

Liang HOU, Xiao-yi LUO, Zi-yang WANG, Jun LIANG, 2020. Representation learning via a semi-supervised stacked distance autoencoder for image classification. *Frontiers of Information Technology & Electronic Engineering*, 21(7):1005-1018. <https://doi.org/10.1631/FITEE.1900116>

# Representation learning via a semi-supervised stacked distance autoencoder for image classification

**Key words:** Autoencoder; Image classification; Semi-supervised learning; Neural network

Corresponding author: Jun LIANG

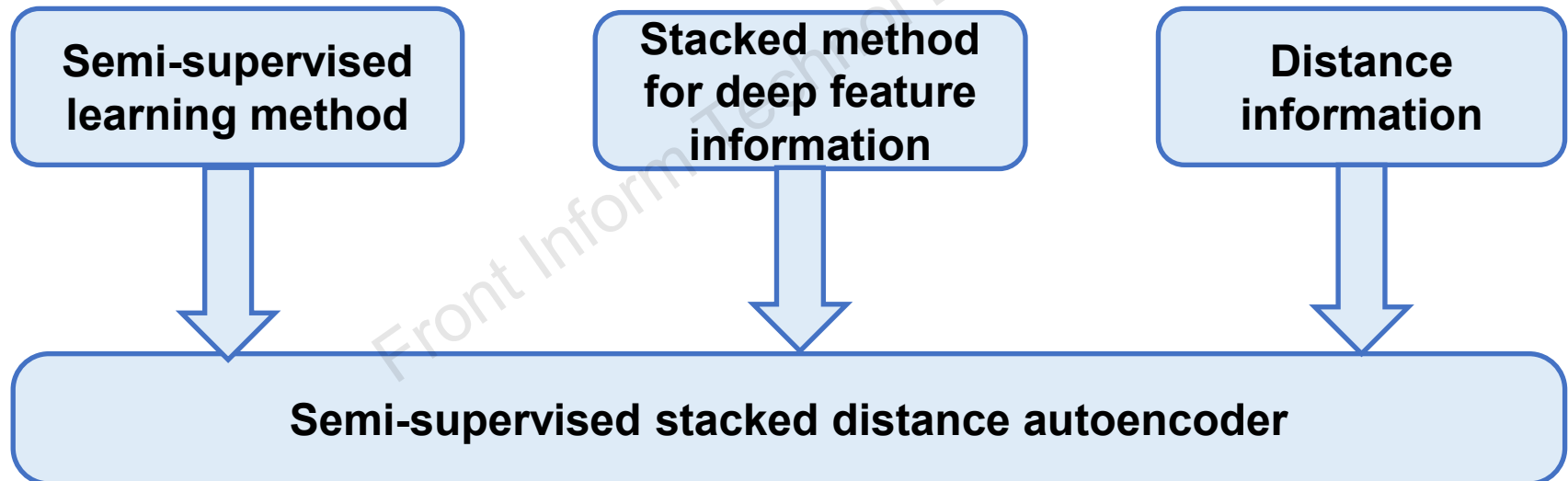
E-mail: [jliang@zju.edu.cn](mailto:jliang@zju.edu.cn)

 ORCID: <https://orcid.org/0000-0003-1115-0824>

# Motivation

---

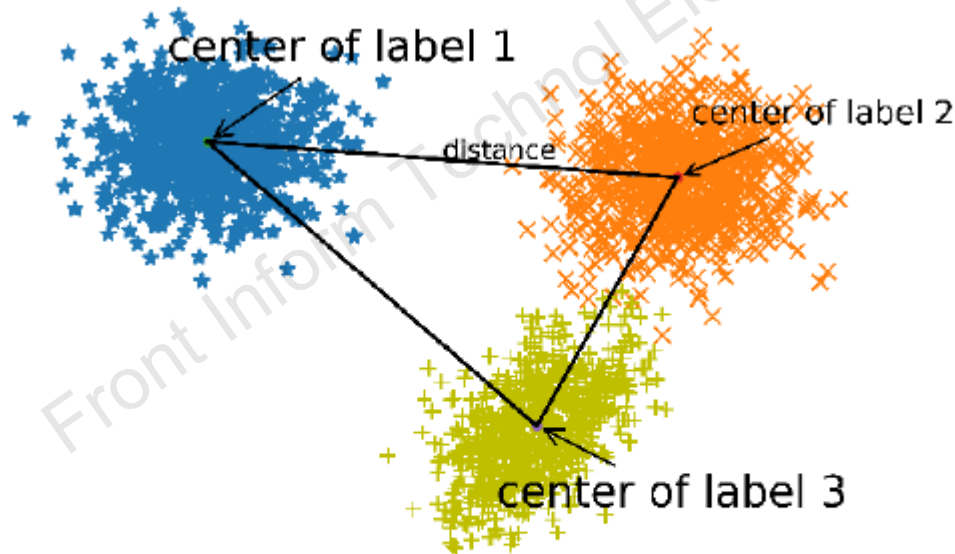
- ❑ To leverage the advantages of unsupervised training methods and class information, a semi-supervised learning method is proposed.
- ❑ To obtain deep information, the stacked autoencoder is proposed.
- ❑ Distance information is introduced into the objective function for improving the classification performance.



# Method

---

- ❑ The “distance” between different classes largely reflects whether the two categories can be easily classified correctly.
- ❑ Stacked semi-supervised learning methods are used to improve the ability of feature extraction.



