Product Image Classification Techniques

Dhaval Wagh, Smita Mahajan

Abstract: ***The excessive amount of e-commerce products within the last few years has become a serious problem for shoppers when searching for relevant product information. This led to the emergence of a recommendation technology that has the capability to discover relevant shopping products that meet the user's preferences. Classification is a machine learning technique that could assist in increasing scalability, creating dynamic user profiles and ultimately improve recommendation accuracy. Many researches have been done in the area of image classification with e-commerce product images. This paper surveys the e-commerce product image classification techniques. This paper examines current practices, problems, and prospects of product image classification. The emphasis is placed on the account of major advanced classification approaches and also the techniques used for improving classification accuracy. This survey involves reviewing the research work done by different professionals and assembling it into one paper.***

Keywords: Classifiers, E-commerce, Image Classification, Machine Learning, Recommendation

I. INTRODUCTION

E-commerce is playing a key role in the global economic growth and cannot be over emphasized the need to keep satisfying customers. In modern times, the e-commerce adoption has led to higher profitability for merchants and brought more satisfaction to consumers [1, 2]. This has also impacted the economy of countries worldwide positively by improving Gross Domestic Product (GDP). It was reported in South Africa that user online retail shopping sales surpassed a trillion Rand (South Africa currency), which has increased to 1.46 trillion Rand in 2016 for the first time. According to the Economist Intelligence Unit (EIU), the real GDP growth of 4.9% for the African continent from 2012 to 2016, which is well above the average global growth [3]. In addition, the collective gross domestic product of Africa continent is anticipated to rise by $1trillion by the year 2020, that is up from US $1.6 trillion in 2010, thanks to the worldwide surge in e-commerce rumored by the PricewaterhouseCoopers (PwC) in South Africa [3,4]. Because of the inherent difficulty in information discovery, the very large quantity of e-commerce information has become a serious challenge for shoppers in recent years. This has led to the emergence of recommendation systems to assist a user in the information discovery. An application that uses image features to filter information from the complete obtainable supply, displays the suitable info supported personal preferences and also kept in the user profile is the content-based image recommendation system [5].

In the e-commerce domain, several ways and techniques are applied to implement recommendation systems. Out of that classification ways are known as a crucial element. Classification ways are found helpful in several recommendation applications like product image retrieval [6], increase scalability [7], product taxonomy browsing [8] and improve overall recommendation accuracy. Product classification is one of the most important processing tasks of content-based recommendation systems. It involves the association of classes with related products from a large number of merchants.

The traditional product data representation method is text tagging. Most of the product classification accuracies rely heavily on text tagging [9–11]. However, product classification supported text tagging has many issues because of that the present analysis efforts have shifted focus to image primarily based product classification models with numerous applications in numerous fields of life. In contrast to text tagging, image based product classification involves the use of images for product representation and classification model. Due to the high dimensionality of extracting image features, limited content analysis, inhomogeneity and other factors often inhibit image-based classification performance [9, 12, 13]. A significant volume of researches have been channeled towards this direction, nevertheless, the number of image-classes investigated and accuracies rumored still leave abundant to be desired for time period applications.

In this paper, we'll shortly discuss the various image classification techniques supported with e-commerce product images and a few of the recent works on these techniques.

II. LITERATURE SURVEY

Classification is a machine learning technique which is used to classify large amount of data based on the various characteristics. An algorithm that implements classification is known as classifier. According to the pattern recognition theory, feature extraction and recognition strategies play key roles within the method of classification. As for the performance, accuracy is one of the common assessments. Support vector machine (SVM) and Artificial neural network (ANN) are two popular classifiers that are applied to e-commerce product image classification and call supporting tasks with some extent of success. Many researchers within the recommendation domain have allotted interesting studies to reinforce the performance of product classification.

Revised Manuscript Received on December 22, 2018.

Dhaval Wagh, Computer Science and Engineering, Symbiosis Institute of Technology, Pune, India.

Smita Mahajan, Computer Science and Engineering, Symbiosis Institute of Technology, Pune, India.

Brian Tomasik [14] has studied the extent to which a standard "bag of visual words" image classifier can be used to tag products with useful information. They explores the feasibility of describing consumer products through supervised image classification. Using Scale Invariant Feature Transform (SIFT) image descriptors at random key points, a hierarchical visual vocabulary, and a variant of nearest neighbor classification, they achieved accuracies between 66% and 98% on 2- and 3-class classification showing that at least for some tasks, tagging products based solely on their images can be done in an automated fashion.

