**Abstract**

Convolutional neural network (CNN) is anartificial neural network that widely used in object recognition and image classification. It is popular by its features of simple structure, few training parameters and great adaptability. This paper will use Colab to train and evaluate the performances of three CNN models (AlexNet, ResNet and VGG16), and do the analyze for the results.

**Introduction**

Convolutional neural network (CNN) is an artificial neural network based on human visual cognitive model. Through its hierarchical feature extraction method, users can obtain the features of various layers from the low level to the high level. CNN is the foundation of deep learning, it plays an important role especially in the field of computer vision, such as image classification, object detection, instance segmentation and scene capture. From the LeNet proposed in the 1990s to the AlexNet proposed in paper "*ImageNet Classification with Deep Convolutional Neural Networks*" in 2012 , CNN had gone through many developments. From VGG, GoogLeNet to ResNet and DenseNet, the layers of convolutional neural network becomes deeper, and the architecture becomes more complex. The solution to the gradient problem is becoming more ingenious, and the performance of CNN model is getting better. The main purpose of this project is to train and evaluate the performance of CNN model based on Colab.

**Related Work**

1. MNIST Dataset

MNIST Dataset is a classic data set in the field of machine learning. This dataset was first proposed by Yan Lecun in the paper Gradient-Based Learning Applied to Document Recognition (the classic CNN network model LeNet-5 was also proposed in this paper). The MNIST dataset consists of 60,000 training samples and 10,000 test samples, each of which is a 28 x 28 pixel grayscale handwritten digital images. These handwritten digital have been dimensioned, and be located in the center of the image. It's worth mentioning that the MNIST dataset is stored in bytes, so the entire dataset needs to be loaded into the numpy array for training and testing.

1. CIFAR-10 Data Set

CIFAR-10 is a widely used data set which are labeled subsets of 80 million tiny images dataset. It consists 50000 training images and 10000 test images. There are ten classes in the dataset, including airplane, automobile, bird, cat, deer, dog, frog, horse, ship and truck; each class contains 6000 images. There is no overlap in the data set.

To compare with the MNIST data set, CIFAR-10 data set has the following differences: first, the CIFAR-10 data set consisted by three-channel color RGB images, while MNIST data set consisted grayscale images. In addition, the size of images in CIFAR-10 data set is 32 × 32, which is slightly larger than images in the MNIST data set. Finally, unlike the handwritten digital, CIFAR-10 samples contains objects in the real world. This means that the noisy of data set would be larger, and the objects have different proportions and features, which makes it difficult to identify them.