Compressed Sensing for Image Classification

Omkar Nitsure
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai, India
omkarnitsure2003@gmail.com

Ojas Karanjkar
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai, India
210070040@iitb.ac.in

***Abstract*—Compressed Sensing is a technique used in image acquisition where specific hardware cameras are used to capture images in a compressed manner. The captured images might be as small as 10% of the original image size. This can be helpful for low-resource compute environments as it requires less power and storage. The compressed measurements can also be used for image classification using some image processing technique or small machine learning models which will further reduce the compute overhead required for larger vision models using full images. In this work, we perform image classification on the MNIST dataset with 10 classes using compressive measurements. This involves sampling the image using a sensing matrix and then using the cluster centers computed on the training data to execute a kmeans-based algorithm for final classification. We achieve a classification accuracy of 75% using this traditional approach. We further learn the sensing matrix on the training data with the objective of maximizing the relative distance between different cluster centers which increases the accuracy to 87%. We also implement the GMLC (smashed filter-based) algorithm to first compute the most likely angle followed by the final classification which gives an impressive accuracy of 70%. We conclude by generating a video with MNIST digits and testing our algorithm on real-classification and it was able to classify the test images in a video of 30 fps with an accuracy of 84%.**

***Index Terms*—Compressed Sensing, Smashed Filter, kmeans , GMLC, sensing matrix, deep learning**

I. DCT Coefficients and analysis

The following figure shows clusters of DCT coefficients for the 10 classes. We ran kmeans algorithm on the clusters but the test set accuracy was only about 20. This clearly is not a good accuracy and thus motivates the need for better methods of classification.

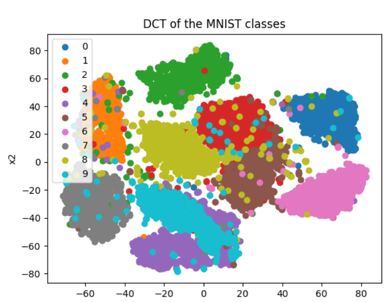


Fig. 1. Clusters of DCT coefficients of various classes

Code is available at https://github.com/omkar-nitsure/Compressed-Sensing-for-Image-Classification

II. Image and Feature clusters of measurements

The following 2 figures show the clusters formed by image pixels and only 390 measurements using a learnt sensing matrix with an objective of maximizing inter-cluster distance. We see that even after using only 50% of the original size of the image we get almost the same information through clever choice of the sensing matrix.

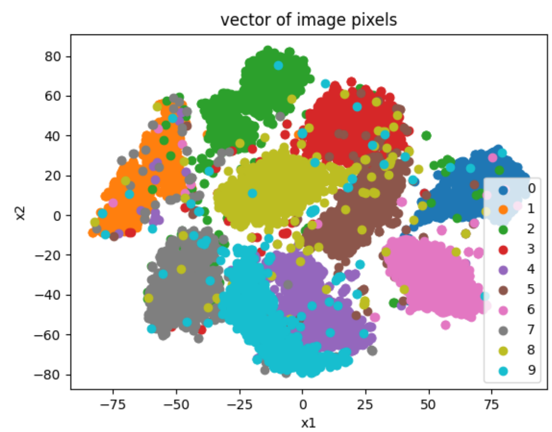


Fig. 2. Clusters of image pixels of the 10 classes

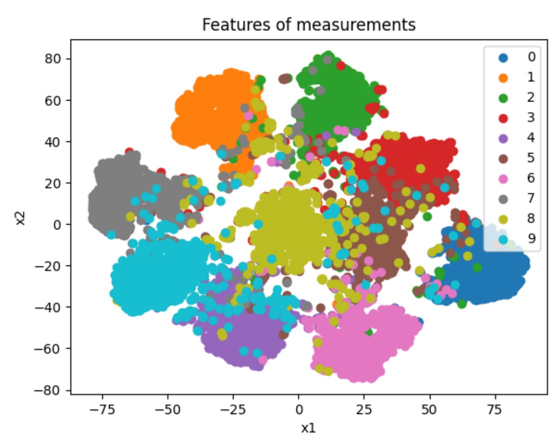


Fig. 3. Measurement clusters using sensing matrix

III. Classification accuracy for 3 classes

We performed image classification on 3 classes from the MNIST dataset by using the compressive measurements obtained through a random sensing matrix. We can see that the accuracy reaches to almost 95% using number of measurements which are only about 7% of number of pixels in the image. Plot of Classification accuracy Vs the number of measurements is given below

