, o modelo utilizou as características de cor e de área da superfície em conjunto com o sistema de inferência *fuzzy* proposto para identificar a colônia de células cancerígenas e determinar se a mesma possuía mais de 50 células. Os resultados experimentais mostram uma correlação muito forte entre as resultados produzidos pelo sistema e as revelados por contagem manual.

O trabalho de Nawgaje & Kanphade (2011) apresenta um sistema de inferência *fuzzy* para a detecção de bordas em imagens microscópicas coloridas de medula óssea. Robusto em relação às condições de variação de luminosidade e considerando a cor dos elementos, para cada *pixel* da imagem verifica o grau de similaridade entre sua cor e as cores do sistema e se o *pixel* pertence ou não à borda. Após o destaque das bordas, a imagem é submetida a outro conjunto de condições, que extrai as partes indesejadas, gerando uma imagem contendo apenas as bordas identificadas.

Objetivando também a detecção de bordas em imagens, Fan & Wang (2011) propuseram um algoritmo multidirecional baseado em morfologia *fuzzy* a fim de lidar com problemas de borrão e imprecisão na localização das fronteiras entre os elementos. No algoritmo, dois limiares foram selecionados para realizar a segmentação e obtenção de bordas. A morfologia *fuzzy* foi adotada para diminuir a perda de informação e elementos estruturais multidirecionais foram usados para detectar as bordas da imagem. Experiências foram realizadas e os resultados mostraram que o algoritmo proposto possui uma capacidade de diminuição de ruído superior aos algoritmos tradicionais.

O trabalho desenvolvido em Wang *et. al.* (2012) apresentou um método de segmentação de *pixels* em imagens coloridas utilizando Máquinas de Vetor de Suporte e o algoritmo *fuzzy c-means*. Utilizando as características de nível de cor e textura do *pixel* como entrada do classificador, foram extraídos modelos de medida de similaridade espacial local. Com tais valores o classificador foi treinado e a imagem colorida foi segmentada. Esta segmentação utilizou ao máximo a informação das cores da imagem e a capacidade do classificador. Testes experimentais demonstraram que o método proposto tem um comportamento eficaz, diminuindo o tempo e aumentando a qualidade de segmentação de imagem coloridas.

### **2.3.3 Segmentação de Componentes Sanguíneos e Classificação Diferencial de Leucócitos utilizando Lógica Fuzzy**

Muitos trabalhos apresentam a utilização da lógica *fuzzy* aplicada à segmentação dos componentes sanguíneos e classificação diferencial de leucócitos. Esta perspectiva mostra-se promissora pela possibilidade de trabalhar em um universo de incertezas e de regras construídas a partir de uma visão especializada.

Uma abordagem baseada em lógica *fuzzy* para segmentação dos núcleos de leucócitos da medula óssea foi desenvolvida por Theera-Umpon (2005), considerando na análise não os *pixels* individualmente, mas um grupo de *pixels* conectados. A proposta segmenta a imagem em duas regiões, núcleo e não-núcleo, utilizando um classificador *fuzzy c-means* e morfologia matemática. O classificador fornece a segmentação gerando vários agrupamentos que são então combinados para formar duas regiões: núcleo e não-núcleo. Os operadores morfológicos de abertura e fechamento são aplicados para a redução de ruídos e melhoria da imagem. Um sistema de classificação, baseado no teorema de Bayes, determina estatisticamente se a célula segmentada é um leucócito. O algoritmo foi utilizado em 431 imagens e os resultados foram comparados com a segmentação manual das imagens por um especialista, mostrando-se 10% melhores para as células consideradas maduras, por apresentarem núcleos mais escuros, enquanto que para células imaturas o erro chegou a 16%.

Em Shitong & Min (2006) foi proposto um novo algoritmo de detecção baseado em uma rede neural *fuzzy* combinando as vantagens da segmentação por limiar e da morfologia matemática com a lógica *fuzzy*. Foi demonstrado que algumas partes dos leucócitos são perdidas durante as operações de abertura e também discutiram os problemas relacionados com a existência de outros componentes e suas sobreposições nas imagens microscópicas de

sangue. Embora tenha apresentado taxas de detecção altas, o método proposto mostrou limitações quanto a manter a integridade do limite do leucócito. Uma versão melhorada do algoritmo foi proposta em Shitong *et. al.* (2007), modificando a equação de estado contendo informações úteis além dos leucócitos e aumentando a integridade do limite da célula. Os índices de acurácia não foram apresentados.

O trabalho apresentado por Ongun *et. al.* (2008) apresenta uma segmentação do núcleo e do citoplasma de leucócitos. Inicialmente, a segmentação é realizada utilizando o algoritmo *fuzzy c-means* modificado como uma técnica de agrupamento para classificar a imagem original em quatro partes: núcleo do leucócito, citoplasma do leucócito, plasma e eritrócitos. Em seguida, o algoritmo foi implementado recursivamente até que fossem eliminadas falsas dispersões ou concentrações devido à ausência de núcleo ou similaridade entre as cores dos *pixels* do citoplasma e do plasma. A minimização em cada iteração foi realizada usando uma cor de *pixel* da vizinhança como uma referência. A saída final mostrou que o método modificado é capaz de extrair as regiões de núcleo e citoplasma de uma forma mais eficiente do que utilizando o método *fuzzy c-means* normal. Os índices de acurácia não foram apresentados.

Em Ghosh *et. al.* (2010) foi proposto um método para segmentação e classificação de leucócitos utilizando divergência *fuzzy* e técnicas de limiarização modificadas. Para segmentação dos leucócitos foi utilizado o algoritmo *watershed* controlado por marcador associado a um operador morfológico. Em seguida, oito descritores geométricos baseados no núcleo leucocitário e um baseado no leucócito foram computados matematicamente e analisados estatisticamente visando diferenciar os cinco grupos leucocitários. No reconhecimento foram estudados os tipos de funções de pertinência *fuzzy Gamma*, Gaussiano e *Cauchy*, sendo este último o que apresentou o melhor resultado. Além deste, Ghosh *et. al.* (2011) apresentaram uma nova abordagem para a medida de divergência com base na extensão *fuzzy* da função de entropia *Reny*. A divergência *fuzzy* utilizando a entropia de *Reny* mostrou uma melhor segmentação dos núcleos leucocitários em imagens de leucemia do que o algoritmo *fuzzy c-means* e o método *Otsu*. A acurácia apresentada para a classificação foi de 83,2%.

Um estudo apresentado por Mohapatra *et. al.* (2011) apresentou o agrupamento, unificação dos segmentos de uma imagem em regiões de interesse, como um procedimento essencial para a correta segmentação de imagem e propôs a integração criteriosa de conjuntos aproximados e conjuntos difusos para o agrupamento e segmentação de leucócitos. Neste estudo, os conjuntos *fuzzy* e aproximados foram adequadamente integrados para alcançar um

melhor desempenho de segmentação. O conceito de associação de conjuntos *fuzzy* mostrou-se eficiente para a manipulação de sobreposição, enquanto que os conjuntos aproximados apresentaram-se como uma solução razoável para lidar com dados incertos, incompletos e imprecisos. Esta combinação sinérgica forneceu ao esquema proposto vantagens sobre outras técnicas conhecidas de agrupamento baseados em segmentação quando da realização de testes em imagens de células. Observou-se que as regiões do citoplasma e núcleo dos leucócitos foram corretamente segmentadas e a análise comparativa revelou que a técnica desenvolvida é robusta para segmentação de imagens sanguíneas microscópicas coradas. Neste estudo foram analisadas 108 amostras coradas pelo método Leishman.

Fatima & Seenivasagam (2011) introduziram um método rápido para segmentação de núcleos celulares em imagens microscópicas fluorescentes utilizando o *fuzzy c-means* seguido por descritores de forma, operações morfológicas e pelo método *watershed*. Foram encontrados índices de acurácia entre 96,45% e 97,36%.

Sansone *et. al.* (2012) propuseram um algoritmo para análise celular subdividido em duas fases: identificação via pré-filtragem Gaussiana e operadores morfológicos e segmentação utilizando agrupamentos *fuzzy*. Os índices de acurácia não foram apresentados.

Um novo esquema de ordenação baseado em interesse por morfologia *fuzzy* em segmentação de imagens de leucócitos sanguíneos, apresentado por Faticahah *et. al.* (2012), propôs a melhoria na precisão da segmentação de núcleos de alta e baixa densidade. No método, dois algoritmos foram desenvolvidos para a segmentação do núcleo e do citoplasma. A morfologia *fuzzy* foi utilizada para a segmentação do núcleo, enquanto que a morfologia matemática binária, utilizando um elemento de estruturação com base no tamanho das células vermelhas, foi utilizada para segmentação do citoplasma. Inicialmente uma imagem de entrada contendo a cor do núcleo é fornecida ao sistema. O modelo de cores HSV é então criado utilizando a imagem de entrada e o elemento estruturante. Após a aplicação de processos de dilatação e erosão *fuzzy*, um processo de abertura *fuzzy* descarta resíduos e segmenta o núcleo do leucócito. Na segmentação do citoplasma, uma imagem é criada a partir da limiarização da imagem original e, aplicando o elemento estruturante, o citoplasma é segmentado. Para avaliar o desempenho do método, 100 amostras de imagens microscópicas de sangue e 10 amostras de imagem microscópica de leucemia foram utilizados. Os resultados de segmentação do método proposto (93,15% para núcleo e 84,43% para citoplasma leucocitário), após comparados com as imagens manualmente segmentadas, mostraram-se superiores quando comparados com outros métodos conhecidos.

Jati *et. al.* (2014) propuseram uma abordagem automática para a segmentação de núcleos leucocitários utilizando uma técnica de limiarização baseada em divergência *fuzzy*. O método proposto minimiza a divergência entre a imagem atual e a ideal a partir de uma nova fórmula de divergência exponencial baseada em entropia *fuzzy*. A fim de lidar com imagens ruidosas, o método define ainda uma função de pertinência baseada na vizinhança do *pixel* analisado. Foram analisadas 110 imagens coradas pelo método Leishman e encontrados índices de acurácia entre 93,90% e 94,93% para imagens com ruído.

O trabalho apresentado por Chaira (2014) propôs um esquema para a segmentação automática de leucócitos em imagens patológicas (células normais e anormais) utilizando as teorias *fuzzy* intuição e intervalar tipo II, preservando a forma dos leucócitos. Foram analisadas 100 amostras de células normais e anormais, mas os índices de acurácia não foram apresentados.

Li *et. al.* (2014) propuseram um sistema de imagem hiperespectral molecular baseado em filtro óptico-acústico sintonizável para observar as manchas de sangue. Uma combinação de algoritmos espaciais e espectrais foi proposta para identificar o citoplasma e núcleo dos leucócitos, integrando os algoritmos *fuzzy c-means* e *K-means* espacial. Em seguida, os parâmetros morfológicos, tais como a área do citoplasma, a área do núcleo, o perímetro, a proporção nuclear, a forma e a solidez foram calculados e avaliados. Os resultados experimentais mostram que o algoritmo proposto tem melhor desempenho que o algoritmo baseado espectral como o novo algoritmo pode usar em conjunto a informação espacial e espectral de células de leucócitos. Foram analisadas 80 amostras coradas pelo método Giemsa.

Todavia, mesmo considerando os resultados positivos obtidos, algumas limitações foram observadas nos trabalhos mencionados: i) As amostras sanguíneas seguem o mesmo padrão de coloração; ii) Segmentação de alguns componentes sanguíneos, em detrimento de outros; iii) Consideráveis taxas de Falso Positivo em decorrência da existência de vizinhança entre as células; iv) Retiradas de Falsos Negativos baseados em áreas com tamanhos pré-definidos; v) Abordagens semi-automáticas para segmentação com limiares pré-determinados para a segmentação ou descarte; vi) Classificação de alguns tipos de leucócitos, desconsiderando outros; vii) Utilização de grande quantidade de descritores, aumentando o custo de processamento; viii) Falta de clareza na definição comportamental dos tipos leucocitários quanto aos descritores que os identificam.

As técnicas desenvolvidas neste trabalho apresentam um esquema automático para segmentação e classificação para leucócitos (núcleos e citoplasma), eritrócitos e plasma

sanguíneos em imagens microscópicas. Usando apenas o canal G da imagem, os três picos mais significativos do histograma são consideradas para o conceito de *tonalidade*. No entanto, considerando que as imagens de lâminas microscópicas apresentam quatro regiões de interesse (núcleo e citoplasma leucocitários, eritrócitos e plasma sanguíneo) e o histograma contém apenas três regiões bem definidas, os métodos tradicionais de segmentação, que utilizam a quantificação e a separação das regiões com base em níveis de cinza, não são suficientes para a correta segmentação de citoplasma leucocitário e eritrócitos. Além de *tonalidade*, o conceito de *proximidade* entre os *pixels* da imagem e o centroide do núcleo leucocitário mais próximo a cada um deles foi desenvolvido objetivando refinar a classificação. Para expressar adequadamente esses conceitos incertos e as relações entre eles, este trabalho apresenta uma abordagem baseada em Sistemas de Inferência *Fuzzy*. A combinação da lógica *fuzzy* com as técnicas de processamento digital de imagens possibilitou um melhor refinamento do processo de segmentação e classificação.

Com o objetivo de abordar as limitações encontradas na literatura, os seguintes aspectos foram considerados: i) Consideração das informações de cor somente nas fases de pré e pós-processamento, e dos níveis de cinza do canal G para todas as amostras na fase de processamento, mesmo que sejam provenientes de vários processos de coloração; ii) Segmentação e classificação de variados componentes sanguíneos, como leucócitos (núcleos e citoplasma), eritrócitos e plasma; iii) Cálculo automático dos suportes dos conjuntos *fuzzy* para as variáveis *tonalidade* e *proximidade*; iv) Identificação de quatro áreas distintas em imagens cujos histogramas tem apenas três regiões bem definidas; v) Utilização da distância euclidiana entre os núcleos leucocitários e os demais *pixels* da imagem, mitigando o problema da vizinhança entre citoplasmas leucocitários e eritrócitos; vi) Redução de falsos negativos para citoplasma leucocitários incorretamente classificados como eritrócitos; vii) Automatização do processo de segmentação; viii) Definição de novos descritores que facilitem a classificação diferencial dos leucócitos; ix) Definição de uma base de imagens cujas amostras sejam provenientes de técnicas de coloração hematológica variadas.

A adaptação das funções de pertinência a cada imagem e a utilização de aspectos como os picos do histograma e as distâncias entre as regiões como suportes dos conjuntos *fuzzy* caracterizam a técnica desenvolvida como automática e robusta, distinguindo-a de outras abordagens.

---

# Capítulo 3

## Processamento Digital de Imagens e Lógica Fuzzy

---

### 3.1 Introdução

Neste capítulo são apresentados conceitos e técnicas de processamento digital de imagens e lógica *fuzzy* relevantes a este trabalho. Para tal, subdivide-se em:

- **Imagens Digitais e Processamento Digital de Imagens:** apresenta os conceitos iniciais sobre imagens digitais, conceitos principais relativos ao processamento digital de imagens e as etapas constituintes deste, descritores de imagens, segmentação e classificação.
- **Lógica Fuzzy:** apresenta os conceitos da lógica *fuzzy* relevantes para este trabalho, como variáveis linguísticas e funções de inferência, e ainda o sistema de inferência *fuzzy* e seus subsistemas.

### 3.2 Imagens Digitais

As imagens, produzidas por uma variedade de dispositivos físicos de captura, tais como câmeras, microscópios, equipamentos de radiografia, radares, equipamentos de ultrassonografia, entre outros, são utilizadas para os mais diversos fins, que vão desde o entretenimento até aplicações específicas, como médicas, tecnológicas e militares. O objetivo principal das análises manual e automática de imagens é extrair informações relevantes para aquela aplicação em foco.

Segundo Gonzalez & Woods (2009),

*“Uma imagem pode ser definida como uma função bidimensional  $f(x,y)$ , em que  $x$  e  $y$  são coordenadas espaciais (plano) e a amplitude de  $f$  em qualquer par de coordenadas  $(x,y)$  é chamada de intensidade ou nível de*

*cinza da imagem nesse ponto. Quando  $x$ ,  $y$  e os valores de intensidade de  $f$  são quantidades finitas e discretas, tal imagem é considerada imagem digital, composta de um número finito de elementos com localização e valor específicos chamados pixels (picture elements)."*

Uma imagem resultante da função bidimensional  $f(x,y)$  descrita acima é usualmente conhecida como imagem em escala ou níveis de cinza, na qual o valor de cada *pixel* é uma única amostra de um espaço de cores que varia desde o preto, menor intensidade, até o branco, maior intensidade (Figura 8).



Figura 8. Convenção dos eixos para representação de imagens digitais (Gonzalez & Woods, 2000).

Imagens em escala de cinza, que podem conter diversos tons em sua composição, são diferentes de imagens binárias em preto e branco, que possuem apenas duas cores. Computacionalmente, as imagens em escala de cinza são armazenadas utilizando-se 8 bits por *pixel*, o que permite 256 intensidades possíveis. As imagens coloridas, por sua vez, apresentam um maior nível de detalhes e um aumento na percepção de suas características pelo olho humano.

Dentre as características de uma imagem, a cor é um poderoso fator que muitas vezes simplifica a identificação de um objeto e sua extração em uma cena. Este é um motivo relevante para a utilização de cor no processamento digital de imagens. Além deste, as imagens coloridas são particularmente importantes nas análises manuais, isto é, realizadas por

seres humanos, visto que estes são capazes de discernir milhares de tons e intensidades de cor, em comparação com apenas duas dúzias de níveis de cinza (Gonzalez & Woods 2009).

Uma imagem colorida é uma imagem multibanda, onde a cor em cada ponto  $(x,y)$  pode ser definida através de três grandezas: luminância, matiz e saturação. A luminância está associada ao brilho da luz, a matiz com o comprimento de onda dominante e a saturação com o grau de pureza (ou intensidade) da matiz. A maioria das cores visíveis pelo olho humano pode ser representada como uma combinação das três cores primárias: vermelho, verde e azul.

Os modelos de cores padronizam a especificação de cores em uma forma amplamente aceita. Em síntese, um modelo é uma especificação de um sistema de coordenadas e um subespaço dentro deste sistema no qual cada cor é representada por um único ponto. São utilizados para classificar as cores e para qualificá-las de acordo com atributos como os citados acima (luminância, matiz e saturação). Dentre os modelos de cores mais utilizados está o modelo RGB ( $R - red$  ou vermelho,  $G - green$  ou verde,  $B - blue$  ou azul).

No modelo de cores RGB cada cor aparece em seus componentes espectrais primários de vermelho, verde e azul, distribuídas em um sistema de coordenadas cartesianas cujo subespaço de interesse é o cubo (Figura 9). Nele, os valores RGB primários estão em três vértices, as cores secundárias, ciano, magenta e amarelo, em outros três vértices o preto está na origem e o branco no vértice mais distante deste. A diagonal entre o preto e o branco é a escala de cinza.

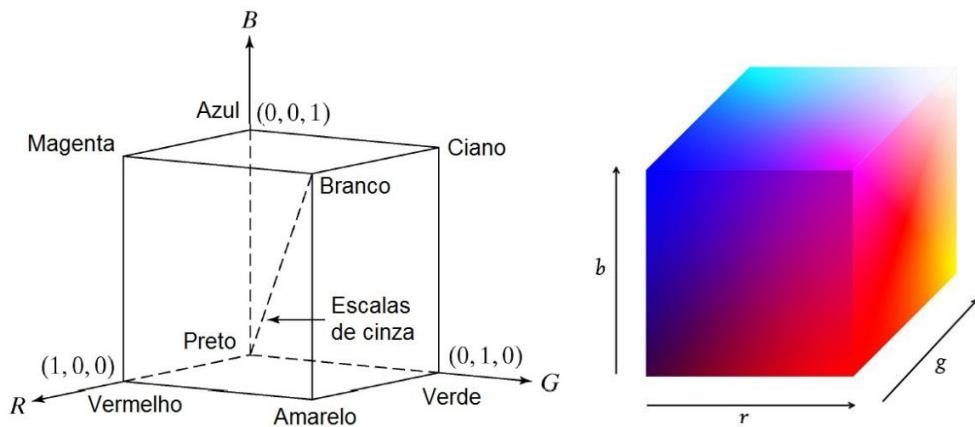
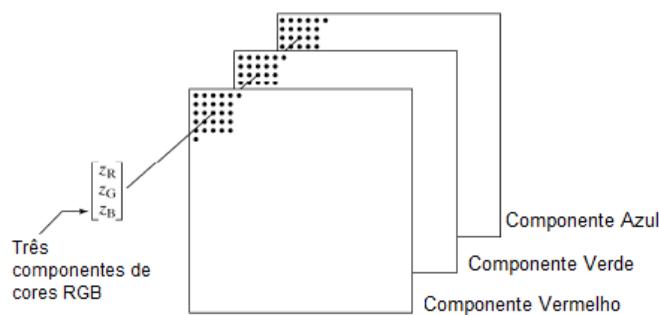


Figura 9. Representações do Modelo RGB (Gonzalez & Woods, 2009).

As imagens representadas no espaço RGB consistem em arranjos de  $x \times y \times 3$  pixels, onde cada pixel é uma tripla correspondente às cores vermelho, verde e azul em uma localização espacial específica (Figura 10).



*Figura 10. Máscaras espaciais para imagens RGB (Gonzalez & Woods, 2009).*

Na Figura 11 podem ser vistas diversas representações de uma mesma imagem: em cores (modelo RGB e seus canais individuais), escalas de cinza e preto e branco.



*Figura 11. Representações de uma imagem: a) cores (modelo RGB), b) escalas de cinza, c) preto e branco, d) componente R, e) componente G, f) componente B. Imagem: Lena.jpg (Gonzalez & Woods, 2000).*

Neste trabalho serão utilizadas as imagens representadas no modelo de cores RGB e o canal G individualmente.

### **3.3 Processamento Digital de Imagens**

Em geral, a imagem pura necessita de transformações que realcem seu conteúdo e propiciem uma extração de informações mais detalhadas e precisas. O Processamento Digital de Imagens realiza estas transformações efetuando a conversão das imagens recém-capturadas em matrizes numéricas e processando-as, utilizando um computador digital para este fim.

Segundo Gonzalez & Woods (2009),

*“O interesse nos métodos de processamento digital de imagens provém de duas áreas principais de aplicação: melhora das informações visuais para a interpretação humana e processamento de dados de imagens para armazenamento, transmissão e representação, considerando a percepção automática por máquinas.”*

Ainda segundo Gonzalez & Woods (2009),

*“A área de processamento digital de imagens envolve processos cujas entradas e saídas são imagens e, além disso, envolve processos de extração de atributos de imagens até – e inclusive – o reconhecimento de objetos individuais.”*

As etapas fundamentais de um sistema de processamento de imagens podem ser definidas de acordo com os seguintes passos: domínio do problema, aquisição da imagem, pré-processamento, segmentação, representação e descrição, reconhecimento e interpretação e o resultado (Figura 12). Durante o processamento da imagem, uma base de conhecimento recebe as informações relevantes que foram extraídas da imagem, podendo ser consultada a qualquer momento.

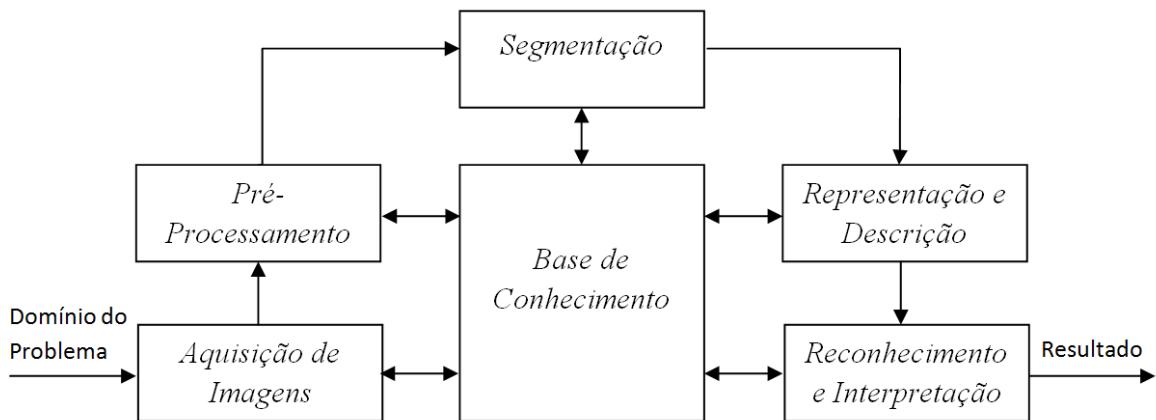


Figura 12. Etapas do Processamento Digital de Imagens (Gonzalez & Woods, 2000)

### 3.3.1 Aquisição de imagens

A fase de *aquisição de imagens* consiste em obter uma representação da informação visual a partir de dispositivos físicos sensíveis a espectros de energia eletromagnética que convertem o sinal elétrico para um formato digital. São exemplos de dispositivos destinados à aquisição de imagens: câmeras digitais, microscópios, tomógrafos, scanners, entre outros.

Neste trabalho serão utilizadas imagens de lâminas sanguíneas pré-fixadas e coradas em lâminas de extensão, cujas aquisições dar-se-ão através da utilização de microscópios ópticos com um fator de aumento de 100x. Desta forma, os componentes sanguíneos poderão ser visualizados e analisados.

### 3.3.2 Pré-processamento

O *pré-processamento* consiste no realce da imagem para enfatizar características de interesse ou recuperar imagens que sofreram alguma perda. Utilizando transformações lineares e não-lineares aplicadas à imagem, as técnicas de pré-processamento têm a função de melhorar a qualidade da imagem e envolvem métodos que operam no domínio espacial, manipulando diretamente os *pixels* da imagem na sua forma original, e no domínio da frequência, manipulando o espectro da imagem. São exemplos de operações de pré-processamento: melhoramento de contraste, delimitação de regiões de interesse, remoção de ruído, reamostragem dos *pixels* em outra escala, redimensionamento, entre outras.

Neste trabalho o pré-processamento será responsável por realizar o tratamento inicial da imagem da lâmina sanguínea, descartando a informação da coloração utilizada e calculando

os valores que serão utilizados durante a etapa de segmentação como suportes dos conjuntos *fuzzy*.

### 3.3.3 Segmentação

A segmentação é a extração ou identificação dos objetos contidos na imagem, separando a imagem em regiões. Em outras palavras, a segmentação refere-se à divisão da imagem em regiões de interesse objetivando possibilitar a análise de forma independente. No agrupamento de regiões são consideradas as características ou propriedades semelhantes entre os *pixels*, como intensidade e textura, ou a perda de uma característica específica para que as regiões segmentadas possuam aspectos que as diferenciem das demais.

Para a escolha do método de segmentação que deve ser utilizado em uma imagem específica, as informações da imagem devem ser previamente consideradas. As técnicas de segmentação podem ser baseadas em similaridades, como limiarização, crescimento de regiões, junção e separação e aglomeração, ou descontinuidades, como detecção de pontos, retas e bordas. A precisão da fase de segmentação poderá determinar o sucesso ou a falha dos procedimentos de análise de imagem.

Na segmentação de componentes sanguíneos em imagens microscópicas, objetiva-se principalmente extrair os eritrócitos e leucócitos da imagem, separando-os da área do plasma sanguíneo. Podem ser encontrados métodos de segmentação celular baseados em limiarização utilizando crescimento de regiões, watershed e Otsu (Wermser *et. al.*, 1984, Liao & Deng, 2002 e Malpica *et. al.*, 1997), baseados no reconhecimento de padrões utilizando máquinas de vetor des, redes neurais, k-means, *fuzzy c-means* e maximização de expectativas (Guo *et. al.*, 2007), baseados em modelos deformáveis utilizando contornos ativos (Glotzos *et. al.*, 2004) e baseados em meta-heurísticas utilizando evolução diferencial, colônia artificial de abelhas e algoritmos genéticos (Nakib *et. al.*, 2007 e Osowski *et. al.*, 2009).

Neste trabalho a etapa de segmentação consiste em um sistema de inferência *fuzzy* responsável por dividir a imagem em quatro regiões contendo os diferentes componentes sanguíneos: plasma sanguíneo, eritrócitos, núcleos e citoplasmas leucocitários.

### 3.3.4 Representação e Descrição

A saída do estado de segmentação é constituída principalmente por dados em forma de *pixels*. Desta forma, faz-se necessário converter os dados para uma forma adequada ao

processamento computacional, realizando assim a representação mais apropriada para posterior descrição dos dados. Para o reconhecimento do objeto é necessário descrever as propriedades das regiões segmentadas (grupos de *pixels*). A descrição é muitas vezes apenas um conjunto de dados que são chamados de descritores do objeto. Um método para descrever os dados também deve ser especificado, de forma que as características de interesse também sejam enfatizadas. Este processo, também chamado de seleção de características, procura extrair propriedades que resultem em alguma informação quantitativa de interesse ou que sejam básicas para a discriminação dos objetos.

As propriedades que representam os objetos podem ser externas, escolhidas quando a atenção primária está voltada para a forma da região, ou internas, escolhidas quando a análise se concentra nas características da própria região, como cor ou textura. A medida de qualquer uma destas propriedades é denominada de *característica* ou *descritor da imagem*. Estas características podem formar um vetor de escalares, denominado *Vetor de Descritor da Imagem* ou *Vetor de Características*, e os objetos podem ser reconhecidos comparando-se simplesmente os seus descritores com os descritores de objetos conhecidos.

Os descritores dos objetos devem apresentar quatro propriedades importantes:

- Devem definir um conjunto completo, ou seja, Dois objetos devem ter os mesmos descritores se e somente se eles tiverem as mesmas características;
- Devem ser congruentes, ou seja, deve-se reconhecer objetos como similares quando estes objetos possuírem descritores semelhantes;
- Devem ter propriedades invariantes, ou seja, deve ser possível reconhecer objetos independente de rotação, escala, posição ou perspectiva;
- Devem ser um conjunto compacto, ou seja, devem representar a essência de um objeto de maneira eficiente.

Na representação e descrição de células as suas principais características são utilizadas para a sua classificação. A maioria dos descritores utilizados com esta finalidade concentra-se em descritores geométricos, como área, raio, convexidade, solidez, perímetro, orientação, circularidade, entre outros, e de textura, como média, variância, entropia, energia, homogeneidade e correlação (Rezatofighi & Soltanian-Zadeh, 2011 e Sabino *et. al.*, 2004).

Neste trabalho são utilizados descritores para a classificação diferencial dos leucócitos em cinco tipos leucocitários: basófilos, eosinófilos, linfócitos, monócitos e neutrófilos.

### 3.3.5 Reconhecimento e interpretação

O último estágio, que envolve reconhecimento e interpretação, atribui um rótulo a um objeto baseado na informação fornecida pelo seu descritor. A interpretação, por sua vez, envolve a atribuição de significado a um conjunto de entidades rotuladas. Assim, dado um conjunto de características relativas a uma região (objeto), pode-se atribuir uma classe essa região selecionada a partir de um conjunto de classes cujas características são conhecidas.

Neste trabalho, os sistemas de inferência *fuzzy* são os responsáveis por, a partir da base de conhecimento, classificar cada um dos componentes sanguíneos e, além disso, classificar diferencialmente os leucócitos.

### 3.3.6 Base de Conhecimento

A base de conhecimento agrega o conjunto de conhecimentos específicos a respeito do domínio do problema. Esse conhecimento pode ser tão simples quanto o detalhamento de regiões de uma imagem na qual pode ser encontrado determinado aspecto relevante, limitando a busca, ou tão complexo como uma lista de possíveis defeitos interconectados que devem ser verificados em uma imagem de um determinado material, por exemplo. Neste trabalho, a base de conhecimento utilizada consiste nas bases de regras *fuzzy* inseridas nos sistemas de inferência para classificação dos componentes sanguíneos.

