

# *Sparse noise minimization in image classification using Genetic Algorithm and DenseNet*

Ibomoiye Domor Mienye

Department of Electrical and Electronic Engineering,  
University of Johannesburg,  
Johannesburg, South Africa

Ikiomoye Douglas Emmanuel

School of Computing Science and Engineering,  
University of Salford,  
Salford, United Kingdom

Priye Kenneth Ainah

Department of Electrical and Electronic Engineering,  
Niger Delta University,  
Wilberforce Island, Nigeria

Ebenezer Esenogho

Centre for Telecommunication,  
Department of Electrical and Electronic Engineering,  
University of Johannesburg,  
Johannesburg, South Africa

**Abstract**—Noise handling is a critical aspect of image processing, which can significantly affect the accuracy of classification and recognition algorithms. In this paper, we propose a technique for improved noise handling in sparse input feature maps where the noise signal is also sparse. The signal-noise relationship is formulated as an optimization problem which is solved by a genetic algorithm. The genetic algorithm is applied to optimize the setting of a non-convexity parameter which yields a more accurate image sparse matrix. The resulting feature map is then classified using a densely connected convolutional network (DenseNet). Lung computed tomography images were used for the experiments. The proposed approach achieves better performance when the classification results are compared with a case in which the input signal has not been denoised using the proposed approach.

**Keywords**—ADMM; DenseNet; genetic algorithm; lung cancer; sparse noise

## I. INTRODUCTION

Noise in image processing can make it difficult to identify salient features in images. The common forms of noise in image processing are ‘salt and pepper’ noise and Gaussian noise [1] – [2]. The presence of noise in an image usually affects the quality of the image. In medical applications, high-quality images enable medical practitioners to make both timely and accurate diagnoses of various ailments. There are several approaches for noise removal/reduction in image processing. Some of these include statistical methods, non-linear filtering, random field, and deep learning neural networks [3]. When considering appropriate noise removal techniques, it is important to ensure that the integrity of the image is not severely compromised [3]. Sparse denoising is the process of removing noise from a sparse representation of an image. The importance of using sparse images is that it can reduce image processing time. There are also various kinds of sparse noise.

Meanwhile, a norm is a function that gives a strictly positive value to a vector or variable. Many norm minimizations are used to achieve sparse representation in image processing. These include the  $L_0$ -norm, the  $L_p$ -norm, the  $L_1$ -norm, and the  $L_{2,1}$ -norm [4]. Unlike the  $l_1$ -norm, the  $l_0$ -norm approach to sparse representation is presented as an NP-hard problem whose solution is difficult to approximate [5]. Therefore, using the  $L_1$ -norm to approximate a sparse representation problem results in an analytical solution, and that is also solvable within polynomial time. Many optimization techniques are also used to perform sparse representation, such as methods that consider constraints called constrained optimization.

The method of constrained optimization involves optimizing an objective function with respect to some variables in the presence of constraints on those variables. There are several constrained optimization methods for sparse representation of images. Some of these include gradient projection sparse reconstruction (GPSR) [6], the interior-point method based sparse representation strategy [7], and alternating direction method (ADM) based sparse representation strategy [8]. This paper proposes an improved alternating direction method of multipliers (ADMM) approach with enhanced noise removal capability to further improve image quality. This is important because while the  $L_1$ -norm approach is effective in obtaining sparse image representations, it is susceptible to the problem of non-convexity. This problem results in a misrepresentation of the final image sparse representation. Therefore, we use a genetic algorithm (GA) to optimize the setting of a non-convexity parameter which yields a more accurate image sparse matrix.

The ADMM is an optimization approach that solves a nonlinear problem by decomposing it into smaller parts and then solves each part sequentially. This ensures that such a problem is easier to solve compared to solving it in its original form. Meanwhile, the  $L_1$ -norm has been used extensively as a regularizer for sparsity. However, it has a problem of trivializing non-zero values of the original signal. This results

in an inaccurate representation of the sparse signal, and consequently, this affects the reconstruction of the original image negatively. Hence, we propose this improvement to using the  $L_1$  regularizer by applying a constrained optimization strategy.

