Front Cover Image Database of Japanese Manga and Typeface Estimation of their Title

Shota Ishiyama[1], Kosuke Sakai[1], and Minoru Mori[1]

Kanagawa Institute of Technology, Atsugi-shi, Kanagawa 243-0292, JAPAN
mmori@ic.kanagawa-it.ac.jp

Abstract. Front cover design of books like Manga is one of the most important factors for appealing contents to users. Fonts used for the title in the front cover are carefully selected among a lot of ones for fitting selected fonts to the content and the design. However, this task to select fonts, that increases attractions of the front cover and are appropriate for human characters pictured in the front cover, is not easy. Few experienced designers or editors can do well. In this paper we try to estimate and recommend appropriate typefaces of fonts that seem to be appropriate for title fonts from an image of front cover of Manga and Light novels. To evaluate our framework, we gathered front cover images of Manga and Light novels, and created database that containing five kinds of images; front cover images with/without title fonts, whole-body images with/without title fonts, and face images. Each image has two types of label encoded from the count number of 5 typefaces used for title fonts. Experimental results using our database show that about 70% of typefaces are correctly estimated and suggest a strong relationship between fonts used for title and front covers.

Keywords: Front cover page · Manga · Light novels · Typeface · DNN.

1 Introduction

Lots of new contents and books of "Manga", comics and graphic novels, and "Light novel", Japanese young adult novels, have been continuously produced and published not even in Japan but in the world. When buying a book of Manga or Light novels, we usually select one on the base of its story, reviews, prices, character designs, and others. At least we all see a front cover page of each book. If a book of Manga or Light novels to buy is decided before, its front cover has nothing to do with sales or selections. But, if not decided, the impression and design of a front cover page seem to be a very important factor for sales on not only Manga books but other many books. Especially many Manga contents feature some human characters and the design of their character influence the popularity. Therefore, the front cover page needs to be designed for drawing attentions of users. One of important elements of the front cover design is the title font. Editors and designers carefully select kinds of fonts to bring out and appeal the charm of the Manga content. Fonts are not only medium of language

but a piece of art and impressions. From a lot of fonts, designers select fonts to fit the targeted Manga content and deliver its attraction to readers. Recently, though the number of amateur Manga and Light novels writers is increasing, they without knowledges about fonts cannot select appropriate fonts to fit their contents and design. If more appropriate fonts for the targeted Manga content or design can be easily selected, more attractive front covers with suitable fonts as its title can be produced and published.

In this paper we propose a framework that estimates suitable typefaces of fonts that fits the design of a front cover of Manga and Light novels and a database that consists of several types of images for this task. In detail, a lot of sets of the front cover image and its font information used as the title are created as the database and we evaluate our framework that estimates typefaces of fonts that fit front cover image or human characters contained in the front cover using our database. For validating relationship between fonts used as the title and human characters, we create whole-body images extracted from the front cover images and face images obtained from the whole-body image. Fig. 1 shows examples of front cover images of Manga and Light novels and fonts used as the titles in these images that we handle in this paper. Fig. 2 shows the overview of our framework that estimates typefaces included in an image using DNN. The paper is organized as follows: Section 2 provides related works about relationships between fonts and design of books or signboards. Our database and proposed method are described in Section 3. Section 4 reports evaluation experiments and discusses experimental results. Section 5 summaries this paper and lists future works.

(a) (b)

Fig. 1: Examples of front cover images of Manga and Light novels. (a) Typeface: Designed, Font name: Hasetoppo [for 17 Hiragana characters], Typeface: Designed, Font name: Takahand [4 Kanji characters] [1]. (b) Typeface: Mincho, Font name: A1 Mincho [for all characters] [2].

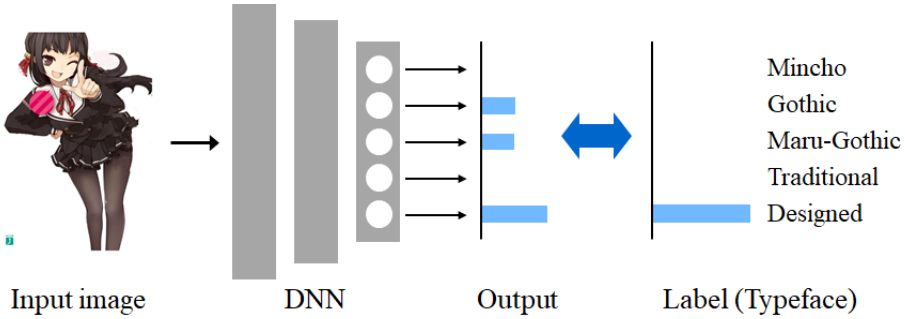

Fig. 2: Overview of our framework.

