Journal of Science and Technology
ISSN: 2456-5660 Volume 06,Special Issue 01, August 2021,

www.jst.org.in DOI: https://doi.org/10.46243/jst.2021.v6.i04.pp15-24

Image-Based Recommendation System using JPEG-Coefficient and RFs Approach

Anshuli Kumari[1], Prof. Nandini Babbar[2]

[1,2]*Dept. of Computer Engineering, NBN Sinhgad School of Engineering, Pune, Maharashtra, India*

To Cite this Article
Anshuli Kumari, Prof. Nandini Babbar, "Image-Based Recommendation System using JPEG-Coefficient and RFs Approach", Journal of Science and Technology, Vol. 06, Special Issue 01, August 2021, pp15-24.
.

Article Info
Received: 15.07.2021 Revised: 24.07.2021 Accepted:10.08.2021 Published: 16.08.2021

Abstract: *Online shopping platforms are expanding at an unstoppable rate all over the world. These platforms mostly depend on search engines, which are still primarily based on the text-base and use keywords matching for finding similar products. However, customers want an interactive platform that would be easy, convenient and reliable for searching related products. In this paper, we have proposed a novel idea of searching for products on an online shopping platform using an image-based approach. In this, a user can provide, select, or click an image, and similar image-based products will be provided to the user. The proposed recommendation system is based on content-based image retrieval and is composed of two major phases; Phase 1 and Phase 2. In Phase 1, the proposed way would find the class or type of the product. In Phase 2, the recommendation system retrieves closely matched similar products. For Phase 1, the approach creates a model of products using Machine Learning (ML). Then the model is used to find the category of the test products. From the ML perspectives, we have used the Random Forests (RF) classifier, and for feature extraction, we have used the JPEG coefficients. The dataset worked upon here includes 20 categories of products. In Phase 1, the evaluation of the proposed model generates a 75% accurate model. For further enhancement of performance, the RF model has been integrated into the Deep Learning (DL) setup achieving 84% accurate predictions. Based on the customized evaluation approach for Phase 2, the proposed recommendation approach achieves 98% correct results, thus demonstrating its efficacy and accuracy for the product recommendation and searching in the daily life routine and practical applications.*

Key Word*:* *Recommendation system, machine learning, random forests, deep learning.*

I. INTRODUCTION

The online retail websites are rapidly growing and their popularity is exponentially increasing. As in the Nielsen Global Connected Commerce Survey (2015) [2], 63% of respondents who bought the products using online services during the past six months. Online shopping customers face the inconvenience of sorting or selecting products from a varied number of available items on the shopping platform. The number of these products is exponentially increasing due to a large number of companies adding to online businesses. On one hand, this business growth is useful but on the other hand, it creates a problem for the user to accurately and optimally buy the desired item.

Thus, these challenges and concerns have raised the demands for recommendation systems to facilitate the users with a convenient and comfortable approach. Traditional Ecommerce based search engines are still struggling because most of these services use text-based searching. Therefore, the traditional search paradigms of the text description can be replaced by the visual search for the product recommendation system. A picture of any product should clarify and fulfil the user's demands for the appearance, usage, brand and type of the desired products.In this paper, we have proposed an idea to search the products effortlessly in an online shopping system using image-based searching techniques. The proposed recommendation system consists of two major phases:

In Phase 1, the proposed approach takes advantage of ML for learning the features of the image or product and generates a learned model. This model is then used to find the category or class of searched products. Once the category or type of the product is identified, in Phase 2, the JPEG feature vectors-based Euclidian distance is used to retrieve the top 20 matching products from the available dataset in the particular class of products. These selected 20 items are then further simplified or processed by the "Struct-Hist" approach for retrieving the top 10 most relevant products.

From the ML concept in Phase 1, we propose the Random Forests (RF) meta-classifier due to its generalization capabilities and excellent performance. For finding the class of the products and the feature extraction from images, we use the JPEG coefficients features. The dataset used here contains product images having 20 categories of products and labels from the Amazon website. In Phase 1, the evaluation of the modelbeing proposed generates a 75% accurate model and for further performance improvements, the RF model is further integrated into the DL setup and achieves 84% accurate predictions. Based on the custom evaluation approach for phase 2, the proposed idea delivers 98% correct recommendations and demonstrates its efficacy. From the implementation aspect of the proposed recommendation model, an online user or customer selects or clicks an image of the product they are looking for and similar products or images are presented to the user.Thus, our used approach is based on learning features from user-based images and recommending or providing similar products based on these images, and this process is more of an assumption.