We will also use a genetic algorithm to optimize the objective function in order to obtain the best possible sparse representation of the original image. The rest of the paper is organized as follows: Section 2 gives a brief review of some applications of the  $L_1$ -norm regularization approach in image classification. Section 3 proposes the approach of optimal selection of the non-convexity parameter for improving the image sparse matrix representation. Section 4 briefly discusses the DenseNet structure used for the image classification process. Section 5 specifies details regarding the image dataset used in the paper, and the performance indices used to test the proposed approach. Section 6 discusses the results obtained and Section 7 concludes the paper.

## II. RELATED WORKS

This section discusses some applications of the  $L_1$ -norm regularization related approaches in image classification. Several regularization methods have been proposed to represent image reconstruction. One of these is the low-rank matrix modeling of local  $k$ -space neighborhoods (LORAKS) [9], and the priority was given to instances involving the low-ranking phase of magnetic resonance imaging (MRI) images. In this work, a rank-deficient matrix was constructed such that its rank was equal to the number of non-zero values. Negligible (near zero) values were approximated using the proposed low-rank method. The constraints considered were phase and support constraints with respect to the image reconstruction process. From the results obtained, the proposed LORAKS method outperformed both conventional  $L_1$ -norm and total variation minimization approaches in terms of the reconstruction error.

Furthermore, in [10], a reconstruction method called  $L_1$ -norm regularization piecewise constant level set approach was proposed. This reconstruction approach was used to improve image results for the fluorescence molecular tomography imaging technique. This technique is based on the fact that fluorescent probes at certain intensities can be used to target specific tumors within the human body. The optimization function was modeled as an  $L_1$ -norm regularization function which is convex. Therefore the Lagrangian-based method was used to optimize the function. The results of this approach were validated by reliable reconstruction results when compared to two other conventional methods.

Additionally, in [11], the ADMM approach was used to solve a convex objective function using a non-convex regularizer. This approach was used to perform 1- and 2-dimensional signal denoising using non-convex penalty functions. The Lagrangian parameter was selected in such a way that its magnitude was greater than the inverse of the tight frame condition parameter. Results obtained for the proposed method were better than the  $L_1$ -norm regularization and reweighted  $L_1$  minimization methods in terms of peak signal-to-noise ratio (PSNR). Also in [12], a mesoscopic fluorescence molecular tomography (MFMT) approach was proposed for performing high-resolution imaging of 3D

biological tissue samples. This approach was based on  $L_1$ -norm regularization. The proposed method achieved superior performance compared to the conjugate gradients (CG) and least square QR (LSQR) methods. Specifically, it achieved a spatial resolution of  $80\mu\text{m}$  for a tissue sample of  $3\text{mm}$  thickness, with accurate locations of fluorophore distribution within the sample. Other applications of the  $L_1$ -norm in image sparsity can be found in [13]-[17]. One common challenge of using the conventional  $L_1$ -norm approach is that it often does not achieve sufficient sparsity [18]. Hence, we propose a method to improve the capability of the  $L_1$ -norm regularization by leveraging the merits of both a genetic algorithm for optimizing the setting of the Lagrangian parameter and non-convex regularization.

## III. PROPOSED METHODOLOGY

The section proposes the approach of optimal selection of the non-convexity parameter for improving the image sparse matrix representation. The signal denoising problem is formulated according to [11]:

$$\arg \min_a [G(a) = \frac{1}{2} \|b - a\|_2^2 + \sum_{i=1}^j \gamma_i \beta((Ta)_i \cdot \eta_i)] \quad (1)$$

where  $\gamma_i > 0$  are regularization parameters,  $\beta$  is a real-valued parameter which attacks non-smoothness of the sparsity matrix,  $\eta_i$  ensures non-convexity of  $\beta$  to enable the convexity of the resulting objective function,  $b$  is the noisy signal resulting from the original input image  $a$ ,  $T$  is an  $m \times n$  tight frame matrix representing the input signal.

