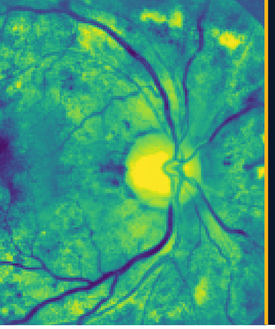


ADAPTIVE HISTOGRAM EQUALIZATION IN DIABETIC RETINOPATHY DETECTION

Daniela A. Parletta danielaangela.parletta@akkodisgroup.com
Giovanna Purgato giovanna.purgato@akkodisgroup.com

ARTIFICIAL INTELLIGENCE & DATA SCIENCE UNIT

DIABETIC RETINOPATHY DETECTION

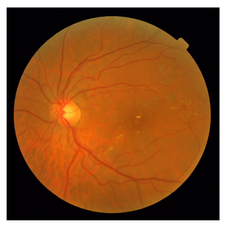
We consider the **Diabetic Retinopathy Detection problem** where, given a raw image of the eye fundus, the objective is to automatically detect the presence of a Diabetic Retinopathy (DR) disease. Such systems are usefull to support the specialists in the diagnosis and the treatment of this pathology.

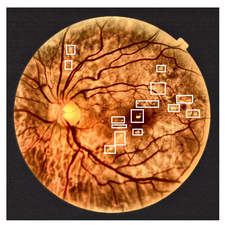
- According to the most recent medical standards, DR can be classified into 5 different stages with 0 and 1 corresponding to healthy subjects, and the other corresponding to the presence of DR ordered by severity.
- Detection is the binary classification version of this problem, where one only want to distinguish between an healthy subject (stage 0 or 1) from a sick subject (stage 2,3 or 4).
- Deep learning methods have showed to be state-of-the-art for this problem.

REMARK — The performances of deep learning methods for DR detection are extremely sensitive to an appropriate equalization of the input images.

IMAGE EQUALIZATION

Image Equalization. Process the image pixel distribution to make it as close as possible to the uniform distribution. This usually results in an improved **perceived contrast**. In the context of DR detection, equalization highlights fine-grained details of the eye fundus, such as micro-aneurysms, a type of lesion that is typical of the advanced DR stages.





On the left, a raw eye fundus image of a subject suffering from a severe diabetic retinopathy. On the right, the equalized image where there are visible and highlighted in the white edged boxes many of the typical lesions caused by the pathology.

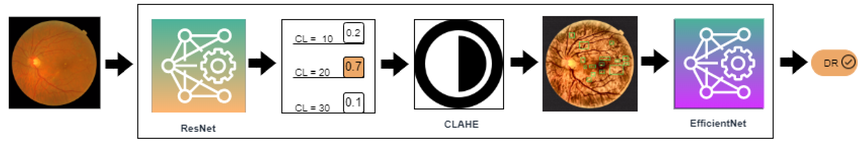
PROBLEM — Many state-of-the-art algorithms for image equalization require input parameters that are tuned ad-hoc through computationally expensive grid-search procedures. Moreover, many of the proposed quality criteria are not directly related to the predictive accuracy of the detection system that uses the equalized images.

Contrast limited adaptive histogram equalization (CLAHE) is one of the most popular algorithms for image equalization. It is based on local equalizations performed at the level of images patches called tiles; the tile size (TS) is a parameter. Additionally it relies on a second parameter, known as clip limit (CL), to limit the level of introduced contrast.

PROPOSED APPROACH

IDEA — We train a deep network to predict, given a raw image and a specific image equalization algorithm, which parameters should be used by the algorithm for the equalization. This network, is trained jointly with the predictive deep learning model used for the DR detection, so its predictions are optimized for the final accuracy.

ADAPTIVE HISTOGRAM EQUALIZATION WITH **CLAHE**



The overall system is composed of a series of two networks, a ResNet and an EfficientNet. After the system is trained in an end-to-end fashion, a raw image is fed to ResNet which selects the CL for the CLAHE. Second, the image is equalized with the tuned CLAHE method; finally, the processed image is fed into EfficientNet for retinopathy classification. The green bounding boxes reported on the equalized image have an illustrative purpose and are not applied in practice: they illustrates the many fine-grained details highlighted by the equalization process.

- The first network is implemented with a relatively simple model in order to not excessively increase the inference time.
- ☹ Training time is increased w.r.t. to a traditional DR detection method.
- ☺ The prediction time is strongly reduced since the computation of the equalization parameters is made through a feedforward pass, instead that an expensive search procedure.
- ☺ The overall accuracy can be even better than a grid-search based system since equalization is optimized for the predictions.

REMARK — Joint training is realized by replacing, at training time only, the hardthresholding implemented by the output layer of the first network with a **Gumbel Soft-max Activation**.

GUMBEL SOFT-MAX ACTIVATION

$$y_{CL,j} = \frac{\exp(\log_2(p_j) + g_j)/\tau}{\sum_{j=1}^{N} \exp(\log_2(p_j) + g_j)/\tau}$$

- This activation is differentiable and then allows for the use of back-propagation.
- It approximates the hard-thresholding at a degree controlled by the temperature $\tau > 0$. Small values of τ lead to very good approximations of the hard-threshold; large values of τ lead to poor approximations but to a greater numerical stability of the gradients.

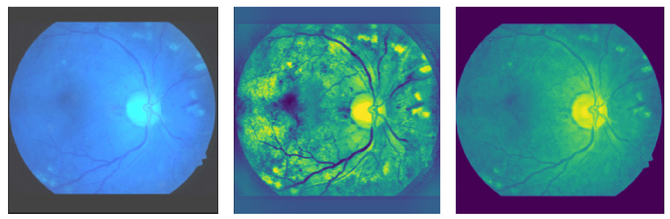
EMPIRICAL RESULTS

Experiments with real-world datasets (MESSIDOR-2, IDRiD) using CLAHE for image equalization.

METHOD	DATA	AC	TIME (S)
No-CLAHE	D1	0.810	**0.799**
Tuned-CLAHE	D1	**0.859**	20.656
Predicted-CLAHE	D1	0.842	1.753
No-CLAHE	D2	0.702	**1.359**
Tuned-CLAHE	D2	0.721	38.844
Predicted-CLAHE	D2	**0.869**	1.879

Method: No-CLAHE corresponds to not making any equalization; Tuned-CLAHE refers to a grid-search procedure to find the parameters for CLAHE; Predicted-CLAHE denotes the proposed method. In **Data**, D1 referes to training on MESSIDOR-2 and testing on IDRiD; D2 the vice-versa. **Ac** stands for accuracy. **Time** is the average inference time across the predictions made on the test set.

LIMITATIONS OF THE ENTROPY BASED CRITERIA



On the left, a raw eye fundus image of a subject suffering from a severe diabetic retinopathy. In the middle, the equalized image with CL = 19 as predicted by our method Predicted-CLAHE. On the right, the equalized image with CL = 2 as tuned by Tuned-CLAHE.
Notice that Tuned-CLAHE, differently from Predicted-CLAHE, miss-classifies this image as that of an healthy subject. This comparison shows that while the first approach is beneficial to improve the perceived quality of the image, it may lead to sub-optimal predictive performance, since it doesn't sufficiently highlight the finer details (e.g. many micro-aneurysms present in the eye fundus of this patient) that are important to detect the pathology.

Daniela A. Parletta Giovanna Purgato

AKKODIS