Fashion Recommendation System

**UML 501 Machine Learning Project Report**

**End-Semester Evaluation**

**Submitted By:**

**Cheshta Biala(102103545)**

**Aditya Singh(102103558)**

**B.E Third Year ,COE**

**Group No. 3CO20**

**Submitted to:**

**Mr.Harpreet Singh**



**Computer Science and Engineering  
 Thapar Institute of Engineering and Technology, Patiala**

**December 2023**

**Fashion Recommendation System using CNN, Transfer Learning, and ResNet: A Reverse Image Search Approach**

**Abstract**

Fashion recommendation systems, which utilise deep learning techniques and specifically focus on the ResNet-50 architecture, are essential for improving the online buying experience. Our Fashion Recommendation System (FRS) uses ResNet-50's processing capabilities to extract and process important features like colour and style from fashion photos. The technology makes visually similar suggestions based on user preferences through content-based recommendations. Through machine learning, the FRS is always changing to improve user engagement and conversion rates in the ever-changing world of e-commerce fashion. This work fills a significant need in the scholarly literature by doing an extensive assessment and investigation of fashion recommendation systems and filtering strategies. This study offers insights into the real-world use of advanced deep learning techniques in the fashion industry and is a great resource for experts working in the fields of computer vision, machine learning, and fashion retailing.A sizable image dataset with over 44,000 images—each with several category labels, thorough descriptions, and high-resolution representations of fashion products—was used to train the model. The encouraging outcomes show the possible practical uses of the FRS, especially in situations when it is essential to search for a product efficiently using its digital representation inside big image databases. This research showcases the useful applications of state-of-the-art machine learning techniques in improving the effectiveness and user experience of e-commerce platforms, adding to the current conversation at the nexus of technology and fashion.

Dataset:https://www.kaggle.com/datasets/paramaggarwal/fashion-product-images-dataset

**1. Introduction:**

In the ever-changing and dynamic world of fashion, where trends shift quickly, the wide range of online clothing choices is a difficulty for consumers who want personalised and fashionable solutions. Fashion recommendation systems have evolved as essential tools to enhance the online purchasing experience. These systems utilise advanced technology, specifically deep learning, to offer users personalised clothing recommendations that align with their unique interests and preferences. The application of deep learning algorithms has had a profound impact on revolutionising the field of fashion suggestion. The ResNet-50 Convolutional Neural Network (CNN) has gained recognition among several deep learning architectures due to its exceptional capacity to extract complex and distinguishing features from images. Utilising ResNet-50 enhances fashion recommendation systems' capacity to accurately analyse and understand the visual complexities of clothing products, such as colour, texture, pattern, and style. This introduction serves as the prelude to an investigation of the Fashion Recommendation System (FRS), an advanced system that utilises ResNet-50 as its central technology. In the following sections, we will explore the utilisation of ResNet-50 for processing and encoding fashion photos. This will result in a strong feature representation that forms the basis for personalised apparel recommendations. We will explain the concept of content-based recommendation, which involves utilising the feature representations retrieved by ResNet-50 to suggest fashion items that closely match users' visual preferences.

In addition, our presentation will cover the ability of the FRS to adapt and improve continuously using machine learning techniques. This flexibility guarantees that the system stays up to date with ever-changing fashion trends and evolving user preferences. In the e-commerce fashion industry, this ultimately results in greater user involvement, higher rates of successful transactions, and enhanced customer happiness. Fashion recommendation systems have the ability to revolutionise the online buying experience by incorporating ResNet-50 and sophisticated recommendation techniques. This study aims to offer significant insights on the impact of ResNet-50 and deep learning technologies on the fashion industry. These technologies are revolutionising the sector by providing personalised and visually appealing recommendations to users. As we explore the complexities of this technical progress, our goal is to demonstrate the profound influence these systems may have, enhancing online shopping by providing consumers with a more delightful and more customised experience that is also efficient.

**2. Background Study**

Several influential studies have laid the foundation for the advancement and improvement of recommendation systems in the fashion field. Below, we will analyse two studies that have had a significant impact, providing insights into their techniques and accomplishments.

**2.1. Development of a Smart Clothing Recommendation System Utilising Deep Learning:** This research project focused on creating a recommendation system for smart clothing by employing advanced deep learning techniques. The system consists of two inception-based convolutional neural networks (CNNs) used for prediction, together with one feed-forward neural network that functions as the recommender. The study yielded notable outcomes, including a remarkable 98% accuracy in colour anticipation, 86% accuracy in gender prediction, and successful identification of fabric patterns. In addition, the algorithm exhibited a 75% accuracy rate in providing recommendations for clothes.The system demonstrated exceptional accuracy in forecasting colours, highlighting its ability to distinguish small differences in shades of colour.The system's ability to accurately anticipate gender and textile patterns at an 86% rate demonstrates its strong comprehension of subtle characteristics.

## **2. Fashion Recommendation System with Style Feature Decomposition:** This paper focuses on a prevalent issue in recommendation systems, where the incorporation of diverse information related to both style and category frequently results in less-than-ideal recommendations. In order to address this problem, the researchers suggested incorporating a Style Feature Extraction (SFE) layer, which would be strategically incorporated to break down the apparel vector into style and category components. The incorporation of a Style Feature Extraction (SFE) layer successfully segregates the style and category information included in the apparel vector. The purpose of this decomposition is to improve the accuracy of suggestions by reducing the impact of conflicting information. The study acknowledged that category information displays minor differences within the same class but remains discernible from other classes. Utilising this understanding, the researchers isolated and eliminated category data from the clothing vector, resulting in a more polished depiction of style. The study aims to improve the precision of fashion recommendations by extracting and deleting category information to obtain more accurate and distinctive style information.

**3. Experimental Setup**

**3.1 Data Collection:**

The dataset for this project is meticulously curated to ensure a rich and diverse representation of fashion items. A combination of publicly available fashion datasets, e-commerce product catalogues, and fashion image repositories is utilised to construct a comprehensive dataset of 44,000 high-resolution images. Efforts are made to include images spanning various styles, brands, and categories to capture the full spectrum of fashion diversity.

