

Detecção de Lesões de Retinopatia Diabética em Imagens de Fundoscopia Utilizando UNet++

Antonio E. P. Alves

*Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará (UFC)
Fortaleza, Brasil
ermeson.alves@lesc.ufc.br*

Débora F. Assis

*Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará (UFC)
Fortaleza, Brasil
debora.ferreira@lesc.ufc.br*

Andressa G. Moreira

*Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará (UFC)
Sobral, Brasil
andressagomes@alu.ufc.br*

Paulo C. Cortez

*Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará (UFC)
Fortaleza, Brasil
cortez@lesc.ufc.br*

Bruno R. Silva

*Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará (UFC)
Fortaleza, Brasil
bruno@lesc.ufc.br*

Resumo—A Retinopatia Diabética (RD) é uma complicação microvascular progressiva do *Diabetes Mellitus* que afeta a retina devido ao excesso de glicose no sangue, causando alterações nos vasos sanguíneos. Em seus estágios iniciais a RD é assintomática, contudo pode levar a problemas visuais como visão turva e pontos cegos nos casos mais avançados. A prevenção dessa doença pode ser realizada pelo controle dos níveis de açúcar no sangue e exames oftalmológicos regulares. Desta forma, a detecção precoce é fundamental para preservar a visão e qualidade de vida do paciente. O avanço das técnicas de visão computacional também expandiu a aplicação de técnicas de segmentação semântica do campo de *Deep Learning* para localização automática das lesões intrarretinianas da Retinopatia Diabética. Neste contexto, este estudo visa comparar o desempenho da arquitetura da rede UNet++, com 4 diferentes redes que substituem o encoder dentro da sua arquitetura, para segmentação automática das lesões da RD em imagens de fundoscopia. São avaliadas a localização e segmentação dos microaneurismas, hemorragias, exsudatos moles e exsudatos duro dos bancos de dados IDRID e DIARETDB1. As análises indicaram que o encoder VGG16 alcançou os melhores resultados de segmentação, em especial para hemorragias, considerando Índice de Jaccard, Coeficiente de Similaridade de Dados e Acurácia.

Index Terms—Retinopatia Diabética, Segmentação Semântica, Fundoscopia, UNet++, Encoders.

I. INTRODUÇÃO

A Retinopatia Diabética (RD) é uma complicação microvascular do *Diabetes Mellitus* que causa anormalidades na retina, em casos mais graves pode levar à cegueira [1]. Segundo informações de *Saeedi et. al* [2], estimou-se que em 2019

Os autores gostariam de agradecer a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Código 001 e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Proc. 313599/2019-0.

um total de 463 milhões de pessoas vivam com diabetes, representando 9,3% da população adulta global (20 a 79 anos). Além disso, estima-se que esse número aumente para 578 milhões (10,2%) em 2030 e 700 milhões (10,9%) em 2045.

Em geral, não há sintomas no estágio inicial da doença, isso pode contribuir para a descoberta em um estado mais avançado, quando há neovascularização estimulada a partir do disco óptico ou retina. As lesões indicativas da RD desde o estágio inicial até a neovascularização são: microaneurismas, hemorragias pontuais e borradas, manchas algodinosas (infartos neuronais localizados causados por fluxo axoplasmático desordenado), exsudatos moles e duros e demais anomalias microvasculares intrarretinianas. [3].

A RD ocorre quando pequenos vasos incham e vazam fluido ou novos vasos sanguíneos anormais crescem, dificultando a visão normal. O estágio inicial da RD é referido como Retinopatia Diabética não Proliferativa (RDNP), pois não há presença de neovascularização (crescimento de novos vasos sanguíneos anormais na retina), e a retina pode apenas conter patologias como pequenas manchas vermelhas escuras ou hemorragias minúsculas formando aglomerados com formato circular [3].

As hemorragias que aparecem nas camadas mais profundas da retina, têm uma forma arredondada. As manchas algodinosas são lesões brancas amareladas na camada de fibras nervosas e também são chamadas de exsudatos moles (ou macios). Os exsudatos duros são tipicamente brilhantes, refletivos e incomuns na RD. Eles aparecem como lesões de cor branca ou creme na retina com tamanhos diferentes. Cada exsudato duro consiste em plasma sanguíneo e lipídios vazados

dos vasos sanguíneos [4].

Para obter o diagnóstico da Retinopatia Diabética, exames com imagens médicas são usados sob avaliação subjetiva de um especialista. A Fundoscopia é um dos exames mais utilizados, tendo em vista ser uma opção mais econômica [5]. As principais características que auxiliam na análise dessas imagens são microaneurismas, hemorragias e manchas algodonosas.

Recentemente, diversos trabalhos foram desenvolvidos para a identificação da retinopatia diabética a partir da detecção ou segmentação das lesões em imagens de fundo do olho. Com base na revisão de Tsiknakis *et al.* [6], é perceptível que o principal objetivo dos trabalhos de segmentação da RD é identificar os quatro tipos de lesões citadas acima.

Dessa forma, este estudo visa avaliar o desempenho da arquitetura de segmentação UNet++, através de diferentes *encoders*, para segmentar diferentes tipos de lesões. Assim, é possível verificar se o *encoder* específico segmenta corretamente todos os tipos de lesão, ou se a avaliação individual dessas lesões por diferentes *encoders* acarreta em uma melhor precisão dos diferentes artefatos para identificar a retinopatia diabética.

A estrutura deste trabalho está disposta da seguinte forma: na seção 2, os trabalhos relacionados são elicitados e discutidos. A metodologia é apresentada na seção 3. Os resultados e discussões são abordados na seção 4. Por fim, na seção 5 são sintetizadas as conclusões e perspectivas de trabalhos futuros.

