

Image Classification of Traditional Indian Clothing

Kelly Butler



Lehenga (the skirt)
with a blouse and
a dupatta

Nehru jacket



Women's mojaris



Saree



Men's kurta



Women's
kurta



Palazzos



Sherwani



Blouse



Gown



Dupatta



Salwar with
dupatta



Dhoti pants



Men's
mojaris

Business Problem

- Can a highly accurate model be created to classify images of traditional Indian clothing for a new e-commerce site in the U.S.?
- This site wants a model that will easily classify the Indian clothing since many of these clothes are quite similar.
- Most machine-learning image-classification models have been trained on western clothing.
- A model exists that does this classification (Rajput and Aneja). However, the images for this model were scraped from the web in 2020 or early 2021 (the paper does not specify) and may be out of date.

Data From Kaggle

15 Class Labels

7500 validation images

7500 test images

91166 training images

Class Label	Number of Images
Women’s Kurtas	11694
Sarees	10791
Blouses	9174
Leggings and Salwars	7787
Men’s Kurtas	6951
Dupattas	6587
Nehru jackets	6491
Lehenga	5753
Gowns	5211
Petticoats	4441
Dhoti Pants	4145
Palazzos	3375
Women’s Mojari	3228
Sherwanis	2992
Men’s Mojari	2546

