

# Research on Mushroom Image Classification Algorithm Based on Deep Sparse Dictionary Learning

Xiyang Guo

School of Electrical Engineering, Southwest Minzu University, Chengdu, 610041, China

---

**Abstract:** The traditional mushroom feature extraction method has low classification efficiency and unsatisfactory effect. Dictionary learning is widely used in image classification. However, the previous work is to learn dictionaries in the original space, which limits the performance of sparse representation classification. In order to solve the problem of spatial redundancy in traditional convolutional neural networks and the weak performance of deep learning in small samples, an improved dictionary learning algorithm, Deep Sparse Dictionary learning (DSDL), is proposed. The input to DSDL is not a matrix gathered from the original grayscale image or a hand-created feature, but rather a relatively deeper feature extraction via a stack autoencoder. Then, a structured dictionary is designed to reconstruct the deep features according to different categories of distinguishing features. In addition, it is necessary to learn the associated structured projection sparse dictionary to ensure that the decoder updates in the direction of the deconvolution operator error is minimal. By utilizing sparse dictionary learning loss functions and autoencoder loss functions, DSDL can simultaneously learn deep latent features and corresponding dictionary pairs. In the testing phase of DSDL, the minimum errors of deep feature and structured projection components for different classes can be directly represented by basic matrix multiplication operations. Experimental results show that the proposed method achieves a good classification effect on mushroom images, which shows the effectiveness of the method.

**Keywords:** Dictionary learning; Sparse representation; Autoencoder; Feature extraction.

---

## 1. Introduction

Mushroom poisoning incidents emerge in society, people pay more and more attention to mushroom food safety, so the accurate identification and classification of mushrooms is very important. Artificial recognition is limited by experience, many mushrooms are very similar in appearance, and automatic recognition of mushroom species with the help of computer image recognition technology has become a research hotspot.

At present, most of the problems of mushroom image classification are solved by deep learning methods, such as neural networks, autoencoders, convolutional neural networks, etc. Convolutional neural network (CNN) has achieved great success in computer image recognition, and its recognition efficiency is constantly improving. However, the feature maps generated by traditional convolutional neural networks still have a large amount of redundancy in spatial dimension, and the processing efficiency is not efficient [1].

As an important machine learning model, dictionary learning aims at sparse representation of the original signal, and the resulting sparse representation is very important to the original data

Strong ability of representation [2]. Dictionary learning can compress most of the redundant information in the original data, and a suitable data dictionary can effectively reduce the redundancy of data representation and improve the discrimination of data, so as to obtain information with certain use value.

By combining a dictionary learning model with a deep neural network, the deep underlying representation of the original image data becomes an input to the dictionary learning model, contributing to more robust and powerful feature descriptions. In 2015, Shen.L et al. proposed a multilevel discriminant dictionary learning algorithm (ML-DDL)[3], and achieved good results in large-scale image

classification. In a paper published in 2016, scholar S.Bahrapour et al. proposed a dictionary learning algorithm based on multimode task-driven [4], which is used in multi-view face recognition, multi-view action recognition and other fields. In the same year, K.Wang et al. proposed a dictionary learning method driven by convolutional neural networks to detect images [5], and enhanced the dictionary's ability to recognize discriminant images mainly by using the depth information of feature mapping. Zhang.H et al. proposed a nonlinear dictionary learning algorithm based on deep neural networks [6], and discussed the improvement of deep neural network features on dictionary learning performance. S. ariyal et al. [7] proposed the deep dictionary learning (DDL) algorithm in 2016 and applied it in tasks such as handwriting font recognition. In 2017, M. lad et al proposed the convolutional dictionary learning algorithm [8], which achieved good results in image super-resolution reconstruction and texture separation. V.shinghal applied robust discriminant deep dictionary learning algorithm to hyperspectral image classification in 2017. In 2017, J.Y. ang and M.H. ang combined the three models of attention model, CRF and dictionary learning and proposed a dictionary learning algorithm for joint conditional random fields [9], applied this method to image visual significance analysis, and used dictionary learning for classification, which could identify and segment interested objects from images. In 2018, V.Shinghal et al. proposed a minimization technique to optimize deep dictionary learning [10]. Y.iu is equivalent to using dictionary learning to motivate deep networks for scene recognition in 2018 [11]. In 2018, M.LAD proposed a multi-layer sparse convolution dictionary learning algorithm (ML-CSC)[12]. J.u et al. proposed a nonlinear dictionary learning algorithm (NDL) and a supervised nonlinear dictionary learning algorithm (SNDL) based on neural networks [13], and SNDL is mainly applied to image classification tasks. In 2019, Tang et al proposed Deep Microdictionary learning

