Journal of Computational Intelligence in Bioinformatics
ISSN 0973-385X Volume 5, Number 3 (2012) pp. 169-174

http://www.ripublication.com/jcib.htm

Automated background correction for batch quantification of images using a new ImageJ macro

Prabha B. Mishra, Kalyani G. Bharadwaj*, Debarshi Chakrabarti, Pranoy Menon, Aurelio S. Lobo, Muralidhara Padigaru, Shashank Rohatagi

Biomarker Discovery, Dept. of Pharmacology, Piramal Healthcare Ltd., 1 Nirlon Complex, Off Western Express Highway, Goregaon (E), Mumbai, 400 063, India.
Email: prabha.mishra@piramal.com
**Corresponding Author Email: kalyani.bharadwaj@piramal.com*

Abstract

ImageJ is a Java-based open source code created by NIH to obtain clinically meaningful information from biological images through elaborate quantification. However, for our purposes of image analysis involving a large volume of images (28,000) with varied background and signal strength, use of ImageJ required laborious manual background correction resulting in significant data loss. Therefore, we designed a new macro to the basic ImageJ such that we achieved significant automated background correction that was individualized for each image with fast analysis time. After analyzing several images including some from an open source database, we conclude that our current ImageJ extension gives us reliable and reproducible results for complex analysis of fluorescent and non-fluorescent images.

Keywords: ImageJ; automation; background correction; batch analysis

Introduction

Scientists studying biological mechanisms are constantly using novel imaging tools to generate visually appealing and versatile images with multiplexed information thereby retrieving a whole array of complex clinically useful information (1). Most studies generate a large number of images to ensure statistical validity making image analysis laborious and time-dependent. Attempts for prompt and valid image analysis resulted in the development of several programs including ImageJ (2). ImageJ is a Java based open source code created by Wayne Rasband at NIH for biological image processing and subsequent analysis (3). ImageJ can display, permit extensive editing,

processing, and analysis of images with elaborate geometric and arithmetic operations for image quantification (4). ImageJ users can incorporate customized macros and/or plug-ins to suit their individual needs (5, 6).

Image analysis in biological settings, especially in the industrial setting, is often limited by the large volume of images captured at different times, resulting in variability, low throughput, and high noise. We came across several images in our dataset either with an overwhelming background or a normal background but weak signal. We used the "Batch Measure" macro of ImageJ (http://rsb.info.nih.gov/ij/macros/BatchMeasure.txt) for batch type analysis and then incorporated the auto threshold functionality for analysis. However, the issues of background and relative signal strength were unresolved thereby requiring manual intervention for each image. This resulted in individual bias, significant data loss, and throughput reduction. To circumvent these issues, we designed our own macro to the "Batch Measure" ImageJ for automated background correction.

Materials and Methods

The background correction in our new macro consists of 3 steps: (1) Calculating mean stained area and the standard deviation for each image (2) Taking the sum (**A**) of this mean area and 3 times the standard deviation (3) Setting "**A**" as the threshold, such that values above threshold were quantified for signal intensity. For florescent images, we used the "Batch Measure" ImageJ, incorporated the auto threshold functionality and the newly designed macro to achieve automated background correction for each image. For non-fluorescent images, we employed the "color deconvolution" plugin for ImageJ, which can analyze only one image at a time (7) along with the "Batch Measure" ImageJ macro and modified it to incorporate our own macro for automated background correction.

We generated fluorescent immunohistochemistry images of pancreas from an *in vivo* mouse model. The resulting enormous dataset of 28,000 images was then subject to analysis using our new macro for color fluorescent images (submitted as "Batch_Measure_color_mean_std.txt" file). The macro for B&W fluorescent images is submitted as "Batch_Measure_bw_mean_std.txt". For simplicity, we show data from 443 pancreas images stained for insulin (β cell area) and DAPI (total pancreas area). All images were analyzed by 4 methods including manual outlining, manual threshold, auto threshold, and our new macro automated background correction technique.

Results and Discussion

Figure 1 gives representative images of mouse pancreas stained for insulin to determine the β-cell area. In figure 1A (left panel), the background signal was too strong to detect signal from the area of interest, which upon application of the auto threshold setting highlighted the entire pancreas instead of β cells alone (middle panel). Figure 1A right panel shows the same image but with the background correction using the new macro threshold setting such that only area above the

computed threshold positive for the stain is highlighted. In another example we observed weak signal strength against a normal background (figure 1B left and middle panels). However, with our current threshold correction technique applied to the same image, we rectified the issue and obtained excellent signal strength (figure 1B, right panel) enabling quantification.