## 3.4 Lógica Fuzzy

Proposta por Zadeh (1973), a lógica *fuzzy* (ou lógica nebulosa) é uma teoria que se propõe a expressar matematicamente as formulações do pensamento humano em linguagem natural sem, contudo, diminuir o poder expressivo das mesmas. Definida como um novo caminho para representar a incerteza, a lógica *fuzzy* permite a representação de conceitos vagos, expressos em linguagem natural, tais como, pequeno, quente, bom, ruim, dentre outros.

Segundo Zadeh (1973):

*“Os seres humanos raramente usam números para resolver problemas. Assim, à medida que a complexidade do sistema aumenta, a habilidade para tornar as proposições precisas diminui até um limiar que está fora do alcance.”*

Como a capacidade de descrição de um modelo matemático para resolução de um problema decresce à medida que o grau de incerteza do mesmo aumenta, fez-se necessário o surgimento de uma teoria que conseguisse tratar tais problemas sem que informações importantes se perdessem durante a manipulação de seus dados (Chiu & Park, 1994).

Neste contexto e segundo Cox (1995), a teoria *fuzzy* é definida como sendo capaz de combinar a imprecisão associada aos eventos naturais e o poder computacional das máquinas para produzir sistemas de resposta inteligentes, robustos e flexíveis.

A lógica *fuzzy* proporciona uma linguagem natural, onde predomina o raciocínio aproximado com proposições imprecisas, utilizando a teoria de conjuntos nebulosos como a principal ferramenta, sendo análoga ao papel da lógica de predicado, utilizada para raciocínio com proposições precisas (Klir & Folger, 1988).

Segundo Lee (1990a):

*“A lógica fuzzy é muito mais próxima do pensamento humano e da linguagem natural do que a lógica tradicional. Basicamente, propicia um significado efetivo na captura de aproximações e informações inexatas do mundo real.”*

A teoria de conjuntos *fuzzy* foi introduzida com o objetivo principal de dar um tratamento matemático a conceitos vagos e subjetivos existentes na comunicação humana. Termos linguísticos subjetivos como “aproximadamente”, “em torno de”, dentre outros, definem conceitos imprecisos que não podem ser tratados adequadamente com os conjuntos convencionais. Formalmente, a Teoria de Conjuntos *Fuzzy* foi concebida como uma generalização da Teoria Convencional dos Conjuntos, fornecendo a instrumentação básica necessária ao estender a definição de operações como pertinência de um elemento, união e intersecção, complemento, leis distributivas, operações algébricas entre conjuntos, entre outros (Boaventura, 2010).

À lógica *fuzzy* cabe representação dos conceitos como palavras, e aos conjuntos *fuzzy* cabe expressar os valores das mesmas. Assim, o conceito de precisão ou imprecisão, dependendo do contexto aplicado, expressar-se-á numericamente indicando a possibilidade, e não a probabilidade, de tal inferência estar ou não correta.

### 3.4.1 Os Conjuntos Fuzzy

Nos conjuntos tradicionais (*crisp*), o conceito de pertinência de um elemento a um conjunto é bem definido. Dado um conjunto  $A$  em um universo  $U$ , os elementos deste universo simplesmente pertencem ou não pertencem àquele conjunto, assumindo apenas um dos dois valores {0, 1}. Isto pode ser expresso pela função característica  $f_A$ , conforme expressão 1.

$$f_A = \begin{cases} 1 & \text{se e somente se } x \in A \\ 0 & \text{se e somente se } x \notin A \end{cases} \quad (1)$$

Os conjuntos *fuzzy*, ao contrário dos conjuntos tradicionais, são capazes de incorporar tanto o conhecimento objetivo, a partir de dados numéricos, quanto o conhecimento subjetivo, a partir de informações imprecisas, suportando assim os modos de raciocínio aproximados e respeitando critérios subjetivos. Vistos como uma generalização dos primeiros, os conjuntos *fuzzy* implementam funções a fim de representar seus elementos, os quais assumem valores reais dentro do intervalo [0,1]. A ideia do conjunto *fuzzy*, portanto, está associada a um grau de pertinência  $\mu$ , que varia de acordo com o elemento em questão. Desta forma, segundo a variação do seu grau de pertinência (da completa exclusão até a total pertinência), um elemento pode pertencer muito ou pouco e até não pertencer a um conjunto *fuzzy*, podendo assumir qualquer um dos valores intermediários (Mattos *et. al.*, 2004).

Em outras palavras, um conjunto *fuzzy* pode então ser definido matematicamente visto que cada possível elemento no conjunto universo possui um valor numérico que representa seu  $\mu$  correspondendo ao grau para o qual aquele elemento é semelhante ou compatível com o conceito representado pelo conjunto (Klir & Folger, 1988).

Assim:

- se  $\mu = 1$ , o elemento pertence ao conjunto, como em um conjunto tradicional;
- se  $\mu = 0$ , o elemento não pertence ao conjunto;
- se  $\mu$  estiver entre o 0 e 1, o elemento pertence ao conjunto com certo grau de pertinência.

### 3.4.2 Conceitos Básicos em Conjuntos Fuzzy

Define-se como *universo de discurso*  $U$  o conjunto genérico de objetos onde a função de pertinência de um elemento  $x$  a um conjunto *fuzzy* é definida. Um universo de discurso pode ser contínuo, quando seus elementos são infinitos, e pode ser discreto, com um número finito de elementos (Lee, 1990b).

Tome-se o exemplo “altura das pessoas”. O universo de discurso, neste caso, é contínuo. Entretanto, no exemplo “posição de uma peça dentro de uma célula de manufatura” tem-se um universo discreto, já que cada peça pode estar apenas em máquinas específicas. Mas mesmo o universo contínuo pode ser representado de maneira discreta a partir da subdivisão de seus valores em um número finito de pontos, sendo esta a melhor forma para simplificar a representação em um computador digital (Lee, 1990b).

Como interfere na definição dos conjuntos *fuzzy*, a representação de  $U$  deve ser previamente definida. Para universos de discurso contínuos, tem-se uma função contínua relacionando o elemento com seu grau de pertinência; para discretizados, são estabelecidos graus de pertinência para cada segmento.

Define-se então um conjunto *fuzzy*  $A$  em um universo  $U$  por uma função de pertinência  $\mu_A(x) = U \rightarrow [0,1]$ , representado por um conjunto de pares ordenados, conforme expressão 2:

$$A = \{x, \mu_A(x)\} \quad x \in U \quad (2)$$

Onde  $\mu_A x$  indica o quanto  $x$  é compatível com  $A$  e a barra é empregada para ligar os elementos  $x$  com seus graus de pertinência em  $A$ .

Podemos concluir então que um conjunto *fuzzy*  $A$  em um universo  $U$  é caracterizado por uma função de pertinência  $\mu$  que, a cada um dos seus objetos, atribui um grau contínuo de associação que varia de  $0 \leq \mu_A(x) \leq 1$ , ou seja, cada elemento  $x$  do conjunto  $A$  é caracterizado por  $[x, \mu_A(x)]$ .

Além de  $U$ , outros conceitos básicos são importantes para o correto entendimento dos conjuntos *fuzzy*, dentre eles: o conjunto suporte, o peso, a normalização, o conjunto de nível e a cardinalidade.

O *conjunto suporte* de um conjunto *fuzzy*  $A$  é o conjunto *crisp* que contém todos os elementos do universo  $U$  que possuem um grau de pertinência superior à zero em  $A$  (expressão 3).

$$\text{sup } A = \{x \in U | \mu_A(x) > 0\} \quad (3)$$

Frequentemente encontra-se na literatura (Klir & Folger, 1988) uma notação definindo conjuntos *fuzzy* como um suporte finito, assumindo que  $x_i$  é um elemento do suporte do conjunto *fuzzy*  $A$  e que  $\mu$  é seu grau de pertinência em  $A$ .

O *peso* de um conjunto *fuzzy* é o maior grau de pertinência alcançado por algum elemento no conjunto. A *normalização*, por sua vez, é encontrada ao dividir cada elemento do conjunto pelo elemento que alcançou o maior peso, tornando-se o maior número igual a 1. Por exemplo, o conjunto de níveis de cinza (30, 50, 80, 100, 70, 40) é normalizado para (0.3, 0.5, 0.80, 1.0, 0.70, 0.40) se cada número for dividido por 100. Portanto, um conjunto *fuzzy* com faixa de pertinência entre 0 e 1, por exemplo, é chamado normalizado quando pelo menos um de seus elementos atinge o grau de pertinência igual a 1 (Kartalopoulos, 1996).

Um *conjunto de nível* ( $\alpha$ -cut) de um conjunto *fuzzy*  $A$  é um conjunto *crisp*  $A_\alpha$  que contém todos os elementos do conjunto  $U$  que possuem  $\mu_A$  igual ou superior ao valor especificado de  $\alpha$ , sendo representado conforme expressão 4.

$$A_\alpha = \{x \in U | \mu_A(x) \geq \alpha\} \quad (4)$$

Apesar de poder ser escolhido aleatoriamente, o valor de  $\alpha$  é normalmente definido a partir do conjunto *fuzzy* em foco.

### 3.4.3 Variáveis Linguísticas

Define-se como *variável linguística* aquela cujos valores são nomes de elementos de conjuntos *fuzzy* (termos linguísticos). A variável linguística admite como termos apenas expressões linguísticas, como “*muito quente*”, “*pequeno*”, “*aproximadamente médio*”, entre outros. Por exemplo, a *velocidade* de um determinado veículo pode ser uma variável linguística assumindo valores *lenta* (*slow*), *média* (*medium*) e *rápida* (*fast*) (Figura 13).

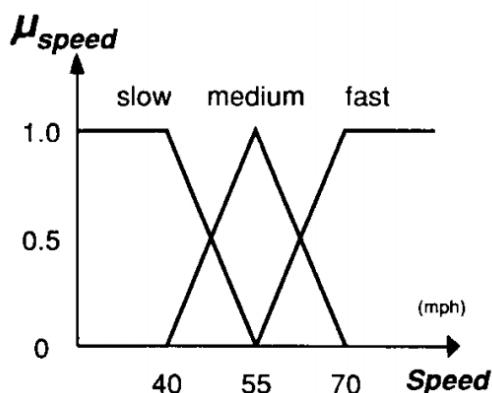


Figura 13. Variável linguística Velocidade (Speed) (Lee, 1990b).

Generalizando, a principal função de uma variável linguística é fornecer, dentro de uma função de pertinência, uma maneira sistemática para uma caracterização aproximada de fenômenos complexos ou mal definidos cujos valores podem ser sentenças em uma linguagem especificada, construídas a partir de termos linguísticos, modificadores, conectivos lógicos e delimitadores (Tanscheit, 2012).

Formalmente, uma variável linguística é caracterizada por uma quíntupla  $(N, T(N), U, G, M)$ , onde:

- $N$ : é o nome da variável;
- $T(N)$ : conjunto de termos de  $N$ , ou seja, o conjunto de nomes dos valores linguísticos de  $N$ ;
- $U$ : universo de discurso;
- $G$ : regra sintática (geralmente uma gramática) para gerar os valores de  $U$  como uma composição de termos  $T(N)$ , conectivos lógicos, modificadores e delimitadores;
- $M$ : regra semântica, para associar a cada valor gerado por  $G$  um conjunto fuzzy em  $U$ .

No caso da variável *temperatura* supracitada, ter-se-ia:

- $N$ : *temperatura*;
- $T(N)$ :  $\{\text{baixa}, \text{média}, \text{alta}\}$ ;
- $U$ : 0 a  $100^\circ$ , por exemplo;
- $G$ : *temperatura não baixa*, por exemplo;
- $M$ : associa o valor acima a um conjunto fuzzy cuja função de pertinência exprime o seu significado.

Assim, a variável linguística *temperatura* poderia ser definida como  $(\text{temperatura}, \{\text{baixa}, \text{média}, \text{alta}\}, [0, 100], G, M)$  e representada pela função de pertinência ilustrada na Figura 14.

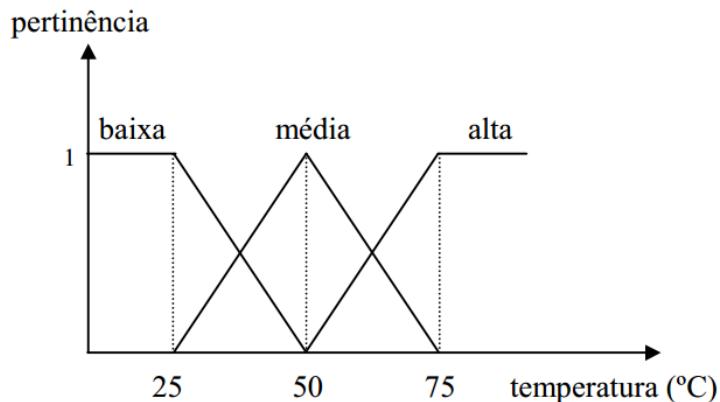


Figura 14. Função de pertinência para a variável linguística temperatura (Tanscheit, 2012).

Observa-se nesta função que abaixo de  $25^\circ$  tem-se a fronteira da negação total ou pertinência 0. Acima de  $75^\circ$  tem-se a fronteira da aprovação total ou pertinência 1. Logo, temperaturas de  $25^\circ$  ou menos com certeza são consideradas *baixas* e de  $75^\circ$  ou mais com certeza são consideradas *altas*. O ponto  $50^\circ$ , por sua vez, determina a *temperatura média* e estende-se um pouco para os dois lados do eixo da *temperatura* “espalhando” os valores das temperaturas no valor *média* desta variável.

Nos pontos intermediários, a *temperatura* é medida interceptando-se a linha vertical com a linha da função de pertinência. A partir de  $25^\circ$ , por exemplo, pode-se começar a pensar que a temperatura está *média*. Para isto, atribui-se um grau de pertinência, conforme a temperatura aumenta, que varia de 0 a 1. Da mesma forma que ele vai tendendo para a temperatura *média*, vai deixando de ter temperatura *baixa*. E à medida que a temperatura progride, afastando-se de  $50^\circ$  e aproximando-se de  $75^\circ$ , a temperatura vai deixando de ser *média* e se aproximando de *alta*. Assim, para  $70^\circ$ , a temperatura será considerada *alta* com grau de pertinência maior que 0,5.

Pode-se concluir então que as variáveis linguísticas são expressas qualitativamente através de termos linguísticos  $G$ , que fornecem os conceitos, e quantitativamente através de valores  $M$ , que fornecem valores aos conceitos.

### 3.4.4 Funções de Pertinência

Cabe à função de pertinência,  $\mu(x)$ , associar os elementos numéricos reais contidos no intervalo  $0 \leq x \leq 1$  aos elementos  $x \in A$ , inferindo desta forma o grau de pertinência do elemento  $x$  no conjunto  $A$ . A pertinência de um elemento em relação a determinado conjunto

deve ser entendida como a "intensidade" com que este elemento está relacionado a tal conjunto (Vaz, 2005).

A definição da função de pertinência é muito subjetiva considerando primordialmente que um conceito qualquer pode apresentar resultados variados para diferentes usuários em diferentes contextos. Além definir o conceito, ou seja, o quanto certo elemento pertence ou não a determinado conjunto, a função de pertinência também visa representar os limites do conjunto. Assim, é de relevante importância mapear corretamente a forma de variação e os limites de um conjunto *fuzzy* através da definição adequada da função de pertinência.

As funções de pertinência podem assumir diversos formatos (Figura 15). Entretanto, as funções triangulares, trapezoidais e gaussianas são as mais comumente empregadas para modelar as distribuições dos dados *fuzzy*.

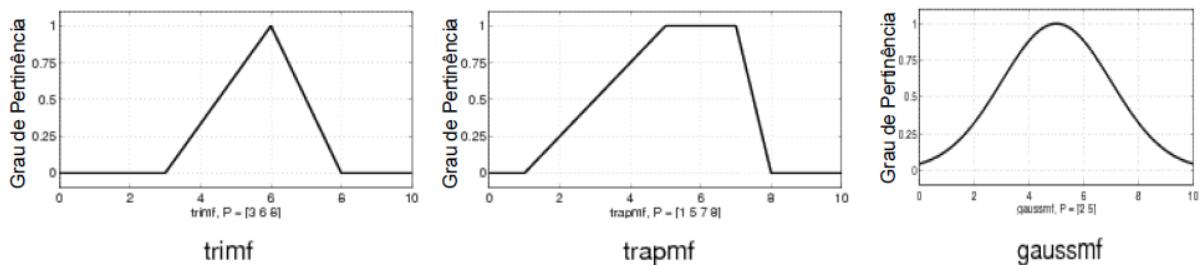


Figura 15. Formatos de funções de pertinência mais utilizados: triangular (*trimf*), trapezoidal (*trapmf*) e gaussiana (*gaussmf*). (MathWorks, 2014).

Para os casos dos conjuntos de dados mais concentrados em uma determinada região, sugere-se o emprego das funções de pertinência trapezoidais ou triangulares. Para os casos onde os conjuntos são mais dispersos, as gaussianas são preferidas. Entretanto, pode-se substituir uma gaussiana por uma função trapezoidal ou triangular quando se deseja um menor custo computacional e o problema em estudo assim o permite. Contudo, estas regras não são exatas. A forma de obtenção dos parâmetros que definem uma função de pertinência varia de acordo com a aplicação, não havendo, portanto, uma metodologia formal para a elaboração de tal função, pelo simples motivo de haver forte dependência com a distribuição dos dados e a natureza do problema (Santos, 2003).

A função trapezoidal é expressa usualmente através de quatro parâmetros que descrevem os quatro vértices do trapézio que representa a função. Dentre eles, dois possuem pertinência igual a 1 e outros dois possuem pertinência igual a 0. Já na função triangular, existe apenas um parâmetro com grau de pertinência igual a 1 (Santos, 2004).

### 3.4.5 Modificadores

As variáveis linguísticas são utilizadas para a criação de regras *fuzzy*. Essas variáveis geralmente necessitam do uso de *modificadores*, ou *hedges*, a fim de alterar a intensidade sobre a qual certo elemento faz ou não parte de um determinado conjunto. Analogamente, esses modificadores atuam na modelagem de conjuntos *fuzzy* da mesma forma que advérbios e adjetivos atuam em um sentença, modificando sua natureza.

Os modificadores pertencem a diversas classes e são matematicamente definidos. As classes e modificadores mais comuns são:

- Intensificadores (*muito*, *extremamente*): reduzem o grau de pertinência dos elementos que pertencem ao conjunto *fuzzy* elevando ao quadrado (*muito*) ou ao cubo (*extremamente*) valor da função de pertinência. Uma vez que as funções tem valor sempre menor que 1, o efeito causado é o de estreitamento, definido como *concentração*.
- Diluidores (*pouco*, *ligeiramente*): aumentam o grau de pertinência dos elementos que pertencem ao conjunto *fuzzy* tomando a raiz quadrada (*pouco*) ou cúbica (*ligeiramente*) do valor da função de pertinência. O efeito causado é o de alargamento, definido como *dilatação*.
- Aproximadores (*em torno*, *aproximadamente*, *quase*): têm a função de alargar ou estreitar uma região *fuzzy*, dependendo do modificador.

Esses modificadores são aplicados aos termos primários  $T(N)$  contidos nos conjuntos que, desta forma, sofrem modificações em seus significados. Estes modificadores, inseridos antes de cada termo, alteram a forma da função de pertinência através das operações matemáticas previamente definidas aplicadas aos seus valores.

Aplicando-se modificadores à variável *estatura*, por exemplo, pode-se determinar os conjuntos *fuzzy* *muito alta*, *ligeiramente alta*, *mais ou menos alta*, *muito baixa*, *pouco baixa*, e assim por diante.

### 3.4.6 Operações

Do mesmo modo que o cálculo proposicional e suas operações básicas de negação, conjunção e disjunção associam-se aos conjuntos *crisp*, associam-se também a lógica *fuzzy* as operações básicas de complemento, interseção e união, sendo chamados então de operadores

clássicos. Estas operações geralmente são definidas em função dos operadores máximo (*max*) e mínimo (*min*), análogos aos operadores *produto* e *soma* da álgebra elementar.

Utilizando as funções *max* e *min* em dois conjuntos fuzzy *A* e *B*, definidos em um conjunto universo *U*, tem-se o *complemento* ou *negação* (*not*) de um conjunto fuzzy *A* normalizado, em relação ao universo de discurso *U*, indicado por  $\neg\mu_A(x)$  e definido, para intervalo de função de pertinência [0,1], conforme a expressão 5.

$$\neg\mu_A(x) = 1 - \mu_A(x) \quad \forall x \in A \quad (5)$$

Assim, se um elemento tem um grau de pertinência de 0.6 em um conjunto *A*, seu grau de pertinência no complemento de *A* será de 0.4.

A operação de *interseção* ou *conjunção* (*and*) de dois conjuntos fuzzy *A* e *B* em relação ao universo de discurso *U*, com suas funções de pertinência  $\mu_A(x)$   $\mu_B(x)$  pode ser definida conforme a expressão 6.

$$\mu_{A \cap B}(x) = \min[\mu_A(x_i), \mu_B(x_i)] \quad (6)$$

A operação de *união* ou *disjunção* (*or*) de dois conjuntos fuzzy *A* e *B* em relação ao universo de discurso *U*, com suas funções de pertinência  $\mu_A(x)$   $\mu_B(x)$  pode ser definida conforme a expressão 7.

$$\mu_{A \cup B}(x) = \max[\mu_A(x_i), \mu_B(x_i)] \quad (7)$$

### 3.4.7 Regras Linguísticas

Os sistemas que implementam sistemas de inferência fuzzy são modelados através de conhecimentos especialistas sobre o problema. Estes conhecimentos, por sua vez, são representados utilizando-se regras linguísticas.

Na forma mais comum de representação, as regras linguísticas são do tipo condição–consequência (*implicação*, *se-então*) frequentemente chamadas de declarações condicionais fuzzy ou simplesmente regras fuzzy. Nestas, um conjunto de condições (proposições *antecedentes* envolvendo variáveis linguísticas) descrevendo uma parcela observável das saídas do processo é associado a consequências (proposições *consequentes* envolvendo descrições linguísticas) que irão manter ou levar o processo às condições de operação desejadas (Shaw & Simões, 1999).

Tome-se como exemplo a seguinte regra linguística:

*Se o aluno estuda muito, então a sua nota é alta.*

Generalizando-a, é possível escrevê-la da seguinte forma:

*Se <antecedente>, então <consequente>.*

Para que a regra seja considerada linguística, os termos *antecedente* e *consequente* devem conter uma ou mais cláusulas linguísticas, como por exemplo, “*aluno estuda muito*”. Quando satisfeito o antecedente (mesmo parcialmente), o processamento do consequente da regra é determinado através de um mecanismo de inferência. O consequente representa um conjunto de ações ou consequências (Rezende, 2005).

A intenção é representar o conhecimento a partir de regras nas quais as condições das implicações são dadas a partir de um conjunto de termos linguísticos associados às variáveis de entrada do processo e as consequências ou saídas expressam-se de modo similar para cada variável.

As regras em um sistema *fuzzy* são expressas na forma de proposições (Haubecher & Tizhoosh, 2000), tais como:

*SE Banda1 é Alto <conectivo1> Banda2 é Alto ENTÃO x é Classe A*

*<conectivo2>*

*SE Banda1 é Médio <conectivo1> Banda2 é Médio ENTÃO x é Classe B*

Onde:

- *Banda1* e *Banda2*: são variáveis de entrada (valores conhecidos);
- *x*: é uma variável de saída (um valor a ser calculado);
- *Alto*: é um valor linguístico com sua respectiva função de pertinência (conjunto *fuzzy*) definida nas Bandas 1 e 2;
- *Médio*: é um valor linguístico com sua respectiva função de pertinência definida nas Bandas 1 e 2;
- *Classe A* e *Classe B*: valores linguísticos com suas respectivas funções de pertinência definidas em x;
- *Conectivo1* e *Conectivo2*: são os operadores lógicos escolhidos para se expressar a inferência desejada. O *conectivo1* está sendo utilizado no processo de agregação da regra e *conectivo2* no processo de combinação das regras. A escolha do método para combinar os conjuntos associados às variáveis do sistema tem grande importância na estrutura do sistema.

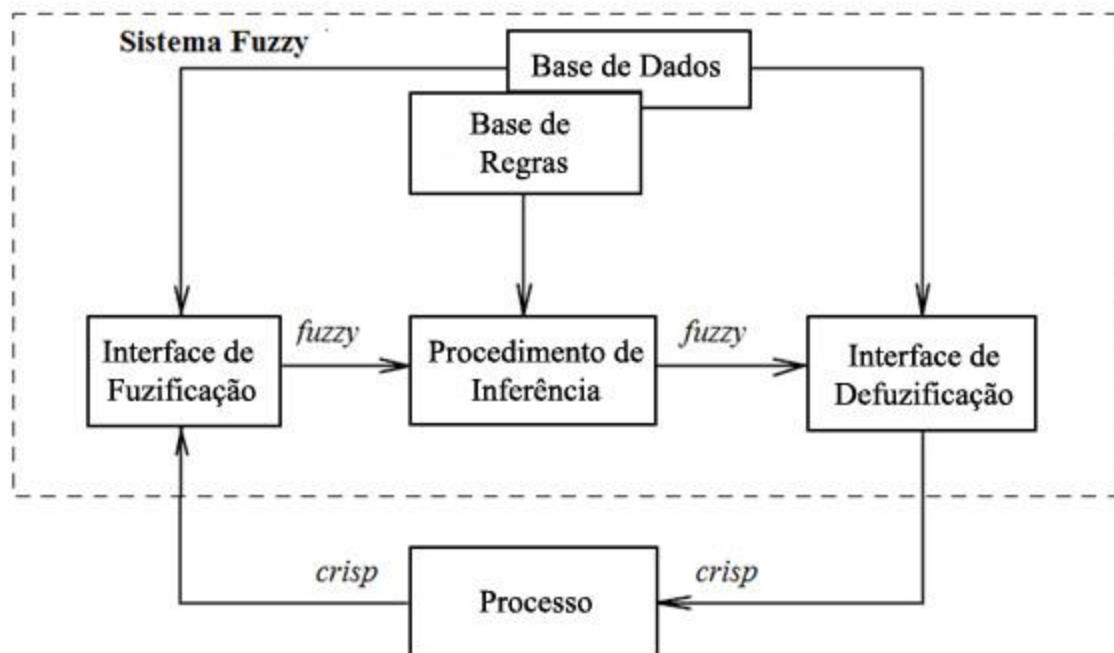
A primeira parte da regra, após o “Se”, também é chamada de premissa ou antecedente da implicação e a segunda parte, posterior ao “então”, de conclusão ou consequente da implicação. Na lógica clássica, se o antecedente da implicação é verdadeiro, o consequente é obrigatoriamente. Na lógica *fuzzy*, analogamente, se o antecedente possui algum grau de verdade, o consequente é verdadeiro para o mesmo grau.

Em um sistema *fuzzy* existem muitas regras que utilizam inferência associativa paralela, isto é, são verificadas concomitantemente, ao invés de em série como nos sistemas tradicionais. Neste caso, em um sistema *fuzzy*, a ordem de execução das regras não é importante, exceto quando uma regra depende dos resultados de execução de outras (Kartalopoulos, 1996).

O conjunto de regras linguísticas de um sistema *fuzzy* é comumente chamado de *Base de Regras*.

### 3.4.8 Sistema de Inferência Fuzzy

O sistema de inferência *fuzzy* é um sistema baseado em regras *fuzzy* e possui cinco componentes básicos: a interface de fuzificação, a base de conhecimento, a base de dados, o procedimento de inferência e a interface de defuzificação, conforme ilustrado na Figura 16.



*Figura 16. Sistema de Inferência Fuzzy (adaptado de Leite, 2009).*

Considerando-se inicialmente entradas não-fuzzy (*crisps*) resultantes de medições ou observações (conjuntos de dados, por exemplo), faz-se necessário mapear estes dados precisos para que sejam realizadas inferências fuzzy.

A *Interface de Fuzzificação* recebe os valores de entrada, condiciona estes valores aos universos de discurso normalizados e fuzifica-os, ou seja, transforma-os em elementos do conjunto que possam se tornar instâncias de variáveis linguísticas. Assim, a interface de fuzzificação está diretamente relacionada às variáveis, seus respectivos universos de discurso e os valores que cada variável pode assumir (conjuntos fuzzy). Através da interface de fuzzificação os valores observados das variáveis de entrada são associados ao respectivo universo de discurso, permitindo uma avaliação do grau de pertinência aos conjuntos fuzzy associados a cada variável. Neste estágio ocorre também a ativação das regras relevantes para uma dada situação. Pinho (1999) cita a necessidade de que especialistas da área estudada sejam consultados durante a atribuição de valores relacionados aos graus de pertinência para cada uma das variáveis em estudo, contribuindo assim para maior precisão nos resultados.

A *Base de Conhecimento Fuzzy* consiste em uma *Base de Regras* e uma *Base de Dados*. A *Base de Regras* pode ser elaborada por especialistas do domínio em forma de sentenças linguísticas e, definindo a estratégia e o controle do sistema, constitui um aspecto fundamental no seu desempenho. Assim, o desempenho será considerado confiável e satisfatório desde que as regras expressem fiel e consistentemente o seu comportamento. Alternativamente ao uso de especialistas, podem ser usados métodos automáticos de extração de regras a partir de dados numéricos que são particularmente úteis em problemas de classificação. A *Base de Dados* armazena as definições necessárias sobre as variáveis linguísticas, normalizações e discretizações dos universos de discurso, as partições fuzzy dos espaços de entrada e saída e as definições das funções de pertinência.

O *Procedimento de Inferência* processa os dados fuzificados de acordo com as regras fuzzy definidas objetivando inferir as ações de controle. Neste componente, as regras são utilizadas para se obter a relação fuzzy previamente modelada. Este componente tem muita importância, visto que, como fornece a saída a partir de cada entrada, é dele que depende o sucesso do sistema. Os conjuntos fuzzy de entrada, relativos aos antecedentes das regras, e o de saída, referentes aos consequentes, podem ser definidos previamente ou gerados automaticamente a partir dos dados.

A *Interface de Defuzificação* processa as saídas fuzzy do procedimento anterior em saídas não-fuzzy através de um escalamento, de modo a compatibilizar os valores normalizados com os valores dos universos de discurso reais das variáveis. Segundo Von

Altrock (1996), a defuzificação consiste na tradução do resultado linguístico do processo de inferência *fuzzy* em um valor numérico. Em outras palavras, como o algoritmo de controle faz com que o processamento das variáveis linguísticas de entrada resulte em um valor da variável linguística de saída, o processo de defuzificação consiste em selecionar um valor numérico específico que represente o resultado *fuzzy* da variável de saída produzido pelo conjunto de regras *fuzzy* (Cox, 1995).

---

# Capítulo 4

## Técnica Desenvolvida

---

### 4.1 Introdução

Neste capítulo são apresentados a contextualização da abordagem desenvolvida, seus conceitos, aspectos relevantes e contribuições. Para tal, subdivide-se em:

- **Contextualização:** Apresenta a relevância da técnica desenvolvida, sua contextualização e ideia geral.
- **Apresentação da Técnica e Etapas Desenvolvidas:** Apresenta a técnica desenvolvida e suas etapas de aquisição, pré-processamento, segmentação, representação e descrição e reconhecimento e interpretação e a base de conhecimento.

### 4.2 Contextualização

Conforme explicitado no primeiro capítulo, este trabalho tem como objetivo apresentar um conjunto de técnicas inteligentes para segmentação automática de componentes sanguíneos e classificação diferencial de leucócitos em imagens microscópicas coradas.

Como a entrada do sistema dar-se-á através de imagens digitais e, a partir deste ponto, um conjunto de técnicas e métodos será aplicado a fim de segmentar e classificar os seus elementos de interesse, um prévio detalhamento deverá ser feito para possibilitar o entendimento do contexto no qual o problema está inserido e o correto ajuste das variantes do processo.