II. TRABALHOS RELACIONADOS

A análise de imagens do fundo do olho tem sido bastante desenvolvida nos últimos anos, principalmente com a aplicação de métodos automáticos de *Deep Learning* para localização de artefatos na retina. Dentre os trabalhos relacionados, destaca-se o estudo realizado por Schweisthal e Lascu [7]. Os autores utilizam a arquitetura U-Net para realizar a segmentação semântica da Retinopatia Diabética com as redes pré-treinadas RESNET-34, INCEPTION-V3 e VGG16 na estrutura de codificação. Os conjuntos de dados considerados foram DIARETDB1 e DIARETDB0. Em relação à análise dos resultados, foram utilizadas as métricas *Intersection-over-Union* (IoU), ou Índice de Jaccard, e o Coeficiente de Similaridade de Dados (Dice).

Qomariah, Tjandrasa e Fatichah [8] propuseram uma nova rede que modifica a U-Net usando unidades residuais com mapeamento de identidade modificado (MResUNet) para segmentar os microaneurismas. Inicialmente, os autores pré-processaram as imagens dos bancos de dados IDRiD e DIARETDB1. Após o treinamento, os resultados indicaram uma complexidade na segmentação dos microaneurismas, dado o tamanho das lesões comparado com a região saudável da imagem.

Xu *et al.* [9] aprimoraram a rede U-Net com um novo modelo denominado *Feature Fusion U-Net* (FFU-Net), na qual a camada de agrupamento na rede é substituída por uma camada convolucional, a fim de reduzir a perda espacial. Ademais, houve a integração do bloco de fusão de recursos

multiescala (MSFF) nos codificadores e a proposta de uma diferente função de perda. As imagens do banco de dados IDRiD foram pré-processadas com corte, filtro gaussiano 3×3 , realce e normalização. Os resultados com a FFU-Net mostraram superioridade à arquitetura U-Net padrão.

Huang, Li, Xiao, Shen, e Xu [10] investigaram as causas patogênicas das lesões da RD e observaram padrões relativos entre elas. A partir disso, propuseram uma nova rede neural de ramificação dupla com a U-Net no início do pipeline de segmentação, seguida por um *Global Transformer Block* e um *Relational Transformer Block* para empregar informações vasculares e de lesão. O estudo é baseado em dois conjuntos de dados, IDRiD e DDR, pré-processados com as técnicas *Resize*, *Random Cropping*, *flips* horizontais e verticais aleatórios e aprimoramento de contraste com CLAHE. Tanto a área sob a curva (AUC) da curva de precisão e revocação (PR) quanto da curva característica de operação de recebimento (ROC) foi utilizada para avaliação do método.

Guo, Yang e Yu (2023) [11] aplicaram o treinamento de redes totalmente convolucionais (FCN) em diferentes escalas para obter máscaras de segmentação e o treinamento de uma CNN para aprender os pesos de fusão das máscaras. As imagens dos conjuntos de dados IDRiD, DDR, E-optha e DIARETDB1 foram normalizadas com filtro gaussiano, ajustado o contraste e intensidade de cores. As métricas de avaliação incluíram a área sob as curvas de precisão e sensibilidade, bem como o coeficiente de similaridade dos dados.

Wang *et al.* [12] elaboraram um método usando CNN com arquitetura colaborativa que compreende uma ramificação contextual e uma ramificação local totalmente automático para segmentação simultânea de hemorragias, exsudatos duros, exsudatos moles e microaneurismas. A validação do método se deu por meio dos *datasets* IDRiD, E-optha e DDR, sob método Otsu para remoção de regiões não teciduais, corte em *pachts* de tamanho 256×256 com sobreposição de 128 *pixels* entre *pachts* adjacentes e aumento de dados.

Jebaseeli, Durai e Peter [13] avaliaram o modelo *Tandem Pulse Cupled Neural Network* (TPCNN) e propuseram *Deep Learning Based A Support Vector Machine* (DLBSVM) para geração automática de vetores de características e classificação e segmentação de vasos sanguíneos respectivamente. Justificaram que a segmentação precisa dos vasos sanguíneos é clinicamente significativa para auxílio no diagnóstico da RD.

Maaliw *et al.* [14] combinaram várias técnicas de pré-processamento e propuseram uma arquitetura de segmentação de imagem DR-UNet, baseada na U-Net. A priori, as imagens dos conjuntos de dados DIARETDB1 e DIARETDB2 foram convertidas em imagens de 1440×900 pixels e normalizadas em imagens de 512×512 pixels. Os autores dividiram a arquitetura proposta em codificador, gargalo e decodificador.

A segmentação de lesões para a identificação da retinopatia diabética é constantemente explorada no âmbito de segmentação de imagens, contudo alguns tipos específicos de lesões ainda não são facilmente identificados, o que pode tornar essa tarefa complexa devido ao tamanho e quantidade das lesões. Dentre os trabalhos acima alguns, como o de

