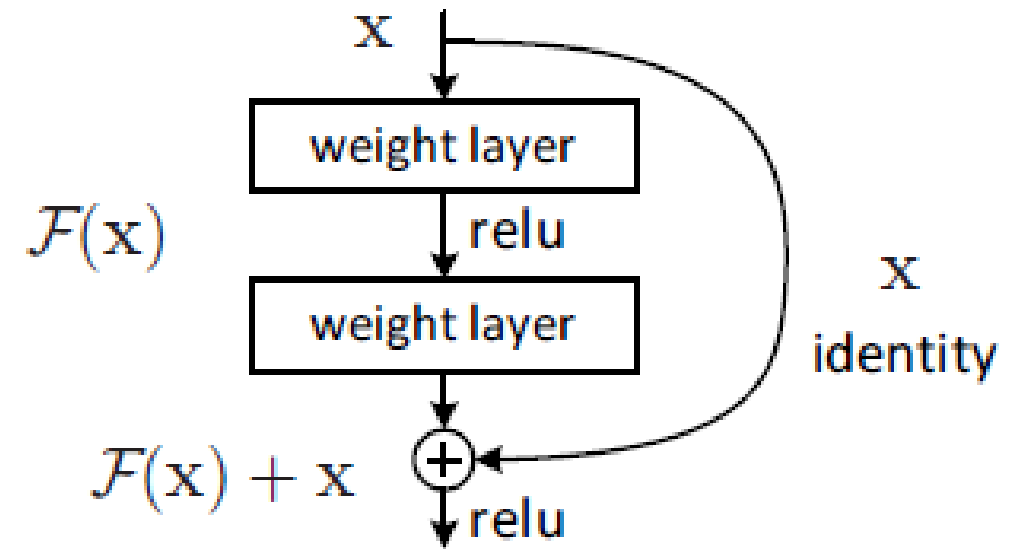


Classification of Hiragana with Plain CNN and ResNet

m5231142 春口 拓磨 (HARUGUCHI Takuma)

Introduction: What is ResNet?

- Residual Network (ResNet) is the convolutional neural network to introduce the residual learning framework.
- The framework prevents performance from getting worse when increasing the number of the layer.
- Therefore, Resnet can have more layers than exiting CNNs.



Problem Statement

- This research topic is to compare the numbers of layers where the accuracy is best for each image datasets about CIFAR-10 and hiragana with the plain CNN and ResNet.

Technical Approach and Models

- Datasets
 - CIFAR-10
 - Hiragana (73 classes) <https://lab.ndl.go.jp/cms/hiragana73>
- Network Model
 - Plain CNN
 - ResNet
- Framework: Pytorch
- Clone and modify the existing implementation of ResNet
 - <https://github.com/a-martyn/resnet>

Modify For the Hiragana Dataset

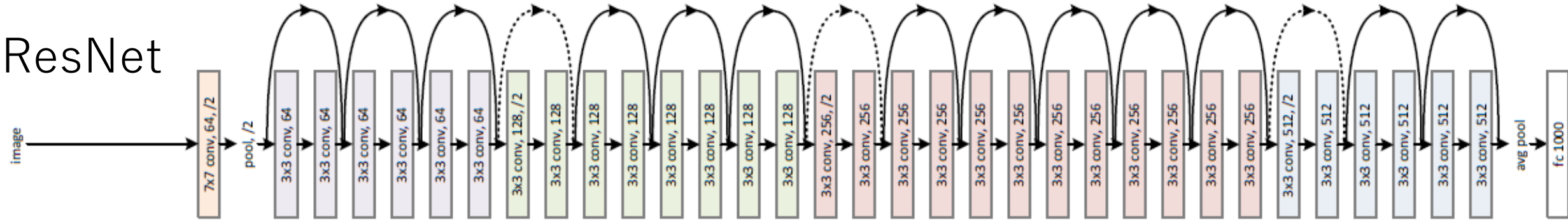
- Resize 48x48 size of hiragana to 32x32
- Delete image transforming of random horizontal flipping and cropping
 - This is because hiragana is character, not objects such as car or bird.
- Change the number of the output layer: 10->73
- Add `torch.nn.DataParallel()` to speed up training

