**JavaFX Convolution Network**

**Introduction**

The purpose of this project is to learn the implementation details of a convolution network application programmed with Java and JavaFx. It is assumed you already know how to install the Java software. Even though this sample implements a network, it is not intended to be reusable. Normally, you would use a library such as Deeplearning4j for building a production neural network. If no such library was available, how would you implement a neural network from scratch?

**Prerequisites**

Understanding neural networks requires knowledge of matrix algebra, partial derivatives, and statistics. Previous experience with linear regression or multilinear regression is also helpful.

The application is based on Java version 18 (JDK 18), which may be downloaded from:

https://jdk.java.net/18/ ->

JDK 18.0.2.1 download

The JavaFx version 18 may be downloaded from:

https://gluonhq.com/products/javafx/ ->

18.0.2 SDK download

You must download the MNIST dataset, which contains thousands of images of handwritten digits 0 to 9. The png image format used by this application was downloaded from: https://github.com/myleott/mnist\_png.

The file structure will look like the following, where the digit images are separated into individual folders 0, 1, ... 9. The advantage of having separate folders is to allow the user to examine specific digits. Another advantage of having separate folders is that we can quickly observe how many images there are for each digit. An equal number of images is read from each folder to perform a network run.

Note that even though the total number of training images is over 40,000, the individual digit folders do not contain the same number of files. The folders for digits 4 and 5 contain only about 2700 images, while the digit 7 folder contains about 6200 images. Therefore, to avoid biased training results, the total number of samples to load should not exceed 25000, 2500 from each folder. This will ensure reading an equal number of images from each traning digit folder.

Similarly for the testing dataset, the digit folders do not contain equal numbers of images. The digit 5 folder contains 892 images, while digit 1 folder contains 1135 images. Therefore, to avoid biased testing results, the total number of samples to load should not exceed 8500, 850 from each folder. This will ensure reading an equal number of images from each testing digit folder.

A picture containing chart

Description automatically generated

After the images are read in, they are combined and randonly shuffled before the network training or testing task begins.

**Code Structure**

The project is a desktop application which assumes a wide screen size. The code wasdivided into front end and back end sections, as shown in the diagram. The front end provides a user interface written in JavaFx, while the back end is launched via a JavaFx concurrent task. There is no server involved.

A picture containing chart

Description automatically generated

**Frontend**

The front end consists of several JavaFx screens. The main program is FxConvoMain, which may be invoked with the windows batch file: FxConvoNet.bat. The startup code sequence outline is:

FxConvoMain

main

Application.launch

start

initLogging

initComponents

tabs.initPanel

fxMenu.initViewer

fxMenu.initMenuBar

initMenuApp

initSettingsMenu

initOptionMenu

initHelpMenu

stage.setScene

stage.show

The Setup menu has options for creating, exporting, and importing the network configuration.

Graphical user interface, text, application

Description automatically generated

The Setup menu -> Import Network Configuration opens the file dialog to read in the network configuration file, which is in json format. Then the network config screen opens in a new tab, Import Config, with all fields populated from the json file. The code sequence outline is:

AppMenu. initSettingsMenu

configView.importConfig

ViewUtil.openFileDialog

FileUtil.getInputStream

JsonUtil.jsonToConfig

configView.setAllLayers

initNetConfigPanel

createGeneralConfig

createInputConfig

createConvoPoolConfig

createInternalConfig

createOutputConfig

The user may create a new network configuration as follows. Click on Setup menu -> Create Network Configuration to open the empty network config screen in a new tab, Create Config.

Graphical user interface, text, application

Description automatically generated

The code sequence outline is:

AppMenu. initSettingsMenu

configView.createAllLayers

initNetConfigPanel

createGeneralConfig

createInputConfig

createConvoPoolConfig

createInternalConfig

createOutputConfig

The Import config panel contains all the basic information to perform a network training run. General parameters includes the directories where the image data is located, number of samples to load, and back propagation parameters. The gradient descent sub-panel includes minmum rate, maximum rate, rate decay parameters, and type of rate decay function: triangle or step.

