Image Classification from Coded Aperture

Jocelyn Ornelas Muñoz, Erica M. Rutter, Roummel F. Marcia

Department of Applied Mathematics, UC Merced

Thursday, December 8, 2022

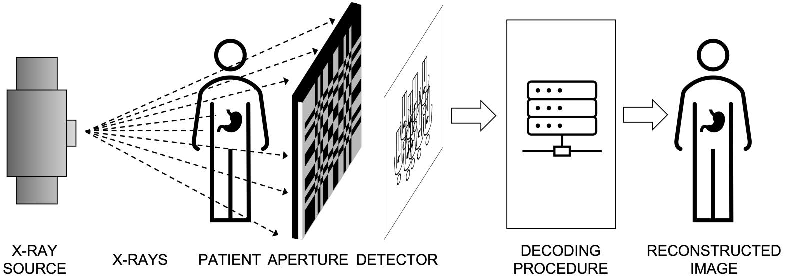
Motivation

In medical imaging, radiography typically uses apertures to modulate the radiation emitted by an x-ray source.

Motivation

In medical imaging, radiography typically uses apertures to modulate the radiation emitted by an x-ray source.

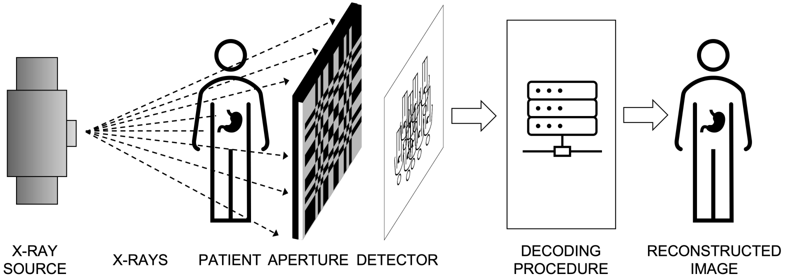
Problem: With more complicated apertures, a decoding procedure is necessary to reconstruct the image.



Motivation

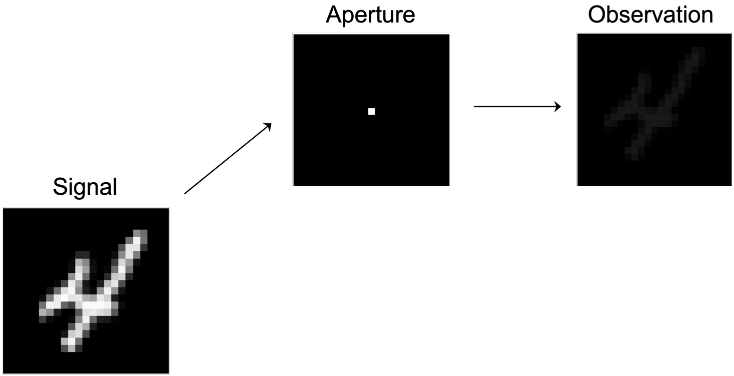
In medical imaging, radiography typically uses apertures to modulate the radiation emitted by an x-ray source.

Problem: With more complicated apertures, a decoding procedure is necessary to reconstruct the image.



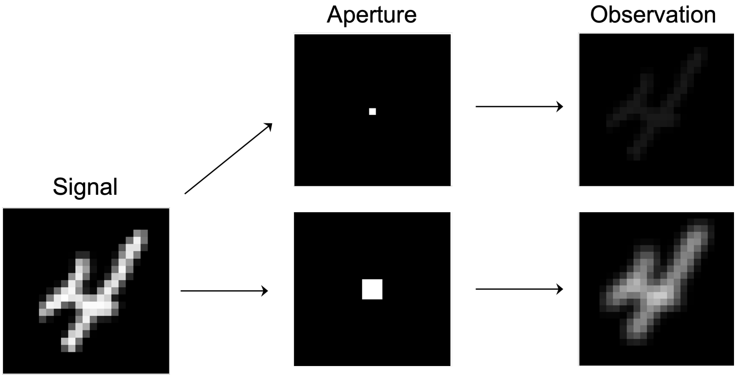
Goal: Classify images from coded observations without reconstructing the image.

