Manga Frame Extraction and its Application to Detecting Inappropriate Images in Manga Books

Satoshi Murakami, Tomoharu Nagao,

***Abstract*— Manga depicts various scenes in multiple areas on a page called "komas", and a story is represented by a sequence of komas. In this paper, we describe a koma extractor using CNN to extract komas as rectangles or polygons from a page image of a manga book. As an application of koma extractor, we also describe an inappropriate image detector that detects whether each koma contains an inappropriate object or not.**

I. IINTRODUCTION

Manga, unlike movies and TV programs that express stories within a fixed rectangular area, depicts different scenes in a series of small areas called " komas", which are located on the page, and expresses the sequence of time flow and the story. There are no rules for the shape, size, and layout of koma, and koma boundaries are sometimes blocked by speech balloons or drawn objects, making it difficult to extract accurate koma shapes using conventional methods. The koma extractor described in this paper uses CNN to estimate arbitrarily placed small regions in a manga page image, and extracts koma from the estimated regions.

II. PROPOSED SYSTEM

In most cases, objects are drawn within the koma without extending out of the koma area. However, when image objects or speech balloons are drawn beyond the koma area, a part of the koma boundary line may be lost. In some cases, large image objects are drawn across multiple komas to make a strong impression, and it is necessary to correctly estimate the koma boundaries when estimating komas with multiple missing parts. As a prior example using deep learning, Ogawa et al. [1] extended the object detection model SSD300 and proposed a method with high robustness and good frame detection accuracy. The koma shape is always approximated as a rectangle because the method selects the one region that has the largest IOU (Intersection Over Union) with the correct rectangular shape region from multiple rectangular proposed regions. As shown in Figure 1, in the case of komas colored blue and orange, a human-shaped object in the orange region is misidentified as being in the next koma, which is surrounded by blue dotted lines and approximated as a rectangle. The koma estimator in this paper correctly estimates the koma shape by estimating the inner and outer regions of the koma pixel by pixel using CNN for semantic segmentation, and then approximating these multiple estimated regions to polygonal polygons in turn using rule-based image processing. The input to the CNN is a pair of images, a manga image and a target mask image generated from the annotation data corresponding to the page (Figure 2), and the output is trained to be the target mask image. The dataset consists of about 4,000 pages images randomly collected from manga, self-made annotation data for each of these images, and program-generated extended images. The average panel recognition accuracy of the ten books in the Manga109 dataset was 96.6% for our method, compared to 96.8% for Ogawa et al. The accuracy is almost the same. Figure 3 shows an example of detecting komas containing inappropriate objects using CNN that recognizes the presence or absence of inappropriate objects in each extracted koma.

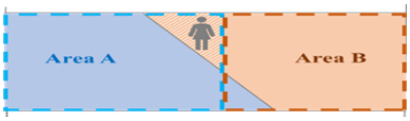

Figure 1. An example where the two komas colored blue and orange are diagonally tangent to each other.

Figure 2. Resized input image (Left), Corresponding target mask image generated from annotation data (Right), "Kitsune no Oyome-chan" Copyright 2019 Batta.

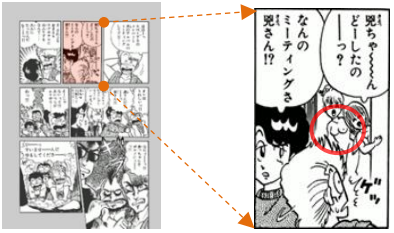

Figure 3. Example of detecting a koma containing an inappropriate bare-breasted object (colored koma), Page image (Left), enlarged koma image (Right), "Koi wa Gu-Cyoki-Pa" Copyright 2014 Kazuhiko Mido.

ACKNOWLEDGMENT

Mr. Batta, the author of "Kitsune no Oyome-Chan," and Mr. Kazuhiko Mido, the author of "Koi wa Gu-Choki-Pa," have kindly given permission to use their images in this paper, and we would like to take this opportunity to express our deepest gratitude to them.

REFERENCES

[1] Toru Ogawa, A. Otsubo, Rei Narita, Yusuke Matsui, Toshihiko Yamasaki, Kiyoharu Aizawa, Object Detection for Comics using Manga109 Annotations, arXiv, 2018.

Satoshi Murakami is a working student of Postdoctoral Program at the Graduate School of Environment and Information Studies, Yokohama National University, Yokohama, Japan. (murakami-satoshi-sv@ynu.jp, sat.murakami@gmail.com).

Tomoharu Nagao is a professor at the Graduate School of Environment and Information Studies, Yokohama National University, Yokohama, Japan. (nagao@ynu.ac.jp).