Treinamento e validação

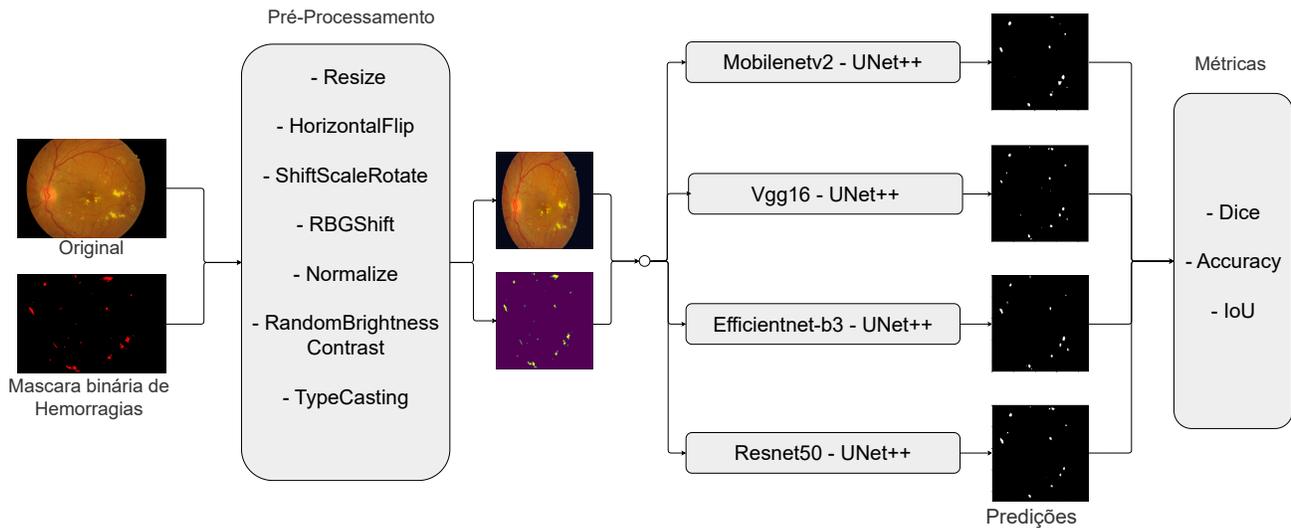


Figura 1: Diagrama da metodologia para um tipo de lesão.

Jebaseeli, Durai e Peter [13] e o de Qomariah, Tjandrasa e Fatichah [8] não consideraram todas as lesões. Outro fator presente nesses trabalhos é a busca por um único modelo para segmentar os 4 tipos de lesões.

III. MATERIAIS E MÉTODOS

O fluxograma apresentando as etapas realizadas é ilustrado na Figura 1.

Inicialmente, os bancos de dados foram divididos em diferentes partes conforme o tipo de lesão, e em cada porção, para fins de treinamento, foi realizado um aumento de dados aplicando diferentes técnicas como rotação, ajuste de contraste, normalização, entre outros. A imagem original, capturada pelo exame de fundoscopia, é submetida a uma etapa de pré-processamento, na qual técnicas de redimensionamento, mudanças de escalas e contraste são utilizadas. Esta etapa é importante para facilitar para os algoritmos de aprendizagem de máquina o reconhecimento dos padrões de cada lesão.

Em seguida, a arquitetura da Unet++ [15] é utilizada para a segmentação das lesões provenientes da retinopatia. Redes leves tais como a MobileNet-v2 [16], EfficientNet-b3 [17] foram utilizadas, visando validar os resultados comparando com duas redes tradicionais, a VGG16 [18] e a ResNet50 [19] através das métricas de coeficiente Dice [20] [21], índice de Jaccard [22] e acurácia.

Nas seções seguintes, serão detalhados os bancos de dados, a rede UNet++, as arquiteturas de *Deep Learning* utilizadas como *encoders* e as métricas de avaliação de segmentação de imagens.

A. Base de dados

IDRID [23]: O conjunto de dados IDRID foi disponibilizado como parte do workshop Diabetic Retinopathy: Seg-

mentation and Grading Challenge, motivado principalmente pela falta de oftalmologistas na população da Índia, e para evoluir as ferramentas de diagnóstico da RD a fim de facilitar a triagem. Os dados são divididos em três partes: uma para aplicação de algoritmos de segmentação semântica, outra para classificação e por fim uma para localização.

Para este artigo foram utilizados os dados de segmentação, que consistem em 81 imagens coloridas de fundoscopia, divididas em 54 imagens para o conjunto de treino e 27 imagens para o conjunto de teste. Também estão disponíveis as máscaras das regiões das lesões de microaneurismas, hemorragias, exsudatos duros, exsudatos moles e disco óptico. Um exemplo de imagem do fundo de olho com retinopatia diabética, e as máscaras para cada tipo de lesão é apresentada na Figura 2.

DIARETDB1 [24]: o banco de dados DIARETDB1 é composto por 89 imagens coloridas de fundoscopia, das quais 84 imagens contêm pelo menos algum sinal da RD, e 5 são consideradas saudáveis. O conjunto de treino é composto por 28 dessas imagens e o conjunto de teste é composto por 61 imagens. Para cada imagem são disponibilizadas uma máscara correspondentes as lesões de microaneurismas, exsudatos duros, exsudatos moles e hemorragias. As imagens foram capturadas com a mesma câmera digital de fundo de campo de visão de 50 graus com configurações de imagem variadas (intensidade de iluminação, velocidade do obturador, abertura, ganho) controladas pelo sistema. Um software foi fornecido para diferentes especialistas a fim de que eles pudessem inspecionar as imagens de fundoscopia e anotar os achados (lesões). Eles foram instruídos a marcar as áreas relacionadas as lesões e relatar o nível de confiança bem como o ponto mais representativo do achado em questão. Os

