

# Read Digits from Natural Images using Convolutional Neural Network

---

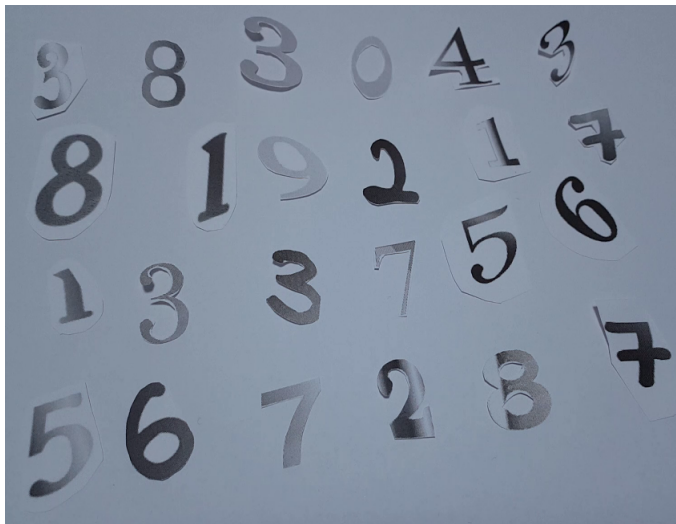
Ramesh Kumar  
Roberto Cai Wu  
January 8, 2018



**Figure:** Digit Recognition [?]

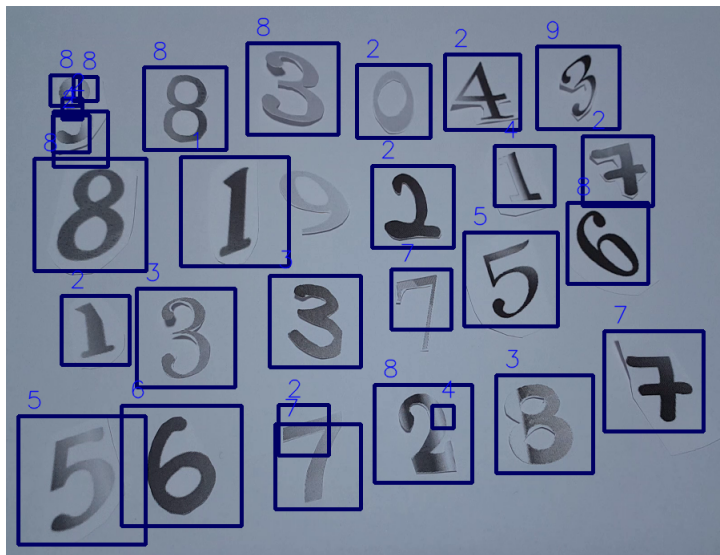
- Task is to read digits from natural images
- We use the MNIST dataset [1], which consists of hand-written digits
- Convolutional neural networks(CNN) for classification of digits
- Computer Vision techniques for detection of digits

## Challenges(1)



**Figure:** Image contains digits with shadows and different fonts

## Challenges(2)



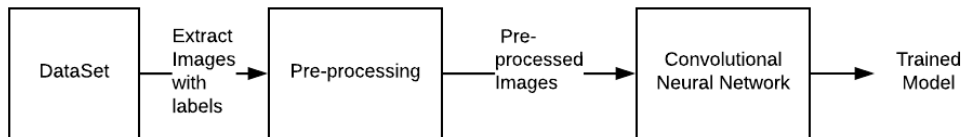
**Figure:** Image with digits and bounding boxes

# Assumptions

- Images contain only digits.
- Background is a solid color and does not change
- Numbers should be distanced enough so that bounding boxes do not overlap.
- Digits are visible to camera, orientations may be varied till  $45^\circ$  or  $60^\circ$

- Camera; we use mobile camera
- Solid background with suitable font color
- Suitable lighting conditions

- Load and Interpret DataSet
- Pre-processing(dataset as well as live camera)
- Convolutional Neural Network
- Post-processing(during camera only)
- Testing and evaluation



**Figure:** Block Diagram of System using Dataset





**Figure:** Block Diagram of System using Camera

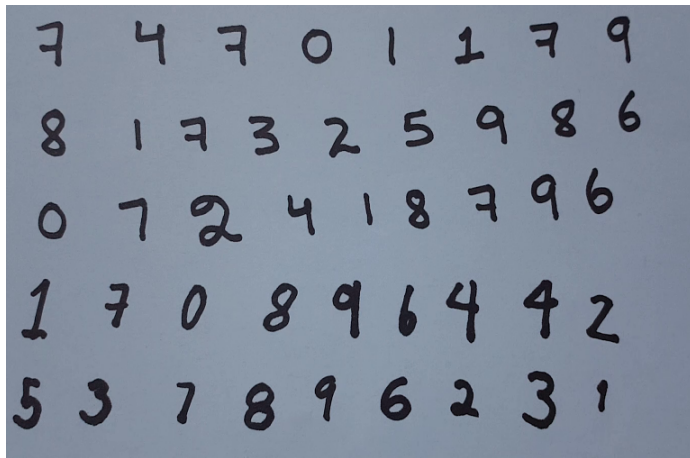
- 10 classes, 1 for each digit
- Digit 1 has label 1, 9 has label 9, and 0 has label 10
- 60,000 digits for training, 10,000 digits for testing
- Digits are arranged in different positions with solid background



