Applied Intelligence
https://doi.org/10.1007/s10489-020-01813-1

Guided autoencoder for dimensionality reduction of pedestrian features

Xuan Li[1] · Tao Zhang[1] · Xin Zhao[1] · Zhengming Yi[1]

Abstract

Autoencoder and other conventional dimensionality reduction algorithms have achieved great success in dimensionality reduction. In this paper, we present an improved autoencoder structure, which was applied it in the field of pedestrian feature dimensionality reduction. The novel method is also verified on Mnist dataset. High-dimensional deep pedestrian features outperform other descriptors while it is challenging for computing capability and memory in existing systems. The dimensionality reduction method we proposed takes advantages of autoencoder and principal component analysis to achieve high efficiency. A novel weight matrix initialization and an improved reconstruction of autoencoder are proposed. Furthermore, by fusing features labeled with the same pedestrian, the proposed structure minimizes the loss after dimensionality reduction. Experimental results demonstrate that our method outperforms traditional dimensionality reduction methods. In the experiment, the pedestrian features were generated by ResNet and Market-1501 data-set. Our method achieves up to 8.834% mAP increment compared to a principal component analysis, when 2048-dimension pedestrian features are reduced to 16-dimension features.

Keywords Autoencoder structure · Principal component analysis · Dimensionality reduction · Feature retrieval

1 Introduction

Data Representation [1–4] has been a longstanding problem in machine learning. Recently, with the rapid development of deep learning, content-based image analysis [30–32] including image classification [5, 6], matching [7, 8] or retrieval [9, 10] shows a better performance via deep learning methods. The deep feature representation obtained through deep learning methods significantly improves the performance of content-based image analysis. But the resulting representation of images is always high-dimensional deep features.

Moreover, a large number of dimensions consume calculation time in the field of image analysis and feature matching. Storing and processing high-dimensional features are expensive. An effective way to compress deep features is deep feature dimensionality reduction, which has been a long-standing research problem in the field of content-based image analysis [11, 12]. Figure 1 shows the pedestrian image of Market-1501 and its 2048-dimension deep feature, the original deep feature is considered to compressed to 32-dimension feature while keeping more information as shown in Fig. 1. Figure 2 shows the example of Mnist dataset and the 16-dimension feature reduced from 784-dimension feature.

Existing dimensionality reduction methods are classified into linear and non-linear methods. Principal Component Analysis (PCA) [13, 14] method and Linear Discriminant Analysis (LDA) [15, 16] method are representatives of the class of linear methods, while non-linear methods including autoencoder [17, 18] and t-distributed Stochastic Neighbor Embedding (t-SNE) [19, 20]. However, when performing dimensionality reduction of high-dimensional deep features using those methods. Some methods are not suitable for addressing the problem, such as t-SNE that applied to the visualization, only available in reducing feature dimensions to 2 or 3 dimensions. some methods could solve the problem would lose most of the original image information. This paper focused on processing dimensionality reduction of

✉ Tao Zhang
zhangtao@tju.edu.cn

Xuan Li
lixuantju@tju.edu.cn

1 School of Electrical and Information Engineering, Tianjin University, No. 92 Weijin Road, Nankai District, Tianjin, China

Published online: 24 July 2020

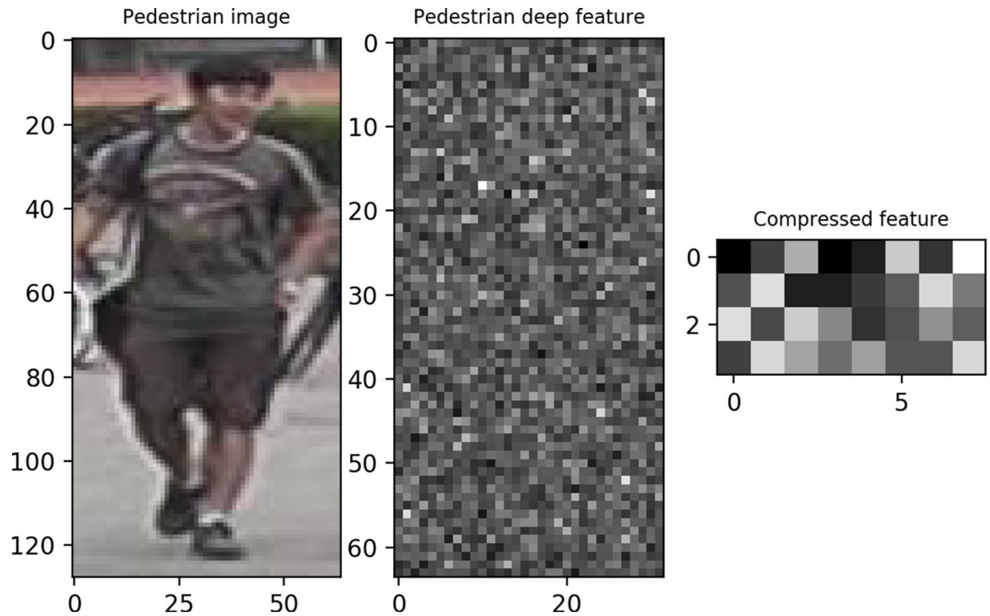

Fig. 1 The example of dimensionality reduction for pedestrian image

deep features and improving the conventional autoencoder structure, according to the principal component analysis of feature matrices, while keeping information about the original features as much as possible.

Guided Autoencoder (GAE) is presented to address the problem of pedestrian features dimensionality reduction. The structure keeps as much information as possible after dimensionality reduction by fusing the deep features of the same person, which performs well in the experiments on the pedestrian and Mnist dataset respectively.

