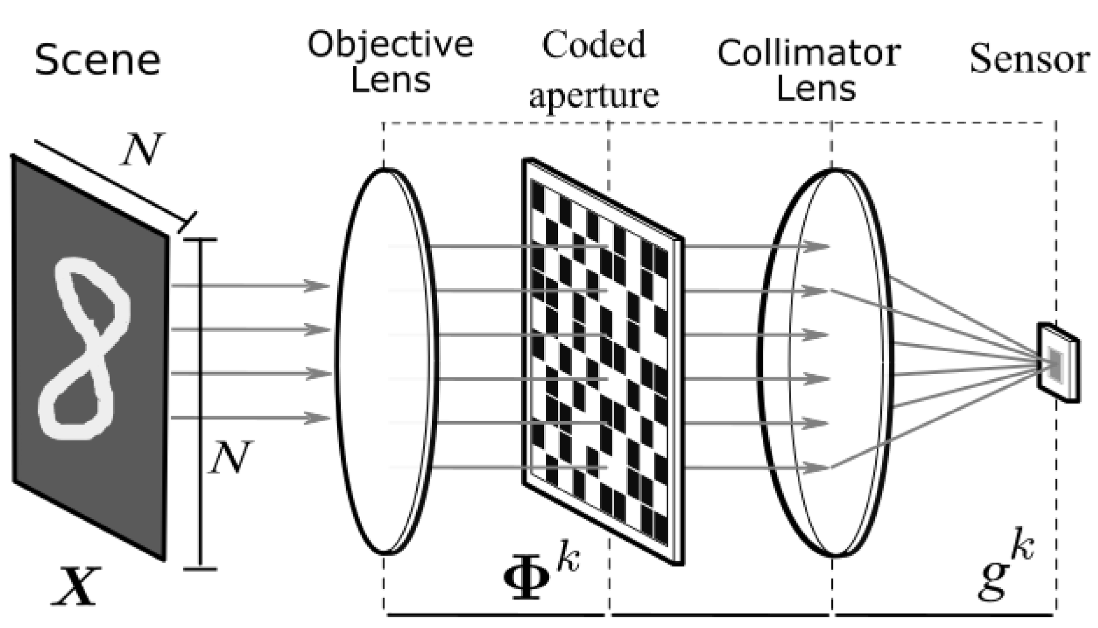
Compressive Classification via Deep Learning using Single-pixel Measurements

Jorge Bacca, Nelson Diaz & Henry Arguello
Universidad Industrial de Santander
jorge.bacca1@correo.uis.edu.co

High-Dimensional Signal Processing Research Group
www.hdspgroup.com

Single Pixel Camera (SPC)



Linear sensing model:

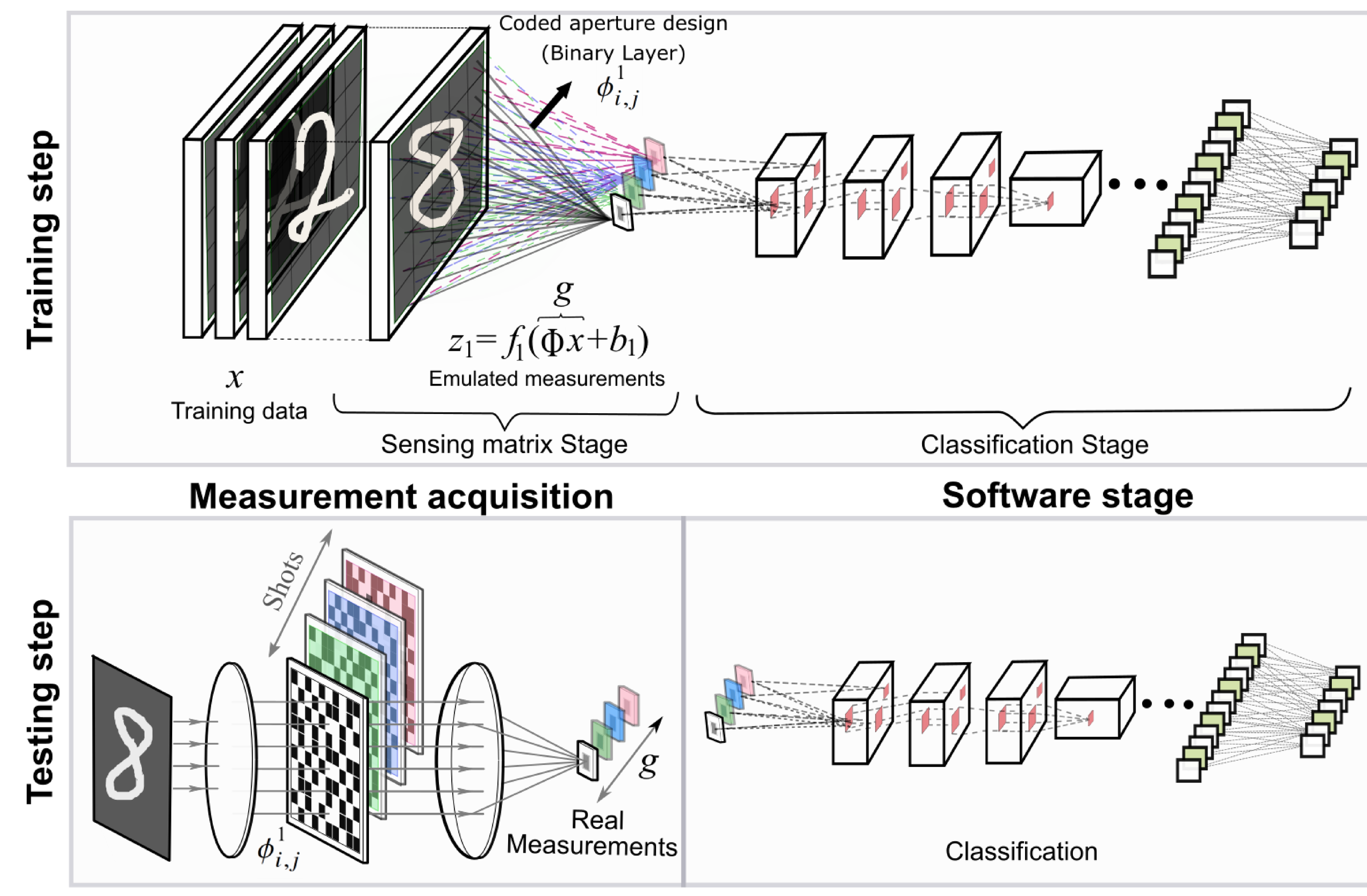
$$(g)^k = \sum_{i=1}^{M}\sum_{j=1}^{N}(\boldsymbol{\Phi}^k)_{i,j}(\boldsymbol{X})_{i,j}, \qquad (1)$$

for $k = 1, \ldots, K$ shots, or in matrix form

$$\mathbf{g} = \boldsymbol{\Phi}\mathbf{x} + \boldsymbol{\omega}, \qquad (2)$$

- $\boldsymbol{\Phi} \in \{-1, 1\}^{K \times MN}$ is a binary sensing matrix.
- $\mathbf{x} \in \mathbb{R}^{MN}$ is the vectorization of the image ($\boldsymbol{X}$)
- $\boldsymbol{\omega}$ is the acquisition noise.
- $\gamma = \frac{K}{MN}$ stands for the sensing ratio.

Coded aperture design and classification



Training step:

With a non-compressed data set $\mathcal{D}_{train} = \{\mathbf{x}_\ell, \mathbf{d}_\ell\}_{\ell=1}^{L}$, the proposed training can be expressed mathematically as

$$\{\boldsymbol{\Phi}, \boldsymbol{\theta}\} = \underset{\boldsymbol{\Phi},\boldsymbol{\theta}}{\operatorname{argmin}} \frac{1}{L}\sum_{\ell=1}^{L}\mathcal{L}(\mathcal{M}_{\boldsymbol{\theta}}(f_1(\boldsymbol{\Phi}\mathbf{x}_\ell + \mathbf{b}_1)), \mathbf{d}_\ell) + \mu \underbrace{\sum_{k=1}^{K}\sum_{n=1}^{MN}(1 + \boldsymbol{\Phi}_{k,n})^2(1 - \boldsymbol{\Phi}_{k,n})^2}_{\text{Impose binary values}}. \qquad (5)$$

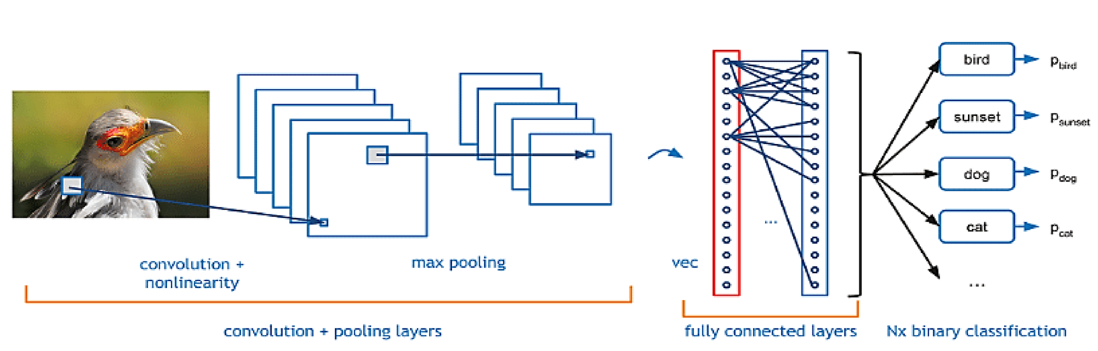
Testing step:

The trained $\boldsymbol{\Phi}$ is used in the SPC to acquire new compressed measurements g and the trained $\mathcal{M}_{\boldsymbol{\theta}}$ is used as a classification operator

$$\mathbf{z} = \mathcal{M}_{\boldsymbol{\theta}}(f_1(\underbrace{\mathbf{g}}_{\boldsymbol{\Phi}\mathbf{x}} + \mathbf{b}_1)). \qquad (6)$$

Image classification

Deep neural networks ($\mathcal{M}_\theta$) have been used for image classification.



With a training set $\mathcal{D}_{train} = \{\mathbf{x}_\ell, \mathbf{d}_\ell\}_{\ell=1}^{L}$, the goal is to learn the parameters $\boldsymbol{\theta}$, such that $\mathbf{d}_\ell = \mathcal{M}_\theta(\mathbf{x}_\ell; \boldsymbol{\theta})$ by solving

$$\boldsymbol{\theta} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \frac{1}{L}\sum_{\ell=1}^{L}\mathcal{L}(\mathcal{M}_{\boldsymbol{\theta}}(\mathbf{x}_\ell), \mathbf{d}_\ell), \qquad (3)$$

• $\mathcal{L}(\mathbf{z}_\ell, \mathbf{d}_\ell) = -[\mathbf{d}_\ell \log(\mathbf{z}_\ell) + (\mathbf{1} - \mathbf{d}_\ell)\log(\mathbf{1} - \mathbf{z}_\ell)]$ stands for the loss classification function.

• $\mathcal{M}_\theta$ is the classification operator (CNN).

After training, $\mathcal{M}_\theta$ can be used as an inference operator to obtain the classification result of a new image $\tilde{\mathbf{x}}$ as

$$\tilde{\mathbf{d}} = \mathcal{M}_\theta(\tilde{\mathbf{x}}). \qquad (4)$$

Simulations

1.Average classification accuracy for different sensing ratios with the MNIST dataset.

The classification network is a modification of the LeNet-5 model.

Sensing Ratio	Shots	Method	Test Accuracy	Training Accuracy
0.25	196	End-to-End	98.48 ±0.04%	99.99 %
		Random+CNN	98.32 ±0.06%	99.99%
		Proposed Method	**98.51 ±0.12%**	**99.99 %**
0.1	78	End-to-End	98.29 ±0.02%	99.99 %
		Random+CNN	97.01 ±0.14%	99.95%
		Proposed Method	**98.30 ±0.32%**	**99.99 %**
0.05	39	End-to-End	**98.09 ±0.04%**	**99.99 %**
		Random+CNN	94.82 ±0.09%	98.78%
		Proposed Method	97.99 ±0.24%	99.98 %
0.01	8	End-to-End	**95.18 ±0.04%**	**97.19 %**
		Random+CNN	58.94 ±0.20%	89.98%
		Proposed Method	89.78 ±0.14%	96.09 %

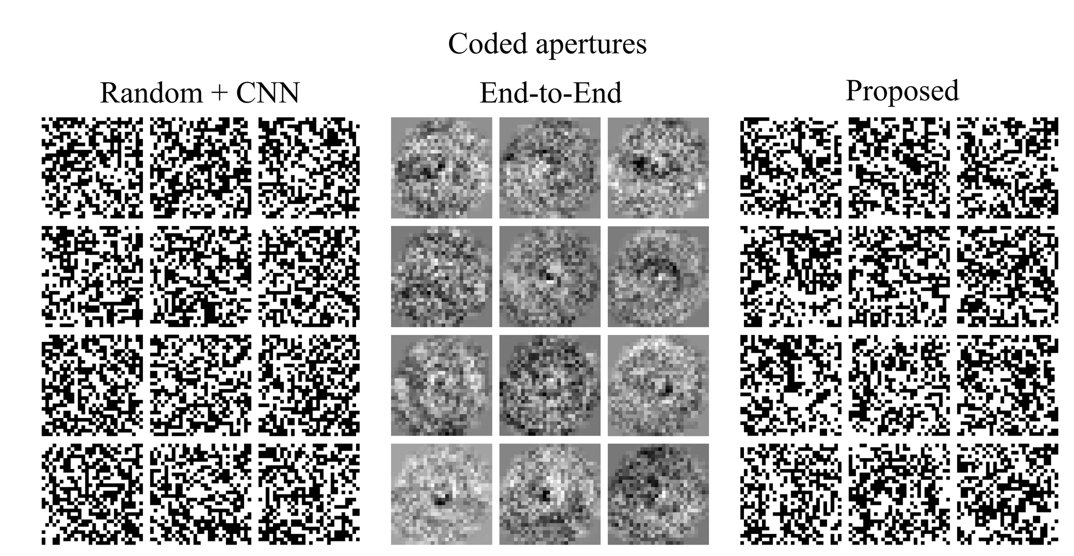
2. Average classification accuracy for different sensing ratios with the CIFAR-10 dataset .

