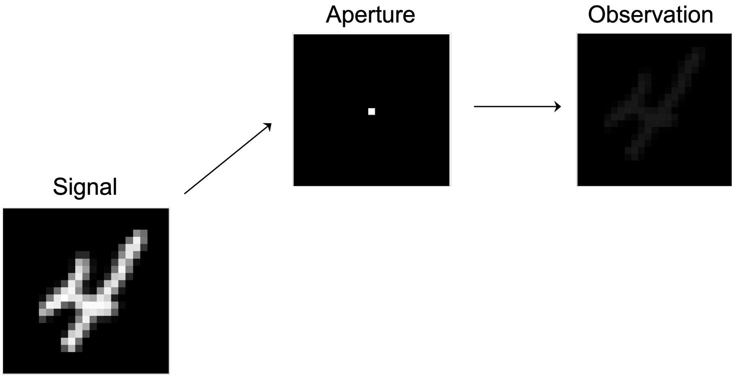
Decoding the Hidden: Direct Image Classification using Coded Aperture Imaging

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

Department of Applied Mathematics, UC Merced
Imaging and Sensing Seminar

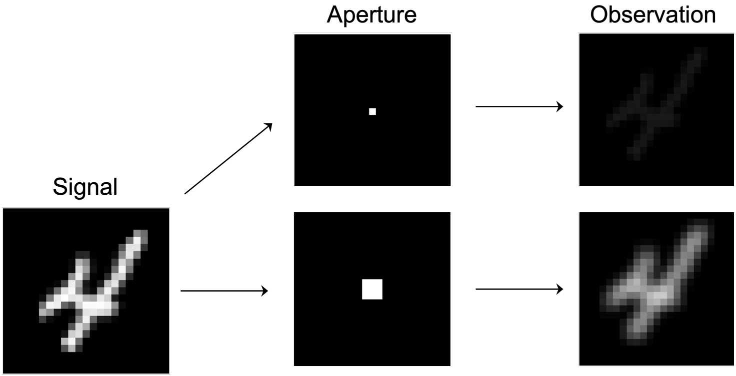
Tuesday, October 10, 2023

Aperture Imaging



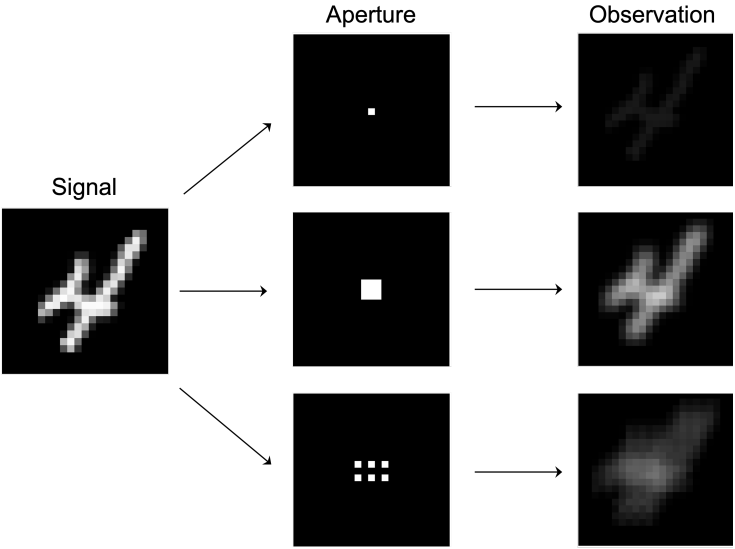
Small pinholes allow little light $\Longrightarrow$ **faint** observatons

Aperture Imaging



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

Aperture Imaging



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

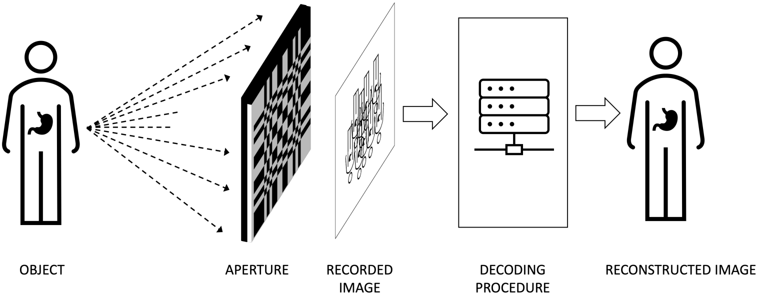
Motivation

In medical imaging, radiography typically uses complex apertures to modulate the radiation emitted by a high energy source and produce high-resolution images.

