iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** I **Month of publication:** January 2023

DOI: https://doi.org/10.22214/ijraset.2023.48656

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

Fashion Advisor: Product Recommendation System

Vaibhav Goel
Maharaja Agrasen Institute of Technology, Delhi, India

Abstract: ***In the fashion domain, predicting compatibility is a significantly difficult task due to its subjective nature. Previous work in this domain focuses on comparing product images rather than a real-world scene. This fails to capture key context like body type, seasons, and other occasions in the scene. This is an important use case which needs to be addressed. Here, we propose a task 'Fashion Advisor' which deals with measuring compatibility between a real-world scene and a product. We use two compatibility scores, global and local where global compatibility considers the overall scene and the local compatibility focuses on finer details of the scene using category guided attention mechanism. There were many different baseline methods compared with the proposed method and the proposed method gives promising results on Fashion and Home Datasets.***
Keywords: Global Compatibility, Local Compatibility, Attention, Resnet

I. INTRODUCTION

In today's world, everyone wants to remain classic. Fashion, an aesthetic expression, has become a significant part of our lives of how we present ourselves to the world. With a rapid change in Fashion trends and living standards, Selecting, and contrasting numerous potential solutions takes a lot of time and effort. Unlike consumers, it is more difficult for merchants to think about what kind of thing consumers want. The Fashion Outfit Recommendation has gained more attention due to the E-Commerce Websites and social media fashion communities.

A survey found that personalised suggestions accounted for almost 35% of Amazon's sales revenue, boosted engagement by 50% when a new visitor first logs in, and increased time spent there by 344% on average (12.9 minutes vs 2.9 minutes for those who did not click recommendations).

Humans are inevitably drawn to things that are more visually appealing. This human tendency has resulted in the evolution of the fashion industry over time. With the development of recommender systems in a variety of disciplines, retail firms are investing in cutting-edge technology to better their bottom line. In the fashion industry, image-based product suggestion is a key task.

Existing work focuses mainly on recommending products that are compatible with an input product image. Although in a realistic setting, input images can be a real-world scene with a person wearing different fashion products (E.g., selfies). The proposed task focuses on recommending complementary products to the given real-world scene. This is a challenging task because of the difficulty involved in handling real-world images. Such images often have varied illumination, clutter and are taken under varying conditions. This is also a useful task since it has many real-world use cases. For e.g., we can give personalised recommendations based on the user's body type and products that suit well for a particular season.

Product Recommendation introduces a choice-driven system to retrieve the best outfit items for the given scene. The proliferation of e-commerce websites in India has resulted in an upsurge in online buying across the country. In the e-commerce market, fashion items are the most sought for. A strong, effective, and efficient fashion recommendation system could enhance revenue and improve client engagement and experience. The ability to recognise many products from a same Product Display page and recommend related things to the customer distinguishes this approach. Specific products from the dataset are shown in Figure 1.



Figure 1: Fashion Products

Image-based product recommendation is a significant task in the fashion domain. Existing work focuses mainly on recommending products that are compatible with an input product image. Although in a realistic setting, input images can be a real-world scene with a person wearing different fashion products (E.g., selfies).

©IJRASET: All Rights are Reserved

II. LITERATURE REVIEW

A. Visually-Aware Fashion Recommendation

Wei-Lin Hsiao et al [1] proposed a Visual Body-aware Embedding (ViBE) that aims to capture the relationship between diverse body shapes and different fashion products. It specifically focuses on identifying different garments that complement well with a person's body shape. Features of clothing and body shapes are extracted using a Resnet-50 model pre-trained on ImageNet and are combined to get a joint representation. This is then used to learn an embedding which measures body-clothing affinity. The proposed approach only recommends garments and there is scope for scaling the model for clothing recommendation.

