Research and implementation of convolution k-means algorithm based on unsupervised learning

Huang Suyu1,a

1School of Computer Science, Wuhan Donghu University, Wuhan, Hubei, China

aemail: huangsy@wdu.edu.cn

Abstract: This paper proposes a framework combining the advantages of unsupervised learning, convolutional neural network and K-means clustering algorithm. When there are few label data available, convolution k-means algorithm is formed by combining k-means algorithm with convolution neural network. Through the learning of convolutional neural network and clustering analysis, as well as the continuous optimization and adjustment of the network, the convolutional k-means algorithm is used to train the deep convolutional network, and the hierarchical function of unsupervised learning technology is used to reduce the dependence on a large number of labeled data. Through the analysis of the experimental results, compared with the supervised learning convolution k-means algorithm, the unsupervised learning convolution k-means algorithm can better represent the data clustering with the increase of the number of filters, and can improve the accuracy of test classification. The unsupervised learning convolution k-means algorithm is better than other unsupervised learning filter methods.

1. Introduction

With the rapid development of information network technology and the widespread popularity of personal computers, the contradiction between the exponential growth of Internet data and the ability of computer to process data has become increasingly prominent. It is more and more difficult for people to get the real useful information, so data mining technology came into being. Data mining is a process of extracting hidden, unknown and potentially valuable data information from a large number of incomplete, fuzzy, noisy and random data. The traditional data mining algorithm uses the data warehouse model to summarize all the data to the central node, and runs the data mining algorithm on the basis of summarizing the data. This method has many limitations, such as the performance bottleneck of the central node, data privacy leakage and so on, which can not adapt to massive data. In order to alleviate these problems, distributed data mining has become a research hotspot. Distributed data mining method is to solve some complex problems that can not be completed in a reasonable time in a single machine environment.

K-means algorithm is a common data mining algorithm, which is widely used in numerical clustering, map clustering, text clustering and image clustering. The inherent disadvantage of K-means algorithm is that it needs to determine the number of clusters K in advance, and the determination of K is easily affected by subjective factors, which will lead to local optimization and the decline of clustering quality.

In essence, deep learning belongs to feature learning, which processes the original data through relatively low-level nonlinear model into higher-level and higher-level expression to achieve the purpose of optimal feature. Deep learning has attracted the attention of scientists all over the world. In
2006, Jeffrey Hinton and Ruslan formally proposed a useful training method in feed forward neural network. This algorithm regards each layer of the network as an unsupervised restricted Boltzmann machine, and can be optimized by using a supervised back propagation algorithm. In 2009, deep Boltzmann machine\cite{1} theory was proposed, which allows two-way connection in the bottom layer.

In large data sets, deep neural networks need to train a large number of data. In most cases, many algorithms can learn deep-seated feature networks from unlabeled image data, but these features are difficult to train and adjust, and are difficult to be effectively applied in practice. Kai et al Proposed a fast alternative method K-means clustering algorithm\cite{2}. This method is a bit fast and easy to implement on a large scale, but the accuracy cannot be well guaranteed. This paper proposes an unsupervised learning deep convolution network training method based on the improved k-means clustering algorithm. This method improves the accuracy of clustering by reducing the number of related parameters in the form of similar filter. The experimental results show that learning the relationship between the layers of deep convolutional neural network can improve the training ability on a small number of labeled data. The experimental results show that the convolutional k-means algorithm has improved the test accuracy in the data set, which is obviously better than other unsupervised learning filter methods.

2. Cluster algorithm and Convolutional neural network

2.1 Introduction to k-means algorithm

Partition based clustering method is a relatively widely used clustering method. K-means algorithm has the advantages of low time complexity, strong scalability and distributed computing, so it has a very wide range of applications in different fields. At the same time, K-means algorithm is also the research object of this paper. The core idea of K-means algorithm is: for a given data set containing N data objects, K-means clustering algorithm first randomly selects K objects Then, according to the given similarity measure, all the data objects in the data set are summarized into the clusters represented by the most similar cluster center points. Then, the center point is updated according to the average value of each clustering data object, and the data object is re divided, and the process is repeated until the data collection is completed. The clustering of data objects in this set does not change or satisfies other termination iteration conditions.

