Comic Book Interpretation based on Deep Neural Networks

Miki Ueno
Information and Media Center, Toyohashi University of Technology

Abstract

The computational interpretation of comics is one of the important topics being studied in the field of artificial intelligence and image recognition. This paper summarizes two types of classification in order to analyze story patterns of four scene comics based on convolutional neural networks(CNNs).

Introduction

The computational interpretation of comics has seen rapid advance in recent years by developing of Convolutional Neural Network(CNN). There are numerous challenging tasks and unknown definitions for the computational interpretation of comics; for instance, recognize drawing by a gray-scaled image, modeling the structure of a story transition. The author focused on four scene comics, which have clear and simple structure, from both aspects of the generation(Ueno, Mori, and Matsumoto 2014) and interpretation(Ueno et al. 2016) of comics by a computational method.

From my previous studies, I found that scene order is important information for interpreting the semantics of scenes and the whole stories in four-scene comics, compared to other types of comics. Therefore, in this study, I constructed CNNs to estimate the orders of the comics and discuss the story structure of four scene comics.

Basic Concept

Comics are popular especially in Japan and in a number of some European countries. Actually, numerous genres and structures of comics exist across the globe. They are regarded as entertainment, expression, and artistic works. Comic stories describe creative fiction, but they can also describe current events clearly so comics are often used in newspapers. In this study, four-scene comics, which are structured with four continuous scenes, are considered. The length of four-scene comics is limited so as to ensure a clear interpretation of the contents. I suggested that, if a story transition pattern can be estimated, it would be helpful for interpreting the semantics of comics by a computational method and useful for developing an interactive application for creators. According to following two types of classification, I suggest that deep neural method can interpret story patterns of four-scene comics.

Two types of Classification

Scene Order

To confirm whether the CNN can recognize the order of transition in comics, ${}_4\mathrm{C}_2$ kinds of two class classification; e.g. the first scene or the second scene, the first scene or the third scene; are carried out. The results suggested that CNNs based on AlexNet(Krizhevsky, Sutskever, and Hinton 2012) can estimate the order of four scene comics. In other words, CNNs obtained specific image features of scene order in four scene comics.

Scene Emotion

To confirm whether the CNN can recognize emotion of comic scene, emotion classification is carried out for 188 fourth-scene images extracted from four-scene comic stories. Human annotator assigned emotional tags to each scene. The results suggested that CNNs can estimate the emotion of four scene comics.

Conclusion

This paper summarize two types classification for analyzing story patterns of four-scene comics utilizing convolutional neural networks. This research is partially supported by DAIKO FOUNDATION.

References

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105.

Ueno, M.; Mori, N.; Suenaga, T.; and Isahara, H. 2016. Estimation of structure of four-scene comics by convolutional neural networks. In *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding, MANPU@ICPR 2016*, 9:1–9:6.

Ueno, M.; Mori, N.; and Matsumoto, K. 2014. 2-scene comic creating system based on the distribution of picture state transition. In *Distributed Computing and Artificial Intelligence, 11th International Conference*, volume 290 of *Advances in Intelligent Systems and Computing*. Springer International Publishing. 459–467.