The classification network is a modification of the AlexNet model.

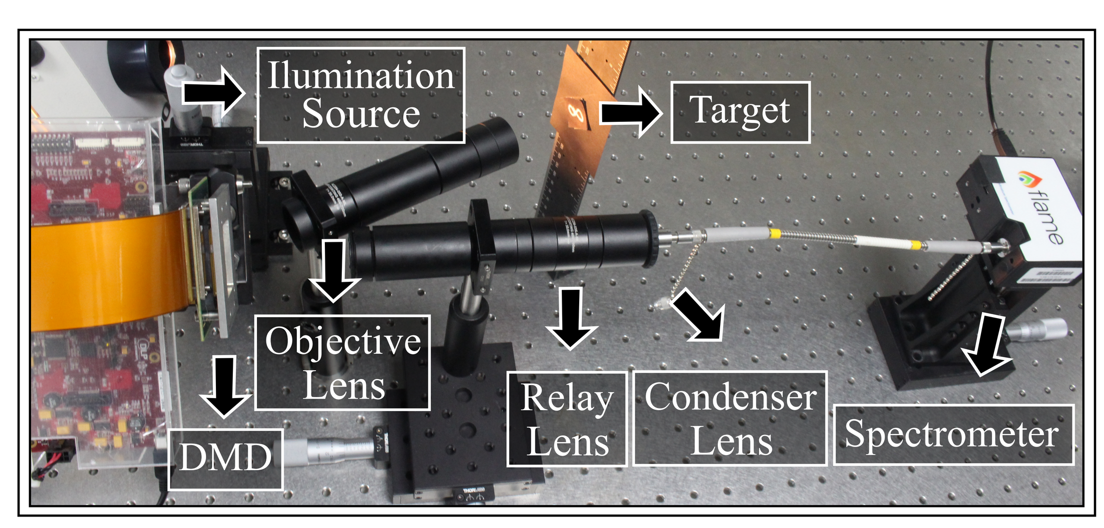
Sensing Ratio	Shots	CNN model	Test Accuracy	Training Accuracy
0.25	768	End-to-End	57.12 ±0.40%	**100 %**
		Random+CNN	45.12 ±0.24%	99.82%
		Proposed Method	**65.17 ± 1.99%**	90.12 %
0.1	307	End-to-End	55.59 ±0.55%	**99.87 %**
		Random+CNN	40.84 ±0.52%	92.55%
		Proposed Method	**59.75 ±2.20%**	89.12 %
0.05	154	End-to-End	54.34 ±0.73%	**91.89 %**
		Random+CNN	34.25 ±0.54%	89.78%
		Proposed Method	**55.78 ±1.20%**	88.79 %
0.01	30	End-to-End	**50.89 ±0.52%**	**90.64 %**
		Random+CNN	30.47 ±0.54%	88.54%
		Proposed Method	47.98 ±2.24%	79.45 %

Experiments

3. Trained coded apertures



4. Real test-bed implementation of the Single pixel imaging with the designed coded apertures.



5. Overall accuracy of the proposed method for the experimental results.

Sensing Ratio (γ)	0.0064	0.0127	0.0383	0.0638	0.1276
Overall accuracy	50 %	72,50 %	75 %	82,50%	95 %

Conclusions

A method to simultaneously learn the binary SPC linear sensing operator and extract non-linear features directly from SPC measurements that are used for classification has been proposed. After the training stage, the trained sensing matrix is employed to acquire the SPC measurements, and the trained classification network is used as an inference operator directly on these compressed measurements.

References

[1] Jorge Bacca, Laura Galvis, and Henry Arguello. Coupled deep learning coded aperture design for compressive image classification. *Optics Express*, 28(6):8528–8540, 2020.

[2] E Zisselman, A Adler, and M Elad. Compressed learning for image classification: A deep neural network approach. *Processing, Analyzing and Learning of Images, Shapes, and Forms*, 19:1, 2018.