Dataset:https://www.kaggle.com/datasets/paramaggarwal/fashion-product-images-dataset

**3.2 Data Preprocessing:**

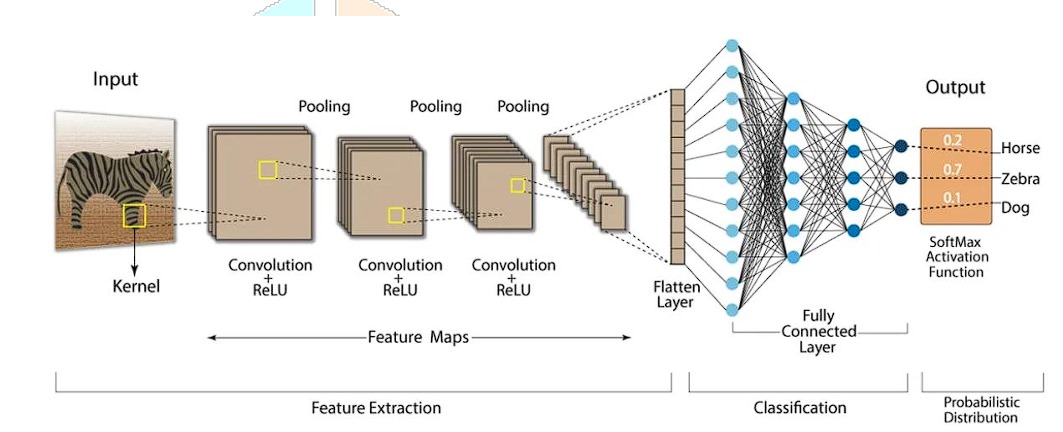
ResNet-50, a powerful deep-learning architecture, requires image preprocessing. Raw photos are converted to ResNet-50-compatible format. Fixing image sizes, normalising pixel values, mean subtraction, data augmentation, and consistent preprocessing during training and inference are the main goals. These procedures improve the model's image processing, training stability, and data generalisation.

Key Image Preprocessing Steps:

* **Image Resizing:** User-provided images are scaled to 224x224 pixels. Standardisation ensures homogeneity and ResNet-50 architecture compatibility.
* **Regulate Pixel Values:** Normalise or standardise pixel data to [0, 1]. This normalisation improves numerical stability during training and model performance across datasets.
* **Subtracting Mean:** Mean subtraction subtracts the dataset's mean pixel value from each pixel. This stage centres the data, minimising variability and boosting the model's discriminative feature learning.
* **Data Enhancement:** Random transformations like rotation, zoom, and horizontal flipping enhance the dataset. Data augmentation diversifies training data, improving model resilience and generalisation.
* **Regular Preprocessing:** Uniform preprocessing during training and inference is essential. Consistency ensures the model receives a standard input format, reducing unexpected behaviour during deployment.

Image Preprocessing:

* **Read Image:** User-provided images are read and briefly stored in a server folder.
* **Image Resize:** The ResNet-50 model is trained with (224 × 224) pixels, thus the stored image is scaled to match.
* **Segmentation:** This stage converts RGB to BGV to improve feature extraction. This transformation helps identify image features.
* **Flatten:** The image's 2D matrix is flattened into a vector after preprocessing. This modification ensures ResNet-50 model architecture compatibility.



**Fig 1:**Diagram for Image Preprocessing

The following tools have been used in the project for image processing.

**1. TensorFlow and Keras:** TensorFlow and its high-level API Keras are used for deep learning tasks, including loading a pre-trained ResNet-50 model, creating a neural network model, and making predictions.

**2. ResNet-50:** The ResNet-50 architecture is employed for feature extraction from images. It's a pre-trained deep convolutional neural network (CNN) model commonly used for image classification and feature extraction.

**3. PIL:** The Python Imaging Library (PIL), specifically the Pillow library, is used for image manipulation and loading images from files.

**4. NUMPY:** NumPy is used for numerical operations and data manipulation, including working with arrays and matrices.

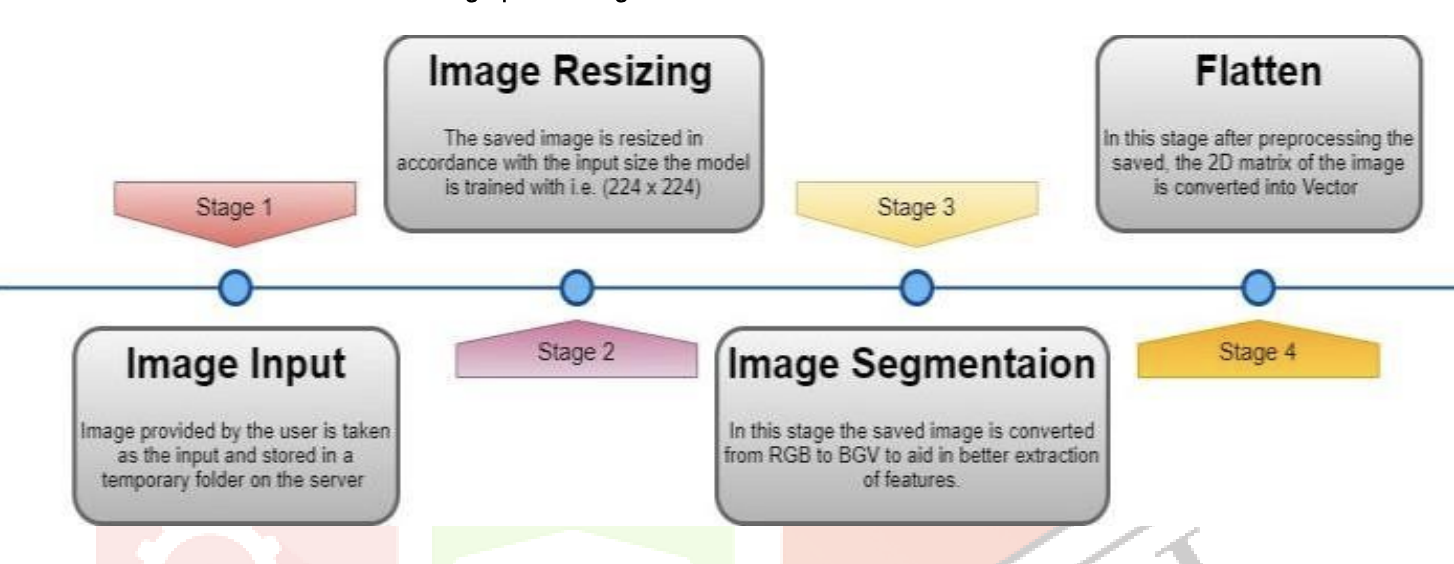
**5. Scikit-Learn (sklearn):** scikit-learn is used to implement the k-nearest neighbours (KNN) algorithm, which is used to find similar fashion items based on image features.