Our two main contributions are as follow:

(i) The structure of GAE is improved via improving weight initialization. A low-dimensional space can be found quickly with good performance through the guidance of the weight initialization.
(ii) The output of the improved structure is reconstructed from the fused features instead of the input. In this case, the features after dimensionality reduction could be seen as fused features.

The relationship between the features of the same person is strengthened by the proposed approach during the dimensionality reduction process.

The rest of the paper is organized as follow. In Section 2, some related works in the area of dimensionality reduction are introduced. In Section 3, the implementation-level of the proposed structure is presented. The proposed method for dimensionality reduction has been discussed in detail. In Section 4, the experimental results are shown and the effectiveness of the proposed method is validated. Finally, in Section 5 we conclude the paper and discuss several research directions for the future.

2 Related work

As a linear dimensionality reduction method, PCA is widely used in image recognition and compression. The greatest variance directions of the data set are found by PCA and

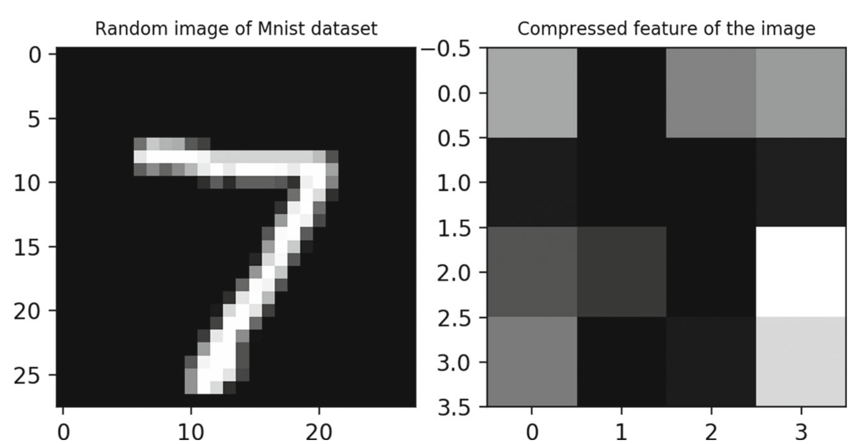

Fig. 2 The example of dimensionality reduction for Mnist dataset

original data is represented according to the principal directions. Tao Li et al. [21] improved 2-Dimensional Principal Component Analysis (2-DPCA) for dimensionality reduction of face recognition and solved the problem that squared F-norm is sensitive to the presence of outliers. A generalized robust metric learning for PCA, namely $\ell_{2,p}$-PCA, was proposed by Wang Q et al. [22] which performs well in image recognition.

As a non-linear dimensionality reduction, the autoencoder algorithm shows its high performance, which contains a symmetrical network structure. Autoencoder is a kind of unsupervised neural network with a symmetrical structure. The output layer of the autoencoder has the same feature dimensionality as the input. It could be used to reduce data space dimensions by minimizing the reconstruction error between its input and output. In 2006, Hinton et al. [23] implemented dimensionality reduction for MNIST data set by using Restricted Boltzmann Machine (RBM) to initialize autoencoders and got art of work result. Based on the denoising autoencoder viewpoint, a generalized autoencoder is proposed by Wei Wang et al. [24], which used a set of instances to reconstruct the original data set. Zhao C. et al. [25] designed a descriptor based on the autoencoder concept and improved the speed of image descriptors significantly without reducing their match ratings noticeably. Petscharnig, S. et al. [26] used deep learning on autoencoder to reduce dimensions for image features and the result proved the method is effective.

Some studies were carried on the relationship between PCA and autoencoder. Baldi and Hornik [27] related principal directions to solutions obtained by training unique autoencoder networks. Sperduti, A. et al. [28]discussed the relationship between liner aotoencoder and principal directions. Jun Ou, Yujian Li. et al. [29] proposed the initialization method called UPSCNNs. It initialized convolutional neural network weights with PCA projection matrices and learns the parameters with unlabeled data.

3 Approach

3.1 Autoencoder

Figure 3 shows the structure of an autoencoder, which consists of two parts: the encoder and the decoder. The encoder part transforms high-dimensional data into low-dimensional data, while the decoder part transforms low-dimensional data into high-dimensional data. The compressed layer value is represented as C_{AE}, which is the low-dimension data required. The whole network learns the identity function $x_{out} = x$ by tuning the weights and bias of each unit, in order to optimize the network parameters.

An encoder and a decoder are represented as (1) and (2) respectively. W and b are weight matrix and bias matrix of encoder respectively. W^T and b' are weight matrix and bias matrix of decoder respectively.

$$E(x) : W(x) + b \tag{1}$$

$$D(x) : W^T(x) + b' \tag{2}$$

Auto-encoder identifies the difference between x and x_{out} as loss function, Mean Absolute Error (MAE) and Mean Square Error (MSE) are widely used loss functions in auto-encoders. MAE reflects the mid-value of data, MSE reflects the average value of data. Equation (3) represents the loss function of auto-encoders.

$$min(f_{loss} : (W^T(Wx + b) + b'), x)) \tag{3}$$

According to what we introduced on auto-encoder above, the output of encoder could be defined by the (4).

$$C_{AE} = W_{AE} * I + b \tag{4}$$

The way of initializing weights of auto-encoder is very important for itself to find the better low-dimensional space. The weights of auto-encoder could be initialized with random or RBM. When the weights initialized with random, the result obtained is always far from the expected value. RBM could calculate the weights and bias according to the latent data structure of input data, so the back propagation could avoid the poor local minimal to some extent. Auto-encoders initialized with RBM could get better desired result.