Para tal explanação, o modelo necessita fortemente do conhecimento de um especialista. O especialista deverá ser proveniente do contexto no qual o problema está inserido e conhecedor das especificidades e peculiaridades do mesmo, bem como da relevância de uma automatização no seu processo. Ao especialista cabe, portanto, a contextualização do domínio do problema, o detalhamento dos conceitos e aspectos do negócio, suas regras, atuações

limítrofes, descritores de regiões de interesse e relevância. Esta atuação desencadeará na construção da *base de conhecimento*, é imprescindível para o correto desempenho da técnica desenvolvida e sendo composta por aspectos relevantes para os corretos resultados no processamento da imagem. Construída tanto autonomamente, através de aspectos inerentes à imagem adquirida, quanto através da inferência dos conhecimentos especialistas, a *base de conhecimento* intervém no processo em muitos momentos. Aspectos como: técnicas de processamento digital de imagens, histogramas, funções de pertinência e regras de inferência fuzzy são intrínsecos e inerentes à *base de conhecimento* da técnica desenvolvida.

É de responsabilidade do modelo, portanto, efetuar o processamento da imagem em todos os aspectos relevantes, interferindo nestes por inferências a partir da *base de conhecimento*. Para tanto, executa sequencialmente uma série de etapas com regras previamente especificadas e delimitadas, cujos objetivos são conhecidos e bem definidos. As etapas perfazem um todo genérico e robusto, ao qual acoplam-se as tomadas de decisão. Estas, por sua vez, são dependentes do domínio do problema e fortemente acopladas às bases de conhecimento. A extração de resultados, finalmente, é dependente de todas estas definições.

### 4.3 Apresentação da Técnica e Etapas Desenvolvidas

De acordo com as etapas do processamento digital de imagens (Figura 12), segue na Figura 17 uma representação gráfica da técnica desenvolvida, suas etapas e, destacados em cada uma, os métodos utilizados neste trabalho.

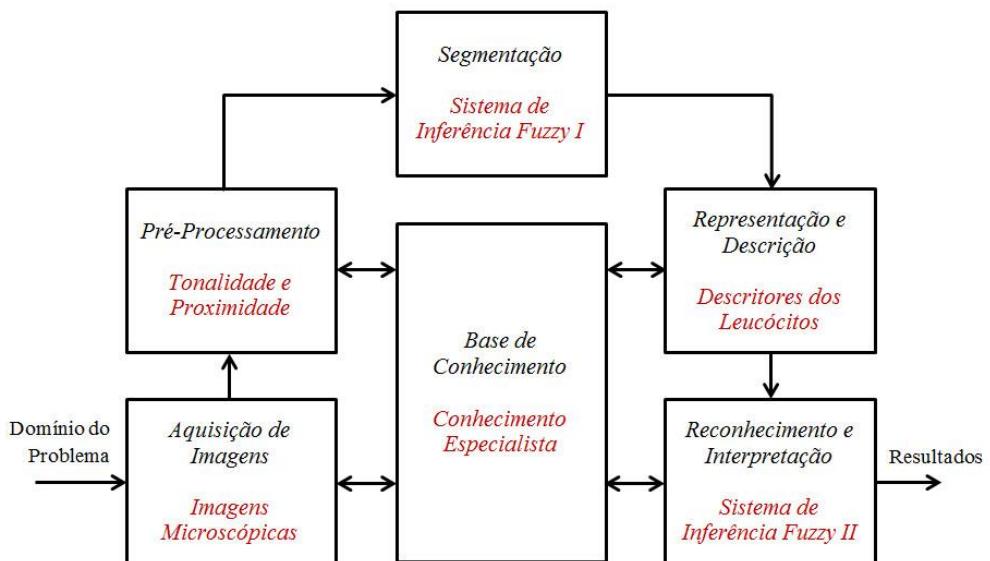


Figura 17. Representação gráfica da técnica desenvolvida neste trabalho.

### 4.3.1 Etapa: Aquisição de Imagens

Esta etapa é a responsável por disponibilizar ao modelo as imagens de entrada que serão processadas. Os esfregaços sanguíneos são preparados, fixados e corados utilizando técnicas de coloração hematológicas variadas. A aquisição da imagem é feita por uma câmera acoplada ao microscópio óptico. Após serem adquiridas, as imagens são inseridas no modelo.

Em esfregaços sanguíneos preparados para análise microscópica, o plasma sanguíneo exibe os tons mais claros da amostra, os eritrócitos e citoplasmas leucocitários exibem tons intermediários e o núcleo leucocitário exibe os tons mais escuros. Considerando que os leucócitos ( $6.000 - 9.000/\text{mm}^3$  de sangue) aparecem com menos frequência do que os eritrócitos ( $4,5 - 6,2$  milhões/ $\text{mm}^3$  de sangue) e que a área ocupada pelo plasma sanguíneo é tão larga quanto ou mais larga que aquela ocupada pelos eritrócitos, o histograma da imagem frequentemente mostra um pico mais escuro, representando quase que imperceptivelmente os núcleos leucocitários, e outros dois picos, representando as áreas mais visíveis de plasma sanguíneo, eritrócitos e citoplasmas leucocitários. Apesar dos tamanhos diferentes das áreas que ocupam, as imagens foco deste trabalho sempre apresentam pelo menos um leucócito, muitos eritrócitos e significantes áreas de plasma sanguíneo.

As imagens utilizadas neste trabalho são provenientes de duas fontes: Hemocentro do Rio Grande do Norte Dalton Cunha (Hemonorte) e *BloodLine Image Atlas*.

As imagens provenientes do Hemonorte, totalizando 415 amostras, foram elaboradas, fixadas, coradas e capturadas com o auxílio de especialistas da área hematológica. As amostras sanguíneas foram preparadas utilizando esfregaços sanguíneos fixados e corados com os métodos Panótico Rápido (150 amostras), Leishman (140 amostras) e Rosenfeld (125 amostras). As imagens foram captadas utilizando o microscópio óptico de imersão LABOMED LX 400 com aumento de 100x e uma câmera iVu 5100 acoplada, com resolução de  $1280 \times 720$ , 96 pixels/polegadas e formato de imagem JPEG. A etapa de aquisição destas imagens foi realizada no primeiro semestre do ano de 2014.

As imagens provenientes do *BloodLine Image Atlas* (*BloodLine Imagem Atlas*, 2010) estão disponíveis para fins educacionais. Muito embora o site disponibilize mais de 800 imagens de lâminas hematológicas que descrevem várias doenças do sangue, optou-se por selecionar apenas as 115 imagens de sangue normal para permitir uma comparação mais precisa entre estas imagens e as demais. Foram selecionadas apenas as imagens classificadas como "sangue normal" e "Área de sangue: denso e fino", todas capturadas por microscópio com fator de aumento de 100x. As colorações utilizadas na preparação das lâminas,

visivelmente diferentes de uma para outra, não foram especificados pelo *site*.

Para que o comparativo entre os resultados da segmentação automática e da segmentação real feita por um especialista pudesse ser feito, todas as 530 imagens utilizadas neste trabalho foram previamente segmentadas de forma manual utilizando o programa Adobe Photoshop CS5 Extended versão 12.0 x64. Todas as segmentações manuais foram previamente aprovadas pelos especialistas da área hematológica que acompanharam o trabalho.

Para cada imagem adquirida, o método proposto segue as etapas seguintes (ilustradas na Figura 18), pré-processamento, segmentação, representação e descrição e reconhecimento e interpretação, interagindo, sempre que for necessário, com a base de conhecimento.

### 4.3.2 Base de Conhecimento

A *base de conhecimento* possui o conhecimento especialista proveniente da área hematológica e necessário aos objetivos do modelo. Foram definidos quatro parâmetros como indispensáveis para a base de conhecimento e que, durante a análise, possibilitarão a tomada de decisões quanto às inferências que devem ser feitas nas imagens:

- Regiões de interesse por cor ou tonalidade;
- Regiões de interesse pela distância euclidiana do núcleo leucocitário;
- Descritores leucocitários para classificação diferencial;
- Bases *Fuzzy* de Regras e de Dados.

#### Regiões de interesse por cor ou tonalidade

As imagens de entrada são compostas por diversas regiões, mas não necessariamente todas devem ser analisadas por não oferecerem relevância significativa ao problema. Como o domínio do problema concentra-se na análise de lâminas de esfregaços sanguíneos corados, na fase de segmentação devem ser considerados os eritrócitos, leucócitos (núcleo e citoplasma) e plasma sanguíneo e na fase de classificação diferencial devem ser considerados somente os leucócitos. Deste modo, a *Base de Conhecimento* deve “saber” quais são as áreas relevantes da imagem que serão analisadas, ou seja, suas regiões de interesse.

Considerando que as imagens a serem analisadas passarão antes por um processo de coloração, pondera-se que o primeiro parâmetro a ser definido na *Base de Conhecimento* é a cor ou tonalidade da região, ou de cada uma das regiões, de interesse. As decisões baseadas

neste aspecto serão tomadas pelo sistema nas etapas de pré-processamento e segmentação partindo da premissa básica de que, independente da coloração hematológica utilizada, as amostras comportam-se dentro de um padrão, qual seja, preservando a tonalidade mais escura para núcleos leucocitários, a mais clara para plasma sanguíneo e a intermediária para citoplasma leucocitário e eritrócitos (Figura 19).



*Figura 18. Exemplo de imagem microscópica de esfregaço sanguíneo contendo um leucócito (púrpura), vários eritrócitos (róseos) e plasma leucocitário (área mais clara da amostra) (Heckner & Freund, 2000).*

No pré-processamento, as técnicas utilizadas para realce deverão ser escolhidas com a finalidade de enfatizar as regiões de interesse e identificar inicialmente as áreas de núcleo leucocitário e plasma sanguíneo. Assim, após estudos feitos com alguns modelos de cores, como YCbCr, RGB e HSV, optou-se pela utilização do modelo RGB e especialmente do canal G visto que o mesmo, quando comparado aos demais canais dos demais modelos, apresentou o núcleo leucocitário com a tonalidade mais escura e o plasma sanguíneo com a tonalidade mais clara.

No *Sistema de Inferência Fuzzy I*, proposto para a etapa de segmentação, a cor relevante será usada como a variável linguística de entrada *tonalidade* e norteará o processo de delimitação das quatro regiões de interesse: eritrócitos, núcleos e citoplasma leucocitários e plasma sanguíneo.

## Regiões de interesse pela distância euclidiana do núcleo leucocitário

Os componentes sanguíneos se destacam uns dos outros nas amostras a partir das cores distintas que assumem quando dos processos de coloração. Porém, os componentes eritrócito e citoplasma leucocitário assumem tonalidades extremamente semelhantes e que não se comportam sempre da mesma forma. Assim, eventualmente os eritrócitos são mais escuros que os citoplasmas e eventualmente ocorre o inverso, o que resulta em histogramas com três regiões bem definidas (núcleo, plasma e demais elementos) e uma quarta região (eritrócito/citoplasma) mal definida e mal posicionada. Por este motivo, além da coloração ou tonalidade dos componentes, neste trabalho a proximidade do citoplasma leucocitário com relação ao núcleo também foi considerada por ser esta é uma poderosa ferramenta de identificação do mesmo, visto que os eritrócitos são anucleados.

No pré-processamento, as distâncias euclidianas entre os *pixels* da imagem e os centróides dos núcleos mais próximos deverão ser calculadas para auxiliar na distinção entre as regiões dos eritrócitos e citoplasmas leucocitários. No *Sistema de Inferência Fuzzy I* proposto para a etapa de segmentação, a distância também será usada como a variável linguística de entrada *proximidade*.

## Descritores dos objetos

A partir de uma imagem muitas propriedades ou características podem ser extraídas. Tais características podem ser gerais ou de domínios específicos. As características gerais são aquelas encontradas comumente, independente da imagem analisada, tais como cor, textura e forma, por exemplo. As características de domínio específico, por sua vez, são dependentes do domínio do problema e da área de aplicação do sistema. Estas características podem ser mensuradas e agrupadas em um vetor de características, denominado vetor de descritores da imagem. Assim, cada objeto da imagem pode ser representado por um elemento ou linha do vetor de descritores da imagem e poderá pertencer ou não a um agrupamento de acordo com estes valores.

Por possuir um campo de atuação muito amplo, o processamento digital de imagens microscópicas lida com objetos que, dependendo do problema, devem ser considerados iguais ou distintos. Portanto, a relevância do descritor é fortemente acoplada ao domínio do problema. Ao mesmo tempo em que um determinado descritor é relevante para um contexto, como a forma é relevante para a diferenciação celular, por exemplo, não é relevante para outro, como a mesma forma não é relevante à identificação de contaminação alimentar por

fungos. E como este conhecimento é externo ao ambiente do modelo em si, deve ser fornecido previamente e incorporado à base de conhecimento.

Assim, os processos de extração de características e configuração do vetor de descritores são fundamentais tanto para a segmentação correta dos objetos relevantes contidos nas imagens quanto para a recuperação, reconhecimento e interpretação destes.

Neste trabalho, os descritores são utilizados para promover a classificação diferencial dos leucócitos em cinco tipos: eosinófilos, basófilos, linfócitos, monócitos e neutrófilos. Os descritores mais comumente utilizados concentram-se em descritores de forma e textura. Neste trabalho, como descritores de forma são utilizados *convexidade do núcleo*, *razão núcleo-leucócito*, *razão núcleo-citoplasma* e *área do leucócito*. Como descritor de textura, tem-se a *energia do citoplasma* (definida como a soma dos quadrados dos elementos na matriz de co-ocorrência). Além destes, são definidos os seguintes descritores específicos: *cantos do núcleo*, *erosões do núcleo*, *diâmetro equivalente* e *granularidade*. Cada descritor será utilizado na etapa de reconhecimento e interpretação como uma variável linguística do *Sistema de Inferência Fuzzy II* proposto para a classificação diferencial dos leucócitos.

### Bases *Fuzzy*

As Bases *Fuzzy* correspondem às *Bases de Regras* e *Bases de Dados* dos Sistemas de Inferência *Fuzzy* propostos para segmentação e classificação. Nestas bases estão contidas as regras *fuzzy* previamente modeladas por especialistas, as variáveis linguísticas de entrada e saída relevantes ao domínio do problema e as funções de inferência dos Sistemas de Inferência *Fuzzy I* e *II* propostos, respectivamente, para a segmentação dos componentes e classificação diferencial dos leucócitos.

#### 4.3.3 Etapa: Pré-Processamento

Na etapa de pré-processamento são calculados os valores que serão utilizados na definição dos suportes dos conjuntos *fuzzy* para as variáveis *tonalidade* e *proximidade*, utilizadas na etapa de segmentação. O algoritmo desta fase pode ser brevemente descrito como:

- Passo 1: Extração do canal *G* da imagem *RGB*.
- Passo 2: Extração do histograma da imagem *G*.

- Passo 3: Cálculo dos três picos mais relevantes do histograma (*PicoEscuro*, *PicoMédio* e *PicoClaro*) que serão usados na definição da variável linguística *tonalidade*.
- Passo 4: Pré-classificação da imagem  $G$  em três regiões, de acordo com os valores de *PicoEscuro* e *PicoClaro* (núcleos leucocitários  $\leq$  *PicoEscuro*, *PicoEscuro* < frente < *PicoClaro* e plasma sanguíneo  $\geq$  *PicoClaro*).
- Passo 5: Cálculo dos centróides das áreas pré-classificadas como núcleos leucocitários após operações de fechamento e preenchimento de regiões.
- Passo 6: Cálculo das distâncias euclidianas (matriz  $D$ ) entre os *pixels* da imagem e o centróide do núcleo leucocitário mais próximo.
- Passo 7: Cálculo dos valores *AltaProx* e *BaixaProx* que serão utilizados na definição da variável linguística *proximidade*.

Detalhando esta etapa e considerando que as imagens microscópicas de lâminas sanguíneas são representadas por matrizes  $A_{m \times n \times 3}$ , nas quais  $m$  e  $n$  correspondem aos dois eixos do plano cartesiano distribuídos nos três canais de cores primárias (modelo RGB), optou-se por utilizar somente o canal  $G$  da imagem por apresentar os núcleos leucocitários, comparado aos canais R e B, com os tons mais escuros de cinza.

Assim, a partir da regra básica de que, independente do método de coloração hematológico utilizado na elaboração da amostra, os núcleos leucocitários e o plasma sanguíneo possuem os tons mais escuros e mais claros respectivamente, os canais de cores RGB da imagem original  $IO = (io_{ij})_{m \times n \times 3}$  (Figura 19a) são separados e somente o canal G,  $G = (g_{ij})_{m \times n}$ , é considerado (Figura 19b). Em seguida, o histograma do canal G,  $HG = (hg_{ij})_{256 \times 1}$ , é construído (Figura 19c) e os três picos mais relevantes do histograma, *PicoEscuro*, *PicoMédio* e *PicoClaro*, são calculados.

Para a definição dos três picos mais relevantes da imagem original, o histograma do canal G é inicialmente considerado como o vetor  $HG$  de 256 posições, cada posição representando um nível de cinza (do 0 ao 255) e contendo a frequência de ocorrência de *pixels* para aquele nível. Ao percorrer o vetor do início ao fim, se o valor da frequência na posição atual for maior que o valor na posição seguinte, a posição seguinte recebe o valor da posição atual. Assim, tem-se uma repetição da frequência maior por um determinado número de posições sempre que houver uma diminuição de frequência na curva do histograma, indicando a presença de vales (Figura 19d). O mesmo procedimento é repetido percorrendo o vetor na

outra direção (a partir do fim) para incluir o caso de o terceiro pico ser menor que o segundo. Ao final, os picos mais relevantes serão os três níveis de cinza cujas frequências mais se repetiram (Figura 19e), com *PicoEscuro* definido como sendo a máxima frequência entre o nível de cinza 0 e a ocorrência do primeiro vale, *PicoMédio* como a máxima frequência entre o primeiro e o segundo vales em *PicoClaro* como a máxima frequência entre o segundo vale e o nível de cinza 255. Estes valores serão usados para definir o suporte do conjunto *fuzzy tonalidade* que será utilizada no *Sistema de Inferência Fuzzy I* na etapa de segmentação.

Considerando os valores *PicoEscuro* e *PicoClaro* como limiares, o canal  $G$  é pré-classificados de acordo com a expressão 8, gerando a matriz  $GC = (gc_{ij})_{m \times n}$ , contendo regiões identificadas como áreas de núcleos leucocitários, de frente (não-núcleos e não-plasma) e de plasma sanguíneo, determinados como 0, 126 e 255, respectivamente.

$$gc_{ij} = \begin{cases} 0, & seg_{ij} \leq PicoEscuro \\ 126, & seg_{ij} > PicoEscuro \wedge g_{ij} < PicoClaro \\ 255, & seg_{ij} \geq PicoClaro \end{cases} \quad (8)$$

Na área pré-classificada como núcleo leucocitário são aplicadas as operações de fechamento (dilatação seguida de erosão), utilizando um elemento estruturante circular, e preenchimento de regiões para conectar objetos próximos que foram desconectados. O centróide de núcleo leucocitário (centro de massa),  $c_{pq}$ , é então calculado e a matriz  $D = (d_{ij})_{m \times n}$  é construída a partir das distâncias euclidianas entre cada *pixel* da imagem e o centróide, de acordo com a expressão 9.

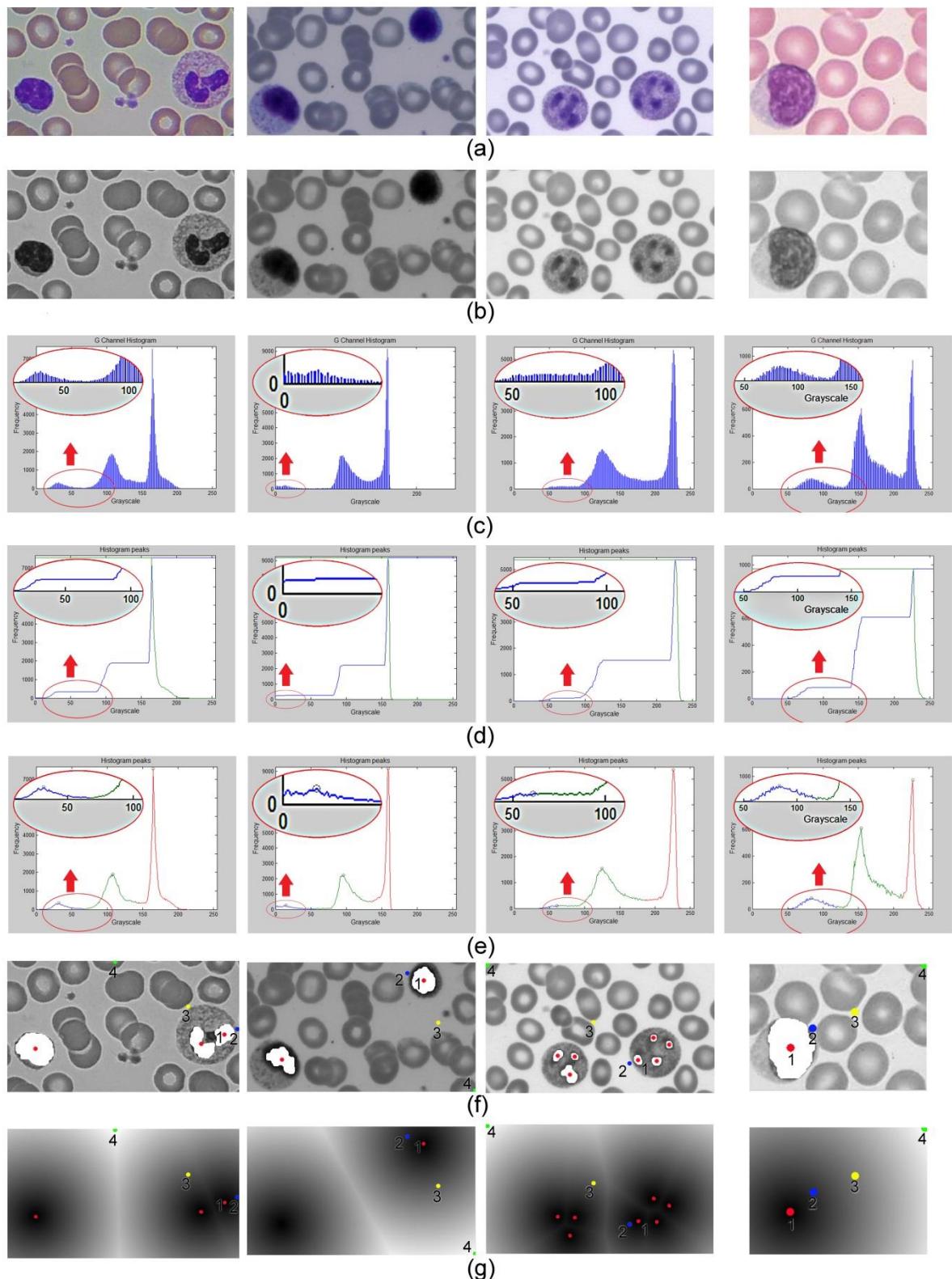
$$d_{ij} = \sqrt{(i - p)^2 + (j - q)^2} \quad (9)$$

Em seguida a matriz  $DP = (dp_{ij})_{n \times m}$ , contendo somente as distâncias euclidianas entre os *pixels* pré-classificados como plasma sanguíneo e o centróide do núcleo leucocitário, também é calculada de acordo com a expressão 10.

$$dp_{ij} = \begin{cases} 0, & segc_{ij} < 255 \\ d_{ij}, & segc_{ij} \geq 255 \end{cases} \quad (10)$$

Se houver mais de uma área pré-classificada como núcleo leucocitário, mais de um centróide será calculado e a distância euclidiana de um ponto qualquer será sempre referente ao centróide do núcleo leucocitário mais próximo a ele.

Finalmente os valores de *AltaProx* e *BaixaProx* são calculados. Considerou-se como *AltaProx* o mínimo valor da matriz *DP*, indicando o plasma sanguíneo mais próximo ao núcleo leucocitário, e como *BaixaProx* como sendo 1/3 da distância euclidiana entre o valor de *AltaProx* e o máximo valor da matriz *D*. Estes valores serão utilizados para definir o suporte do conjunto *fuzzy* para a variável *proximidade* utilizada no *Sistema de Inferência Fuzzy I* na etapa de segmentação (Figuras 19f e 19g).



*Figura 19. a) Imagens originais RGB $m \times n \times 3$  (amostras coradas com os métodos Panótico Rápido, Leishman, Rosenfeld e com coloração não especificada, respectivamente); b) Canais G; c) Histogramas dos canais G com aumento (escalas de cinza x frequência); d) Picos dos histogramas: repetição das frequências mais altas; e) Picos dos histogramas: PicoEscuro –*

*azul, PicoMédio–verde, PicoClaro–vermelho; f) Limiarização dos núcleos leucocitários (branco), centróides (1-vermelho), AltaProx (2-azul), BaixaProx (3-amarelo) e máximo valor da matriz D (4-verde); g) Distâncias Euclidianas de cada pixel em relação ao centróide do núcleo leucocitário mais próximo (mesmos pontos 1, 2, 3 e 4).*

#### 4.3.4 Etapa: Segmentação

Na fase de segmentação, o *Sistema de Inferência Fuzzy I* classifica os *pixels* de cada região da imagem considerando as variáveis linguísticas *fuzzy* de entrada *tonalidade* e *proximidade*, suas respectivas funções de pertinência e a base de regras *fuzzy*. Nesta fase, o algoritmo pode ser descrito como:

- Passo 1: Construção das variáveis linguísticas *fuzzy* de entrada *tonalidade* e *proximidade* utilizando os valores *PicoEscuro*, *PicoMédio*, *PicoClaro*, *AltProx* e *BaixaProx* calculados na fase de pré-processamento.
- Passo 2: Fuzificação da matriz G considerando os termos linguísticos da variável *fuzzy tonalidade* (escuro, médio, claro) e suas respectivas funções de pertinência.
- Passo 3: Fuzificação da matriz D, considerando os termos linguísticos da variável *fuzzy proximidade* (alta e baixa) e suas respectivas funções de pertinência.
- Passo 4: Aplicação da *Base de Regras Fuzzy*.
- Passo 5: Agregação das saídas e defuzificação da variável *fuzzy* de saída *classe* (citoplasma leucocitário, núcleo leucocitário, eritrócitos e plasma sanguíneo) utilizando o método da Média dos Máximos (MoM) (Tanscheit, (2012)).

Desta forma, a etapa de segmentação consiste em um *Sistema de Inferência Fuzzy* que, utilizando funções de pertinência para fuzificação de dados *crisp* e um conjunto de regras *fuzzy* bem definidas, classifica a imagem em quatro regiões distintas: núcleo e citoplasma leucocitários, eritrócitos e plasma sanguíneo. A região do leucócito é resultante da conjunção entre as regiões do núcleo e citoplasma leucocitários,

Os componentes sanguíneos se destacam uns dos outros nas amostras porque eles exibem diferentes cores durante os processos de coloração. Considerando que este trabalho se concentra especificamente no canal *G* da amostra, diferentes níveis de cinza identificam esses componentes. No entanto, os eritrócitos e os citoplasmas leucocitários apresentam níveis de

cinza muito similares e nem sempre se comportam da mesma forma. Assim, em algumas imagens os eritrócitos apresentam tons mais escuros do que os citoplasmas leucocitários e em outras imagens ocorre o inverso, resultando em histogramas com três regiões bem definidas (núcleo leucocitário, plasma sanguíneo e outros elementos) e uma quarta região mal definida e mal posicionada (eritrócitos/citoplasmas leucocitários). Por esta razão, além dos níveis de cinza, a *proximidade* entre o núcleo e o citoplasma leucocitários foi também considerada, uma vez que esta é uma ferramenta poderosa de identificação, tendo em vista que os eritrócitos são anucleados. Uma abordagem *fuzzy* foi utilizada para identificar estes conceitos imprecisos de tonalidade e proximidade, devido ao alto grau de incerteza que envolve a segmentação destas áreas. Um *Sistema de Inferência Fuzzy* é proposto, utilizando o Método Mamdani de três fases apresentado por Mamdani (1974) (Figura 20).

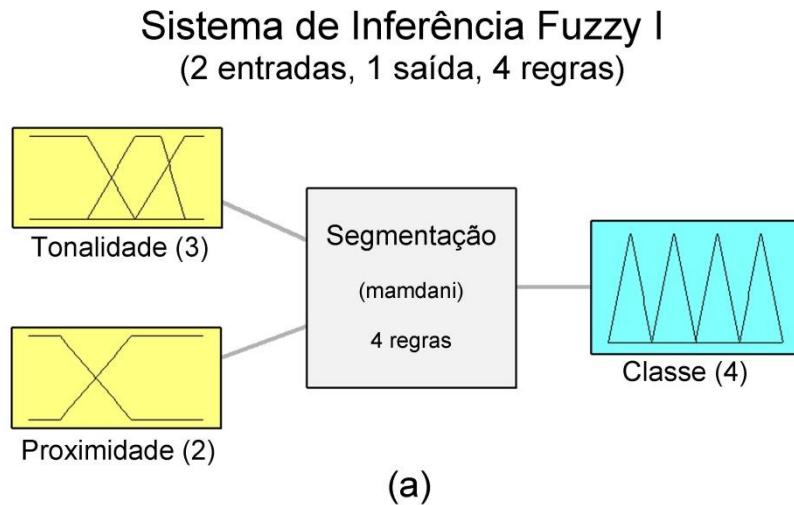
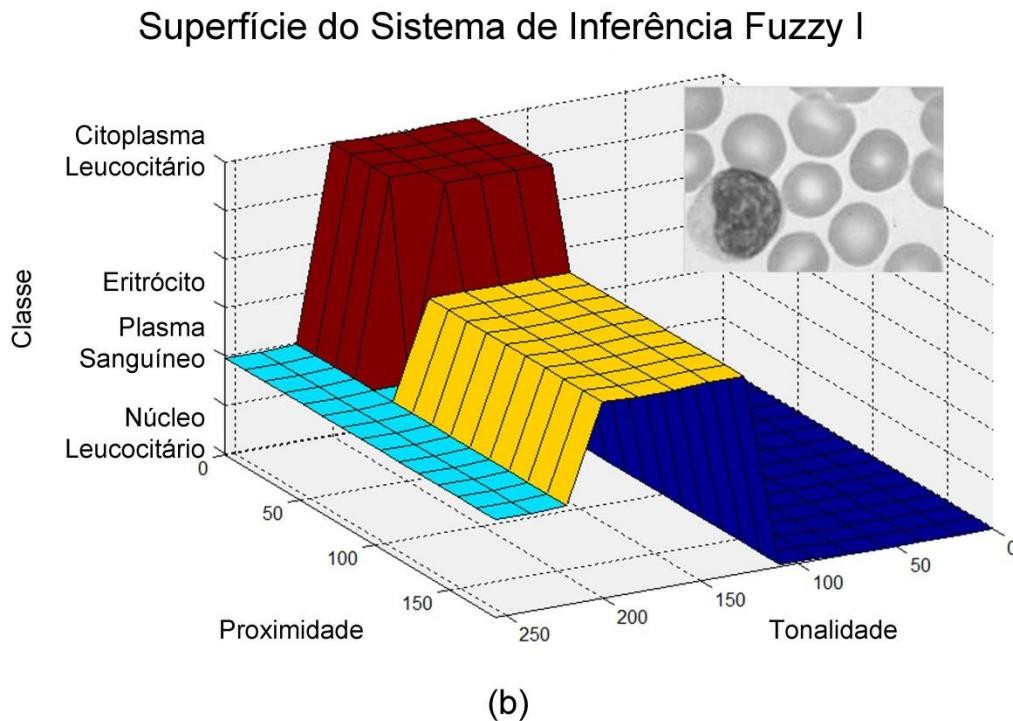


Figura 20. Sistema de Inferência Fuzzy I – Variáveis linguísticas fuzzy de entrada e saída.