Concerning the non-convexity of the penalty parameter  $\beta$ , the following assumptions are made:

1. The first derivative of  $\beta$  over the range of positive real numbers approaches unity i.e.  $\beta(a, 0^+) \rightarrow 1$
2. Based on the assumption in 1, therefore:
$$\beta(a) = |a| \quad (2)$$
3.  $\beta'_{a \rightarrow \infty}(a, \eta) = -\eta \quad \forall a > 0 \quad (3)$

$\eta$  controls the degree of non-convexity of  $\beta$ . Depending on the sparsity matrix, the value of  $\beta$  is varied to keep  $G(a)$  as convex as possible. To ensure that this is the case, we also enforce the condition that  $\beta$  is symmetrical about the origin:

$$\beta(-a, \eta) = \beta(a, \eta) \quad (4)$$

Also, according to [20], we consider a proximity operator with respect to the resulting noisy signal  $b$ , this is defined as:

$$\Gamma_\beta(b, \gamma) = \arg \min_{b \in \mathbb{R}} \left[ \frac{1}{2} (b - a)^2 + \gamma \beta(a, \eta) \right] \quad (5)$$

To represent the original problem  $G(\cdot)$  in a more relaxed form, we specify a Lagrangian operator according to the following form:

$$\left(\frac{c}{a}\right)^{1/2} - b \leq x \quad (6)$$

In terms of the Lagrangian, we define the problem in Equation (4) subject to the following non-linear constraint:

$$|a(x + b)^2| \geq c \quad (7)$$

Therefore, the Lagrangian of the denoising problem is formulated as:

$$\mathcal{L}(\arg \min_{a,x} G(a, x)) = [\frac{1}{2} \|b - a\|_2^2 + \sum_{i=1}^J \gamma_i \beta ([Ta]_i \cdot \eta_i) + (\|a(X + b)\|_2^2 - c)] \quad (8)$$

With respect to Equation (8), the regularization parameter setting for the non-convex behavior of the problem is dynamically optimized using a genetic algorithm as specified in Algorithm 1. The ADMM approach is also implemented using GA for decomposing the convex problem  $G(\cdot)$ . The algorithm is shown in Algorithm 2.

Algorithm 2 ensures that the original problem  $G(\cdot)$  has a solution that is as close as possible to the optimal solution. It is important to note that convergence of the ADMM algorithm to the optimal solution depends on the degree of convexity of  $G(\cdot)$  [21]. Therefore, we choose to perform the steps in Algorithm 1 with respect to the Lagrangian in Equation (8). This ensures optimal convexity before the problem is decomposed using the ADMM approach. Meanwhile, the block diagram of the proposed methodology for the image classification is shown in Figure 1.

---

**Algorithm 1**

---

Initialize  $G(a), \gamma_i, \beta, \eta_i$   
for  $i = 1, 2, \dots$   
set rank of  $\mathcal{L}(G(a, x))$  equal to the rank of  $G(a)$   
optimize  $\arg \min_a G(a)$  while iteratively varying  $\gamma, \beta$  and  $\eta$  for optimal convexity  
end for  
end

---



---

**Algorithm 2**

---

Input  $a, b, T, c$   
Initialize  $\gamma, \beta, \eta$   
Repeat:  
 $a, \gamma, \beta, \eta \rightarrow \Gamma_{\beta, opt}(b, \gamma)$   
Until convergence