Compared with other clustering algorithms, K-means algorithm is more flexible and efficient in dealing with large data or data sets. Especially when dealing with convex clustering, the result is better. When the clusters in the result are dense and there is a big difference between clusters, the clustering effect of K-means algorithm is obviously better than other algorithms.

2.2 Convolutional neural network

Convolutional neural network can use different combination of neurons and learning rules. Convolutional neural network is composed of input layer, hidden layer and output layer. The input layer of convolutional neural network is usually input in matrix form. The hidden layer can be divided into volume layer, pool layer, and full connection layer. Convolution layer is the most important part of convolution network. It extracts the input data.

After the convolution layer, we need to talk about the pooling layer. After the data features are successfully extracted, the generated feature map needs to go to the pooling layer for feature selection and information filtering. Using the same filling makes the convolution layer and pooling layer have the same size feature map. Image features can be extracted by multiple convolution layers and pooling layers, and then the image features extracted by full connection layer can be classified. Finally, it is connected to the output layer. In the image classification problem, logic function can be used to output classification label. Such a relatively complete convolutional neural network framework came out. Convolutional neural network is a multi-layer neural network. In many components of the network, a single layer is composed of two-dimensional plane, and then the plane is composed of many individual neurons.
With the progress of science and technology, deep learning gradually affects our lives. It is mainly used in image recognition, speech recognition, etc. In image recognition \[^4\], firstly, the input image information is processed and recognized by multi-level convolution neural network. At the bottom layer, convolution starts from the original image, and then describes the local edge and problem features of the image. After that, the information from the upper layer is integrated and abstracted to a higher level through the middle layer. Second, in the aspect of language recognition \[^5\], the first step is to input the sound signal into the computer. Then, the digital sampling is converted into an easy-to-handle electrical signal, and the electrical signal is input into the neural network for further processing after anesthesia.

3. Convolution K-means clustering algorithm based on unsupervised learning

3.1 Data initialization
The usual K-means clustering algorithm needs some considerations, especially the initialization of centroid. Although it is very common to initialize k-means algorithm with randomly selected samples from data, the results are often unsatisfactory. Images may be too densely grouped in some areas. Therefore, initializing k-means algorithm with randomly selected image blocks leads to a large number of centroids leaning together, which leads to many clusters in these centroids becoming almost blank. Because a single cluster includes all the points in a dense area. On the contrary, it is better to initialize the centroid randomly from the normal distribution, and then normalize it to unit length. Because the data has been whitened, it is hoped that the important part of the data has been re scaled or distributed as a sphere. Then the result of randomly selecting the center of mass on a sphere is obviously optimized. Another small adjustment of the improved algorithm is to use the damping update of the cluster center. This form of damping has little effect on the larger clusters, and only prevents the smaller clusters from being pulled too far in a single iteration.

3.2 Feature extraction and classification
After data preprocessing, unsupervised learning algorithm is used to extract features from unlabeled data. The process is implemented by K-means algorithm, sparse self coding, sparse constrained Boltzmann machine and Gaussian mixture unsupervised learning method.

3.2.1 Traditional K-means learning filter
The filter is learned by K-means algorithm. The classical k-means algorithm is used to find the clustering center with the minimum distance between the points in Euclidean space. Usually, these points are randomly extracted image blocks, and the center point is the filter used to encode the image. At this time, the k-means algorithm learns the dictionary algorithm from the data vector and looks up the dictionary. The algorithm is as follows:

\[
s_j^{(i)} = \begin{cases} D(j)^T w(i) & \text{if } j = \arg\max |D(j)^T w(i)| \\ 0 & \text{otherwise} \end{cases}
\]

\[
D = WS^T + D
\]

\[
D(j) = \frac{D(j)}{\|D(j)\|_2}
\]

The traditional K-means clustering algorithm is to learn and train the centroid of each cluster in the image block. In convolutional network, the filter is applied to the image by convolution. Most of the image blocks have the same direction after training by K-means clustering algorithm, and are parallel to each other. After convolution, redundant feature graphs will be generated in the position not far away. Therefore, in order to change this phenomenon, some small improvements are needed to convolution K-means clustering algorithm.

