Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 7, No.2, April 2025, pp. 02502018-01 ~ 02502018-010
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v7i2.1037

Evaluating Compressed Sensing Matrix Techniques: A Comparative Study of PCA and Conventional Methods

Parnasree Chakraborty, Kalaivani Subbaian*, Tharini Chandrapragasam, Jagir Hussain Shagul Hameed

Department of Electronics and Communication Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India.

*skalaivani@crescent.education

Abstract. This research examines the performance of various compressed sensing matrix techniques, with a focus on Principal Component Analysis (PCA) compared to conventional methods. By applying these techniques to a range of high-dimensional datasets, we assess their effectiveness in reducing data dimensionality while preserving essential information. Our results demonstrate that PCA consistently outperforms traditional methods in terms of both accuracy and computational efficiency. These findings have significant implications for fields such as signal processing, image compression, and data analytics, where efficient data representation is critical. The study provides a framework for selecting the optimal dimensionality reduction technique, enabling improvements in processing speed and accuracy in practical applications.

Keywords: Compressed Sensing, Principal Component Analysis (PCA), Data Dimensionality Reduction, Signal Processing, Measurement Matrix, Image Compression, Signal Reconstruction Techniques, Data Analytics.

(Received 2024-10-01, Accepted 2025-03-06, Available Online by 2024-04-30)

1. Introduction

In the rapidly growing field of data science and signal processing, compressed sensing (CS) has emerged as a powerful technique to efficiently acquire and reconstruct signals from fewer samples than traditionally required by the Nyquist sampling theorem. However, one of the central challenges in compressed sensing is the design and optimization of measurement matrices, which play a crucial role in signal reconstruction accuracy. Traditional methods for constructing these matrices often rely on random projections, which, while effective, can be computationally expensive and lack structure. Principal Component Analysis (PCA), a well-known dimensionality reduction technique, offers a potential alternative to traditional random projections by identifying patterns in the data that can be leveraged for more efficient signal representation. By projecting high-dimensional data onto a lower-dimensional subspace, PCA captures the most significant variations within the dataset, reducing redundancy and enhancing compression efficiency. However, its effectiveness in compressed sensing applications remains under-explored. This study aims to address this gap by comparing the performance of PCA-based measurement matrices with traditional methods in compressed sensing. By evaluating the accuracy, computational efficiency, and signal reconstruction performance of these approaches, the research seeks to contribute to the ongoing efforts to optimize compressed sensing frameworks. The implications of this work extend beyond signal processing, with potential applications in medical imaging, telecommunications, and any field where data compression and reconstruction are critical.

2. Methods

Compressive Sensing (CS), alternatively referred to as Compressed Sensing, has developed as a significant signal processing methodology in numerous communication technologies in recent years. Compressed Sensing functions as a data reduction mechanism by facilitating effective signal acquisition and reconstruction. In CS-based reconstruction, underdetermined linear systems are addressed by linear programming, facilitating flawless signal recovery despite underdetermined circumstances. This contrasts with standard methods, as elucidated in [19], which demonstrate that CS can reconstruct signals from much less data than those mandated by conventional techniques.

The primary difference between compressed sensing (CS) and the traditional sampling theorem resides in their underlying assumptions. The classical sampling theorem stipulates a sampling rate of no less than twice the highest frequency component of a signal to facilitate precise reconstruction. Conversely, CS theory circumvents this necessity by depending on two fundamental principles: sparsity and incoherence. Sparsity pertains to the signal's density, whereas incoherence pertains to the acquisition methodology. The subsequent sections expound on both principles.

