Journal of Clinical and
Medical Images

Research Article

ISSN: 2640-9615 | Volume 6

Automatic Semantic Annotation of News Images in Mobile Internet of Things and Construction of Semantic Internet of Things System

Wang[2], Ling[3] and Chen H[1*]

[1]School of News and Communication, Pingdingshan University, Pingdingshan, 467000, Henan, China

[2]School of Engineering, Lishui University, Lishui, 323000, Zhejiang, China

[3]School of Management, Wuzhou University, Wuzhou, 543003, Guangxi, China

[*]Corresponding author:

Haijun Chen,
School of News and Communication, Pingdingshan University, Pingdingshan, 467000, Henan, China,
E-mail: drangelachen@126.com

Received: 09 Aug 2022
Accepted: 19 Aug 2022
Published: 25 Aug 2022
J Short Name: JCMI

Citation:

Chen H, Automatic Semantic Annotation of News Images in Mobile Internet of Things and Construction of Semantic Internet of Things System. J Clin Med Img. 2022; V6(15): 1-10

Keywords:

Mobile Internet of things; News images; Automatic semantic annotation; Visual features

1. Abstract

With the continuous development of mobile and semantic IoT (Internet of Things), the number of news images is increasing sharply. Automatic semantic annotation technology is an important tool to manage and retrieve news images. Therefore, the study is focused on the automatic semantic annotation of news images. First, the problems of the existing automatic annotation methods of news images are analyzed and studied, and then an automatic semantic annotation method of news images based on the fusion of weight and features is proposed. The proposed method extracts the colour features, texture features, and shape features of news images to determine their stability by using the standard deviation of features, and then calculates the weighting coefficient to realize automatic annotation of the news images. The results of the study in the corel5k dataset show that the accuracy of the weight-feature fusion annotation model is 60% for the colour feature, 62% for the texture feature, and 67% for the mixture of colour, shape, and texture. The accuracy of the fusion of weight and feature annotation method is 7% higher than that of the single feature annotation method. The maximum accuracy of the fusion of weight and feature annotation method is 91.2%. The study provides a reference for the research of automatic annotation of news images.

2. Introduction

The semantic web is a network that takes the machine as the center and connects with the computer or other electronic equipment to make it have high intelligence and sharp logical judgment ability. A sound semantic network is like a human brain, and can insight into the user's intention. And it can achieve the user's goal independently. A semantic network is a graph structure that represents knowledge in the form of patterns of interconnected nodes and arcs and its classification of networks are Definitional networks, Assertional networks, Implicational networks, Executable networks, Learning networks, Hybrid networks. The classifications of semantic network are Definitional networks, also known as a generalization or subsumption hierarchy, emphasise the relationship between a concept type and a newly defined subtype. Assertional networks are intended to validate propositions. As models of the conceptual structures underlying natural language semantics, some assertional networks have been proposed. Implicational networks connect nodes using implication as the principal relation. Executable networks incorporate mechanisms such as marker passing or associated routines. Learning networks construct their representations by learning from examples. Hybrid networks integrate two or more of the preceding strategies in a single network or in distinct but closely interconnected networks. With the development of the Internet and multimedia information technology, more and more cameras and other electronic devices are used in our lives. And various images appear accordingly and are applied unprecedentedly. Also, the development of the Internet makes them spread faster and faster [1, 2] . The large-scale increase of news images has a higher demand for the management, retrieval, and storage. The automatic annotation can make users understand the content of news images directly, and manage, retrieve, and classify massive news images based on semantics. Automatic annotation is the process by which a computer system automatically assigns metadata in the form of captioning or keywords to a digital image and also defined as automatic image tagging or linguistic indexing. For the news images generated and accumulated by the speed of TB every day, the semantic automatic annotation method of news images can help users locate the target image quickly, improve the utilization rate of news images, and provide convenience for the public to obtain news images[3]. The studies on automatic semantic annotation of news images have made great achievements in China and foreign countries. Albukhitan et al. (2020) [4] pointed out that a great deal of digital content is produced, which makes people need to establish an image analysis and retrieval system based on image processing and machine learning. They also explained that machine learning can bridge the semantic gap in image retrieval, and proposed an automatic annotation framework of images, which can acquire training images from social media, and combine supervised and unsupervised machine learning to retrieve images. Besides, the research also emphasizes that the performance of the annotation system can be optimized with the increase of the number of images continuously, and suggests that the hardware of the machine learning algorithm should be upgraded to improve the speed of image retrieval. Tarek (2018) [5] proposed a semantic annotation framework for the Arabic language based on the deep learning model, which uses different models to annotate a set of given Arab documents and ontology. The results show that different vectors and matching can be done by using different Arabic language embedding models, which proves that the framework has good performance. Dutta et al. (2018) [6] pointed out that advanced digital capture technologies contribute to the explosive growth of network images. To retrieve the required images from a large number of images, users are more interested in text queries than other visual image-query methods. The processing system for textual image queries will be used to identify the picture category for the selection of a subset of images to search for data responsive to the query. The query can be used by an output type selector, which is part of the query processing system, to identify the query's output type. Visual image-query is the method of searching image databases for items that are relevant to a user query and the searching problem is reconstructed as an information retrieval challenge. Semantic annotation of images is considered an important step to realize image processing and retrieval. Semantic annotation marks the existing images on the network, which can interpret the images and make the search process easier. To annotate the image effectively, many image interpretation techniques are produced to explore semantic concepts of images. However, due to the complexity and diversity of images, establishing an effective annotation method of images is still very challenging, and gains much attention. Ma et al. (2019) [7] argued that automatic annotation of images is one of the basic problems of computer vision and machine learning, and provided a set of text tags to describe the semantics of a given image. In the last decade, a large number of annotation technologies of images are put forward, and many achievements are achieved in the studies on the annotation of various datasets. However, most of the research is still limited to the quantitative results of the data, ignoring the key factors related to the attributes and evaluation indicators of the datasets, which have a considerable impact on the research results. Here, after the 10 most advanced annotation methods of images are evaluated by CNN (convolutional neural networks), a new quantitative method is proposed after the specific deviation of the dataset, the evaluation criteria of each tag on each image, and the influence of changing the number and types of the tags, are analyzed. In content-based image retrieval systems, color and texture characteristics are essential. Color descriptors include color histograms, color moments, and Color Coherence Vector (CCV) features, whereas texture descriptors include discrete wavelet transform features and co-occurrence matrices.

