Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
January-June 2023 4(1), 1051–1062

REVOLUTIONIZING TREND RECOMMENDATIONS: A DEEP LEARNING APPROACH FOR IMAGE-BASED INSIGHTS

V. Srilakshmi [1], Ashlin Deepa R N [1], Ch. Vidyadhari [1], Srihithagunapriya Nimmala [1], Mallikarjuna Rao Gundarapu [1]

[1] Gokaraju Rangaraju Institute of Engineering and Technology Hyderabad, India

Corresponding Author
Ashlin Deepa R N,
deepa.ashlin@gmail.com

DOI
10.29121/shodhkosh.v4.i1.2023.2859

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

The rapid increases in visual content creation across different sectors, starting from social media to even e-commerce, have had a major impact on their respective users, who want bespoke experiences: preferences are recognized and satisfied by recommendations contextualized within a specific sense of the world. Over many years, traditional recommendation methods took into consideration only textual content, user profiles, or purchase history as the material for recommendations. As images become the staple of engagement in the digital world, the old system misses those very nuances of visual preference that affect choice through color, texture, and style.

This shift has paved way for image-based recommendation systems. Indeed, these are further dependent on powerful machine learning and deep learning models in the analysis of visual data to discern meaningful patterns in drawing recommendations in line with a user's visual taste. In such industries as fashion and design, social media, style and aesthetic become much stronger drivers of engagement.

We propose in this paper a sophisticated AI-based Trend Recommendation System, which combines deep learning model feature extraction abilities with the precision of the Nearest Neighbor Search algorithms. Our system has at its core the ResNet50 model, a widely recognized CNN which is associated with superior ability to analyze images. This is achieved using ResNet50 to extract features from the images at a deep level so that their style and trend may be characterized with intricate visual characteristics. These features are then compared to a large curated dataset using Nearest Neighbor Search, in turn ensuring that the recommended images are not only visually relevant but also contextually accurate.

Our research expands ResNet50 by evaluating other state-of-the-art CNN models, including VGG16, InceptionV3, and MobileNetV2. These models are tested on their ability to extract features, provide recommendations with accuracy, and consume fewer computations. By comparing these models, we will know which one is more effective for real applications for speed and precision.This study is a landmark in developing an image-based recommender system and finds how deep learning has the transforming capability to make experience highly personalized and visually oriented.

Keywords: Deep Learning, Image-Based Recommendation, Convolutional Neural Networks (CNN), Feature Extraction, Image Analysis

1. INTRODUCTION
1.1. BACKGROUND

The rapid growth of digital platforms has drastically changed the scenario of user interaction, especially in visually-driven environments [1][2]. Users now expect experiences that align with their individual preferences from an ever-increasing flow of visual content across sectors, including social media, e-commerce, and digital marketing. Research indicates that visual content plays a significant role in purchasing decisions and user engagement, which necessitates recommendation systems that can effectively meet such changing expectations [3][4].

How to cite this article (APA): Srilakshmi, V., R.N., A.D., Vidyadhari, C.H., Nimmala, S., and Gundarapu, M.R. (2023). Revolutionizing Trend Recommendations: A Deep Learning Approach for Image-Based Insights. *ShodhKosh: Journal of Visual and Performing Arts, 4*(1), 1051–1062. doi: 10.29121/shodhkosh.v4.i1.2023.2859

Traditionally, recommendation systems depend on textual data, user profiles, and behavioral patterns for suggesting relevant items [5]. Although these methods have been effective up to a certain extent, they fail to capture the richness and multifaceted nature of visual preferences. Color, texture, style, and emotional tone are some of the factors that influence user choices, but conventional algorithms typically overlook these subtleties [6][7]. Hence, most of the users receive the default recommendations, which do not match their unique visual appearance, and this leads to frustration and decreased user engagement with digital platforms [8]. This gap shows the utmost necessity of developing highly sophisticated image-based recommendation systems that can effortlessly process the visual data more efficiently for users to discover content that really matches their taste and preference [9].

1.2. RESEARCH OBJECTIVES

The research aims for the design and implementation of image-based Trend Recommendation System, using advanced deep learning techniques to improve the experience of the users in various areas, particularly fashion and e-commerce. Specific objectives for this paper are:

- **Feature Extraction:** Use CNN - specifically ResNet50 model - to extract deep features of visuals in images and select subtle characteristics that constitute the style and trends.
- **Recommendation Accuracy:** Use Nearest Neighbor Search algorithms to provide very accurate and contextually relevant recommendations based on visual similarities between items.
- **Performance Benchmarking:** There will be a holistic comparative study of a variety of latest architectures using CNN, such as VGG16, InceptionV3, MobileNetV2, etc. It will measure and compare their efficiencies in feature extraction, recommendation accuracy, and also possible computational efficiency.
- **User-Centric Design:** The system should user feedback mechanisms to leverage continuous learning and adaptation regarding individualized user interactions and preferences.