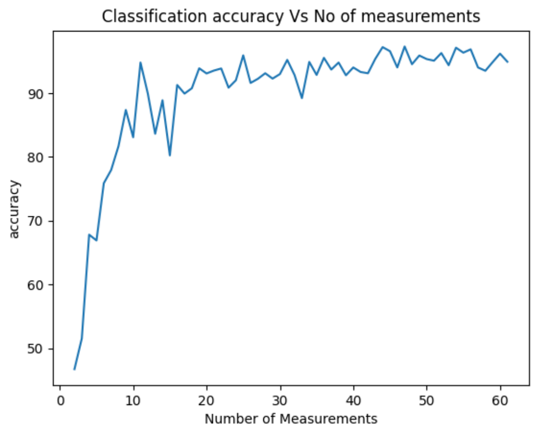


Fig. 4. Classification accuracy Vs No of measurements for 3 classes

IV. Manifold Cluster Center distances

The following plot depicts the mean cluster centroid distance among different classes as a function of number of measurements. We can see that the inter-cluster distance increases with more number of measurements which is a good sign as it suggests that having more and more measurements will help to improve classification accuracy.

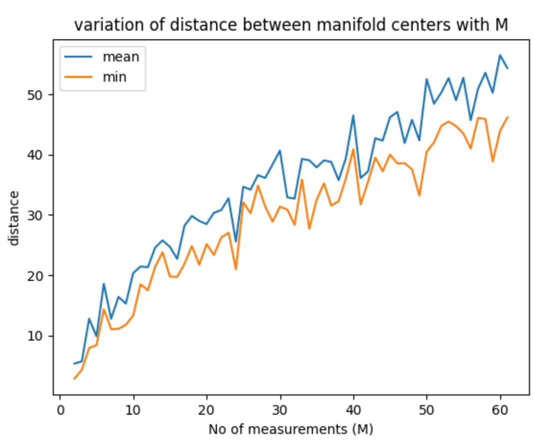


Fig. 5. Inter-centroid distance Vs No. of measurements

V. Classification using a random Sensing Matrix

The sensing matrix was initialized by samples drawn from a standard normal distribution. We then used that matrix to do classification with all the 10 classes of the MNIST dataset. The following plot depicts the classification accuracy as a function of number of measurements.

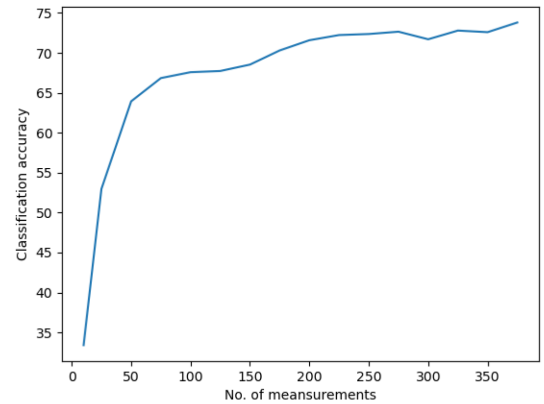


Fig. 6. Inter-centroid distance Vs No. of measurements

VI. Learning the Sensing matrix using data

We propose using the data specific to a certain application to learn the sensing matrix with the objective of maximizing the inter-cluster distance by defining a kernel per class. The exact mathematical formulation is as follows -

1) The kernel function is defined as follows:

$$K_c(f_\theta(x), e_c) = \exp\left[-\frac{\frac{1}{n}||W_c f_\theta(x) - e_c||_2^2}{2\sigma^2}\right]$$

2) The correct class label should have the largest kernel value. Thus, at test time the class label is evaluated as follows:

$$\arg\max_c K_c(f_\theta(x), e_c)$$

3) The sensing matrix was learnt to give the correct value of the kernel function using the following objective of **binary cross entropy**:

$$L(x, y) = -\sum_c y_c \log(K_c) + (1 - y_c)\log(1 - K_c)$$

VII. Classification accuracy comparisons

The classification accuracy of both the random and learnt sensing matrix for different number of measurements ranging in 10-400 were compared. We can see that the classification accuracy is consistently better for learnt sensing matrix than the random sensing matrix.

