Handwritten Character Classification Using Convolutional Neural Net

By: Chris Corona and Nolan Walker

Motivation

Computer vision

- Video / image classification
- Object recognition
- Recommendation systems
- Natural language processing

Convolutional Neural Network

A neural net that is especially good with images because it has a built-in filtering mechanism.

- Assign importance to various aspects in an image
- Captures spatial and temporal dependencies
- Learn high-level features

Convolutional Layer

Apply kernel to obtain convolved feature

- Reduce image into form that is easier to process, without losing features
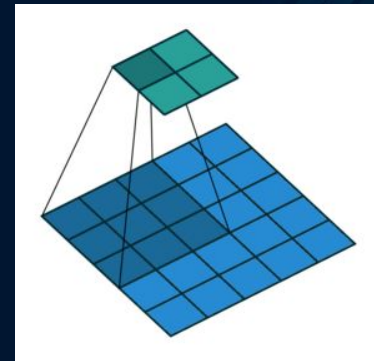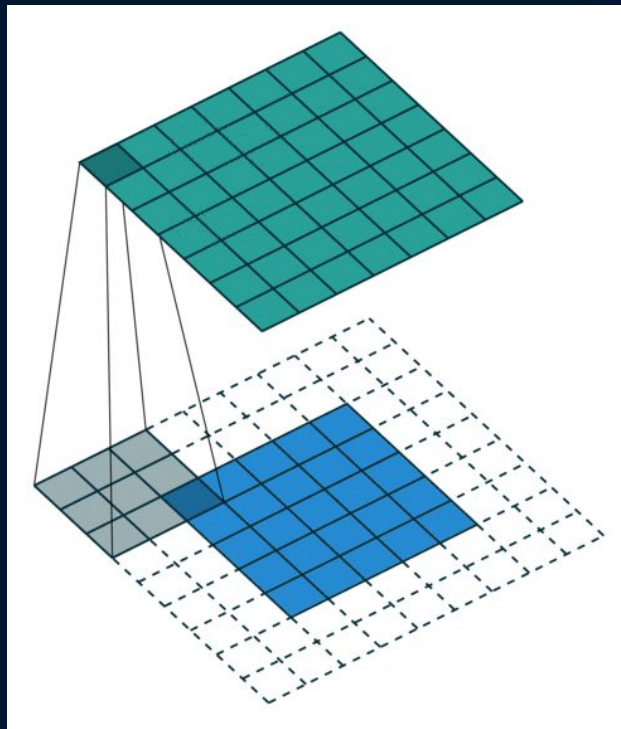

Image

Convolved Feature

Padding and Stride

Change in dimensionality from choice of padding and stride

- Padding for same / increased dimensionality
- No padding for reduced dimensionality
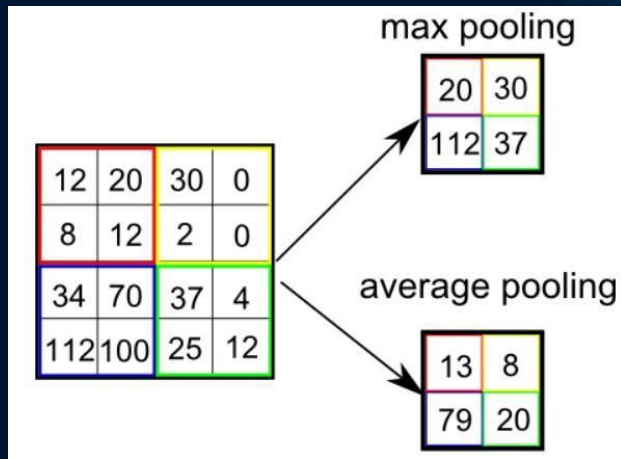
- Stride for more reduction