3.2 GAE

Figure 4 shows the whole structure of GAE. (1) The weight matrix of the novel autoencoder structure is initialized by the guided way, that is, GAE identifies directions where the data variation in the data is maximal. (2) GAE establishes reconstruction set by feature fusion, subsequently, the structure around features of the same pedestrian are learned from the fused feature of the same pedestrian.

3.2.1 Initialization of GAE

The weight matrix of GAE is initialized by the guidance of the greatest variance directions of the data set, which is calculated by means of PCA. The procedure of PCA could be described as below.

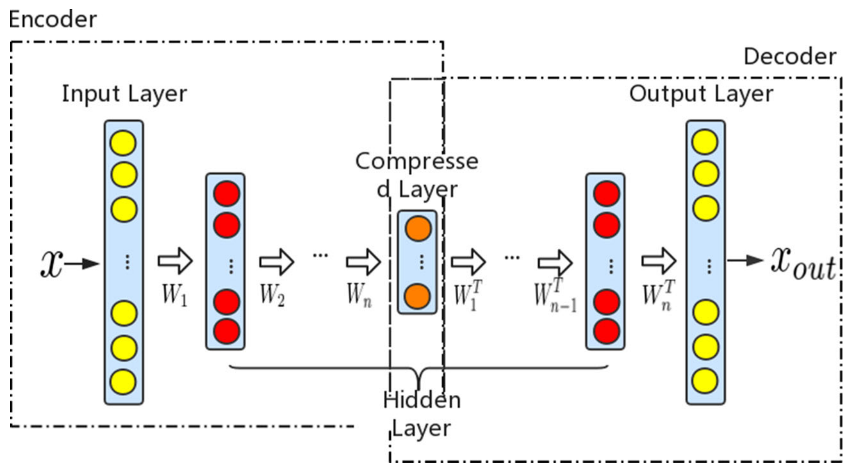

Fig. 3 The structure of an autoencoder

1. Suppose the original data set has m deep features and each feature is a n-dimension vector, the original data set represented as matrix represented with (5).

$$I = \begin{pmatrix} \alpha_{00} & \cdots & \alpha_{0n} \\ \vdots & \ddots & \vdots \\ \alpha_{m0} & \cdots & \alpha_{mn} \end{pmatrix} \tag{5}$$

2. Calculate the mean of each row of the matrix represented with (5) separately, and subtract the corresponding mean from each row of the matrix respectively.

3. Calculate the co-variance matrix of the result.

$$C = \begin{pmatrix} cov(\alpha_0, \alpha_0) & cov(\alpha_0, \alpha_1) & \cdots & cov(\alpha_0, \alpha_n) \\ \vdots & \vdots & \vdots & \vdots \\ cov(\alpha_n, \alpha_0) & cov(\alpha_n, \alpha_1) & \cdots & cov(\alpha_n, \alpha_n) \end{pmatrix} \tag{6}$$

The diagonal line of the matrix are variances of different features, and the off-diagonals are the co-variances. Co-variance represents the degree to which two variables change at the same time.

4. Calculate eigenvalues λ represented with (7) and corresponding eigenvectors represented with (8) of the co-variance matrix C.

$$\lambda = \begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_n \end{pmatrix} \tag{7}$$

$$\omega = \begin{pmatrix} \omega_{00} & \cdots & \omega_{0n} \\ \vdots & \ddots & \vdots \\ \omega_{n0} & \cdots & \omega_{nn} \end{pmatrix} \tag{8}$$

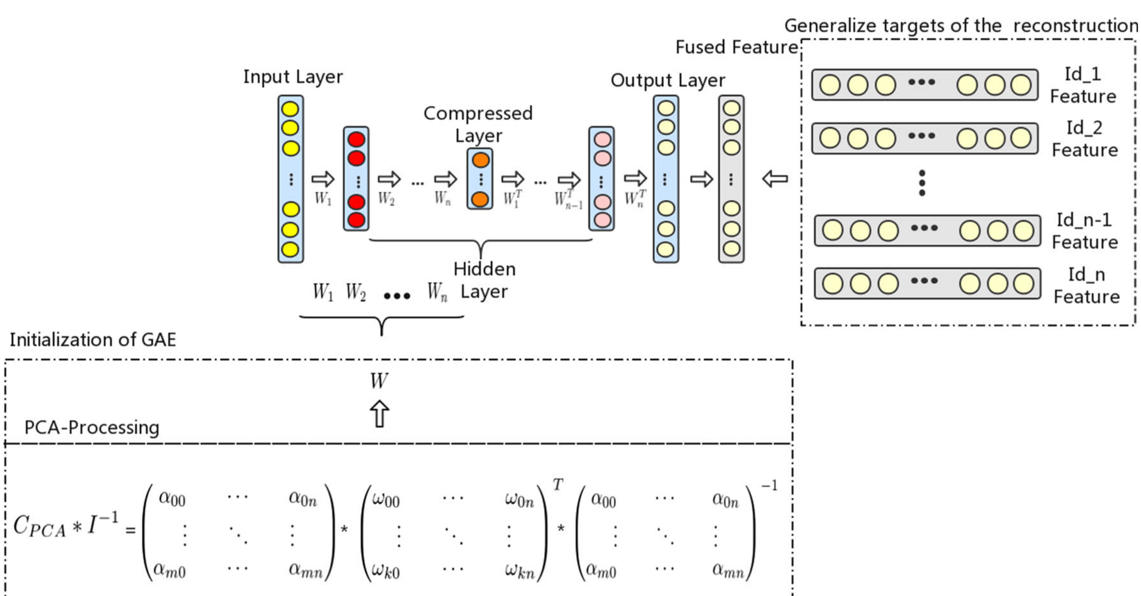

Fig. 4 The structure of GAE

PCA assumes that the directions with the largest variances on the matrix C are the most "important". λ measure the amount of variation retained by each principal component, and it can be used to determine the number of principal components to retain after PCA.

