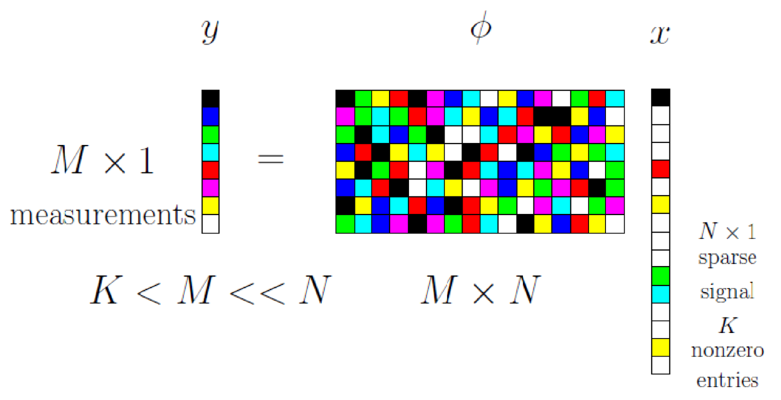
DIRECT INFERENCE ON COMPRESSIVE MEASUREMENTS USING CONVOLUTIONAL NEURAL NETWORKS

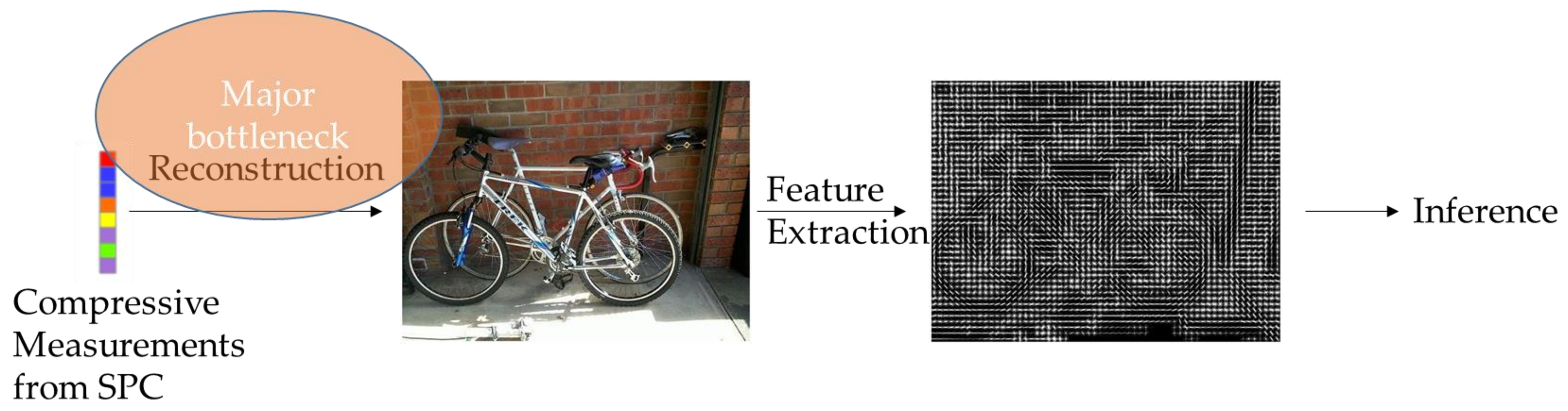
Suhas Lohit, Kuldeep Kulkarni, Pavan Turaga
School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ
School of Arts, Media and Engineering, Arizona State University, Tempe, AZ

Compressive Sensing (CS)



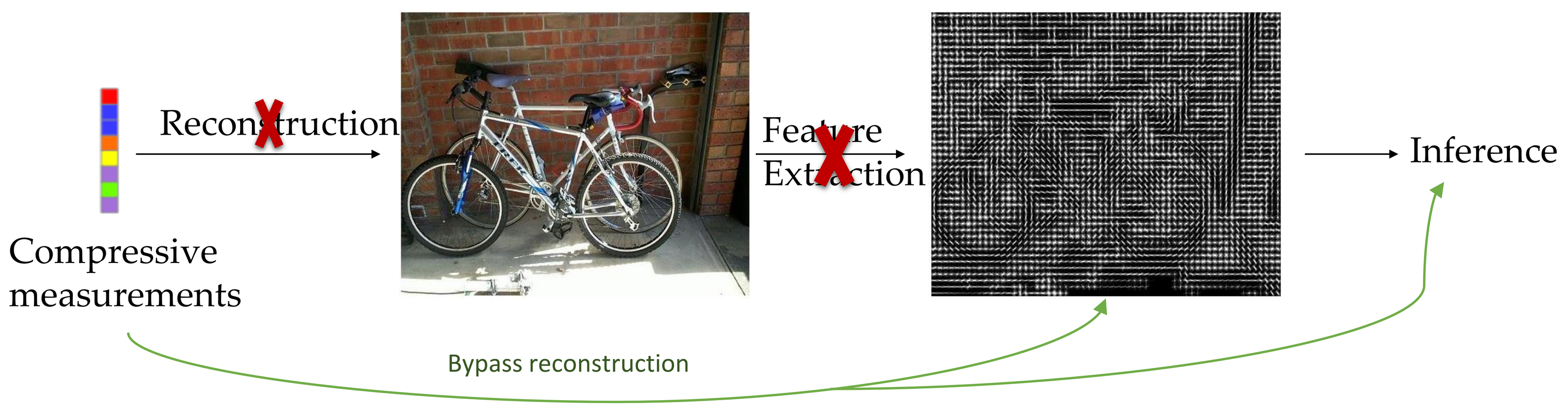
- The Single-Pixel Camera (SPC) is a popular example of a compressive imager.

