**Enhanced CBIR System for**

**Commercial Applications**

# Project ID:31139

*B.Tech. Project Report submitted for fulfillment of the requirements for the*

*Degree of Bachelor of Technology Under BijuPattnaik University of Technology*

*Submitted By*

# V.Litesh Kumar Silla Suhani Param Jyoti Sahu Nikhil Patnaik



*2024 – 2025*

*Under the guidance of*

**Prof. Charulata Palai**

**Roll No. 202110124**

**Roll No. 202113343**

**Roll No. 202140150**

**Roll No. 202110154**

**NIST INSTITUTE OF SCIENCE & TECHNOLOGY (Autonomous)**

**Palur Hills, Berhampur, Odisha – 761008, India**

**ACKNOWLEDGEMENT**

We would like to take this opportunity to thank all those individuals whose invaluable contribution in a direct or indirect manner has gone into the making of this project a tremendous learning experience for me.

It is my proud privilege to epitomize my deepest sense of gratitude and indebtedness to my faculty advisor, **Prof. Charulata Palai**, for her valuable guidance, keen and sustained interest, intuitive ideas, and persistent endeavor. Her guidance and inspirations enabled me to complete my report work successfully.

We give our sincere thanks to **Dr. Kunj Bihari Swain**, Project Co-Ordinator, for giving us the opportunity and motivating us to complete the project within the stipulated period of time and providing a helping environment.

We acknowledge with immense pleasure the sustained interest, encouraging attitude, and constant inspiration rendered by **Prof. (Dr.) Sukant K. Mohapatra (Chairman), Prof. (Dr.) P. Rajesh Kumar (Principal), Prof. (Dr.) Brojo Kishore Mishra (HOD**), School of Computer Science) N.I.S.T. Their continued drive for better quality in everything that happens at N.I.S.T. and selfless inspiration has always helped us to move ahead.

V.Litesh Kumar (202110124)

Silla Suhani (202113343) Param Jyoti Sahu (202140150) Nikhil Patnaik (202110154)

ACKNOWLEDGEMENT 2

**Table of Contents**

TABLE OF CONTENTS… 3

ABSTRACT 4

1. INTRODUCTION 5
2. PROBLEM STATEMENT 8
3. OBJECTIVE 9
4. LITERATURE SURVEY 12
5. METHODOLOGY 16

5.1-Flowchart 23

5.2- Algorithm 23

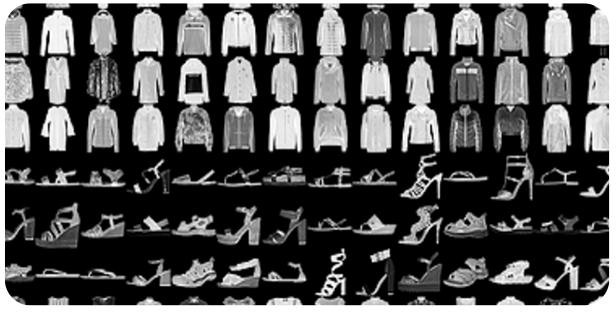
1. SIMULATION 25
2. FUTURE ADD ONS… 28
3. CONCLUSION 33
4. REFERENCES… 34

**ABSTRACT**

The exponential growth of e-commerce has driven the need for sophisticated tools to improve product discovery and enhance user engagement. Traditional text-based search methods often fall short when users can only provide visual cues, highlighting the importance of Content-Based Image Retrieval (CBIR) systems. This project introduces a CBIR framework tailored for fashion products, leveraging ResNet-50, a deep learning architecture for extracting detailed visual features, to enable image- based search functionality. The system is built and tested on the Fashion-MNIST dataset, a well-known collection of grayscale images representing ten categories of fashion items, such as T-shirts, dresses, and bags. To better suit real-world application, the images are resized and processed to extract features that capture the unique visual characteristics of each category. Retrieval accuracy is evaluated through cosine similarity metrics, ensuring high relevance and precision in search results. Extensive testing across diverse fashion items demonstrates the model’s effectiveness, positioning CBIR as an ideal solution for bridging the gap between user intent and accurate product recommendations. By advancing visual search capabilities, this CBIR system promises to transform user experience on e- commerce platforms, fostering a more intuitive and visually-driven approach to product discovery that enhances both customer satisfaction and operational efficiency.

1. **INTRODUCTION**

In recent years, the advancement of image-based search technologies has become a key area of interest in artificial intelligence (AI) and deep learning, especially within the e-commerce sector. Content-Based Image Retrieval (CBIR) systems provide a novel solution to traditional text-based search limitations by enabling users to find visually similar products based solely on image input. This approach is particularly beneficial in fashion and retail, where product discovery often relies on visual cues rather than textual descriptions. The primary objective of a CBIR system is to match a user-provided image with relevant items from a product catalog, enhancing the search experience by bypassing the need for detailed textual information. Although intuitive for users, CBIR presents a complex challenge for machines, requiring robust feature extraction to capture the nuances of various product categories. This project leverages ResNet-50, a powerful deep learning architecture, to extract and analyze these visual features, enabling accurate and efficient retrieval in a structured e-commerce environment.



*Significance of Content-Based Image Retrieval(CBIR)*

The significance of CBIR systems in e-commerce is profound, as they address key challenges in product search and discovery, especially in visually-driven domains like fashion retail. By allowing users to search based on image similarity rather than textual metadata, CBIR provides a more intuitive method for discovering products,

especially when descriptive details may be absent or inadequate. This visual approach improves search accuracy, allowing customers to find similar products efficiently, while reducing dependency on manual tagging, which can often be inconsistent. Enhanced by deep learning models like ResNet-50, CBIR systems are capable of recognizing detailed visual patterns and features within images, enabling them to capture subtle nuances between product types. For businesses, this leads to higher engagement and conversion rates, as customers are more likely to find relevant items quickly, thereby improving overall user experience and satisfaction. Furthermore, CBIR accelerates search processes, making it ideal for large-scale e- commerce platforms where rapid and accurate retrieval is essential. Through these advantages, CBIR holds the potential to transform online shopping, bridging the gap between user intent and product discovery, fostering business growth, and revolutionizing the way users interact with e-commerce platforms.

*Deep Learning Approach*

Recent advances in deep learning have enabled the creation of sophisticated models for image-based search and retrieval, such as Content-Based Image Retrieval (CBIR). Convolutional Neural Networks (CNNs) are central to this approach, as they excel at extracting detailed visual features that are essential for identifying similarities between images. In this project, we use ResNet-50, a deep CNN architecture that is highly effective for extracting robust features from images. ResNet-50’s design incorporates residual learning, which allows the model to maintain performance and stability even with very deep layers, avoiding the vanishing gradient issue common in other architectures. Pretrained on the ImageNet dataset, ResNet-50 is equipped to recognize a wide range of patterns and textures, which enhances its suitability for diverse fashion items in the CBIR system.

For measuring similarity between product images, feature vectors extracted by ResNet-50 are compared using cosine similarity. This method efficiently quantifies the resemblance between images, enabling the system to provide highly relevant search results. By leveraging deep feature representations, this CBIR system minimizes reliance on textual metadata and manual tagging, making it highly adaptable to visual search needs. The deep learning approach used here exemplifies how CNNs, such as ResNet-50, can transform product search by offering users a precise, intuitive way to explore visually similar items in e-commerce, thereby enhancing both search accuracy and user experience.

1. **PROBLEM STATEMENT**

Developing an effective Content-Based Image Retrieval (CBIR) system for e- commerce presents a unique set of challenges that span both technical and operational aspects. Traditional search methods rely heavily on text-based queries and metadata, which can be limited or inconsistent. CBIR, however, aims to allow users to search through visual inputs, directly leveraging image content to retrieve similar products. This visual approach requires a model that can extract and compare intricate visual features across diverse product categories. Key objectives include improving search accuracy by minimizing dependency on textual data, enhancing efficiency by speeding up search response times, and ultimately supporting business growth by increasing customer satisfaction and conversion rates. Each of these objectives demands a system that accurately matches user-provided images with similar items in the catalog, overcoming the limitations of metadata and text-based search.

A major challenge in implementing CBIR for e-commerce lies in the need to interpret a wide variety of visual elements with precision. Unlike object detection or classification, CBIR requires an advanced, context-sensitive understanding of image features that account for differences in texture, color, and shape. For example, two shirts might share similar colors but differ in texture or style, which the system must recognize to provide accurate search results. Additionally, speed is crucial in large- scale e-commerce platforms, where users expect instant, relevant results. To meet these demands, the CBIR model must be capable of efficient, high-speed processing without compromising accuracy, making it essential to deploy deep learning architectures like ResNet-50 that can handle complex visual data with high precision. These challenges highlight the need for a robust CBIR solution that aligns with the evolving requirements of e-commerce, facilitating an improved, visually- driven search experience.

1. **OBJECTIVE**

The primary objective of this project is to develop an efficient Content-Based Image Retrieval (CBIR) system for e-commerce, which enhances search accuracy, speed, and user satisfaction by leveraging advanced deep learning techniques for visual feature extraction. This system is designed to allow users to find visually similar products by uploading an image, bypassing limitations associated with traditional text-based search. The following objectives guide the development of an effective, scalable, and user-centric CBIR system:

# Develop a Robust Deep Learning Model for Visual Search

The first goal is to construct a deep learning model that can accurately analyze visual data and extract detailed features essential for recognizing similarities between products. Using ResNet-50, a convolutional neural network (CNN) with proven effectiveness in visual feature extraction, the objective is to capture the unique characteristics of each product, such as texture, color, and shape, ensuring precise image-based retrieval.

# Enhance Search Accuracy and Contextual Relevance

A critical objective is to ensure high search accuracy by accurately matching user- provided images with relevant products in the catalog. This involves leveraging deep learning features that reduce dependency on metadata, which is often inconsistent. By focusing on visual similarity, the CBIR system aims to return results that closely resemble the user’s input, improving contextual relevance and search reliability.

# Optimize the Model for Efficiency and Real-World Scalability

Given the demands of real-time e-commerce environments, optimizing the model for speed and scalability is essential. Techniques such as dimensionality reduction

and efficient indexing of feature vectors may be employed to enhance performance without compromising accuracy. This objective ensures that the CBIR system is suitable for deployment on large-scale platforms, where quick response times are crucial for user engagement.

# Improve User Experience and Drive Business Growth

Another objective is to enhance user experience by providing a seamless and intuitive search process that meets users’ visual discovery needs. By offering precise and relevant product suggestions, the system can increase customer satisfaction and potentially boost conversion rates, aligning with business goals of improved product visibility and customer retention.