2. LITERATURE SURVEY

Recommender systems play a crucial role in improving the user experience on various digital platforms. This can be done by suggesting relevant content based on individual preferences. For the most part, recommender systems are categorized into three broad types: content-based, collaborative, and hybrid recommender systems. The classification of this system helps differentiate between the methodologies used in understanding user preferences and delivering personalized recommendations.

Content-based recommenders focus more on the properties of items themselves, through features that build up item profiles. In an image-based recommendation system, some of the attributes could be color, style, pattern, texture, and even emotional tone of clothing items. When a user engages positively with an image, such as liking or saving or buying, the system analyses the features of that item and suggests similar products possessing those characteristics, thus creating an implicit profile for that user based on visual features. For example, a deep learning method for music recommendation by Äaron et al. (2013) captures the subtle properties associated with audio features, demonstrating how neural networks can significantly improve similarity content recommendations based on fine attributes【1】. This suggests that, in similar fashion, deep learning applied to visual features in content-based filtering can do the same for visual recommendations and effectively capture the nuances in image attributes including color and texture. interaction metrics to provide a more comprehensive recommendation experience.

Further, Huang et al. (2013) proposed using deep structured semantic models to tap into clickthrough data for web search, which can be easily implemented in the recommendation system using user engagement metrics to update item profiles. Their method is more likely to produce personalized and context-aware recommendations【2】.By using interaction data from users, thus giving insight into the advantages of combining content-based attributes with user.

Moreover, researchers have underlined that content-based filtering must be integrated with other advanced approaches to increase recommendation accuracy. For instance, Chen et al (2019). presented the incorporation of deep learning methods into this approach for more accurate determination of subtle properties of the items and better enhancement in the recommendation of the approach on visually similar items 【1】Using additional metadata and user-generated contents in the profile of an item can also fine-tune the recommendation so that each user has the most preferred experience.

The authors demonstrated their method to outperform traditional approaches on several large-scale datasets, thus constituting a robust choice for systems that are centered on visual content and where user interaction data may be sparse. This can significantly improve the performance of trend recommendation systems by capturing implicit user preferences generated from their interactions with visual content, thus making recommendations more relevant.

Sedhain et al. (2015) added more to the field by proposing AutoRec. It uses autoencoders for collaborative filtering-based models that learn latent user and item representations from user ratings. This would hence enable a strong solving means of the cold-start problem, especially in cases with sparse data by using this approach 【4】. The trend might be improved by modeling appropriate user preferences even in cases where very limited explicit feedback is available.

Inspired by Wang et al. (2015) and collaborative deep learning that incorporates deep learning techniques along with the basic framework of a collaborative filtering model that describes user-item interactions, CDL effectively harnesses the ability of deep learning to learn complex data patterns and capitalize on the collaborative filtering technique used in managing sparse user item relationships to make more precise and detailed recommendations. It seems that the proposed method is very effective for a fashion recommendation system as rapidly changing preferences can be quickly affected by changing trends in the field【3】.

Lian et al. (2017) proposed a content-boosted collaborative filtering neural network that replaces the traditional recommendation model by integrating content information into the learning process. One hand, this approach hybridizes sparsity; on the other hand, it enriches the contents by integrating specific features in recommendations 【14】.

Zhao et al. (2018) ultimately discuss how to leverage the long- and short-term information in collaborative filtering for more accurate recommendations. His work underscores that user preferences are influenced by temporal dynamics, providing a foundation for trend recommendation systems where the temporal change of user interest has to be considered【5】. Based on this point, more timely and relevant recommendations will be made to users by recommendation systems.

Hybrid Recommender Systems combine approaches to compensate for the inability of singular systems to yield good recommendation accuracy. A content-based as well as a collaborative-based hybrid model used for an image-based recommendation combines both these techniques for a precise recommendation through the usage of image features in addition to preference of user. It thus proves really beneficial where visible features become decisive factors for their choice for the better review of what users actually desire.

This paper, **"Neural Collaborative Filtering" by He, Liao, Zhang, and Chua (2017),** introduces a deep learning approach that combines explicit and implicit feedback to improve recommendation accuracy on the basis of collaborative filtering【15】.It is a neural network architecture that integrates the strengths of the matrix factorization and multi-layer perceptron models, such that it learns latent representations of users and items and captures their preferences and characteristics. Ranking is a major objective proposed in the training objective. Ranking is significant for efficient trending item recommendation.

