(a) Rs=0.995, Ps=0.998, Fs=0.997

(b) Rs=0.978, Ps=0.985, Fs=0.981

(c) Rs=0.554, Ps=0.957, Fs=0.702

(d) Rs=0.964, Ps=0.912, Fs=0.937

(e) Rs=0.995, Ps=0.996, Fs=0.997

(f) Rs=1.000, Ps=0.971, Fs=0.985

FIGURE 8. Examples of mathematical symbol and expression detection results for the proposed method. (a)–(f) show cropped regions from page images in different articles. The components in cyan, magenta and yellow denote the TP, FP and FN components, respectively. The sub-captions show the recall, precision and F-measure values for each page.

We created two types of subsets of the training dataset that consisted of a fixed number of document page images. The document page images in the first subset were collected in an article-wise manner from the full dataset. This means that all page images in the determined articles were included in the subset. The second subset was a collection of the same number of page images that were randomly selected from the full dataset. We expect that a diversity of font faces and mathematical notation styles is kept by the random selection. We adjusted the number of images in the subsets and repeated the experiment for U-Net training and evaluation.

Figure 10 shows the results of the experiments. The vertical and horizontal axes denote the symbol-based F-measure value F_s and number of pages in the subset, respectively. Whereas the F-measure value rapidly decreased when the number of pages in the article-based subset