

# Aplicação de Técnicas de Aprendizado Profundo na Classificação da Retinopatia Diabética

Arthur Carneiro Kaizer<sup>1</sup>, Alexei Manso Correa Machado<sup>1,2</sup>

<sup>1</sup>Departamento de Ciência da Computação

Pontifícia Universidade Católica de Minas Gerais (PUC Minas), Belo Horizonte, Brasil

<sup>2</sup>Departamento de Anatomia e Imagem - Faculdade de Medicina

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brasil

ac.kaizer@sga.pucminas.br, alexeimcmachado@gmail.com

**Abstract.** *Diabetic retinopathy (DR) is a degenerative condition of the retina caused by diabetes. Early diagnosis is crucial to prevent disease progression and preserve vision. However, traditional diagnostic methods require a thorough analysis of retinal images by experts, making the process time-consuming and subjective. This study proposes a deep learning model for the automated classification of DR, based on the analysis of different convolutional neural networks and image processing techniques. The results indicate that the combination of preprocessing using CLAHE and the ResNet50 with the ADAMW optimizer achieved the best performance, with an accuracy of 0.83, ROC-AUC of 0.87, and a Kappa coefficient of 0.85.*

**Resumo.** *A retinopatia diabética (DR) é uma condição degenerativa da retina causada pelo diabetes, sendo o diagnóstico precoce crucial para evitar a progressão da doença. No entanto, os métodos tradicionais de diagnóstico exigem a análise minuciosa das imagens por especialistas, tornando o processo demorado e subjetivo. Este estudo propõe um modelo de deep learning para a classificação de DR, baseado na análise de diferentes redes convolucionais e técnicas de processamento de imagens. Os resultados indicam que a combinação do pré-processamento utilizando CLAHE e a ResNet50 com o otimizador ADAMW alcançou o melhor desempenho, com uma acurácia de 0,83, ROC-AUC de 0,87 e coeficiente Kappa de 0,85.*

## 1. Introdução

A retinopatia diabética (RD) é uma condição degenerativa associada à diabetes, caracterizada por alterações microvasculares na retina que podem levar à cegueira se não forem identificadas e tratadas precocemente. O diagnóstico precoce da RD é essencial para prevenir sua progressão, mas o método tradicional exige análise detalhada por especialistas, tornando-o custoso e demorado. A complexidade da detecção da RD, devido à variabilidade das lesões em tamanho e forma, ressalta a necessidade de soluções automatizadas para auxiliar especialistas na identificação precisa da doença. A aplicação de redes neurais profundas, especialmente redes neurais convolucionais (CNNs), tem se mostrado promissora na automação do diagnóstico, permitindo a identificação eficiente dos padrões característicos da RD. No entanto, os métodos de aprendizado supervisionado ainda precisam de aprimoramentos para aumentar a precisão e diferenciar estágios específicos da doença com maior eficácia.

Diversos métodos podem ser encontrados na literatura para a classificação da RD utilizando aprendizado profundo. [Momeni Pour et al. 2020] desenvolveram um modelo baseado na arquitetura EfficientNet-B5, enquanto [Sudarmadji et al. 2020] combinaram redes neurais com algoritmos genéticos, alcançando alta precisão na classificação. Outros trabalhos, como os de [Saeed et al. 2021], exploraram transferência de aprendizado para aprimorar modelos pré-treinados, enquanto [Patra and Singh 2022] empregaram técnicas de expansão e balanceamento de dados para melhorar a sensibilidade do modelo. Recentemente, [Kalyani et al. 2023] introduziram redes de cápsulas (*CapsNet*) como alternativa às CNNs tradicionais, obtendo resultados promissores na base Messidor. Esses estudos demonstram a diversidade de abordagens para a classificação da RD, destacando avanços em CNNs e novas arquiteturas como as redes de cápsulas.

Este estudo tem como objetivo realizar uma comparação entre diferentes modelos de aprendizado profundo para a classificação automatizada da retinopatia diabética, avaliando o desempenho de distintas arquiteturas de CNNs, bases de dados e técnicas de processamento de imagens. A análise das abordagens mais eficazes e robustas contribuirá para o desenvolvimento de soluções mais precisas e eficientes no diagnóstico da RD, explorando estratégias de pré-processamento e balanceamento de dados para melhorar a acurácia dos modelos.

## 2. Materiais e Métodos

### 2.1. Conjunto de Dados

A fim de propor um modelo generalizado para a classificação multi estágio de RD, optou-se por utilizar dois conjuntos de dados no desenvolvimento dos modelos de aprendizagem profunda.

