Robust Predictable Discriminative Dictionary Learning for Face Recognition (Postprint)

Authors: Zhang Jian, Mi Jianxun

Date: 2020-09-28T00:00:00+00:00

Abstract

We propose a Predictable Discriminative K-SVD Network (DKSVDN) for face recognition. The model constructs a novel dictionary structure comprising a class label dictionary and a descriptive dictionary to balance discriminative and reconstructive performance. The corresponding sparse coding vector consists of a label coding vector and a descriptive coding vector. To address the low time efficiency of sample sparse coding, we employ a collaborative training approach that leverages predictive neural networks and discriminative dictionary learning models to accelerate predictive sparse coding. Furthermore, a dream-like training method is specifically introduced for DKSVDN to enhance model robustness when training set diversity is insufficient. Comparative experiments on mainstream face datasets demonstrate the superior performance of the proposed model.

Full Text

Preamble

Robust Predictable Discriminative Dictionary Learning for Face Recognition

Zhang Jian a,b, Mi Jianxun a,b†
(a. School of Computer Science & Technology; b. Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts & Telecommunications, Chongqing 400065, China)

Abstract: This paper proposes a predictable discriminative K-SVD network model (DKSVDN) for face recognition. The model constructs a novel dictionary structure comprising a class label dictionary and a descriptive dictionary to balance discriminative and reconstructive performance. The corresponding sparse coding vector consists of a label coding vector and a descriptive coding vector. To address the low time efficiency of sparse coding for samples, a predictive neural network is co-trained with the discriminative dictionary learning model to accelerate sparse code prediction. Additionally, a novel dream-like training method is introduced specifically for DKSVDN to enhance model robustness when training set diversity is insufficient. Comparative experiments on mainstream face datasets demonstrate the model' s superior performance.

Keywords: dictionary learning; sparse representation; face recognition; neural network

0 Introduction

In recent years, demand for biometric-based identification technologies has grown substantially. Face recognition, as a crucial biometric technology, has garnered widespread attention due to its non-contact nature, rapid identification, and ease of deployment, rapidly becoming one of the most popular research directions in computer vision. Face recognition techniques [1-3] initially relied on geometric feature comparison, subsequently introducing data-driven algorithmic models including machine learning-based methods such as Principal Component Analysis and Eigenfaces [4-11], as well as deep learning approaches like Residual Neural Networks and GoogleNet [12-17], continuously improving recognition accuracy. However, real-world environments present numerous uncertain factors, making the challenges faced by face recognition algorithms in practical applications more complex. Issues such as illumination, random noise, facial expressions, pose variations, and occlusions from accessories are often inadequately represented in training data. Insufficient sampling of real-world scenarios severely impacts the performance of data-driven face recognition algorithms, representing a critical problem that must be addressed.

Dictionary learning is a generative model based on compressed sensing theory [18], first proposed by Olshausen and Field [19] in 1996. Prior breakthroughs in biological research on mammalian visual mechanisms revealed that many neurons along the visual pathway exhibit selectivity for specific stimuli in primary and intermediate vision, such as color, texture, orientation, size, and even object images from different viewpoints. Based on these findings, Olshausen and Field designed an energy-based dictionary learning model. Through training on sample images, dictionary learning models demonstrated that natural images have highly compact sparse representation coefficients on overcomplete dictionaries—that is, images can be reconstructed through linear combination of a few dictionary atoms, likely representing the encoding strategy employed by the V1 region of the visual cortex. Additionally, literature [20] proposed a binary alternating optimization structure to solve the model, comprising sparse coding and dictionary construction phases. This structure established the foundation for dictionary learning and has been universally adopted by subsequent algorithms. Aharon et al. [21] proposed the renowned K-SVD method by exploring connections between vector quantization and dictionary learning algorithms, improving the energy-based dictionary learning model by generalizing the K-means algorithm. K-SVD innovatively introduced Singular Value Decomposition (SVD) in the dictionary construction phase to minimize reconstruction error. Unlike existing dictionary learning algorithms limited to batch processing, Online Dictionary Learning (ODL) [22] proposed an online dictionary training method, offering advantages for today' s massive training datasets. ODL employs $_0$ norm constraints for sparsity and uses second-order optimization in the dictionary construction phase, enabling stochastic and mini-batch optimization. The dictionary learning framework achieved tremendous success in image reconstruction [23], drawing many researchers' attention to image classification tasks.