Aperture Imaging



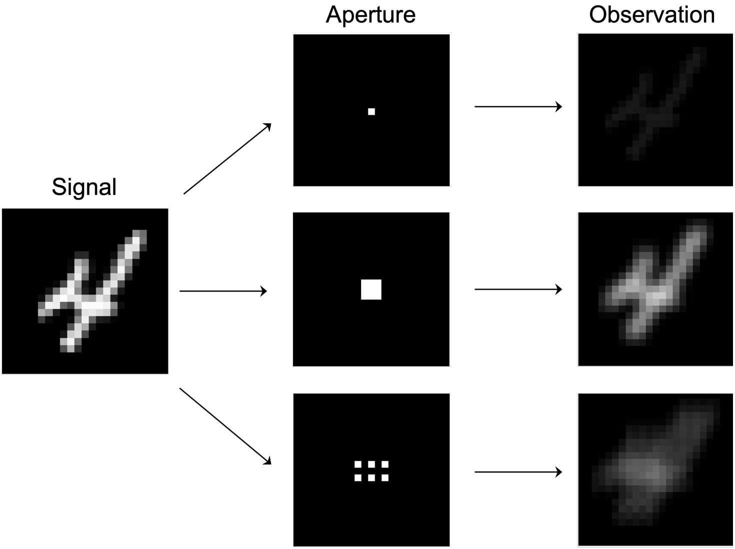
Small pinholes allow little light $\implies$ **faint** observatons

Aperture Imaging



Larger pinholes allow more light, but decrease resolution $\implies$ **blurry** observations

Aperture Imaging



Multiple small pinholes $\implies$ **overlapping** observations

Modified Uniformly Redundant Array (MURA)

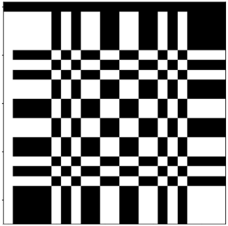
A MURA pattern A consists of specified openings that has a corresponding decoding pattern G[1].

[1] Gottesman and Fenimore (1989)

Modified Uniformly Redundant Array (MURA)

A MURA pattern A consists of specified openings that has a corresponding decoding pattern G[1].

Let p be a prime number and $A = \{A_{ij}\}_{i,j=0}^{p-1}$ be the binary aperture array. Set



$$A_{ij} = \begin{cases} 0 & \text{if } i = 0 \\ 1 & \text{if } j = 0, i \neq 0 \\ 1 & \text{if } C_i C_j = +1 \\ 0 & \text{otherwise} \end{cases}$$

where

$$C_i = \begin{cases} +1 & \text{if } i \text{ is a quadratic residue modulo } p \\ -1 & \text{otherwise} \end{cases}$$

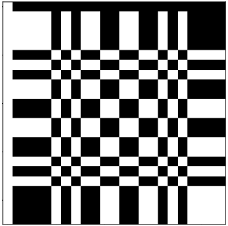
[1]Gottesman and Fenimore (1989)

Modified Uniformly Redundant Array (MURA)

A MURA pattern A consists of specified openings that has a corresponding decoding pattern G[1].

The decoding function G is constructed as follows:

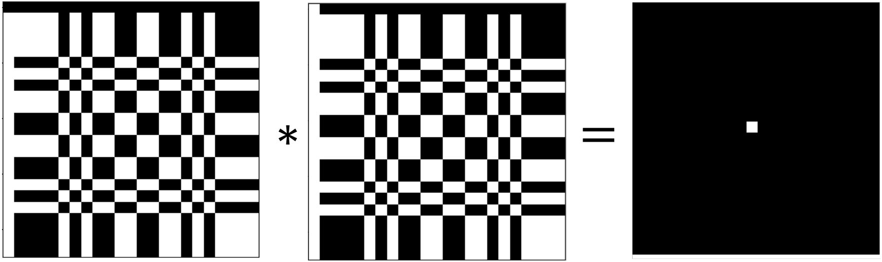
$$G_{ij} = \begin{cases} +1 & \text{if } i + j = 0 \\ +1 & \text{if } A_{ij} = 1, i + j \neq 0 \\ -1 & \text{if } A_{ij} = 1, i + j \neq 0 \end{cases}$$



[1] Gottesman and Fenimore (1989)

Modified Uniformly Redundant Array (MURA)

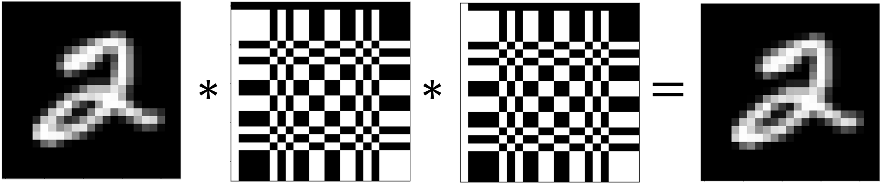
A MURA pattern A consists of specified openings that has a corresponding decoding pattern G[1].