Experiments

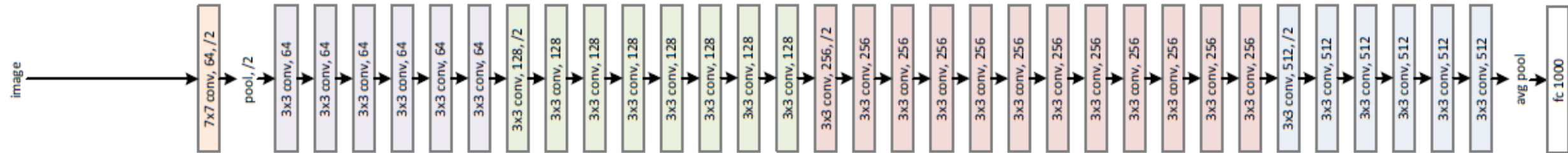
- Train the plain net and ResNet for CIFAR-10 and hiragana dataset
- Compare the numbers of layers and test error for each network.
- Testing Network Model
 - Plain Network (CNN): Layer = 20, 32, 44, 56
 - ResNet: Layer = 20, 32, 44, 56

The Structure of Networks

ResNet



Plain Network



Results: Test Error of CIFAR-10

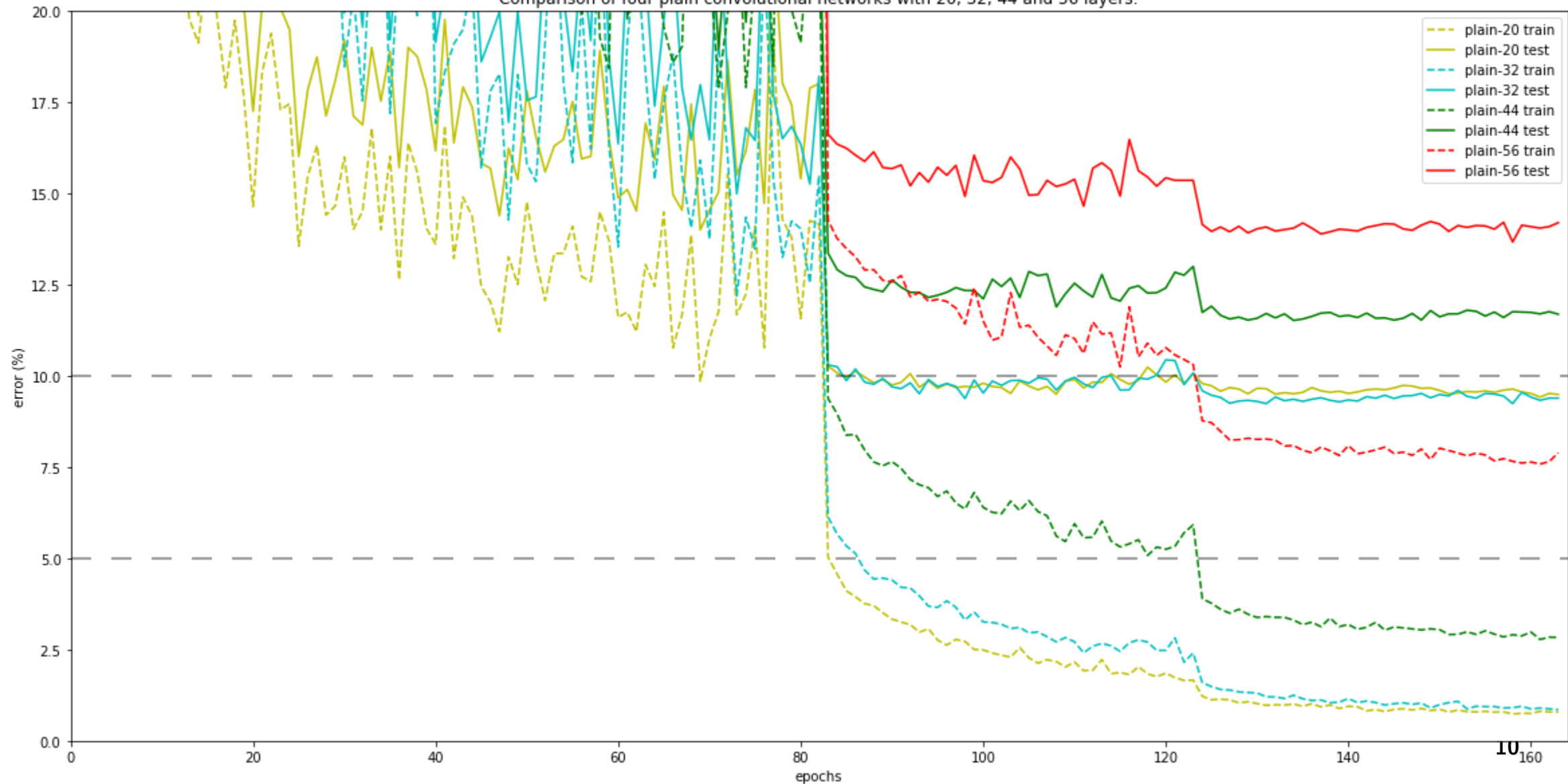
| Model | Test Error | Model | Test Error |
|------------|------------|----------|------------|
| PlainNet20 | 0.0942 | ResNet20 | 0.0842 |
| PlainNet32 | 0.0924 | ResNet32 | 0.0790 |
| PlainNet44 | 0.1152 | ResNet44 | 0.0716 |
| PlainNet56 | 0.1367 | ResNet56 | 0.0770 |

