Research on the Restoration of Sparse Image based on Deep Learning

Luqi He*, Jiaming Gong, Ning Wang, Yunxin Kuang, Canyao Zhao, Yimei Peng
College of Data Science, Guangzhou Huashang College, Guangzhou, Guangdong, 511300, China
**Corresponding Author.*

Abstract: This article mainly explores the sparse image information restoration technology based on deep learning. Through in-depth analysis of image damage and noise characteristics, this paper proposes a novel method aimed at restoring sparse information from images with damage and noise. This paper introduces the characteristics of image damage and noise, and proposes a deep learning based solution to address these issues. By using deep neural networks, the model proposed in this article can effectively learn sparse information in images and recover it from damage and noise. This paper elaborates on the implementation process of the proposed model, including core elements such as the design of loss functions and the selection of optimization algorithms. Specifically, the loss function proposed in this article can effectively measure the accuracy of model predictions and help the model better learn sparse information in images. In addition, this paper provides a comprehensive review and outlook on relevant work, summarizes the shortcomings of current research work, and proposes future research directions. In summary, the sparse image information restoration technology based on deep learning proposed in this article can effectively improve image quality and provide new ideas and directions for research in related fields.

Keywords: Deep Learning; Sparse Information; Loss Function; Optimization Algorithm

1. Introduction

With the continuous development of digital image processing technology, images play an increasingly important role in various application fields. However, due to the imperfections of the imaging system, errors in the transmission process, noise and other factors, images often have some degradation [1]. This will not only reduce the image quality, but also adversely affect the subsequent image processing tasks. Therefore, it is of great practical significance to study how to restore degraded images. Traditional image restoration methods are mainly based on the sparse representation theory. They remove noise and restore the original image by finding the most sparse representation in the image. However, due to the complexity of images and the diversity of noise, these methods often face some challenges. In recent years, the rapid development of deep learning technology provides a new solution for image restoration [2].

In recent years, deep learning technology has made remarkable progress in the field of image restoration and editing. Traditional image restoration methods are mainly based on various image processing techniques, such as denoising, interpolation, repair, etc., such as using Wiener filtering and constrained least squares filtering to restore images [3]. However, these methods often find it difficult to handle complex image damage and noise types. The image restoration method based on depth learning can achieve more effective image restoration by training the depth neural network to learn the characteristics of image damage and noise. In image editing, deep learning technology also shows great potential. For example, the advantages of the GaussNewton method and gradient descent method of neural networks was extracted, which have significant repair effects on blurred images and fast running speed [4]. And the other method like using a mind evolution algorithm to obtain the initial weights and thresholds of a BP neural network before training it [5]. By training deep neural networks to learn the distribution of image features, editing tasks such as super-resolution reconstruction, style conversion, and semantic segmentation of images can be achieved. However, existing deep learning based image restoration and editing methods still have problems such as high computational complexity and insufficient interpretability, which require further research and improvement. This article mainly explores the sparse image information restoration technology based on deep learning. Firstly, it introduces the basic theories of sparse representation and deep learning, and then elaborates on the sparse image information restoration algorithm based on deep learning in detail. Finally, it summarizes and prospects the relevant work.

2. Overview of Sparse Representation Theory

2.1 Sparse Representation Theory

The theory of sparse representation is a new signal representation and image processing method. A review of algorithms based on sparse models shows that, its basic idea is to use linear prediction of a few compressible signals to reconstruct and process signals [6]. This theory does not rely on the traditional Shannon Nyquist sampling law and high-precision reconstruction performance, which proposed by [7], so it is widely used in signal compression, image quality improvement, image classification and other fields. Its representation diagram is shown in Figure 1:

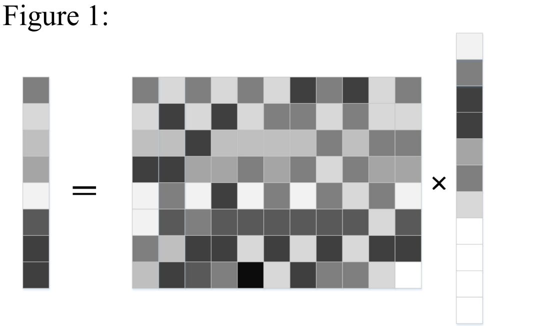

Figure 1. Schematic Diagram of Sparse Representation

The signal y can be showed in a dictionary D as:

$$y \approx D \times c \tag{1}$$

In the above formula, the elements of the coefficient vector C are mostly zeros, with only a few non-zero values. Such a C with sparse pattern can effectively express the basic structure of data and obtain a more compact representation of data by combining it with dictionary D. Sparsity plays an important role in data representation. Dictionary learning can obtain sparse coding. A general optimization model for obtaining sparse coding through dictionary learning can be written as follows:

$$\underset{D_i\{c_i\}}{argmin} \sum_{i=0}^{K} ||y_i - D_{c_i}||_2^2 + \lambda_i ||c_i||_0 \tag{2}$$

$y_i \subset RZ$ represents a group of signals that need to be estimated, c_i corresponds to the exiguous coding of signal y_i, and D represents the dictionary. $\{\lambda_i\}$ represents a set of balance factors used to balance the fidelity of the objective function with the sparsity of the coefficient vector. In the process of dictionary learning, an overcomplete dictionary D is usually used, as the redundancy of the dictionary helps to discover exiguous results. Another option is to use an orthogonal dictionary, which can strike a balance between algorithm performance and computational cost. Research has shown that using orthogonal dictionaries can develop efficient algorithms that can achieve performance comparable to redundant dictionaries on several issues in the field of image restoration.

The sparse representation theory can effectively extract hyperspectral ground information.

$$\min ||\alpha||_0 \quad s.t. x = D\alpha \tag{3}$$

In the above expression, α For sparse coefficient vectors, D is an overcomplete dictionary, x is a reconstructed signal, and norm l_0 is used to count the number of non zero elements in vector α. For a given dictionary, each signal can be linearly expressed using a few primitives in the dictionary. Therefore, sparse representation theory can explore the sparsity of spatial distribution and arrangement of ground objects, and high spectral correlation can further assist in sparse information mining of ground objects, which has its unique advantages. Common sparse representation algorithms include methods based on linear transformations, such as Fast Fourier Transform, Karhunen-LoeveTransform(KLT), etc. And methods based on neural networks, such as autoencoders, deep neural networks, etc. [8].

In summary, sparse representation is a method of representing signals or data using a small number of non-zero elements. To obtain a sparse representation, it is necessary to solve an optimization problem with the goal of minimizing reconstruction errors and limiting the number of non-zero elements. The commonly used algorithms for solving sparse representations include base tracking, matching tracking, and orthogonal matching tracking.

2.2 Orthogonal Matching Pursuit

Orthogonal matching pursuit is an improved matching pursuit algorithm that eliminates selected basis functions in the dictionary through orthogonalization operations at each iteration, thereby ensuring that non zero elements in sparse representations are more sparse and accurate. Specifically, in each iteration, the orthogonal matching tracking selects and updates the basis function that best matches the reconstruction error, and orthogonalizes the already selected basis functions to eliminate their interference in the next iteration. This algorithm has high sparse representation accuracy and computational efficiency.

Orthogonal Matching Pursuit(OMP) [9] is an algorithm for solving sparse representation optimization problems. Its core idea is to search for an optimal solution with sparsity through iterative updates. Unlike base tracking, orthogonal matching tracking selects atoms orthogonal to the current residual in each iteration and updates the direction of the solution.

Specifically, the orthogonal matching pursuit algorithm transforms the optimization problem into a least squares problem, and finds the optimal solution by solving this least squares problem. The basic steps of this algorithm are as follows:

(1) Convert the optimization problem into a least squares problem. In this problem, the objective function is the L1 norm, and the constraint is the square Euclidean distance. Convert these two functions into the standard form of the least squares problem.

(2) Use a least squares algorithm such as Least Residual, Least Absolute Deviation, etc. to solve this least squares problem.