Motivation

In medical imaging, radiography typically uses complex apertures to modulate the radiation emitted by a high energy source and produce high-resolution images.

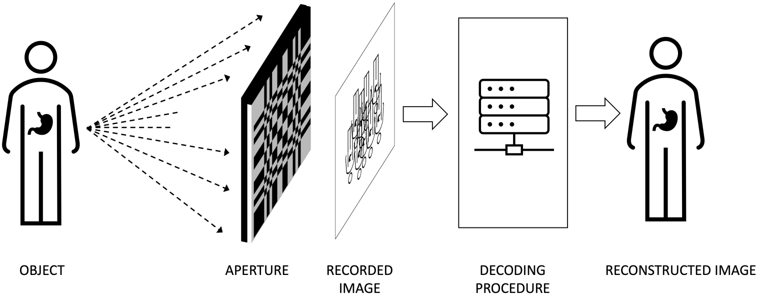
Problem: Complicated apertures require a decoding procedure to reconstruct the image.



Motivation

In medical imaging, radiography typically uses complex apertures to modulate the radiation emitted by a high energy source and produce high-resolution images.

Problem: Complicated apertures require a decoding procedure to reconstruct the image.



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

Modified Uniformly Redundant Array (MURA)

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

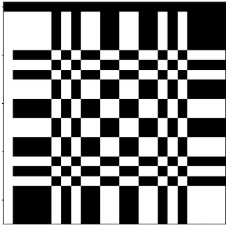
[1] Gottesman and Fenimore (1989)

Modified Uniformly Redundant Array (MURA)

A MURA pattern $\mathbf{A}$ consists of specified openings that has a corresponding decoding pattern $\mathbf{G}$[1].

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

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



where

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

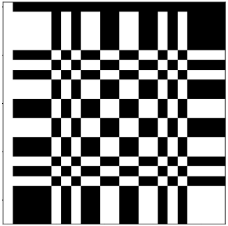
[1] Gottesman and Fenimore (1989)

Modified Uniformly Redundant Array (MURA)

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

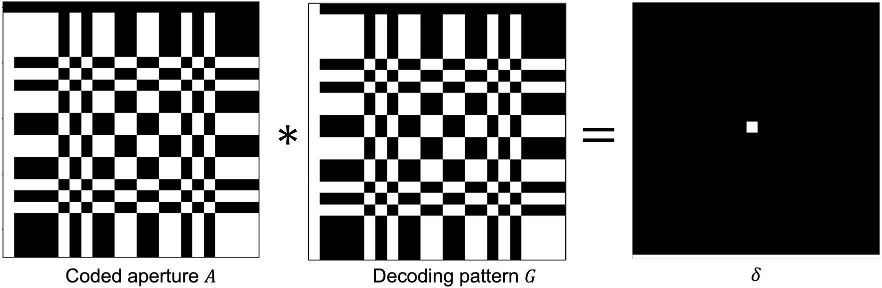
The decoding function **G** is constructed as follows:

$$\mathbf{G}_{ij} = \begin{cases} +1 & \text{if } i+j=0 \\ +1 & \text{if } \mathbf{A}_{ij}=1, i+j \neq 0 \\ -1 & \text{if } \mathbf{A}_{ij}=1, i+j \neq 0 \end{cases}$$



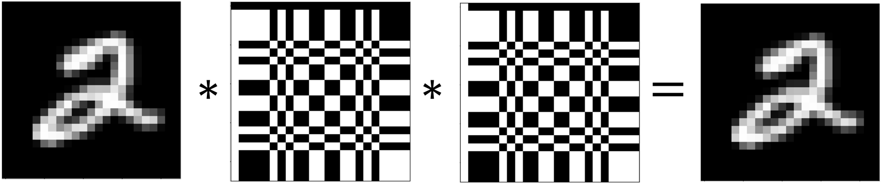
[1]Gottesman and Fenimore (1989)

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

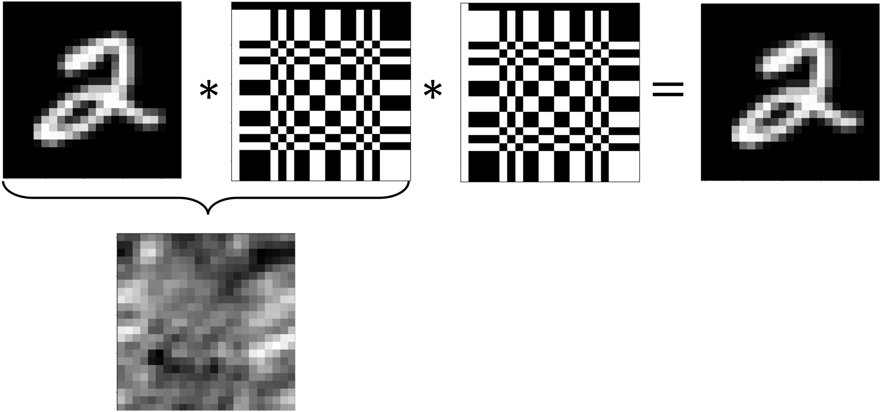


[1]Gottesman and Fenimore (1989)