---

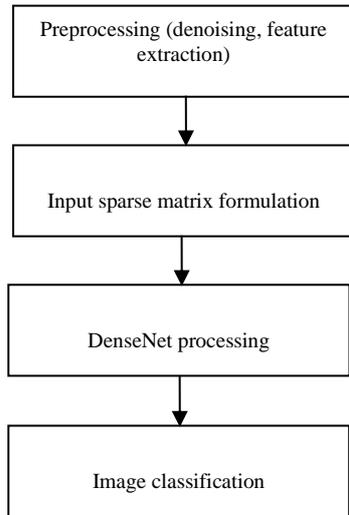


Fig. 1. Proposed methodology for medical image classification

#### IV. IMAGE CLASSIFICATION

The process of image classification is implemented using a Densely Connected Convolutional Network (DenseNet) [19]. This neural network structure has the following advantages:

- Learned features from previous layers are propagated to successive layers throughout the network
- The number of training parameters is significantly reduced
- Learning gradient is maintained all through the learning stage
- Features used during the training of various stages of the network are reused throughout the learning process

#### V. DATASET AND PERFORMANCE INDICES

The image dataset used in this paper consists of 422 lung computed tomography (CT) images obtained from The Cancer Imaging Archive (TCIA) supported by the University of Arkansas for Medical Sciences (UAMS), under collection name: NSCLC-Radiomics [22]. To evaluate the performance of the proposed method, we use some performance indices which include peak signal-to-noise ratio (PSNR), average sparsity, recognition accuracy, and computational time. Average sparsity is the ratio of non-zero elements to the total number of sparse features [14]. This means that the smaller the value of average sparsity, the better the diversity of the sparse representation [23]. Mathematically, it is represented as:

$$\text{Average sparsity} = \frac{1}{F_S} \sum_{n=1}^{F_S} \frac{\sum_{p=1}^P \|v_{s,p}\|}{P \times R} \quad (9)$$

$F_S$  is the total number of sparse features,  $P$  is the total number of patches in the image,  $v_{s,p}$  is the sparse feature vector of patch  $p$ ,  $R$  is the total resolution of the image represented by  $P$ . The PSNR measures the quality of the reconstructed image in dB by measuring the peak signal to noise ratio compared to the original image. Therefore, the higher the PSNR, the better the quality of reconstruction.

$$\text{PSNR} = 10 \log_{10} \left( \frac{f_{max}^2}{MSE} \right) \quad (10)$$

$f_{max}$  is maximum fluctuation in the input image data type,  $MSE$  is mean square error.

#### VI. RESULTS AND DISCUSSIONS

In this work, we use the genetic algorithm to find the optimal setting of the non-convexity parameter,  $\eta$  for which the peak signal-to-noise ratio (PSNR) of the resulting sparsity matrix is maximum, i.e. the non-convex regularization was performed by using the genetic algorithm to optimize the non-convexity parameter  $\eta$ . The genetic algorithm has the following settings for the optimization process:

Number of particles	100
Mutation rate	0.7
Feature evaluations	5,000
Search space dimension	50

TABLE I. PSNR VALUES FOR 10 SETTINGS OF CROSSOVER RATE ( $C_r$ ) FOR GA

$C_r$	0.45	0.38	0.58	0.63	0.25	<b>0.70</b>	0.84	0.92	0.75	0.69
$\eta$	0.16	0.24	0.36	0.29	0.44	<b>0.35</b>	0.31	0.53	0.66	0.51
PSNR (dB)	11.3	12.5	16.1	15.7	12.9	<b>21.8</b>	18.4	10.1	12.2	13.3

TABLE II. VARIATION BETWEEN CROSSOVER RATE AND AVERAGE SPARSITY OF INPUT SPARSE MATRIX

$C_r$	0.45	0.38	0.60	0.63	0.25	<b>0.70</b>	0.84	0.92	0.72	0.55
$\eta$	0.16	0.24	0.32	0.29	0.44	<b>0.35</b>	0.31	0.53	0.61	0.57
Av. sparsity	0.121	0.270	0.018	0.167	0.036	<b>0.012</b>	0.388	0.432	0.274	0.094

TABLE III. COMPARISON BETWEEN PROPOSED SND-ADMM AND CONVENTIONAL ADMM FOR AVERAGE SPARSITY, COMPUTATIONAL TIME, AND RECOGNITION ACCURACY

