

Comparação de modelos de segmentação utilizados na identificação de linfócitos maduros e imaturos em lâminas sanguíneas

Comparison of segmentation models used to identify mature and immature lymphocytes in blood film

Avelino Alonso¹

 <https://orcid.org/0000-0002-3769-8276>

Matheus Marinho¹

 <https://orcid.org/0000-0002-4383-625X>

Clara Medeiros¹

 <https://orcid.org/0000-0002-3769-8276>

Carmelo Bastos-Filho¹

 <https://orcid.org/0000-0002-0924-5341>

¹ Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil. E-mail: agaj@ecomp.poli.br, mblm@ecomp.poli.br, anaclara.vita@upe.br, carmelofilho@ecomp.poli.br

DOI: 10.25286/rep.v7i2.2207

Esta obra apresenta a Licença Creative Commons Atribuição-Não comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: AG, Alonso Junior; M Marinho; ACAMV, FRAGOSO DE MEDEIROS; CJA Bastos Filho. Comparação de modelos de segmentação utilizados na identificação de linfócitos maduros e imaturos em lâminas sanguíneas. Revista de Engenharia e Pesquisa Aplicada, v.7, n. 2, p. 13-22, 2022.

RESUMO

A Leucemia Linfocítica Aguda (LLA) atinge cerca de 75 mil pessoas por ano, dentre estas 80% são crianças, sendo uma doença altamente invasiva e fatal o diagnóstico rápido é de grande importância, os métodos tradicionais de diagnóstico são caros e demorados, portanto, a utilização de métodos de segmentação de imagem utilizando inteligência artificial, podem auxiliar na detecção dos elementos de interesse em lâminas de sangue, os linfoblastos. Este trabalho comparou os modelos: Segnet, Mobilenet Segnet, Vgg Segnet, Resnet50 Segnet, Vgg Unet, Resnet50 Unet, Mobilenet Unet, FCN 8, FCN 32 e FCN 32 Mobilenet, pela precisão por pixel e tempo de execução. Utilizou-se a base de dados ALL-IDB, contendo lâminas de sangue de pacientes saudáveis e possíveis LLA. Como resultado, observou-se que as redes MobileNet desempenharam melhor, dentre elas, destacou-se a Mobilenet Unet onde o resultado da precisão média das classes foi de 83,4%.

PALAVRAS-CHAVE: Leucemia; Segmentação; LLA; Redes neurais;

ABSTRACT

Acute Lymphocytic Leukemia (ALL) affects about 75 thousand people a year, among these 80% are children, being a highly invasive and fatal disease, rapid diagnosis is of great importance, traditional diagnostic methods are expensive and time-consuming, therefore, the use of image segmentation methods using artificial intelligence can help detect the elements of interest in blood slides, the lymphoblasts. This work compared the models: Segnet, Mobilenet Segnet, Vgg Segnet, Resnet50 Segnet, Vgg Unet, Resnet50 Unet, Mobilenet Unet, FCN 8, FCN 32 and FCN 32 Mobilenet, by pixel precision and execution time. The ALL-IDB database was used, containing blood slides from healthy patients and possible ALL. As a result, it was observed that MobileNet networks performed better, among them, Mobilenet Unet stood out, where the result of the mean average precision of the classes was 83.4%.

KEY-WORDS: Leukemia; Segmentation; ALL; Neural networks;

1. Introdução

A leucemia é uma doença maligna dos glóbulos brancos. Tem como principal característica o acúmulo de células doentes na medula óssea, que substituem as células sanguíneas saudáveis.

A medula óssea é o local de fabricação das células sanguíneas e ocupa a cavidade dos ossos. Nela são encontradas as células que dão origem aos glóbulos brancos (leucócitos), os glóbulos vermelhos (hemácias ou eritrócitos) e às plaquetas.

Na leucemia, uma célula sanguínea que ainda não atingiu a maturidade sofre uma mutação genética que a transforma em uma célula cancerosa. Essa célula anormal não funciona de forma adequada, multiplica-se mais rápido e morre menos do que as células normais. Dessa forma, as células sanguíneas saudáveis da medula óssea vão sendo substituídas por células anormais cancerosas [5].

Além disso, esses leucócitos malignos podem fluir pelo sangue e causar danos a outras partes do corpo, como: fígado, rim, baço e até mesmo cérebro, o que pode levar a outras formas fatais de câncer [1].

Existem mais de 12 tipos de leucemia, sendo que os quatro primários são leucemia mieloide aguda (LMA), leucemia mieloide crônica (LMC), leucemia linfocítica aguda (LLA) e leucemia linfocítica crônica (LLC).

Dentre as leucemias citadas, destaca-se a leucemia linfocítica aguda (LLA) por seu maior risco ser em crianças de até 5 anos e por este motivo foi escolhida como foco da pesquisa.

A LLA geralmente é diagnosticada pela realização de um teste de hemograma completo. Neste teste, o médico irá verificar o número de leucócitos, se está muito alto, alguns sinais de células de leucemia. Porém, comumente o paciente pode ser submetido a outro procedimento, aspiração de medula óssea [2], seguido por um exame microscópico de distensão sanguínea [1]. Todos estes métodos que podem levar ao diagnóstico de leucemia são caros e também executados de formas manuais, os quais dependem totalmente dos especialistas médicos e são suscetíveis a erros humanos [6].

Com o intuito de diminuir os erros que podem acontecer durante a análise de uma lâmina de um paciente com possível LLA e democratizar este diagnóstico (diminuição de custos do exame), pode-se utilizar processos computacionais para facilitar o diagnóstico. Isto é, como um auxílio em diagnósticos médicos, pode-se utilizar processos computacionais na identificação de células doentes em uma lâmina de sangue.

