Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2022.030107
ISSN 2790-2692 Vol. 3, Issue 1: 76-83

Clothing Feature Recognition and Classification Based on Convolutional Neural Network

Fawaz Asghari*

Islamic Azad University, Iran

**corresponding author*

Keywords: Convolutional Neural Network, Feature Recognition, Clothing Classification, Attribute Prediction

Abstract: With the increasing number of people buying clothing online, the experience of customers in the process of purchasing clothing products online becomes particularly important. Commodity search, as one of the steps in the purchasing process, plays a pivotal role in a good shopping experience. However, whether the clothing picture establishes the correct product label and the accuracy of the label also affects the process of commodity search. This paper mainly studies the clothing feature recognition and classification based on convolutional neural network. Firstly, this paper analyzes the training process of convolutional neural network, and builds the clothing classification attribute prediction network model based on Xception. The experimental results show that the model is better than the common CNN model in clothing image classification and attribute prediction.

1. Introduction

The rapid development of e-commerce has driven the development of the clothing industry. However, the domestic demand market of clothing consumption is expanding from quantity expansion to quality differentiation and diversification, and the clothing design is also expanding from general to diversified direction. People's consumption concept has changed from the common demand to the pursuit of individuality, and their consumption mode of clothing also tends to high-quality consumption, individual consumption and differential consumption [1-2]. In recent years, the offline competition of traditional clothing has extended to online, resulting in the emergence of a large number of original online clothing stores in apparel e-commerce, and the emergence of "Internet celebrity fashion" driven by "Internet celebrity economy", which promotes the formation of stylization of e-commerce clothing [3]. The popularity of the Internet and mobile devices has accelerated the development of the clothing industry. Content-based online recommendation system has become a common way to meet consumers' demand for clothing.

Efficient and accurate identification and classification of clothing images and establishment of corresponding visual labels play an important role in online recommendation system. It can provide consumers with a more convenient way to retrieve their favorite clothing products [4-5]. However, most of the current clothing image will have to rely on artificial classification and annotations, this manual annotation exist obvious drawback is that consumes a lot of artificial cost and low efficiency, and manual annotation subjectivity is strong, each person to give the clothing image information understanding is different, annotation keywords also differ in thousands ways [6]. Therefore, in this paper, style and style recognition and classification of e-commerce images are carried out based on convolutional neural network.

In recent years, the rise of deep learning provides a new research direction for feature extraction and recognition and classification technology. Deep learning is used in a layered model for the simulation of human learning, the use of the feature of its hierarchical structure, the model of the previous layer can draw the image of the underlying characteristics of lower and into the middle can extract less abstract characteristics, through the way of learning step by step using characteristics of the bottom, middle, to the abstract and complex data representation. If labeled data is available, deep learning can be used for supervision problems [8-9]. As one of the representative algorithms of deep learning, convolutional neural network (CNN) has made outstanding achievements in image classification and object detection. Its layer-by-layer increasing pattern and end-to-end structure can be used to learn specific image representation and is widely used in the field of clothing: Some scholars divided the body into four parts (head, body, arm and leg) by posture estimation and used P-CNN partition to extract fabric, texture and color features, and then fused the features. The results were input into Gaussian mixture model to realize the classification of clothing style [10-11]. When observing clothes, the first is the whole and the second is the details, so it is crucial to make a detailed description of clothes. The above method only extracts a few features of clothes, and identifies and classifies based on a single attribute, so it cannot describe the characteristics of clothes well [12-13].

In this paper, clothing images are collected from various e-commerce platforms, clothing image sample data sets are established, the network structure of deep convolutional neural network is optimized and improved, network parameters are fine-adjusted, and the style recognition and classification of e-commerce clothing based on deep convolutional neural network is achieved using scientific and objective methods.