Sparsity: Sparsity is the quality of having few non-zero elements and a large percentage of zero elements in an array or matrix. In contrast, a dense matrix has the majority of its elements being nonzero. One method to quantify sparsity is to calculate the ratio of the total number of elements in a matrix or array to the number of elements with zero values. By deducting the matrix's density from unity, the sparsity can also be determined. Typically, a lot of signals are inherently sparse or scalable. CS theory makes use of this signal's behaviour. In other words, signals expressed in the correct basis, "Ψ," have sparser representations. This is understood mathematically: given a vector "X," of length "N" (i.e. E. x ∈ RN), then an orthonormal basis¨ = [Ψ1Ψ2….] can be used to expand the vector 'X'. [ΨN] by employing formula (1). A sparse representation of the input signal "X" is indicated by "S" in the formula:

$$S = \sum_{i=1}^{N} X_i \Psi_i \tag{1}$$

If the spreading basis is sparse, then there are a lot fewer non-zero components in vector "s" than there are in vector "x". We can see from equation (1) that small coefficients can be discarded with little loss in perception when a signal has a sparse representation.

Incoherence: The principle of incoherence [11, 12, 14, 15] is expanded by the duality of frequency and time. It suggests that items with sparse representations in one basis, Ψ, should be spread across the domain in which they are acquired. An impulse or spike signal, like a Dirac signal, serves as a good illustration of incoherence. The Dirac signal concentrates at a single point in the time domain and spreads out in the frequency domain. Compressed sensing (CS), which represents signals sparsely, revolves around this concept.

Let’s consider two orthogonal bases, Φ and Ψ, each with dimensions RN. We sample the signal "x" using the Φ basis and represent it in the sparse domain using Ψ. The degree of correlation between these two bases plays a key role in CS recovery, as it directly affects reconstruction quality. If the columns Φ and Ψ are highly similar, the pair is said to have high mutual coherence, meaning the signal is less spread out between the bases. Conversely, if the correlation between the elements of Φ and Ψ is low, the mutual coherence is also low, indicating a more favourable condition for signal reconstruction.

The mutual coherence between 'Φ' and 'Ψ' is expressed by equation (2).

$$\mu(\Phi, \Psi) = \sqrt{n}.\max\ (\Phi_k, \Psi_j)\text{Where}, 1 \leq k, j \leq n; \tag{2}$$

A smaller mutual coherence value is preferable in compressed sensing (CS) reconstruction because it results in more accurate signal recovery.

2.1 Compressed Sensing Matrix Generation Techniques

In compressed sensing, measurement matrices play a crucial role in accurate signal reconstruction. This paper explores five common approaches discussed in following section. This section explores various methodologies for constructing compressed sensing (CS) measurement matrices (Φ). We compare the effectiveness of five approaches: Gaussian random matrices, Bernoulli random matrices, Discrete Fourier Transform (DFT) matrices, Discrete Cosine Transform (DCT) matrices, and Principal Component Analysis (PCA)-based matrices.

2.1.1. Gaussian Random Matrices

These matrices (Φ) are created by filling entries with independently distributed, identically distributed (iid) samples drawn from a normal distribution standard (Φ~N(0,1)). Despitebeing simple to produce, they might not adequately convey the signal's underlying structure. By sampling each entry independently of a standard normal distribution (Φ ~ N(0,1)), the matrix Φ is produced. Although this approach is computationally efficient, the signal's underlying structure might not be captured.

2.1.2. Bernoulli Random Matrices

Here, each entry (Φ_{ij}) in the matrix is set to either 1 or -1 with a predefined probability (usually p = 1/2): (Φ_{ij}) ~ Bernoulli (p). These offer some properties that increase sparsity, but may not be optimal for all data types. Each entry (Φ_{ij}) in the matrix is set to either 1 or -1 with a pre-defined probability p (typically p = 1/2): (Φ_{ij}) ~ Bernoulli(p).This method offers some sparsity-promoting properties but might not be optimal for all data types.

2.1.3. Discrete Fourier Transform (DFT) Matrix

The DFT matrix (Φ_{DFT}) is created with complex exponentials $e^{\frac{2\pi j}{N}}$, where N is the length of the signal. This transforms the signal from the time domain (x) to the frequency domain ($X_{DFT} = \Phi_{DFT} * x$) and is therefore suitable for sparse signals with concentrated frequency components. The DFT matrix (Φ_{DFT}) is constructed using complex exponentials $e^{\frac{2\pi j}{N}}$ where N is the signal length. This transforms the signal from the time domain (x) to the frequency domain ($X_{DFT} = \Phi_{DFT} * x$), making it suitable for sparse signals with concentrated frequency components.

