

UNIVERSIDADE REGIONAL DE BLUMENAU
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
CURSO DE CIÊNCIA DA COMPUTAÇÃO – BACHARELADO

VONCELL: UM PROTÓTIPO PARA RECONHECIMENTO DE
CÉLULAS DO SISTEMA IMUNE DO TIPO LINFÓCITO E
NEUTRÓFILO

JULIANO MUELOSCHAT

BLUMENAU
2016

JULIANO MUELOSCHAT

**VONCELL: UM PROTÓTIPO PARA RECONHECIMENTO DE
CÉLULAS DO SISTEMA IMUNE DO TIPO LINFÓCITO E
NEUTRÓFILO**

Trabalho de Conclusão de Curso apresentado ao curso de graduação em Ciência da Computação do Centro de Ciências Exatas e Naturais da Universidade Regional de Blumenau como requisito parcial para a obtenção do grau de Bacharel em Ciência da Computação.

Prof. Aurélio Faustino Hoppe, Mestre - Orientador

**BLUMENAU
2016**

**VONCELL: UM PROTÓTIPO PARA RECONHECIMENTO DE
CÉLULAS DO SISTEMA IMUNE DO TIPO LINFÓCITO E
NEUTRÓFILO**

Por

JULIANO MUELOSCHAT

Trabalho de Conclusão de Curso aprovado
para obtenção dos créditos na disciplina de
Trabalho de Conclusão de Curso II pela banca
examinadora formada por:

Presidente: _____
Prof. Aurélio Faustino Hoppe, Mestre – Orientador, FURB

Membro: _____
Prof. Dalton Solano dos Reis, Mestre – FURB

Membro: _____
Prof. Daniel Theisges dos Santos, Mestre – FURB

Blumenau, 08 de dezembro de 2016.

Dedico este trabalho aos meus professores,
que sempre serviram como fonte de inspiração
e admiração em minha caminhada.

AGRADECIMENTOS

A minha mãe, que sempre prestou todo carinho e atenção ao longo dos anos.

A minha noiva e biomédica Mônica Diel, pelas contribuições neste trabalho.

Ao meu amigo Arnaldo Uber Junior, pelas contribuições neste trabalho.

Aos meus colegas de graduação Luís, Jean, Thainara, Filipe, Philip, Daniel e Camila, que tornaram esses anos de estudos mais divertidos.

Aos meus professores, que ao longo da minha trajetória sempre me transmitiram seus conhecimentos e valores.

Ao meu orientador, Aurélio Faustino Hoppe, por sua motivação, postura e dedicação para a conclusão deste trabalho.

Nascido para perder, viva para vencer.

Lemmy Kilmister

RESUMO

Este trabalho apresenta um protótipo que realiza o reconhecimento de leucócitos típicos do tipo linfócito e neutrófilo através de imagens capturadas a partir de um microscópio óptico. O fluxo do protótipo foi dividido em duas fases: realce do campo e descoberta de células e, isolamento e extração de características do núcleo e citoplasma para inferir a classificação da célula. A primeira fase é a etapa de realce do campo, que utiliza a imagem de entrada, convertendo-a de RGB para o sistema de cores HSV. O canal V é usado para delimitar a região do campo, enquanto que o canal S apresenta o realce dos núcleos das células. A partir disto, o canal S sofre várias etapas de eliminação de ruídos para separar o núcleo do restante da imagem. Após extrair o fundo da imagem, os núcleos sofrem crescimento de região, criando-se assim a região das células. As regiões das células são recortes da imagem de entrada. A segunda fase separa o núcleo e citoplasma. Eles são utilizados na extração e cálculo das características morfológicas da célula. Tais características são usadas para inferir a classificação da célula através de uma rede neural do tipo MLP. A rede neural utilizou 22 células, 7 linfócitos e 15 neutrófilos, para realizar o treinamento, sobre uma base de 40 imagens com 89 células. Os resultados obtidos demonstram que o protótipo obteve um percentual de 94,7% de acerto na contagem total de células. Enquanto que na contagem diferencial, obteve 95% de acurácia na classificação de linfócitos e de 94,4% em neutrófilos.

Palavras-chave: Leucócitos. Visão computacional. Aprendizagem de máquina.

ABSTRACT

This work presents a prototype that performs the recognition of typical leukocytes of the lymphocyte and neutrophil type through images captured from an optical microscope. The flow of the prototype was divided into two phases: field enhancement and cell discovery, and isolation and extraction of characteristics of the nucleus and cytoplasm to infer cell classification. The first phase is the field enhancement step, which uses the input image, converting it from RGB to the HSV color system. The V-channel is used to delineate the region of the field, whereas the S-channel presents the enhancement of the nuclei of the cells. From this, the S channel undergoes several noise elimination steps to separate the core from the remainder of the image. After extracting the bottom of the image, the nuclei undergo regional growth, thus creating the region of the cells. Cell regions are clippings of the input image. The second phase separates the nucleus and cytoplasm. They are used in the extraction and calculation of the morphological characteristics of the cell. These characteristics are used to infer the classification of the cell through an MLP neural network. The neural network used 22 cells, 7 lymphocytes and 15 neutrophils, to perform the training, on a base of 40 images with 89 cells. The results show that the prototype obtained a percentage of 94.7% accuracy in the total cell count. While in the differential count, it obtained 95% accuracy in lymphocyte classification and 94.4% in neutrophils.

Key-words: Leukocytes. Computer vision. Learning machine.

LISTA DE FIGURAS

Figura 1 - Esfregaço sanguíneo	18
Figura 2 - Célula do tipo neutrófilo	19
Figura 3 - Célula tipo linfócito	19
Figura 4 - Resultado obtido com método Otsu.....	21
Figura 5 - Resultado obtido com método Mean-Shift	21
Figura 6 - Centro de massa e vetor Mean-Shift.....	22
Figura 7 - Aplicação erosão.....	23
Figura 8 - Aplicação dilatação.....	24
Figura 9 - Neurônio artificial.....	25
Figura 10 - Enquadramento do reconhecimento de leucócitos	26
Figura 11 - Extração dos grupos.....	28
Figura 12 - Fluxograma de processamento empregado no sistema.....	29
Figura 13 - Diagrama de casos de uso	32
Figura 14 – Fluxo principal do protótipo	33
Figura 15 - Diagrama de atividades para delimitação do campo e descoberta de células.....	33
Figura 16 - Imagem de entrada.....	34
Figura 17 - Transformação RGB para HSV	35
Figura 18 - Separação dos canais HSV	35
Figura 19 - Limiarização do canal de V e seleção da componente	36
Figura 20 - Delimitação da região do canal S	37
Figura 21 - Limiarização do canal S.....	38
Figura 22 - Operações morfológicas	38
Figura 23 - Validações geométricas	39
Figura 24 - Aplicação do crescimento de região isonômico	40
Figura 25 - Imagem final da fase de realce do campo.....	41
Figura 26 - Criação da lista de componentes conexas.....	42
Figura 27 - Crescimento uniforme de região.....	42
Figura 28 - Imagem da fase final de descoberta.....	43
Figura 29 - Diagrama de atividades do tratamento de células.....	44
Figura 30 - Resultado da clusterização.....	45
Figura 31 - Resultado do isolamento do núcleo	45

Figura 32 - Lista de imagens da seleção de componentes.....	46
Figura 33 - Pré-processamento do canal H.....	47
Figura 34 - Extração do citoplasma.....	48
Figura 35 - Características do núcleo	48
Figura 36 - Características do Citoplasma.....	49
Figura 37 - Tela principal do protótipo	50
Figura 38 - Ação para carregar a imagem	51
Figura 39 - Parametrizações do protótipo.....	51
Figura 40 - Casos descartados	52
Figura 41 - Leucócitos típicos	53
Figura 42 - Tonalidades na lâmina	53
Figura 43 - Casos inconsistentes	54
Figura 44 - Teste experimental.....	55
Figura 45 - Extração de citoplasmas com ruídos.....	56

LISTA DE QUADROS

Quadro 1 - Resultado do reconhecimento de leucócitos Lohita.....	27
Quadro 2 - Características dos trabalhos correlatos	29
Quadro 3 - Carregamento da imagem	34
Quadro 4 - Transformação RGB para HSV	34
Quadro 5 - Separação dos canais HSV	35
Quadro 6 - Busca da maior componente	36
Quadro 7 - Inversão binária e subtração	37
Quadro 8 - Operações morfológicas	38
Quadro 9 - Validações geométricas.....	39
Quadro 10 - Crescimento de região isonômico	40
Quadro 11 - Criação da lista de componentes conexas	41
Quadro 12 - Crescimento uniforme de região	42
Quadro 13 - Extração da imagem-região.....	43
Quadro 14 - Algoritmo Mean-Shift	44
Quadro 15 - Isolamento do núcleo	45
Quadro 16 - Pré-processamento a partir do canal M.....	46
Quadro 17 - Seleção da componente e reconstrução da célula.....	47
Quadro 18 - Subtração do citoplasma.....	47
Quadro 19 - Treinamento da rede neural.....	49
Quadro 20 - Utilização da rede neural.....	50
Quadro 21 - Imagens com lâmpada branca	65
Quadro 22 - Imagens com lâmpada amarela	66

LISTA DE TABELAS

Tabela 1 - Resultados da descoberta de células.....	54
Tabela 2 - Resultados da extração de núcleos	55
Tabela 3 - Resultados da extração de citoplasmas	56
Tabela 4 - Resultado sintético da classificação	56
Tabela 5 - Resultado analítico da classificação	57

LISTA DE ABREVIATURAS E SIGLAS

API - Application Programming Interface

EE - Elemento Estruturante

HSV - Hue, Saturation, Value

MLP - MultiLayer Perceptron

RF - Requisitos Funcionais

RGB - Red, Green, Blue

RNA - Rede Neural Artificial

RNF - Requisitos Não Funcionais

UML - Unified Modeling Language

SUMÁRIO

1 INTRODUÇÃO.....	15
1.1 OBJETIVOS.....	16
1.2 ESTRUTURA DO TRABALHO.....	16
2 FUNDAMENTAÇÃO TEÓRICA.....	17
2.1 COLETA E ANÁLISE DA AMOSTRA DE SANGUE.....	17
2.2 SEGMENTAÇÃO DE IMAGENS.....	19
2.2.1 LIMIARIZAÇÃO.....	20
2.2.2 CLUSTERIZAÇÃO.....	21
2.3 MORFOLÓGIA MATEMÁTICA.....	22
2.4 REDE NEURAL.....	24
2.5 TRABALHOS CORRELATOS.....	25
2.5.1 An extensible computer vision application for blood cell recognition and analysis.....	26
2.5.2 Avaliação da proliferação celular baseada em PCNA-ciclina por análise morfológica aplicada a linfomas malignos não-hodgkin.....	27
2.5.3 Sistema para contagem automática de células sanguíneas através de visão computacional.....	28
2.5.4 Comparação entre trabalhos correlatos.....	29
3 DESENVOLVIMENTO DO PROTÓTIPO.....	31
3.1 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO.....	31
3.2 ESPECIFICAÇÃO.....	31
3.2.1 Técnicas e ferramentas utilizadas.....	31
3.2.2 Diagrama de casos de uso.....	32
3.2.3 Fluxo principal do protótipo.....	32
3.2.4 Operacionalidade da implementação.....	50
3.3 RESULTADOS E DISCUSSÕES.....	51
3.3.1 Montagem da base de teste.....	52
3.3.2 Delimitação do campo e descoberta de células.....	54
3.3.3 Extração do núcleo.....	54
3.3.4 Extração do citoplasma.....	55
3.3.5 Classificação.....	56
4 CONCLUSÕES.....	59

4.1 LIMITAÇÕES.....	61
4.2 EXTENSÕES	61
REFERÊNCIAS	62
APÊNDICE A – BASE DE IMAGENS	65

1 INTRODUÇÃO

A humanidade desde seu surgimento busca soluções para as mais diversas patologias existentes que comprometem constantemente a saúde de parte da população. Dessa forma, desenvolveram-se ferramentas de reconhecimento dos diversos componentes que formam o ser humano a fim de estabelecer os estados de saúde e os causadores das doenças, para que os tratamentos possam ser efetuados aumentando a qualidade de vida dos indivíduos (HALL, 2011; HEGENBERG, 1998).

Nesse contexto, surgiram os laboratórios, que nas últimas décadas focaram suas atenções para os processos de análises e resultados, visando desenvolver metodologias e programas que auxiliem os processos, prevenindo e minimizando a ocorrência de falhas. Atualmente, os laboratórios de análise não medem esforços para compreender as causas de seus erros: nas análises de processos do trabalho, na avaliação da qualidade e redução de falhas (SCIACOVELLI et al, 2011).

Estes esforços buscam a diminuição na taxa de erros laboratoriais nas etapas que envolvem a análise da amostra. Sendo que os erros laboratoriais ocasionalmente impactam no diagnóstico, desse modo, sempre há necessidade de criar metodologias que confirmem a qualidade dos resultados obtidos. Logo, se indicadores de qualidade baseados em evidências são desenvolvidos, implementados e monitorados, pode-se descobrir pontos fracos e identificar soluções adequadas (PLEBANI, 2009).

Segundo Campestrini (1991), dentre os processos que são utilizados para diagnosticar a saúde do paciente existe uma gama imensa de exames laboratoriais. No entanto, o exame mais comumente solicitado é o hemograma, visto que fornece informações relevantes sobre o estado de saúde do paciente, proporcionando o diagnóstico através do reconhecimento e quantificação das células do sistema de defesa do corpo.

Segundo Failace (2003), o hemograma está presente em cerca de 48% das listas de exames realizados por pacientes. Essa demanda denota que a contagem de leucócitos é indispensável no diagnóstico e controle evolutivo de diversas patologias. Assim, destaca-se a importância na qualidade durante todo processo do exame. E, para isso, uma coleta adequada, um esfregaço bem executado e corado corretamente, contribui para que o profissional possa realizar a contagem e diferenciação do esfregaço sanguíneo.

A contagem dos leucócitos é realizada utilizando o esfregaço sanguíneo corado que é observado no microscópio por um profissional. Este observa em campos distintos da lâmina as células, diferenciando e contando-as (NAOUM; NAOUM, 2008).

Neste âmbito, a visão computacional que enfoca suas tecnologias para extração de informações a partir de imagens pode contribuir para a contagem e diferenciação das células de um esfregaço sanguíneo observado através do microscópio. A partir das imagens que são capturadas em um microscópio, estas imagens fornecem dados de alto nível, sendo a visão computacional responsável em manipular estes dados.

Diante desse contexto, este trabalho apresenta o desenvolvimento de um protótipo para o reconhecimento de leucócitos, do tipo linfócito e neutrófilo, através de técnicas de visão computacional para extração de informações e a utilização de uma rede neural do tipo MLP para classificação das células.

1.1 OBJETIVOS

O objetivo deste trabalho é o desenvolvimento de um protótipo de reconhecimento de células do sistema imune do tipo linfócitos e neutrófilos utilizando visão computacional.

Os objetivos específicos do trabalho são:

- a) efetuar a segmentação das células existentes em uma lâmina de sangue;
- b) realizar a extração de características morfológicas das células;
- c) classificar as células de acordo com o seu tipo: linfócitos e neutrófilos.

1.2 ESTRUTURA DO TRABALHO

Este trabalho está dividido em quatro capítulos. O primeiro capítulo apresenta os objetivos e a motivação para desenvolvimento do trabalho. O segundo capítulo trata da fundamentação teórica do trabalho, explicando os principais conceitos e técnicas utilizadas no desenvolvimento do protótipo. No terceiro capítulo são descritos a arquitetura do trabalho através de diagramas, o detalhamento da implementação do protótipo e os resultados obtidos nos testes realizados. Por fim, são apresentadas as conclusões e limitações do trabalho, assim como sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

A seção 2.1 descreve a coleta e análise da amostra de sangue. A seção 2.2 aborda técnicas de segmentação de imagens. A seção 2.3 dedica-se a explicar o conceito da morfologia matemática e suas principais operações. Na seção 2.4 são apresentados conceitos de Rede Neural e o algoritmo *backpropagation*. Por fim, a seção 2.5 apresenta os trabalhos correlatos.

2.1 COLETA E ANÁLISE DA AMOSTRA DE SANGUE

O ser humano está constantemente susceptível a inúmeras doenças e para combatê-las o sistema imunológico será responsável pela proteção da saúde. No que se refere a proteger-se, o corpo humano apresenta vários recursos. Inicialmente existe a barreira física que se constitui principalmente da pele, impedindo os microrganismos se adentrarem. Caso algum agente consiga abrir uma brecha e invadir o organismo, ele se depara com uma rede gigantesca de proteção dos tecidos internos, sendo os leucócitos os principais encarregados de eliminar qualquer intruso. Essas células protetoras são, portanto, indicativos do estado de saúde dos pacientes pode ser obtidos através da quantificação e diferenciação no hemograma (PARSLOW et al., 2004; LORENZI et al., 2003). Elas, também podem ser quantificadas a fim de se obter informações sobre o estado de saúde. Assim, o exame hemograma concede um relatório numérico dos elementos sanguíneos: eritrócitos, leucócitos e plaquetas, constando descrições de suas características (WILLIAMSON; SNYDER, 2013, p. 200). A quantidade desses componentes encontrados indica a resposta do organismo a vários tipos de doenças (LAB TEST ONLINE BR, 2016).

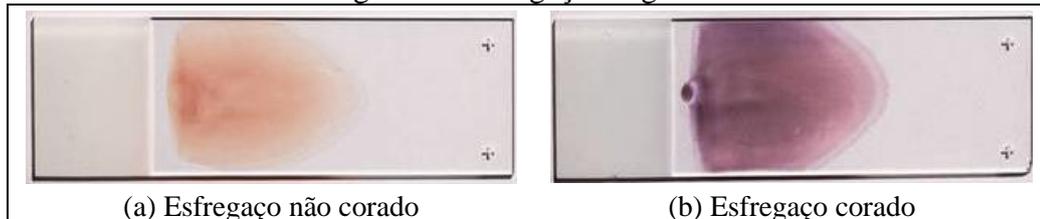
Para a realização do hemograma, qualquer vaso sanguíneo pode ser utilizado para a realização da punção. Entretanto, a região preferencial dos profissionais é a fossa anticubital da área anterior do braço, devido à localização que facilita o procedimento. Esse material coletado é separado em um tubo contendo anticoagulante e em um esfregaço (lâmina com sangue), onde ambos são utilizados para a análise dos leucócitos (MELO; SILVEIRA, 2015).

A avaliação dos leucócitos é realizada de duas formas: manual através do esfregaço sanguíneo, uma lâmina de vidro com o sangue do paciente, cujo preparo inclui adição de corantes que realçam as estruturas para a observação e diferenciação no microscópio óptico; e automatizada, que utilizada de equipamentos eletrônicos que disponibilizam a contagem (MONTEIRO, 2005, p. 24).

O tubo de sangue contendo anticoagulante é utilizado na análise automatizada, na qual, através de métodos químicos, a contagem diferencial é realizada. Já o esfregaço, é realçado

com corante e segue para a observação diferencial realizada por um profissional. Essas etapas possibilitam que a contagem e diferenciação dos leucócitos sejam mais completas e fidedignas (MAKRIS; GREAVES, 1998). Dentre os leucócitos existentes, os neutrófilos e linfócitos são de interesse desse trabalho (FAILACE; FERNANDES, 2015). A Figura 1 apresenta a imagem de um esfregaço.

Figura 1 - Esfregaço sanguíneo



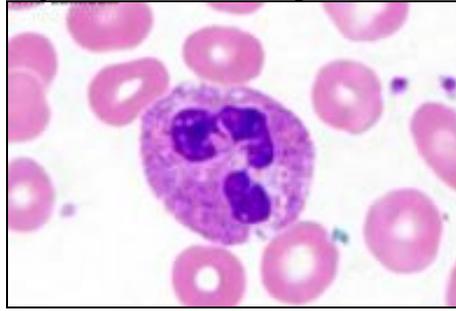
Fonte: Leukos (2013).

A análise manual do esfregaço assegura um melhor resultado, através da revisão dos dados apresentadas na contagem automatizada do hemograma, tanto na morfologia quanto na quantificação das linhagens leucocitárias. Constatando assim, através da contagem total e diferencial, a resposta do organismo a vários tipos de doenças (MAKRIS; GREAVES, 1998). Normalmente em um esfregaço sanguíneo os linfócitos e neutrófilos juntos representam cerca 85% do total de leucócitos (FAILACE, 2003, p. 178; LAB TEST ONLINE BR, 2016).

Dentre os leucócitos, os neutrófilos são identificados como células que possuem o núcleo segmentado e citoplasma com grânulos finos. De modo geral, possuem a função de englobar e destruir as bactérias que invadem o corpo. Por outro lado, os linfócitos, são células pequenas com núcleo redondo e quantidade pequena de citoplasma, dentre suas funções, pode-se destacar sua ação de combater agentes virais. Sendo assim, ao realizar um exame de sangue, a contagem alterada destes tipos indica uma possível ação patológica no organismo do paciente. Dessa forma, fica mais evidente a necessidade de um exame que avalie as respostas do organismo (RAPAPORT, 1990).