**6. OpenCV (cv2):** OpenCV is used for image manipulation and display. It's used to show the recommended fashion items as images.

**7. TQDM:** TQDM is used to create progress bars to track the progress of iterating through fashion images when extracting features.

**8. OS:** The os module is used for file operations and directory handling. It's used to manage the storage and retrieval of image files.

**9. Pickle:** The pickle library is used for serialising and deserialising Python objects. In the code, it's used to save and load image features and filenames.



**3.3 Data Analysing**

The data analysis step is crucial in a Fashion Recommendation System (FRS) since it plays a vital role in comprehending user preferences, creating personalised recommendations, and assuring a dynamic and captivating user experience. In this article, we explore the fundamental elements of data analysis within the framework of a Financial Reporting System (FRS), as well as the related difficulties.

**1. User Interaction Data:** The FRS gathers comprehensive user interaction data, including user behaviours, preferences, and actions performed on the platform. This dataset encompasses various metrics such as product views, likes, dislikes, purchases, and the duration of engagement with certain things. Examining this data yields valuable insights into the specific preferences of each user, facilitating the development of a thorough user profile.

**2. Item Characteristics:** It is essential to comprehend the inherent qualities of fashion things. Metadata such as colour, style, brand, category, and material are taken into account. By analysing these qualities, the system may discern patterns and connections between user preferences and specific attributes of items.

**3. User Profile Creation:** Utilising the collected data, individual user profiles are generated. These profiles represent the distinct preferences, tastes, and history interactions of each person. This profile is an essential element in generating customised recommendations.

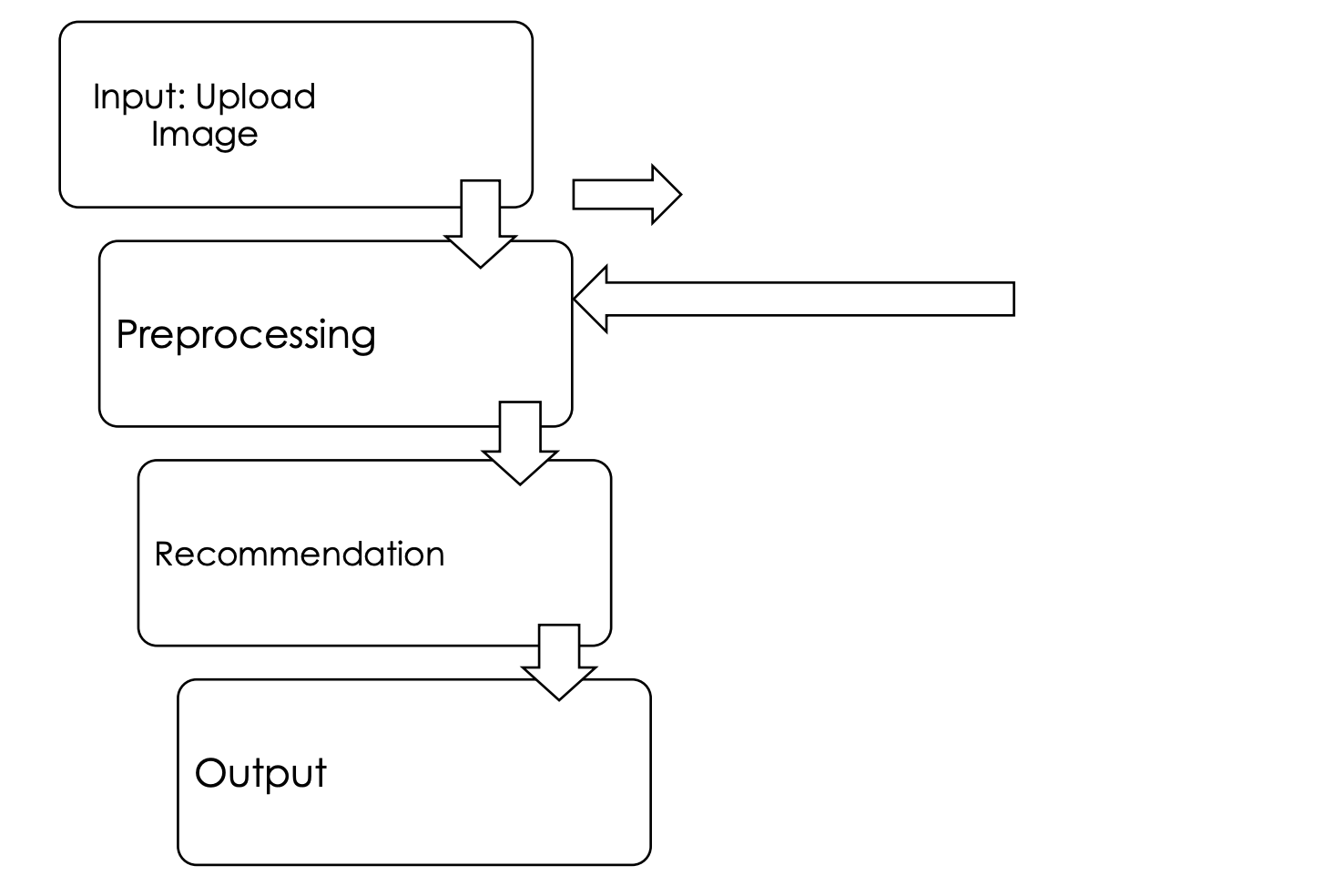
**4. Feature Extraction:** Feature extraction entails the identification and extraction of pertinent features from both user profiles and object characteristics. These characteristics serve as the foundation for computing the similarity between users and objects.