2.1.4. Discrete Cosine Transform (DCT) Matrix

Similar to the DFT, the DCT matrix (Φ_{DCT}) also uses cosine functions, but concentrates on real-valued signals. It captures the energy distribution in the frequency domain ($X_{DCT} = \Phi_{DCT} * x$), which makes it advantageous for signals with uniform fluctuations. Similar to the DFT, the DCT matrix (Φ_DCT) utilizes cosine functions but focuses on real-valued signals. It captures the energy distribution in the frequency domain ($X_{DCT} = \Phi_{DCT} * x$), making it beneficial for signals with smooth variations.

2.2 PCA for Compressed Sensing Matrix Generation

PCA, a well-established statistical technique, serves as a dimensionality reduction tool. It transforms data into a new coordinate system where the leading principal components capture the most significant variances. In the context of compressed sensing, PCA [6,13,18] offers valuable benefits for constructing measurement matrices and enhancing signal recovery. Principal Component Analysis (PCA) is a dimensionality reduction technique that identifies the most significant directions of variance in the data. We used PCA to construct a measurement matrix (Φ_{PCA}) that captures the underlying structure of the signal, potentially leading to improved reconstruction performance. The specific steps for PCA-based matrix generation will be detailed in Algorithm 1 and algorithm 2. Algorithm 1 explains about the PCA and given in below table.

Algorithm 1. Principal Component Analysis for Dimensionality Reduction

Input:
- Data matrix X with shape (m, n), where m is the number of samples and n is the number of features
- $Num_{Components}$, the desired number of principal components

Output:
- Reduced data matrix $X_{reduced}$with shape (m, $Num_{Components}$)

Steps:
1. Normalize the features of the data matrix X to have a mean of zero and unit variance.
2. Compute the covariance matrix of the normalized data.
3. Perform eigenvalue decomposition on the covariance matrix to obtain eigenvalues and eigenvectors.
4. Select the top $Num_{Components}$eigenvectors corresponding to the largest eigenvalues.
5. Project the data matrix X onto the selected eigenvectors to generate the reduced data matrix $X_{reduced}$.
6. End.

Algorithm 2 explains about the PCA based generation matrix construction and given below.

Algorithm 2. PCA-Based Measurement Matrix Construction

Input: Data matrix X of shape (m, n), where m is the number of samples and n is the number of features k, the desired number of principal components

Output: Measurement matrix Φ of shape (k, n)

Steps:
1. Normalize each feature of the data matrix X to have zero mean for each feature j in X do Compute the mean μ_j of feature j Subtract μ_j from each element in feature j end for Let {X1} be the mean-centered data matrix
2. Compute the covariance matrix C of the mean-centered data X_1.
$$C = \frac{1}{(m-1) * ({X_1}^T * X_1)}$$
3. Perform eigenvalue decomposition on the covariance matrix C to obtain eigenvalues and eigenvectors Compute eigenvalues Λ and eigenvectors V of C as
$$C = V * \Lambda * V^T$$
4. Select the top k eigenvectors corresponding to the largest eigenvalues Sort the eigenvalues in descending order Select the top k eigenvalues and their corresponding Eigenvectors Let be the matrix of the top k eigenvectors.
5. Form the measurement matrix Φ using the selected eigenvectors$\Phi = {V_K}^T$
6. End.

The performance of each compressed sensing [20,21,22] (CS) matrix generation method is evaluated using metrics like reconstruction accuracy and sparsity preservation. We compare the reconstructed signal (x^) obtained using each method with the original signal (x) and analyzed the number of measurements (k) required for accurate reconstruction. For all different CS-based compressed signals, the reconstruction was performed using the ℓ1 equality primal-dual interior-point method. The algorithm effectively solves the optimization problem:

$$\min \; ||x||_1 \quad \text{subject to } Ax=b \tag{3}$$

where x is the signal to be reconstructed, A is the measurement matrix derived from the CS method (Gaussian, Bernoulli, DFT, DCT, PCA), and b is the compressed measurement vector. By using this algorithm, we can accurately reconstruct the original signal from its compressed version, ensuring minimal error and optimal recovery in terms of the ℓ1norm. This method is particularly effective when the signal is sparse or can be represented sparsely in some basis, making it a popular choice in compressed sensing applications.