Mairal et al. [24] proposed task-driven dictionary learning, introducing supervised information into traditional dictionary learning for handwritten digit recognition. Discriminative K-SVD (D-KSVD) [5] extended K-SVD for face recognition, training a dedicated classifier simultaneously with dictionary learning and demonstrating excellent performance across multiple face datasets. Jiang et al. [25] proposed LCKSVD (Label Consistent K-SVD) as an improvement over D-KSVD, adding a label regression term for stronger discriminative information. While these methods embed classification information in sparse coding, another approach constructs discriminative dictionaries based on the principle that samples are more similar to those from the same class. Classification relies on reconstruction errors from each class-specific sub-dictionary, assigning samples to the class with minimal error [10, 11, 26]. This approach originates from Sparse Representation-based Classification (SRC) [10], which directly uses all training samples to construct a discriminative dictionary. SRC achieved surprising results in classification tasks but suffers from obvious drawbacks: the dictionary size required for optimal performance is often excessively large. Subsequently, DLSI (structural incoherent dictionary learning method) [27] improved upon SRC for more compact dictionaries.

Researchers have also proposed methods that make both coding vectors and dictionaries discriminative. For instance, FDDL (Fisher discrimination dictionary learning) [28] is a typical method where both components are discriminative. FDDL applies Fisher discriminant criteria to learn a structured dictionary while simultaneously applying them to coding coefficients to maximize between-class distance and minimize within-class distance, achieving excellent discriminative performance. Wang et al. [29] proposed a model (DLSPC) that learns both class-specific and common dictionaries, using intra-class dictionaries to capture the most discriminative details for each class while a common dictionary preserves shared elements. DLSPC also constrains coding such that samples are expressed only on their own class dictionary while suppressing expression on other class dictionaries, using reconstruction errors on each sub-dictionary for classification. However, since dictionary learning models depend on sparse coding, additional computation is required to obtain coding coefficients in practical applications, limiting large-scale deployment. LeCun et al. [30] attempted to use neural networks to predict sparse coding coefficients with promising results. Wang et al. [31] proposed deep $_0$ encoders using deep network structures to approximate $_0$ norm solutions for sparse coding.

This paper proposes a predictable discriminative K-SVD network that embeds label information into the K-SVD algorithm through a novel dictionary structure, enabling sparse coding to directly contain class labels. Unlike other dictionary learning classification methods, this eliminates the need for additional computation to determine class. The dictionary structure separates detailed descriptive information from essential features, providing stronger interpretability and enabling sample generation. Simultaneously, a predictive neural network module is trained to improve computational efficiency for sparse coding. Furthermore, a dream-like training method is introduced that oversamples the training dataset to generate virtual samples for training the predictive network, improving prediction accuracy. Experimental results on mainstream face image databases demonstrate excellent performance in face image recognition.

1 Related Work

In classical sparse representation problems, a sample vector x can be represented as a linear combination of prototype elements: $x = D\alpha$, where D represents the dictionary matrix with each column being a dictionary atom, and α is the sparse representation coefficient vector on dictionary D.

Since dictionary D is overcomplete, the sparse representation problem has infinitely many solutions. To obtain the desired solution, appropriate sparsity constraints must be imposed. The solution should obviously be:

min $\|\alpha\|_0$ subject to $x = D\alpha$ (1)

where $\|\cdot\|_0$ is the $_0$ norm representing the number of non-zero elements in vector α.

Sparse representation solves the optimal coding problem on a specific dictionary, yet excellent sparse coding also requires a suitable dictionary. Pre-designed fixed dictionaries are commonly used, but designing a good preset dictionary often consumes substantial time and effort with significant random factors affecting results. Moreover, to comprehensively cover required features, preset dictionaries are generally set to be over-redundant. To overcome these limitations, dictionary learning algorithms have been proposed to learn the most appropriate dictionary from training data.

