

# CS-4641-Outfit-Picker

---

## Introduction:

This project focuses on creating an AI-generated wardrobe through an algorithm capable of identifying the type and color of clothing in an image. Users will input the clothing items they have in their closet, and the model will then suggest outfit combinations by applying complementary color pairings and matching the compatibility of different clothing items, based on the season and gender. This takes away the stress of picking an outfit each morning by providing the user with inspiration. To achieve this, we will use a Kaggle dataset that includes features such as clothing types and their corresponding colors. Given the dataset's extensive size, we divide it into training and testing sets to evaluate the algorithm's accuracy in detecting and labeling clothing. [Link to the dataset:](#)

## Problem Definition:

One of the first challenges of the day is often getting dressed which can cause a certain level of frustration, so our model aims to alleviate this frustration by simplifying the process. Additionally, it can also assist individuals with visual impairments by acting as a guide to help them avoid clashing outfits. The initial stages of the project involved testing using supervised learning algorithms like Random Forest Classifiers and Support Vector Machine (SVM) models to determine which performs better at labeling clothing items. We further developed it by utilizing PCA (principal component analysis) for "reducing the dimensionality" and grouped clothing items into clusters and reduced the complexity of the features [1]. Lastly, we used neural networks for recommending complementary colors to learn the relationships between colors and fashion trends from the dataset.

## Methods:

For data processing, PCA and K-means clustering were used as they are effective in dimensionality reduction and clustering, respectively. PCA played a significant part in reducing the dataset to its most meaningful features by eliminating irrelevant variables while retaining the ones that have the highest variance. Moreover, K-means was applied to group the clothing items in clusters based on features like color, category, and style. The elbow method was used to determine the optimal number of clusters helping in making the data well-structures for further analysis.

For supervised learning, Random Forest Classifier was initially experimented with because it can handle high-dimensional data and gives robust predictions. However, the model was overfitting, mainly when classifying items with multiple output labels, which led us to explore other approaches. Next, we implemented a Support Vector Machine (SVM), which performed better in distinguishing

items within the MasterCategory. Its accuracy did go down for smaller classifications like baseColor and articleType.

Lastly, we used a Neural Network (NN) for classifying article types from uploaded images. The NN was designed to learn the relationships between tops, bottoms, and footwear by analyzing the extracted features. Image data was preprocessed to make sure it was compatible with the pre-trained models, and the dominant colors were identified using K-means clustering. These colors were matched with their complements to suggest cohesive outfit combinations. Throughout our research, we found that neural nets were the optimal choice for classifying clothing. In particular, ResNet18 was found to be the best choice, particularly when trying to cluster given a large number of options; in our case, the model had to choose between 30 article types and hence, we found it to be the best for our use case.

## Results and Discussion:

### Analysis

We used Principal Component Analysis (PCA) for the task of reducing the dimensionality of the dataset while still retaining the most information/variance. By doing this, we were able to drop columns of the data that were not useful in our recommendation system, keep only the data points/images that would contribute to an outfit, and reduce the data into two principal components that held significant meaning in terms of "Apparel". Each string datapoint was transformed into a float that accurately represented its relationship to other points in its vicinity. Image data was transformed into an average of the color pixels in the image and one-hot encoding was used to turn categorical features into binary indicators. After the numeric features were scaled, the K-means algorithm was used to cluster the data. The ideal number of clusters (k) were determined using the elbow method as seen in the graph. The algorithm effectively deduced clusters based on color, category, and style which are color coded. In this way, the data was sorted in clusters and ready to be used by NN and Random Forest. Another algorithm we used to categorize the clothing items was Support Vector Machines. The data was split into training and testing to reduce overfitting. The model had relatively high accuracy classifying items in the MasterCategory, however this accuracy dropped when classifying items into baseColour and articleType.