Wang-Cheng Kang [2] used an end to end learning approach based on the Siamese-CNN framework to build the Visually aware fashion recommendation system, which returns all possible visually similar images in the result. Later they used a GAN (Generative Adversarial Network) to maximize the user preferences by generating more novel fashion items. Currently, this system recommends items that are like query items, so, In the future, other than improving the quality of recommended images, they want to make it more personalized for users on both visual and non-visual forms of data.

B. Fashion Outfit Complementary Item Retrieval

Mane et al. [3] worked on textual data to identify functionally similar as well as complementary fashion items. An instance was treated as a quadruplet of the form ¡ a, c, s, n ¿ where a is the anchor item, c is the complementary item, s is the similar item and n is the negative item. Features are extracted from each of the items' title information using the Universal Sentence Encoder [3]. They use these features to learn a Quadruplet Loss which differentiates similar items and complimentary items and similar and dissimilar items by a larger margin. Yen-Liang Lin et al. [4] introduced a new method for retrieving similar items to complete a perfect outfit. It suggests a choice-driven technology to retrieve the best outfit items for the given attire. Let us say, for a pant, shirt, and shoe combination, it finds a black lathered hand-bag. This model classifies items with a category-based attention selection mechanism and performs indexing and searching to find out enhancing items which can be fit in the outfit.

C. Hierarchical Fashion Graph Network for Personalised Outfit Recommendation

In Xing Chen Li [8] has somehow tried to unify the fashion compatibility modelling and the personalised outfit recommendation. A hierarchical structure was constructed by mapping the user-outfit and outfit-items interactions, which is Hierarchical Fashion Graph Network using Graph Neural network. The graph consists of three levels - User, Outfit and Items represented by nodes having connections at cross level. The nodes are assigned with a unique ID and are initialised with embeddings. The node embeddings are refined by aggregating the transferred information from lower to higher levels in the Hierarchical graph convolution by three embedding propagation steps as information propagation across items, from items to outfits and from outfits to users. The model then predicts the score related to the personalised recommendation and compatibility prediction. The graph model does not incorporate any textual feature which can help in refining the node embeddings [6].

III. DATASET

A. Kaggle Fashion Dataset

The pictures are from the Kaggle Fashion Product Images Database. After the inventory has been classified and embeddings have been produced, the result is used to produce suggestions. The final dataset consists of about 50000 images that we collected from the internet by scraping the web and various clothing retail sites as shown in Figure 2. To train the models, a very specialised dataset was gathered. To create the feature vector and graphics that are specific to the classes of these features, the stated specific features must be present. To remove duplicates and irrelevant photos, data is automatically curated and cleaned. Images that were not specific to the class they were intended to belong to for training are removed using a manual check. The training and validation datasets can be created from the photos, and a well-represented portion can be set aside for the test.



Figure 2: Kaggle Fashion Dataset

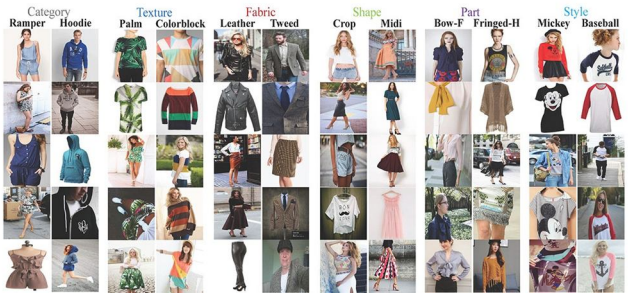


Figure 3: Deep Fashion Dataset

B. Deep Fashion Dataset

A huge database of clothing called Deep Fashion is well-liked by researchers. It has approximately 800,000 unique fashion images, ranging from loose consumer photos to perfectly staged shop photos. It is richly annotated with information about clothing. More than 300,000 cross-pose/cross-domain image pairs are also included. A selection of the photos from the Deep Fashion Attribute Prediction is specially utilised. There are 1,000 different clothing qualities, 50 different clothing categories, and 289,222 different clothing photos as shown in Figure 3. Each image has bounding box and clothing type annotations.

