Analytics for Hemoglobin count based on image data

Sam Mathew
Senior Process Engineer
Gyan Data Private Limited
Chennai - 600 113 India
sam@gyandata.com

About the Client

En'Urga Inc. is an Indiana (US) based company leading the spray patternation technology and other optical diagnostic equipments for single and two-phase reacting flows. Located at the Purdue Research Park, En'Urga Inc. specializes in R&D and proof-of-concept (PoC) applications as well as customization and calibration of optical diagnostic equipment. The application areas include measurement of temperatures, gas concentrations, emissivity and particulate characteristics.

Motivation

Data analysis has been practiced in medicine ever since its inception; for diagnostics as well as cure through clinical trials. Computer vision based study though has made its way into medicine only recently, especially for diagnostics. This is evident from the last decade accounting for 87.5% of the publications in computer vision-based analytical chemistry (CVAC) procedures [1]. Hemoglobin count is essential for detecting anemia in time but is limited by phlebotomy and lab testing, both being great challenges in times of epidemics that beg safe and efficient infrastructure. CVAC procedures are now being explored for Hemoglobin prediction from images of eye conjunctiva.

Problem

CVAC procedures are still in their nascent stage and reasonable conclusion can only be derived through controlled experiments. The project envisioned building a statistical model for Hemoglobin prediction from images of eye conjunctiva taken on cattle in the field and in the sheep-pens.

The initial experiments were carried out by the research team at Purdue. Over 100 images for sheep's palpebral conjunctiva were analyzed and the RGB (red-green-blue) channels from the region of interest (ROI; in this case the eye conjunctiva) was extracted. **Gyan Data** was tasked with building a predictive model for Hemoglobin, based on the RGB data and a few derived quantities without any knowledge of the actual application. During the course of the interaction, the real problem definition surfaced once the client shared more details. Our research revealed the necessity for controlled environment during image acquisition, currently missed out in the project.

Solution

The initial scope for Gyan Data's engagement was to apply different predictive models on just over 100 data points, with 10 variables in the input space and one predictor variable (Hemoglobin level). Initial attempts focused on applying a host of model building techniques as simple as ordinary least-square (OLS) to more advanced kernel-Principal Component Regression (kPCR) with leave-one-out-cross-validation (LOOCV) for the small data-set. These efforts were preceded by data visualization and exploration that together confirmed the inadequacy of the naive regression methods.

Based on the variable names and investigating through the limited client's inputs, we could explore published research done for similar applications. This revealed that during image acquisition, uneven lighting severely impacts the R, G, B channels of the image of palpebral conjunctiva. Consequently we suggested controlled lighting and a photographic standard for future experiments.

Based on the approach followed in the published literature [2], an OLS model was built with derived features that are based on RGB channels in the conjunctiva image scaled by a photographic standard, that was available in all images. In order to have uniform lighting though, this approach could only be applied on an even fewer data points. Recommendations were subsequently provided for future controlled experiments.

References

[1] Luis Fermn Capitn-Vallvey, Nuria Lpez-Ruiz, Antonio Martnez-Olmos, Miguel M. Erenas, and Alberto J. Palma. Recent developments in computer vision-based analytical chemistry: A tutorial review. *Analytica Chimica Acta*, 899(Supplement C):23 – 56, 2015.

[2] Selim Suner, Gregory Crawford, John McMurdy, and Gregory Jay. Non-invasive determination of hemoglobin by digital photography of palpebral conjunctiva. *The Journal of emergency medicine*, 33(2):105–111, 2007.