EffiSegNet: Rede de segmentação automatizada para suporte ao diagnóstico preciso de queilite actínica e carcinoma espinocelular de lábio em imagens clínicas.

Palavras-Chave: Redes neurais, Fotografias clínicas, Segmentação.

Autores(as):

Hugo Abritta Reis, FOP - UNICAMP

Vitor Simbalista Teixeira Soares, ICT - UNIFESP

Anna Luíza Damaceno Araújo, FMUSP

Ivan José Correia Neto, FOP - UNICAMP

Cristina Saldivia Siracusa, FOP - UNICAMP

Matheus Cardoso Moraes, ICT - UNIFESP

Alan Roger Santos Silva, FOP - UNICAMP

Márcio Ajudarte Lopes (Orientador), FOP - UNICAMP

INTRODUÇÃO:

A queilite actínica (QA) é uma desordem potencialmente maligna do vermelhão do lábio que acomete principalmente indivíduos de pele branca submetidos a exposição solar crônica. Dentre as características clínicas da QA, está incluída a perda de definição entre a pele e o vermelhão do lábio. A presença de úlceras persistentes e de difícil cicatrização pode ser indicativa de transformação maligna em casos de QA (**Figura 1**). O objetivo deste trabalho foi implementar uma arquitetura de segmentação de imagens (EffiSegNet) para lesões em lábio.

Figura 1. Imagens clínicas de QA e carcinoma espinocelular de lábio (CECL).

METODOLOGIA:

Este estudo retrospectivo desenvolveu um modelo de segmentação com base em 414 fotografias digitais de 315 pacientes com QA e CECL, atendidos no Serviço de Diagnóstico Oral (OROCENTRO) da Faculdade de Odontologia de Piracicaba (FOP/UNICAMP). A segmentação manual foi realizada por dois especialistas (H.A.R. e I.J.C.N) utilizando o software Aperio Image Scope e a mesa digitalizadora Huion Inspiroy H1060P. Uma terceira especialista (A.L.D.A) revisou as segmentações para reduzir vieses, sendo as discordâncias resolvidas por consenso.

O modelo escolhido para a tarefa de segmentação foi o EffiSegNet-B0[1] baseado em técnicas de aprendizado por transferência, utilizando pesos pré-treinados no conjunto de dados ImageNet. Devido à natureza específica das imagens clínicas, em que transformações como espelhamento vertical ou alterações de escala poderiam comprometer a relação anatômica dos elementos, não foram aplicadas técnicas de aumento de dados (data augmentation).

A estratégia de validação cruzada 10-fold[2] foi utilizada para estimar o intervalo de confiança das métricas de teste. O conjunto de dados original (n = 414) foi dividido em 10 subconjuntos de 41 ou 42 imagens cada. Em cada iteração, os dados foram particionados de forma que 90% (372 imagens) fossem usados para treinamento e os 10% restantes (42 imagens) para teste, garantindo que todas as imagens fossem utilizadas pelo menos uma vez na fase de teste ao longo dos 10 folds. A divisão 90-10 foi escolhida por ser uma estratégia eficaz para maximizar o desempenho e a capacidade de generalização do modelo, especialmente considerando o tamanho limitado do conjunto de dados.

RESULTADOS E DISCUSSÃO:

O modelo atingiu uma notável acurácia de segmentação superior a 95%, com valores de sensibilidade (recall) e precisão de 88% e 90%, respectivamente, demonstrando sua capacidade de identificar corretamente as regiões relevantes nas imagens. Os coeficientes IoU e Dice (F1-score) indicam um desempenho favorável na detecção de regiões afetadas, atingindo 80% e 89%, respectivamente (**Tabela 1**).

Tabela 1. Comparação dos resultados de segmentação do modelo EffiSegNet-B0 (presente estudo) com os resultados do estudo de Vezakis.

