Effects of Common Image Manipulations on Diagnostic Performance in Digital Pathology – Human Study

Platisa Ljiljana[1] MSc, Van Brantegem Leen[2] PhD, Kumcu Asli[1] MSc, Marchessoux Cédric[3] PhD, Vansteenkiste Ewout[1] PhD, and Philips Wilfried[1] PhD

[1]*TELIN-IPI-iMinds, Ghent University, Ghent, Belgium (ljiljana.platisa@telin.ugent.be)*
[2]*Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium*
[3]*Healthcare Division, Barco N.V., Kortrijk, Belgium*

A very recent work of Ref.[1] studied the effects of image manipulation and image degradation on the perceived attributes of image quality (IQ) of digital pathology slides. However, before any conclusions and recommendations can be formulated regarding specific image manipulations (and IQ attributes), it is necessary to investigate their effects on the diagnostic performance of clinicians when interpreting these images.

In this study, 6 expert pathologists interpreted digital images of H&E stained animal pathology samples in a free-response (FROC) experiment. Participants marked locations suspicious for viral inclusions (inclusion bodies) and rated them using a continuous scale from 0 (low confidence) to 100% (high confidence). The images were the same as in Ref.[1]: crops of digital pathology slides of 3 different animal tissue samples, all 1200×750 pixels in size. Each participant viewed a total of 72 images: 12 non-manipulated (reference) images (4 of each tissue type), and 60 manipulated images (5 for each reference image). The extent of artificial manipulations was adjusted relative to the reference images using the HDR-VDP metric [2] in the luminance domain: added Gaussian blur ($\sigma_b=3$), decreased gamma (-5%), added white Gaussian noise ($\sigma_n=10$), decreased color saturation (-5%), and JPG compression (libjpeg 50). The images were displayed on a 3MP medical color LCD in a controlled viewing environment.

Preliminary analysis assessing the change in the number of positive markings in the reference and manipulated images indicates that blurring and changes in gamma, followed by changes in color saturation, could have an effect on diagnostic performance. This largely coincides with the findings from Ref.[1], where IQ ratings appeared to be most affected by changes in color and gamma parameters. Importantly, diagnostic performance appears to be content dependent; it is different across tissue types. Further data analysis (including JAFROC) is ongoing and shall be reported in the conference talk.

References

[1] Platisa L., Van Brantegem L., Vander Haeghen Y., Marchessoux C., Vansteenkiste E., and Philips W., "Psycho-visual evaluation of image quality attributes in digital pathology slides viewed on a medical color LCD display," SPIE MI 2013 (in press).

[2] Mantiuk R., Daly S., Myszkowski K., and Seidel H. P., "Predicting visible differences in high dynamic range images - model and its calibration," in Human Vision and Electronic Imaging X, IS&T/SPIE's 17th Annual Symposium on Electronic Imaging (2005), Rogowitz B. E., Pappas T. N., and Daly S. J., eds., 5666, pp. 204–214, 2005.