ソandン

082340544 HIRAO Momona

ソ and ン

- A kind of Japanese Katakana
- Difficult to distinguish

My interest

What features make ソ and ン different?

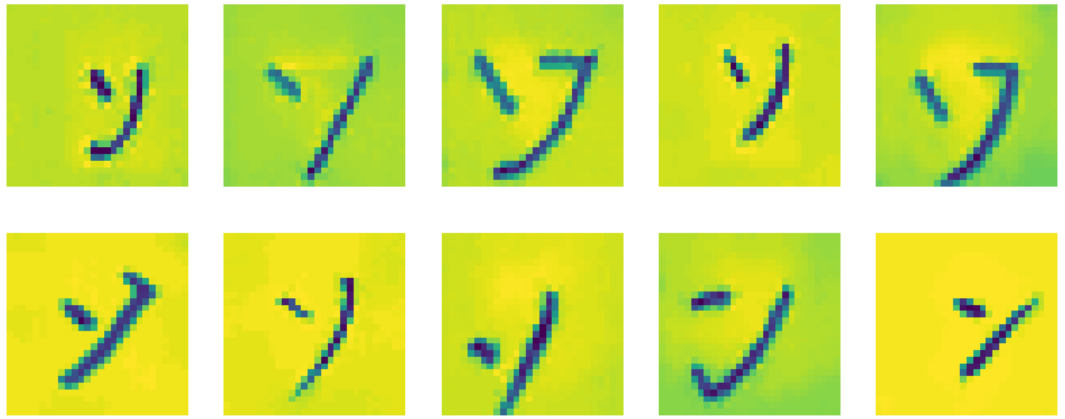
I tried to

1. **classify** images
2. **generate** images of them

Datasets

Used handwritten and printed images from **the ETL Character Database** (http://etlcdb.db.aist.go.jp)

1. Extracted only ソ and ン images
2. Shaped images
   - Resized to 28 × 28
   - Reduced background noise
3. Shaffled them by characters
4. Separated to training and test data sets