Although domestic and foreign scholars do some research on the automatic semantic annotation of news images, the accuracy of the automatic semantic annotation of news images is still low. To improve the performance of automatic semantic annotation of news images, the method of combining the bottom visual features and the fusion of the color features, texture features, and shape features of images is studied.

3. Automatic Annotation of News Image Based on Fusion Features

3.1. Basic concepts and process of semantic annotation of images

Semantic annotation of images is the process of automatically generating relevant semantic words to explain the content of the image. To facilitate the process of semantic annotation, automatic annotation can be regarded as the language description process of the machine. On convenience grounds, the image training sets with tags are set as A, the image training sets without tags are set as B, the semantics of the tag corresponding to the image training set B is C $\{c_1, c_2, c_3,..., c_n\}$, and Ci is the keyword. The process of automatic annotation of images is as follows: a certain mode is used to calculate the randomly selected image Ai, find the keywords describing the image, and then annotate it. The semantic automatic annotation model of images is in Figure 1, which shows that the standard automatic annotation model of images can be divided into three parts: underlying visual feature extraction, generation annotation model, and the annotation of the images. The process of semantic annotation falls into three stages: text-based manual annotation stage, content-based automatic annotation stage, and semantic-based annotation stage. Figure 1 shows the process of semantic annotation [8-11]. The technique of manually identifying areas in a picture is known as manual annotation.

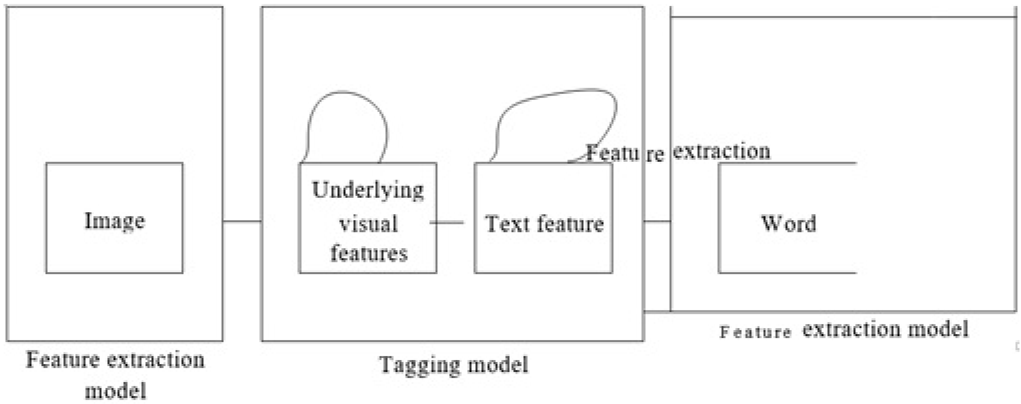


Figure 1: The process of semantic annotation

3.2 Feature extraction

The colour feature is a feature of the surface of the object in a news image based on pixel point feature. Colour features are easy to extract, have less dependence on the size and direction of the image, and high robustness. Also, they are the most intuitive attributes of objects. Only using the colour features of objects can distinguish semantic categories [12]. The principle of the colour moment is to use an image matrix to reflect the distribution of colours. The Principle of the color moments are mostly employed in image retrieval applications as characteristics for colour indexing in order to compare how similar two images are based on colour and it is addressed in the revised manuscript.The common colour moments have a first moment, second moment, and third moment. The first moment describes the average colour of the image, the second moment describes the colour distribution of the image, and the third moment describes the symmetry of the colour distribution of the image. The colour moment has the advantages of low dimension, fast speed, and high accuracy [13]. The colour moments in RGB colour space are used to extract the colour features of news images. The first moment, second moment, and third moment are calculated in equations (1) ~ (3).

$$M_{i1} = \frac{1}{N}\sum_{i=1}^{N} P_{ij} \quad (1)$$

$$M_{i2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(P_{ij} - M_{i1}\right)^2} \quad (2)$$

$$M_{i3} = \sqrt[3]{\frac{1}{N}\sum_{i=1}^{N}\left(P_{ij} - M_{i1}\right)^3} \quad (3)$$

In equations (1) ~ (3), i is the components of the colour model, i 1 is the components of R, i 2 is the components of G, and i 3 is the components of B. pij is the probability that the pixel of the colour component i is j , and N is the number of pixels.