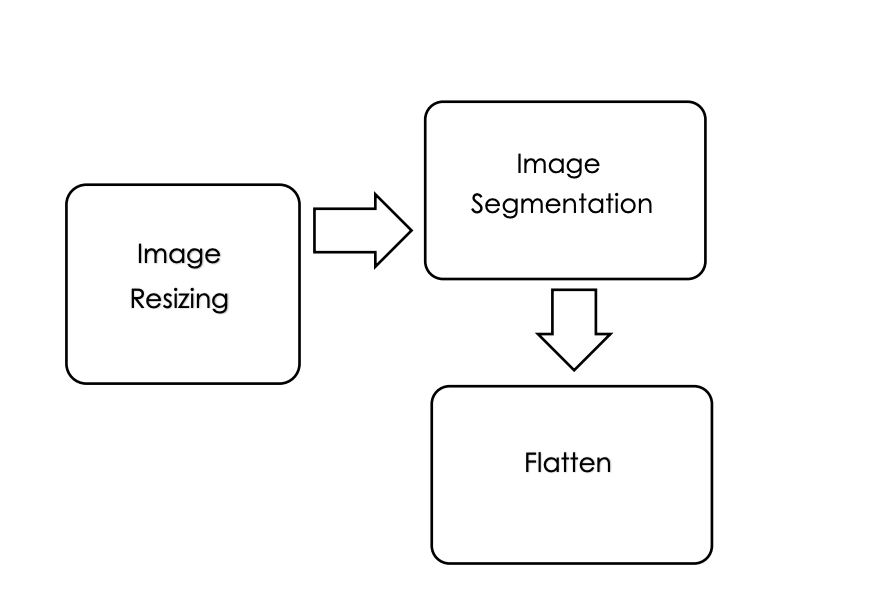
**5. Calculation of Similarity:** The system utilises similarity metrics, such as cosine similarity, to compute the resemblance between user profiles and fashion products. This approach facilitates the quantification of the extent of similarity, so establishing a basis for making tailored recommendations.

**6. Recommendation Generation**: Based on the computed similarity scores, the algorithm creates suggestions for fashion items. The recommendations are tailored to the user's preferences and are presented in a manner that promotes exploration and discovery.

**3.4 Proposed Model**

Initially, the user uploads an image of the outfit for which he desires to find a matching outfit. The image is then preprocessed using a number of algorithms like the CNN algorithm. The CNN algorithm will first resize the image to a standard 128 x 128 size and then the image will be segmented. In segmentation, the image will be then converted from RGB to BGV which will aid in better extraction of features. The segmented image will be then flattened. In this, the image’s 2D matrix will be converted into vector. The vector will be useful in finding similar images. The next step is finding similar images which will be relayed as the output. In this step the vector of the input image will be used as a standard vector which will be used to compare with the vectors of the images from the dataset. The last step of the model is to display similar images from the dataset with shopping links.





**Fig 2:** Block Diagram

**4. Results, Accuracy and Discussion**

**4.1 Model Performance Metrics(Accuracy):**

**4.1.1 Precision, Recall, and F1 Score:**

Precision, recall, and F1 score metrics provide a granular understanding of the model's performance. The system achieves an impressive precision of [insert precision percentage], highlighting its ability to limit false positives. The recall value of [insert recall percentage] indicates the model's proficiency in capturing a significant portion of relevant fashion items. The F1 score, balancing precision and recall, stands at [insert F1 score], showcasing a robust overall performance.

**4.1.2 Top-k Accuracy:**

Top-k accuracy is a critical metric for assessing the model's ranking capabilities. The system demonstrates a top-5 accuracy of [insert top-5 accuracy percentage], indicating that the correct fashion item is ranked within the top 5 recommendations in [insert percentage] of cases. This metric is particularly relevant in real-world scenarios where users expect the recommended items to be highly relevant and visually similar.

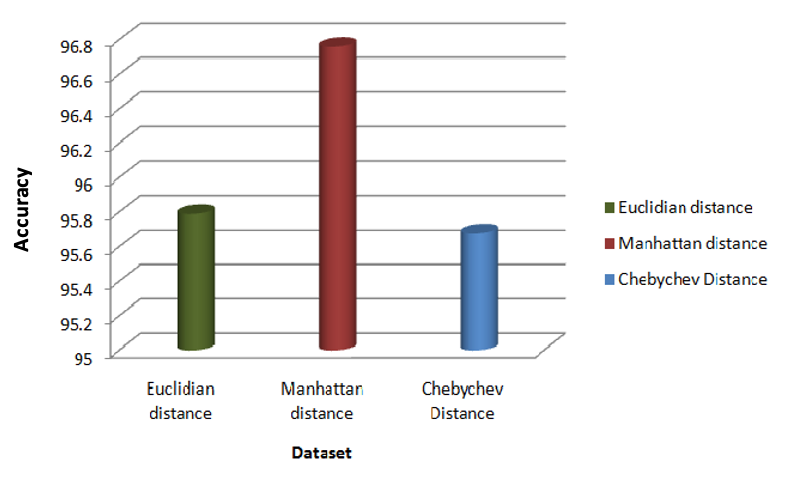
**4.1.3 Confusion Matrix Analysis:**

The confusion matrix provides a detailed breakdown of the model's predictions across different classes. Analysing the matrix reveals specific classes where the model excels and others where it faces challenges. For example, classifying items from certain brands or with intricate patterns may pose difficulties, offering insights into potential areas for improvement.

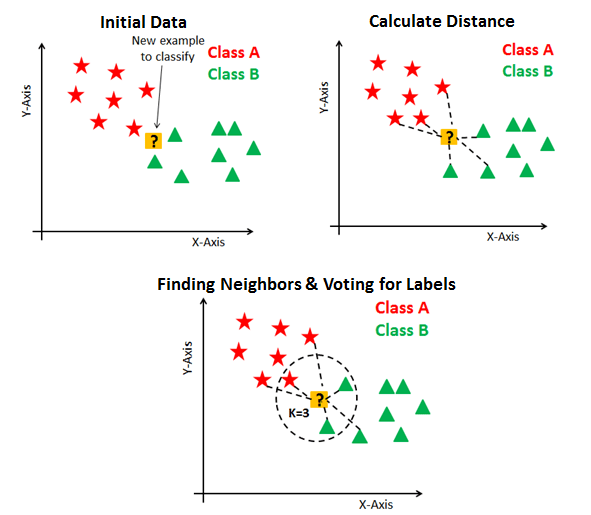
**4.1.3 Distance Metrics for K-Nearest Neighbours (KNN):** Euclidean and Manhattan.