Results: Test Error of Hiragana

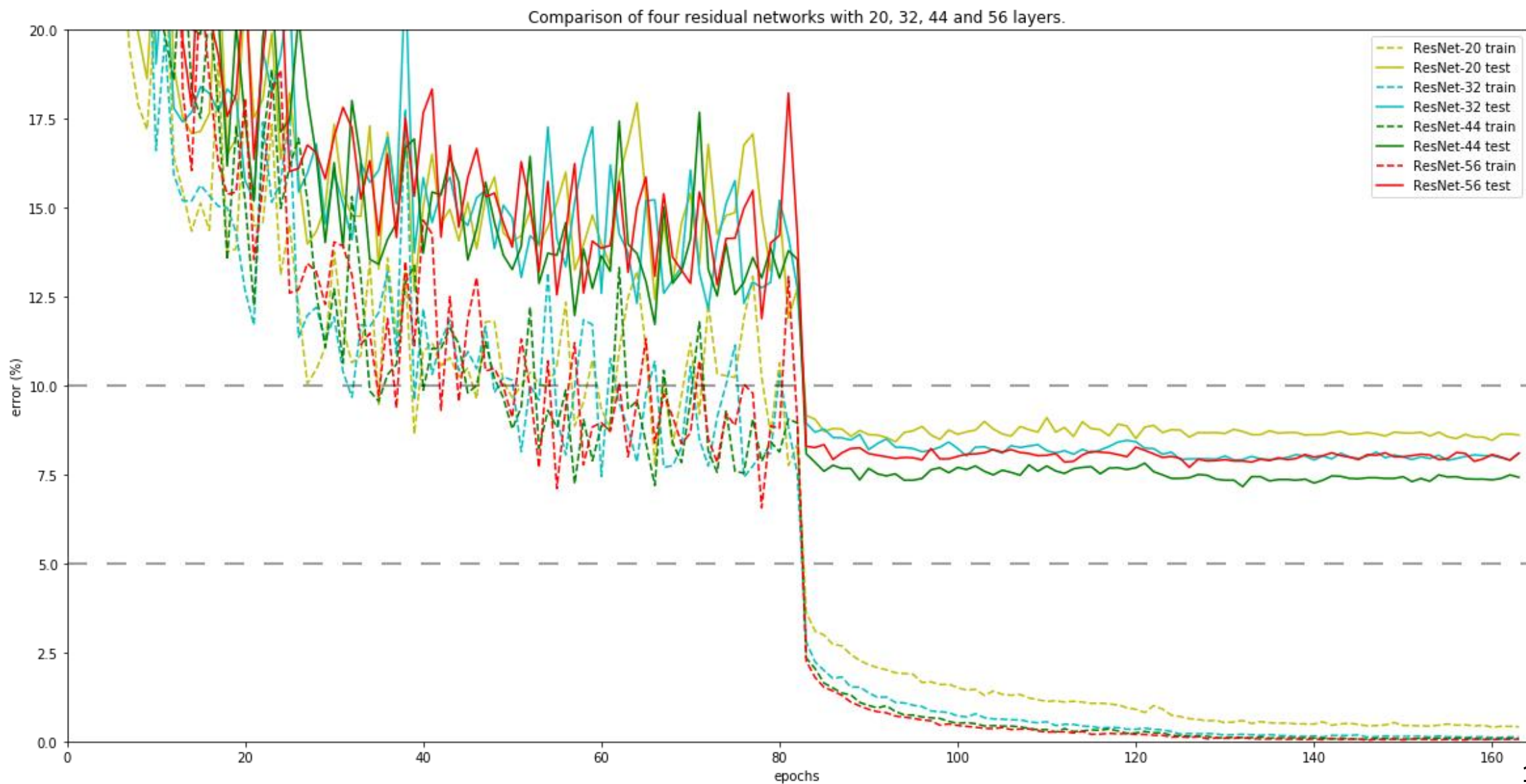
| Model | Test Error | Model | Test Error |
|------------|------------|----------|------------|
| PlainNet20 | 0.003125 | ResNet20 | 0.003563 |
| PlainNet32 | 0.011250 | ResNet32 | 0.005687 |
| PlainNet44 | 0.023688 | ResNet44 | 0.010625 |
| PlainNet56 | 0.077938 | ResNet56 | 0.007000 |