Since news images have the characteristics of high visual similarity, colour features cannot be used to describe and distinguish news images [14- 15]. Therefore, in the research of automatic semantic annotation of news images, it is necessary to extract other features of news images. The texture feature of an image describes the arrangement of the components of each object in the image, and it is a periodically changing visual element. There are three commonly used texture feature extraction methods, which are statistics-based texture feature extraction method, signal processing-based texture feature extraction method, and structure-based feature extraction method [16-18]. The research object of the texture feature extraction method based on statistics is the gray value of the current pixel, and the extracted texture feature is the first or higher derivative statistical information of gray. The research object of the texture feature extraction method based on signal processing is the image that changes in the time domain or frequency domain, and the extracted texture feature is the stable feature in the changes. Texture features extracted by the structure-based feature extraction method are texture elements. Based on the above analysis, SIFT (Scale Invariant Feature Transform), Gabor filter, and sparse coding are used as the texture feature extraction method. The extraction process is shown in Figure 2.

SIFT (Scale Invariant Feature Transform) - The scale-invariant feature transform (SIFT) is a computer vision feature identification algorithm used to find and describe local features in images. Gabor filters are bandpass filters that are used in image processing to extract features and analyse texture. The shape of the news image is important content in describing the news image. The method with a constant size and changing features is used to extract the shape features of the news image. To obtain the feature vector with a constant size and/changing features, the extreme value of the scale space needs to be tested to find the direction of the selected feature points [19]. The process of shape feature extraction is divided into four steps:

1) All images in the scale space are searched, and the maximum and minimum values of the scale space are detected by using Gaussian differential function to obtain the scale and the extreme

point, which is invariable.

2) Gaussian blur is applied to the image to find the position of the feature points with less change.

3) The orientation of feature points is determined by the position of local feature points.

4) Local gradients are measured for feature points of different scales, which need to be described [20-22].

Based on the above, the colour, texture, and shape features of the news images are obtained, but the results are unsatisfactory in some dimensions. Therefore, the features of each dimension are fused to obtain their weight. In probability statistics, the standard deviation represents the volatility of data, and the standard deviation of each dimension of the news images is positively related to the stability of the feature. Suppose that the dataset of the news images has x classes, each class has y graphs, and the feature dimension is Z. Then the standard deviations of class a and the dimension b are shown in equation (4) [23, 24].

$$S_{ab} = \sqrt{\frac{1}{Y-1}\sum_{i=(b-1)Y+1}^{a^*Y}\left(X_{ij}-X_{ab}\right)^2}, i \in [(a-1)Y+1, a^*Y], j = a \in X, b \in Z \quad (4)$$

The weight matrix of each class and dimension is shown in equation (5).

$$R = (r_{ab}) = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1y} \\ r_{21} & r_{22} & \cdots & r_{2y} \\ \cdots & \cdots & \cdots & \cdots \\ r_{x1} & r_{x2} & \cdots & r_{xz} \end{bmatrix} \quad (5)$$

In equation (5), $r_{ab} = \frac{1}{1+S_{ab}}, 1 \le a \le X, 1 \le b \le Z.$

The process of semantic annotation of the news images based on weight fusion is shown in Figure 3.

Figure 3 shows that the colour moment method is used to extract colour features of remote sensing images based on RGB space, the texture features of remote sensing images are extracted by mixed methods, and the shape features of remote sensing images are extracted by SIFT. And then the weight fusion of the extracted visible features of remote sensing images is carried out [25]. Support Vector Machine (SVM) classifier is used to classify the regression models from the group of data. Also, it is known as non-probabilistic, binary linear classifier. It majorly used to separates the data into classes.