2 Related Works

Shinhara et al. analyzed that how genres of English books effect selections of colors of their front covers and fonts used as their book title [3]. They split the genres of English books into 32 kinds of groups, and font types were split into 6 typefaces. They validated the performance of genre estimation from font typefaces. Their experimental results showed that there was the relationship between the genre and font typeface. Also, books with similar colored front covers are in the same or similar genres and font type groups. In [5], they validated a relationship between fonts used in signboards and impressions of shops with these signboards. 7 signboards with each different type of fonts were prepared as test images. Users selected genres of eateries or restaurants from the impression based on appearances of each font. Experimental results showed that users often imagined different genres of eatery from different types of fonts. The tendency on the font selection on the signboard of real eateries was very similar to that obtained in their experiments. From these researches, we can say that impressions felled from designed linguistic substance like front covers of books and signboards has a strong relationship between fonts selected in them. On books of Manga and Light novels, the design of front covers seems to have some degree of relationship with fonts used in these front covers. Therefore, as one of design analyses, the estimation result of font from front cover pages enables us to easily select fonts that suit human character pictured in front cover pages for automatic or semi-automatic book design.

3 Proposed Method

This section describes our database that consists of front faces from Manga and Light novels, typefaces used in front faces, several kinds of label information for each data. And we explain a DNN model as our estimation framework used in our experiments.

3.1 Database of front cover images

For validating the typeface estimation on front cover images from Manga and Light novels, we gathered hundreds of front cover images from many books of them. Here, it's too difficult to obtain detailed information of font such as a concrete font name and a kind of typeface used in front covers only by observing font appearances. Therefore, we selected front cover images containing title fonts whose detailed information as their concrete font name and typeface information. Such books are introduced in several books that describe about designing and editing works of Manga and Light novels [6–11]. In this paper, to evaluate not only the relationship between fonts and the whole image of front cover but also that between fonts and whole-body image of human character contained in the front cover or a face of such human characters, we gathered only front cover images that involve human characters. The number of collected images is finally 581, and the number of unique titles is 227. As mentioned above, to investigate the relationship between fonts used in the front cover and whole bodies or faces too, we extracted a part of human characters and a part of faces from each front cover image. On front cover images and whole-body images from human characters, we erased font parts by using Adobe Photoshop. No Face images involve font images. Finally, we created five types of dataset that are front cover images with/without fonts, whole-body images with/without fonts, and face images. Fig. 3 shows examples of each type of images.

3.2 Typefaces of fonts used in title

This paper discusses only Japanese fonts as targets. Japanese fonts consist of more than 1,000 kinds. And designers sometimes originally adjust the shape of existing font design for suiting fonts to the design of front covers. Therefore, numbers of kinds of Japanese fonts exist. Moreover, some different kinds of fonts have very similar shape. From such reasons, estimating the concrete font name seems to be impossible. In this paper, we try to estimate not the detailed information such as concrete font name but the typeface name of each font as the general information. On the basis of Japanese font analysis in [12, 13], we classify each font into 5 typefaces; "Mincho", "Gothic", "Maru-Gothic", "Traditional", and "Designed". And we estimate a class of typefaces of fonts used from each image such as front cover, whole body, and face. Fig. 4 shows examples of 5 typefaces mentioned above.

3.3 Label

A label information of each image contains the name of typeface used as title fonts and their number of counts. Here, some titles have multiple fonts and typefaces as shown in Fig. 1. Therefore, the label of several image has counts on multi typefaces among five typefaces mentioned above. To clarify such conditions on title fonts used in front cover of books, we investigated 581 images we gathered. Table 1 shows the number and ratio of images including only one typeface and

(a) Front cover image with title fonts

(b) Front cover image with no title fonts

(c) Body image with title fonts

(d) Body image with no title fonts

(e) Face image (enlarged)

Fig. 3: Examples of front cover and several part images. JTC Janken font (designed typeface) is used [5].