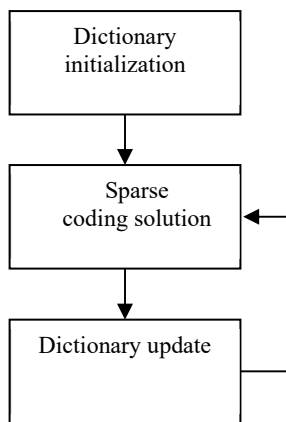
coding Network (DDLCN)[14], which can be used for image representation and classification.

Although the above networks have great innovations in algorithms, they are basically trained on large data sets. In the actual mushroom identification, small batch data need to be classified, and more network parameters will cause the above network can not be fully trained or produce overfitting. The data set used in this paper is small. In order to reduce the influence caused by overfitting and solve the spatial redundancy problem in traditional convolutional neural networks, a mushroom image classification and recognition method under small samples, namely Deep sparse dictionary learning (DSDL), is proposed. This method extracts relatively deeper features through the stack autoencoder, and then designs a structured dictionary according to different categories of discriminant features to reconstruct the deep features.

## 2. Fundamental

### 2.1. Basic theory of sparse dictionary learning

In the traditional sparse representation classification method, the original training sample is generally used as the dictionary, and the sparse representation matrix is approximates to the ideal sparse representation through various constraints, and only the sparse coefficient is required to solve, so the training and solving process is simple. However, due to the simple structure of the dictionary, the representation coefficient can not reach the ideal sparse structure, which is not conducive to the classification of complex images. In order to approach the optimal sparse representation and obtain better classification results, some scholars put forward dictionary learning methods to obtain discriminative dictionaries. Compared with traditional sparse representation classification methods, dictionary learning methods not only learn sparse representation coefficients, but also learn dictionaries. The step of the dictionary learning model is to obtain the initial dictionary, which can be obtained by matrix decomposition, or the training set can be used as the initial dictionary. Then the sparse coefficient is solved, and then the dictionary and sparse coefficient are updated alternately. The algorithm flow is shown in Figure 1.



**Figure 1.** The flow chart of sparse dictionary learning

The following is a brief introduction of classic dictionary learning algorithms such as MOD[15], ODL[16] and K-SVD[17].

#### 1) Optimal direction dictionary learning

The optimal direction algorithm (MOD) was first proposed by Engan in 1999. It selects a direction with the fastest error reduction while using the least square method to update the dictionary and sparse coefficient alternately. Its objective function is as follows:

$$\min \sum_i \|x_i\|_0, \text{ s. t. } \|T - DX\|_F^2 \leq \varepsilon \quad (1)$$

Where T represents the training sample, D is the complete dictionary, X is the sparse coefficient,  $T_0$  is the sparse threshold,  $\|\cdot\|_F$  is the Frobenius norm of the matrix. The dictionary updating process of MOD algorithm is simple, but it faces NP-hard problem and cannot obtain global optimal solution.

#### 2) Online dictionary learning

Online dictionary learning (ODL) solves the non-convex problem of MOD algorithm, that is,  $\ell_1$  norm is used instead of  $\ell_0$  norm in MOD model, and its objective function is as follows:

$$\min_{D, X} \|T - DX\|_F^2 + \lambda \sum_i \|x_i\|_1 \quad (2)$$

Where  $\lambda$  is the weight parameter of the constraint term of the  $\ell_1$  norm. In the solution, the dictionary D and the representation coefficient A are updated by alternate iteration.