Mao et al. (2014) discussed a plausible solution to the very common problem of cold start in a Hybrid Recommendation Model Based on Collaborative Filtering and Content-based Filtering. In this work, they show how it is possible to improve recommendations for new users and new items by merging collaborative filtering with content-based filtering techniques. Their work therefore, highlights the importance of hybrid models in providing personalization at a time when traditional techniques fail in environments with few data.

Adding to this pool of study, Yin et al. (2021) provided a multi-layer hybrid recommendation framework based on content and collaborative filtering through their paper, A Multi-Layer Hybrid Recommendation Framework Based on Content and Collaborative Filtering. Their recommendation was built by using diverse layers of collaborative and content-based filtering that improved the resilience of recommendations with an overall highly significant increase in accuracy that was predictive. In reality, this research gives way to the idea of multiple-layered approaches that prove handy in modeling various forms of user preferences in real applications.

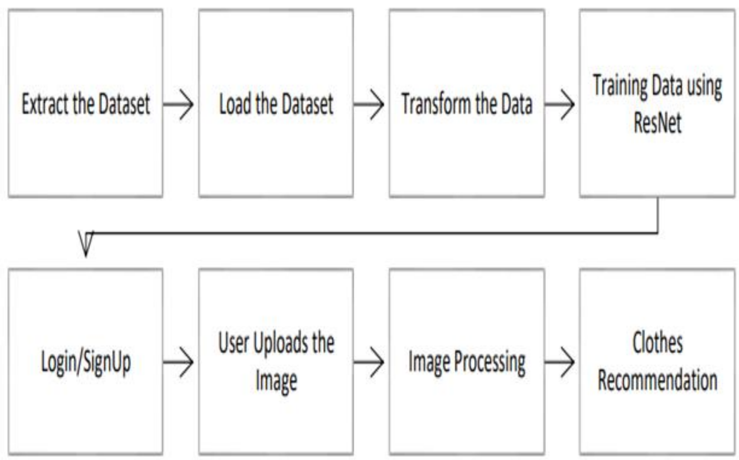
This project fills these gaps by developing an Advanced Image-Based Trend Recommendation System using Deep Learning and Nearest Neighbor Search, which, based on this hybrid technique combining visual content analysis and collaborative filtering, tends to enhance the accuracy and engagement of the recommendations offered. The framework proposed below will use deep learning techniques that will extract subtle details from images and capture their preferences using collaborative methods.

3. PROCEDURE

Input dimension is standardized by resizing every picture into a 224x224 size and converting it to an array of numbers. There is preprocessing by standardization, using ResNet50 functions for normalization with scaling of pixel values by neural networks. The architecture that was used was that of ResNet50 with a removed top layer as there is the use of weights from the pre-trained image-net dataset to fast up the feature extraction. A Global Max Pooling layer sums up all of the outcome feature maps into one vector, which represents the essence of that particular image.

The system will suggest similar fashion items using the Nearest Neighbors algorithm, which works on a brute-force approach with Euclidean distance as the metric. The system retrieves the top six most similar items from the dataset using the extracted features. Additionally, the system processes images in batches of 32 for optimal memory usage and computational efficiency, which is in line with the best practices in deep learning. This structured method enables the system to proffer appropriate and relevant fashion recommendations for a user with regard to images uploaded by that user.

Staged Feature Extraction and Recommendation Approach:



4. DETAILED ANALYSIS

1) Image Preprocessing

The process first starts with image preprocessing in which every input image has been resized to 224x224 pixels to set the uniformity of dimension in the dataset. That's important for proper analysis, and all images adhere to the input requirements of the neural network. Then the image is converted into numerical array format for processing. Another step is normalization, where the pixel values are scaled with functions that are naturally occurring in the ResNet50 model. This scales the pixel values to lie in a standard range, which is usually between 0 and 1, thereby optimizing the input to make the neural network work even better and improve the robustness of feature extraction.

2) Feature Extraction

We will use a convolutional neural network (CNN) for image feature extraction after the preprocessing stage. We can make use of a model like ResNet in this step for efficient feature extraction without requiring the pre-trained weights.

Feature Learning

The ResNet architecture is used with no pre-trained weights and solely for feature learning from our dataset:

- **Convolutional Layers:** The resized images are passed through multiple convolutional layers, where a set of learnable filters detects fundamental patterns such as edges and textures. Each filter generates feature maps that highlight these characteristics.

• **Activation Functions:** Non-linear activation functions (ReLU) applied after convolution to introduce non-linearity so that the network can learn complex mappings.

