SIIM CMIMI24
Oct 21-22 | Boston, MA
Conference On Machine Intelligence In Medical Imaging

Adversarial Domain Adaptation for Robust Glaucoma Classification

Homa Rashidisabet, PhD Candidate, University of Illinois Chicago; R. V. Paul Chan, MD, MSc, FACS; Thasarat S. Vajaranant, MD, MHA; Darvin Yi, PhD

Introduction/Background

While deep learning models have shown promising results in automated glaucoma prediction, they lack robustness when faced with a data domain shift. In real-world medical settings, discrepancies among images from different hospitals and patient populations create a domain shift, adversely impacting model performance and generalization across diverse datasets.

Methods/Intervention

We propose a Domain Adaptation (DA) method to address the performance degradation of deep learning models under data shift in glaucoma classification. Implementing Ganin et al., 2015, our deep learning architecture incorporates an adversarial learning component to eliminate domain-specific information, enabling the model to learn invariant features across data domains. The DA model is jointly trained on glaucoma classification and domain classification tasks, minimizing differences between domains while optimizing for the main task. We validated the DA method against the state-of-the-art ResNet-50 model using fundus images from LAG (n=4854), REFUGE (n=249), and University of Illinois Chicago (UIC) dataset (n=711) as both source and target data.

Results/Outcome

Using the UIC dataset as the source, the source-only model achieves 71.7% accuracy on LAG and 82.0% accuracy on REFUGE. The DA model improves these by 12.3% and 6.0%, achieving 84.0% and 88.0% accuracy, respectively. The train-on-target models, which serve as a reference, represent the upper bound on DA performance, while the source-only model without adaptation indicates the lower bound. When swapping source and target, the source-only model achieves 72.0% and 70.5% accuracy on UIC. The DA model improves these by 12.5% and 7.0%, achieving 84.5% and 77.5% accuracy on UIC.

Conclusion

Our proposed DA method effectively addresses the performance degradation of deep learning models when there is a shift in data. By learning domain-invariant features and unlearning domain-specific information, the DA method significantly improves the performance of the state-of-the-art ResNet-50 model in glaucoma classification.

Statement of Impact

Glaucoma, affecting eighty million people worldwide, is a leading cause of blindness. Deep learning has improved glaucoma prediction using fundus images, but performance degrades with data shifts from varied sources. Addressing this degradation is crucial to ensure clinical reliability and avoid patient risks. We propose a method to maintain performance across diverse datasets by adapting to data variations, validated on three datasets.

Method	Source: UIC Target: LAG	Source: UIC Target: REFUGE	Source: LAG Target: UIC	Source: REFUGE Target: UIC
Source only	71.7% [69.1%, 69.7%]	82.0% [77.5%, 87.5%]	72.0% [69.7%, 74.8%]	70.5% [68.9%, 73.1%]
Domain Adaptation	**84.0% [83.9%, 84.3%]**	**88.0% [85.0%, 92.5%]**	**84.5% [83.2%, 86.6%]**	**77.5% [75.6%, 79.8%]**
Train on target	96.2% [96.1%, 96.4%]	94.0% [92.5%, 97.5%]	80.0% [78.2%, 81.5%]	80.0% [78.2%, 81.5%]

Table 1. Glaucoma classification accuracy results for different source and target domains. The first row shows the performance with no adaptation (lower bound), while the last row represents training with target domain data and known labels (upper bound).

Keywords

Glaucoma classification; Domain adaptation; Computer vision; Fundus analysis