Newly Scraped Test Data



500 images of
each type



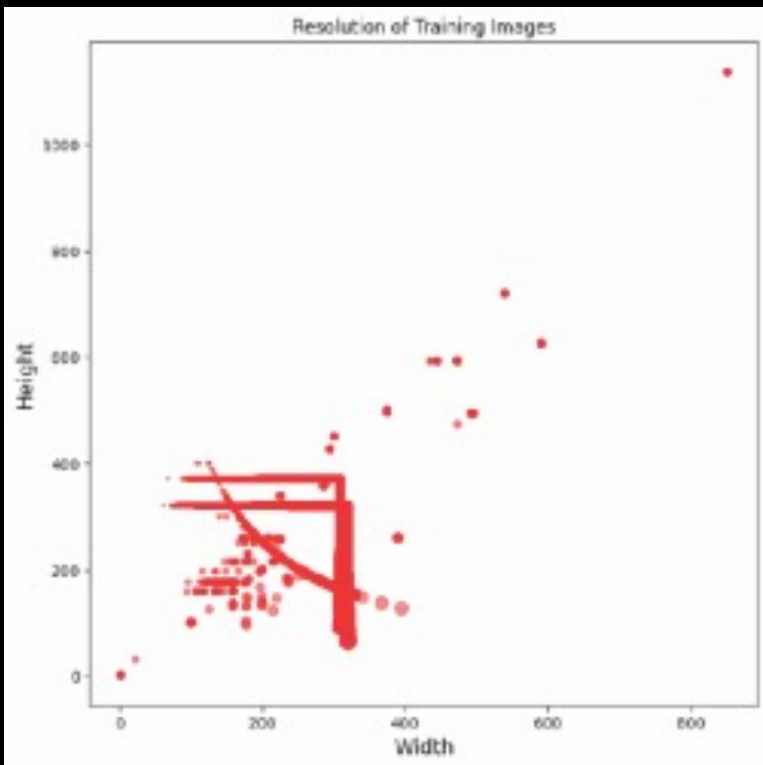


Image resolution in the training set

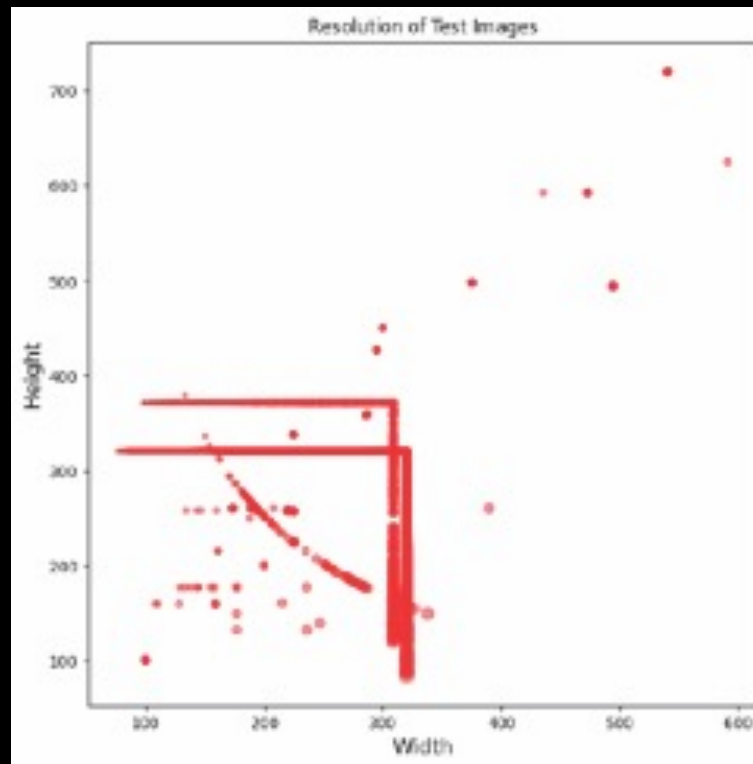


