

Rede Neural Convolutacional para o Diagnóstico de Leucemia

Luis H. S. Vogado, Rodrigo M. S. Veras, Flavio H. D. Araujo,
Romuere R. V. Silva, Kelson R. T. Aires

¹Departamento de Computação,
Universidade Federal do Piauí - Teresina, Brasil

lhvogado@gmail.com, {rveras, flavio86, romuere, kelson}@ufpi.edu.br

Abstract. *Leukemia is a type of cancer that affects the production of blood cells in the bone marrow which makes it challenging to coagulate blood and fight infection. In this work, we propose a method for the automatic diagnosis of leukemia using Convolutional Neural Networks (CNNs). We use pre-trained CNNs and learning transfer techniques in constructing the proposed method. We employed the Deeply Fine Tuning Modified (DFTM) technique combined with data augmentation operations to fine-tune the pre-trained model. To train and test the proposed method, we used a set of 2304 images from 14 different image databases. The proposed method reached an accuracy of 98.84%, and when compared to other works, we observed greater robustness and consistency in the results. We conclude that the fine-tuning is more robust the classification of heterogeneous images when compared to the features extraction through CNNs.*

Resumo. *A leucemia é um tipo de câncer que afeta a produção de células sanguíneas na medula óssea o que dificulta a coagulação do sangue e o combate a infecções. Nesse trabalho propomos um método para o diagnóstico automático de leucemia utilizando Redes Neurais Convolutacionais (CNNs). Nós utilizamos CNNs pré-treinadas e técnicas de transferência de aprendizagem na construção do método proposto. Empregamos a técnica Deeply Fine Tuning Modified (DFTM) combinada com operações de aumento de dados para refinar um modelo pré-treinado. Para treinar e testar o método proposto, utilizamos um conjunto de 2304 imagens de 14 bases diferentes. O método proposto atingiu acurácia de 98,84% e quando comparado com outros trabalhos, observamos maior robustez e consistência nos resultados. Ao final, concluímos que o ajuste fino é mais robusto a classificação de imagens heterogêneas quando comparado com a extração de características através de CNNs.*

1. Introdução

Leucemia é uma das doenças mais perigosas de acordo com o *American Cancer Society*¹, tendo sido estimados cerca de 61.780 novos casos e 22.840 mortes em 2019. A leucemia é uma doença sem origem definida que afeta a produção de glóbulos brancos na medula óssea. Com a presença da enfermidade, células jovens ou blastos são produzidos de forma anormal, substituindo células sanguíneas normais (glóbulos brancos, vermelhos e plaquetas). Conseqüentemente o indivíduo sofre com problemas no transporte de oxigênio e

¹<https://cancerstatisticscenter.cancer.org/!/cancer-site/Leukemia>

no combate a infecções. Dentre as formas de diagnóstico da leucemia, temos a punção lombar, mielograma, hemograma e citometria de fluxo [Agaian et al. 2016].

No intuito de auxiliar médicos no diagnóstico de doenças como a leucemia, são desenvolvidos Sistemas de Auxílio ao Diagnóstico (*Computer-aided Diagnosis - CAD*). Usualmente, esses sistemas utilizam técnicas de processamento de imagens e aprendizado de máquina para classificar imagens como saudáveis ou doentes, provendo assim um diagnóstico suplementar. Eles também podem auxiliar no atendimento de mutirões e na prevenção de fadiga por parte do especialista, o que pode gerar diagnósticos errôneos.

Neste trabalho, propomos um método composto pela utilização de uma Rede Neural Convolucional (*Convolutional Neural Network - CNN*) pré-treinada e técnicas de transferência de aprendizado e aumento de dados para diagnosticar leucemia através de imagens de lâminas de sangue. No desenvolvimento da abordagem, utilizamos a técnica de ajuste fino *Deeply Fine Tuning Modified (DFTM)*. Avaliamos o método em 2304 imagens pertencentes a 14 bases de dados públicas obtidas de diferentes hospitais e com características distintas de cor, contraste e textura. Os resultados obtidos são comparados com trabalhos existentes no estado da arte.

O restante do artigo está organizado da seguinte forma. Trabalhos relacionados são apresentados na Seção 2. O método proposto, as bases de dados utilizadas nos experimentos e a técnica de aumento de dados são apresentados na Seção 3. Na Seção 4 são apresentados os resultados obtidos, assim como a discussão entorno da performance obtida. Por fim, na Seção 5, apresentamos as conclusões e os trabalhos futuros.