Pooling Layer

Way to reduce a layer's complexity, downsample the number of features conveyed

- Max
- Average
- Global: at the end

Activation Function

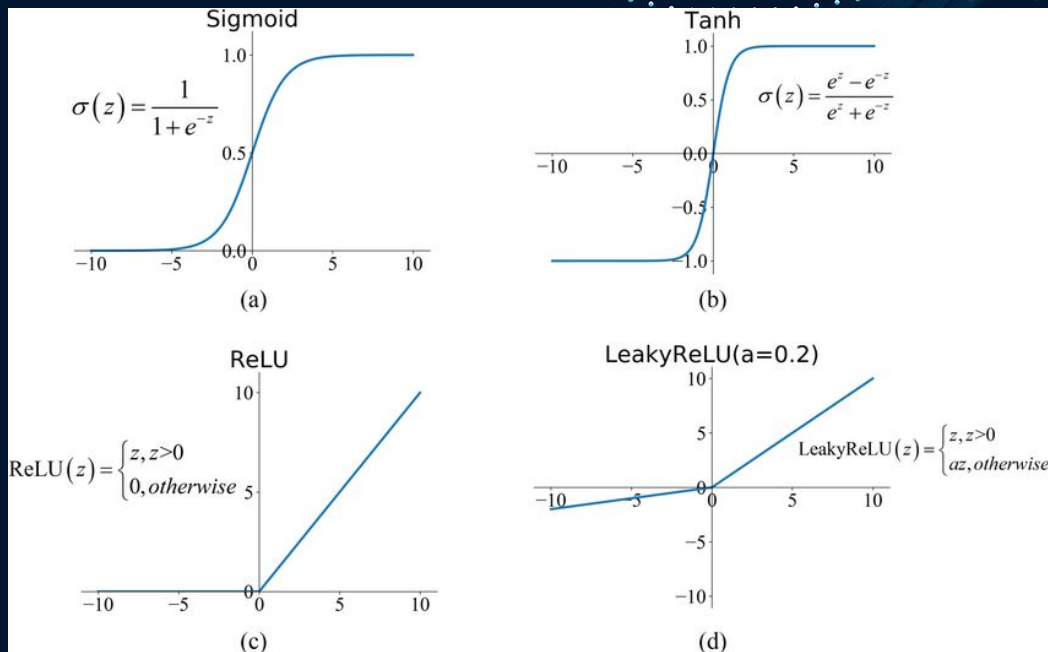Introduce non-linearity to the model

- Sigmoid
- Tanh
- ReLu
- LeakyReLu

Sigmoid

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

(a)

Tanh

$$\sigma(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

(b)

ReLU

$$ReLU(z) = \begin{cases} z, z > 0 \\ 0, otherwise \end{cases}$$

(c)

LeakyReLU(a=0.2)