#### 3) K-SVD dictionary learning

Although its objective function is the same as the MOD algorithm, the dictionary solving process is different from the MOD algorithm. K-SVD adopts the way of updating the dictionary column by column, that is, fixing other column dictionary atoms and updating only one column atom at the same time, so it has the characteristics of fast convergence. Updates to the current dictionary and presentation coefficients are also done using alternate iterations. K-SVD has been widely used in pattern recognition and image classification because of its fast convergence. In addition, with the increasing demand for classification or recognition tasks, researchers have mainly studied and improved dictionaries and sparse representation coefficients, and many new dictionary learning methods based on the above models have emerged. Therefore, the traditional single-layer dictionary learning method model can be summarized as follows:

$$\min_{D, X} \|T - DX\|_F^2 + \lambda \|X\|_p + \psi(D, X, Y) \quad (3)$$

Where, Y is the label matrix,  $p=0,1,2$ , representing the type of  $\ell_p$  norm constraint, and  $\psi(D, X, Y)$  is the discriminant term. Different single-layer dictionary learning models can be constructed by changing the structure of  $\psi(D, X, Y)$ , so that discriminative dictionaries and sparse representation coefficients adapted to different visual tasks can be learned.

## 2.2. Stack autoencoder

The stack autoencoder, or SAE, is formed by superimposing one AE on another AE, and an autoencoder AE consists of two parts: encoding and decoding. The autoencoder AE maps the input to the hidden layer, while the decoder maps the hidden layer feature representation back to the original data. For a given input vector  $x$ , it tries to learn a hidden layer feature representing  $h_{w,b(x)}$  such that  $h_{w,b(x)} \approx x$  forces the output to equal the input, where W is the network

weight parameter and  $b$  is the bias parameter.

There are several extensions to the basic autoencoder AE architecture, one of which is the stack autoencoder SAE. It has multiple hidden layers, and the hidden layer features of the upper layer are represented as the input of the current layer, which is stacked layer by layer to form a deep network, as shown in Figure 2.

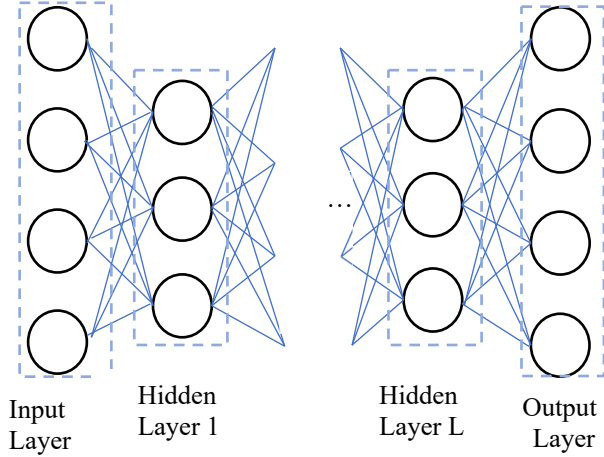


Figure 2. The structure of stack autoencoder

SAE is an unsupervised training architecture, which has two stages, encoding and decoding. In the coding stage, each layer of autoencoder needs to be encoded from front to back along the network, while in the decoding stage, each layer of decoder needs to be decoded from back to front by the network. In the stack self-coding neural network, the greedy training method is used to obtain the network parameters, that is, the original sample is first input to train the first layer of the neural network, so as to obtain the parameters of the first layer network. Then the original sample is converted into a vector  $X$  composed of the activation values of the hidden layer through the first layer network with the parameters determined, and  $X$  is continued as the input of the second layer of the neural network. The parameters of the second layer network are obtained by training. Similarly, the parameters of each subsequent layer of the network are trained in this way. Generally speaking, in order to obtain better training results, fine-tuning steps need to be added to the above training process, that is, after the training process is completed, the backpropagation method is used to optimize and adjust the parameters of each layer, because the training of each layer is to obtain the parameters of the current layer by fixing the parameters of other layers, and the parameters at this time are not optimal for the whole network. In practice, the fine-tuning method is generally used when the parameters are trained close to the convergence state. If the randomly initialized weight parameters are fine-tuned directly, the results will be poor because the parameters only converge to the local optimal. DDL is similar to SAE in coding part but in opposite training direction. DDL reconstructs the original signal  $X=D*Z$  by learning dictionary  $D$  and depth feature  $Z$ . SAE learns the weight parameter  $W$  so that  $W*X=Z$  to obtain the output signal, where  $X$  is the feature representation and  $Z$  is the output signal.