Lastly, Neural Networks were used for the task of creating outfit combinations by learning the complex relationship between tops, bottoms, and footwear. Initially, images were preprocessed to ensure compatibility with pre-trained models when passed into the neural networks. Outputs of the NN were converted to probabilities using softmax, and the top 3 class predictions were extracted and used for interpretable results. The dominant color of images was extracted using K-means, and their complementary colors were used to match certain items with other clothing pieces. The RGB of the complementary colors were found by subtracting 255 by the dominant color that was found in the image. Doing so would allow for a more cohesive outfit.

The Neural Network approach resulted in a moderate success rate when classifying the article types into the 30 categories. The model showed an accuracy of 45% on article types within topwear, 70% for bottomwear, and 53% on article types within shoes. In classifying the categories, the models performed at rates of 91%, 94%, and 92% respectively for topwear, bottomwear, and shoes. This shows that the model could classify categories well but struggled when it had to define more specific article types. As such, one of the next steps our team identified would be to train the model for a lesser number of groups. Many of the categories included within the classification were similar, resulting in miscategorizations. This was supported with the confusion matrix; for instance, casual shoes were often misclassified as sports shoes. Similarly, there were numerous instances where inner vests were classified as T-shirts. The current approach within the model removed any category that occurred less than 30 times within the dataset; however, potentially changing this to a larger value would have been beneficial, as 30 instances may not be sufficient to properly identify a category. The current implementation of the CNN only unfroze the last layer; for further experimentations, we could try unfreezing more layers to be able to test for the optimal performance.

Finally, the K Means algorithm was able to classify the dominant color after applying a mask to the image. The mask would effectively remove all of the background utilizing CV2 preprocessing. The K-means algorithm would find the rgb of the dominant color with an accuracy of 69%. It is also important to note that the ground truths listed these colors as broader categories; however, the RGB is much more specific. As a result, there was a tolerance implemented into the evaluations to account for this.

Much of the misclassifications could potentially be attributed to the imbalance of clothing types within the dataset as well. For instance, the initial dataset had an extreme imbalance when comparing groups such as T Shirts, with 2906 instances, versus something like sweatshirts with only 263 instances. Utilizing a more balanced dataset with less variance could have allowed for results within training to be further optimized, resulting in better testing evaluations. Another possible approach would be to train neural nets on footwear, apparel, and topwear separately so that the network could learn more meaningful relationships between articles of clothing within each category. This would allow the model to achieve higher accuracy and precision by category, instead of trying to discriminate between categories when the image is uploaded.

## Comparison

PCA helped with feature extraction which significantly improved the performance of the NN. While our SVM algorithm improved the overfitting issue with the Random Forest, it did not completely resolve it. In NN's, the computational load during training was significantly greater, and for Random Forest, overfitting was still prevalent but reduced due to compact representation of the data. Hence, the team decided to move forward with more complex models like ResNet-18 to extract feature data with convolution which was not possible with our previously attempted models. The use of K-means in clustering and outfit selection was effective as this algorithm is straightforward to implement and computationally inexpensive compared to NN. In fact it was more ideal for low-level tasks such as

color identification, providing high efficiency with minimal effort. SVM is a supervised algorithm that was well suited for article classification since there were labels provided in the dataset. It could serve as an alternative to NN since the hyperplanes found by SVM serve to classify at high-dimensional feature spaces. However, SVM might struggle with high dimensional data as compared with deep learning methods. On the other hand, K-means is lightweight and was used for the color detection task, yet would struggle with learning decision boundaries and applying it to classification tasks. For our outfit suggestion algorithm, K-means complemented PCA by doing the low-level work and preparing the data to be processed by PCA. In terms of categorizing items, SVM did not perform as accurately as PCA and often classified similar clothing into wrong categories. Although it is useful in reducing overfitting, PCA provided better analysis for the purposes of our project.

## Next Steps

For future implementation of the application we plan on building an API that can process images in large quantities for better scalability. We also want to provide the users with a personalized experience by allowing them to train their own models based on their wardrobe and personal style preferences. We can also enhance the color recommendation algorithm using color theory models which would refine the complementary matching beyond just RGB subtraction. Lastly, we can make the model more robust by incorporating additional features like aesthetics, seasonal adaptations, and occasion-based recommendations.