Os neutrófilos formam um exército de fagócitos circulantes, preparados para responder imediatamente e em grande número sempre que necessário. São os leucócitos mais abundantes: no adulto normal representam entre 50% a 75% do total circulante. Possuem um núcleo que se cora em púrpura e é dividido em 2 a 5 segmentos ou lóbulos, conectados por estreita faixa ou filamento de material nuclear. A cromatina é heterogênea e o citoplasma é azul-pálido e cheio de fina granulação que se cora em lilás. Esses grânulos são ditos neutrofílicos, pois adquirem o componente ácido e básico do corante (BAIN, 1998; PARSLOW et al., 2004). Na Figura 2 é apresentada uma célula do tipo neutrófilo.

Figura 2 - Célula do tipo neutrófilo

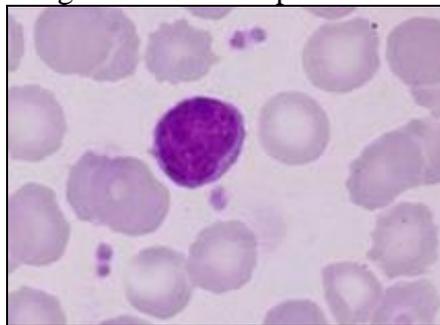


Fonte: Pathpedia (2016).

Esse exército celular possui como principal função fagocitar partículas, em especial bactérias. Quando o organismo é acometido por alguma doença bacteriana, os neutrófilos iniciam uma resposta que pode ser percebida pelo aumento de seu número no sangue. Assim a importância de sua contagem é compreendida, quando sua alteração é observada pode auxiliar o melhor diagnóstico (PARSLOW et al., 2004; LORENZI et al., 2003).

Os linfócitos seguem em segundo lugar em quantidade no sangue: aproximadamente 20% a 45%. São células menores do que os neutrófilos, possuem um contorno arredondado ou irregular e citoplasma claro, azul-pálido e são agranulares. A cromatina nuclear é condensada, sendo mais densa e homogênea em linfócitos menores, ficando com uma coloração mais suave em linfócitos maiores (BAIN, 1998). A Figura 3 apresenta uma célula do tipo linfócito.

Figura 3 - Célula tipo linfócito



Fonte: Ricci (2016).

Os linfócitos são encontrados no sangue em proporções que variam de acordo com as condições fisiológicas e realizam várias funções na resposta imune, dentre elas, está envolvida em resposta a antígenos virais. Da mesma forma que as outras células, quando é identificado a quantidade destas células, são discriminadas no hemograma para que se possa contribuir no diagnóstico de patologias (LORENZI et al., 2003).

2.2 SEGMENTAÇÃO DE IMAGENS

A segmentação de imagens possui diversas aplicações práticas, como equipamentos de tomografia, localização de objetos em imagens de satélite, radares de trânsito e

reconhecimento facial (ATAIDE; PIMENTEL 2008). Para encontrar uma solução através da segmentação de imagens, é necessário encontrar um fluxo empírico e prever ajustes para superar dificuldades encontradas como falta de contraste, bordas irregulares, sombra, elementos sobrepostos e outros fatores que podem gerar um resultado de falso. Logo, deve-se encontrar uma segmentação de imagens que procure diferenciar os pixels um dos outros e do fundo da imagem, permitindo assim interpretar pixels contíguos e agrupá-los em regiões (SOARES, 2008, p. 51-52). Sendo assim, Petrou e Bosdogianni (1999, p. 527) apontam a segmentação de imagens como uma das principais áreas de pesquisa do processamento de imagens.

Em imagens complexas, a segmentação mostra-se como uma das atividades mais árduas da área de processamento de imagens, sendo que sua precisão indica um eventual sucesso ou erro de apuração de análise computadorizada. Um ponto importante neste processo é aplicá-lo unicamente para a identificação dos objetos ou regiões de interesse, não havendo obrigação de segmentar uma imagem acima do nível de detalhe necessário para o reconhecimento destes elementos (GONZALES; WOODS, 2008, p. 689).

Comumente existe uma dificuldade do processo de segmentação de imagem que é o fato da variação de estruturas contidas nas imagens. Logo, é necessário identificar estas características a partir da geometria, forma, topologia, textura, cor e brilho, sendo escolhidas aquelas que possibilitam melhor a identificação das características desejadas (SOARES, 2008, p. 51).

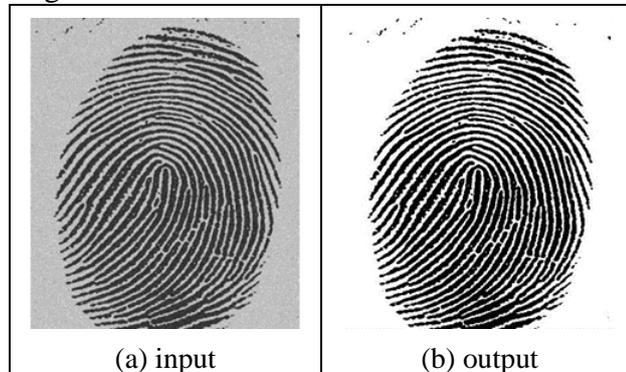
Assim, neste trabalho destacam-se duas abordagens, a limiarização e a morfologia matemática. Onde Russ (2011, p. 395) aponta o processo de limiarização como uma técnica simples, e que pode separar regiões de interesse do restante da imagem. E, conforme Gonzales e Woods (2008, p. 650) e, Costa e Cesar Jr. (2000, p. 255), a morfologia matemática insere uma abordagem para resolução de inúmeros problemas de pré-processamento de imagens, filtros de ruídos, detecção de formas, associação de padrões entre outros problemas que a morfologia matemática pode ser aplicada.

2.2.1 LIMIARIZAÇÃO

O processo de limiarização pode ser comparado a uma problemática relacionada à decisão estatística, onde o objetivo é minimizar a taxa média de erro ocasionada na classificação dos pixels em duas classes, nas cores preto e branco, que pode ser resolvida baseada na técnica de Otsu (GONZALES; WOODS, 2008). O método de Otsu calcula um histograma e, a partir disto, têm-se a divisão entre duas classes, onde cada classe possuirá seus

níveis de cinza com as características de sua média e desvio padrão. Através deste procedimento, é obtida uma nova imagem destacando a classe de interesse do restante da imagem (GAZZIRO, 2013). Na Figura 4 é apresentada o resultado utilizando o método de Otsu.

Figura 4 - Resultado obtido com método Otsu



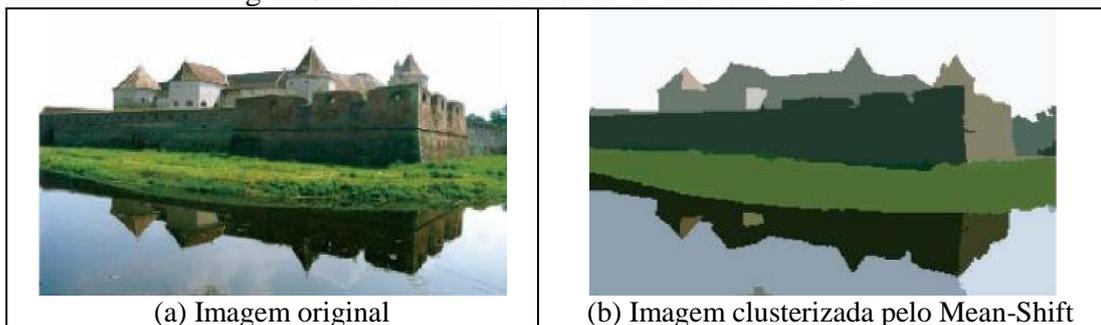
Fonte: Gonzales e Woods (2008, p. 743).

O método de limiarização utiliza a igualdade da intensidade dos pixels, ou seja, a divisão de uma imagem é realizada sobre regiões similares de acordo com um limiar. (GONZALES; WOODS, 2008, p. 712). O conceito por trás deste algoritmo consiste em definir um valor de critério, percorrer os pontos da imagem e separar os pixels que possuem valor maior que este limiar como sendo pontos de interesse, enquanto os outros são considerados pontos do plano de fundo (GONZALES; WOODS, 2008, p. 760).

2.2.2 CLUSTERIZAÇÃO

Uma das abordagens da segmentação de imagem é o agrupamento de regiões, conhecido também por clusterização. Uma das técnicas de clusterização conhecidas é o Mean-Shift (deslocamento pela média) também conhecido como estimação do gradiente de uma função de densidade. Trata-se de um algoritmo de rastreamento utilizado para achar modas (regiões mais densas) em um conjunto de dados, manifestando uma função de densidade de probabilidade (RÉZIO, 2008, p. 33). Na Figura 5 é apresentada o resultado da técnica do Mean-Shift.

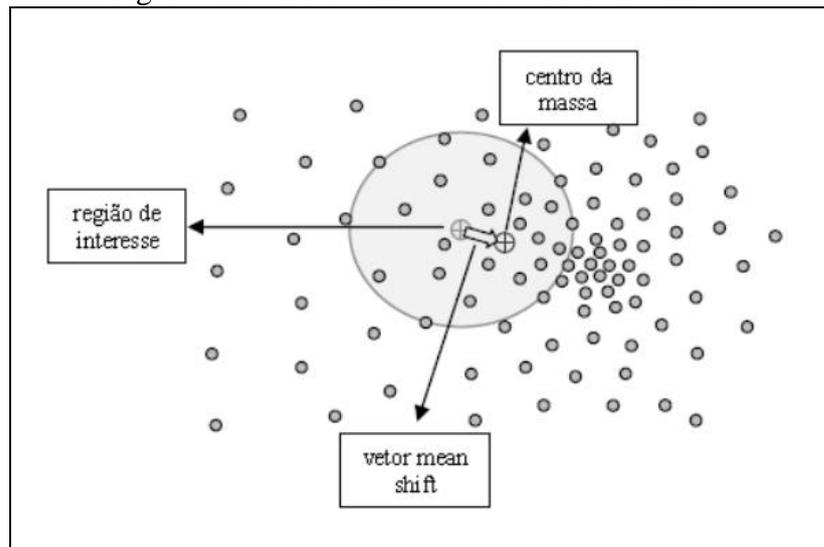
Figura 5 - Resultado obtido com método Mean-Shift



Fonte: Rézio (2008).

O Mean-Shift, através do histograma de cores, detecta a região com maior densidade de pontos de uma determinada cor e calcula uma média ponderada para encontrar o centro de massa desta região mais densa. Em seguida, é traçado um vetor entre o centro da região de interesse e o centro de massa obtido, conhecido como vetor de Mean-Shift que aponta sempre na direção do maior incremento na densidade (RÉZIO, 2008, p. 35), conforme pode ser visto na Figura 6.

Figura 6 - Centro de massa e vetor Mean-Shift



Fonte: Rézio (2008, p. 34).

A estimação de densidade do gradiente realizada por este algoritmo é não-paramétrica, ou seja, onde as probabilidades da função de densidade são extraídas diretamente do conjunto de dados, sendo essa a vantagem deste algoritmo, pois não é necessária uma avaliação da imagem por inteiro gerando um bom desempenho computacional. Esse processo repete até que não se encontrem regiões com maior densidade do que a região já obtida, ou seja, o centro de massa coincide com o centro da região ou a diferença entre eles é insignificante (RÉZIO, 2008, p. 35).

2.3 MORFOLÓGIA MATEMÁTICA

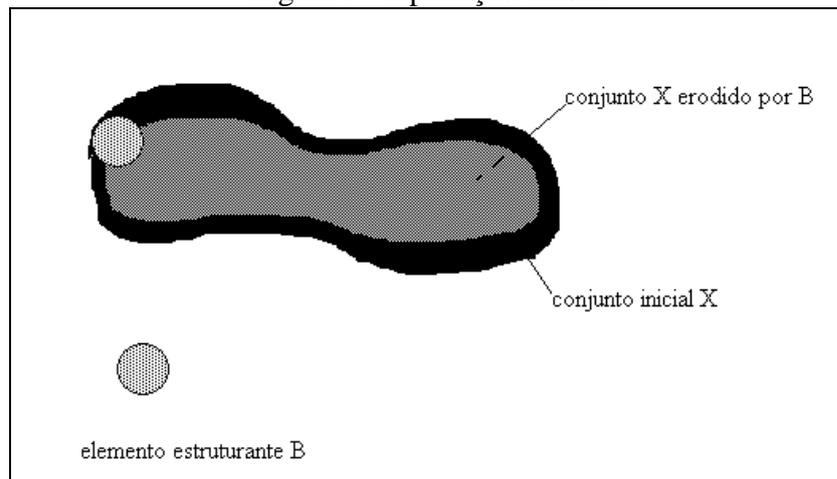
A morfologia matemática tem origem no trabalho dos matemáticos franceses J. Serra e G. Matheron publicado no meio dos anos 1960 e tem sido aplicada a uma série de problemas de processamento de imagens (COSTA; CESAR JR., 2000, p. 255). Dentre as operações morfológicas, a erosão e dilatação se destacam neste trabalho.

Segundo Facon (1996, p. 2), o princípio da morfologia matemática é extrair uma informação referente à geometria de um conjunto da imagem, aplicando uma transformação com outro conjunto denominado de Elemento Estruturante (EE). De acordo com sua forma e

tamanho é possível testar se o EE está ou não contido no conjunto desconhecido da imagem (FACON, 1996, p. 2). Ainda segundo Gonzales e Woods (2000, p. 370), a linguagem da morfologia matemática é fortemente baseada na teoria dos conjuntos, fornecendo dessa forma uma abordagem poderosa para diversos problemas em relação ao processamento de imagens.

Segundo Costa e Cesar Jr. (2000, p. 260), a erosão é comumente aplicada no pré-processamento de imagens, pois pode ser utilizada no processo de redução de ruídos. A Figura 7 mostra a aplicação da erosão em uma imagem.

Figura 7 - Aplicação erosão

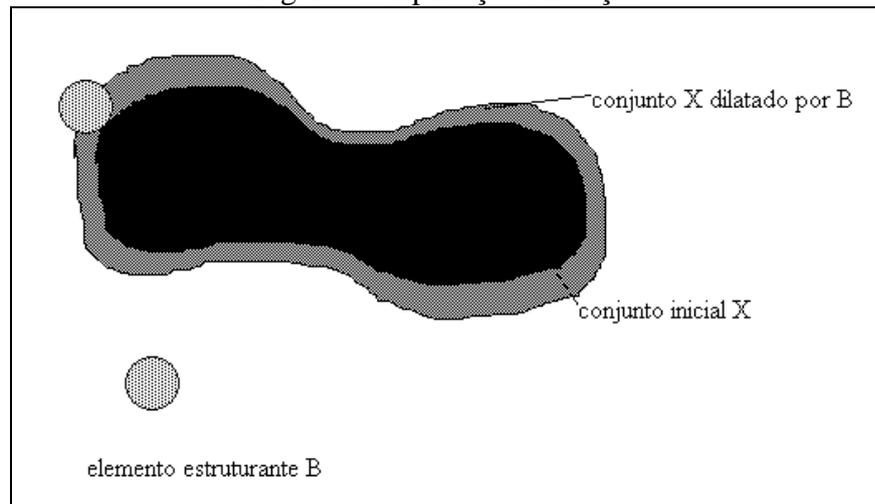


Fonte: Facon (1996).

Pode se dizer que a erosão funciona como se o EE deslizesse internamente pelo conjunto X, removendo a região deste até o seu ponto central. Essa característica de diminuição dos agrupamentos de pixels da imagem faz com que pequenas regiões com tamanho inferior ao do EE sejam eliminadas, o que é útil na remoção de pequenos ruídos granulares na imagem. A erosão também possui as características de aumentar os furos presentes na imagem além de separar um conjunto de pixels que está ligado por uma região estreita com espessura inferior ao tamanho do EE.

De forma semelhante à operação de erosão, a operação de dilatação morfológica envolve uma imagem binária a ser processada e um EE (COSTA; CESAR JR., 2000, p. 255). De acordo com (GONZALES; WOODS, 2008, p. 655), de modos opostos à operação de erosão, a operação de dilatação aumenta ou engrossa os objetos da imagem, sendo uma das aplicações mais simples da dilatação é o preenchimento de lacunas. A Figura 8 demonstra aplicação da dilatação em uma imagem.

Figura 8 - Aplicação dilatação



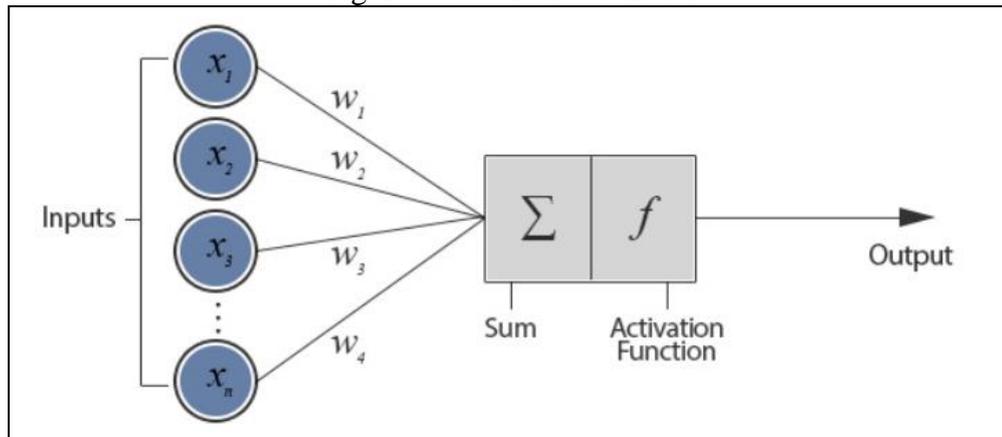
Fonte: Facon (1996).

Por analogia, é possível dizer que a dilatação funciona como se o EE deslizesse pelo contorno do conjunto X , promovendo um crescimento da dessa região pelo seu ponto central. Com essa característica de expansão das regiões, ela possui os efeitos de preenchimento espaços vazios com tamanho inferior ao EE, além de conectar agrupamentos de pixels distintos que estão separados por uma distância inferior ao tamanho do EE.

2.4 REDE NEURAL

Uma Rede Neural Artificial (RNA) pode ser definida como um conjunto de neurônios artificiais simulando o funcionamento do cérebro humano, onde esta rede neural é treinada para alcançar um processo inteligente esperado a partir de um conjunto de informações de entrada e saída (WU et al., 2007). O funcionamento de uma rede neural geralmente envolve três etapas: treinamento, validação e utilização. Durante a fase de treinamento os parâmetros podem mudar para serem ajustados, a etapa de validação fixa os parâmetros e é verificado se o sistema é capaz de resolver determinado problema. Atingindo um percentual aceitável de acerto a rede pode ser empregada na utilização do problema a qual foi treinada (KLEIN; MARTINS, 2006). A Figura 9 apresenta o neurônio artificial de uma RNA.

Figura 9 - Neurônio artificial



Fonte: Klein e Martins (2006).

Um das características mais relevantes nas RNAs é a capacidade de aprender através de valores de entrada e de generalizar sobre os padrões do conjunto de dados de treinamento, tornando possível a interpretação de outros padrões similares, mas não necessariamente idênticos aos utilizados durante o aprendizado (MACIEL, 2005, p. 20). Uma rede neural padrão do tipo Perceptron é um modelo limitado a separação de dados lineares. Já uma rede neural do tipo MultiLayer Perceptron (MLP) utiliza o aprendizado supervisionado através do algoritmo *backpropagation* que pode separar dados não lineares (KLEIN; MARTINS, 2006).

O algoritmo *backpropagation* é um processo de aprendizado supervisionado que pode ser dividido em duas etapas. A primeira chamada de *forward* que consiste no estabelecimento do valor de saída da rede ao se apresentar uma entrada. A segunda denominada *backward* utiliza a comparação entre a saída desejada e a saída calculada para atualizar os pesos da rede. Os pesos são atualizados a partir da camada de saída para a última camada escondida e assim sucessivamente até chegar a primeira camada escondida (KLEIN; MARTINS, 2006). Segundo Haykin (2001, p. 183) o modelo MLP tem sido aplicado com nível satisfatório para resolver problemas difíceis através do seu treinamento supervisionado com o algoritmo de retropropagação de erro (*backpropagation*).

2.5 TRABALHOS CORRELATOS

Esta seção tem como objetivo investigar na literatura trabalhos relacionados ao uso de visão computacional aplicada nas áreas da saúde. A seguir serão apresentados três trabalhos voltados ao reconhecimento de células utilizando processamento de imagens digitais, empregando técnicas de segmentação, extração de características e classificação. A seção 2.5.1 descreve o sistema desenvolvido por Priyankara e Silva (2006). Na seção 2.5.2 é apresentado a tese de Weber (1997). Já a seção 2.5.3 descreve um sistema de contagem de

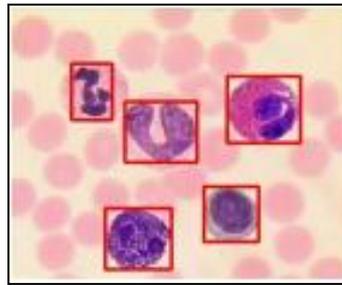
células sanguíneas (FEDECHEN et al., 2012). Por fim, a seção 2.5.4 faz uma pequena comparação entre as características mais importantes dos trabalhos correlatos apresentados.

