

# SOM Visualization of Handwritten Digits

Bahadir Bagci

November 2025

## 1 Introduction

This mini-report focuses on applying and interpreting a Self-Organizing Map (SOM) to the MNIST-784 dataset. In this dataset, each image represents a single digit between 0 and 9, stored as 784 numerical values corresponding to the  $28 \times 28$  pixel grid. The goal of this study is not to build a classifier but to explore how the SOM organizes high-dimensional visual data into a two-dimensional map. As a reminder, the SOM—an unsupervised learning method—clusters samples based on similarity, allowing visually similar digits to appear near each other on the grid.

We focus on showing the network structure and interpreting the resulting Neighbor Weight Distances and Sample Hits Map.

## 2 SOM Structure

The SOM architecture used in this experiment consisted of 784 input neurons (one per pixel) and a  $20 \times 20$  hexagonal output grid (400 neurons in total). Each neuron in the output layer learns a weight vector representing a group of similar digit patterns. The structure can be seen in Figure 1.

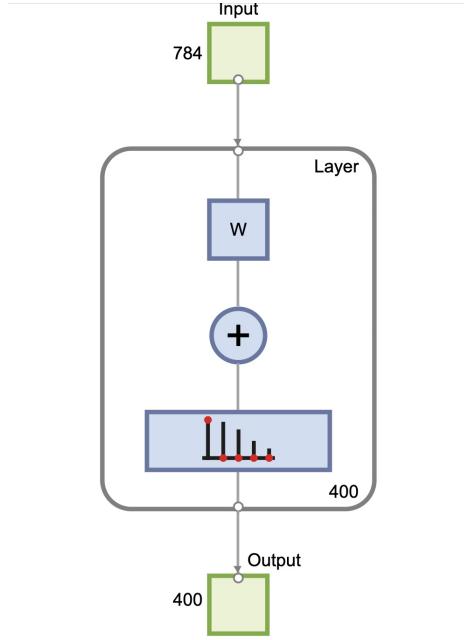


Figure 1: SOM Architecture

### 3 Training and Visualization

Figure 2 shows the U-Matrix (neighbor weight distances). In this figure, the yellow borders highlight areas where neighboring neurons differ the most - which is desirable - while the darker red zones show regions where the neurons have learned similar features. The map forms several smooth regions separated by lighter borders, meaning that the SOM grouped visually similar digits close to each other. Digits such as 1 and 7, or 0 and 8, which share comparable shapes, are likely mapped to neighboring areas.

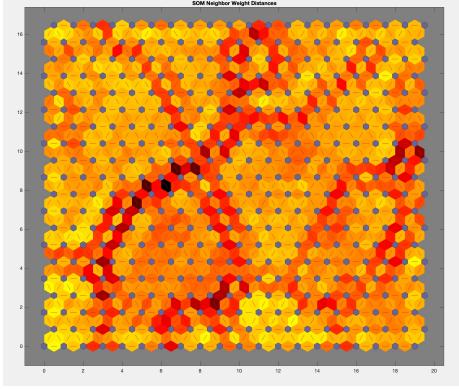


Figure 2: Neighbor Distances

Figure 3 shows the Sample Hits map. The numbers inside each neuron show how many samples were mapped to it. Some areas - especially in the lower-left corner - contain more samples than others. These regions indicate that many similar digit images were grouped into neighboring neurons. This often happens with digits such as 5, which have a consistent writing style. In contrast, regions with fewer samples may correspond to digits that vary more in shape.

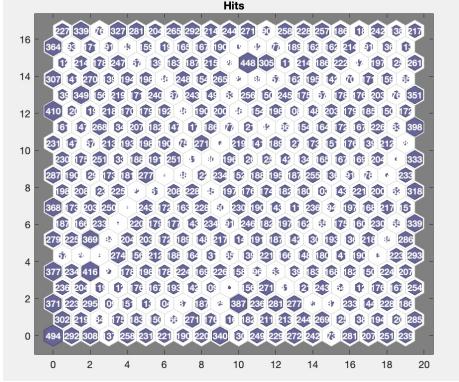


Figure 3: Sample Hits Map

## 4 Conclusion

In this short experiment, the Self-Organizing Map gave a general view of how the handwritten digits are organized on a two-dimensional grid. The map shows several regions that likely group together digits with similar shapes. Since the model is unsupervised, it does not show which digits appear in each region, so the interpretation is based only on the visible patterns in the map.