2. Garment Image Feature Classification Based on Convolutional Neural Network

2.1. Convolutional Neural Network Training Process

(1) Back propagation

In the process of neural network training, it is necessary to adjust the parameter optimization network in the neuron, and the optimal network is gradually generated in the process of constant parameter adjustment [14]. At present, most neural networks adopt the back propagation algorithm as the parameter optimization algorithm in the network structure [15].

The essential idea of back propagation algorithm is based on iteration of gradient descent algorithm. Since there is an error between the final calculation result of the neural network and the annotation result of the actual image dataset, the loss function can be used to calculate the error value between the two, and then the parameter values in the network can be iterated in reverse according to the error value until the error value reaches the predicted range and the network converges [16].

The back propagation theory relies on the chain rule, which is a derivative rule of compound functions in calculus. The derivative formula of compound functions is as follows:

$$h(x) = f(g(x)) \tag{1}$$

$$h'(x) = f'(g(x))g'(x) \tag{2}$$

The principle of the back propagation algorithm is to iterate the weight parameters of neurons layer by layer from the output end to the input end of the neural network [17]. The update method of parameters is the inverse of the error value calculated by the loss function, and the parameters are updated along the negative direction of the gradient. The specific formula is:

$$w = w - \eta \frac{\partial E}{\partial w}$$
$$b = b - \eta \frac{\partial E}{\partial b} \tag{3}$$

Where η represents the learning rate, E represents the image loss value, and the score represents the gradient of the weight w and bias B, respectively.

In order to iterate the parameters of neurons in each layer of the neural network model, the backpropagation algorithm uses the chain rule to differentiate the image errors, and then calculates the gradients of the weights w and B on the input connections of the current neurons.

The derivative of the chain rule is a recursive relationship from back to back. According to the given recursion relation, the neuron update gradient of each layer can be calculated layer by layer from back to front, and the gradient of each layer is equal to the product of the gradient of the latter layer and the local partial derivative of the current layer. The network parameters are updated layer by layer from back to front by layer calculation [18].

(2) Prevent overfitting

Overfitting phenomenon is a common phenomenon in the process of network training. When the accuracy performance results in the training set gradually improve, while the accuracy data indicators in the testing set continue to decline, it is called overfitting phenomenon. During supervised training, data sets need to be divided, one part as training set and the other part as testing set. The goal is to minimize the value of loss function obtained in the training process of neural network. In order to alleviate the overfitting phenomenon in the process of network training, the following methods can be adopted:

The amount of clothing image data can be increased by obtaining new clothing images and data on the original images in the data set. Among them, the commonly used data augmentation methods include: adjusting the size of the target object in the clothing image, the brightness of the overall image, the saturation of the overall image and other ways. With the increasing amount of clothing image data, the learning ability of the network structure will also be improved. By constantly learning the parameters in the network, the accuracy of the final image classification task will be improved.

The method of reducing the magnitude of the characteristic variables while preserving the characteristic variables in the network structure is called regularization. The function of regularization parameter is to balance multiple feature targets in a network model, so that the fitting curve of the function is smoother and the generalization ability of the network structure is stronger.

2.2. Xception Feature Network Construction

Whether based on the traditional picture feature extraction method or the convolutional neural network method, the clothing key point detection model generally extracts the feature based on the clothing picture, and then obtains the coordinate position of the key point according to the extracted

picture features. In order to improve the diversity of the network extracted features and supplement the uniqueness of the picture feature information, this paper proposes a method to integrate the prior information of clothing types with the picture feature information, so as to improve the diversity of the network extracted features and improve the detection effect of the clothing key point detection model.

Into the range of prior information is the clothing category information as a word, and then through the sole hot coding way to deal with category information is sparse expression vector, then through word embedded method to convey the sparse vector into expression vector, and the information, the feature vector which is a category of extracting it and to extract the image feature information fusion, Then, the positions of clothing key points are obtained by using the fused features, and the end-to-end clothing key point detection model integrated with the prior information of clothing types is finally built.