2.5.1 An extensible computer vision application for blood cell recognition and analysis

Priyankara e Silva (2006) propuseram um sistema especializado em reconhecimento células de sangue que pudesse ser utilizado em países que não conseguem fazer investimentos necessários para aquisição de equipamentos laboratoriais para automatização de exames. O sistema desenvolvido foi denominado de Lohitha e, utiliza visão computacional e rede neural como solução para a contagem e classificação de glóbulos brancos.

O esfregaço sanguíneo realça as cores dos glóbulos brancos, com isso o Lohitha aplica um histograma RGB para remover o fundo e ficar com as áreas de interesse em torno das células. Após a remoção do fundo, é extraída uma imagem binária e outra imagem com a borda em torno das células. Em seguida é realizada a remoção dos ruídos através da mediana e traçada uma região quadrangular em torno da célula. Posteriormente, foi aplicado um método de saturação com valores estatísticos para validar as células candidatas e eliminar componentes não interessados. Por fim, foi utilizada uma rede neural com cerca de 100 imagens para classificar as células. A Figura 10 demonstra o sistema Lohitha aplicado em uma imagem de interesse.

Figura 10 - Enquadramento do reconhecimento de leucócitos



Fonte: Priyankara e Silva (2006, p. 8).

Segundo Priyankara e Silva (2006), Lohitha é uma aplicação de domínio muito específico e possui um método eficaz para o reconhecimento e contagem de células de sangue como uma alternativa para a contagem manual feita por profissionais e principalmente em lugares com baixos recursos operacionais e carência de profissionais especializados. No Quadro 1 os resultados obtidos na contagem diferencial de leucócitos pelo sistema Lohitha que utilizou uma rede neural treinada com 130 células.

Quadro 1 - Resultado do reconhecimento de leucócitos Lohita

Output Input	Input cell identified (as a percentage)				
	Neutrophils	Eosinophils	Basophils	Lymphocytes	Monocytes
Neutrophils	85%	14%	0%	0%	1%
Eosinophils	13%	80%	0%	0%	7%
Basophils	0%	0%	12%	61%	27%
Lymphocytes	0%	0%	3%	82%	15%
Monocytes	0%	2%	3%	18%	77%

Fonte: adaptado de Fedechen et al (2012, p. 6).

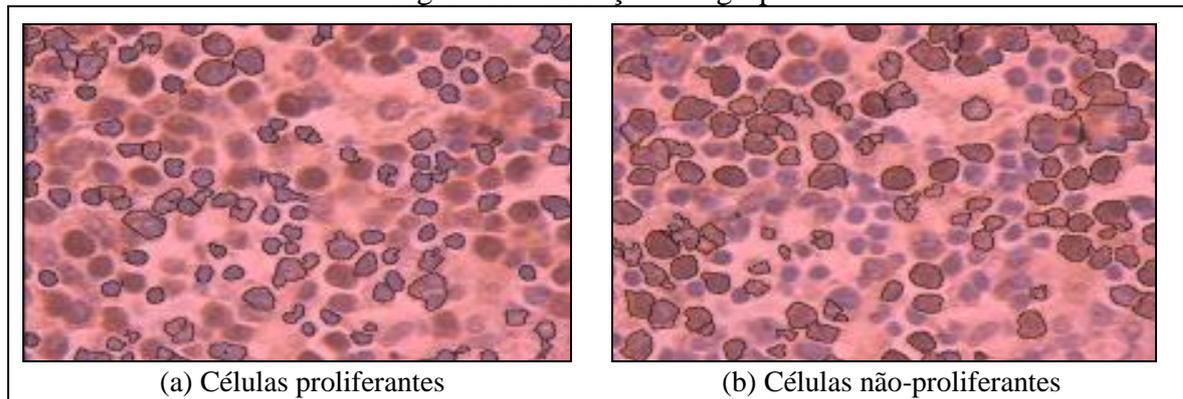
Entre os resultados obtidos, observa-se um percentual satisfatório em quatro tipos de leucócitos. No entanto, os basófilos tiveram 12% de acurácia, que segundo Priyankara e Silva (2006) isto pode estar relacionado ao treinamento da rede neural, que para este tipo foi utilizado 5 basófilos.

2.5.2 Avaliação da proliferação celular baseada em PCNA-ciclina por análise morfológica aplicada a linfomas malignos não-hodgkin

A proposta de Weber (1997) é a automatização do processo de análise microscópica de células de tumores malignos do sistema linfático por meio de visão computacional. Além de realizar a contagem automática das células de interesse que são coloridamente distintas. Visto que o tradicional método de contagem é realizado por profissionais que analisam amostras de maneira manual, tornando-o um método lento, repetitivo, cansativo e sujeito a falhas humanas.

Com uma imagem adquirida de uma lâmina, corada com o anticorpo monoclonal anti-PCNA-Ciclina, observa-se duas famílias de células, as azuladas e as marrom-avermelhado. A partir disso, são estimadas as quantidades de células proliferadas (marrons) sobre a totalidade (marrons + azuladas). A segmentação inicial entre o fundo e os objetos de interesse baseia-se na média das intensidades, uma vez que o fundo é mais claro que as células. Após a separação dos elementos de interesse do fundo, operações morfológicas (reconstrução, erosão e dilatação) removem o restante dos resíduos. Com a nova imagem aplica-se novamente uma média das intensidades para que seja possível classificar as células entre o grupo azul e grupo avermelhado. Na Figura 11 é apresentada a separação das células.

Figura 11 - Extração dos grupos



Fonte: Weber (1997).

Com as técnicas utilizadas por Weber (1997) foi possível distinguir as células não proliferantes das proliferantes, tornando possível a viabilidade da utilização de sistemas baseados em processamento de imagem no reconhecimento destas células.

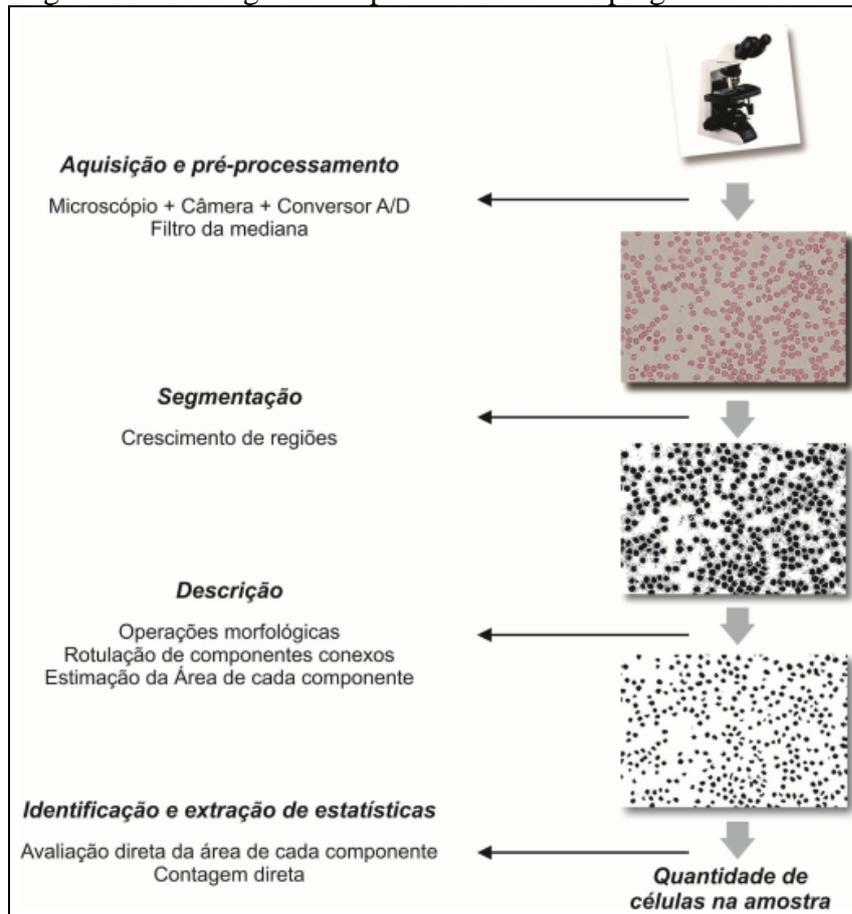
2.5.3 Sistema para contagem automática de células sanguíneas através de visão computacional

O trabalho de Fedechen et al. (2012) teve como finalidade o reconhecimento de células sanguíneas de uma amostra do esfregaço de sangue. O objetivo do sistema é reconhecer as células e realizar a contagem automática através de processamento de imagens utilizando técnicas de visão computacional. O sistema foi implementado na linguagem C++ e utiliza biblioteca OpenCV.

Para alcançar o objetivo, Fedechen et al. (2012) utilizaram imagens capturadas por uma câmera sob o microscópio de uma amostra biológica de células do esfregaço sanguíneo. Estas imagens coloridas foram submetidas a um pré-processamento utilizando um filtro de mediana para eliminar ruídos. Em seguida, foi empregado um algoritmo de inundação em busca da separação das áreas de interesse e do fundo da imagem, cuja inundação mostrou maior robustez do que uma limiarização direta, principalmente quando observada a diferença de iluminação.

Na etapa subsequente, foi obtido uma imagem binária onde as células são destacadas do fundo. Em seguida, foi aplicado uma sequência de operações morfológicas de erosão e dilatação para que as células sejam diferenciadas de outras componentes e ruídos. Por fim, cada um dos elementos foi reconhecido pelo algoritmo de rotulação de componentes conexas e calculada área para validar elementos dentro de uma faixa aceitável. A Figura 12 mostra o fluxo de processamento aplicado na imagem de interesse.

Figura 12 - Fluxograma de processamento empregado no sistema



Fonte: Fedechen et al (2012, p. 6).

As imagens processadas pelo sistema de Fedechen et al. (2012) foram encaminhadas para um especialista humano para fazer a contagem manual. Os resultados apresentados pelo sistema e pelo especialista humano foram submetidos à análise estatística pelo teste do Chi-Quadrado, concluindo-se que o sistema apresentou resultados satisfatórios na contagem de células sanguíneas a partir de imagens microscópicas.

2.5.4 Comparação entre trabalhos correlatos

O Quadro 2 apresenta as características mais relevantes dos trabalhos apresentados nesta seção.

Quadro 2 - Características dos trabalhos correlatos

Trabalho	Fedechen et al. (2012)	Lohitha (2006)	Weber (1997)
Características			
Público alvo?	Especialista	Especialista	Especialista
Técnicas de classificação?	Contagem e validação de componentes	Rede neural	Matemática morfológica
Categoria de células?	Hemácias	Hemácias/Leucócitos	Linfomas

Fonte: elaborado pelo autor.

A partir do Quadro 2, é possível observar que os trabalhos destinam-se a profissionais qualificados, onde é necessário um material previamente preparado e cuja aquisição das imagens precisa ser de um ambiente controlado. Sendo assim, em nenhum momento o profissional é substituído, onde os sistemas podem servir como apoio e contribuir de maneira qualitativa.

Percebe-se que todos os trabalhos são voltados a exames de laboratório e que o material coletado para obtenção das informações é o sangue do paciente. Em uma escala temporal dos trabalhos, nota-se que o sangue do paciente continua servindo como importante fonte de informações no que diz ao estado de saúde do paciente e motivação para desenvolvimento de sistemas que geram resultados baseados em evidência.

3 DESENVOLVIMENTO DO PROTÓTIPO

Neste capítulo são demonstradas as etapas do desenvolvimento do protótipo. Na seção 3.1 é demonstrado os principais requisitos. A seção 3.2 apresenta a especificação. A seção 3.3 apresenta de forma detalhada a implementação. Por fim, a seção 3.4 demonstra os resultados dos testes, sugestões e melhorias do protótipo.

3.1 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

Os Requisitos Funcionais (RF) e Requisitos Não Funcionais (RNF) do protótipo de reconhecimento de células típicas do tipo linfócitos e neutrófilos são:

- a) permitir que o usuário selecione uma imagem previamente existente no computador a fim de ser analisada (RF);
- b) utilizar técnicas de realce e segmentação a fim de gerar uma nova imagem, a partir da original, que mostre possíveis componentes de interesse (RF);
- c) realizar a extração do núcleo e citoplasma das células (RF);
- d) realizar o cálculo da área e circularidade (RF);
- e) classificar as células entre linfócitos e neutrófilos através de características morfológicas inferidas em uma rede neural (RF);
- f) ser desenvolvida na linguagem de programação Java (RNF);
- g) utilizar o ambiente de desenvolvimento Eclipse (RNF);
- h) ser desenvolvida utilizando a biblioteca JavaCV (RNF).

3.2 ESPECIFICAÇÃO

A especificação do protótipo foi representada através de diagramas da Unified Modeling Language (UML), utilizando a ferramenta StarUML. Neste trabalho foram desenvolvidos os diagramas de casos de uso e de atividades. A seguir são apresentadas as técnicas e ferramentas utilizadas e a operacionalidade da implementação.

3.2.1 Técnicas e ferramentas utilizadas

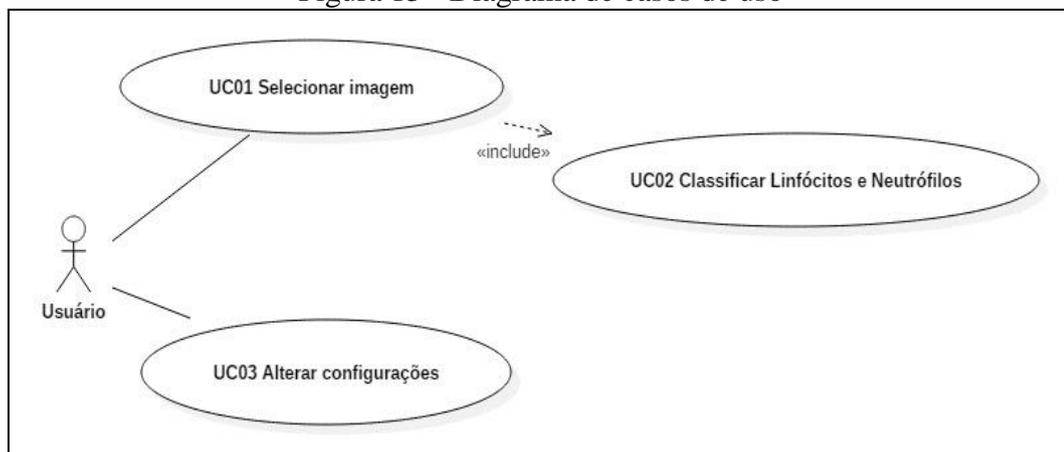
O protótipo foi desenvolvido utilizando a linguagem Java em sua versão 8 no ambiente de desenvolvimento Eclipse versão Neon no sistema operacional Windows 10. A tecnologia utilizada na implementação foi a biblioteca de visão computacional JavaCV. Ela disponibiliza as funções do OpenCV, desenvolvidas em C/C++, para a linguagem Java, utilizando a tecnologia JNI presente no Java para acesso aa DLL's nativas.

O OpenCv fornece diversas soluções como carregamento de imagens, conversões para outros sistemas de cores e separação de canais, segmentação, operações morfológicas, extração de contornos, cálculo de área, cálculo da *bounding box*, implementação de algoritmos como por exemplo, de bordas de Canny, entre outras funcionalidades. Outro recurso utilizado no protótipo é a biblioteca Encog em sua versão para JAVA para classificação das células através de rede neural do tipo MLP.

3.2.2 Diagrama de casos de uso

Esta seção apresenta o diagrama de casos de uso do protótipo, exibido na Figura 13. Identificou-se apenas um ator, denominado *Usuário*, o qual utiliza todas as funcionalidades do protótipo.

Figura 13 - Diagrama de casos de uso



Fonte: elaborado pelo autor.

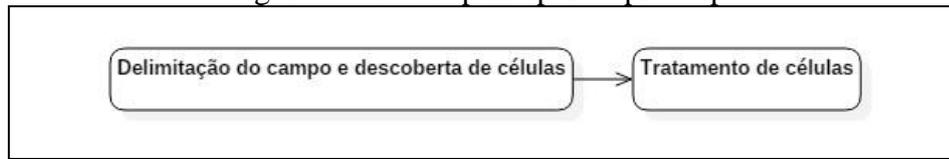
No UC01 - *Selecionar Imagem* o protótipo deve exibir uma janela para que o usuário possa selecionar uma imagem do seu computador. Após o carregamento da imagem, é executado o caso de uso UC02 - *Classificar Linfócitos e Neutrófilos* que responsável por realizar as etapas de processamento de imagens para identificar linfócitos e neutrófilos típicos.

No caso de uso UC03 - *Alterar Configurações* o usuário pode realizar a manutenção dos parâmetros de configuração de sistema que são empregados nas etapas de segmentação, extração de características e classificação das células.

3.2.3 Fluxo principal do protótipo

Para melhor compreensão, o fluxo do protótipo foi dividido em duas fases distintas, onde inicialmente é tratado a delimitação do campo e descoberta de células, e, posteriormente a fase de tratamento de células, conforme mostra a Figura 14.

Figura 14 - Fluxo principal do protótipo



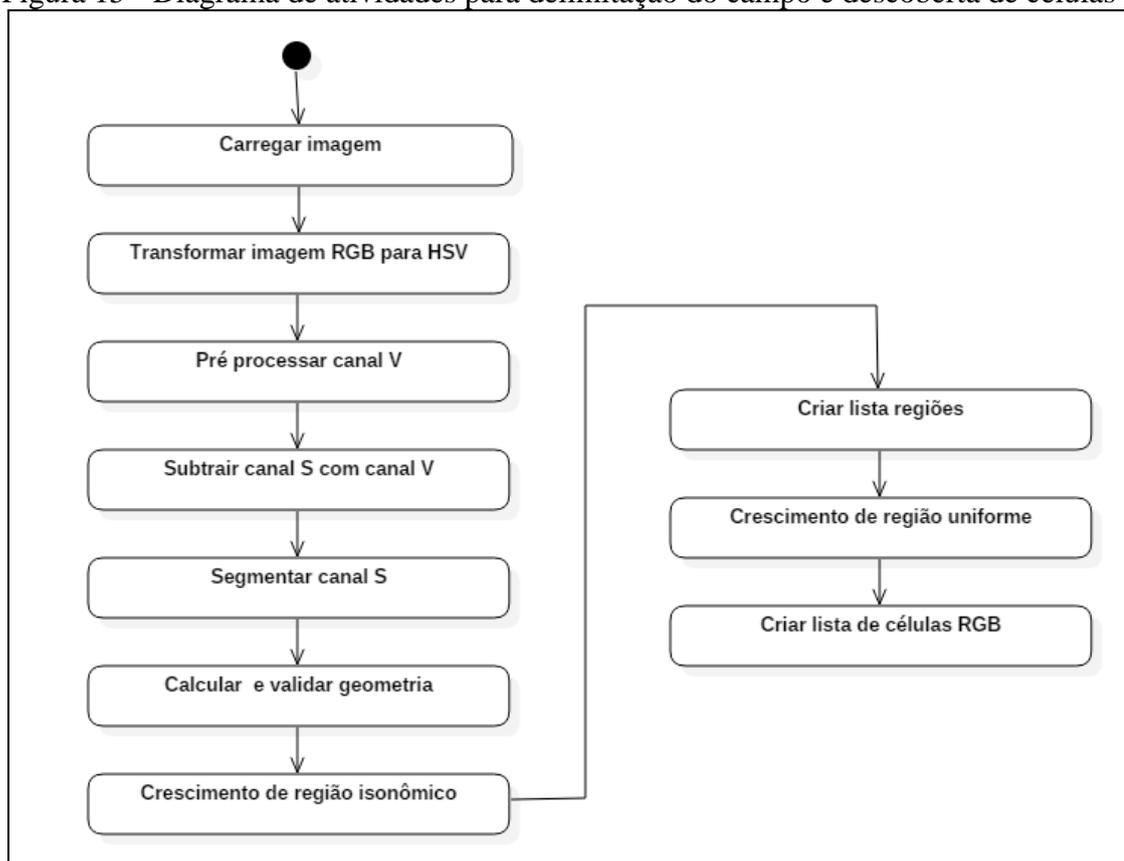
Fonte: elaborado pelo autor.

O fluxo inicia pela delimitação da região do campo, que define a região principal onde será realizado o processamento da imagem para descoberta das células. Esta etapa é descrita na seção 3.2.3.1. A partir da descoberta da região das células, inicia-se a etapa de tratamento e classificação das células, sendo detalhada na seção 3.2.3.2.

3.2.3.1 Delimitação do campo e descoberta de células

Nesta seção é detalhado o processo de delimitação do campo onde a imagem de entrada passa por diversas etapas para descobrir as células. A Figura 15 apresenta o diagrama de atividades com etapas existentes neste processo.

Figura 15 - Diagrama de atividades para delimitação do campo e descoberta de células



Fonte: elaborado pelo autor.

Primeiramente, o usuário seleciona a imagem para ser analisada pelo sistema. As imagens foram capturadas por uma câmera fotográfica digital a partir de um microscópio óptico. O Quadro 3 apresenta a codificação para o carregamento da imagem no sistema.