### 2.3. 2.3 Deep convolutional dictionary learning model

By combining the dictionary learning model with the deep neural network, the deep potential representation of the

original image data becomes the input of the dictionary learning model, which solves the spatial redundancy problem of the feature map, and helps to improve the sparse reconstruction of data and more robust feature description under small samples. Therefore, a new dictionary learning method called Deep Sparse Dictionary Learning (DSDL) is proposed.

Stack autoencoders are widely used to extract depth features from data in an unsupervised manner. By designing a certain encoder layer and a decoder layer, the reconstruction error between the decoded output and the original input is minimized, so it is often used for data compression and denoising tasks. The deep sparse dictionary learning method proposed in this paper uses convolution operators to simultaneously learn depth features and corresponding discriminant dictionary pairs. As can be seen from Figure 3, the overall architecture is as follows:

1. Stack autoencoder. The input training image  $Y$  is sent to the encoder layer to form a depth feature  $Z=h(Y;\theta_e)$ , and then the reconstructed  $\hat{Y}=g(\hat{Z};\theta_d)$ , where  $\theta_e$  and  $\theta_d$  are the encoder and decoder parameters, respectively, and  $h$  and  $g$  are the mapping functions implemented by the convolutional and deconvolution layers, respectively.
2. Discriminating dictionary pair learning. The first dictionary,  $P$ , serves as an analysis dictionary to get a relatively low-dimensional representation, and the second dictionary,  $D$ , trains the construction of the image. It is worth noting that they are all structured dictionaries with columns that are consistent with labels.

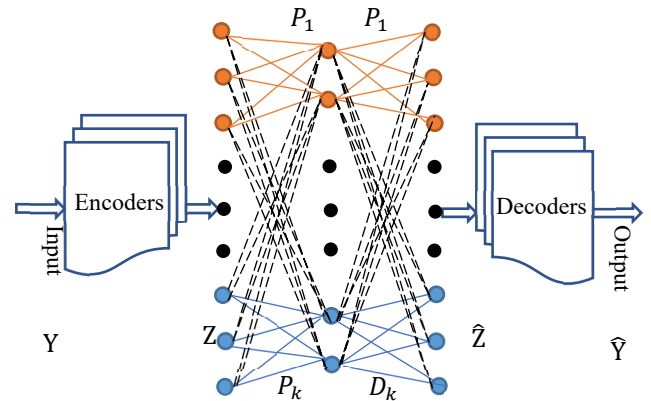


Figure 3. Deep sparse dictionary learning model

The DSDL proposed in this paper aims to train both depth features and sparse dictionaries. Therefore, the objective function of DSDL is shown in equation (4). Where  $Z$  represents a submatrix that does not include the KTH  $Z$ . The first is the autoencoder loss, which minimizes the error between the original training image and the reconstructed image. Sparse reconstruction is performed by passing the depth feature  $Z$  of the training image to the first analysis dictionary  $P$  for potential representation of  $X$ , and subsequently to the discriminant construction dictionary  $D$ .

$$\text{Min}_{\{P,D,X,\theta\}} \sum_{K=1}^C \left\{ \|Y_K - \hat{Y}_K\|_F^2 + \lambda_1 (\|D_K X_K - Z_K\|_F^2) \right. \\ \left. + \|P_K Z_K - X_K\|_F^2 + \lambda_2 \|P_K \hat{Z}_K\|_F^2 \right\} \quad (4)$$

$$\text{s.t. } Z=h(Y;\theta_e), \hat{Y}=g(\hat{Z};\theta_d), \theta = \{\theta_e, \theta_d\},$$

In equation (4), the parameters  $\lambda_1$  and  $\lambda_2$  maintain a

balance between the dictionary representation and the discriminant terms (the non-convex optimization problem in (4) can be solved by alternate direction method), then, the network layer shown in Figure 3 is used to design the relevant variables, and finally, the BP algorithm [18] is used to update the network parameters containing the dictionary pairs. For each  $k$  class, we can use a random sub-gradient descent algorithm to update  $(D_k^{(t)}, P_k^{(t)})$  on the  $t$  iteration until convergence:

$$\begin{aligned} D_k^{(t)} &= D_k^{(t-1)} - \mu \frac{\partial \mathcal{L}}{\partial D_k^{(t-1)}}, \\ P_k^{(t)} &= P_k^{(t-1)} - \mu \frac{\partial \mathcal{L}}{\partial P_k^{(t-1)}}, \end{aligned} \quad (5)$$