II. LITERATURE REVIEW

In [3], the proposed approach depicts the human sense and thinking of the relationships between objects based on their appearances. The approach is based on the human perception or imagination of visual connections between products. For image-based recommendation, the authors in [4] aim to recommend images using Tuned Perceptual Retrieval (TPR), Complementary Nearest Neighbour Consensus (CNNC), Gaussian Mixture Models (GMM), Markov Chain (MCL), and Texture Agnostic Retrieval (TAR). The authors report that the CNNC, GMM, TAR, and TPR are easy to train. In [5], the authors focus on learning similarity between different products by using CNN for various products in a single image.

III. CLASSIFICATION AND FEATURES

Here are some classification methods and the features we have used in the proposed recommended approach.

*A.*RANDOM FOREST(RF) CLASSIFIER

According to an RF algorithm by Leo Breiman [6], more number of trees in the forest, higher is the performance of the classification model. The RFbelongs to the ensemble classifier models, and are primarily brought in use for problems including classification and regression. The RF algorithm operates by producing decision trees in the training phase. These trees output the class labels, and the final decision on the selection of the class is dependent on the majority voting of the trees. The below given Figure 1 depicts the flow of the RF classifier. A sample'n' from 'N' samples is given as input to the RF classifier. Firstly, the model generates several trees by using feature vector subsets. The sample being tested is assigned to that class which evaluates to the highest voting scores from the trees.

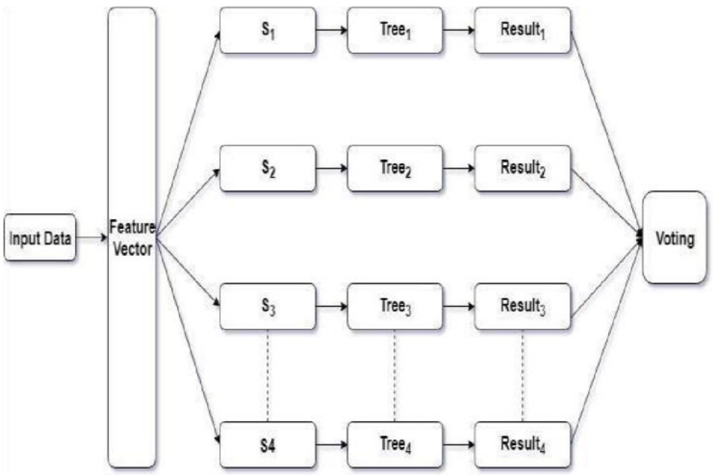


Figure 1: A general depiction of the RF Classifier Model.

B. CONVOLUTION NEURAL NETWORKS (CNN)

CNN has proved to be an easy-to handle and a very innovative tool of Deep Learning to study about the image feature sets and depict relationships in low-level features to higher-level objects in images. Due to its general architecture, it contains interconnected layers. It has repeated convolutional blocks, Rectified Linear Units (ReLU), and pooling layers. Convolutional layers performs the task of convoluting the input with the help of applied set of filters. These filters are learned during the training phase. The non-linear behaviour in the data is modelled by the ReLU layer [7]. The pooling layer samples the input given and combines the image class into a single entity.

C. JPEG-COEFFICIENT FEATURES

JPEG Coefficient algorithm is developed by the Joint Photographic Expert Group (JPEG). This algorithm converts the original input image into the YCbCrcolor space for compression purposes in order to reduce the size of the image along with maintaining its optimal quality. The image is divided into either 8×8 or 16×16 pixel blocks, depicted as example in Figure 2 below. Further, Discrete Cosine Transform (DCT) method is applied to the 8×8 pixel window which generates 64 values. Then with the help of the JPEG quantization process, the high-frequency values of an image are eliminated, and the low-frequency details are stored for further use. Before the recommendation task, we extract the JPEG image features of each image in our data set. For feature detection and feature vector construction, we have applied the JPEG Coefficient algorithm.