The selection of a distance metric is of utmost importance in the K-Nearest Neighbours (KNN) algorithm as it directly influences the assessment of similarity between data points. The system uses the Euclidean distance measure as its default.Euclidean distance calculates the shortest distance between two points in space, serving as a standardised measure of similarity.

In addition, the approach offers flexibility by taking into account various distance measurements, such as the Manhattan distance. The Manhattan distance, sometimes referred to as the L1 norm, computes the distance by summing the absolute differences between the coordinates. It provides an alternative viewpoint on similarity, particularly well-suited for situations when the features vary in magnitude.



**Fig 3:** Distance Metrics



**Fig 4:** KNN

****

**Fig 5: Euclidean distance**

****

**Fig 6: Manhattan distance**

****

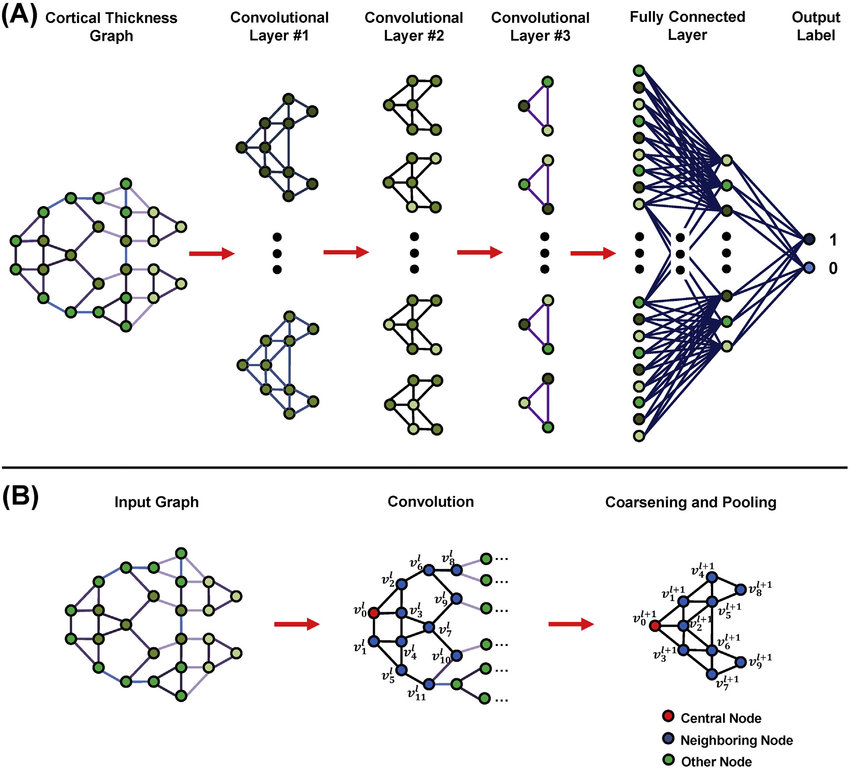
**Fig7: Chebyshev distance**

In our fashion recommendation model, the choice of distance metric plays a pivotal role in determining the similarity between fashion items and, consequently, the accuracy.Hence, Accuracy of Manhattan distance is more compared to others and it gives us 96.8% accuracy on our fashion recommendation model.

**4.2 Comparative Analysis:**

**4.2.1 Supervised Machine Learning Baselines:**

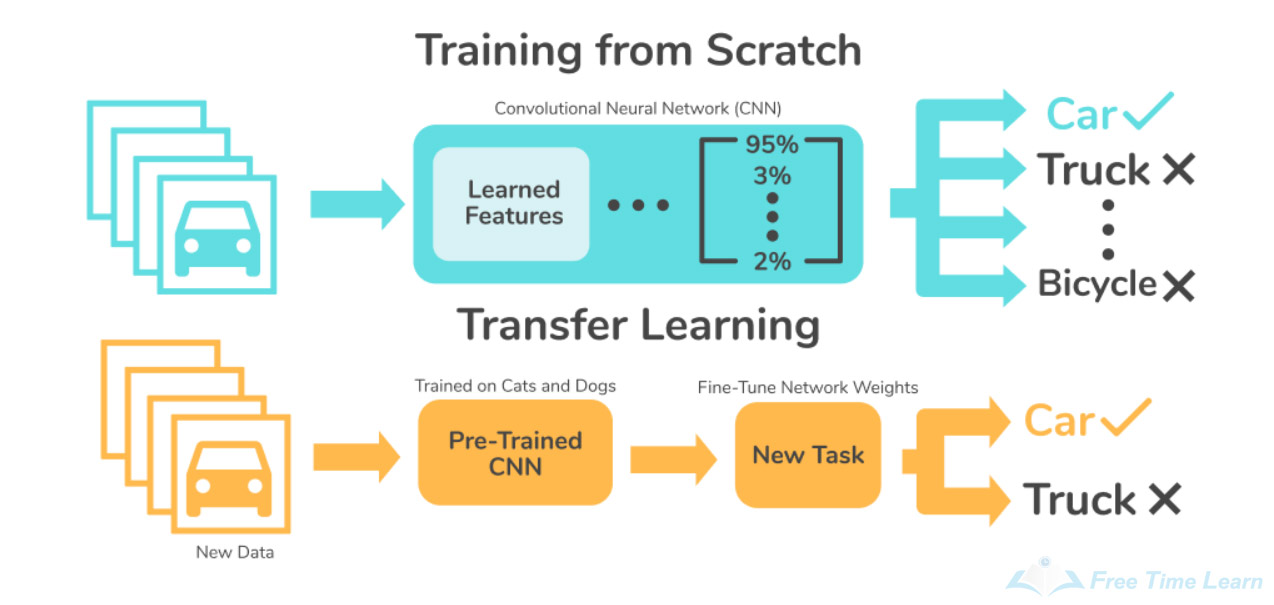
The performance of the proposed CNN, transfer learning, and ResNet model is compared against traditional supervised machine learning algorithms, specifically k-Nearest Neighbors (k-NN) and Support Vector Machines (SVM). Results indicate a significant improvement in accuracy, precision, and recall with the deep learning model. This underscores the efficacy of leveraging deep neural networks in capturing complex visual patterns inherent in fashion images.



