

# SEGMENTAÇÃO DE LESÕES RETINIANAS ASSOCIADAS À RETINOPATIA DIABÉTICA POR MEIO DE REDES NEURAIAS PROFUNDAS

Segmentation of retinal lesions associated with diabetic  
retinopathy using deep neural networks

Rafaela Menezes<sup>1</sup>

Ismael Lima<sup>2</sup>

Carlos Santos<sup>3</sup>

1 Instituto Federal Farroupilha (IFFar), *Campus* Alegrete, RS, Brasil  
rafaelamenezes.ti@gmail.com - <https://orcid.org/0009-0008-8177-1080>

2 Instituto Federal Farroupilha (IFFar), *Campus* Alegrete, RS, Brasil  
ismaellima123@outlook.com.br - <https://orcid.org/0009-0006-6848-199X>

3 Instituto Federal Farroupilha (IFFar), *Campus* Alegrete, RS, Brasil  
carlos.santos@iffarroupilha.edu.br - <https://orcid.org/0000-0001-8362-8984>

## RESUMO

A retinopatia diabética é uma das principais causas de perda de visão, sendo diagnosticada mediante a investigação da presença de lesões na retina, tais como exsudatos duros e algodinosos, microaneurismas e hemorragias. A identificação e o tratamento dessas lesões, em fases iniciais, podem evitar a perda de visão. O diagnóstico é complexo e necessita de profissionais especializados e infraestrutura para atender à crescente demanda. Diante disso, este trabalho propõe um método baseado em redes neurais profundas para realizar a segmentação de lesões associadas à retinopatia diabética. O método proposto demonstrou ser eficaz na identificação das lesões, obtendo na etapa de teste acurácias de 99,91%, 99,96%, 99,51% e 99,98%, para exsudatos duros, exsudatos algodinosos, hemorragias e microaneurismas, respectivamente.

**PALAVRAS-CHAVE:** lesões retinianas; segmentação; redes neurais profundas.

## ABSTRACT

Diabetic retinopathy is one of the main causes of vision loss and is diagnosed by investigating the presence of retinal lesions, such as hard exudates, soft exudates, microaneurysms and hemorrhages. Identification and treatment of these lesions in the early stages can prevent vision loss. The diagnosis is complex and requires specialized professionals and infrastructure to meet the growing demand. In this context, this work proposes a method based on deep neural networks to segment lesions associated with Diabetic Retinopathy. The proposed method demonstrated to be effective in identifying lesions, obtaining accuracies of 99.91%, 99.96%, 99.51% and 99.98% in the test stage, for hard exudates, soft exudates, hemorrhages and microaneurysms, respectively.

**KEYWORDS:** retinal lesions; segmentation; deep neural networks.

# 1 Introdução

Na área da oftalmologia investiga-se cada vez mais formas de aprimorar e acelerar o diagnóstico de doenças graves, como no caso da Retinopatia Diabética (RD). A RD é uma doença comumente diagnosticada em pacientes diabéticos. Nessa doença aparecem lesões no olho que podem causar a diminuição da visão ou até mesmo a perda total. As lesões tipicamente são classificadas como Exsudatos Duros (EX), Exsudatos Algodonosos (SE), Hemorragias (HE) e Microaneurismas (MA) (Santos *et al.*, 2022).

Caso o diagnóstico da doença seja realizado precocemente, o paciente terá menos chances de ter perda total da visão. Os exames oftalmológicos que identificam essas lesões são realizados por médicos especializados, porém a quantidade desses profissionais é limitada, o que resulta na demora no diagnóstico (Santos *et al.*, 2022). Segundo a INTERNATIONAL DIABETES FEDERATION (2024), em 2000, o número global estimado de adultos com diabetes era de 151 milhões. Em 2009, houve um crescimento de 88%, chegando a 285 milhões pessoas com diabetes. A INTERNATIONAL DIABETES FEDERATION (2024) estima que haverá 600 milhões de pessoas com diabetes em 2035 e, mais de 700 milhões de pessoas, em 2045.

Uma das soluções para a detecção e diagnóstico da RD é utilizar técnicas computacionais para analisar as imagens de fundo do olho e auxiliar os médicos especialistas na realização de diagnósticos mais rápidos e precisos. Redes neurais profundas têm sido utilizadas na literatura para auxiliar no desenvolvimento de sistemas informatizados que possam auxiliar no diagnóstico médico. Assim, o objetivo deste trabalho é propor um método para auxiliar na segmentação de lesões retinianas associadas à RD (Li, 2019).