In dictionary learning problems, beyond basic sparse representation solving, an optimal dictionary matrix D must be learned from the training set. Denoting the training set as $X = [x_1, x_2, \cdots, x\ ]$ ^{m×n}, the dictionary learning problem can be expressed as:

min_{D,A} ‖X - DA‖_F^2 subject to ‖α_i‖$_0 \leq T_0$ i (4)

or equivalently:

min_{D,A} ‖X - DA‖_F^2 subject to ‖α_i‖$_0 \leq T_0$ i (5)

where $A = [\alpha_1, \alpha_2, \cdots, \alpha\ ]$ ^{k×n} is the sparse coding matrix and $\|\cdot\|$_F is the Frobenius norm for calculating reconstruction error. This paper primarily considers equation (5). Since $_0$ norm problems are non-convex without closed-form solutions, alternating iteration strategies are widely adopted for $_0$ norm-based dictionary learning:

First, fix dictionary D and compute the optimal sparse representation coefficient matrix A based on the current dictionary. Second, fix sparse coefficient matrix A and optimize dictionary D to minimize reconstruction error. Alternate these two steps until convergence.

After training, the final dictionary D remains unchanged during testing for encoding new samples, whose coding solves equations (2) or (3).

K-SVD, as one of the most renowned dictionary learning methods, has received extensive attention. In the sparse coding phase, K-SVD uses matching pursuit or orthogonal matching pursuit to solve coefficient vectors. In the dictionary update phase, the matrix multiplication term DA in equation (5) is decomposed into K rank-1 matrices. During updating, only one matrix is updated at a time while fixing the remaining K-1 matrices, optimizing sequentially to minimize reconstruction error. The reconstruction error term can be rewritten as:

$\|E_k - d_k\alpha^k\|_F^2$ (6)

where α^k represents the k-th row of coefficient matrix A. E_k denotes the reconstruction error when expressing samples with the remaining dictionary atoms after removing the k-th atom. Using SVD decomposition of the error matrix $E_k = U\Delta V^T$, the first principal component (first column of U) is taken as dictionary atom d_k, with the corresponding first row of the singular value matrix Δ as the coefficient. However, to maintain sparsity, K-SVD employs a small trick: the final dictionary atom d_k and sparse coefficient α_k use the first row and first column values from Δ. K-SVD demonstrates good performance across different datasets and signal/image reconstruction tasks.

2 Proposed Algorithm

K-SVD is an $_0$-constrained dictionary learning algorithm whose framework contains two sub-steps: sparse coding update and dictionary update. Through alternating iteration until convergence, the entire optimization process implicitly forms an auto-encoder, where the sparse coding sub-step serves as a non-parametric encoder and the dictionary reconstruction part acts as a linear decoder.

This paper proposes a novel discriminative dictionary learning model while constructing a feed-forward neural network to predict discriminative sparse coding coefficients, improving algorithmic time efficiency.

2.1 Algorithm Description

This paper proposes the Discriminative K-SVD Network model (DKSVDN), which incorporates discriminative information into the original K-SVD and improves sparse coding computational efficiency. Its objective function is expressed as:

min_{D,A,F} ‖Y - Dα‖_F^2 + λ_1‖α‖$_0$ + λ_2‖A - f_F(X)‖_F^2 (7)

where the first term is the discriminative dictionary learning module, the second term is the sparsity constraint, and the third term is the predictive neural network module. X is the input sample matrix, A is the target sparse matrix, and F represents all parameters of the feed-forward neural network.

2.1.1 Discriminative Dictionary Learning Module Sparse representation coefficients possess naturally excellent features that demonstrate good robustness in both image reconstruction and classification tasks. However, reconstruction and discriminative performance are often difficult to balance in traditional dictionary learning models. This paper designs a novel sparse coding and dictionary structure that maintains both properties. The model embeds class labels into sparse coding, making the code consist of two parts: label coding and descriptive coding. Correspondingly, the dictionary is also divided into two sub-dictionaries: a label dictionary and a descriptive dictionary.