(a) Mincho

(b) Gothic

(c) Maru-Gothic

(d) Traditional

(e) Designed

Fig. 4: Examples of each typeface. Each character expresses Japanese Hiragana.

that containing multiple typefaces for their labels. Table 2 provides the number and ratio of images on the main typeface of title fonts used in each image. Table 1 shows that there is a certain number of images with multiple typefaces and we cannot ignore such images. Table 2 gives that designed typefaces are often selected for the front cover of Manga and Light novels and traditional typefaces rarely are used. They seem to be common and acceptable because of characteristics of such genres. On the other hand, the count number of images containing the Mincho typeface as main title fonts is more than the sum of images including Gothic or Maru-Gothic. This is unexpected at least for us.

Table 1: Number of front cover images containing single or multiple fonts.

Num. of font typefaces	Num. of images	Ratio [%]
Single	505	86.9
Multiple	76	13.1

Table 2: Number of images for each typeface.

Typeface	Num. of images	Ratio [%]
Mincho	211	36.3
Gothic	91	15.7
Maru-Gothic	76	13.1
Traditional	20	3.4
Designed	183	31.5

On the basis of labels mentioned above, we encode each label information into two other types of label and adopt our evaluation method for each type of label. The first one is so-called one-hot encoding; a typeface with the most counts has one and the other typefaces have zero. This label has one-hot vector and this type of label are usually used for the multi-class classification task. In this paper we call this type of label "hard label". In the evaluation step, when a typeface with the highest probability in an output is same as that with one hot value, the estimation of typeface is regarded as correct. The other type of label consists of ratio values for each typeface; each ratio is computed by dividing count number of each typeface by the total count number among all the typefaces. So, if an original label has counts for multiple fonts, this kind of label has values on multiple positions. We call this type of label "soft label" in this paper. In the evaluation, we compare a typeface of the highest probability in an output with that with the highest label value. If it's same, the estimation is regarded as correct. Table 3 shows an original label information and 2 kinds of label that are encoded from the original one in our experiment.

Table 3: Examples of original label and 2 kinds of labels used in our data.

Typeface	Original	Hard	Soft
Mincho	0	0	0
Gothic	6	1	0.6
Maru-Gothic	0	0	0
Traditional	0	0	0
Designed	4	0	0.4

3.4 Estimation Model

As a framework for estimating typefaces used in each front cover image, we use Deep Neural Networks (DNN). In this paper, we exploit a pre-trained DNN model trained by using many data and fine-tune such a model because few training images tend to obtain a model that over-fits training images. As a pre-trained DNN model, we use the VGG16 model [14] that were trained by the ImageNet dataset that contains 14 million images of 1,000 classes. We exploit only convolution layers of this pre-trained DNN model as a feature extraction part and trained new dense layers as a classification part by the use of training data mentioned in 3.1 in the training process. Moreover, 3 convolution layers in VGG16 were re-trained in the fine-tuning process. In the training using data with hard or soft labels, dense layers with the Softmax function were used in the final output layer. Fig. 5 shows the structure of the DNN model used in our experiments.

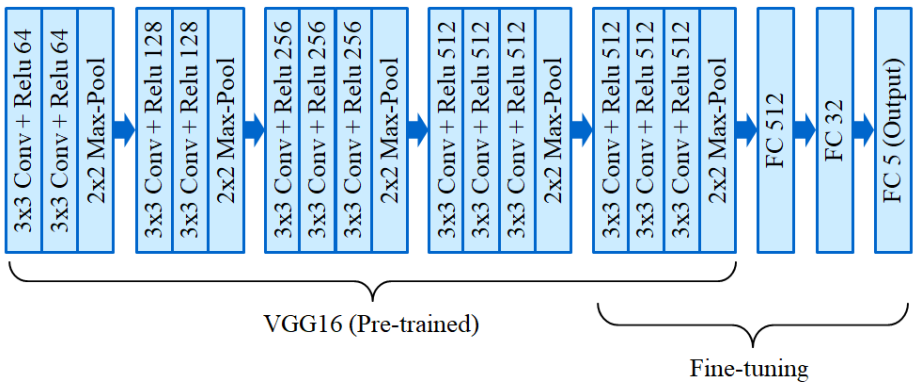

Fig. 5: The structure of our DNN model.

4 Experiments

In this section we describe evaluation experiments for validating and analyzing the performance of our approaches to estimate typefaces of title fonts in the front cover of Manga and Light novels.

4.1 Experimental Set-up

As experimental data, we split each data of front cover, whole body, and face 581 images into 4 sets. 436 images from 3 sets were used as training data and 145 images from another set were used as test one. We adopted cross-validation; therefore, each set were used as test data once in the rotation and 4 tests were totally carried out. Each result shown below was calculated as the average value among 4 tests.

