

Detecção de Leishmaniose Visceral em Humanos Utilizando Redes Neurais

Viviane Barbosa Leal Dias¹, Armando Luz Borges¹, Clésio de Araujo Gonçalves^{2,3}, Ana Carolina Landim Pacheco⁴, Romuere Rodrigues Veloso e Silva^{1,2}

¹Sistemas de Informação, Universidade Federal do Piauí, Picos, Brasil

²Ciência da Computação, Universidade Federal do Piauí, Picos, Brasil

³Departamento de Informática, Instituto Federal do Sertão Pernambucano, Ouricuri, Brasil

⁴Ciências Biológicas, Universidade Federal do Piauí, Picos, Brasil

Abstract. *Leishmaniasis is a serious disease transmitted by infected mosquitoes, which can be fatal without adequate treatment. A quick and efficient diagnosis is essential. Although Computer Vision techniques can help with diagnosis, they are often expensive and time-consuming due to high computational requirements. This study aims to train low computational cost convolutional neural networks to assist in the diagnosis of Visceral Leishmaniasis. Our results were compared with four previous works, and our method showed significant and promising results. This demonstrates that low-cost convolutional neural networks can be an effective approach for the automated diagnosis of visceral leishmaniasis in humans.*

Resumo. *A leishmaniose é uma doença grave transmitida por mosquitos infectados, que pode ser fatal sem tratamento adequado. Um diagnóstico rápido e eficiente é essencial. Embora as técnicas de Visão Computacional possam ajudar no diagnóstico, elas costumam ser caras e demoradas devido aos altos requisitos computacionais. Este estudo tem como objetivo treinar redes neurais convolucionais de baixo custo computacional para auxiliar no diagnóstico da Leishmaniose Visceral. Nossos resultados foram comparados com quatro trabalhos anteriores, e nosso método mostrou resultados significativos e promissores. Isto demonstra que redes neurais convolucionais de baixo custo podem ser uma abordagem eficaz para o diagnóstico automatizado de leishmaniose visceral em humanos.*

1. Introdução

A Leishmaniose é uma doença parasitária composta por diversas formas clínicas que podem variar tanto nos sintomas quanto na intensidade da doença. Essas doenças podem ser causadas por parasitas do gênero *Leishmania*, que são transmitidos aos seres humanos

através da picada de flebotomíneos fêmeas infectados, pequenos insetos encontrados em regiões tropicais e subtropicais do mundo. De acordo com [Gonçalves et al. 2022] a manifestação mais severa da doença é a Leishmaniose Visceral (LV) e pode ser letal nos casos em que não é tratada corretamente.

Segundo [Borges et al. 2023] a LV apresenta evolução crônica, e pode acarretar em comprometimento dos órgãos internos e do sistema imunológico. [Kumar and Nylén 2012] informam que os sintomas comuns são febre de longa duração, astenia, emagrecimento, anemia, aumento do fígado e do baço, além de uma diminuição acentuada no número de plaquetas sanguíneas do paciente. Embora esses sintomas possam demorar semanas ou até meses para realizarem manifestação, existe a possibilidade de agravamento da doença, então surge a necessidade de diagnóstico na fase inicial.

O diagnóstico da LV é realizado através de avaliações microscópicas na biópsia de medula óssea do paciente. Conforme [Silva et al. 2014] o exame parasitológico visa identificar a presença de amastigotas e é o método mais confiável de identificação da doença. Realizar a análise manual dos exames pode ser exaustivo e demandar muito esforço dos profissionais da saúde [e Silva et al. 2016], principalmente quando se tem uma grande quantidade de dados para avaliar.

Tendo em vista os fatos citados anteriormente, e considerando a importância de identificar a doença em estágios iniciais, é essencial o uso de métodos acessíveis e eficazes para a realização de diagnóstico da LV. Em virtude disso, este trabalho visa aplicar técnicas de Visão Computacional (VC) em conjunto com o uso de Redes Neurais Convolucionais (CNNs) que demandem baixo custo computacional, ou seja, que não exija um grande investimento em termos de poder de processamento e custos. Essa abordagem visa realizar a classificação automática e contribuir para o diagnóstico da LV de forma mais acessível e eficiente em termos de recursos financeiros e computacionais.

O restante do artigo segue a seguinte estrutura: A seção 2 apresenta os trabalhos e estudos relacionados encontrados na literatura. A seção 3 apresenta a metodologia desenvolvida para a condução do trabalho. A seção 4 apresenta os resultados obtidos com o desenvolvimento e aplicação da metodologia proposta. Por fim, a seção 5 apresenta breves conclusões a respeito de todo o estudo realizado.