As fases do Método Mamdani são: fuzificação, procedimentos de inferência e defuzificação, visando classificar de cada imagem em quatro regiões após o processamento individual de cada *pixel* (Figura 21).



*Figura 21. Sistema de Inferência Fuzzy I – Superfície do Sistema.*

A fase de fuzificação consiste em mapear as entradas numéricas *crisp* para os conjuntos *fuzzy*, representados pelas variáveis linguísticas de entrada, através das funções de pertinência. O processo de inferência *fuzzy* é responsável por aplicar a base de regras *fuzzy* para os valores de entrada fuzificados, inferindo assim o valor fuzificado à saída correspondente. A fase de defuzificação é usada para associar um valor numérico *crisp* ao valor *fuzzy* de saída obtido no processo de inferência *fuzzy*.

Na entrada do *Sistema de Inferência Fuzzy I*, cada *pixel* tem associado à ele o seu nível de cinza ( $g_{ij}$ ) e a distância euclidiana entre ele e o centróide do núcleo mais próximo ( $d_{ij}$ ). Estes valores são fuzificados através das funções de pertinência, as regras *fuzzy* são aplicados, os resultados são agregadas e o valor final é defuzificado, obtendo-se a classificação do *pixel* em termos da classe a qual este pertence (Figura 22).

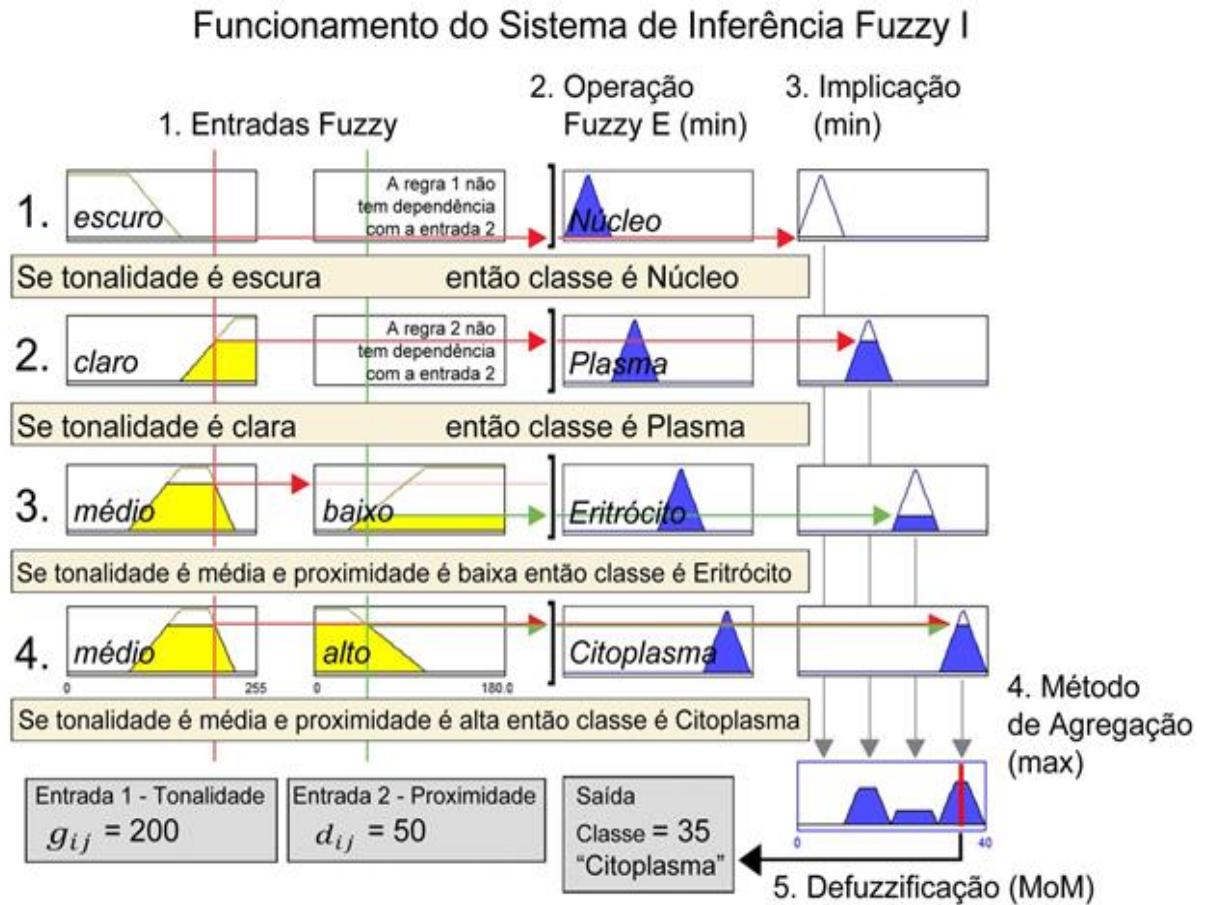


Figura 22. Sistema de Inferência Fuzzy I – funcionamento do sistema para um pixel considerando  $g_{ij} = 200$  e  $d_{ij} = 50$ .

O Sistema de Inferência Fuzzy I proposto possui três variáveis linguísticas, duas variáveis de entrada (*tonalidade* e *proximidade*) e uma variável de saída (*classe*), e suas respectivas funções de pertinência (Figura 23).

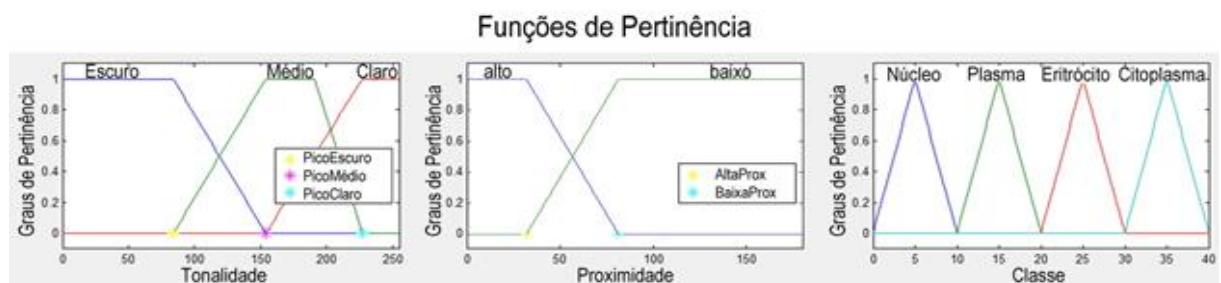


Figura 23. Sistema de Inferência Fuzzy I – Exemplos de Funções de Pertinência para valores PicoEscuro, PicoMédio, PicoClaro, AltaProx e BaixaProx previamente definidos na etapa de pré-processamento para uma imagem específica.

A variável linguística *fuzzy* de entrada *tonalidade* refere-se aos valores de níveis de cinza do histograma da imagem *G*. Três termos linguísticos foram definidos para *tonalidade*: *escura*, *média* e *clara*. Os suportes dos seus respectivos conjuntos *fuzzy* são representados pelos valores dos três picos mais significantes encontrados no histograma do canal *G* na etapa de pré-processamento (*PicoEscuro*, *PicoMédio* e *PicoClaro*). Funções de pertinência trapezoidais são definidas para as tonalidades *escura*, *média* e *clara*, quando mais de um valor assume a pertinência máxima. Qualquer valor abaixo do *PicoEscuro* possui pertinência máxima ao conjunto *fuzzy tonalidade escura*, e qualquer valor acima do *PicoClaro* possui pertinência máxima ao conjunto *fuzzy tonalidade clara*. Finalmente, qualquer valor entre *tonalidade média* e  $((PicoClaro - PicoMédio)/2) + PicoMédio$  possui máxima pertinência ao conjunto *fuzzy tonalidade média*. O universo do discurso é caracterizado pelos valores do canal *G* da imagem.

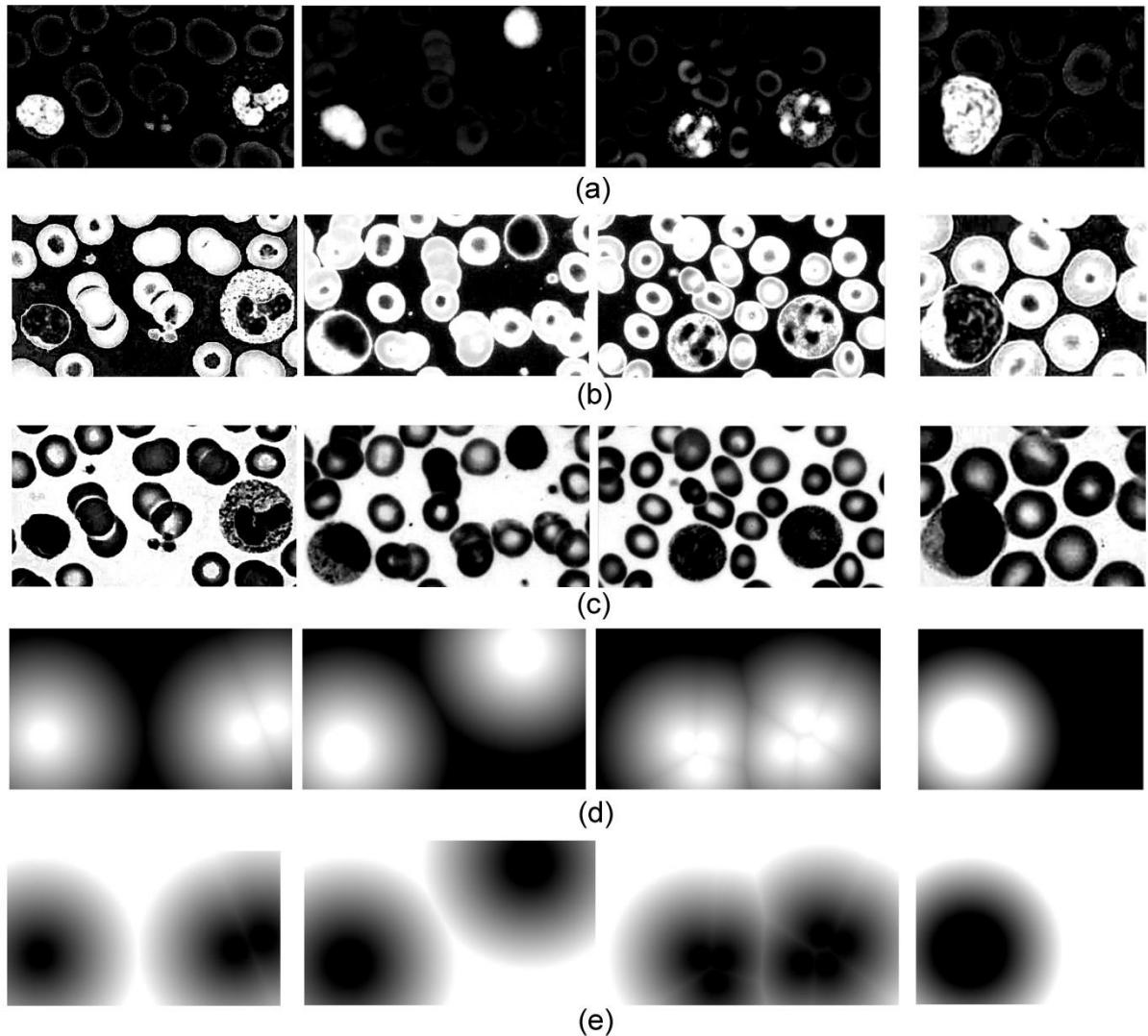
A variável linguística *fuzzy* de entrada *proximidade* refere-se às distâncias euclidianas de cada *pixel* em relação ao centróide da área do núcleo leucocitário mais próximo a ele. Dois termos linguísticos foram definidos para *proximidade*: *alta* e *baixa*. Os suportes dos seus respectivos conjuntos *fuzzy* são representados pelos valores *AltaProx* e *BaixaProx* encontrados nas distâncias euclidianas da matriz *D*, na etapa de pré-processamento. Funções de pertinência trapezoidais são definidas para as proximidades *alta* e *baixa*. Qualquer valor abaixo de *AltaProx* possui pertinência máxima ao conjunto *fuzzy proximidade alta*, e qualquer valor acima do *BaixaProx* possui pertinência máxima ao conjunto *fuzzy proximidade baixa*. O universo do discurso é caracterizado pelos valores da matriz *D*.

Os conceitos que definem os suportes dos conjuntos *fuzzy* para as variáveis linguísticas *tonalidade* e *proximidade* são bem delimitados e, independentemente da imagem, seguem as mesmas definições. No entanto, os valores resultantes da aplicação desses conceitos (*PicoEscuro*, *PicoMédio*, *PicoClaro*, *AltaProx* e *BaixaProx*), os quais representam os valores a serem aplicados às funções de pertinência, são adaptativos, uma vez que refletem a aplicação do conceito para a imagem específica, resultando em valores diferentes e específicos para cada imagem.

A variável linguística *fuzzy* de saída *classe* refere-se às classificações finais da imagem após a segmentação promovida pelo *Sistema de Inferência Fuzzy I*. Quatro conjuntos *fuzzy* foram definidos: núcleo leucocitário (*Núcleo*), citoplasma leucocitário (*Citoplasma*), eritrócitos (*Eritrócitos*) e plasma sanguíneo (*Plasma*).

Na entrada do sistema e para cada *pixel* individualmente, as matrizes *G* e *D*, respectivamente, são fuzificadas, obtendo-se as matrizes contendo os graus de pertinência de

cada um dos seus elementos aos conjuntos *fuzzy tonalidade escura* (Figura 24a), *tonalidade média* (Figura 24b), *tonalidade clara* (Figura 24c), *proximidade alta* (Figura 24d) e *proximidade baixa* (Figura 24e).



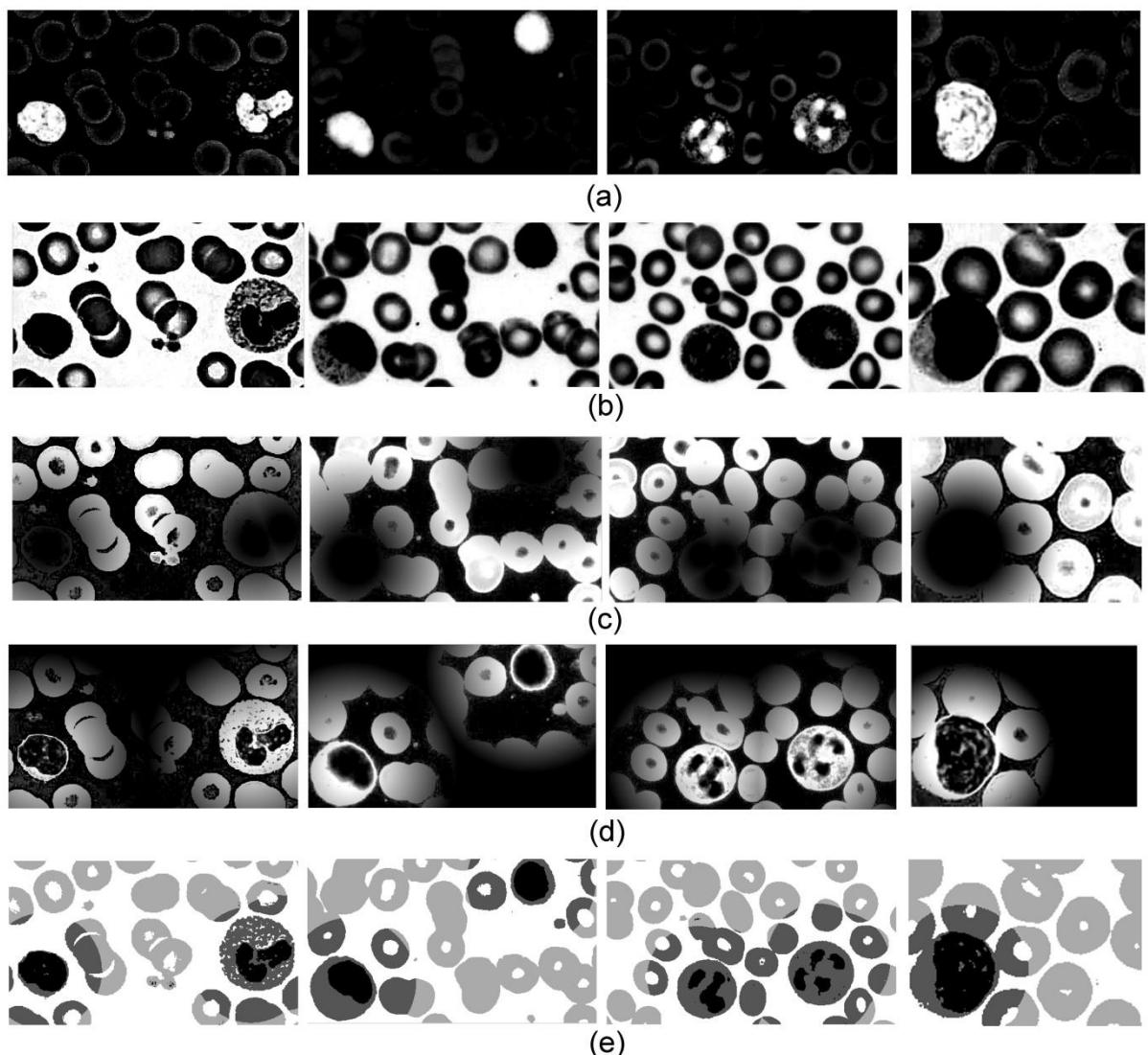
*Figura 24. Processo de fuzzificação (amostras coradas com os corantes Panótico Rápido, Leishman, Rosenfeld e sem coloração especificada, respectivamente): a) fuzzificação para tonalidade escura; b) fuzzificação para tonalidade média; c) fuzzificação para tonalidade clara; d) fuzzificação para proximidade alta; e) fuzzificação para proximidade baixa.*

Após o processo de fuzzificação, as seguintes regras fuzzy são aplicadas às matrizes fuzzificadas:

- Regra 1: Se *tonalidade é escura* então *classe é Núcleo*;
- Regra 2: Se *tonalidade é clara* então *classe é Plasma*;

- Regra 3: Se *tonalidade* é média e *proximidade* é baixa então classe é Eritrócito;
- Regra 4: Se *tonalidade* é média e *proximidade* é alta então classe é Citoplasma.

A agregação das saídas da variável *fuzzy classe* é então realizada (Figuras 25a, 25b, 25c e 25d) e, em seguida o resultado é defuzificado pelo método da Média dos Máximo (MoM) (Figura 25e). Neste método de defuzificação, a saída determinística (não-fuzzy) é obtida pela média dos dois elementos extremos no universo e que correspondem aos maiores valores das funções de pertinência (média dos valores de máxima pertinência). Depois de defuzificação, a imagem é classificada nas quatro regiões de interesse mencionadas anteriormente: *Núcleo*, *Citoplasma*, *Eritrócitos* e *Plasma*.

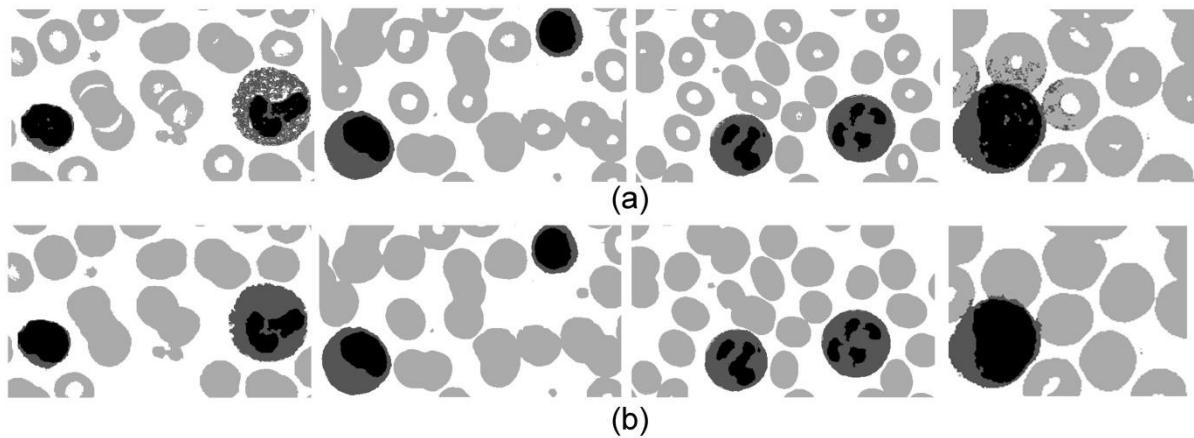


*Figura 25. Regras fuzzy (amostras coradas com os corantes Panóptico Rápido, Leishman, Rosenfeld e sem coloração especificada, respectivamente): a) Regra fuzzy 1 – classe Núcleo; b) Regra fuzzy 2 – classe Plasma; c) Regra fuzzy 3 – classe Eritrócito; d) Regra fuzzy 4 – classe Citoplasma; e) Variável de saída Classe.*

Após a saída do *Sistema de Inferência Fuzzy I*, a classificação é refinada através de um pós-processamento com o objetivo de remover eritrócitos erroneamente classificados como citoplasmas (falsos positivos). Finalmente, a imagem é segmentada. O algoritmo pode ser descrito como:

- Passo 1: Remoção de falsos positivos para citoplasma leucocitários cujos valores RGB foram classificadas incorretamente como eritrócitos.
- Passo 2: Eliminação de citoplasma leucocitários com áreas consideradas pequenas.
- Passo 3: Preenchimento de regiões nas áreas classificadas.
- Passo 4: Segmentação da imagem original em quatro regiões distintas, leucócitos (núcleo e citoplasma), eritrócitos e plasma sanguíneo, de acordo com as respectivas classificações.

Detalhando o pós-processamento da etapa de segmentação, refinamentos são feitos na classificação *fuzzy*. Em primeiro lugar, a região classificada como *Citoplasma* é verificada a fim de determinar se existem *pixels* cujos valores RGB também estão presentes na região classificada como *Eritrócito*. Em caso positivo, este *pixel* é considerado como um falso positivo para *Citoplasma* e será então classificado como *Eritrócito* (Figura 26a). Em seguida, são removidos os *Citoplasmas* cujas áreas forem consideradas pequenas (menos de  $\frac{1}{4}$  da área dos demais) e é feita a operação de preenchimento de regiões. Finalmente a imagem original *IO* original é segmentada em quatro: leucócitos (núcleo e citoplasma), eritrócitos e plasma sanguíneo, de acordo com suas respectivas classificações (Figura 26b). A imagem segmentada contendo a classe *Leucócitos (IL)* será resultante da operação de conjunção (*AND*) entre as classes *Núcleo* de e *Citoplasma*. Esta classe é relevante, tendo em conta a importância da segmentação dos leucócitos para a análise diferencial deste componente e para que sejam realizadas análises comparativas entre os resultados da segmentação obtidos pelo método proposto e os outros resultados relatados na literatura, tendo em vista que, na maioria dos casos, eles fornecem índices de acurácia para a segmentação dos leucócitos como um todo e não individualmente para núcleo e citoplasma leucocitários.



*Figura 26. Pós-processamento (amostras coradas com os corantes Panótico Rápido, Leishman, Rosenfeld e sem coloração especificada, respectivamente): a) Descarte de falsos positivos para citoplasma; b) Classificação final.*

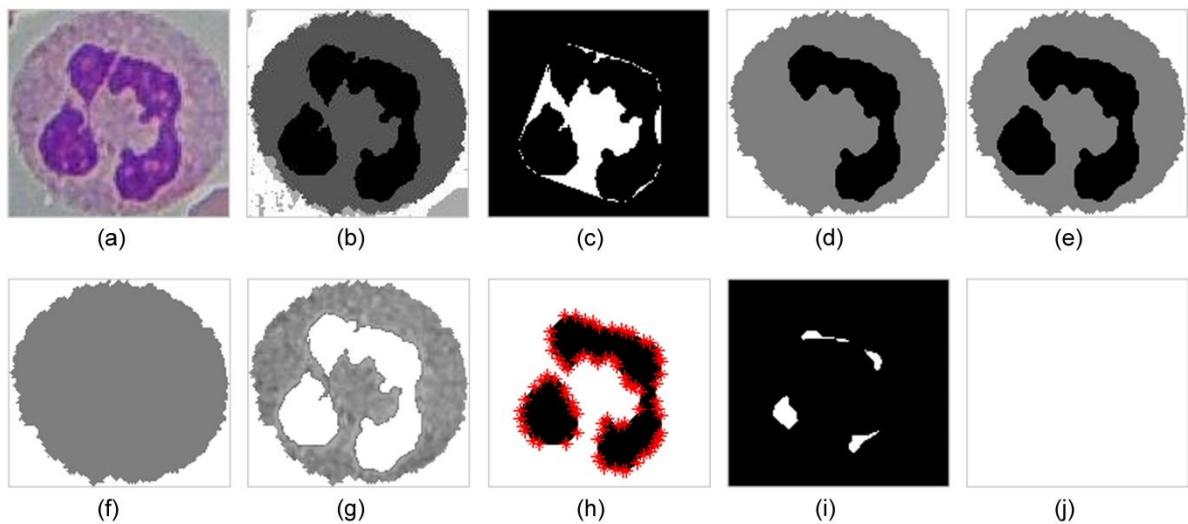
#### 4.3.5 Etapa: Representação e Descrição

Após a etapa de segmentação (Figuras 27a e 27b) ocorre a extração de características dos leucócitos segmentados. A etapa de Representação e Descrição consiste em salientar as características dos leucócitos para que possam ser utilizadas na classificação diferencial dos mesmos nos cinco tipos principais: Basófilos, Eosinófilos, Linfócitos, Monócitos e Neutrófilos.

Os descritores utilizados na literatura concentram-se em descritores de forma e textura. Como descritores de forma são utilizados *convexidade do núcleo*, *razão núcleo-leucócito*, *razão núcleo-citoplasma* e *área do leucócito*. Como descritor de textura, tem-se a *energia do citoplasma*. Além destes, são definidos os descritores *cantos do núcleo*, *erosões do núcleo*, *diâmetro equivalente* e *granularidade*. Os descritores utilizados neste trabalho podem ser definidos como:

- *Convexidade do núcleo*: após calcular a área do menor polígono convexo que contém a região do núcleo segmentado (função `bwconvhull`, MathWorks (2014)), este descritor reflete o percentual desta área que não é núcleo (Figura 27c).
- *Razão núcleo-leucócito*: descritor que reflete o percentual de ocupação da área do maior componente do núcleo segmentado em relação à área total do leucócito (Figura 27d).

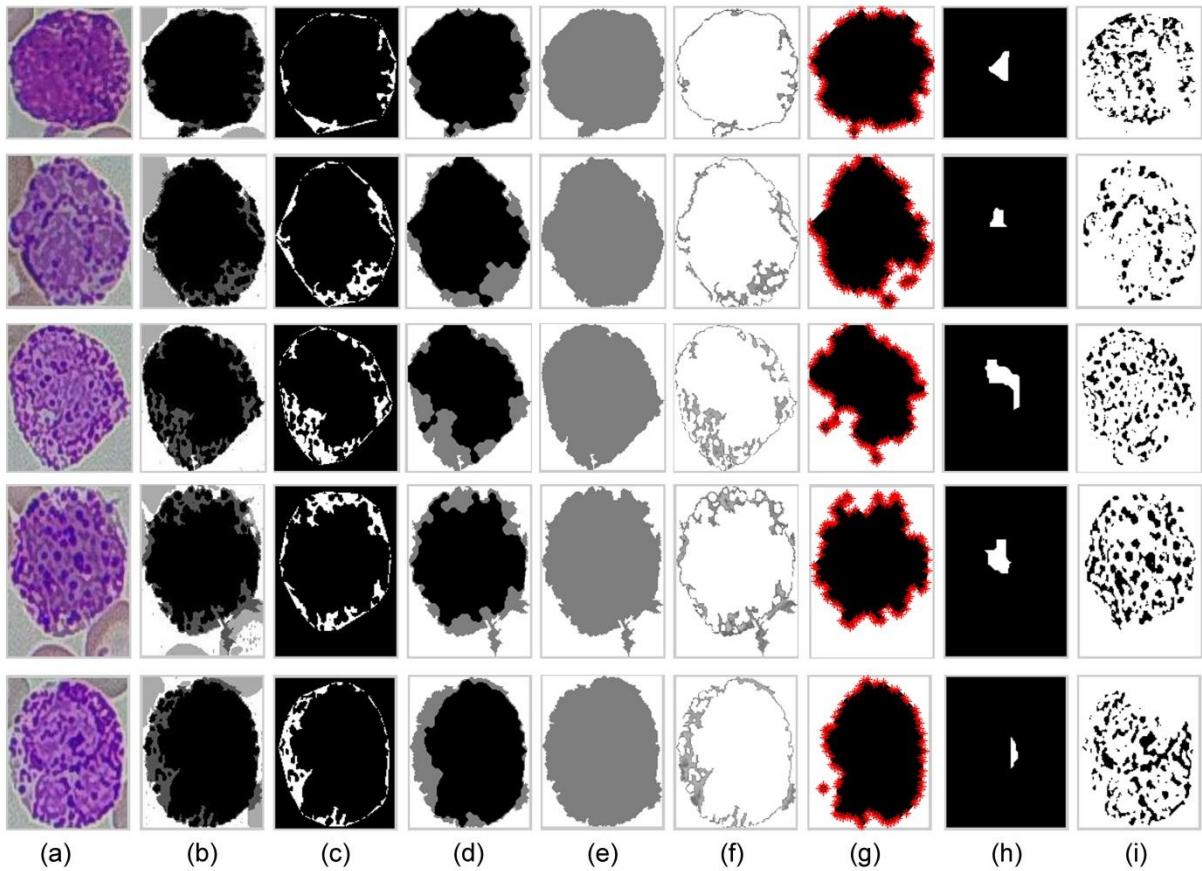
- *Razão núcleo-citoplasma*: após serem efetuadas operações de abertura e fechamento no núcleo segmentado para o descarte de áreas pequenas, este descritor reflete o percentual de ocupação da área do citoplasma em relação à área total do núcleo (Figura 27e).
- *Área do leucócito*: descritor que reflete a área total do leucócito segmentado (Figura 27f).
- *Energia do citoplasma*: após criar a matriz de co-ocorrência,  $MC_{ij}$ , da imagem do citoplasma segmentado, que consiste na frequência em que um *pixel* com valor de nível de cinza  $i$  ocorre horizontalmente adjacente a um *pixel* com o valor  $j$ , a energia do citoplasma, definida como a soma dos quadrados dos elementos na matriz de co-ocorrência (função `graycoprops` (`imagem`, '`Energy`'), MathWorks (2014)), é calculada (Figura 27g).
- *Cantos do núcleo*: descritor que reflete a quantidade total de cantos (função `corner`, MathWorks (2014)) do núcleo segmentado (Figura 27h).
- *Erosões do núcleo*: descritor que reflete a quantidade máxima de erosões (função `imerode`, MathWorks (2014)) que podem ser aplicadas ao núcleo segmentado antes que o mesmo seja completamente erodido (Figura 27i).
- *Diâmetro equivalente*: descritor que reflete o diâmetro do círculo (função `regionprops` (`imagem`, `{'EquivDiameter'}`)), MathWorks (2014)) com a mesma área da região segmentada como leucócito (Figura 27f).
- *Granularidade*: após classificar a imagem original do leucócito (em tons de cinza) em quatro regiões utilizando o método de limiarização *Otsu* multinível (Huang at al, 2012), cujos limiares são os três maiores picos do histograma, e selecionar apenas a região mais escura, são descartados os elementos desta região que apresentam as maiores áreas. Este descritor reflete a quantidade de pontos que sobram após estes descartes (Figura 27j).