# Implement and Evaluate with Rigorous Performance Metrics

To assess the effectiveness of the CBIR model, this project aims to implement comprehensive evaluation metrics, such as cosine similarity scores, to quantify the quality of search results. This objective ensures that the model’s accuracy, relevance, and efficiency are thoroughly validated, providing a benchmark to compare with traditional search systems and identify areas for improvement.

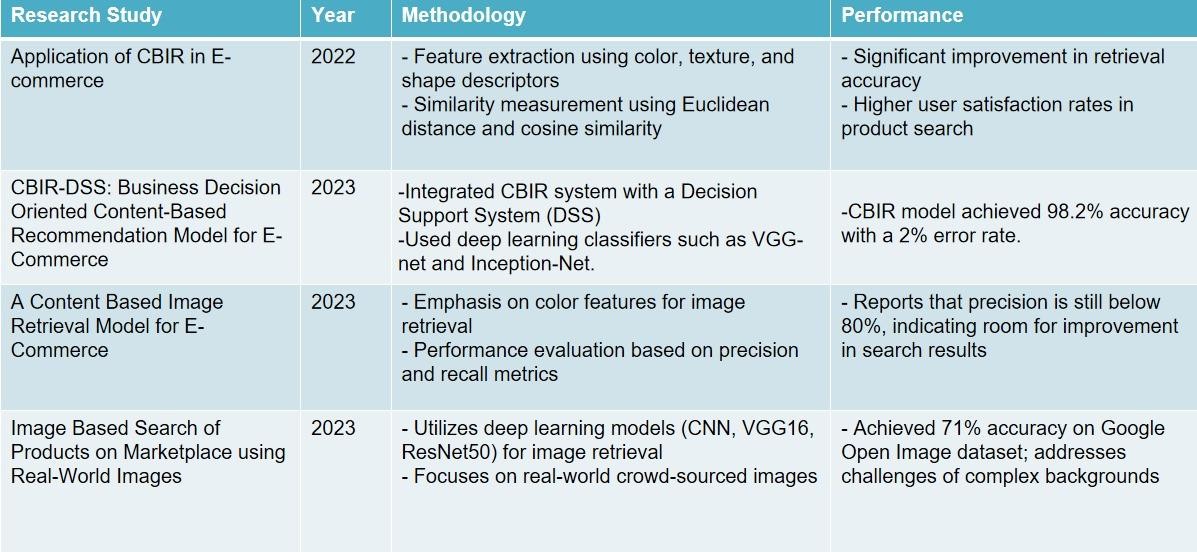
# Explore Practical Applications and Future Extensions

The final objective is to explore practical applications for the CBIR system and potential enhancements. Applications may include fashion recommendation engines, user personalization in e-commerce, and interactive search tools for mobile platforms. Future extensions could focus on expanding the model's applicability across other product categories and further improving visual feature extraction for enhanced retrieval accuracy. This objective highlights the CBIR system's adaptability and potential impact in various domains.

1. **LITERATURE SURVEY**

A literature survey is a foundational step in developing an effective Content-Based Image Retrieval (CBIR) system, especially when aiming to enhance search capabilities in the e-commerce sector. Traditional search methods, which rely on metadata and textual descriptions, often suffer from limitations in accuracy due to inconsistent tagging and insufficient detail. Recent advancements in deep learning, particularly the use of Convolutional Neural Networks (CNNs), have revolutionized CBIR by providing a means to analyze and categorize images based on their visual features rather than textual metadata. Among CNN architectures, ResNet-50 has proven effective for its ability to extract detailed and high-level visual features, helping to overcome challenges such as intraclass variations, occlusion, and lighting inconsistencies. Studies in computer vision suggest that high-performing models for feature extraction can also enhance CBIR systems, as visual similarity and matching rely on accurate image classification. CBIR systems go beyond traditional image classification, focusing on retrieving relevant images from a large database based on feature similarity, which adds complexity to the task. The literature reveals that combining CNNs with similarity metrics, like cosine similarity, can lead to highly relevant search results, thus bridging the gap between visual input and retrieval accuracy. Research further highlights the importance of optimizing CBIR models for both efficiency and scalability. Techniques such as dimensionality reduction and vector indexing contribute to managing large datasets and ensuring quick retrieval times. The application of these optimized CBIR systems in e-commerce platforms has shown that by accurately matching products based on visual similarity, companies can increase customer satisfaction and potentially boost conversion rates. As CBIR evolves, combining CNNs with traditional image retrieval techniques and advanced optimization methods appears to be the path forward, making CBIR not

only a viable alternative to text-based search but a significant improvement in terms of precision, user engagement, and adaptability across diverse product categories.



The evolution of **Content-Based Image Retrieval (CBIR)** has been driven by the increasing need for more efficient and accurate image search techniques, particularly in domains such as **e-commerce, healthcare, and surveillance**. Traditional CBIR systems relied on **handcrafted feature extraction methods**, utilizing techniques like **color histograms, edge detection, texture descriptors (GLCM, LBP), and shape analysis** to identify similarities between images. While these methods were computationally lightweight and interpretable, they struggled with capturing deep

semantic relationships within images. This limitation became more evident when dealing with complex datasets where variations in lighting, pose, and background caused significant retrieval errors.

The introduction of **deep learning-based CBIR** has revolutionized image retrieval by allowing models to learn **hierarchical feature representations** automatically, significantly improving retrieval accuracy. Deep **Convolutional Neural Networks (CNNs)** have replaced traditional methods by extracting **robust, high-dimensional feature vectors** that capture intricate details such as **textures, patterns, and structural elements**. Unlike handcrafted methods, CNNs can automatically learn the most discriminative features, making them highly effective in distinguishing between visually similar images.

Among CNN architectures, **ResNet-50** has emerged as one of the most widely used models for CBIR due to its ability to handle deep feature extraction efficiently. The **residual learning framework** of ResNet-50 addresses vanishing gradient issues, allowing the network to learn deeper representations without degradation in performance. By extracting **feature vectors** from the last convolutional layer, the system can measure similarity between images using metrics such as **cosine similarity or Euclidean distance**. This approach enables highly accurate and scalable image retrieval. However, despite its effectiveness, **ResNet-50 has some limitations**, including high computational costs and potential overfitting when applied to large-scale datasets.

To overcome these challenges, researchers have developed **more advanced deep learning architectures** that improve upon ResNet-50. One such model is **EfficientNet**, which introduces a novel **compound scaling technique** that balances network depth, width, and resolution. Compared to ResNet-50, EfficientNet achieves **higher accuracy with fewer parameters**, making it ideal for large-scale CBIR applications. Another breakthrough in deep learning for CBIR is **Vision**

**Transformers (ViT)**, which processes images as **a sequence of patches** instead of using convolutional operations. Unlike CNNs, ViTs are capable of capturing **long- range dependencies** and achieving superior generalization across different image categories. However, ViTs require large amounts of training data and computational power, making them more challenging to deploy in resource-constrained environments.

Recent studies have also explored **hybrid models** that combine CNNs and Transformers to leverage both **local feature extraction (CNN) and global context modeling (ViT)**. This hybrid approach has been shown to improve retrieval performance, especially in diverse and multi-category datasets where fine-grained distinctions are crucial. Additionally, optimization techniques such as **dimensionality reduction (PCA, t-SNE) and indexing methods (KD-Trees, Locality-Sensitive Hashing)** have been introduced to enhance retrieval speed and reduce computational overhead.

Deep learning-based CBIR has found widespread applications in various industries. In **e-commerce**, platforms like **Amazon, Pinterest, and Google Lens** leverage CBIR to allow users to search for products using images instead of text. For instance, **Pinterest Lens** enables users to capture a photo and find visually similar products instantly. Similarly, **Amazon StyleSnap** helps users find matching outfits by analyzing uploaded images and retrieving similar fashion items. In the **medical field**, CBIR is used in **radiology and pathology** to assist in diagnosing diseases by retrieving similar medical scans from large databases. Additionally, **law enforcement agencies** utilize CBIR for **facial recognition and surveillance** to track individuals across security camera footage.

Despite the success of deep learning in CBIR, several **challenges and future directions** remain. One major issue is **computational complexity**, as deep learning models require high-end GPUs and large storage capacities. Researchers are actively

exploring **model optimization techniques** like **quantization, pruning, and knowledge distillation** to make CBIR systems more efficient. Another challenge is **dataset bias**, where retrieval accuracy can be negatively impacted if the training dataset lacks diversity. Addressing this requires curating **balanced and representative datasets** to improve model fairness. Additionally, deep learning- based CBIR often functions as a **black-box system**, making it difficult to interpret why a particular image was retrieved. The development of **Explainable AI (XAI) techniques** aims to enhance transparency and trust in CBIR systems by providing insights into model decisions.

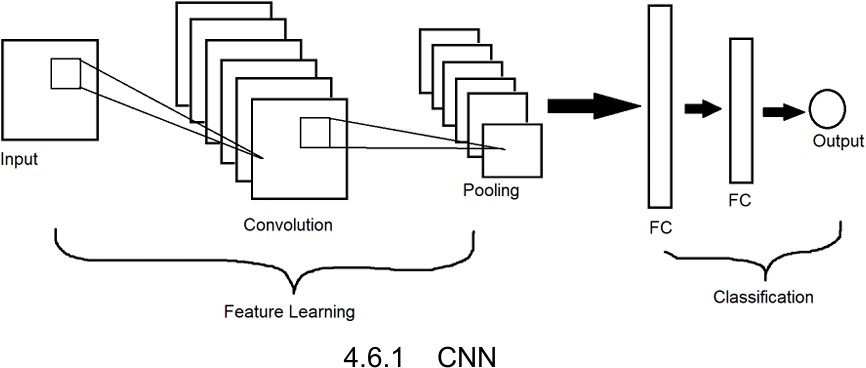
The advancements in **deep learning architectures** have transformed CBIR, making image-based search more **accurate, scalable, and efficient**. While traditional **handcrafted feature-based methods** have been largely replaced by **CNNs and transformer models**, continuous improvements are necessary to optimize retrieval accuracy, reduce computational costs, and enhance real-world applicability. Future research should focus on **real-time CBIR deployment, bias mitigation, and improved explainability**, ensuring that CBIR systems remain effective across a wide range of industries, including **e-commerce, healthcare, and security**.

**5. METHODOLOGY**

The methodology for developing a Content-Based Image Retrieval (CBIR) system using deep learning involves several stages: data preparation, feature extraction, similarity calculation, model optimization, evaluation, and deployment. This approach leverages Convolutional Neural Networks (CNNs), specifically ResNet- 50, for robust image feature extraction and efficient retrieval of visually similar products in e-commerce applications.