O conjunto de dados Indian Diabetic Retinopathy Image Dataset (IDRID) [Porwal et al. 2018] é composto por 516 imagens de retina, acompanhado por anotações detalhadas que identificam lesões características dessas condições, além de indicar a gravidade das mesmas, realizadas por especialistas médicos.

O APTOS 2019 Blindness Detection [Karthik 2019] possui imagens de retina classificadas em cinco níveis de gravidade de cegueira (0 a 4), onde 0 representa a ausência de retinopatia diabética e as classes subsequentes representam retinopatia diabética leve, moderada, grave e proliferativa, respectivamente. Cada imagem foi avaliada por oftalmologistas especializados, que atribuíram um rótulo indicando o nível de gravidade da retinopatia diabética presente na imagem. A figura 1 apresenta um exemplo de uma imagem em cada classe.



**Figure 1. Imagens de fundo de olho para cada grau de RD.**

A base de dados foi dividida aleatoriamente em conjuntos de treinamento (70%), validação (15%) e teste (15%), mantendo uma distribuição similar à base de dados orig-

inal. Ao se analisar a distribuição de cada conjunto (treinamento, teste e validação), observa-se que as classes "Sem RD" e "RD Moderada Não Proliferativa" são majoritárias.

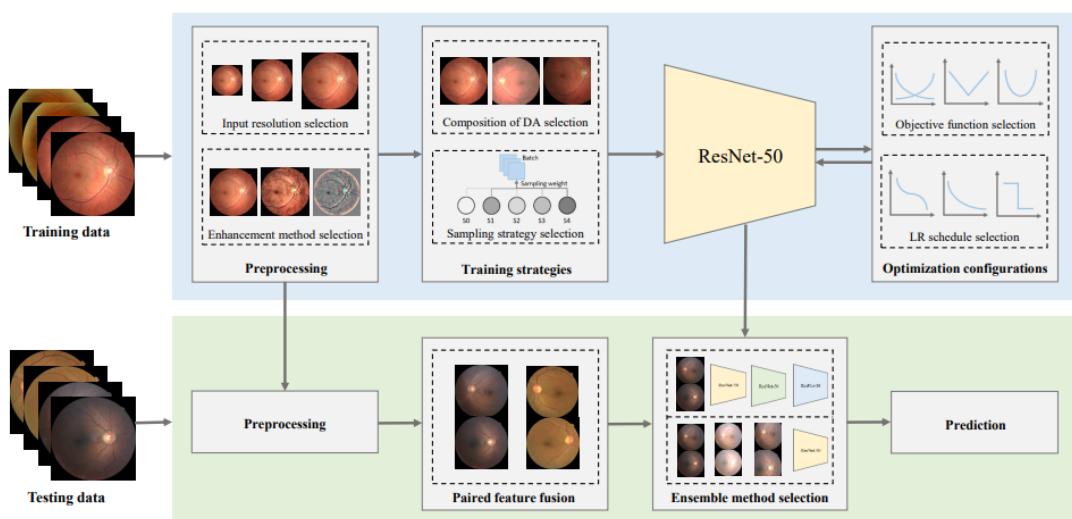
## 2.2. Pré-processamento de Dados

Foram aplicados os seguintes métodos de pré processamento de dados:

- **Recorte Circular:** Essa técnica foi utilizada com objetivo de concentrar a análise nas regiões de interesse das imagens de fundo de olho, como a retina. São removidas informações irrelevantes do entorno da imagem, como a área ao redor do olho.
- **Redimensionamento:** após o recorte circular, as imagens são redimensionadas para um tamanho uniforme de 256 x 256 *pixels*. Essa etapa garante que todas as imagens tenham a mesma resolução, o que facilita o processamento e a análise pelos modelos de aprendizado de máquina.
- **Processamento de Imagem usando CLAHE:** para mitigar os desafios associados ao aumento do contraste em imagens médicas, como o risco de aumento de ruído de fundo e perda de informações úteis, adotou-se a abordagem proposta por Zuiderveld et al. [Lee and Ke 2021] que emprega a Equalização de Histograma Adaptativo com Limite de Contraste (CLAHE).