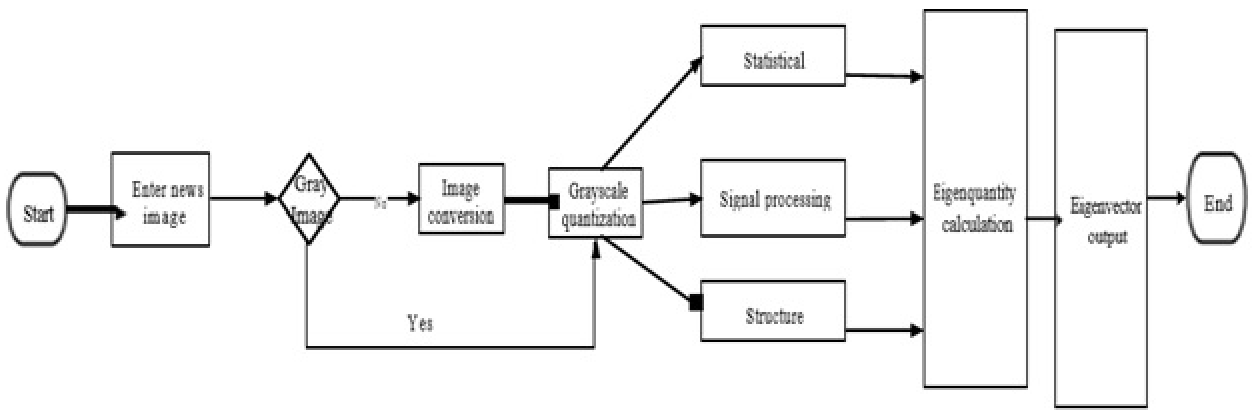


Figure 2: Feature extraction

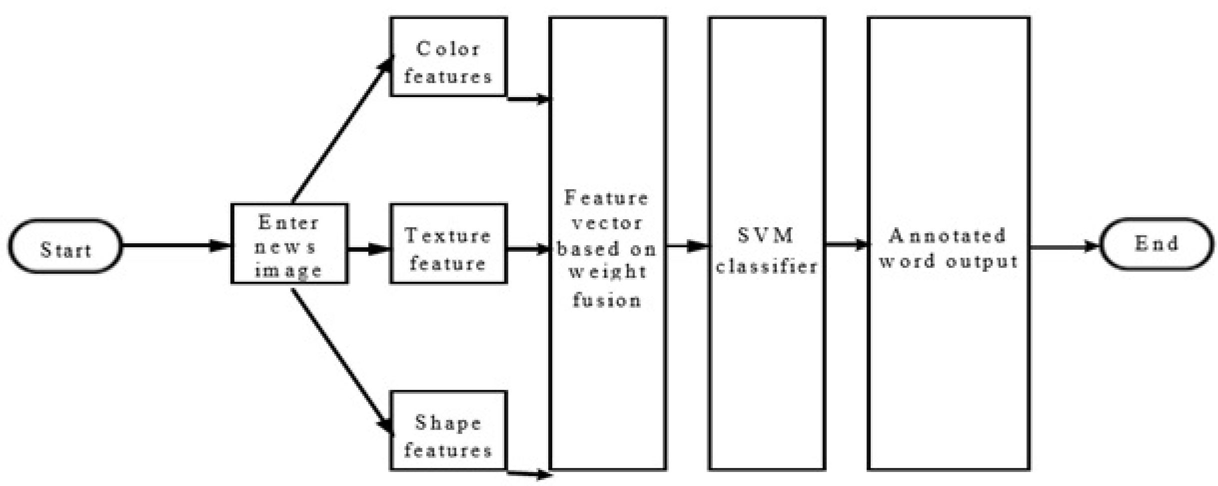


Figure 3: Semantic annotation of the news images based on weight fusion

3.3 Experimental methods and datasets

The effectiveness of weight determination and weight fusion is tested respectively to verify the effectiveness of the automatic annotation method. And the performance of the framework is tested by Corel5k, which is widely used in the field of automatic annotation of images. The Corel5k dataset is shown in Figure 4. The Corel image dataset covers multiple images with different themes and consists of several CDs. Each CD contains 100 images of the same size, which can be converted into a variety of formats. Each CD represents a semantic theme, such as buses, dinosaurs, and beaches. Corel5k image dataset is usually divided into three types: 4000 images as the training set, 500 images as the validation set to estimate model parameters, and the remaining 500 images as the test set to evaluate algorithm performance. After the optimal model parameters are found by using the validation set, 4000 training sets and 500 validation sets are mixed to form a new training set.

Each image in the dataset is tagged with 1-5 words, and 374 tagging words in the training set and 263 tagging words in the test set are used [26, 27].

Figure 4: Corel5k dataset (partial)

4. Deep Learning Annotation Method of News Images Based on the Fusion of Multiple Features

4.1 Improvement of DBN

DBN (Deep Brief neural network) is a generated model, which consists of multiple dominant neurons and recessive neurons. Dominant neurons are used to receive input and recessive neurons are used to extract features. After the weights between the neurons are trained, the whole neural network can generate training data according to the maximum probability. DBN stands for Deep Belief Network which is used as a generative graphical model, and it can be thought of as a combination of simple, unsupervised networks like limited Boltzmann machines. Boltzmann machines are stochastic and generative neural networks that can describe and solve challenging combinatorial problems by learning internal representations and is also known as a stochastic Hopfield network with hidden units. Restricted Boltzmann Machine (RBM) is a generative stochastic artificial neural network capable of learning a probability distribution across a collection of inputs .DBN is composed of several limited Boltzmann machines. And the network structure of a typical DBN is shown in Figure 5, which reflects that there are many multi-layer nodes in DBN [28, 29]. The nodes in each layer are not connected, while the nodes of the two adjacent layers are fully connected. The nodes in the bottom layer of the network are observable variables, and the nodes in other layers are hidden variables. The connection between the top two layers is undirected, and the connection between the other layers is directed. The training of DBN can be divided into two steps: 1. Training each layer of RBM (Restricted Boltzmann Machine) network independently and unsupervised. When the feature vector is mapped to different feature spaces, the feature information is obtained as much as possible; 2. Setting up BP (Back propagation) network at the last layer of DBN. The output eigenvector of RBM is used as its input eigenvector, and the entity relation classifier is trained. Moreover, the weight in its layer is the best for the vector mapping of the layer, and not for the feature vector mapping of DBN. BP propagates the error to each layer of RBM from top to bottom, and adjusts the whole DBN. The training of RBM network model can be regarded as the initialization of the weight parameter of deep BP network, which makes up for the shortcoming of long-time training of BP network [30, 31].