**Fig. 3** Illustration of the distances between three labels  
“distance” means the distance between each two class centers

# Main results

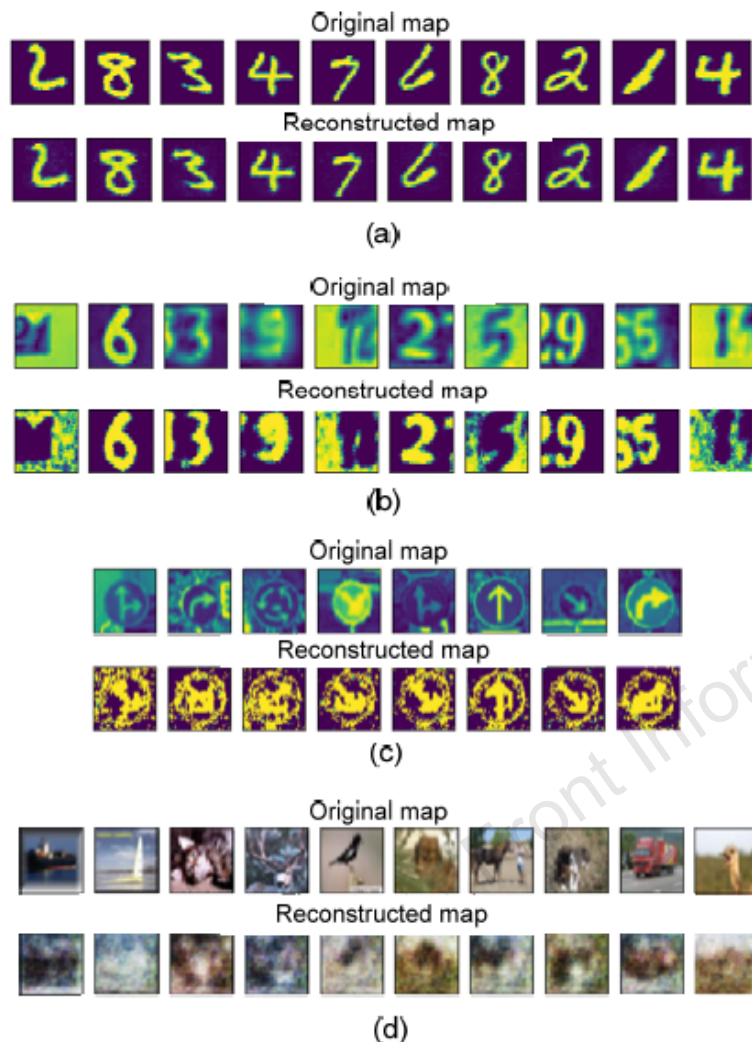


Table 3 Experimental results of the MNIST, SVHN, GTSRB, DS, and CIFAR-10 datasets with 50% labeled/unlabeled data from the original data

Model	Classification accuracy (unsupervised)				
	MNIST	SVHN	GTSRB	DS	CIFAR-10
AE	0.9271	0.6564	0.7958	0.9293	0.6149
Sv-AE	-	-	-	-	-
DAE	0.9554	0.7289	0.8205	0.9630	0.6208
StackedDAE	-	-	-	-	-
SAE	0.9606	0.7364	0.8260	0.9667	0.6294
DSAE	-	-	-	-	-
StackedDSAE	-	-	-	-	-

Model	Classification accuracy (supervised)				
	MNIST	SVHN	GTSRB	DS	CIFAR-10
AE	-	-	-	-	-
Sv-AE	0.9636	0.7619	0.8355	0.9745	0.6315
DAE	0.9507	0.7523	0.8212	0.9730	0.6287
StackedDAE	0.9567	0.7541	0.8254	0.9759	0.6345
SAE	-	-	-	-	-
DSAE	0.9722	0.7614	0.8350	0.9831	0.6463
StackedDSAE	0.9759	0.7662	0.8361	0.9856	0.6549

Model	Classification accuracy (semi-supervised)				
	MNIST	SVHN	GTSRB	DS	CIFAR-10
AE	-	-	-	-	-
Sv-AE	-	-	-	-	-
DAE	0.9601	0.7486	0.8337	0.9798	0.6319
StackedDAE	0.9733	0.7584	0.8357	0.9825	0.6427
SAE	-	-	-	-	-
DSAE	0.9730	0.7694	0.8447	0.9835	0.6447
StackedDSAE	<b>0.9764</b>	<b>0.7723</b>	<b>0.8459</b>	<b>0.9868</b>	<b>0.6591</b>

The best results are in bold

Fig. 12 Original and reconstructed maps of the distance autoencoder: (a) MNIST dataset; (b) SVHN dataset; (c) GTSRB dataset; (d) CIFAR-10 dataset

# Conclusions

---

In this paper, we proposed a semi-supervised stacked distance autoencoder model that can extract the features of high-dimensional data based on the distance information for a semi-supervised image classification problem. As we know, the traditional autoencoder is an unsupervised method, which does not use label information. We extend this principle to other major autoencoder models including the sparse autoencoder. From the experiments described above, we argue that the labels' distance information has an effect on the classification performance. The proposed semi-supervised distance autoencoder models are evaluated on several datasets, and the experimental results showed that by considering the distance information, more robust features can be generated. The better features contribute to improved classification results.