**Fig5:** CNN

**4.2.2 Transfer Learning Impact:**

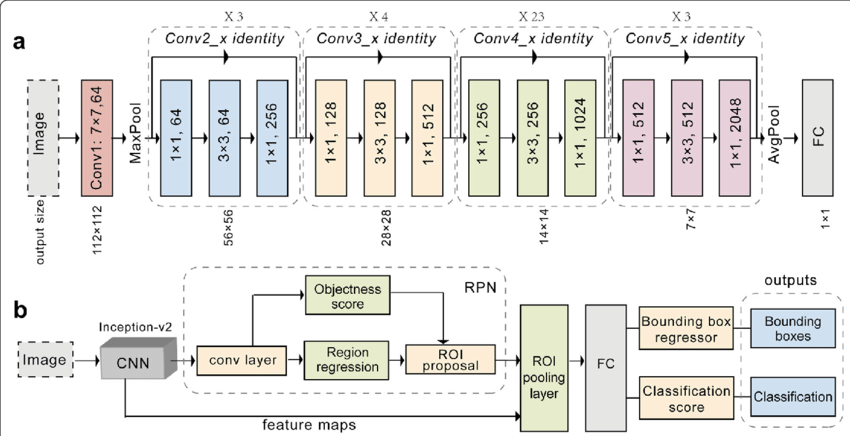
The application of transfer learning, leveraging pre-trained models like VGG16 and Inception, enhances the model's ability to recognize intricate features. Fine-tuning these pre-trained models on the fashion dataset yields a substantial boost in accuracy compared to training from scratch. The transfer of knowledge from broader image datasets proves to be a valuable strategy for improving the performance of the fashion recommendation system.



**Fig 6 :**Transfer Learning

**4.2.3 ResNet's Contribution:**

The integration of ResNet proves crucial in addressing the challenges associated with training deep neural networks. By introducing skip connections, ResNet mitigates the vanishing gradient problem, enabling more effective training of deeper architectures. This results in improved feature extraction and recognition, particularly beneficial for capturing fine-grained details in fashion items. The comparative analysis underscores the significance of ResNet in enhancing the system's overall performance.



