JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Product Categorization in Fashion and Lifestyle Commerce using Machine Learning

[1]Munmi Dutta, [2]Nilotpal Hazarika, [3]Amrita Ganguly

[1]Research Scholar, [2]Senior Software Engineer, [3]Professor
[1]Department of Electrical Engineering,
[1]Assam Engineering College, Guwahati, India

Abstract : In this fashion world the latest trends and products keep on changing. As such automatically categorizing and searching through a huge data of images according to the patterns is a tough work. A machine learning (ML) technique is proposed in this paper for automatically identifying the patterns from the images. This approach is for the e-Commerce portal to automatically organize the products in the proper category of the catalog without human input. It is also beneficial for customers to search a specific product when text description for the target product is not present.

IndexTerms - machine learning (ML), convolutional neural network (CNN), deep neural networks (DNN).

I. INTRODUCTION

Fashion industry deals with a large amount of data as fashion trends keep on changing as such manual sorting of the products in different categories is tedious and time consuming. Normalizing and cleaning of products has been possible in the consumer electronics space with the use of stock keeping units (SKUs). Poorly organized catalogs and the poor description of products and quantifying product details, the problem of organizing products as SKUs in categories such as fashion (for example fashion apparel and accessories) and lifestyle (such as interior design items) has been largely unsolved. Site management becomes cumbersome and critical if one has to build an aggregate fashion commerce application that has to combine several such poorly tagged catalogs. SKUs for fashion and lifestyle preparation are possible with the use of deep learning (DL). With the advancement in deep convolutional neural networks (DCNN) have led to implementation evolution in image classification [1] [2]. Deep networks have the ability to intricate all levels of features from low to high and show the classification in the output layer in high accuracy depending on the number of stacked layers used [1] [3] [4] [5]. In this paper convolutional neural network (CNN) on product image is applied to automatically organize the products in the proper category of the catalog.

Product categories are the structural foundation of every e-Commerce site. In e-Commerce, we have a variety of products and hence assignment of the right category to all of these products is tedious work. Using machine learning (ML) at this point reduces the labor and is a good way to improve the product category search. Category structures that are presently being used in e-Commerce have to deal with a lot of categories under which a search might fall. The category manager who is responsible to assign the right product always finds it difficult and time consuming to do the same. The objective of the product search algorithm using ML is that the database automatically assigns the categories and also filters the categories since there are thousands of categories for different products and to show the most relevant ones with minimum search time.

One of the major challenges which every e-Commerce site faces is that every online store has a different category structure. Product categories too range from groceries to fashion, furniture etc. There could be different industries for varying different product categories too. Stores for the same industry may have a drastically varying different category system. All these challenges have to be addressed to make a robust system of product categorization. One way to deal with this would be to train a ML model for every single customer for every single unique category structure present in the database. It will be a difficult task for operations to maintain these models but it will allow for the entire system to have highly specific models that are specifically trained for these categories. Second option is to use one ML model that recommends from a broad set of product categories in the first stage and then in the next stage only recommend categories that are highly specific to specific stores [6] [7] [8]. For this a separate matching algorithm is required at the end. The main ML that is proposed here recommends the general category. This model is more flexible, easy to deploy, works for stores with relatively smaller databases and is more effective.

The structure of the rest of the paper is as follows. In section II a brief review of some of the methods adopted by others for image classification is given. Section III the image dataset used and the steps used for the preparation of the dataset is presented. The proposed method algorithm using CNN for the object classification for category classification in e-Commerce sites is also presented in this section. Section IV introduces the experiments and the result obtained. Finally, the conclusions of this paper are described.