Modified Uniformly Redundant Array (MURA)



Modified Uniformly Redundant Array (MURA)

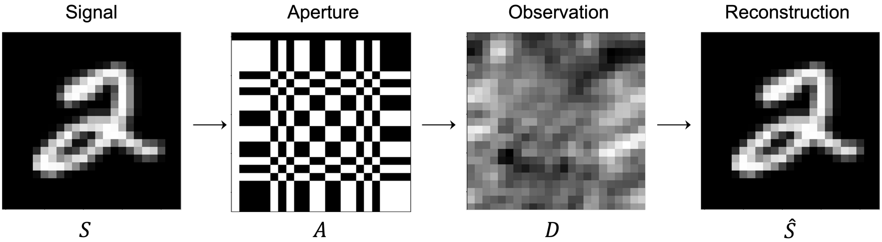


Coded observation

Coded observations appear irrecognizable, but MURAs are 50% open patterns[1] $\implies$ 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

$$\mathbf{D} = \mathbf{S} * \mathbf{A} + \mathbf{B}$$

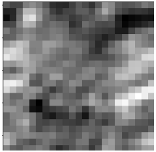
where **B** is background noise. The MURA reconstruction is given by

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

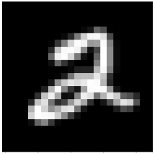
where **G** is the decoding pattern.

Proposed Method

Goal: Given coded observation data $\mathbf{D}$, **classify** the original unknown image $\mathbf{S}$ using two different methods.



Method I: Classify raw coded data $\mathbf{D}$ using a convolutional neural network (CNN).



Method II: Classify reconstructions, $\hat{\mathbf{S}} = \mathbf{D} * \mathbf{G}$ (Used for comparison with **Method I**).

Why classify coded observations?

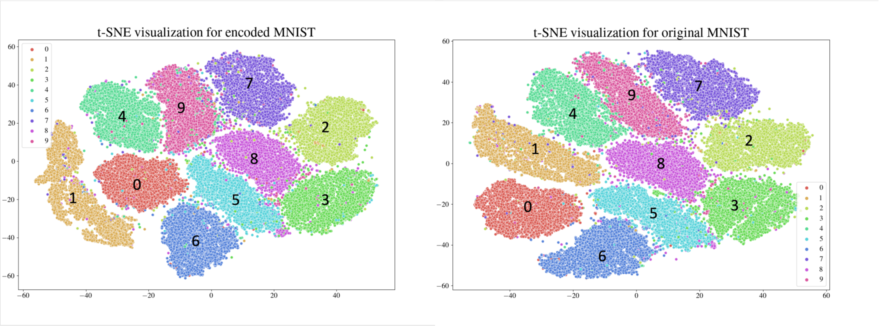


Figure 1: t-SNE visualizations for encoded MNIST (left) and the original MNIST (right) dataset.

Experiment Set Up

Datasets:

MNIST Handwritten Digits

- 28×28 pixels (grayscale)
- $70,000$ total images
- 80% training, 10% validation, 10% testing

CIFAR10

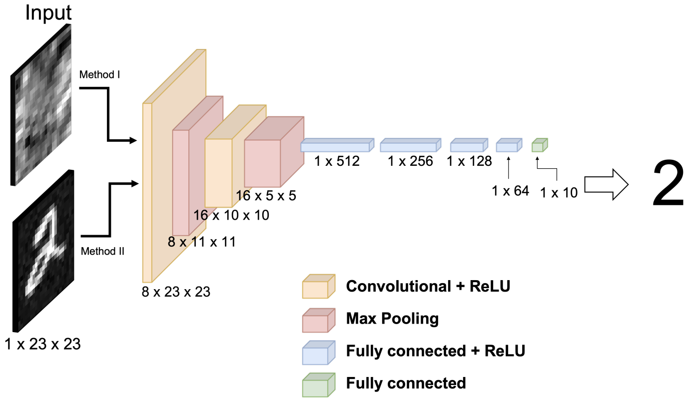
- 32×32 pixels (color images converted to grayscale)
- 10 classes: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck
- $60,000$ total images
- 80% training, 10% validation, 10% testing