Para a realização de processos computacionais em lâminas de sangue, utiliza-se a Patologia digital (PD), que é o processo de conversão de lâminas de vidro de histologia em imagens digitais. Normalmente este processo é feito através da acoplagem de uma câmera de alta definição no microscópio [26].

A Patologia digital e a digitalização de procedimentos na área de saúde, possibilita a utilização de técnicas de análise digitais neste âmbito, como a utilização do aprendizado de máquina.

O aprendizado de máquina é um campo da inteligência artificial (IA) que consiste em um módulo de extração de recursos que extrai as características de um determinado dado, no caso de imagens podem-se destacar como principais características: formas, bordas e texturas. Após a extração das características, um módulo de classificação utiliza os dados extraídos para tentar classificar as imagens.

A principal limitação do aprendizado de máquina é a incapacidade de extrair características que diferenciam o conjunto de dados de treinamento. Para este desafio, costuma-se usar técnicas de aprendizado profundo (DL, *Deep learning*) [4]. Na utilização em imagens, o desafio é ainda mais difícil, para este, as principais arquiteturas que utilizam estas técnicas de DL, são as Redes Neurais Convolucionais (CNN).

Para a classificação de imagens são utilizados diversos métodos e classificadores, como por exemplo, SVM (*Support Vector Machine* ou Máquina de Vetores de Suporte) na classificação de patologias [20], KNN (*K Nearest Neighbor*) utilizados em processamento de grandes bases de dados na área médica [21], a utilização de aprendizado de máquina (ML, *Machine Learning*) na detecção de anomalias em imagens de histopatologia da próstata [22], além da utilização de aprendizado de máquina na classificação de imagens de lâminas de vidro de histologia [3].

Este trabalho utilizou um conjunto de dados de imagens microscópicas de amostras de sangue, indicando pacientes com LLA e saudáveis, como objeto de análise, realizado por uma profissional biomédica. Através desta, foi possível a geração de um novo conjunto de dados, que contém as imagens originais (sem tratamento) e a imagem equivalente em forma de anotação (imagem utilizada para modelos de segmentação), através do novo conjunto de dados, foram avaliados e comparados modelos de CNN com base na sua precisão por pixel e tempo de treinamento.

2. Fundamentação Teórica

2.1 Leucemia Linfocítica Aguda (LLA)

O termo câncer, se refere a mais de 100 diferentes tipos de doenças malignas que se caracterizam pelo crescimento anormal de células, tendo em comum a invasão desordenada de tecidos e órgãos [14]. As células tendem a ser imaturas, ocasionando a formação de tumores por se infiltrar de forma incontrollável, podendo ser localizadas em diversas regiões do corpo [14].

Dentre os variados tipos de doenças malignas, tem-se a leucemia, na qual uma célula sanguínea que ainda não completou sua diferenciação, sofre mutações que a modificam, transformando-a em cancerígena [15]. Dentre as leucemias existentes, a leucemia linfocítica aguda (LLA), destaca-se características como a disseminação e agressão ao sistema imunológico, causada pela produção acelerada e desregulada nos leucócitos imaturos (linfoblastos). Desta forma, o corpo prioriza a produção destas células, diminuindo cada vez mais a produção de células saudáveis (plaquetas, hemácias e etc), ocasionando sintomas como anemias, infecções associadas e hemorragias [16].

A LLA é o câncer mais comum entre crianças e adolescentes, corresponde a cerca de 80% dos casos de leucemias na infância, e apenas 20% das leucemias do adulto [17]. Apesar de afetar diversas faixas etárias, a LLA é mais comum em crianças entre 2 a 5 anos, com maior acometimento da raça branca e pequena predominância para indivíduos do sexo masculino, porém, podendo aparecer também em idosos acima dos 60 anos [15].

De acordo com o Instituto Nacional de Câncer (INCA) [14], epidemiologicamente, a LLA ocorre na frequência de 1 : 25 mil crianças e adolescentes entre 0 a 15 anos, tendo no mundo, a cada ano, cerca de 75 mil novos casos de LLA em

crianças e adolescentes. Ainda de acordo com o INCA estima-se que para cada ano do triênio 2020/2022, sejam diagnosticados no Brasil, cerca de 11 mil novos casos nesses próximos 3 anos (cerca de 3.700 novos casos ao ano), onde as regiões Sudeste e Nordeste irão apresentar os maiores números de casos novos [14].

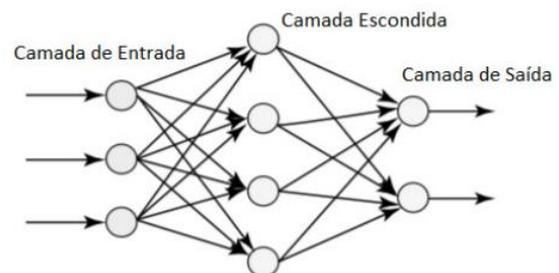
2.2 Redes Neurais Artificiais

O cérebro humano é composto por cerca de 10 bilhões de células que são denominadas neurônios, estes, estão constantemente interagindo uns com os outros através de impulsos que são chamados de sinapses. Cada neurônio pode receber um ou mais sinapses e, dependendo da forma de recebimento, podem ou não passar este impulso para um outro neurônio, formando uma cadeia de impulsos. Todo este conjunto é como o cérebro humano realiza suas funções [10].