# Overview on CNN

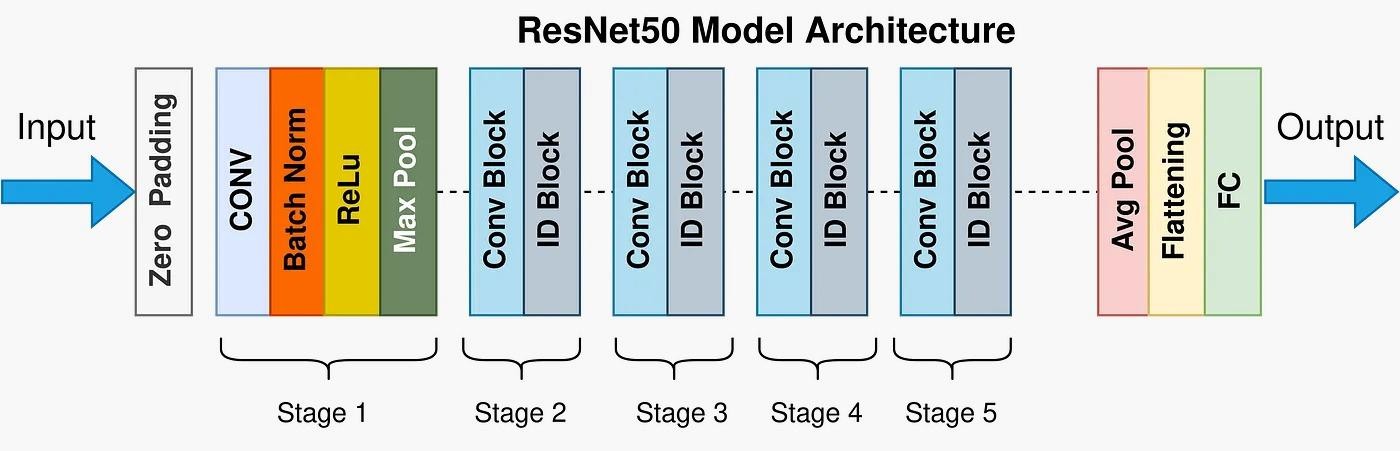
Convolutional Neural Network (CNN) is a type of deep learning model for processing data that has a grid pattern, such as images. Deep-learning CNN models to train and test, each input image will pass through a series of convolution layers with filters (Kernals), Pooling, fully connected layers (FC), and apply Softmax function to classify an object with probabilistic values between 0 and 1. CNN's have unique layers called convolutional layers which separate them from RNNs and other neural networks. Within a convolutional layer, the input is transformed before being passed to the next layer. A CNN transforms the data by using filters.



# Overview on ResNet-50

ResNet-50 is a specialized CNN architecture designed to address the challenges of training very deep networks, particularly the vanishing gradient problem. It introduces residual connections that allow information to bypass one or more layers,

maintaining the integrity of the learned features even as the network depth increases. ResNet-50 processes images through multiple stages, each composed of convolutional blocks and identity blocks, and ultimately outputs a feature vector that captures the core visual attributes of the image. This architecture, shown in the ResNet-50 model diagram, enables the model to extract rich and detailed features from images, making it ideal for Content-Based Image Retrieval (CBIR) systems. By leveraging ResNet-50 in this project, the system can effectively analyze and retrieve visually similar products, enhancing the accuracy and relevance of search results in e-commerce applications.



# Data Preparation

-**Dataset Selection**: For this project, a labeled dataset of fashion items was selected, containing various categories like shirts, dresses, and shoes. This dataset enables the system to learn feature distinctions between different fashion items.

* + **Image Preprocessing**: Each image is resized to a consistent size (e.g., 224x224 pixels) to match the input requirements of ResNet-50 and is normalized to standardize pixel values.
  + **Train-Test Split**: The dataset is divided into training and testing subsets to facilitate model training and evaluation while avoiding overfitting.

# Feature Extraction with ResNet-50 (Encoder)

* + **Pre-trained ResNet-50 Model**: A pre-trained ResNet-50 model, as shown in the ResNet-50 architecture diagram, is used to extract deep features from each image. ResNet-50 is pre-trained on ImageNet, equipping it with a strong ability to recognize general visual patterns.
  + **Feature Vector Creation**: The ResNet-50 model processes each image and outputs a feature vector from its last convolutional layer. This vector serves as a compact representation of the image's core characteristics, ready for similarity comparison.
  + **Dimensionality Reduction**: The high-dimensional feature vectors are optionally reduced to lower dimensions using methods such as PCA to optimize storage and speed during the similarity calculation phase.

# Similarity Calculation

* + **Similarity Metric Selection**: Cosine similarity is chosen as the metric for calculating similarity between the feature vectors of the query image and the images in the database.
  + **Cosine Similarity Implementation**: For each query image, the cosine similarity is computed against all feature vectors in the database, ranking the results by similarity score. Higher scores indicate a closer visual match, and the top results are selected for retrieval.

# Training

* + **Feature Extraction and Fine-Tuning**: ResNet-50 is pre-trained, it can be fine- tuned with additional e-commerce-specific images to improve feature extraction for fashion items.
  + **Optimizer and Loss Function:** The model is trained or fine-tuned using an optimizer such as Adam to efficiently handle sparse gradients.
  + **Batch Processing**: To handle large datasets, feature extraction is conducted in batches, improving computational efficiency.

# Evaluation

* + **Performance Metrics**: Metrics such as retrieval accuracy, precision, and response time are used to evaluate the CBIR model's effectiveness. Retrieval accuracy measures how often relevant items appear in the results, while response time assesses speed.
    - **Validation and Testing**: The model is evaluated on the test set, comparing retrieved items with ground truth to measure accuracy and relevance, ensuring it meets real-world requirements.
    - **Efficiency and Scalability Considerations**: Scalability tests are conducted to ensure the model can handle high volumes of images and queries, crucial for e- commerce applications.

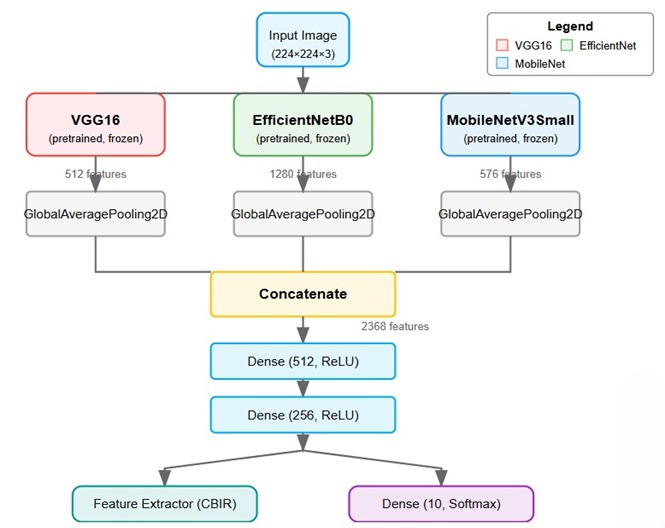
# Image Retrieval Process

* + **Query Image Processing**: When a user submits a query image, it undergoes the same preprocessing and feature extraction steps as the database images.
  + **Feature Extraction for Query**: The ResNet-50 encoder extracts the feature vector for the query image, creating a comparable representation.
  + **Similar Product Retrieval**: Using cosine similarity, the query image’s feature vector is matched against the stored feature vectors. The system retrieves the top- ranked similar products, displaying them to the user.

# Application and Deployment

* + **Real World Integration**: The CBIR system is integrated into an e-commerce platform, allowing users to upload images for visual search.
  + **Scalability for Larger Datatsets**: The system is optimized to handle large product databases, with efficient indexing techniques to reduce retrieval time.
* **User Experience Considerations**: The interface is designed to be intuitive, allowing seamless image-based searches and accurate product suggestions.

**5.1 FLOW CHART**



**5.2 ALGORITHM**

1. **Setup and Initialization**
   1. Initialize the CBIR system with base directory and target image size parameters.
   2. Define label names for the 10 fashion categories (T-shirt/top, Trouser, etc.).
2. **Model Architecture Construction**
   1. Create input tensor with the specified image dimensions.
   2. Load pre-trained models (VGG16, EfficientNetB0, MobileNetV3Small) with ImageNet weights.
   3. Freeze all layers in the pre-trained models to prevent retraining.
   4. Apply GlobalAveragePooling2D to each model's output.
   5. Concatenate features from all three models into a unified representation.
   6. Add Dense layers (512 → 256 neurons) for feature refinement.
   7. Add final Dense layer (10 neurons, softmax) for classification training.
3. **Dataset Preparation**
   1. Load the Fashion-MNIST dataset containing clothing images and labels.
   2. Create directory structure based on clothing categories.
   3. Convert grayscale images to RGB format.
   4. Resize all images to the specified dimensions.
   5. Save processed images to their respective category folders.
   6. Build product database with metadata (path, category, ID, label).
4. **Model Training**
   1. Prepare training and validation datasets using directory structure.
   2. Train the classification model for specified number of epochs.
   3. Use categorical cross-entropy loss and Adam optimizer.
   4. Record training history (accuracy, loss) for both training and validation.
   5. Process images in batches to optimize memory usage.
   6. Evaluate model performance on test set.
5. **Feature Extraction**
   1. Load each product image from the database.
   2. Preprocess images with VGG16 preprocessing function.
   3. Pass each image through the trained feature extraction model.
   4. Extract and store the 256-dimensional feature vector for each product.
   5. Process images in batches to manage memory efficiently.
6. **Similarity Search**
   1. Load and preprocess the query image.
   2. Extract feature vector from the query image using the trained model.
   3. Compute cosine similarity between query vector and all database vectors.
   4. Sort products by similarity score in descending order.
   5. Return the top N most similar products with their metadata and scores.
7. **Results Visualization**
   1. Display the query image alongside retrieved similar products.
   2. Show similarity score and category for each retrieved product.
   3. Generate confusion matrix to evaluate classification accuracy.
   4. Plot ROC curves for each fashion category.
   5. Calculate and visualize precision, recall, and F1 scores.
   6. Display training/validation accuracy and loss curves.