$$LeakyReLU(z) = \begin{cases} z, z > 0 \\ az, otherwise \end{cases}$$

(d)
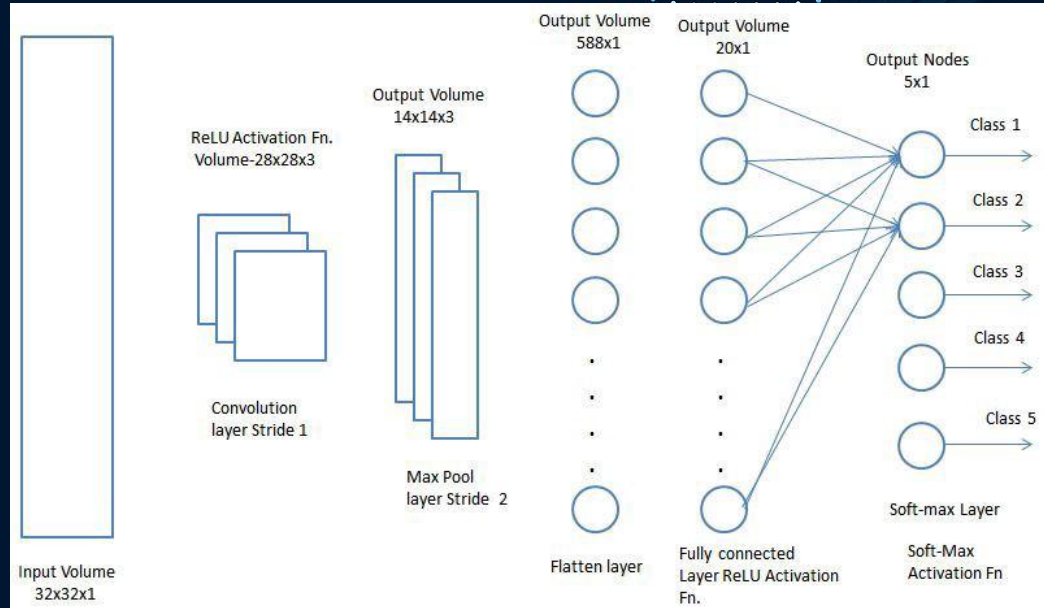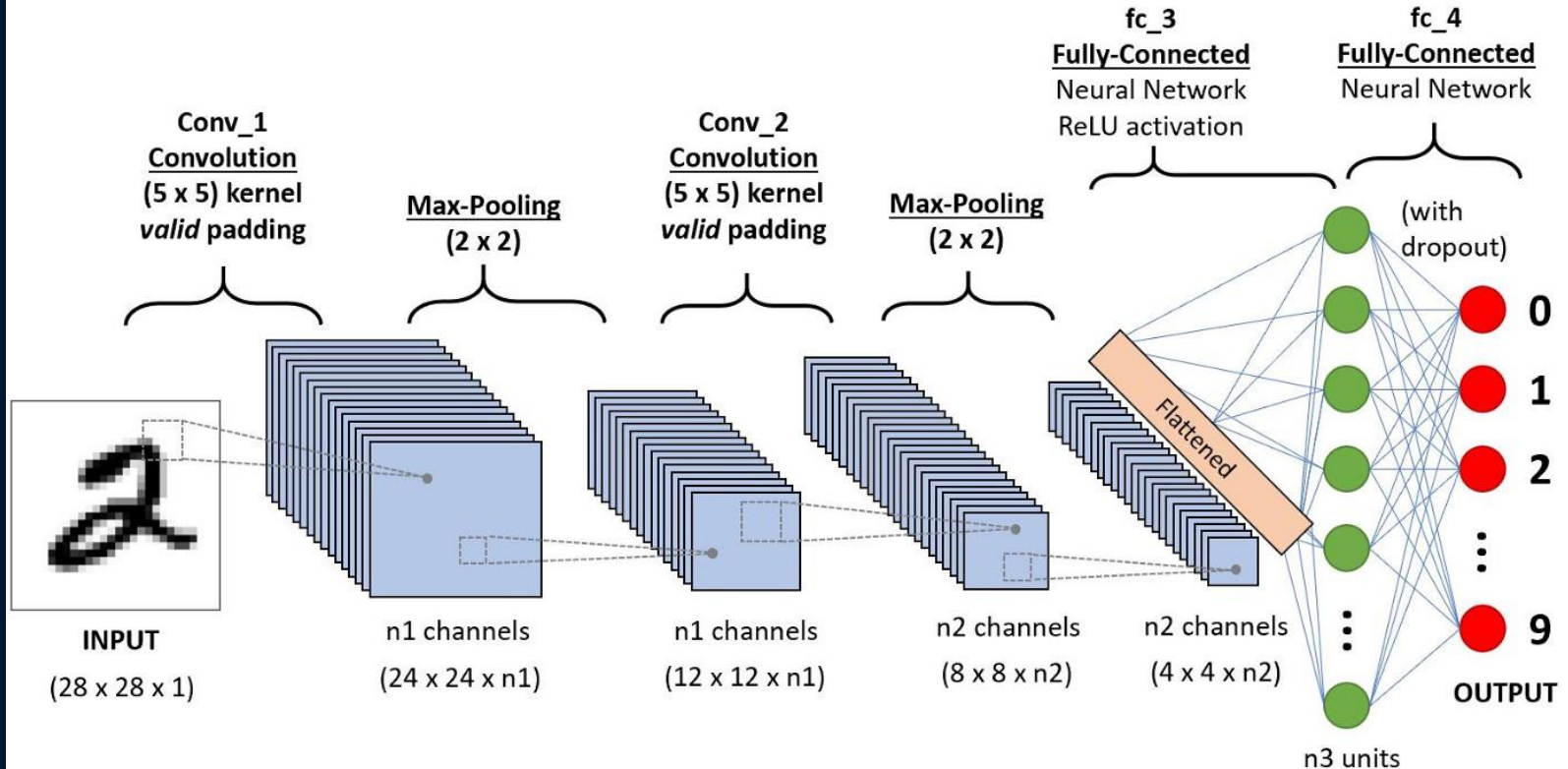
Classification Layer