Gantt Chart (also on uploaded on the repository):

<https://docs.google.com/spreadsheets/d/14Gmxkaln59VNtca28qT3HHAYiilfrBQZosTXbp7RQnU/edit?usp=sharing>

Contribution Table:

Ava: Writeup & pca

Priya: Writeup & pca

Devansh: Neural network/Kmeans

Nisarg : Neural network/Kmeans

Tahsin: pca/video

Appendix:

Figure 1: Complete Confusion Matrix over all Article Types (Neural Network Result)

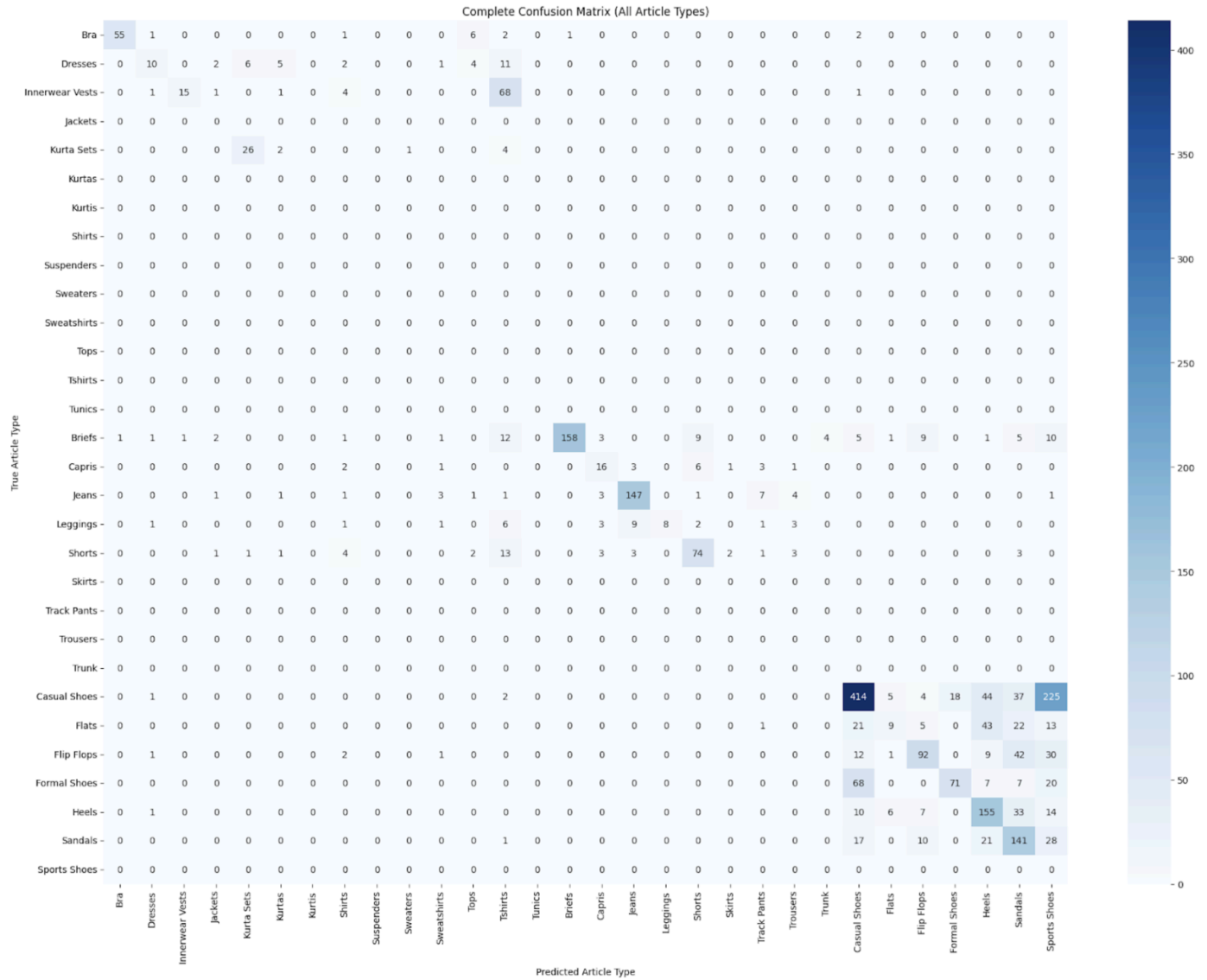


Figure 2: Category level confusion matrix (Neural Network Result)