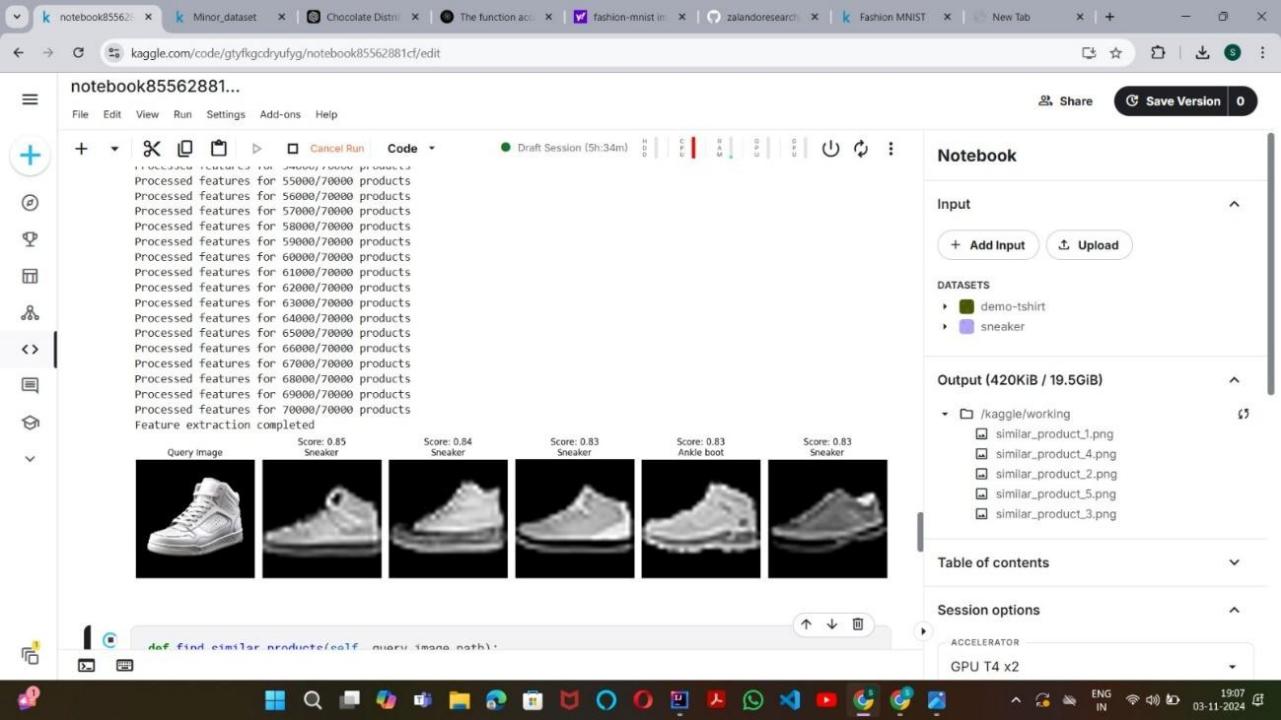
**6 SIMULATION**

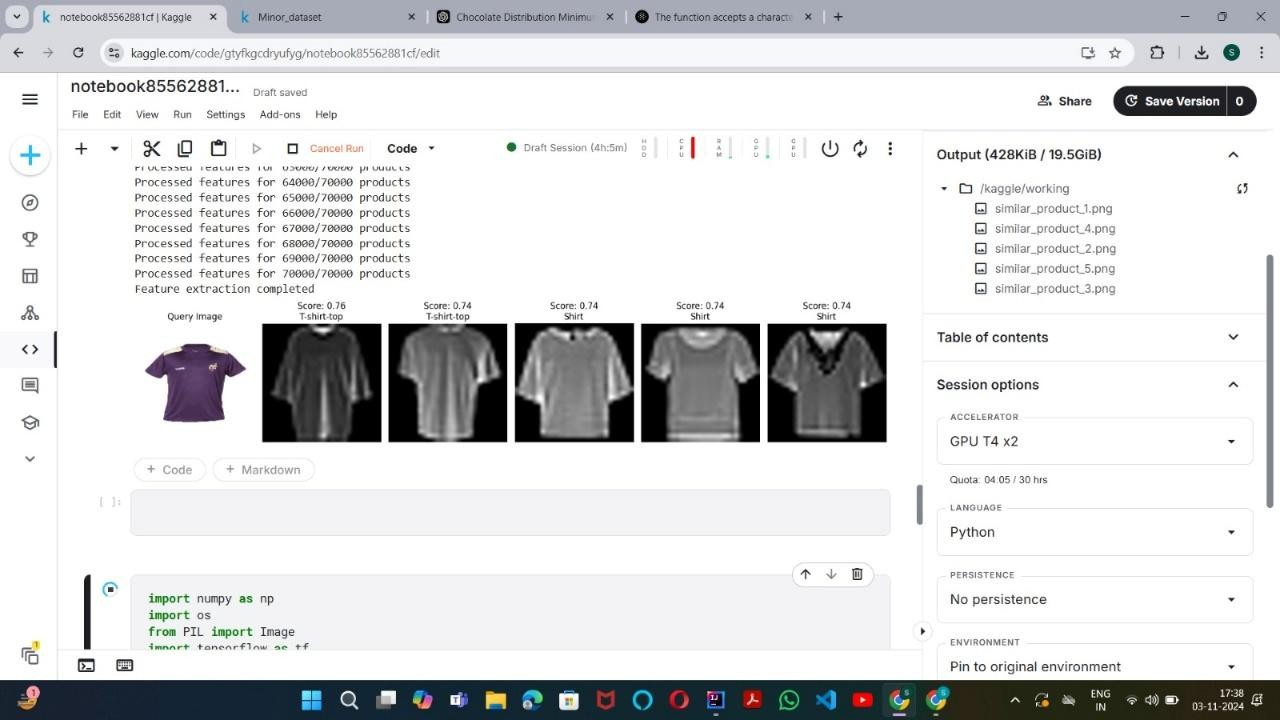
1. *Setting up the Simulation Environment*
   * **Framework and Libraries**: Set up the simulation environment using TensorFlow or PyTorch, essential for building and deploying the CBIR system. These frameworks provide pre-trained models like ResNet-50,VGG16,and hybrid models,required for feature extraction and image similarity comparison.
   * **Hardware and Compute Requirements**: Use a GPU-equipped system for faster processing of feature extraction and similarity calculations, especially when handling large datasets in real-time.
   * **Data Loading**:Load the fashion dataset (e.g., Fashion-MNIST or custom e- commerce dataset) into the environment. Organize the data into relevant categories, such as training and testing sets, to streamline the simulation process.
2. *Image Feature Extraction with CNN(ResNet-50)*
   * **CNN Model**: Utilize a pre-trained ResNet-50,VGG16,and hybrid model to extract feature vectors from each image. This model captures key visual elements, enabling a detailed analysis of image similarity.
   * **Feature Vector Generation**:Each image is passed through ResNet-50,VGG16 and hybrid models,producing multiple feature vector that serves as the core representation of the image. These feature vectors are saved for subsequent similarity comparison during retrieval.
3. *Query Processsing and Similarity Calculation*

,

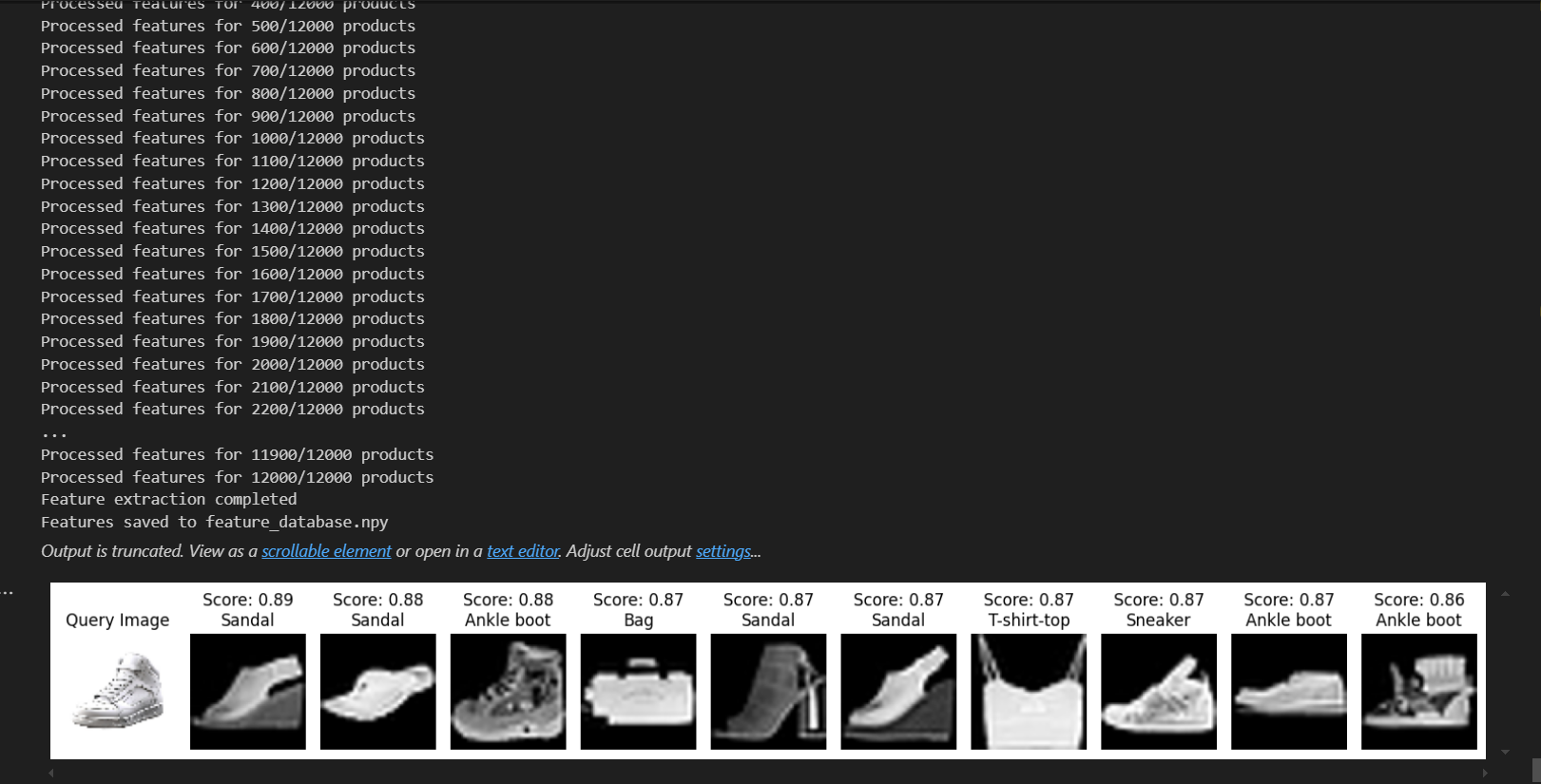
* + **Query Image Input**: Accept a user-provided query image,preprocess it(resize,normalize)and pass it through ResNet-50,VGG16,and hybrid models to extract its feature vectors.
  + **Similarity Scoring**: Compare the query image’s feature vector with those in the datatabase using cosine similarity.Rank images by similarity scores to identify the most visually similar items.