Journal of Science and Technology
ISSN: 2456-5660 Volume 06,Special Issue 01, August 2021,

www.jst.org.in *DOI: https://doi.org/10.46243/jst.2021.v6.i04.pp15-24*

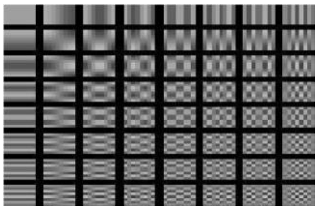


Figure 2: Image showing 8 × 8 pixel JPEG block matrix.

IV. THE PROPOSED MODEL

The approach we have proposed here is based on the clicking and selection of the images by the user and learning features from those images and recommending similar products based on these images. The proposed system uses Computer vision and ML and based on the assumption shown in Figure 3 it is composed of two major phases: Phase 1 and Phase 2.

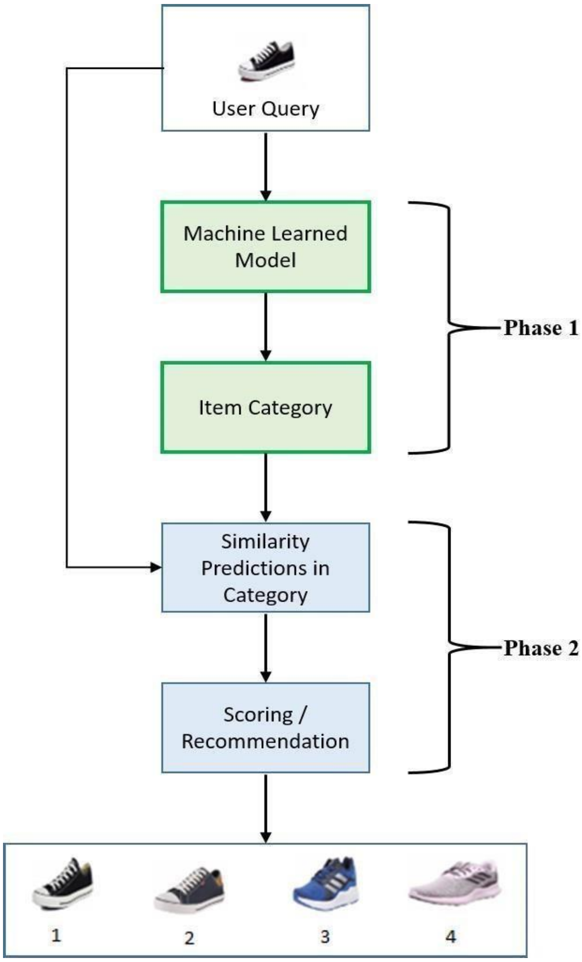


Figure 3: The flow of the proposed recommendation system where Green represents 'Phase 1' which provides the category of the image searched forand the blue colour represents the 'Phase 2' which closely identifies similar products.

Journal of Science and Technology
ISSN: 2456-5660 Volume 06,Special Issue 01, August 2021,

www.jst.org.in DOI: https://doi.org/10.46243/jst.2021.v6.i04.pp15-24

A.'PHASE 1' - CATEGORY FINDING

In this phase the proposed approach models and learns the class/category of the product based on the image characteristics selected or given by the online retailer or customer. Once the category is selected by Phase 1, Phase 2 retrieves the closely matched similar products from the corresponding category. The phase 1 is based on the training and testing paradigm of the ML. For this phase, the approach takes advantage of the ML for learning various aspects of features of the image or product characteristics and produces a learned model. The below Figure 4 shows the first phase of the recommendation system, which is the detailed explanation of the green blocks in Figure 3.