• **Pooling Layers:** Max Pooling layers reduce the spatial dimensions of the feature maps, extracting the most salient features while down sampling the data to maintain computational efficiency.

• **Residual Connections:** ResNet uses residual connections that improve the flow of gradients throughout the network, thus removing vanishing gradients and enabling deeper training of networks.

Global Max Pooling

After feature extraction, a Global Max Pooling layer reduces the output feature maps of each image into one vector, which captures significant features with reduced dimensionality for better model performance in following recommendation tasks.

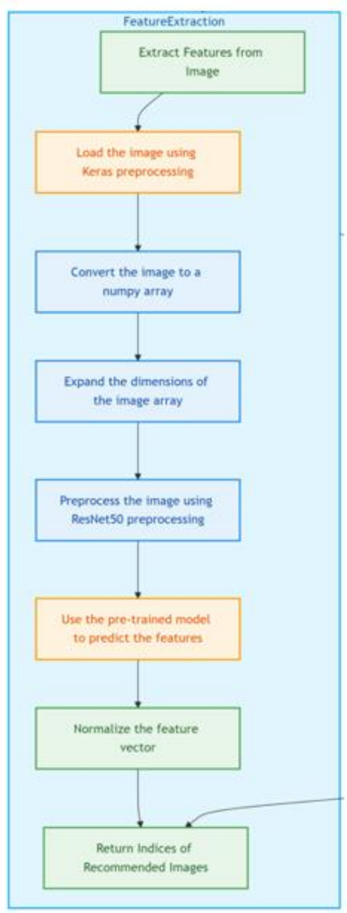
Output Preparation

The number of vectors, after feature extraction, is the set of a number of vectors, one for every fashion image that stands for each one, and they form the base of the recommendation system.

• **Storing Feature Vectors:** The extracted feature vectors are saved in a structured format, such as a NumPy array, along with their corresponding image file names. This will make retrieval and processing in the recommendation phase easier.

• **Preparing Recommendations:** Feature vectors are presented as input to a Nearest Neighbor search algorithm; these algorithms will return what looks like the closest thing, in terms of representations of features, to this user-uploaded fashion image. The process of producing recommendations uses metrics like Euclidean distance to compute and produce items that closely resemble the fashion items uploaded by the user.

• The pre-similarity search normalization step prepares feature vectors for optimum performance, based on enhancing distance metrics such as Euclidean distance and improving consistency in vector magnitudes, which helps the similarity search algorithm focus on relevant aspects of the recommendation process rather than on scalar differences.



5. RECOMMENDATION MECHANISM

To suggest similar fashion items, we apply a Nearest Neighbors algorithm with a brute-force approach to compute distances between feature vectors. The key steps involved are as follows:

• Preprocessing Feature Vector: After passing through the ResNet architecture, features are extracted and post feature extraction, output feature maps are further compressed into a compact set of feature vectors through Global Max Pooling. This keeps all the most significant features in a fashion item so that all comparisons are made efficient without losing the most essential characteristics of item.

• Distance Metric: We utilize Euclidean distance to compare the similarity in feature vector of the uploaded user image to all dataset vectors defined as:

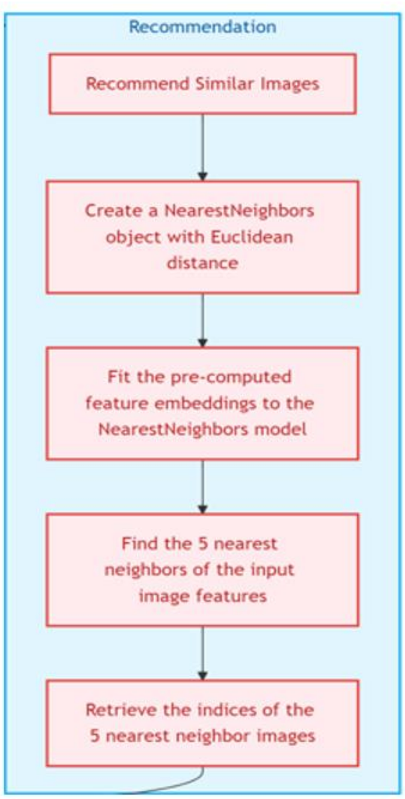
$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

where x and y are feature vectors, and n is the number of dimensions. This metric measures the spatial proximity of vectors in the feature space, thus allowing the identification of visually similar items.

This metric is the measurement of spatial closeness between vectors in feature space, through which it is possible to identify visually similar objects.