2. Trabalhos Relacionados

Historicamente, os sistemas de diagnóstico de leucemia utilizaram características de cor, textura e forma. Nesse contexto, destacamos os trabalhos de [Patel and Mishra 2015] e [Singhal and Singh 2016]. A principal deficiência existente nesses trabalhos é o tamanho do conjunto de dados utilizados nos experimentos. Em [Patel and Mishra 2015], os autores combinaram as características de tradicionais com atributos estatísticos para posterior classificação com a *Support Vector Machine (SVM)* [Cortes and Vapnik 1995]. Ao todo, os autores utilizaram 27 imagens para avaliar a abordagem proposta. Enquanto em [Singhal and Singh 2016], o conjunto de dados apresentado foi de 260 imagens. A textura foi o principal descritor utilizado neste trabalho e o classificador foi o SVM. A acurácia obtida por ambos os trabalhos ainda é de 93,75% e 93,80%.

A utilização de abordagens que envolvem técnicas de aprendizado profundo vem se popularizando ao longo dos anos. Diante disso, surgiram aplicações em diversas subáreas do processamento de imagens e visão computacional [Ronneberger et al. 2015, Bejnordi et al. 2018]. Dentre diversas técnicas de aprendizado de máquina, as CNNs são consideradas o tipo de modelo mais bem sucedido. Suas camadas convolucionais dispostas de forma sequencial e profunda são capazes de abstrair dados em alto nível, funcionando como caixa preta, pois não se sabe claramente como é realizado seu aprendizado [Lei et al. 2018]. No entanto, considerando seus altos índices de acurácia, as CNNs possibilitam o desenvolvimento de soluções robustas e simplificadas para problemas conhecidos na literatura.

O trabalho proposto por Thanh et al. [Thanh et al. 2018], descreve uma arquitetura para o diagnóstico de leucemia. A CNN proposta possui sete camadas, cinco são

convolucionais e duas são totalmente conectadas. Os autores utilizaram o aumento de dados na base ALL-IDB 1 com 108 imagens. Foram aplicadas operações de rotação, translação, rotação, desfoque e equalização de histograma. Como resultado, o aumento de dados proporcionou 1188 a serem utilizadas nos experimentos. Os autores dividiram o conjunto em 70% para treinar a CNN e 30% para testes. A acurácia obtida foi de 96,60%.

Os autores em Shafique et al. [Shafique and Tehsin 2018] propuseram um modelo de CNN para diagnosticar diferentes subtipos de leucemia linfóide aguda. A arquitetura proposta foi pré-treinada e continha sete camadas, sendo quatro convolucionais e três totalmente conectadas. A base de dados utilizada nos experimentos foi a ALL-IDB 2 com 260. Os autores avaliaram entradas de diferentes sistemas de cor e aplicaram o aumento de dados para realizar o treino da CNN. Sendo assim, a base foi aumentada para 760 imagens. Os resultados apresentados foram significativos para o problema. A CNN proposta obteve 99,50% de acurácia para o diagnóstico entre imagens normais e doentes e 96,06% para o diagnóstico dos subtipos da leucemia.

Em Rehman et al. [Rehman et al. 2018], os autores propuseram uma abordagem utilizando CNNs na classificação dos diferentes tipos de leucemia linfóide aguda (ALL). A arquitetura AlexNet foi utilizada na etapa de classificação e sua camada de saída foi ajustada para classificar as 4 classes apresentadas no trabalho (L1, L2, L3 e normal). Para validar a abordagem, os autores compararam o resultado obtido com técnicas de classificação propostas na literatura. A técnica validação utilizada foi o *holdout*. Os autores também compararam seus resultados com trabalhos da literatura e observamos que apesar da acurácia de 97,78%, outros trabalhos superaram seus resultados.

Na metodologia proposta por Vogado et al. [Vogado et al. 2018], os autores utilizaram a técnica de transferência de aprendizado para extrair características da última camada totalmente conectada de três CNNs. Os autores analisaram as características extraídas e devido a alta dimensionalidade do vetor, realizaram a seleção de atributos utilizando a técnica razão de ganho. O tamanho final do vetor de características foi estabelecido de forma empírica. No entanto, os autores concluíram que, com 5% das características, a metodologia já alcança uma excelente acurácia. Para classificar o vetor extraído, foi utilizado o SVM com núcleo *Radial Basis Function* (RBF). Neste trabalho, a metodologia de validação empregada foi o *k-fold cross-validation*. Contudo, os autores não buscaram uma otimização do classificador para melhorar seus resultados, optando pela utilização da configuração padrão do SVM.

Na Tabela 1, listamos alguns trabalhos da literatura e seus resultados. Uma característica observada nesses trabalhos é a existência de apenas uma base de dados nos experimentos. Apenas no trabalho de Vogado et al. [Vogado et al. 2018], os autores utilizaram mais de uma base em seus experimentos. O uso de diferentes bases proporciona ao sistema maior confiabilidade ao lidar com a entrada de imagens com características distintas das apresentadas no treinamento.

3. Materiais e Métodos

Durante a construção do método proposto, utilizamos conjuntos de técnicas e operações que buscam maximizar a sua acurácia. As bases de dados utilizadas nos experimentos propostos foram obtidas de diferentes fontes e contam com uma alta taxa de heterogeneidade, tornando a tarefa de classificação mais desafiadora.

