

## A Study on Improving Free-Hand Sketch Recognition of Infants Using Deep Learning

Mi-Hwa Song<sup>1</sup>

<sup>1</sup>School of Information and Communication Science, Semyung University

**Article History:** Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

**Abstract:** Due to its unique characteristics, infant paintings have a significantly lower recognition rate than adult images. According to the study of infant art, infant paintings have many features that are different from adult images, such as the appearance of many self-centered and exaggerated expressions. In this paper, we will introduce a method to improve the recognition rate of such children's drawings by utilizing deep learning. Create a pre-processor that generalizes the unique characteristics of the child to improve the low recognition rate of the infant figure, and primarily refine the data. High accuracy was obtained as a result of securing and executing 80 adult sketches for each of 250 classified items using CNN, which is often used for image recognition. Through this research, it is expected that it will be possible not only to improve the cognitive ability of infant figures, but also to measure learning ability and child development through infant drawings, and to utilize it in child psychotherapy through emotion recognition.

**Keywords:** Free-Hand Sketch, CNN, Hand Drawn Sketch, Pattern Recognition, Machine Learning

### 1. Introduction

The desire for expression, which is one of the basic human desires, can be found in various art activities. Among them, Free-hand Sketch is a method that human beings have long used as a general means of expressing their intentions that can be freely used without using tools such as persons, compasses, and paints. Moreover, in recent years, due to image recognition and processing technology accompanying the development of artificial intelligence technology, the results of human artistic activities have become systematically well recognized and can be utilized. However, unlike adult freehand sketches, toddler drawings still have a significantly lower recognition rate. This is because the characteristics of infant paintings are displayed in a wide variety of forms, unlike adults, making it difficult to find clear patterns. In this study, in order to increase the recognition rate of such infant images, the characteristics of infant art and children's art studied in art education are classified, and the images of infants are combined with the data of adult images that have already shown high accuracy. Consider a plan that can solve the low accuracy when is included.

### 2. Related Works

#### Features of children's drawings

Infants, unlike ordinary adults, often express their thoughts in their own images, which is often difficult to understand. From the point of view of these infant arts, Lowenfeld is a private fire (2-4 years old), conscious, who draws the main characteristics according to the developmental stage of the infant in a meaningless manner when self-expression is started. Conductive tableware (4-7 years old) trying to make a simple expression process, the graphic period (7-9 years old) when the concept of things is formed, the dawn period (9-11 years old) who is interested in realistic expression of patterns Was classified into 6 stages: a doctor's actual machine (11 to 13 years old) who actually expresses reasonably, and an adolescent (13 to 16 years old) who engages in creative activities [1]. In this study, the main research subjects were images aged 4 years or older who started conscious expression activities by unconscious attempts. Expressions of infant art have characteristics that are commonly displayed. The sloppy expression of the shape that the foot is attached to the head, the anthropomorphic expression that humanizes not the person but the subject, the self-centered form of the section chief and reduction, and the transparent expression that draws the invisible are time. Typical examples are the coexistence expression displayed on the screen as space, and the use of base line and sky line to display the concept of space [2].

#### Deep Learning and Convolutional Neural Networks

Deep Learning is a type of machine learning that learns a lot of data to find patterns, and is a method using a network structure using a neural network with multiple layers. With the development of computer performance and the emergence of a lot of data, sufficient learning is possible, showing great performance in image recognition, speech recognition, intelligent robots, and natural language processing. A convolutional neural network (CNN) is a representative image recognition artificial neural network for deep learning, a convolutional layer that extracts features of an image in the intermediate layer between the input layer and the output layer, and a pooling layer that

reduces the feature maps obtained from the convolutional layer. It is composed of a total bonding layer that combines several units of each side.

**Dataset**

Utilize two types of datasets, cybertron dataset [3] and rendered dataset [4], for the recognition of children's drawings. The cybertron dataset is 80 sketch images in each of 250 categories, and consists of a total of 20,000 images. The Rendered dataset consists of a total of 125 categories, with 500 to 700 images in each category. Forty-seven categories were selected from a total of 125 categories (75481 images), and the total number of data used was 20,000. Let CNN learn the categories and classify them accurately. Divide the test data of the data set into 20% and the train data into 80% and train them. The sample images for each of the two datasets are the same as in Figure 1.

**Feature extraction and classification**

Most of the pictures drawn by toddlers can be easily examined by adults. However, the process of letting a computer know what this picture looks like is not easy. Before training sketch images with CNN (Convolutional Neural Networks), the data must be in a data format that is easy to process. Therefore, the image was resized to a certain size (64 \* 64) and converted to 24-bit RGB format. Therefore, one image is shown in 3 \* 64 \* 64 (all 12288 elements). I made a model with 3 synthetic gopchuns, valid functions (ReLU), and 3 maximum pool layers when building a CNN model. Then I placed two previous bond layers so that I ended up with 5 classes. Figure 1 shows a sketch image of each dataset. Figure 2 shows the system structure for infants and figure recognition. An attractive feature of CNNs is that they serve as a display of useful features for inputting the output of the inner layer.