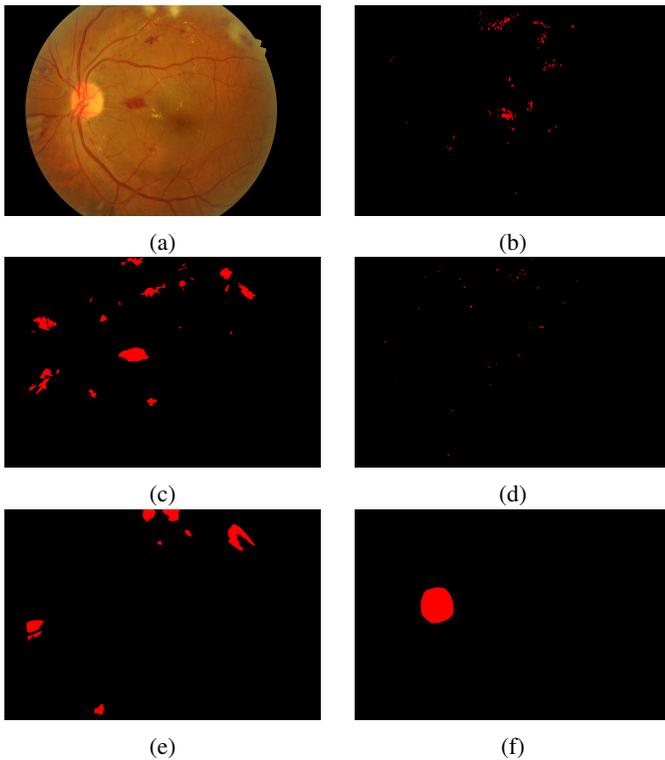


Figura 2: Imagens de Exemplo do *dataset* IDRID. (a) Imagem de Fundoscopia com lesões da RD. De (a)-(f): máscaras das lesões. O canal de cor vermelho é o único representativo. (b) Exsudatos Duros. (c) Hemorragias. (d) Microaneurismas. (e) Exsudatos Moles. (f) Disco Óptico.

níveis de confiança das máscaras, $\{< 50\%, > 50\%, 100\%\}$, representaram a certeza da decisão de que um achado marcado está correto. A Figura 3 apresenta exemplos de imagens do banco de dados DIARETDB1 com as máscaras de cada tipo de lesão e as quatro áreas conforme as marcações dos especialistas.

A divisão dos conjuntos de dados em treino e teste nesse trabalho é baseada na divisão original de ambos, com o acréscimo de um terceiro conjunto para validação usando 10% dos dados de treino.

B. Pré-processamentos

Antes de enviar as imagens como entrada das redes, é realizado um redimensionamento para o tamanho de 224×224 , que é um tamanho adequado para não haver erros de concatenação de recursos nas “conexões de salto” que une a rede codificadora com a decodificadora.

A implementação do trabalho foi realizada com a biblioteca dedicada Pytorch [25]. Os métodos de aumento de dados são implementados com auxílio da biblioteca *Albumentations* [26] e incluem *Flips* Horizontais aleatórios com probabilidade 0,5, *ShiftScaleRotate* (Deslocamento Horizontal e/ou Vertical, Mudança de Escala e Rotação aleatoriamente), *RGBShift* (Mudança aleatória dos valores dos canais de cores RGB),

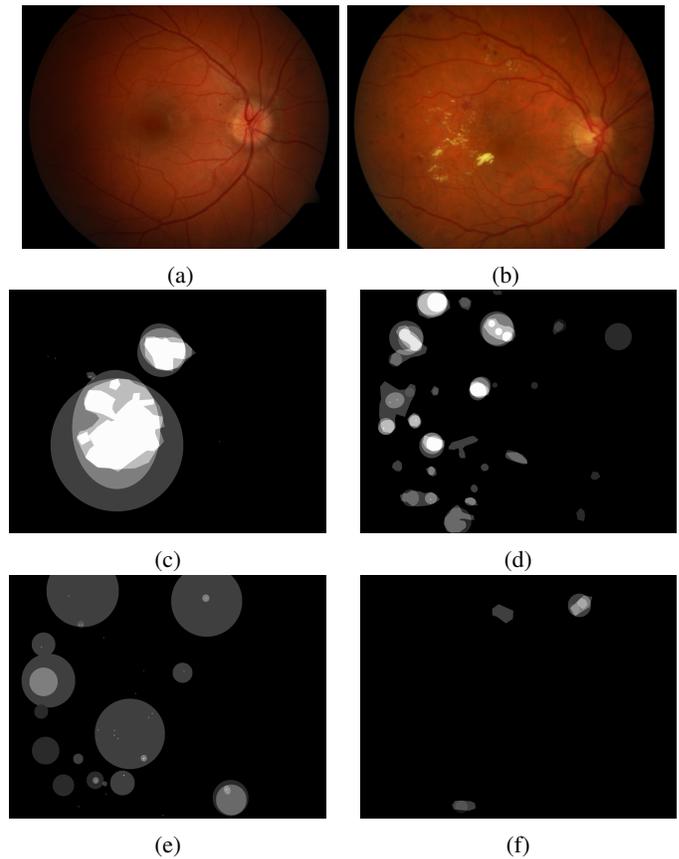


Figura 3: Imagens do *dataset* DIARETDB1. (a) Imagem saudável. (b) Imagem com lesões da RD. De (c) a (f): máscaras das lesões da imagem com lesões da RD. (c) Exsudatos Duros. (d) Hemorragias. (e) Microaneurismas. (f) Exsudatos Moles.

Normalização dos Dados através da média e desvio padrão dos valores das imagens, *RandomBrightnessContrast* (Mudança aleatória no brilho e contraste da imagem) e *TypeCasting* para mudança da estrutura de dados. Essa etapa de aumento de dados é realizada somente na etapa de treinamento. E os mesmos parâmetros são usados a cada treino.