Jia Shijie [15] have we proposed a fast supervised image classifier which is based on class-specific Pyramid Histogram of Words (PHOW) descriptor and Image-to-Class distance (PHOW/I2C). Firstly in the training phase, the local features are represented as soft-voting PHOW descriptors and densely sampled, and then with the help of means and variances of distribution of each visual word in each labeled class, the class-specific descriptors are built. Then for online testing, the normalized chi-square distance is calculated between the descriptor of query image and each class specific descriptor. The class label corresponding to the least I2C distance is taken as the winner. Experiments demonstrate the effectiveness and quickness of our method in the tasks of product classification and achieved 84% for 30 product classes.

For the task of product classification over large number categories, Shi-jie Jia [16] employed kernel-based SVM classifier combining multiple features, including one global descriptor and three complimentary local descriptors. , They employ one global descriptor (GIST) with three complimentary local descriptors (shape descriptor PHOG, appearance descriptor PHOW, and texture descriptor PLBP) as the image representation set. Furthermore, they investigate the ways to combine discriminative features with diverse type of kernels for Support Vector Machine (SVM) classifier. Experiments on the merchandise image dataset (PI 100) showed the performance improved considerably by features fusion.

For the task of visual-based automatic product image classification for e-commerce, JIA Shi-jie [17] constructs a set of support vector machine (SVM) classifiers with different model representations. Each base SVM classifier is trained with either various types of features or completely different special levels. The likelihood outputs of those SVM classifiers are concatenated into feature vectors for training another SVM classifier with a Gaussian radial basis function (RBF) kernel. Ten SVM classifiers are derived; each is trained on one spatial level of PGIST (pyramid of Gist), PHOW (pyramid histogram of words), PHOG (pyramid histogram of gradient) and PLBP (pyramid of local binary pattern), respectively. Then the outputs of these SVM classifiers with chi-square kernel are combined for the final classification result. This scheme achieves state-of-the-art average accuracy of 86.9% for product image classification on the public product dataset PI 100, which indicates a good direction for implementing automatic product image classification in practice.

Anitha Kannan [18] have initiated a study into the relatively unexplored classification setting involving text and image signals, where the image signals are less discriminative compared to text-based signals, and explored however image signals may be wont to complement text classifiers. They proposed a novel algorithm Confusion Driven Probabilistic Fusion++ (CDPF++) which learns a number of three-way image classifiers focused only on those confusing categories of the text signal so as to capture the area of the discriminating surface that the dominant text classifier is unable to capture. Through a variety of experiments on datasets from a major Commerce search engine's (Bing Shopping) catalog, they observed a 12% (absolute) improvement in CDPF++'s precision at 100% coverage compared to classifiers that solely use textual description of products; and a 16% (absolute) improvement in recall at 90% precision over the same baseline.

Shi-jie Jia [19] have presented two layers of SVM classifier to combine each spatial level of pyramid histogram of words (PHOW) and pyramid histogram of orientated gradients (PHOG) descriptors. In the 1st layer, every of six chi-square kernel SVM classifiers is trained with a special level of pyramid bar chart of words (PHOW) and pyramid bar chart of directed gradients (PHOG) descriptors, and so the chance outputs of those SVM classifiers square measure concatenated into feature vectors for coaching another SVM classifier with a Gaussian RBF kernel. Experimental results compared 2 Direct SVMs and 2 layers-SVMs stacking on ninety four categories to report an overall accuracy of eighty one.2% and 79.1% severally. This means that it's necessary to use a mixture of many base classifiers to create effective use of the complementary options, that provides how forward for the improvement of product classification.

He Zhang [20] have adopted SVM classifier combined with PHOG (Pyramid of Histograms of Orientation Gradients) descriptors to implement product image classification. The support vector machine maps the input vectors into a high-dimension house, during which a grievous bodily harm margin supper hyper plane is ready up to classify the samples and PHOG will flexibly represent the special layout of native image type. The experimental result indicates the common accuracies increase with the coaching samples and showed the effectiveness of the algorithmic program.

S.A. Oyewole [21] have proposed color image classification framework that integrates linear and radial basis function (LaRBF) kernels for SVM. The first paramount objective of this study is to extract color features of 100 image classes obtained from the PI100 database using the histogram of oriented gradient (HOG) algorithm in five different color models, which are RGB, oRGB, XYZ, HSV and YUV. The second important objective of this study is to develop a customized kernel based on the linear kernel and the RBF kernel to improve the classification of e-commerce product images using SVM. They customized LaRBF kernel based on linear and RBF kernels.