IV. PROPOSED METHOD

A. Pre-processing

The datasets used for Shop the Look tasks cannot be directly applied. If the product is present in the scene image, it will make the model biased. The model might only learn to recommend products which appear in the scene. So, for training a generalised model, the products must be cropped out from the scene image. Given the scene image and bounding box surrounding the product, we take the portion of the scene image which appears in the top, bottom, left and right regions of the bounding box and consider that region which has the maximum area [5]. All the missing/NaN values are being handled which images are being broken, or the json files bounding boxes coordinates are not cleared properly. The specific features listed will be required to form the feature vector and extracted images specific to the classes of these features. An automated curating and cleaning of data is done to eliminate duplicates and non-relevant images. A manual check is done to eliminate images that were not specific to the class they were meant to belong to for training. We can use the images as the training and validation dataset and separate out a well-represented section for the test.

B. Modelling

In this project, we suggest a model that makes use of a convolutional neural network and a recommender system backed by neighbours. According to the image the neural networks are initially trained, after which a database is made for the objects in the inventory and an inventory is chosen for the purpose of providing suggestions. Based on the supplied image, the nearest neighbour's algorithm is utilised to locate the most pertinent products, and suggestions are provided.

C. Feature Extraction

In this step, the Resnet-50 model is used to extract the scene's and the product photos' features. The feature map derived from an intermediate Resnet block is utilised to compute the embedding after that. After pre-processing the data, transfer learning from ResNet50 is used to train the neural networks. In order to fine-tune the network model to address the current issue, more layers are included in the final layers which replaces it from the ResNet50 weights and architecture. The ResNet50 architecture is depicted in the Figure 4.

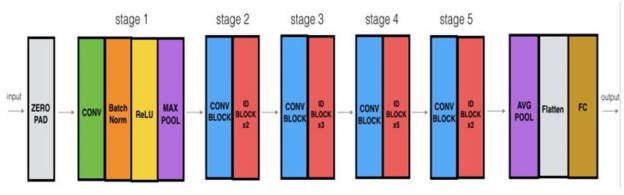


Figure 2. ResNet50 architecture

Figure 4: Resnet Architecture

To implement the proposed model as referred in Figure 5, we needed to build models to initially classify the categories of features. Using these models, we form the fashion vector i.e. a vector with all categories of features classified for each image. With these vectors, we build a style profile for a user based on the input images. We match this profile with fashion vectors of images in repository to form the best suited recommendations.

Also, trying another Network architecture, we use Keras wrapper for TensorFlow and VGG16 model as the base of the network. The base of our network is VGG16. On top of the network, we add a customised dense layer of 512 neurons with RELU activation. For the final classification, we use a SoftMax layer of n output neurons where n is the no of subclasses for each category of feature. We use a pretrained VGG16 model on ImageNet as the starting weights / parameters. With these stacked CNNs, each input image entered is classified into one of the classes for each category , which are then together combined to form a Fashion Vector for that image.