C. Algoritmos baseados em Deep learning

A UNet++ é uma rede baseada na U-Net [27] com modificações que mostraram superioridade com relação a versões anteriores da arquitetura. Para a segmentação semântica, a U-Net teve um papel importante por introduzir a lógica de uma rede totalmente convolucional para combinar uma sub-rede codificadora, responsável por capturar o contexto (ou campo receptivo) da imagem de entrada, com uma sub-rede decodificadora, que gera uma imagem com maior resolução a partir dos mapas de características [7]. Os autores da UNet++ afirmam que “O que distingue a UNet++ da U-Net são os caminhos de salto reprojados que conectam as duas sub-redes e o uso de supervisão profunda” [27].

Tais caminhos de salto tratam da conectividade das sub-redes em cada nível da arquitetura. Na UNet++, a saída da

rede codificadora em determinado nível passa por blocos de convolução densos para então servir de entrada para a rede decodificadora, os autores afirmam que a profundidade dos blocos convolucionais dependem do nível.

O presente estudo atua na *backbone* da UNet++, ou sub-rede codificadora, utilizando quatro redes pré-treinadas no conjunto de dados Imagenet, comparando o resultado de cada uma ao segmentar lesões da RD.

Utilizou-se 3 canais de entrada para a rede conseguir lidar com imagens coloridas, com uma taxa de aprendizado igual a $1e10^{-3}$ e função de perda Dice [28]. As diferentes configurações para o modelo UNet++ foram treinadas por 100 épocas com auxílio da unidade de processamento gráfico NVIDIA T4, do ambiente de programação Google Colaboratory. O repositório para o notebook python desse trabalho pode ser acessado em <https://github.com/ermeson-alves/UNetpp-RD>.

D. Métricas

Para avaliar o desempenho do modelo ao realizar segmentação semântica, métricas tradicionais de desempenho de algoritmos de *Deep Learning* podem não ser o ideal. Até mesmo métricas que parecem adequadas podem sofrer de determinados vieses estatísticos, sendo necessário reunir medidas apropriadas de desempenho [29].

Tais métricas devem ser capazes de, dado uma máscara binária com regiões de interesse e fundo, avaliar o quão bem o modelo está gerando a máscara binária com base em seu treinamento. Assim, medidas de modelos de classificação não se saem bem tendo em vista a grande quantidade de verdadeiros negativos gerados [29].

As métricas utilizadas nesse estudo têm o cálculo baseado em uma matriz de confusão, como ilustra a Tabela I. Trata-se uma matriz de correspondência das classes verdadeiras, rotuladas nas linhas, com as classes previstas, rotuladas nas colunas. Como foi considerado uma lesão de cada vez, a máscara da verdade fundamental (padrão ouro) possui apenas duas classes por pixel: *background* ou lesão.

Tabela I: Exemplo de matriz de confusão

	Lesão	Background
Lesão	VP	FN
Background	FP	VN

Nos valores da tabela, VP (verdadeiro positivo) referem-se ao número de predições para o tipo lesão que são de fato lesão, VN (verdadeiro negativo) é o número de predições como *background* que são de fato *background*, FP (Falso Positivo) consiste no número de predições que o modelo classificou como lesão, mas se tratam de *background*, e o FN (Falso Negativo) corresponde ao número de predições que o modelo classificou como *background*, mas se tratam de uma lesão. Com base nisso, as métricas utilizadas podem ser definidas da seguinte forma:

1) *Acurácia*:

$$Acc = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

2) *Coefficiente Dice*:

$$DSC = \frac{2VP}{2VP + FP + FN} \quad (2)$$

3) *IoU (Índice de Jaccard)*:

$$IoU = \frac{VP}{VP + FP + FN} \quad (3)$$

A respeito da métrica IoU, foram utilizados dois tipos de cálculo: IoU por imagem (cálculos da pontuação IoU para cada imagem, seguidos da média de todos os cálculos individuais), o IoU do *dataset* (agrega-se união e interseção em todo conjunto de dados e em seguida calcula-se IoU).

A justificativa para o cálculo do IoU por imagem e o IoU do *dataset* é porque para conjuntos de dados com imagens “vazias” (imagens sem classe de destino), uma grande lacuna pode ser observada entre os dois tipos de cálculo. “Imagens vazias influenciam muito em IoU por imagem e pouco em IoU por *dataset*” [30].

Vale ressaltar que embora a acurácia não seja ideal em tarefas de segmentação semântica, ela pode ser utilizada em conjunto com as métricas de coeficiente Dice e IoU, as quais são úteis para avaliar a segmentação e localização de objetos em imagens, para comparar os resultados com os trabalhos relacionados na literatura.

IV. RESULTADOS

As análises são realizadas separadamente por tipo de lesão, para a fase de treinamento é considerado apenas uma lesão específica por vez contra a imagem de fundo. Após a etapa de treinamento, para um dos *encoders*, os valores de pesos aprendidos pela UNet++ foram salvos em um arquivo de *checkpoint*. Ao total, foram criados 16 arquivos de *checkpoints* para cada conjunto de dados, sendo 4 arquivos correspondentes aos *encoders* para cada um dos 4 tipos de lesões avaliadas. Esses arquivos foram utilizados para realizar os testes.

As métricas IoU por Imagem, IoU do *dataset*, Coeficiente Dice e Acurácia são calculadas para cada um dos *encoders* e os resultados são agrupados por lesão. Os resultados correspondentes ao banco de dados IDRIR dessas quatro métricas considerando os *encoders* MobileNet-v2, VGG16, EfficientNet-b3 e ResNet50 são apresentados na Tabela II.