Flatten final output from convolutional layers and send to regular neural net

- Learn combinations of high-level features
- Possibly non-linear functions
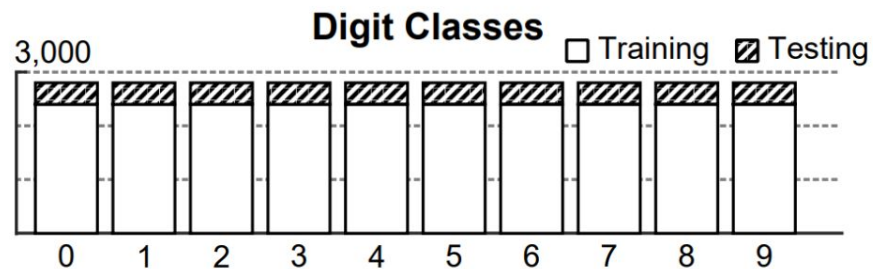- Softmax layer

CNN Sequence

EMNIST Dataset

The ***Extended*** MNIST (EMNIST) dataset is the successor to the popular MNIST dataset
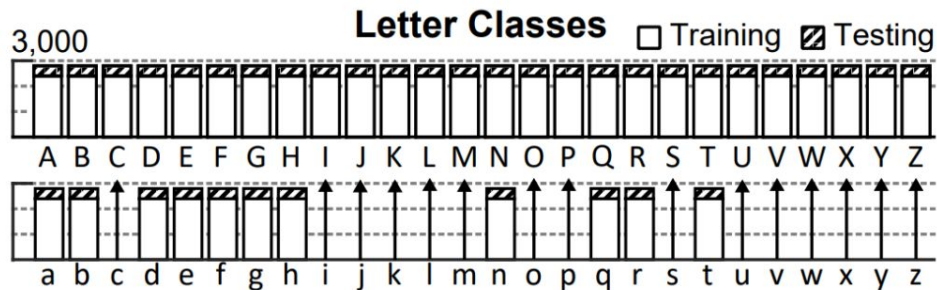
- 131,600 images, 28 x 28 pixels in grayscale
  - 112,800 train / 18,800 test
- Pre-processed
  - Size-normalized, centered, and fixed-size image
- 47 balanced classes
  - Digits: 0-9
  - Uppercase letters: A-Z
  - Lowercase letters: a, b, d, e, f, g, h, n, q, r, t
- More challenging than MNIST

EMNIST Dataset

EMNIST Balanced Dataset

Digit Classes — Training, Testing
3,000
0 1 2 3 4 5 6 7 8 9

47 Classes, 131,600 Samples

Letter Classes — Training, Testing
3,000
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
a b c d e f g h i j k l m n o p q r s t u v w x y z