Assume Y = [y_1, y_2, ···, y] ^{c×n} is the label matrix, where each column y_i represents the label vector of training sample x_i. The label vector is a c-dimensional vector where c is the number of classes. If x_i belongs to class j, then the j-th dimension of y_i takes value 1 while other dimensions are 0. The sparse coding vector is specifically represented as α_i = [y_i; d_i], where y_i is the 0-1 label vector and d_i is the descriptive code. The label vector y_i corresponds to the label sub-dictionary D_l, which contains the same number of atoms as classes. The corresponding descriptive dictionary is D_d. The label dictionary D_l captures the most essential features of each class, typically representing class-specific characteristics at the class center. The descriptive dictionary D_d stores features unrelated to class information—generic features that may appear in any class—for describing detailed image information. Therefore, the dictionary is specifically represented as D = [D_l, D_d], and the discriminative dictionary learning module is defined as:

min_{D,A} ‖X - DA‖_F^2 = ‖X - [D_l, D_d][Y; D]‖_F^2 (8)

2.1.2 Predictive Neural Network Module Given an input sample image x, the predictive neural network module rapidly predicts the sample's sparse coefficients. Unlike traditional sparse coding methods that use iterative optimization —resulting in unpredictable and time-consuming computation during testing— the predictive neural network module employs a parametric model: a neural network. Combined with nonlinear activation functions, neural networks possess powerful approximation capabilities while significantly reducing computation time, providing reasonably accurate sparse representation coefficients within a predictable timeframe for subsequent tasks.

This paper uses a multi-layer feed-forward fully connected network f_F(·) as the predictive module. To obtain sparsity, the activation function adopts a soft-thresholding function. Therefore, the predictive neural network module expression is:

f_F(x) = $\Phi(W_2\Phi(W_1x + b_1) + b_2)$ (9)

where $\Phi(\cdot)$ represents the soft-thresholding activation function.

In summary, the final objective function is:

min_{D,A,F} ‖X - [D_l, D_d][Y; D]‖_F^2 + λ_1‖[Y; D]‖$_0$ + λ_2‖A - f_F(X)‖_F^2 (10)

subject to ‖α_i‖$_0 \leq T_0$ i, where T_0 is a fixed threshold. For classification using sparse coding, the label coding portion is first extracted, and the dimension with the maximum value in the label coding indicates the predicted class of the test image. Notably, for better prediction performance and to avoid overfitting, this paper adopts a co-training approach for the predictive neural network module and discriminative dictionary learning module.

2.2 Algorithm Optimization and Solution

The objective function is clearly non-convex. This paper solves the model using an alternating optimization strategy with the following procedure:

a) Fix dictionary D and use Orthogonal Matching Pursuit (OMP) [32] to search for optimal sparse representation coefficients on the current dictionary. More specifically, the label vector y is known supervisory information and remains fixed, with the optimization target being the optimal descriptive coefficient part d.

b) Use the sparse representation coefficients obtained in step a) as fixed values, and update dictionary atoms individually using the K-SVD algorithm, including both label dictionary D_l and descriptive dictionary D_d. Simultaneously, use the sparse coefficient matrix as target values and the training sample matrix as input to update the predictive neural network module' s parameter set F via stochastic gradient descent (SGD) until convergence or reaching a fixed number of iterations.

Alternate steps a) and b) until convergence or reaching a fixed number of iterations. Additionally, the predictive neural network module' s predictions can serve as initial values (warm start) for optimizing sparse representation to accelerate convergence in step a).

2.3 Dream-like Training Method

Unknown Pattern in Known Class (UPKC) is a widespread and critical challenge in real-world face recognition. UPKC encompasses various facial variations including expression changes, occlusion variations, etc. Unlike random information in broad facial variations, UPKC focuses on regular facial variation patterns that are often predictable and imaginable for the human brain and can be generalized across classes. For example, if a unique pattern exists in one class —such as a person wearing a scarf with a particular pattern—the human brain can easily imagine faces from other classes wearing the same scarf. However, for statistical machine learning algorithms, since this pattern is missing in other classes, the scarf may be learned as a class-specific feature, causing misclassification. Similar patterns include wearing sunglasses, different degrees of smiling, etc. Since collecting all patterns for all individuals is impossible, UPKC is a challenge that must be addressed.

To improve DKSVDN’ s robustness against UPKC, this paper designs a novel dream-like training method. The concept closely resembles human dreaming processes and functions. Literature [33] first introduced a similar idea for training Helmholtz machines, and later for Deep Belief Networks (DBN) [34]. As mentioned, DKSVDN contains a specially structured dictionary composed of class-specific and descriptive dictionaries. Correspondingly, a sample’ s sparse coding on this dictionary is divided into class label vector segments and descriptive vector segments. The dream-like training method fully leverages DKSVDN’ s structural characteristics, reasonably inferring UPKC’ s sparse coding vectors in the sparse coding space and generating virtual training samples using the dictionary to train the predictive neural network module, thereby improving prediction accuracy.