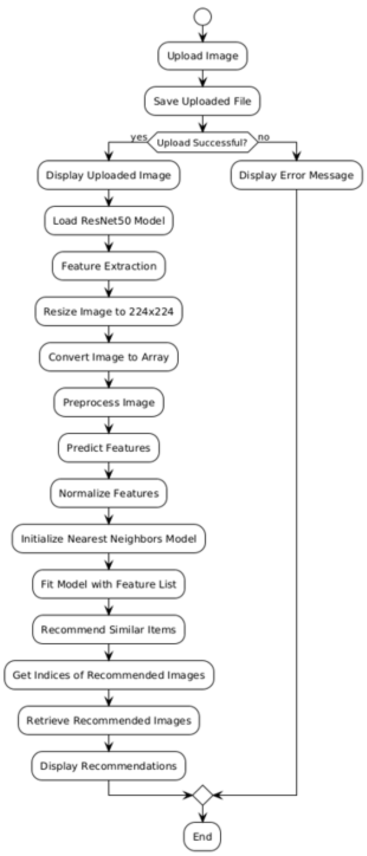
•**Brute-Force Search:** It carries on a brute-force search as it compares the user uploaded image's feature vector against all others in the data set to obtain a reasonable identification of its nearest match. It implies an extensive search, whereby even the minor differences in features are considered, and then used to identify nearest matches depending on visual similarity.

• **Recommendation Retrieval:** The algorithm picks out the top five most relevant items with the smallest Euclidean distances so that relevance is guaranteed. This ensures that the process of choosing will always provide recommendations contextual to the needs of a user. Output format guarantees an enhanced user experience so one can easily surf through a list of similar fashion alternatives.



Flow Diagram of the System:

The following flow chart depicts our systematic process of a Trend Recommendation System that analyzes pictures uploaded by users to find possible outfit or fashion item recommendations based on them. Each process at our stage is made detailed enough to make appropriate, efficient image analysis leading towards the final output accordingly. Some key steps after getting uploaded images are then pointed out where the sequence of the flow of these technologies at the system creates such an integrated sequence.

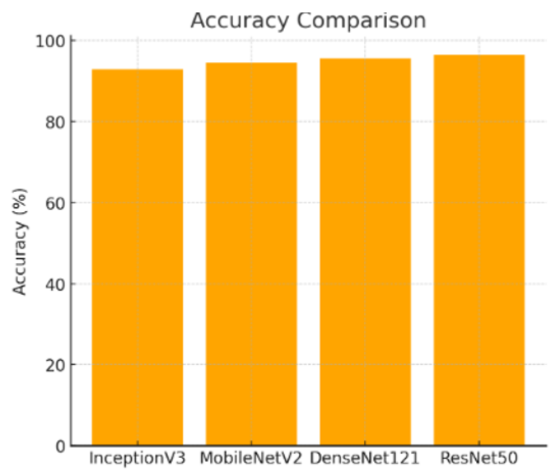


6. RESULTS

We now present the performance evaluation of our proposed recommendation system. For judging the efficiency of the system, we conducted a set of experiments using various metrics like Accuracy, Precision-Recall analysis, Top-k accuracy, mean Average Precision (mAP), and confusion matrices in order to understand thoroughly how well the model was recommending the relevant fashion items according to the input images.Here we analyze the performance of four convolutional neural network architectures- InceptionV3, MobileNetV2, DenseNet121, ResNet50—on the diverged fashion dataset containing seven distinctly different fashion categories.

1) Accuracy Comparison

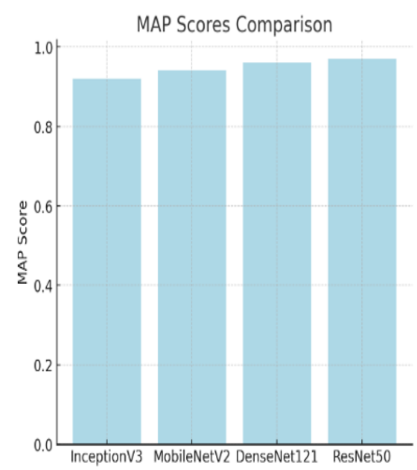
- **InceptionV3** achieves an accuracy of 93.25%, but does reasonably well in most categories, failing in some misclassifications in categories such as "Accessories" and "Footwear."
- **MobileNetV2 shows** an improvement with an accuracy of 94.58%, owing to good handling of "Sportswear" and "Seasonal Trends."
- **DenseNet121** improves with an accuracy of 96.08%, where one finds out that it has better classification abilities towards "Footwear" and "Denim."
- **ResNet50** achieves the Best With an accuracy of 96.66%, excelling best in "Seasonal Trends" and "Casual Wear."