Utilizando as redes neurais biológicas como inspiração, surgem as Redes Neurais Artificiais, que são modelos computacionais de aprendizagem de máquina. Estas, possuem a capacidade de realizar reconhecimento de padrões que podem ser identificados em sons, imagens, vídeos e outros [10].

Problemas mais complexos exigem uma arquitetura mais robusta para que sejam resolvidos [10]. Por isto, utiliza-se modelos com dezenas de neurônios em formas de camadas, chamados de *Multilayer Perceptron* (MLP).

Figura 1 – Arquitetura *Multilayer Perceptron*



Fonte: [10]

A Figura 1 representa uma arquitetura simples de uma MLP. Apresentando 3 tipos de camadas: camada de entrada, camada escondida e camada de saída. A camada de entrada é onde os dados são inseridos nestas redes (imagens, sons e etc) transformados em valores numéricos. A camada escondida, que pode conter uma ou mais camadas, nesta camada os dados são processados para depois serem transferidos para a camada de saída que na saída, resultará em algum valor que representa o dado processado.

2.2.1 Redes Neurais Convolucionais

A Aprendizagem Profunda (DL), é uma área de aprendizagem de máquina que se inspira como a maioria dos seres vivos interpretam as informações visuais.

Na DL, utiliza-se múltiplas camadas de aprendizagem com o intuito de que cada camada assimile representações simples das informações extraídas pelo modelo. Desta forma, o modelo consegue extrair características que definem as informações.

O modelo mais utilizado para aprendizagem de características de informações visuais é a rede neural convolucional (CNN, Convolutional Neural Network). Estas redes utilizam camadas de convolução e subamostragem, que são as camadas de abstração automática de características de imagens. Após essas camadas é utilizada uma MLP completamente conectada, comumente responsável pela realização da função de classificação.

2.3 Segmentação de Imagens

A segmentação de uma imagem consiste em subdividi-la em várias regiões ou objetos que fazem parte de sua composição através de algoritmos de segmentação.

A segmentação possibilita o isolamento das áreas de interesse em uma imagem como por exemplo formas, linhas, curvas e bordas [23]. Por ser uma das etapas fundamentais na análise de imagens digitais, a segmentação tem diversas aplicações como por exemplo na agricultura [18], reconhecimento óptico de caracteres, imagens médicas [19].

Gonzalez e Woods [23], dividem os algoritmos de segmentação em duas propriedades básicas de valores de intensidade: descontinuidade e similaridade.

A categoria de descontinuidade aborda a divisão de uma imagem com base nas mudanças repetidas de intensidade, como as bordas, por exemplo. Em relação à similaridade, a abordagem é baseada na divisão de uma imagem em regiões que são semelhantes de acordo com alguns critérios que são predefinidos.

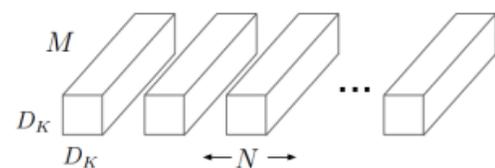
Neste trabalho foram utilizados alguns modelos *backbones* para segmentação baseados em diversas arquiteturas, são elas: ResNet, VGG e MobileNet.

ResNet é uma rede convolucional produzida pela empresa *Microsoft*, que ficou conhecida após a vitória no concurso ILSRVC (*ImageNet Large Scale*

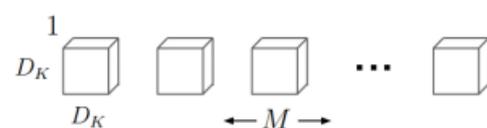
Visual Recognition Challenge) do ano de 2015. Quando as redes profundas começam a convergir, o problema de degradação pode aparecer: com o aumento da profundidade da rede, a precisão fica saturada e, em seguida, degrada rapidamente. Inesperadamente, tal degradação não é causada pelo *overfitting*, e adicionar mais camadas ao modelo profundo ocasionaria maior erro de treinamento. Desta forma, a ResNet foi construída por blocos residuais, estes blocos tem uma entrada x que passa por uma série de operações de convolução-relu-convolução. O resultado da operação $f(x)$ é adicionado à entrada original x .

A VGG (*Visual Geometry Group*), utiliza camadas convolucionais 3 x 3 empilhadas umas sobre as outras com profundidade crescente ao invés de filtros grandes. Em algumas variações da VGG, consiste em duas camadas totalmente conectadas com 4096 canais cada, que são seguidas por outra camada totalmente conectada com 1000 canais para prever 1000 rótulos. A última camada totalmente conectada usa a camada softmax para fins de classificação [24].

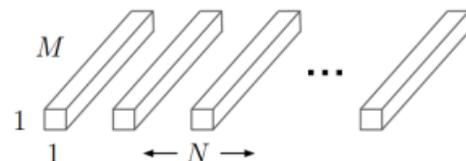
Figura 2 – Conceito de convolução distintas



(a) Standard Convolution Filters



(b) Depthwise Convolutional Filters



(c) 1 x 1 Convolutional Filters called Pointwise Convolution in the context of Depthwise Separable Convolution

Fonte: [27]

A MobileNet possui uma proposta diferente, de ser uma rede neural profunda, porém leve, sendo uma das redes de principais interesses para

sistemas móveis de visão computacional. Este modelo é baseado em um conceito de *Depthwise Separable Convolution*, que é uma forma de convolução fatorada [27].

A comparação entre os diversos tipos de convolução pode ser visualizada na Figura 2. A convolução em profundidade, *Depthwise* é feita através de uma única convolução em cada canal de cor, ela apenas filtra os canais de entrada e não os combina para criar novas *features* [27].