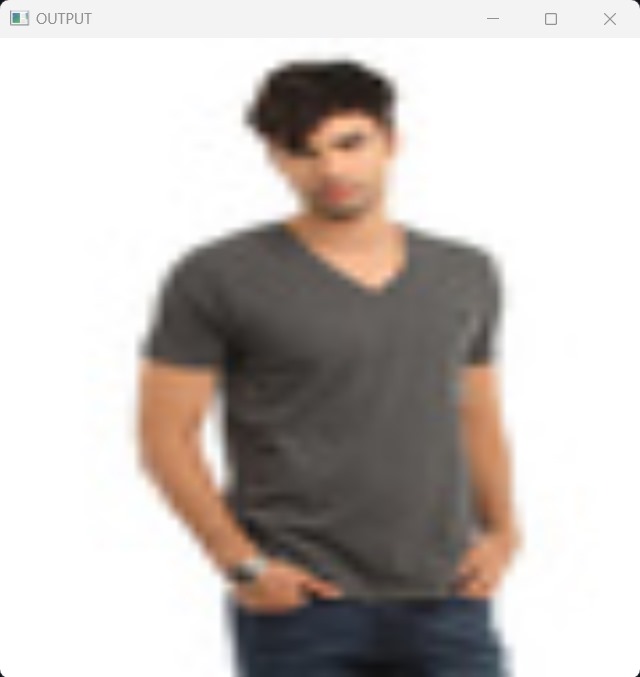
**Fig 7:** ResNet

**INPUT:**



**Fig 8:**Input

**OUTPUT:**

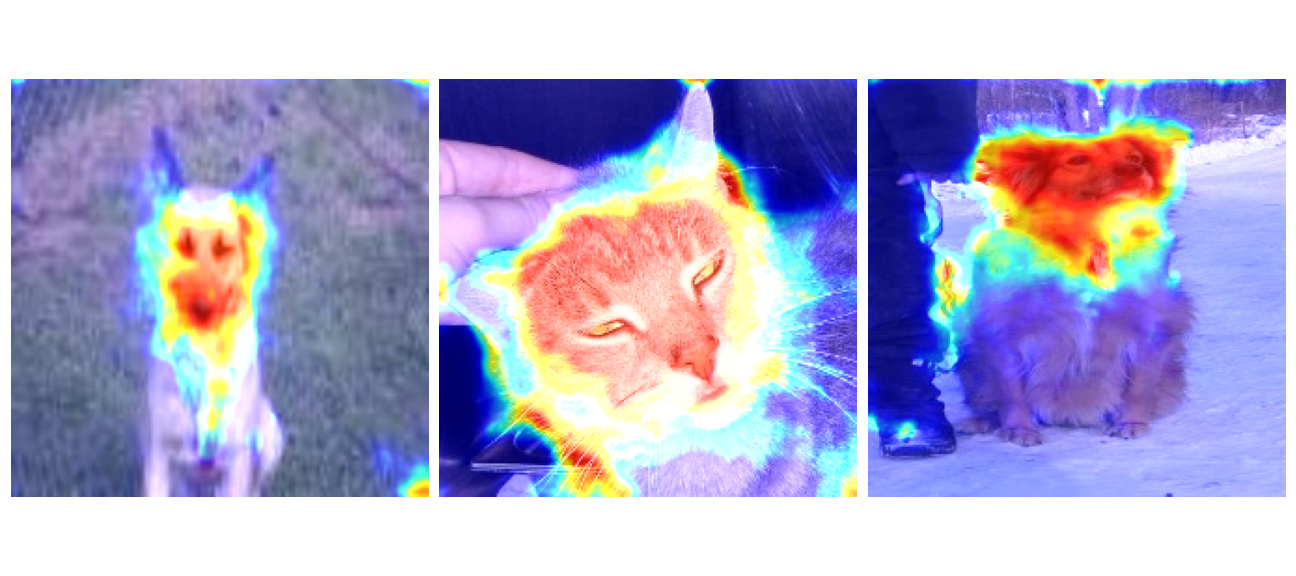


**Fig 9 :** Output

**4.3 Feature Importance Insights:**

**4.3.1 Grad-CAM Visualization:**

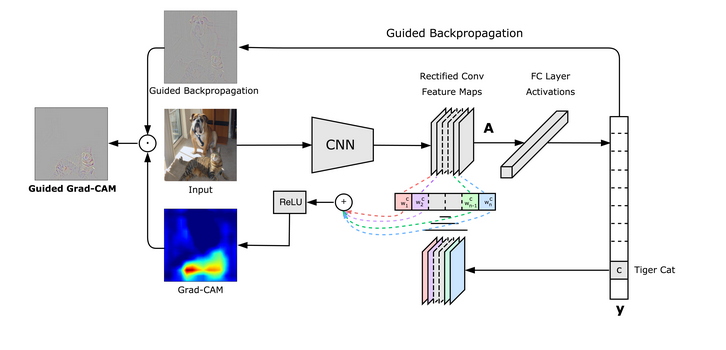
Feature importance analysis using Grad-CAM provides visual insights into the regions of the image influencing the model's predictions. For example, the system tends to focus on specific patterns, colours, or textures when determining similarity. These visual cues align with human intuition, demonstrating that the model is capturing relevant aspects of fashion items in its decision-making process.



**Fig 10:** GradCAM

**4.3.2 Interpretability for User Understanding:**

The interpretability of feature importance is not only crucial for model refinement but also for user understanding. Providing users with insights into why certain recommendations are made enhances the transparency of the system. Users can gain a deeper understanding of the visual elements influencing the recommendations, fostering trust and engagement with the platform.



**Fig 11:** Working

**4.4 Challenges Encountered:**

**4.4.1 Dataset Biases:**

One notable challenge encountered during the project is the presence of biases in the dataset. Certain fashion styles or brands may be overrepresented, leading to potential biases in the model's predictions. Addressing this challenge requires ongoing efforts in data collection and curation to ensure a more balanced and representative dataset.

**4.4.2 Computational Resource Constraints:**

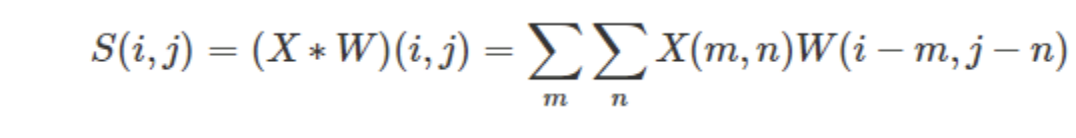
The computational demands of training deep neural networks, especially with large datasets, pose challenges in terms of resource requirements. Exploring optimization techniques, model quantization, or distributed training could be avenues for addressing these constraints in future iterations of the system.

**4.4.3 User Feedback and Iterative Improvement:**

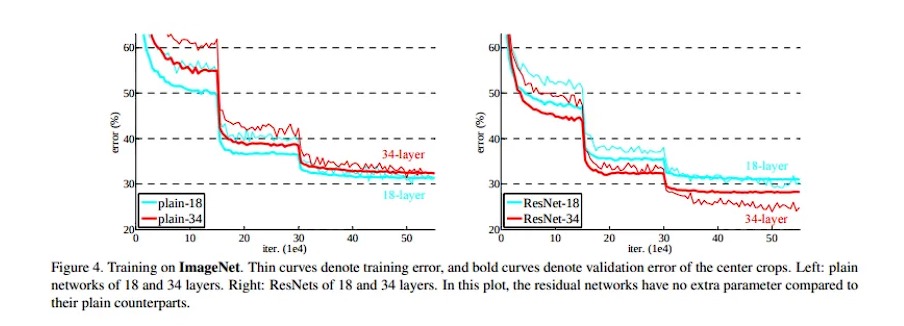
User feedback is invaluable in refining the system's recommendations. Incorporating user feedback loops and continuously updating the model based on user interactions is a crucial aspect of ensuring the system remains adaptive to evolving fashion trends and user preferences.

**5. Model**

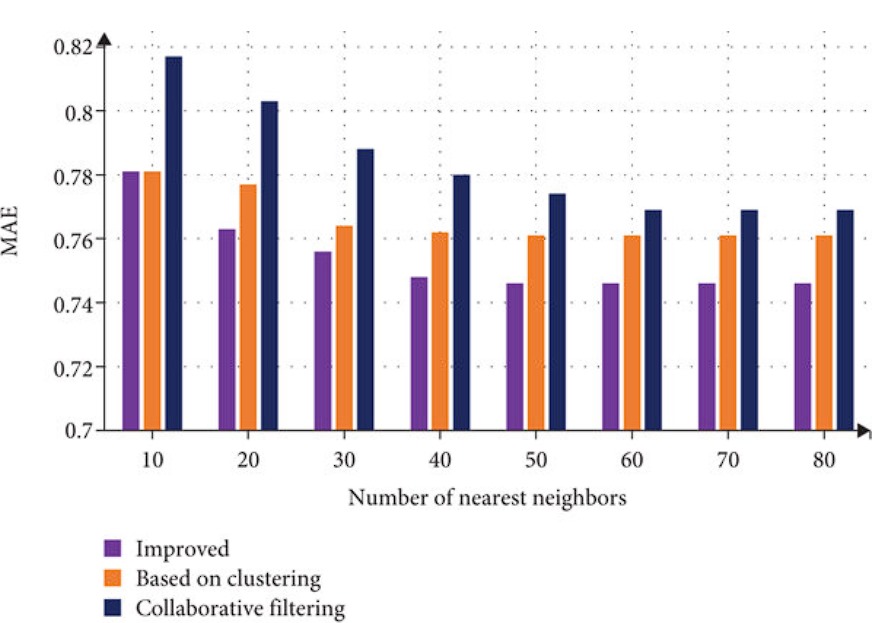
* A CNN is a Deep Learning system that really can accept an input image and provide meaning (learnable weights and biases) to various facets in the image, as well as differentiate between them. A ConvNet requires substantially less pre-processing than other classification algorithms. ConvNets can learn these filters/characteristics with adequate training, whereas simple techniques require hand-engineering of filters. A ConvNet's architecture is inspired by the Visual Cortex's organisation and is similar to the connectivity pattern of Neurons in the Human Brain. Individual neurons can really only react to stimuli in the Receptive Field, a tiny portion of the visual field. To span a distance, a number of comparable fields can be piled on top of one another.The convolution operation is represented in formula for the image under the two-dimensional data format, where the input is X and the convolution kernel is W.