Each data with hard or soft label were trained and tested as a multi-class classification problem. Thus, the cross entropy was used as their cost function for their training.

On the basis of preliminary experiment results, we set several experimental set-ups as follow. The number of epochs is 300. The batch size is 1 because of few training data. The initial learning late is 2.0e-5. Adam was used as the optimizer. The number of dense layers is 3, and the numbers of units in each dense layer are 512, 32, and 5, respectively. The cross-entropy was used as the loss function. As data augmentation, basic techniques as rotation, horizontal shift, vertical shift, and horizontal flip were adopted.

Each image is normalized into 256×256 pixels as a default size. Front cover images and body images are normalized into several sizes and details are described later.

4.2 Experimental Results

First of all, we describe experimental results using data with hard labels and default settings. Typeface classification rates within the top 1 and 2 for each image of the front covers with or without title fonts, whole body with or without title fonts, and face are shown in Table 4. Here, top 2 means the cumulative classification rate within the top 2. The reason we show the top 2 result is that some images contain multi fonts, so only one candidate output as an estimated typeface is not seemed to be so fair. Also, multiple candidates seem to be helpful and useful for designers and editors to select suitable fonts for front cover design. Table 4 gives that significant differences between images with title fonts and without title fonts. The difference of about 10% between 2 kinds of front-cover images can be regarded as the advantage that DNN models have learned the typeface information from not only the design but also fonts themselves directly. Therefore, about 63% given by front cover images without title fonts seems to be a standard estimated accuracy. Result obtained by face images is almost same as that by front cover images without title and this is a little surprise. One reasons of this result is that a facial impression may affect the selection

of fonts used for the title design in front covers. Fig. 6 shows an example that the typeface used as title font was correctly estimated from an only face image. This example seems to provide the impression of font shape is similar to that derived from the human character's face. Other reason seems that front cover images was normalized into too small ones. In the next experiment, we validate accuracies using images with another size or aspect ratio. On the other hand, the estimation using whole-body images gives lower rates than other types of images. The lower results obtained by whole-body images seem to be caused by variations of composition in whole-body images (See Appendix Fig. 9 and Fig. 10). The difference between 2 kinds of whole-body images, about 4%, is smaller than that between front cover images, about 10%. This reason is that some original whole-body images have no title fonts. Also, as shown in Table 1, several data used in the experiment have multiple typefaces. Therefore, obtaining high rates is very difficult. The fact that all the rates within the top 2 are under 90% despites of a 5 class-classification problem indicates that this task is not very easy.

Table 4: Estimation rates [%] on hard labels.

	Font cover w/title	Front cover w/o title	Body w/title	Body w/o title	Faces
Top 1	72.5	62.8	62.3	58.5	63.0
Top 2	87.9	84.3	79.9	78.8	80.7

(a) Face images (enlarged) (b) Original whole image

Fig. 6: Example of a face image that a typeface of title fonts used in front cover is correctly estimated [15]. Typeface: Mincho. Font name: Marumei Old.

Next, as mentioned above, we validate other image sizes or ratios that are different from default sizes. Table 5 shows estimation rates for every type of images with several image sizes or ratios. On the basis of image sizes in our database, 1.0, 1.5 and 2.0 as ratios of an image height to an image width were used for front cover images. 1.0, 2.0, and 4.0 as ratios of an image height to an image width were assessed for whole-body images. Only face image was normalized with keeping an image ratio is 1.0. Table 5 provides that bigger images have obtained better results. In particular, the increase of rates for front cover images, about 10%, are significant. This suggests that the default size is too small for such images and often lose important information including font one. And normalized images with similar ratio to original's one have almost same rates as enlarged square images. This gives that the characteristics as crucial information for each font class are retained if an original ratio of each image is changed.

Table 5: Estimation rates [%] on hard labels in several image sizes and ratios.

Image height	Image width	Font cover w/title	Front cover w/o title	Body w/title	Body w/o title	Faces
256	256	72.5	62.8	62.3	58.5	63.0
384	256	76.9	67.1	-	-	-
384	384	78.5	65.6	64.9	63.3	62.5
512	128	-	-	60.4	58.7	-
512	256	79.0	70.6	63.3	62.3	-
512	512	81.8	70.1	65.2	62.5	65.4
768	512	80.0	68.0	-	-	63.5
768	768	80.4	72.6	68.5	65.2	63.5
1,024	256	-	-	68.3	64.2	-
1,024	512	81.4	70.4	65.9	64.4	-

Then, we compared experimental results using images with hard labels and soft ones. Table 6 shows estimation rates for 5 kinds of image with each label. Rates with soft labels are obtained using same image sizes/ratios that provided the best rates on hard labels. Compared to results with hard labels, all the results on soft labels are a little lower than those on hard labels. These results suggest that the expression by soft label encoding is not appropriate to our experimental data that have multi classes.