Para a geração de novas *features*, utiliza-se a uma convolução chamada *Pointwise*, para calcular uma forma linear da anterior, *Depthwise*. Esta rede apresenta no total 28 camadas, considerando as duas convoluções como camadas separadas [25].

3. Trabalhos relacionados

Scotti [7] apresenta diferentes métodos para medir com precisão, propriedades de células brancas em imagens de microscópio. Dentre os principais métodos abordados neste trabalho estão: remoção de fundos indesejados, estimativa de diâmetro médio e segmentação de imagens auto-adaptáveis. Através da análise, afirma que com a utilização dos métodos descritos é possível extrair características suficientes dos glóbulos brancos para o diagnóstico de leucemia aguda.

Em [8], os autores apresentam uma CNN para detecção automatizada de leucemia linfoblástica aguda e seus subtipos. Dado ao número limitado de dados de treinamento, para reduzir o overtraining (situação onde o modelo é "super treinado", isto é, converge para dados de treinamento mas não consegue generalizar para teste), os autores utilizaram técnicas de aumento de dados, como rotação de imagens e espelhamento. Os mesmos relatam que o modelo treinado sem a utilização deste tipo de técnica tende a não generalização do modelo.

Kolokolnikov e Samorodov [9] desenvolvem um estudo comparativo de técnicas de AD para classificação de glóbulos brancos. Durante o processo, vários conjuntos de dados são criados com o objetivo de balancear os dados de cada classe. Os autores apresentam resultados da avaliação do modelo e apontam dois métodos que fornecem maior precisão característica do modelo de classificação.

4. Metodologia

Esta seção apresenta o conjunto de dados utilizados neste trabalho, os modelos utilizados para análise, o pré-processamento dos dados, bem como técnicas para melhorar os resultados obtidos.

4.1 ALL-IDB

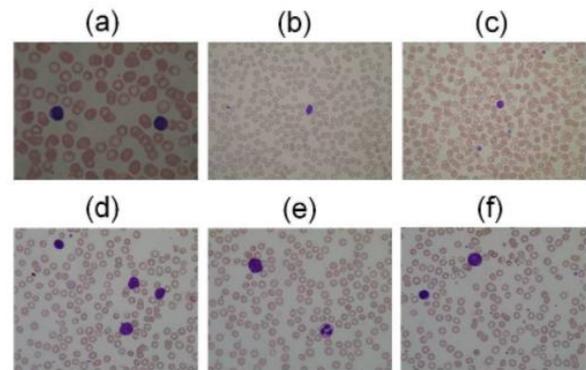
O ALL-IDB [11] é um conjunto de dados público e gratuito de imagens microscópicas de amostras de sangue, que foi criado especificamente para a avaliação e comparação de algoritmos de segmentação e classificação de imagens.

Para cada imagem do conjunto de dados, a classificação e posição dos linfoblastos foi identificada e avaliada por oncologistas especialistas.

Para a montagem do conjunto de dados, foi utilizado um microscópio óptico de laboratório acoplado a uma câmera de alta resolução para garantir a qualidade da imagem. Desta forma foram gerados dois conjuntos de dados: (i) *ALL_IDB1* e (ii) *ALL_IDB2*.

O primeiro conjunto (i) é composto por 108 imagens, onde os linfócitos foram marcados por especialistas. As imagens são tiradas com diferentes ampliações do microscópio.

Figura 3 – Exemplos de imagens em ALL_IDB1

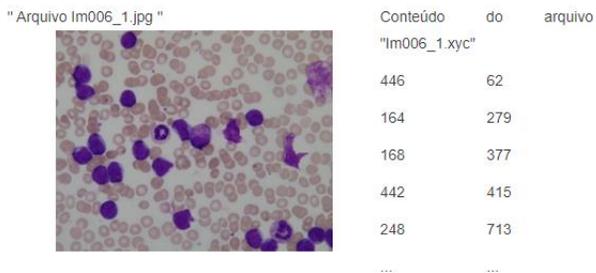


Fonte: [11]

Na Figura 3 é apresentado seis imagens contidas neste conjunto de dados, onde três são células referentes a pacientes considerados não LLA (imagens de a - c) e outras três são lâminas referentes a prováveis pacientes LLA (imagens de d - f).

Cada imagem deste conjunto é associada a um arquivo de classificação, isto é, para cada arquivo de imagem (lado esquerdo da Figura 4) há um arquivo de texto relatando as coordenadas dos centróides de prováveis linfoblastos (lado direito da Figura 4).

Figura 4 – Imagem original e Arquivo de anotação da imagem

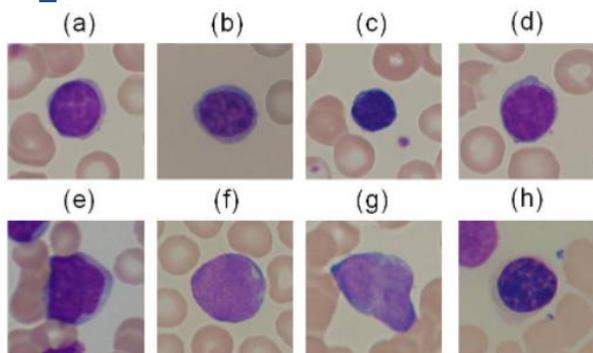


Fonte: [11]

Além dos arquivos de anotação, que relatam coordenadas da imagem, os arquivos são nomeados com a notação "ImXXX_Y.jpg" onde XXX é um contador inteiro de 3 dígitos e Y é um dígito booleano, onde 0 se nenhuma célula de blasto estiver presente e igual a 1 se pelo menos uma célula de blasto estiver presente na imagem. Portanto, todas as imagens marcadas com Y = 0 são de indivíduos classificados como saudáveis e todas marcadas com Y = 1 são de pacientes com possível LLA.