2. Trabalhos Relacionados

Nesta seção serão apresentados e discutidos trabalhos relacionados a classificação automática de LV através de técnicas de processamento de imagens. Irão ser abordados os objetivos primordiais de cada trabalho, as técnicas utilizadas, e os resultados de cada projeto. A Tabela 1 apresenta os trabalhos relacionados que foram escolhidos através de três principais critérios: (1) Esse trabalho se trata de LV? (2) Esse trabalho utiliza técnicas para classificação de imagens? (3) Os métodos utilizados funcionam mesmo com baixo custo computacional? Com base nas respostas, os trabalhos selecionados foram os seguintes:

Para auxiliar no diagnóstico de LV, o estudo conduzido por [Borges et al. 2023] desenvolveu um sistema que extrai informações de imagens captadas no exame parasitológico de pacientes. Duas técnicas de classificação de imagens foram empregadas: uma em que as redes CNN pré-treinadas seguiram o tamanho de entrada padrão de cada

Table 1. Abordagem utilizada para levantamento dos trabalhos relacionados

Trabalho	Critério 1	Critério 2	Critério 3
[Borges et al. 2023]	Sim	Sim	Não
[Gonçalves et al. 2022]	Sim	Sim	Não
[Górriz et al. 2018]	Sim	Não	Sim
[Isaza-Jaimes et al. 2020]	Sim	Não	Sim
Este Trabalho	Sim	Sim	Sim

rede, e outra em que as mesmas redes foram redimensionadas para o tamanho de 75x75. Além disso, foram utilizados diferentes formatos de cor na etapa de pré-processamento para cada teste. Quatro CNNs foram avaliadas: Densenet201, VGG16, ResNet152V2 e InceptionResnetV2. A melhor acurácia foi de 99,2%, obtida pelo modelo Densenet201 com tamanho de entrada padrão e com a base de dados em LAB, é importante ressaltar que o objetivo principal do estudo foi a classificação das imagens, e que as CNNs utilizadas exigem um custo computacional considerável de *hardware*.

[Gonçalves et al. 2022] propõe realizar a classificação automática de LV. Para isso, foi utilizada uma base de dados que incluía 151 imagens, sendo 79 delas positivas e 72 negativas. Foram desenvolvidos dois métodos: o primeiro, que consistiu em uma CNN baseada na arquitetura LeNet, obtendo uma acurácia máxima de 78,7% e um recall máximo de 64%; e o segundo, que utilizou um extrator de características através de diversas CNNs pré-treinadas, incluindo a InceptionResNetV2, InceptionV3, NASNet-Large, ResNet152V2 e Xception, tendo a NASNetLarge alcançado o melhor resultado, com recall de 75% e acurácia de 64%. Mesmo com resultados encorajadores, os autores utilizaram CNNs que precisam de um *hardware* mais exigente em termos de recursos.

A técnica utilizada por [Górriz et al. 2018] propõe fazer uso da arquitetura U-net para a identificação de parasitas *Leishmania* em imagens microscópicas, com o objetivo de classificá-los em amastigotas, promastigotas ou parasitas aderentes. Para isso, foi utilizada durante o experimento uma base de dados composta por 45 imagens, sendo 37 para treinamento e 8 para teste. O F1-Score obtido foi de 77,7%, o que representa um resultado relevante. No entanto, a base de dados apresenta um número limitado de amostras.

O estudo conduzido por [Isaza-Jaimes et al. 2020] descreveu uma abordagem computacional para a realização da detecção automática de *Leishmania*, a técnica escolhida pelos autores propõe utilizar um conjunto de dados aberto composto por somente 45 imagens. A metodologia proposta alcançou uma taxa de reconhecimento de cerca de 80% dos parasitas, o que é um valor considerável, entretanto esse trabalho não utilizou durante sua metodologia técnicas de aprendizado profundo para a identificação dos parasitas.

Ao analisar o levantamento do estado da arte, pode-se concluir que existem lacunas nos trabalhos mostrados. Foram apresentados trabalhos que realizavam apenas a segmentação de LV, e trabalhos que até realizavam a classificação utilizando técnicas de VC, mas que não utilizavam técnicas de aprendizado profundo, ou que necessitam de um elevado consumo de investimentos computacionais para funcionar. Levando em consideração as lacunas existentes na revisão do estado da arte, existe a necessidade

de propor neste trabalho um método capaz de classificar automaticamente LV utilizando CNNs que não precisem de um custo computacional muito elevado para ser eficaz.

