

Research on Text Detection of Network Advertising Images Based on Deep Learning

Pokwai To^{1, a}

¹ School of Astronautics, Beihang University, Beijing, 100191, China

Keywords: Text detection, Network advertising images, Deep learning

Abstract. The text detection of advertising images is of great significance in current economic life. The traditional method of text detection often omits words or wrongly detects the text. In this paper, the deep learning algorithm is introduced to detect the text in the online advertising images. Based on the algorithm of restricted Boltzmann machine, this paper improves the accuracy and speed of the traditional algorithm.

Introduction

Currently, the words often appear in the advertisement images with very important information. How to make the computer to detect the text in the image is a hot direction in the field of computer vision. The text in the image is affected by various factors, such as brightness, shadow, font, size, arrangement and so on. It is very important to accurately locate the text area for subsequent text extraction and recognition. With the popularity of the network, information dissemination is more and more dependent on the Internet, compared to traditional advertising, network advertising spread fast, wide audience, low prices, with obvious advantages. Network advertising than traditional advertising and more interactive, and can be targeted for transfer to the target audience, at the same time, timely updates faster, more flexible, more and more emphasis on business, especially the information industry related companies, such as some games mostly use the Internet website publicity, through pictures and other forms of advertising to create a large number of the spam, which greatly affect the user's normal use. According to the user's browsing and performance of different forms of online advertising interference, precision, implantable, text link, picture, animation and video forms. In network advertisement, advertising images have attracted much attention because of its rich and attractive appearance, these advertisements in benefits for businesses but also affected the normal use of ordinary users, but also increased the burden of commercial image search engine retrieval. In the specific application of many such as license plate extraction, text in natural scene recognition and image retrieval and illegal text in the network monitoring system detected in the picture to purify the network environment, text positioning is the key process. In the modern economy, the advertising industry is a highly independent information industry. In this special industry, as the product of information and advertising carrier, by strengthening the means to constantly promote a large number of information transmission in various fields of social and economic life, the commodity economy and the prosperity of the market plays an important role in promoting. With the development of science and technology and information technology, especially the development of computer technology, network technology and storage technology, multimedia information data in modern technology has been the use of various means large acquisition and production of various types of. In order to be able to detect a large number of ads in the advertising information, advertising companies usually use manual detection methods. Manually through the web page advertising pictures, identify the contents of the text in the picture. Through the statistics of the number of ads, put the location of such information to analyze whether the current advertising is correct. Advertising companies in advertising, advertising may be expected to launch a big gap. For example, advertising requirements in the layout of a specific location, the emergence of advertising refresh frequency and the emergence of other types of advertising. Due to the large number of advertising, wide range and complex content, the efficiency is low and the cost is high. In view of the problem of text detection in advertising images, this paper puts forward a system for automatic

detection of online advertising image detection. By comparing and identifying with a given image, select the desired image in the selected page. In the selected advertisement picture, we need the content of the advertisement text, and then we can identify the position of the text by comparing with the content in the text library. As the text in the advertisement may be processed by the arts, all the recognized texts are selected as the selected text area.

Deep Learning and Restricted Boltzmann Machine

Deep Learning. The concept of deep learning comes from the study of artificial neural network. Multilayer perceptron with multiple hidden layers is a kind of deep learning structure. In depth learning, a more abstract high-level representation of attribute categories or features is formed by combining low-level features. The concept of deep learning was proposed in 2006 by Hinton. This paper proposes an unsupervised training algorithm based on deep belief network, which brings the hope to solve the problem of deep structure optimization. In addition, the convolutional neural network proposed. It is the first truly multi-layer structure learning algorithm, which uses the spatial relative relation to reduce the number of parameters to improve the training performance. Deep learning is a new machine learning field in the study of the motivation is to establish and simulate the human brain to analyze neural network learning mechanism, it imitates human brain to interpret the data, such as image, sound and text. In many cases depth, it is sufficient to represent any function with a given target accuracy. However, the price is that the number of nodes in the graph may become very large. The theoretical results confirm the fact that the number of nodes that are needed to grow exponentially with the size of the input. We can see the depth structure as a factorization. Most randomly selected functions cannot be represented efficiently, either with deep or shallow architecture. However, many can be effectively expressed by the depth of the structure cannot be used to express the effectiveness of shallow architecture. The existence of a compact and deep representation means that there is a structure in a function that can be represented. If there is no structure, it will not be possible to generalize well.

Restricted Boltzmann Machine. The restricted Boltzmann machine is a special topological structure of Boltzmann machine. The principle of the origin of Boltzmann machine in statistical physics, is a kind of modeling method based on the energy function, which can describe the interaction between the variables of higher order learning algorithm of Boltzmann machine is very complex, but the model and learning algorithm of physical explanation, more complete and strict mathematical statistics theory foundation. Boltzmann machine is a symmetric random feedback type coupling two value unit neural network, consisting of visible layers and a plurality of hidden layer nodes, the network is divided into visible units and hidden units, with the visible units and hidden units to express the learning model of random network and random environment. The restricted Boltzmann machine is a stochastic generative neural network which can learn the probability distribution of the input data set. Deep trust network and deep Boltzmann machine, which is composed of many layers of neurons. It has been used in many machine learning tasks, which can solve some complex problems and improve learning performance to some extent. The deep neural network is composed of a number of restricted Boltzmann machine stacks, and there is no connection between the visible and hidden neurons in the restricted Boltzmann machine. The deep neural network is trained by restricted Boltzmann machine, which is based on the hierarchical unsupervised greedy pre-training method. The results obtained are as the initial value of the supervised learning probability model, and the learning performance is greatly improved. Unsupervised feature learning is a statistical modeling of the complex hierarchical structure and massive data sets. Through unsupervised pre training network to obtain higher order abstract features, and provides better initial weights, the weights defined in the favorable range of global training, training in the use of local information by layer between the layers, focus on the characteristics of the training data itself, can reduce the risk of overfitting learning objectives, and to avoid cumulative error deep neural network transfer problem of long. restricted Boltzmann machine due to the structural element representation ability and easy reasoning has been used successfully as the depth of the use of the neural network,

and has received extensive attention in recent years. As a practical application, the restricted Boltzmann machine learning algorithm has shown the superior performance of learning. The restricted Boltzmann machine learning occupies the core position in the study of deep neural network.