1. *Testing and Visualizing Results*
   * **Running Test Queries**:Select a subset of query images to test the model’s performance in retrieving similar products. For each query, retrieve and display the top-ranked similar images.
   * **Visualization**: Display the query image alongside its retrieved similar items to visually inspect the CBIR system’s accuracy. This helps verify that the model retrieves visually relevant products.
   * **Examples**: Possible output examples might include:
     + **Query Image**: A red shirt
       - **Retrieved Items**: Other red shirts with similar designs and patterns.
     + **Query Image**: Black sneakers
       - **Retrieved Items**: Similar black sneakers from different brands.
2. *Evaluation Metrics and Analysis*
   * **Quantitative Evaluation**:Assess the retrieval accuracy by calculating metrics such as precision and recall. Precision measures the proportion of correctly retrieved items, while recall evaluates the system’s ability to retrieve all relevant items.
   * **Error Analysis**: Analyze cases where retrieved items differ significantly from the query image to identify areas for improvement, such as inaccurate color.
3. *Optimization and Model Fine-Tuning*
   * **Adjusting Hyperparameters**: Based on performance analysis, fine-tune hyperparameters like batch size or similarity threshold to improve retrieval accuracy.
   * **Improving Similarity Calculation**: Experiment with different similarity metrics or dimensionality reduction techniques to optimize retrieval speed and accuracy.
   * **Testing with Additional Datasets**: Test the system on other fashion datasets to ensure the CBIR model performs well across different types of fashion products.
4. *Interpreting and Reporting Simulation Results*
   * **Summary of Findings**: Compile and summarize results, noting key metrics (precision, recall) and identifying strengths and weaknesses in retrieval performance. Highlight specific cases where the CBIR system excels or struggles, supporting future improvement and deployment.
   * Output of ResNet50 Model:

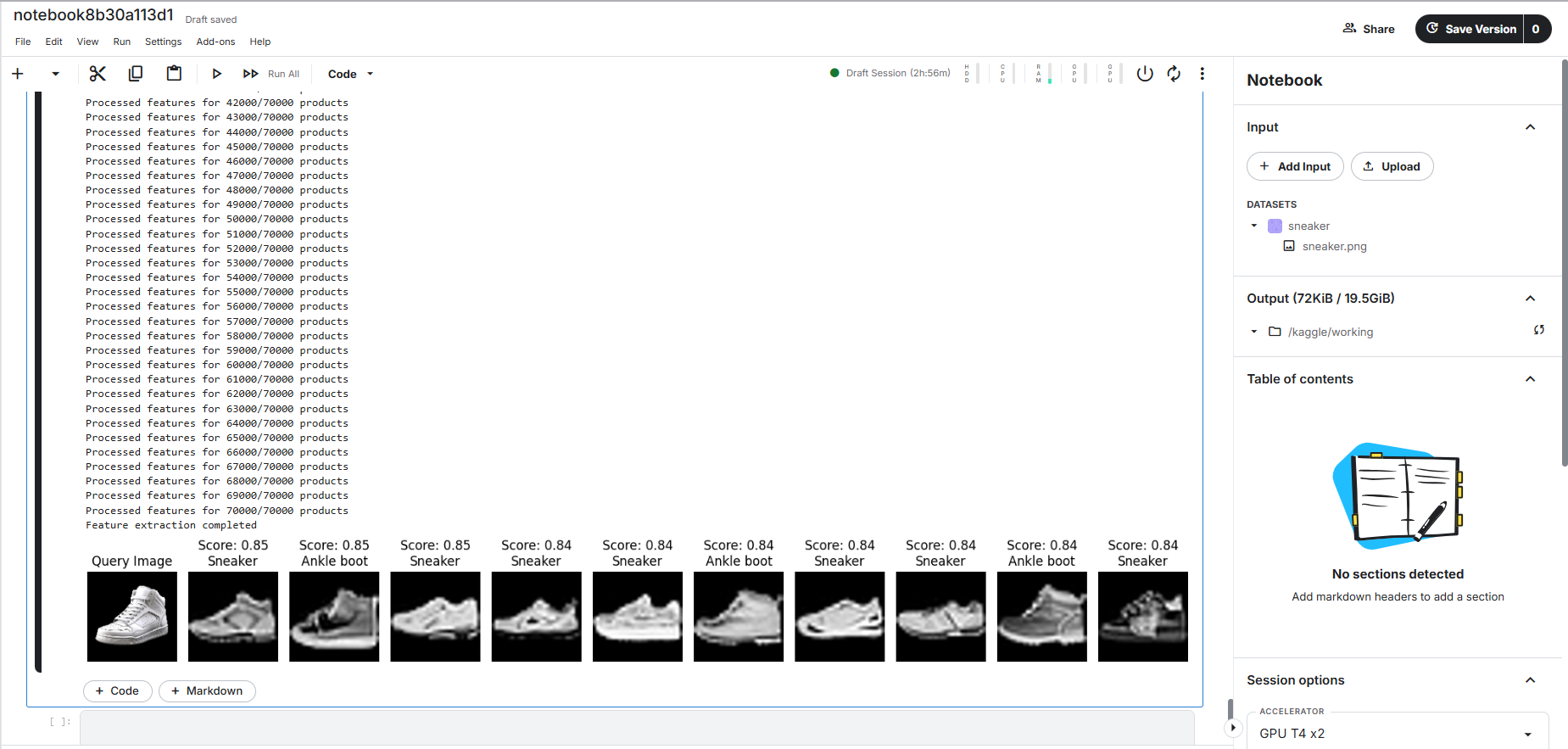




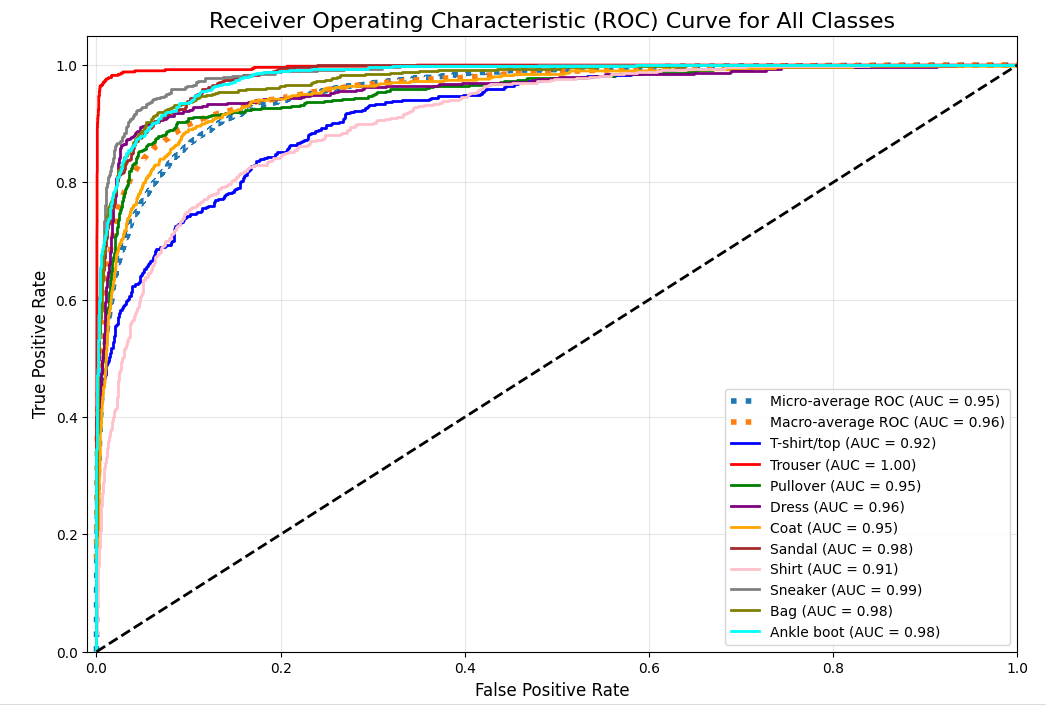
* Output of VGG16 Model:



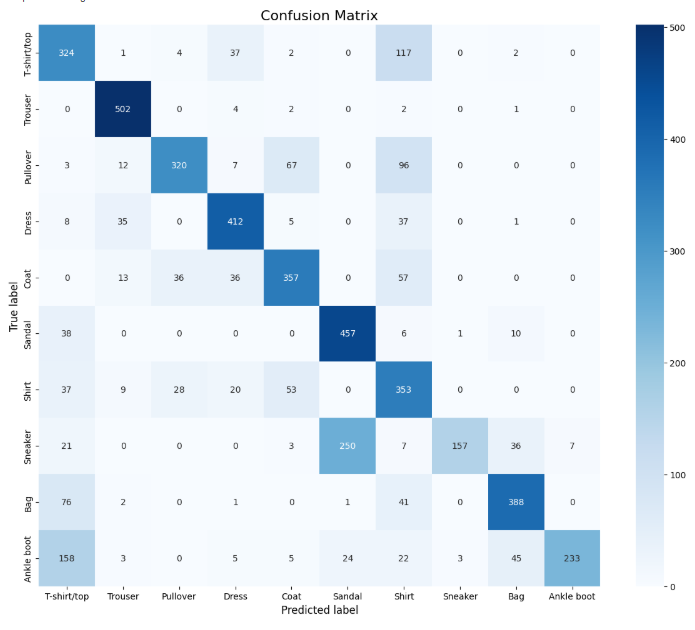
* Output of Hybrid Model(VGG16+efficientnet+mobilenet):