3. Metodologia

Esta seção descreve a metodologia usada para eficientemente classificar LVs, mesmo com CNNs de baixo custo. As etapas incluem (1) aquisição de dados, (2) pré-processamento, (3) aumento de dados e (4) classificação, conforme ilustrado na Figura 1.

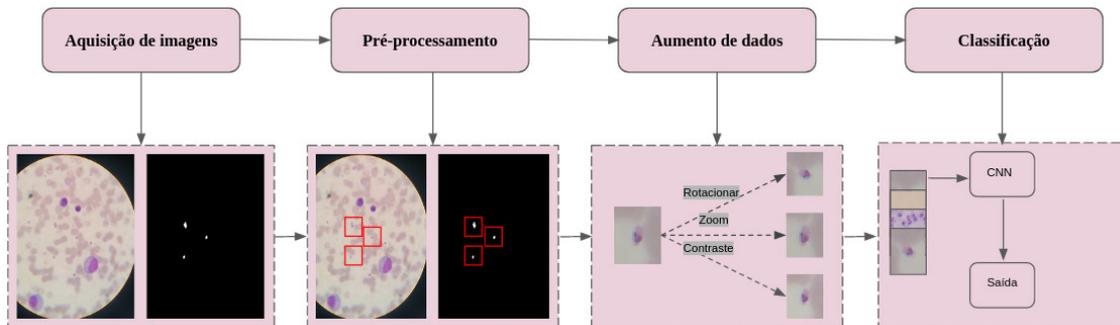


Figure 1. Fluxograma da metodologia utilizada para classificação de LV.

3.1. Base de Dados

A captura da base de imagens utilizada neste estudo foi realizada em conjunto com o Instituto de Doenças Tropicais Natan Portella e o Centro de Inteligência em Agravos Tropicais Emergentes e Negligenciados (CIATEN), utilizando o microscópio Olympus Cba em conjunto com uma câmera digital. O conjunto adquirido é composto por 150 imagens, das quais 72 são negativas, ou seja, não possuem amastigotas, e as outras 78 são positivas, apresentando a presença desses parasitas. Importante ressaltar que as imagens não possuem uma padronização nas suas dimensões, isso torna necessário o uso de pré-processamento das imagens antes da etapa de classificação.

3.2. Pré-processamento

Durante a etapa de pré-processamento, foram utilizadas as máscaras binárias da base positiva de dados para realizar a identificação dos eixos de cada região de interesse. A busca foi realizada utilizando a coordenada central de cada amastigota, e para padronizar o tamanho das imagens, através dessa coordenada foi obtida uma máscara 75×75 de cada parasita presente nas imagens originais como mostra a Figura 1. Com todos os parasitas devidamente recortados, a nova base de imagens positiva resultou em 559 imagens. Para balanceamento dos dados, foram realizados 559 recortes aleatórios de tamanho 75×75 na base de dados negativa, de forma que a base de dados ficou com um total de 1118 imagens. A divisão das imagens foi realizada da seguinte forma: 60% para treino, 10% para validação e 30% para teste.

3.3. Aumento de dados

Durante a metodologia, foram utilizadas técnicas de ampliação de dados para aprimorar o treinamento da CNN. As técnicas utilizadas consistem em rotação aleatória vertical e

horizontal, zoom aleatório numa escala de 0% até 10%, e contraste aleatório por um fator de até 0,2. Essa metodologia foi utilizada a fim de expandir a diversidade dos dados de treinamento, já que as amastigotas não contém padrões definidos de formato ou tamanho.

3.4. Classificação de imagens

Neste trabalho foram utilizadas CNNs como ferramentas para extrair atributos e capacitar com precisão a classificação de imagens em relação a presença de LV. As CNNs foram escolhidas através da ausência de trabalhos que utilizem métodos acessíveis em termos computacionais para classificação de LV.

Durante a metodologia deste trabalho, foram utilizadas quatro modelos pré-treinados de CNNs: MobileNet (MN)[Howard et al. 2017]; MobileNetV2 (MNV2) [Sandler et al. 2019]; ResNet50 (RN50) [He et al. 2015]; e ResNet50V2 (RNV2) [He et al. 2016]. Esses métodos foram escolhidos devido à eficiência de treinamento com custos computacionais mais baixos e tempos de inferência curtos, como é mostrado na Tabela 2. Outro método utilizado foi a transferência de aprendizado, essa técnica foi utilizada em decorrência do número pequeno de imagens na base de dados e da significativa otimização no processo de treino das redes.

Table 2. Parâmetros das CNNs utilizadas neste trabalho.