Quadro 3 - Carregamento da imagem

```

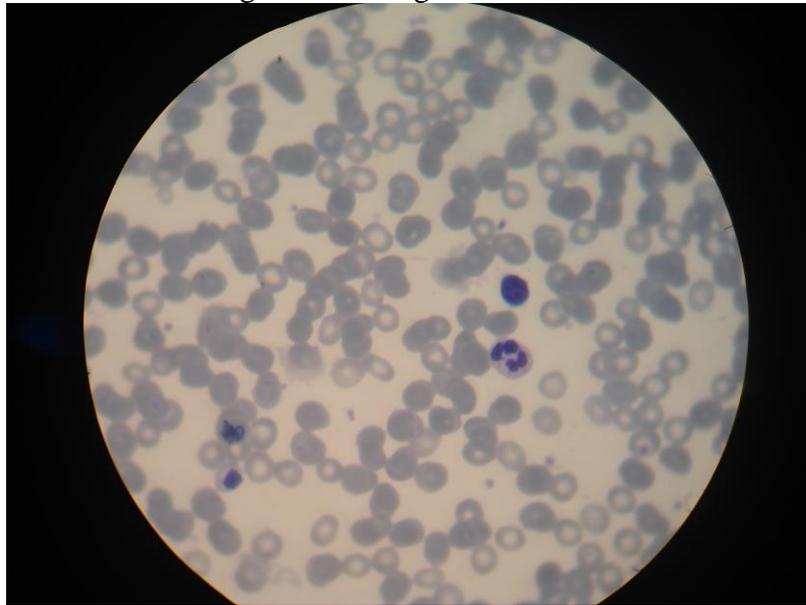
1. public static IplImage loadImage(String path) {
2.     File file = new File(path);
3.     if (file.exists()) {
4.         return cvLoadImage(path, CV_LOAD_IMAGE_COLOR);
5.     }
6.     return null;
7. }

```

Fonte: elaborado pelo autor.

Após o usuário selecionar o caminho da imagem, o método `cvLoadImage` carrega a imagem da lâmina de sangue em memória. O resultado do carregamento pode ser visualizado na Figura 16.

Figura 16 - Imagem de entrada



Fonte: elaborado pelo autor.

A imagem de entrada é base para o início das etapas de pré-processamento, segmentação, validação e classificação as células. A seguir a imagem é convertida para o modelo de cores HSV. O sistema de cores HSV foi escolhido, pois ele obteve sucesso em apresentar realce do núcleo das células entre seus canais. O Quadro 4 apresenta a codificação responsável por realizar a transformação do sistema RGB para HSV.

Quadro 4 - Transformação RGB para HSV

```

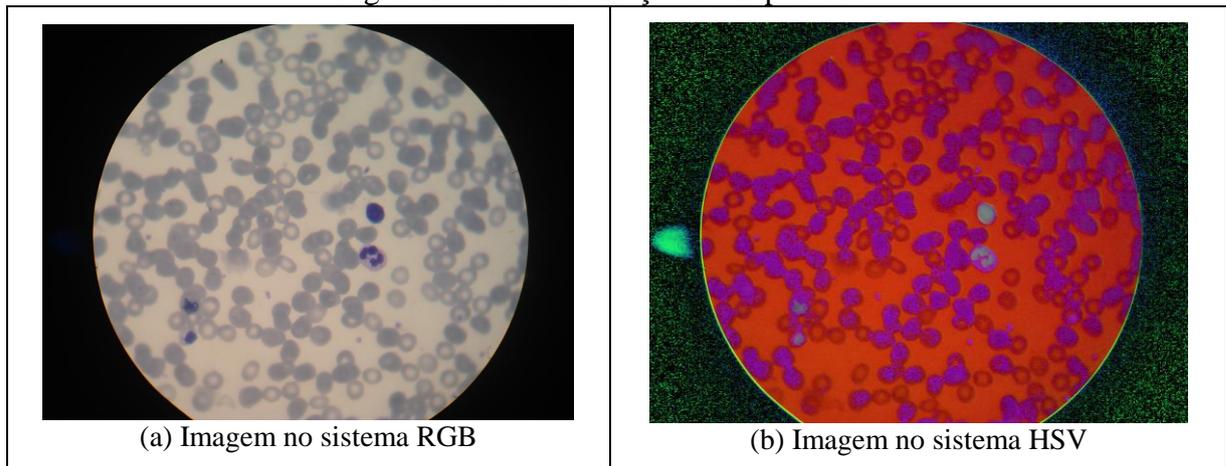
1. IplImage hsv = cvCreateImage(src.cvSize(), IPL_DEPTH_8U, 3);
2. cvCvtColor( src, dst, CV_BGR2HSV );

```

Fonte: elaborado pelo autor.

Na codificação acima é criado uma nova imagem através do método `cvCreateImage`, que utiliza como parâmetro as dimensões da imagem de entrada. A seguir o método `cvCvtColor` realiza a conversão da imagem de entrada em uma nova imagem no padrão de cores HSV através do parâmetro `BGR2HSV`. O resultado deste processamento pode ser visualizado na Figura 17.

Figura 17 - Transformação RGB para HSV



Fonte: elaborado pelo autor.

Para realçar as informações contidas na imagem, ela terá seus canais separados. Os canais do modelo HSV são denominados: canal H (matiz), canal S (saturação) e canal V (valor). O Quadro 5 apresenta a codificação para realizar a separação dos canais.

Quadro 5 - Separação dos canais HSV

```

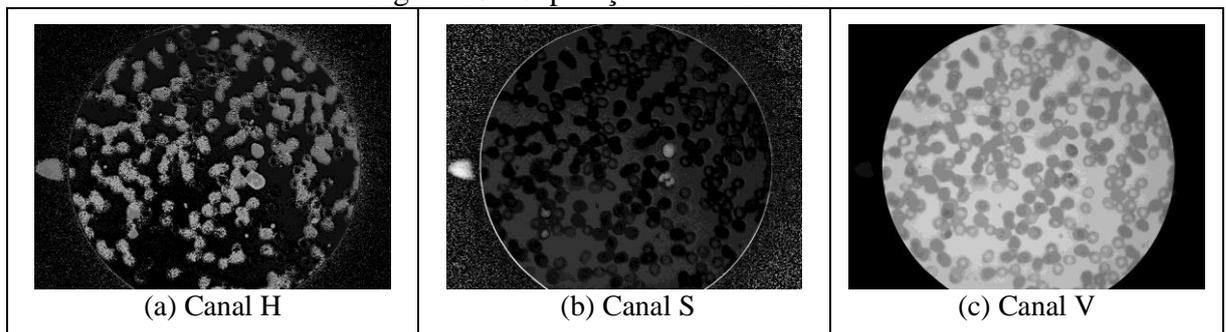
1. IplImage hue = cvCreateImage(hsv.cvSize(), IPL_DEPTH_8U, 1);
2. IplImage sat = cvCreateImage(hsv.cvSize(), IPL_DEPTH_8U, 1);
3. IplImage val = cvCreateImage(hsv.cvSize(), IPL_DEPTH_8U, 1);
4. cvSplit( src, hue, sat, val, null );

```

Fonte: elaborado pelo autor.

No Quadro 5 são criadas três imagens através do método `cvCreateImage`. O método `cvSplit` é quem divide os canais. Para isso, são passados como parâmetro de entrada a imagem HSV e as três imagens criadas, que recebem a informação dos canais: H, S e V. A Figura 18 apresenta o resultado da separação dos 3 canais.

Figura 18 - Separação dos canais HSV



Fonte: elaborado pelo autor.

A partir da Figura 18 pode-se observar que cada canal apresenta uma informação distinta e, que todos possuem alguma informação relevante para realizar a extração do núcleo, citoplasma e da região que contem as células a serem analisadas. A seguir é apresentada a seção que detalha o uso do canal de V.

3.2.3.1.1 Delimitação da região do campo

Após a separação dos três canais, o canal V é utilizado para delimitar a região do campo da imagem, sendo limiarizado para destacar a região de interesse. O Quadro 6 apresenta a codificação que realiza a limiarização da imagem e seleção da maior componente.

Quadro 6 - Busca da maior componente

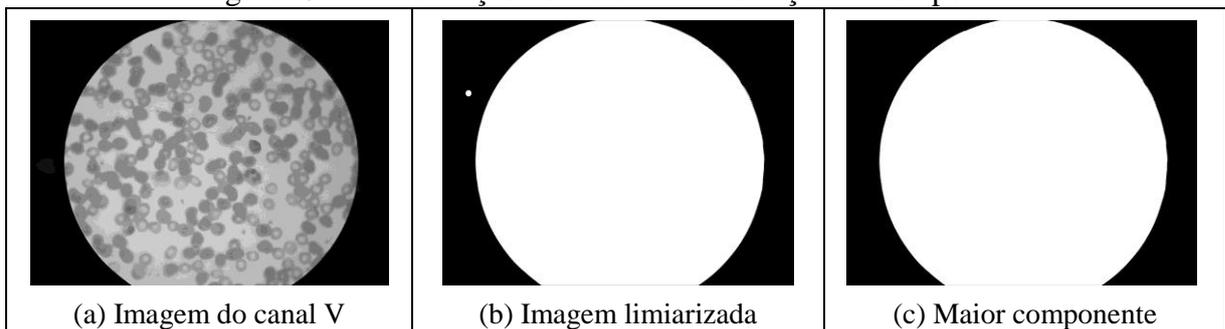
```

1. cvThreshold(src, dst, :param, 255, CV_THRESH_BINARY);
2. cvFindContours(src, storage, contours,
                 Loader(sizeof(CvContour.class),
                       CV_RETR_CCOMP,
                       CV_CHAIN_APPROX_SIMPLE,
                       cvPoint(0,0));
3. for (ptr = contours; ptr != null; ptr = ptr.h_next()) {
4.     if( getArea(ptr) > bigArea){
5.         bigArea = getArea(ptr);
6.         bigCountourPtr = ptr;
7.     }
8. }
```

Fonte: elaborado pelo autor.

No quadro acima, a imagem é processada pelo método `cvThreshold`, onde o parâmetro `:param` tem um valor baixo para que os pixels da região do campo formem uma grande componente. O parâmetro `255` é referente à cor destino, e o parâmetro `CV_THRESH_BINARY` é o tipo de limiarização. Após o processo de limiarização, a região do campo torna-se a maior componente da imagem. Eventualmente, podem existir ruídos na parte externa. Logo, para eliminar estes elementos ruidosos é realizada a busca pela maior componente, presumindo-se que ela represente a região do campo. O resultado deste processamento pode ser visualizado na Figura 19.

Figura 19 - Limiarização do canal de V e seleção da componente



Fonte: elaborado pelo autor.

Após encontrar o maior componente da imagem é criada uma nova imagem com este elemento. Esta imagem será utilizada como máscara para delimitar a região do campo. Ou seja, ela será utilizada no processo de subtração que deixará apenas a região de interesse. Para isso, é necessário inverter a informação de cor binária antes de realizar a subtração com o canal S. O Quadro 7 apresenta a codificação que realiza a inversão da cor binária da imagem.

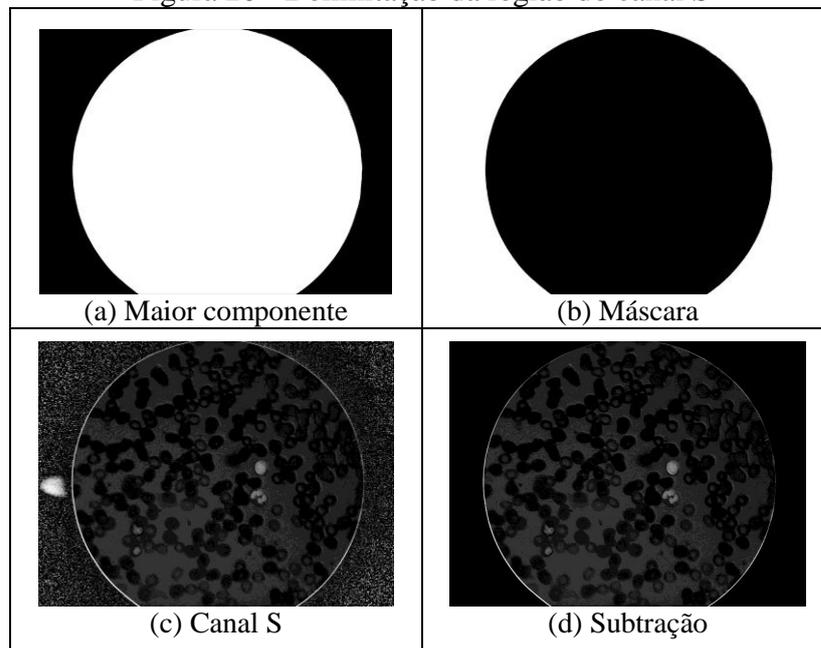
Quadro 7 - Inversão binária e subtração

```
1. cvNot(channel_v, masc);
2. cvSub(channel_s, masc, dst);
```

Fonte: elaborado pelo autor.

No quadro acima, a imagem do canal V (limiarizada) é utilizada como parâmetro para o método `cvNot` que realiza a negação dos valores binários da cor, transformando a imagem destino em uma cópia de cor invertida. O canal S é usado como parâmetro para o método `cvSub` que realiza a subtração com a máscara da região do campo. O resultado deste processamento pode ser visualizado na Figura 20.

Figura 20 - Delimitação da região do canal S



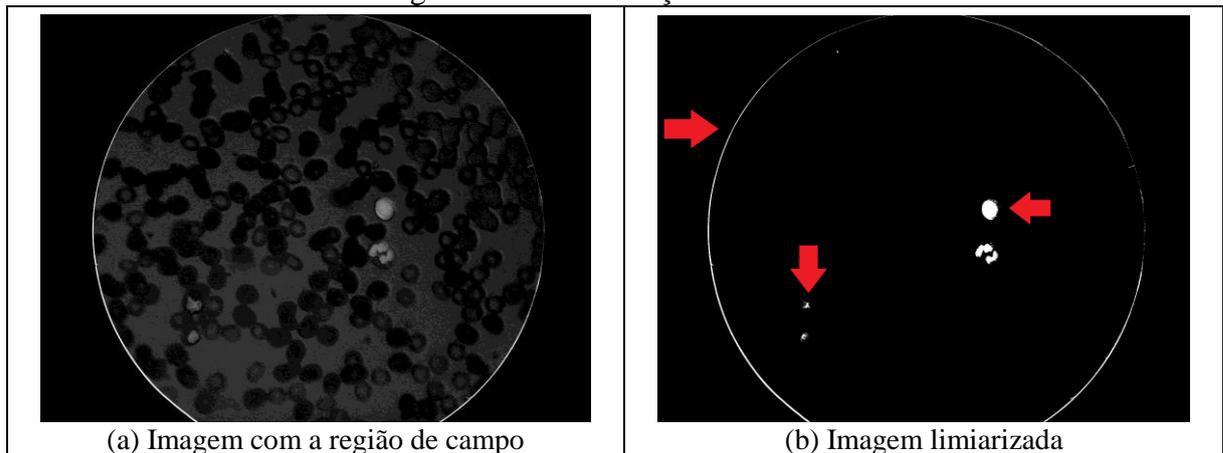
Fonte: elaborado pelo autor.

Após realizar o processo de inversão de cores, a imagem resultante, de forma binária, delimita a região do campo. Essa imagem serve para eliminar os ruídos externos à região do canal S que possui o realce dos núcleos. O canal S é usado como parâmetro para o método `cvSub` que realiza a subtração com a máscara da região do campo, obtendo-se assim o canal S com eliminação total da região externa (Figura 20d).

3.2.3.1.2 Segmentação do canal S

Após realizar o processo de subtração, percebe-se que os núcleos estão em tons mais claros em relação ao restante da imagem. A partir disso, aplica-se uma limiarização para tentar realçá-los. O resultado deste processamento pode ser visualizado na Figura 21.

Figura 21 - Limiarização do canal S



Fonte: elaborado pelo autor.

Após a limiarização, a Figura 21b apresenta as eventuais componentes desta etapa: arco da região do campo, ruídos de diversos tamanhos e o núcleo das células. Neste ponto, são realizadas as operações morfológicas de erosão e dilatação para diminuir os ruídos. O Quadro 8 apresenta a codificação que realiza as operações morfológicas.

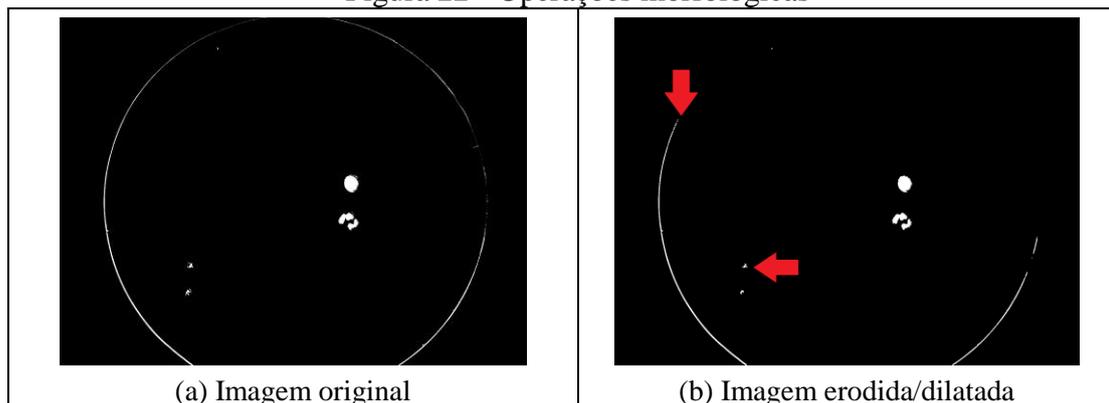
Quadro 8 - Operações morfológicas

```
1. cvErode(src, dst, null, 2);
2. cvDilate(src, dst, null, 2);
```

Fonte: elaborado pelo autor.

Na codificação acima, a imagem passa pelo processo de erosão (`cvErode`) e dilatação (`cvDilate`). Estas operações removem parte dos ruídos e o resultado deste processamento pode ser visualizado na Figura 22.

Figura 22 - Operações morfológicas



Fonte: elaborado pelo autor.

Após realizar as operações morfológicas, entre as componentes da imagem, ainda restam ruídos e, para conseguir distinguir as componentes de interesse dos ruídos, realiza-se uma validação por características geométricas. O Quadro 9 apresenta a codificação que realiza a validação de componentes.

Quadro 9 - Validações geométricas

```

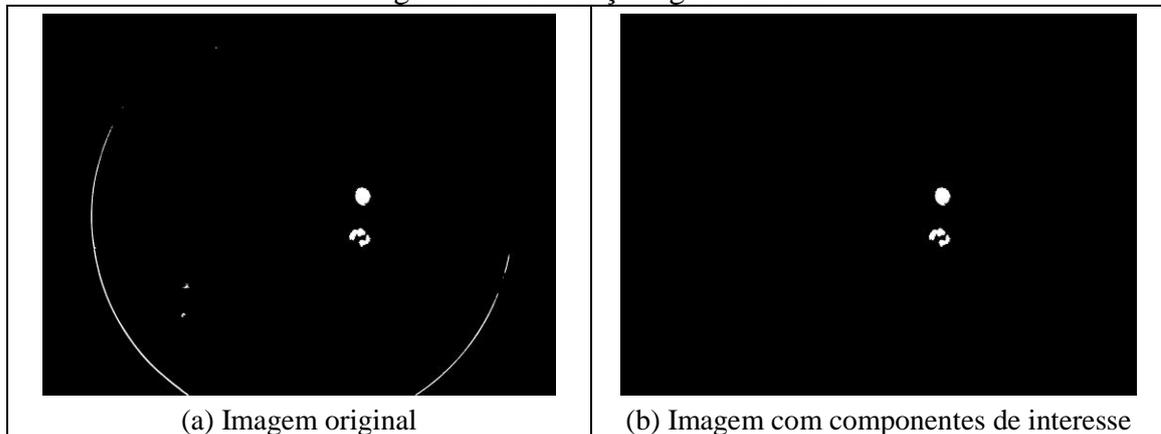
1.cvFindContours(src);
2.for (ptr = contours; ptr != null; ptr = ptr.h_next()) {
3.  if( (getArea(ptr) >= CNT_LEUCOCITY_MIN_AREA) &&
4.      (getArea(ptr) <= CNT_LEUCOCITY_MAX_AREA) &&
5.      (getCircularity(ptr) >= CNT_LEUCOCITY_MIN_CIRCULARITY)) {
6.      cvDrawContours(dst, ptr, CV_FILLED);
7.  }
8.}

```

Fonte: elaborado pelo autor.

Na codificação acima cada componente `ptr` da lista `contours` é usado para calcular área e circularidade do contorno da componente, onde os elementos que forem validados são adicionados em uma nova imagem destino. Ao realizar a etapa de validação, presume-se que só fiquem na imagem as componentes de interesse. O resultado deste processamento pode ser visualizado na Figura 23.

Figura 23 - Validações geométricas



Fonte: elaborado pelo autor.

Após criar a imagem com as componentes de interesse, percebe-se que podem existir casos de núcleos com lóbulos desconexos na mesma célula. Isto normalmente ocorre em neutrófilos. Em linfócitos típicos existe apenas um núcleo condensado. Para tratar núcleos desconexos na mesma célula é realizado o crescimento isonômico de região, onde é verificada a geometria de cada componente. Aquelas que tiveram características de núcleos segmentados passam pelo crescimento de região. No Quadro 10 é apresentada a implementação que realiza o crescimento de região isonômico.

