Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2022.030304
ISSN 2790-2692 Vol. 3, Issue 3: 28-35

Image Multi-feature Fusion with Restricted Boltzmann Machine and Neural Network

Yan Wang*

Jinzhong University, Jinzhong, China

clinical20222022@163.com

**corresponding author*

Keywords: Restricted Boltzmann Machine, Neural Network, Image Multi-Feature Fusion, Dataset

Abstract: Benefiting from the growing mass of big data and the rapid development of computing efficiency, a series of advanced models of machine learning and deep learning have been proposed one after another, and have achieved great success in many application scenarios. The research purpose of this paper is to fuse the image multi-features of the restricted Boltzmann machine and the neural network in the experiment, using the convolutional restricted Boltzmann machine, to establish a model and investigate the three datasets in the ship, oil, and machine. Image fusion performance evaluation results on the test set. The results show that the model can improve the performance of image fusion to a certain extent.

1. Introduction

It is the basic application of image processing. Image object extraction refers to identifying and interpreting meaningful objects from a sequence of images. Image scene classification refers to identifying different types of scene images, that is, based on the scene characteristics of the image the correct technique for classifying images in a dataset [1]. Object extraction and scene classification from high-resolution remote sensing images are research hotspots in the fields of computer vision and remote sensing. The deep learning model has the ability to automatically extract features from massive images, making it a trend to apply image methods based on restricted Boltzmann machines and neural networks to the field of image multi-feature fusion processing.

With the rapid arrival of the era of mobile Internet, there are more and more systems using biometrics for security authentication.Moradikia M presents a multi-eigen decomposition method for reconstructing synthetic borehole radar images from unsampled data in range and azimuth. This problem is addressed using a link-based radar image reconstruction-decomposition technique, in which features of interest are simultaneously processed and decomposed. Unlike traditional methods, the proposed technique provides multi-section images as well as synthetic borehole radar images [2]. Cuevas E studied that multi-dimensional segmentation proved to be superior to one-dimensional segmentation methods based only on grayscale information. The transformation algorithm is the most common algorithm for multi-feature segmentation. Despite the interesting results, the transformation is computationally expensive, which is an obstacle for segmental visualizations where the feature map contains multi-dimensional features. The proposed method considers a two-dimensional feature map including the grey scale value and area difference of each pixel in the image. In this case, two sets of features are distinguished: valid data and invalid data [3]. The research of image multi-feature fusion method is a key step in the research.

This paper studies restricted Boltzmann machine; neural network and image multi-feature fusion of restricted Boltzmann machine and neural network, including image feature extraction and image fusion. In the experiment, the convolution restricted Boltzmann machine is used to build the model and investigate the image fusion performance evaluation results on the test set of the ship, oil and engine datasets.

2. Research on Image Multi-feature Fusion Based on Restricted Boltzmann Machine and Neural Network

2.1. Restricted Boltzmann Machines

Restricted Boltzmann Machines, as an energy-based generative model, are widely used in data dimensionality reduction, data encoding and deep neural networks. The structure of a standard restricted Boltzmann machine is a symmetric bipartite graph consisting of binary-type visible layer nodes and hidden layer nodes, so it can only process binary input data [4-5]. To solve this drawback, Gaussian-Bernoulli Restricted Boltzmann Machines are designed to deal with real-type data (eg, image data). Different from standard RBM, the visible layer of GRBM consists of real nodes. Thanks to the powerful unsupervised data encoding ability of GRBM, GRBM is often regarded as an efficient unsupervised representation learning model, and many representation learning models based on GRBM have been proposed one after another. However, the existing GRBM-based representation learning models all ignore the point-to-point affinity information in the original input data, which can be used to improve the expressiveness of the data representations learned by the model in the hidden layer [6-7].