3.2.2 Learning by convolution means filter
Because many filters generated by adjacent positions are useless, the standard k-means algorithm will extract random modules from the input image, and the size of these modules matches the size of the
Thus, change the size of the window to delete useless image blocks, and use the remaining image blocks for clustering. Size the filter to about half the size of the window. Then the input image is randomly selected, and the center of K-means clustering algorithm is used to convolute the filter, and the similarity measure outside all positions is obtained by calculation. The active image block is similar to the centroid of K-means clustering. Finally, a specific image block is extracted from the window and classified into the corresponding centroid. The improved algorithm is as follows

\[
S_j^{(i)} = \begin{cases} 
D^{(i)}W^{(i)}_{(x,y)} & \text{if } (j, x, y) = \arg \max |D^{(i)}W^{(i)}_{(m,n)}| \\
0 & \text{otherwise}
\end{cases}
\]

\[
D = W_{(x,y)}S^T + D
\]

When the redundant filter is removed, the remaining filters can better learn the features. In contrast, the filters trained by convolution k-means algorithm are more diversified.

4. Experimental results and analysis

In the whole experiment, stl-10 data set is used. The data set contains 96 × 96 RGB images of multiple categories. Each category in the data set has 600 images for training and 900 images for testing. In the experiment, although the data is not down sampled, the dimension of the first layer is effectively reduced by increasing the step size slightly. The dimension can be effectively reduced to K × 2 × 2 by maximum pool processing after convolution layer. Finally, the correction linear cell activation function is used. In order to compare the effectiveness of the filter trained by K-means algorithm and convolution k-means algorithm, a large pool is set to reduce the dimension to K × 2 × 2, and 0.1 learning rate and 0.9 momentum rate are used in the experiment.

![Figure 1 relationship between precision and number of filters](image)

Figure 1 relationship between precision and number of filters
In the experiment, the k-means algorithm and convolution k-means algorithm are compared. In the relationship between the precision and the number of filters, the size of the filter is set to $3 \times 11 \times 11$, and the number of filters is changed. It can be seen that increasing the number of filters can improve the performance of the two algorithms. The filtering performance of convolution k-means algorithm is better than that of K-means algorithm. By changing the size of the filter, it is found that the accuracy of K-means algorithm and convolution k-means algorithm improves with the increase of the size of the filter. It can be seen that under the same conditions, the accuracy of convolution k-means algorithm is significantly better than the traditional K-means algorithm. Relationship between precision and filter size as shown in Figure 2.

5. Conclusion
Conclusion this paper proposes a framework combining the advantages of convolutional neural network, unsupervised learning and K-means clustering algorithm. When there is little tag data available, the framework and convolutional neural network are used together, and the K-means clustering algorithm is modified. By comparing other unsupervised learning algorithms of convolution k-means algorithm, the experimental results show that with the increase of filter number, the supervised algorithm will lose accuracy and does not match the training set, while the unsupervised learning algorithm convolution k-means algorithm can better represent data clustering. The accuracy of convolution K-means clustering and other unsupervised learning algorithms in stl-10 data set is compared. The accuracy of convolution k-means algorithm in stl-10 data set is achieved. The results show that convolution k-means algorithm is better than other unsupervised learning filter methods in stl-10 dataset.

Acknowledgments
Project of Hubei Natural Science Foundation in 2018. (Project Number:2018CFC877).

Reference
[1] Krizhevsky A, Sutskever I, Hinton G. (2012) Image Net Classification with Deep Convolutional Neural Networks. NIPS. Curran Associates Inc.
[2] Jia Deng, Wei Dong, Richard Socher, et al. (2009) Image Net: a Large-Scale Hierarchical Image Database. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 20-25 June, Miami, Florida, USA. IEEE
[3] Cai Weihong, Liu Zhen. (2005) Research on Intrusion Detection Based on density clustering algorithm. Computer engineering and application, 152-154.
[4] Min Lequan, Kan Zhiqiang, Wang Jingtao. (2011) A new CNN template for processing gray images of deformable convex objects. Journal of Guangxi Normal University: 56-60.
[5] Zhang Jianhua. (2015) Application Research of speech recognition based on deep learning. Beijing University of Posts and telecommunications.