Where  $\mu$  is the learning rate (usually set to a smaller value in experiments), such as 0.001, and  $\mathcal{L}$  relative to each class-specific dictionary pair  $D_K$  and  $P_K$  can be calculated as:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial D_K} &= (D_K X_K - Z_K) X_K^T, \\ \frac{\partial \mathcal{L}}{\partial P_K} &= \lambda_1 (P_K Z_K - X_K) Z_K^T + \lambda_2 P_K \bar{Z}_K \bar{Z}_K^T, \end{aligned} \quad (6)$$

The gradient in (6) is then passed to the stack autoencoder and used in standard backpropagation to calculate the network parameter gradient  $\partial \mathcal{L} / \partial \theta$ . Because the columns of both the analysis dictionary and the synthesis dictionary are the same as the columns of the typical class, there is no need to update the unrelated subdictionaries represented by the dashed lines in Figure 3 when updating the corresponding dictionary for class  $k$ .

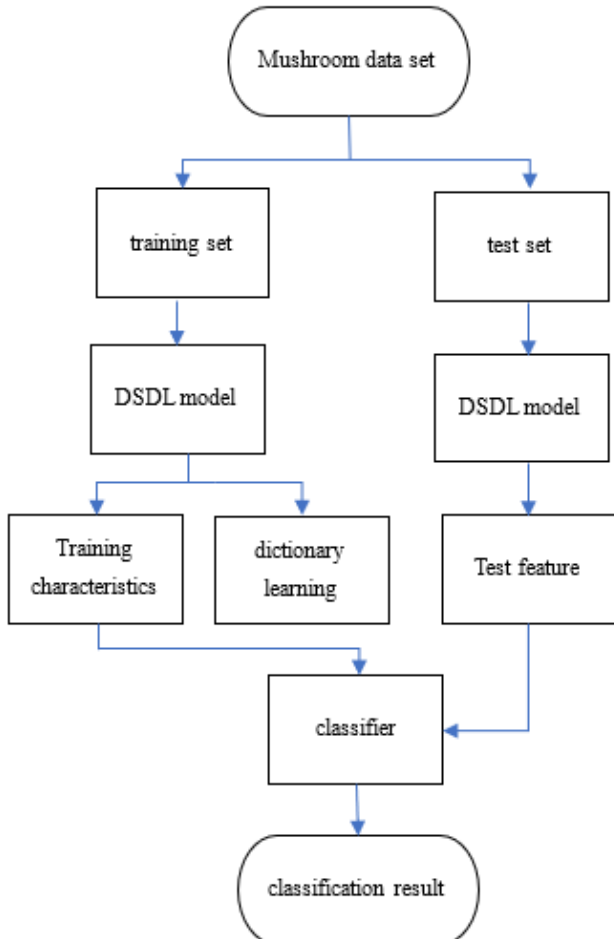


Figure 4. The flow chart of classification of DSDL

As shown in Figure 4, for such a complete image classification system, it is necessary to divide all samples in mushroom image library into training set and test set, use the training set to train DSDL model, obtain the depth sparse features and dictionary of the training set, and then use the dictionary as the network parameters of a new DSDL model and import the test set at the same time to obtain the depth features of the test set. The initialization parameters of the DSDL model used in the training set and the DSDL model used in the test set should be consistent, and then the deep sparse features of the training set should be trained on the classifier. Finally, the deep sparse features of the test set should be input into the trained classifier and the classification label should be predicted to obtain the final classification result.

### 3. Analysis and Discussion

Algorithm running environment: PyCharm2023; Acceleration environment: cuda\_11.7; Operating system: Windows 10; Deep learning framework :Pytorch 2.1.1; Language: Python 3.7.

#### 3.1. Data set description

The data set used is from the Kaggle Public Mushroom Dataset, which has a total of 9533 images grouped into 12 categories. The data set has two main characteristics: 1. There is a serious imbalance in the number of samples, for example, the number of milk mushrooms is as high as 1563, while the number of larch mushrooms, scarlet wet umbrella and milk boletus is less than 400.