In this paper, the fusion of a priori knowledge about the clothing categories EFLAN network structure is based on the coordinate point detection model regression methods to fixing the clothing point, the network structure is mainly divided into two link, the first link is through the convolutional neural network for image feature extraction, the second link of clothing categories by means of word embedded feature extraction.

Image feature extraction: Given the input photo, Xception by convolution convolution neural network for feature extraction, get the F features of the image, and then to get the characteristics of the figure to the global average pooling, the final output is obtained by the connection operation again, because it is based on the regression method of clothing point positioning, therefore return number for numerical prediction of key point number 2 times.

Clothing category feature extraction: Given the clothing category corresponding to an input image, the clothing category is encoded by the method of unique hot coding, and then the feature of the clothing category is obtained by the dot product with the word embedding matrix, where the word embedding matrix is the network parameter that needs to be trained. Finally, the softmax function is used to normalize the learned features.

Feature fusion: The method of feature fusion is product operation. The final network output is obtained by one-to-one multiplication between the features extracted from the picture and the features extracted from the clothing category, which is the corresponding coordinate point position of the clothing key point.

3. Clothing Image Classification Experiment

3.1. Data Set Selection

The DeepFashion dataset is a large-scale clothing dataset proposed by scholars at the Chinese University of Hong Kong. It consists of more than 800,000 images, which are buyer show and seller show photos in different scenes, different shooting angles and different types of clothing. The images are sourced from domestic and foreign e-commerce sites and direct Google searches. The DeepFashion dataset is made up of four subsets, They are Category and Attribute Prediction Benchmark, In-shop Clothes Retrieval Benchmark, Consumer-shop Clothes Retrieval Benchmark and Fashion Landmark Detection Benchmark.

3.2. Evaluation Criteria

For category classification, this paper adopts the standard TOP-K classification accuracy as the evaluation criterion. Simply speaking, topk refers to the most likely K results in the prediction results, which is represented by data as the topk categories in the output probability value. If these

K categories contain true labels, it is considered that the prediction is correct; if they do not, it is considered as a prediction error. The top-K classification accuracy is to calculate the proportion of the number of TOP-K predicted correct images in all test images. Obviously, the larger K is, the more relaxed the requirements for the prediction results will be, and the final top-K accuracy will be higher. The experiment in this paper tested the classification accuracy under top-3 and TOP-5 respectively.

For attribute prediction, TOP-K recall was used as the evaluation criterion, and the number of matched attributes in TOP-K was counted by ranking the scores of 1000 attributes. For example, when k is 5, if only one of the first five attributes is truly possessed, and the garment has a total of two attributes, then the recall rate is 50%. The final result of the experiment is to calculate the average of the recall rates of multiple samples. Similarly, the recall rates of Top-3 and TOP-5 are counted.

4. Analysis of Experimental Results

For category classification and attribute prediction, this paper compares with the methods of other two scholars.

Table 1. Experimental results of clothing category classification

	FashionNet	Corbiere	Xception
Top-3	83.57	87.04	91.56
Top-5	91.31	93.17	96.04

As shown in Table 1, is the accuracy of different network models in clothing classification.

Table 2. Clothing texture prediction contrast

	FashionNet	Corbiere	Xception
Top-3	38.67	54.23	58.59
Top-5	50.36	65.18	67.46

As shown in Table 2, are the prediction results of different network models in clothing texture attributes.