O segundo conjunto (ii) é composto por 260 imagens, onde as dimensões das imagens são menores quando comparadas com o primeiro conjunto.

Figura 5 – Exemplos de imagens do conjunto ALL_IDB2



Fonte: [11]

A Figura 5 apresenta exemplos de imagens contidas neste conjunto, onde as células saudáveis de pacientes não LLA (de a - d), e as células de prováveis linfoblastos de pacientes LLA (de e - h) estão representadas.

Este conjunto também utiliza a mesma forma de notação nos nomes dos arquivos, isto é

"ImXXX_Y.jpg". Desta forma, é possível analisar as células de forma individual.

Neste trabalho foi utilizado apenas o primeiro conjunto de dados (ALL_IDB1).

4.2 Aperfeiçoamento do conjunto de dados

O primeiro conjunto de dados (ALL_IDB1) passou por uma etapa de re-validação, que foi feita manualmente por uma profissional da área de biomedicina.

Nesta etapa os dados foram analisados individualmente, onde a cada imagem do conjunto de dados, foi marcado digitalmente as células de interesse, isto é, células linfóides imaturas e maduros. Desta forma, o novo conjunto de dados criado após esta etapa apresenta marcações indicando não só as áreas de interesse para possível LLA, mas também de linfócitos, permitindo então, a classificação dos dois tipos de células.

As marcações foram efetuadas digitalmente através do programa GIMP [12].

4.3 Pré-processamento

Os dados reais de hoje são altamente suscetíveis a ruídos, falta de dados e sua origem em várias fontes. Consequentemente, a qualidade dos dados pode ser comprometida levando a resultados de baixa qualidade.

Portanto, para isso, é necessária uma fase de pré-processamento dos dados, que consiste em uma fase de preparação dos dados a ser aplicada para geração de melhores resultados.

Esta etapa foi responsável pelo processo de criação de imagens de anotação (ou máscaras) das áreas de interesse para a segmentação e aumento de dados.

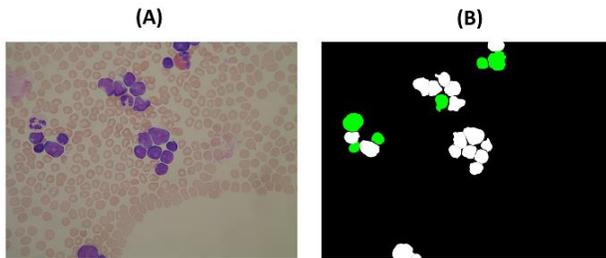
A partir dos pontos indicando os centróides de cada região de interesse (linfócitos maduros e linfoblastos), foram geradas imagens, que representavam cada classe, isto é, cada componente que componha a imagem, são estes: (i) Pixels que representavam o Fundo da lâmina (região de não interesse), (ii) Pixels que representavam um linfócito maduro, (iii) Pixels que representavam um linfoblasto.

Todos os pixels foram coloridos em tons de cinza de forma que pudessem representar a sua classe. Os pixels que representavam uma região de não interesse na classificação, foram coloridos de preto absoluto, RGB (Red Green Blue) todos zero

(RGB: 0, 0, 0). Os pixels, que representavam um linfócito maduro, foram coloridos em RGB, todos em RGB: 1, 1, 1. Já os pixels que representavam um linfoblasto, foram representados através do RGB: 2, 2, 2. A coloração de cada cor foi escolhida para facilitar a representação (todos os pixels, tanto vermelho, quanto verde e azul, sendo o mesmo número). Este processo foi realizado manualmente, utilizando o programa de edição de imagens GIMP [12].

Esta etapa é importante para que os modelos de classificação consigam distinguir o que cada pixel representa, portanto, permitindo a classificação de células LLA e células saudáveis.

Figura 6 – Etapa de pré-processamento



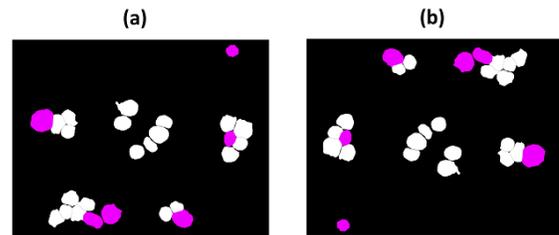
Fonte: Os autores.

Ao final do processo tem-se como resultado duas imagens: a imagem original (Figura 6 - A) e a imagem mascarada (Figura 6 - B).

Com as imagens originais e a imagem com sua respectiva máscara (indicando cada classe), o conjunto de dados foi dividido em dois subconjuntos, um de treino e outro de teste. O conjunto de treino é o responsável para que o modelo aprenda os padrões que podem indicar o tipo de classificação de cada pixel da imagem. Já o de teste, é o conjunto que o modelo não tem acesso, e servirá como conjunto avaliador, isto é, de onde serão calculadas as métricas como precisão de cada pixel, para avaliar a eficácia dos modelos. Estes conjuntos foram divididos em 75% e 25% para treino e teste, respectivamente.

Para extrair uma melhor capacidade de generalização dos modelos, foram aplicadas no conjunto de treinamento, técnicas de aumento de dados. Sendo elas: (i) inversão no eixo X, (ii) inversão no eixo Y e (iii) rotação de no máximo 35 graus.