## 2 Material e Métodos

Para desenvolvimento do método foi utilizada uma rede neural convolucional U-Net (Ronneberger; Fischer; Brox, 2015) com o auxílio das bibliotecas Tensorflow, Keras, Numpy, Pandas e Matplotlib na ferramenta Jupyter Notebook do ambiente de desenvolvimento integrado Anaconda. Para codificação foi utilizada a linguagem Python. O modelo foi treinado por 50 épocas com taxa de aprendizagem de 0,001 e 2 lotes por época. As métricas de desempenho utilizadas para avaliar o modelo foram o Coeficiente Dice, a Interseção sobre União (IoU) e a Acurácia.

## 2.1 Conjunto de dados

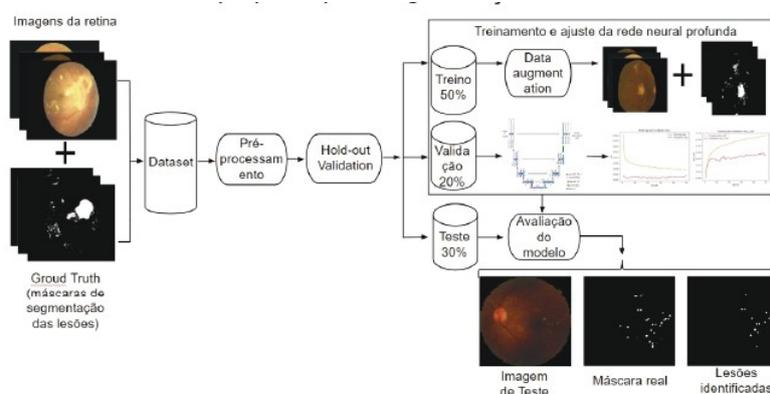
Para realização dos experimentos foi utilizado o conjunto de dados público de imagens da retina Dataset for Diabetic Retinopathy (DDR) (Li, 2019), que fornece três tipos de anotações em nível de imagem, nível de pixel para segmentação e caixa delimitadora das lesões. Todas as imagens foram catalogadas por oftalmologistas conforme a Classificação Internacional de Retinopatia Diabética.

Este conjunto de dados contém 12.522 imagens com resolução variável. Mesmo que possua grande quantidade de imagens das retinas, apenas uma quantidade pequena possui as lesões rotuladas, sendo: 570 imagens com Microaneurismas rotulados; 601 imagens com Hemorragias rotuladas; 486 imagens com Exsudatos Duros rotulados; e, 239 imagens com Exsudatos Algodonosos rotulados.

## 2.2 Método proposto

O método proposto possui as seguintes etapas: (1) Pré-processamento; (2) Hold-out Validation; (3) Data Augmentation; (4) Treinamento e Avaliação do Modelo. Na etapa de pré-processamento, as imagens foram redimensionadas para a proporção de 64 x 64 pixels. Em consequência desse redimensionamento, as imagens perderam a qualidade dos detalhes que apresentavam originalmente. Esta etapa foi realizada em função da limitação de memória da GPU utilizada nos experimentos. Após o pré-processamento, as imagens passaram para a etapa de Hold-out Validation. Na Figura 1 é apresentado o método proposto.

FIGURA 1 – Método proposto para segmentação das lesões retinianas.



Fonte: elaborado pelos autores (2023).

### 2.2.1. Hold-out Validation

O *Hold-out Validation* é a técnica de separação dos dados na qual é definido um percentual para cada conjunto de dados (treino, validação e teste) para criação de amostras. Este método possui o objetivo de comprovar que uma amostra criada pode ter significância para representar um todo, sendo importante no sentido de utilizar uma etapa de Teste em que o modelo não possui conhecimento a priori dos dados deste conjunto, a fim de garantir a capacidade de generalização do modelo, uma vez que a rede neural profunda é treinada e ajustada nos conjuntos de treino e validação apenas (Santos *et al.*, 2022).

O conjunto de dados foi dividido na proporção de 50% das imagens para o conjunto de treino, 20% para o conjunto de validação e 30% para o conjunto de teste (Santos *et al.*, 2022). É importante destacar que todas as imagens do conjunto de dados possuem *Ground Truth* (rótulos) das lesões de fundo do olho, para que seja possível avaliar as previsões realizadas pelo modelo de acordo com os rótulos das lesões certificadas por médicos especialistas.