Para cada métrica considerando o tipo de lesão, o melhor valor está destacado em negrito. Percebe-se que, para o banco de dados IDRID, os melhores resultados são obtidos usando a arquitetura VGG16 como *encoder*, alcançando resultados superiores principalmente na tarefa de segmentar hemorragias e exsudatos moles. Além disso, a complexidade para segmentar o disco óptico mostrou-se mínima em relação à segmentação das lesões da RD, isto é obtido devido o tamanho dessa região. Considerando apenas as lesões, os melhores resultados de segmentação são atribuídos aos exsudatos moles. Por outro

Tabela II: Resultados para o banco de dados IDRID

Métricas (%)	Microaneurismas			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	5,69	8,70	4,36	3,90
iou dataset	6,08	9,31	5,12	4,33
acurácia	99,85	99,86	99,87	99,78
dice	1,46	7,04	9,75	8,31
Métricas (%)	Exsudatos Duros			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	38,10	37,25	36,69	34,91
iou dataset	52,38	49,22	51,80	49,21
acurácia	99,36	99,34	99,38	99,32
dice	68,75	65,97	68,25	65,96
Métricas (%)	Exsudatos Moles			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	38,58	44,85	32,66	39,76
iou dataset	45,64	50,98	38,77	45,87
acurácia	99,75	99,77	99,76	99,75
dice	62,68	67,53	55,87	62,89
Métricas (%)	Hemorragias			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	26,94	28,99	28,90	20,96
iou dataset	32,89	36,04	31,01	23,44
acurácia	98,96	99,11	99,02	99,10
dice	49,50	52,98	47,34	37,98
Métricas (%)	Disco Óptico			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	93,07	92,91	93,07	92,30
iou dataset	93,15	93,22	93,06	92,38
acurácia	99,88	99,88	99,87	99,86
dice	96,46	96,49	96,41	96,04

lado, os resultados do *encoder* Efficientnet-b3 para microaneurismas foram os menores, devido a maior complexidade na identificação. Isso ocorre por conta do formato dessas lesões, conforme mostrado na Figura 2-d). As regiões são muito pequenas e acabam dificultando o aprendizado do modelo. Independente do tamanho da imagem, os microaneurismas sempre são representados em pequenas parcelas de pixels, com relação a toda a imagem.

Destaca-se que, mesmo a VGG16 tendo sido superior para microaneurismas, exsudatos moles e hemorragias, o *encoder* MobileNet-v2 foi melhor para exsudatos duros. Isso indica que a cada lesão específica pode exigir uma rede diferente para codificação.

Avaliando o banco de dados DIARETDB1, as demais redes de codificação se destacaram. Verifica-se que os melhores resultados são obtidos usando o *encoder* EfficientNet-b3, alcançando valores superiores para cada uma das métricas. Diferente do banco de dados IDRID, o DIARETDB1 não possui a máscara segmentada com a região do disco óptico, no entanto, essa região não é considerada tão importante para identificar a RD, visto que as principais características que acusam essa doença são as lesões, presentes nas demais regiões da retina — como os microaneurismas, exsudatos e hemorragias. Os resultados das análises individuais agrupadas por lesão do banco de dados DIARETDB1 são apresentados na Tabela III.

Comparando os resultados obtidos nos dois bancos de dados

Tabela III: Resultados o banco de dados DIARETDB1

Métricas (%)	Microaneurismas			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	2,13	2,74	1,73	1,99
iou dataset	10,98	03,58	04,47	05,27
acurácia	91,76	49,97	81,87	82,84
dice	19,78	06,91	08,56	10,01
Métricas (%)	Exsudatos Duros			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	43,42	65,66	60,33	64,24
iou dataset	53,95	56,92	55,48	57,05
acurácia	98,90	99,00	98,93	98,93
dice	70,09	72,55	71,36	72,65
Métricas (%)	Exsudatos Moles			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	45,58	24,45	62,71	55,16
iou dataset	16,63	15,53	19,01	18,47
acurácia	99,53	99,38	99,70	99,68
dice	28,51	26,88	31,95	31,19
Métricas (%)	Hemorragias			
	MobileNet-v2	VGG16	Efficient-b3	ResNet50
iou imagem	27,59	36,84	38,10	41,85
iou dataset	44,36	44,82	44,24	34,21
acurácia	98,66	98,77	98,49	98,55
dice	61,46	61,90	61,34	50,99

avaliados, verifica-se que o *encoder* VGG16 apresenta melhores resultados para hemorragias. Enquanto para os microaneurismas, a rede VGG16 apresenta melhores valores apenas no banco de dados IDRID, pois para DIARETDB1 o *encoder* MobileNet-v2 resulta em melhores resultados. Destaca-se ainda que para exsudatos moles no banco de dados DIARETDB1 os melhores resultados são com a rede Efficientnet-b3.

Contudo, ressalta-se que em ambos os banco de dados são encontradas limitações na detecção das lesões. Os microaneurismas e hemorragias por vezes são muito menores que a imagem original e em alguns casos se confundem com vasos sanguíneos e com a própria retina. Os exsudatos também apresentam certa complexidade na tarefa de segmentação.

Essa complexidade em detectar alguns tipos de lesões são vistas também em estudos como [14], [11] e [8], no qual indicaram que tiveram limitações para a detecção das lesões, seja por problemas de contraste da imagem, segmentar uma área maior do que a da lesão, ou por imagens ruidosas que levaram a erros de segmentação e classificação.

Uma sumarização dos trabalhos relacionados com os melhores resultados do método comparativo desenvolvido neste estudo é apresentado na Tabela IV. Alguns trabalhos apresentaram valores superiores de IoU e coeficiente Dice aos obtidos neste estudo, como os autores Schweisthal e Lascu [7]. Contudo, esses autores obtiveram esses valores ao combinar todas as lesões e simplificar a tarefa de segmentação para binária: com ou sem RD.