(3) In each iteration, the direction of the solution is updated based on the residual information of the current optimal solution and gradually approaches the optimal solution. Specifically, the orthogonal matching tracking algorithm compares the current residual with all available atoms, selects atoms that are orthogonal to the residual, and updates the direction of the solution.

The advantage of orthogonal matching tracking algorithm is that it can converge to the optimal solution faster. Because each iteration selects atoms that are orthogonal to the current residual and updates the direction of the solution. At the same time, the orthogonal matching tracking algorithm can also be combined with other optimization algorithms to further improve the optimization effect. The pseudo code (Input: sensing matrix $\cdot\phi(M \times N)$, Measurement vector $y(N \times 1)$, Sparse degree K , Output: Reconstructed signal $\hat{x}$) is as follows:

(1) Initialization margin $r_0 = y$, Number of generations selected $n = 0$, Reconstruction signal $x_0 = 0$,Index set $s_0 = 0$

(2) Calculate the inner product of each column of the margin and measurement matrix as: $g^n = \phi^T \cdot r^{n-1}$

(3) Find the element with the highest absolute value in g^n

(4) Update atomic combinations $\phi_{s_n} = \phi_{s_{n-1}} \cup \{\varphi_k\}$, and update new index set $s_n = s_{n-1} \cup \{k\}$

(5) Calculating the approximate solution of signals using the least squares method: $x_n = (\phi_{s_n}^T \phi_{s_n})^{-1} \phi_{s_n}^T y$

(6) Calculate and update margin: $r_n = y - \phi_{x_n}$

(7) Update Iterations: $n - n + 1$, If the iteration conditions are met, then $\hat{x} = x_n$, If the conditions for iteration are not met, return to (2) and continue with generation selection;

3. Overview of Deep Learning Algorithms

Deep learning is a machine learning method based on neural networks, which learns complex patterns and patterns of data by constructing multi-layer neural networks. Deep learning has a wide range of applications in the field of image processing, among which the most commonly used techniques include convolutional neural networks (CNN).

3.1 CNN

Convolutional neural network (CNN) continuously updates the model through model iteration, rule optimization, weight correction, and other methods to enhance the robustness of the model and complete tasks more efficiently. The number of neural networks is proportional to the model effect within a certain range, that is, the more neural networks there are, the more obvious the effect. When used in image processing, neural networks first extract features from the image, which is completed by convolutional layers. After the pooling layer removes the repeatedly recognized features, the image information is transformed into a matrix or vector through fully connected layer feature calculation to obtain the target results of image digital feature extraction and image calculation, for subsequent model iteration and weight optimization. The process is shown in figure 2:

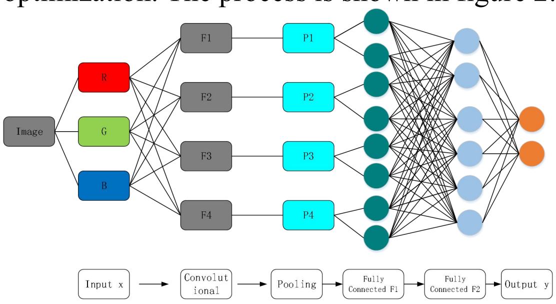

Figure 2. The Diagram of Convolutional Neural Network

CNN model has good fault-tolerant ability, parallel processing ability, and self-learning ability, which can handle problems in complex environmental information, unclear background knowledge, and unclear inference rules. This method combines feature extraction with multi-layer perceptrons to quickly and accurately classify patterns in images. It has been successfully applied in various tasks such as pattern recognition, object identification, and image classification. Unlike traditional methods, it can handle defects and distortions in the sample and has superior generalization abilities. Overall, convolutional neural networks are a powerful tool for image recognition and classification.

3.2 Loss-Function

The loss function(LF) evaluates the accuracy of a model's predictions compared to the actual values. This article uses mean square error (MSE) as the LF. Specifically, we calculate the mean square error between the network output and the target output and use it as the value of the LF. By minimizing the LF, the model parameters can be optimized to improve the image restoration effect

The advantage of MSE is to square the difference values, so larger error values have a greater impact on fit, which helps to capture the prediction error of the model more sensitively.