Quadro 10 - Crescimento de região isonômico

```

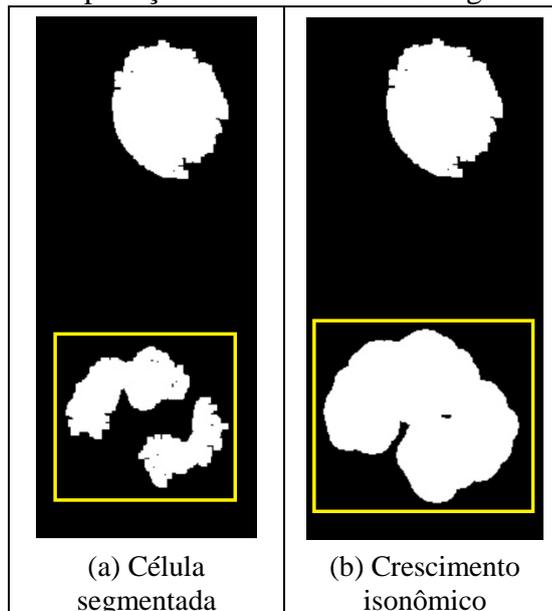
1. cvFindContours (src);
2. for (ptr = contours; ptr != null; ptr = ptr.h_next()) {
3.   cvDrawContours (dst, ptr, CV_FILLED);
3.   if ( ((getCircularity(ptr) <= CNT_CELL_SEGMENTED_MAX_CIRCULARITY) ||
4.         ((getCircularity(ptr) >= CNT_CELL_SEGMENTED_MAX_CIRCULARITY &&
5.           getArea(ptr) < CNT_SMALL_CELL_SEGMENTED_MAX_AREA))) ) {
6.
7.     if (getArea(ptr) <= CNT_SMALL_CELL_SEGMENTED_MAX_AREA) {
8.       cvDrawContours (dst, ptr, PARAM_SIZE_BIG_CONTOUR);
9.     } else {
10.      cvDrawContours (dst, ptr, PARAM_SIZE_SMALL_CONTOUR);
11.    }
12.  }
13. }

```

Fonte: elaborado pelo autor.

No quadro acima, primeiramente, encontra-se os contornos através do método `cvFindContours`. E, na sequência, cada componente da lista é desenhado mantendo sua forma original através do parâmetro `CV_FILLED`. Em seguida, é verificado se a componente atende ao critério de núcleo segmentado. Quando o núcleo é segmentado, é realizado um contorno com espessura maior que aumenta o tamanho da componente. O tamanho da espessura varia, sendo maior para componentes pequenos e menor para componentes maiores. Esse controle é necessário a fim de evitar que núcleos maiores recebam uma espessura muito grande e se conectem a células próximas. A Figura 24 apresenta o crescimento de região isonômico para tratar desses casos, a imagem foi ampliada para melhorar a visualização, pois a imagem resultante mantém o dimensionamento original da imagem entrada.

Figura 24 - Aplicação do crescimento de região isonômico

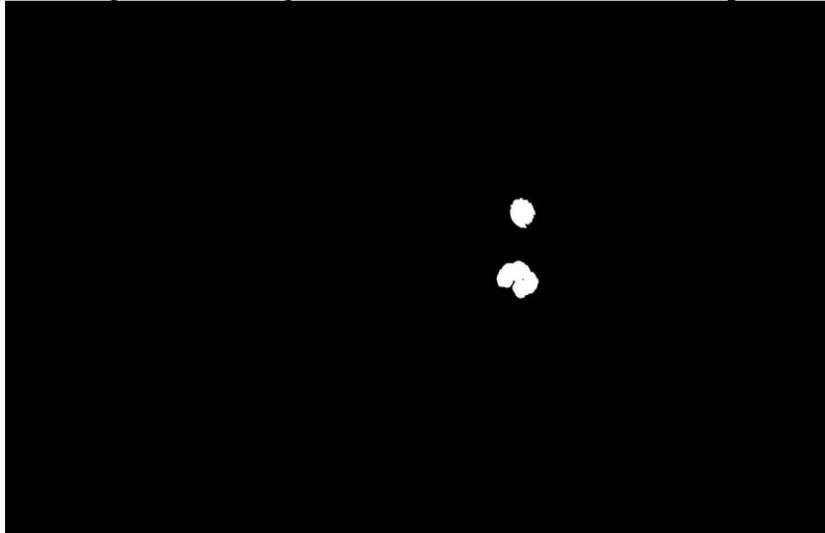


Fonte: elaborado pelo autor.

Nota-se que os núcleos desconexos apresentados na Figura 24a sofreram crescimento de região e agora formam uma nova componente conexa, Figura 24b, enquanto os conexos

continuam do mesmo tamanho. O critério analisado para efetuar o crescimento de região isonômico é baseado nas características geométricas através dos cálculos da área e do grau de circularidade da componente. Depois de realizar o crescimento de região isonômico, tem-se uma imagem binária com os núcleos realçados, conforme mostra a Figura 25.

Figura 25 - Imagem final da fase de realce do campo



Fonte: elaborado pelo autor.

3.2.3.1.3 Isolamento de componentes

Depois de realizar a etapa de crescimento de região, presume-se que a imagem tenha apenas as componentes que representam o núcleo da célula. Logo, é necessário encontrar o restante da célula, no caso, o citoplasma. Tendo conhecimento de que uma característica típica do citoplasma é estar ao redor do núcleo da célula, aplicou-se novamente crescimento de região a partir das componentes encontradas. Para isso, criou-se uma lista de imagens, onde em cada imagem existe apenas uma componente conexa. Isso foi feito, pois o crescimento proposto pode conectar elementos da imagem de entrada. O Quadro 11 apresenta a implementação que realizada a criação da lista de componentes conexos.

Quadro 11 - Criação da lista de componentes conexas

```

1. cvFindContours (src);
2. for (ptr = contours; ptr != null; ptr = ptr.h_next()) {
3.   component = IplImage.create();
4.   cvDrawContours (component, ptr, CV_FILLED);
5.   list.add (component);
6. }

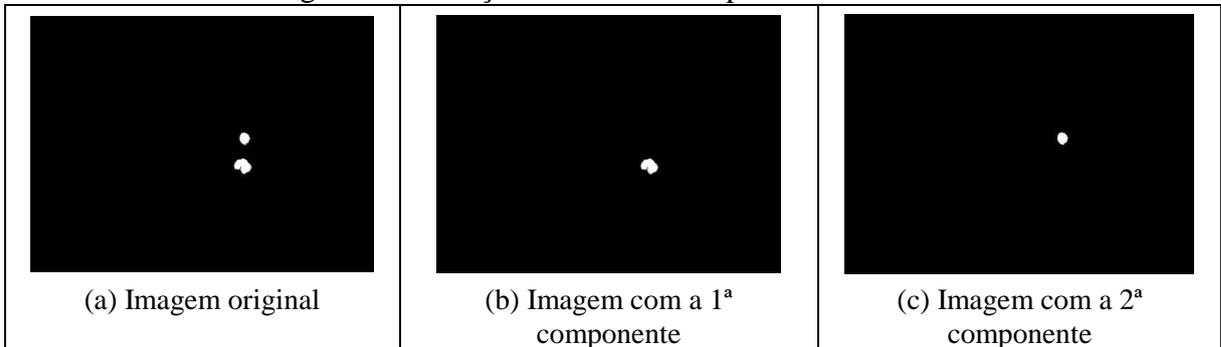
```

Fonte: elaborado pelo autor.

Na implementação acima é utilizado o método `cvFindContours` para descobrir o contorno das componentes que representam as células. Posteriormente, cada componente é desenhada através do método `cvDrawContours`, mantendo todas as características da

componentes através do parâmetro `CV_FILLED`, que preenche o interior do contorno. O resultado da criação da lista é apresentado na Figura 26.

Figura 26 - Criação da lista de componentes conexas



Fonte: elaborado pelo autor.

A Figura 26a demonstra a imagem de entrada e as Figura 26b e Figura 26c as componentes isoladas. Após a criação desta lista, realiza-se o crescimento de região, que neste ponto, trata de maneira igual cada elemento. A expansão é proporcionalmente igual a sua dimensão. O Quadro 12 apresenta como é realizado o crescimento uniforme de região.

Quadro 12 - Crescimento uniforme de região

```

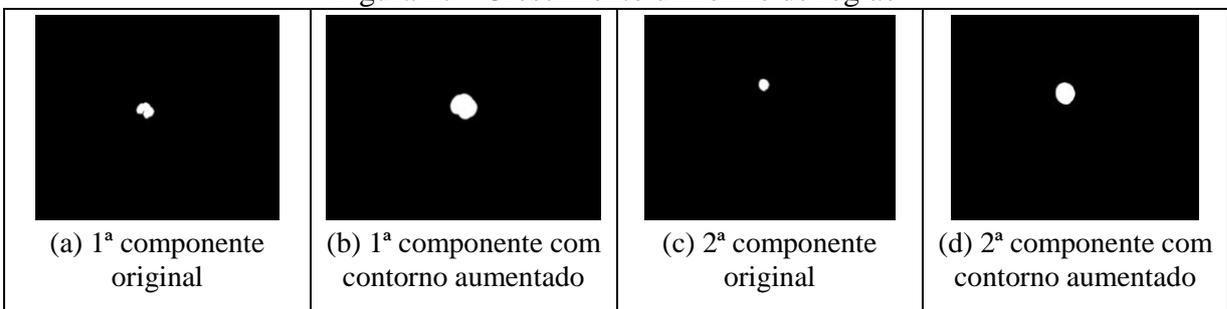
1. for(IplImage src: list){
1.   cvFindContours(src);
3.   for(ptr = contours; ptr != null; ptr = ptr.h_next()) {
4.     component = IplImage.create();
5.     cvDrawContours(component, ptr, CV_FILLED);
6.     cvDrawContours(component, ptr, PRM_SIZE_UNIFORM);
6.     newList.add(component);
7. }

```

Fonte: elaborado pelo autor.

O Quadro 12 mostra que a lista de imagens é percorrida e, que para cada uma delas, encontra-se o contorno a fim de aumentá-lo. Isto é feito através do parâmetro `PRM_SIZE_UNIFORM`. O resultado deste crescimento pode ser observado na Figura 27.

Figura 27 - Crescimento uniforme de região



Fonte: elaborado pelo autor.

Conforme apresentado na Figura 27b e Figura 27d, as componentes tiveram sua região expandida, logo, presume-se que nesta região exista uma célula completa. Essa lista de células binárias é usada como máscara para subtrair a imagem de entrada, gerando assim uma nova

lista de células coloridas no modelo RGB. Neste momento, também é realizado o redimensionamento da imagem para o tamanho de sua *bounding box*. O Quadro 13 apresenta a implementação dessa etapa.

Quadro 13 - Extração da imagem-região

```

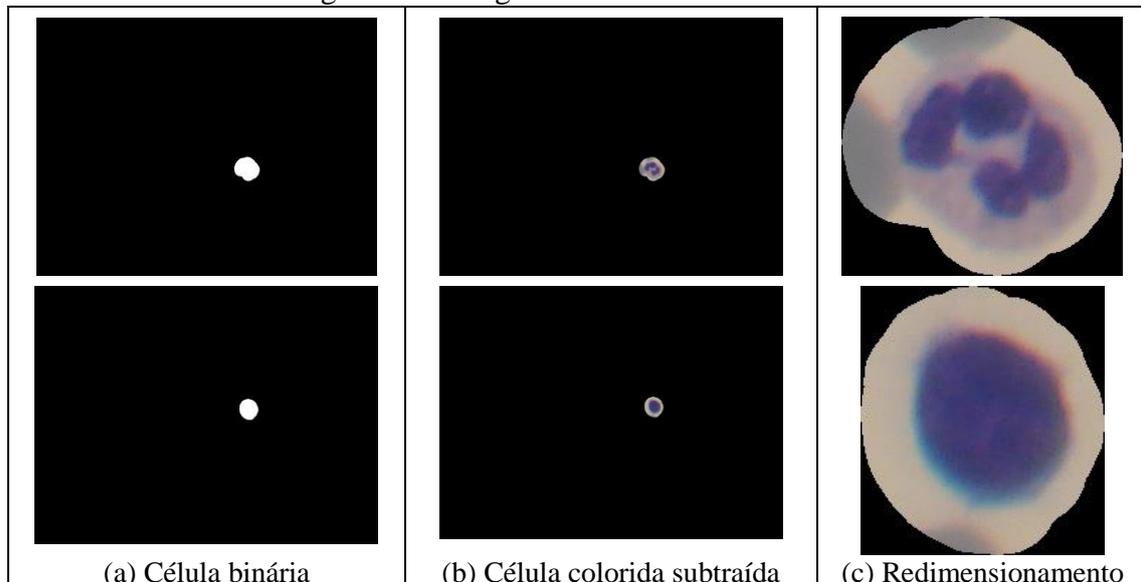
01. for(IplImage src: list){
02.   cvFindContours(src);
03.   for(ptr = contours; ptr != null; ptr = ptr.h_next()) {
04.     src = IplImage.create();
05.     resultImage = IplImage.create();
06.     cvDrawContours(resultImage, ptr, CV_FILLED));
07.     cvNot(resultImage, resultImage);
08.     cvSub( img_in resultImage, src);
09.     CvRect BBox = cvBoundingRect(ptr);
10.     cvSetImageROI(src, BBox);
11.     IplImage resize = IplImage.create();
12.     cvCopy(src, resize);
13.     cvResetImageROI(src);
14.     newList.add(resize.clone());
15. }

```

Fonte: elaborado pelo autor.

A extração de cada célula colorida ocorre através da subtração de cada célula binária. Em seguida é calculado a *bounding box* através do método `cvBoundingRect`. Onde os pontos da *bounding box* são usados para delimitar a região e realizar o redimensionamento através do método `cvSetImageROI`. A nova lista contempla basicamente recortes da imagem de entrada, porém cada região contempla uma célula, conforme mostra a Figura 28.

Figura 28 - Imagem da fase final de descoberta



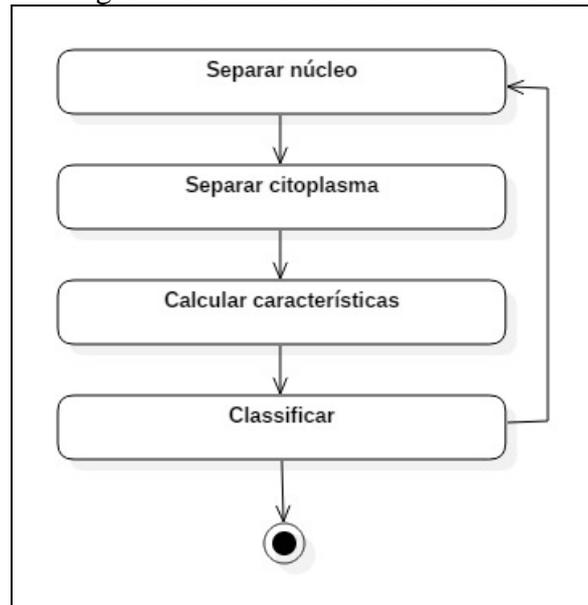
Fonte: elaborado pelo autor.

Por fim, se obtém uma lista de células, Figura 28c, que serve como entrada para a etapa de tratamento.

3.2.3.2 Tratamento de células

Nesta seção são detalhadas as etapas do tratamento das células descobertas, que são: separação do núcleo, separação do citoplasma, a extração de características e o uso das informações adquiridas para realizar a classificação através da rede neural do tipo MLP. A Figura 29 apresenta o diagrama de atividades com etapas existentes neste processo.

Figura 29 - Diagrama de atividades do tratamento de células



Fonte: elaborado pelo autor.

Nas próximas seções são detalhados os procedimentos realizados em cada uma das etapas que compõem o tratamento de células.

3.2.3.2.1 Separação do Núcleo

A separação do núcleo começa pela segmentação da lista de imagem-região através da clusterização. Isso faz com que as cores do núcleo se tornem mais homogêneas e por consequência tem-se uma delimitação melhor para a extração do núcleo. Ao qual, possui tons que se sobressaem do restante da imagem. A seguir o Quadro 14 apresenta a codificação para realizar a clusterização.

Quadro 14 - Algoritmo Mean-Shift

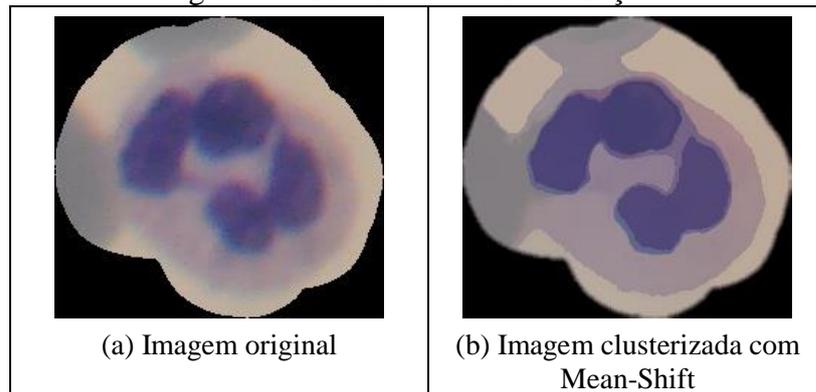
```

01. for(IplImage src: list){
02.   cvPyrMeanShiftFiltering(src, src, 20, 20);
03. }
  
```

Fonte: elaborado pelo autor.

O método de clusterização, `cvPyrMeanShiftFiltering`, aplica o algoritmo Mean-Shift utilizando uma janela de tamanho 20 e um raio de cor com tamanho 20. O resultado pode ser visualizado na Figura 30.

Figura 30 - Resultado da clusterização



Fonte: elaborado pelo autor.

A partir do resultado da clusterização, percebe-se que a região do núcleo fica homogeneamente delimitada. Em seguida é realizado o mesmo processo executado na etapa de segmentação do canal S. O Quadro 15 apresenta o código responsável pelo isolamento do núcleo.

Quadro 15 - Isolamento do núcleo

```

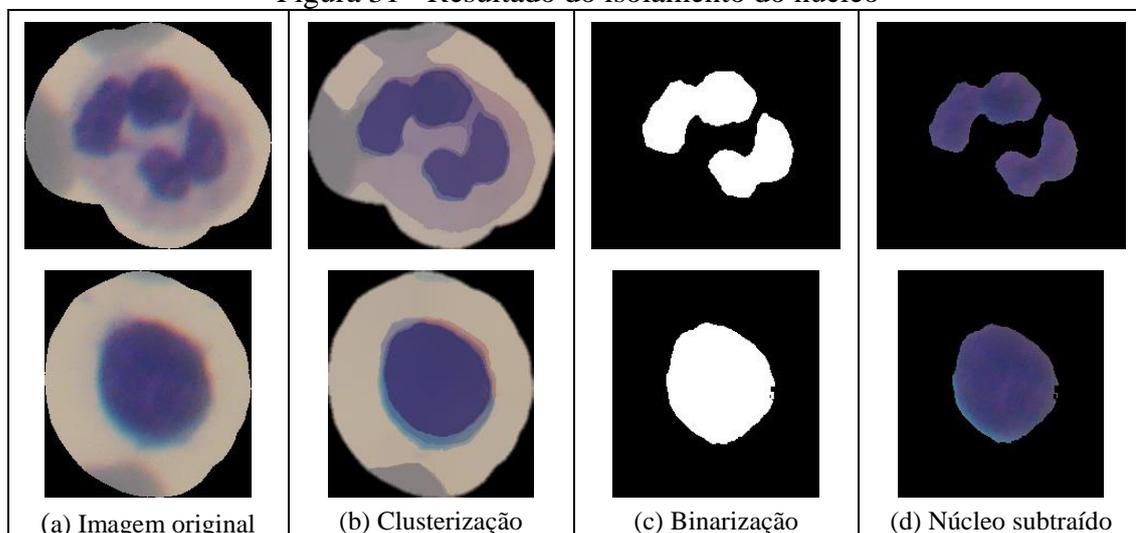
01. IplImage s = IplImage.create( src.cvSize(), src.depth(), CV_8UC1 );
02. IplImage dst = IplImage.create();
03. cvCvtColor(src, dst, CV_BGR2HSV);
04. cvSplit(dst, null, s, null, null);
05. cvThreshold(s, dst, PARAM_THRESHOLD_NUCLEO);
06. IplImage msk_nucl = IplImage.create( src.cvSize(), src.depth(), 3 );
07. cvCvtColor(dst, msk_nucl, CV_GRAY2BGR);
08. cvNot(msk_nucl, msk_nucl);
09. dst = IplImage.create( src.cvSize(), src.depth(), 3 );
10. cvSub(img_in, msk_nucl, dst);

```

Fonte: elaborado pelo autor.

A implementação utiliza a mesma lógica apresentada na seção 3.2.3.1.2, porém a imagem primeiramente é clusterizada. O resultado obtido é isolamento do núcleo da célula pode ser visto na Figura 31.

