ResNet Application for Claissification of Kushusiji Recognition

Chen Steve
Bartow Hanvos Kent School Ningbo Campus
Email: stevechen2018@gmail.com

Abstract: **Object detection based on human resource is gradually being impossible under present circumstance, when the magnitude, quantity and complexity of information is continuously**

growing. With the rapid development of deep learning and computation source, computer program is able to recognize objective in a short time with higher accuracy. Objective Detection with base on Deep Neural Networks is not only used in healthcare diagnosis, face recognition, security and a lot of other fields. In this paper, the problem we faced is recognizing a large quantity of Kuzushiji, a kind of old Japanese written character that is hardly used in modern Japanese society. For they are so different from modern Japanese Language, it is hardly read by not professional person. By the model described in thus paper, this problem can be solved. Due to the huge input dataset and more than 4000 classification, the final outcome has got an accuracy higher than 76% which ranks in the top 20% around the whole competition.

I. Introduction

The problem we faced in Kaggle is Kuzushiji Recognization. Kuzushiji is a kind of written characters in ancient Japan and there are very few fluent readers of Kuzushiji today (only 0.01% of modern Japanese natives). In order to learn about this culture and prevent this character from being lost, a facilitated tool to recognize Kuzushiji is important. The model we use in this problem is Faster R-CNN and ResNet on Tensorflow platform. The part which I'm responsible in our team is argument regulation and coding for Faster R-CNN and ResNet.

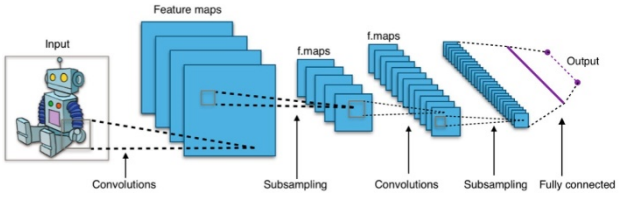
Both Faster R-CNN and ResNet are neural networks that serve for different functions. A neural network contain input layer, output layer and hidden layers. The whole network must contain fully connect layer and activation layer. Each layer has multiple neurons that connected to the next layer along with assigned weights. To update the weights which goal is to minimizing error function also known as loss function is gradient descent. In order to calculate the gradient, based on the chain rule in calculus, every time of training the whole network have evolved both forward propagation and backward propagation.

When we increase the complexity and depth of a simple, plain neural network architecture, the accuracy we get not always improved as we wished. And with the depth increased, we will suffer the diminish of gradient, which make backpropagation hard to proceeding. In this case, a new form of neural network is taken in our model. We use the residual network and successfully solves the problem of gradient disappearance. Residual network is used to classify the characters in our task on the basis of loctations has determined by Faster-RCNN.

II. Material and methodology

Residual Network

Residual Network in our model is for recognizing characters in distinguished region. ResNet can achieve a higher accuracy than the Convolution neural network in identifying characters. Convolution neural network is usually include convolution layer and maxpooling or average pooling or subsampling layers . Convolution Neural Network is forward network. After the network comes to the fully connected layer to output as shown in following figure 1. Each set of elements in previous layer correlate with the next layer, and the derivative from the last layer pass on the previous layers and finally achieves the gradient descent to find a proper weight and bias.

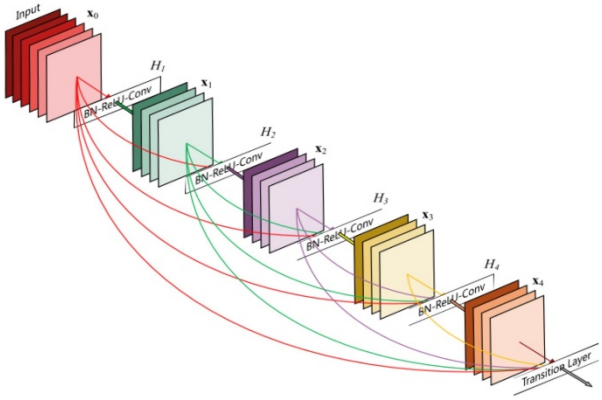


Main formula for Fully Connect-ed Layer:

y = weight * x + bias

The simple neural network can solve simple problems. However, when the complexity of hidden layers increases, vanishing gradient or exploding gradient may take place, which lead to a worse result than that of previous models.

Using residual network is the resolution we chose to solve the shortage of other objective detection network architectures. Comparing with other Convolution Neural Network, the core idea of ResNet is introducing a so-called "identity shortcut connection that skips one or more layers.

