Image Classification of Traditional Indian Clothing

Scenario

A new e-commerce site in the U.S. wants to sell traditional Indian clothing. This business wants a model that will easily classify the Indian clothing since many of these clothes are quite similar. However, most machine-learning image-classification models have been trained on western clothing. A model exists that does this classification (Rajput and Aneja).[1] However, the images for this model were scraped from the web in 2020 or early 2021 (the paper does not specify). The question is whether this model will work well for images currently on e-commerce sites.

Data

To train the model, the Indo fashion dataset on Kaggle,[2] was used because these images were used in the previously discussed model. This dataset contains 91166 training images, 7500 validation images, and 7500 test images. The images are labeled as one of fifteen types of clothing: blouses, dhoti pants, dupattas, gowns, men's kurtas, women's kurtas, leggings and salwars, lehenga, men's mojaris, women's mojaris, Nehru jackets, palazzos, petticoats, sarees, and sherwanis. Pictures of each of these types of clothing are shown in the Appendix. Both the test and validation sets have 500 images of each class. The training set has the following distribution:

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 mojaris	3228
Sherwanis	2992
Men's mojaris	2546

[1] Rajput, P.S., Aneja, S., *IndoFashion : Apparel Classification for Indian Ethnic Clothes*, submitted 6 Apr 2021, https://arxiv.org/abs/2104.02830

[2] https://www.kaggle.com/datasets/validmodel/indo-fashion-dataset, accessed 10 Aug 2023

New images were scraped from the web using Beautiful Soup between 18-30 September 2023 from the following sites: Utsav Fashion (https://www.utsavfashion.com), KarmaPlace (https://www.karmaplace.com), Mojari (https://mojari.com.pk), Rajwadi (https://www.rajwadi.com), Flipkart (https://www.flipkart.com), and Amazon.in (https://www.amazon.in). For each category, the images were randomized and then the first 500 were used for the new test set. Most of these sites are ones from which the original data had been scraped, but I did not use as many different sites as Rajput and Aneja did.

Exploratory Data Analysis

There is a wide range of image sizes in the datasets. Graphs of height versus width for the training set, the original test set, and the new test set are in Figures 1, 3 and 4, respectively. While most heights and widths in the training set fall between 300 and 400 pixels, there are seven images which have a width greater than 800 pixels. An example of one of these is shown in Figure 2. The original test set has similar sizes to that of the training set except that it does not have any images with widths greater that 800 pixels. However, the new test set is quite different. It has one image with a width of 5000 and a height of 2617 (see Figure 5), which has been left off the graph in order to see the other sizes better. Also, while there are many images in the 300-to-400-pixel range, there are also many images in the 600-pixel range. The average size of the new test images is greater than those from the original data.

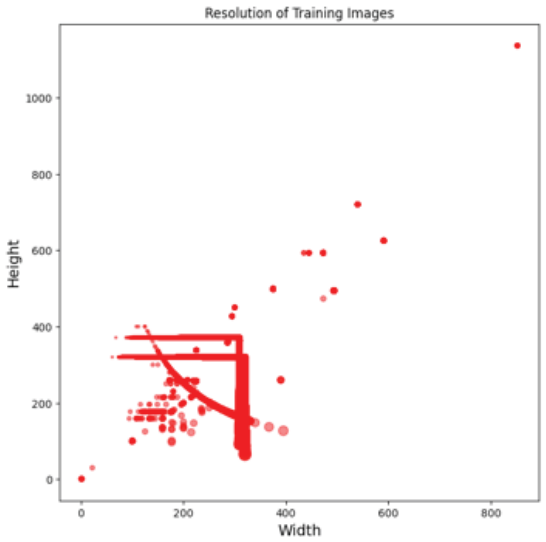

Figure 1. Image resolution in the training set.

Figure 2. One of the largest images in the training set.

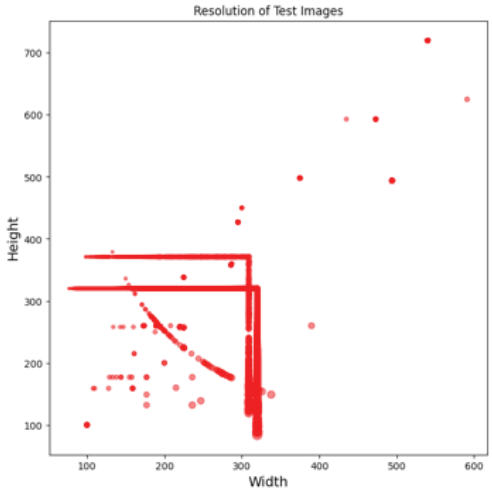

Figure 3. Image resolution in the test set.