	$C_r$	$\eta$	Av. sparsity	Comp. time (secs)	Recognition accuracy
Proposed method (sparse noise denoising ADMM)	0.45	0.16	0.121	46.5	76.02±0.56
	0.38	0.24	0.270	53.8	77.43±0.48
	0.58	0.36	0.018	51.7	73.67±0.43
	0.63	0.29	0.167	44.3	75.43±0.41
	0.25	0.44	0.036	41.5	71.07±0.74
	<b>0.70</b>	<b>0.35</b>	<b>0.012</b>	<b>36.9</b>	<b>89.65±0.55</b>
	0.84	0.31	0.388	42.1	81.73±0.31
	0.92	0.53	0.432	48.2	69.42±0.68
	0.75	0.66	0.274	52.4	74.83±0.51
	0.69	0.51	0.094	50.6	72.17±0.53
Without sparse noise denoising (ADMM only)	-	0.16	0.263	55.2	69.12±0.66
	-	0.24	0.417	59.6	71.23±0.41
	-	0.36	0.219	60.2	68.27±0.49
	-	0.29	0.378	50.4	70.35±0.48
	-	0.44	0.116	49.4	61.92±0.81
	-	0.35	0.217	47.1	74.66±0.39
	-	0.31	0.458	51.6	71.72±0.44
	-	0.53	0.528	54.2	61.88±0.62
	-	0.66	0.389	61.9	68.94±0.31
	-	0.51	0.215	58.4	69.52±0.52

To demonstrate the effectiveness of the GA to optimize the non-convexity parameter, we examined the denoising capability of Algorithm 1 by using various settings of the GA crossover rate and observing the effect on PSNR. The results are shown in Table 1.

From the results in Table 1, the mutation or crossover rate was varied and the effect on the non-convexity parameter was observed. There was no clear trend regarding the relationship between  $C_r$  and  $\eta$ . However, it was established that a variation in the crossover rate resulted in changes in the non-convexity parameter. Also, this variation affected the PSNR. From the results obtained, the mutation rate with the highest PSNR was 0.70, and this was used in the parameter setting for the GA. As a further investigative step to determine the significance of the crossover rate in the preprocessing step of the image classification process, we examined the effect of its variation on the average sparsity of the sparse matrix that formed the input to the DenseNet. A better sparse matrix is indicated by a lower value of average sparsity. The results are shown in Table 2.

From Table 2, it is important to note that there is a relationship between the PSNR and the average sparsity for the sparse matrix. In other words, for the optimal crossover rate setting of 0.70, it is observed that the PSNR and average sparsity values were the best compared to the other nine values used. Ten different settings were chosen for each case to provide diversity in the analysis. Also, it can be observed that in Table 2, slightly different values were used in some cases to ensure that better values of  $C_r$  and  $\eta$  did not exist within the experimental limits of this paper. We also compared the performance of the DenseNet in terms of classification of the CT images for the specified dataset for the 10 settings specified in Table 1. The performance indices used

were the average sparsity, computational time, and recognition accuracy. The results are summarized in Table 3.

From Table 3, it can be seen that the proposed method outperforms the conventional ADMM method in terms of average sparsity, computational time, and recognition accuracy. In particular, we observe that the use of GA to select the setting for the non-convexity parameter is beneficial for achieving improved computational time and recognition accuracy, even in the case where there is no sparse denoising. It can be seen that the conventional ADMM method achieves better performance with  $\eta = 0.35$ . The proposed method achieves superior performance for all performance indices considered. To further show how good the proposed method performs, it is compared with two other well-known CNNs for medical image recognition, and the results are shown in Table 4.