CIFAR-10 Plain Net

Comparison of four plain convolutional networks with 20, 32, 44 and 56 layers.

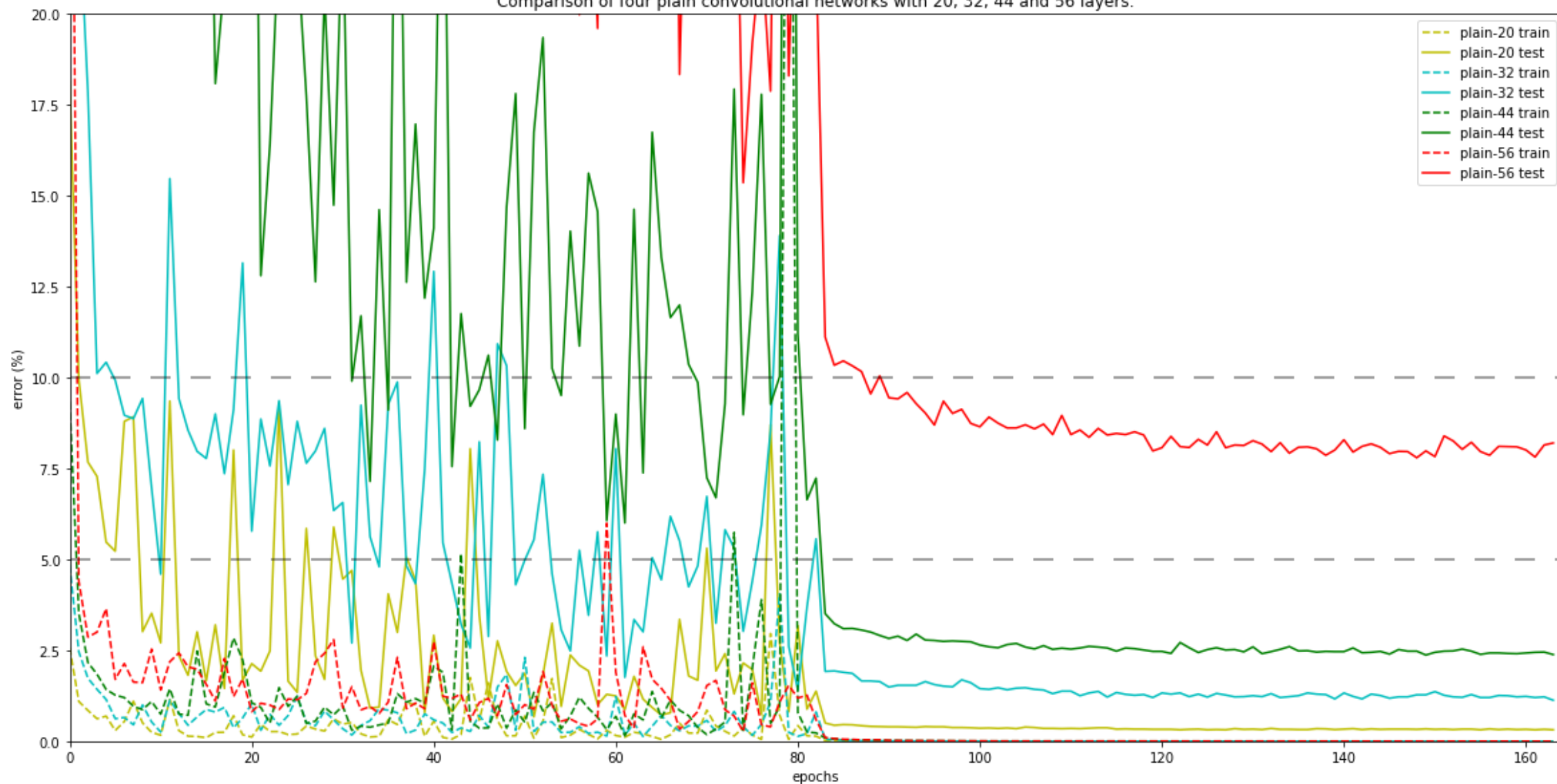


CIFAR-10 ResNet



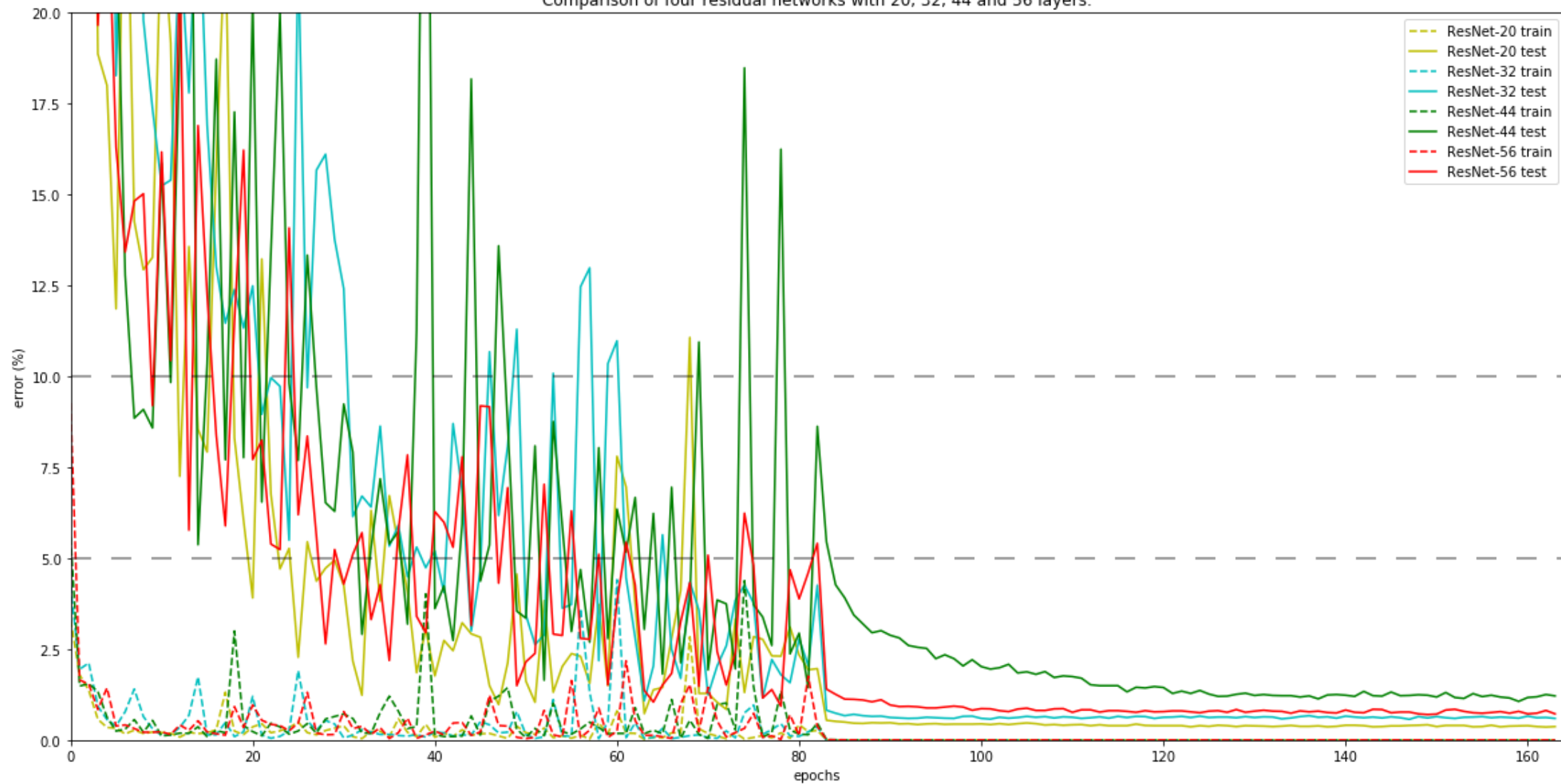
Hiragana Plain Net

Comparison of four plain convolutional networks with 20, 32, 44 and 56 layers.



Hiragana ResNet

Comparison of four residual networks with 20, 32, 44 and 56 layers.



Discussion: The summary of Hiragana Result

- **Similar to CIFAR-10:**
- More the number of the layer in the plain net is, the worse the test error is.
- In the ResNet, the increasing of layers does not strongly worsen the test error.
- **Different from CIFAR-10:**
- The best model of hiragana is PlainNet20 while the best of CIFAR-10 is ResNet44.