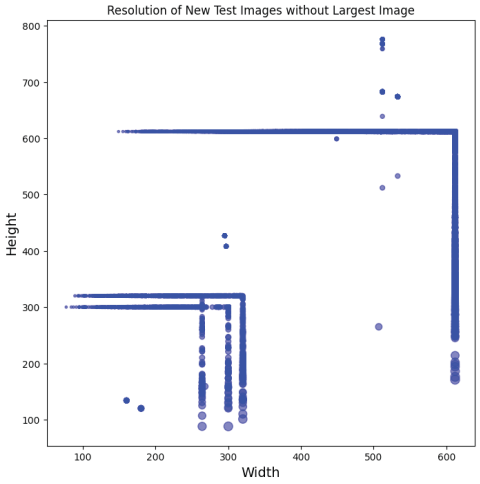

Figure 4. Image resolution in the new test set.

Figure 5. The image with the largest size from the new test set.

Methodology

A convolutional neural network (CNN) was trained to classify the images in the Indo fashion dataset using Tensorflow/Keras. The training data came from the training set in the Kaggle dataset. The model was evaluated on the test set from Kaggle and the new test set using accuracy, precision, recall, F1 score, and confusion matrices.

Modeling and Metrics

A number of CNN architectures were tried on different sizes of images and grayscale versus RBG; I found that 70-by-70-pixel grayscale images gave the best results. The model that gave the highest accuracy has two convolutional layers with 32 3 by 3 filters, one convolutional layer with 64 3 by 3 filters and then 0.2 dropout, the last convolutional layer with 128 3 by 3 filters followed by 0.4 dropout, a dense layer with 140 nodes and then 0.45 dropout. The loss was calculated as sparse categorical cross entropy with the Adam optimizer (see full model on next page). The images were resized to 70 by 70 pixels and converted to grayscale by Keras' image_dataset_from_directory.

Sequential()
Conv2D(32, (3,3), activation='relu', input_shape=(70, 70, 1), padding='same')
BatchNormalization()
MaxPooling2D(pool_size=(2, 2))

Conv2D(32, (3,3), activation='relu', padding='same')
BatchNormalization()
MaxPooling2D(pool_size=(2, 2))

Dropout(0.2)

Conv2D(64, (3,3), activation='relu', padding='same')
BatchNormalization()
MaxPooling2D(pool_size=(2, 2))

Dropout(0.3)

Conv2D(128, (3,3), activation='relu', padding='same')
BatchNormalization()
MaxPooling2D(pool_size=(2, 2))

Dropout(0.4)

Flatten()

Dense(140, activation='relu')

Dropout(0.45)

Dense(15, activation='softmax')
compile(loss='SparseCategoricalCrossentropy', optimizer='adam', metrics=['accuracy'])

Table 1. Convolutional neural network that gave the best accuracy.

The performance metrics for this model can be found in Table 2. While accuracy is a reasonable metric given that the data is not imbalanced, I have included the macro values for precision, recall, and the F1 score, all of which are close to the value of the accuracy, as expected. The accuracy of this model is 86.23% using the test data from Kaggle and 63.91% using the new test set. I would have liked to get a higher accuracy; however, Rajput and Aneja started with a ILSVRC 2012-pre-trained ResNet-50 backbone and got an accuracy of 88.43%, which is not much better than mine. Also, my model was trained only on the training data they provided.

	Accuracy	Precision	Recall	F1 Score
Original test dataset	86.23 %	85.98 %	84.68 %	84.61 %
New test dataset	63.91 %	68.59 %	63.91 %	63.59 %

Table 2. Performance metrics for the CNN model.

While 86.23% accuracy good, I wanted to understand why multiple models got stuck around that range. After more exploration of the data, I found error or confusions with the data. An analysis of the product titles for the images in the data from Kaggle shows that some images are mislabeled, and others have product titles that include more than one label. Some examples from the training data: an image labeled as a gown has a product title of "Women's Cotton Green & Red Geometric Print Saree with Blouse Piece (Sarita-510_Green & Red_Free Size)"; the label of dhoti pants is given to

an image that has the product title of "Men's Nehru Jacket (VASMJ011_44_Multicolored_44)"; and an image labeled as a saree has the description, "Women's Net Semi-stitched Lehenga Choli (67-IM5V-TOTL_Multicolored_Free Size)". All of these labels are simply incorrect.