Tabela 1. Estado da arte.

Abordagem	Ano	Descritor	Classificador	Qt. de imagens	Acurácia(%)
[Patel and Mishra 2015]	2015	Forma, textura, cor e estatística	SVM	27	93,75
[Singhal and Singh 2016]	2016	Textura	SVM	260	93,80
[Thanh et al. 2018]	2018	Deep features	CNN	1188	96,60
[Rehman et al. 2018]	2018	Deep features	CNN	330	97,78
[Shafique and Tehsin 2018]	2018	Deep features	CNN	760	99,50
[Vogado et al. 2018]	2018	Deep features	SVM	1268	99,76

3.1. Método Proposto

O método proposto é composto pela utilização da CNN pré-treinada Vgg-19 [Simonyan and Zisserman 2014] e técnicas de ajuste fino e aumento de dados para diagnosticar leucemia. Na Figura 1, apresentamos a estrutura do método, incluindo as etapas de transferência de aprendizado e ajuste fino da arquitetura.

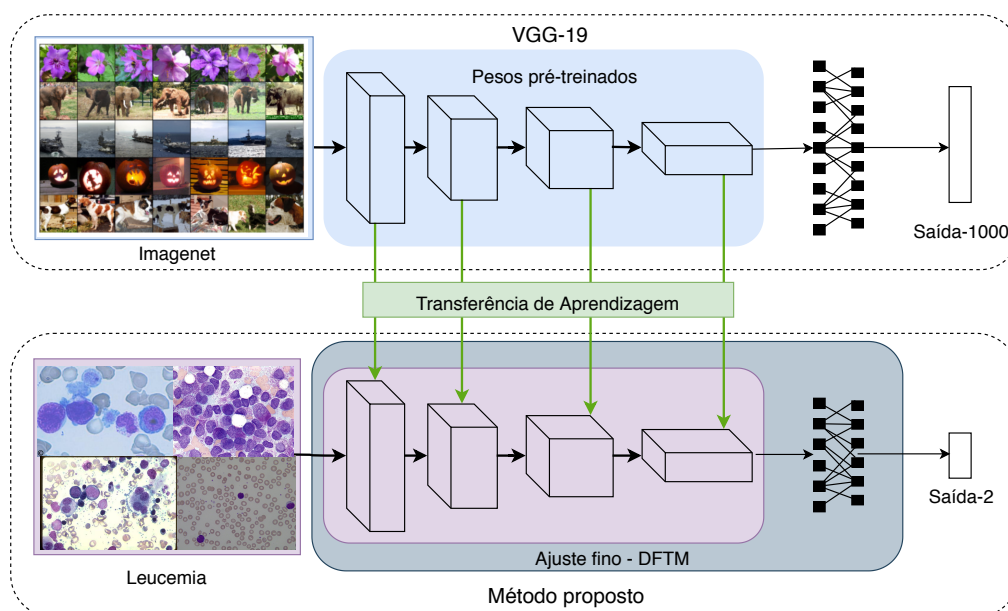


Figura 1. Fluxograma exemplificando a transferência de aprendizagem e ajuste fino utilizados no desenvolvimento do método proposto.

Durante o desenvolvimento do método proposto, avaliamos a utilização de diversas arquiteturas de CNNs e diferentes formas de realizar o ajuste fino. CNNs como a Vgg-16 [Simonyan and Zisserman 2014], InceptionV3 [Szegedy et al. 2016] e ResNet [He et al. 2016] foram consideradas na avaliação. No entanto, após inúmeros experimentos, constatamos que a Vgg-19 apresentou melhores resultados e por isso, foi selecionada dentre as CNNs. Dentre as técnicas de ajuste fino, utilizamos o DFTM. Esta técnica proporcionou o melhor comprometimento entre acurácia, complexidade do modelo e tempo de execução.

3.2. Transferência de Aprendizagem

As Redes Neurais Convolucionais são o modelo de rede de aprendizado profundo mais conhecidos e utilizados atualmente. O conceito de CNNs foi originado da grande capacidade de redes neurais com várias camadas realizarem abstrações de alto nível. Conseguindo assim, modelar um grande conjunto de dados através de transformações lineares

e não lineares. Consequentemente, amplificando a sua capacidade de aprendizado e proporcionando uma grande performance na classificação de dados [Bengio et al. 2013].

A utilização de CNNs apresenta dois desafios, o primeiro é a necessidade de grande poder computacional e o segundo é na obtenção de um largo conjunto de dados para treinar os milhares de parâmetros pertencentes a rede. Porém, ainda existem formas de se utilizar a capacidade da CNN. A principal é através do uso de pesos pré-treinados em bases de dados de propósito geral. Utilizando esses pesos é possível diminuir a necessidade de retrainar todos os parâmetros, assim como utilizar informações já existentes na rede pré-treinada [Yosinski et al. 2014]. Ao todo, existem duas formas de utilizar a capacidade de aprendizado das CNNs através da transferência de aprendizado. A primeira compreende a extração de características utilizando as camadas totalmente conectadas como vetores de características a serem submetidos aos classificadores estabelecidos no estado da arte como por exemplo a SVM.