## 2.3. Arquitetura do Modelo

Para a classificação da retinopatia diabética, foi utilizada a ResNet50 como base devido à sua profundidade e conexões residuais, que mitigam problemas de degradação do desempenho. Além disso, foram analisadas as arquiteturas DenseNet201, que maximiza o fluxo de informações entre camadas, e ViT Large Patch32 224, uma abordagem baseada em *Vision Transformers*, capaz de capturar padrões complexos em imagens de alta resolução. A Figura 2 ilustra a arquitetura da ResNet50 utilizada como referência para comparação.



**Figure 2. Arquitetura da ResNet50 utilizada como base para comparação com as demais arquiteturas, adaptada de [Huang et al. 2023]**

## 2.4. Avaliação do Modelo

A matriz de confusão é utilizada para avaliar o desempenho do modelo, classificando corretamente os casos em Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN). Além disso, o monitoramento da *loss* ao longo do treinamento auxilia na detecção de *overfitting* e *underfitting*, permitindo ajustes nos hiperparâmetros para otimização do modelo.

## 3. Experimentos e Resultados

O método proposto por [Huang et al. 2023] foi adaptado para incluir pré-processamento de imagens, ajustes na arquitetura da rede e extração de métricas de desempenho. Os experimentos foram conduzidos no Google Colab com GPU T4, utilizando imagens redimensionadas para 256x256 e 512x512 *pixels*. Como não houve variação significativa nos resultados, optou-se pela resolução de 256x256 para otimizar o desempenho computacional.

Os modelos ResNet50, DenseNet201 e ViT Large Patch32 224 foram treinados e avaliados, conforme mostrado na Tabela 1. A ResNet50 obteve o melhor desempenho, especialmente com o otimizador ADAMW, que apresentou acurácia superior ao SGD ao longo do treinamento. Os experimentos compararam os efeitos dos otimizadores SGD e ADAMW na ResNet50. A Figura 3 exibe a evolução da perda durante o treinamento.

**Table 1. Comparação de Modelos**

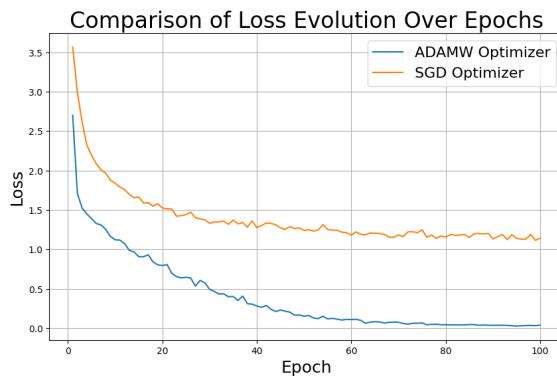
Modelo	Acurácia	Área curva ROC	Kappa
ViT Large Patch32 224	0.69	0.80	0.56
DenseNet201	0.79	0.87	0.85
Resnet50 com otimizador SGD	0.80	0.89	0.85
Resnet50 com otimizador ADAMW	0.83	0.87	0.85

A análise das matrizes de confusão (Figuras 4 e 5) revela que ADAMW apresentou melhor desempenho na classe "Mild NPDR", enquanto o SGD obteve maior precisão na classe "No DR". O ADAMW demonstrou melhor sensibilidade na identificação de estágios iniciais da RD, tornando-se a escolha mais adequada para esse cenário. O otimizador apresentou ainda um maior equilíbrio na distribuição de erros, enquanto o SGD mostrou maior precisão em classes específicas. Dessa forma, o ADAMW parece ser a melhor opção para a detecção precoce da RD devido à sua maior sensibilidade.

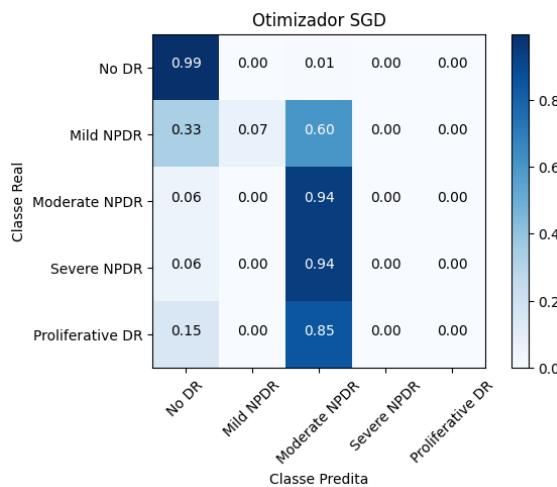
## 4. Conclusão

Este estudo avaliou diferentes modelos de aprendizado profundo para a classificação automatizada da retinopatia diabética, explorando técnicas de pré-processamento, arquiteturas de redes neurais convolucionais e otimizadores. A combinação da técnica CLAHE com a arquitetura ResNet50 e o otimizador ADAMW obteve os melhores resultados, alcançando 83% de acurácia, AUC de 0.87 e coeficiente *Kappa* de 0.85. O ADAMW destacou-se especialmente na detecção de estágios iniciais da doença.