II. Related Work

A CNN is an Artificial Neural Network that has been so far been popularly used for analyzing images. Although image analysis has been a wide use of CNNs they can be also used for other data analysis or classification problems as well. Most generally we take CNN as an ANN that has some type of specialization for being able to pick out or detect patterns and make sense of that. Training the neural network to extract the right features from the input images is the main work. This pattern detection is what makes CNN so useful for image analysis. CNN is such a form of ANN, what differentiate it from just a standard perceptron or multilayer perceptron (MLP), well a CNN has hidden layers called convolutional layers and these layers are precisely what makes the CNN. CNN also has other non-convolutional layers as well. Convolutional layers are able to detect patterns in images. CNN consists of an input layer, hidden layers and an output layer. Hidden layer consists of multiple convolutional layers, pooling layers, fully connected layers and normalization layers (as shown in Fig. 1). Training the neural network to extract the right features from the input images is the main work.

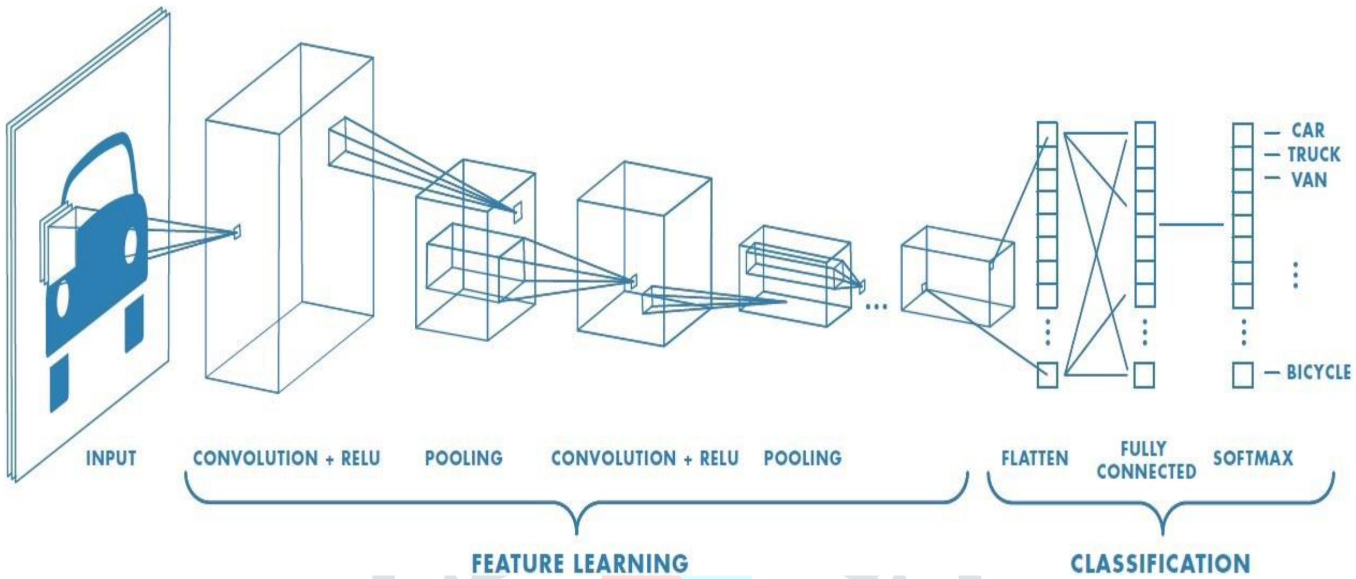

Fig. 1. Representation of layers within a CNN [9].

The proposed method for product categorization is related to CNN. Lot of research work till date has been conducted in the field of image detection and classification. Some of the previous work related to image classification using CNN are as follows:

In [10] they have presented MSURU, it is an image recognition system for commerce search engines (mainly for Facebook Marketplace). The model architecture used by them is ResNeXt-101 32×4d. It includes 101 layers, 32 groups and with group widths C=4 [11]. MSURU was predicted to perform 16% better than the state of the art system used at Facebook [12].

In [13] the authors have explained the different CNN architectures like LeNet-5(1998), AlexNet-2012, ZFNet, VGGNet, GoogLeNet, ResNet, DenseNet, CapsNet, SENet. The model description of each of the CNN architecture was given along with the training details. They also compared each of the architecture performances.