*Figura 27. Segmentação e descritores leucocitários: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.*

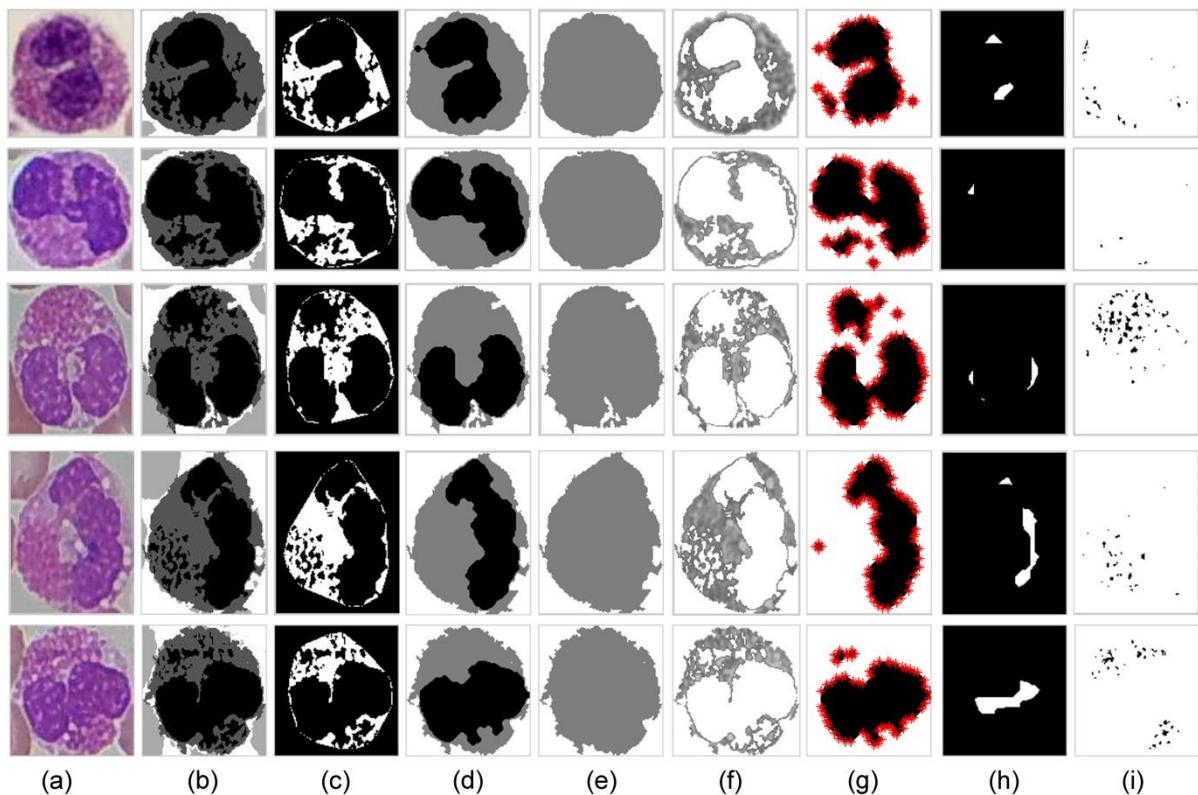
Para cada um dos descritores propostos observa-se que, dependendo do tipo leucocitário, o comportamento é variável.

Os Basófilos (Figuras 28a e 28b) apresentam uma área variável de pontos não-núcleos dentro do polígono convexo (*Convexidade do núcleo* – Figura 28c), a maior parte do leucócito ocupada pelo núcleo (*Razão núcleo-leucócito* – Figura 28d), uma pequena parte de citoplasma comparada com a área do núcleo (*Razão núcleo-citoplasma* – Figura 28d), uma grande área ocupada pelo leucócito com diâmetro equivalente consequentemente alto (*Área do leucócito e Diâmetro equivalente* – Figura 28e), valor alto para a energia do citoplasma (*Energia do citoplasma* – Figura 28f), grande quantidade de cantos ou esquinas no núcleo (*Cantos do núcleo* – Figura 28g), núcleos com áreas consideráveis resultando em um número alto de erosões (*Erosões do núcleo* – Figura 28h) e granularidade alta (*Granularidade* – Figura 28i).



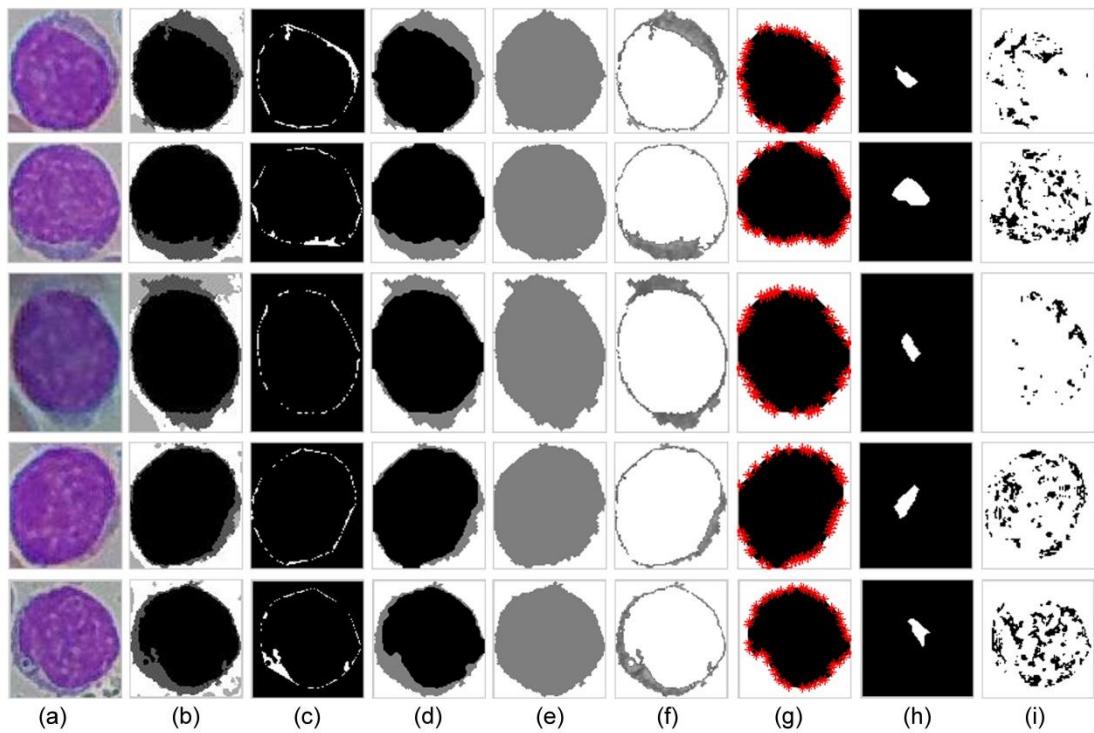
*Figura 28. Basófilos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.*

Os Eosinófilos (Figuras 29 e 29b) apresentam uma grande área não-núcleo dentro do polígono convexo (*Convexidade do núcleo* – Figura 29c), uma parte variável do leucócito ocupada pelo núcleo (*Razão núcleo-leucócito* – Figura 29d), uma grande parte de citoplasma quando comparada com a área do núcleo (*Razão núcleo-citoplasma* – Figura 29d), uma grande área ocupada pelo leucócito com diâmetro equivalente consequentemente alto (*Área do leucócito e Diâmetro equivalente* – Figura 29e), valor para a energia do citoplasma variável (*Energia do citoplasma* – Figura 29f), grande quantidade de cantos ou esquinas no núcleo (*Cantos do núcleo* – Figura 28g), núcleos com áreas variáveis resultando em números baixos e altos de erosões (*Erosões do núcleo* – Figura 29h) e granularidade alta (*Granularidade* – Figura 29i).



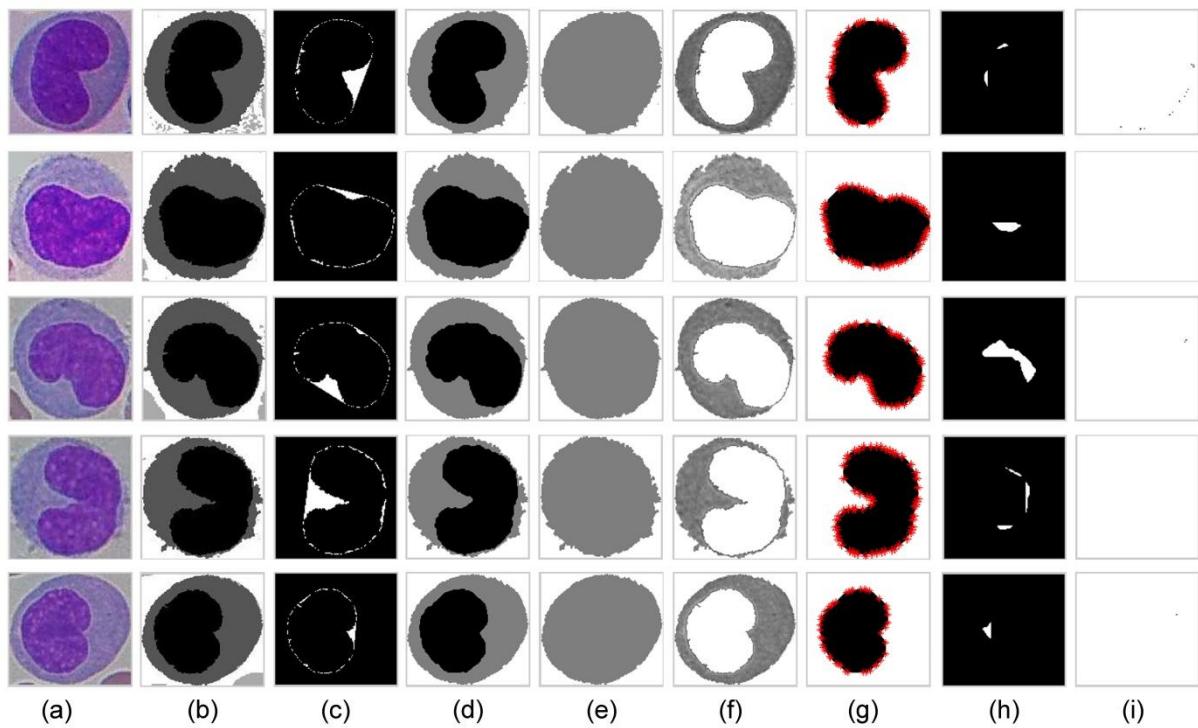
*Figura 29. Eosinófilos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.*

Os Linfócitos (Figuras 30a e 30b) apresentam uma pequena área não-núcleo dentro do polígono convexo (*Convexidade do núcleo* – Figura 30c), uma grande parte do leucócito ocupada pelo núcleo (*Razão núcleo-leucócito* – Figura 30d), uma parte variável de citoplasma quando comparada com a área do núcleo (*Razão núcleo-citoplasma* – Figura 30d), uma pequena área ocupada pelo leucócito com diâmetro equivalente consequentemente pequeno (*Área do leucócito e Diâmetro equivalente* – Figura 30e), valor para a energia do citoplasma alto (*Energia do citoplasma* – Figura 30f), pequena quantidade de cantos ou esquinas no núcleo (*Cantos do núcleo* – Figura 30g), núcleos com áreas grandes porém com área total pequena resultando em um número variável de erosões (*Erosões do núcleo* – Figura 30h) e granularidade variável (*Granularidade*– Figura 30i).



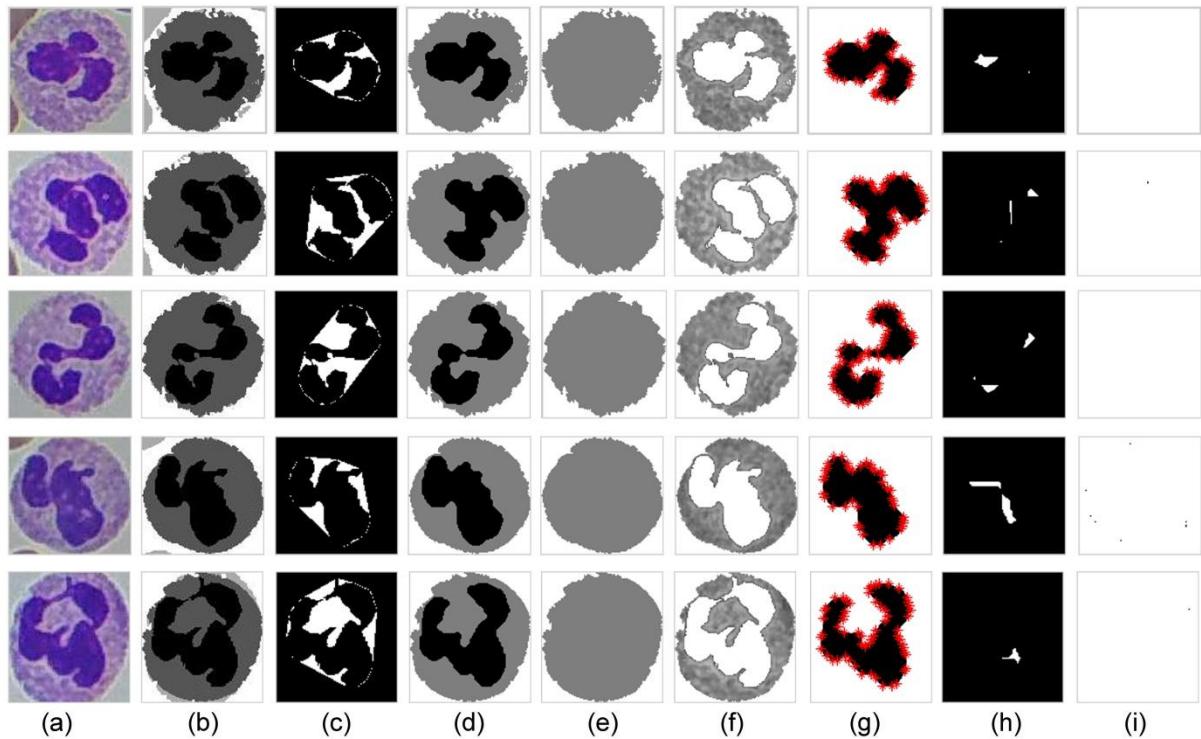
*Figura 30. Linfócitos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.*

Os Monócitos (Figuras 31a e 31b) apresentam uma pequena área não-núcleo dentro do polígono convexo (*Convexidade do núcleo* – Figura 31c), uma parte variável do leucócito ocupada pelo núcleo (*Razão núcleo-leucócito* – Figura 31d), uma grande parte de citoplasma quando comparada com a área do núcleo (*Razão núcleo-citoplasma* – Figura 30d), uma grande área ocupada pelo leucócito com diâmetro equivalente consequentemente alto (*Área do leucócito e Diâmetro equivalente* – Figura 31e), valor para a energia do citoplasma alto (*Energia do citoplasma* – Figura 31f), quantidade de cantos ou esquinas no núcleo variável (*Cantos do núcleo* – Figura 31g), números variáveis de erosões (*Erosões do núcleo* – Figura 31h) e granularidade variável (*Granularidade* – Figura 31i).



*Figura 31. Monócitos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.*

Os Neutrófilos (Figuras 32a e 32b) apresentam uma grande área não-núcleo dentro do polígono convexo (*Convexidade do núcleo* – Figura 32c), uma pequena parte do leucócito ocupada pelo núcleo (*Razão núcleo-leucócito* – Figura 32d), uma grande parte de citoplasma quando comparada com a área do núcleo (*Razão núcleo-citoplasma* – Figura 32d), uma área variável ocupada pelo leucócito mas com diâmetro equivalente constantemente alto (*Área do leucócito e Diâmetro equivalente* – Figura 32e), valor para a energia do citoplasma baixo (*Energia do citoplasma* – Figura 32f), quantidade de cantos ou esquinas no núcleo variável (*Cantos do núcleo* – Figura 32g), núcleos com áreas pequenas resultando em números baixos de erosões (*Erosões do núcleo* – Figura 32h) e granularidade baixa (*Granularidade* – Figura 32i).



*Figura 32. Neutrófilos e seus descritores: a) Leucócito segmentado (RGB); b) Leucócito classificado (núcleo em preto e citoplasma em cinza escuro); c) Convexidade do núcleo; d) Razão núcleo-leucócito; e) Razão núcleo-citoplasma; f) Área do leucócito e diâmetro equivalente; g) Energia do citoplasma; h) Cantos do núcleo; i) Erosões do núcleo; j) Granularidade.*

Para cada leucócito segmentado (imagem *IL*), um vetor de características é calculado de tal forma que cada descritor tenha o seu respectivo valor representado em uma posição do vetor. Na etapa seguinte, cada valor do vetor de características será utilizado como uma variável linguística de entrada do *Sistema de Inferência Fuzzy II*.

#### 4.3.6 Etapa: Reconhecimento e Interpretação

Na fase de Reconhecimento e Interpretação, o *Sistema de Inferência Fuzzy II* classifica diferencialmente o leucócito de acordo com o seu tipo leucocitário, considerando como variáveis linguísticas *fuzzy* de entrada o vetor de características (descritores), suas respectivas funções de pertinência e a base de regras *fuzzy*. Nesta fase, o algoritmo pode ser descrito como:

- Passo 1: Construção das variáveis linguísticas *fuzzy* de entrada *convexidadeN*, *razãoNL*, *razãoNC*, *ÁreaL*, *EnergiaC*, *CantosN*, *ErosõesN*, *diâmetroL* e *granularidade* utilizando os valores do vetor de características da imagem *IL* calculado na fase anterior.
- Passo 2: Fuzificação do vetor de características da imagem *IL* considerando os termos linguísticos das variável de entrada *fuzzy* e suas respectivas funções de pertinência.
- Passo 3: Aplicação da *Base de Regras Fuzzy*.
- Passo 4: Agregação das saídas e defuzificação da variável *fuzzy* de saída *tipo* (*Basófilo*, *Eosinófilo*, *Linfócito*, *Monócito* e *Neutrófilo*) utilizando o método da Média dos Máximos (MoM).

O *Sistema de Inferência Fuzzy II* (Figura 33) tem como objetivo identificar e categorizar os leucócitos segmentados, baseado na extração de suas características, em um dos cinco tipos leucocitários básicos: basófilo, eosinófilo, linfócito, monócito ou neutrófilo. Este sistema também utiliza o Método Mamdani de três fases: fuzificação, procedimentos de inferência e defuzificação.

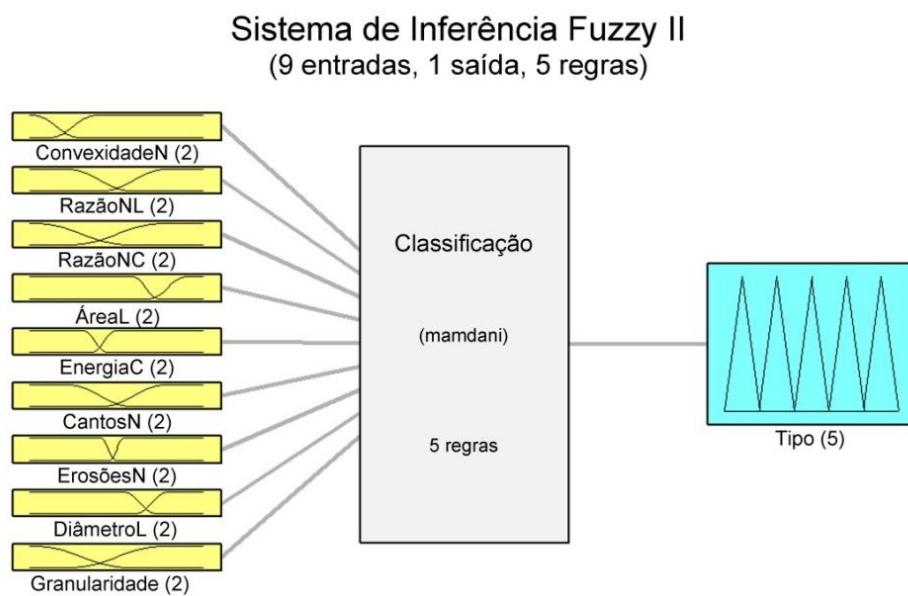


Figura 33. Sistema de Inferência Fuzzy II.

Na entrada do *Sistema de Inferência Fuzzy II*, cada imagem de leucócito previamente segmentado (*IL*) tem associado a ela o vetor de características. Estes valores são fuzificados

através das funções de pertinência, as regras *fuzzy* são aplicadas, os resultados são agregados e o valor final é defuzificado, obtendo-se a classificação do leucócito de acordo com o tipo leucocitário ao qual pertence.

O *Sistema de Inferência Fuzzy II* possui dez variáveis linguísticas, nove variáveis de entrada e uma variável de saída (*tipo*). Para cada variável linguística *fuzzy* de entrada, *convexidadeN*, *razãoNL*, *razãoNC*, *ÁreaL*, *EnergiaC*, *CantosN*, *ErosõesN*, *diâmetroL* e *granularidade*, foram mapeados dois conjuntos *fuzzy*: *baixo* e *Alto* (Figura 34).

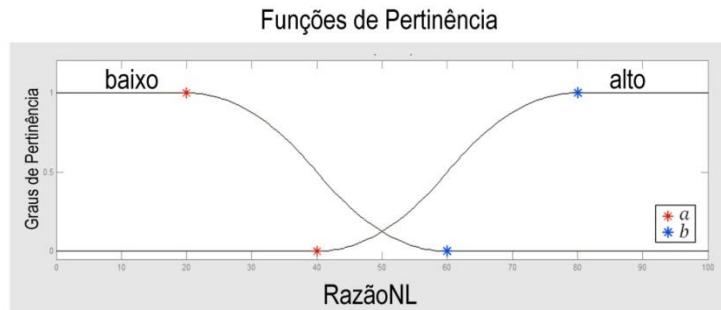


Figura 34. Função de Pertinência para as variáveis de entrada fuzzy (neste exemplo, RazãoNL) do Sistema de Inferência Fuzzy II.

O universo de discurso de cada variável é composto pelas imagens contendo os leucócitos segmentados e as funções de pertinência de cada uma são definidas na forma Z, curva polinomial assimétrica aberta para a esquerda (conforme expressão 11), para os conjuntos *fuzzy* *baixo* e na forma S, curva polinomial assimétrica aberta para a direita (conforme expressão 12), para os conjuntos *fuzzy* *alto*.

$$f(x, a, b) = \begin{cases} 1 - 2 \left( \frac{x-a}{b-a} \right)^2, & x \leq a \\ 2 \left( \frac{x-b}{b-a} \right)^2, & a \leq x \leq \frac{a+b}{2} \\ 0, & \frac{a+b}{2} \leq x \leq b \end{cases} \quad (11)$$

$$f(x, a, b) = \begin{cases} 0, & x \leq a \\ 2 \left( \frac{x-a}{b-a} \right)^2, & a \leq x \leq \frac{a+b}{2} \\ 1 - 2 \left( \frac{x-b}{b-a} \right)^2, & \frac{a+b}{2} \leq x \leq b \\ 1, & x \geq b \end{cases} \quad (12)$$

Após as análises feitas em amostras de leucócitos do banco de imagens, os melhores valores para os suportes dos conjuntos *fuzzy*, suas variáveis linguísticas e respectivas funções de pertinência foram definidos conforme apresentado na Tabela 1. Os valores são definidos conforme as expressões 11 e 12 e suas variáveis  $a$  e  $b$ .

<b>Variável Linguística</b>	<i>baixo</i>		<i>alto</i>	
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
<i>ConvexidadeN</i>	0	30	10	40
<i>RazãoNL</i>	20	60	40	80
<i>RazãoNC</i>	0	60	20	80
<i>ÁreaL</i>	7000	9000	8000	11000
<i>EnergiaC</i>	0,30	0,45	0,35	0,50
<i>CantosN</i>	20	60	40	80
<i>ErosõesN</i>	8	10	9	11
<i>DiâmetroL</i>	80	110	90	120
<i>Granularidade</i>	0	30	10	40

*Tabela 1.* Limiares das funções de pertinência utilizadas na fuzificação do sistema de inferência fuzzy utilizado na etapa de classificação.

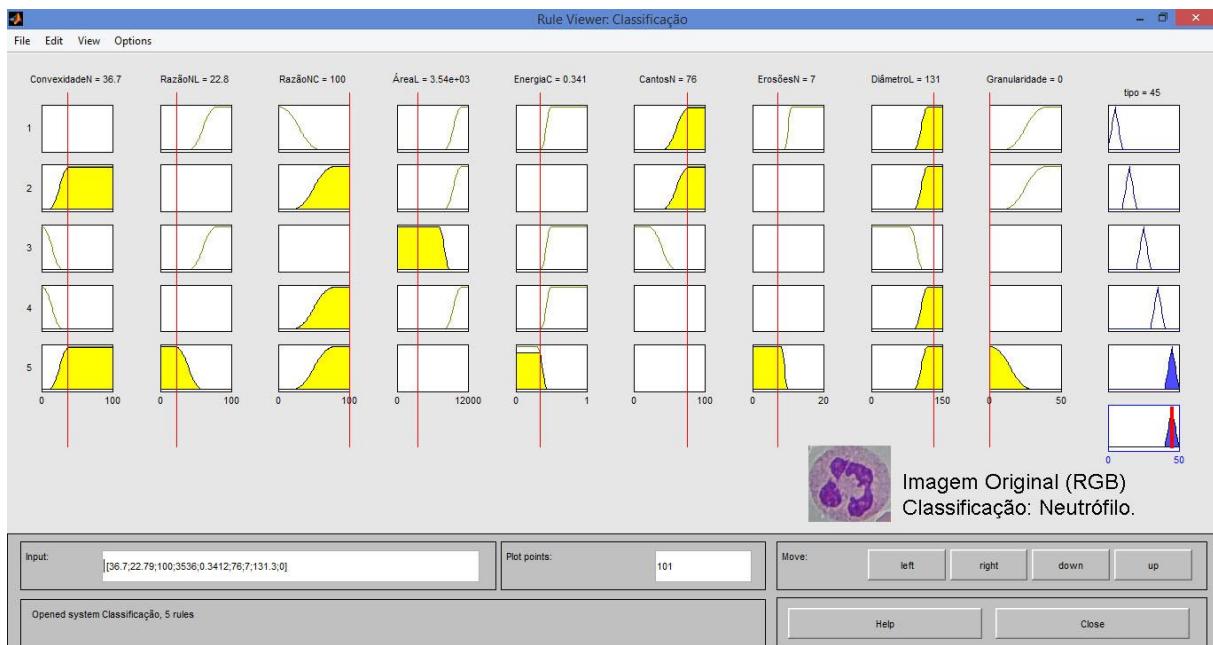
A variável linguística *fuzzy* de saída *tipo* refere-se às classificações finais da imagem após a classificação diferencial do leucócito promovida pelo *Sistema de Inferência Fuzzy II*, quais sejam: *Basófilo*, *Eosinófilo*, *Linfócito*, *Monócito* e *Neutrófilo*.

Para a definição das variáveis de entrada, cada matriz *IL* contendo um leucócito segmentado é fuzificada, obtendo-se resultados contendo os graus de pertinência de cada leucócito aos conjuntos *fuzzy* *ConvexidadeN baixa*, *ConvexidadeN alta*, *RazãoNL baixa*, *RazãoNL alta*, *RazãoNC baixa*, *RazãoNC alta*, *ÁreaL baixa*, *ÁreaL alta*, *EnergiaC baixa*, *EnergiaC alta*, *CantosN baixa*, *CantosN alta*, *ErosõesN baixa*, *ErosõesN alta*, *DiâmetroL baixa*, *DiâmetroL alta*, *Granularidade baixa* e *Granularidade alta*.

Após o processo de fuzificação, as seguintes regras *fuzzy* são aplicadas aos valores fuzificados:

- Regra 1: Se *RazãoNL* é *alta* e *RazãoCN* é *baixa* e *CantosN* é *alto* e *ErosõesN* é *alta* e *DiâmetroL* é *alto* e *ÁreaL* é *alto* e *Granularidade* é *alta* e *EnergiaC* é *alta* então *tipo* é *Basófilo*;
- Regra 2: Se *ConvexidadeN* é *alta* e *RazãoCN* é *alta* e *CantosN* é *alto* e *DiâmetroL* é *alto* e *ÁreaL* é *alta* e *Granularidade* é *alta* então *tipo* é *Eosinófilo*;
- Regra 3: Se *ConvexidadeN* é *baixa* e *RazãoNL* é *alta* e *CantosN* é *baixo* e *DiâmetroL* é *baixo* e *ÁreaL* é *baixa* e *EnergiaC* é *alta* então *tipo* é *Linfócito*;
- Regra 4: Se *ConvexidadeN* é *baixa* e *RazãoCN* é *alta* e *DiâmetroL* é *alto* e *ÁreaL* é *alto* e *EnergiaC* é *alta* então *tipo* é *Monócito*;
- Regra 5: Se *ConvexidadeN* é *alta* e *RazãoNL* é *baixa* e *RazãoCN* é *alta* e *ErosõesN* é *baixa* e *DiâmetroL* é *alto* e *Granularidade* é *baixa* e *EnergiaC* é *baixa* então *tipo* é *Neutrófilo*;

Aplicadas as regras *fuzzy* é feita a agregação das saídas e a variável *fuzzy* de saída tipo é defuzificada pelo método Média dos Máximos (MoM). Finalmente, a imagem é classificada em um dos cinco tipos leucocitários. Para a imagem mostrada na Figura 27, a classificação final pode ser observada na Figura 35.



*Figura 35. Classificação final da imagem mostrada na Figura 27 utilizando o Sistema de Inferência Fuzzy II: Neutrófilo.*

---

# Capítulo 5

## Análises de Resultados

---

### 5.1 Introdução

Neste capítulo são apresentados os experimentos realizados, os resultados encontrados e discussões. Para tal, subdivide-se em:

- **Experimentos:** Apresenta as análises feitas após a aplicação das técnicas desenvolvidas nas imagens previamente elaboradas.
- **Resultados:** Apresenta os índices de *Gold Standards* e Acurácia encontrados na Segmentação dos componentes sanguíneos e classificação diferencial dos leucócitos.
- **Discussões:** Análise dos resultados obtidos.

### 5.2 Experimentos

Foram analisadas 530 imagens microscópicas de esfregaços sanguíneos de diferentes tamanhos e colorações hematológicas contendo componentes sanguíneos normais, dentre os quais 742 leucócitos incluindo todos os cinco tipos leucocitários analisados por este trabalho, e vizinhança entre citoplasmas leucocitários e eritrócitos. Estas colorações diferem de amostra para amostra, tornando este um conjunto de imagens heterogêneo e com um alto grau de dificuldade nas tarefas de segmentação e classificação (Figura 36).

As imagens foram processadas em quatro grupos: amostras utilizando um processo de coloração bem definido (grupo 1 – G1, grupo 2 – G2, grupo 3 – G3) e aquelas que utilizaram um processo de coloração não especificado (grupo 4 – G4). Os primeiros três grupos contêm as 415 imagens provenientes do Hemonorte, cujo processo de aquisição foi detalhado no capítulo anterior (G1 – 150 amostras – Panótico Rápido; G2 – 140 amostras – Leishman e G3 – 125 amostras – Rosenfeld) e o quarto grupo (G4) contém as 115 imagens provenientes do *BloodLine Imagem Atlas* cujas colorações não foram especificados (Figura 36).