 - Note that the best model in CIFAR-10 of the original repository is ResNet56. The cause of the difference may be adding `torch.nn.DataParallel`.
- In both the plain net and ResNet of the hiragana, the number of the layer in the best model is the smallest: PlainNet20 and ResNet20
- The train errors of all models in the hiragana reached to zero while the errors of plain nets in CIFAR-10 do not.

Interpretation of the Result

- Why “The best model in the hiragana is PlainNet20”?
 - It should mean hiragana as dataset is too simple image for CNN to classify it.
 - In fact, the best test error of the hiragana is 0.003125 while the best of the CIFAR-10 is 0.0770. The accuracy of hiragana classification is very high.
 - In addition, the train errors of all models in the hiragana reached to zero. It means that the networks cannot learn the features from the training dataset since the hiragana is more simpler than CIFAR-10.
- **Even if the number of the class is large, it seems not to worse the accuracy of CNN. The complexity of each image seems to affect the accuracy.**
 - Hiragana: 73 classes, CIFAR-10: 10 classes

Reference

- GitHub
 - <https://github.com/a-martyn/resnet>
 - <https://github.com/s1230038/resnet>
 - CIFAR-10 Test Commit: f87b5c49c1314fb0cbcd8012ccbfb15e491f8b92
 - Hiragana Testing Commit: 4bbf87236909c505d241386995ddbc5083745e0c
- He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- https://deepage.net/deep_learning/2016/11/30/resnet.html

Thank You for Listening