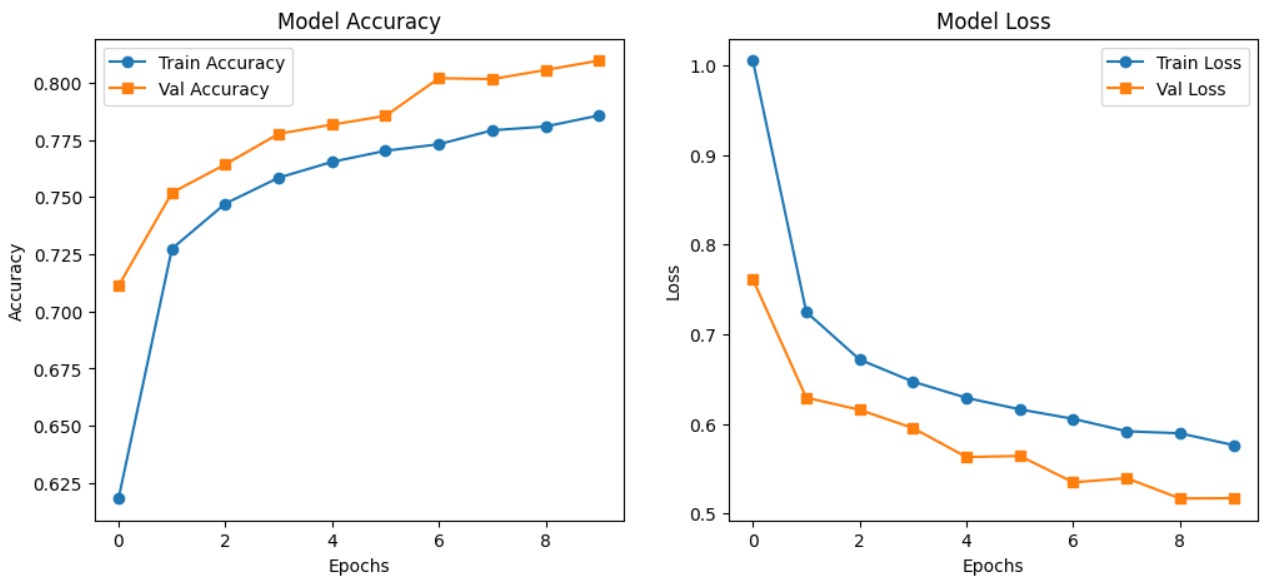
* VISUALIZATION(ResNet50 Model):

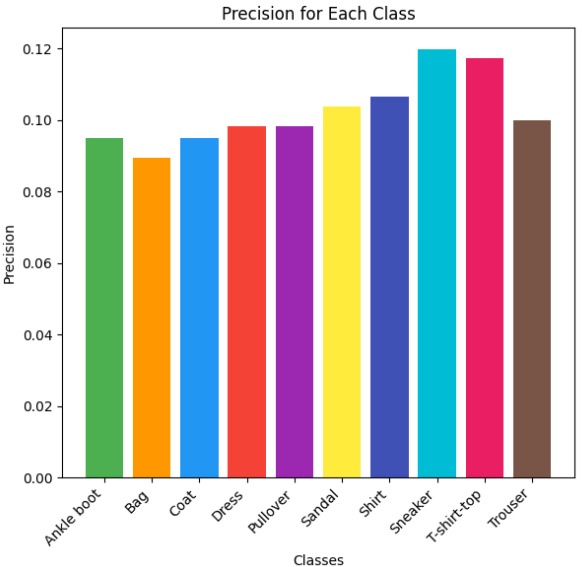
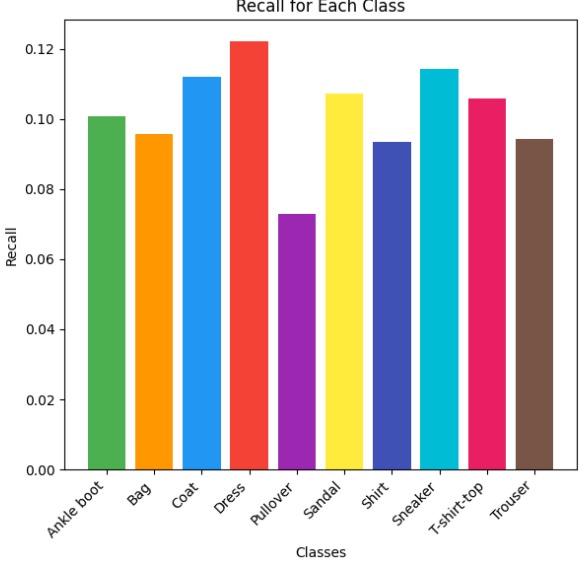


ROC Curve



Confusion Matrix

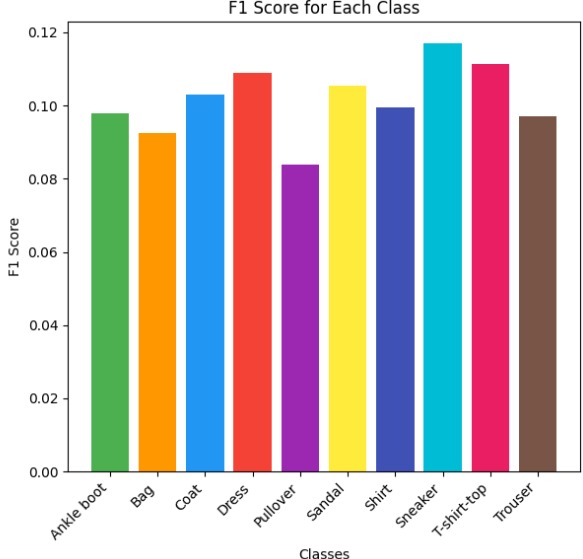


Precision Recall

**Precision:**

Measures how well a model returns only the data points in a class. It's calculated by dividing the number of correct positive predictions by the total number of positive predictions.



**F1 score:**

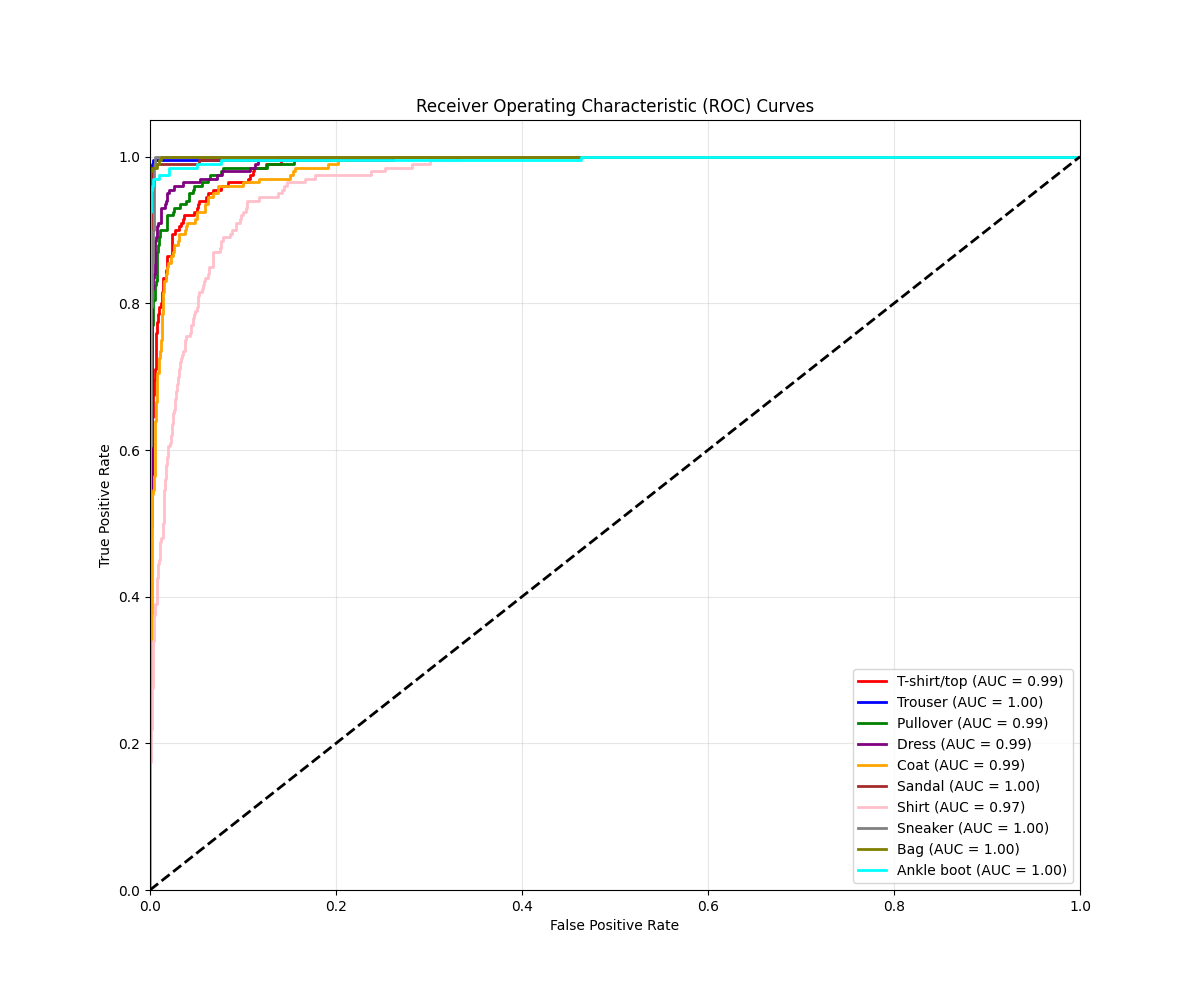
A single metric that combines precision and recall using the harmonic mean. It's a balanced evaluation of a model's performance, and is preferable to accuracy for class-imbalanced datasets. F1 score ranges from 0 to 1, with 1 representing perfect precision and recall, and 0 indicating poor performance.