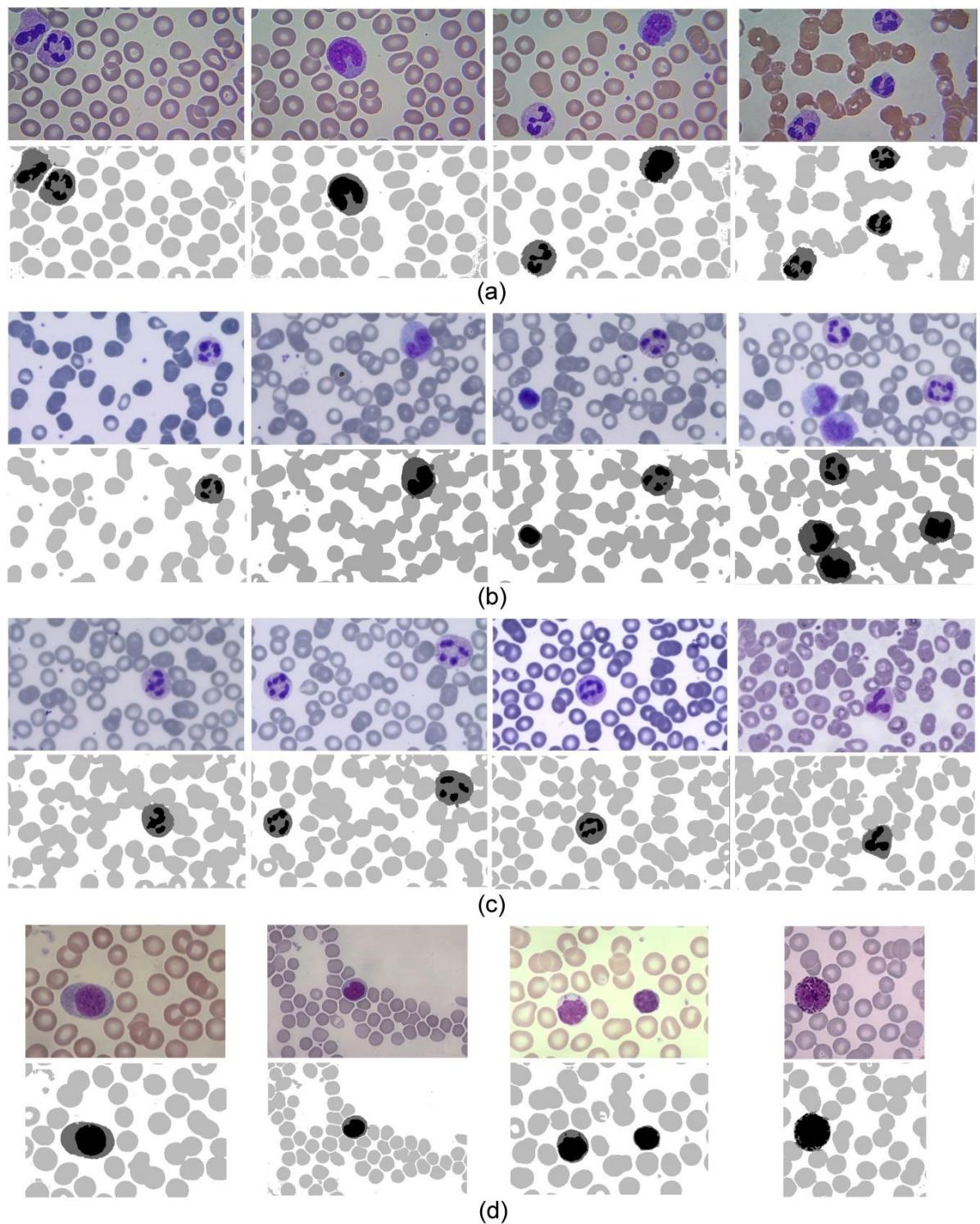
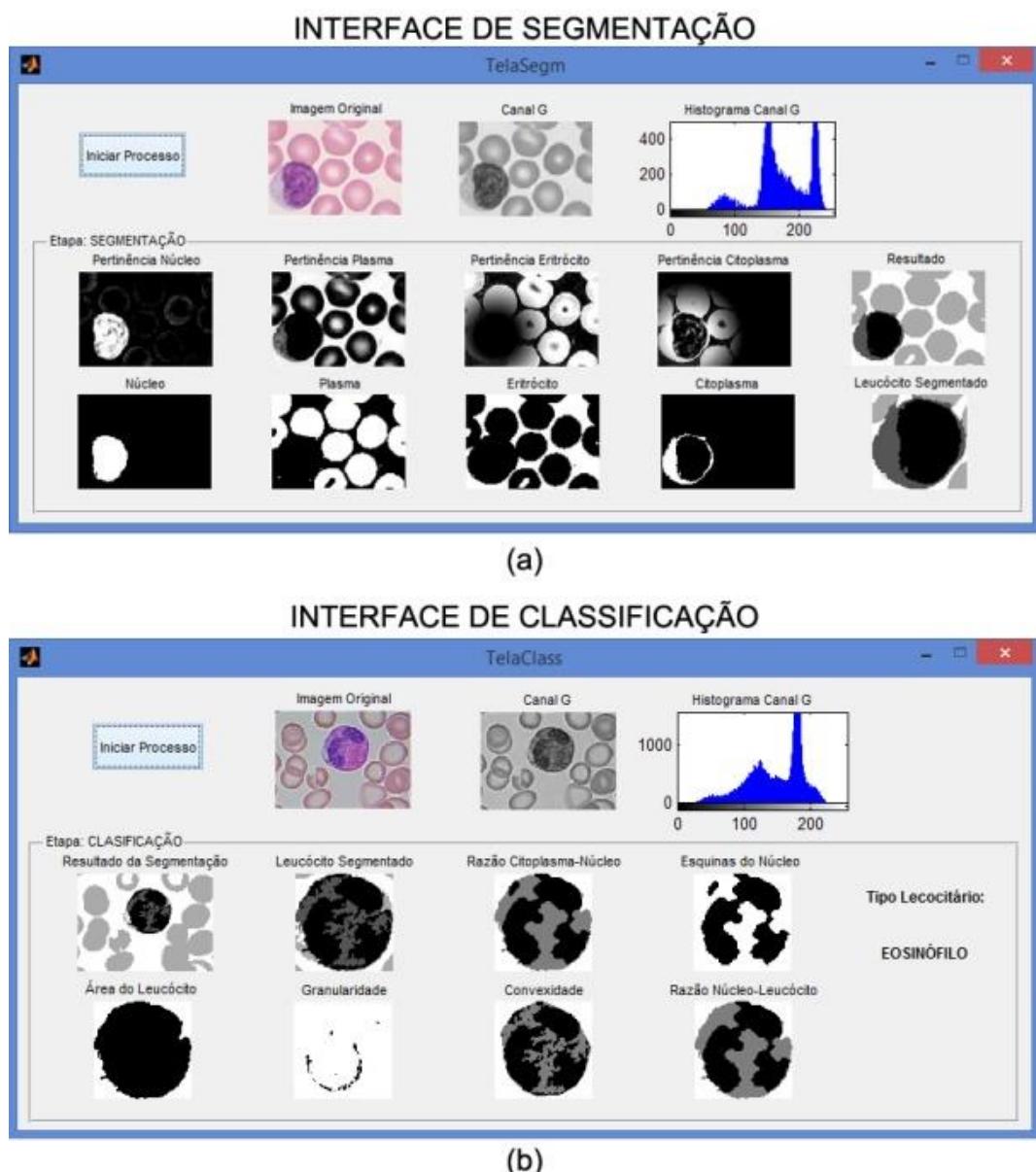


Figura 36. Exemplos de imagens originais (RGB) e segmentadas pelo Sistema de Inferência Fuzzy I proposto, cujas colorações hematológicas são, respectivamente: a) G1: Panótico Rápido; b) G2: Leishman; c) G3: Rosenfeld; d) G4: coloração não especificada.

Para avaliar a técnica desenvolvida, foi desenvolvido um sistema computacional utilizando o software MathWorks Matlab R2014a (Mathworks, 2014) contendo os *Sistemas de Inferência Fuzzy I e II*. O sistema consiste em um algoritmo que processa o banco de imagens em lote para que os percentuais possam ser calculados para cada grupo de imagens processado. Também foram desenvolvidas duas interfaces para processamento individual de cada imagem e visualização dos resultados (Figura 37).



*Figura 37. Interfaces desenvolvidas para o processamento individual das imagens: a) Interface de Segmentação; b) Interface de Classificação.*

## 5.3 Resultados

Os resultados são apresentados separadamente para a segmentação de componentes sanguíneos e classificação diferencial de leucócitos.

### 5.3.1 Segmentação

Todas as imagens foram processadas e os seus resultados na etapa de segmentação foram comparados com a segmentação manual, previamente elaborada e aprovada por especialistas, para os cálculos dos índices de *Gold Standards*. Considerando que os componentes foram identificados em 100% dos casos, a análise feita avaliou a área encontrada para cada componente pela segmentação automática comparada com a área segmentada manualmente por especialistas. As métricas consideradas para os índices *Gold Standards* foram Verdadeiro Positivo (TP), Falso Positivo (FP), Verdadeiro Negative (TN) e Falso Negativo (FN). Além dessas métricas, o índice de Acurácia (AC) também foi calculado conforme Aghajari e Damayanti (2011), de acordo com a expressão 13.

$$AC = ((TP + TN) / (TP + FN + TN + FP)) * 100 \quad (13)$$

Os resultados obtidos para as segmentações são apresentados na Tabela 2.

Componente	Grupo	TP (%)	TN (%)	FP (%)	FN (%)	AC (%)
<b>Leucócito</b>	G1	94,41	99,91	0,09	5,59	97,16
	G2	96,80	99,81	0,19	3,20	98,31
	G3	94,35	99,85	0,15	5,65	97,10
	G4	93,91	99,41	0,59	6,09	96,66
	Média	94,87	99,75	0,25	5,13	97,31
<b>Núcleo Leucocitário</b>	G1	92,83	99,93	0,07	7,17	96,38
	G2	91,80	99,88	0,12	8,20	95,84
	G3	83,10	99,91	0,09	16,90	91,50
	G4	91,24	99,57	0,43	8,76	95,41
	Média	89,74	99,82	0,18	10,26	94,78
<b>Citoplasma Leucocitário</b>	G1	85,12	99,81	0,20	14,88	92,46
	G2	89,48	99,73	0,27	10,52	94,60
	G3	85,46	99,72	0,28	14,54	92,59
	G4	79,94	98,85	1,15	20,06	89,40
	Média	85,00	99,53	0,47	15,00	92,26

	G1	95,75	94,35	5,65	4,25	95,05
	G2	98,75	92,63	7,37	1,25	95,69
<b>Eritrócito</b>	G3	99,29	92,24	7,76	0,71	95,77
	G4	98,04	92,05	7,95	1,96	95,05
	<b>Média</b>	<b>97,96</b>	<b>92,82</b>	<b>7,18</b>	<b>2,04</b>	<b>95,39</b>
	G1	93,83	96,08	3,92	6,17	94,96
<b>Plasma Sanguíneo</b>	G2	92,09	99,07	0,93	7,91	95,58
	G3	91,77	99,53	0,47	8,23	95,65
	G4	89,08	99,05	0,95	10,92	94,07
	<b>Média</b>	<b>91,69</b>	<b>98,43</b>	<b>1,57</b>	<b>8,31</b>	<b>95,06</b>

*Tabela 2. Resultados percentuais (Gold Standards) encontrados para a segmentação dos componentes sanguíneos para os grupos de imagens G1, G2, G3, G4 e média dos resultados para todos os grupos: Verdadeiro Positivo (TP), Verdadeiro Negativo (TN), Falso Positivos (FP), Falso Negativo (FN) e Acurácia (AC).*

### 5.3.2 Classificação

Após a segmentação e a extração dos descritores, os leucócitos foram classificados, utilizando o *Sistema de Inferência Fuzzy II* proposto, em um dos cinco tipos leucocitários. Os resultados obtidos pelo classificador *fuzzy* desenvolvido neste trabalho podem ser vistos na matriz de confusão apresentada na Tabela 3.

Tipos	Basófilo	Eosinófilo	Linfócito	Monócito	Neutrófilo	AC
Basófilo	33	1	0	0	0	97,05
Eosinófilo	2	154	3	2	1	95,06
Linfócito	1	0	183	2	0	98,39
Monócito	0	6	1	106	1	92,98
Neutrófilo	0	6	1	0	239	97,15
AC Média						<b>96,36</b>

*Tabela 3. Matriz de confusão, precisão e acurácia geral para as classificações dos cinco tipos de leucócitos.*

## 5.4 Discussões

Diferentes grupos de imagens foram processados neste trabalho objetivando a

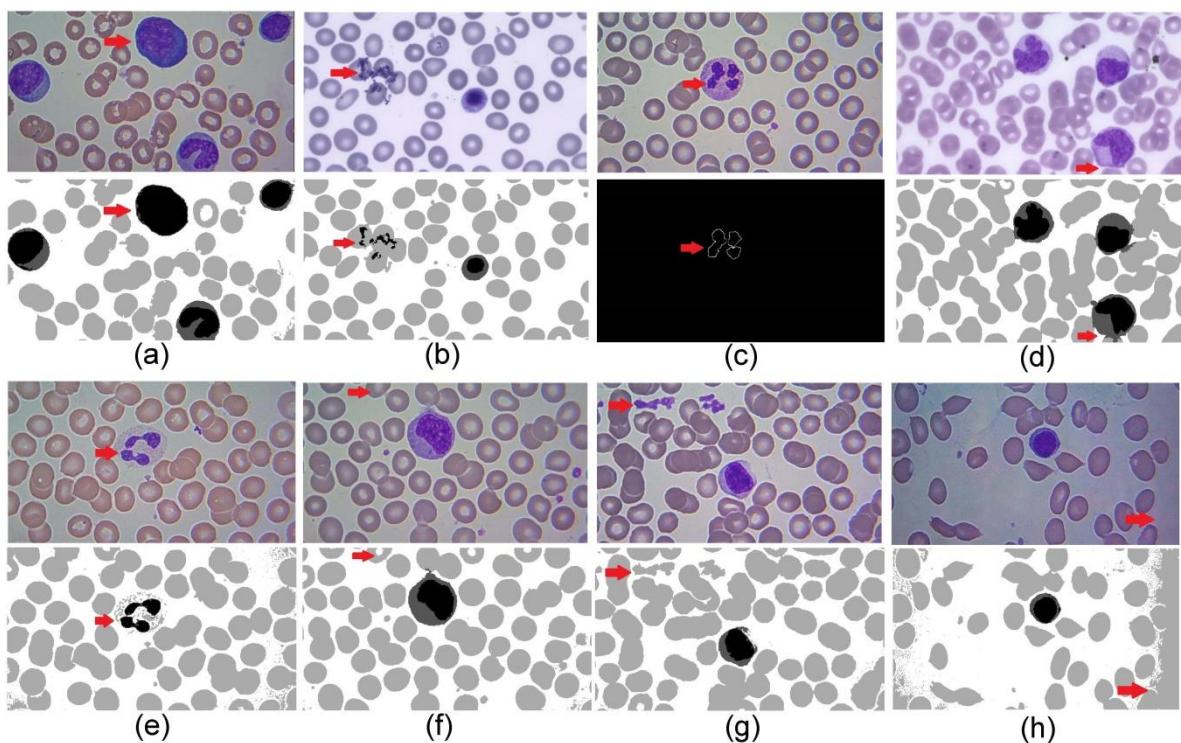
segmentação automática dos componentes sanguíneos e classificação diferencial dos leucócitos com um alto nível de independência da técnica de coloração hematológica aplicada à amostra. Taxas significativas de TP e TN foram obtidas, indicando os percentuais de sucesso na identificação da presença e da ausência, respectivamente, dos componentes nas amostras, enquanto que as taxas de FP e FN indicaram os percentuais de erros na identificação da presença e da ausência dos componentes, quando em comparação com as suas respectivas segmentações manuais.

As discussões são apresentadas separadamente para segmentação e classificação.

#### **5.4.1 Segmentação**

A análise das 530 imagens, especificamente para leucócitos (núcleos e citoplasmas), mostrou altos índices de precisão na segmentação desses componentes, demonstrando um alto nível de relevância para a porcentagem de acertos em relação aos erros obtidos na segmentação. Para os núcleos leucocitários, observou-se que o valor obtido para FP foi proveniente da semelhança entre as colorações do núcleo e citoplasma leucocitários presentes em algumas amostras (Figura 38a), fato este que também interferiu na taxa de FN para citoplasmas leucocitários, e da presença de artefatos resultantes do processo de preparação (Figura 38b). Já os valores de FN foram provenientes principalmente da incapacidade de identificar a borda dos núcleos leucocitários existentes na segmentações manuais (Figura 38c). Quanto aos citoplasmas leucocitários, foram observados valores de FP quando os eritrócitos adjacentes aos citoplasmas leucocitários foram incorretamente identificados (Figura 38d) e de FN principalmente como resultados dos processos de coloração de citoplasmas leucocitários, que, dependendo da preparação da amostra ou o tempo de exposição ao corante reagente, não foram corados de forma significativa (Figura 38e).

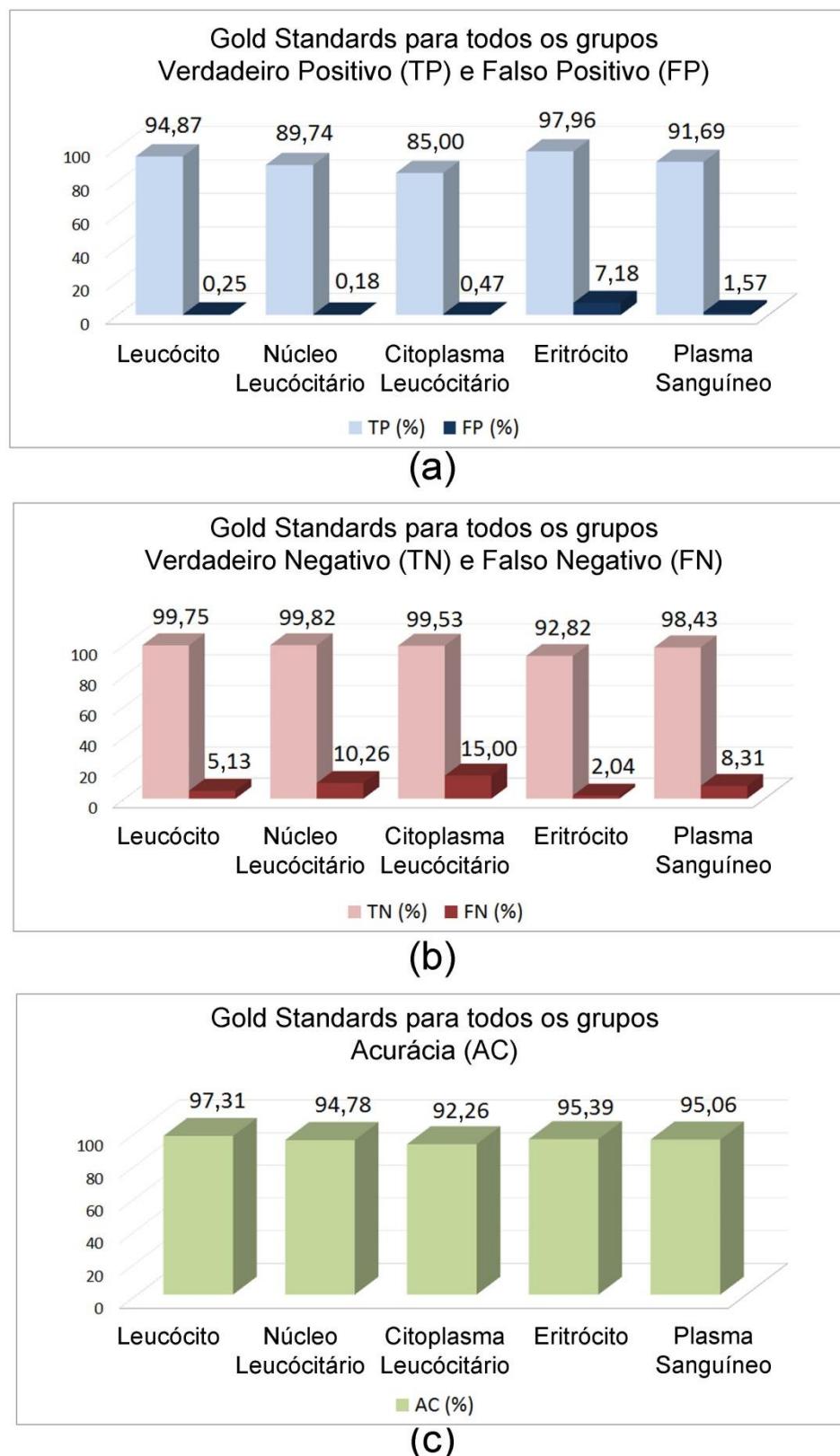
No que diz respeito à classificação de eritrócitos, a taxa de FN foi proveniente principalmente da operação de preenchimento de regiões que não foi adequadamente realizada para os eritrócitos localizados nas extremidades das imagens segmentadas (Figura 38f). Em tais casos, os centros dos eritrócitos foram incorretamente classificados como plasma sanguíneo, o que também refletiu negativamente sobre a taxa de FP destes. A taxa de FP foi observada quando da presença de plaquetas, componentes sanguíneos não cobertos por este trabalho (Figura 38g), artefatos e perda do foco do microscópio nas bordas da imagem (Figura 38h). Neste último caso, o plasma sanguíneo, que exibiu um tom mais escuro, foi incorretamente classificado como eritrócito, impactando negativamente em sua taxa de FN.



*Figura 38. Exemplos de imagens com resultados negativos: a) FP para núcleos leucocitários e FN para citoplasmas leucocitários - Semelhança entre as suas cores; b) FP para núcleos leucocitários - presença de artefatos; c) FN para núcleos leucocitários - borda da segmentação manual; d) FP para citoplasmas leucocitários - eritrócitos adjacentes; e) FN para citoplasmas leucocitários - não foram corados de forma significativa; f) FN para eritrócitos - localizados nas bordas; g) FP para eritrócitos - presença de plaquetas; h) FP para eritrócitos - perda de foco do microscópio.*

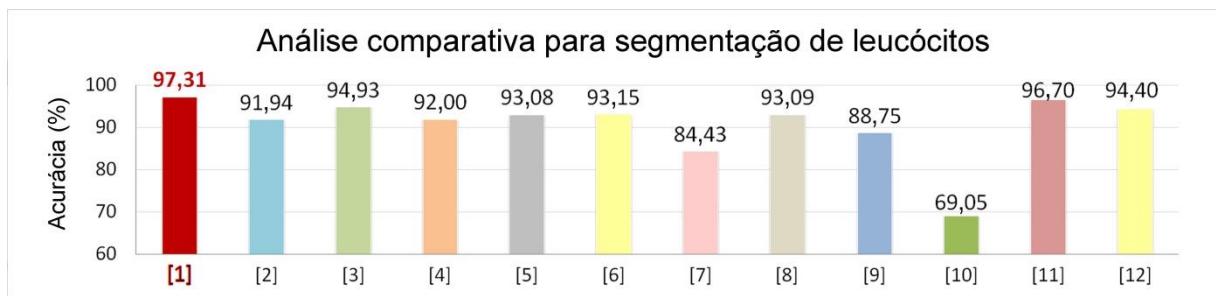
Os resultados obtidos para os grupos G1, G2 e G3 foram semelhantes para a segmentação de leucócitos, eritrócitos e plasma sanguíneo, e apresentaram pequenas variações na segmentação dos componentes dos leucócitos (núcleo e citoplasma), em decorrência das colorações aplicadas às amostras. Os grupos G1, G2, G3 apresentaram resultados semelhantes aos obtidos pelo grupo G4, demonstrando um elevado grau de independência em relação à técnica de coloração hematológica aplicada à amostra.

As taxas de TP, TN, FP, FN e AC obtidas para a segmentação demonstram a capacidade do método para obter êxito no que se propôs a realizar (Figuras 39a, 39b e 39c).



*Figura 39. Percentuais médios obtidos para as métricas Gold Standards e Acurácia da técnica desenvolvida para segmentação. a) Verdadeiro Positivo (TP) e Falso Positivo (FP); b) Verdadeiro Negativo (TN) e falso negativo (FN); c) Acurácia (AC).*

Os resultados médios de acurácia obtidos para todos os componentes mostram a relevância da técnica desenvolvida na segmentação e, quando comparada a outros métodos utilizados especificamente para a segmentação de leucócitos, e aplicados os mesmos valores de medição (Figura 40), também se mostrou viável.



*Figura 40. Análise comparativa para a segmentação de leucócitos entre os percentuais de acurácia (AC) utilizando a técnica desenvolvida e os percentuais disponíveis na literatura, sendo [1] Técnica Fuzzy Proposta, [2] Zheng et. al. (2014), [3] Jati et. al. (2014) – com ruído, [4] Putzu & Ruberto (2013), [5] Ramesh et. al. (2012), [6] Fatichah et. al. (2012) - núcleo leucocitário, [7] Fatichah et. al. (2012) – citoplasma leucocitário, [8] Rezatofighi & Soltanian-Zadeh (2011), [9] Ko et. al. (2011) – núcleo leucocitário, [10] Ko et. al. (2011) - citoplasma leucocitário, [11] Hamghalam & Aytollahi (2009) e [12] Ramoser et. al. (2005).*

Quando os resultados obtidos para segmentação de leucócitos utilizando a técnica desenvolvida foram comparados aos descritos em outras abordagens encontradas na literatura, observou-se que os demonstrados neste trabalho eram iguais ou superiores em termos de número de amostras processadas e maior diversidade dos processos de coloração utilizados nas preparações das amostras (Tabela 4).

Método de Segmentação de Leucócitos	Todas as Amostras	Corantes	AC (%)
<b>Técnica Fuzzy Proposta</b>	<b>530</b>	<b>Fast Panoptic, Leishman, Rosenfeld e não especificada</b>	<b>97,31</b>
Zheng <i>et. al.</i> (2014)	54	Coloração Rápida	91,94
Jati <i>et. al.</i> (2014)	110	Leishman	93,90 – 94,93 (com ruído)
Cuevas <i>et. al.</i> (2013)	517	Não especificada	98,26
Putzu & Ruberto (2013)		Não especificada	92,00
Ramesh <i>et. al.</i> (2012)	320	Wright	93,08 93,15
Faticahah <i>et. al.</i> (2012)	100	Não especificada	(núcleo leucocitário) 84,43 (citoplasma leucocitário)
Rezatofighi & Soltanian-Zadeh (2011)	150	Wright-Giemsa	93,09
Ko <i>et. al.</i> (2011)	260	Não especificada	88,75 (núcleo leucocitário) 69,05 (citoplasma leucocitário)
Hamghalam & Aytollahi (2009)	30	Giemsa	96,70
Khashman (2009)	360	Não especificada	99,17
Zamani & Safabakhsh (2006)	30	Não especificada	95
Ramoser <i>et. al.</i> (2005)	534	Não especificada	94,40
Jiang <i>et. al.</i> (2003)	8	Não especificada	91,2
Sinha & Ramakrishna (2003)	115	May-Grünwald-Giemsa	80

*Tabela 4. Análise comparativa entre índices de acurácia (AC) para segmentação de leucócitos utilizando a técnica fuzzy proposta e os descritos na literatura.*

Outros resultados podem ser visualizados no Apêndice A.

#### 5.4.2 Classificação

Analizando as 742 imagens contendo leucócitos segmentados provenientes dos grupos G1, G2 e G3, foram encontrados percentuais de acurácia de 97,05%, 95,06%, 98,39%, 92,98% e 97,15% para as classificações de basófilos, eosinófilos, linfócitos, monócitos e neutrófilos, respectivamente. Estes altos índices resultantes demonstraram uma capacidade

relevante do método para a classificação diferencial e, quando comparado com outras abordagens encontradas na literatura para o mesmo problema e utilizando as mesmas taxas de medição (Tabela 4), o método proposto também se mostrou relevante. (Tabela 5).

Método	Quantidade de Leucócitos Analisados					Coloração	Classificação Diferencial (AC %)								
	B	E	L	M	N		B	E	L	M	N				
Técnica Fuzzy Proposta	34	162	186	114	246	Panótico Rápido, Leishman e Rosenfeld	97,06	95,07	98,39	93,0	97,16				
Rajendran <i>et. al.</i> (2011)	não especificado					não especificado	95,0								
Rezatofighi & Soltanian-Zadeh (2011)	50	19	29	24	28	Gismo-Right	98,64	94,74	93,10	95,83	100,0				
Ramesh <i>et. al.</i> (2012)	5	56	298	65	830	Wright	100,0	66,67	88,51	80,17	98,25				
Hamghalam & Aytollahi (2009)	90					Giemsa	não especificado								
Hiremath <i>et. al.</i> (2010)	0	0	34	12	29	Giemsa	0	0	98,0	96,50	97,50				
Mircic & Jorgovanović (2006)	0	200				não especificado	86,0								
Putzu & Ruberto (2013)	245					não especificado	75,10								
Huang <i>et. al.</i> (2012)	não especificado					não especificado	98,0	98,0	84,30	93,30	81,30				
Xie <i>et. al.</i> (2010)	28	34	64	40	64	não especificado	89,29	94,12	98,48	75,5	90,63				
Ramoser <i>et. al.</i> (2005)	32	79	110	103	210	não especificado	87,50	98,0	92,80	88,40	78,90				

*Tabela 5. Análise comparativa entre taxas de acurácia (AC) para classificação diferencial de leucócitos (Basófilos – B, Eosinófilos – E, Linfócitos – L, Monócitos – M, Neutrófilos – N) encontradas nos trabalhos descritos na literatura e resultantes da técnica desenvolvida neste trabalho para os grupos G1, G2 e G3.*

---

# Capítulo 6

## Conclusões

---

### 6.1 Introdução

Neste capítulo são apresentadas as conclusões, as contribuições, os trabalhos futuros e as publicações. Para tal, subdivide-se em:

- **Conclusões:** Apresenta um fechamento do trabalho contendo as conclusões a partir dos objetivos específicos elencados anteriormente.
- **Contribuições:** Apresenta as contribuições.
- **Trabalhos Futuros:** Apresenta as perspectivas para trabalhos futuros.
- **Publicações:** Apresenta as publicações.

### 6.2 Conclusões

A detecção de componentes sanguíneos e a diferenciação entre eles é um importante tópico da área hematológica e possibilita ao profissional a análise do paciente e a detecção de anomalias e doenças. Faz-se necessário que sistemas computacionais auxiliem neste trabalho oferecendo ferramentas de apoio que diminuam o custo e aumentem a precisão. Tais ferramentas devem ter uma inteligência incorporada que contenha o conhecimento especializado necessário às tarefas decorrentes destas análises.

Este trabalho apresentou um conjunto de técnicas utilizando Sistemas de Inferência *Fuzzy* capaz de segmentar, em uma imagem microscópica de esfregaço sanguíneo, as áreas de núcleo e citoplasma leucocitários, eritrócitos e plasma e classificar diferencialmente os leucócitos em cinco tipos leucocitários: basófilos, eosinófilos, linfócitos, monócitos e neutrófilos.

Durante a execução do trabalho, o estado da arte foi estudado para que as limitações pudessem ser detectadas e trabalhadas. Visando minimizar as taxas de erros na segmentação e

classificação, foi incorporado ao conjunto de técnicas etapas de pré e pós-processamento que mitigaram as limitações previamente identificadas.

Visando aumentar a precisão das análises, as técnicas foram desenvolvidas para que a segmentação fosse automática e adaptativa para que, a cada imagem, um mesmo conceito pudesse ser automaticamente adaptado ao novo contexto sem a interferência do usuário. Desta forma, o conhecimento especialista foi incorporado nas bases de conhecimento agregando valor às técnicas e aumentando as taxas de sucesso nos experimentos. Com este conhecimento foram definidas funções de pertinência *fuzzy* capazes de realizar as tarefas de segmentação e classificação a partir de um olhar especializado.

Além disso, um banco de imagens real foi construído com o auxílio dos especialistas da área hematológica utilizando variadas técnicas de coloração para que os resultados pudessem ter um caráter intrínseco de fidelidade, generalidade e realidade.