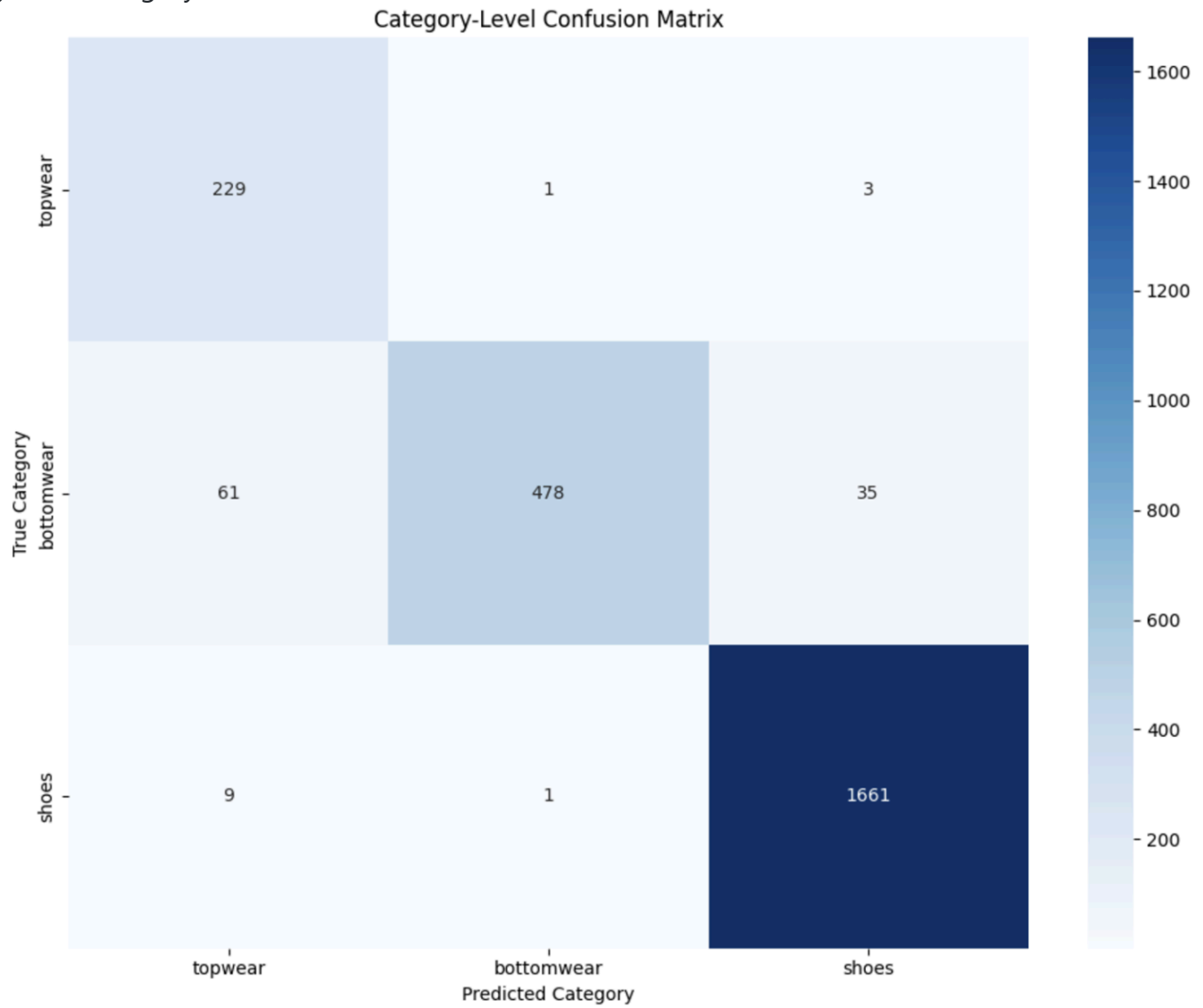


Figure 3: Confusion Matrix for topwear (Neural Network Result)



