**Dataset Eyes-defy-anemia**

The dataset Eyes-defy-anemia contains 218 images of eyes, in particular conjunctivas, which can be used for research on the diagnosis/estimation of anemia based on the pallor of conjunctiva.

The same images can be effectively used to study segmentation algorithms of conjunctivas or exposed parts of the sclera and iris. All images of the dataset are accompanied by segmented elements (palpebral, forniceal and palpebral + forniceal conjunctivas) useful both to directly correlate the pallor with the value of Hb, to assess the performance of segmentation algorithms. Each image is accompanied by essential information such as the value of Hb measured in the laboratory, age, and sex of the patient, listed in xlsx files.

The images were captured with a Samsung S6 smartphone, equipped with a particular device that magnifies images, standardizes lighting, and eliminate the influence of ambient light. So, all the images were captured at the same distance and with the same white LED light. The description of the acquisition set is included in our papers linked to the dataset.

IMPORTANT NOTE: The segmentation of the parts of the eye was carried out by experienced personnel according to their own interpretation: every user of the images of the dataset can use our segmentations or use the whole images of the eye and segment them as he sees fit.

The acquisition and preparation of this dataset has required a lot of work without any remuneration. We provide it free of charge, but we ask those who intend to use our dataset the courtesy to quote the following papers (thanks in advance):

- R. Maglietta, M. E. Griseta, F. Clemente, A. Guarini, G. Dimauro, Advances in the Automated Diagnosis of Anemia: a shared Eyes-defy-anemia dataset and a novel non-invasive system with machine learning, under submission

- G. Dimauro, D. Caivano, P. Di Pilato, A. Dipalma, M.G. Camporeale, A Systematic Mapping Study on Research in Anemia Assessment with Non-Invasive Devices. Appl. Sci. 2020, 10, 4804. DOI 10.3390/app10144804

- G. Dimauro and L. Simone, Novel Biased Normalized Cuts Approach for the Automatic Segmentation of the Conjunctiva, Electronics 2020, 9, 997. DOI 10.3390/electronics9060997

The dataset Eyes-defy-anemia is divided in two folders, one for Italian patients and one for Indian patients.

The folder Italy contains 123 folders and the file Italy.xlsx. The name of the folder, from 1 to 123, refers to the number contained in the “Number” field of the file Italy.xlsx. Within each numbered folder there are 4 files:

- file name.jpg representing the original eye;

- forniceal\_palpebral.png: image depicting the two conjunctives, segmented by hand;

- forniceal.png: image depicting the forniceal conjunctiva, segmented by hand;

- palpebral.png: image depicting the palpebral conjunctiva, segmented by hand.

The exception are folders labelled as 1, 35, 54, 58, 75 and 109; in fact in these images of the eye the forniceal conjunctiva is not exposed, so in these folders only available 2 files:

- file name.jpg representing the original eye;

- palpebral.png: image depicting the palpebral conjunctiva, segmented by hand.

For patient 93 the level of Hgb has not been recorded, however we consider it useful to make available to the scientific community the 4 images as listed above.

All the images of the eye were acquired in Bari (Italy).

The India folder contains 95 folders and the India.xls file.

The folder name, from 1 to 95, includes the number that corresponds to the “Number” field of the file India.xlsx.

Within each numbered folder there are 4 files:

- file name.jpg representing the original eye;

- forniceal\_palpebral.png: image depicting the two conjunctivas, segmented by hand;

- forniceal.png: image depicting the forniceal conjunctiva, segmented by hand;

- palpebral.png: image depicting the palpebral conjunctiva, segmented by hand.

All images of the eye were acquired in Karapakkam, Chennai (India).

All the images obtained after segmentation are RGB images of equal size (1067 x 800 pixels).