The dream-like training method combined with DKSVDN proceeds as follows:

a) Collect and record samples’ sparse codings on the current dictionary.

b) Split the descriptive coding segments from sparse codings and collect them into a descriptive coding pool.

c) For all class coding vectors, randomly select descriptive coding vectors from the pool for concatenation to generate reasonably inferred sparse coding vectors.

d) Generate virtual samples using the inferred sparse coding vectors combined with the dictionary via equation (1).

e) Use virtual samples and their corresponding inferred sparse coding vectors as a training set to train the predictive neural network.

Dream-like training utilizes the generative dictionary learning model for sample synthesis in the sparse domain. Compared with existing oversampling methods, dream-like training fully considers sample reasonableness, making the virtual sample space distribution closer to the real sample space and effectively improving the predictive neural network' s accuracy. Dream-like training can be inserted after any iteration in the DKSVDN model optimization process.

3 Experiments

3.1 Datasets and Experimental Setup

3.1.1 Dataset Description To verify DKSVDN' s performance across different face recognition scenarios, experiments were conducted on mainstream face databases: the AR face database [35] and the Extended Yale B face database [8].

The AR face database collected facial images of 126 individuals under various conditions, containing over 4,000 images total. The dataset is divided into two subsets photographed at different times. In the AR dataset, each person (i.e., each class) has 26 facial images photographed under different illumination conditions, including 12 images with occlusions (specifically sunglasses and scarf occlusions) and 14 images without occlusion but with expression and pose variations. Compared with Extended Yale B, the AR dataset contains more interference factors beyond illumination, including expression and pose variations and facial occlusions, making it more realistic and challenging.

In our experiments, we used a subset of the AR face database consisting of 3,120 images from 120 individuals. For each class, 13 images were randomly selected as training data and the remaining 13 composed the test set. All images were uniformly cropped to 40$×50*pixelsandnormalizedwithℓ*_{2}$ norm. [Figure 1: see original paper] shows sample face images from the AR dataset.

The Extended Yale B face database consists of 2,414 frontal face images from 38 individuals photographed under varying illumination conditions and expressions. Besides illumination, variations include expression changes. Each individual has approximately 59 to 64 images. For convenience, we selected 31 individuals from Extended Yale B, each with 64 face images (totaling 1,984 images), removing 7 individuals with fewer than 64 images. For each class, 32 images were randomly selected as training data and the remaining 32 composed the test set. All images were uniformly cropped to 40$×50*pixelsandnormalizedwithℓ*_{2}$ norm. [Figure 2: see original paper] shows sample face images from the Extended Yale B dataset.

3.1.2 Parameter Settings In the predictive coding module, a feed-forward neural network serves as the prediction module with a two-layer structure containing 1,500 and 500 neurons respectively. The number of dictionary atoms is 500, with label dictionary atoms corresponding to 100 for the AR database and 31 for the Extended Yale B database. The T_0 value is set to 15.

3.1.3 Classification Scheme The proposed DKSVDN model does not require training an additional classifier; class information can be directly read from the label vector portion of sparse coding. Since the predictive neural network module cannot obtain perfectly binary label vectors, the test sample's class is determined as the class corresponding to the dimension with the maximum value in the expression vector, mathematically expressed as:

class(x) = argmax_j ŷ_j (11)

where ŷ is the predicted label vector and ŷ_j represents the value of the j-th dimension. Using the trained model, sparse coding vectors can be obtained in two ways: directly predicted by the predictive neural network module, or using the predicted sparse coding vector as an initial value for further optimization via the K-SVD algorithm to obtain more precise sparse coding. Both computation methods are evaluated in subsequent experiments.