5. Sort eigenvalues λ in descending order and choose the P eigenvectors that correspond to the k largest eigenvalues. Construct the projection matrix P from the selected k eigenvectors.

$$P = \begin{pmatrix} \omega_{00} & \cdots & \omega_{0n} \\ \vdots & \ddots & \vdots \\ \omega_{k0} & \cdots & \omega_{kn} \end{pmatrix} \tag{9}$$

6.Multiply the matrix I by the transpose of the matrix P and get the k-dimension matrix C_{PCA}. C_{PCA} is the low-dimension feature set after dimensionality reduction. Equation (10) shows the multiplication. I is the original data before dimensionality reduction, P^T could be seen as a weight matrix of PCA.

$$C_{PCA} = I * P^T \tag{10}$$

Based on what we analyzed about PCA above, PCA could be considered as a way of matrix transformation, which could be represented with (11).

$$C_{PCA} = I * W_{PCA} \tag{11}$$

Comparing (4) with (11), if the bias matrix of the autoencoder is set to zero matrix and the back propagation of autoencoder is not used, the autoencoder and PCA are both matrix transformation.

If the encoder has only one hidden layer, the compressed layer is the output of the input layer. In this case, the PCA process could be considered as a special situation of encoder and the output of encoder isC_{PCA}. The weight matrix could be represented as (12).

$$W_{AE} = C_{PCA} * I^{-1} \tag{12}$$

GAE proposed here uses W_{AE} from PCA to initialize the weight matrix of an autoencoder structure. GAE was guided by maximal variation in features to find a better result. The greatest variance directions of the data set might learn the relationship among different pedestrian more quickly. Figure 5 shows the initialization of GAE. When autoencoder has one more hidden layer, a series of weight matrixes would be obtained in order to initialize the autoencoder. C^1_{PCA} is the compressed data after first dimensionality reduction of by PCA, I^1 is the input of the first dimensionality reduction by PCA, $C^1_{PCA} * I^{1^{-1}}$ is used to initialize the first layer of the autoencoder structure. Our experiments show that the initialization of GAE has a better performance than Random initialization and RBM initialization involved in pedestrian features and Mnist dataset dimensionality reduction.

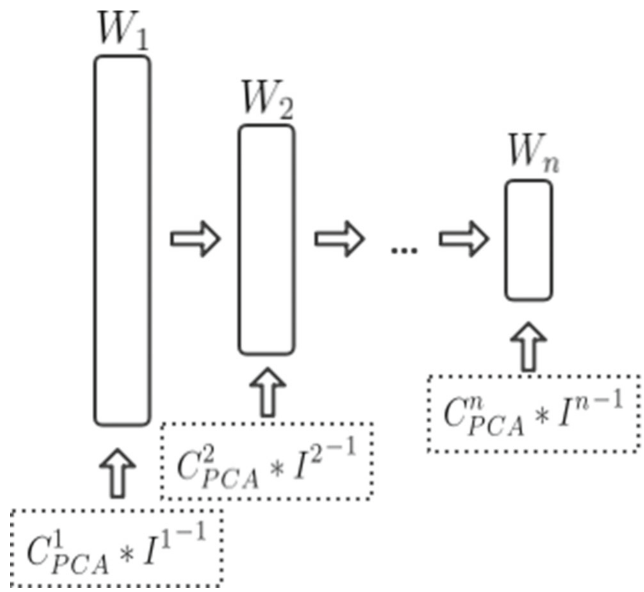

Fig. 5 The initialization of GAE

3.2.2 Reconstruction of GAE

GAE generates a fused feature for each pedestrian separately and gets a fused feature set finally, after which GAE reconstructs features of each pedestrian separately from the corresponding fused feature.

The distance between features of the same pedestrian is expected smaller than the distance between features of the different pedestrian. For better performance, GAE learns the low-dimension data space while fusing the features from the same pedestrian. Each item in the set obtained by GAE represents its counterpart to some extend. In this way, GAE trains the parameters by minimizing the difference between the output and the certain fused feature instead of the original feature.The features of the same pedestrian would be as close as possible during dimensionality reduction. Therefore, GAE fuses features from the same pedestrian while reducing dimensions.

Figure 6 shows how to generate the fused feature of one pedestrian in GAE. First, calculate the mean value of the features, the mean value of features could be seen as a fake point as point 1 in Fig. 6. However, there may be existing some abnormal points colored in green in Fig. 6 which may affect the mean value evaluation of the features. To reduce the interference of extreme points, in the second step, the 3 points nearest to point 1 are found according to the distance metric, the 3 points are colored in bright yellow. Finally, calculate the mean of the bright yellow points and get the

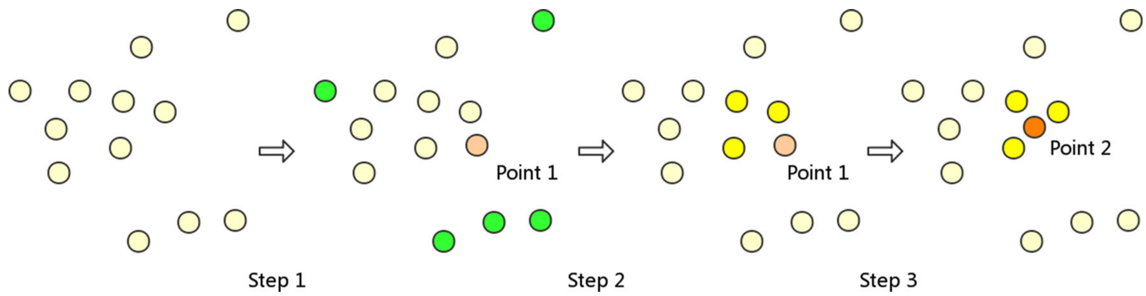

Fig. 6 The generation of fused feautres

new fake point 2 replacing fake point 1 as the final fused feature point.

