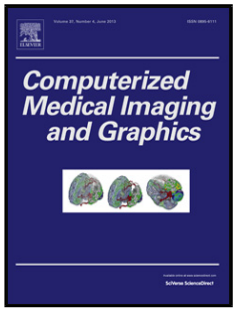

Title: Efficient multi-kernel multi-instance learning using weakly supervised and imbalanced data for diabetic retinopathy diagnosis

Author: Peng Cao Fulong Ren Chao Wan Jinzhu Yang Osmar Zaiane

PII: S0895-6111(18)30479-8
DOI: https://doi.org/doi:10.1016/j.compmedimag.2018.08.008
Reference: CMIG 1583

To appear in: *Computerized Medical Imaging and Graphics*

Received date: 1-1-2018
Revised date: 9-7-2018
Accepted date: 22-8-2018

Please cite this article as: Peng Cao, Fulong Ren, Chao Wan, Jinzhu Yang, Osmar Zaiane, Efficient multi-kernel multi-instance learning using weakly supervised and imbalanced data for diabetic retinopathy diagnosis, *<![CDATA[Computerized Medical Imaging and Graphics]]>* (2018), https://doi.org/10.1016/j.compmedimag.2018.08.008

Efficient multi-kernel multi-instance learning using weakly supervised and imbalanced data for diabetic retinopathy diagnosis

Peng Cao[a], Fulong Ren[a,b], Chao Wan[c,*], Jinzhu Yang[a,b], Osmar Zaiane[d]

[a]*Computer Science and Engineering, Northeastern University, Shenyang, China*
[b]*Key Laboratory of Medical Image Computing of Ministry of Education, Northeastern University, Shenyang, China*
[c]*Department of Ophthalmology, the First Hospital of China Medical University, Shenyang, China*
[d]*Computing Science, University of Alberta, Edmonton, Alberta, Canada*

Abstract

Objective: Diabetic retinopathy (DR) is one of the most serious complications of diabetes. Early detection and treatment of DR are key public health interventions that can significantly reduce the risk of vision loss. How to effectively screen and diagnose the retinal fundus image in order to identify retinopathy in time is a major challenge. In the traditional DR screening system, the accuracy of micro-aneurysm (MA) and hemorrhagic (H) lesion detection determines the final screening performance. The detection method produced a large number of false positive samples for guaranteeing high sensitivity, and the classification model was not effective in removing false positives since the suspicious lesions lack label information.

Methods: In order to solve the problem of supervised learning in the diagnosis of DR, we formulate weakly supervised multi-class DR grading as a multi-class multi-instance problem where each image (bag) is labeled as healthy or abnormal and consists of unlabeled candidate lesion regions (instances). Specifically, we proposed a multi-kernel multi-instance learning method based on graph kernel. Moreover, we develop an under-sampling from instance level and over-sampling from bag level to improve the performance of the multi-instance learning in the diagnosis of DR.

Results: Through empirical evaluation and comparison with different baseline

*Corresponding author.
Email address: caopengneu@gmail.com (Chao Wan)