EMNIST Dataset

Baseline Results

In 2017, Gregory Cohen created an instructive baseline with his paper *EMNIST: An extension of MNIST to handwritten letters*

- OPIUM classifier
  - Accuracy 78.02%
  - 10,000 hidden nodes

- Linear classifier (not shown)
  - Accuracy 50.93%

Classifier Performance on the EMNIST Balanced Datasets
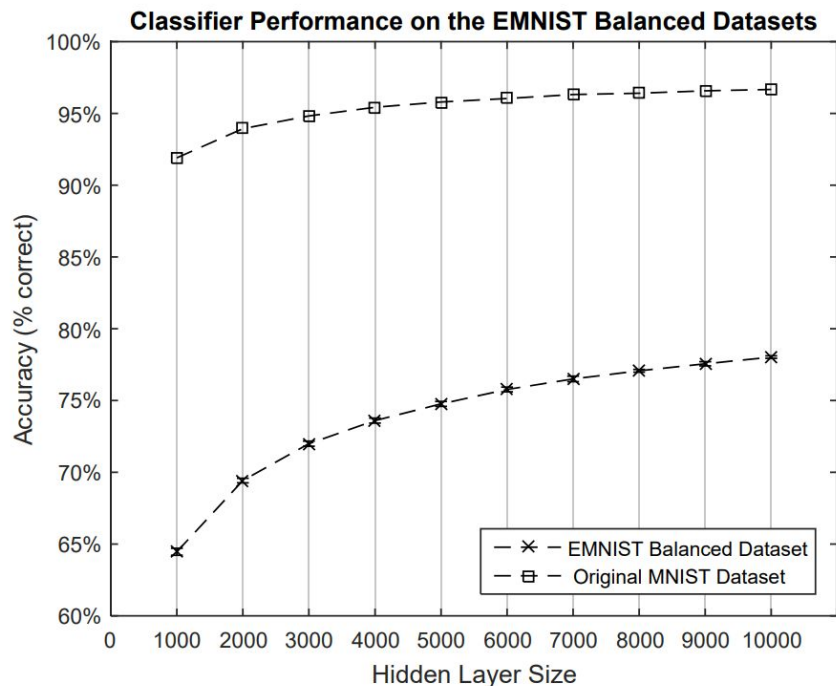
Our Results

Using a very basic convolutional neural net, we were able to exceed the Cohen et al results from 2017

- CNN classifier
  - Accuracy 87.22%
  - 3 convolutional layers
  - ReLu activation
  - 2 max pooling layers

Our Results

```
layers = [
    imageInputLayer([28,28,1])

    convolution2dLayer(3,8,'Padding','same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer(2,"Stride",2)

    batchNormalizationLayer
    convolution2dLayer(3,16,'Padding','same')
    reluLayer

    maxPooling2dLayer(2,"Stride",2)

    convolution2dLayer(3,16,'Padding','same')
    batchNormalizationLayer
    reluLayer

    fullyConnectedLayer(47)
    softmaxLayer
    classificationLayer
    ];
```

Results

Validation accuracy:	87.22%
Training finished:	Max epochs completed

Training Time

Start time:	26-Apr-2022 15:36:53
Elapsed time:	85 min 11 sec

Training Cycle

Epoch:	20 of 20
Iteration:	17620 of 17620
Iterations per epoch:	881
Maximum iterations:	17620

Validation

Frequency:	50 iterations

Other Information

Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	0.01

Demonstration

Need volunteers to suggest characters

Characters will be drawn in MS Paint

Use the trained net to make live predictions

Can we get 8 out of 9 correct?   (≈ 88%)

References

Cohen, G., Afshar, S., Tapson, J., Schaik, A. (2017). EMNIST: an extension of MNIST to handwritten letters. *2017 International Joint Conference on Neural Networks*. 2921-2926.

https://www.nist.gov/itl/products-and-services/emnist-dataset

https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53