

- ✓ MNIST is a database of handwritten digits.
- ✓ MNIST stands for Mixed National Institute of Standards and Technology database.
- ✓ MNIST contains two sets (training set “60000” examples) and (test set “10000” examples”)
- ✓ For files available on the MNIST database:

train-images-idx3-ubyte.gz: training set images (9912422 bytes)

train-labels-idx1-ubyte.gz: training set labels (28881 bytes)

t10k-images-idx3-ubyte.gz: test set images (1648877 bytes)

t10k-labels-idx1-ubyte.gz: test set labels (4542 bytes)

How to extract images from the MNIST file using MATLAB:

1. You must download the for files that mentioned above from
<http://yann.lecun.com/exdb/mnist/index.html>
2. Unzip the files that you downloaded.
3. train images files train-images-idx3-ubyte file consists of a 28 by 28 row wise image pixel.
4. To read this file using MATLAB

```
Xtrain = fopen('train-images-idx3-ubyte','r','b'); % to open the binary file
```

```
MN=fread(Xtrain,1,'int32')
```

```
MN =
```

```
2051
```

```
Im_num= fread(Xtrain,1,'int32')% to read the number of images
```

```
Im_num =
```

```
60000
```

Note: for this file the data pixels of the image stored from byte 16th to the end

```
[800 (16+28*28)].
```

Since the data images started at byte 16th you can find it and start reading the data by using the following commands:

```
fseek(Xtrain,16,'bof');
```

```
img= fread(Xtrain,28*28,'uchar'); % each image has 28*28 pixels in unsigned byte format
```

```

Y=zeros(28,28);

for i=1:28

Y(i,:)=img((i-1)*28+1:i*28);

end

imshow(Y); % to show the first image in the training file

```

Reading labels:

To read the labels

```

TrainI_l = fopen('train-labels-idx1-ubyte','r','b'); % first we have to open the binary file

MagicNumber=fread(TrainI_l,1,'int32')

MagicNumber =

        2049

Labels_num= fread(TrainI_l,1,'int32')% Read the number of labels

Labels_num=

        60000

The first label is stored in 8th byte of file in unsigned byte format:

fseek(TrainI_l,8,'bof');

fread(TrainI_l,1,'uchar')

ans =

5

```

You can use the following functions to read the MNIST images and labels

```

IM = LoadMNISTImages (' train-images.idx3-ubyte');

Labels = LoadMnistLabels(train-labels.idx1-ubyte');

Labels = Labels' % to get the label transpose

```

```

function labels = loadMNISTLabels(filename)
fp = fopen(filename, 'rb');
assert(fp ~= -1, ['Could not open ', filename, '']);

magic = fread(fp, 1, 'int32', 0, 'ieee-be');
assert(magic == 2049, ['Bad magic number in ', filename, '']);

numLabels = fread(fp, 1, 'int32', 0, 'ieee-be');

labels = fread(fp, inf, 'unsigned char');

assert(size(labels,1) == numLabels, 'Mismatch in label count');

fclose(fp);

end

```

```

function images = loadMNISTImages(filename)
fp = fopen(filename, 'rb');
assert(fp ~= -1, ['Could not open ', filename, '']);

magic = fread(fp, 1, 'int32', 0, 'ieee-be');
assert(magic == 2051, ['Bad magic number in ', filename, '']);

numImages = fread(fp, 1, 'int32', 0, 'ieee-be');
numRows = fread(fp, 1, 'int32', 0, 'ieee-be');
numCols = fread(fp, 1, 'int32', 0, 'ieee-be');

images = fread(fp, inf, 'unsigned char');
images = reshape(images, numCols, numRows, numImages);
images = permute(images,[2 1 3]);

fclose(fp);

% Reshape
images = reshape(images, size(images, 1) * size(images, 2), size(images, 3));
% Convert the image into double and rescale it to [0,1]
images = double(images) / 255;
images = images';

end

```

Also, you can show some examples from the MNIST by writing the following code

```

Figure
Colormap(Gray)

```

```
for 1=1:50 % to show the first 50 images from the chosen file
X=reshape(images(:,i),[28,28]);
Imshow(X);
Title(num2str(labels(i)))
end
```