Figure 7 shows some targets of reconstructions for conventional autoencoders in Mnist dataset. Each item is reconstructed from itself. Figure 8 shows the fused features, which are the targets of reconstruction for GAE. The different items of the same class are reconstructed from the same one which is the fused feature via GAE.

In total, GAE could be summarized as follows: Firstly, GAE initializes weight matrix based on principal component analyzing as a guided way, thus original features could be reduced to low-dimensional features as expected at a faster rate. Secondly, GAE fuses features of each id and trains its parameters by minimizing the reconstruction errors between its output and the fused features, hence, GAE could get a low-dimension and fused features with minimal loss of information.

4 Experiment

Here ResNet and Market-1501 are used to generate the original pedestrian features. The training set and the testing set have 1w features separately, while the query set has 100 features.

In order to ensure the reliability of the experimental results, the experiment was repeated more than 20 times, the training set, testing set, query set are regenerated respectively from the Market-1501 dataset during each experiment. The performance of mean average precision(mAP) is evaluated in the task of pedestrian feature and Mnist dataset retrieval. Table 1 shows the mAP comparison of different dimensionality reduction methods at top1, top5, top10, top20, and top50 respectively.

Table 1 shows that GAE outperforms server other methods for deep feature dimensionality reduction. GAE shows better results than LDA method when 2048-dimension features are reduced to 16-dimension features and 8-dimension features, meaning GAE is more stable than LDA. When deep feature dimensions are reduced to 64 or 32, our model even shows better results than the original features because of feature fusion. VAE shows unsatisfied results which are similar when the features reduced to different dimensions because the latent vector of VAE is a probability distribution learned by encoder and decoder.

Some visualization experiments are adopted in Fig. 9 to compare GAE with server other dimensionality reduction methods. Features of each pedestrian id with dimensionality reduction are colored in the same color. Compared with the results of GAE, the two-dimension features of each pedestrian id produced by PCA and LDA are discrete. Most of two-dimension features coded by Random initialization are similar in in both dimensions, while most of two-dimension features coded by RBM initialization autoencoder are similar a

Fig. 7 Random targets of reconstruction for conventional autoencoder

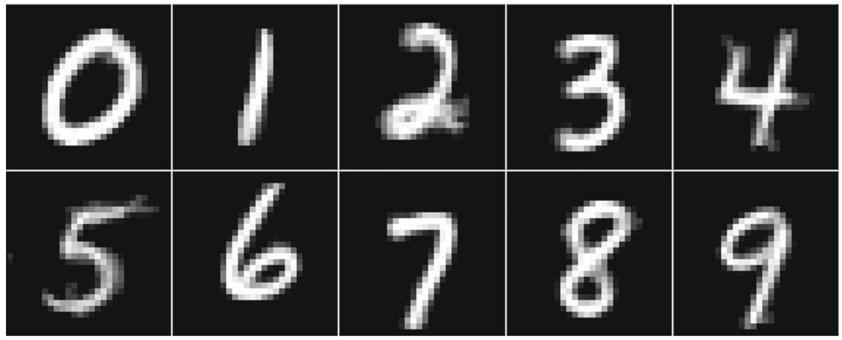

Fig. 8 Fused targets of reconstruction for GAE

certain dimension. For the dimensionality reduction of deep pedestrian features, the features of different pedestrians are similar and the classes of pedestrians is huge, it would be infeasible for autoencoder to learn the differences among classes for pedestrian dimensionality reduction. If it shows similar characters in different dimensions, which is different from a dataset like Mnist, the greatest variance directions of the data set might learn the subtle difference among different pedestrians. GAE is guided by principal variation of dataset to initialize the weight matrix. In this case, GAE

Table 1 Results of different dimensionality reduction methods for pedestrian features

	Top1	Top5	Top10	Top20	Top50
Original feature (2048-dim)	0.0956	0.4368	0.6866	0.8351	0.8844
The dimension of deep features is reduced to 64					
PCA	0.0956	0.4295	0.6751	0.8257	0.8758
kernel-PCA	0.0956	0.4299	0.6761	0.8254	0.8758
LDA	0.0956	0.4447	0.7349	0.9089	0.9484
autoencoder with Random initialization	0.0956	0.2269	0.2737	0.2989	0.3359
VAE	0.0956	0.096	0.0962	0.0963	0.0967
GAE	**0.0956**	**0.4355**	**0.6887**	**0.8696**	**0.9146**
The dimension of deep features is reduced to 32					
PCA	0.0956	0.4214	0.6597	0.8089	0.86
kernel-PCA	0.0956	0.4224	0.6598	0.8074	0.8595
LDA	0.0956	0.4293	0.6861	0.8476	0.8926
autoencoder with Random initialization	0.0956	0.1664	0.1931	0.2057	0.2195
VAE	0.0956	0.096	0.0962	0.0963	0.0967
GAE	**0.0956**	**0.4246**	**0.6808**	**0.8436**	**0.8908**
The dimension of deep features is reduced to 16					
PCA	0.0956	0.383	0.5523	0.722	0.7769
kernel-PCA	0.0956	0.3828	0.5811	0.72	0.7763
LDA	0.0956	0.3891	0.5954	0.7532	0.8032
autoencoder with Random initialization	0.0956	0.1107	0.1146	0.117	0.1236
VAE	0.0956	0.0967	0.0971	0.0975	0.0977
GAE	**0.0956**	**0.4228**	**0.6381**	**0.7907**	**0.8427**
The dimension of deep features is reduced to 8					
PCA	0.0956	0.3074	0.4426	0.5458	0.6096
kernel-PCA	0.0956	0.3007	0.4415	0.5445	0.6073
LDA	0.0956	0.2847	0.4017	0.4934	0.5528
autoencoder with Random initialization	0.0956	0.1046	0.1082	0.11	0.1168
VAE	0.0956	0.0958	0.0965	0.0966	0.0971
GAE	**0.0956**	**0.3239**	**0.4793**	**0.5939**	**0.6589**