VIII. Classification Accuracy after random rotations

We defined a function to rotate an input image by an angle chosen randomly in 0-360 and then implemented smashed filter-based GMLC algorithm [1] for first computing the correct angle of rotation and then estimating the correct class. The same experiment was repeated for the learnt matrix. It can be observed that for enough number of samples, the classification accuracy for random sensing matrix is better than the learnt one. The plots are as follows -

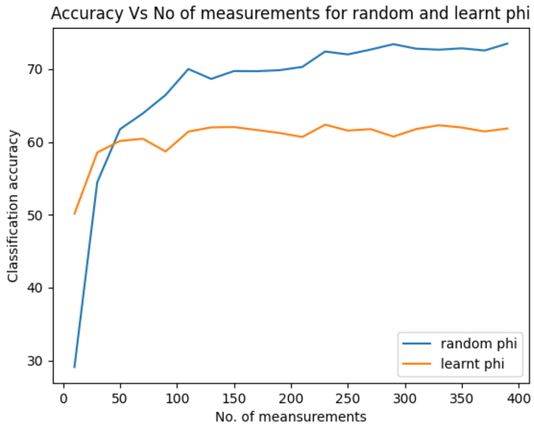


Fig. 7. Accuracy Vs no of samples

IX. Classification Accuracy after adding noise

Random normal noise was added to the data as no sensor will have perfect acquisition. We can see that the classification accuracy is better for the learnt ϕ even after adding a lot of noise. The results for various noise variances for both random and learnt sensing matrix are as follows -

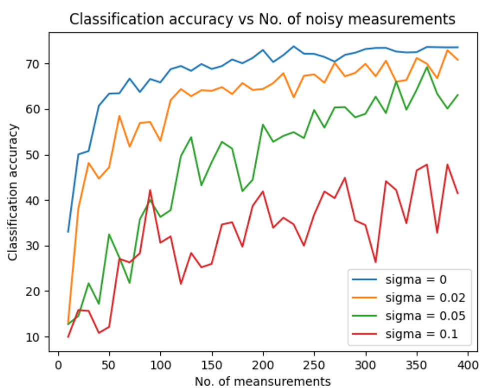


Fig. 8. random sensing matrix

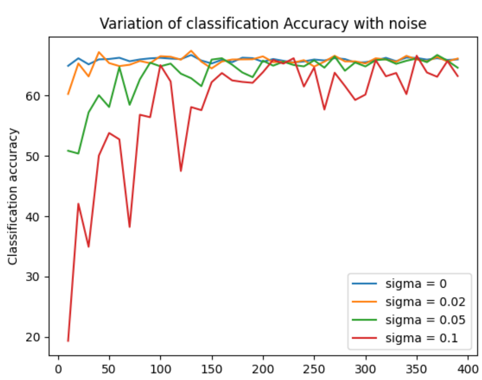
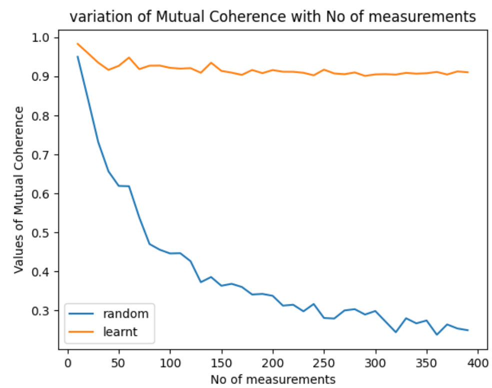


Fig. 9. learned sensing matrix

X. Mutual Coherence values for sensing matrices

Property of mutual coherence of a matrix is defined as the largest value of dot product between any 2 columns of a matrix. Smaller value of coherence allows for good reconstruction of signals of larger values of sparsity. Thus we want the mutual coherence to be as small as possible. Here we can see that the coherence values are better for the random matrix and thus randomly initialized sensing matrix will do a better image reconstruction than the learnt one. Following plot compares the coherence values of both random and learnt matrix for different number of measurements -



XI. Real time classification in video

We generated a video of 12 randomly sampled images from the test split of MNIST at a rate of 30 fps and then used our model to predict the labels of digits being displayed in the video in real-time using both random and learnt sensing matrix. The classification accuracies were as follows -

- Random sensing matrix - 9/12 correct predictions
- Learnt sensing matrix - 10/12 correct predictions

References

[1] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "The smashed filter for compressive classification and target recognition," in *Computational imaging V*, vol. 6498, pp. 142–153, SPIE, 2007.