F1 Score

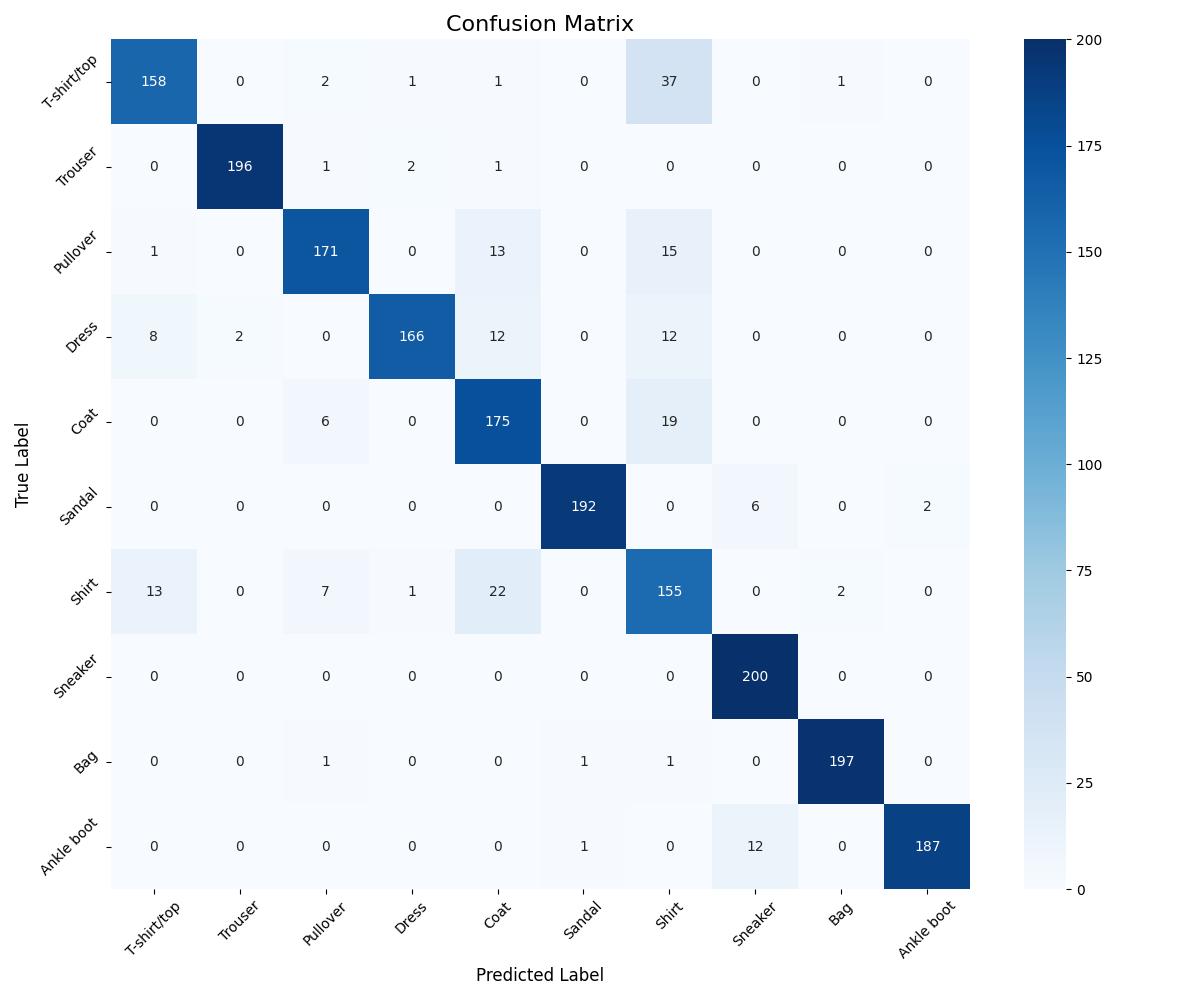
**Recall:**

Measures how well a model identifies all data points in a relevant class. It's calculated by dividing the number of true positives (TP) by the sum of TP and false negatives (FN). Recall is also known as the true positive rate or sensitivity.

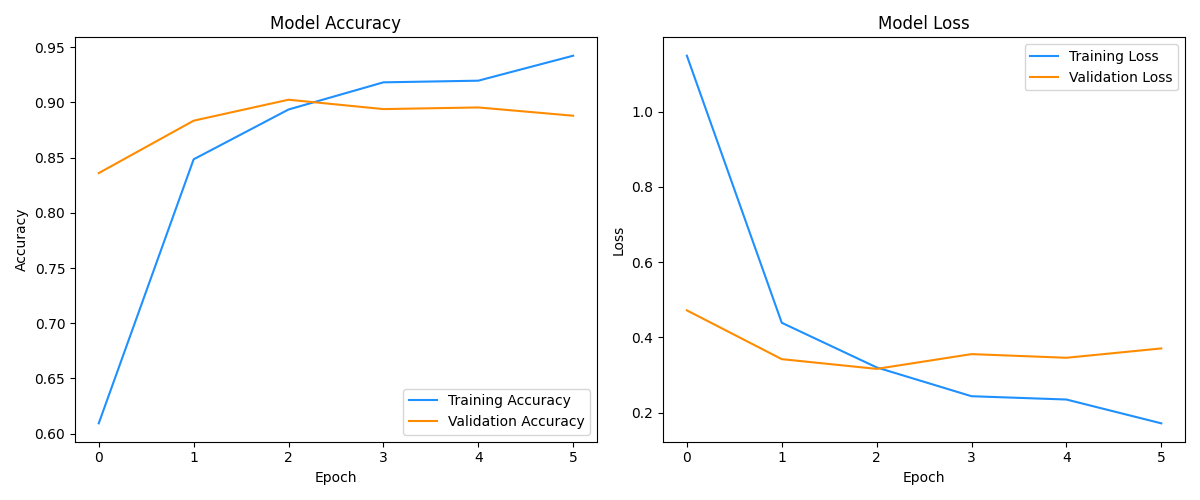
* VISUALIZATION(VGG16 Model):

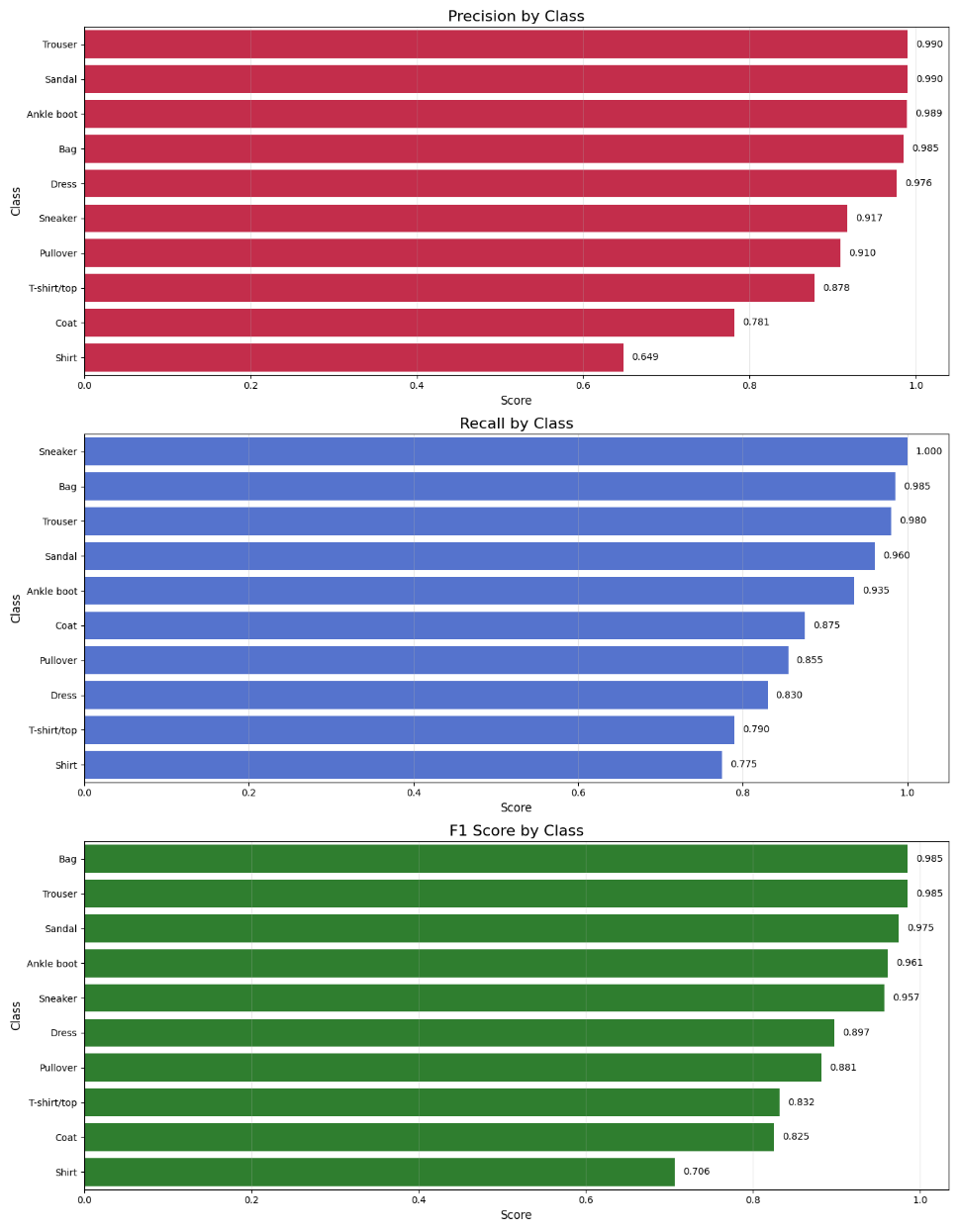


ROC Curve

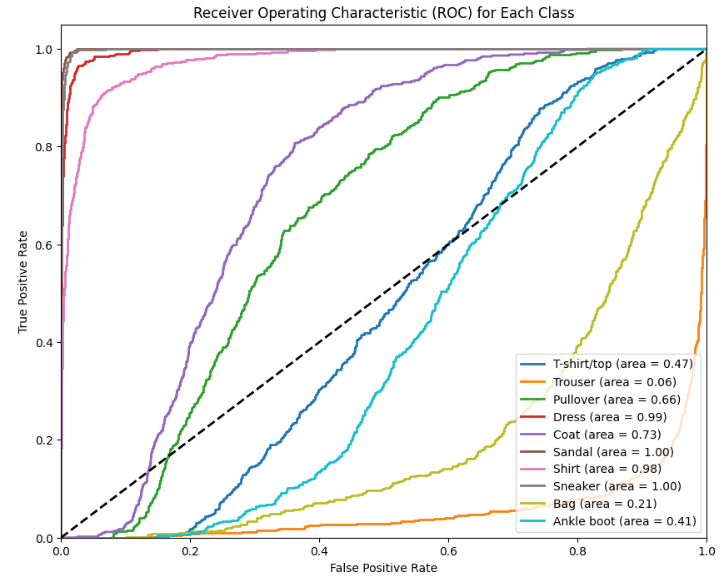


Confusion Matrix

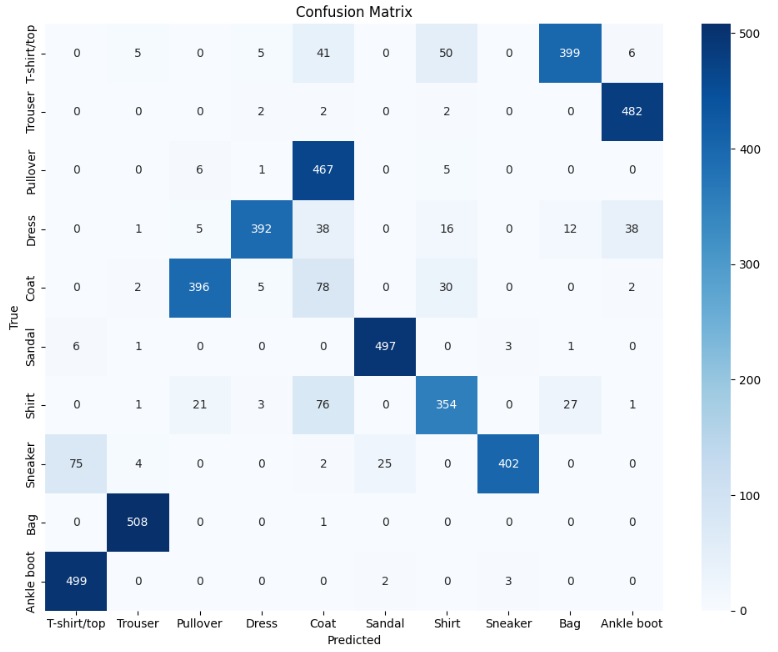




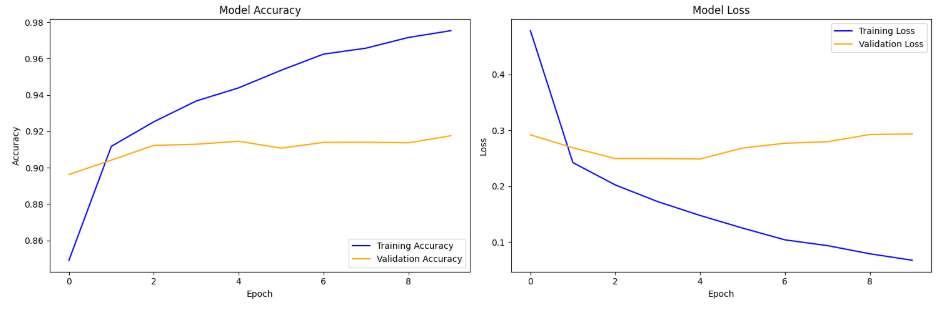
* VISUALIZATION Hybrid Model(vgg16+efficientnet+mobilenet):