Com a técnica de *Hold-out Validation* foi possível que a rede neural fosse treinada e ajustada com os conjuntos de treino e validação sem o conhecimento do conjunto de teste. Assim, foi possível avaliar precisamente a capacidade de generalização e previsão do modelo sem problemas relacionados ao ajuste na identificação das lesões.

Ao completar esta etapa, os conjuntos de treino, validação e teste ficaram com 383, 149 e 225 imagens, respectivamente. Após esta etapa foram criadas novas imagens artificialmente por meio de *Data Augmentation*, explicada a seguir.

### 2.2.2. Data Augmentation

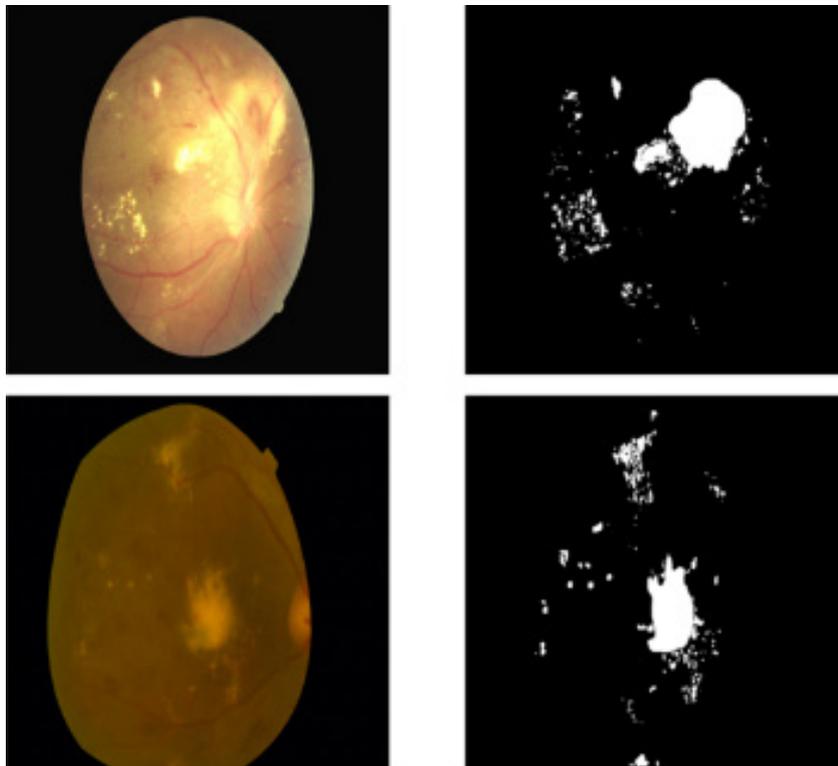
A etapa de *Data Augmentation* foi utilizada com o propósito de aumentar a quantidade de imagens e anotações das lesões do conjunto de dados utilizado nos experimentos. Novas imagens foram geradas aplicando as operações de *Horizontal Flip*, *Vertical Flip*, *Elastic Transform*, *Grid Distortion*, *Optical Distortion*, *Coarse Dropout* disponíveis na biblioteca *Albumentations* (Buslaev *et al.*, 2018).

A partir do uso da biblioteca *Albumentations* foi possível gerar novas imagens: (1) espelhando a imagem original horizontalmente (*Horizontal Flip*); (2) espelhando a imagem original verticalmente (*Vertical Flip*); (3) aplicando distorções na estrutura da imagem, por exemplo, efeito de ondas/ondulações sobre a

imagem (*Elastic Transform*); (4) aplicando distorções na grade (*grid*) da imagem, causando o efeito de variações na distribuição ao longo da imagem (como se a imagem vista de cima tivesse algumas "montanhas", em que algumas regiões estão mais "próximas" e outras mais distantes (*Grid Distortion*); (5) aplicando distorções ópticas, que imitam distorções causadas por lentes na fotografia (*Optical Distortion*); e, (6) eliminação aleatória de regiões retangulares da imagem para simular a oclusão e tamanhos variados de objetos encontrados em configurações do mundo real (*Coarse Dropout*).

Para fins de comparação, o conjunto de Treino do *dataset* DDR possui 383 imagens com lesões rotuladas, e no final da etapa de *Data Augmentation*, esse mesmo conjunto de Treino passou a ter 6.589 imagens com lesões rotuladas. Destaca-se que os conjuntos de Validação e Teste do *dataset* DDR não foram alterados com o *Data Augmentation*. A Figura 2 exibe uma comparação entre a imagem e máscara sem alteração e uma imagem e máscara com alteração (gerada na etapa de *Data Augmentation*).