3.2 Experimental Results

3.2.1 Face Recognition Comparison algorithms include DLSI, LCKSVD, FDDL, FDDL-LCSRC [36], N-PCA [37], and RCSRC [38]. On the AR face dataset, DLSI sets 5 intra-class dictionary atoms per class, with other settings consistent with literature [25]. LCKSVD dictionary size is set to 500 according to the number of classes in the AR database. N-PCA uses a Nearest Neighbor (NN) classifier. Detailed results are shown in . Evidently, with a smaller discriminative dictionary size, DKSVDN' s classification performance significantly outperforms LCKSVD and DLSI. Additionally, time efficiency comparisons show DKSVDN' s prediction time is substantially faster.

On the Extended Yale B face dataset, DLSI sets 15 intra-class dictionary atoms per class, with other settings consistent with literature [25]. LCKSVD dictionary size is set to 496 according to the number of classes in Extended Yale B. N-PCA uses NN classification. presents the experimental results on Extended Yale B, showing excellent performance in both classification accuracy and time efficiency.

3.2.2 Dream-like Training Experiments Given that the dream-like training method targets scenarios with numerous and dispersed patterns, we validated its enhancement effect on DKSVDN using the AR face database. The AR database contains three patterns: frontal face, wearing sunglasses, and wearing scarf. Since the dataset is divided into two sessions, we used part of Session 1 as the training set. Session 1 contains the first 13 images per person (7 frontal faces, 3 sunglasses, 3 scarves), with Session 2 having the same composition. Individuals in Session 1 were divided into three groups: complete pattern group, missing-sunglasses group, and missing-scarf group. The complete group uses all images, the missing-sunglasses group uses frontal faces and scarf images, and the missing-scarf group uses frontal faces and sunglasses images. Other images from Session 1 and all Session 2 images serve as the test set. To illustrate the grouping details clearly, with 100 classes numbered 1-100, the training set composition is shown in .

shows the recognition performance of the predictive neural network module on the test set for DKSVDN models with and without dream-like training. Evidently, dream-like training significantly improves DKSVDN' s recognition rate, particularly for sunglasses and scarf occlusion patterns, achieving substantially better performance than the original DKSVDN model.

4 Conclusion

The predictable discriminative KSVD model DKSVDN employs a dictionary structure combining label and descriptive dictionaries. Correspondingly, sparse coding vectors based on this dictionary contain both label vector segments and descriptive vector segments, enabling rapid classification using this specially structured sparse coding. Simultaneously, the model trains a predictive neural network module for sparse coding prediction, effectively solving the low time efficiency problem of traditional dictionary learning. Furthermore, the dream-like training method effectively enhances DKSVDN' s robustness under insufficient sample diversity. Experimental comparisons on mainstream face datasets demonstrate that the DKSDN model and dream-like training method exhibit excellent performance in complex environments.

References

[1] Wu Jinyi, Zhou Delong. Survey of face recognition [J]. Application Research of Computers, 2009, 26 (9): 3205-3209.

[2] Trigueros D S, Meng L, Hartnett M. Face recognition: From traditional to deep learning methods [J]. arXiv preprint arXiv: 181100116, 2018.

[3] Masi I, Wu Y, Hassner T, et al. Deep face recognition: A survey [C]// 2018 31st SIBGRAPI conference on graphics, patterns and images (SIBGRAPI). IEEE: 471-478.

[4] Kang Lipan, Chen Fangfu, Fan Zizhu. New extended sparse representation for face recognition [J]. Application Research of Computers, 2016, 33 (3): 937-939.

[5] Zhang Q, Li B. Discriminative K-SVD for dictionary learning in face recognition [C]// 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010: 2691-2698.

[6] Shen Tao, Li Keqing, Xia Yu. Single sample face recognition based on robust sparse representation [J]. Application Research of Computers, 2018, 35 (11): 3491-3496.

[7] Zhao Wen, Wu Xiaojun. Robust recognition of discriminative low-rank representation with dictionary learning [J]. Application Research of Computers, 2017,

34 (10): 3157-3161.

[8] Lee K-C, Ho J, Kriegman D J. Acquiring linear subspaces for face recognition under variable lighting [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27: 684-698.

[9] Turk M, Pentland A. Face recognition using eigenfaces [C]// Proceedings 1991 IEEE computer society conference on computer vision and pattern recognition. 586-591.

[10] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2009, 31: 210-227.

[11] Mi J-X, Sun Y, Lu J. Robust supervised sparse representation for face recognition [J]. Cognitive Systems Research, 2020.