The HOG features extracted in completely five different color models were wont to train SVM that utilizes four existing kernels and therefore the projected LaRBF kernel. Experimental results show that the proposed LaRBF kernel based SVM gave the best average accuracy in the RGB color model that is 83.5%. This results extremely promising for coming up with sensible applications of product image classification, recommendation and alternative e-commerce systems.

ANNs have the power to model nonlinear relationships between a collection of input variables within the user profile and corresponding preferences of the user. Large number of product data often caused over fitting when an ANN is integrated with any e-commerce applications such as recommendation systems. To overcome this over fitting problem, Bashiri M [22] introduced cross-validation which is considered to be one of the most effective methods to ensure over fitting does not occur. Here, on the market information is typically partitioned off into 3 sets that are training, testing and validation. The coaching set is employed to regulate the association weights, the testing set is employed to envision the performance of the network at numerous stages of learning and training is stopped once the error at intervals the testing set can increase. The validation set is employed to judge the performance of the model once training has been with success accomplished.

In Bonnett C.[23], an object recognition pipeline for e-commerce products was developed employing a combination of convolutional neural nets (Deep Learning) and natural language processing (NLP). In his work 85%, 86% and 86% classification accuracy was realized with image, text and a hybrid of both respectively on 10 classes. The researcher finally achieved 94% classification accuracy by using transfer learning and fine-tuning with neural networks. However, the number of 10 classes considered is still very limited.

Deep convolutional neural networks (CNNs) have shown an excellent success in single-label image classification, it's vital to notice that real world pictures typically contain multiple labels, that may correspond to totally different objects, scenes, actions and attributes in a pictures. Traditional approaches to multi-label image classification learn independent classifiers for each category and employ ranking or thresholding on the classification results. These techniques, although working well, fail to explicitly exploit the label dependencies in an image. Jiang Wang [24] has utilized recurrent neural networks (RNNs) to address this problem. Combined with CNNs, the projected CNN-RNN framework learns a joint image-label embedding to characterize the linguistics label dependency in addition with the image-label relevance, and it are often trained end-to-end from scratch to integrate each info during a unified framework. The proposed framework combines the advantages of the joint image/label embedding and label co-occurrence models by employing CNN and RNN to model the label co-occurrence dependency in a joint image/label embedding space. Experimental results on public benchmark datasets demonstrate that the projected design achieves higher performance than the progressive multi-label classification models.

Jung-Woo Ha [25] demonstrate a successful report on a deep learning-based item categorization method, i.e., deep categorization network (DeepCN), in an e-commerce website DeepCN is an end-to-end model using multiple recurrent neural networks (RNNs) dedicated to data attributes for generating features from text data and absolutely connected layers for classifying item classes from the generated features. They evaluate DeepCN on large-scale real-world data including more than 94 million items with approximately 4,100 leaf categories from a Korean e-commerce website. Experiment results show our methodology improves the categorization accuracy compared to the model which uses single RNN in addition as a regular classification model uses unigram-based bag-of-words.

Petar Ristoski [26] presented an approach that leverages neural language models and deep learning techniques in combination with standard classification approaches for product matching and categorization. During this approach they use structured product information as management for coaching feature extraction models ready to extract attribute-value pairs from matter product descriptions. To minimize the need for lots of data for supervision, they use neural language models to produce word embedding from large quantities of publicly available product data, which boost the performance of the feature extraction model, so resulting in higher product matching and categorization performances. Furthermore, they use a deep Convolutional Neural Network to produce image embedding from product images, which further improve the results on both tasks. Results shows that the text embedding improves product categorization significantly and image embedding is used as a weak signal for product matching and robust signal for product categorization.

Alexander Schindler [27] presented an empirical study of applying deep Convolutional Neural Networks (CNN) to the task of fashion and attire image classification to boost meta-data enrichment of e-commerce applications. This study focuses on the task of fashion image classification. This study focuses on the task of fashion image classification. Five completely different CNN architectures were analyzed using clean and pre-trained models. The models were evaluated in three completely different tasks person detection, product and gender classification, on two small and huge scale datasets. These tasks is evaluated on two completely different datasets on additional 2 different scales. First, massive analysis is performed on a smaller scale dataset and therefore the best playacting models are then applied to large scale datasets. The results are comparable to the product classification task in the sense that pre-trained and fine-tuned models provide the highest accuracies with a best performing value of 88%.