A segunda forma de transferência de aprendizado é o ajuste fino ou *fine-tuning*. É uma técnica que exige grande poder computacional quando comparado com a técnica anterior, uma vez que você irá retrainar a CNN com a sua base de dados após realizar a transferência dos pesos pré-treinados, tornando-a mais específica ao problema abordado. De acordo com [Tajbakhsh et al. 2016] e [Izadyyazdanabadi et al. 2018], existem dois tipos de ajuste fino, o primeiro é denominado *Shallow Fine Tuning* (SFT), que consiste no congelamento das camadas do topo da CNN, geralmente os conjuntos de camadas convolucionais. As camadas do topo são consideradas mais gerais e guardam informações de forma, textura e cor, enquanto as camadas localizadas ao final da rede são mais específicas. Em tese, esse tipo de ajuste fino irá prover uma maior especialização nas últimas camadas.

O segundo ajuste fino é denominado *Deeply Fine Tuning* (DFT), no qual toda a rede é ajustada, desde as camadas do topo até a saída. Essa técnica exige mais processamento e uma grande quantidade de dados, mas garante um refinamento mais profundo com a base utilizada. Diante disso, a técnica consegue ser mais eficaz e está mais próxima a um treinamento completo da rede.

Em trabalhos da literatura, observamos que redes com menor número de camadas e parâmetros permitem uma melhor acurácia em bases de dados menores e problemas binários [Araujo et al. 2018]. Consequentemente, em nossos experimentos, desenvolvemos o que denominamos de *Shallow Fine Tuning - Modified* (SFTM) e *Deeply Fine Tuning - Modified* (DFTM). Nessas duas metodologias, nós substituímos as camadas totalmente conectadas com 4096 elementos por novas com menor dimensionalidade. Essas duas técnicas permitem a criação de um modelo mais enxuto e que tende a ser retrainado mais rapidamente. Além disso, em modelos menos complexos em termos de quantidade de parâmetros, estes tendem a sofrer menos com o *overfitting*. Na Figura 2, apresentamos exemplos do funcionamento de cada técnica de ajuste fino descrita neste trabalho.

Em [Zhang et al. 2017], os autores substituíram as camadas totalmente conectadas por novas com 1024 e 256 elementos respectivamente. Devido a semelhança das imagens de leucemia com imagens de câncer cervical apresentadas em seu trabalho, utilizamos o tamanho dessas camadas como ponto de partida na construção da abordagem proposta. Através de experimentos, constatamos que camadas totalmente conectadas com 1024 e

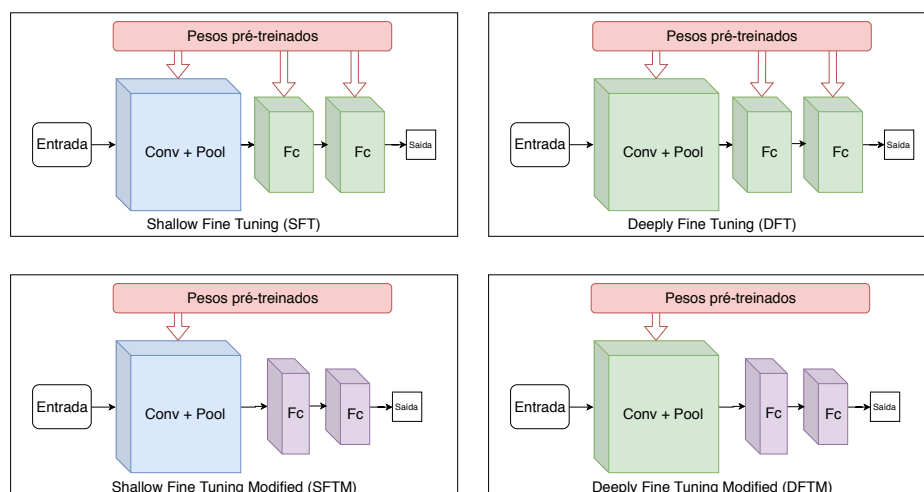


Figura 2. Ilustração simplificada de cada técnica de ajuste fino. A cor azul, predominante em algumas camadas, representa o congelamento dos parâmetros durante o treinamento, enquanto a cor verde representa que a camada é treinável. A cor roxa existe apenas nas técnicas SFTM e DFTM, uma vez que as camadas com essas cores não recebem a transferência de aprendizado, ou seja, seus parâmetros inicialmente são definidos de forma randômica.

