Research on Convolutional Neural Network Image Recognition Algorithm Based on Computer Big Data

Yuanyi Chen

Abstract. Aiming at the problems of slow convergence and low recognition accuracy of the existing image big data recognition methods, a convolution neural network based image big data recognition method is proposed. In the concept of machine vision, image recognition refers to the ability of software to distinguish characters, positions, objects, actions and handwriting in pictures. Computers can apply machine vision techniques, combined with artificial intelligence and cameras to image recognition. Describes the convolution layer, pool layer, full connection layer and classification layer in the model. MNIST data sets are simulated to verify the superiority of the proposed method. Convolutional neural network has achieved good results in the field of image recognition, but its network structure has a great impact on the effect and efficiency of image recognition. In order to improve the recognition performance, a new convolution neural network structure is designed and implemented by repeated use of smaller convolution kernels. The results show that compared with other similar algorithms, the false recognition rate of this method is low, and the lower average square error of classification can be realized in a short time. This paper expounds the basic flow of digital recognition system from the aspects of data extraction and preprocessing, feature extraction and selection, classifier design and so on[1].

Keywords: Mnist; Deep Learning, Classifier Model Applied Analysis, Recognition Accuracy

1. Background information
The human brain receives a lot of information every moment and can process and save it[2]. Image recognition is not an easy task, a good way is to apply metadata to unstructured data. Hiring experts to manually label music and film libraries may be a daunting task, but some challenges are almost impossible, such as teaching driverless car navigation systems to distinguish passers-by from a variety of motor vehicles, or to label and classify millions of videos or photos that users send to social media every day[3]. The next time you use it, you can read it out and use it. Artificial neural network is an artificial intelligence method which imitates the structure and unit of human brain. Face recognition has become one of the research hotspots in the field of computer vision and machine learning. It is
also one of the most successful applications in the field of image analysis and understanding. Convolutional neural network is a further improvement of artificial neural network. It is a special multilayer perceptron designed to recognize two-dimensional images. The original image does not need too much preprocessing to learn the invariant features of the image. The face recognition system mainly compares a detection image with the training image in the database to obtain the recognition results. In the common face recognition algorithms, due to the influence of human facial expressions, ornaments and environmental factors. In the concept of machine vision, image recognition refers to the ability of software to distinguish characters, positions, objects, actions and handwriting in pictures. Computers can apply machine vision techniques, combined with artificial intelligence and cameras to image recognition. The most important step is how to describe and select appropriate face image features. In order to improve the accuracy of face recognition and retrieval. The complexity of face recognition is that facial features change with external and internal conditions. Image recognition algorithm generally uses machine learning method to simulate the recognition of human brain[4]. According to this method, we can teach computers to distinguish visual elements in images. The computer relies on a large database. By recognizing the pattern of data presentation, the image can be understood and then related labels and categories can be formed.

2. CNN details

2.1. Constructional layers
Convolution layer is the local connection of precursor layer characteristics by simulating simple cells. Different feature maps are extracted by convolution operation of the input image, and then each unit of the current layer is connected with local blocks of the feature map of the front layer by weight. Local weighting and nonlinear transformation[5]. Through multi-layer convolution and downsampling, 128-dimensional feature vectors can be obtained for every 7 consecutive frames of images. The number of units in the output layer is the same as the number of video actions, and each unit of the output layer is fully connected to the 128-dimensional feature vector. The 128-dimensional feature vector is classified by linear classifier, and all the trainable parameters in DCNN model are initialized randomly, and then trained by online BP algorithm.

![Figure 1. The common neural network.](image)

2.2. Local perception
According to the forward transmission process of BP network signal, we can easily calculate the output of network nodes[6]. For example, for the net input marked as a red node in the figure above, it is equal to the accumulation of the product of all the node values of the upper layer of neurons connected to the red line and the weights represented by the red line. The calculation process will also be 102 times slower. The essence of convolution layer is to solve the problem of computational explosion. It is implemented by limiting the connection between the hidden unit and the input unit: each hidden unit can only connect part of the input unit. The input region size of each implicit unit connection is called the receptive field of that neuron (receptive field). Since the neurons in the convolutional layer are three-dimensional and deep. The parameters of the convolutional layer contain a series of filters. Each filter trains a depth. the area connected to the input image is the same at different depths of the output unit, but the parameters (filters) are different. although each output unit
only connects part of the input. The value is calculated by the weight and the dot product of the input plus bias, which is the same as the common neural network, as shown in the figure below. Such a calculation process, many books call it convolution. In fact, for digital filtering, the coefficients of its filters are usually symmetric. Otherwise, the calculation of convolution requires reverse folding and then multiplication and accumulation. Does the above neural network weights satisfy symmetry? I think the answer is no! Therefore, the above call it convolution operation, is obviously biased. But this is not important, just a noun. However, signal processing people, in the first contact with convolutional neural networks, brought some misunderstandings.