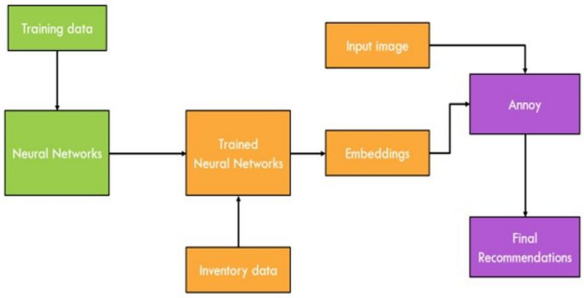


Figure 1. Block diagram of proposed system

Figure 5: Proposed Model Block Diagram

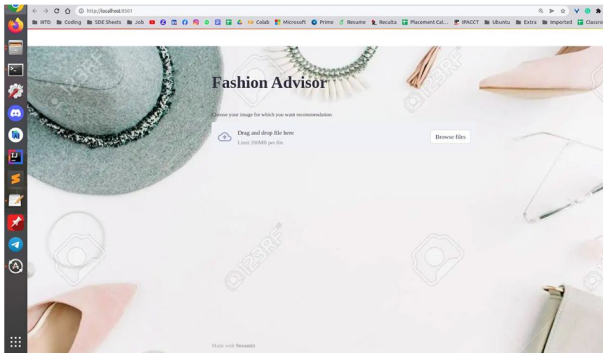


Figure 6: Application Front Page

D. Recommendation Generation

Our recommended method makes use of Sklearn Nearest Neighbors to produce recommendations. This enables us to locate the input image's closest neighbours. The Cosine Similarity measure is the similarity metric employed in this project. From the database, the top 5 recommendations are pulled, and their photos are displayed. The application recommendations are shown in Figures 6, 7 and 8.

We use embedding generation to represent images/products such that similar ones are grouped together whereas dissimilar ones are moved away so that in order to retrieve products that are like the products present in the query image. There are various ways to calculate similarity between images after they are converted into a vector with n-dimensions. Cosine similarity and Euclidean distance are two examples.

V. EVALUATION

Transfer learning is utilised to overcome the limitations of the tiny Fashion dataset. As a result, we pre-train the classification models on the Deep Fashion dataset, which contains 44,441 garment photos. On the dataset, the networks are trained and evaluated. Through the trained model and networks, the results show that the model is trained accurately with low error, loss and achieving high f-score as shown in table 1.

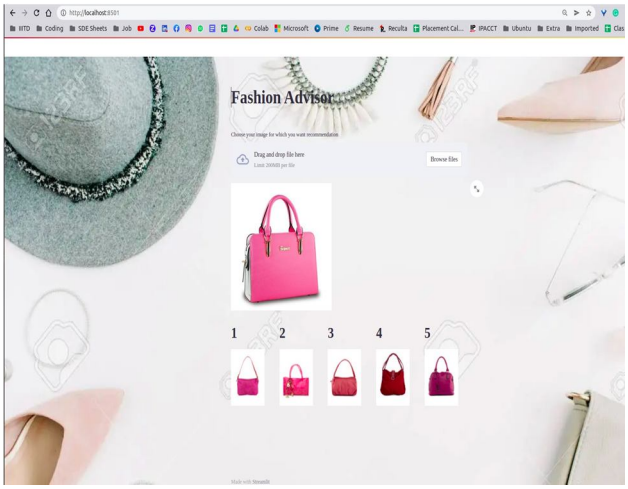


Figure 7: Recommendations by Application for a purse

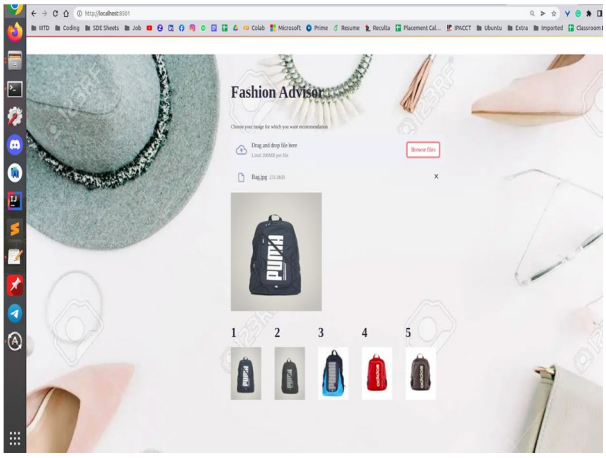


Figure 8: Recommendations by Application for a Bag

We observed that the baseline models that were pretrained on Image-Net features such as Res-net and VGGNet performed better, which revealed that visual compatibility is distinct from visual similarity which makes it important to learn the notion of compatibility from the data. The model was trained with the images, which made it more effective and thus, boosting the performance of the model. Our proposed method deals with both global and local appearance thus considering the key context of the scene. Using category guided attention enhances the ability of the model to identify key details of the scene image and compare it with the product. It takes a decision by incorporating the scene, product, and the category of the product as well. The accuracy and loss graphs with respect to epochs observed while training is shown in the Figures 9 and 10.