Devido à diversidade de imagens e banco de dados, em problemas de segmentação de lesões de retinopatia diabética os banco de dados públicos não possuem um padrão de imagem e de lesões especificadas, o que dificulta a comparação com trabalhos na literatura. O presente trabalho analisou 4

principais lesões características da RD, enquanto os autores como Qomariah, Tjandrasa e Faticah [8] e Jebaseeli, Durai e Peter [13] avaliaram apenas lesões específicas, os primeiros avaliaram Microaneurismas, enquanto os demais avaliaram lesões nos vasos sanguíneos para identificar a retinopatia diabética.

V. CONCLUSÃO

Este trabalho apresenta uma contribuição ao comparar diferentes redes codificadoras dentro da arquitetura UNet++ para segmentação de lesões da Retinopatia Diabética, tema importante e atual para a área de pesquisa. Os resultados obtidos permitem concluir que diferentes redes apresentam melhores desempenhos na segmentação de específicas lesões, evidenciando a complexidade de segmentar tais artefatos devido à variedade de tamanhos e de formas de imagens das lesões causadas pela retinopatia diabética.

Dessa forma, verifica-se que para diferentes lesões e também para diferentes conjuntos de dados, a melhor configuração para UNet++ varia de acordo com a lesão, com a VGG16 se destacando ao segmentar hemorragias e microaneurismas, a MobileNet-v2 se destacando para exsudatos duros e microaneurismas, e a EfficientNet-b3 sendo mais eficiente para exsudatos moles. Com isso, é possível selecionar as melhores ferramentas para auxiliar no diagnóstico assistido por computador da Retinopatia Diabética, o que acarreta uma redução de dificuldade no diagnóstico precoce, evitando assim maiores complicações, como a cegueira.

Para trabalhos futuros, sugere-se realizar comparações entre diferentes técnicas de pré-processamento de imagens, com o intuito de avaliar o impacto na segmentação das lesões. Além disso, é recomendado ampliar a quantidade de arquiteturas de segmentação de imagens para comparar sua eficácia na identificação das principais lesões associadas a retinopatia diabética.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. E com apoio do Laboratório de Engenharia de Sistemas de Computação (LESC) - UFC.

Tabela IV: Comparação de desempenho do método proposto com os trabalhos relacionados.

Referência	Bases	Pré-processamentos	Métodos	Lesões	Resultados
Schweisthal e Lascu [7]	DIARETDB	Aumento de dados, normalização	U-Net com: RESNET34, VGG16 e INCEPTIONV3 na codificação	Hemorragias, Exsudatos (moles e duros) e Microaneurismas	IoU: 98,64%, Dice: 99,15%
Qomariah, Tjandrasa e Faticah [8]	IDRID, DIARETDB	CLAHE, corte	MResUNet	Microaneurismas	Acc: 99,76%, Sen: 61,96%, Esp: 99,81%
Xu <i>et al.</i> [9]	IDRID	Corte, <i>denoising</i> , normalização, e aprimoramento de contraste	FFU-Net e MSFF	Hemorragias, Exsudatos (moles e duros) e Microaneurismas	Sen: 87,55%, IoU: 84,14%, Dice: 91,38%
Huang, Li, Xiao, Shen, e Xu [10]	IDRID e DDR	Redimensionamento, <i>Random Cropping</i> , <i>Random Flips</i> e CLAHE	Bloco transformador de relação (RTB) para incorporar mecanismos de atenção	Hemorragias, Exsudatos (moles e duros) e Microaneurismas	AUC_PR: 90,24%, AUC_ROC: 99,80%
Guo, Yang e Yu [11]	IDRID, DDR, DIARETDB, E-optha	Normalização (Filtro Gaussiano), e aprimoramento de contraste	Multiplas redes FCN com escalas de entrada diferentes	Hemorragias, Exsudatos, Microaneurismas e Manchas algodonosas	AUC_PR: 76,96%, Dice: 71,20%
Wang <i>et al.</i> [12]	IDRIR, DDR e E-Optha	Método Otsu aumento de dados	CLC-Net, SE-ResNeXt-50	Hemorragias, Exsudatos (moles e duros) e Microaneurismas	AUC_PR: 82,7%,
Jebaseeli, Durai e Peter [13]	STARE, DRIVE, HRF, REVIEW e DRIONS	CLAHE apenas cor verde	TPCNN, DLBSVM	Vasos Sanguíneos	Sens: 80,88%, Esp: 99,80%, Acc: 99,49%
Maaliw <i>et al.</i> [14]	DIARETDB	Redimensionamento, Normalização, e Aumento de Dados	DR-UNet	Hemorragias, Exsudatos (moles e duros) e Microaneurismas	IoU: 87,1% e Dice: 84,5%
Metodologia proposta	IDRID DIARETDB1	Aumento de dados, Normalização, <i>Random flips</i> , <i>RandomRGBShift</i> e aprimoramento de contraste	UNet++ com: Mobilenetv2, VGG16, Resnet50 e Efficientnetb3 no <i>backbone da rede</i>	Hemorragias, Exsudatos (moles e duros) e Microaneurismas	IoU: 65,66% Dice: 72,65%, Acc: 98,93%