512 apresentaram melhores resultados. Assim como, a inclusão do *dropout* entre as totalmente conectadas com valores de 0,5 e 0,6. A camada de classificação utilizada possui apenas 1 neurônio saída, sendo do tipo *sigmoid*, uma vez que possuímos um problema de classificação binária.

A partir de experimentos iniciais, definimos de forma empírica a configuração da rede, assim como, a sua taxa de aprendizado. Como observamos em trabalhos do estado da arte, para realizar um ajuste fino, o valor da taxa de aprendizado do ajuste deve ser menor que os valores geralmente utilizados ao se treinar uma CNN do zero. Portanto, utilizamos 0,001 e 0,8 para a taxa de aprendizado e o *momentum*, respectivamente. A função de perda ou *loss function* utilizada durante o refinamento foi a *binary crossentropy*. Essa função tem como principal objetivo quantificar a *loss* (ou perda) gerada pelos parâmetros treinados até o momento. Sendo assim, através dessa função, é possível analisar o quão distante a CNN está da predição ideal.

Para otimizar os parâmetros, de forma que o custo da CNN seja reduzido, utilizamos o algoritmo de otimização *Stochastic Gradient Descent* (SGD). Para treinar a CNN proposta, utilizamos 50 épocas e conjunto de treinamento de tamanho 32. Na execução de todos os experimentos, utilizamos um computador equipado com um processador Intel® Xeon™ @ 3.60GHz com 24GB de memória RAM e também uma Placa NVIDIA TITAN XP 12GB GDDR5X com 3840 Núcleos.

3.3. Bases de dados

Um dos desafios na criação de um sistema de auxílio médico é a habilidade de diagnosticar a doença em bases de dados com características distintas. Sendo assim, é necessário refinar a capacidade de generalização do sistema. Para avaliar e auxiliar no desenvolvimento do método proposto neste trabalho, utilizamos 14 bases de dados públicas originalmente

coletadas de diferentes hospitais (Tabela 2). As imagens possuem Constatamos que, dentre as bases obtidas, apenas 3 são balanceadas com imagens de ambas as classes, são elas: UFG, ALL-IDB1 e ALL-IDB2 [Labati et al. 2011].

Nas bases de dados apresentadas, observamos dois padrões de imagens, algumas possuem apenas um leucócito e outras possuem múltiplos leucócitos. No entanto, bases como a UFG² e Bloodline, possuem esses dois padrões. Apesar da existência desses padrões, as bases de dados possuem características distintas de cor, textura e contraste. As bases utilizadas possuem imagens com diferentes resoluções. Diante disso, realizamos o redimensionamento de todas as imagens para o tamanho 224x224, sendo esse o padrão da CNN utilizada. O sistema de cor utilizado no pré-treino da CNN foi o RGB, sendo assim, não alteramos os padrões de cores das imagens.

Tabela 2. Bases de dados utilizadas no trabalho e a quantidade de imagens

Databases	Saudáveis	Doentes	Total	Referência
ALL-IDB 1	59	49	108	[Labati et al. 2011]
ALL-IDB 1 (Crop)	0	510	510	[Labati et al. 2011]
ALL-IDB 2	130	130	260	[Labati et al. 2011]
Leukocytes	149	0	149	[Sarrafzadeh and Dehnavi 2015]
CellaVision	109	0	109	[Rollins-Raval et al. 2012]
Atlas	0	88	88	-
Omid et al. 2014	154	0	154	[Omid Sarrafzadeh 2014]
Omid et al. 2015	0	27	27	[Sarrafzadeh et al. 2015]
ASH	0	96	96	[Madhukar et al. 2014]
UFG	57	64	121	-
Bloodline	0	204	204	[Vale et al. 2014]
ONKODIN	0	78	78	[Böhm 2008]
CellaVision 2	100	0	100	[Zheng et al. 2018]
JTSC	300	0	300	[Zheng et al. 2018]
Total	1058	1246	2304	-

3.4. Aumento de Dados

CNNs que possuem milhões de parâmetros necessitam de uma grande quantidade de dados para serem treinadas. Consequentemente, até mesmo para refinar uma CNN de pequeno porte, milhares de imagens são necessárias. Diante deste desafio, trabalhos propostos na literatura utilizaram a técnica de aumento de dados com o intuito de solucionar este problema [Krizhevsky et al. 2012]. Esta técnica consiste na criação de um novo conjunto de imagens utilizando operações de rotação, translação, iluminação e, até mesmo, a combinação dessas técnicas. Consequentemente, são utilizadas variações em cada uma dessas técnicas, permitindo a multiplicação do conjunto de dados em até 20 vezes. O aumento de dados tem como principais objetivos a redução no *overfitting* da rede assim como o aumento da acurácia durante a etapa de treinamento [Perez and Wang 2017].