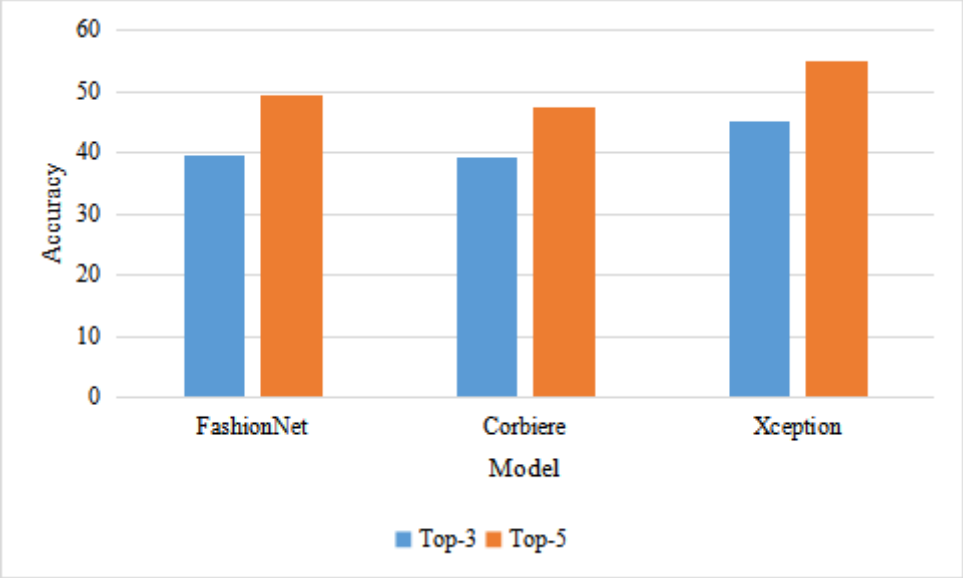

Figure 1. Clothing fabric prediction and comparison results

As shown in Figure 1, are the prediction results of different network models in clothing fabric properties.

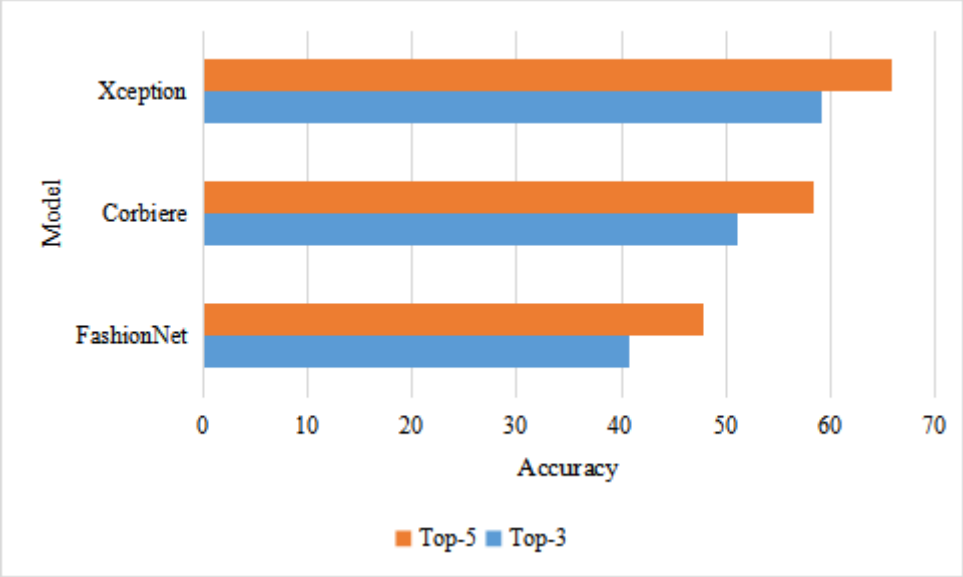

Figure 2. Clothing shape prediction and comparison results

As shown in Figure 2, are the prediction results of different network models in clothing shape attributes.