Modelo	Tamanho (MB)	Parâmetros	Tempo (ms) por etapa de inferência (CPU)	Tempo (ms) por etapa de inferência (GPU)
MobileNet	16	4,3M	22,6	3.4
MobileNetV2	14	3,5M	25,9	3.8
ResNet50	98	25,6M	58,2	4.6
ResNet50V2	98	25,6M	45,6	4.4

Durante o treinamento das redes, foi utilizada a estratégia de congelar as camadas convolucionais a fim de preservar os extratores de características já previamente estabelecidos por meio da etapa de transferência de aprendizagem. Essa abordagem visa evitar possíveis modificações que poderiam comprometer a qualidade dos resultados obtidos.

A etapa utilizada após a classificação foi o *fine tuning*. Essa técnica realiza o descongelamento de todas as camadas convolucionais para que o modelo possa ser integralmente treinado com o conjunto de dados específico mencionado anteriormente. O uso dessa abordagem tem a capacidade de aprimorar significativamente o desempenho da rede, à medida que os extratores de recursos se adaptam aos padrões do conjunto de dados, fortalecendo o seu poder de generalização.

Os parâmetros utilizados para o treinamento da rede têm impacto direto em sua performance. Neste estudo, foram realizadas 40 épocas de treinamento, direcionando tal esforço exclusivamente à camada de classificação e adotando uma taxa de aprendizado de 0,000001. O uso de uma taxa de aprendizado reduzida se justifica pela perspectiva de conferir maior acurácia ao processo de ajuste dos pesos. Na fase de *fine tuning*, por sua vez, optamos por parâmetros ainda mais específicos, incluindo uma taxa de aprendizado ainda menor, dividindo o valor anterior por 5.

3.5. Avaliação de desempenho

Os resultados foram avaliados através do uso das seguintes métricas populares para a classificação de imagens: Perda, Acurácia, Precisão, Recall, F1Score e Kappa [Fleiss et al. 2013]. Essas métricas serão brevemente apresentadas, a seguir.

A perda é uma métrica responsável por avaliar a diferença entre as previsões feitas pelo modelo e os valores reais. A acurácia é utilizada para avaliar desempenho de classificação, e pode ser definida como a proporção de previsões corretas feitas pelo modelo em relação ao total de previsões. Precisão é a proporção de previsões positivas corretas em relação ao total de previsões positivas feitas pelo modelo. Recall pode ser definido como a taxa de previsões positivas corretas em relação ao total de amostras positivas reais. Já o F1Score representa a média harmônica entre a precisão e o recall, e leva em conta tanto as previsões positivas corretas quanto a capacidade do modelo em identificar todas as amostras positivas. O Kappa avalia a concordância entre as predições do modelo e os valores verdadeiros, considerando a probabilidade de acerto ao acaso.

4. Resultados

Neste trabalho, o valor de entrada utilizada em cada CNN foi de 224×224 , todas as imagens estavam em formato RGB. A Tabela 3 apresenta a média dos resultados obtidos com a base de teste antes da aplicação do *fine tuning*, e a Tabela 4 apresenta a média dos resultados obtidos com a mesma base após o uso de *fine tuning*.

Table 3. Métricas obtidas através das CNNs.

CNN	Perda	Acc	Precisão	Recall	F1Score	Kappa
MN	0,080	0,984	0,984	0,984	0,984	0,968
MNV2	0,065	0,986	0,989	0,984	0,986	0,973
RN50	0,106	0,970	0,968	0,973	0,970	0,941
RNV2	0,033	0,994	0,994	0,994	0,994	0,989

Table 4. Métricas obtidas com uso de fine tuning.

CNN	Perda	Acc	Precisão	Recall	F1-Score	Kappa
MN	0,076	0,984	0,984	0,984	0,984	0,968
MNV2	0,064	0,986	0,989	0,984	0,986	0,973
RN50	0,129	0,954	0,972	0,936	0,953	0,909
RN50V2	0,034	0,994	0,994	0,994	0,994	0,989

Ao realizar uma análise nas tabelas apresentadas, é perceptível que o melhor resultado obtido foi utilizando a RN50V2. É notável também que na maioria das arquiteturas, o uso de *fine tuning* não influenciou tanto nos resultados. Na MN e na MNV2, houve uma leve queda na perda após o *fine tuning*, porém as demais métricas obtiveram o mesmo resultado. Na RN50 a taxa de perda aumentou, e as outras métricas caíram consideravelmente. A RN50V2 trouxe uma pequena queda na perda, mas as demais métricas não obtiveram mudanças nos resultados.