Text Detection Algorithm of Network Advertising Images Based on Restricted Boltzmann Machine

Image Pretreatment. In this paper, we need to carry on a series of pretreatment to the digital image, including the image gray transformation two value processing text block localization and the text edge detection and so on. We will transform the color image to grayscale pixel contains only the luminance information, and the background of redundant information in digital images and other noise removal, so as to provide convenience for the subsequent block of text segmentation and text extraction. The noise of digital image text extraction and recognition is mainly due to the illumination factors, effective image information leads to color and other factors of interference generated by the noise removal mainly used filtering methods, including mean filter and median filter, the noise removal in digital image pixel. The background and foreground color image by digital image by gray processing after the separation, the general method of gray pixel threshold for separation, including the global threshold separation method of local threshold separation method and dynamic variable threshold separation method. Due to the large amount of calculation and the low efficiency of the dynamic threshold separation method, the separation effect of the global thresholding method is poor. Therefore, the algorithm uses the local optimal threshold separation method, through the calculation of the different pixels in the digital image of the gray level gradient value of the pixel separation. The two value image processing features of the region, through the difference changes, the digital image gray color variation and texture features were detected, so as to realize the separation of the block of text and image background. In this algorithm, the first order and two order differential operators are used to calculate the gray value of the image pixels so as to realize the edge detection in different regions.

Layer Optimization. We can divide the gray image of the original digital image into several layers. Due to the central character of the digital image, we can get the final image layer by calculating the distance of each image pixel. In this algorithm in order to improve the accuracy of the layer, the layer selection process, the first layer in accordance with the size of the center distance is divided into two sub sets the connected domain analysis operation using pixel noise removal, then select the minimum distance of image layers as the final selected results, this layer contains to the block of text extraction and recognition, character segmentation can be carried out on it. In this paper, supervised learning. Given a training set of R , we mark it in the training set R contains 3 kinds of picture styles. One is a text free picture (marked as 0), one is a text containing the picture (marked as -1), and the other is a picture containing only text (marked as 1). Then, the two-layer restricted Boltzmann machine is used to study the distribution characteristics of the training set R samples, so as to get the parameters of each node and the node with minimum energy function of restricted Boltzmann machine. At the top level, we use the classifier to classify the image style according to the information. In the two-layer, the restricted Boltzmann machine, compared with the single layer limited Boltzmann machine, is enhanced, and the generalization ability and classification accuracy are improved. In the leaning process of training, the purpose is to learn the characteristics of the data and get the minimum energy function parameters.

Words Detection. Text segmentation is the process of removing the text information from the text outside the text area, and then dividing the text into a small area with a single character. In this paper, we use the projection method and the region generation method to extract the characters in the text area, in which the horizontal projection is used to segment the text, and the region generation method is used to segment the text. In the area of the image obtained after threshold segmentation in may also contain some noise pixels, to obtain segmentation results are also included in the pixel part does not belong to the text area, so the need for these noise pixels are removed after processing, segmentation

can be obtained after the final text image recognition. At present, the optical character recognition technology has been improved. On the writing materials were scanned and converted into computer code can be unified identification and storage has achieved great success. Therefore, the relevant literature and the vast majority of the use of the OCR software modules have been commercialized to convert code character recognition, but because of the low-resolution text area and complex background, so the video text effect after the two values and print scan binarization result is large, and because of the OCR system the current is mainly for scanning the printed text, so the recognition effect is not ideal. To solve this problem, we can take two kinds of programs, the traditional scheme can be used to segment the text, with the contrast of the text in the library to compare the characteristics of the word will identify the word and sort. We propose a method using deep learning of OCR, through repeated modification to expand the content of the machine learning ability. Machine learning methods can increase the sample library of feature parameters, after a period of learning machine can automatically identify text content. Even if the background interference, after a period of time to learn the correct results can be obtained. At present, some systems use a dictionary method to improve the recognition effect, and achieved some results, but in general, text recognition conversion, especially the low-resolution text image recognition, still need further research, in order to improve the recognition rate.

Conclusion

The detection of the words in the advertisement image is very valuable. Although the human eye can easily identify these words, the acquisition of these words is also facing great difficulties for the computer. The existing work has made useful attempts to solve these problems and solved the problems to some extent. However, the results are still far away from the practical application, needing the coordinated efforts form the scholars at home and abroad.

References

- [1] Ye Dan, Zhou Jing. Research on Image Spam Text Area Detection Method [J]. Designing Techniques of Posts and Telecommunications, 2013(12): 76-79.
- [2] Meng Li, Li Yuelong, Chen Gang. TV Advertisement Retrieval Based on Shot Change and Text Detection [J]. Journal of Jilin University (Science Edition), 2012, 50(6): 1185-1191.
- [3] Wu Hui, Zhao Yuqian, Li Xianghua, Zou Beiji, Zou Runmin. Text detection from natural scene images based on color clustering [J]. Journal of Central South University (Science and Technology), 2015, 46(6): 2098-2103.
- [4] Fei Mengyuan, Chuang Yuelong, Guo Wenping. Image text region detection based on Android [J]. Information Technology, 2016 (6): 103-106.