2. The image resolution size is inconsistent, including multiple resolutions of the image. Firstly, the image size is unified to  $300 \times 384$  by preprocessing, and then the image is enhanced by random cropping, color distortion, image rotation, etc., and the data set is divided into training set, verification set and test set according to the ratio of 6:2:2.

#### 3.2. Training Strategy

Before training DSDL algorithm, proper pre-training and fine-tuning strategies can minimize the reconstruction error between  $Y$  and  $\hat{Y}$  without getting a trivial all-zero solution. We only use all the training images to train the encoder and decoder. The encoder and decoder are then initialized with the updated parameters. In the fine-tuning stage, the gradient descent method (4) is applied. After the whole network is trained, the reconstruction error of image classification is calculated using the parameters of the encoder layer and the dictionary pair:

$$\text{Label}(y) = \underset{k}{\text{argmin}} \| h(y; \theta_e) - D_K P_K h(y; \theta_e) \|_2 \quad (7)$$

Where  $y \in R^p$  is a vectorized test image, as shown in Table 1.1, the steps of mushroom image classification by DSDL algorithm are as follows:

Table 1. The step of DSDL

DSDL algorithm flow:

a) Input: training image  $Y \in R^{p \times n}$ , the atomic number of each class  $d$ , parameter  $\{\lambda_1, \lambda_2\}$ , test image  $y \in R^{p \times n}$ .

b) By training the autoencoder network, initialize the whole network with the pre-trained parameters.

c) Update dictionary and network parameters by BP algorithm.

d) Using the dictionary for  $\{D, P\}$  and the encoder parameter  $\theta_e$ , the  $k$  reconstruction error is calculated by

formula (7).

- e) Implement the result label of the test image through (7).
- f) Output: Label of the test image: Label (y).

In Table 1, it can be observed that all the training images share an autoencoder network that generates the depth features used for dictionary pair learning. In addition, structured dictionary pairs are trained to represent a specific class of trained images to aid in image classification. The stack autoencoder is introduced into the dictionary pair learning framework, which Bridges the gap between deep feature representation and discriminative dictionary learning. DSDL introduces a common unsupervised autoencoder architecture that does not require a third-party pre-trained model, such as ResNet. Only a limited number of layers are used to explore in an unsupervised manner, resulting in features with better depth than the original features, so that the proposed deep sparse dictionary learning can be carried out in the potential feature space and show its performance.

### 3.3. Experimental results and analysis

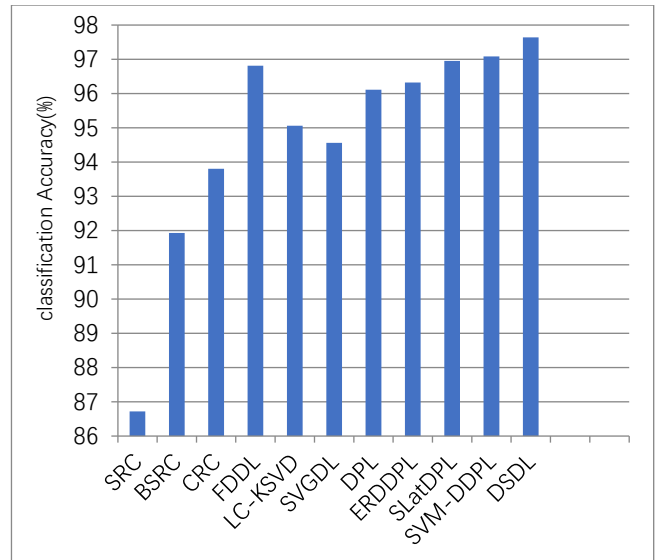
We use a publicly available mushroom image dataset to evaluate the proposed DSDL image classification algorithm. The proposed algorithm is compared with other image classification methods of dictionary learning. That is, classification based on sparse representation (SRC)[19], classification based on Block sparse representation (BSRC)[20] (in which the block sparse coding problem uses the CVX toolbox [21]), classification based on cooperative representation (CRC)[22], and LC-KSVD (LC-KSVD2) In[23]), FDDL[24], Support vector Guided Dictionary learning (SVGDL)[25], dictionary pair learning (DPL)[26], Efficient and robust Discriminant Dictionary pair learning (ERDDPL)[27], SLatDPL[28], and Support vector Machine embedded discriminant Dictionary pair learning (SVM-DDPL)[29]. In the experimental comparison, we trained the correlation algorithm with 20 iterations respectively, updated the learning rate  $\eta = 10^{-3}$  of the correlation variable, and took the average value of the classification results for 10 runs. Parameters  $\lambda_1 \times \lambda_1$  from  $\{0.01, 0.02, 0.03, 0.04, 0.05\} \times \{0.001, 0.005, 0.01, 0.05\}$  grid. In the experiment, ReLU was used as the activation function, and the parameters of the autoencoder, such as the convolution kernel size mentioned in [30,31], were first used, and then fine-tuned on the mushroom dataset used.