FIGURA 2 – Comparação entre uma imagem original e uma imagem após a aplicação de *Data Augmentation*.



Fonte: elaborado pelos autores (2023).

### 2.2.3. Arquitetura da rede

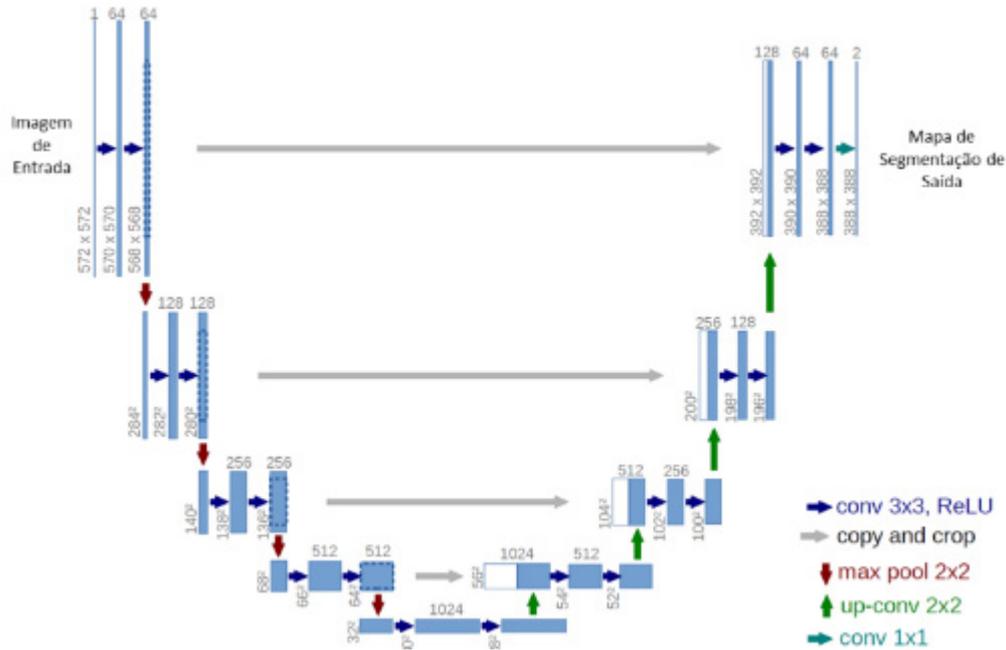
A arquitetura da rede utilizada foi a U-Net, que se trata de uma rede neural totalmente convolucional em formato de U. Originalmente foi desenvolvida para analisar imagens biomédicas e contém a capacidade de ser treinada do início ao fim com poucas imagens. Essa rede neural é dividida em duas partes, uma de contração (lado esquerdo), que armazena as informações relevantes e uma de expansão (lado direito) simétrico, que proporciona uma segmentação com perfeição. Desse modo, essa arquitetura se diferencia de outras redes neurais totalmente convolucionais, as quais não possuem a combinação entre os mapas de características da parte de contração com os seus devidos correspondentes simétricos da parte de expansão (Ronneberger; Fischer; Brox, 2015).

A parte de contração é uma arquitetura de uma Rede Neural Convolucional (CNN) que possui aplicações de duas convoluções  $3 \times 3$  seguidas de uma função de ativação de Unidade Linear Retificada (ReLU) e uma operação de subamostragem com ativação máxima de  $2 \times 2$ , para reduzir a dimensão da imagem. Segundo Silva (2019), em cada uma dessas etapas, o número de canais de características é duplicado.

A parte de expansão consiste em um levantamento do mapa de características acompanhado de um sobreamostragem  $2 \times 2$  que realiza a metade do número de canais de características que se constitui da união com o mapa de características correspondentemente cortado do caminho de contratação e duas convoluções  $3 \times 3$ , cada uma contínua por uma função de ativação ReLU. Essa etapa é importante em função da perda de pixels de borda em cada convolução na parte de contração. Na última camada é aplicada uma convolução  $1 \times 1$  para mapear cada vetor de características para o número desejado de classes (Silva, 2019).

Neste trabalho, a U-Net foi utilizada para gerar a segmentação de lesões retinianas em imagens de RD. Na Figura 3 é apresentada uma ilustração da arquitetura e as camadas que constituem a rede (Ronneberger; Fischer; Brox, 2015).

FIGURA 3 – Arquitetura e camadas que constituem a U-Net utilizada no método proposto.