Com relação as arquiteturas CNNs, a RN50V2 se mostrou mais confiável que as demais. Com o valor de perda equivalente a 3% mostrou eficiência com acurácia, precisão, recall e F1Score superiores a 99%. O valor do Kappa também foi significativo, sendo este superior a 98%. Através desses resultados, é possível afirmar que métodos que utilizam baixo custo computacional podem obter resultados consideráveis, mesmo utilizando pouco poder de processamento de *hardware*.

Table 5. Comparação das métricas deste trabalho com o estado da arte.

Métodos	Acc	Precisão	Recall	F1-Score	Kappa
[Borges et al. 2023]	0,992	1,00	0,984	0,992	0,984
[Gonçalves et al. 2022]	0,787	0,941	0,640	0,761	0,581
[Górriz et al. 2018]	-	0,757	0,823	0,777	-
[Isaza-Jaimes et al. 2020]	-	0,787	-	-	-
Este Trabalho	0,994	0,994	0,994	0,994	0,989

Os resultados obtidos destacaram em relação a estudos anteriores. As métricas encontradas superaram quase todas as métricas do estudo mais próximo [Borges et al. 2023], evidenciando nossa eficácia na classificação. A abordagem utilizada demonstrou uma capacidade notável de identificar com precisão casos positivos e negativos. Em suma, os resultados deste trabalho representam um avanço significativo em relação a estudos anteriores, com perspectivas promissoras em aplicações médicas e de VC.

5. Conclusão

É possível concluir que o uso de CNNs de baixo custo computacional é um método eficaz e confiável para a classificação automática de LV. O trabalho proposto obteve um resultado superior aos trabalhos relacionados em termos de custo computacional, e mesmo utilizando CNNs que conseguem funcionar com menos investimento computacional, ainda alcançou resultados superiores em quase todas as métricas dos trabalhos da literatura.

O método proposto pode ser tão confiável, e mais acessível que os citados da literatura, pois possibilita gerar resultados sobre classificação de LV utilizando menos custo computacional que os trabalhos do estado da arte levariam, além disso, com o uso do método proposto o investimento em custo de recursos seria menor. Logo, pode-se afirmar que este trabalho cumpriu com seu objetivo, e consegue ser eficiente em termos de tempo, processamento, recursos e desempenho.

Como trabalhos futuros, pretendemos automatizar a detecção de LV, implementando um detector juntamente com o uso de CNNs. Além disso, pretendemos também realizar classificação utilizando CNNs com um custo computacional ainda menor, de forma que essa classificação possa ser realizada utilizando até mesmo *smartphones*.

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) <https://www.gov.br/cnpq/pt-br>, e ao Dr. Carlos Henrique Nery Costa pelo laboratório que nos proporcionou as amostras.

References

- Borges, A. L., de Araújo Gonçalves, C., Dias, V. B. L., de Andrade, N. B., Aguiar, B. G. A., and e Silva, R. R. V. (2023). Visceral leishmaniasis detection using deep learning techniques and multiple color space bands.
- e Silva, R. R. V., de Araujo, F. H. D., dos Santos, L. M. R., Veras, R. M. S., and de Medeiros, F. N. (2016). Optic disc detection in retinal images using algorithms committee with weighted voting. *IEEE Latin America Transactions*, 14(5):2446–2454.
- Fleiss, J. L., Levin, B., and Paik, M. C. (2013). *Statistical methods for rates and proportions*. John Wiley & Sons.
- Gonçalves, C. d. A., Borges, A. L., Dias, V. B. L., de Andrade, N. B., Aguiar, B. G. A., and Negligenciados, P. (2022). Método automático para detecção de leishmaniose visceral em humanos.
- Górriz, M., Aparicio, A., Raventós, B., Vilaplana, V., Sayrol, E., and López-Codina, D. (2018). Leishmaniasis parasite segmentation and classification using deep learning. In *International Conference on Articulated Motion and Deformable Objects*, pages 53–62. Springer.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Identity mappings in deep residual networks.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications.
- Isaza-Jaimes, A., Bermúdez, V., Bravo, A., Sierra Castrillo, J., Hernández Lalinde, J. D., Fossi, C. A., Flórez, A., and Rodríguez, J. E. (2020). A computational approach for leishmania genus protozoa detection in bone marrow samples from patients with visceral leishmaniasis.
- Kumar, R. and Nylén, S. (2012). Immunobiology of visceral leishmaniasis. *Frontiers in immunology*, 3:251.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2019). Mobilenetv2: Inverted residuals and linear bottlenecks.
- Silva, J., Zacarias, D., Figueiredo, L., Soares, M., Ishikawa, E., Costa, D., and Costa, C. (2014). Bone marrow parasite burden among patients with new world kala-azar is associated with disease severity. *The American journal of tropical medicine and hygiene*, 90.