

Detecção de Glaucoma em Imagens de Fundo de Olho através de Triplet Loss

Lisle Faray de Paiva, Gabriel Sousa Silva Costa, Geraldo Braz Junior

¹Vision and Image Processing Laboratory - Núcleo de Computação Aplicada
Universidade Federal do Maranhão
Av. dos Portugueses, 1966 - Vila Bacanga, São Luís - MA, Brasil

{lisle, geraldo, gss.costa}@nca.ufma.br

Resumo. *O glaucoma, a maior causa de cegueira irreversível do mundo, ocorre quando há um aumento da pressão intraocular, danificando as fibras do nervo óptico. Seu diagnóstico se dá através de imagens de retinografia. Tais imagens são difíceis de distinguir um olho normal de um olho doente, o que prejudica a detecção automática. Este trabalho, visando melhorar a distinção entre as classes doente e saudável, propõe uma metodologia de detecção de glaucoma capaz de diferenciar eficientemente as classes através do mapeamento de espaços euclidianos.*

1. Introdução

O glaucoma é a segunda principal causa de cegueira em todo o mundo, perdendo somente para a catarata. De acordo com [Tham et al. 2014], estima-se que em 2020, 76 milhões de pessoas serão afetadas pela doença e, em 2040, os números poderão chegar à 111,8 milhões de pessoas. Seu diagnóstico pode ser dado a partir da análise de imagens de fundo de olho, uma vez que a quantidade de perda de fibras do nervo óptico tem um efeito direto na configuração da rima neural.

Imagens de fundo de olhos saudáveis são muito similares a imagens de olhos glaucomatosos o que pode dificultar uma Rede Neural de as distinguir. Neste trabalho analisamos a função de custo Triplet Loss para detecção de glaucoma mapeando espaços euclidianos onde distâncias correspondem à medida de similaridade entre as classes (glaucoma e normal) e então utiliza-se os espaços feitos (embeddings) como vetores de características.

2. Metodologia

Este trabalho apresenta uma metodologia para o diagnóstico de glaucoma utilizando embeddings como vetores de características. Primeiramente passa-se as imagens da base de dados por uma rede neural convolucional sendo 80% da base para treino e 20% dividido entre validação e teste que irá aprender como diferenciar as classes e então gerar os embeddings. Com os embeddings gerados, utiliza-se o AutoSklearn para a classificação.

A primeira etapa da metodologia consiste do treinamento da arquitetura para gerar os embeddings euclidianos. O treinamento começa alimentando a base de dados RIM-ONE à arquitetura *Deep* que então, baseado em [Schroff et al. 2015], utiliza a função de loss Triplet para aprender a diferenciar as classes. Uma vez que os embeddings são gerados, eles são passados para o Auto-Sklearn os classificar, uma vez que o Auto-Sklearn automaticamente encontra o melhor algoritmo para um determinado *dataset* e otimiza seus parâmetros.

3. Resultados

Os embeddings são gerados a partir de uma rede neural convolucional treinada para diferenciar as classes utilizando a função de loss Triplet testando diversos tamanhos de espaço. Na Figura 1 pode-se ver, na esquerda, o espaço euclidiano antes do treinamento onde as duas classes estão misturadas entre si, já na direita pode-se observar uma melhor diferenciação de ambas as classes.

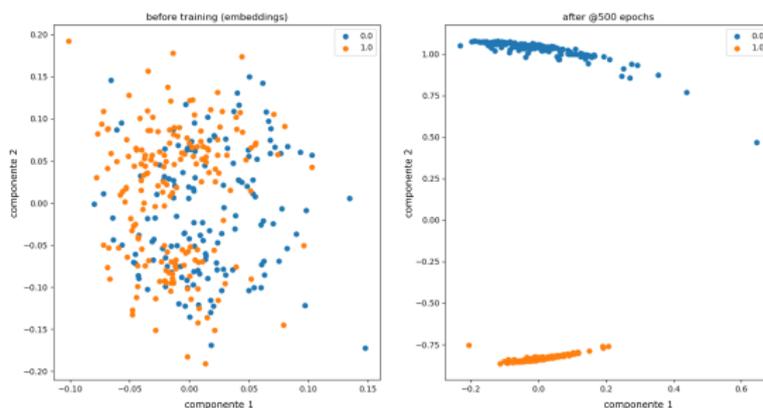


Figura 1. Apresentação das features geradas, para toda a base de teste, antes e depois da rede neural, usando PCA para apresentação em 2D

Utilizando os embeddings como vetores de características, os classificamos através do autoklearn. Na Tabela 1 pode-se observar os 3 tamanhos de espaços que tiveram melhores métricas ordenados por maior acurácia.

Tabela 1. Resultados da classificação dos embeddings

Qtd. Embedding	Sensibilidade	Especificidade	Precisão	Acurácia	F1 Score
22	95.12%	82.14%	88.64%	89.86%	91.76%
18	88.89%	83.33%	90.91%	86.96%	89.89%
58	88.89%	83.33%	90.91%	86.96%	89.89%

4. Conclusão

A metodologia proposta obteve como melhor resultado a sensibilidade de 95,12% e acurácia de 89,96% o que indica não só o sucesso na diferenciação entre as classes como também a melhora na identificação de cada classe, uma vez que, como visto na Figura ?? antes do treinamento cada classe estava muito espalhada entre seus indivíduos e, após o treinamento, não só foi possível discernir melhor cada classe como também os indivíduos de mesma classe se agruparam. Em trabalhos futuros pretende-se avaliar opções da arquitetura da rede neural utilizada, já que foi utilizada uma rede simples e também realizar validação cruzada para confirmar os resultados.

Referências

- Schroff, F., Kalenichenko, D., and Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823.
- Tham, Y.-C., Li, X., Wong, T. Y., Quigley, H. A., Aung, T., and Cheng, C.-Y. (2014). Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*, 121(11):2081–2090.