Figura 7 – Aumento de dados aplicados ao treino



Fonte: Os autores.

Cada imagem é replicada e submetida a uma das técnicas de aumento de dados, onde a imagem original (Figura 7 - a) continua no processo de treino, e além dela, há uma inclusão de uma segunda imagem criada a partir da primeira (Figura 7 - b), neste caso a técnica utilizada foi a inversão no eixo Y.

4.4 Modelos

Com o conjunto de dados pré-processados, a próxima etapa é seleção, preparação e treinamento dos modelos. Para este trabalho, foram utilizadas arquiteturas já criadas e disponíveis gratuitamente através da biblioteca de rede neural de código aberto escrita em Python, Keras [13].

Através desta biblioteca é possível utilizar modelos pré-criados e que já foram utilizados e avaliados pela comunidade, a Tabela 1 apresenta os modelos que serão comparados neste trabalho.

Os modelos representados na Tabela 1, foram selecionados com base na sua maior utilização no âmbito de segmentação de imagens digitais da biblioteca Keras [13], além da fácil implementação através de códigos em Python. Trazendo maior agilidade a processos de pesquisa e comparação de modelos.

Tabela 1 - Modelos de segmentação comparados no trabalho*

NOME DO MODELO	BASE DO MODELO	MODELO DE SEGMENTAÇÃO
Segnet	Vanilla CNN	Segnet
MobileNet Segnet	MobileNet	Segnet
Vgg Segnet	VGG 16	Segnet
Resnet50 Segnet	Resnet-50	Segnet
Vgg Unet	VGG 16	U-Net
Resnet50 Unet	Resnet-50	U-Net
MobileNet Unet	MobileNet	U-Net
FCN 8	Vanilla CNN	FCN 8
FCN 32	Vanilla CNN	FCN 32
FCN 32 MobileNet	MobileNet	FCN 32

Fonte: Os autores.

Os modelos serão avaliados através de duas métricas: (i) Precisão do pixel e (ii) Tempo de treinamento.

A Precisão do pixel (i), é a métrica de avaliação mais amplamente usada para modelos de segmentação. É definido como a precisão da previsão em pixels (Equação 1).

$$(1) P_{acc} = \frac{\sum_{i=0}^K (p_{ii})}{\sum_{i=0}^K \sum_{j=0}^K (p_{ij})}$$

Na equação acima, K representa o número total de pixels na imagem de teste, e p_{ii} é o pixels previstos como classe i , e a parte denominador é representada como p_{ij} , o número de pixels da classe i previsto como classe j .

Já o tempo de treino (ii), é o tempo de execução em segundos que a máquina efetuou o treinamento do modelo, este tempo pode estar associado a custo, já que em alguns sistemas de processamento na nuvem o usuário paga por tempo de execução.

4.5 Proposta

Neste trabalho é proposto uma análise comparativa de modelos de segmentação de imagens utilizados em uma base de dados contendo imagens sanguíneas indicando possíveis pacientes saudáveis e com LLA.

Foram utilizados modelos de segmentação disponíveis na biblioteca digital Keras [13], apresentando resultados satisfatórios em suas segmentações utilizando modelos de simples implementação e renomados na literatura.

Durante o processo comparativo, os autores também desenvolveram uma base de dados contendo imagens de mascaramento, indicando 3 classes distintas em cada imagem por coloração de pixel (fundo da lâmina, células LLA e células de linfócitos maduros). As imagens foram geradas e validadas por uma profissional da área da hematologia. Com isso, promovendo meios para facilitar futuros trabalhos no âmbito da segmentação de imagens de LLA.

Além disto, indicando quais os melhores modelos foram avaliados para esta tarefa.

5. Resultados

Para definir qual modelo teve um melhor desempenho, todos foram treinados utilizando a mesma base de treino e foram comparadas utilizando as métricas de precisão do pixel (Pixel Acc.) e tempo de treino.

Para avaliar os modelos a priori, todos foram treinados utilizando as mesmas

configurações iniciais em um modelo de busca exaustiva, ou seja, todos os modelos foram treinados e avaliados manualmente, sendo comparados pelas métricas citadas anteriormente.

Após o treinamento de todos os modelos, apresentaram os seguintes resultados (Tabela 2).

Tabela 2 - Resultados dos modelos

NOME DO MODELO	TEMPO DE TREINO (s)	Pixel Acc.* (A)	Pixel Acc.* (B)	Pixel Acc.* (C)
Segnet	267	99%	62%	2%
Mobilenet Segnet	395	99.5%	77.8%	61.6%
Vgg Segnet	443	98.3%	51%	48.3%
Resnet50 Segnet	360	99%	63%	3%
Vgg Unet	485	99%	67%	9%
Resnet50 Unet	424	99%	68%	18%
Mobilenet Unet	440	99%	83.6%	63.1%
FCN 8	335	99.1%	69%	1.5%
FCN 32	683	98.4%	44%	7%
FCN 32 Mobilenet	918	99.1%	55.8%	3%

*Métrica de precisão do pixel de cada classe (A, B e C).

Fonte: Os autores.

Na tabela 2, observa-se cada modelo, sendo representados em cada linha, trazendo informações sobre cada modelo através das colunas, dentre elas: Nome do modelo (primeira coluna), Tempo de duração do treino daquele modelo em segundos (segunda coluna), seguido da precisão de pixel para cada classe do modelo, sendo representados da terceira, quarta e última coluna, respectivamente. As classes existentes para classificação são: (A) Classe referente a o que não era região de interesse (fundo da lâmina e outros), (B) Classe dos linfoblastos e (C) Classe dos linfócitos maduros.