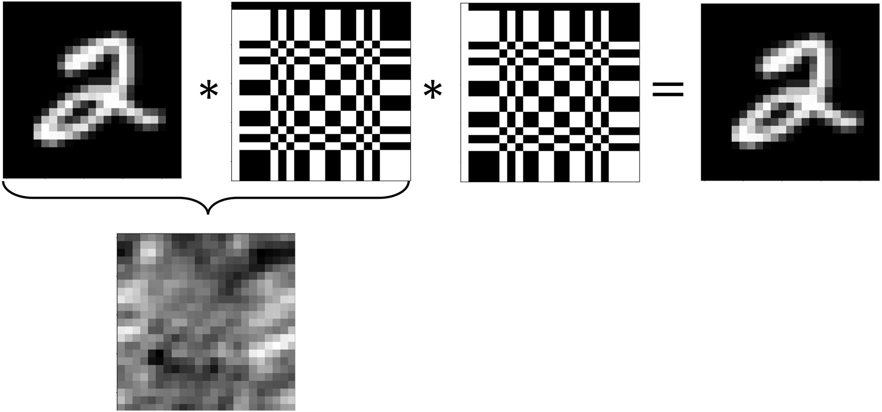
Coded aperture A Decoding pattern G δ

[1] Gottesman and Fenimore (1989)

Modified Uniformly Redundant Array (MURA)



Modified Uniformly Redundant Array (MURA)

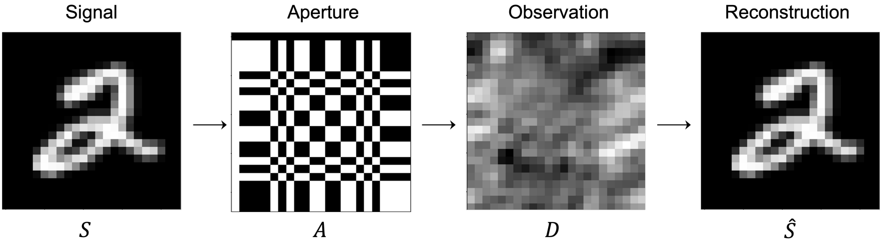


Coded observation

Coded observations appear irrecognizable, but MURAs are 50% open patterns[1] $\Longrightarrow$ decoded observations are much brighter than those from small pinhole cameras.

[1] Gottesman and Fenimore (1989)

MURA aperture imaging



The observation D is given by

$$D = S * A + B$$

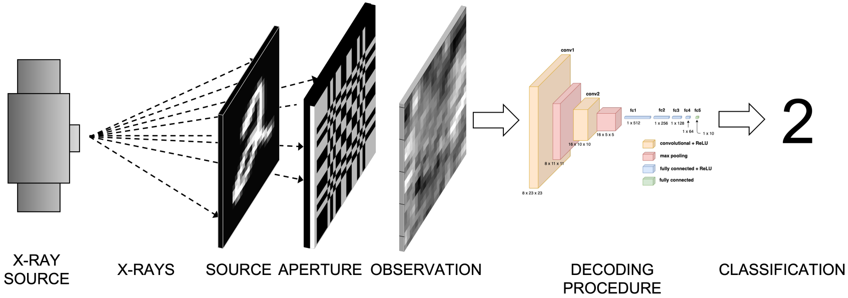
where B is background noise. The MURA reconstruction is given by

$$\hat{S} = D * G$$

where G is the decoding pattern.

Proposed Method

Goal: Classify handwritten digits from coded observations using a convolutional neural network (CNN) **without reconstructing the image**.



