

Deep Learning for Early Detection of Diabetic Retinopathy

Anurag J. Vaidya, Department of Biomedical Engineering, Bucknell University, PA

Joshua V. Stough, Department of Computer Science, Bucknell University, PA

Introduction: Diabetic retinopathy (DR)— a leading cause of blindness— is a diabetes complication which causes damage to retinal blood vessels [1]. DR can be treated non-invasively if diagnosed early enough [2]. However, early diagnosis requires a medical examination, which may not be possible in some demographical regions. Much previous work has largely focused on determining the presence of DR and not the severity [3]. Thus, the goal of this study was to develop a single neural network that could 1) detect presence or absence of DR 2) perform early detection of DR 3) perform multi-class classification on DR severity.

Materials and Methods The APTOS DR dataset used contains of 3,681 fundus images with a 0-4 rating [4]. The “0” rating indicates no DR and rating of 4 signifies severe proliferative DR. For all models, images were resized to 256x256 pixels and normalized. For the first model (binary classifier), the image labels were changed to “0” if the image had no DR, or “1” otherwise. For the second model (early detector), images with no DR were labelled “0,” “1” if they had early DR (corresponding to “1” and “2” from original dataset), or “2” if they had severe DR (corresponding to “3” and “4” from original dataset). For the final model (severity classifier), the original labels were maintained. For all models, we performed a five-fold cross validation study. The same neural network architecture was used for all three experiments: three convolution layers, one global average pooling layer, and three fully connected layers. With each convolutional layer, the image size was halved. All models trained for 40 epochs. Training and validation loss and accuracy were tracked. For the binary classifier, ROC curves were created for each split. For the early detector and severity classifier, confusion matrices were created for each split.

Results and Discussion: Across all five splits, the binary classifier provided average validation accuracy of 96.2% and average area under the curve (AUC) of 0.992 (Figure 1A). Our binary classifier achieved higher AUC score than current literature models [3]. The early detector provided an average validation accuracy of 86%. The confusion matrix suggests that the early detector is able to accurately predict early onset of DR, however, is unable to differentiate between early onset and severe DR. Finally, the severity classifier model achieved an average validation accuracy of 79.4% across all five splits. From Figure 1C, we see that the model is able to identify healthy subjects very accurately but cannot classify the more severe cases. This may happen because deeper in the model, the aperture of the convolutional layers is larger than the minute features distinguishing early onset DR from severe DR.

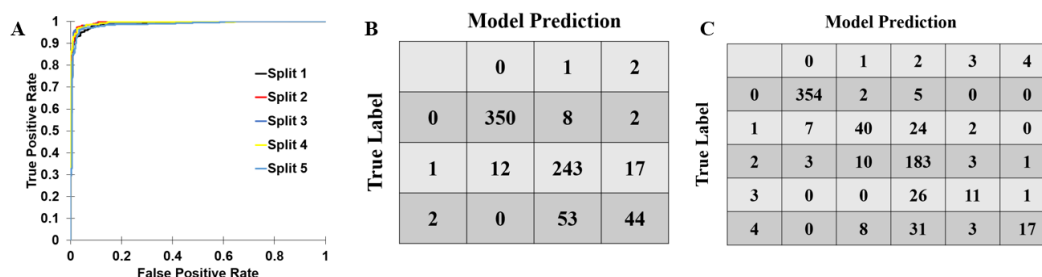


Figure 1: (A) ROC curves for binary classifier model for all five splits. Example confusion matrices of one stratified split for (B) early detector and (C) severity classifier

Conclusions: Our study presents a single neural network run on three experiments (binary classifier, early detector, and severity classifier). Our binary classification model provides better predictive capabilities than state-of-the-art literature models. The early detector and severity detector do not provide as high accuracy. Future works include studying residual networks in which channels from initial convolutional layers are fed to deeper layers in the model to better predict the early onset of DR and severity of DR.

References:

- [1] Viswanath, K. Comm. Eye Health, 2003; 16(46): 21–24.
- [2] NHS UK, Diabetic Retinopathy Treatment, 2018
- [3] Gargeya, R. PubMed, 2017; 124(7):962-969
- [4] <https://www.kaggle.com/c/aptos2019-blindness-detection/leaderboard>