2.2. Introduction to Neural Networks

Neuroscientists in the United States have created a machine that can simulate human perception—the perceptron. It was originally applied to solve some simple problems, such as recognizing some English letters [8-9]. Since the single-layer neural network system represented by the single-layer perceptron at that time could not solve the simple XOR and other linear inseparable problems, the development of the artificial neural field has been stagnant. Later, the concept of neural network back-propagation algorithm and multi-layer perceptron was proposed, and the problem was solved. The multi-layer perceptron can form a neural network by interconnecting a large number of neurons, and the basic structure mainly includes three layers: input layer, hidden layer and output layer. In the multi-layer perceptron network, the parameter gradient can be calculated by the back-propagation algorithm, and it is easy to use the optimization algorithm based on the gradient information for optimization. Adding nonlinear activation functions can build complex nonlinear systems. The objective function or loss function is optimized by adjusting the weights in the neurons [10-11].

2.3. Image Multi-Feature Fusion with Restricted Boltzmann Machine and Neural Network

(1) Image feature extraction

Reflection spectral characteristics refer to the fact that electromagnetic waves are incident on different ground objects, and the ground objects reflect different energies, so this reflection can be described by reflectivity [12-13]. The reflectivity is the expression form of a function. The variables of this function are the incident angle, the electrical properties of the ground objects such as conductivity, dielectric value, the surface smoothness of the ground objects, etc. Generally speaking, when the incident electromagnetic wave wavelength is a When the value is constant, the lighter the color of the ground object in the grayscale image of the remote sensing image, the greater the reflectivity, and the darker the color tone, which means the smaller the reflectivity of the ground object, and the grayscale of the remote sensing image. Image interpretation is an important step in remote sensing interpretation. The characteristic of emission spectrum means that the various particles such as atoms, molecules, etc. that make up the material are doing random Brownian motion all the time. The surrounding environment radiates electromagnetic waves, and the existence of emissivity is a measure of the radiation ability of ground objects. The definition of emissivity is also for a specific wavelength, and the emissivity is determined by the characteristics of the ground objects, which include the surface smoothness, temperature, dielectric constant and other factors [14-15].

(2) Fusion of Images

Image fusion refers to orienting the multi-source remote sensing images collected by scientific researchers from different channels and ways to the same target. The favorable information is extracted and merged into a unified and high-quality image, which aims to improve the information efficiency of the image, and at the same time, it also improves the spatial clarity and spectral resolution of the original remote sensing technology image, which helps to better display the ground scenery. For details, for detection, the method with more applications and more mature technology is the fusion technology of infrared light and visible light. The two or more images to be fused need to ensure that the registration operation has been completed and the image bit width must be consistent, if the conditions are not met [16-17]. Image fusion technology also includes three levels from low-level to high-level, followed by data information level (image level) integration. The most widely researched and applied methods are based on Pixel-level fusion algorithms, most of the algorithms we use at present are still based on fusion at this level. The integration of the data information level (image level) is also the process of integrating the image by directly processing the data information obtained from the sensor. The subtle signals that cannot be extracted by other levels of fusion methods are also the technical cornerstone of advanced image fusion. The technical limitations of low-pixel layer image fusion cannot be ignored, because the signal is relatively redundant because the pixels are operated on, so the computer needs to technically process a large amount of various data information at the same time. It will take quite a long time, and the results will also increase significantly, and it is impossible to complete real-time data processing; in addition, when completing various data communications, due to the large amount of various data, it is easy to cause signal noise. In addition, if the image cannot be strictly registered to complete the image fusion, the merged image will be blurred, or the target and details will be unclear and inaccurate [18]. The spatial domain algorithm and time-varying domain calculation based on the image-level fusion algorithm are widely used and the method with high precision is the wavelet transform method.

3. Investigation and Research on Image Multi-feature Fusion Based on Restricted Boltzmann Machine and Neural Network

3.1. Research Content

By comparing the differences between the three models of DBN, DBM, and RBM in the SateUite-2000 dataset and the Vaihigen dataset. Image fusion performance evaluation results of the three models on the Levir dataset. Finally, the evaluation results of image fusion performance on the test set of three datasets of ship, oil and aircraft are investigated and analyzed.