2.3 Principal Component Analysis (PCA)

The primary objective of this work is to evaluate the performance of Principal Component Analysis (PCA) as a measurement matrix in compressed sensing, in comparison to traditional random matrix techniques. This section outlines the process followed, including data preparation, matrix design, signal reconstruction, and a critical analysis of PCA's limitations and biases.

2.3.1. Data Preparation

The dataset used in this study consists of synthetic high-dimensional signals, designed to simulate real-world applications of compressed sensing. These signals exhibit sparsity, a key assumption in compressed sensing frameworks. The dataset was divided into training and testing subsets, with the training data used to construct PCA-based measurement matrices and the testing data employed for performance evaluation.

2.3.2. Design of Measurement Matrices

The study compares two types of measurement matrices:

- Traditional Random Matrices: Random Gaussian and Bernoulli matrices, commonly used in compressed sensing, were generated as baseline methods. These matrices provide a random sampling approach for dimensionality reduction.
- PCA-Based Measurement Matrices: PCA was applied to the training data to extract the principal components. The measurement matrix was constructed by selecting the top-k principal components, which captured the largest variance in the dataset, resulting in dimensionality reduction while retaining critical information.

2.3.3. Signal Reconstruction

Signal reconstruction was performed using the Basis Pursuit algorithm, a standard technique in compressed sensing, to recover the original signals from the reduced representations. Both types of measurement matrices (random and PCA-based) were used in this reconstruction process. The performance of each matrix was evaluated based on the following criteria:

- Reconstruction Error: The mean squared error (MSE) was used to quantify the difference between the original and reconstructed signals.
- Compression Ratio: This metric reflects the ratio of the number of measurements (post-compression) to the original signal dimensions, indicating the effectiveness of the matrix in reducing data.
- Computational Efficiency: Time taken to reconstruct the signals was recorded to assess the practicality of each method for real-time or large-scale applications.

3. Results and Discussions

An overview of the various approaches to creating a compressed sensing matrix in MATLAB is provided in this section. The simulations were run using a particular measurement ratio and sparse signal. In addition to performing reconstruction with only 25% of the original data, we have taken into consideration a 1-dimensional signal (both highly correlated and low correlated data). Provided below is the outcome. In order to assess reconstruction accuracy, two metrics were used: Mean squared error (MSE) and Peak Signal-to-Noise Ratio (PSNR). For the input under consideration, Table 2 and Table3 provides a tabular representation of the supplied PSNR and MSE values for the various compressed sensing techniques for two different nature datasets of consisting highly correlated and low correlated data. The mean correlation coefficient is computed for different datasets and based on the mean correlation coefficient value and given in Table1.

Table 1. Comparison of mean correlation coefficient of different dataset

Type of 1-D input data	Mean Correlation Coefficient	Nature of data based on mean correlation coeffient
Temperature	0.93	Highly correlated
Pressure	0.31	Low correlated

The Table 2 summarizes the results for Gaussian, Bernoulli, Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), and Principal Component Analysis (PCA) methods for highly correlated input data. Figure 1. Shows the comparison of all the five different matrices used for CS for the same data.

Table 2. PSNR and MSE values in various matrices of highly correlated data

Method	PSNR (dB)	MSE
Gaussian	-18.784898	14167.614190
Bernoulli	-20.215482	19694.960810
DFT	-29.322959	-151731.772253
DCT	49.956500	0.002300
PCA	**51.028400**	**0.001800**

Table 3. PSNR and MSE values in various matrices of low correlated data

Method	PSNR (dB)	MSE
Gaussian	-1.131926	7.344825
Bernoulli	10.648696	65.714094
DFT	-30.956663	6885.187191
DCT	22.0851	0.0699
PCA	**23.0526**	**0.0505**