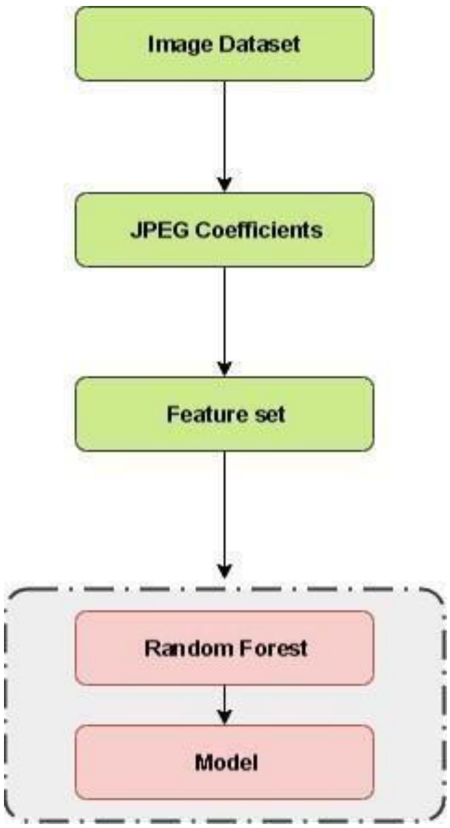


Figure 4: Phase 1 of the recommendation model using a machine learning approach.

B.'PHASE 1' - DEEP LEARNING FEATURES INTEGRATION

The structure of the Deep Learning-based RF (DL-RF) approach is composed of five different Convolution Layers shortly termed as CL.The CL is then again followed by 3 fully connected layers whereineach layer uses a kernel for filtering purposes. The Kernel coefficients are calculated in an incremental way during the training phase. Various features from DL are obtained from the seventh (C7)layer and RF finds about the class or category distribution or category of the images based on these DL features only.

www.jst.org.in DOI: https://doi.org/10.46243/jst.2021.v6.i04.pp15-24

C. 'PHASE 2' - IMAGE BASED RECOMMENDATION

For the given image, the image category is found out by Phase 1. The query image is then searched in the corresponding category in Phase 2. As shown in Figure 5, the category images are loaded with the query image. The JPEG features are extracted from all the images falling in a particular category as well as for the query image. This further put both the category images and the query image in the same vector space. The next step is to find the similarity between the feature vector of the query image and the feature vectors of the category images. For similar products selection and vector matching, we use the Euclidean distance method.

The scores based on their similarity are then sorted in the ascending order, and the most relevant 20 images are selected as the possible featured pictures for the recommendation system. Then these 20 images are further studied and analysed by image-based similarity matching process named Structural Similarity Index (SSIM) which includes color histograms in the matching process and it is represented as the ''Struct- Hist'' matching process. So, a total of 20 items are retrieved based on the Euclidian distance and then the 10 most similarly related products out of these are selected by the Struct-Hist approach. The above two used approaches always produce very accurate results.

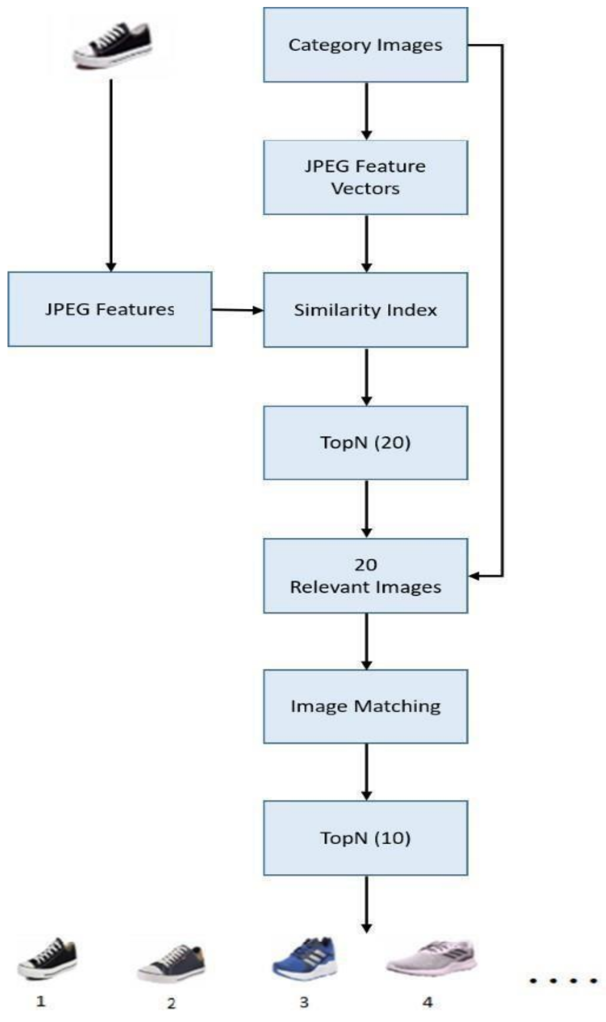


Figure 5: Second phase of the recommendation system whichclosely identifies the similar images from that of the query image after the category of the product is found
in Phase 1.