3.2. Convolutional Restricted Boltzmann Machines

N_w represents the size of the convolution kernel, and * represents the convolution operation. Assuming that the unit states of all nodes in the model are binary variables, the corresponding energy function of the model is defined as:

$$E(v,h;\theta)=-\sum_{k=1}^{k}h^k\bullet(\overline{w}*v)-\sum_{k=1}^{k}b_k\sum_{i,j=1}^{N_h}h_{ij}^k-a\sum_{i,j=1}^{N_v}v_{ij} \tag{1}$$

Therefore, the activation probabilities of the input layer unit v^i and the hidden layer unit h_{ij}^k are expressed as:

$$p(v_{ij}=1|h)=g\left(\left(\sum_k w^k*h^k\right)+a\right) \tag{2}$$

4. Analysis and Research of Image Multi-feature Fusion Based on Restricted Boltzmann Machine and Neural Network

4.1. Image Multi-Feature Generation Shape Experiment

In order to better compare the differences between the three models of DBN, DBM, and RBM, we use the SateUite-2000 dataset and the Vaihigen dataset. The training time and average Euclidean distance metric for these three models are shown in Table 1 and Figure 1:

Table 1. Mean Euclidean distance measures

	Satellite-2000			Vaihigen		
	Training time: (s)	Training set	Test set	Training time: (s)	Training set	Test set
DBN	4957.51	6.3241	7.2542	6351.41	7.3617	8.3249
DBM	5137.15	5.3614	5.1241	6987.47	6.3817	7.6281
RBM	1368.07	4.3971	5.0628	3510.91	5.0149	6.3214