Graphical user interface, text, application, email, website

Description automatically generated

The Input layer panel has input fields for image pixel height and width which is necessary for matrix sizing, described in the back end section.

Graphical user interface, application, Teams

Description automatically generated

The Convolution/Pool Layer panel contains sizing information, and options to choose the action function. The Pool activation function is set to None and cannot be changed.

Graphical user interface, application

Description automatically generated

The Internal Layer panel has fields for activation function and output node size. The input node size is determined by the previous layer output matrix size.

Graphical user interface, text, application, email

Description automatically generated

The Output Layer panel has fields for activation function and output node size, which equals 10 for the digit images.

Graphical user interface, application

Description automatically generated

After the entering the parameters, the user may export the network configuration. The code sequence outline is:

AppMenu. initSettingsMenu

configView.exportConfig

ViewUtil.openCreateFileDialog

FileUtil.getOutputStream

JsonUtil.configToJson

The Process menu -> Display sample item opens a file dialog to allow the user to select an image to examine.

Graphical user interface, text

Description automatically generated with medium confidence

The code sequence outline is:

AppMenu.initOptionMenu

pixelView.displayImage

imageFileDialog

initData

ImageDataUtil.loadImageData

viewData

imageView.setImage

ImagePane(imageMatrix)

The actual image is shown on the left of the screen. On the Normalized Image Grid, the pixel values are displayed after being normalized 0.0 to 1.0, where 0 is black and 1 is white. Values less than 0.25 are not displayed.

Chart

Description automatically generated

The Process menu -> Train the Network item opens the network task screen under a new tab, Train.

Graphical user interface, text

Description automatically generated

Graphical user interface, application, table

Description automatically generated

The buttons at the top of the network task screen allow the user to run the network task, cancel a running task, and to take a snapshot of the performance charts.



Currently, there are three screens that display network performance information. The Cumulative Performance Plot and Performance Hitogram are updated periodically as the task is running. The Summary Confusion Matrix is available after the run has completed.

Table

Description automatically generated

Chart, bar chart, histogram

Description automatically generated

Chart, waterfall chart

Description automatically generated

After the training run has completed, the user may export the training model parameters, which includes all weights and biases from each layer. See the Setup menu -> Export Training Model fit Parameters.

**Graphical user interface, text, application, chat or text message

Description automatically generated**

The training model parameters may be imported before running the test cases. The testing dataset should include a different set of images than the training dataset.

**Backend**

The back end implements a convolution network for image classification on the MNIST dataset, which consists of several thousand images of handwritten digits 0 to 9.

Informative references:

ref: https://www.analyticsvidhya.com/blog/2020/02/mathematics-behind-convolutional-neural-network/

ref: https://towardsdatascience.com/deriving-the-backpropagation-equations-from-scratch-part-2-693d4162e779

eta triangle function:

ref: https://www.jeremyjordan.me/nn-learning-rate/

**Model Notation**

In each layer:

X = input matrix (e.g. matrix of image pixel values)

W = weight matrix

B = bias matrix

Z = W \* X + B

S = activation function (e.g. sigmoid, RELU, softmax)

Y = S(Z)

Y = output matrix

L = loss function (e.g. sum square error, log likelihood)

**Matrix Sizes**

Input Layer:

(Rows, columns) of input image = (n, n)

n = image size

Convolution Layer:

filter matrix = (f,f),

f = filter size

let nf = n - f + 1

convolution output matrix = ( nf , nf )

Pool Layer:

pool matrix = (p,p)

let nfp = (n-f+1) / p = nf / p

pool output matrix = ( nfp , nfp )

number of pool output nodes:

pOut = nfp \* nfp \* numFilters

Internal Layer:

internal layer input matrix = (pOut,1)

internal layer output matrix = (iOut, 1)

Output Layer:

"output layer" input matrix = (iOut, 1)

"output layer" output matrix = (oOut, 1)

Example:

Input Layer:

input image = (28, 28), n = 28

1st Convolution Layer:

filter matrix = (5,5), f = 5

nf = 28 - 5 + 1 = 24,

convolution output matrix = (24 , 24)

numFilters = 20

1st Pool Layer:

pool matrix = (2,2), p = 2

nfp = nf/2 = 24/2 = 12,

pool output matrix = (12 , 12)

pool output nodes:

pOut = 12 \* 12 \* 20 = 2880

2nd Convolution/Pool Layer:

filter matrix = (5,5), f = 5

nf = 12 - 5 + 1 = 8,

convolution output matrix = (8 , 8)

numFilters = 50

2nd Pool Layer:

pool matrix = (2,2), p = 2

nfp = nf/2 = 8/2 = 4,

pool output matrix = (4 , 4)

pool output nodes:

pOut = 4 \* 4 \* 50 = 800

Internal Layer:

internal layer input matrix, (pOut,1) = (800, 1)

internal layer output matrix, (iOut, 1) = (500, 1)

Output Layer:

output layer input matrix, (iOut, 1) = (500, 1)

output layer output matrix, (oOut, 1) = (10, 1)

**Forward Propagation**

During forward propgation, each layer begins with input matrix X, we calculate Z = W \* X + B, and then Y = S(Z). Y becomes the matrix X for the following layer. Each layer may have different weights W, biases B, and activations S.

The example in this application follows the propagation sequence:

Input layer -> 1st Convolution layer -> 1st Pool layer -> 2nd Convolution layer -> 2nd Pool layer -> Internal layer -> Output layer

The forward propgation code sequence is:

ConvoNetTrain.fit

fitBatch

trainAllLayers

convoLayer.trainForward

poolLayer.trainForward

MTX.listToSingleCol

internalLayer.trainForward

outputLayer.trainForward

updateEval

outputLayer.setActualY

outputLayer.updateBatchLoss

**Back Propagation**

Training Objective: iterate over multiple inputs, update weight and bias to minimize the loss. We update the weights and biases going in the reverse direction, from output layer backward to input layer.

Input layer <- 1st Convolution layer <- 1st Pool layer <- 2nd Convolution layer <- 2nd Pool layer <- Internal layer <- Output layer

The back propgation code sequence is:

ConvoNetTrain.backProp

outputLayer.batchLossFn

outputLayer.backProp

internalLayer.backProp

poolLayer.backProp

convoLayer.backProp

Back propagation involves relating the change in weights W and biases B to the change in output value Y. At each iteration either the predicted digit is correct, or it is the wrong digit. The difference between predicted Y and actual Y turns out to be related to the change in the loss function L with respect to the change in Z. Recall that Z = W \* X + B. Therefore, we can relate derivatives of L to derivatives of W and B.

Back propagation begins at the Output layer. We calculate the matrix partial derivative of loss as follows:

predicted y = softmax(z)

parital loss with respect to z:

Normally, calculating dL/dZ is a two step process:

dL/dZ = (dL/dY)\*(dY/dZ)

However, for the sofmax activation function, the result is simply the difference between predicted and actual y values.

dL/dZ[k] = (predicted y[k]) - (actual y[k])

where k is the class index (0 to 9 for digit classification)

Next find dL/dX for back propagation:

dL/dX = (dL/dZ)\*(dZ/dX)

Note that dL/dX for a layer is set equal to dL/dY on the previous layer. This allows back propagation as follows:

Output layer

dL/dW = change in loss due to change in weight

dL/dB = change in loss due to change in bias

eta = gradient descent rate

Weight correction at each iteration from i to i+1:

W(i + 1) = W(i) - eta\*(dL/dW)

Bias correction at each iteration from i to i+1:

B(i + 1) = B(i) - eta\*(dL/dB)

How to calculate dL/dW?

dL/dW = (dL/dY)\*(dY/dZ)\*(dZ/dW)

How to calculate dL/dB?

dL/dB = (dL/dY)\*(dY/dZ)\*(dZ/dB)

Find dL/dX for back propagation:

dL/dX = (dL/dY)\*(dY/dZ)\*(dZ/dX)

How to back propagate to previous layer?

dL/dY previous layer = dL/dX current layer,

Then apply dL/dW calculation as above to update weight and bias in previous layer,

Then calculate dL/dX in previous layer for back prop input (dL/dY) to next previous layer

**Conclusion**