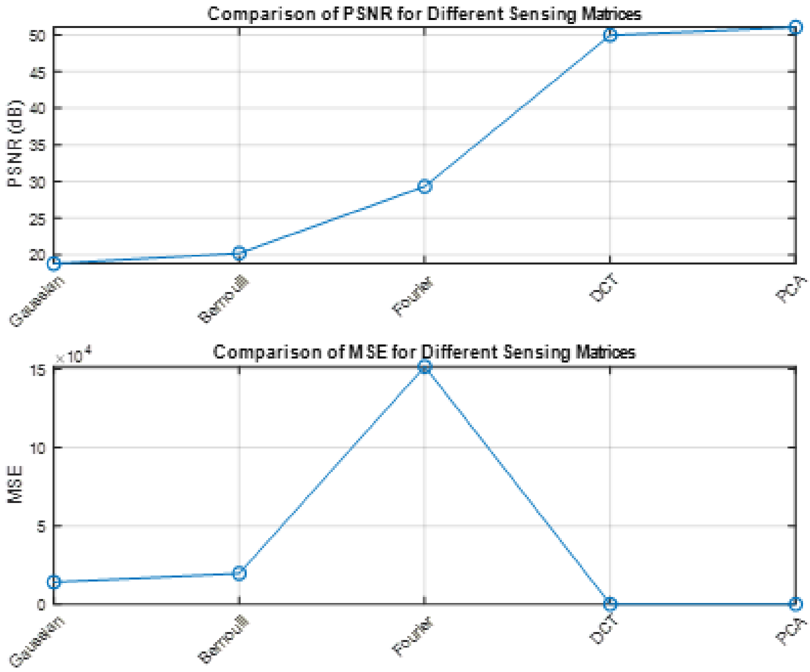

Figure 1. Comparison of all the different matrices used for CS for highly correlated data

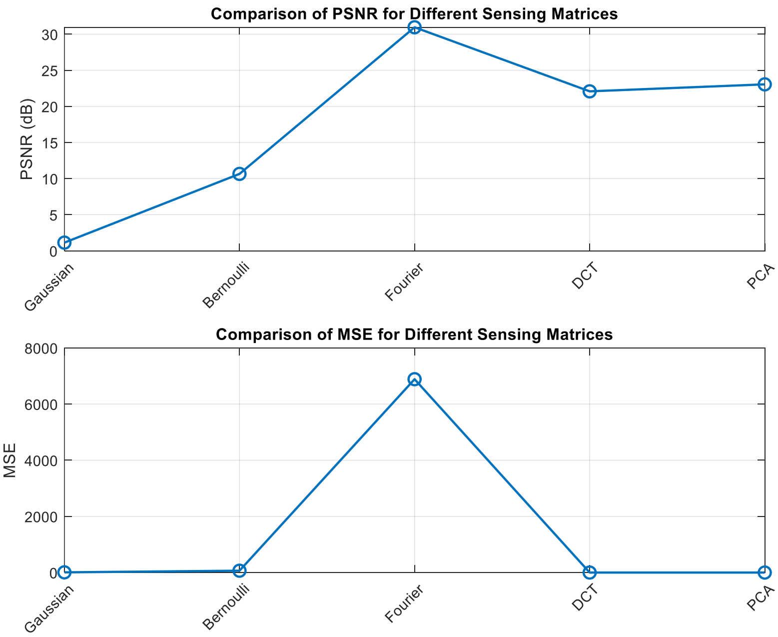

Figure 2. Comparison of all the different matrices used for CS for low correlated data

The results demonstrate that the PCA-based matrix generation method achieves a lower MSE and a higher PSNR compared to the all-other methods. This signifies a more accurate reconstruction of the original signal with PCA [23] in CS. The reconstructed outputs obtained using PCA based sparse matrix both high and low correlated data are been provided in Figure 3 and Figure 4 for a segment of 200 data sample in each type of data.