Figura 31 - Resultado do isolamento do núcleo

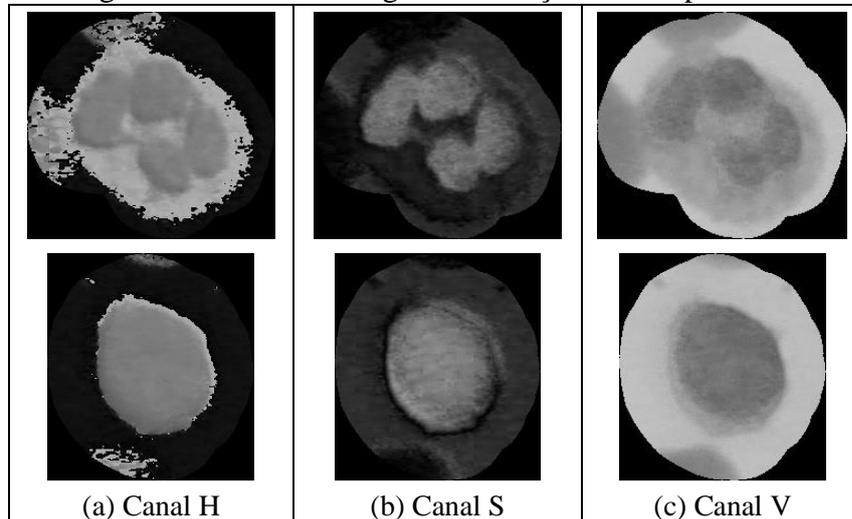


Fonte: elaborado pelo autor.

3.2.3.2.2 Separação do citoplasma

A separação do citoplasma começa pelo pré-processamento da imagem para o sistema de cores HSV, conforme mostra a Figura 32.

Figura 32 - Lista de imagens da seleção de componentes



Fonte: elaborado pelo autor.

A partir da transformação para o modelo HSV, percebe-se que o canal H realça melhor a célula em relação aos outros canais. Diante disso, o canal H é submetido a operações morfológicas e seguido pela limirização e validações geométricas. O Quadro 16 apresenta o código desta etapa.

Quadro 16 - Pré-processamento a partir do canal M

```

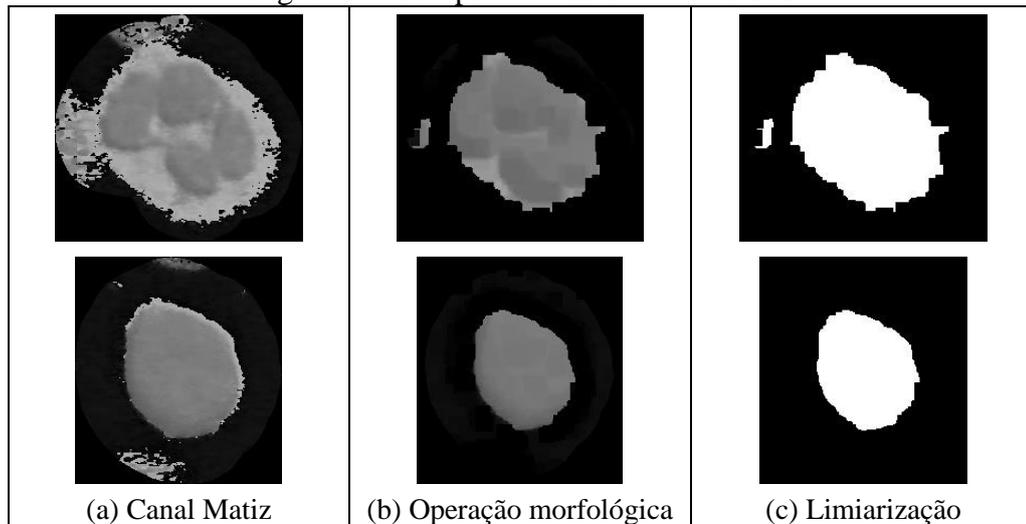
1. cvErode(h,h,null,1);
2. cvDilate(h,h,null,1);
3. cvErode(h,h,null,2);
4. cvDilate(h,h,null,1);
5. cvErode(h,h,null,2);
6. cvDilate(h,h,null,2);
6. cvErode(h,h,null,3);
6. cvErode(h,h,null,3);
7. dst = IplImage.create( src.cvSize(), src.depth(), 1 );
8. cvThreshold(h, dst, PARAM_THRESHOLD_CYTOPLASM, CNT_COLOR_WHITE,
CV_THRESH_BINARY);

```

Fonte: elaborado pelo autor.

Primeiramente, aplicou-se o filtro morfológico de erosão para eliminar pequenos ruídos, seguido por uma sequência morfológica e, novamente por mais duas erosões com elementos estruturantes maiores. Este processo tem como objetivo de desmembrar hemácias conectadas à célula. O resultado pode ser visualizado na Figura 33.

Figura 33 - Pré-processamento do canal H



Fonte: elaborado pelo autor.

A partir da Figura 33, pode-se perceber que algumas regiões indesejadas ainda estão presentes na imagem. Para eliminá-las, optou-se por verificar e deixar a maior componente conexa. Depois disto, realizou-se uma operação de dilatação para recuperar possíveis regiões do citoplasma entorno da componente, conforme mostra o Quadro 17.

Quadro 17 - Seleção da componente e reconstrução da célula

```
1. dst = findBigComponent(dst);
2. cvDilate(dst, dst, null, 3);
3. cvDilate(dst, dst, null, 3);
3. cvDilate(dst, dst, null, 1);
```

Fonte: elaborado pelo autor.

Ao ter as células isoladas em grandes componentes. Basta realizar uma operação de subtração da imagem origem, seguida por uma subtração do núcleo. Dessa forma, obtém-se a região do citoplasma. O Quadro 18 demonstra a implementação responsável por esta etapa.

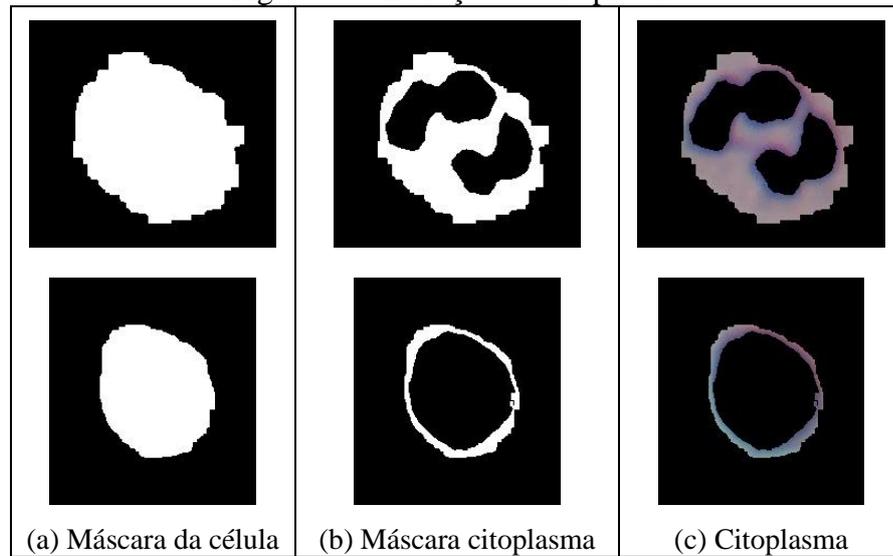
Quadro 18 - Subtração do citoplasma

```
1. cvCvtColor(dst, msk_cito, CV_GRAY2BGR);
2. dst = IplImage.create( src.cvSize(), src.depth(), 3 );
3. cvNot(msk_nucl, msk_nucl);
4. cvNot(msk_cito, msk_cito);
5. cvAdd(msk_nucl, msk_cito, msk_cito);
```

Fonte: elaborado pelo autor.

Primeiramente é criada uma imagem com 3 canais, em seguida é executado o método `cvCvtColor` que transforma a imagem com 1 para 3 canais de cor. Isto é necessário para realizar a operação de subtração. Após a conversão, aplica-se o método `cvNot` para inverter a imagem binária. Em seguida, realiza-se a subtração do núcleo. Dessa forma, tem-se a região que representa o citoplasma, conforme mostra a Figura 34.

Figura 34 - Extração do citoplasma

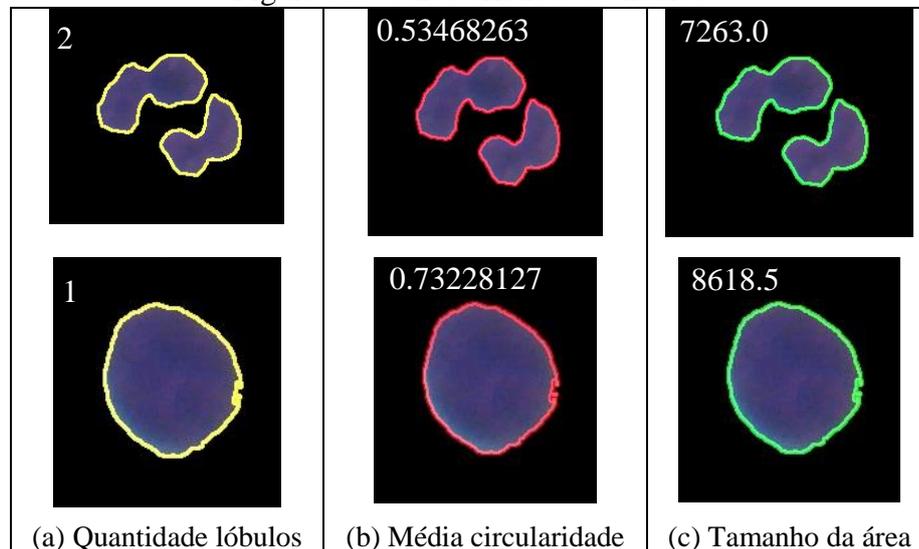


Fonte: elaborado pelo autor.

3.2.3.2.3 Classificação

Por fim, a etapa de classificação utiliza as características que foram extraídas do núcleo e citoplasma da célula. Entre as características extraídas do núcleo estão: quantidade lóbulos, média de circularidade do núcleo, tamanho da área do núcleo. A seguir a Figura 35 apresenta as características extraídas do núcleo.

Figura 35 - Características do núcleo

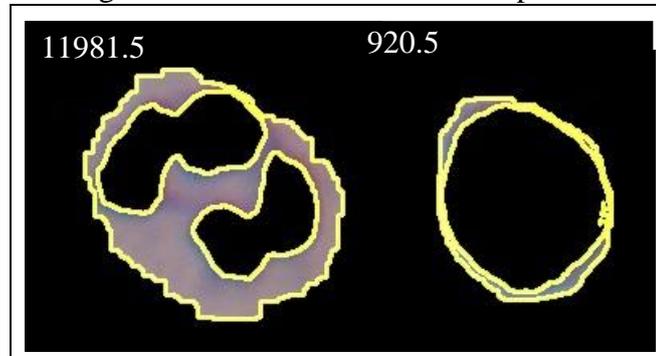


Fonte: elaborado pelo autor.

Como pode ser observado na Figura 35a, a quantidade de lóbulos é um critério importante na classificação, embora possam existir neutrófilos onde os lóbulos fiquem próximos fazendo com que computacionalmente seja identificada uma única componente. Outra característica importante é a média de circularidade do núcleo ao qual tende a ser mais alta em linfócitos e menor em neutrófilos. Por último, o tamanho da área do núcleo é usado

para avaliar sua regularidade, no entanto, não diferencia diretamente um linfócito de um neutrófilo. Já o citoplasma tem como única característica relevante neste momento, o tamanho de sua área como pode ser visto na Figura 36.

Figura 36 - Características do Citoplasma



Fonte: elaborado pelo autor.

As características extraídas são utilizadas na rede neural. O no Quadro 19 apresenta a implementação do carregamento e treinamento da rede neural.

Quadro 19 - Treinamento da rede neural

```

01. File cellsFile = loadData();
02. VersatileDataSource source = new CSVDataSource(cellsFile,
        false, CSVFormat.DECIMAL_POINT);
03. data = new VersatileMLDataSet(source);
04. data.defineSourceColumn("nucleo-amount", 0);
05. data.defineSourceColumn("nucleo-circularity", 1);
07. data.defineSourceColumn("nucleo-length", 2);
08. data.defineSourceColumn("citoplasm-length", 3);
09. ColumnDefinition outputColumn = data.defineSourceColumn("type", 4);
10. data.analyze();
11. data.defineSingleOutputOthersInput(outputColumn);
12. EncogModel model = new EncogModel(data);
13. model.selectMethod(data, MLMMethodFactory.TYPE_FEEDFORWARD);
14. data.normalize();
15. model.holdBackValidation(0.3, true, 1001);
16. model.selectTrainingType(data);
17. model.getTrainingDataset();
18. model.getValidationDataset();

```

Fonte: elaborado pelo autor.

No quadro acima, o método `loadData` retorna o arquivo que serve como base de dados para o treinamento da rede. As colunas dos dados de entrada são definidos através do método `defineSourceColumn`, enquanto que a coluna com a saída é definido por `defineSingleOutputOthersInput(outputColumn)`. O treinamento é realizado através do método `model.getTrainingDataset()`. O Quadro 20 apresenta a utilização da rede neural.

Quadro 20 - Utilização da rede neural

```

01. public String sort(Cell cell) {
02.     NormalizationHelper helper = data.getNormHelper();
03.     MLData input = helper.allocateInputVector();
04.     helper.normalizeInputVector(cell, input.getData(), false);
05.     MLData output = bestMethod.compute(input);
06.     String leucocityType =
07.         helper.denormalizeOutputVectorToString(output)[0];
08.     return leucocityType;
09. }

```

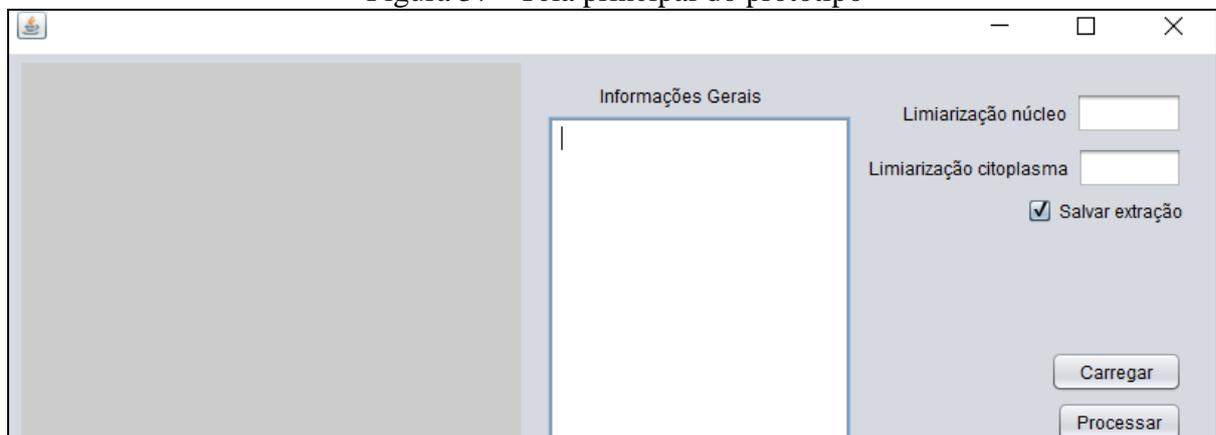
Fonte: elaborado pelo autor.

Através do método `sort` é passado como parâmetro `cell` contendo as características da célula. Logo, o método `bestMethod.compute(input)` é o responsável por todo processamento para inferir a classificação da célula.

3.2.4 Operacionalidade da implementação

O protótipo possui um painel no canto esquerdo para visualização da imagem escolhida. Ao lado direito do painel, tem-se o campo de informações gerais para apresentar o resultado do processamento. As parametrizações estão localizadas do canto direito, para definição de limiares e geração de imagens além de um botão para selecionar a imagem de entrada e um botão para iniciar o processamento. A Figura 37 mostra a tela principal do protótipo, onde pode-se visualizar os itens descritos.

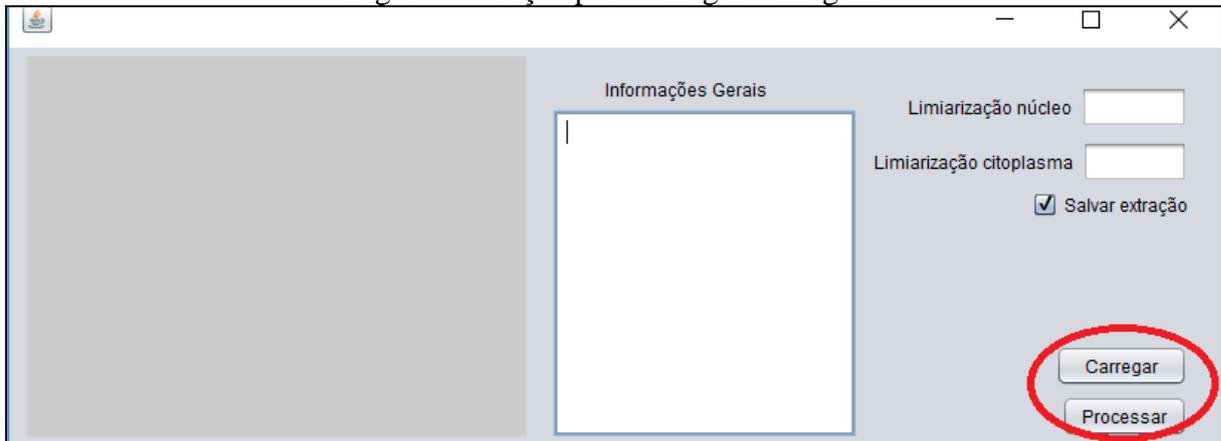
Figura 37 - Tela principal do protótipo



Fonte: elaborado pelo autor.

Para carregar uma imagem, o usuário deve acessar o botão `Carregar` e, para realizar o processamento, clicar no botão `Processar`, conforme mostra a Figura 38.

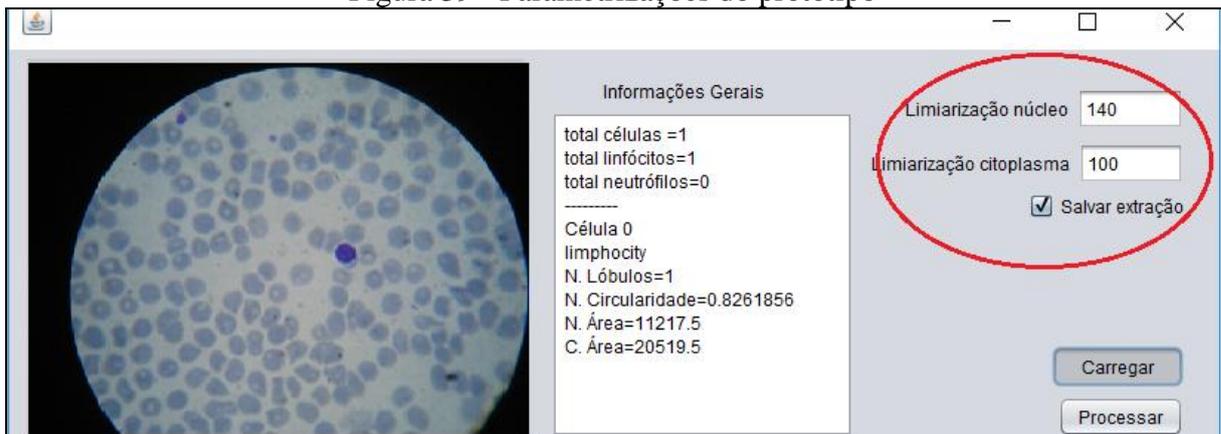
Figura 38 - Ação para carregar a imagem



Fonte: elaborado pelo autor.

Para realizar a parametrização do protótipo, foi disponibilizado um campo Limiarização do núcleo e um campo Limiarização do citoplasma. Além disso, existe uma opção para salvar a extração das células em arquivo, conforme mostra a Figura 39.

Figura 39 - Parametrizações do protótipo



Fonte: elaborado pelo autor.

O campo central *Informações gerais* é responsável pela saída das informações apuradas pelo protótipo. Nele é apresentado a quantidade de células encontradas, a contagem diferencial e um detalhamento individual de cada célula descoberta na imagem.

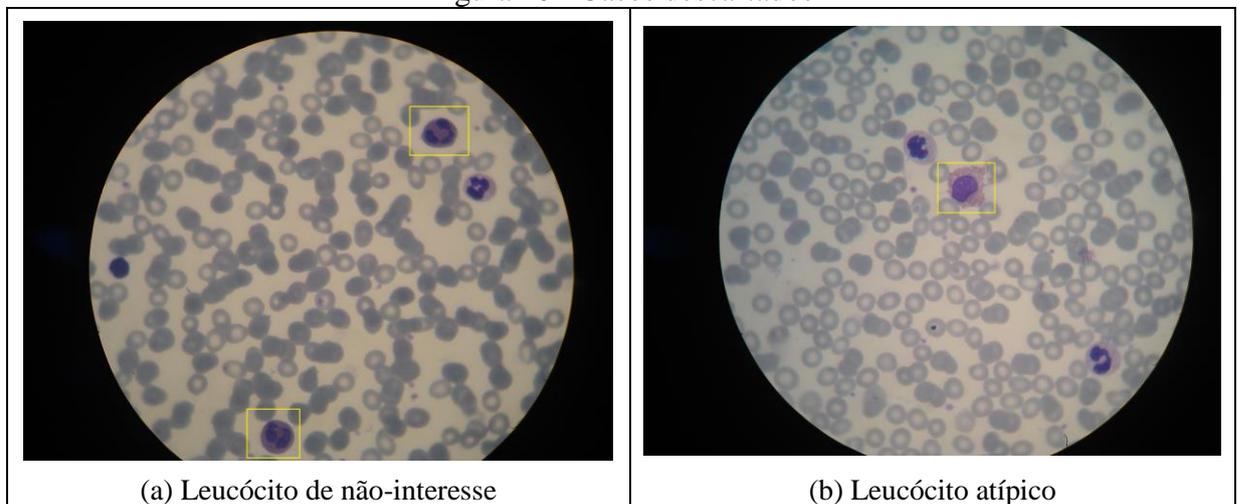
3.3 RESULTADOS E DISCUSSÕES

Esse capítulo apresenta os resultados obtidos na execução do protótipo. A seção 3.3.1 mostra o processo de montagem da base de testes. A seção 3.3.2 demonstra os resultados obtidos na etapa de Delimitação do campo e descoberta de células. Na seção 3.3.3 têm-se os resultados referente extração do núcleo. A seção 3.3.4 apresenta os resultados da extração do citoplasma. Por fim, a seção 3.3.5 demonstra os resultados obtidos na etapa de classificação.