It should be noted that MSE is greatly affected by outliers, as the difference in the square of outliers is amplified. When using MSE for model evaluation, attention should be paid to the handling of outliers and the robustness of the model.

In summary, MSE is a commonly used fit metric used to evaluate the differences between model predictions and actual observations. A smaller MSE value indicates a better fit of the model.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 \qquad (4)$$

In formula (4), n represents the number of samples, Y_i represents the true value, $\widehat{Y}_i$ representing the predicted value.

MSE is a measure of the average error in a parameter estimation. It is calculated by taking the expected value of the squared difference between the estimated and true values. A lower MSE indicates a better fit between the data and the predictive model. This makes it a useful tool for evaluating the accuracy of a model and the degree of variation in the data.

3.3 Algorithm Optimization

The optimization algorithm is used to adjust the parameters of the model in time of the cultivation approach to minimize the LF. This article uses the Adam optimization algorithm for training. By using the Adam optimization algorithm, better training results can be obtained with fewer iterations.

Usually, we break down the dataset into small batches for training, which is the usually used Mini batch (MB) SGD. The MB SGD algorithm can bring fast learning rate, it isn't consistently able to achieve the true global good point when reaching the optimal point, but constantly hovers around the optimal point. Another drawback of the MB SGD algorithm is that it requires manual selection of a suitable learning late. Higher values may overfit, while lower values may result in slower convergence speed.

The Momentum optimizer is a gradient. In the current iteration step t, the Momentum optimization algorithm can be written as the following formula:

$$v_{dw} = \beta v_{dw} + (1-\beta)dW \quad (5)$$

$$v_{db} = \beta v_{db} + (1-\beta)db \quad (6)$$

$$W = W - \alpha v_{dw} \quad (7)$$

$$b = b - \alpha v_{db} \quad (8)$$

The main concept is to minimize the fluctuations in the gradients by regularizing the parameters in the network.

In formulas (5) and (6), vdw and vdb represent the accumulated momentum of the LF's gradient during the first t-1 iterations. The parameter β is used to control the accumulation rate and is typically set to 0.9. The Momentum optimizer uses a weighted average approach to minimize gradient fluctuations.

In formula (5) and (6), DW and db represent the gradients obtained through backpropagation of the LF. Formulas (7) and (8) are used to update the weight and bias vectors of the network, with α being the learning rate. The Momentum optimization algorithm addresses the issue of large fluctuations in the update amplitude of MB SGD, resulting in faster convergence of the network.

4. Sparse Image Information Restoration Technology Based on Deep Learning

The sparse image information restoration technology based on deep learning mainly combines deep learning technology with sparse representation theory, and achieves the goal of image restoration through the learning and optimization of neural networks. This algorithm learns complex image features and patterns by constructing a deep neural network, and utilizes sparse representation theory to restore the original image. Specifically, this algorithm uses convolutional neural networks (CNN) as the basic framework and adds some additional structures to the network to achieve image restoration purposes.

The advantage of a sparse image information restoration algorithm based on deep neural networks is that it can more accurately restore the features and details of the original image, and has stronger robustness and generalization ability. However, this algorithm requires a large amount of training data and computing resources, and also faces issues such as how to choose appropriate network structures and parameters. The sparse image information restoration algorithm based on deep neural networks is an effective image restoration method, and its implementation principle is as follows:

Establishing a deep neural network model: A convolutional neural network is used to establish the model. As CNN is particularly suitable for processing image data, the CNN model is used as the basis for image restoration. The CNN model consists of multiple convolutional and pooling layers, and the combination of these layers enables CNN to have strong feature extraction capabilities.

Adopting orthogonal matching tracking method: In the deep neural network model, the orthogonal matching tracking algorithm is used to barely represent the image. The OMP algorithm learns the exiguous representation of an image by selecting the atoms most relevant to the current residual in each iteration, until the preset number of iterations is reached or more relevant atoms cannot be found.

Define LF: Use mean square error as the LF. Specifically, for each training sample (y_i, x_i), the LF is defined as:

$$L(W) = \left||y_i - x_i|\right|_2^2$$
$$= ||Ax\_i - y\_i||\_2\text{^}2 \quad (9)$$

In formula (5), A is the measurement matrix, y is the observed image, and x is the original image.