Analisando os resultados, levando em consideração a precisão baseada em classes, observa-se a não eficácia em algumas redes na classificação da Classe C (linfócitos maduros), isto se deve, possivelmente, pela falta de componentes indicando margens limites entre as células, visto que algumas redes utilizam a técnica de Segmentação Semântica de imagens, portanto, os modelos classificaram de forma errada todos os linfócitos maduros, como linfoblastos.

Ainda sobre os modelos, nota-se o destaque dos modelos baseados na *Mobilenet*, onde, a maioria, com exceção da *FCN 32 Mobilenet*, apresentam uma precisão satisfatória em comparação aos outros modelos propostos.

Os modelos baseados em Redes Completamente Convolutivas (FCN, Fully

Convolutional Network), apresentam um treinamento bem mais custoso (maior tempo de treino), comparado a outras redes, devido a sua complexidade visto que a última camada da rede é substituída por operações convolutivas.

Todos os modelos treinados na Tabela 2, utilizaram as configurações iniciais de época definidas como padrão pela biblioteca digital Keras [13], uma época e com a largura da entrada (*input width*), e altura da entrada (*input height*) em 1696 e 1344, respectivamente.

Com o objetivo de melhorar os resultados obtidos, uma nova etapa de treinamento foi realizada, utilizando os três modelos mais bem avaliados, levando em conta a média de precisão das classes. As redes selecionadas para a rodada de testes com configurações aprimoradas foram as: *Mobilenet Segnet*, *Mobilenet Unet* e *Vgg Segnet*.

Os modelos foram novamente treinados utilizando novas configurações, duas épocas para cada modelo e a mesma configuração quanto a altura e largura de entrada, além da utilização de aumento de dados para todos os elementos. Estas configurações aumentam o tempo de execução para cada treinamento, por isso não foram utilizadas na primeira etapa de busca exaustiva.

Tabela 3 - Resultados dos modelos classificados

NOME DO MODELO	TEMPO DE TREINO (s)	Pixel Acc.* (A)	Pixel Acc.* (B)	Pixel Acc.* (C)
Mobilenet Segnet	926	99,3%	76,3%	45,1%
Mobilenet Unet	961	99,6%	81,4%	69,2%
Vgg Segnet	932	99,2%	63,4%	43,8%

*Métrica de precisão do pixel de cada classe (A, B e C).

Fonte: Os autores.

Através da Tabela 3, observa-se os modelos que obtiveram um melhor resultado com base em sua precisão por classe, além de um tempo de treinamento satisfatório. Destaca-se o modelo *Mobilenet* baseado na *Unet*, este obteve as melhores precisões na média das classes propostas para a segmentação.

Os modelos *Mobilenet Segnet* e *Vgg Segnet*, apresentaram um resultado inferior na segmentação das classes B e C quando comparados ao *Mobilenet Unet*, isto deve-se à grande semelhança das classes em questão.

Vale ressaltar a distribuição decrescente em relação a precisão dos modelos, onde a Classe A apresenta uma maior precisão em comparada a Classe B e a Classe C, isto se deve à distribuição não

equilibrada das classes, onde a Classe C é em menor quantidade visto a menor presença em imagens de sangue em pacientes saudáveis e a Classe A, a maior aparição, onde o fundo da imagem representa uma maior porcentagem da imagem quando comparado aos objetos das demais classes.

É importante observar também a precisão da Classe A nos três modelos, onde obtiveram uma precisão de mais de 99%, indicando a facilidade dos modelos de classificar o fundo da imagem, podendo ser utilizado como uma forma de remoção de fundo para classificação direta apenas das classes B e C.

6. Conclusões

Este trabalho teve como objetivo aplicar modelos de segmentação pré-existentes de fácil acesso através da biblioteca digital Keras [13], além de avaliá-los na segmentação de imagens de lâminas de sangue, contendo Linfócitos maduros e Linfoblastos, indicando regiões de interesses nas lâminas.

Para tanto, foi gerado um conjunto de dados contendo imagens de mascaração de lâminas de sangue, indicando os elementos de interesse em cada lâmina com suas respectivas classificações (Linfócitos maduros e imaturos) e validado por uma especialista da área.

Dos modelos avaliados o modelo *Mobilenet Unet* apresentou um melhor desempenho, obtendo uma precisão para identificar linfócitos imaturos (linfoblastos) de 81,4% e 69,2% para identificação de linfócitos maduros (células saudáveis). Além do tempo de processamento, que para estas configurações foi de 961 segundos.

É importante observar que a principal métrica utilizada na avaliação dos modelos é a precisão por pixel em cada classe, esta medida é calculada através da comparação direta de cada pixel, isto é, todos os pixels da imagem segmentada pelo modelo, são comparados à anotação da imagem original, caso não sejam iguais, aquele pixel é contabilizado como um erro e assim por diante.

Desta forma, estas métricas podem não representar o nível de eficácia do modelo em sua aplicação em um cenário real, visto que alguns pixels classificados de forma errada podem não apresentar uma distinção notável visivelmente.

Além disto, a utilização apenas da precisão como métrica comparativa pode mascarar resultados quando o número de representação das classes não for equilibrado.

Portanto, é proposto como trabalhos futuros a utilização de outras métricas de classificação como *F1 score*. Também a criação de um novo modelo, com

base nos já conhecidos para segmentação de imagem, proposto para segmentação de imagens médicas.