Neste trabalho, durante a etapa de refinamento da CNN, utilizamos a técnica de aumento de dados randômica provida pela API Keras. Uma vez que a base de dados possui classes praticamente balanceadas, o aumento foi aplicado igualmente no conjunto de treino. O intervalo de rotação escolhido foi de 40°, já o intervalo de translação vertical, horizontal, cisalhamento e zoom foi de 0,2. Também utilizamos a inversão na horizontal e na vertical, visto que as imagens de núcleo não possuem assimetria. A técnica de preenchimento por reflexão foi aplicada para substituir *pixels* em preto resultantes das técnicas

²<https://hematologia.farmacia.ufg.br>

de rotação e de translação. Por fim, realizamos a normalização dos *pixels* da imagem de entrada para 0 e 1.

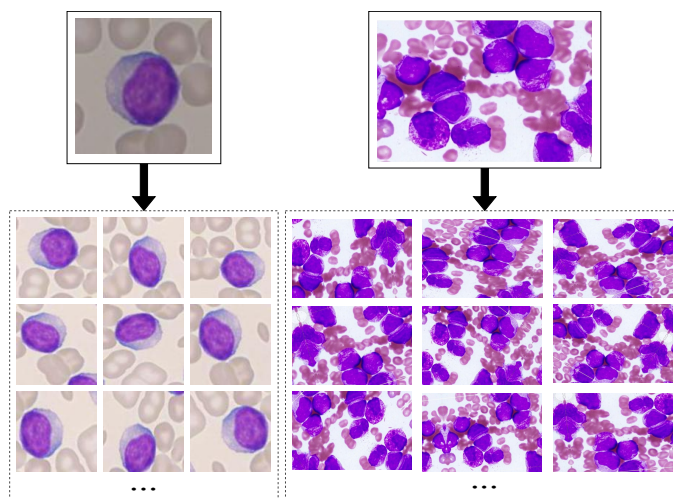


Figura 3. Exemplos da aplicação das operações de aumento de dados em imagens com a presença (direita) e sem a presença (esquerda) de leucemia.

4. Resultados e Discussão

A partir das técnicas de ajuste fino apresentadas, realizamos experimentos com o objetivo de estabelecer a melhor abordagem a ser utilizada para o problema apresentado. Dentre as métricas calculadas durante os experimentos, consideramos a acurácia e a *loss*. O método de validação utilizado neste trabalho é o *k-fold cross validation*, com o valor de *k* igual a 5. Através de avaliação empírica, efetuamos experimentos em diferentes arquiteturas, no entanto, iremos apresentar apenas a Vgg-19, pois esta apresentou os melhores resultados.

Os resultados da Tabela 3 representam os experimentos executados com a arquitetura Vgg-19 nas abordagens de refinamento. A partir de seus resultados, observamos que as técnicas DFT e DFTM obtiveram os melhores valores de acurácia com respectivamente 98,82% e 98,84%. No entanto, apesar de resultados próximos, a DFTM foi 10 segundos mais rápida que a DFT e a consideramos uma técnica mais eficiente, visto que proporciona uma diminuição na complexidade da arquitetura através da redução na quantidade de parâmetros treináveis.

Tabela 3. Comparação entre as técnicas de ajuste fino.

Técnica de ajuste fino	Parâmetros treináveis	Tempo de treino (s)	loss	Acurácia
SFT	119.549.953	1016,99	0,1714	93,86% ± 0,51
SFTM	26.216.449	1017,07	0,2139	92,54% ± 0,73
DFT	139.574.337	1073,40	0,0372	98,82% ± 0,32
DFTM	46.240.833	1064,54	0,0472	98,84% ± 0,46

A partir do resultado obtido pelo método proposto, realizamos uma comparação com trabalhos do estado da arte na Tabela 4. Os valores de acurácia apresentados nessa tabela foram obtidos nos trabalhos originais.

Primeiramente, observamos que a quantidade de imagens utilizadas em todos os trabalhos é inferior a apresentada em nossos experimentos. Em [Patel and Mishra 2015]

Tabela 4. Comparação entre o método proposto e os trabalhos do estado da arte.

Abordagens	Qt de Imagens	Acurácia
[Patel and Mishra 2015]	27	93,75%
[Singhal and Singh 2016]	260	93,80%
[Thanh et al. 2018]	1188	96,60%
[Rehman et al. 2018]	330	97,78%
[Shafique and Tehsin 2018]	760	99,50%
[Vogado et al. 2018]	1268	99,76%
Método Proposto	2304	98,84%

e [Singhal and Singh 2016], os autores utilizaram técnicas de extração de características baseadas em textura, forma e cor. Consequentemente, partindo dos resultados, os métodos obtiveram 93,75% e 93,80%, respectivamente. A pequena quantidade de imagens utilizadas nos experimentos desses trabalhos atestam que apesar de utilizarem apenas a mesma base, falta robustez caso imagens de bases novas fossem utilizadas como validação.