Since the original Boltzmann machine annotation model limits the visible layer to binary variables, the features of news images are used as the visible layer input of DBN, and each neuron is a real number. Therefore, the traditional constrained Boltzmann machine needs to be improved. The improved Boltzmann machine, like the traditional Boltzmann machine, is still composed of a visible layer and a hidden layer. The node in the visible layer is defined $v \in R^I$, and the node $v \in \{0,1\}^1$ in the hidden layer is a binary variable. The energy function is shown in equation (6), the state of the node in the visible layer is shown in equation (7), and the state of the node in the hidden layer is shown in equation (8).

$$E(v,h|\theta) = \sum_{i=1}^{I} \frac{(v_i - b_i)^2}{2\sigma^2} - \sum_{i=1}^{I}\sum_{j=1}^{J} v_i W_{ij} h_j - \sum_{j=1}^{J} a_j h_j \quad (6)$$

$$p = (V_i = X|h) = \frac{1}{\sqrt{2\pi S_i}} exp \frac{\left(x - S_i \sum_{j=1}^{j} h_j W_{ij}\right)^2}{(-2S_i^2)} \qquad (7)$$

$$p\left(h_j = 1|V\right) = \sigma\left(\sum_{i=1}^{I} W_{ij}\frac{V_i}{S_i}\right) \qquad (8)$$

News images contain rich content. It is difficult to make full use of all the information contained in the news images by using a single feature for automatic semantic annotation, which reduces the accuracy of remote sensing annotation of news images. G_1 represents the colour feature vector, G_2 is the texture feature vector, and G_3 is the shape feature vector. $R = (r_1, r_2, r_3)$ represents the weight fusion. The feature description of the news images is shown in equation (9).

$$G = r_1 g_1 + r_2 g_2 + r_3 g_3 \qquad (9)$$

The improved automatic annotation of news images is shown in Figure 6. The characteristics of the news images are taken as the input of the model, and then the Boltzmann machine network is trained upward layer by layer. Because of the feature vector $v \in R^I$ of the news images, equation (6) is used to calculate the energy function. Equation (7) is used to calculate the hidden layer, and the calculated value of the hidden layer is taken as the input vector of the Boltzmann machine. And the equations of the energy function of other Boltzmann machines remain unchanged, and the steps are continued to the last layer of DBN. In the last layer of DBN, the output feature vector of the Boltzmann machine is input into the classifier, and the category of the news images is used as the annotation word to complete the annotation of the news images. Then, the hidden layer is restored, and the BP network structure is used to optimize the annotation model. After the features are transformed layer by layer, the features of the news images input to the network are transformed from the initial space to the new feature space, forming the ideal features suitable for annotation. The network structure is shown in Figure 6.

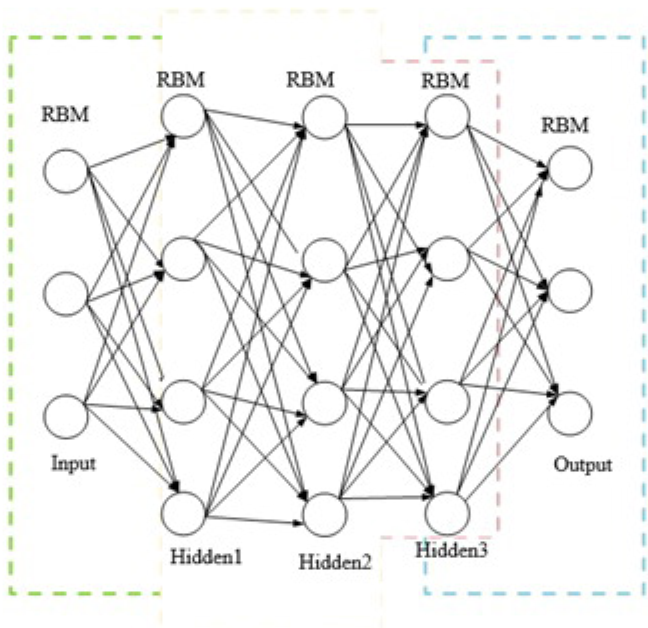


Figure 5: Structure of DBN

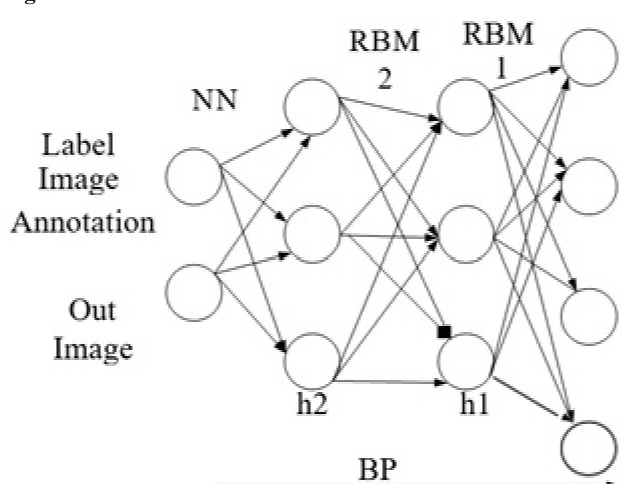


Figure 6: Automatic annotation of the news images

4.2 Experimental methods

The effectiveness of the automatic annotation method based on the fusion of multiple features is verified from two aspects. The first is the influence of DBN on the accuracy of automatic annotation of news images, and the second is the performance of the improved annotation method based on DBN. For the first, the following parameters are set: the number of visible node layers in the first layer Boltzmann machine and the feature dimension of the news images are set the same; the number of hidden layers is set as 2,3,4, and the number of hidden layer nodes is set as 200, 300, 400, and 500 respectively. For the second, a group of comparative experiments are carried out, and the following parameters are set: 300, 600, 900 images are extracted respectively from the 4000 images of Corel5k image dataset as three training sets, and the experimental methods described in Section 2 and section 3 are employed.

5. Test of the Performance of the Framework

5.1. Experiment on Automatic Semantic Annotation of News Images Based on Weight

Figure 7 shows the Relationship between weight and annotation accuracy.