Figure 4: Confusion matrix for bottomwear

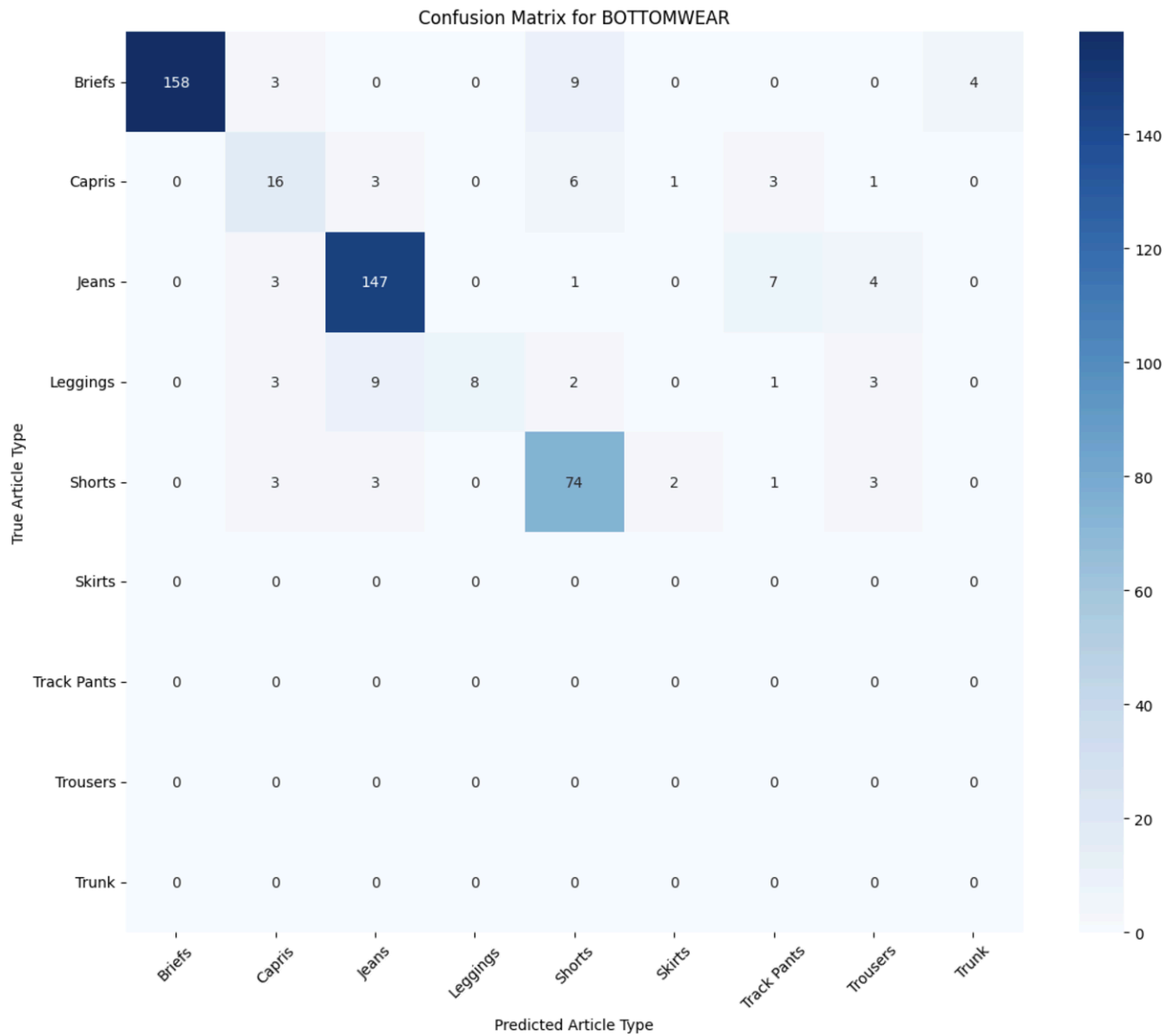




Figure 5: Confusion matrix for shoes