Os trabalhos de [Thanh et al. 2018], [Rehman et al. 2018], [Shafique and Tehsin 2018] e [Vogado et al. 2018], utilizaram técnicas baseadas em CNNs para a extração de características e classificação. Dentre esses métodos, [Vogado et al. 2018] apresentou a melhor acurácia e a maior base de dados dentre os trabalhos do estado da arte. Ao compararmos os resultados do método proposto com os demais, observamos que a acurácia da abordagem proposta em [Vogado et al. 2018] foi superior. Já em [Thanh et al. 2018], [Rehman et al. 2018] e [Shafique and Tehsin 2018], os resultados apresentados foram excelentes, no entanto, as imagens utilizadas nos experimentos possuem apenas uma base de dados, enquanto o método de [Vogado et al. 2018] e o método proposto possuem diversas bases de dados.

A transferência de aprendizagem utilizada pelos autores em [Vogado et al. 2018] consistiu na extração de características utilizando a última camada totalmente conectada. A partir dos experimentos realizados, os autores afirmaram que com pelo menos 5% das características selecionadas a partir da razão de ganho, o classificador SVM apresenta resultados acima de 99%.

Considerando que os melhores resultados foram obtidos pela abordagem proposta em [Vogado et al. 2018], utilizamos o código original dessa abordagem para realizar uma comparação com o método proposto no mesmo conjunto de dados apresentado neste trabalho. Os resultados obtidos são ilustrados na Tabela 5.

Tabela 5. Comparação entre a abordagem proposta e o trabalho proposto em [Vogado et al. 2018] com o mesmo conjunto de dados.

Abordagem	Acurácia
[Vogado et al. 2018]	92,70%
Método Proposto	98,84%

Podemos observar na Tabela 5, que os resultados obtidos por [Vogado et al. 2018] com as mesmas bases de dados apresentadas não são superiores ao método proposto neste trabalho. Tendo obtido apenas 92,70%, enquanto o proposto alcançou 98,84%. Observamos que o trabalho de [Vogado et al. 2018] utilizou algumas das bases de dados apresentadas neste trabalho. Ainda sim, percebe-se uma redução na acurácia da abordagem

devido a existência de bases de dados provenientes de atlas (ASH, UFG, Bloodline e ONKODIN), ou seja, apresentam imagens de diversos microscópios diferentes com resoluções, texturas e características de cor heterogêneas.

5. Conclusão

Neste trabalho, apresentamos um método para o diagnóstico de leucemia em imagens lâminas de sangue. Através de técnicas de transferência de aprendizado e ajuste fino, foi possível desenvolver um método que possibilite esse diagnóstico de forma precisa e robusta. Observamos que, dentre duas abordagens de transferência de aprendizado, o ajuste fino é considerado superior a extração de características.

Os resultados obtidos apresentaram valores acima de 98% e quando comparados com trabalhos do estado da arte, se mostraram superiores. Uma das principais vantagens do método proposto é utilizar a capacidade da própria CNN para classificar imagens de diversas bases de dados heterogêneas. Diante disso, podemos atestar este método como possível solução para um sistema de auxílio médico, uma vez que abrange o maior número de imagens possíveis.

Para trabalhos futuros utilizaremos outro método de validação com o intuito de avaliarmos o experimento com bases completas e não de um conjunto de dados aleatório. Isso tem por objetivo, assemelhar o sistema de auxílio com sistemas do mundo real, onde imagens de diferentes lugares são avaliadas em um mesmo sistema. Além disso, procuraremos investigar se a CNN refinada gerada como resultado do método proposto neste trabalho é capaz de extrair características mais representativas que os modelos pré-treinados na Imagenet.

Referências

- [Agaian et al. 2016] Agaian, S., Madhukar, M., and Chronopoulos, A. T. (2016). A new acute leukaemia-automated classification system. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 0(0):1–12.
- [Araujo et al. 2018] Araujo, F. H., Silva, R. R., Medeiros, F. N., Parkinson, D. D., Hexemer, A., Carneiro, C. M., and Ushizima, D. M. (2018). Reverse image search for scientific data within and beyond the visible spectrum. *Expert Systems with Applications*, 109:35 – 48.
- [Bejnordi et al. 2018] Bejnordi, B. E., Mullooly, M., Pfeiffer, R. M., Fan, S., Vacek, P. M., Weaver, D. L., Herschorn, S., Brinton, L. A., van Ginneken, B., Karssemeijer, N., Beck, A. H., Gierach, G. L., van der Laak, J. A. W. M., and Sherman, M. E. (2018). Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Modern Pathology*, 31:1502–1512.
- [Bengio et al. 2013] Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1798–1828.
- [Böhm 2008] Böhm, J. (2008). Pathologie-websites im world wide web. *Der Pathologe*, 29(3):231–242.
- [Cortes and Vapnik 1995] Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

