Experimenting with IndiaFashion: Indian Ethnic Apparel Classification model

Sachin Patel
MSc in Computing (Data Analytics)
Dublin City University
sachin.patel3@mail.dcu.ie

Sneha Yadav
MSc in Computing (Data Analytics)
Dublin City University
sneha.yadav4@mail.dcu.ie

Abstract – Cloth categorization is a crucial challenge faced by e-commerce websites, where efficient and accurate classification of products enhances user experience and boosts sales. Among various clothing styles, Indian Ethnic apparel holds a unique cultural significance and is exclusively worn by the Indian population. With the rapid growth of the e-commerce sector and technological advancements in recent decades, the demand for tailored image classification models to categorize Indian Ethnic apparel has surged. This paper presents an in-depth exploration and benchmarking of an Indian Ethnic apparel classification model. The authors extended the IndoFashion dataset by introducing additional classes and demonstrated the training and evaluation of models on the augmented dataset. The study showcases the performance of the best model, achieving an impressive accuracy of 88.33%. The outcomes of this research hold considerable significance, not only for e-commerce platforms but also for the advancement of image classification models in various applications. The ability to accurately categorize Indian Ethnic Apparel can significantly enhance product recommendations and personalized shopping experiences for users. Moreover, the insights gained from this study can serve as a valuable resource for further development and refinement of image classification models for diverse domains.

***Keywords*: classification, image classification, clothes classification, transfer learning for clothes classification, Indian ethnic wear classification, apparel recognition, fashion categorization**

1. Introduction

Apparel plays a significant role in our daily lives, serving as a means to identify individuals based on their dressing styles, reflecting diverse arts and cultural influences. With India's vast population of over 1.3 billion, the country boasts a rich variety of formal clothing options, each influenced by factors such as regional climates. However, when individuals seek to purchase formal apparel from retail or e-commerce stores, they often encounter categorizations primarily focused on Western clothing styles. This disparity highlights the pressing need to incorporate a comprehensive categorization system for Indian Ethnic apparel within these stores.

In an effort to enhance the user experience, e-commerce stores strive to offer a well-designed product classification system that considers the perspective of customers, leading to a streamlined buying experience. Fashion e-commerce websites, in particular, have witnessed a surge in popularity for online clothing shopping, making it imperative for e-commerce companies to effectively categorize their extensive range of clothing products. While datasets like FashionMNIST and DeepFashion [7] have been developed in recent years for clothing categorization, they predominantly focus on Western apparel, such as shirts, jeans, jackets, and shorts. Such databases fail to address the substantial stylistic differences between traditional Indian clothing, including Kurtis, sarees, kurtas, lehengas, dhotis, and Western clothing.

The primary objective of this paper is to train a model using a more extensive dataset by augmenting the existing IndoFashion dataset with an additional class dedicated to Indian Ethnic apparel which is not included in the existing dataset. The classification of Indian ethnic apparel poses a significant challenge due to its diverse range of styles, colours, and patterns. In addressing this challenge, Convolutional Neural Networks (CNNs) have demonstrated promising results in image classification tasks. Consequently, there has been a growing interest in developing an image classification model for Indian ethnic apparel classification.

Within the fashion industry, accurate classification of apparel holds particular importance as it can significantly improve search, recommendation, and inventory management systems. This paper delves into the experimenting with the models used for Indian ethnic apparel classification, exploring various approaches and techniques employed for benchmarking and evaluating the performance of these models. By studying the effectiveness of these models, we aim to pave the way for improved clothing categorization, leading to enhanced user experiences and more tailored e-commerce offerings.

2. Related Work

The datasets Atlas [2], Generative Indian Fashion, and IndoFashion[1] are among the few that include Indian ethnic clothes. The Atlas[2] dataset, which focuses on classifying fabrics, is inadequate for deep model training because of its high imbalance and absence of more than 40 pictures for several class categories. Generative Indian Fashion [3], which likewise has a few (12K) images and class divisions, does not address the problem of textile classification. While the IndoFashion dataset is a sizable collection of 106K photos from several e-commerce websites that represent 15 different categories of clothing, this research suggests the first automated technique to classify Indian ethnic clothing.

The IndoFashion dataset, which includes 106K photos of 15 categories covering the most popular Indian ethnic garments, was created by Shivangi Aneja [1] et al. The authors evaluated the existing datasets for ethnic clothing and discovered that there were no large-scale datasets available for ethnic datasets. In the study, the authors of the IndoFashion dataset picked ResNet models which were trained on Imagenet (ILSVRC 2012) for developing a classification model on the IndoFashion dataset. In the study, the sample images in the dataset were normalised to 128x256 pixels before providing it to the network and enhancing the images to minimise overfitting and model performance. In the study, It was found that applying geometrical and colour augmentation alone or in combination increased the model's performance[1]. According to this paper, It was also discovered that highly large and sophisticated models like resnet-101 were not required for the dataset because the model that performed the best was resnet-50, which had an accuracy of 88.93%[1]. The Adam optimizer was used in the study with a learning rate of 1e-3, decay when validation loss reaches a plateau for 5 consecutive epochs, and early termination after 25 consecutive validation loss epochs. For the small ResNet-18 model, for example, training with the entire dataset (91K photos) increased classification accuracy by an absolute 10.33% (from 77.19% to 87.52%) compared to training with only 10% of the dataset (9K images), taking advantage of the dataset's significant diversity. This indicates that the model's performance increases significantly when the training data is large. Even yet, there were several categories where the model struggled to pick the right one. For instance, because the samples for women's kurtas and women's gowns were so similar, the model got them mixed up. It was concluded that by including an equal number of photographs in each category, this can be avoided and the dataset's quality raised.