We can run the web server, by executing the command stream lit with the main recommender app. You can upload an image and then the model will recommend the top 5 similar products using the k Nearest Neighbours Algorithm [7].

Method	AUC	Precision	Recall
ResNet-50	0.9522	0.9443	0.9366
VGG-16	0.9057	0.8922	0.8821
Proposed Method	0.8573	0.8311	0.8197

Table 1: Results on Test Dataset

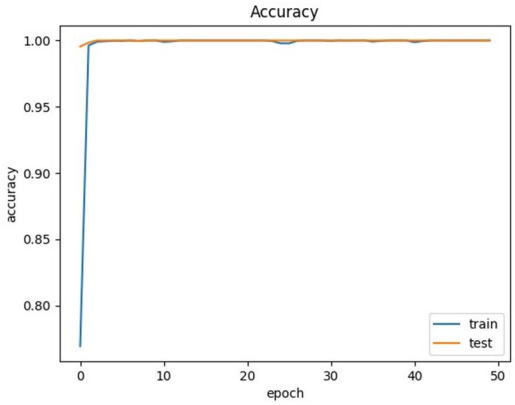


Figure 9: Accuracy vs Epochs

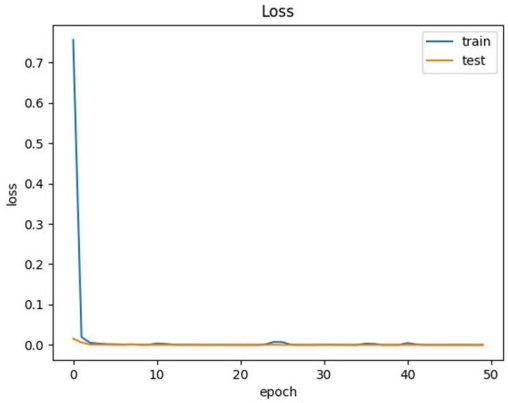


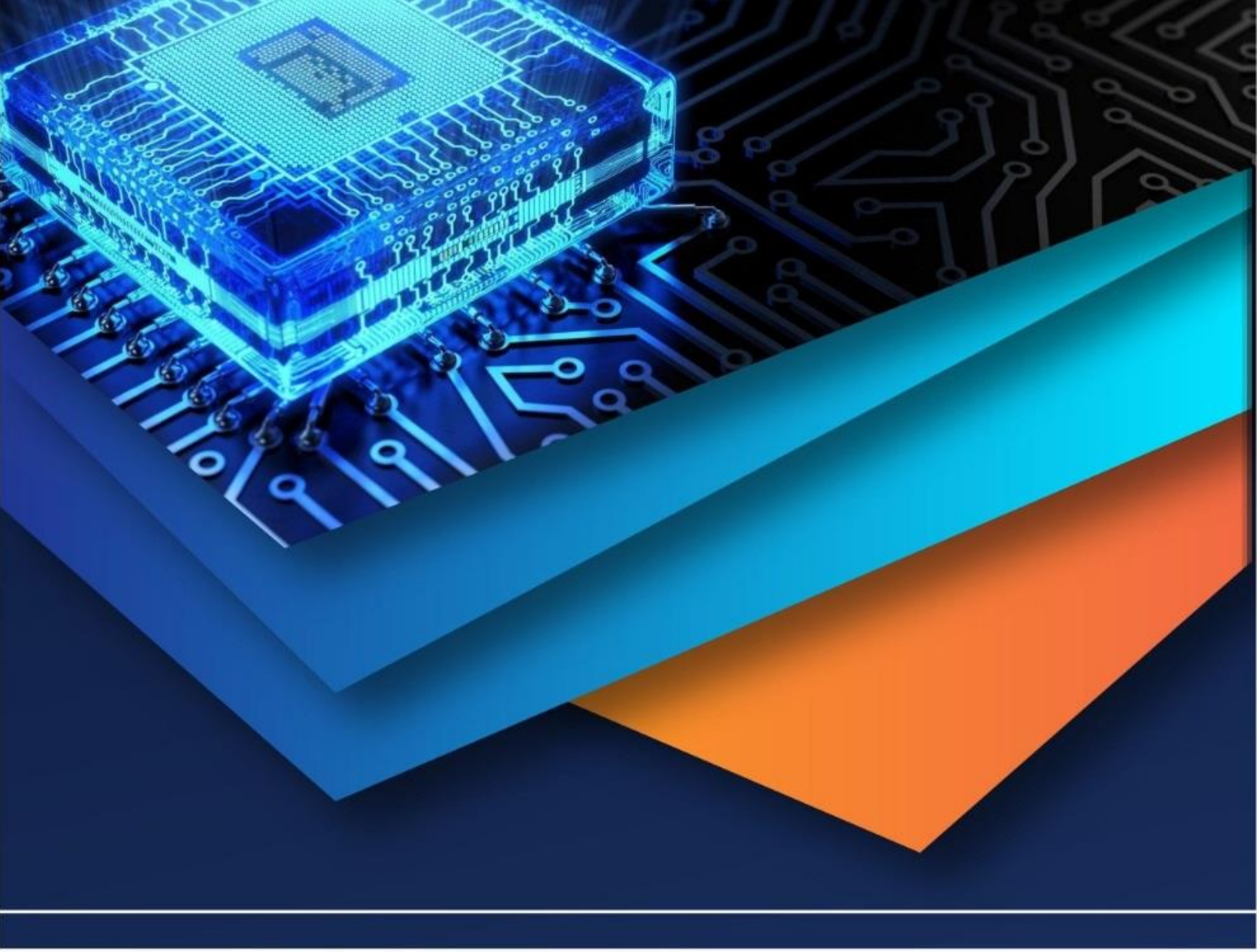
Figure 10: Loss vs Epochs

VI. CONCLUSION

We proposed a revolutionary framework for fashion recommendation that is data-driven, visually connected, and simple effective recommendation systems for creating fashion product photos in this research. The proposed method is divided into two stages. Initially, our suggested solution extracts picture features using a CNN classifier, for example, by allowing users to upload any random fashion image from any E-commerce website and then creating similar images to the uploaded image depending on the features and texture of the input image. It is critical that such research continue in order to improve recommendation accuracy and the entire experience of fashion discovery for both direct and indirect consumers.

REFERENCES