The bold entries are the results of the GAE which is presented in this paper

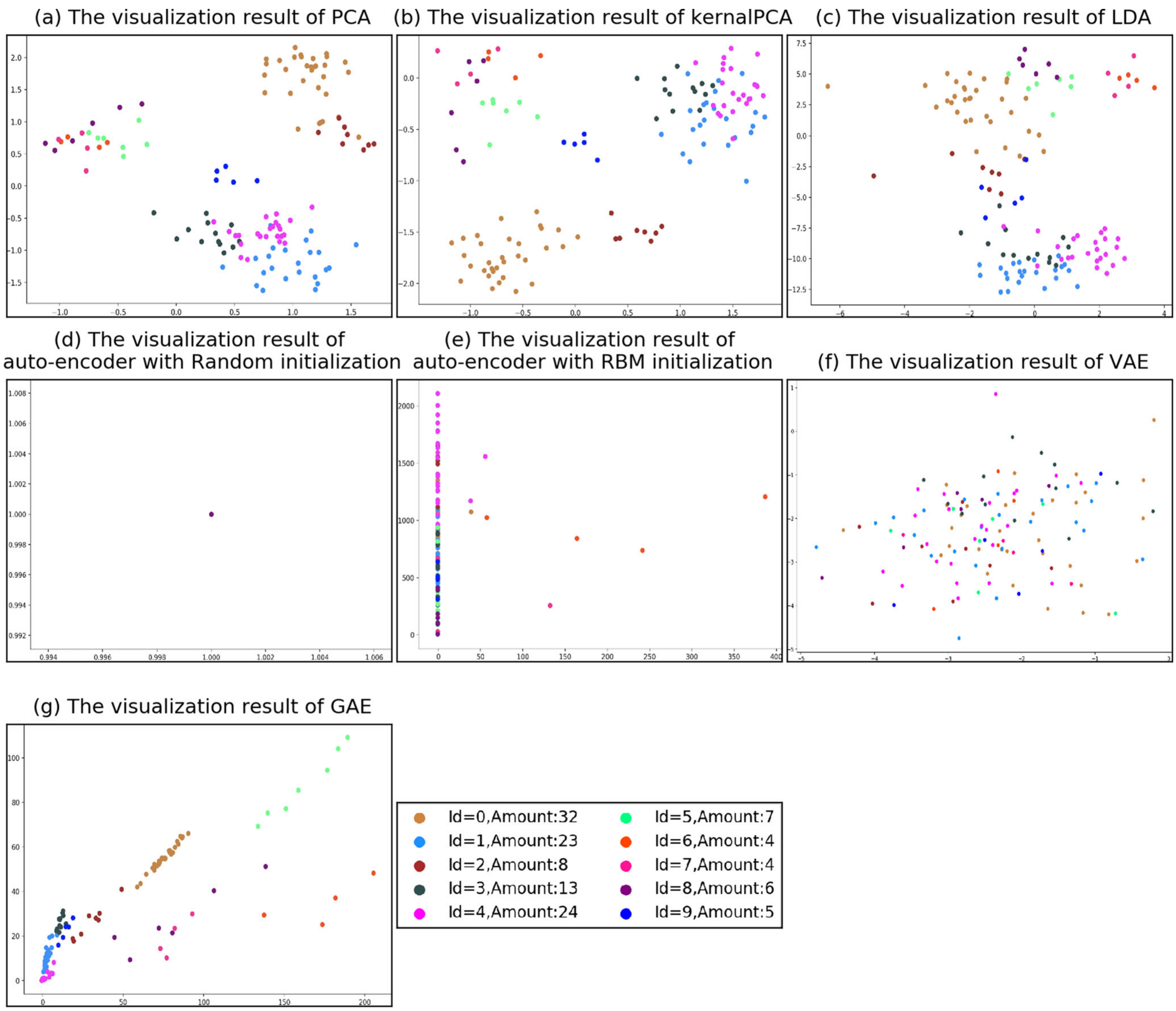

Fig. 9 The visualization results of GAE and server other dimensionality reduction methods

could learn the connection among different dimensions of pedestrian features, and could also tell the slight differences of each pedestrian.

Figure 10 shows the comparision of mAP at top2000 for the Mnist dataset retrieval when the features are reduced to 8-dimension. The original features could not be reduced to more than 9-dimension by LDA, because the dimensions after reduction by LDA should smaller than the classes of the dataset. GAE could calculate a convergence faster and more stable via the novel initialization and the fused objects of reconstruction. Compared to other methods with good performance, GAE could avoid the locally optimal solution effectively and get better performance.