REFERÊNCIAS

- [1] N Cheung and P Mitchell. Wong tylancet. *Diabetic retinopathy*, 376(9735):124–136, 2010.
- [2] Pouya Saeedi, Inga Petersohn, Paraskevi Salpea, Belma Malanda, Suvi Karuranga, Nigel Unwin, Stephen Colagiuri, Leonor Guariguata, Ayesha A Motala, Katherine Ogurtsova, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the international diabetes federation diabetes atlas. *Diabetes research and clinical practice*, 157:107843, 2019.
- [3] Judith Lechner, Olivia E O’Leary, and Alan W Stitt. The pathology associated with diabetic retinopathy. *Vision research*, 139:7–14, 2017.
- [4] S Jeyalakshmi, D Padmapriya, Divya Midhunchakkkravarthy, and Ali Ameen. Detection of hard exudate from diabetic retinopathy image using fuzzy logic. In *Intelligent Computing and Innovation on Data Science: Proceedings of ICTIDS 2019*, pages 543–550. Springer, 2021.
- [5] Alan W Stitt, Timothy M Curtis, Mei Chen, Reinhold J Medina, Gareth J McKay, Alicia Jenkins, Thomas A Gardiner, Timothy J Lyons, Hans-Peter Hammes, Rafael Simo, et al. The progress in understanding and treatment of diabetic retinopathy. *Progress in retinal and eye research*, 51:156–186, 2016.
- [6] Nikos Tsiknakis, Dimitris Theodoropoulos, Georgios Manikis, Emmanouil Ktistakis, Ourania Boutsora, Alexa Berto, Fabio Scarpa, Alberto Scarpa, Dimitrios I Fotiadis, and Kostas Marias. Deep learning for diabetic retinopathy detection and classification based on fundus images: A review. *Computers in biology and medicine*, 135:104599, 2021.
- [7] Brigitte Schweisthal and Mihaela Lascu. Semantic segmentation of diabetic retinopathy imaging. In *2022 International Symposium on Electronics and Telecommunications (ISETC)*, pages 1–4. IEEE, 2022.
- [8] Dinial Qomariah, ITS Nopember, H Tjandrasa, and C Faticah. Segmentation of microaneurysms for early detection of diabetic retinopathy using mresunet. *Int. J. Intell. Eng. Syst.*, 14:359–373, 2021.
- [9] Yifei Xu, Zhuming Zhou, Xiao Li, Nuo Zhang, Meizi Zhang, and Pingping Wei. Ffu-net: Feature fusion u-net for lesion segmentation of diabetic retinopathy. *BioMed Research International*, 2021, 2021.
- [10] Shiqi Huang, Jianan Li, Yuze Xiao, Ning Shen, and Tingfa Xu. Rtnet: Relation transformer network for diabetic retinopathy multi-lesion segmentation. *IEEE Transactions on Medical Imaging*, 41(6):1596–1607, 2022.
- [11] Tianjiao Guo, Jie Yang, and Qi Yu. Diabetic retinopathy lesion segmentation using deep multi-scale framework. *Biomedical Signal Processing and Control*, page 105050, 2023.
- [12] Xiyue Wang, Yuqi Fang, Sen Yang, Delong Zhu, Minghui Wang, Jing Zhang, Jun Zhang, Jun Cheng, Kai-yu Tong, and Xiao Han. Clc-net: Contextual and local collaborative network for lesion segmentation in diabetic retinopathy images. *Neurocomputing*, 2023.
- [13] T Jemima Jebaseeli, C Anand Deva Durai, and J Dinesh Peter. Retinal blood vessel segmentation from diabetic retinopathy images using tandem pcnn model and deep learning based svm. *Optik*, 199:163328, 2019.
- [14] Renato R. Maaliw, Zoren P. Mabunga, Maria Rossana D. De Veluz, Alvin S. Alon, Ace C. Lagman, Manuel B. Garcia, Luisito Lolong Lacatan, and Rhowel M. Dellosa. An enhanced segmentation and deep learning architecture for early diabetic retinopathy detection. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0168–0175, 2023.
- [15] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*, pages 3–11. Springer, 2018.
- [16] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *CoRR*, abs/1704.04861, 2017.
- [17] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.
- [18] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.
- [20] Thorvald A Sorensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on danish commons. *Biol. Skar.*, 5:1–34, 1948.
- [21] Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
- [22] Paul Jaccard. The distribution of the flora in the alpine zone. 1. *New phytologist*, 11(2):37–50, 1912.
- [23] Prasanna Porwal, Samiksha Pachade, Ravi Kamble, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabudde, and Fabrice Meriaudeau. Indian diabetic retinopathy image dataset (idrid): a database for diabetic retinopathy screening research. *Data*, 3(3):25, 2018.
- [24] Tomi Kauppi, Valentina Kalesnykiene, Joni-Kristian Kamarainen, Lasse Lensu, Iiris Sorri, A. Raninen, R. Voutilainen, Hannu Uusitalo, Heikki Kälviäinen, and Juhani Pietilä. Diaretdb1 diabetic retinopathy database and evaluation protocol. volume 2007, 01 2007.
- [25] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library.
- [26] Alexander Buslaev, Vladimir I. Iglovikov, Eugene Khvedchenya, Alex Parinov, Mikhail Druzhinin, and Alexandr A. Kalinin. Alumentations: Fast and flexible image augmentations. *Information*, 11(2), 2020.
- [27] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.
- [28] Carole H Sudre, Wenqi Li, Tom Vercauteren, Sebastien Ourselin, and M Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*, pages 240–248. Springer, 2017.
- [29] Dominik Müller, Iñaki Soto-Rey, and Frank Kramer. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes*, 15(1):1–8, 2022.
- [30] Pavel Iakubovskii. Segmentation models pytorch. https://github.com/qubvel/segmentation_models.pytorch, 2019.