Data obtained with the new macro was compared with other methods of quantification. Here we report pancreatic β-cell mass (pancreas weight (mg) **x** ratio of insulin stained area to total pancreas area) for all 4 methods (Figure 2). When compared with β-cell mass reported in literature (~2 mg) (8; 9), the percent error for the manual outlining, manual threshold, auto threshold, and the new macro methods was 9.9 ± 47.2, 146 ± 29, 32805 ± 5698, and 221.4 ± 30.5 (all values are mean ± SD) respectively. Despite a high percent error with the new macro method, the variation was very low compared to the manual outline technique, indicating inconsistency and operator bias in the latter method. The high percent error in the new macro method can be attributed to a higher estimation of the pancreas weight during tissue harvest resulting in deviation from normal. The auto threshold method however showed a significantly higher β cell mass compared to the other 3 methods (figure 2, $p<0.05$) with a markedly high percent error and variance suggesting the unreliability of the former. The manual threshold method is closest to the new macro method in computing the β-cell mass (figure 2) and also demonstrated a similar percent error.

To further confirm the validity of our macro; we obtained endometrial cancer images stained for DAB-labeled estrogen receptor α (ERα) from the Human Protein Atlas database (www.proteinatlas.org) (10, 11). Per the database, these images were categorized into 6 groups based on stain intensity (representative images depicted in Figure 3). We downloaded these images, obtained RED channel images for the DAB stain, and processed them with our new macro for non-fluorescent images (submitted as "Batch_Measure_color_deconvolution plugin_mean_std.txt" file) to estimate stained area, which was plotted against the manual call-outs from the database (Figure 4). Spearman's rank correlation coefficient for this dataset was found to be 0.86 ($p<0.001$), thereby suggesting a good association between our quantification technique and the manual call-outs from the database.

Thus, with this simple modification to the existing ImageJ, we were able to achieve several goals: (1) Automated background correction for each image without manual intervention (2) Quick processing times (1 hour for 28,000 images as opposed to 1900 hours by the manual outline method) (3) Conversion of images to grayscale before processing thereby maintaining uniformity and enabling analysis of RGB and B&W images (4) Background correction for each RGB channel possible; useful for multiple protein staining. In conclusion, our newly designed macro allows for rapid and automated analysis of large image datasets with better precision that is independent of operator bias.

Acknowledgments

The authors wish to thank Dr. Jyothi Subramanian for her critical input on the statistical analyses of the data.