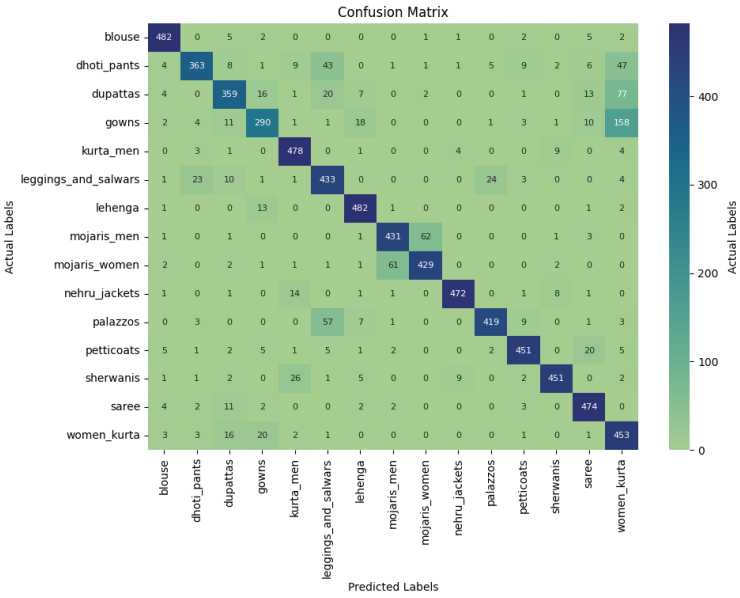

Figure 6. Confusion matrix for original test set.

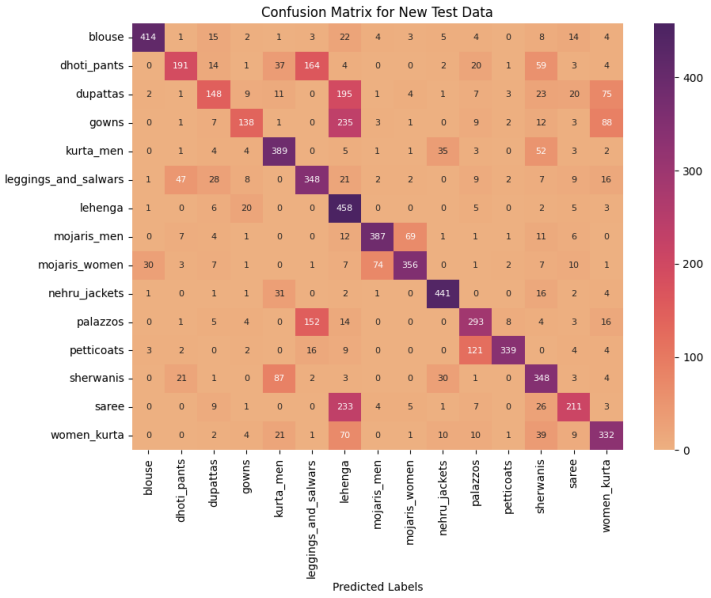

Figure 7. Confusion matrix for new test set.

The confusion matrices are of particular interest given that it is the only real way to check what the model is doing. Figure 6 shows this for the original test data and Figure 7 for the new test set. Note that the biggest misclassification in Figure 6 is gowns and dupattas labeled as women's kurtas. Again, the data is at least partially at fault here. Figure 8 shows how an image from the training set labeled a gown and an image from the test set labeled a women's kurta look alike. A kurta is a top that can be long but is worn with salwars or leggings (pants), while a gown is an

Figure 8. A women's kurta from the training set on the left and on the right a gown from the test set incorrectly labeled a women's kurta.

outfit on its own. Similarly, Figure 9 shows an image from the training set labeled a dupatta and an image from the test set labeled a women's kurta. A dupatta is a long scarf that is normally worn with a kurta. Both images show a woman wearing a kurta and dupatta. The other misclassifications occur also between similar items of clothing or items of clothing that are worn together.

Figure 9. A women's kurta from the training set on the left and on the right a dupatta from the test set incorrectly labeled a women's kurta.

While the model predicts the new test set with much lower accuracy, the new test set's confusion matrix is interesting in that the items most confused with one another are not the same as those for the original test set. Gowns, dupattas, and sarees are mislabeled more often as lehengas, rather than gowns and dupattas being confused for women's kurtas as in the original test set.

Actionable 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 would need 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; Figure 5 is a good example of this, having a model wearing leggings posed very differently than in the original data.
- Many clothes are shown in conjunction with one another. An example is a man wearing a sherwani and dhoti pants. That image should be presented to customers when they search for a sherwani or dhoti pants. However, since sherwanis are only for men, the dhoti pants should be categorized as men's dhoti pants. This model did not distinguish between men's and women's for these.

Appendix

Lehenga (the skirt) with a blouse and a dupatta

Saree

Gown

Dupatta

Blouse

Women's Mojaris

Men's Mojaris

Nehru Jacket

Women's Kurta

Men's Kurta

Sherwani

Petticoat

Salwar (with dupatta)

Dhoti Pants

Palazzos