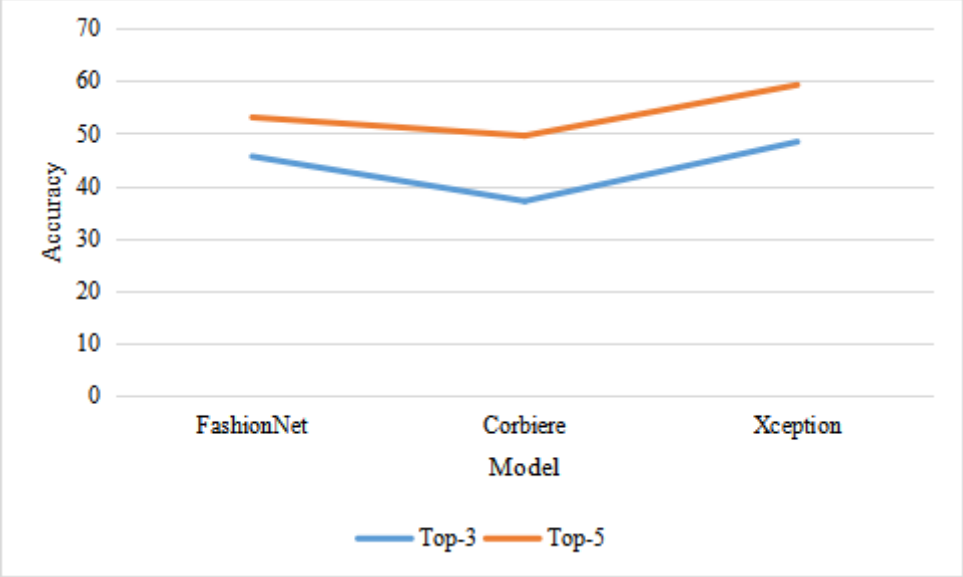

Figure 3. Clothing part of the prediction and comparison results

As shown in Figure 2, are the prediction results of different network models in clothing part attributes.

From the above data, it can be seen that the accuracy of the network model proposed in this paper in clothing classification and attribute prediction has been significantly improved. In terms of category prediction, the accuracy of TOP-3 and TOP-5 has exceeded 90%, and the accuracy of TOP-5 has exceeded 96%. For attribute prediction, clothing attributes are too detailed, especially the prediction of clothing fabrics, which is very difficult. As can be seen from the table, the accuracy of attribute prediction of the four categories has been significantly improved, among

which the prediction of texture and shape has a better performance, with the top-3 accuracy of texture attributes reaching 58.59%. The top-5 accuracy reached 67.46%, the top-3 accuracy of shape features reached 59.23%, and the top-5 accuracy reached 65.93%, which also conforms to the characteristics of feature extraction by convolutional neural network. In terms of convergence speed, the network converges after the fourth round.

5. Conclusion

With the continuous progress of social culture and the gradual rise of online shopping, buying clothes online has become people's life habits. The common way of clothing search is usually to input one or more keywords to describe clothes. Therefore, accurate classification and attribute annotation of clothes are very important. This paper focuses on three tasks: clothing key point detection, clothing category classification and clothing attribute prediction. This paper mainly explores the method of using convolutional neural network for clothing classification and attribute prediction. There are still many shortcomings and the following aspects need to be further improved: Due to the lack of clothing image data sets at present, in terms of the selection of data sets, this paper only conducts research experiments on the most widely used DeepFashion data set, which should be extended to other data sets in the future.

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] Stepchenko A M. Land-Use Classification Using Convolutional Neural Networks. Automatic Control and Computer Sciences, 2021, 55(4):358-367. https://doi.org/10.3103/S0146411621040088

[2] Anand R, Shanthi T, Dinesh C, et al. AI based Birds Sound Classification Using Convolutional Neural Networks. IOP Conference Series Earth and Environmental Science, 2021, 785(1):012015. https://doi.org/10.1088/1755-1315/785/1/012015

[3] S Zorzi, E Maset, A Fusiello. Full-Waveform Airborne LiDAR Data Classification Using Convolutional Neural Networks. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(10):8255-8261. https://doi.org/10.1109/TGRS.2019.2919472

