Deep learning based binary image quality algorithm for low-cost fundus imaging system in remote care settings

Taylor Shagam, BS[1]; Minhaj Alam, PhD[1]; Simon Bello, PhD[1]; Laura Tracewell, OD[1]; Bryan Rogoff, OD[1]; Patricia Sha, OD[1]; Nolleisha Graves, COA[1]; Dorothy Hitchmoth, OD[2]; Mary K. Durbin, PhD[1]; Niranchana Manivannan, PhD[1]

[1]Carl Zeiss Meditec Inc, Dublin, CA, USA; [2]Dr. Dorothy L. Hitchmoth, PLLC., New London, NH, USA

Poster #3544717

PURPOSE

An image quality (IQ) algorithm is important in assisting medical assistants with little to no experience in ophthalmology to judge whether an image is of sufficient quality for clinical diagnosis. In this study, we developed a binary IQ classification for providing real-time feedback in a remote care or primary care setting.

METHODS

Data:

- 5173 images acquired using VELARA™ 200 (Zeiss, Dublin, CA) camera from a retrospective study were graded by 2 graders and adjudicated by an optometrist.
- Images were acquired from healthy and diseased subjects with various retinal pathologies and graded for quality of readable clinical information on a 1-5 scale (1-very poor and 5-excellent). Images were annotated as good quality if the fovea, optic nerve, and superior and inferior arcades were all visible (no more than 5% of the image is unreadable).
- The ground truth was determined by converting the gradings for binary: 0-Insufficient if IQ is <=2 (65.6% of 5173 images) and 1-Sufficient if IQ is >2 (34.4%).
- Small pupils, incorrect fixation, lack of focus and other artifacts were the main causes of insufficient IQ.

Algorithm:

- The dataset was split into three: i) training- 3646, ii) validation- 928 and iii) hold-out test- 599. The training set was augmented using flip and brightness adjustments. The deep learning architectures used to train the IQ classification are shown in Figure 1.
- ImageNet weights were preloaded and sigmoid activation with binary cross entropy loss were used. The data was resampled and reweighted to account for the class imbalance.
- For hyperparameter tuning, all the networks were trained using Adam with cyclic learning rate scheduler and Stochastic Gradient Descent (SGD) with Nesterov momentum. SGD with Nesterov was chosen for the final model as it performed better.
- Sensitivity, specificity and execution time were compared to select the final IQ model.

taylor.shagam@zeiss.com

Disclosures: TS (E), MA (C), SB (E), LT (E), BR (E), PS (E), NG (C), DH (C), MD (E), NM (E) : Carl Zeiss Meditec, Inc.

RESULTS

- The performance of all networks in hold-out test set is shown in Figure 1. VGG-16-provided high sensitivity, specificity with lower execution time and was selected as the final model.
- VGG-16 achieved 99% sensitivity, 91% specificity and 220ms execution time in i5-10400H CPU. Figure 2 shows some of the example results from IQ algorithm.

A

Model	Sensitivity	Specificity	Execution time
VGG -16	99%	91%	220ms
Inception Resnet V2	98%	91%	578ms
Inception V3	97%	90%	432ms
ResNet50	97%	91%	290ms
EfficientNet B0	96%	89%	195ms

B

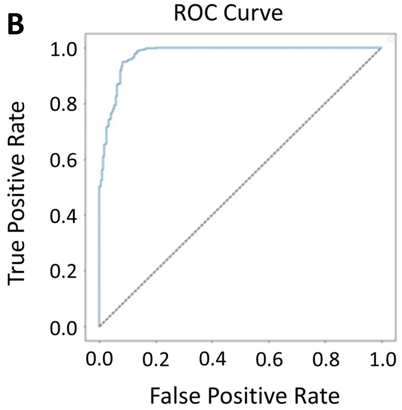

Figure 1.
A) Performance of the various networks.

B) Receiver operator characteristic curve for the final VGG-16 model.

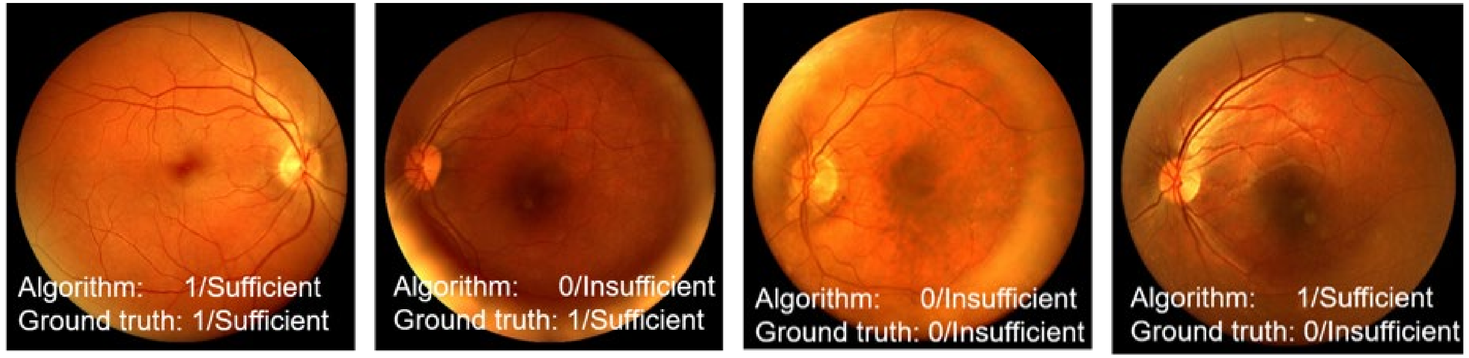

Figure 2. Sample results from proposed algorithm. The images from left to right show examples of true positive, false negative, true negative, and false positive.

CONCLUSIONS

We developed an image quality algorithm with 99% sensitivity and 91% specificity with an execution time to provide real-time feedback to the operator on whether to retake the fundus images.