Image resolution in the test set

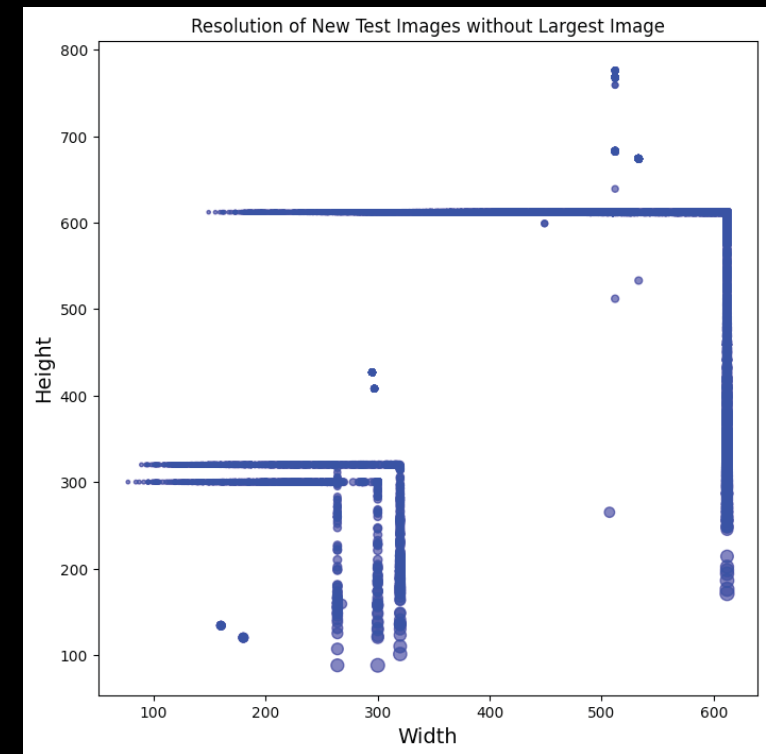
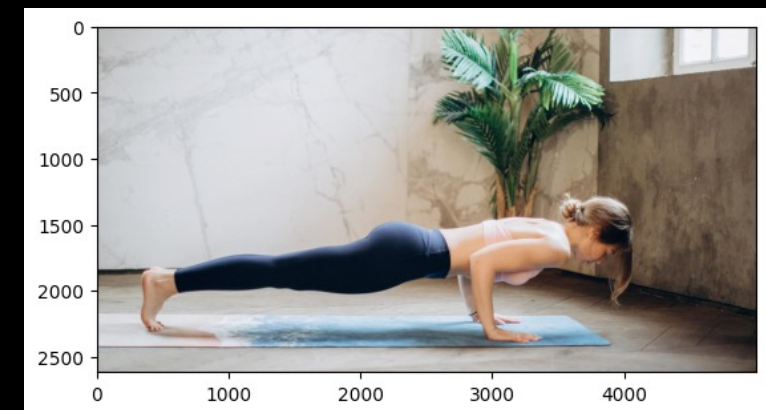
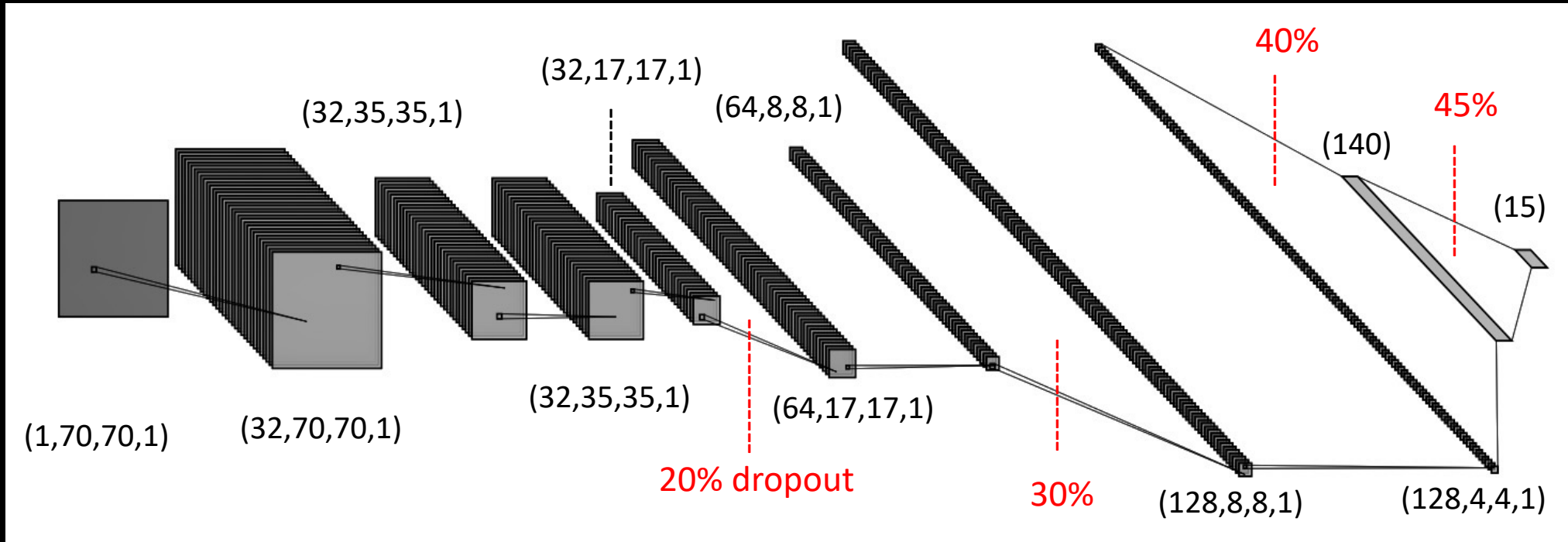