Traditional Pipeline – Reconstruct-then-infer



- Recovering x from y is ill-posed but possible if x is sparse and MR (M/N) is sufficiently large.
- Most algorithms are iterative in nature and are computationally expensive. The reconstruction quality is also poor at low measurements rates of 0.1.

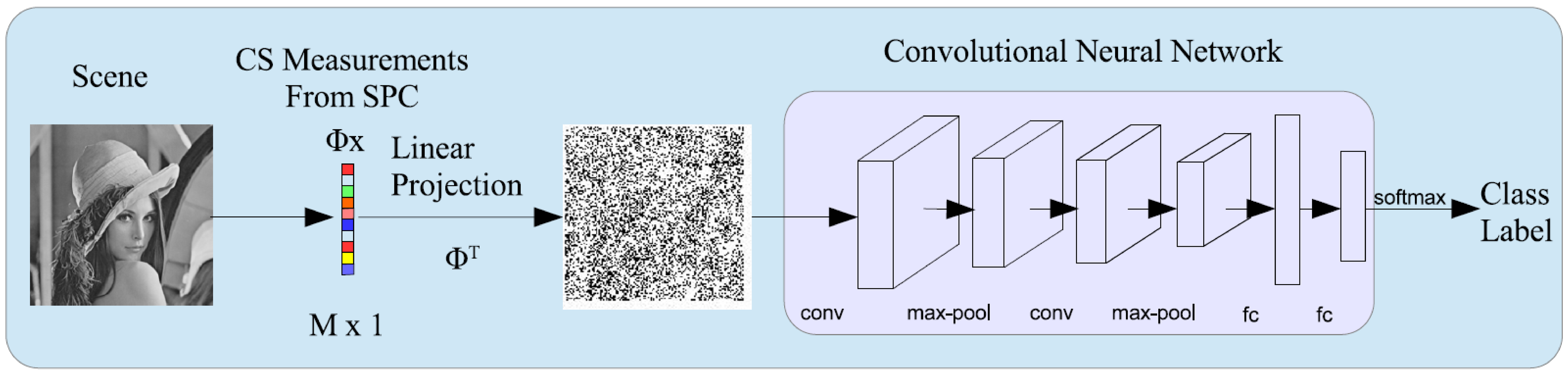
Reconstruction-free Feature Extraction/Inference



- Dimensionality-reduced matched filters – Smashed Filters [1]
  - Not robust to input variations.
  - Johnson – Lindenstrauss lemma is used to perform detection directly in the compressed domain.
  - Computationally much faster than reconstruct-then-infer paradigm.
- Dimensionality reduced correlation filters – Smashed Correlation Filters [2][3]
  - Utilizes J-L lemma to extract features directly without reconstruction.
  - More robust to input variations but cannot handle changes in pose and lighting since the features are still linear.
  - Although faster than reconstruct-then-infer, still computationally inefficient since the test image needs to be correlated with the template filter for each class.

Direct Inference Using Convolutional Neural Networks

- Project measurements back to the pixel space, which allows us to use the same CNN architectures designed for image recognition.
- Train a deep network on the "pseudo-images" to output the class labels.
- Computationally more efficient than smashed correlation filters since a single forward pass is sufficient to determine the class label.
- Possible to learn linear projection step (currently fixed to Φ^T) jointly with the remaining layers.



Experimental Results

MNIST Hand-written digit database

- Grayscale images of hand-written digits (0 - 9)
- Image size = 28 x 28 (784 pixels)
- 50000 training images, 10000 testing images
- Φ is a random Gaussian matrix of size **m x 784**
- CNN architecture based on LeNet-5 [4]

Measurement Rate (MR)	Number of Measurements (m)	Test Error (%)	
		Smashed Correlation Filters [3]	Our Method
1 (Oracle)	784	13.86	0.89
0.25	196	27.42	1.63
0.10	78	43.55	2.99
0.05	39	53.21	5.18
0.01	8	63.03	41.06

ImageNet Database

- RGB images belonging to 1000 classes
- 1.2 million training images and 50000 test images of size 256 x 256
- Φ is a low rank column permuted Hadamard matrix (approximating a Bernoulli matrix) of size **m x 65536**. Measurements are computed using Fast Walsh-Hadamard Transform.
- CNN architecture is based on AlexNet [5] – consists of 5 convolutional layers and 2 fully connected layers.

Measurement Rate (MR)	Number of Measurements (m)	Accuracy (%)
1 (Oracle)	65536	56.88
0.25	16384	39.22
0.10	6554	29.84

[1] Mark A Davenport, Marco F Duarte, Michael B Wakin, Jason N Laska, Dharmpal Takhar, Kevin F Kelly, and Richard G Baraniuk, "The smashed filter for compressive classification and target recognition," in Electronic Imaging. International Society for Optics and Photonics, 2007, pp. 64980H–64980H
[2] K. Kulkarni and P. Turaga, "Reconstruction-free action inference from compressive imagers," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. PP, no. 99, 2015.
[3] Suhas Lohit, Kuldeep Kulkarni, Pavan Turaga, Jian Wang, and Aswin C. Sankaranarayanan, "Reconstruction-free inference on compressive measurements," in 4th Intl. Conf. on Computational Cameras and Displays, held in conjunction with IEEE CVPR, June 2015.
[4] Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick ´ Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
[5] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "Imagenet classification with deep convolutional neural networks," 2012, pp. 1097–1105.