3.3.1 Montagem da base de teste

As imagens utilizadas foram capturadas a partir de um ambiente controlado. Ao qual, foram adquiridas através de uma câmera digital sobre o microscópio óptico, que possui controle de luz e foco e, com auxílio de um profissional biomédico. A maior dificuldade encontrada foi na definição do foco para conseguir uma fotografia do campo da lâmina de boa qualidade. Sendo que das 300 imagens capturadas, 200 foram descartadas por falta de foco, e das 100 restantes, foram selecionadas 40 imagens que tinham os tipos de leucócitos desejados: linfócito e neutrófilo. Na Figura 40 são apresentados alguns exemplos de imagens que foram descartadas.

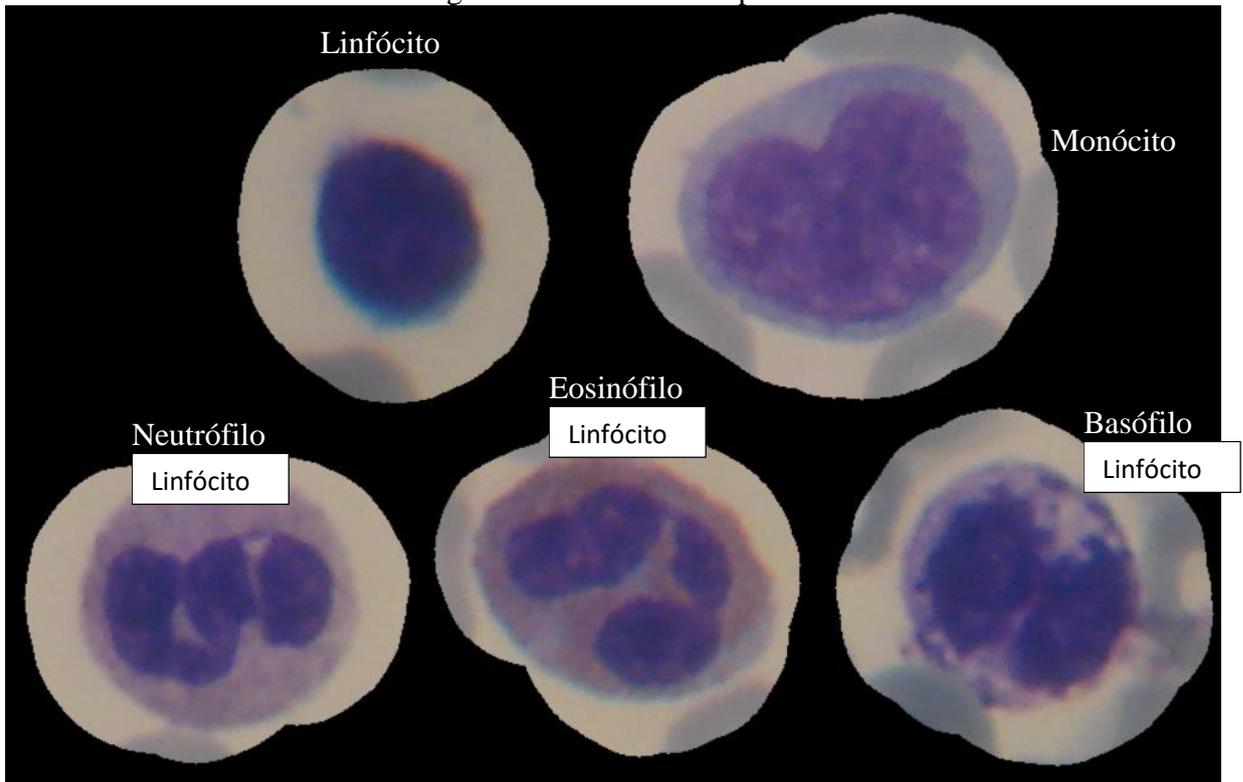
Figura 40 - Casos descartados



Fonte: elaborado pelo autor.

No esfregaço sanguíneo da Figura 40a, em destaque o tipo eosinófilo, além monócitos e basófilos que não são classificados neste trabalho. A Figura 40b apresenta um caso atípico de leucócito, que também foram descartados. Para melhor entendimento dos tipos encontrados na fase de aquisição, a Figura 41 apresenta um exemplar de cada tipo leucócito típico capturado na fase de aquisição, sendo que os tipos desprezados são: basófilo, eosinófilo e monócito. Ressalta-se que os dois leucócitos da parte superior são familiares agranulados e os três da parte inferior são familiares granulados.

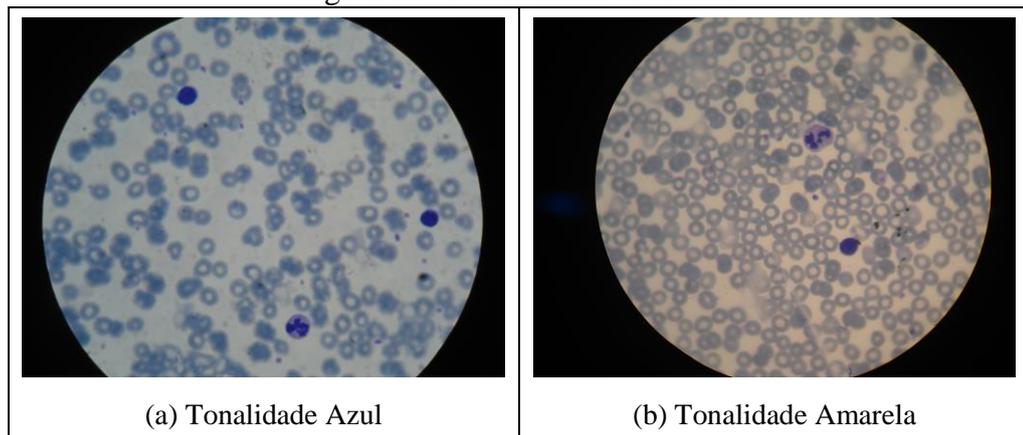
Figura 41 - Leucócitos típicos



Fonte: elaborado pelo autor.

Ao realizar os primeiros testes, percebeu-se que a cor da lâmpada influenciava nos resultados. Conforme pode ser visualizado na Figura 42, a cor da lâmpada define uma tonalidade macro sobre as imagens e, dependendo da cor, é necessário ajustar os parâmetros de execução.

Figura 42 - Tonalidades na lâmina



Fonte: elaborado pelo autor.

Embora possam existir outros fatores que podem influenciar na cor predominante da lâmina, a cor da luz do microscópio foi a mais relevante. É possível controlar a intensidade de luminosidade de modo que o material de estudo fique com luz e foco adequados, tornando-os mais perceptíveis. As imagens adquiridas foram de equipamentos com uma objetiva de

imersão em 100x. Devido aos inúmeros equipamentos no mercado, cada um possui suas próprias características e, o profissional utiliza o que considera ser o mais adequado.

3.3.2 Delimitação do campo e descoberta de células

O primeiro teste teve como objetivo verificar a eficiência do protótipo em delimitar o campo e descobrir as células. A Tabela 1 descreve os resultados obtidos para descobrir as células. Nela, tem-se a quantidade de imagens utilizadas, a quantidade de células reais e a quantidade de células encontradas pelo protótipo, quantidade de imagens com resultados inconsistentes e seu percentual de acurácia.

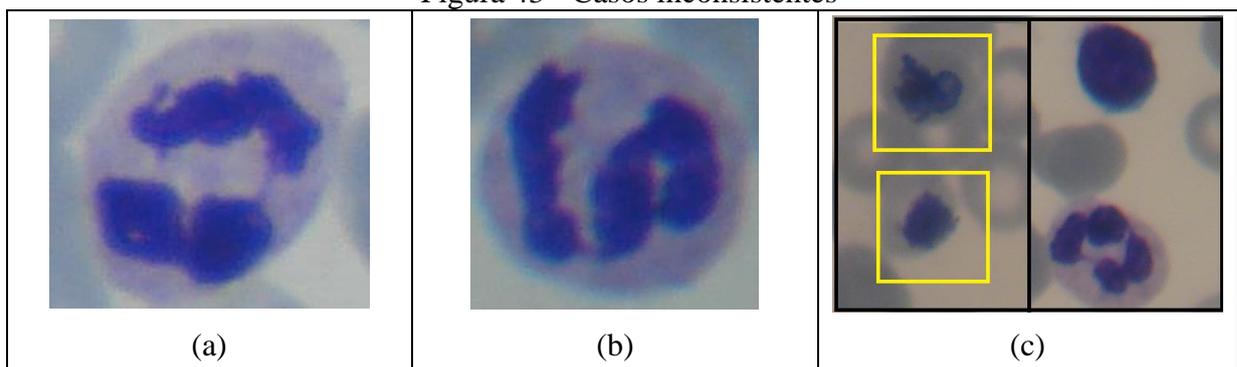
Tabela 1 - Resultados da descoberta de células

Quantidade de imagens	Quantidade de células reais	Quantidade de células encontradas	Quantidade de imagens com resultado inconsistente	Percentual de imagens com resultado consistente
40	89	94	6	94,7%

Fonte: elaborado pelo autor.

A partir da Tabela 1, observa-se que seis imagens resultaram em uma análise inconsistente, que quantificadas representam sete células encontradas erroneamente (Figura 43a e Figura 43b). Isso aconteceu porque dois neutrófilos duplicaram-se, devido à etapa de crescimento de região isonômico não ter conseguido conectar os lóbulos destes núcleos. Já as outras três componentes, Figura 43c, o protótipo considerou como tipo neutrófilo em função de sua morfologia e, outros dois neutrófilos, foram classificados como linfócitos.

Figura 43 - Casos inconsistentes



Fonte: elaborado pelo autor.

3.3.3 Extração do núcleo

O segundo teste teve por objetivo a extração do núcleo da célula. Conforme pode ser observado na Tabela 2, toda célula teve a região do núcleo extraída com ressalvas para cinco casos.

Tabela 2 - Resultados da extração de núcleos

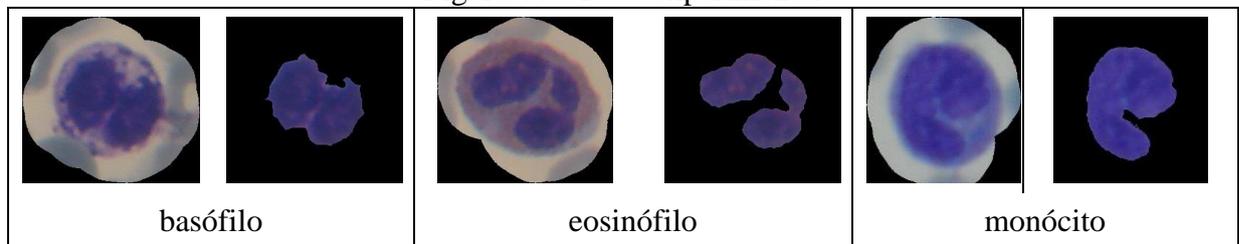
Quantidade de imagens	Quantidade de células reais	Quantidade núcleos encontrados
40	89	94

Fonte: elaborado pelo autor.

Como pode ser observado na tabela acima, das 89 células reais, o protótipo encontrou 94 células. Onde, houve 3 casos que foram considerados erroneamente como possíveis células. Já os outros 2 casos são neutrófilos desconexos, sendo assim, para avaliação de extração de núcleo são considerados 91 casos válidos. Logo, se obtém um percentual de 96,8% de acurácia na extração e validação dos núcleos.

Em um teste realizado para verificar se o protótipo conseguiria extrair núcleos de outros tipos de leucócitos, obteve-se sucesso conforme pode ser visualizado na Figura 44.

Figura 44 - Teste experimental



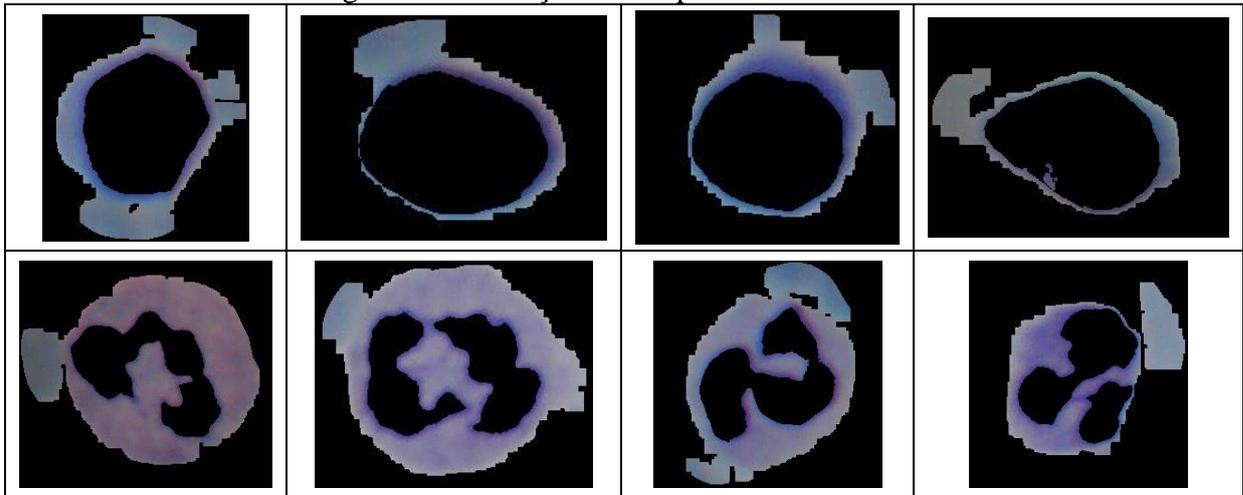
Fonte: elaborado pelo autor.

O monócito pertence à família dos agranulados e possui características morfológicas similares aos linfócitos, porém o núcleo é maior. Pode ter uma circularidade menor e possui a região do citoplasma maior. Já os tipos basófilos e eosinófilos pertencem à família dos granulados (neutrófilos). O basófilo tem o tamanho similar ao linfócito, porém entre suas principais características, estão os grânulos azul-roxo presentes no citoplasma. Já o eosinófilo é similar ao neutrófilo, porém este possui grânulos avermelhados na região do citoplasma.

3.3.4 Extração do citoplasma

A principal motivação para extrair o citoplasma é a utilização da sua área para servir como parâmetro para a rede neural inferir a classificação da célula. A Figura 45 apresenta os casos que a segmentação não conseguiu eliminar todos os ruídos.

Figura 45 - Extração de citoplasmas com ruídos



Fonte: elaborado pelo autor.

A partir da Figura 45, observa-se que em todos os casos, o ruído é parte de uma hemácia que esta conexas ao leucócito no esfregaço sanguíneo. Em virtude disto, foi verificada a quantidade citoplasmas extraídos que apresentam ruídos similares aos da Figura 45. A seguir, a Tabela 3 apresenta os casos irregulares.

Tabela 3 - Resultados da extração de citoplasmas

Quantidade de células reais	Quantidade extrações com ruído	Percentual extrações ideais	Percentual extrações com ruído relevante
89	23	79,53%	20,47%

Fonte: elaborado pelo autor.

A Tabela 3 demonstra que cerca de 20% das extrações de citoplasma possuem ruídos médio-grande. Isto pode prejudicar a etapa de classificação, pois tipicamente linfócitos possuem um citoplasma pequeno.

3.3.5 Classificação

No último teste, é verificado a acurácia do protótipo em classificar as células. Para isso, um profissional biomédico classificou as 40 imagens em 89 células. Depois disso, foi realizado o treinamento e validação da rede neural do tipo MLP. A rede neural utilizada é do tipo *Perceptron* e foi treinada com 8 linfócitos e 15 neutrófilos. A Tabela 4 apresenta os resultados de maneira sintética.

Tabela 4 - Resultado sintético da classificação

Profissional		Protótipo		
Neutrófilos	Linfócitos	Neutrófilos	Linfócitos	Classificação trocada
51	38	54	40	2

Fonte: elaborado pelo autor.

A Tabela 4 demonstra que a rede neural conseguiu inferir a classificação da maioria dos casos, sendo que erros ocorrem apenas em neutrófilos. Estes erros são provenientes de

dois neutrófilos classificados como linfócito, dois neutrófilos com núcleo desconexo que acabaram duplicando e três ruídos classificados como neutrófilos. A seguir a Tabela 5 apresenta de forma analítica a classificação cruzada.

Tabela 5 - Resultado analítico da classificação

Amostra	Profissional		Protótipo			% acerto
	Neutrófilo	Linfócito	Neutrófilo	Linfócito	Inconsistência	
001	2		2			100%
002	5		6		1-ND	83,3%
003		1		1		100%
004	1	1	2	1	1-ND	66,7%
005	1	2	1	2		100%
006	1	2	2	2	1-IN	75%
007	1	2	1	2		100%
008	3	1	3	1		100%
009	3	1	3	1		100%
010	1	1	1	1		100%
011		1		1		100%
012	1		1			100%
013		1		1		100%
014		2		2		100%
015		1		1		100%
016		1		1		100%
017	2	2	2	2		100%
018	1		1			100%
019	1	1	1	1		100%
020	1	1	1	1		100%
021	3		3			100%
022	1	1	1	1		100%
023		1		1		100%
024	2		2			100%
025	1	1	1	1		100%
026	1	1	1	1		100%
027	2		2			100%
028	2	1	2	1		100%
029	1	1		2	1-NL	50%
030	1	2	1	2		100%
031	1	2	1	2		100%
032	4		4			100%
033	1	1	3	1	2-IN	50%
034	1	1	1	1		100%
035	2	1	1	2	1-NL	66,7%
036	1	1	1	1		100%
037	1	1	1	1		100%
038	1		1			100%
039		1		1		100%
040	1	1	1	1		100%

ND – Neutrófilo com núcleo desconexo. IN – Inválido, ruído classificado como neutrófilo.

NL – Neutrófilo classificado como linfócito.

Fonte: elaborado pelo autor.

A partir da Tabela 4 e Tabela 5, percebe-se que o protótipo alcançou um percentual satisfatório de acertos, considerando que a rede neural foi treinada somente para classificar linfócitos e neutrófilos, logo, eventuais ruídos geralmente são classificados como neutrófilos. No que se refere à contagem diferencial, dos 51 neutrófilos reais, foram encontrados 54, obtendo-se 94,4% de acurácia. Já entre os linfócitos reais 38, foram encontrados 40, obtendo assim 95% de acurácia.

4 CONCLUSÕES

A proposta deste trabalho foi desenvolver um protótipo baseado em evidências para auxiliar o processo de contagem total e diferencial de leucócitos do tipo linfócito e neutrófilo em virtude desses tipos representarem cerca de 85% dos leucócitos encontrados no hemograma. Foi utilizada a linguagem Java para construção do protótipo, utilizando a biblioteca JavaCV e Encog para facilitar o trabalho de processamento de imagens e classificação.

As imagens foram obtidas em um ambiente bem controlado, contudo mostrou-se uma tarefa árdua, onde das 300 imagens obtidas, somente 100 tiveram um foco satisfatório, onde foram selecionadas 40 imagens que se enquadraram na proposta deste trabalho. Inicialmente, percebeu-se que as cores das lâmpadas interferiam nas técnicas de segmentação, onde soluções baseadas em cores e histogramas tiveram de ser descartadas devido às particularidades que cada célula trazia e criava novos cenários a serem atendidos.

O primeiro objetivo que era realizar a segmentação das células teve uma série de desafios, onde primeiramente foi delimitar a região do campo do microscópio, pois cada imagem proporcionava uma nova forma e tamanho diferente. Realizando a conversão do modelo RGB para HSV, utilizou-se o próprio canal V para servir de máscara e conseguir eliminar componentes externos. Enquanto que o canal S limiarizado teve um nível satisfatório de realce para atender a descoberta da região do núcleo das células.

Em relação ao núcleo das células, o principal desafio era detectar lóbulos desconexos e uni-los sem conectá-los com outras células, onde este item foi parcialmente atendido. Porém, com poucos casos onde o resultado foi indesejável. O crescimento de região uniforme, cujo propósito é delimitar a região ao redor do núcleo para agregar o citoplasma também foi atendido parcialmente, apesar de ser empírico e eventualmente não contemplar todo o citoplasma.