Miguel G. Vieira [28] united the hierarchical classification method through deep learning applied to ecommerce images to attain a significant improvement in the classification accuracy and compare the results to both

of classifications alone. Here the classification approach is presented considering two datasets (studio and general purpose e-commerce images) individually. In both datasets, the images are divided into several categories, each one labeled following classes commonly found in e-stores. The first dataset covers images from traditional e-commerce websites, with clean, studio-like images, while the second is formed by images take in a domestic-like fashion, with no special treatment. Deep learning classifiers were trained to classify the images for each dataset separately, using a hierarchical approach to improve the classification rates.

Hessel Tuinhof [29] developed a two-stage deep learning framework that recommends fashion images based on other input images of similar style. For that purpose, a neural network classifier is employed as a data-driven, visually-aware feature extractor. This approach is tested on the publicly available Fashion dataset. In this work, they followed a different, data-driven approach, where customer preferences are automatically extracted from available information on the customer. More specifically, their focus on fashion products and develop a method that only requires a single input image to return a ranked list of similar style recommendations. The projected two-stage approach uses a CNN classifier to extract options that area unit used as input for similarity recommendations. Combined with additional ancient content-based recommendation systems, our framework will facilitate to extend robustness and performance.

S.A. Oyewole [30] have proposed an enhanced product image classification architecture and Eigen-based product image feature extraction formula that delivers a good product image illustration for big product categories. They proposed an enhanced product image classification architecture which has data acquisition pre-processing, feature extraction, dimensionality reduction and ensemble of machine learning methods as components. Core amongst these components is the Eigenvector based fusion algorithm that is meant to obtain dimensionality reduced Eigen Color Feature from the histogram of oriented gradient based color image representative features. The ensembles of Artificial neural network and Support vector machine were trained with the Eigan Color feature to classify product pictures from the PI100 corpus into one hundred categories and their classification accuracies were compared. They obtained a progressive classification accuracy of 87.2% with the artificial neural network ensemble that is a powerful result in comparison to existing results. Also the final classification model obtained from this work can be easily integrated with any other decision supporting applications in e-commerce domain to improve its quality.

Various techniques for product image classification are depicted. The review of techniques emphasized that hierarchal framework in product image classification techniques and how they are used to improve the classification accuracies.

III. CONCLUSION

This paper surveys the various classification techniques that are part of machine learning and are used in classifying e-commerce product images. At first the various classification methods that are being currently used in product image classification were studied. This involved studying the available research. Based on that research this paper was written listing the various classification techniques in use. The summarization of major advanced classification approaches and the techniques used for improving classification accuracy are discussed.

Future analysis will focus on the event of additional discriminating feature extraction methods. Additionally, researchers will perform additional intensive comparative studies on alternative state-of-the art machine learning techniques to further enhance the performance of the product image classification techniques reported in this paper.

REFERENCES

1. Baier D, Daniel I, Frost S. Naundorf R. “Image data analysis and classification in marketing”, Advanced Data Analytics Classification 2012;6:253–76.
2. Bhattacharya S, Das R. “Facilitating consumer satisfaction by content based product classification”, ICBPEM, proceedings of international conference at national institute of technology, Rourkela, Springer; 2014.
3. PwC South Africa. South African retail and consumer products outlook: 2012– 2016, last assessed on 29 of June 2017 from <http://www.pwc.co.za/en/publications/retail- and-consumer outlook.html>.
4. McKinsey Global Institute. Lions on the move, the progress and potential of African economies; 2010.
5. Olugbara OO, Ojo SO, Mphahlele MI. “Exploiting image content in location based shopping recommender systems for mobile users”. In: J Information Technology Decision Making 2010;9:759–78.
6. Vailaya A, Figueiredo MAT, Jain AK, Zhang H-J. “Image classification for content-based indexing”. IEEE Trans Image Process 2001;10(1):117–30.
7. Jain YK, Wadekar S. “Classification-based retrieval methods to enhance information discovery on the web”. Int J Manag Inform Technol (IJMIT) 2011;3(1):33–44.
8. Pohs W. “Building taxonomy for auto-classification”. Bulletin of the American Society for Information Science and Technology, 2013;39(2):34–8.
9. Kannan A, Talukdar PP, Rasiwasia N, Ke Q. “Improving product classification using images”. In: ICDM; 2011. p. 310:319.
10. Agrawal R, Srikant R. “On integrating catalogs”. In: Proceedings of the 10th international conference on world wide web(WW-10). New York: ACM Press; 2001. p. 603–612.
11. Chakrabarti S, Sarawagi S, Godbole S. “Cross-training: learning probabilistic mappings between topics”. In: Proceedings of the Ninth ACM SIGKDD international conference on knowledge discovery and data mining Washington; 2003. p. 24–7.
12. Han Y, Choi SM. “A content recommendation system based on category correlations”. Fifth international multi-conference on computing in the global information technology; 2010. p. 66–70.
13. Mishra G, Kar J, Nath SS, Chakraborty S, Dey N. “A survey of image classification methods and techniques”. In: International conference on control, instrumentation, communication and computational technologies (ICCICCT), vol. 978, no. 1; 2014. p. 554–7.
14. Tomasik B, Thiha P, Turnbull D. “Tagging products using image classification”. In: Proceedings of the 32 international ACM SIGIR conference on research and development in information retrieval. Boston, MA, USA; 2009. p. 792–3.