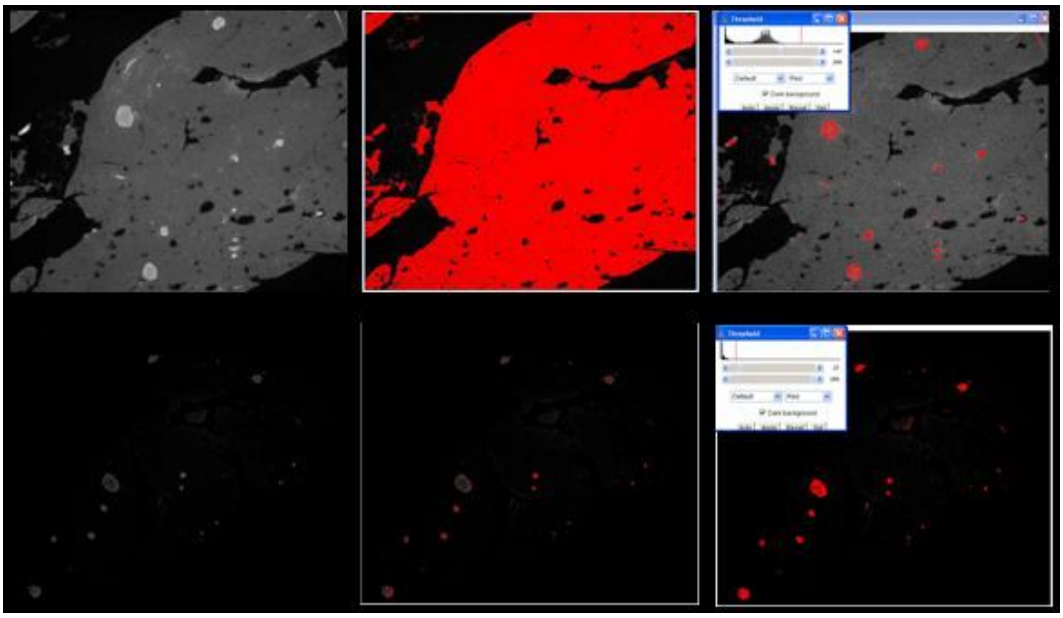

Figure 1: Representative images of pancreas sections stained for the islets with insulin. The top panel shows a representative image of pancreas with normal signal intensity but against a high background on the left panel; middle panel is highlighted for quantifiable area based on auto threshold, which as seen considers the entire pancreas area due to the background; right panel is after adjusting threshold as mean + (3xSD), thereby highlighting only the insulin stained area. The bottom panel depicts a representative pancreas image showing no background interference but with low signal intensity on the left panel; upon auto threshold application, the islets are highlighted insufficiently thereby resulting in data loss in the middle panel; right panel is after the application of the new threshold setting adjustment.

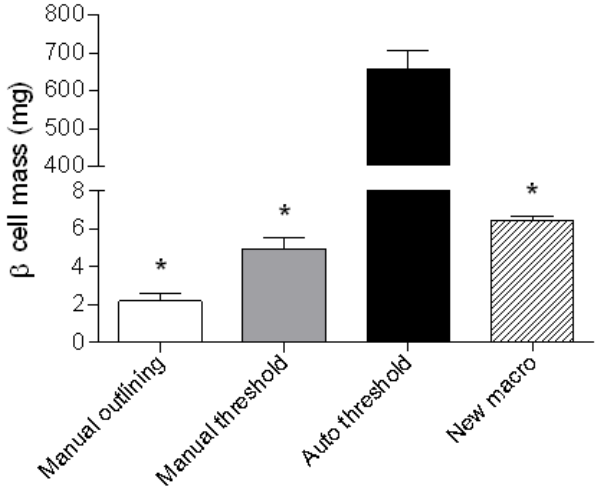

Figure 2: β cell mass calculated for pancreas sections by 4 different methods. Data are mean ± SD, n=6, * $p<0.05$ compared to the auto threshold method.

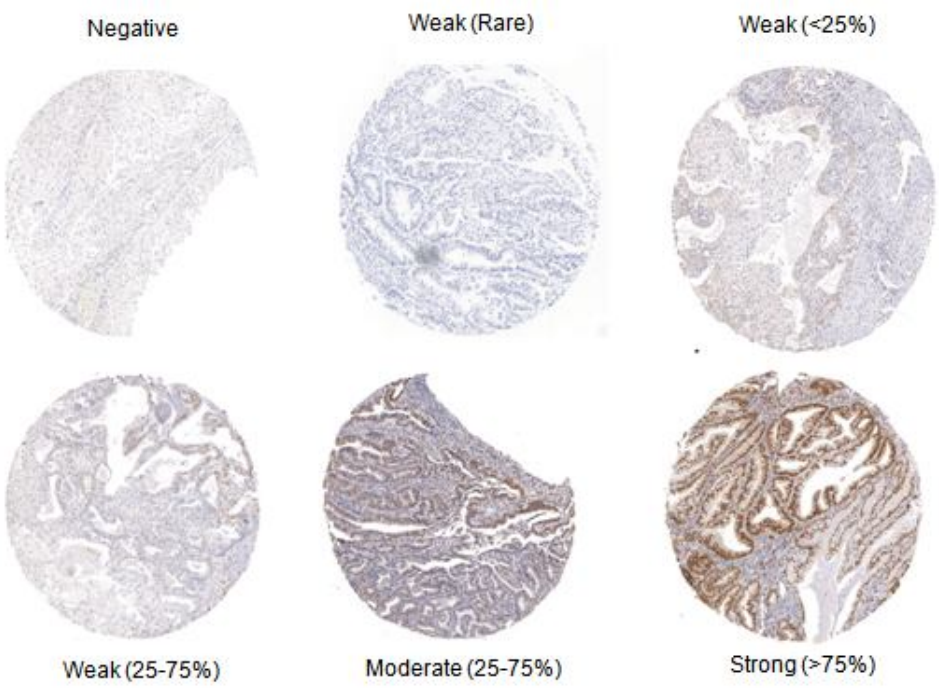

Figure 3: Representative images of human endometrial cancer stained for DAB-labeled estrogen receptor α (ERα) categorized into 6 groups based on stain intensity obtained from www.proteinatlas.org.