Final Remarks: Many papers only focus on the image classification/identification in general point of view. That is classification related work is done only on boardly different classes, but very less work has been done for apparel or category classification for an e-Commerce site. Keeping this point in view, we have employed the proposed CNN model to create an application for category classification in e-Commerce sites. In addition, we have included many of the Indian attire in the dataset.

III. Model

The proposed model acts as an aggregate over the existing direct e-Commerce as well as online marketplace. To provide a good consumer experience and to provide discovery and search experience to consumers, it is essential to have good robust technology.

A. Challenges

The challenges for building a robust solution for identification of category for images may include differentiation of single item versus a pack of item from images as shown in Fig. 2 and Fig. 3, images in different poses as shown in Fig. 4, confusing categories in images as shown in Fig. 5, a sample image of watch taken from training dataset is shown in Fig. 6, a sample image of handbag taken from training dataset is shown in Fig. 7, a sample image of blouse taken from training dataset is shown in Fig. 8. and a sample image of sherwani taken from the training dataset is shown in Fig. 9.

Fig. 2. Image of pant (single item).

Fig. 3. Image of pants (pack item).

Fig. 4. Image of pant (in backside pose).

Fig. 5. Image of party wear salwar (looks like a gown).

Fig. 6. Watch image from training dataset.

Fig. 7. Handbag image from training dataset.

Fig. 8. Blouse image from training dataset.

Fig. 9. Sherwani image from training dataset.

B. Importance

Wrong search results shown to customers on the basis of search engines is an important factor to be resolved. Proper categorization would have not led to such a disaster in the display of search products. As shown in Fig. 10 though in the search box “red shirt” is typed a blazer is being displayed along with other red shirts. This is because at present search works on the search keywords of the product the customer typed to the matching keywords of the search index.