Fonte: Ronneberger; Fischer; Brox (2015).

#### 2.2.4. Avaliação do modelo

As métricas aplicadas para a validação dos resultados foram o Coeficiente Dice, a Interseção sobre União (IoU) e a Acurácia. Segundo Silva (2019), o Coeficiente Dice é frequentemente utilizado como medida para desempenho de segmentação, no qual indica a proporção de pixels compartilhados entre duas amostras em relação ao total de pixels. A equação do Coeficiente Dice é dada pelo dobro de verdadeiros positivos (VP) comuns em duas amostras, dividido pela soma do número total de amostra, somado aos falsos positivos (FP) e falsos negativos (FN), em que os valores compreendidos por essa medida vão de 0% (nenhuma correspondência entre os volumes) a 100% (total correspondência entre os volumes).

Os VP representam os pixels corretamente apontados como sendo lesão, os FP correspondem aos pixels erroneamente classificados como lesão e os FN aos pixels classificados de forma errada como não sendo lesão, conforme apresentado na Equação 1.

$$\text{Coeficiente Dice} = \frac{2VP}{2VP+FP+FN}$$

Já a Acurácia, por sua vez, calcula o total de acertos baseado em todas as instâncias segmentadas, calculado a partir de todos os casos (segmentados corretamente e incorretamente). A equação da Acurácia é expressa através da razão do acerto total (dado por VP + VN) pelo total de casos, conforme apresentado na Equação 2 (Silva, 2019).

$$\text{Acurácia} = \frac{VP+VN}{VP+VN+FP+FN}$$

A métrica IoU é uma medida que quantifica a sobreposição entre uma caixa delimitadora prevista e uma caixa delimitadora real. Desempenha um papel fundamental na avaliação da precisão da localização de objetos, sendo calculada usando a área de interesse sobre a área de união (Oliveira *et al.*, 2022), conforme a Equação 3.

$$\text{IoU} = \frac{VP}{VP+FP+FN}$$

Essas métricas têm o intuito de medir o desempenho do modelo proposto e auxiliar na identificação de problemas durante as etapas de treinamento, validação e teste. Elas também contribuem para possíveis ajustes na arquitetura da rede neural convolucional ou nos hiperparâmetros aplicados durante os experimentos com o objetivo de aumentar a capacidade preditiva do modelo.

### 3 Resultados e Discussão

Nos experimentos do método proposto foram utilizadas as imagens com as lesões rotuladas do conjunto de dados público de retinopatia diabética DDR. Este conjunto de dados foi dividido em subconjuntos de treino, com 6.589 imagens; validação, com 149 imagens; e, teste, com 225 imagens.

Inicialmente, as imagens foram redimensionadas para 64 x 64 pixels para que fossem criadas novas imagens artificialmente (*Data Augmentation*). Em seguida, foi iniciada a etapa de treinamento da rede neural convolucional U-Net e a segmentação das lesões retinianas. Essa última etapa foi executada separadamente para cada lesão. As Tabelas 1 e 2 apresentam os resultados obtidos com o método proposto com a rede neural profunda U-Net nas etapas de validação e teste no conjunto de dados DDR.

TABELA 1 – Resultados obtidos com o método proposto com a U-Net na etapa de validação no conjunto de dados DDR.

	Exsudatos duros	Exsudatos algodonosos	Hemorragias	Microaneurismas
IoU	0,79075	0,72556	0,70061	0,81200
Acurácia (%)	99,91	99,90	99,84	99,97

Fonte: elaborado pelos autores (2024).

TABELA 2 – Resultados obtidos com o método proposto com a U-Net na etapa de teste no conjunto de dados DDR.

	Exsudatos duros	Exsudatos algodonosos	Hemorragias	Microaneurismas
IoU	0,69803	0,89172	0,61334	0,87993
Acurácia (%)	99,76	99,96	99,51	99,98