Figure 7 shows the accuracy of automatic semantic annotation of news images is 62% after it is weighted by using colour features, while the accuracy is 57% without the weight; the accuracy is 67% after it is weighted by using texture features, and 68% without the weight; the accuracy is 64% after it is weighted by using shape features, and 60% without the weight. From the above data, it is found that the accuracy of the automatic annotation of news images by using texture features is the highest. After it is weighted, the accuracy is improved. It is also found in Figure 7 that the annotation accuracy of the fusion of multiple feature is higher than that of a single feature.

Figure 8 shows the comparison between the annotation method based on the fusion of multiple features and the annotation method with a single feature in the case of different numbers of news images in the training set.

Figure 8 shows that when the number of training images is 300, the annotation accuracy of color feature is 60%, and the annotation accuracy of the texture feature is 62%. The annotation accuracy of the fusion of color, shape, and texture is 67%. When the number of training images is 500, the annotation accuracy of color feature is 67%, and the annotation accuracy of the texture feature is 68%. The accuracy of the fusion of color, shape, and texture is 73%. When the number of training images is 1500, the annotation accuracy of color feature is 87%, and the annotation accuracy of the texture feature is 89%. The accuracy of the fusion of color, shape, and texture is 94%. To sum up, with the increase of the number of news images used for training, the annotation accuracy of a single feature is improved by 27%, and that of the fusion feature is always high, with an increase of 17%.

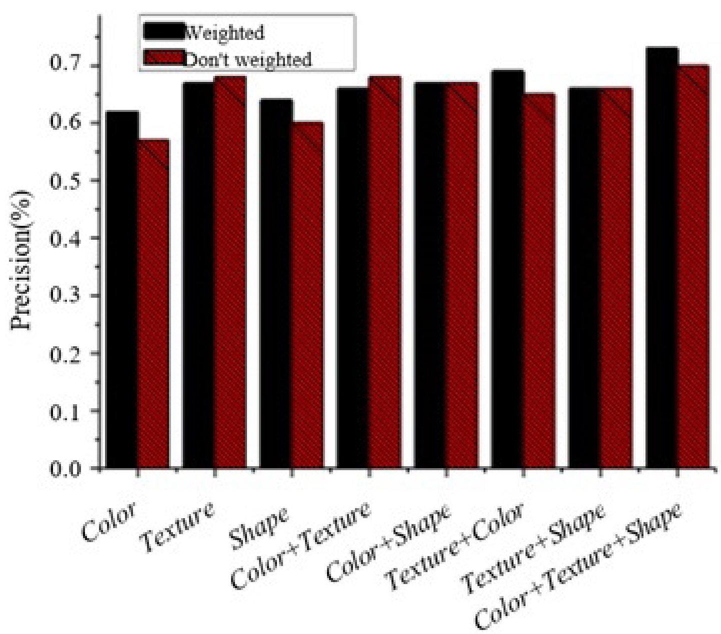


Figure 7: Relationship between weight and annotation accuracy

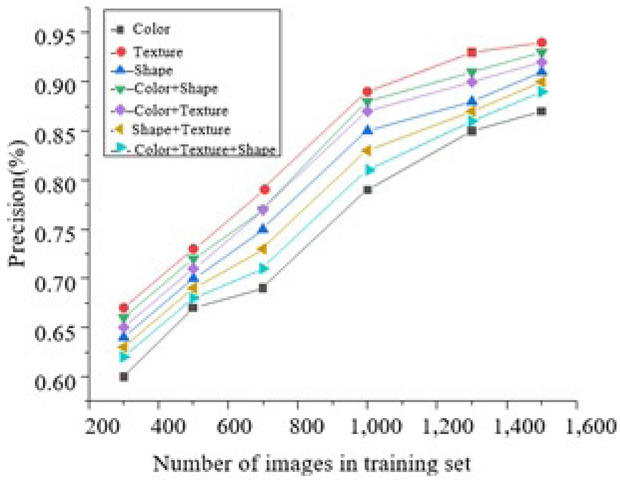


Figure 8: Relationship between the number of training images and annotation accuracy

5.2. Test of the annotation method of news images based on feature fusion

The influence of DBN parameters on the annotation is shown in Figure 9.

Figure 9 shows that when the number of hidden layers is 2 and the number of hidden layer nodes is 200, the annotation accuracy of the model is 81.8%. When the number of hidden nodes is 300, the annotation accuracy of the model is 86.9%, and the annotation accuracy of the model is the highest. When the number of hidden layers is 3 and the number of hidden layer nodes is 200, the annotation accuracy of the model is 92.1%. When the number of hidden nodes is 300, the annotation accuracy of the model is 91.6%. When the number of hidden layers is 4 and the number of hidden layer nodes is 200, the annotation accuracy of the model is 89.5%. When the number of hidden nodes is 300, the annotation accuracy of the model is 87.8%. In a word, the annotation accuracy of the tagged network achieves the maximum when the number of hidden layers is 3 and the number of hidden layer nodes is 200. On a whole, with the increase of the number of hidden layers and hidden layer nodes, the annotation accuracy ofthe network declines slowly.

Figure 10 shows the comparison of the accuracy of the semantic annotation of news images between the weight feature fusion method and the fusion method of multiple features.

Figure 10 shows that the accuracy of the deep learning annotation method based on the fusion of multiple features is higher than that based on the weight feature fusion method, because the method can reveal the distribution features of high-dimensional feature data by active learning layer by layer, which can improve the performance of semantic annotation.