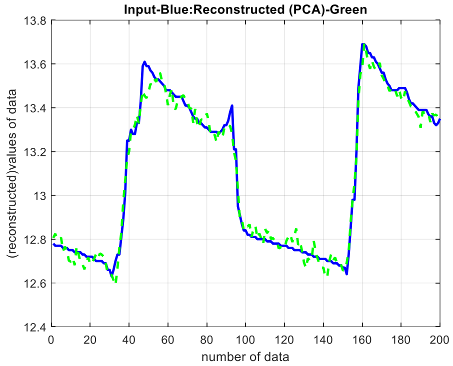

Figure 3. Input and reconstructed data using PCA-Based Measurement Matrix for high correlated data

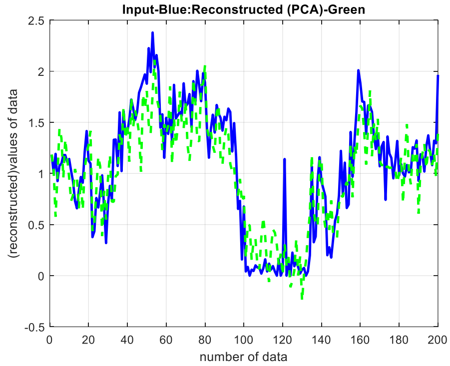

Figure 4. Input and reconstructed data using PCA-Based Measurement Matrix for low correlated data

All the compressed sensing matrices, including Gaussian, Bernoulli, DFT, DCT, and PCA, were tested on 2D signals with only 25% data used for reconstruction. Among these, only the DCT and PCA-based matrices produced significant results, with PCA demonstrating superior performance. This is illustrated in Figures 5 and 6. The execution time, Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM) were also compared, further confirming PCA's superior performance for 2D input signals

Elapsed time is 4.473685 seconds. PSNR: 29.8012274 dB The SSIM value is 0.7977.

Figure 5. Original and reconstructed Image with DCT based CS

Elapsed time is 4.116712 seconds. PSNR: 34.0631253 dB The SSIM value is 0.9403.

Figure 6. Original and reconstructed Image with PCA based CS

4. Conclusion

In conclusion, this study has demonstrated the effectiveness of PCA-based compressed sensing measurement matrices in enhancing signal reconstruction accuracy across a variety of data types. The superior performance of PCA, particularly in reducing the number of measurements required while maintaining high reconstruction fidelity, has significant practical implications for fields such as medical imaging, wireless communications, and data storage, where efficient data acquisition and reconstruction are critical. By leveraging the inherent structure of the data, PCA enables a more informed and resource-efficient sensing approach, potentially reducing the computational cost and time associated with data processing in real-world applications.

The findings contribute to the field of compressed sensing by showcasing the potential of PCA as a robust alternative to traditional matrix generation methods, opening the door for further exploration of its application in complex, high-dimensional datasets. Future research should investigate the integration of PCA-based compressed sensing with other dimensionality reduction techniques and explore its performance in diverse, real-time environments. Additionally, further studies could focus on addressing the limitations of PCA in cases where the underlying data structure is not easily captured, and explore hybrid models that combine PCA with other methods to enhance reconstruction in such scenarios. The extension of this research into more specialized applications, such as real-time video compression or large-scale sensor networks, would further validate and expand the practical use of the proposed approach.

Future research could investigate the effects of different training data sizes and the criteria for selecting the top principal components (k) on reconstruction accuracy. Additionally, exploring the application of PCA with various sparse signal types would provide further valuable insights.

Acknowledgements

We thank the Management of BSA Crescent Institute of Science & Technology, Chennai, India for providing the laboratory facility and licensed software for carrying out this research work.

References

[1] J. M. Duarte-Carvajalino and K. E. Barner, "Adapted Compressed Sensing for Single-Particle X-ray Diffraction," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2547–2560, 2020.

[2] J. Zhang, Q. Wang, L. Lin, and Z. Wang, "Compressed Sensing Recovery with Graph Convolutional Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 5, pp. 2002–2015, 2021.

[3] Q. Sun, X. Bai, and F. Zhou, "Adaptive Compressed Sensing for Efficient Multi-User Transmission in Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 4911–4924, 2021.