[12] Neto J B C, Marana A N, Ferrari C, et al. Deep Learning from 3DLBP Descriptors for Depth Image Based Face Recognition [C]// 2019 International Conference on Biometrics (ICB). IEEE, 2019: 1-7.

[13] Prasad P S, Pathak R, Gunjan V K, et al. Deep learning based representation for face recognition [M]// ICCCE 2019. Springer, Singapore, 2020: 419-424.

[14] Bashbaghi S, Granger E, Sabourin R, et al. Deep learning architectures for face recognition in video surveillance [M]// Deep learning in object detection and recognition. City: Springer, 2019: 133-154.

[15] Goswami G, Agarwal A, Ratha N, et al. Detecting and mitigating adversarial perturbations for robust face recognition [J]. International Journal of Computer Vision, 2019, 127 (6-7): 719-742.

[16] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.

[17] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[18] Baraniuk R G. Compressive sensing [lecture notes] [J]. IEEE signal processing magazine, 2007, 24: 118-121.

[19] Olshausen B A, Field D J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images [J]. Nature, 1996, 381: 607-609.

[20] Olshausen B A, Field D J. Sparse coding with an overcomplete basis set: A strategy employed by v1? [J]. Vision research, 1997, 37: 3311-3325.

[21] Aharon M, Elad M, Bruckstein A. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Trans on Signal Processing, 2006, 54: 4311-4322.

[22] Mairal J, Bach F, Ponce J, et al. Online dictionary learning for sparse coding [C]// Proceedings of the 26th annual international conference on machine learning. 2009: 689-696.

[23] Ni Hao, Ruan Ruolin, Liu Fanghua. Image super-resolution based on online dictionary learning with two regularization parameters [J]. Application Research of Computers, 2016, 33 (3): 277-281.

[24] Mairal J, Bach F, Ponce J. Task-driven dictionary learning [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2012, 34 (4): 791-804.

[25] Jiang Z, Lin Z, Davis L S. Label consistent k-svd: Learning a discriminative dictionary for recognition [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2013, 35: 2651-2664.

[26] Mi J-X, Fu Q, Li W. Adaptive class preserving representation for image classification [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 7427-7435.

[27] Ramirez I, Sprechmann P, Sapiro G. Classification and clustering via dictionary learning with structured incoherence and shared features [C]// 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010: 3501-3508.

[28] Yang M, Zhang L, Feng X, et al. Fisher discrimination dictionary learning for sparse representation [C]// 2011 International Conference on Computer Vision. IEEE, 2011: 543-550.

[29] Wang D, Kong S. A classification-oriented dictionary learning model: Explicitly learning the particularity and commonality across categories [J]. Pattern Recognition, 2014, 47 (2): 885-898.

[30] Gregor K, LeCun Y. Learning fast approximations of sparse coding [C]// Proceedings of the 27th International Conference on International Conference on Machine Learning. 399-406.

[31] Wang Z, Ling Q, Huang T S. Learning deep 0 encoders [C]// Thirtieth AAAI Conference on Artificial Intelligence. 2016.

[32] Pati Y C, Rezaiifar R, Krishnaprasad P S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition [C]// Proceedings of 27th Asilomar conference on signals, systems and computers. IEEE: 1993: 40-44.

[33] Dayan P, Hinton G E, Neal R M, et al. The helmholtz machine [J]. Neural computation, 1995, 7: 889-904.

[34] Hinton G E, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets [J]. Neural computation, 2006, 18 (7): 1527-1554.

[35] Martinez A M. The ar face database [J]. CVC Technical Report24, 1998.

[36] Wang Wei, Zhu Zongjiu, Lu Jun. Sparse Representation for Face Recognition Based on Dictionary Learning and Locality Constraint [J]. Computer Knowledge and Technology, 2018, 14 (5): 3.

[37] Mi J-X, Zhang Y-N, Lai Z, et al. Principal component analysis based on nuclear norm minimization [J]. Neural Networks, 2019, 118: 1-16.

[38] Mi Jianxun, Lin Zhikai. Sparse representation classification via L2-norm based reconstruction sample constraint for face recognition [J]. Application Research of Computers, 2020, 37 (4): 5.

Note: Figure translations are in progress. See original paper for figures.