Métricas	**Trabalho atual**	**Intervalo de confiança 95% (Trabalho atual)**	**KVasir-SEG [1]**
Acurácia	0,959 ± 0,003	[0,956 - 0,961]	n/a
Sensibilidade (Recall)	0,88 ± 0,02	[0,87 - 0,90]	0,9368
Precisão	0,90 ± 0,02	[0,89 - 0,92]	0,9475
IoU	0,80 ± 0,02	[0,78 - 0,81]	0,8794
Dice (F1-score)	0,89 ± 0,01	[0,89 - 0,90]	0,9304

IoU: *Intersection over Union.*

Os bons resultados obtidos podem ser visualmente confirmados pelas máscaras de predição geradas no conjunto de teste (**Figura 2**). Embora alguns contornos apresentem pequenos artefatos e irregularidades, as máscaras preditas exibem uma sobreposição significativa com os padrões ouro, especialmente nos quatro primeiros exemplos. No último exemplo, a segmentação foi um pouco menos precisa, provavelmente devido à iluminação limitada e à presença de um artefato incomum (o dedo enluvado de um profissional de saúde). Ainda assim, os lábios acometidos pela lesão foram corretamente isolados do restante da imagem.

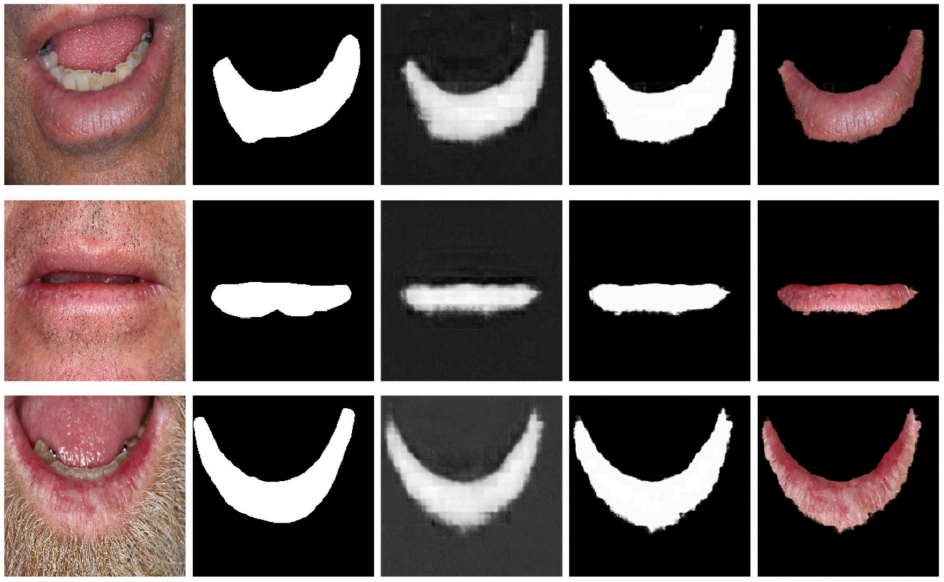

Figura 2. Resultado visual da segmentação. Da esquerda para a direita: imagem original, máscara original, saída bruta, máscara predita e região segmentada.

CONCLUSÕES

O EffiSegNet foi aplicado a um novo conjunto de dados focado em lesões labiais, explorando sua adaptabilidade e utilidade clínica. Os resultados corroboram com a hipótese de que arquiteturas compactas são capazes de, simultaneamente: (i) reduzir a complexidade computacional das tarefas e (ii) preservar sua eficácia mesmo em cenários desafiadores. A aplicação de arquiteturas convolucionais mais eficientes e enxutas, como o EffiSegNet, pode trazer ganhos significativos para o prognóstico de diversas condições clínicas.

BIBLIOGRAFIA

1. Vezakis IA, Georgas K, Fotiadis D, Matsopoulos GK. EffiSegNet: Gastrointestinal Polyp Segmentation through a Pre-Trained EfficientNet-based Network with a Simplified Decoder. 2024.

2. Berrar D. Cross-Validation. In: Ranganathan S, Gribskov M, Nakai K, Schönbach C, editors. Encyclopedia of Bioinformatics and Computational Biology. Elsevier; 2019. p. 542–5. https://doi.org/10.1016/B978-0-12-809633-8.20349-X