[4] Al A. Traffic Sign Classification Using Convolutional Neural Networks and Computer Vision . Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(3):4244-4250. https://doi.org/10.17762/turcomat.v12i3.1715

[5] Anatya S, Mawardi V C, Hendryli J. Fruit Maturity Classification Using Convolutional Neural Networks Method. IOP Conference Series: Materials Science and Engineering, 2020, 1007(1):012149 (7pp). https://doi.org/10.1088/1757-899X/1007/1/012149

[6] Park G S, An H S, Lee S M. Environmental noise classification using convolutional neural networks for hearing aids. Journal of Rehabilitation Welfare Engineering & Assistive Technology, 2019, 13(4):297-303. https://doi.org/10.21288/resko.2019.13.4.297

[7] Shri S J, Jothilakshmi S. Crowd Video Event Classification using Convolutional Neural Network. Computer Communications, 2019, 147(Nov.):35-39. https://doi.org/10.1016/j.comcom.2019.07.027

[8] Ahishali M, Kiranyaz S, Ince T, et al. Dual and Single Polarized SAR Image Classification Using Compact Convolutional Neural Networks. Remote Sensing, 2019, 11(11):1340. https://doi.org/10.3390/rs11111340

[9] Ayta U C, A Güne, Ajlouni N. A novel adaptive momentum method for medical image classification using convolutional neural network. BMC Medical Imaging, 2022, 22(1):1-12. https://doi.org/10.1186/s12880-022-00755-z

[10] Sharma R. ECG Classification using Deep Convolutional Neural Networks and Data Analysis . International Journal of Advanced Trends in Computer Science and Engineering, 2020, 9(4):5788-5795. https://doi.org/10.30534/ijatcse/2020/236942020

[11] Haut J M, Paoletti M E, Plaza J, et al. Active Learning With Convolutional Neural Networks for Hyperspectral Image Classification Using a New Bayesian Approach . IEEE Transactions on Geoscience and Remote Sensing. 2018, 56(11):6440-6461. https://doi.org/10.1109/TGRS.2018.2838665

[12] Nawaz Z, Zhao C, Nawaz F, et al. Role of Artificial Neural Networks Techniques in Development of Market Intelligence: A Study of Sentiment Analysis of eWOM of a Women's Clothing Company . Journal of Theoretical and Applied Electronic Commerce Research, 2021, 16(5):1862-1876. https://doi.org/10.3390/jtaer16050104

[13] Alotaibi G, Awawdeh M, Farook F F, et al. Artificial intelligence (AI) diagnostic tools: utilizing a convolutional neural network (CNN) to assess periodontal bone level radiographically—a retrospective study . BMC Oral Health, 2022, 22(1):1-7. https://doi.org/10.1186/s12903-022-02436-3

[14] Kruthiventi S, Ayush K, Babu R V. DeepFix: A Fully Convolutional Neural Network for Predicting Human Eye Fixations. IEEE Transactions on Image Processing, 2017, 26(9):4446-4456. https://doi.org/10.1109/TIP.2017.2710620

[15] Acharya U R, Fujita H, Lih O S, et al. Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network. Knowledge-Based Systems, 2017, 132(sep.15):62-71. https://doi.org/10.1016/j.knosys.2017.06.003

[16] Paoletti M E, Haut J M, Plaza J, et al. A new deep convolutional neural network for fast hyperspectral image classification. Isprs Journal of Photogrammetry & Remote Sensing, 2017, 145PA (NOV.):120-147. https://doi.org/10.1016/j.isprsjprs.2017.11.021

[17] Evo I, Avramovi A. Convolutional Neural Network Based Automatic Object Detection on Aerial Images. IEEE Geoscience & Remote Sensing Letters, 2017, 13(5):740-744. https://doi.org/10.1109/LGRS.2016.2542358

[18] Haenssle H A, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Annals of Oncology, 2018, 29(8):1836-1842. https://doi.org/10.1093/annonc/mdy520