A partir do momento em que foi realizada a etapa de descoberta e a criação da lista de células, o método de extração de núcleo atendeu bem aos tipos linfócitos e neutrófilos, e possivelmente aos demais leucócitos. Para melhorar a extração do núcleo foi aplicado uma clusterização na imagem, isso fez com que o serrilhado fosse menor e conseqüentemente teve um aumento no grau de circularidade dos núcleos, o que ajudou a melhorar a diferenciação deste atributo entre os tipos deste trabalho.

Já a extração do citoplasma foi o maior desafio enfrentado neste trabalho, a clusterização aplicada no núcleo tornou o citoplasma mesclado com outras regiões. Por isso,

foi descartada a utilização da clusterização para extração do citoplasma. Alguns modelos de cores foram testados, entre eles: Luv, Lab, YCbCr e HLS, porém todos eles não tinham a região do citoplasma realçada e nem mesmo o próprio modelo HSV. Além do próprio modelo RGB ter sido descartado em virtude de que a solução baseou-se em cores e as intensidades variam muito em cada célula.

Como não se obteve um canal com o realce do citoplasma, foi realizada uma solução baseada na célula como um todo, para realizar posteriormente a subtração do núcleo. Neste caso, foi utilizado o canal H do modelo HSV, onde este canal sofreu uma série de operações morfológicas com a intenção de remover ruídos conexos, principalmente hemácias. O núcleo teve um papel importante nesta etapa para que a célula ficasse condensada e não sofresse alterações por conta das operações morfológicas. E, por fim, é realizada a subtração do núcleo, originando assim a parte do citoplasma. Contudo, cerca de 20% dos casos possuem ruídos médio-grandes, logo, considerou-se que esta etapa foi atendida parcialmente. Como primeiro objetivo deste trabalho a extração da célula e suas respectivas partes, núcleo e citoplasma, considerou-se o objetivo parcialmente atendido.

O segundo objetivo de extrair as características morfológicas da célula foi atendido com sucesso. Sendo que o grau de circularidade do núcleo atendeu de maneira satisfatória em vários momentos de validação morfológica do protótipo. Além da circularidade, a quantidade de lóbulos de uma célula serviu de maneira adequada para distinguir os tipos de células. O cálculo de área, núcleo e citoplasma, contribuem para eliminar ruídos.

O terceiro e último objetivo, classificar as células, foi atendido. Contudo, o sucesso deste objetivo é intimamente ligado aos objetivos anteriores. A rede neural do tipo MLP utiliza as características morfológicas extraídas anteriormente para inferir a classificação da célula. Além de resolver a problemática da enorme variação de características morfológicas de cada imagem, consegue resolver e inferir a classificação com nível satisfatório de acurácia.

Por fim, conclui-se que protótipo desenvolvido pode realizar com nível satisfatório a contagem total e a diferenciação de leucócitos do tipo linfócito e neutrófilo além de servir como base para outros trabalhos que vierem a classificar outros tipos de leucócitos. Também é possível concluir que para um protótipo deste gênero possa ser utilizado em um laboratório, torna-se necessário que se consiga classificar todos os leucócitos e com um percentual fidedigno de acurácia e, além disto, estudar como seria empregada a fase de aquisição das imagens, visto que a contagem manual feita por um profissional é um processo prático comparado a fase de aquisição das imagens.

4.1 LIMITAÇÕES

A primeira limitação prevista pelo protótipo é a variação de equipamentos, lâmpadas, cores e tudo que vier impactar nas características de uma imagem. Na base de dados utilizada neste protótipo, conseguiu-se encontrar uma solução que atenda aos casos, no entanto, é necessário informar manualmente os limiares do núcleo e citoplasma. O valor dos limiares são maiores para microscópios de lâmpada branca, enquanto que os de lâmpada amarela tiverem valores menores.

A base de dados da rede neural é um arquivo editável, embora seja possível melhorar o *dataset*, o protótipo contempla apenas os tipos: linfócito e neutrófilo e, atende somente aos quatro atributos pré-definidos: quantidade lóbulos, média de circularidade dos núcleos, tamanho do núcleo e tamanho do citoplasma.

4.2 EXTENSÕES

Abaixo estão listadas algumas sugestões de extensões para o protótipo:

- a) classificar outros tipos: basófilo, eosinófilo e monócito;
- b) utilizar uma rede Bayer e Kohonen para classificação e verificar acuracidade;
- c) verificar se os métodos de segmentação Watershed, Mumford-Shah, JSEG, e RHSEG podem melhorar o realce do citoplasma para realizar a extração;
- d) melhorar a união lóbulos separados;
- e) avaliar segmentação de imagens adquiridas de outros equipamentos;
- f) automatizar a definição dos limiares do núcleo e citoplasma;
- g) avaliar a segmentação de imagens em um esfregaço sanguíneo sem corante de realce.

Outras sugestões para tornar o protótipo aplicável como método interativo ao ensino:

- a) transformar em uma aplicação para dispositivo móvel para que o dispositivo possa ser utilizado diretamente sobre o microscópio;
- b) transformar em uma aplicação *real-time* para os resultados possam ser visualizados em tempo real, sem a necessidade de capturar uma foto;
- c) transformar em uma aplicação de realidade aumentada para que os componentes de interesse possam oferecer algum conteúdo, com textos, objetos 3D e rótulos.

REFERÊNCIAS

- CAMPESTRINI, Selma. **Laboratório & Sangue de rotina**. 2. ed. Curitiba: Champagnat, 1991. 102 p.
- COSTA, Luciano F.; CESAR JR., Roberto M. **Shape analysis and classification: theory and practice**. Boca Raton: CRC Press, 2000.
- FACON, J. **Morfologia matemática: teoria e exemplos**. Curitiba: Champagnat, 1996.
- FAILACE, Renato. **Hemograma: manual de interpretação**. Porto Alegre: Artmed, 2003. 298p.
- FAILACE, Renato; FERNANDES, Flavo. **Hemograma: manual de interpretação**. 6. ed. [S. l.]: Artmed, 2015.
- FEDECHEN, Emerson. A. et al. **Desenvolvimento de um sistema para contagem automática de células sanguíneas através de visão computacional**. Gaspar, 2012. Disponível em: <http://www.inf.furb.br/seminco/2012/slides/2012-09-04_21-15_artigo.pdf>. Acesso em: 16 mar. 2016.
- GAZZIRO, Mario. **Quadro segmentação de imagens**. São Carlos, 2013. <http://wiki.icmc.usp.br/images/b/bb/Otsu_e_derivadas.pdf>. Acesso em: 2 de abr. de 2016.
- GONZALES, Rafael C.; WOODS, Richard E. **Processamento de imagens digitais**. 1. ed. São Paulo: Edgard Blucher, 2000.
- _____. **Digital image processing**. 3rd ed. Upper Saddle River: Pearson Prentice Hall, 2008.
- HALL, John E. **Tratado de fisiologia médica**. 12 ed. Rio de Janeiro : Elsevier, 2011. 3p.
- HAYKIN, Simon. **Redes neurais: princípios e prática**. Porto Alegre: Bookman, 2001.
- HEGENBERG, Leonidas. **Doença: um estudo filosófico**. Rio de Janeiro: Editora FIOCRUZ, 1998. 137 p. Disponível em: <<http://books.scielo.org/id/pdj2h/pdf/hegenberg-9788575412589-03.pdf>>. Acesso em: 23 nov. 2016.
- LAB TEST ONLINE BR. **Contagem diferencial de leucócitos**. Disponível em: <<http://www.labtestsonline.org.br/understanding/analytes/differential/tab/test/>> Acesso em: 3 abr. 2016.
- LEUKOS. **Leukocytes archives**. Disponível em: <http://www.leukos.us/archives/11-01-2013_11-30-2013.asp> Acesso em: 24 nov. 2016.
- LORENZI, Therezinha F. et al. **Manual de Hematologia Propedêutica e Clínica**. 3. ed. Rio de Janeiro: Guanabara Koogan S.a., 2003.
- KLEIN, Carlos; MARTINS, Joel. **Implementação de rede neural em hardware de ponto fixo**. 2006. 87 f. Monografia - Curso de Engenharia Elétrica, Universidade de Brasília, Brasília, 2006. Disponível em: <http://bdm.unb.br/bitstream/10483/915/1/2006_JoelMartins_CarlosKlein.pdf>. Acesso em: 24 nov. 2016.
- MACIEL, Josias. **Análise de um sistema de crédito cooperativo através de redes neurais (MLP) com a utilização do algoritmo levenberg marquardt**. 2005. 87 f. Dissertação (Mestrado em Ciências) – Curso de Pós-graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Curitiba.

MAKRIS, Michael; GREAVES, Michael. **O sangue nas doenças sistêmicas**. [S. l.]: Manole, 1998. 101 p.

MELO, Márcio A. W.; SILVEIRA, Cristina M. **Laboratório de Hematologia: Teorias, técnicas e atlas**. [S. l.]: Editora Rubio, 2015. 280 p.

MONTEIRO, Felipe G. **Comparação dos resultados de hemogramas do contador eletrônico ABX PENTRA 60® com a microscopia**. 2005. 90 f. Dissertação (Mestrado) - Curso de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/7399/000543631.pdf>>. Acesso em: 3 abr. 2016.

NAOUM, Paulo Cesar; NAOUM, Flávio Augusto. **Interpretação laboratorial do Hemograma**. 2008. São José do Rio Preto. Disponível em: <http://www.ciencianews.com.br/arquivos/ACET/IMAGENS/Artigos_cientificos/Interphemo.pdf>. Acesso em: 24 nov. 2016.

PARSLOW, Tristram G. et al. **Imunologia médica**. 10. ed. Rio de Janeiro: Guanabara Koogan S.a., 2004. 16 p.

PATHPEDIA. **Blood cells**. Disponível em: <http://www.pathpedia.com/education/eatlas/histology/blood_cells/Images.aspx?00fa6c0f-5e6c-4c1b-a8b9-dfd2e32174f3>. Acesso em: 3 abr. 2016.

PETROU, Maria; BOSDOGIANNI, Panagiota. **Image processing: the fundamentals**. New York: John Wiley & Sons, 1999.

PLEBANI, Mario. **Exploring the iceberg of errors in laboratory medicine**. Padova, 2009. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0009898109001351>>. Acesso em: 20 mar. 2016.

PRIYANKARA G.P.M; SILVA, R.K.O.H. **An extensible computer vision application for blood cell**. Sri Lanka, 2006. Disponível em: <https://people.csail.mit.edu/oshani/publications/Lohitha_research.pdf>. Acesso em: 16 mar. 2016.

RAPAPORT, Samuel I. **Hematologia**. 2. ed. São Paulo: Roca, 1990. 157, 161 – 170 p.

RÉZIO, Ana C. C. **Reconhecimento e rastreamento de objetos**. 2008. 67f. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Departamento de Computação, Universidade Católica de Goiás, Goiânia. Disponível em: <<http://aldeia3.computacao.net/greenstone/collect/trabalho/index/assoc/HASH0147.dir/doc.pdf>>. Acesso em: 26 nov. 2016.

RICCI, Francesca. **La leucemia linfática crônica**. Disponível em: <http://www.malattiedelsangue.org/wp/?page_id=126>. Acesso em: 3 abr. 2016.

SCIACOVELLI, Laura et al. Quality Indicators in Laboratory Medicine: from theory to practice. **Clin Chem Lab Med**, New York, p.836-843, 2011. Disponível em: <http://www.centroricercabiomedica.net/crb/pdf/Lavori/2011_CCLM_QI_1.pdf>. Acesso em: 30 mar. 2016.

SEGMENTAÇÃO DE IMAGENS APLICADA A JOGOS. Indaiatuba: Reverte, n. 6, 17 nov. 2008. Disponível em:

<<http://www.fatecid.com.br/reverte/index.php/revista/issue/view/10>>. Acesso em: 26 nov. 2016

BAIN, Barbara J. **Células sanguíneas: consulta rápida**. Porto Alegre: Artmed, 1998. 118p.

SOARES, Heliana B. **Análise e classificação de imagens de lesões de pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte**. 2008. 154 f. Tese (Doutorado em Engenharia Elétrica e Computação) – Universidade Federal do Rio Grande do Norte, Natal. Disponível em:

<http://bdtd.bczm.ufrn.br/tesesimplificado/tde_arquivos/19/TDE-2008-12-03T063915Z-1651/Publico/HelianaBS_da_capa_ate_cap4.pdf>. Acesso em: 26 nov. 2016.

RUSS, John C. **The image processing handbook**. 6th ed. Boca Raton: CRC Press, 2011.

WEBER, Andréa. **Avaliação da proliferação celular baseada em PCNA-c1iclina por análise morfológica aplicada a linfomas malignos não-Hodgkin**. 1997. Dissertação (Mestrado) - Curso de Informática Industrial, Ciências Exatas e da Terra, Universidade Tecnológica Federal do Paraná, Curitiba, 1997. Disponível em:

<<http://www.ppgia.pucpr.br/~facon/Books/2006ERI2006CapituloNo3.pdf>>. Acesso em: 2 mar. 2016.

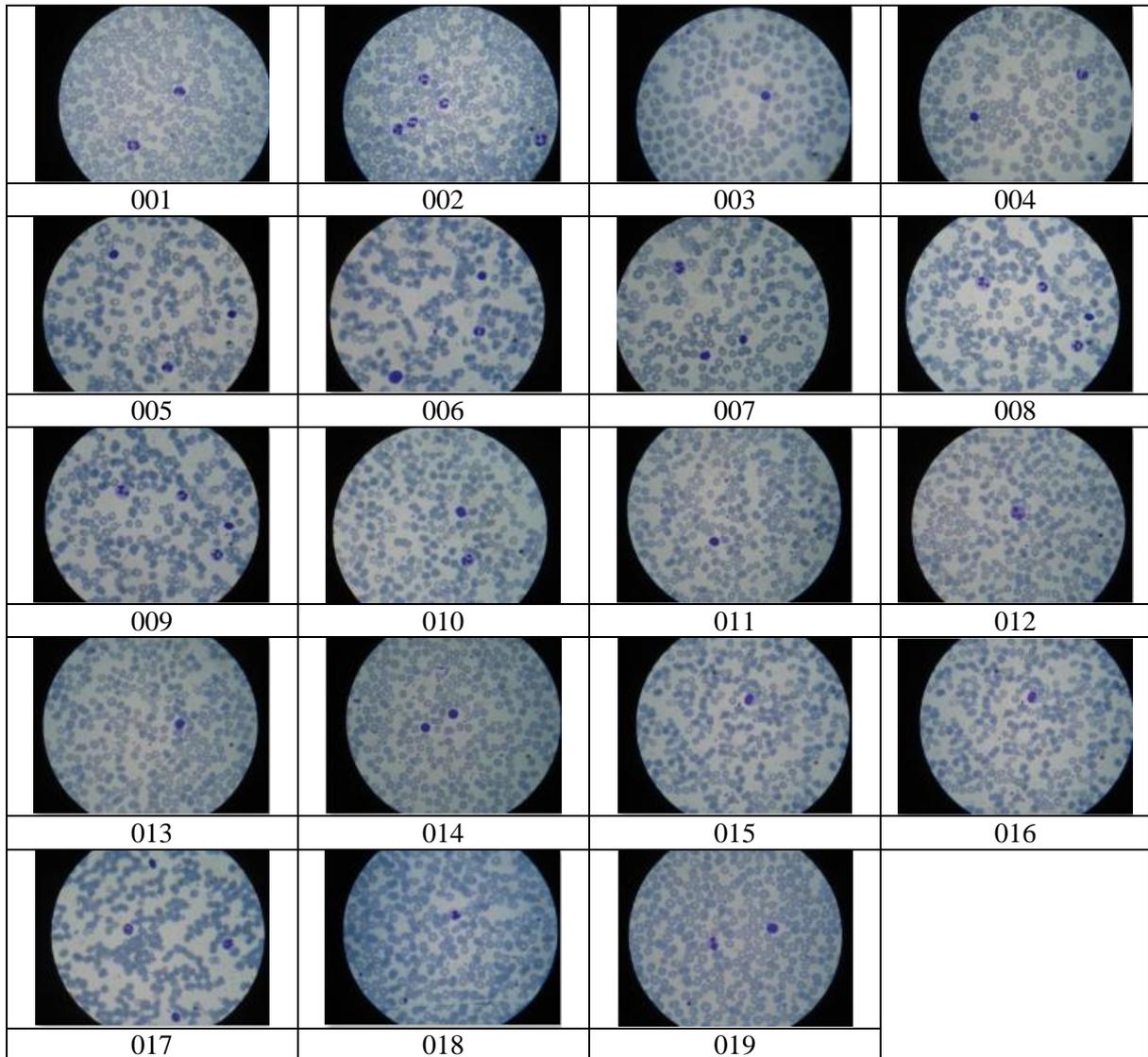
WILLIAMSON, Mary A.; SNYDER, L. M. **Interpretação de Exames laboratoriais**. Rio de Janeiro: Guanabara Koogan, 2013. 1004 p.

WU, Stephen G. et al. A leaf recognition algorithm for plant classification using probabilistic neural network. In: IEEE INTERNATIONAL SYMPOSIUM ON SIGNAL PROCESSING AND INFORMATION TECHNOLOGY, 7th, 2007, Giza. **Proceedings...** [S.l]: IEEE, 2008. p. 11-16.

APÊNDICE A – Base de imagens

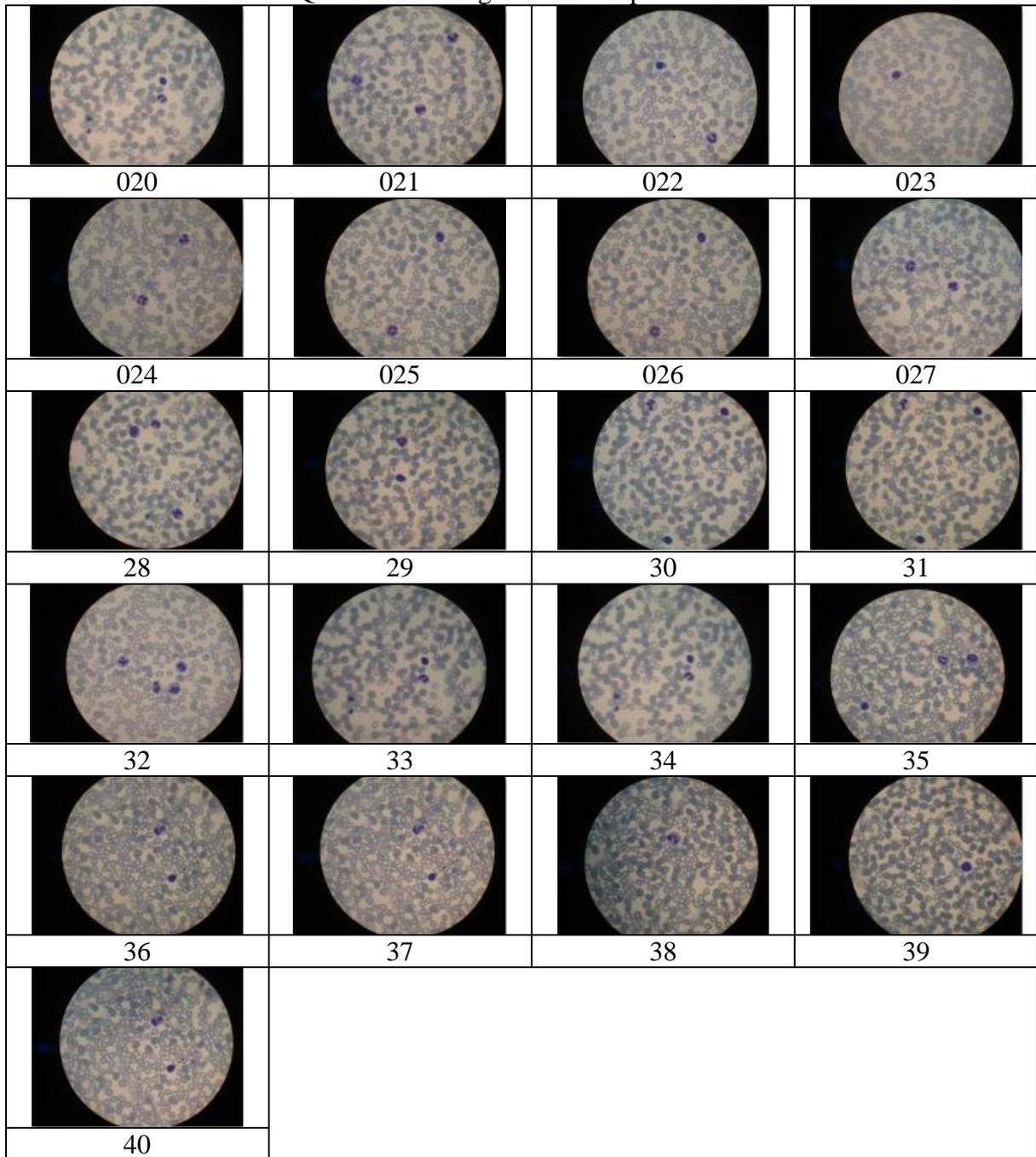
A seguir são apresentadas as imagens que compõem a base de dados utilizada neste trabalho.

Quadro 21 - Imagens com lâmpada branca



Fonte: elaborado pelo autor.

Quadro 22 - Imagens com lâmpada amarela



Fonte: elaborado pelo autor.