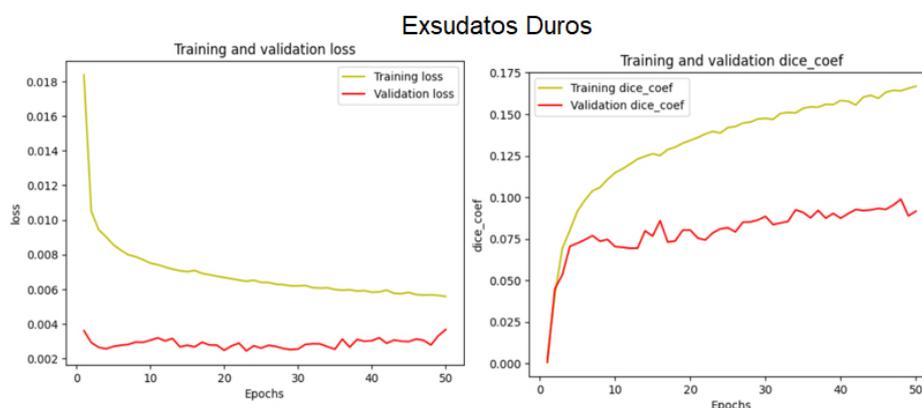
Fonte: elaborado pelos autores (2024).

Considerando os resultados apresentados nas Tabelas 1 e 2, pode-se constatar que os resultados foram satisfatórios no conjunto de dados DDR, alcançando na etapa de Teste a Acurácia de 99,76%, 99,96%, 99,51%, e 99,98% para os exsudatos duros, exsudatos algodonosos, hemorragias e microaneurimas, respectivamente. Estes resultados sinalizam que o método proposto alcançou resultados promissores na segmentação das lesões de fundo do olho.

Já a métrica IoU na etapa de Teste os valores obtidos foram, respectivamente, 0,69803, 0,89172, 0,61334 e 0,87993 para exsudatos duros, exsudatos algodonosos, hemorragias e microaneurimas. Com esses resultados é possível verificar que o desempenho do método proposto na segmentação das lesões exsudatos duros e hemorragias foi inferior em relação às lesões exsudatos algodonosos e microaneurimas.

Os resultados obtidos são atribuídos às etapas utilizadas no método proposto. As Figuras 4, 5, 6 e 7 apresentam os gráficos da curva de aprendizado do método proposto durante as etapas de treinamento e validação, aplicando o conjunto de dados DDR ao longo de 50 épocas.

FIGURA 4 – Gráfico de aprendizado do método proposto durante as etapas de treinamento e validação – exsudatos duros.

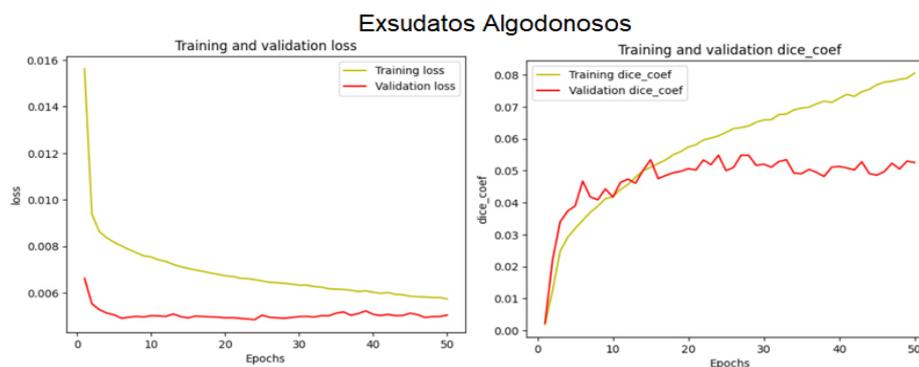


Fonte: elaborado pelos autores (2024).

No gráfico da esquerda, a linha em amarelo apresenta o resultado da função de perda na etapa de treinamento (Training loss), e a linha em vermelho expressa o resultado da função de perda na etapa de validação (Validation loss), conforme acontece o treinamento do modelo (Epochs).

Em três dos quatro gráficos da função de perda o Training loss localizou-se no valor 0,006 (Figuras 4, 5 e 7). Apenas as lesões hemorragias encontrou-se no valor 0,012 (Figura 6), o que sugere uma maior perda de informação durante a etapa de treinamento dessa lesão.

FIGURA 5 – Gráfico de aprendizado do método proposto durante as etapas de treinamento e validação – exsudatos algodonosos.

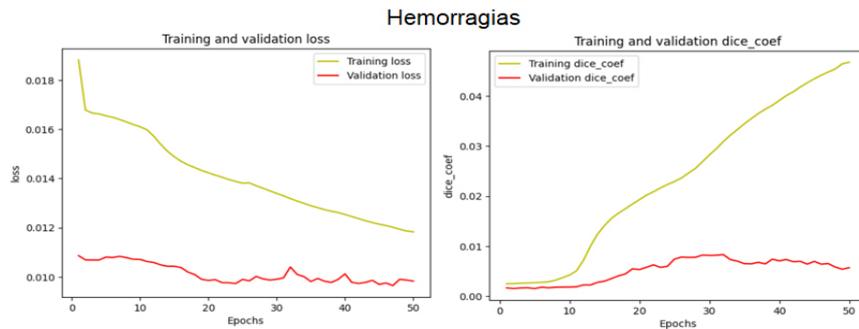


Fonte: elaborado pelos autores (2024).