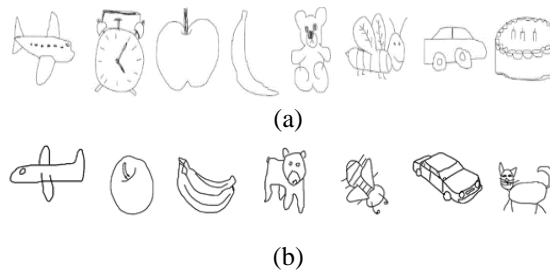


Figure 1. (a) cybertron dataset[3] sample sketches (b) rendered dataset[4] sample sketches

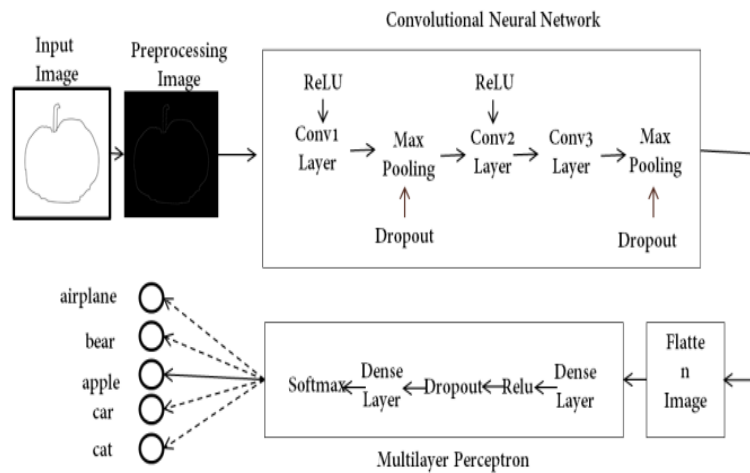


Figure 2 Structure of the system for infant figure recognition

**3. Result**

Table 1 shows the results of classifying about 20,000 images in each data set into CNNs. Looking at the classification results, the cybertron dataset showed better results than the rendered dataset. The accuracy of the model test set was 99.6%. Changes in training loss and testing loss with epoch progression of the models in Fig. 3 and Fig. 4 can be confirmed.

Table 1: Comparison of 2 types of Dataset

cybertron data set[3]			
data	categories	accuracy	loss
Test=14000 Train=6000 Total=20000	250	0.996	0.02
rendered 256*256 data set[4]			
data	categories	accuracy	loss
Test=14352 Train=6150 Total=20502	47	0.976	0.08

Figure 3. Cybertron dataset: loss graph of learning set and test set

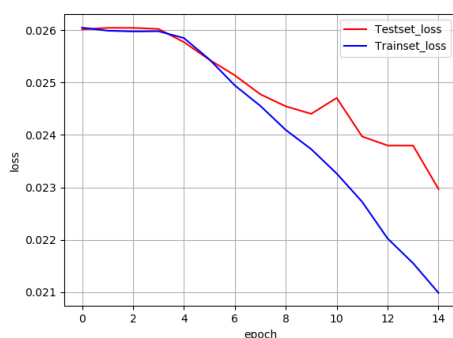
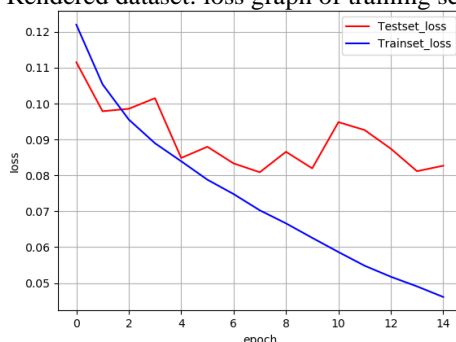


Figure 4. Rendered dataset: loss graph of training set and test set



The error for the training set continues to decrease as the number of epochs increases. It looks like you finished learning before overfitting occurred.

#### 4. Conclusion

In order to increase the picture recognition rate of infants, this study proposed a method for improving infant picture recognition using CNN of deep learning, and the experiment accuracy was 0.996 (99.6%), which shows that there is a significant improvement effect on children's picture recognition. there was. Through this, it is expected that not only the improvement of children's ability to recognize pictures, but also the measurement of learning ability and the use of children's psychological therapy through emotional recognition will be possible.

#### References

1. Grandstaff, L. J. (2012). Children's Artistic Development and the Influence of Visual Culture (Doctoral dissertation, University of Kansas)
2. Coates, E., & Coates, A. (2011). The subjects and meanings of young children's drawings. Exploring children's creative narratives, 86-110.

3. Eitz, M., Hays, J., & Alexa, M. (2012). How do humans sketch objects? *ACM Transactions on graphics (TOG)*, 31(4), 1-10.
4. Sangkloy, P., Burnell, N., Ham, C., & Hays, J. (2016). The sketchy database: learning to retrieve badly drawn bunnies. *ACM Transactions on Graphics (TOG)*, 35(4), 1-12.
5. Wang, F., Kang, L., & Li, Y. (2015). Sketch-based 3d shape retrieval using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1875-1883).
6. Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2008). Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys (Csur)*, 40(2), 1-60.
7. Fu, L., & Kara, L. B. (2009, January). Recognizing network-like hand-drawn sketches: a convolutional neural network approach. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Vol. 49026, pp. 671-681).