Figures 9 and 10 demonstrate that different weight initialization methods affect the performance of autoencoders when autoencoders were applied to dimensionality reduction. It could be inferred that the network structure of autoencoder acquire a locally optimal solution when the loss is not equal to zero. The weights initialized in different ways might provide different directions for dimensionality reduction.

As we can see, the proposed GAE outperforms the other four models. GAE reduces the original feature dimensionalities with minimal information loss, and produces a better visualization of codes after dimensionality reduction.

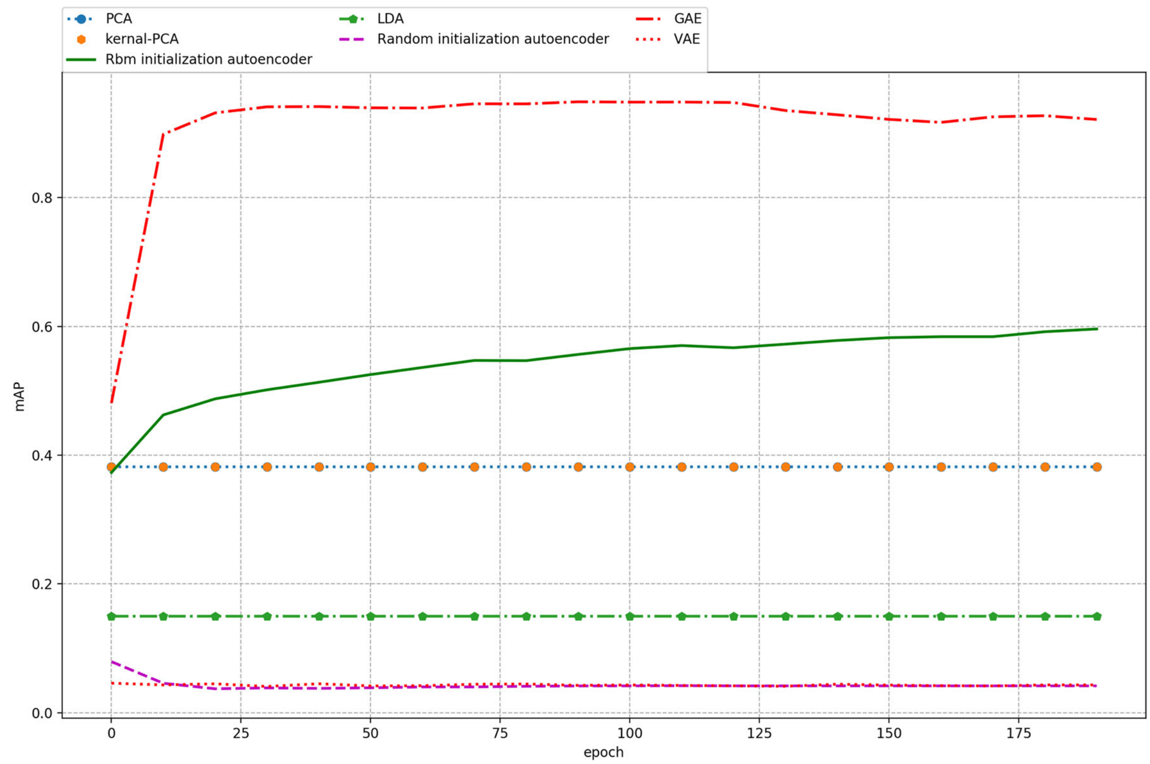

Fig. 10 The comparision of mAP at top2000 for the Mnist dataset retrieval when the features are reduced to 8-dimension. The dimensionality reduction methods are Random initialization autoencoder, RBM initialization autoencoder, VAE, PCA, LDA, kernel-PCA and GAE respectively

5 Conclusion and future work

During dimensionality reduction, autoencoder might get a poor local optimal solution as the final result owing to large initial weights. In the field of deep pedestrian feature dimensionality reduction, the features are high-dimension data and the structure around features is similarity, thus it is infeasible to train autoencoder with many hidden layers and large weights.

In this paper, GAE is proposed to address the problem of dimensionality reduction for pedestrian features with minimal loss of information. At first, GAE is initialized with the principal components of all features. To fit the dimensionality reduction of deep feature space, the weight matrix is initialized according to maximal variation in the space. Then GAE is trained by minimizing the reconstruction error between the output and the reconstruction set. The input of GAE is reconstructed by fused features and the low-dimensional features of the same pedestrian are closer than those of conventional dimensionality reduction methods. Experimental results demonstrate that GAE presented outperforms conventional algorithms in the field of pedestrian features dimensionality reduction..

There are several future research directions along with this topic. First, labels are required to fuse features in GAE. An unsupervised or semi-supervised structure might be proposed to process large volumes of data efficiently. Secondly, GAE is used to reconstruct according to features after fusion. In the future, fused features could be generated by clustering features of the same pedestrian for better feature fusion.

