[[1]](#footnote-1)

Image Classification using different datasets on Convolutional Neural Networks

Vaishali Shukla, Thanuja Yarremreddy, Pruthvi Kasani

*Abstract*—In this report, we try to analyze CNN models using different datasets like Cifar 10, Cifar 100 and Fashion MNIST dataset. In the end, we try to compare the results and come up with a conclusion.

# Introduction

T

HIS document explains why we had chosen this problem statement and talk about the solutions that might be helpful. So, initially we will talk about the CNN Model before going to the problem statement and others.

Convolutional Neural Networks (CNNs) is the most popular neural network model being used for image classification problem. The big idea behind CNNs is that a local understanding of an image is good enough. The practical benefit is that having fewer parameters greatly improves the time it takes to learn as well as reduces the amount of data required to train the model. Instead of a fully connected network of weights from each pixel, a CNN has just enough weights to look at a small patch of the image. It’s like reading a book by using a magnifying glass; eventually, you read the whole page, but you look at only a small patch of the page at any given time.

# Problem Statement

The problem of Image Classification goes like this: Given a set of images that are all labeled with a single category, we are asked to predict these categories for a novel set of test images and measure the accuracy of the predictions. There are a variety of challenges associated with this task, so we try to overcome them and come up with a good model with a decent accuracy.

# Related Work

This portion of the document also explains why the CNN model is preferred compared to the other models. It also depends on number of factors Parameters, Network etc. Basic idea of CNN is that the filters work and scan the complete feature matrix along with the dimensional reduction. This is one of the primary reasons why CNN is apt and fits the network for image classification and processing.

Later, we have decided that CNNs are effective for image classification as the concept of dimensionality reduction suits the huge number of parameters in an image.

# Datasets

In this project, we have used different kinds of three Datasets namely Cifar 10, Cifar 100 and Fashion MNIT datasets. We will discuss more about in detail below:

## Cifar 10:

The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly selected images from each class. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class.

The classes are completely mutually exclusive. There is no overlap between automobiles and trucks. "Automobile" includes sedans, SUVs, things of that sort. "Truck" includes only big trucks. Neither includes pickup trucks.

## Cifar 100:

The CIFAR-100 dataset consists of 60000 32x32 color images in 100 classes, with 600 images per class. There are 500 training images and 100 testing images per class. There are 50000 training images and 10000 test images. The 100 classes are grouped into 20 super classes. There are two labels per image - fine label (actual class) and coarse label (superclass).

## Fashion-MNIST:

Fashion-MNIST is a dataset of Zalando's article images—consisting of a training set of 60,000 examples and a test set of 10,000 examples. Each example is a 28x28 grayscale image, associated with a label from 10 classes.

Each image is 28 pixels in height and 28 pixels in width, for a total of 784 pixels in total. Each pixel has a single pixel-value associated with it, indicating the lightness or darkness of that pixel, with higher numbers meaning darker. This pixel-value is an integer between 0 and 255. The training and test data sets have 785 columns. The first column consists of the class labels (see above) and represents the article of clothing. The rest of the columns contain the pixel-values of the associated image.

Labels of this dataset include:

* 0 T-shirt/top
* 1 Trouser
* 2 Pullover
* 3 Dress
* 4 Coat
* 5 Sandal
* 6 Shirt
* 7 Sneaker
* 8 Bag
* 9 Ankle boot

# Preprocessing

In this step, we will try to process the data before fitting it into the model.

* Initially, we import the required libraries and the dataset from the keras.
* Later, we divide the data as testing and training data.
* Now, we try to fetch the total number of classed as well as the unique labels.
* We now print the shapes of testing and training data.
* Next part is to visualize the dataset. So, we print the 15 rows and 15 columns of images in form a matrix.
* We convert the variables y\_train, y\_test from the decimal values into binary i.e., one-hot encoding.
* Now, we reshape the model to have a single channel.
* We print the shape to check if it has the single channel.

We follow the same process for all the three datasets but they differ is few steps (for example we just reshape and normalize the data from the dataset cifar 100)

# Model

In this step, we will train the model, evaluate it and predict it with the testing data.

* Build the network.
* First created a Con2D to perform convolution process.
* Create 32 filters where each filter consists of 3\*3 matrix with an activation function of RELU.
* Input shape is dimension of image which is 28\*28 needed for Input layer.
* Increase the depth of the network by adding more layers.
* Create a fully connected artificial neural network which uses Dense
* Since we have 10 classes, and dataset has 10 classes, so output has to have 10 values.
* Now, we compile model and SGD optimizer.
* Use fit method to train the model using training data.
* Final evaluation of the model using testing data and get the score.
* Feed the x\_data into the model and determining what are the predicted classes going to be.
* Comparing the predicted classes to the true label(y\_test)
* Returning all the binary value to decimal to compare with predicted classes.
* Reshaping x\_test to plot the matrix for prediction
* Confusion Matrix to summarize all the results in one location.
* Later, we can also add more layers and do a dropout to see the prediction results.

A screenshot of a video game

Description automatically generated with medium confidence

A picture containing qr code

Description automatically generated

# References

* Neural Network Projects with Python by James Loy
* Kaggle competiton, [online] Available: https://www.kaggle.com/c/cifar-10/leaderboard. 1.Kaggle competiton, [online] Available: https://www.kaggle.com/c/cifar-10/leaderboard.
* Goodfellow, Y. Bengio and A. Courville, Deep learning, MIT Press, pp. 164-365.
* D. Kingma and J. Ba, "Adam: A method for stochastic optimization", arXiv preprint arXiv:1412.6980, 2014.
* M. D. Zeiler, "ADADELTA: an adaptive learning rate method", arXiv preprint arXiv:1212.5701, 2012.
* E. Hinton et al., "Improving neural networks by preventing coadaptation of feature detectors", 2012/07/03 arXiv preprint.
* Jaswal, D., Vishvanathan, S., Soman K.P.: Image classification using convolutional neural networks. Int. J. Adv. Res. Technol. 3(6) (2014)

## Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 2, pp. 84–90, 2017.

* S. Zhang, R. Benenson, M. Omran, J. Hosang, and B. Schiele, “Towards reaching human performance in pedestrian detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 973–986, 2018.
* J. Lu, G. Wang, and J. Zhou, “Simultaneous feature and dictionary learning for image set based face recognition,” IEEE Transactions on Image Processing, vol. 26, no. 8, pp. 4042–4054, 2017.
* Q. Zheng, M. Yang, J. Yang, Q. Zhang, and X. Zhang, “Improvement of generalization ability of deep CNN via implicit regularization in two-stage training process,” IEEE Access, vol. 6, pp. 15844–15869, 2018.
* Meshkini, K.H., Ghassemian, H.: Texture classification using shearlet transform and GLCM. In: Iranian International Conference of Electrical Engineering (2017)
* Bhatnagar, S.H., Ghosal, D., Kolekar, M.H.: Classification of fashion article images using convolutional neural networks. In: Fourth International Conference on Image Information Processing (ICIIP) (2017)
* Lin, K., et al.: Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. Frontiers Plant Sci. (2019)
* Madhavi, K.V., Tamikodi, R., Sudha, K.J.: An innovative method for retrieving relevant images by getting the top-ranked images first using interactive genetic algorithm. In: 7th International Conference on Communication, Computing and Virtualization (2016)

1. [↑](#footnote-ref-1)