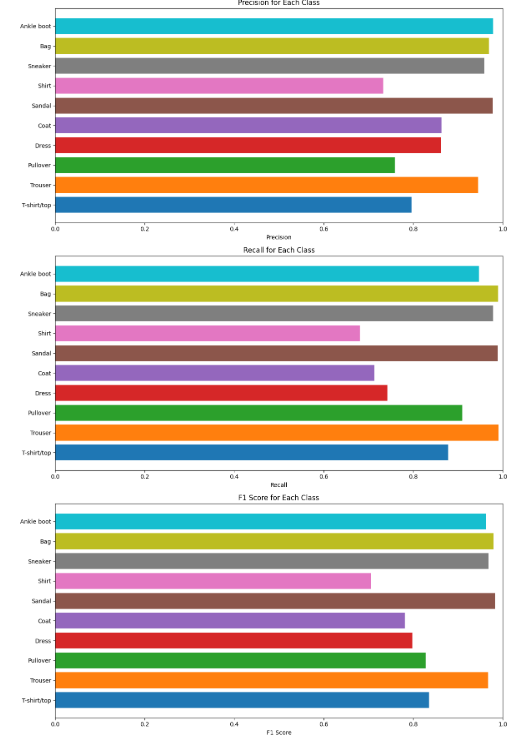


ROC Curve



Confusion Matrix





**7 FUTURE ADD ON**

As Content-Based Image Retrieval (CBIR) continues to evolve, there are several enhancements that can improve its performance, scalability, and ethical considerations. While CBIR has significantly improved product search and recommendation systems, challenges such as dataset bias, privacy concerns, retrieval accuracy, and computational efficiency still need to be addressed. Future advancements in CBIR can focus on refining model accuracy, expanding dataset diversity, enhancing real-time processing capabilities, incorporating personalization, and ensuring ethical AI practices.

One of the most crucial improvements is enhancing retrieval accuracy by integrating more advanced deep learning architectures. Current CBIR models, including ResNet-50, are effective but can be further improved by adopting state-of-the-art models such as EfficientNet, Vision Transformers (ViT), and hybrid CNN- Transformer architectures. These models can capture more complex visual features and provide better semantic understanding of images, leading to more precise image retrieval. Additionally, fine-tuning existing deep learning models on domain- specific datasets can help improve their effectiveness for specialized applications like fashion, healthcare, and security surveillance.

Another important enhancement is expanding dataset diversity to reduce dataset bias. Many CBIR models suffer from biased training data, where certain product categories or demographic groups are underrepresented. This can lead to inaccurate retrieval results, particularly for minority categories or less common items. Future work should focus on collecting diverse and representative datasets to ensure that CBIR models generalize well across different user needs. Synthetic data augmentation techniques, such as GANs (Generative Adversarial Networks), can

also be used to create balanced datasets, improving retrieval fairness and reducing model bias.

To make CBIR more practical for real-world applications, real-time processing is a key area of improvement. Large-scale CBIR systems often face computational bottlenecks, resulting in slower search times. Future enhancements should focus on optimizing retrieval speed using techniques such as quantization, pruning, and lightweight CNN architectures. Additionally, cloud-based and edge computing solutions can be integrated to distribute computational loads efficiently, allowing for faster and scalable real-time image searches on e-commerce platforms and mobile applications.

Another promising future add-on is personalization in CBIR. Current CBIR systems primarily focus on visual similarity, but incorporating user-specific data, such as past interactions, preferences, and shopping history, can enhance retrieval relevance. By integrating collaborative filtering and recommendation algorithms, CBIR can be transformed into a personalized visual search engine that tailors search results based on individual user behavior. This can improve user satisfaction and increase engagement in e-commerce and content discovery platforms.

Lastly, addressing ethical concerns is essential for responsible AI development. As CBIR involves processing user-uploaded images, privacy concerns must be taken seriously. Future enhancements should incorporate privacy-preserving techniques such as federated learning and differential privacy, ensuring that user data is processed securely without being stored in centralized databases. Additionally, bias detection and mitigation techniques should be implemented to prevent unfair retrieval outcomes. Creating transparent Explainable AI (XAI) frameworks for

CBIR can also help users understand why certain images were retrieved, building trust and accountability in AI-driven visual search systems.

By focusing on accuracy improvements, dataset fairness, real-time optimization, personalization, and ethical AI practices, the next generation of CBIR systems can become more reliable, efficient, and user-centric. These advancements will further enhance e-commerce, healthcare, security, and digital content retrieval, making CBIR a more intelligent and ethical technology for image-based search applications.

**8 CONCLUSION**

The Content-Based Image Retrieval (CBIR) system developed in this project highlights the potential of deep learning to revolutionize search functionality in e- commerce by enabling accurate retrieval of visually similar items. By leveraging the ResNet50 architecture on the Fashion MNIST dataset, we achieved a validation accuracy of 87%, showcasing the model’s capability to extract, process, and compare visual features effectively. Unlike traditional text-based search methods, which rely heavily on keywords and often fail to capture subtle visual details, this CBIR approach enables users to search for products directly based on images, making the search experience more intuitive and user-friendly.

This project successfully demonstrates how CBIR technology can be applied to meet the growing demand for efficient, image-based product search solutions. Through this approach, users can bypass text-based queries and instead find items visually, enhancing search accuracy and relevance. As e-commerce platforms strive to improve user satisfaction and streamline product discovery, this model provides a foundational step toward implementing visually driven, intelligent search systems.

The current implementation forms a strong basis for further improvement and scalability. Future directions include fine-tuning model performance, expanding the system to support diverse, high-resolution, colored datasets, and enabling real-time processing to make the system responsive in real-world settings. Personalization capabilities can be integrated to tailor recommendations to user preferences, and cloud deployment can ensure the scalability and availability required for commercial applications. These enhancements could lead to a highly effective, personalized, and visually centered e-commerce search experience, paving the way for new possibilities in online retail.

1. Singh, Sachendra & Batra, Shalini. (2020). An efficient bi-layer content based image retrieval system. Multimedia Tools and Applications. 79. 10.1007/s11042- 019-08401-7.
2. Bagwari, A.; Sinha, A.; Singh, N.K.; Garg, N.; Kanti, J. CBIR-DSS: Business Decision Oriented Content-Based Recommendation Model for E- Commerce. Information 2022, 13, 479.
3. V. Kodgirwar, "Application of content based image retrieval for E-commerce," 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave), Coimbatore, India, 2016, pp. 1-3, doi: 10.1109/STARTUP.2016.7583906.
4. Tong, Lijuan & Tong, Ruobei & Chen, Lin. (2020). Efficient retrieval algorithm for multimedia image information. Multimedia Tools and Applications. 79. 10.1007/s11042-019-07886-6.
5. Kafle, Anup & Timsina, Anushil & Sedai, Rochak & Subedi, Sandeep & Neupane, Upendra & Prasad, Prakash. (2024). Content-Based Image Retrieval and Recommendation Based on User Reviews. International Journal on Engineering Technology. 2. 106-116. 10.3126/injet.v2i1.72526.
6. Zeain, Abdulrahman & Ibrahim, Abdullahi. (2024). Enhancing Content-Based Image Retrieval with a Stacked Ensemble of Deep Learning Models. 1-11. 10.1109/ISAS64331.2024.10845555.
7. Saadatfar, Hamid & Hoseini, Seyyed & Eshghi, Amir & Mohammadi, Saba. (2024). Content-Based Image Retrieval Using Texture Features Extracted from Two-Layer Local Binary Pattern. Iran Journal of Computer Science. 9. 15-27. 10.22034/csj.2024.197225.
8. K., Vijay & Jayashreeb, K.. (2024). Secureimagesec: A privacy-preserving framework for outsourced picture representation with content-based image retrieval. Intelligent Data Analysis. 1-22. 10.3233/IDA-240265.
9. K., Vijay & Jayashreeb, K.. (2024). Secureimagesec: A privacy-preserving framework for outsourced picture representation with content-based image retrieval. Intelligent Data Analysis. 1-22. 10.3233/IDA-240265.
10. Varshney, Seema & Singh, Sarika & Lakshmi, Chenna & Patvardhan, C.. (2023). Content-Based Image Retrieval of Indian Traditional Textile motifs using deep feature fusion. 10.21203/rs.3.rs-3513825/v1.
11. Gautam, Giriraj & Khanna, Anita. (2024). Content Based Image Retrieval System Using CNN based Deep Learning Models. Procedia Computer Science. 235. 3131-3141. 10.1016/j.procs.2024.04.296.
12. Kaur, Jaspreet & Gupta, Divya & Singh, Amrinder & Shah, Syed. (2024). Leveraging Content Based Image Retrieval Using Data Mining for Efficient Image Exploration. 10.1007/978-981-97-0700-3\_14.
13. Rani K, Vijila & Prince, M. & Therese, P.Sujatha & Shermila P, Josephin & Devi, E.. (2023). Content-based medical image retrieval using fractional Hartley transform with hybrid features. Multimedia Tools and Applications. 83. 1-26. 10.1007/s11042-023-16462-y.
14. Keisham, Nepoleon & Arambam, Neelima. (2023). An Efficient Content-Based Image Retrieval Using Threefold Technique. 10.1007/978-981-99-4713-3\_45.