Table 6: Estimation rates [%] on hard and soft labels.

Label	Font cover w/title	Front cover w/o title	Body w/title	Body w/o title	Faces
Hard	81.8	72.6	68.5	65.2	65.4
Soft	78.2	72.1	66.3	63.9	64.6

Finally, we analyzed error results and causes. Main mis-estimations are that Gothic or Maru-gothic typefaces were classified into Designed one. Some kinds of Designed fonts are very similar to Gothic or Maru-Gothic fonts. To solve this kind of errors, we need to introduce another scheme or information.

5 Conclusions

In this paper we have created the database that consists of about 580 images of front cover of Manga and Light novels books, and have proposed the framework for estimating typefaces of title fonts designed in such images for recommending the selection of suitable fonts for the title design. Our database contains five types of images; two kinds of whole front cover images of each books, two kinds of whole-body images extracted from a human character pictured in front cover images, and a character face image that are a part of a human character one. The difference between two kinds for front cover images and whole-body ones is with/without title fonts. In our experiments, we have estimated a kind of typeface of title fonts used in front covers from each type of images. We have exploited the pre-trained DNN model for the font typeface estimation and criteria for each label. Experimental results have shown that the design of front cover of books and human characters contained in the front cover have strong correlation with font typefaces selected for their front cover. The top rate of estimating kinds of typefaces using images without title fonts was about 72% on front cover image and about 65% for whole-body and face images. From these results we can say that estimating typefaces of title fonts in the front cover will enable us to ease the selection of suitable fonts for front cover design by providing candidates in the near future.

Future works are to gather more samples of cover front pages of Manga and Light novels for increasing training samples, correct the imbalance of data size among typefaces, and estimate not only kinds of typeface but concrete font names.

References

1. Suzuki, D., Uru, D.: OniAni. vol.1, Media Factory (2010)
2. Haruba, N.: The Quintessential Quintuplets. vol.7, Kodansha (2019)
3. Shinhara, Y., Karamatsu, T., Harada, D., Yamaguchi, K., Uchida, S.: Serif or sans: visual font analytic on book covers and online advertisements. In Proceedings of International Conference on Document Analysis and Recognition (2019)
4. Imakawa, S., Kodoi, T.: What kind of restaurant does the characters on the signboards conjure up?, Graduate School of Education Bulletin, Hiroshima University, vol.65, pp.249–256 (2016)
5. Akime, J., Morisawa, H.: Otome-Ge no Koryakutaishou ni Narimashita..., vol.2, Ascii Media Works (2012)
6. BNN Editorial Department.: The design of Light Novels, BNN Shin-sha (2018)
7. Sannoumaru, S., Yuzuki R.: Logotype! Ritto-sha, A5 edition (2015)

8. MdN Editorial Department.: The Graphics Design of Manga & Anime, MdN EXTRA vol.1, MdN Corporation (2014)
9. MdN Editorial Department.: The Graphics Design of Manga & Anime, MdN EXTRA vol.2, MdN Corporation (2015)
10. Nichibou Editorial Department.: The New Comics Design, Japan Publicaitons Inc (2020)
11. Nichibou Editorial Department.: The New Comics Design 2, Japan Publicaitons Inc (2021)
12. Morisawa + DESIGNNING Editorial Department.: Type and Font Book, Mainichi Communications (2010)
13. Date C.: Moji-no-Kihon, Graphics-sha (2020)
14. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In Proceedings of International Conference of Learning Representation (2015)
15. Natsumi, K., Shizuki, M.: Hazakura ga Kita Natsu, Ascii Media Works (2008)

Appendix

Other examples of 2 kinds of front cover images, 2 kinds of whole-body images, and face images in our database are shown below.

(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Fig. 7: Examples of front cover images with title fonts.

Fig. 8: Examples of front cover images with no title fonts.

Fig. 9: Examples of whole-body images with title fonts.

Fig. 10: Examples of whole-body images with no title fonts.

Fig. 11: Examples of face images (enlarged).