Image resolution in the new test set



Image Sizes





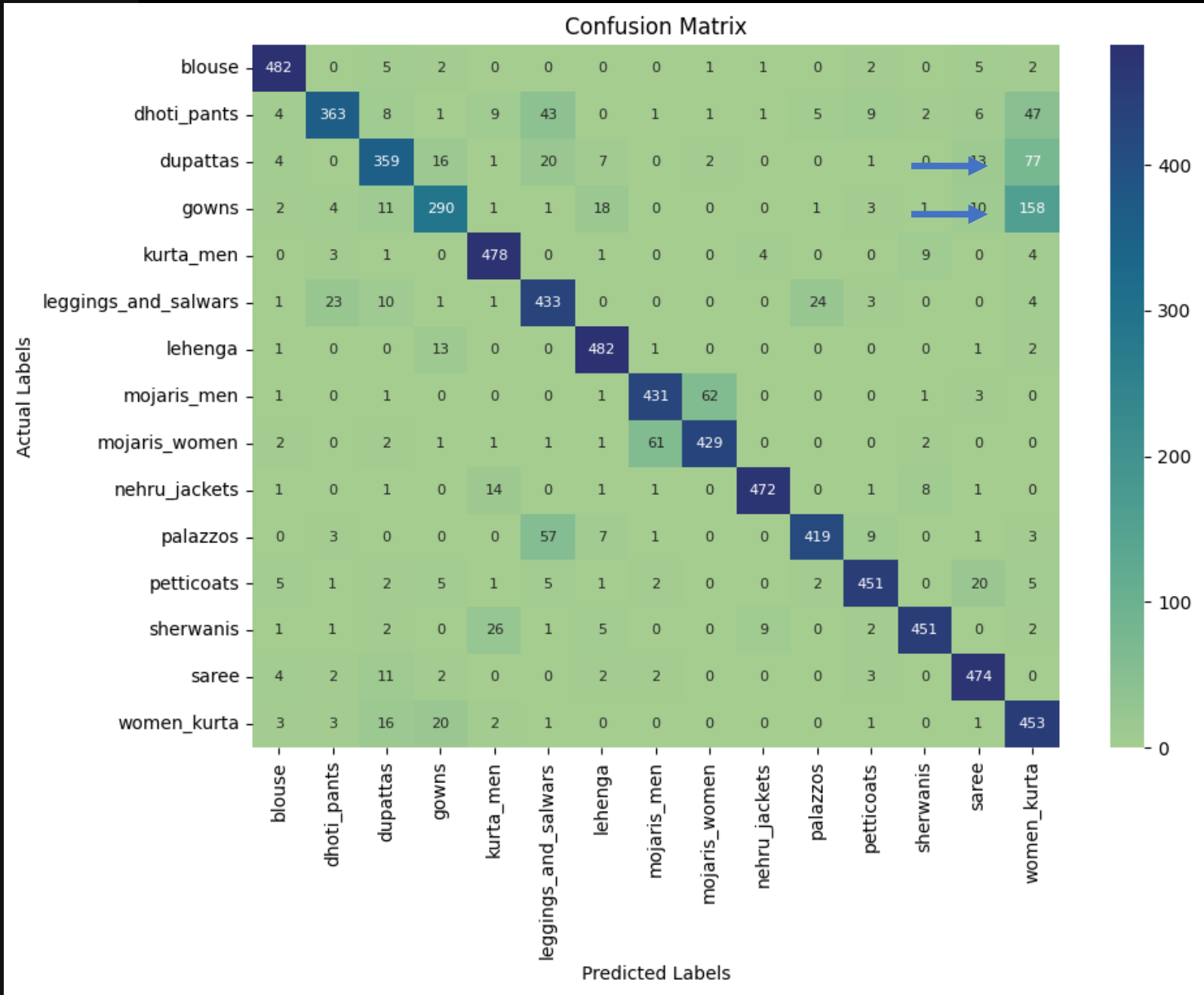
Convolutional Neural Network New Model

Metrics

With Original Test Set	Accuracy	Precision	Recall	F1 Score
ResNet 50	85.05 %	85.91 %	85.05 %	85.10 %
New Model	86.23 %	85.98 %	84.68 %	84.61 %
With New Test Set				
New Model	63.91 %	68.59 %	63.91 %	63.59 %