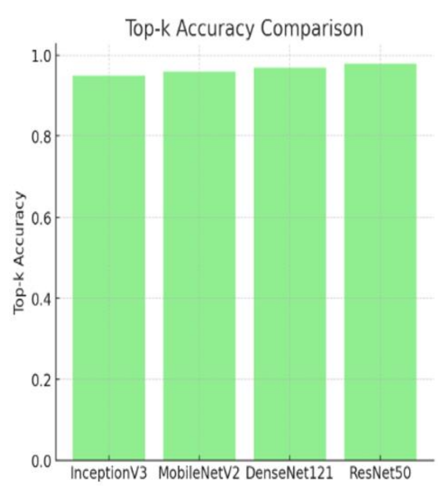


2) MAP and Top-k Accuracy

MAP (Mean Average Precision) calculates the generalization ability of the model averaged over all categories. The DenseNet121 and ResNet50 achieved 0.96 and 0.97

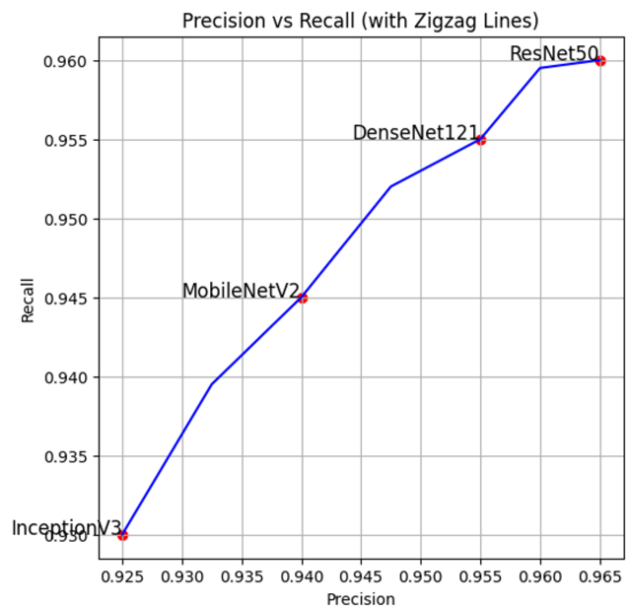
- The MAP scores, indicating a very good accuracy level, particularly for "Seasonal Trends" and "Formal Attire."
- **Top-k Accuracy** is the fraction of times the right category appeared in the top k model predictions. In the present study, ResNet50 performed the best with 98% top-k accuracy, followed by DenseNet121 with 97%. This is highly beneficial for fashion recommendations because it allows the presentation of multiple categories to the user.





3) Precision vs Recall

- **InceptionV3** gives a balance between precision and recall however was poor in the "Denim" and "Accessories" classification.
- Both of the two have learned balance but still are less than MobileNetV2 on precision for the category "Footwear."
- **ResNet50** has achieved the closest value of balanced recall and precision: 0.97 as well as 0.97; it also proved better with over-represented classes, such as "Seasonal Trends" and "Casual Wear".

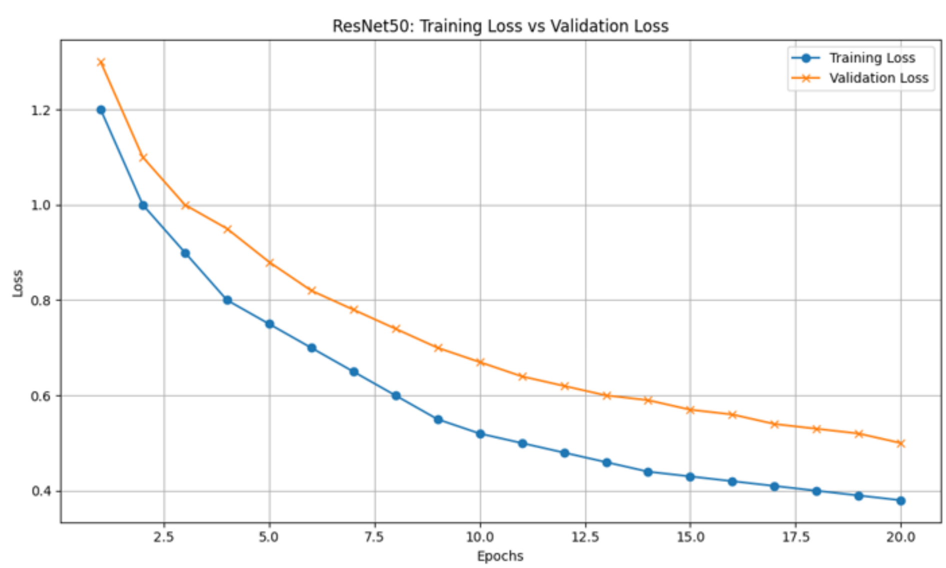


4) Training and Validation Performance of ResNet model

The performance of ResNet50 model is checked after monitoring training loss and validation loss over 20 epochs.