Some images of the created dataset

Got **4800** training data and **1219** test data

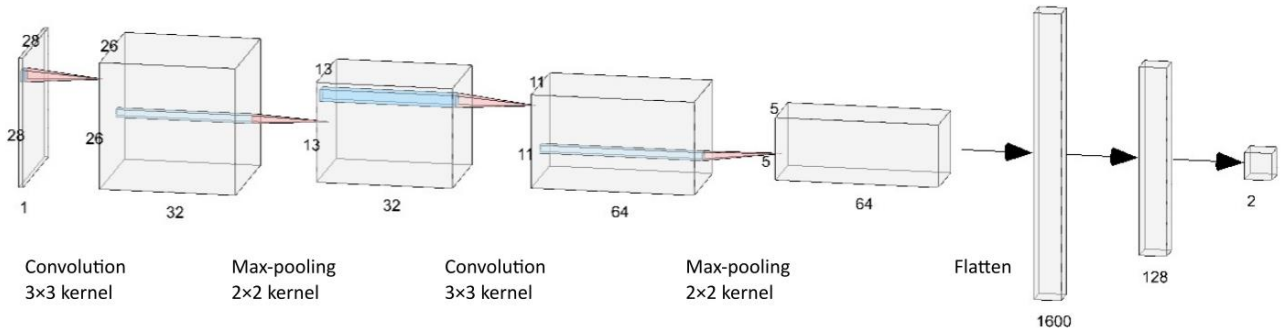
Classify images

CNN with 2 convolution layers + dense layers

Since the amount of data was small, I used CNN which can extract features

I tried very simple CNN, as I did in Homework, but

- Used Batch normalization
- Changed optimizer from SGD to Adam



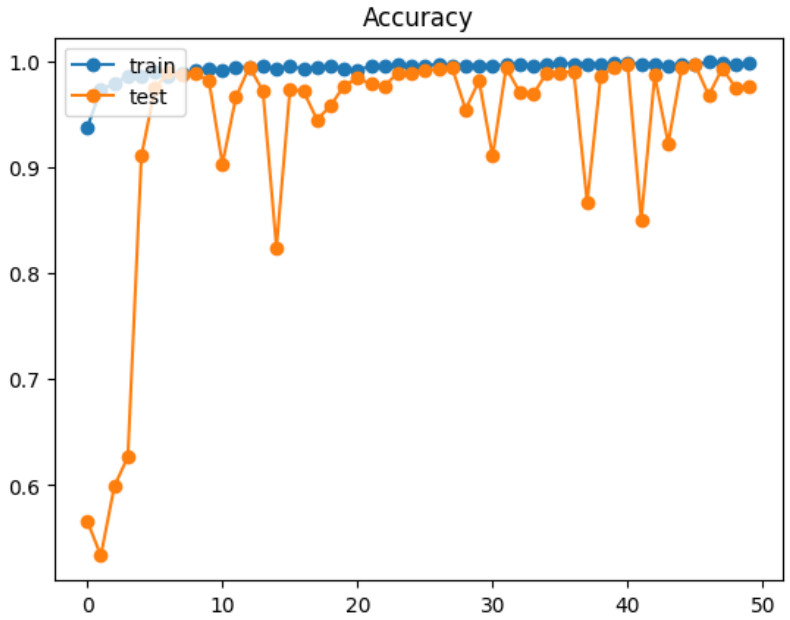
Structure of the CNN model

Classify images

Accuracy

was not stable?

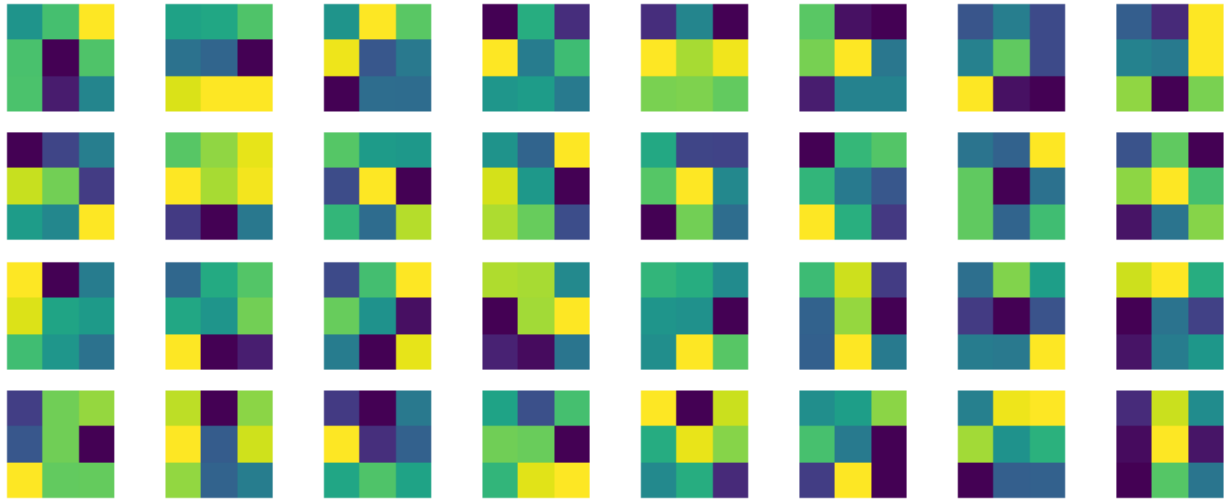
→ because of the number of epochs?



Weights of the first convolutional layer

Look like there are a few differences

→ Did not train well / kernel was small?