15. Jia Shijie, Kong Xiangwei, Jin Guang. “Automatic fast classification of product-images with class-specific descriptor”. In: Journal of Electronics (China); 2010. Vol.27 No.6.
16. Jia S, Gu Y, Zou J. “Product image classification with multiple features combination”. In: International conference on E-business intelligence, Atlantis Press; 2010. p. 469–75.
17. Jia S, Kong X, Man H. “Automatic product image classification with multiple support vector machine classifiers”. J Shanghai Jiatong Univ (Sci) 2011;16 (4):391–4.
18. Anitha Kannan, Partha Pratim Talukdar, Nikhil Rasiwasia, Qifa Ke. “Improving Product Classification Using Images”. In: IEEE 11th International Conference on Data Mining (ICDM): 2011.p. 310-319.
19. Jia S, Kong X, Hong M. “Automatic product images classification with two layer of SVM classifier”. In: Green communications and networks. Dordrecht: Springer; 2012. p. 209–17.
20. Zhang 20, Sha Z. “Product classification based on SVM and PHOG descriptor”. International Journal of Computer Science and Network Security, 2013;13(9):1–4.
21. S.A. Oyewole, O.O. Olugbara, E. Adetiba, T. Nepal. “Classification of Product Images in Different Color Models with Customized Kernel for Support Vector Machine”. In: IEEE Third International Conference on Artificial Intelligence, Modelling and Simulation; 2015.p.153-158.
22. Bashiri M, Geranmayeh AF. “Tuning the parameters of an artificial neural network using central composite design and genetic algorithm”. Scientia Iranica 2011;18(6):1600–8.
23. Bonnett C. Classifying e-commerce products based on images and text; 2016. Last assessed on 21 of June 2017 from <http://cbonnett.github.io/Insight. html>.
24. Jiang Wang, Yi Yang, Junhua Mao, Zhiheng Huang., et al. “CNN-RNN: A Unified Framework for Multi-label Image Classification”. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016.p.2285-2294.
25. Jung-Woo Ha, Hyuna Pyo, Jeonghee Kim. “Large-Scale Item Categorization in e-Commerce Using Multiple Recurrent Neural Networks”. In: 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining 2016 (KDD16);p.107-115
26. Petar Ristoski, Petar Petrovski, Peter Mika, Heiko Paulheim. “A Machine Learning Approach for Product Matching and Categorization”. In: IOS Press; 2016.p.1-20.
27. Alexander Schindler, Thomas Lidy, Stephan Karner, Matthias Hecker. “Fashion and Apparel Classification using Convolutional Neural Networks”. In: Forum Media Technology (FMT2017), St.Poelten, Austria; 2017.
28. Miguel G. Vieira, Jander Moreira. “Classification of E-commerce-related Images Using Hierarchical Classification with Deep Neural Networks”. In: IEEE Workshop of Computer Vision; 2017.p.114-119.
29. Hessel Tuinhof, Clemens Pirker, and Markus Haltmeier. “Image Based Fashion Product Recommendation with Deep Learning”. In: arXiv preprint arXiv:1805.08694.
30. S.A. Oyewole, O.O. Olugbara. “Product image classification using Eigen Colour feature with ensemble machine learning”. In: Egyptian Informatics Journal, Elsevier; 2018.p.1-18.

AUTHORS PROFILE

Dhaval Wagh has received B.E. [Computer Science and Engineering] from Indore Institute of Science and Technology, Indore (M.P.) in 2017 & pursuing M.Tech.[Computer Science and Engineering] from Symbiosis Institute of Technology, Pune. His research interest includes Image Processing and Machine Learning.

Smita Mahajan has received her M.E (Information Technology) from Mumbai University, pursuing Ph.D. from Symbiosis International University, Pune in the area of Artificial Intelligence. She is specialized Computer Networks. She works as Asst. Professor in Symbiosis Institute of Technology under Symbiosis International University.