2.3. CNN development prospects
In recent years, deep learning technology has broken through the three-layer structure based on the artificial neural network model. Its capacity to express itself has increased substantially, and success in the field of computer vision. The deep learning network mainly includes convolution neural network, cyclic neural network, self-encoder and long-term memory network. Constructional neural networks are particularly suitable for processing two-dimensional data. There are many successful applications in the field of image processing and recognition. CNN will expose its defects in the face of accurate spatial relations. For example, the position of the mouth in the face image is placed on the forehead, CNN will still recognize it as a face. Another major problem with CNN is the inability to understand new ideas. Hackers can confuse CNN judgment by making some subtle changes. It is also the first network to solve important business applications. cyclic and recurrent networks are suitable for processing time-dependent sequential data. Computer hardware is evolving, The deep learning neural network is very impressive in computer vision, Breakthrough progress has been made in image retrieval and image classification. In 2012, For the first time, Hinton professor at the University of Toronto applied deep neural networks to image classification tasks, The classification error rate of the image is greatly reduced. Hyperautomation is a collection of machines, software packages and automation tools for delivering work. Superautomation contains not only a rich combination of tools, but also all the steps of automation itself (discovery, analysis, design, automation, measurement, monitoring and re-evaluation). The main focus of hyperautomation is to understand the scope of action of automation steps, their correlation with each other, and their combination and coordination. The data set it USES is Image Net. The team led by Professor Li Feints consists of 16 convolution/full connection layers in the VGG-16 structure and 19 convolution/full connection layers in the VGG-19 structure. All have achieved good application in recognition. Compared to traditional methods. Deep learning can achieve higher recognition accuracy. However, there are still some problems, such as long training and testing time and large dictionary storage, which need further study and improvement. In order to reduce its scale, the VGG deep learning network is deeply analyzed and the trend starts with robot process automation (RPA. But robotic process automation alone is not hyperautomation. It requires a combination of tools to help replicate the human part of the task process.

![Figure 2. The common neural network.](image-url)
3. Image object classification

Given a set of images all marked with a single category, we ask for a new set of test images to predict what category they are? And measure the accuracy of the prediction. This task involves a variety of challenges, including viewpoint change, scale change, in-class change, image deformation, image occlusion, illumination conditions, background clutter, etc. The image recognition technology based on neural network mainly introduces neural network algorithm, which is more powerful than the traditional image recognition method. It is a new image recognition technology. IBP neural network image recognition method does not need to extract image features, but only needs to input image data in neural network recognizer. Network training and testing using MATLAB show that the neural network image recognition system without feature extraction has strong anti-interference ability. This paper introduces the structure and algorithm of image recognition technology based on neural network. Computer vision researchers have proposed a data-driven approach to solve this problem. Instead of trying to specify each image category of interest directly in the code, they provide the computer with many examples of each image class, and then develop learning algorithms to see these examples and understand the visual appearance of each class. In other words, they first accumulate the training dataset of labeled images and then provide it to the computer to familiarize it with the data.

![Convolutional neural network structure.](image)

4. Common CNN architecture

Machine learning algorithm has important theoretical and practical significance for artificial intelligence system. In the past ten years, deep learning algorithms and models represented by deep neural networks have made remarkable achievements in feature selection and learning tasks. Breakthrough progress has been made in many fields, such as image recognition, speech recognition, machine translation and so on. The data processing ability in some tasks even surpasses the artificial level. The deep neural network model shows excellent performance in AI applications, but there are still some difficult problems in feature sparsity, scale invar, channelization co variate offset. Focusing on the above technical problems, this paper has carried out relevant research on the basis of previous work and has achieved some meaningful results. Main research work and innovation points are as follows: 1. The connection algorithm is extended to improve the mask generation strategy, where the mask generation function depends on the output value of the neurons in the previous layer. LeNet-5 is the most commonly used one, it is the simplest network architecture. It has 2 convolutional layers and 3 fully connected layers (total 5 layers, this nomenclature is common in neural networks, and this number represents the sum of convolutional and fully connected layers). Average-Pooling layer, which we now call the sub-sampling layer, has some trainable weights (now not common when designing CNN networks). The network architecture has about 60,000 parameters. This network architecture has become a standard "template": stack-type convolution and pooling layers with one or more fully connected layers as the end of the network.
5. Common CNN architecture

References
[1] Zhuang Fuzhen, LUO Ping, He Qing, et al. Research progress in Transfer Learning [J]. Journal of Software, 2015, 26(1): 2639.
[2] Li Yandong, HAO Zongbo, LEI Hang. [J]. Journal of Computer Applications, 2016, 36(9):2508-2515.
[3] Gonzalez, Woods, Edings, etc. Digital image processing :MATLAB [M]. Electronic Industry Press, 2005.134-140.
[4] Long Mingsheng. Research on transfer Learning Problems and Methods [D]. Doctoral dissertation. Beijing: Tsinghua University, 2014.
[5] Grauman K, Leibe B. Visual object recognition[J]. Synthesis lectures on artificial intelligence and machine learning, 2011, 5(2): 176-181.
[6] Andreopoulos A Computer Vision,Tsotsos J K.50Years of object recognition: Directions forwarder[J].&Image Understanding, 2013, 117(8): 827 — 891.