Using the Adam optimization algorithm for training, the specific parameter settings are as follows [10]: $\beta 1 = 0.9, \beta 2 = 0.999$, $\alpha = 0.001$, which means the initial learning rate is 0.001, t is the current number of training steps, $m(t)$ and $v(t)$ are the first and second moment estimates, respectively, $rms(t)$ is the root sign of the estimated value $v(t)$ (abbreviated as "root mean squared" in RMSprop), rms_min is the minimum value maintained during training, $u(t)$ is the small batch normalized version of the unmodified gradient vector $g(t)$.

$$m\_hat(t) = m(t)/(1-\beta 1\text{^}t) \quad (10)$$

$$v\_hat(t) = v(t)/(1-\beta 2\text{^}t) \quad (11)$$

$$W = W - \alpha \cdot m\_hat(t)/(sqrt(v\_hat(t)) + \varepsilon)(12)$$

In formula (8), ε is a small quantity to eliminate mistakes produced by modulo by zero.The Training of the Model: To train the model, we used a training dataset. By minimizing the LF $L(W)$, the Adam optimization algorithm is used to update the weight matrix W, until convergence or preset training steps are reached.

The test of the Model: We used a test dataset to test the trained model, calculated some metrics such as the errors and accuracy that occurred during testing to assess the performance of the model.

The Application of the model: After the model training is completed, this model can be used to restore sparse image information. Using the observed image y as input, the original image x is solved using a deep neural network model and an orthogonal matching tracking algorithm.

5. Conclusions

Sparse image information restoration technology based on deep learning is an emerging image restoration method that combines deep learning technology with sparse representation theory to obtain more accurate image restoration results. This article introduces the basic theories of sparse representation and deep learning, elaborates on the principle and implementation process of sparse image information restoration algorithms based on deep learning, and analyzes and compares their advantages and disadvantages.

Acknowledgments

This project is sponsored by the Youth Academic Program of Guangzhou Huashang University under grant 2023HSQX040. Project Name: Research on Cloud Computing AI Empowering the Development of Audiovisual Industry.

References

[1] Li Junshan, Yang Yawei, Zhang Jiao, et al. 2018, Research progress on degraded image restoration methods. Liquid Crystal and Display, 33 (08): 676-689

[2] Li Junfei, Xu Liming, Wang Yang, et al. 2023, A review of research on citation classification of scientific and technological literature based on deep learning technology. Frontiers of Data and Computing Development, 5 (04): 86-100

[3] Zhang Yanhong, Qin Fengqing, Jiang Li, et al. 2022, Research on Gaussian Blurred Image Restoration with Constrained Least Squares Filter. Volkswagen Technology, 24 (09): 10-13

[4] Shi Ying, Qi Hui. 2019, An Image Restoration Method Based on LMBP Neural Network. Journal of Taiyuan Normal University (Natural Science Edition), 18 (02): 64-67

[5] Mo Siyu, Zhou Qing. 2018, Image Restoration Method Based on MEA-BP Neural Network. Electronic Design Engineering, 26 (21): 185-189

[6] Li Qingquan, Wang Huan. 2019, Overview of Optimization Algorithms Based on Sparse Representation Theory. Surveying and Mapping Geographic Information, 44 (04): 1-9.

[7] Shan G. 2020, The study of signal reconstruction: an investigation independent of Shannon-Nyquist theorems. Science China (Information Sciences), 63(10):279-281.

[8] Tao Wenbin, Qian Yurong, Zhang Yiyang, et al. 2022, A review of deep clustering algorithms based on autoencoders. Computer Engineering and Applications, 58 (18): 16-25

[9] Yue X, Junyu W, Lei Y. 2023, Orthogonal matching pursuit algorithm based on weighted cosine similarity for sound source localization. Physica Scripta, 98(9).

[10] Zhao Xiaohong, Heng Chunni, Du Shujuan. 2020, Medical image convolutional super-resolution reconstruction for residual networks and Adam optimization algorithm. Adhesive, 44 (12): 78-85