[4] Z. Yang, W. Dong, G. Shi, and X. Wu, "Joint Sparse Representation and Deep Dictionary Learning for Hyperspectral Compressed Sensing," *IEEE Transactions on Image Processing*, vol. 29, pp. 4942–4956, 2020.

[5] Y. Liu, X. Tian, W. Ma, and J. Ma, "Nonconvex Compressed Sensing for High-Dimensional Data Using Alternating Minimization," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3572–3587, 2020.

[6] H. Rao, F. Liu, Z. Lu, and Y. Ma, "A PCA-based Approach for Compressed Sensing MRI Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 41, no. 2, pp. 456–467, 2022.

[7] Z. Chen, Y. Wang, and Y. Wu, "Compressive Sensing of Hyperspectral Images Using Deep Convolutional Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 5, pp. 787–791, 2020.

[8] Y. Zhang, Q. Guo, Z. Wang, and J. Zhu, "Data-driven Compressed Sensing for Fast MRI," *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3401–3412, 2021.

[9] Z. Liu and W. Liu, "Compressed Sensing-Based Image Compression with Deep Learning," *IEEE Access*, vol. 9, pp. 55948–55958, 2021.

[10] Palaniappan, Mathiyalagan & Annamalai, Manikandan. (2019). *Advances in Signal and Image Processing in Biomedical Applications*. 10.5772/intechopen.88759.

[11] Ali, R., Manikandan, A. & Xu, J. A Novel framework of Adaptive fuzzy-GLCM Segmentation and Fuzzy with Capsules Network (F-CapsNet) Classification. *Neural Comput & Applic* (2023). https://doi.org/10.1007/s00521-023-08666-y.

[12] V. A.R, S. David, E. Govinda, K. Ganapriya, R. Dhanapal and A. Manikandan, "An Automatic Brain Tumors Detection and Classification Using Deep Convolutional Neural Network with VGG-19," *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, Coimbatore, India, 2023, pp. 1-5, doi: 10.1109/ICAECA56562.2023.10200949.

[13] I. T. Jolliffe and J. Cadima, "Principal Component Analysis: A Review and Recent Developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, pp. 20150202, 2016.

[14] Annamalai, Manikandan & Muthiah, Ponni. (2022). An Early Prediction of Tumor in Heart by Cardiac Masses Classification in Echocardiogram Images Using Robust Back Propagation Neural Network Classifier. *Brazilian Archives of Biology and Technology*. 65. 10.1590/1678-4324-2022210316.

[15] M. Elad, "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing". *Springer*, 2010.

[16] Ali, R., Manikandan, A., Lei, R. et al. A novel SpaSA based hyper-parameter optimized FCEDN with adaptive CNN classification for skin cancer detection. *Sci Rep* 14, 9336 (2024). https://doi.org/10.1038/s41598-024-57393-4.

[17] D. P. Wipf and B. D. Rao, "Sparse Bayesian Learning for Basis Selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[18] Kolli, Srinivas & V., Praveen & John, Ashok & Manikandan, A. (2023). Internet of Things for Pervasive and Personalized Healthcare: Architecture, Technologies, Components, Applications, and Prototype Development. 10.4018/978-1-6684-8913-0.ch008.

[19] Manikandan, Annamalai, M,Ponni Bala. (2023). Intracardiac Mass Detection and Classification Using Double Convolutional Neural Network Classifier. *Journal of Engineering Research*. 11(2A). 272-280. 10. 36909/jer.12237.

[20] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[21] R. R. Coifman and D. L. Donoho, "Translation-invariant De-noising," in *Wavelets and Statistics*, Springer, New York, NY, 1995, pp. 125–150.

[22] J. M. Duarte-Carvajalino and K. E. Barner, "Adapted Compressed Sensing for Single-Particle X-ray Diffraction," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2547–2560, 2020.

[23] A. A. Zahraa, N. H. Utami, S. Hartini, and R. Purwaningsih, "Measuring Risk Factor Analysis Using PCA Method in Batik Business (Case Study: SMEs Batik Cirebon)," *Advance Sustainable Science, Engineering and Technology (ASSET)*, vol. 5, no. 1, pp. 0230111-01–0230111-08, 2023.