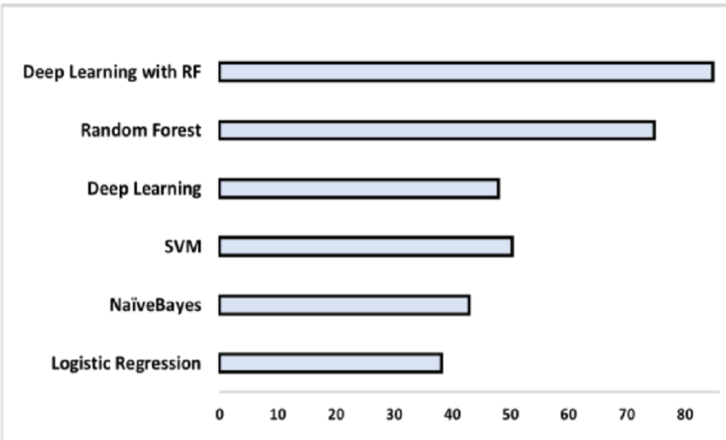
V. EXPERIMENTAL EVALUATION

Here we have presented the dataset, performance evaluation of the two phases: Phase 1 and Phase 2 of the proposed recommendation system along with its parameters.

*A.*DATASET AND FEATURES

Here, we have used the Amazon product image data. We selected a dataset of 3.5 million products that consists of 20 categories. Then we randomly chose 100 images for each class. Datasets origin is referenced from [8] .

B.'PHASE 1'- AND EVALUATION ANALYSIS OF PERFORMANCE



The features retrieved by the RF are based on JPEG coefficients. Some of the reasons for choosing the JPEG Coefficients as feature vectors is because they are computationally feasible and efficient due to the DCT transformation. Also, because of the pre-comparative analysis in our work. In the preliminary experiments, we compared HOG, Color Layout, and auto-correlogram features and from there we obtained the highest performance for the JPEG features in all cases. Also, the JPEG feature calculation is much faster compared to all other features.

For performance analysis in Phase 1, we focused on the 10folds cross-validation as a training and testing paradigm for the validation of the class retrieval. It uses 90% data for training and 10% data for testing. For 90% of training data, the model is created and tested on the 10% testing data.

For the calculation parameters of Phase 1, we use the Precision, Recall, and the F-measure. The precision and recall are mostly trusted for accuracy when the classes are not balanced as we have here in our case of dataset. The Fmeasure takes both the Precision and Recall for calculation which can be relied upon. Its equation is mentioned below: $Precision\ Recall,\ F - measure = 2\ (1)\ Precision + Recall.$

Figure 6: Graph showing Phase 1 Performance evaluation based on Precision where the X-axis hows the % Precision values. Deep Learning is the approach where features are extracted using JPEG and classified using Deep learning. 'Deep learning with RF' is where the features are extracted by the Deep learning and then they are learned by the random forest classifier.

C. 'PHASE 2'-EVALUATIONASPECTS

The evaluation process of Phase 2 of this system is not as easy as compared to that of Phase 1. Hence, we adopted the Auto correlogram vector-based Euclidian distance method for evaluation of Phase 2. For its evaluation, 10 similar images and the query images are calculated. Then each retrieved image (vector) is compared to the query image by three approaches. They are:

1. Subjective similarity
2. Euclidian distance-based similarity
3. Cosine similarity.

The presented approach goes through two steps. First retrieving 20 images, and then selecting 10 images using Struct-Hist approach which outperforms the others and shows its efficacy for the image-based product recommendations. We believe that the combination of the above two steps always produces accurate and satisfactory recommendation results with 98% accuracy as shown in the Figure 7 below.

Figure 7: Example retrievals.