From the results obtained in Table 4, it can be seen that the proposed method generally performs better than the other two methods with respect to average sparsity, computational time, and recognition accuracy. However, the performance of the R-

CNN and MDNet can also be improved if they are integrated with the proposed denoising framework. Furthermore, the proposed pre-processing algorithm improves the classification accuracy of the DenseNet compared to a case where the DenseNet alone was used to classify the images. Another observation that is important to note is the fact that improved performance has been observed with the same DenseNet structure used in the proposed approach and conventional ADMM. In other words, improved image recognition results can be achieved not only by improving the structure of the neural network but also by improving the preprocessing stage of the image recognition process.

TABLE IV. PERFORMANCE COMPARISON BETWEEN THE PROPOSED METHOD, REGION-BASED CNN (R-CNN) AND MULTI-DOMAIN CNN (MDNET)

	$C_r$	$\eta$	Av. sparsity	Comp. time (secs)	Recognition accuracy
Proposed method	0.45	0.16	0.121	46.5	76.02±0.56
	0.38	0.24	0.270	53.8	77.43±0.48
	0.58	0.36	0.018	51.7	73.67±0.43
	0.63	0.29	0.167	44.3	75.43±0.41
	0.25	0.44	0.036	41.5	71.07±0.74
	<b>0.70</b>	<b>0.35</b>	<b>0.012</b>	<b>36.9</b>	<b>89.65±0.55</b>
	0.84	0.31	0.388	42.1	81.73±0.31
	0.92	0.53	0.432	48.2	69.42±0.68
	0.75	0.66	0.274	52.4	74.83±0.51
	0.69	0.51	0.094	50.6	72.17±0.53
R-CNN	-	-	0.312	49.8	71.74±0.41
	-	-	0.217	55.1	70.18±0.31
	-	-	0.119	56.9	67.86±0.85
	-	-	0.114	46.9	69.38±0.73
	-	-	0.098	48.2	68.42±0.55
	-	-	0.109	50.8	61.49±0.61
	-	-	0.401	55.6	70.45±0.42
	-	-	0.591	59.6	63.03±0.45
	-	-	0.304	60.3	71.34±0.49
MDNet	-	-	0.113	59.3	69.83±0.36
	-	-	0.293	48.4	72.42±0.69
	-	-	0.210	54.6	72.63±0.63
	-	-	0.108	54.3	70.58±0.48
	-	-	0.173	45.1	72.94±0.52
	-	-	0.052	46.3	69.11±0.56
	-	-	0.058	48.5	81.17±0.44
	-	-	0.099	42.1	71.47±0.59
	-	-	0.453	51.8	66.18±0.27
-	-	0.292	57.9	73.17±0.59	
-	-	0.101	56.5	70.88±0.53	

## VII. CONCLUSION

In this paper, a method is proposed to improve medical image recognition tasks. It involves an improved ADMM approach with enhanced noise removal capability to improve the image quality which leads to better classification results when the DenseNet is applied. To achieve this we used a genetic algorithm to optimize the setting of a non-convexity parameter which yields a more accurate image sparse matrix. From the results obtained, we have established that improving the preprocessing stage of the image recognition process can significantly improve the overall recognition accuracy and performance. In particular, we see that using GA as a tuning approach for the non-convexity parameter improves the noise removal capability of ADMM. While it has been established that DenseNet is effective in image recognition applications, this paper has demonstrated the importance of optimal tuning of denoising parameters. This has been further established by the results obtained in Table 3, in which the results gotten using conventional ADMM with the same DenseNet structure were inferior to those obtained using the GA-optimized non-convexity parameter. Future research will investigate the effect of GA parameter tuning on other forms of norm regularization by comparing their performance on selected image recognition tasks.

