Multimedia Tools and Applications (2024) 83:89837–89870
https://doi.org/10.1007/s11042-024-18968-5

Grading the severity of diabetic retinopathy using an ensemble of self-supervised pre-trained convolutional neural networks: ESSP-CNNs

Saeed Parsa[1] · **Toktam Khatibi**[2]

Received: 23 October 2023 / Revised: 22 December 2023 / Accepted: 13 March 2024 /
Published online: 2 April 2024

Abstract
Diabetic retinopathy (DR) is a common eye disorder that can lead to vision problems and blindness, necessitating accurate grading for effective treatment. While various artificial intelligence (AI) systems have been developed, surpassing human analysis in detecting DR, deep neural networks require large annotated datasets to learn the complex patterns and relationships necessary for grading, which are often limited in availability, to learn the intricate patterns and relationships required for accurate grading. However, such datasets are often limited in availability, requiring significant investments of human resources and time for the labeling process. To address these challenges, we propose ESSP-CNNs, a framework that harnesses popular CNN architectures (VGGNet, AlexNet, and ResNet). Our approach employs self-supervised learning, specifically the Bootstrap Your Own Latent (BYOL) technique, to pre-train neural networks on a vast unlabeled dataset. Additionally, we employ deep ensemble learning to construct a robust model for DR grading. Our methodology encompasses three main components: preprocessing fundus images, BYOL-based pre-training, and ensemble model construction. We conduct experiments and comparisons using the EyePACS and IDRiD datasets, with BYOL pre-training on EyePACS to enable the CNN models to acquire meaningful representations of fundus images, while IDRiD is used for severity grading. The performance of the proposed framework is further confirmed through thorough validation using the Messidor dataset. Through extensive experimentation on the IDRiD and Messidor datasets, ESSP-CNNs achieve notable accuracies of 71.84% and 75.42%, specificities of 88.76% and 87.13% along with AUC of 86.02% and 86.54%, respectively. The experimental results validate the effectiveness of our methodology in grading the severity of DR, with the ensemble model built from pre-trained CNNs yielding promising outcomes. Moreover, we compare our methodology against other state-of-the-art methods in DR grading, and our results demonstrate its satisfactory performance, surpassing previous alternatives in accurately assessing DR severity.

Keywords Diabetic retinopathy · Deep learning · Convolutional Neural Networks · Self-supervised learning · Ensemble learning · Color fundus images

Extended author information available on the last page of the article