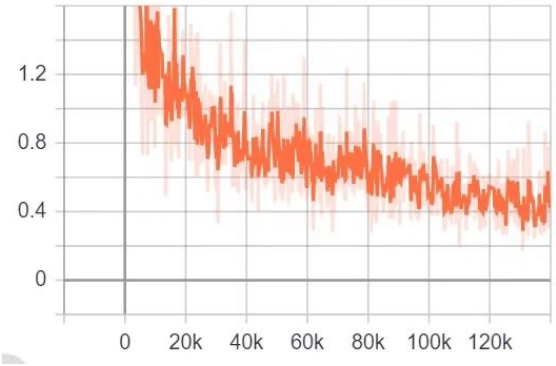


Main Formula for Residual Network:

y = F(x, {W}) + x

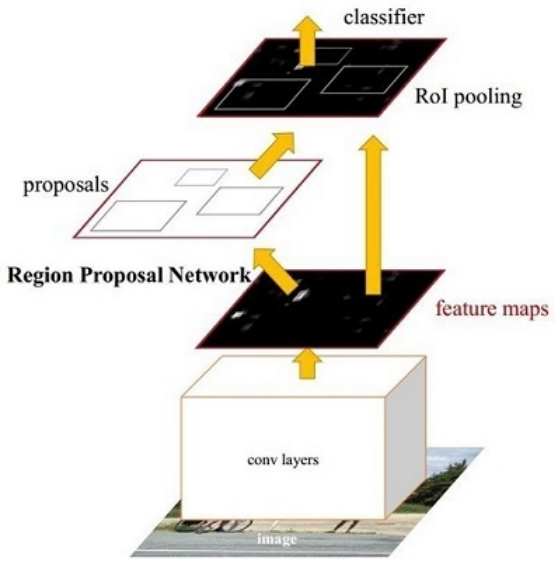
The main difference between simple neural network and residual network is a skip connection. The left one is convolution connection and the right one is the skip connection.

The Loss of recognizing characters



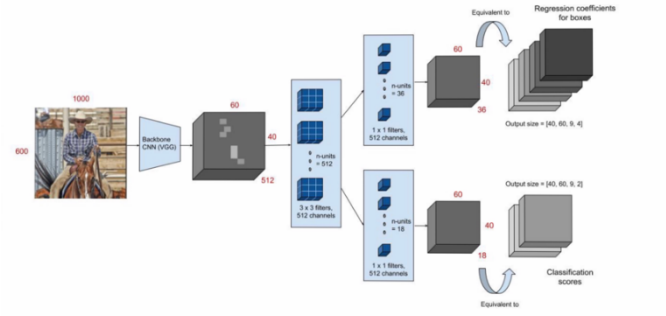
Above figure shows the loss function of ResNet change when we are training ResNet to recognizing the Characters. All other code can be checked in my github.

For the localization of the characters in the pictures, we use Faster RCNN .

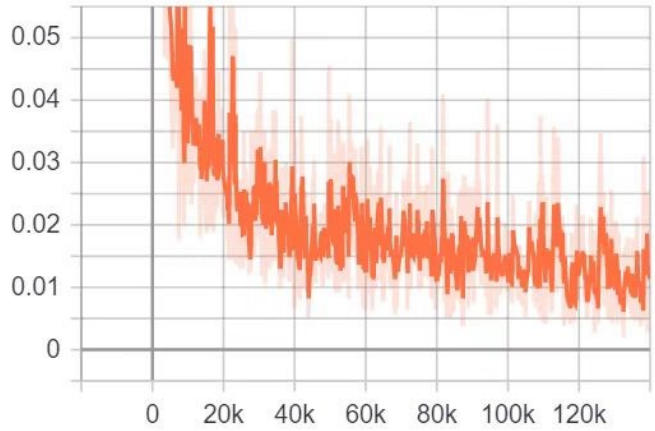


Faster R-CNN in our model is to distinguish regions contain characters. Faster R-CNN is in the family of R-CNN (Regions with CNN) and literally, Faster R-CNN can have a higher efficiency. This network has input layer, inner layers and an output layer. The inner layers here include convolutional layer, activation layer, pooling

layer, inception model and fully-connected layer,



The Loss of recognizing regions:



The diagram above shows how the loss of the Faster R-CNN changes over time.

Evaluation and Discussion

To identify the position of each characters, we used UNet at first but this solution was canceled for a better efficiency of Residual Network. The UNet can also serve for the use as a character-position recognizer, but the architecture of this network is a convolutional network, which requires a intensive operation, causing a lower efficiency than other networks that are not based on convolutional network. This idea carry out this whole project that the efficiency and accuracy are all what we pursue. The model we made now cannot achieve an accuracy that is over 95% and the operation time is too long to be a model that can be used to spread the culture of Kuzushiji

III. Reference

Jesucristo. “Kuzushiji Recognition Complete Guide.” Kaggle. Kaggle, 22 Oct. 2019. Web. 9 Nov. 2019. [https://www.kaggle.com/jesucristo/kuzushiji-recognition-complete-guide/notebook]

Shilpa Ananth. “Faster R-CNN for Object Detection” TowardDataScience, 10 Aug. 2019. Web. 9 Nov. 2019. [https://towardsdatascience.com/faster-r-cnn-for-object-detection-a-technical-summary-474c5b857b46]

Prakash Jay. “Understanding and Implementing Architectures of ResNet and ResNet for state-of-the-art Image Classification: From Microsoft to Facebook [Part 1]”, Facebook AI Research. 8 Feb. 2018. Web. 9 Nov. 2019. [https://medium.com/@14prakash/understanding-and-impleme...resnet-and-resnext-for-state-of-the-art-image-cf51669e1624]

markable performance in application of object detection; it is suitable for recognizing a Kuzushiji character. A basic form of neuron network includes an input layer, several hidden layers, and an output layer. By using thi