The annotation results of some images are shown in Table 1.

Table 2 shows the comparison of annotation results of different algorithms in the Corel5k dataset.

Tables 1 and 2 show that the annotation accuracy of the proposed algorithm is higher than that of SVM,

NN, and DBN. When the visual features of the images are extracted, the feature learning strategy of the reconstructed DBN model can directly optimize the visual features, and the fine-tuning network is used to construct a stronger feature expression network, which proves that the improved DBN and framework is feasible and effective in learning the semantics of visual features and images.

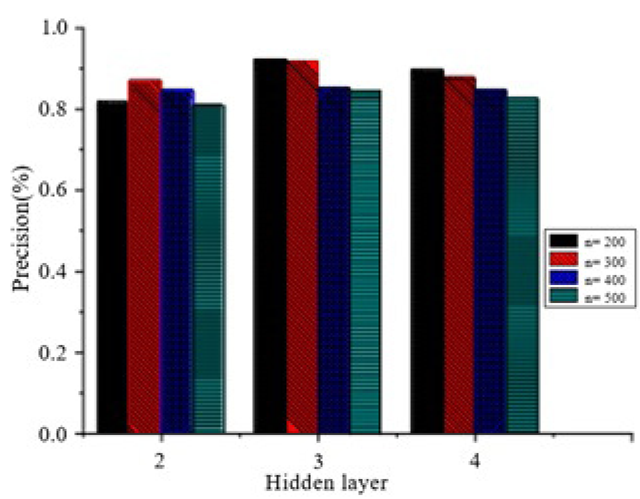


Figure 9: The influence of DBN parameters on the annotation (n = nodes)

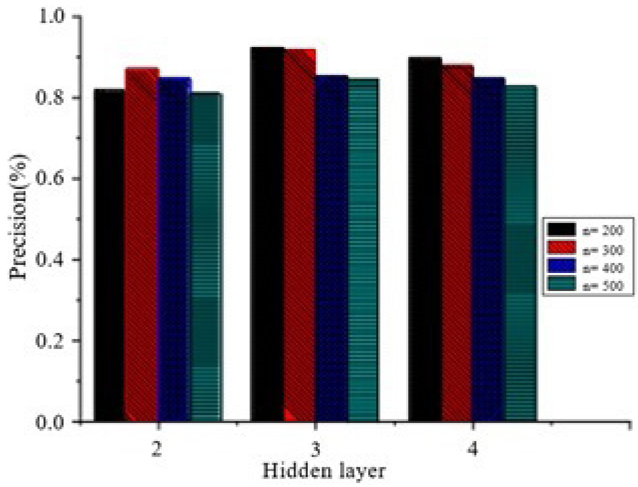


Figure 10: The comparison of the accuracy of the semantic annotation of news images between weight feature fusion method and fusion method of multiple features

Table 1: Annotation results of some images

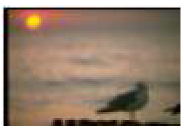
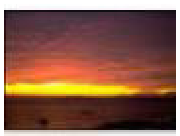
Original images	Annotation results	Manual annotation results
	The red color The sun: The bird	The sun: The bird The cloud
	The cloud	The cloud
	The cloud The trees The mountains The red color	The cloud The trees

Table 2: Comparison of annotation results of different algorithms

Algorithms	Accuracy
SVM(Support Vector Machine)	76.20%
NN(Neural Networks)	75.30%
DBN(Deep Brief Networks)	86.30%
The annotation method of the fusion of multiple features	91.20%

6. Conclusion

The problem of low accuracy of automatic semantic annotation caused by the complex structure of news images is studied. First, the fusion of multiple features is proposed to describe the news images, which improves the accuracy of the single feature description method. And a new method of automatic annotation of news images by using weight feature fusion is put forward. This method extracts the color, texture, and shape features of news images, calculates the differences of the annotation of each class and dimension, and determines the stability of the data to obtain the coefficient matrix. Then, the visual features of the news images are weighted to get a reasonable feature descriptor, and the classifier is used to complete the semantic annotation of the news images. The method proposed solves the problem of low accuracy of automatic annotation of news images. However, it is limited to study the problem of the low accuracy of semantic automatic annotation by fusing the color features, texture features, and shape features of news images. In the future, this algorithm will be used in more feature research to improve its applicability.

References

1. He Xing, Zhang Hansi, Bian Jiang. User-centered design of a web-based crowdsourcing-integrated semantic text annotation tool for building a mental health knowledge base. Journal of Biomedical Informatics. 2020.
2. Hailiang Li, Bin Zhang, Yu Zhang, Linfeng Wei. A semi-automated annotation algorithm based on weakly supervised learning for medical images. Biocybernetics and Biomedical Engineering(2). 2020.
3. Pranab Gajanan Bhat, Badri Narayan Subudhi, T. Veerakumar, Vijay Laxmi, Manoj Singh Gaur. Multi-Feature Fusion in Particle Filter Framework for Visual Tracking. IEEE Sensors Journal(5). 2020.
4. Saeed Albukhitan, Ahmed Alnazer, Tarek Helmy. Framework of Se-

mantic Annotation of Arabic Document using Deep Learning. Procedia Computer Science(C). 2020.

5. Tarek Helmy. A Generic Framework for Semantic Annotation of Images. International Journal of Image and Graphics(3). 2018.

6. Dutta Ayushi, Verma Yashaswi, Jawahar CV. Automatic image annotation: the quirks and what works. Multimedia Tools and Applications(24). 2018.

