# MNIST Digit Classification Using LSTM RNNs: Representing Images as Time Series Using Hilbert Curves

Beckett Hyde - Spring 2023 - Blake Hamilton

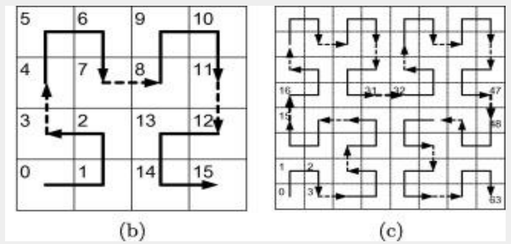
## Problem Space

Handwritten digit labeling and image labeling tasks are commonly solved using Machine Learning techniques, particularly Convolutional Neural Networks (CNNs). We attempted to recognize digits using a different method: transforming 2D images into time series data using Hilbert curves, then analyzing them using Long Short-Term Memory Recurrent Neural Networks (LSTM RNNs).

## Approach

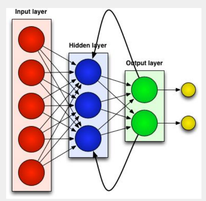
### Hilbert Curves

Hilbert space-filling curves are a way to map every point in $N$x$N$ to a point in $M$ where $|M|=|N|^2$ that may better preserve relative location information. Relative distance is inherently two-dimensional in images; by winding through the space, close pixels both horizontally and vertically stay close in the image.




### Recurrent Neural Networks

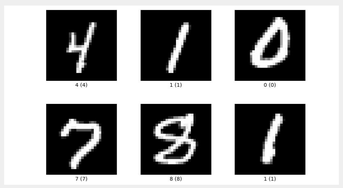
We treated the stream of dimension-reduced digits as a standard time series and fed it into the network the standard way, training them on a set of one-hot-encoded solution values (0,1,2,etc).




### Long Short Term Memory (LSTM)

We treated the stream of dimension-reduced digits as a standard time series and fed it into the network the standard way training them on a set of one-hot-encoded solution values (0,1,2,etc).

## Dataset



We used the Keras MNIST Digits Classification Dataset. This dataset includes 70,000 images. Each entry has 784 features, as each digit image is a square with 28 pixel edge lengths. For some trials, we scaled the images to 14-pixel edges. Each feature has a number range between 0 to 255, with the values corresponding to the pixel darkness (as the images are greyscale). The labels are the standard base ten digits. This dataset is a popular handwritten digit set, and it has been used in many applications. 10,000 images were set aside as a testing set.

## Results

### Baseline Results

The baseline tests were completed for the project using a Convolutional Neural Network (CNN). Using this conventional approach, an accuracy of 0.97 was achieved. This is as expected, and will serve as the standard to compare our alternative method Recurrent Neural Network against.

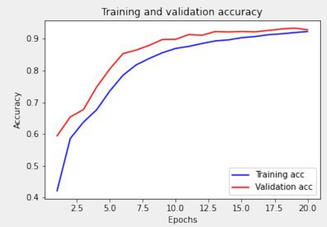
### Discussion

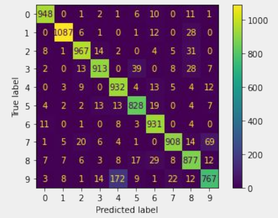
An accuracy of 0.97 was achieved for both the baseline approach as well as the Naive time series approach, while a slightly lower 0.92 accuracy was achieved using Hilbert Curves. We were not able to conclusively find any reason why this was the case.

It is possible that the greater smoothness of the curve derived from th enaive approach made it easier for the network to learn the dynamics. It is also possible that exponential information loss hit the Hilbert map harder than the naive one.

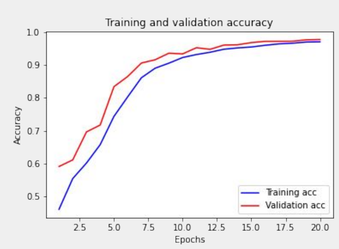
### RNN Results

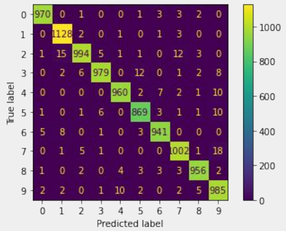
After processing the image data using Hilbert curves and training using a Long Short Term Memory model, we achieved a 0.92 accuracy on the test data.. Our Neural Network consisted of an LSTM layer, a dropout layer, Relu layer, and a SoftMax layer. We used 20 epochs and a batch size of 32.




After getting an accuracy we were happy with using Hilbert Curves, we investigated the difference between using Hilbert curves and taking a naive approach, simply flattening the images into one-dimensional arrays row by row. We kept all other pieces of the process (neural network, image compression) the same for consistency, Oddly enough, this naive approach worked better than the version using Hilbert curves. It resulted in an accuracy of 0.97.




## Discussion

A big discovery we have made is that, in determining the model's accuracy, we needed to only analyze the predicted label near the end of each sample in the time series.A big discovery we have made is that, in determining the model's accuracy, we needed to only analyze the predicted label near the end of each sample in the time series. The RNN was trying to predict digits based on only a fraction of the actual image data because of internal memory limits. To solve this problem, we had to compress the images from 28 x 28 data points to 14 x 14 data points. Overall, the investigation was very successful, as we were able to get a Recurrent Neural Network to accurately predict handwritten digits using image data transformed into time series data.