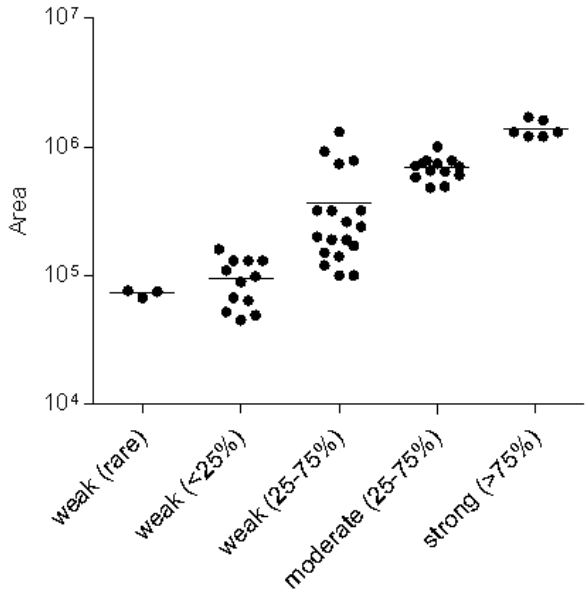

Figure 4: Quantification of intensity density using our macro method for the images obtained from www.proteinatlas.org. Data are presented as a scatter plot with their mean indicated as the horizontal line. Spearman's correlation coefficient r=0.87 ($p<0.001$).

References

[1] Laurinavicius, A., Laurinaviciene, A., Dasevicius, D., Elie, N., Plancoulaine, B., Bor, C., and Herlin, P., 2012, "Digital image analysis in pathology: Benefits and obligation," Anal. Cell. Pathol. (Amst), 35(2), pp. 75-78.

[2] Prasad, K., and Prabhu, G. K., 2011, "Image analysis tools for evaluation of microscopic views of immunohistochemically stained specimen in medical research-a review," J. of Med. Syst., Epub ahead of print.

[3] Rasband, W. S., 1997-2011, "ImageJ," U.S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/

[4] Abramoff, M. D., Magalhaes, P. J., and Ram, S. J., 2004, "Image processing with ImageJ," Biophot. Int., 11(7), pp. 36-42.

[5] Abramoff, M. D., Kalmann, R., de Graaf, M. E., Stilma, J. S., and Mourits, M. P., 2002, "Rectus extraocular muscle paths and decompression surgery for Graves orbitopathy: mechanism of motility disturbances," Invest. Ophthalmol. Vis. Sci., 43(2), pp. 300-307.

[6] Meijering, E., Jacob, M., Sarria, J. C., Steiner, P., Hirling, H., and Unser, M., 2004, "Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images", Cytometry. Part A, 58(2), pp. 167-176.

[7] Ruifrok, A. C., and Johnston, D. A., 2001, "Quantification of histochemical staining by color deconvolution," Anal. Quant. Cytol. Histol., 23(4), pp. 291-299.

[8] Kulkarni, R. N., Almind, K., Goren, H. J., Winnay, J. N., Ueki, K., Okada, T., and Kahn, C. R., 2003, "Impact of genetic background on development of hyperinsulinemia and diabetes in insulin receptor/insulin receptor substrate-1 double heterozygous mice," Diabetes, 52(6), pp. 1528-1534.

[9] Wong, V. S., Oh, A. H., Chassot, A. A., Chaboissier, M. C., and Brubaker, P. L., 2011, "R-spondin1 deficiency in mice improves glycemic control in association with increased beta cell mass," Diabetologia, 54(7), pp. 1726-1734.

[10] Uhlen, M., Bjorling, E., Agaton, C., Szigyarto, C. A., Amini, B., Andersen, E., Andersson, A. C., Angelidou, P., et al., 2005, "A human protein atlas for normal and cancer tissues based on antibody proteomics," Mol. Cell. Proteomics, 4(12), pp. 1920-1932.

[11] Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., et. al., 2010, "Towards a knowledge-based Human Protein Atlas," Nat Biotechnol, 28(12), pp. 1248-1250.