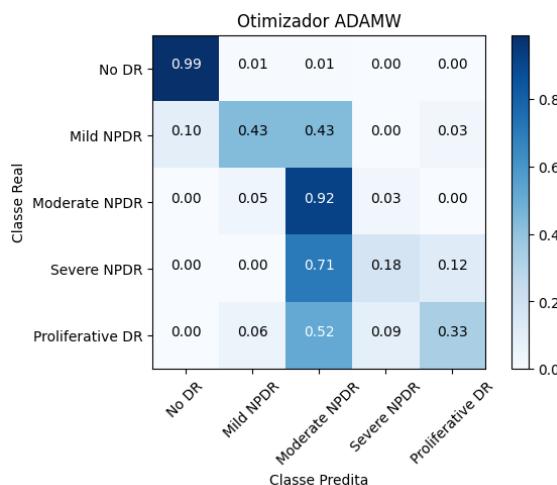
Alguns estudos encontrados na literatura reportam resultados superiores. Por exemplo, [Momeni Pour et al. 2020] obtiveram acurácia acima de 90% com EfficientNet-



**Figure 3. Função de Perda da ResNet50 usando os otimizadores SGD e ADAMW**



**Figure 4. Matriz de confusão - ResNet50 com SGD**



**Figure 5. Matriz de confusão - ResNet50 com ADAMW**

B5, e [Patra and Singh 2022] melhoraram a sensitividade com técnicas avançadas de balanceamento. As diferenças de desempenho podem estar relacionadas ao uso de bases de dados maiores e estratégias mais sofisticadas de aumento de dados.

Apesar de não ter superado outros métodos reportados na literatura, este trabalho apresenta uma análise comparativa sólida do uso de redes neurais profundas para o diagnóstico através de imagens. Como trabalhos futuros, recomenda-se explorar bases de dados mais diversificadas, técnicas de interpretabilidade de modelos e abordagens para lidar com conjuntos de dados pequenos, visando maior aplicabilidade clínica e robustez dos modelos desenvolvidos.

**Agradecimentos** — AMC Machado agradece à FAPEMIG pelo auxílio financeiro através dos projetos APQ-02753-24 e APQ-06556-24.

## References

- Huang, Y., Lin, L., Cheng, P., Lyu, J., Tam, R., and Tang, X. (2023). Identifying the key components in resnet-50 for diabetic retinopathy grading from fundus images: a systematic investigation. *Diagnostics*, 13(10):1664.
- Kalyani, G., Janakiramaiah, B., Karuna, A., and Prasad, L. N. (2023). Diabetic retinopathy detection and classification using capsule networks. *Complex & Intelligent Systems*, 9(3):2651–2664.
- Karthik, Maggie, S. D. (2019). Aptos 2019 blindness detection.
- Lee, C.-H. and Ke, Y.-H. (2021). Fundus images classification for diabetic retinopathy using deep learning. In *Proceedings of the 13th International Conference on Computer Modeling and Simulation, ICCMS ’21*, page 264–270, New York, NY, USA. Association for Computing Machinery.
- Momeni Pour, A., Seyedarabi, H., Abbasi Jahromi, S. H., and Javadzadeh, A. (2020). Automatic detection and monitoring of diabetic retinopathy using efficient convolutional neural networks and contrast limited adaptive histogram equalization. *IEEE Access*, 8:136668–136673.
- Patra, P. and Singh, T. (2022). Diabetic retinopathy detection using an improved resnet50-inceptionv3 structure. In *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–6. IEEE.
- Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., and Meriaudeau, F. (2018). Indian diabetic retinopathy image dataset (idrid).
- Saeed, F., Hussain, M., and Aboalsamh, H. A. (2021). Automatic diabetic retinopathy diagnosis using adaptive fine-tuned convolutional neural network. *IEEE Access*, 9:41344–41359.
- Sudarmadji, P. W., Deviani Pakan, P., and Yefrenes Dillak, R. (2020). Diabetic retinopathy stages classification using improved deep learning. In *2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, pages 104–109.