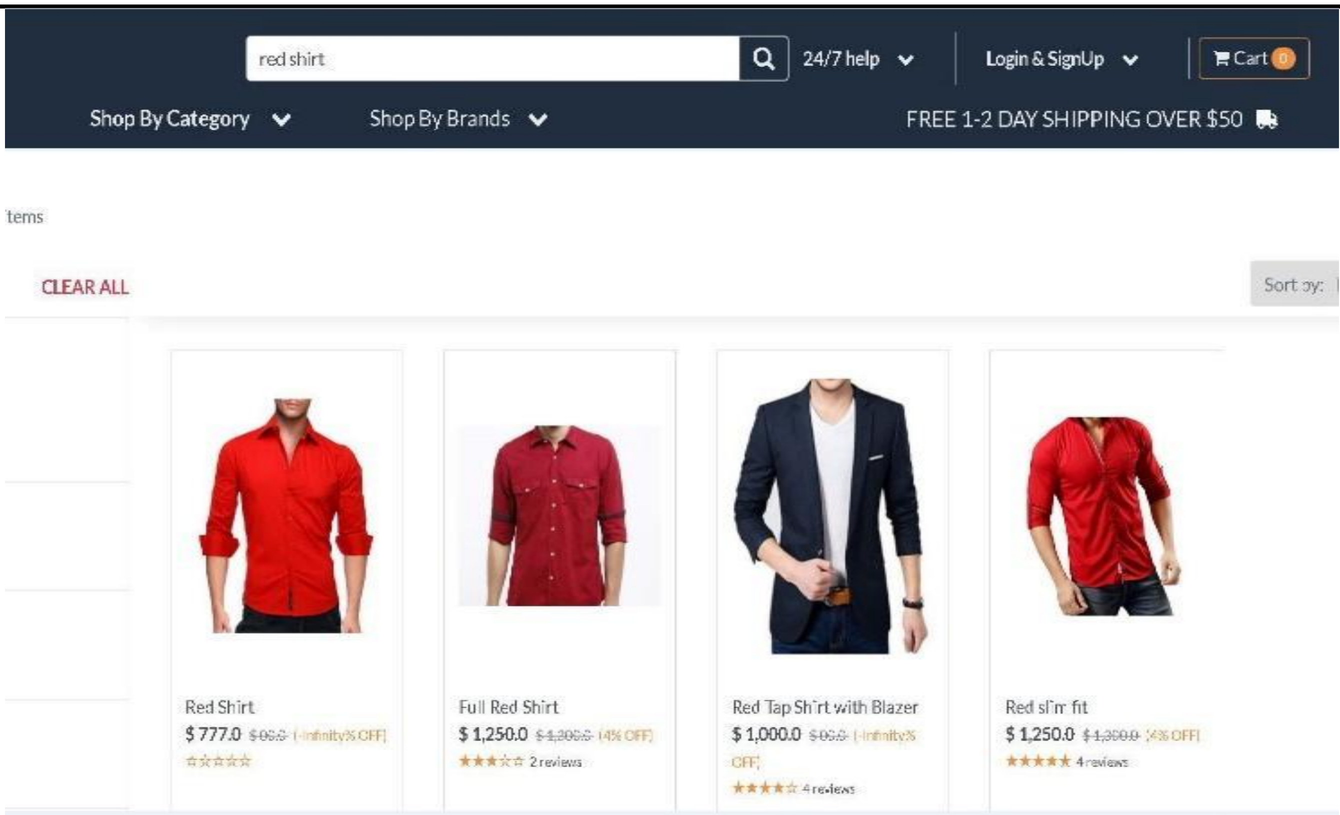

Fig. 10. Display of wrong products along with the correct one.

C. Input Representation

The dataset consists of 45000 images taken using mobile and professional cameras. The background used for the images is not fixed. The images were captured using iPhone XS, iPhone 12, OnePlus A6000, Sony A7III, Canon 5D Mark3 and Canon 5D Mark 4.

Fifteen classes were considered for classification in the fashion commerce including some of the Indian dress types. The classes are blouse, lehenga, formal pant, dhoti, hand bag, heel, dress, saree, salwar, sherwani, shirt, shoe, shorts (man), tops and watch. For each class 3000 images were taken. A total of 45000 images were collected for all the classes. Data augmentation is done in some of the classes due to shortage of data collected. The images were mostly in white background but some are in different backgrounds and in different poses. The data set was divided in the ratio of 8:2 for training set and test set for the creation of the ML model.

D. Setup

The training was done on Ram 16GB, Core i7 5^{th} Generation, HDD 2TB, GPU RAM 4GB. Creating the dataset was tedious work to get the right format data. We have not used any dataset available online like FMNIST and Kaggle as our dataset contains Indian attire too. Training and testing the model was a hard work and time taking process.

E. Model Architecture

In this paper the normalization of the catalog is given more preference. The steps adopted to create the ML model for the class identification are shown in Fig. 11.

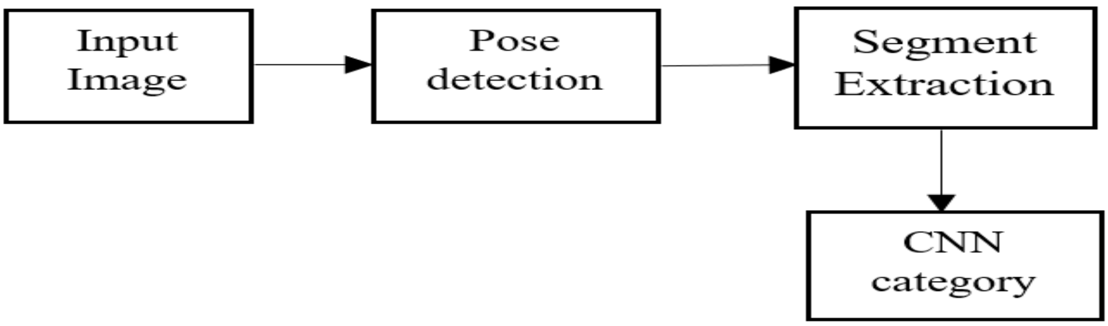

Fig. 11. Block diagram for class classification.

For image classification the most popular among the ML techniques is DNN. In the case of image classification CNN is the best for its ability to recognise the features like lines, edges, shape and colours and to train the ML from it [8].