Experiments

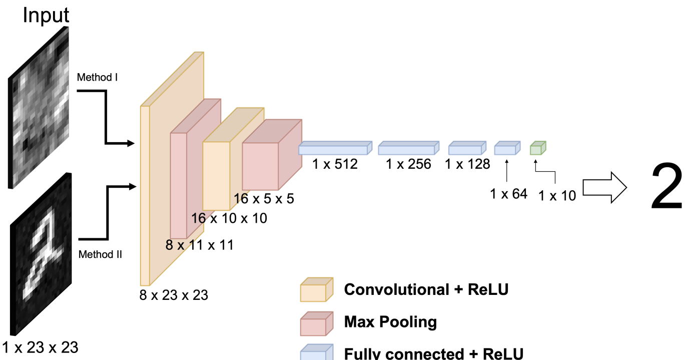
Experiment 1: Classify original images **S** for baseline comparison

Experiments

Experiment 1: Classify original images **S** for baseline comparison
Experiment 2: Classify encoded images **D** (Method I)



Experiment 1: Classify original images $\mathbf{S}$ for baseline comparison
Experiment 2: Classify encoded images $\mathbf{D}$ (Method I)
Experiment 3: Classify reconstructed images $\hat{\mathbf{S}}$ (Method II)



Results: MNIST Dataset

	Testing Data							
	Noiseless		10dB		5dB		1dB	
Training data	I	II	I	II	I	II	I	II
Noiseless	97.53	99.29	96.53	98.90	93.81	97.14	82.73	88.80
10dB	97.94	99.09	97.61	99.03	95.93	98.60	88.17	95.77
5dB	97.77	99.09	97.44	99.14	96.44	98.93	91.91	97.17
1dB	96.70	98.80	96.54	98.84	95.91	98.63	93.86	97.94
Original	-	96.36	-	86.86	-	64.04	-	36.90

Table 1: MNIST dataset classification accuracy using encoded data (Method I) and reconstructed data (Method II). Rows indicate the training noise level, and columns indicate the testing noise level.

Results: CIFAR10 Dataset

	Testing Data							
	Noiseless		10dB		5dB		1dB	
Training data	I	II	I	II	I	II	I	II
Noiseless	40.30	55.78	27.72	36.18	15.38	17.48	**11.93**	11.65
10dB	37.97	50.25	36.55	48.78	33.33	38.85	**26.37**	21.77
5dB	36.33	44.00	35.97	43.25	34.22	40.98	30.65	35.52
1dB	33.25	40.38	33.07	40.17	33.05	38.67	31.75	34.87
Original	-	54.43	-	34.75	-	17.75	-	11.90

Table 2: CIFAR10 dataset classification accuracy using encoded data (Method I) and reconstructed data (Method II). Rows indicate the training noise level, and columns indicate the testing noise level.

Why are the CIFAR10 results not great?

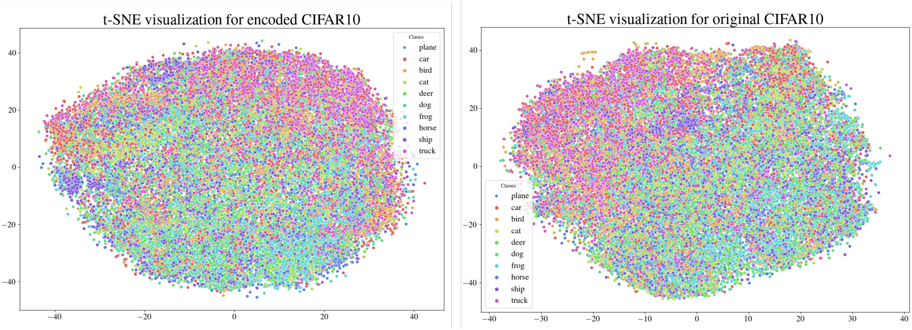


Figure 2: t-SNE visualization of CIFAR data.

The t-SNE visualization did not exhibit distinct separability. Thus, the classification problem may pose a greater challenge.

Summary

- Method II often yields better classification results compared Method I, but it is not directly applicable when the decoder **G** is unknown.
- Method I provides reasonable classification results **without reconstruction** or knowledge of the encoding and decoding arrays.
- Dimensionality reduction allows us to **infer the difficulty** of classification
- More **sophisticated models** may be better suited to "difficult" data

Gracias!