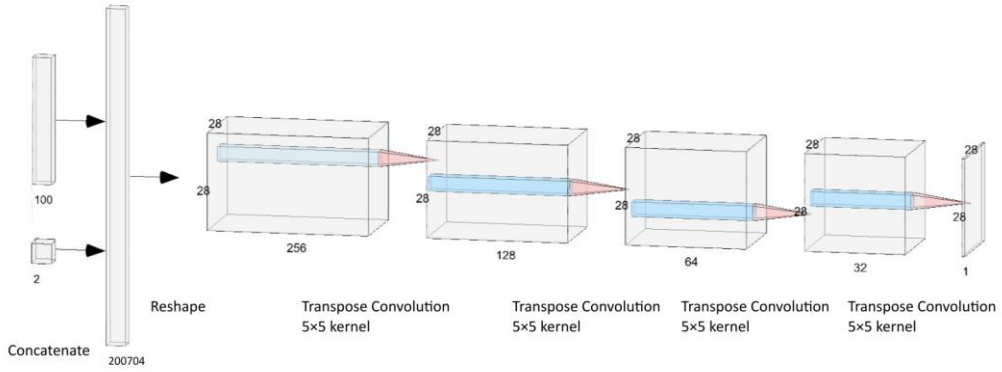


Weights of the first convolutional layer

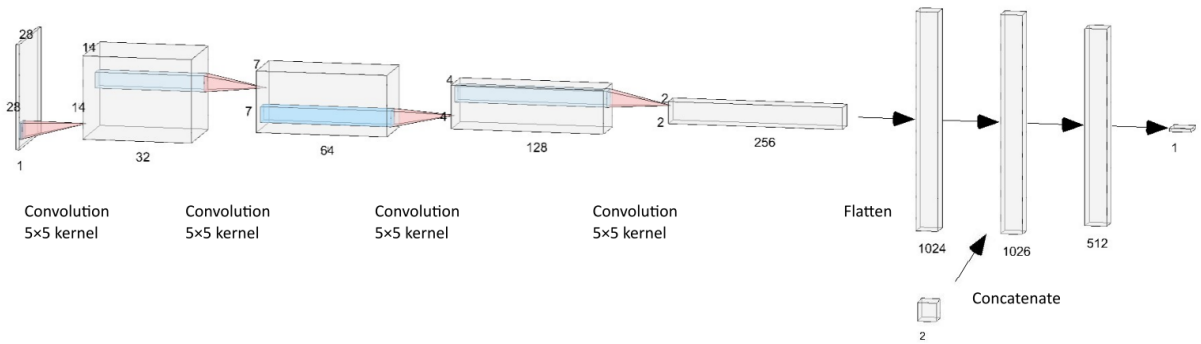
Generate images

CGAN (Conditional Generative Adversarial Nets)

Since I wanted to get images between ソ and ン, I used the CGAN instead of a standard GAN.



Structure of the generator model

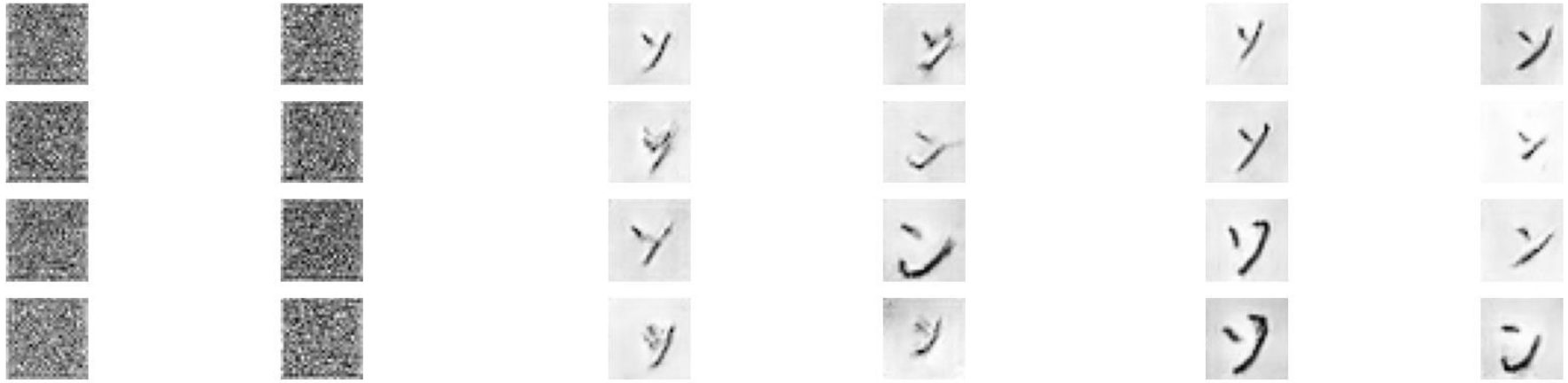


Structure of the discriminator model

Generate images

I mainly referred to https://qiita.com/o93/items/96ca7dbcc82a8dd873ad

And adjusted the model based on https://github.com/soumith/ganhacks



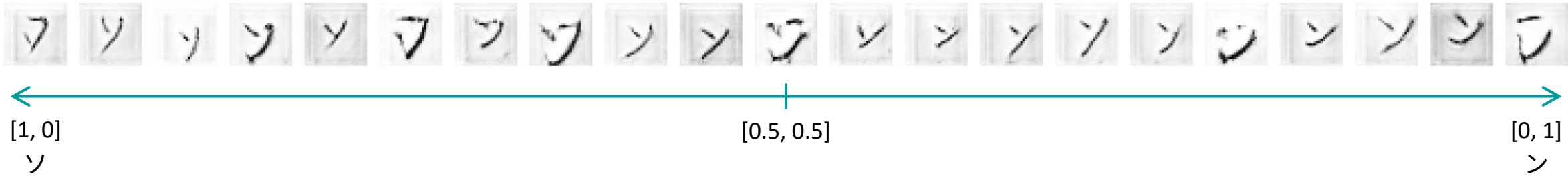
Epoch 0 Epoch 20 Epoch 40

Generate images

Result:

I gave labels and noise vectors to the trained generator and got below

```
labels = np.array([[1-i*0.05, i*0.05] for i in range(20 + 1)])
```



Generated images

Conclusion

- likely to be judged by はらい

  What will happen if I use images drawn with a ballpoint pen?

- Bending at the beginning of writing makes it easier to distinguish

Generated images with annotations