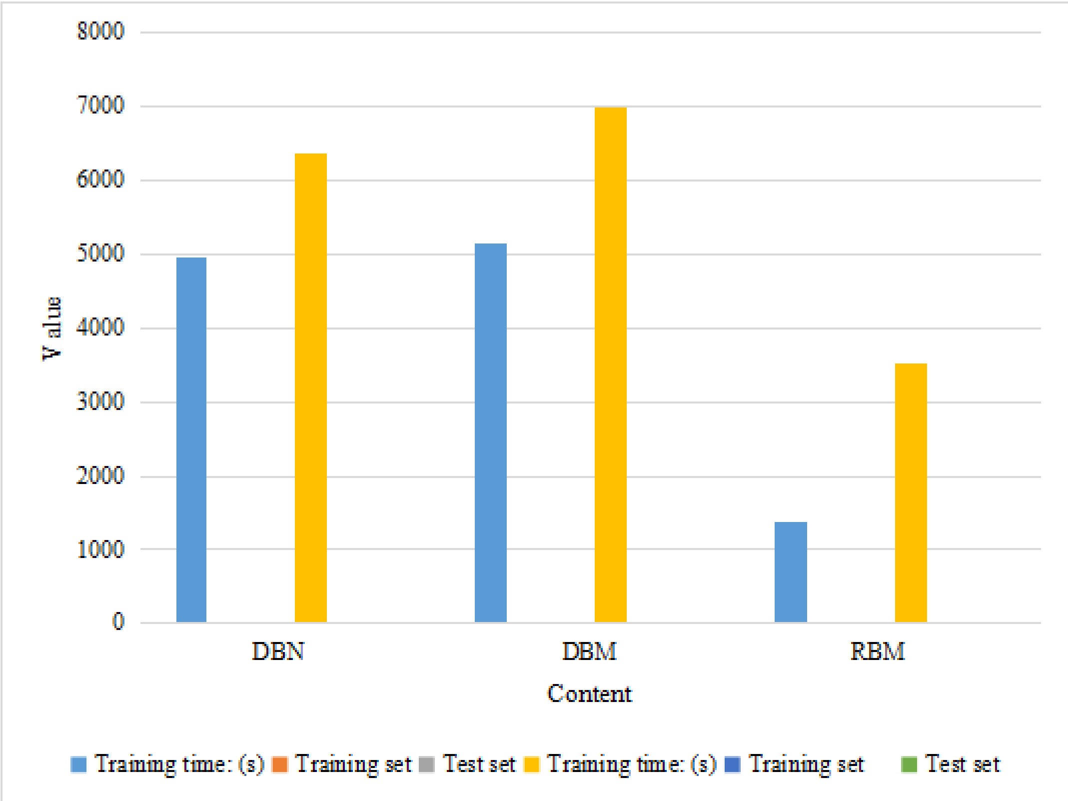

Figure 1. Training time and results of different models

The Restricted Boltzmann Machine results are better in terms of the ability to generate shapes. It can be seen in the Satellite-2000 dataset that when the aircraft target is small and the restricted Boltzmann machine generates the best shape, it is in line with reality.

4.2. Datasets

The Levir dataset contains two types of images: RGB images and labeled images. In order to better solve the problem of image multi-feature fusion, this experiment uses a three-stage method to create a new dataset. Three datasets are made by category, and the specific data are shown in Table 2 and Figure 2:

Table 2. Analysis of targets in LEVIR base on scale

Scale(pixel)	Scale<I000	Scale 1000-1500	Scale 1500-2000
Levir-oil drum	204	401	164
Levir-ship	281	105	135
Levir- airplane	701	492	301

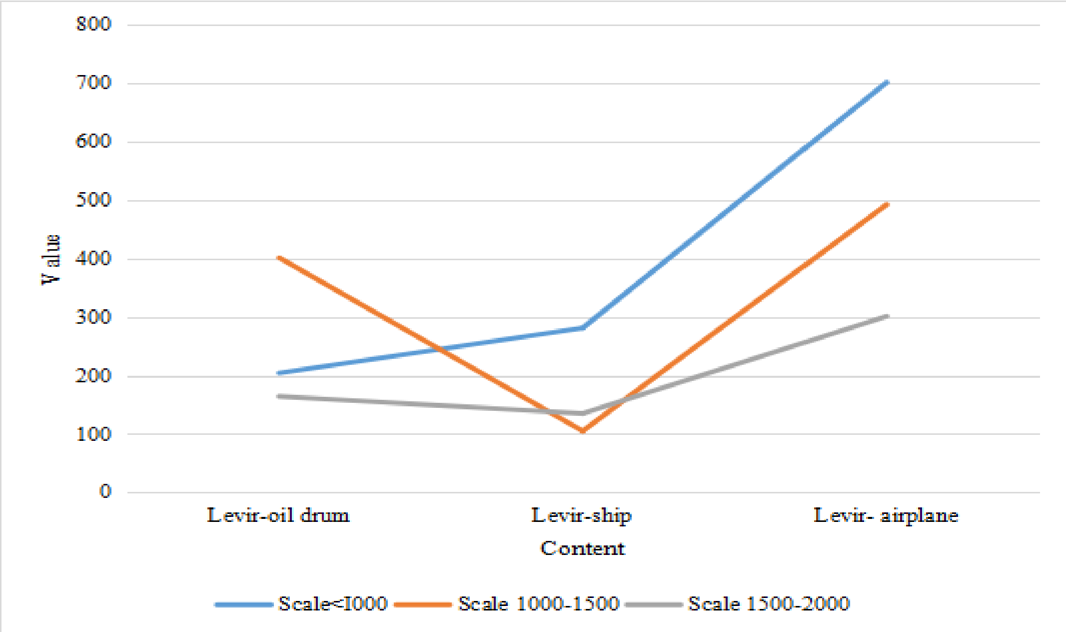

Figure 2. Results comparison plots of the three datasets

4.3. Dataset Experimental Results

In order to verify the image multi-feature fusion performance of the proposed model, the image fusion performance of the three models on the test of the Levir dataset, the specific evaluation results are shown in Table 3 and Figure 3:

Table 3. Evaluation of the segmentation results for the different models

	CV		CRBM-CV		CRBM-E-CV	
	Avg Global acc	Avg IOU	Avg Global acc	Avg IOU	Avg Global acc	Avg IOU
Levir-oil drum	71.361	61.084	98.241	95.327	99.541	96.142
Levir-ship	72.311	71.193	93.367	96.315	90.236	97.364
Levir-airplane	89.365	84.314	96.946	93.019	97.015	94.015

In order to solve the problem of poor classification performance due to lack of training samples in the task and insufficient training of the model, a multi-feature image classification model based on restricted Boltzmann machine is proposed to achieve the purpose of distinguishing image categories. The model can improve the performance of image fusion to a certain extent.

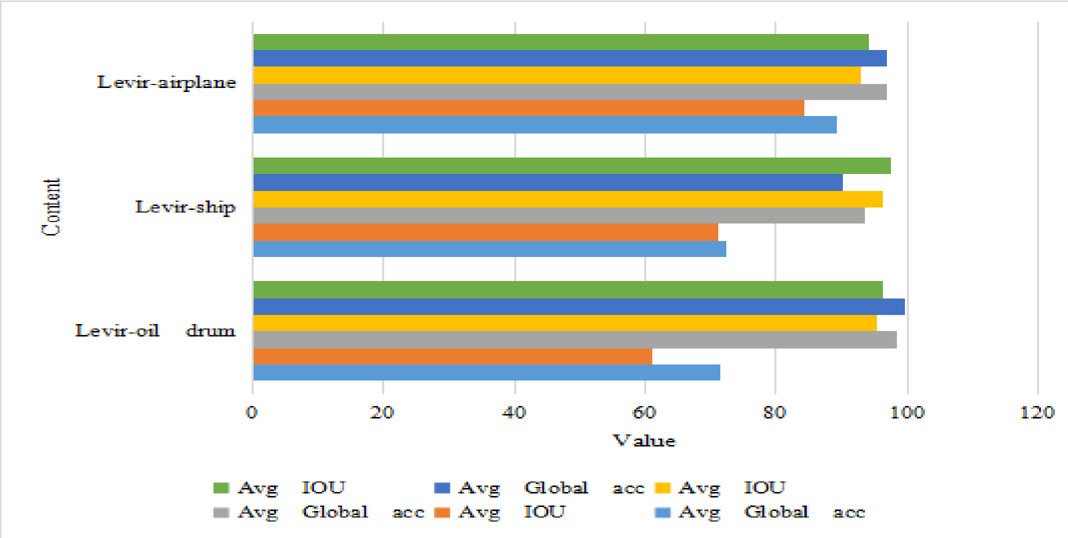

Figure 3. Graphical representation of data

5. Conclusion

Artificial intelligence technology represented by machine learning and deep learning has developed vigorously and achieved remarkable achievements, bringing great convenience to human life. Among many advanced learning models, deep neural networks have achieved great success in various application scenarios due to their powerful layer-by-layer coding capabilities. A layer-by-layer encoding model of data based on restricted Boltzmann machines was proposed, which accelerated the development of deep neural networks and opened the curtain of deep learning. The proposal of restricted Boltzmann machine greatly enriches the application scenarios of restricted Boltzmann machine, and also promotes the development and prosperity of image multi-feature fusion of restricted Boltzmann machine and neural network.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Vasamsetti S, Setia S, Mittal N, et al. Automatic underwater moving object detection using multi-feature integration framework in complex backgrounds. Computer Vision, IET, 2018, 12(6):770-778. https://doi.org/10.1049/iet-cvi.2017.0013

[2] Moradikia M, Samadi S, Cetin M. Joint SAR Imaging and Multi-Feature Decomposition from 2-D Under-Sampled Data Via Low-Rankness Plus Sparsity Priors. Computational Imaging