[1] W.-L. Hsiao and K. Grauman, "Vibe: Dressing for diverse body shapes," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 11059–11069.
[2] W.-C. Kang, C. Fang, Z. Wang, and J. McAuley, "Visually-aware fashion recommendation and design with generative image models," in 2017 IEEE International Conference on Data Mining (ICDM). IEEE, 2017, pp. 207–216.
[3] M. R. Mane, S. Guo, and K. Achan, "Complementarysimilarity learning using quadruplet network," arXiv preprint arXiv:1908.09928, 2019.
[4] Y.-L. Lin, S. Tran, and L. S. Davis, "Fashion outfit complementary item retrieval," 2020.
[5] X. Li, X. Wang, X. He, L. Chen, J. Xiao, and T.- S. Chua, "Hierarchical fashion graph network for personalised outfit recommendation," in Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022, pp. 159–168.
[6] V. D. M. Ernani, E. A. Nogueira, and D. Guliato, ContentBased Filtering Enhanced by Human Visual Attention Applied to Clothing Recommendation 2022 in Proceedings of the International Conference on TOOLS with Artificial Intelligence, pp. 644–651.
[7] L. Chen, F. Yang, and H. Yang. Image-based product recommendation system with convolutional neural networks, 2021.
[8] Praveen P., Rama B(2021). "An Optimised Clustering Method To Create Clusters Efficiently" 2021 Journal Of Mechanics Of Continua And Mathematical Sciences , ISSN (Online) : 2454 -7190 15(1), pp 339-348.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)