The experimental results of DSDL in mushroom data set and other 10 algorithms are listed in Table 2, where the best results are shown in bold. The experimental analysis is as follows:

**Table 2.** Comparison of classification effects of different models

Model	Year	Accuracy
SRC	2009	86.72
BSRC	2011	91.93
CRC	2011	93.80
FDDL	2011	96.81
LC-KSVD	2013	95.06
SVGDL	2014	94.56
DPL	2014	96.11
ERDDPL	2021	96.32
SLatDPL	2021	96.95
SVM-DDPL	2022	97.08
DSDL	2023	97.64

As can be seen from Table 2, the classification accuracy of DSDL is 0.56% higher than that of SVM-DDPL and 0.69% higher than that of SLatDPL. These data fully demonstrate the effective use of stack autoencoder in deep dictionary learning networks. Compared with the traditional dictionary learning algorithm, the classification accuracy of DSDL is improved by 0.81%~10.92%, which further indicates that the depth features obtained by the algorithm are conducive to mushroom image classification.



**Figure 5.** Histogram of classification results of each algorithm on mushroom data set

It can be seen intuitively from Figure 5 that the DSDL algorithm proposed in this paper has excellent performance compared with other algorithms. In summary, from the experimental results, it can be analyzed that:

1. The proposed DSDL algorithm is slightly improved than other methods and achieves higher classification accuracy than traditional dictionary-based image classification methods. Traditional dictionary learning methods may not be able to extract robust representations. In contrast, DSDL learns both dictionary pairs and depth autoencoders, which is more robust for image classification.

2. In general, DSDL Bridges the mutual framework of traditional discriminant dictionary learning and deep learning. The traditional dictionary learning method presupposes that images are linear and can be represented linearly by several images of the same class, while DSDL makes images more linearly correlated by converting images to deep space. Therefore, the algorithm is more robust in image classification based on dictionaries.

## 4. Conclusion

In this paper, we propose a small sample mushroom image classification recognition algorithm, that is, deep sparse dictionary learning, which combines dictionary learning with autoencoder framework. The proposed DSDL enables us to simultaneously learn a reconstruction dictionary and an analysis dictionary of the image's deep features, which contribute to a robust representation. The experimental results show that DSDL algorithm is better than common dictionary learning methods in classifying mushroom data set from Kaggle. In terms of application, software development can be combined with mobile devices and embedded devices in the

future, which will help improve the accuracy of classification of edible mushrooms and poisonous mushrooms, and further reduce the occurrence of accidental ingestion of poisonous mushrooms.

