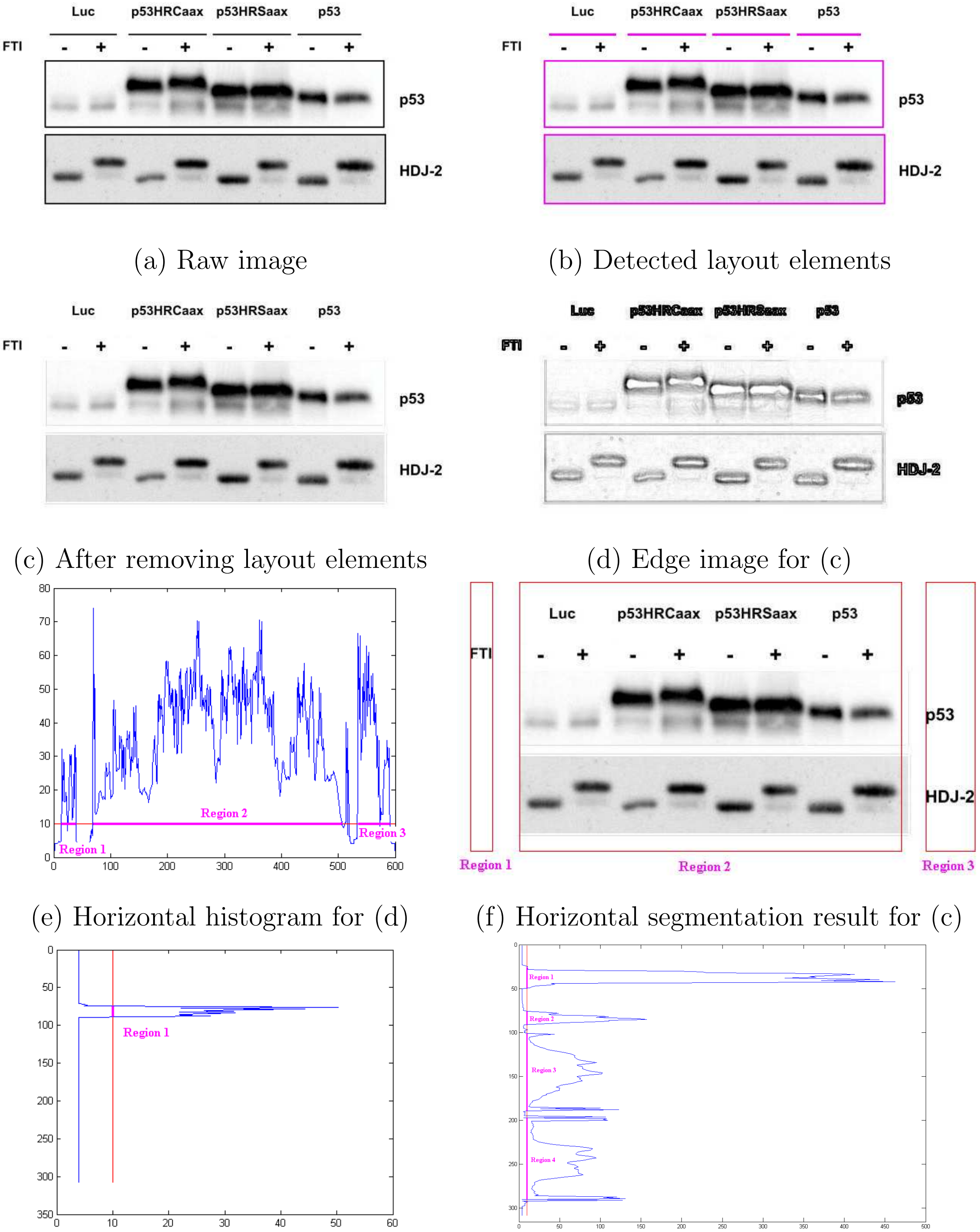
A New Pivoting and Iterative Text Detection Algorithm for Biomedical Images: Appendix B

A Text Region Detection Example using our Algorithm

In Figure 1, we show the working of our algorithm on a sample image. Given a raw input image (a), our algorithm first detect the layout elements such as lines and panel boundaries for layout purpose (b). After removing these layout elements from the raw image (c), we apply a Sobel edge detector to derive the edge image (d). We then construct the horizontal histogram for the edge image (e). Through segmenting the horizontal histogram, we detect three horizontal text regions in the image, which are marked Region 1 to 3 (e). According to these horizontal ranges, we can correspondingly detect three text regions in the image (f). Starting with Region 1 in (f), we construct a vertical histogram, shown in (g-1). We proceed identically for Regions 2 and 3 in (f), obtaining vertical histograms as shown in (g-2) and (g-3), respectively. Segmenting the vertical histogram in (g-1), we obtain a vertical text region range. Then, the horizontal region range marked "Region 1" in (e) and the vertical region range marked "Region 1" in (g-1), are combined to construct a refined text region for Region 1 in (f), shown in (h-1). We proceed similarly for Regions 2 and 3 to receive refined image text regions (h-2 and h-3). Text regions shown in (h-1) and (h-3) can not be further refined and are flagged by the algorithm.

The subregion shown in (h-2) can be further refined through an additional round of our PTD algorithm (i-1). A horizontal histogram for Region 1 in (h-2) detects further image text segments (j-1). A similar result is obtained for region 2 in (h-2), see (i-2) and (j-2). Finally, we implemented a naive intensity based image block detection procedure, which marks non-text region at the end of each iteration. Running this procedure helps us remove the large none-areas inside Region 3 and Region 4 in (h-2). These areas are marked in pink in (k), and are subtracted in the final output image (l).