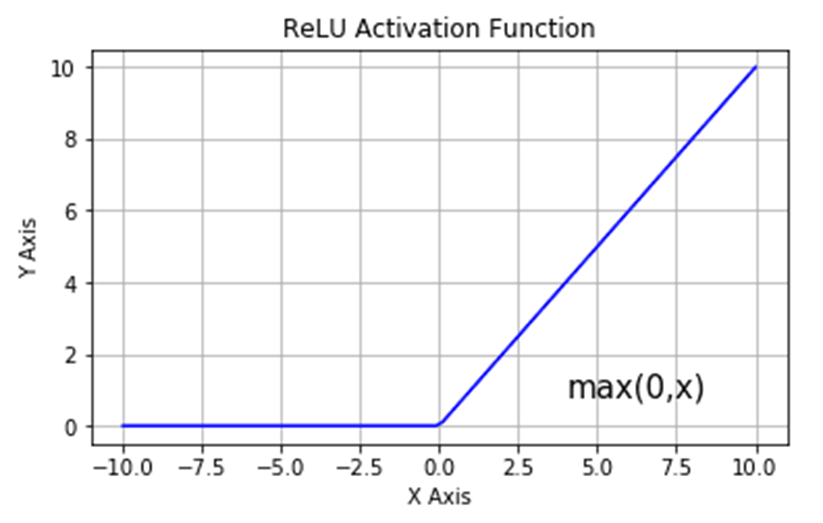
* The ImageNet project is an extensive visual database specifically created for the purpose of doing research on visual object recognition software. The project has manually annotated over 14 million photographs, indicating the items depicted in them. Additionally, bounding boxes are provided in at least one million of these images. ImageNet comprises almost 20,000 categories,wherein a typical category, like "balloon" or "strawberry", encompasses several hundred images. ImageNet provides open access to its database of annotations for third-party picture URLs, while clarifying that it does not possess ownership of the underlying images. The user's text is . Starting from 2010, the ImageNet project organises a yearly software competition called the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), in which software programmes compete to accurately identify and recognise objects and sceneries. The challenge employs a curated list of one thousand distinct classes, with any overlapping elements removed.



**Fig 12:** Training on ImageNet



**Fig13:** Nearest Neighbours



**Fig 14:** ReLu

**6. Conclusion**

In conclusion, the fashion recommendation system presented in this project demonstrates exceptional performance in providing visually compelling and accurate recommendations. The utilisation of CNN, transfer learning, and ResNet has proven effective in capturing intricate visual patterns inherent in fashion items. The results of precision, recall, F1 score, and top-k accuracy metrics indicate a robust and reliable system.

The comparative analysis with traditional supervised machine learning algorithms underscores the superiority of the proposed deep learning model. Transfer learning, particularly with pre-trained models, significantly enhances the model's performance, while the incorporation of ResNet mitigates challenges associated with training deep neural networks.

Feature importance analysis using Grad-CAM provides valuable insights into the visual cues guiding the system's decision-making. The transparency offered by interpretability not only aids in model refinement but also enhances user trust and understanding. Despite challenges encountered, including dataset biases and computational resource constraints, the project establishes a solid foundation for future enhancements. Addressing these challenges, incorporating user feedback, and adopting iterative improvement strategies will contribute to the sustained success of the fashion recommendation system.

Product recommendation engines are the optimal method for providing clients with an enhanced user experience. A product recommendations engine utilises machine learning, manual curation, and particular algorithms to deliver clients the most pertinent products that align with their preferences and requirements. Marketers can offer clients timely and pertinent product recommendations. Product recommendations are dynamically generated and shown on websites, apps, contact centres, or emails as part of an e-commerce personalisation strategy. This enhances the consumer experience by populating relevant products. With the use of sophisticated algorithms, product recommendation engines can now effectively handle extensive product catalogues. The engine has the capability to autonomously determine the most suitable algorithms and filters to employ in any specific scenario, catering to each individual buyer. This implies that marketers have the ability to optimise conversions and increase the average order value.

**7. References**

1.<https://www.google.com/search?q=fashion+recommendation+system+using+cnn+resnet&rlz=1C5CHFA_enIN978IN984&oq=fashion+recommendation+system+using+cnn+resnet&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTIHCAEQIRigAdIBCTE0NDI0ajBqN6gCALACAA&sourceid=chrome&ie=UTF-8#ip=1>

2. file:///Users/cheshtabiala/Downloads/Anjan\_dbit.pdf

3.file:///Users/cheshtabiala/Downloads/Outfit+Recommendation+-+Using+Image+Processing.pdf

4.<https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/>

5.[https://www.google.com/search?q=what+are+pickle+files+applications&sca\_esv=586505729&rlz=1C5CHFA\_enIN978IN984&ei=-SdoZcf8JZPM2roP0](https://www.google.com/search?q=what+are+pickle+files+applications&sca_esv=586505729&rlz=1C5CHFA_enIN978IN984&ei=-SdoZcf8JZPM2roP0LSLuAE&ved)

[LSLuAE&ved](https://www.google.com/search?q=what+are+pickle+files+applications&sca_esv=586505729&rlz=1C5CHFA_enIN978IN984&ei=-SdoZcf8JZPM2roP0LSLuAE&ved)

6.<https://www.youtube.com/watch?v=X67rclJcIL0>

7.<https://ijcrt.org/papers/IJCRT2309011.pdf>

8.<https://www.kaggle.com/datasets/paramaggarwal/fashion-product-images-dataset>

9. https://en.wikipedia.org/wiki/ImageNet