Além disso, efetuar a análise dos modelos de forma manual por profissionais da área de hematologia baseados na classificação visual de cada imagem, interpretando e contabilizando a quantidade de células classificadas corretamente ou não. Desta forma, pode-se avaliar no cenário real, qual modelo foi mais eficaz em suas classificações.

Referências

- [1] **The American Society of Hematology**. Acessado em: 28 de Junho, 2021. Disponível em: <https://www.hematology.org/>
- [2] **Curesearch for Childrens Cancer Research**. Acessado em: 23 de Junho, 2021. Disponível em: <https://curesearch.org>.
- [3] GURCAN, M.N.; BOUCHERON L. E.; CAN A.; MADACHUSHI A.; RAJPOOT N. M.; YENER B. "Histopathological image analysis a review". 2009.
- [4] H. LEE et al. "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations". 2009, pp. 609–616. *26th Annual International Conference on Machine Learning*.
- [5] **Instituto Nacional de Câncer Ministério da Saúde**. Acessado em: 16 de Junho, 2021. Disponível em: <https://www.inca.gov.br/>
- [6] Mrazek C, Lippi G, Keppel MH, et al. **Errors within the total laboratory testing process, from test selection to medical decision-making - A review of causes, consequences, surveillance and solutions**. *Biochem Med (Zagreb)*. 2020; 30(2):020502. doi:10.11613/BM.2020.020502.
- [7] F. Scotti. "Robust Segmentation and Measurements Techniques of White Cells in Blood Microscope Images". 2006.
- [8] S. Shafique and S. Tehsin. "Acute lymphoblastic leukemia detection and classification of its sub-types using pretrained deep convolutional neural networks". 2018.
- [9] G. Kolokolnikov and A. Samorodov. "Comparative study of data augmentation strategies for white blood cells classification". 2019.
- [10] A. Abraham. **Artificial Neural Networks**. American Cancer Society, 2005. ISBN 9780471497394.
- [11] R. D. Labati, V. Piuri, and F. Scotti, "All-idb web site". University of Milan, Department of Information Technologies, Disponível em: <http://www.dti.unimi.it/fscotti/all>.
- [12] **GIMP - The Free & Open Source Image Editor**. Disponível em: <https://www.gimp.org/>.
- [13] Chollet, F. et al., 2015. **Keras**. Disponível em: <https://github.com/fchollet/keras>.
- [14] **INCA - Instituto Nacional de Câncer José Alencar Gomes da Silva**. Estimativa 2020: incidência de câncer no Brasil/Instituto Nacional de Câncer José Alencar Gomes da Silva. Rio de Janeiro : INCA, 2019.
- [15] MOREIRA et al. **Avaliação dos aspectos citológicos e laboratoriais da leucemia linfóide aguda**. *Revista Eletrônica Acervo Saúde - REAS*. Vol.13(5).
- [16] AMARAL E JUVENALE. **Leucemia linfóide aguda em pacientes infanto-juvenis**. *Braz. J. Hea. Rev. Curitiba*, v. 3, n. 3, p.4770-4784.
- [17] DUTRA, R. A.; ABRAHÃO, C. A.; LOPES, F. M.; ROCHA, R. F. S.; ROSA JUNIOR, S. P. **A importância do hemograma no diagnóstico precoce da leucemia**. *Revista Eletrônica Acervo Saúde / Electronic Journal Collection Health*. Vol. 12 (7), p. 1-8, e3529, Jun. 2020.
- [18] SARATH, D. S.; SILVA, G. G.; PERUCA, R. D.; MACHADO, B. B.; ROEL, A. R.; PISTORI, H., **Quantificação automática da área foliar na cultura da soja usando segmentação de imagens coloridas**. X Congresso Brasileiro de Agroinformática, 2015.
- [19] MARCOMINI, K. D.; CARNEIRO, A. A.; SCHIABEL, H., **Application of Artificial Neural Network Models in Segmentation and Classification of Nodules in Breast Ultrasound Digital Images**, *International Journal of Biomedical Imaging*, vol. 2016, 13 pages, 2016.
- [20] A. M. Takakura, D. R. Pereira, F. A. Silva, M. A. Pazoti, L.L. Almeida, H. M. Sapia. **Uso do aprendizado de máquina no diagnóstico médico de patologias**. *Colloquium Exactarum*, v. 10, n.1 , Jan-Mar. 2018, p. 78–90.
- [21] W. Xing e Y. Bei, "Medical Health Big Data Classification Based on KNN Classification Algorithm," *IEEE Access*, vol.8, pp. 28808-28819, 2020, doi: 10.1109.
- [22] C. A. C. Seabra, "Prostate cancer biochemical recurrence prediction after radical prostatectomy using machine learning analysis of histopathology". 2019.

Disponível em: <http://hdl.handle.net/10451/40425>

[23] GONZALEZ, R. C.; WOODS, R. E. **Processamento Digital de Imagens**. Edgard Blucher, 2010.

[24] K. Simonyan, A. Zisserman. **Very Deep Convolutional Networks for Large-Scale Image Recognition**. Computer Vision and Pattern Recognition. arXiv:1409.1556. 2014.

[25] A. G. Howard, Z. Menglong, Bo Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam. **MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications**. 2017.

[26] Sellaro TL, Filkins R, Hoffman C, Fine JL, Ho J, Parwani AV, Pantanowitz L, Montalto M. **Relationship between magnification and resolution in digital pathology systems**. J Pathol Inform 2013; 4 : 21.

[27] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam. **MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications**. Computer Vision and Pattern Recognition (cs.CV). arXiv:1704.04861