Já a *Validation loss* de exsudatos duros e exsudatos algodonosos situou-se no valor 0,004, ao passo que as hemorragias no valor de 0,010, e microaneurismas no

valor 0,002. Estes resultados demonstram que, na etapa de validação, as hemorragias obtiveram a maior perda, enquanto que os microaneurismas a menor perda.

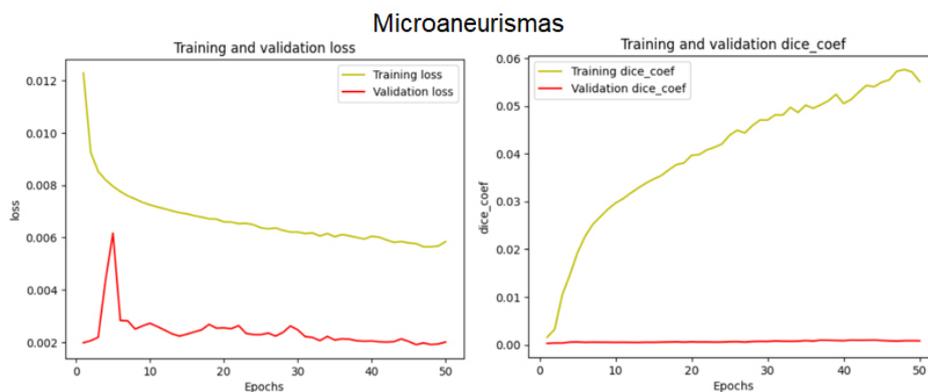
FIGURA 6 – Gráfico de aprendizado do método proposto durante as etapas de treinamento e validação – hemorragias.



Fonte: elaborado pelos autores (2024).

Por outro lado, no gráfico da direita, a linha amarela mostra o resultado da função do Coeficiente Dice na etapa de treinamento (*Training dice\_coef*) e a linha vermelha exibe o resultado da função do Coeficiente *Dice* na etapa de Validação enquanto segue-se o Treinamento do modelo (*Epochs*).

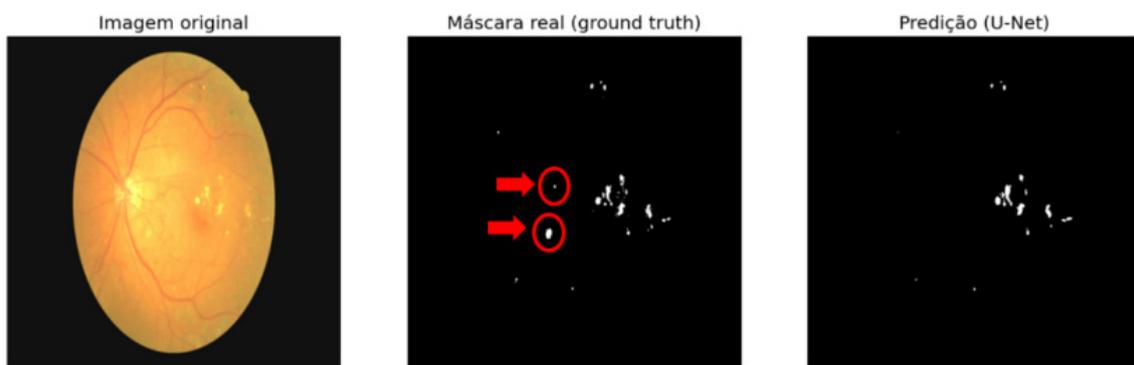
FIGURA 7 – Gráfico de aprendizado do método proposto durante as etapas de treinamento e validação – microaneurismas.



Fonte: elaborado pelos autores (2024).

No *Training dice\_coef*, as lesões hemorragias e microaneurismas situam-se no valor 0,05, sendo as lesões com o menor valor. Já os exsudatos algodonosos ficam com valor de 0,08, e os exsudatos duros com o valor de 0,175. A Figura 8 apresenta um exemplo de segmentação de exsudatos duros realizados pelo método proposto em uma imagem do conjunto de Teste do conjunto de dados DDR.