**Figure:** Example images from MNIST dataset [1]

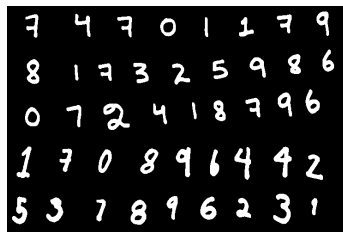
- Resize image to 1080x720 pixels
- Convert to gray scale
- Apply Gaussian filter
- Use a binary thresholding
- Canny edge detection
- Find contours
- Draw bounding box around contours

## Pre-processing from camera(2)

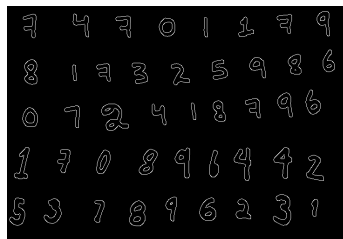


**Figure:** Image taken from live camera

## Pre-processing from camera (3)

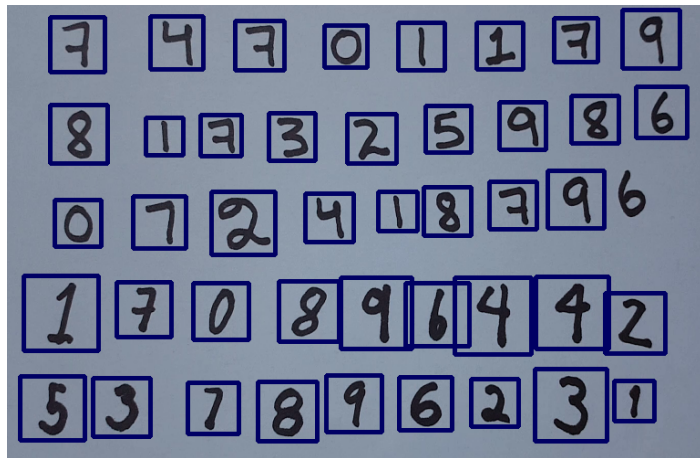


(a) Binary threshold image



(b) Canny edge detector

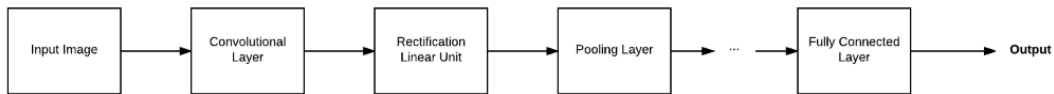
## Pre-processing from camera (4)



**Figure:** Bounding boxes drawn over original image

# Convolutional Neural Network(CNN)

- State-of-the-art shows CNN performs better as compare to other approaches[2]
- Extracts features from the images and classify them
- Three type of layers
  - Convolutional: Extract low-level and high-level features
  - Pooling: Reduce amount of parameters and computations
  - Fully Connected: Neurons are fully connected



**Figure:** Basic Architecture of CNN



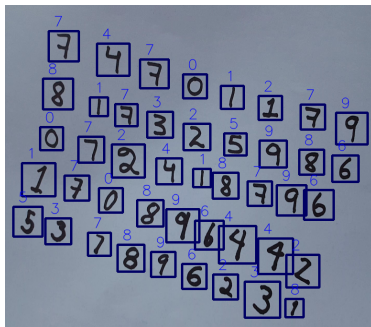
# Our CNN Model



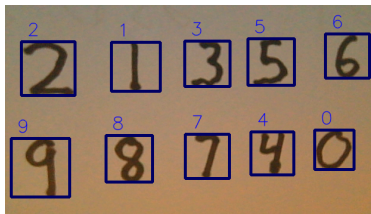
**Figure:** Model used to train MNIST dataset

- Print/write numbers on a sheet of paper (different size, font, color, and orientation)
- Test the images of digits from live camera under different conditions (light and perspective)
- Use test set to compute accuracy of model, gives us 99.13%

## Testing & Evaluation(2)



(a) Detection of digits with different orientations



(b) Detection of digits



Yann LeCun and Corinna Cortes.

**MNIST handwritten digit database.**

2010.



Pierre Sermanet, Soumith Chintala, and Yann LeCun.

**Convolutional neural networks applied to house numbers digit classification.**

In *ICPR*, pages 3288–3291. IEEE Computer Society, 2012.



Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y Ng.

**Reading digits in natural images with unsupervised feature learning.**

In *NIPS workshop on deep learning and unsupervised feature learning*, volume 2011, page 5, 2011.



Ian J Goodfellow, Yaroslav Bulatov, Julian Ibarz, Sacha Arnoud, and Vinay Shet.  
**Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks.**