Confusion Matrix

Original Test Data



Comparing Training and Test Images

Training set:
Women's
Kurta



Test set:
Gown

Predicted label:
Women's Kurta

Comparing Training and Test Images

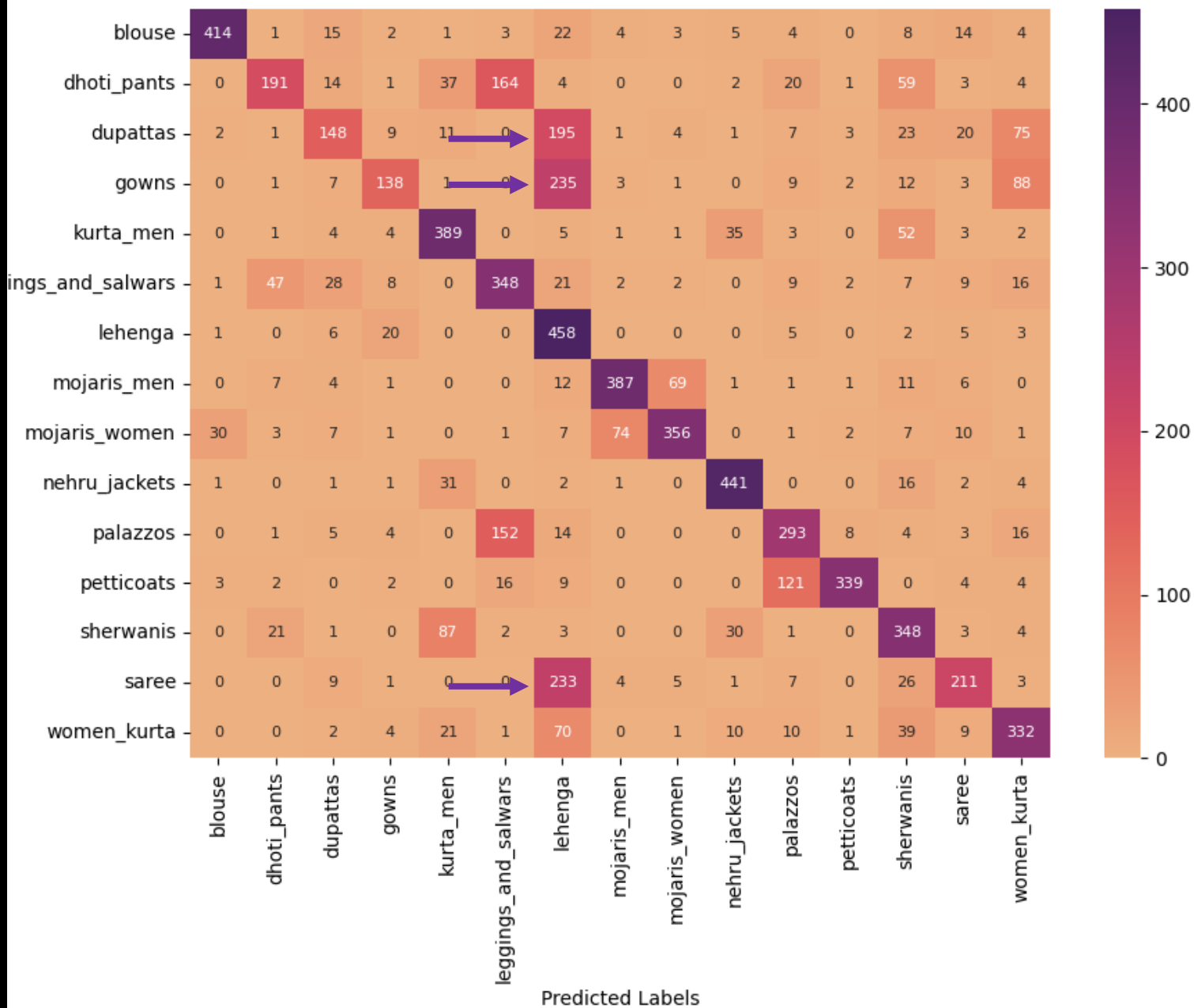


Training set
Women's Kurta



Test set
Dupatta
Predicted:
Women's Kurta

Confusion Matrix for New Test Data



Confusion Matrix Newly Scraped Test Data



Mis-labeled Images

Men's Mojari

Training Image
Number

56696

Product Title

Ladies Jutti Women Jutti Women Mojari Tan



56958

Womens Mojari



Business Insights

- Since the accuracy with newly acquired data was 63.91% significantly lower than the 86.23% for the original test set, a new model needs to be trained on current images from other e-commerce sites.
- It is imperative that the data used in training a new model be correctly categorized. The product title may be a better way of determining the category than the label used in searching for the image.
- Much of the change in images may not reflect a change in the fashion of clothing in each category but a rather a change in the way they are presented.
- Many clothes are shown in conjunction with one another. A multi-class multi-label model may be a better approach.