Training pipeline

Dataset: MNIST Handwritten Digits

- 28×28 pixels (grayscale)
- $60,000$ training images
- $10,000$ testing images

Training pipeline

Dataset: MNIST Handwritten Digits

- 28×28 pixels (grayscale)
- $60,000$ training images
- $10,000$ testing images

Approach:

1. Resize MNIST digits to 23×23 pixels

Training pipeline

Dataset: MNIST Handwritten Digits

- 28×28 pixels (grayscale)
- $60,000$ training images
- $10,000$ testing images

Approach:

1. Resize MNIST digits to 23×23 pixels
2. Generate aperture of size 23×23 (in C++)

Training pipeline

Dataset: MNIST Handwritten Digits

- 28×28 pixels (grayscale)
- $60,000$ training images
- $10,000$ testing images

Approach:

1. Resize MNIST digits to 23×23 pixels
2. Generate aperture of size 23×23 (in C++)
3. Convolve images with aperture using series of $1D$ Fast Fourier Transforms (FFT) and $1D$ inverse FFT to get coded observations (in C++)

Training pipeline

Dataset: MNIST Handwritten Digits

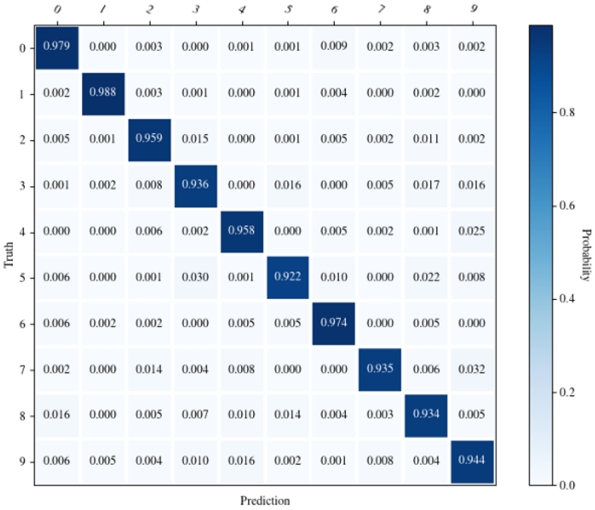
- 28×28 pixels (grayscale)
- $60,000$ training images
- $10,000$ testing images

Approach:

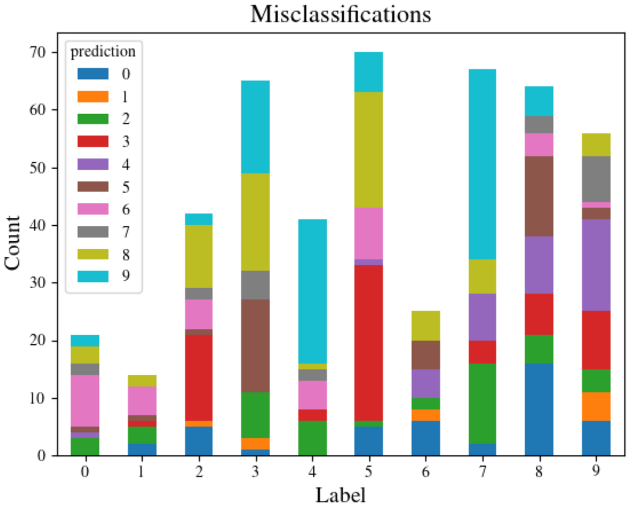
1. Resize MNIST digits to 23×23 pixels
2. Generate aperture of size 23×23 (in C++)
3. Convolve images with aperture using series of $1D$ Fast Fourier Transforms (FFT) and $1D$ inverse FFT to get coded observations (in C++)
4. Train a 2-layer CNN

Results

CNN performance on coded observations test set.
Accuracy: **95.35**%



Results



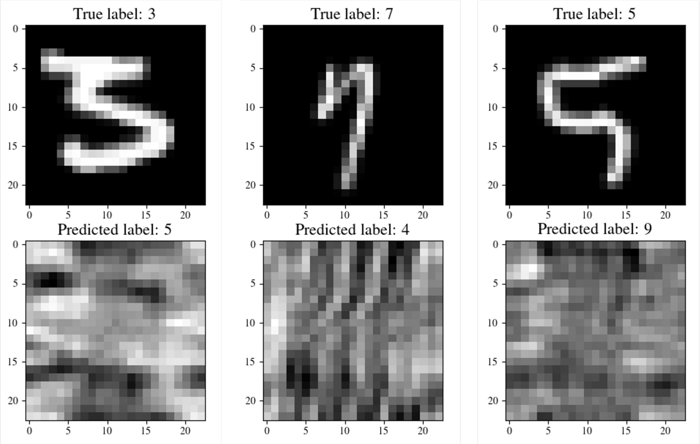
Digits 3, 5 and 7 were the most common misclassifications.

Results

Most common misclassifications:

True image

Observation



Gracias!