- [He et al. 2016] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778. IEEE Computer Society.
- [Izadyyazdanabadi et al. 2018] Izadyyazdanabadi, M., Belykh, E., Mooney, M., Martirosyan, N., Eschbacher, J., Nakaji, P., Preul, M., and Yang, Y. (2018). Convolutional neural networks: Ensemble modeling, fine-tuning and unsupervised semantic localization for neurosurgical cle images. *Journal of Visual Communication and Image Representation*, 54:10–20.
- [Krizhevsky et al. 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc.
- [Labati et al. 2011] Labati, R. D., Piuri, V., and Scotti, F. (2011). All-idb: The acute lymphoblastic leukemia image database for image processing. In *18th IEEE International Conference on Image Processing (ICIP)*, pages 2045–2048.
- [Lei et al. 2018] Lei, D., Chen, X., and Zhao, J. (2018). Opening the black box of deep learning. *CoRR*, abs/1805.08355.
- [Madhukar et al. 2014] Madhukar, M., Agaian, S., and Chronopoulos, A. T. (2014). Automated screening system for acute myelogenous leukemia detection in blood microscopic images. *IEEE Systems Journal*, 8(3).
- [Omid Sarrafzadeh 2014] Omid Sarrafzadeh, Hossein Rabbani, A. T. H. U. B. (2014). Selection of the best features for leukocytes classification in blood smear microscopic images. In *Proc. SPIE*, volume 9041, pages 9041 – 9041 – 8.
- [Patel and Mishra 2015] Patel, N. and Mishra, A. (2015). Automated leukaemia detection using microscopic images. *Procedia Computer Science*, 58:635–642.
- [Perez and Wang 2017] Perez, L. and Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *CoRR*, abs/1712.04621.
- [Rehman et al. 2018] Rehman, A., Abbas, N., Saba, T., ur Rahman, S. I., Mehmood, Z., and Kolivand, H. (2018). Classification of acute lymphoblastic leukemia using deep learning. *Microscopy Research and Technique*, pages 1–8.
- [Rollins-Raval et al. 2012] Rollins-Raval, M., Raval, J., and Contis, L. (2012). Experience with cellvision dm96 for peripheral blood differentials in a large multi-center academic hospital system. *Journal of Pathology Informatics*, 3(29):1–9.
- [Ronneberger et al. 2015] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham. Springer International Publishing.
- [Sarrafzadeh and Dehnavi 2015] Sarrafzadeh, O. and Dehnavi, A. M. (2015). Nucleus and cytoplasm segmentation in microscopic images using k means clustering and region growing. *Advanced Biomedical Research*, pages 79–87.
- [Sarrafzadeh et al. 2015] Sarrafzadeh, O., Rabbani, H., Dehnavi, A. M., and Talebi, A. (2015). Detecting different sub-types of acute myelogenous leukemia using dictionary learning and sparse representation. In *ICIP*, pages 3339–3343. IEEE.

- [Shafique and Tehsin 2018] Shafique, S. and Tehsin, S. (2018). Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks. *Technology in Cancer Research Treatment*, 17:1–7.
- [Simonyan and Zisserman 2014] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556.
- [Singhal and Singh 2016] Singhal, V. and Singh, P. (2016). *Texture Features for the Detection of Acute Lymphoblastic Leukemia*, pages 535–543. Springer Singapore, Singapore.
- [Szegedy et al. 2016] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 2818–2826.
- [Tajbakhsh et al. 2016] Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., and Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, 35:1299–1312.
- [Thanh et al. 2018] Thanh, T. T. P., Vununu, C., Atoev, S., Lee, S.-H., and Kwon, K.-R. (2018). Leukemia blood cell image classification using convolutional neural network. *International Journal of Computer Theory and Engineering*, 10(2):54–58.
- [Vale et al. 2014] Vale, A. M. P. G., Guerreiro, A. M. G., Neto, A. D. D., Cavalvanti Junior, G. B., de Sá Leitão, V. C. L. T., and Martins, A. M. (2014). Automatic segmentation and classification of blood components in microscopic images using a fuzzy approach. *Revista Brasileira de Engenharia Biomédica*, 30:341 – 354.
- [Vogado et al. 2018] Vogado, L. H. S., Veras, R. M. S., Araújo, F. H. D., e Silva, R. R. V., and Aires, K. R. T. (2018). Leukemia diagnosis in blood slides using transfer learning in cnns and SVM for classification. *Engineering Applications of Artificial Intelligence*, 72:415–422.
- [Yosinski et al. 2014] Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, NIPS'14*, pages 3320–3328, Cambridge, MA, USA.
- [Zhang et al. 2017] Zhang, L., Lu, L., Noguees, I., Summers, R. M., Liu, S., and Yao, J. (2017). Deeppap: Deep convolutional networks for cervical cell classification. *IEEE Journal of Biomedical and Health Informatics*, 21(6):1633–1643.
- [Zheng et al. 2018] Zheng, X., Wang, Y., Wang, G., and Chen, Z. (2018). Fast and robust segmentation of white blood cell images by self-supervised learning. *Micron*, 107:55–71.