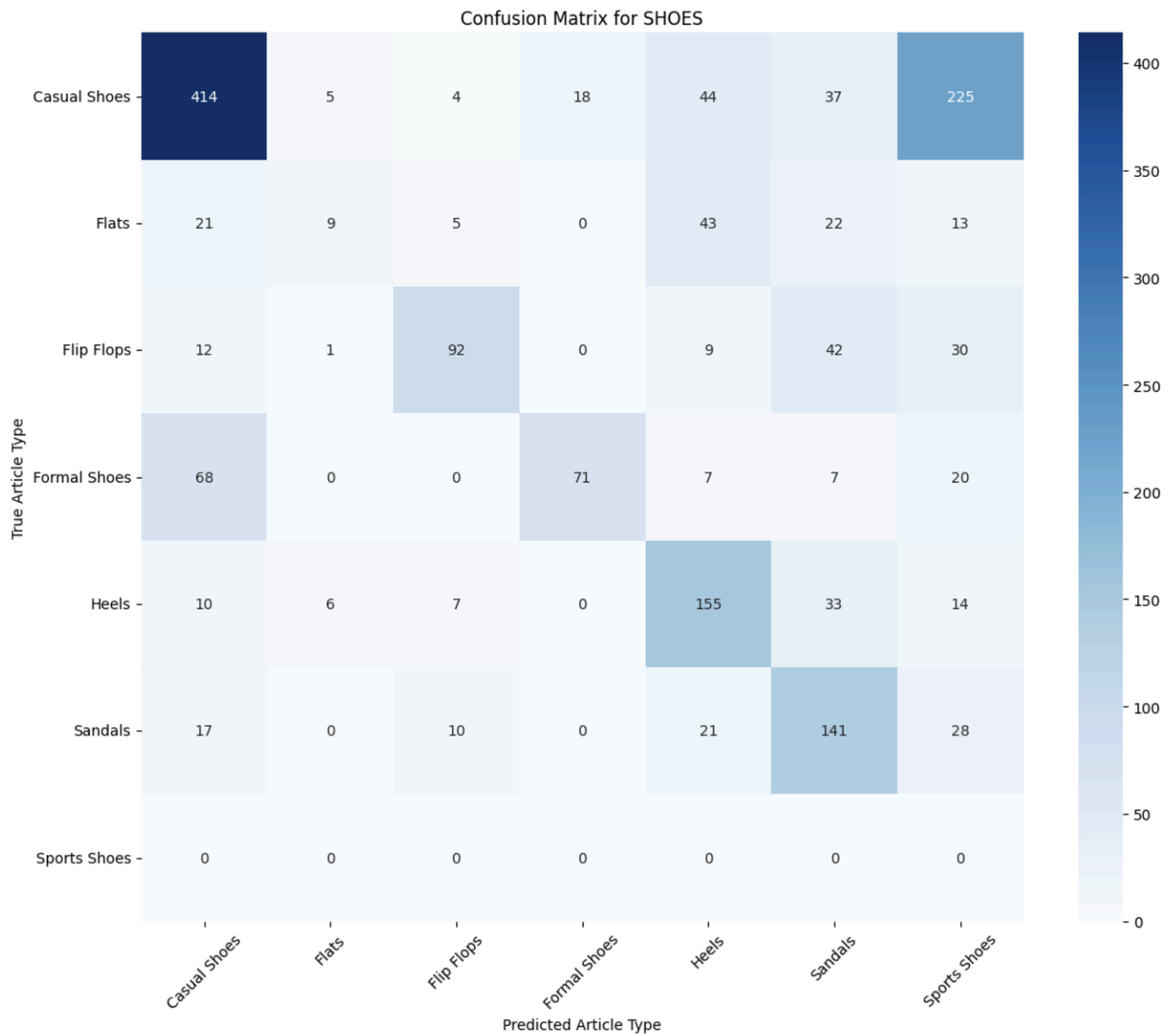


Figure 6: Clothing Classification and Outfit Generation

Predicting class for 9990.jpg:



Top 3 predictions:

1. Tshirts – 88.51%
2. Sweatshirts – 7.83%
3. Sweaters – 1.75%

Uploaded article details:

Filename: 9990.jpg  
Article: Tshirts (88.51%)  
Color: (153, 154, 155)  
Category: topwear

Recommended complementary pieces:

Filename: 26996.jpg  
Article: Track Pants (97.66%)  
Color: (121, 93, 89)  
Category: bottomwear

Filename: 9216.jpg  
Article: Casual Shoes (98.46%)  
Color: (123, 94, 94)  
Category: shoes

Track Pants  
97.7%



Casual Shoes  
98.5%



Predicting class for 8951 (1).jpg:



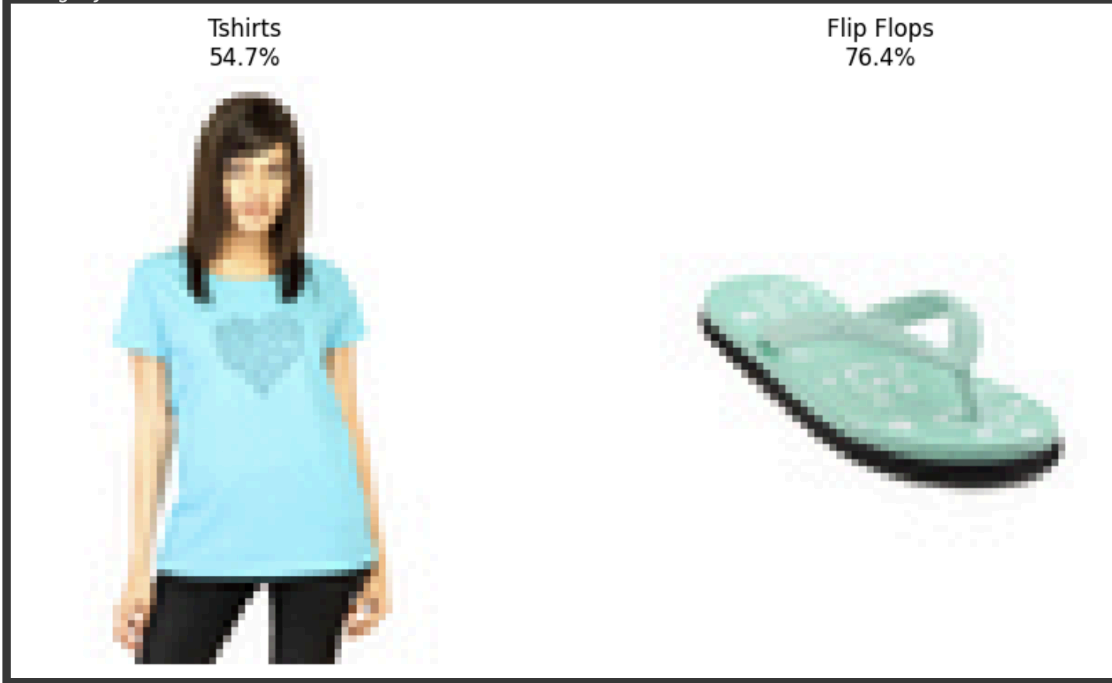
Top 3 predictions:

1. Track Pants – 69.85%
2. Trousers – 27.34%
3. Tshirts – 1.73%

## Recommended complementary pieces:

Filename: 12546.jpg  
Article: Tshirts (54.73%)  
Color: (145, 179, 174)  
Category: topwear

Filename: 19881.jpg  
Article: Flip Flops (76.43%)  
Color: (144, 184, 172)  
Category: shoes



## References:

- [1] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," Philosophical Transactions of the Royal Society A Mathematical Physical and Engineering Sciences, vol. 374, no. 2065, pp. 20150202–20150202, Mar. 2016, doi: <https://doi.org/10.1098/rsta.2015.0202>.