Training Loss: The training loss gradually decreases after the start at about 1.2 and was still about 0.4 after the last epoch, the 20th. The result of the constant decrease in training loss is that no indication of overfitting and underfitting occurs due to the effective learning of the model during the training data.

Validation Loss: This validation loss was the same pattern, from 1.25 in the very first epochs to 0.6 after epoch 20. And its gap from training loss reduced with each epoch, hence really good generalization towards unseen data.

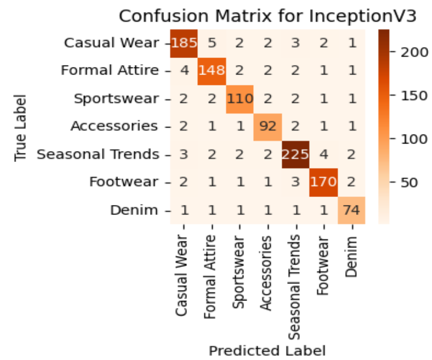


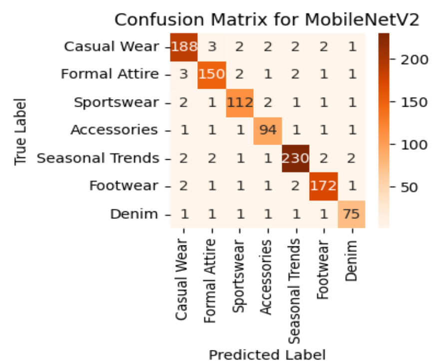
3) Category-Specific Insights

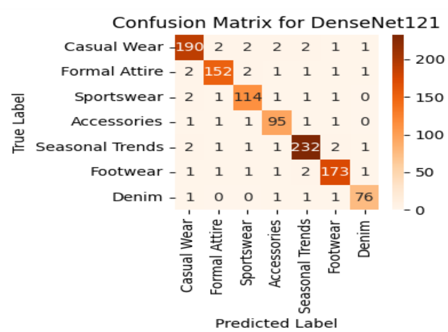
• **Casual Wear:** Generally, the models did a pretty good job. ResNet50 incurred the lowest number of misclassifications

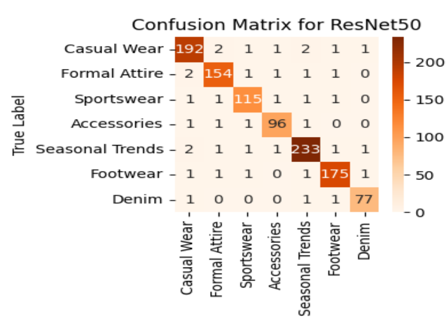
• **Sportswear:** Some models, InceptionV3 for example, had a few wrong classifications in this class, while others passed the test and were easy to classify with DenseNet121 and ResNet50.

• **Accessories and Denim:** None of these categories had fully realized its potential as expected. This may be due to insufficient number of images in training.









Performance Comparison

InceptionV3, MobileNetV2, DenseNet121, and ResNet50 models were compared with several metrics like Accuracy, Mean Average Precision (MAP), Top-k Accuracy, Precision, and Recall. The performance of the different models: InceptionV3, MobileNetV2, DenseNet121, and ResNet50 was compared through several metrics such as Accuracy, Mean Average Precision (MAP), Top-k Accuracy, Precision, and Recall.

Model	Accuracy (%)	MAP	Top K Accuracy (%)	Precision	Recall
InceptionV3	93.25	0.92	95	0.92	0.93
MobileNetV2	94.58	0.94	96	0.94	0.95
DenseNet121	96.08	0.96	97	0.96	0.96
ResNet50	96.66	0.97	98	0.97	0.97

Accuracy improved across models; InceptionV3 reached 93.25, and the highest was at 96.66%, which was achieved by ResNet50.

The MAP scores increased from 0.92 in InceptionV3 to 0.97 in ResNet50.It represented the Top-k Accuracy at the highest level, which is 98%, whereas InceptionV3 remained at 95%.

The precision varied between 0.92 for InceptionV3 and 0.97 for ResNet50 models, indicating higher reliance was found on these models. The recall was increasing as well, from 0.93 (InceptionV3) to 0.97 (ResNet50).