FIGURA 8 – Segmentação de Exsudatos Duros em uma imagem da retina do conjunto de Teste do dataset DDR por meio do método proposto.



Fonte: elaborado pelos autores (2024).

Pode-se observar na Figura 8 três imagens. A imagem da esquerda, é uma imagem da retina do dataset DDR que não foi utilizada durante o Treinamento e ajuste da rede neural convolucional. No centro, a imagem do *Ground Truth* dos exsudatos duros presentes nessa mesma imagem. E à direita, a imagem da predição realizada pelo método proposto.

Mesmo que o método proposto não tenha conseguido segmentar apropriadamente todas as lesões presentes na imagem ao centro, é evidente a capacidade de predição do modelo, que conseguiu segmentar grande parte dos exsudatos duros presentes na retina, mesmo se tratando de microlesões.

## 4 Considerações finais

Esse trabalho teve como objetivo apresentar um novo método constituído por uma rede neural totalmente convolucional para realizar a segmentação de lesões retinianas associadas à retinopatia diabética.

O método é composto por etapas de Pré-processamento, *Hold-out Validation* e *Data Augmentation*. Foram realizados experimentos para treinamento e ajuste

de hiperparâmetros do modelo. Em seguida, foi realizada uma etapa de avaliação do método proposto a fim de verificar a capacidade de generalização do modelo quanto à segmentação das lesões retinianas.

O método proposto alcançou na etapa de validação, a acurácia de 99,91%, 99,90%, 99,84% e 99,97% para exsudatos duros, exsudatos algodonosos, hemorragias e microaneurismas, respectivamente. Já na etapa de teste, obteve-se a acurácia de 99,76%, 99,96%, 99,51%, e 99,98% para exsudatos duros, exsudatos algodonosos, hemorragias e microaneurismas, respectivamente.

Em trabalhos futuros pretende-se realizar experimentos com novas arquiteturas de U-Nets e explorar novas técnicas de aumento de dados a fim de verificar a melhoria da precisão do método nas segmentações das lesões.

## Agradecimento

Este estudo foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – Brasil, por meio da concessão de bolsa PIBIC – EM.

Agradecemos ao Laboratório de Ensino, Pesquisa, Extensão e Produção em Informática Lab Quest (LEPEP Lab Quest) do Instituto Federal Farroupilha Campus Alegrete pelo apoio e colaboração para o sucesso do projeto.

## Referências

BUSLAEV, A. et al. Alumentations: fast and flexible image augmentations. *In: COMPUTER VISION AND PATTERN RECOGNITION*, 2018, Cornel. 2018. Disponível em: <https://arxiv.org/abs/1809.06839>. Acesso em: 22 nov. 2024.

INTERNATIONAL DIABETES FEDERATION. **Facts & figures**, 2024. Disponível em: <https://idf.org/about-diabetes/diabetes-facts-figures/>. Acesso em: 14 out. 2024.

LI, T. et al. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. **Information Sciences: an international journal**, v. 501, n. C, p. 511-522, 2019. Disponível em: <https://dl.acm.org/doi/10.1016/j.ins.2019.06.011>. Acesso em: 14 out. 2024.

OLIVEIRA, M, V. et al. PPM-UNET: uma rede neural convolucional para a segmentação de rins em imagens de TC. *In: SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO APLICADA À SAÚDE*, 22., 2022, Teresina. **Anais eletrônicos**[...]. Teresina: Sociedade Brasileira de Computação, 2022. Disponível em: <https://sol.sbc.org.br/index.php/sbcas/article/view/21638>. Acesso em: 09 ago. 2024.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional networks for biomedical image segmentation. *In: MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION (MICCAI 2015)*, 2015, Cornell. 2015. Disponível em: <https://arxiv.org/abs/1505.04597v1>. Acesso em: 09 ago. 2024.

SANTOS, C. et al. A new approach for detecting fundus lesions using image processing and deep neural network architecture based on YOLO Model. **Sensors**, v. 17, n. 22, p. 1-37, 2022. Disponível em: <https://www.mdpi.com/1424-8220/22/17/6441>. Acesso em: 09 ago. 2024.

SILVA, W. A. **Segmentação do disco óptico em imagens retinoscópicas utilizando a Rede Neural Convolutio-  
nal U-Net. 2019.** Monografia (Trabalho de Conclusão de Curso) - Universidade Federal do Piauí, Piauí, 2019.  
Disponível em: <https://bit.ly/46JHzqf>. Acesso em: 09 ago. 2024.