Os resultados encontrados foram comparados às segmentações manuais, também validadas por especialistas, de forma que pudessem expressar valores corretos e não perspectivas de sucesso. Os experimentos mostraram resultados expressivos, aumentando o campo de pesquisa na área e contribuindo com a inclusão uma nova abordagem *fuzzy* para o problema de segmentação e classificação celular.

## 6.3 Contribuições

Em geral, os resultados obtidos em todas as métricas e análises aplicadas foram significativos, encorajando o uso das técnicas desenvolvidas neste trabalho. Como contribuições principais podem ser citadas:

- (i) A segmentação e classificação dos componentes sanguíneos com um alto nível de independência da técnica de coloração hematológica aplicada à amostra;
- (ii) A identificação automática na segmentação, para cada imagem específica, dos suportes dos conjuntos *fuzzy* utilizados, o que não é comum em lógica *fuzzy*, cujos onde os suportes são muitas vezes obtidos heuristicamente;
- (iii) A identificação de vários componentes na segmentação, como núcleo e citoplasma leucocitários, eritrócitos e áreas de plasma sanguíneo.
- (iv) A minimização das taxas de Falsos Positivos e Falsos Negativos na segmentação, especialmente em decorrência da vizinhança entre ás células e de descartes de áreas com tamanhos pré-definidos, respectivamente;

- (v) A definição de descritores baseados na tonalidade da amostra e na proximidade entre núcleos e citoplasmas leucocitários na segmentação, permitindo a diferenciação entre eritrócitos e citoplasmas leucocitários, cujos tons de cinza são extremamente semelhantes;
- (vi) A construção de funções de pertinência *fuzzy* automáticas e adaptativas, cujos valores variam automaticamente para cada amostra inserida na aquisição;
- (vii) A identificação de quatro regiões distintas na segmentação considerando histogramas que mostram apenas três regiões bem definidas;
- (viii) A identificação na classificação diferencial de comportamentos individuais e coletivos que definem os tipos leucocitários e permitem a extração de características que identifiquem os cinco tipos leucocitários;
- (ix) A construção de novos descritores na classificação diferencial com alta independência da coloração utilizada na amostra e utilização, no total, de poucos descritores;
- (x) A utilização na classificação diferencial de poucos descritores, reduzindo o custo computacional do processamento;
- (xi) A construção de uma base de imagens utilizando diversas colorações hematológicas para fins de pesquisa nos problemas de análise celular;
- xii) O mapeamento do conhecimento especialista para a construção das bases de conhecimento capazes de reconhecer automaticamente os componentes sanguíneos e diferenciar os leucócitos independente da coloração hematológica.

## 6.4 Trabalhos Futuros

Como perspectivas de trabalhos futuros, pode-se citar:

- (i) A construção de sistemas computacionais para a análise de imagens microscópicas de esfregaços sanguíneos com forte independência da coloração hematológica utilizada;
- (ii) A anotação e ampliação do banco de imagens para pesquisa pública;
- (iii) O desenvolvimento de módulos computacionais inteligentes que possam ser embarcados em dispositivos de baixo custo para a análise celular;
- (iv) O desenvolvimento de outras soluções e a aplicação ao mesmo banco de imagens para análises comparativas;
- (v) A extensão dos descritores para as análises de outros tipos de células e tecidos;
- (vi) O desenvolvimento de métodos de contagem celular.

## 6.5 Publicações

Vale, AMPG.; Guerreiro, AMG; Dória Neto, A. D.; Cavalcanti Junior, G. B.; Leitão, V. C. L. T. S.; Martins, A. M. *Auto-segmentation and classification of blood components in microscopic images using a fuzzy approach*. Revista Brasileira de Engenharia Biomédica. v. 30(4): 1-14, 2014.

---

# Referências

---

- Amabis, J. M.; Martho, G. R. (2008) *Biologia das células: origem da vida, citologia, histologia, embriologia*, 2<sup>a</sup> edição, Editora Moderna.
- Boaventura, I. A. G. (2010) *Números Fuzzy em Processamento de Imagens Digitais e Suas Aplicações na Detecção de Bordas*. Tese de Doutorado – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos.
- Boaventura, I. A. G.; Volpe, V. M.; Silva, I. N.; Gonzaga, A. (2006) *Fuzzy Classification of Human Skin Color in Color Images*, IEEE International Conference on Systems, Man and Cybernetics, 8 – 11 Oct., 2006, Taipei, Taiwan.
- Bonventi Jr., W.; Reali Costa, A. H. (2000) *Comparação entre métodos de definição de conjuntos nebulosos de cores para classificação de pixels*. International Joint Conference IBERAMIA'2000 and SBIA'2000, Workshop Proceedings, Atibaia, SP, Brazil, pp.105-110.
- Bradbury, S. (1989) *An Introduction to the Optical Microscope*, Royal Microscopy Society, microscopy Handbooks 01, Oxford Science Publications.
- Biomedicina Brasil (2011) *Esfregaço sanguíneo*. Disponível em <http://www.biomedicinabrasil.com/> 2011/03/esfregaco-sanguineo.html e acessado em 02 de março de 2012.
- BloodLine Imagem Atlas (2010) *BloodLine Image Atlas*. Disponível em <http://image.bloodline.net/> e acessado em 10 de janeiro de 2010.
- CEB (2012) *Tecido Sanguíneo*. Disponível em <http://www.iceb.ufop.br/decbi/histologia> e acessado em 10 de março de 2012.
- Chaira, T. (2014) *Accurate segmentation of leukocyte in blood cell images using Atanassov's intuitionistic fuzzy and interval Type II fuzzy set theory*. Micron. 2014, 61: 1-8.

- Chan, Y; Tsai, M.; Huang, D.; Zheng, Z.; Hung, K (2010) *Leukocyte nucleus segmentation and nucleus lobe counting*. BMC Bioinformatics 2010, 11:558.
- Chiu, C.; Park, C.S. (1994) *Fuzzy cash flow analysis using present worth criterion*. The Engineering Economist, vol 39(2): 113-138.
- Cox, E. (1995) *Fuzzy logic for business and industry*. Massachusetts: Charles River Media Inc.
- Cuevas, M. D; Manzanares, M.; Zaldivar, D.; Perez-Cisneros, M. (2013) An improved computer vision method for white blood cells detection. Computational and Mathematical Methods in Medicine. 2013. 14 p.
- Dorini L, Minetto R, Leite N. (2007) *White blood cell segmentation using morphological operators and scale-space analysis*. Brazilian Symposium on Computer Graphics & Image Processing (SIBGRAPI); 2007; 294-304.
- EMBRAPA (2012) *Aplicações da microscopia óptica*. Disponível em [www.cnpab.embrapa.br](http://www.cnpab.embrapa.br) e acessado em 25 de abril de 2012.
- Fan, S.; Wang, H. (2011) *Multi-direction Fuzzy Morphology Algorithm for Image Edge Detection*. Journal of networks, VOL. 6, NO. 6.
- Fatichah, C; Tangel, M. L; Widyanto, M., R.; Dong, F.; Hirota, K. (2012) *Interest-Based Ordering for Fuzzy Morphology on White Blood Cell Image Segmentation*. Journal of Advanced Computational Intelligence and Intelligent Informatics Vol.16 No.1.
- Fatima, M. M.; Seenivasagam, V. (2011).*A fast fuzzy-c means based marker controlled watershed segmentation of clustered nuclei*. Proceedings of International Conference on Computer, Communication and Electrical Technology.
- FCF (2012) *Coloração em Hematologia*. Disponível em <http://www.fcf.usp.br> e acessado em 10 de abril de 2012.
- Figueiró, T. (2005) *Automatic Detection of Blood Cells on Color Images using Image Matching and Flood Map*. SIBGRAPI 2005, Natal, RN.

- GENOMA (2012) *Microscopia Óptica*. Disponível em [http://genoma.ib.usp.br/educacao/Microscopia\\_2.pdf](http://genoma.ib.usp.br/educacao/Microscopia_2.pdf) e acessado em 24 de abril de 2012.
- Glotzos, D.; Spyridonos, P.; Cavouras, D.; Ravazoula, P.; Dadioti, P.A.; Nikiforidis, G. (2004) *Automated segmentation of routinely hematoxylin-eosin-stained microscopic images by combining support vector machine clustering and active contour models*. Anal. Quant. Cytol. Histol., 26: 331–340.
- Gonzalez, R. C.; Woods, R. E (2000) *Processamento Digital de Imagens*. 2<sup>a</sup> edição, Editora Edgard Blücher, São Paulo.
- Gonzalez, R. C.; Woods, R. E (2009) *Processamento Digital de Imagens*. 3<sup>a</sup> edição, Editora Pearson, São Paulo.
- Goodhew, P. J; Humphreys , J.; Beanland, R.(2001) *Electron microscopy and analysis*. 3<sup>a</sup> edição, Editora Taylor & Francis, Londres.
- Ghosh, M.; Das, D.; Chakraborty, C.; Ray, A. K. (2010) *Automated leukocyte recognition using fuzzy divergence*. Micron, 41: 840–846.
- Ghosh, M.; Das, D.; Ray , A. K.; Chakraborty, C. (2011) *Development of Renyi's entropy based fuzzy divergence measure for leukocyte segmentation*. J. Med. Imaging Health Inform., 1: 334–340.
- Gulati, G.; Song, J. M.; Florea, A. D.; Gong, J. (2013) *Purpose and criteria for blood smear scan, blood smear examination, and blood smear review*. Annals of Laboratory Medicine. 33: 1-7.
- Guo N, Zeng L, Wu Q. (2007) *A method based on multispectral imaging technique for white blood cell segmentation*. Computers in Biology and Medicine. 2007; 37(1): 70-76.
- Hamghalam M, Ayatollahi A. (2009) *Automatic counting of leukocytes in Giemsa-stained images of peripheral blood smear*. ICDIP 2009: Proceedings of the International Conference on Digital Image Processing; 2009 mar 7-9; Bangkok, TH. 2009. p. 13-16.
- Hamghalam M.; Motameni, M.; Kelishomi A.E. (2009) *Leukocyte segmentation in giemsa-stained image of peripheral blood smears based on active contour*. ICSPS 2009:

- Proceedings of International Conference on Signal Processing Systems; 2009 mai 15-17; Singapore. 2009. p. 103-106.
- Haubecker, H.; Tizhoosh, H. R. (2000) *Fuzzy image processing*. Computer Vision and Applications: A Guide for Students and Practitioners. San Diego: Academic Press, cap. 16, p. 541–576.
- Heckner F, Freund M. (2000) *Hematologia: Microscopia Prática*. 9 edição. São Paulo: Editora Santos, São Paulo. p13 – 17.
- Hiremath PS, Bannigidad P, Geeta S. (2010) *Automated identification and classification of white blood cells (leukocytes) in digital microscopic images*. IJCA, Special Issue on RTIPPR; 2010; 2:59–63. Published By Foundation of Computer Science.
- Huang, D.; Hung, K.; Chan, Y. A (2012) A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images. *The Journal of Systems and Software*. 85: 2104– 2118.
- Jiang, K.; Liao, Q.; Dai, S. (2003) *A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering*. International Conference on Machine Learning and Cybernetics. 2003 Nov 2-5. 5: 2820-2825.
- Jati A, Singha G, Mukherjee R, Ghoshb M, Konara A, Chakraborty C, et. al. (2014) *Automatic leukocyte nucleus segmentation by intuitionistic fuzzy divergence based thresholding*. *Micron*. 2014; 58: 55-65.
- Jiji, G. W.; Ganesan, L. (2010) *A new approach for unsupervised segmentation*. *Applied Soft Computing*. 10, 689–693.
- Kartalopoulos, S. V. (1996) *Fuzzy Logic, understanding Neural Networks and Fuzzy Logic*. 1 ed. Piscataway: IEEE Press, p. 121-151.
- Khashman, A. (2009) *Blood Cell Identification Using Emotional Neural Networks*. *Journal of Information Science And Engineering* 25, 1737.
- Klir, G. J; Folger, T. A. (1988) *Fuzzy Sets, Uncertainty, and Information*. 1 ed. New Jersey: Prentice Hall.

- Ko BC, Gim JW, Nam JY. (2011) *Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake*. Micron. 2011; 42: 695-705.
- Kumar, B.R.; Joseph, D. K.; Sreenivas, T.V. (2002) *Teager energy based blood cell segmentation*. 14 th International Conference on Digital Signal Processing; 2002. vol. 2:619–622.
- Kuse, M.; Sharma, T.; Gupta, S. (2010) *A classification scheme for lymphocyte segmentation in H&E stained histology images*. ICPR'10 Proceedings of the 20th International conference on Recognizing patterns in signals, speech, images, and videos. 2010 Aug 23-26. Istanbul, Turkey. p. 235-243.
- Lee, C.C. (1990a) *Fuzzy Logic in Control Systems: Fuzzy Logic Controller - Part I*. IEEE Transactions on Systems, Man and Cybernetics. v. 20, n. 2, p. 404-418.
- Lee, C.C. (1990b) *Fuzzy Logic in Control Systems: Fuzzy Logic Controller - Part II* IEEE Transactions on Systems, Man and Cybernetics. v. 20, n. 2, p. 419-430.
- Leite, L. C. M. (2009) *Geração e simplificação da base de conhecimento de um sistema híbrido fuzzy-genético*. Dissertação (Mestrado em Engenharia) – Universidade do Estado do Rio de Janeiro, Rio de Janeiro.
- LFF (2012) *Microscópio Óptico*. Disponível em <http://fap01.if.usp.br/~lff/mo.html> e acessado em 25 de abril de 2012.
- Li Y.; Chung, F.; Wang, S. (2008) *A robust neuro-fuzzy network approach to impulse noise filtering for color images*. Applied Soft Computing 8, 872–884.
- Li, Q.; Wang, Y.; Liu, H.; He, X.; Xu, D.; Wang, J.; Guo, F. (2014) *Leukocyte cells identification and quantitative morphometry based on molecular hyperspectral imaging technology*. Computerized Medical Imaging and Graphics. Vol 38, Issue 3, April 2014, Pages 171–178.
- Liao, Q.; Deng, Y. (2002) *An accurate segmentation method for white blood cell images*. Proceedings of IEEE International Symposium on Biomedical Imaging.

- Lin, S.; Chen, H.; Lin, Y. (2011) Automatic Counting Cancer Cell Colonies using *Fuzzy Inference System*. *Journal of Information Science And Engineering* 27, 749-760.
- LVP (2012) *Laboratório Virtual de Patologia*. Disponível em <http://anatpat.unicamp.br/biinflparacoco4.html> e acessado em 21 de abril de 2012.
- Maliska, A. M. (2007) *Microscopia Eletrônica de Varredura*. Universidade Federal de Santa Catarina, UFSC.
- Malpica, N.; Solrzano, C.O.; Vaquero, J.J.; Santos, A.; Vallcorba, I.; Garca-Sagredo, J.M.; Del Pozo, F. (1997) *Applying watershed algorithms to the segmentation of clustered nuclei*. *Cytometry*, 28: 289–297.
- Martinez, M.B., Tadei, C.R. ; Trabulsi, L.R., Alterthum, F. (2005) *Microbiologia*. 4<sup>a</sup> Ed, São Paulo: Atheneu, 718 p.
- MathWorks (2014). MATLAB User's Guide. R2014b Documentation – *Fuzzy Logic Toolbox*. Product Documentation. 2014. Disponível em <http://www.mathworks.com/help/fuzzy-foundations-of-fuzzy-logic.html#bp78170-2> e acessado em 10 de outubro de 2014.
- Mattos, M. C.; Nassar, S. M.; Souza, E. C.; Strada, P. F. (2004) *O Raciocínio Fuzzy no Desenvolvimento de um Sistema de Controle da Assistência Respiratória em Neonatos*. Anais e Programação do IX CBIS, Universidade do Extremo Sul Catarinense - UNESC, Criciúma, SC.
- Mircic, S.; Jorgovanovic, N. (2006) *Automatic classification of leukocytes*. *J. Autom. Control*, 16: pp. 29–32.
- Mohapatra, S.; Patra, D.; Kumar, K. (2011) *Unsupervised Leukocyte Image Segmentation Using Rough Fuzzy Clustering*. International Scholarly Research Network. ISRN Artificial Intelligence. Volume 2012.
- Moura, R. A. (1997) *Técnicas de Laboratório*. 3<sup>a</sup> Ed. São Paulo: Atheneu.

- Nawgaje, D. D.; Kanphade, R. D. (2011) *Implementation of Fuzzy Inference System For White Blood Cell Cancer Detection Using DSP TMS320C6711*. International Journal of Engineering Science and Technology (IEST). Special Issue Feb 2011.
- Nazlibilek, S.; Karacor, D.; Ercan, T.; Sazli, M. H.; Kalender, O.; Ege, Y. (2014) *Automatic segmentation, counting, size determination and classification of white blood cells*. Measurement 2014, 55: 58–65.
- Nakib, A.; Oulhadj, H.; Siarry, P. (2007) *Microscopic image segmentation with two-dimensional exponential entropy based on hybrid microcanonical annealing*. Proceedings of International Conference on Machine Vision Applications.
- Ongun, G.; Halici, U.; Leblebicioglu, K.; Atalay, V.; Beksac, M.; Beksak, S. (2008) *A modified fuzzy clustering for white blood cell segmentation*, Proceedings of the 3rd International Symposium on Biomedical Engineering, pp. 356–359.
- Osowski S.; Siroic, R.; Markiewicz, T.; Siwek, K. (2009) *Application of Support Vector Machine and Genetic Algorithm for Improved Blood Cell Recognition*. IEEE Transactions On Instrumentation and Measurement, 58: 7.
- Osuna, V.; Cuevas, E.; Sossa, H. (2013) *Segmentation of blood cell images using evolutionary methods*. Proceedings of EVOLVE – A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation II, Springer Berlin Heidelberg.
- Otsu, N. (1979) *A threshold selection method from gray level histograms*. IEEE Trans. Syst. Man Cyber., 9, pp. 62–66.
- Pan, C.; Park, D.; Yang, Y.; Yoo, H. (2012) *Leukocyte image segmentation by visual attention and extreme learning machine*. Neural Comput. Appl., 21: 1217–1227.
- Pelczar Jr, M. J., Chan, E. C. S., Krieg, N. R. (2005) *Microbiologia: Conceitos e Aplicações*. 2<sup>a</sup> Ed, v. 1, São Paulo: Pearson, 524p.
- Pinho, A. F. (1999) *Uma contribuição para a resolução de problemas de programação de operações em sistemas de produção intermitentes flow-shop: A consideração de*

- incertezas.* Dissertação (Mestrado em Engenharia) – Universidade Federal de Itajubá, Itajubá.
- Putzu L, Ruberto C. (2013) *White blood cells identification and classification from leukemic blood image.* IWBBIO 2013: Proceedings of the International Work-Conference on Bioinformatics and Biomedical Engineering; 2013 March 18-20; Granada, ES. 2013. p. 99-106.
- Rajendran, S.; Arof, H.; Mokhtar, N.; Mubin, M.; Yegappan, S.; Ibrahim, F. (2011) *Dual modality search and retrieval technique analysis for leukemic information system.* Sci. Res. Essays, 6: 247–255.
- Ramesh N, Dangott B, Salama ME, Tasdizen T. (2012) *Isolation and two-step classification of normal white blood cells in peripheral blood smears.* J Pathol Inform; 3. 13p.
- Ramoser H, Laurain V, Bischof H, Ecker R. (2005) *Leukocyte segmentation and classification in blood-smear images.* Proceedings of the IEEE Engineering in Medicine and Biology Conference; 2005 sep 1-4; Shanghai, CN. 2005. p. 3371-3374.
- Rezatofighi, S. H.; Soltanian-Zadeh, H.; Sharifian R.; Zoroofi R. A. (2009) *A new approach to white blood cell nucleus segmentation based on gram-schmidt orthogonalization.* Digital Image Processing, 2009 International Conference on, 2009, pp. 107-111
- Rezatofighi, S. H.; Soltanian-Zadeh, H. (2011) *Automatic recognition of five types of white blood cells in peripheral blood.* Computerized Medical Imaging and Graphics 35, 333–343.
- Rezende, S. O. (2005) *Sistemas Inteligentes: Fundamentos e Aplicações.* Barueri: Manole.
- Richetto, M. R. S. (2007) *Sistema de tratamento de imagens digitais para auxílio a exames laboratoriais,* Dissertação de Mestrado, Universidade de Taubaté, São Paulo.
- Sabino D. M. U.; Costa L. F.; Rizzatti E. G.; Zago M. A. (2004) *A texture approach to leukocyte recognition.* Real-Time Imaging; 2004; 10(4):205-216.
- Sansone M, Zeni O, Esposito G. (2012) *Automated segmentation of comet assay images using gaussian filtering and fuzzy clustering.* Med Biol Eng Comput. 2012; 50(5): 523-32.

- Santos, A. P.; Bandeira, M. B.; Siqueira, L. O. (2009) *Comparação entre diversos métodos de contagem diferencial de leucócitos em pacientes leucopênicos*. Revista Brasileira de Hematologia. Hemoter, vol.31, n.3, pp. 203-205.
- Santos, D. B. L. M. (2004) *Procedimento para a Construção dos Conjuntos Fuzzy utilizados em Controladores Semaforicos*. Dissertação de Mestrado, Publicação TU. Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 109 p.
- Santos, W. P. (2003) *Análise de imagens digitais em patologia usando morfologia matemática e lógica nebulosa*. Dissertação de Mestrado. Centro de Tecnologia e Geociências. Universidade Federal de Pernambuco.
- Saraswat, M.; Arya, K. (2014) *Supervised leukocyte segmentation in tissue images using multi-objective optimization technique*. Eng. Appl. Artif. Intell., 31 (2014), pp. 44–52.
- Saraswat, M.; Arya, K.; Sharma, H. (2013) *Leukocyte segmentation in tissue images using differential evolution algorithm*. Swarm Evol. Comput., 11: 46–54.
- Shaw, I. S.; Simões, M. G. (1999) *Controle e Modelagem Fuzzy*. 1.ed. São Paulo: Edgard Blücher Ltda.
- Shitong, W.; Min, W. (2006) *A new detection algorithm (nda) based on fuzzy cellular neural networks for white blood cell detection*. IEEE Trans. Inf. Technol. Biomed., 10: 5–10.
- Shitong, W.; Chung, K. F.; Duan, F. (2007) *Applying the improved fuzzy cellular neural network IFCNN to white blood cell detection*, Neurocomputing, 70: 1348–1359.
- Sinha, N., Ramakrishnan, A.G. (2003) *Automation of differential blood count*. Proc. Conf. Convergent Technologies for Asia-Pacific Region, TENCON 2003, vol. 2, pp. 547–551.
- Souza, M. H. L., Elias, D. O. (2005) *Princípios de Hematologia e Hemoterapia*. Centro de Estudos Alfa Rio, Rio de Janeiro.
- Tan, K. S.; Isa, N. A. M. (2011) *Color image segmentation using histogram thresholding - Fuzzy C-means hybrid approach*. Pattern Recogn 44, 1, 1-15.

- Tanscheit, R. (2012) *Sistemas Fuzzy*. Disponível em <http://www2.ica.ele.puc-rio.br/Downloads/41/LN-Sistemas%20Fuzzy.pdf> e acessado em 05 de maio de 2012.
- Theera-Umporn, N.; Dhompongsa, S. (2007) *Morphological granulometric features of nucleus in automatic bone marrow white blood cell classification*. IEEE Trans. Inform. Technol. Biomed., 11: 353–359.
- Theera-Umporn, N. (2005) *Patch-Based White Blood Cell Nucleus Segmentation Using Fuzzy Clustering*. ECTI Transactions on Electrical Eng., Electronics and Communications. Vol III. Nº 1.
- Theml, H.; Diem, H.; Haferlach, T. (2004) *Color Atlas of Hematology. Practical Microscopic and Clinical Diagnosis*. 2nd revised edition, Thieme Stuttgart, New York.
- Tortora, G. J., Funke, B. R., Case, C. L. (2003) *Microbiologia*, 6a ed, São Paulo: Artmed, p.267-393.
- UNIP (2012) *Microscopia Óptica*. Disponível em <http://biomedicinaemacao-unip.blogspot.com.br/2012/02/microscopia-optica.html> e acessado em 25 de abril de 2012.
- Vaz, A. M. (2005) *Estudos das Funções de pertinência para conjuntos fuzzy utilizados em Controladores Semafóricos Fuzzy*. Dissertação de Mestrado, Faculdade de Tecnologia, Universidade de Brasília, DF.
- Vendruscolo, A.; Takaki, I.; Bersani-Amado, L. E.; Dantas, J. A.; Bersani-Amado, C. A.; Cuman R. K. N. (2006) *Antiinflammatory and antinociceptive activities of zingiber officinale roscoe essential oil in experimental animal models*. Indian Journal of Pharmacology, 38(1): pp. 58-59.
- Von A. (1996) *Constantin. Fuzzy logic and neuroFuzzy applications in business and finance*. New Jersey: Prentice Hall PTR.
- Wang, X.; Zhang, X.; Yang, H.; Bu, J. (2012) *A pixel-based color image segmentation using support vector machine and fuzzy C-means*. Neural Networks 33, 148–159.

- Wermser, D.; Haussmann, G.; Liedtke, C. (1984) *Segmentation of blood smears by hierarchical thresholding.* Comput. Vis. Graph. Image Process., 25: 151–16.
- Wu, J.; Zeng, P.; Zhou, Y.; Olivier, C. (2006) *A novel color image segmentation method and its application to white blood cell image analysis.* Proceedings of International Conference on Signal Processing.
- Xie, C.; Zhu, J.; Chen, X.; Mi, L.; Nishida, N.; Springer, T. A. (2010) *Structure of an integrin with an alphaI domain, complement receptor type 4.* EMBO J. 2010 Feb 3; 29(3):666-79.
- Zadeh, L.A. (1973). *Outline of a New Approach to the Analysis of Complex Systems and Decision Processes.* IEEE Trans. on Systems Man & Cybernetics, Vol.3: 28-44.
- Zamani, F.; Safabakhsh, R. (2006) *An unsupervised GVF snake approach for white blood cell segmentation based on nucleus.* Proceedings of International Conference on Signal Processing.
- Zheng X, Wang Y, Wang G, Chen Z. (2014) *A novel algorithm based on visual saliency attention for localization and segmentation in rapidly-stained leukocyte images.* Micron. 56: 17-28.