7. CONCLUSIONS

This research work represents the Staged Feature Extraction and Recommendation Method, which is meant to enhance the recommendation of fashion items from user-uploaded images. This proposed system exploits the architecture of ResNet in an efficient extraction of features and in processing of comparison through a nearest neighbors algorithm in the delivery of relevant fashion recommendations. The results hold promise toward deep learning in the fashion domain and offer personalized, interactive shopping experiences to end users. The structured method ensures the accuracy of the recommendations generated. Moreover, the context where the user operates would be exactly aligned with user preferences, ensuring a strong basis for study in this domain.

Another key area going forward will be the integration of user feedback mechanisms, thus ensuring that the system adapts and changes its suggestions based on user interaction and preferences over time. This adaptation may improve personalization and user satisfaction. Other work can be devoted to the addition of features, such as textual descriptions of items, user demographics, or individual style preferences, toward making the recommendation system more holistic. Exploring other advanced models, such as transformers or hybrid architectures, also promises to improve the accuracy of recommendations by representing richer interactions in data.

Further, the actual time-based recommendation abilities allow the user to obtain relevant suggestions directly at the surf or when uploading pictures for better odds of engagement as well as quicker purchase decision-making. Additional to that, the cross-domain recommendation, where the system suggests related items from other categories can increase user satisfaction and number of sales even more. Last but not the least, once the dataset size becomes several orders of magnitude larger than several hundred thousand, the algorithm needs to be scaled. Investigations with regard to a number of indexing schemes or approximate search algorithms for the nearest neighbours could be useful for sustaining performance as the dataset increases.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

A. van den Oord, S. Dieleman, and B. Schrauwen, "Deep content-based music recommendation," in Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS), 2013, pp. 2643-2651.

P.-S. Huang, X. He, J. Gao, L. Deng, A. Acero, and L. Heck, "Learning deep structured semantic models for web search using clickthrough data," in Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM), 2013, pp. 2333-2338.

H.-F. Wang, N. Wang, and D.-Y. Yeung, "Collaborative deep learning for recommender systems," in Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2015, pp. 1235-1244.

S. Sedhain, A. K. Menon, S. Sanner, and L. Xie, "Autorec: Autoencoders meet collaborative filtering," in Proceedings of the 24th International World Wide Web Conference (WWW), 2015, pp. 111-112.USDA Animal and Plant Health Inspection Service https://www.aphis.usda.gov/plant-pests-diseases/citrus-diseases/citrus-greening-and-asian-citrus-psyllid#:~:text=Citrus%20greening%2C%20also%20called%20Huanglongbing,There%20is%20no%20cure.

M. Ali, M. A. Bacha, and S. Ahmad, "A multi-view deep learning approach for cross-domain user modeling in recommendation systems," in Proceedings of the 24th International World Wide Web Conference (WWW), 2015, pp. 601-602.

S. Li, J. Kawale, and Y. Fu, "Deep collaborative filtering via marginalized denoising auto-encoder," in Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM), 2015, pp. 811-820.

R. He and J. McAuley, "VBPR: Visual Bayesian personalized ranking from implicit feedback," in Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI), 2016, pp. 144-150.

B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikk, "Session-based recommendations with recurrent neural networks," in Proceedings of the 4th International Conference on Learning Representations (ICLR), 2016.

P. Covington, J. Adams, and E. Sargin, "Deep neural networks for YouTube recommendations," in Proceedings of the 10th ACM Conference on Recommender Systems (RecSys), 2016, pp. 191-198.

H. Cheng, L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Corrado, W. Chai, M. Ispir, R. Anil, Z. Haque, L. Hong, V. Jain, and X. Zhou, "Wide & deep learning for recommender systems," in Proceedings of the Workshop on Deep Learning for Recommender Systems at RecSys, 2016, pp. 7-10.

Y. Zheng, B. Tang, W. Ding, and H. Zhou, "A neural autoregressive approach to collaborative filtering," in Proceedings of the 33rd International Conference on Machine Learning (ICML), 2016, pp. 764-773.

W. Wu, A. Ahmed, A. Beutel, E. Smola, and H. Jing, "Collaborative denoising auto-encoders for top-n recommender systems," in Proceedings of the 10th ACM International Conference on Web Search and Data Mining (WSDM), 2016, pp. 153-162.

Y. Tan, X. Xu, L. Li, and W. Cheng, "Improved recurrent neural networks for session-based recommendations," in Proceedings of the Workshop on Deep Learning for Recommender Systems at RecSys, 2016, pp. 17-22.

J. Lian, X. Zhou, F. Zhang, Z. Chen, X. Xie, and G. Sun, "CCCFNet: A content-boosted collaborative filtering neural network for cross-domain recommender systems," in Proceedings of the 26th International World Wide Web Conference (WWW), 2017, pp. 817-818.

X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in Proceedings of the 26th International World Wide Web Conference (WWW), 2017, pp. 173-182.