References

1. Zhang Y, Zhang Z, Zhang Z et al (2020) Deep self-representative concept factorization network for representation learning[C]. In: Proceedings of the 2020 SIAM international conference on data mining. Society for Industrial and Applied Mathematics, pp 361–369
2. Zhang Z, Li F, Zhao M et al (2017) Robust neighborhood preserving projection by nuclear/l2, 1-norm regularization for image feature extraction[J]. IEEE Trans Image Process 26(4):1607–1622
3. Zhang Y, Zhang Z, Li S et al (2018) Unsupervised nonnegative adaptive feature extraction for data representation[J]. IEEE Trans Knowl Data Eng 31(12):2423–2440
4. Ye Q, Li Z, Fu L et al (2019) Nonpeaked discriminant analysis for data representation[J]. IEEE Trans Neural Netw Learning Sys 30(12):3818–3832
5. Mahmood A, Bennamoun M, An S et al (2019) Deep image representations for coral image classification. IEEE J Ocean Eng 44(1):121–131
6. Li Chen, Wang Kai, Xu N (2017) A survey for the applications of content-based microscopic image analysis in microorganism classification domains. Artif Intell Rev 51:1–70

7. Yi Hou, Hong Z, Zhou S (2017) BoCNF: efficient image matching with bag of ConvNet features for scalable and robust visual place recognition. Auton Robot 42:1–17
8. Yang Wei, Zhong Liming et al (2017) Lin predicting CT image from MRI data through feature matching with learned nonlinear local descriptors. IEEE Trans Med Imaging PP:977
9. Tsochatzidis L, Zagoris K et al (2017) Arikidis computer-aided diagnosis of mammographic masses based on a supervised content-based image retrieval approach. Pattern Recogn 71:106–117
10. Lu X, Zheng X, Li X (2017) Latent semantic minimal hashing for image retrieval. IEEE Transactions on Image Processing 26:355–368
11. Carcenac M, Redif S (2015) A highly scalable modular bottleneck neural network for image dimensionality reduction and image transformation. Appl Intell 44:557–610
12. Yang YB, Li YN, Pan LY, Li N, He GN (2013) Image retrieval based on augmented relational graph representation. Appl Intell 38:489–50
13. Hosoya H, Hyvärinen A (2016) Learning visual spatial pooling by strong PCA dimension reduction. Neural Comput 28:1249–1264
14. Wold S, Esbensen K, Geladi P (1987) Principal component analysis. Chemometrics and Intelligent Laboratory Systems 2:37–52
15. Shu X, Lu H (2014) Linear discriminant analysis with spectral regularization. Appl Intell 40:724–731
16. Laohakiat S, Phimoltares S, Lursinsap C (2017) A clustering algorithm for stream data with LDA-based unsupervised localized dimension reduction. Journal 381:104–123
17. Gao G, Liu L, Wang L, Zhang Y (2019) Fashion clothes matching scheme based on Siamese Network and AutoEncoder. Multimedia Systems 25:593–602
18. Kim M, Lee W, Cho D-H (2017) A novel PAPR reduction scheme for OFDM system based on deep learning. IEEE Commun Lett 22:510–513
19. Van Der Maaten L (2014) Accelerating t-SNE using tree-based algorithms. The Journal of Machine Learning Research 15:3221–3245
20. Zhou H, Wang F, Tao P (2018) T-distributed stochastic neighbor embedding method with the least information loss for macromolecular simulations. Journal of Chemical Theory and Computation 14:5499–5510
21. Li T, Li M, Gao Q, Xie D (2017) F-norm distance metric based robust 2DPCA and face recognition. Neural Netw 94:204–211
22. Wang Q, Gao Q, Gao X, Nie F (2017) $\ell_{2,p}$-Norm based PCA for image recognition. IEEE Transactions on Image Processing 27:1336–1346s
23. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313:504–507
24. Wang W, Huang Y, Wang Y, Wang L (2014) Generalized autoencoder: a neural network framework for dimensionality reduction. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 490–497
25. Zhao C (2016) An autoencoder-based image descriptor for image matching and retrieval
26. Petscharnig S, Lux M, Chatzichristofis S (2017) Dimensionality reduction for image features using deep learning and autoencoders. In: Proceedings of the 15th international workshop on content-based multimedia indexing, pp 1–6
27. Baldi P, Hornik K (1989) Neural networks and principal component analysis: Learning from examples without local minima. Neural Networks 2:53–58
28. Sperduti A et al (2003) Linear autoencoder networks for structured data. International Workshop on Neural-Symbolic Learning and Reasoning
29. Ou J, Li Y, Shen C (2018) Unlabeled PCA-shuffling initialization for convolutional neural networks. Appl Intell 48:4565–4576
30. Zhang Y, Zhang Z, Qin J et al (2018) Semi-supervised local multi-manifold Isomap by linear embedding for feature extraction[J]. Pattern Recogn 76:662–678
31. Xiao Q, Dai J, Luo J et al (2019) Multi-view manifold regularized learning-based method for prioritizing candidate disease miRNAs[j]. Knowl-Based Syst 175:118–129
32. Wang H, Yang Y, Liu B et al (2019) A study of graph-based system for multi-view clustering[J]. Knowl-Based Syst 163:1009–1019

Xuan Li is a PhD in School of Electronic Information Engineering from Tianjin University, Tianjin, China. Her current interests are machine learning, computer vision and evolutionary computation.

Tao Zhang received the M.S. degree in School of Electronic Information Engineering from Tianjin University, Tianjin, China, in 2001, the Ph.D. degree from the same University in 2004. He is currently an associate professor of the School of Electrical and Information Engineering, Tianjin University, China. His current interests include image, video, and acoustic signal processing, auditory model, speech enhancement and hardware/software partitioning.

Xin Zhao received the Ph.D. degree in School of Electrical and Information Engineering from Tianjin University, Tianjin, China, in 2020. He is currently a lecturer of the School of Electrical and Information Engineering, Tianjin University, China. His research interests are evolutionary computation, machine learning and image processing.

Zhengming Yi graduate student at School of Electrical and Information Engineering from Tianjin University, Tianjin, China. His research interest covers computer vision and person reidentification.