The image size 224x224 is taken for all the images in the dataset. The ML model proposed in this paper consists of 8 layers including 5 convolutional layers and 3 fully connected. With the first convolutional layer kernel size of 5x5 is taken. In all the second, third, fourth and fifth convolutional layers kernel size of 3x3 is taken. At the end softmax function is added to get the output class. For the improvement of our ML model performance, collection of larger datasets can be done, more powerful model's techniques can be applied and can use better methods for preventing overfitting. To increase the performance of the ML model fine-tuning and re-training of the ML model is done.

IV. RESULTS

After a lot of trial and error of changing batch size and number of epochs and with the help of early stopping by using this architecture an accuracy of 93% and average confidence of 86% was achieved. The confusion matrix obtained in the creation of the ML model is shown in Table 1. Adding more data with a deeper model will help in getting a higher prediction confidence score. Adding extra layer results in attaining high accuracy but pays an adverse effect on the training time by elongating it. Removal of a single convolutional layer has an adverse impact on the overall performance of the network. Our work shows that a good result could be achieved on a challenging dataset with the application of DCNN.

Table 1. Confusion matrix

Classes	**Confusion matrix**														
	blouse	***lehenga***	***Formal pant***	***dhoti***	***dress***	***handbag***	***heel***	***saree***	***salwar***	***sherwani***	***shirt***	***shoe***	***shorts***	***tops***	***watch***
blouse	548	30									10			12	
lehenga	25	558									2			15	
Formal pant			565	15									20		
dhoti			12	553					5	30					
dress					562				12					26	
handbag	2					568							18	12	
heel							585						15		
saree					28			566	6						
salwar					34				551					15	
sherwani			14	18						562			6		
shirt	16									17	545			22	
shoe							19					76	5		
shorts				19									555	26	
tops	28	21						9						542	
watch							36								564

V. CONCLUSION

Manual labor for categorization will be reduced to an extent by using this ML model. This model will replace the manual work for maintaining a precise and accurate category structure. Thus, it will facilitate the user navigator to have a peaceful experience of shopping in an online store.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, vol. 25, 2012.

[2] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," Neural computation, vol. 1, no. 4, pp. 541-551, Dec. 1989.

[3] M. D. Zeiler and R. Fergus. "Visualizing and understanding convolutional neural networks," in European conference on computer vision, Sep. 2014, pp. 818-833.

[4] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," arXiv preprint arXiv:1312.6229, 2013.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.

[6] B. Moghaddam and M. H. Yang, "Learning Gender with Support Faces," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 707-711, May 2002.

[7] S. Zoghbi, G. Heyman, J. C. Gomez, and M. F. Moens, "Fashion Meets Computer Vision and NLP at e-Commerce Search," International Journal of Computer and Electrical Engineering, vol. 8, no. 1, pp. 31-43, Jan. 2016.

[8] A. F. M. Agarap, "An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification," arXiv preprint arXiv:1712.03541, 2017.

[9] Mathworks.com: Convolutional Neural Network (2018). https://www.mathworks.com/content/mathworks/www/en/discovery/convolutional-neural-network.html. Accessed 20, Apr. 2018.

[10] Y. Tang, F. Borisyuk, S. Malreddy, Y. Li, Y. Liu, and S. Kirshner, "MSURU: Large Scale E-commerce Image Classification with Weakly Supervised Search Data," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery \& Data Mining, Jul. 2019, pp. 2518-2526.

[11] S. Xie, R. B. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated Residual Transformations for Deep Neural Networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1492-1500.

[12] D. Mahajan, R. B. Girshick, V. Ramanathan, K. He, M. Paluri, Y Li, A. Bharambe, and L. van der Maaten, “Exploring the Limits of Weakly Supervised Pretraining,” in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 181-196.

[13] F. Sultana, A. Sufian, and P. Dutta, “Advancements in Image Classification using Convolutional Neural Network,” in 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Nov. 2018, pp. 122-129.