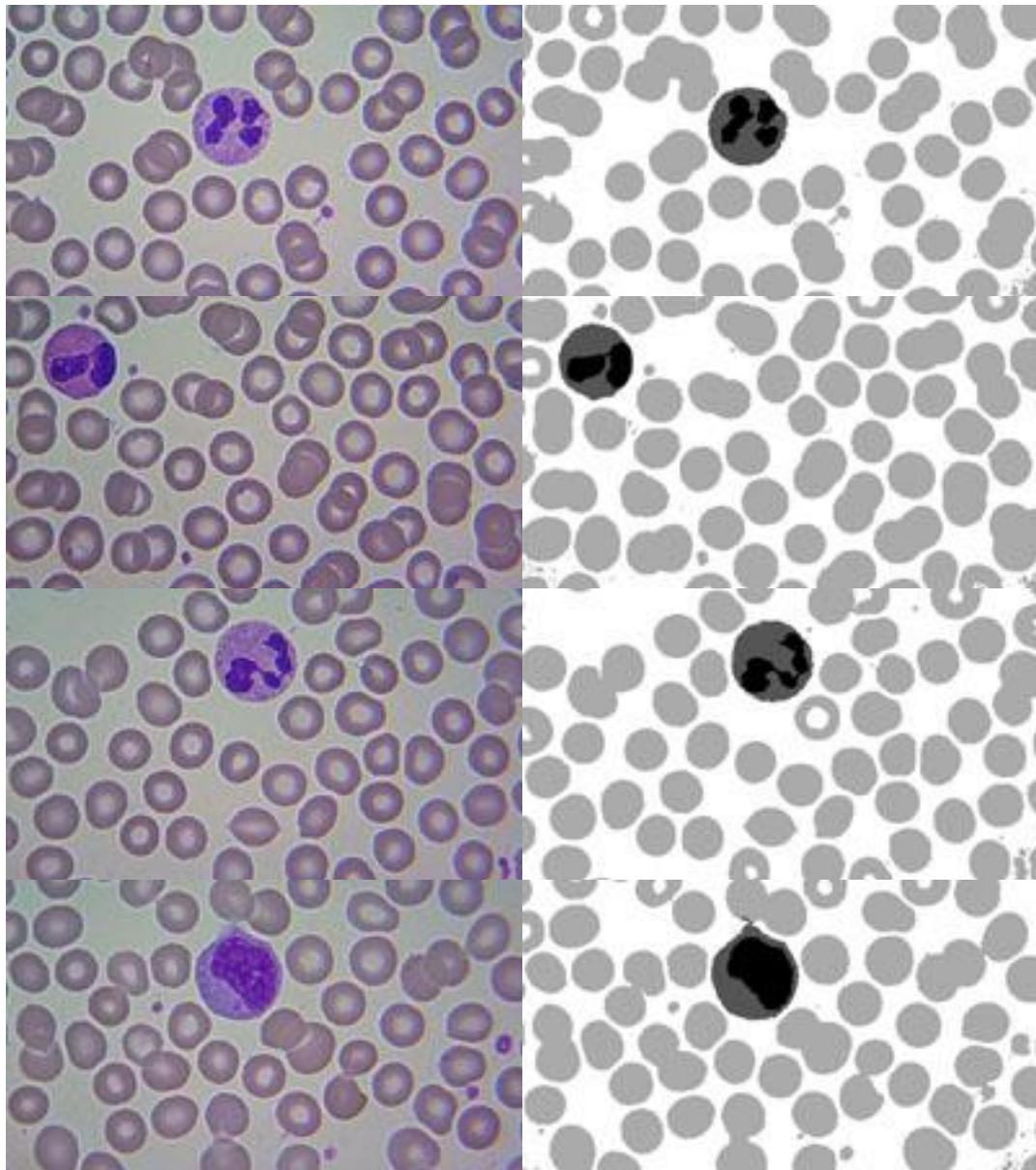
---

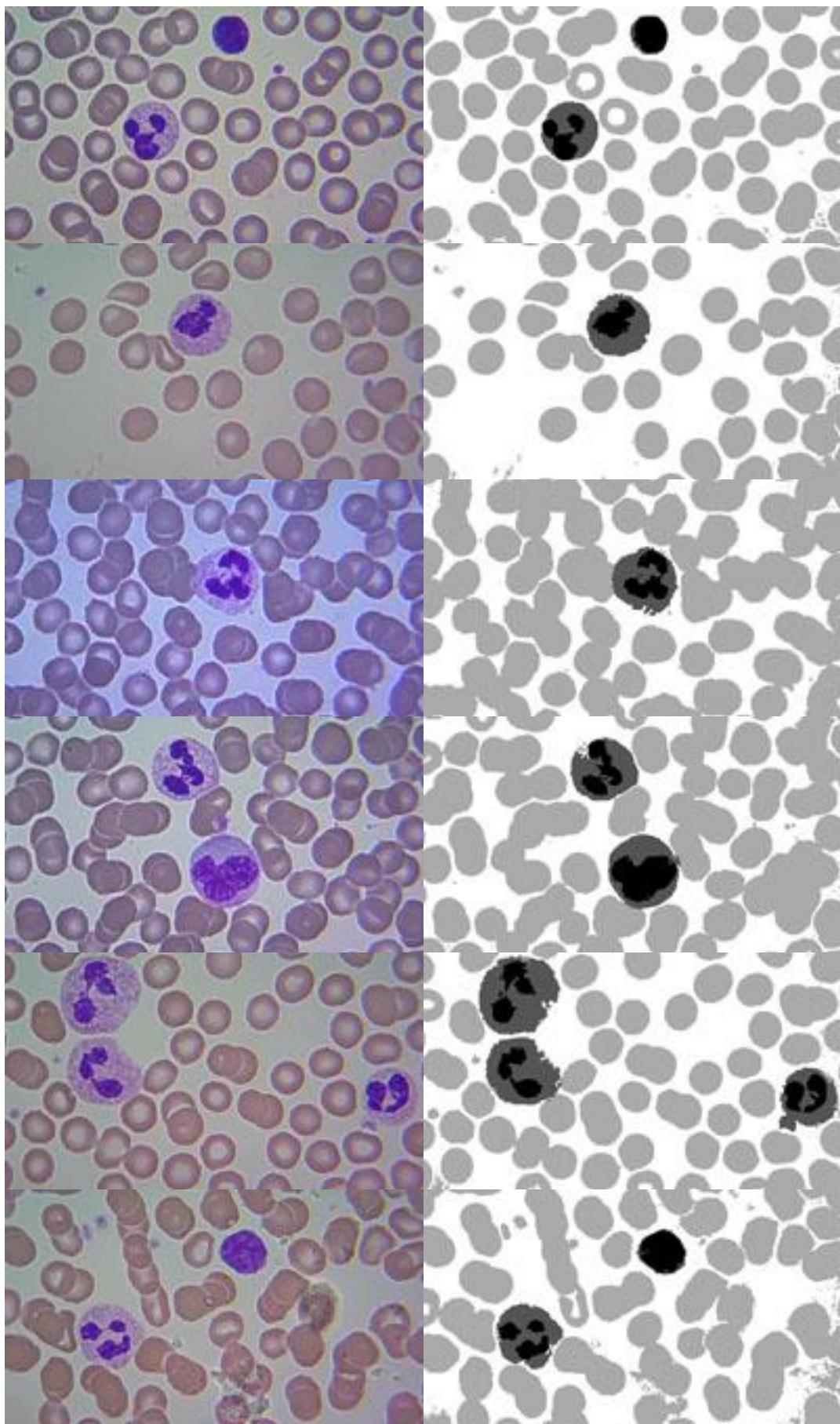
# Apêndice A

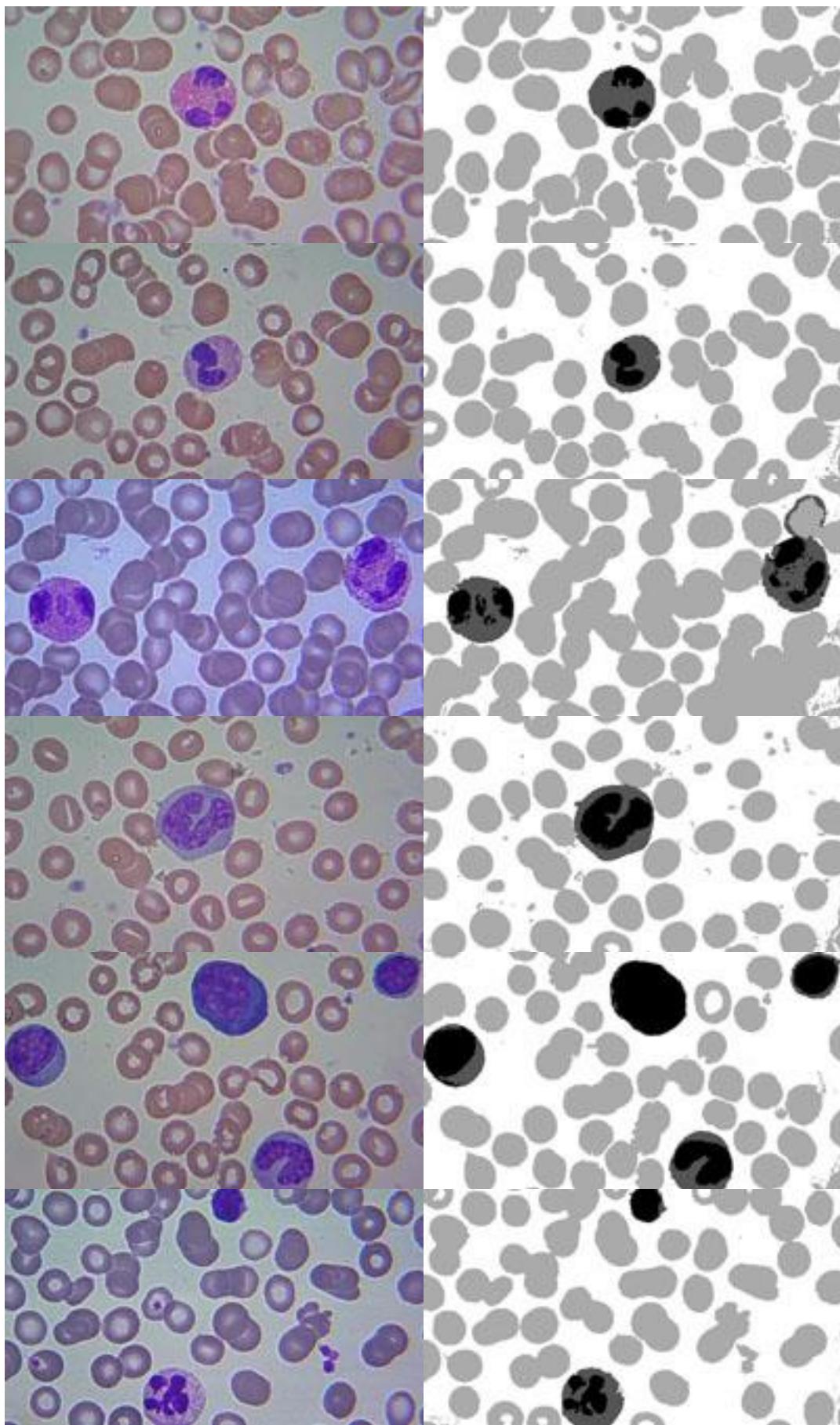
---

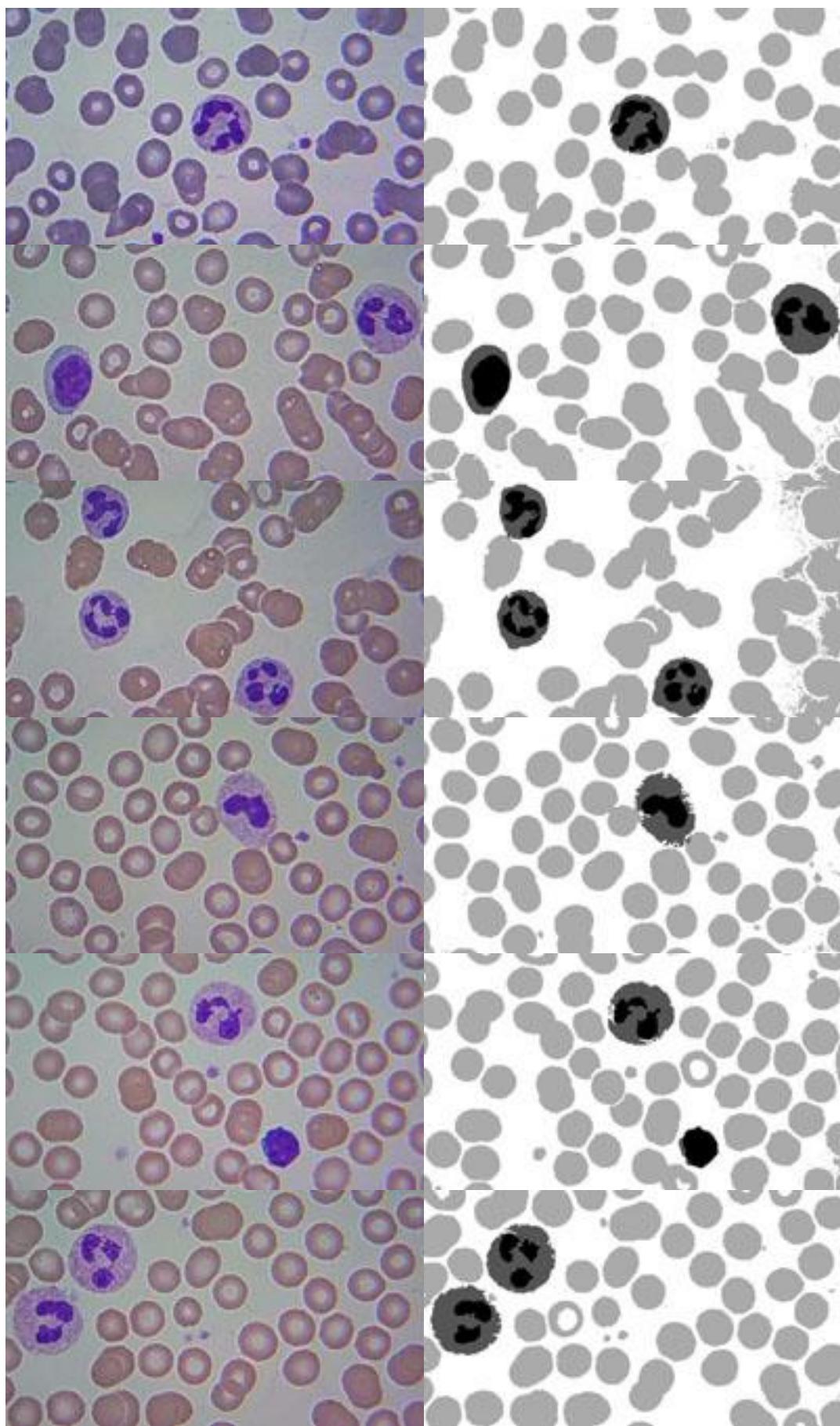
Segue algumas imagens processadas na segmentação e os respectivos resultados.

## 1. Método de Coloração: Panótico Rápido

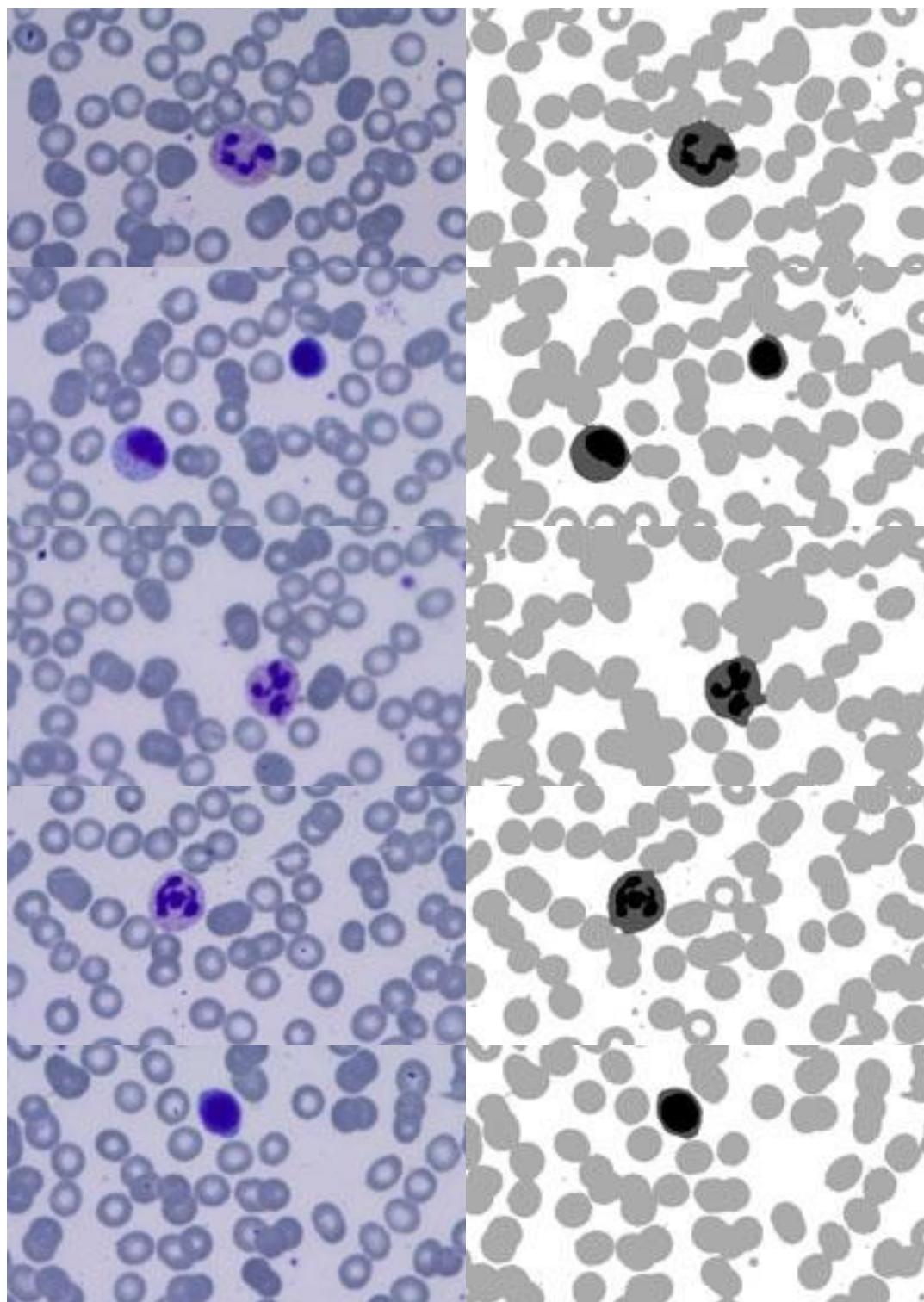


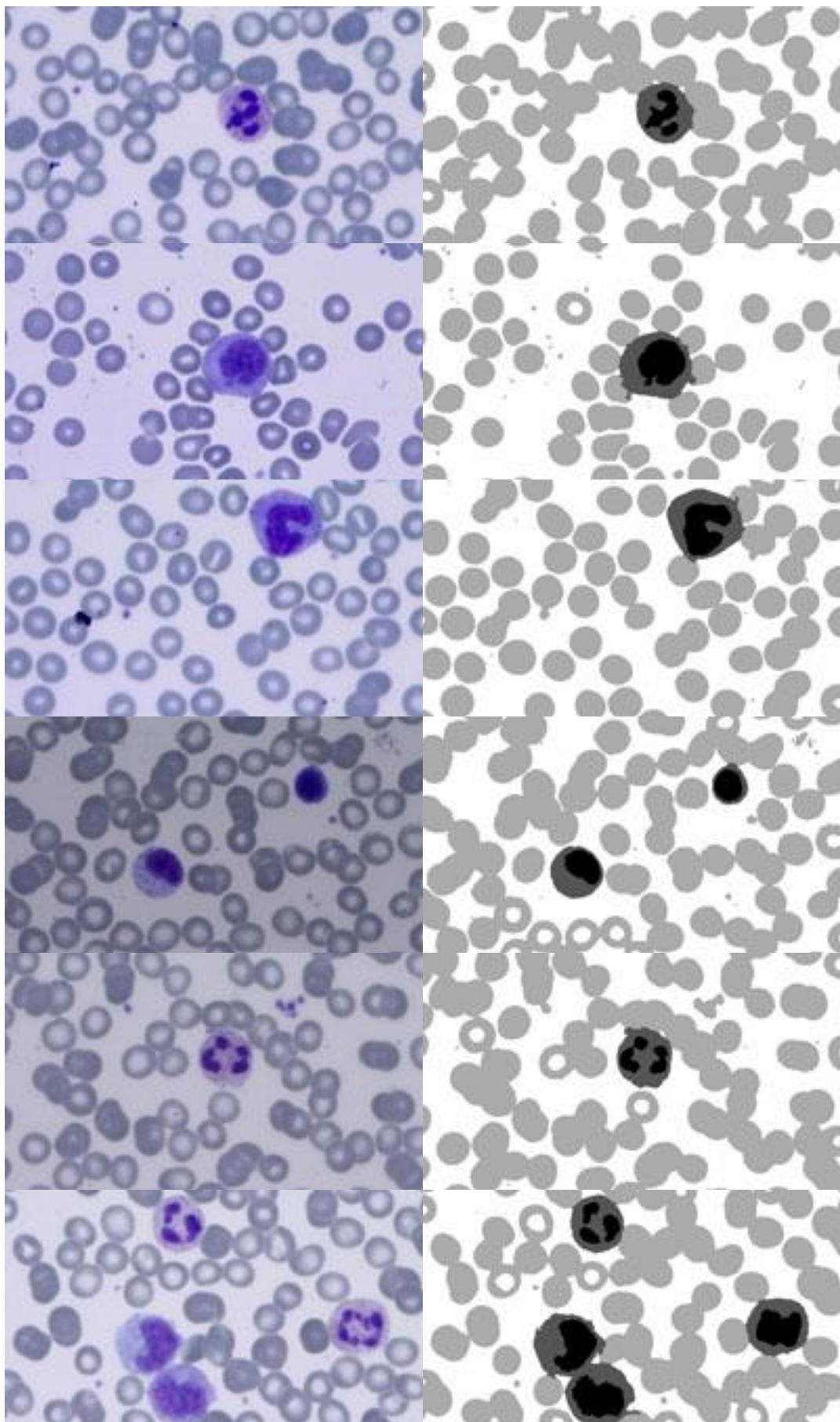


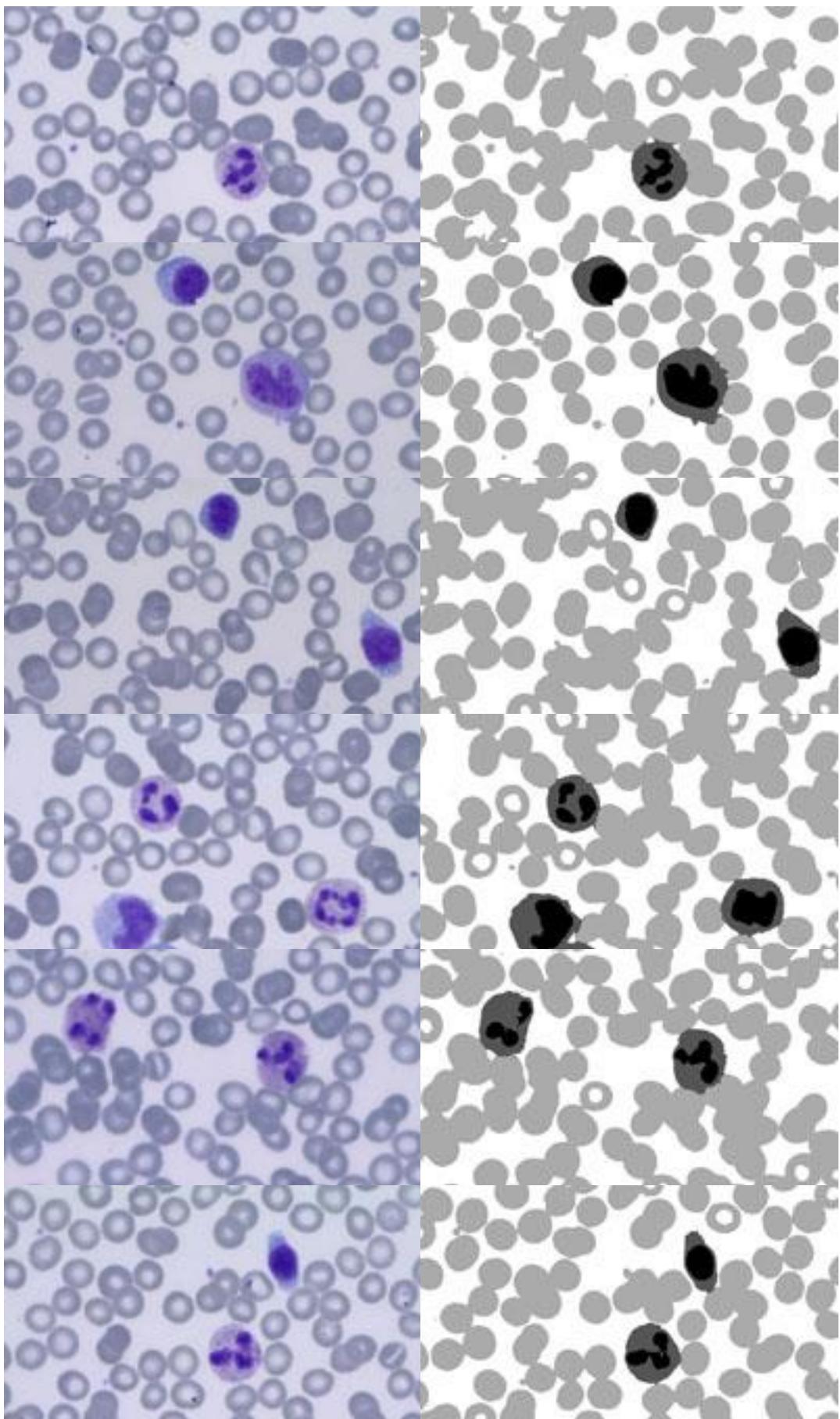


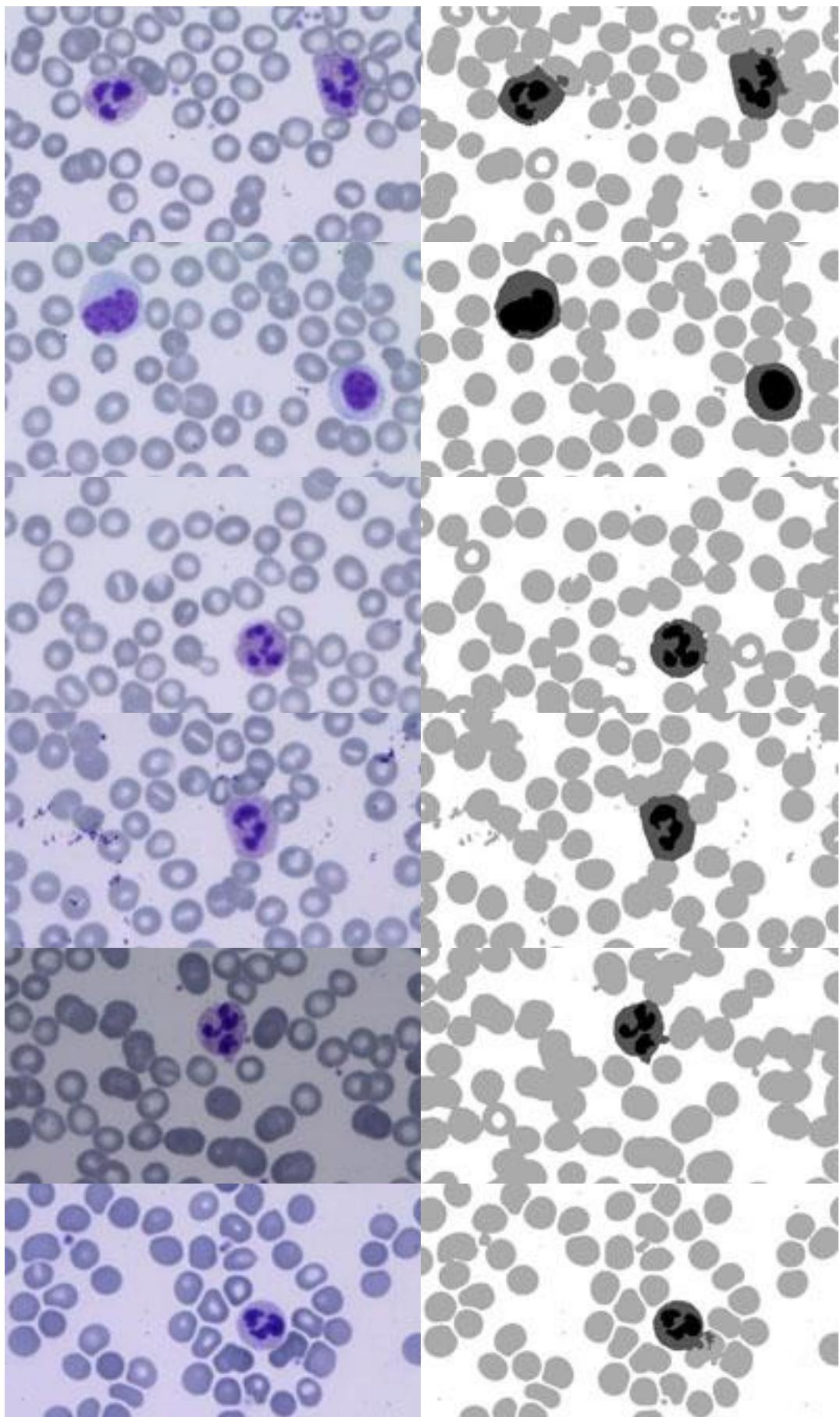


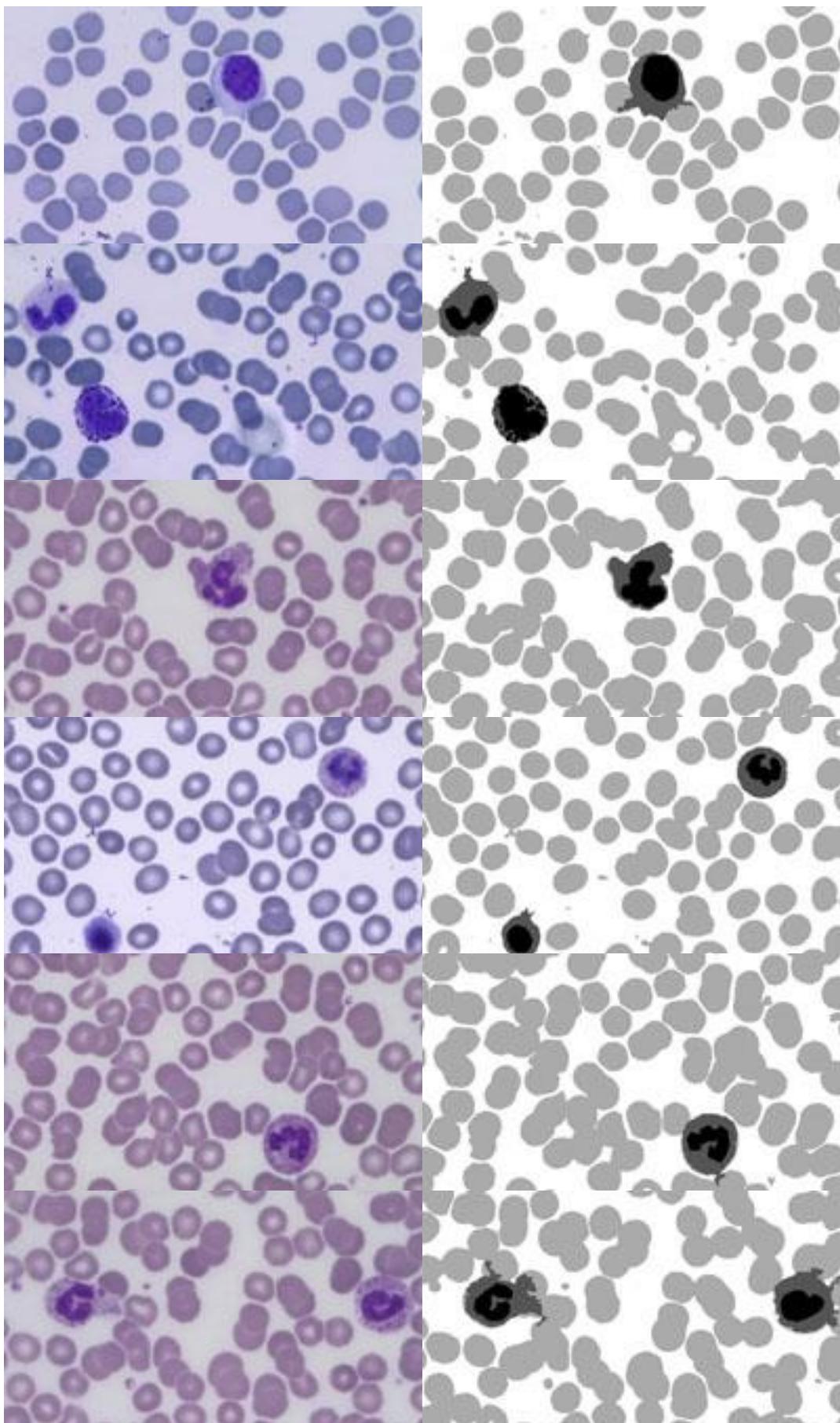
## 2. Método de Coloração: Rosenfeld e Leishman

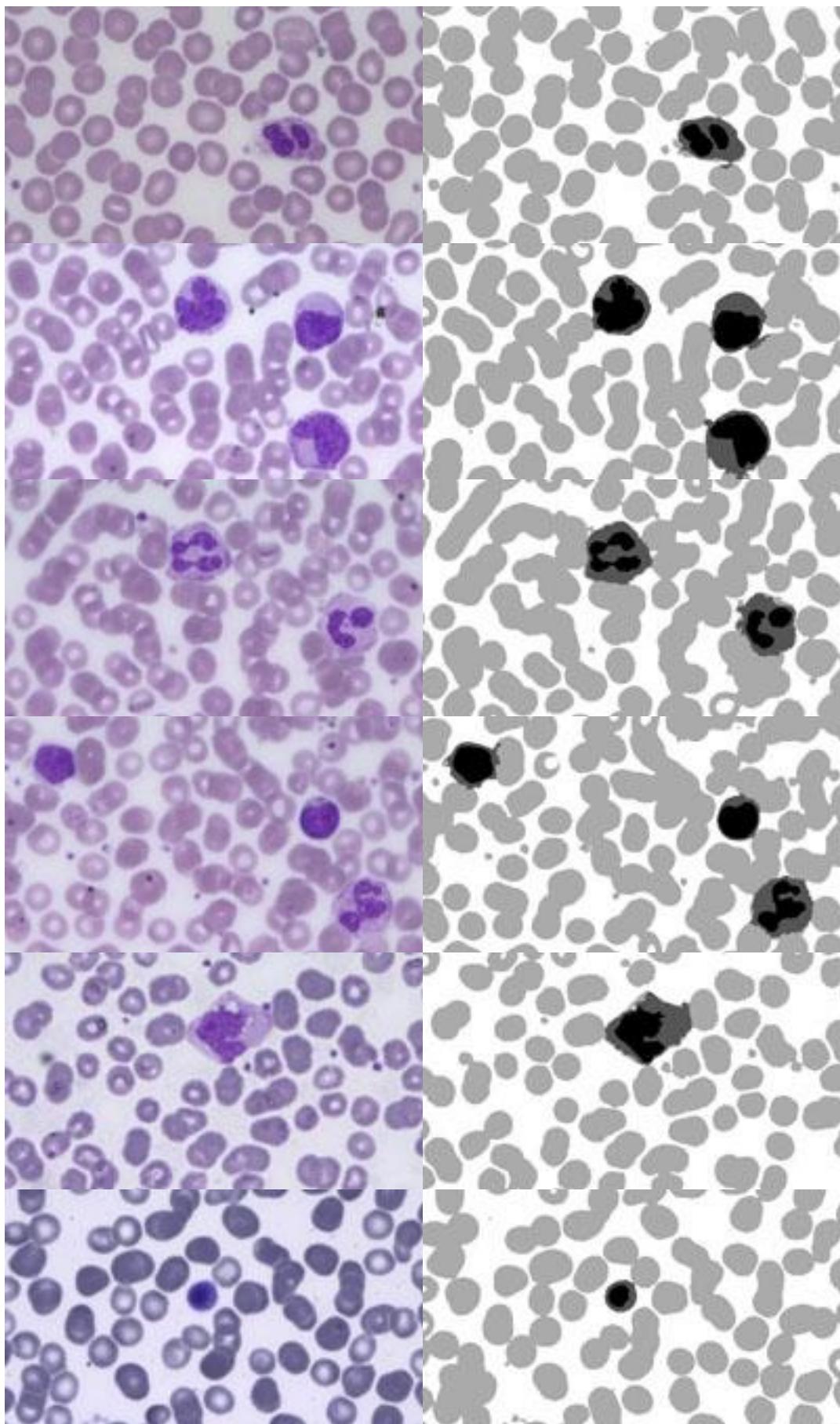












### 3. Método de Coloração: não especificado