## References

- [1] Chen Y, Kalantidis Y, Li J, et al. Multi-Fiber Networks for Video Recognition [J]. 2018 (1): 364-380.
- [2] Unnikrishnan P, Govindan V K, Kumar S D M . Enhanced sparse representation classifier for text classification[J]. Expert Systems with Application, 2019, 129(SEP.):260-272.
- [3] L.Shen, G.Sun, Q.Huang, S.Wang, Z.Lin, E.Wu.Multi-level discriminative dictionary learning with application to large scale image classification[J]. IEEE Transactions on Image Processing, 2015, 24(10): 3109-3123.
- [4] S.Bahrapour, N.M.Nasrabadi, A.Ray, W.K.Jenkins. Multimodal task-driven dictionary learning for image classification[J]. IEEE transactions on Image Processing,2015, 25(1):24-38.
- [5] Mairal, F. Bach, J. Ponce, G. Sapiro. Online dictionary learning for sparse coding[C]. Proceedings of the 26th annual international conference on machine learning. 2009:689-696.
- [6] H.Zhang, Liu H, Song and R.F.Sun. Nonlinear dictionary learning based deep neural networks[C]. 2016 International Joint Conference on Neural Networks (IJCNN). IEEE,2016: 3771-3776.
- [7] S.Tariyal, H. Aggarwal, A.Majumdar. Greedy deep dictionary learning for hyperspectral image classification [C].2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). IEEE, 2016: 1-4.
- [8] V.Papayan, Y.Romano, J.Sulam, M.Elad. Convolutional dictionary learning via local processing[C]. Proceedings of the IEEE International Conference on Computer Vision. 2017: 5296-5304.
- [9] J.Yang, M.H.Yang. Top-down visual saliency via joint CRF and dictionary learning[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 39(3): 576-588.
- [10] V.Singhal, A.Majumdar. Majorization minimization technique for optimally solving deep dictionary learning[J]. Neural Processing Letters, 2018, 47(3): 799-814.
- [11] Y.Liu, Q.Chen, W.Chen, I. Wassell. Dictionary learning inspired deep network for scene recognition[C]. Thirty-Second AAAI Conference on Artificial Intelligence. 2018.
- [12] J.Sulam, V.Papayan, Y.Romano, M.Elad. Multilayer convolutional sparse modeling and dictionary learning[J]. IEEE Transactions on Signal Processing, 2018,66(15): 4090-4104.
- [13] J. Hu, Y.P. Tan. Nonlinear dictionary learning with application to image classification[J]. Pattern Recognition, 2018, 75:282-291.
- [14] H.Tang, H.Wei, W.Xiao.Deep Micro-Dictionary Learning and Coding Network[C].2019 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2019: 386-395.
- [15] Engan K, Aase S O, Husoy J H. Frame based signal compression using method of optimal directions (MOD)[C]//ISCAS'99. Proceedings of the 1999 IEEE International Symposium on Circuits and Systems VLSI (Cat. No. 99CH36349). IEEE, 1999, 4: 1-4.
- [16] 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. ACM, 2009: 689-696.
- [17] Aharon M, Elad M, Bruckstein A. K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation[J]. IEEE Transactions on Signal Processing, 2006, 54(11):4311-4322.
- [18] Krizhevsky A, Sutskever I, Hinton G(2012)Imagenet classification with deep convolutional neural networks.
- [19] Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y (2009) Robust face recognition via sparse representation. IEEE Trans Pattern Anal Mach Intell 31(2):210-227.
- [20] Elhamifar E, Vidal R (2011) Robust classification using structured sparse representation.
- [21] Michael G, Stephen B (2013) Cvx: Matlab software for disciplined convex programming, version 2.0 beta.
- [22] Zhang, L, Yang M, Feng X (2011) Sparse representation or collaborative representation: which helps face recognition.
- [23] Jiang Z, Lin Z, Davis LS (2013) Label consistent k-svd: learning a discriminative dictionary for recognition. IEEE Trans Pattern Anal Mach Intell 35(11):2651-2664.
- [24] Yang M, Zhang L, Feng X, Zhang D (2011) Fisher discrimination dictionary learning for sparse representation.
- [25] Cai S, Zuo W, Zhang L, Feng X, Wang P (2014) Support vector guided dictionary learning.
- [26] Gu S, Zhang L, Zuo W, Feng X (2014) Projective dictionary pair learning for pattern classification.
- [27] Wang Y, Du H, Zhang Y, Zhang Y (2021) Efficient and robust discriminant dictionary pair learning for pattern classification. Digital Signal Process 118:103227.
- [28] Zhang Z et al (2021) Twin-incoherent self-expressive localityadaptive latent dictionary pair learning for classification. IEEE Trans Neural Netw Learn Syst 323:947-961.
- [29] Dong J, Yang L, Liu C, Cheng W, Wang W (2022) Support vector machine embedding discriminative dictionary pair learning for pattern classification. Neural Netw 155:498-511.
- [30] Ji P, Zhang T, Li H, Salzmann M, Reid I (2017) Deep subspace clustering networks.
- [31] Zhang J, et al. (2019) Self-supervised convolutional subspace clustering network, 5468-5477.