D. TIME COMPLEXITY OF VARIOUS STEPS

Table 1 drawn below shows the time taken by different processes we have used in the proposed recommendation system. The time given here is represented in seconds and it is the average of several runs. As seen in the Table 1, the average time taken for calculating and learning the JPEG features by RF is 35 seconds. The time utilized for finding the category of the query image by the JPEG and RF is 0.03 seconds. The time for calculating and learning the Deep features by RF is 1000 seconds. Time taken to find the category of the query image by the Deep Learning features and RF is 0.3 seconds. The average time taken to find the top 20 related items in 'Phase 2' is 1 second. And lastly, the time taken to retrieve the top 10 images out of 20 associated items in 'Phase 2' is 0.003 seconds.

Table No. 1: Some parameters with their respective Time Complexities.

Parameters	**Avg. Time (Seconds)**
Time for calculating and learning the JPEG features by RF	**35**
Time to find the category of the query image by the JPEG and RF	**0.03**
Time for calculating and learning the Deep features by RF	**1000**
Time to find the category of the query image by the Deep features and RF	**0.3**
Time to find the top 20 related items from the dataset	**1**
Time to retrieve the 10 related items out of 20 already retrieved	**0.003**

VI. CONCLUSION

We presented an image-based product recommendation system comprising of two phases. In Phase 1, the proposed model finds out the class or type of the products being searched for. In Phase 2, the proposed recommendation system retrieves the category of the closely matched similar products. We used Computer Vision and ML in phase1 of learning the class of a product, using the RF classifier. For feature extraction from images, we used the JPEG coefficients as image features. In the evaluation of Phase 1, the system generates a 75% accurate model. But for further performance enhancements, the RF model is further integrated and passed into the DL setup and achieves 84% correct evaluations and predictions. In Phase 2, the combination of two steps used in the proposed approach of finding the 20 similar items based on the Euclidian distance and from that 20 we found 10 most relevant products using Struct-Hist approach, was believed to be the reason of the production of such efficient and accurate recommendations producing a 98% accurate model. The Struct-Hist approach thus overpowers and performed with great results than the other methods and shows its efficiency for the image recommendation. The article proposed here contributes not only to the recommendation-based systems but also the algorithm presented here in the article can be used for general computer vision problems and hence providing an interactive platform to the customers based on image-based selection while shopping online.

www.jst.org.in *DOI: https://doi.org/10.46243/jst.2021.v6.i04.pp15-24*

REFERENCES

[1] F. Ullah et al.: "Image-Based Service Recommendation System: A JPEGCoefficient RFs Approach", date of publication December 25, 2019, date of current version January 7, 2020. Digital Object Identifier 10.1109/ACCESS.2019.2962315

[2] Nielsen. (2015). Green Generation: Millennials Say Sustainability is a Shopping Priority. The Nielsen Global Survey of Corporate Social Responsibility and Sustainability. Accessed: May 3, 2019. [Online]. Available: https://www.nielsen.com/us/en/insights/article/2015/greengenerationmillen nials-say-sustainability-is-a-shopping-priority/

[3] R. He and J. McAuley, ''Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering,'' in Proc. 25th Int. Conf. World Wide Web, 2016, pp. 507–517.

[4] V. Jagadeesh, R. Piramuthu, A. Bhardwaj, W. Di, and N. Sundaresan, ''Large Scale Visual Recommendations From Street Fashion Images,'' in Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, pp. 1925–1934, 2014.

[5] S. Bell and K. Bala, ''Learning visual similarity for product design with convolutional neural networks,'' ACM Trans. Graph., vol. 34, no. 4, p. 98, 2015

[6] L. Breiman, ''Random forests,'' Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ''ImageNet classification with deep convolutional neural networks,'' in Proc. Adv. Neural Inf. Process. Syst., 2012.

[8] Amazon Dataset. Accessed: Sep. 3, 2019. [Online]. Available at: https://drive.google.com/drive/folders/1eaCPVwj9ORkVHA4JYn0QFXkH milukm

[9] Ganesh Shinde and Altaf Mulani, "A Robust Digital Image Watermarking using DWT-PCA", International Journal of Innovations in Engineering Research and Technology (IJIERT), Vol. 6 Issue 4 April 2019.

[10] Kulkarni P.R., Mulani A.O. and Mane P. B., "Robust Invisible Watermarking for Image Authentication", In Emerging Trends in Electrical, Communications and Information Technologies, Lecture Notes in Electrical Engineering, vol 394, pp 193-200, Springer, Singapore, 2017.