7. Yanchun Ma, Yongjian Liu, Qing Xie, Lin Li. CNN-feature based automatic image annotation method. Multimedia Tools and Applications(3). 2019.

8. Kharmyssov C, Ko MWL, Kim JR. Automated segmentation of optical coherence tomography images. Chinese Optics Letters(1). 2019.

9. Hodneland Erlend, Dybvik Julie A, Wagner Larsen Kari S, Šoltészová Veronika, Munthe Kaas Antonella Z, Fasmer Kristine E, et al. Automated segmentation of endometrial cancer on MR images using deep learning. Scientific Reports(1). 2021.

10. Thomas Czerniawski & Fernanda Leite. Automated segmentation of RGB-D images into a comprehensive set of building components using deep learning. Advanced Engineering Informatics. 2020.

11. Z. Chen, M. Rigolli, D. Vigneault, F. Contijoch. Automated Multi-Chamber Segmentation And Imaging Plane Re-slicing Of Cardiac CT Images Via Deep Learning. Journal of Cardiovascular Computed Tomography(3). 2020.

12. Anamaria Radoi,Mihai Datcu. Multilabel Annotation of Multispectral Remote Sensing Images using Error-Correcting Output Codes and Most Ambiguous Examples. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing(7). 2019.

13. Kaleem Razzaq Malik,Muhammad Asif Habib,Shehzad Khalid,Mudassar Ahmad,Mai Alfawair, Awais Ahmad,Gwanggil Jeon. A genericmethodology for geo-related data semantic annotation. Concurrency and computation: practice and experience(15). 2018.

14. Gu H, He Z, Wei D, Elhanan G, Chen Y. Validating UMLS Semantic Type Assignments Using SNOMED CT Semantic Tags. Methods of information in medicine. 2018; 57(1): 43-53.

15. Malik KR, Habib MA, Khalid S, Ahmad M, Alfawair M, Ahmad A, Jeon G. A generic methodology for geo-related data semantic annotation. Concurrency and Computation: Practice and Experience(30). 2018.

16. Sleimi Amin, Sannier Nicolas, Sabetzadeh Mehrdad, Briand Lionel, Ceci Marcello, Dann John. An automated framework for the extraction of semantic legal metadata from legal texts. Empirical Software Engineering(3). 2021.

17. Björn Technau. The Multi-Component Model for the semantic analysis of slurs. Pragmatics and Society(2). 2020.

18. Ganglong Fan, Jian Shen. Application of Big Data Computing Model Based on Semantic Data Analysis Technology. IOP Conference Series Materials Science and Engineering. 2020; 750: 012071.

19. Yong Fang, Cheng Huang, Yu Su, Yaoyao Qiu. Detecting malicious JavaScript code based on semantic analysis. Computers & Security(93). 2020.

20. Gong Y, Lu N, Zhang J. Application of deep learning fusion algorithm in natural language processing in emotional semantic analysis. Concurrency and Computation: Practice and Experience(31). 2019.

21. Vo AD, Nguyen QP, Ock CY. Semantic and syntactic analysis in learning representation based on a sentiment analysis model. Applied Intelligence(3). 2020.

22. Teplitzky BA, McRoberts M, Ghanbari H. Deep learning for comprehensive ECG annotation. Heart rhythm. 17(5 Pt B): 881-8.

23. Lahiani A, Gildenblat J, Klaman I, et al. Generalizing Multistain Immunohistochemistry Tissue Segmentation using End-to-End Color Deconvolution Deep Neural Networks. IET Image Processing(7). 2019.

24. Chunde Yang, Qingshui Liu1, Zhangyong Li, Wei Wang, Xinwei Li. End-to-end bone age assessment based on Attentional Region Localization. Journal of Physics: Conference Series(1). 2021.

25. Twinanda AP, Yengera G, Mutter D, Marescaux J, Padoy N. RSD-Net: Learning to Predict Remaining Surgery Duration from Laparoscopic Videos Without Manual Annotations. IEEE transactions on medical imaging. 2019; 38(4): 1069-78.

26. Yulei Niu, Zhiwu Lu, Ji-Rong Wen, Tao Xiang, Shih-Fu Chang. Multi-Modal Multi-Scale Deep Learning for Large-Scale Image Annotation. IEEE Transactions on Image. 2019.

27. Hao Cui, Jian Li, Qingwu Hu, Qingzhou Mao. Real-Time Inspection System for Ballast Railway Fasteners Based on Point Cloud Deep Learning. IEEE Access(8). 2020.

28. Gao Chaoqian; Hu, Chen; Li, Tianping. Target Tracking Based on Camshift Algorithm and Multi-feature Fusion. Journal of Physics: Conference Series(3). 2020.

29. Yang Feng, Cong Rikun, Xing Mengmeng, Ding Bo. Study on AO classification of distal radius fractures based on multi-feature fusion. Journal of Physics: Conference Series(1). 2021.

30. Ruihua Li, Haojie Gu, Bo Hu, Zhifeng She. Multi-feature Fusion and Damage Identification of Large Generator Stator Insulation Based on Lamb Wave Detection and SVM Method. Sensors(17). 2019.

31. Yunhua Wu, Nan Yang, Zhiming Chen, Bing Hua. Multi-Feature Fusion Based Relative Pose Adaptive Estimation for On-Orbit Servicing of Non-Cooperative Spacecraft. Journal of Harbin Institute of Technology (New Series)(6). 2019.