## REFERENCES

- [1] J. J. Madhura and D. R. R. Babu, "An effective hybrid filter for the removal of Gaussian-impulsive noise in computed tomography images," in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2017, pp. 1815–1820.
- [2] S. Gupta and R. K. Sunkaria, "Real-time salt and pepper noise removal from medical images using a modified weighted average filtering," in *2017 Fourth International Conference on Image Information Processing (ICIIP)*, 2017, pp. 1–6.
- [3] P. Sun, D. Wang, V. C. Mok, and L. Shi, "Comparison of Feature Selection Methods and Machine Learning Classifiers for Radiomics Analysis in Glioma Grading," *IEEE Access*, vol. 7, pp. 102010–102020, 2019.
- [4] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A Survey of Sparse Representation: Algorithms and Applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.
- [5] E. Amaldi and V. Kann, "On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems," *Theoretical Computer Science*, vol. 209, no. 1–2, pp. 237–260, Dec. 1998.
- [6] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, Dec. 2007.
- [7] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2009.
- [8] J. Yang and Y. Zhang, "Alternating direction algorithms for L1-problems in compressive sensing," *SIAM journal on scientific computing*, vol. 33, no. 1, pp. 250–278, 2011.
- [9] J. P. Haldar, "Low-Rank Modeling of Local  $k$ -Space Neighborhoods (LORAKS) for Constrained MRI," *IEEE Transactions on Medical Imaging*, vol. 33, no. 3, pp. 668–681, Mar. 2014.
- [10] Y. An *et al.*, "A Novel Region Reconstruction Method for Fluorescence Molecular Tomography," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1818–1826, Jul. 2015.
- [11] A. Parekh and I. W. Selesnick, "Convex Denoising using Non-Convex Tight Frame Regularization," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1786–1790, Oct. 2015.
- [12] F. Yang, M. S. Ozturk, L. Zhao, W. Cong, G. Wang, and X. Intes, "High-Resolution Mesoscopic Fluorescence Molecular Tomography Based on Compressive Sensing," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 1, pp. 248–255, Jan. 2015.
- [13] M. Thiel and D. Omeragic, "High-Fidelity Real-Time Imaging With Electromagnetic Logging-While-Drilling Measurements," *IEEE Transactions on Computational Imaging*, vol. 3, no. 2, pp. 369–378, Jun. 2017.
- [14] Y. Yu, Z. Sun, W. Zhu, and J. Gu, "A Homotopy Iterative Hard Thresholding Algorithm With Extreme Learning Machine for Scene Recognition," *IEEE Access*, vol. 6, pp. 30424–30436, 2018.
- [15] W. Liu, Z. Zhang, X. Chen, S. Li, and Y. Zhou, "Dictionary Learning-Based Hough Transform for Road Detection in Multispectral Image," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2330–2334, Dec. 2017.
- [16] F. Xiaoyu, L. Qiusheng, and S. Baoshun, "Compressed Sensing MRI With Phase Noise Disturbance Based on Adaptive Tight Frame and Total Variation," *IEEE Access*, vol. 5, pp. 19311–19321, 2017.
- [17] M. Mirzaei, A. Asif, and H. Rivaz, "Combining Total Variation Regularization with Window-Based Time Delay Estimation in Ultrasound Elastography," *IEEE Transactions on Medical Imaging*, vol. 38, no. 12, pp. 2744–2754, Dec. 2019.
- [18] J.-F. Cai, S. Osher, and Z. Shen, "Split Bregman Methods and Frame Based Image Restoration\*," *Multiscale Modeling & Simulation; Philadelphia*, vol. 8, no. 2, pp. 337–369, 2009.
- [19] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269.
- [20] P. L. Combettes and J.-C. Pesquet, "Proximal thresholding algorithm for minimization over orthonormal bases," *SIAM Journal on Optimization*, vol. 18, no. 4, pp. 1351–1376, 2007.
- [21] J. Eckstein, and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 3, pp. 293–318, 1992.
- [22] University of Arkansas for Medical Sciences (UAMS), "TCIA Collections," *TCIA Collections*. [Online]. Available: <https://www.cancerimagingarchive.net/>. [Accessed: 08-Nov-2019].
- [23] I. D. Mienye, Y. Sun, and Z. Wang, "Improved sparse autoencoder based artificial neural network approach for prediction of heart disease," *Informatics in Medicine Unlocked*, vol. 18, p. 100307, Jan. 2020, doi: 10.1016/j.imu.2020.100307.