IEEE Transactions on, 2019, 5(1):1-16. https://doi.org/10.1109/TCI.2018.2881530

[3] Cuevas E, Becerra H, Luque A, et al. Fast multi-feature image segmentation. Applied Mathematical Modelling, 2021, 90(5):742-757. https://doi.org/10.1016/j.apm.2020.09.008

[4] Nanjun, Paoletti, Mercedes, et al. Feature Extraction with Multiscale Covariance Maps for Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(2):755-769. https://doi.org/10.1109/TGRS.2018.2860464

[5] Maksimovic-Moicevic S, Lukac Z, Temerinac M. Objective estimation of subjective image quality assessment using multi-parameter prediction. Image Processing, IET, 2019, 13(13):2428-2435. https://doi.org/10.1049/iet-ipr.2018.6143

[6] Fuqiang H E, Luo H, Yao X, et al. Self-adaptive Bridge bare rebar detection algorithm based on local image segmentation and multi-feature filtering. Journal of Applied Optics, 2020, 41(3):508-515. https://doi.org/10.5768/JAO202041.0302004

[7] Kapadia M R, Dr C. Content Based Medical Image Retrieval System for Accurate Disease Diagnoses Using Modified Multi Feature Fused Xception Model. Indian Journal of Computer Science and Engineering, 2021, 12(1):89-100. https://doi.org/10.21817/indjcse/2021/v12i1/211201179

[8] Xu S, Amira O, Liu J, et al. HAM-MFN: Hyperspectral and Multispectral Image Multiscale Fusion Network with RAP Loss. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(7):4618-4628. https://doi.org/10.1109/TGRS.2020.2964777

[9] Kumar M, Singh P. FPR using machine learning with multifeature method. IET Image Processing, 2018, 12(10):1857-1865. https://doi.org/10.1049/iet-ipr.2017.1406

[10] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4):834-848. https://doi.org/10.1109/TPAMI.2017.2699184

[11] Fazlzadeh A, Khoshmaram A, Feyzipour A. How Quality, Value, Image, and Satisfaction Create Loyalty at an Iran Telecom. International Journal of Business & Management, 2018, 6(8):580-580.

[12] Elad M, Datsenko D. Corrigendum: Example-Based Regularization Deployed to Super-Resolution Reconstruction of a Single Image. Computer Journal, 2018, 53(7):1131-1131. https://doi.org/10.1093/comjnl/bxp097

[13] Costa P, Galdran A, Meyer M I, et al. End-to-end Adversarial Retinal Image Synthesis. IEEE Transactions on Medical Imaging, 2018, 37(99):781-791. https://doi.org/10.1109/TMI.2017.2759102

[14] Ghamisi P, Yokoya N, Li J, et al. Advances in Hyperspectral Image and Signal Processing: A Comprehensive Overview of the State of the Art. IEEE Geoscience & Remote Sensing Magazine, 2018, 5(4):37-78. https://doi.org/10.1109/MGRS.2017.2762087

[15] Dong, Nie, Roger, et al. Medical Image Synthesis with Deep Convolutional Adversarial Networks. IEEE Transactions on Biomedical Engineering, 2018, 65(12):2720-2730. https://doi.org/10.1109/TBME.2018.2814538

[16] Selvaraju R R, Cogswell M, Das A, et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. International Journal of Computer Vision, 2020, 128(2):336-359. https://doi.org/10.1007/s11263-019-01228-7

[17] Deal M M, Nolet G. Nullspace shuttles. Geophysical Journal International, 2018, 124(2):372-380. https://doi.org/10.1111/j.1365-246X.1996.tb07027.x

[18] Alsmadi, Mutasem K. A hybrid Fuzzy C-Means and Neutrosophic for jaw lesions segmentation. Ain Shams Engineering Journal, 2018, 9(4):697-706. https://doi.org/10.1016/j.asej.2016.03.016