Dealing with Scale and Rotational Invariance of Digit Identification

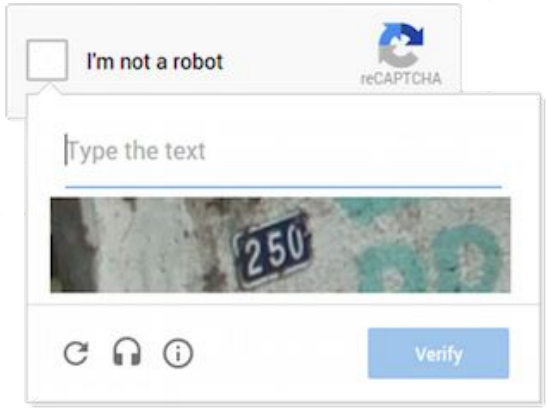
Scott Chow and Robert Cyprus

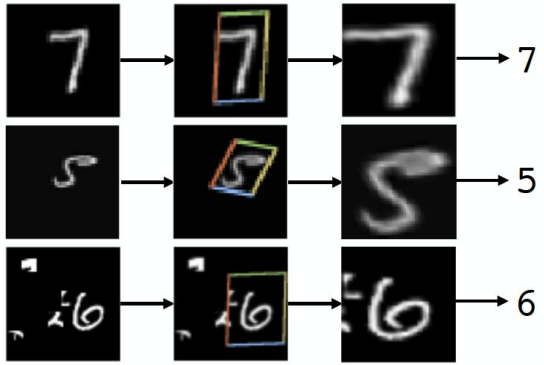
Problem

Task: Identify digits and characters from images

Issue:

- Words can be rotated or distorted in images.
- Standard Convolutional Neural Nets (CNN) do not deal with these distorted characters.





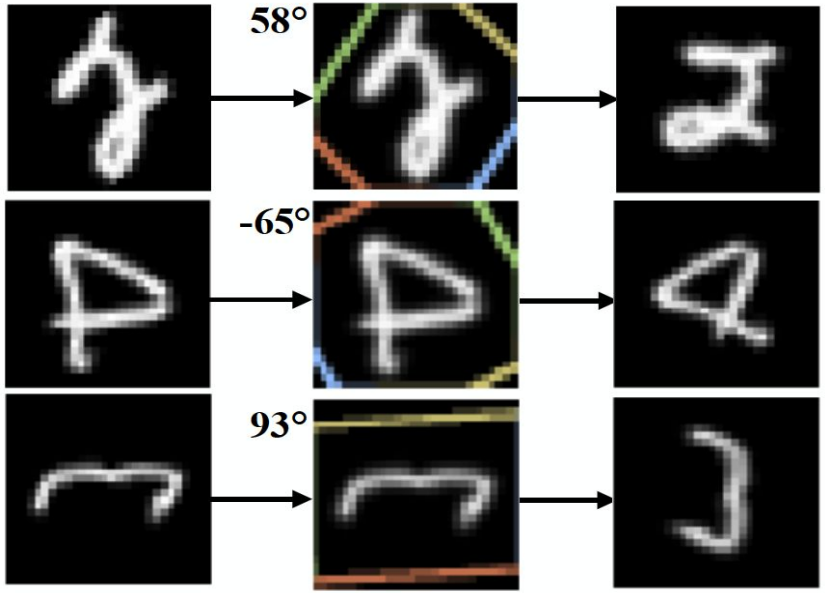
Solution

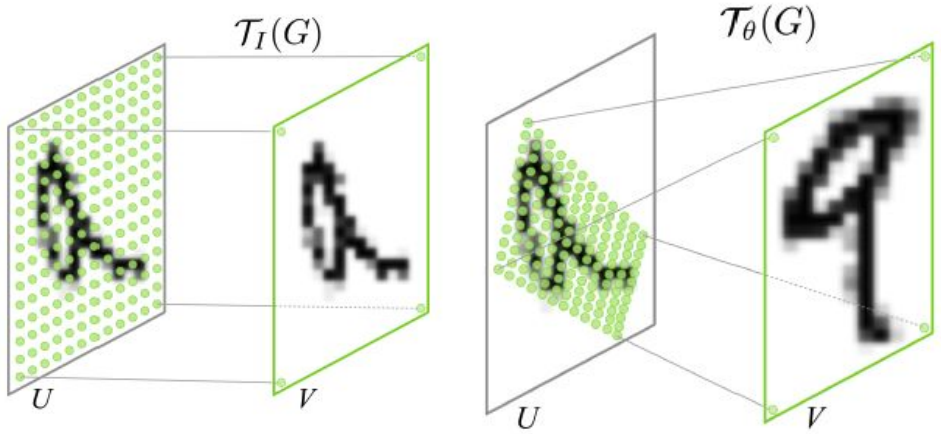
- Investigate methods for making CNNs scale- and rotational-invariant
  - i.e. Output the same result for an image even if it has been distorted
- Potential Approaches
  - Spatial Transformer Networks (STNs) - Jaderburg et al.
  - Tiled Convolutional Neural Networks (Tiled-CNN) - Le et al.
  - Moment Based Features - Singh et al.

Primary Text of Interest

"Spatial Transformer Networks" by Google DeepMind, published Feb 2016

- By: Max Jaderberg et al.





Our Project Plan

- Review Google DeepMind's code from "Spatial Transformer Networks"
- Implement an STN in series with a CNN for digit identification from the MNIST dataset, as demonstrated by DeepMind
- Compare results with standard CNN only model

Stretch Goals

- Investigate other methods of scale/rotational invariance:
  - Tiled-CNN
  - Moment-based Features