(a) Raw image

(b) Detected layout elements

(c) After removing layout elements

(d) Edge image for (c)

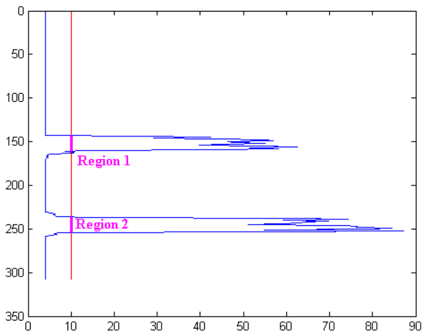
(e) Horizontal histogram for (d)

(f) Horizontal segmentation result for (c)

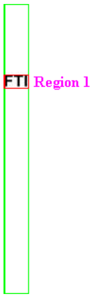
(g-1) Vertical histogram for Region 1 in (f)

(g-2) Vertical histogram for Region 2 in (f)

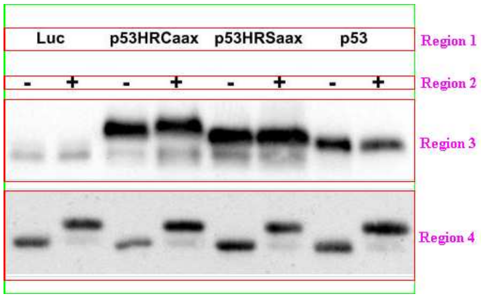
(to be continued on the next page)



(g-3) Vertical histogram for Region 3 in (f)



(h-1) Segmentation result for Region 1 in (f)

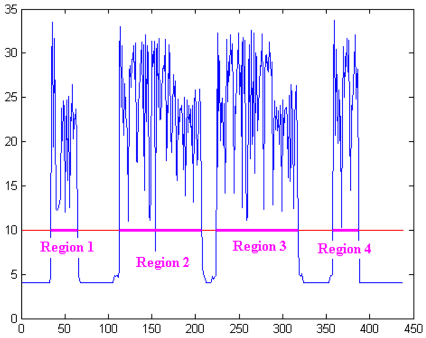


(h-2) Segmentation result for Region 2 in (f)

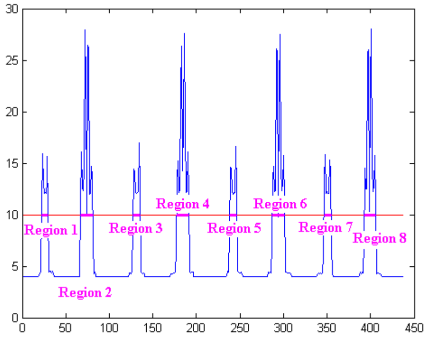
p53
Region 1
HDJ-2
Region 2

(h-3) Segmentation result for Region 3 in (f)

(Image boundaries are indicated in green.)

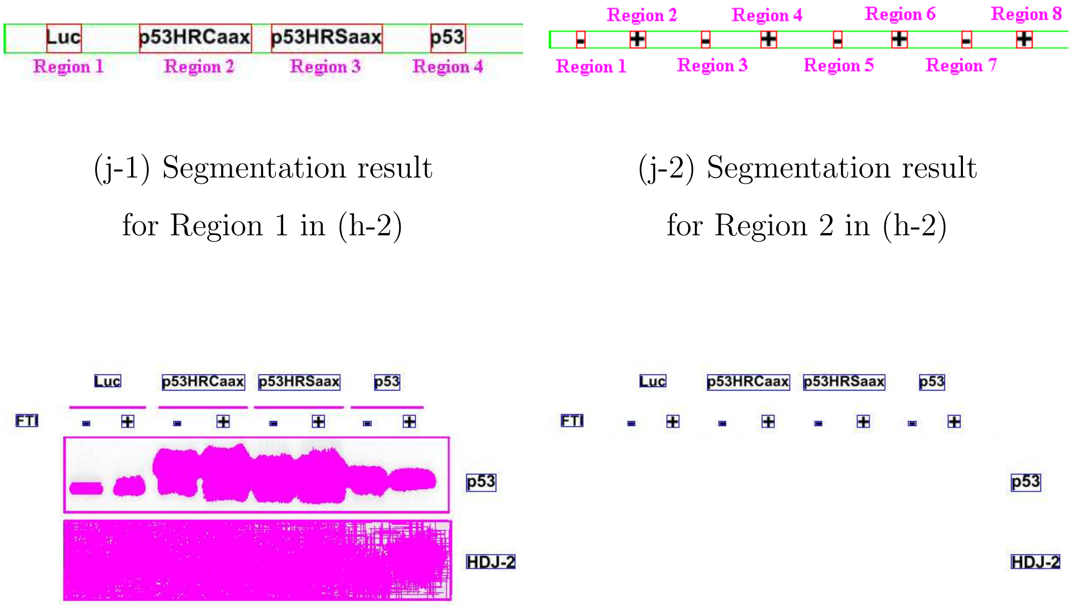


(i-1) Horizontal histogram for Region 1 in (h-2)



(i-2) Horizontal histogram for Region 2 in (h-2)

(to be continued on the next page)



(j-1) Segmentation result for Region 1 in (h-2)

(j-2) Segmentation result for Region 2 in (h-2)

(k) The text detection result after all iterations

(l) The final text detection result

Figure 1: A step-by-step text detection example.