The Atlas product categorization dataset, created by Venkatesh Umaashankar [2], primarily focuses on 52 apparel articles and includes information about their name, price, image, and category path. The taxonomy in the Atlas dataset mostly focuses on Western clothing and a little portion of Indian ethnic clothing. The taxonomy tree employed has a maximum depth of three tiers and a total of 52 category paths. The product title, image, and price of the products were extracted from data from Indian E-commerce stores using the web scraping tools Scrapy and selenium, and the data was then put in JSON format.

There were both regular and zoomed photos, and the CNN model was used to remove the noisy images. They employed an image classification model based on Resnet34. Additionally, they generated sequences using an attention-based encoder-decoder neural network architecture. They employ a convolution neural network (CNN) in encoder to produce fixed-size vectors. The photos were reduced in size to a single dimension before being input into the encoder, which employs 101 layered Residual Networks that have already been trained to do the ImageNet classification task. They combined the Attention in the Decoder with the LSTM Network of the RNN. LinearSVM, CNN, Resnet34-based image classification, and Attention-based Seq to Seq models were used for the training, and the f-scores of the benchmark models were compared. Although the classification model performs better than Seq to Seq, Seq to Seq can do better as the number of categories increases. It was found that none of the more recent category pathways produced by the Seq-to-Seq model were always valid. Only two of the five new category routes that their model produced in the example—which are shown in Table 2—were found to be valid. Thus, from newly constructed category paths, the category pathways that could be used to enrich the taxonomy must be manually selected out. Therefore, it is appropriate to go with the classification model.

DeepFashion2[4], proposed by Liu et al. [1], is an extensive fashion dataset that addresses the limitations of its predecessor, DeepFashion[5]. It features over 300,000 images spanning 13 categories and 7 major attribute groups, catering primarily to western clothing. The dataset's comprehensive annotations and fine-grained labels have made it a benchmark for fashion-related tasks such as instance-level retrieval and attribute prediction.

The motivation to utilize the IndoFashion dataset for further study lies in its unique focus on Indian ethnic apparel. While both Atlas and DeepFashion2 provide valuable insights into the fashion domain, they are primarily tailored towards western clothing. In contrast, IndoFashion caters specifically to the intricacies of Indian ethnic wear, encompassing draping styles, regional design preferences, and diverse fabric choices.

In this study, we are using the IndoFashion dataset and the original CNN models, ResNet18, ResNet50, and ResNet101, to investigate the impact of incorporating additional classes into the dataset. The IndoFashion dataset, originally introduced by Shivangi et al., comprises a diverse collection of Indian ethnic apparel images. We have expanded this dataset by adding three additional classes. Our aim is to examine how the introduction of these new classes affects the classification performance of the selected models. By leveraging the same CNN models used in the original IndoFashion research, we seek to analyze whether the models demonstrate improved accuracy, robustness, or encounter any potential challenges in handling the augmented dataset. This analysis will offer valuable insights into how the existing models perform when faced with an enriched fashion taxonomy that includes new Indian ethnic clothing categories. Ultimately, our study aims to contribute to the advancement of research in the domain of Indian ethnic wear classification and its practical applications in the fashion industry.

3. Dataset

The IndoFashion dataset is an open-source dataset available on GitHub. The authors of the IndoFashion dataset used Selenium and BeautifulSoup for collecting images from various E-commerce websites that feature a wide range of Indian Ethnic Apparels. While the authors initially selected 15 major classes, the dataset has the potential for further expansion by incorporating additional classes. As a result in this study, it was decided to augment the existing IndoFashion dataset with image samples from one unique class, "pagdi," and two subclasses (Anarkali kurta and A-Line Kurta) derived from the existing class "Women kurta." This augmentation aims to explore how the inclusion of one new class and two sub-classes impacts the performance of the models trained on the IndoFashion dataset.

3.1 Dataset Collection - In the Data Collection section, image samples for the three new classes were acquired using automated tools, specifically Selenium and Chrome Webdriver. These tools facilitated the collection of images from various online platforms, including Flipkart[6], Amazon[7], Myntra[8], Utsav Fashion[9], Bodyline[10], and Google[11], among others.
However, it is important to note that the gathered image samples contained some irrelevant images that did not belong to the specified categories. As a result, a manual effort was necessary to curate the dataset and remove the irrelevant image samples, ensuring the dataset's accuracy and relevance for the classification task. This manual cleaning process was essential to maintain the dataset's integrity and quality for subsequent model training and evaluation.

3.2 Dataset Statistics - The IndoFashion dataset comprises a total of 110,000 images, representing 18 unique cloth categories for the cloth classification task. To ensure a robust evaluation, we meticulously split the dataset into three subsets: training, validation, and test sets. To maintain fairness in our evaluation, we carefully distributed the classes equally among the validation and test sets.
In our efforts to enhance the dataset, we conducted data augmentation techniques, resulting in an extended dataset that now includes a comprehensive representation of 18 distinct cloth categories. These categories are outlined in Table

1, providing a detailed overview of the diversity within the dataset.

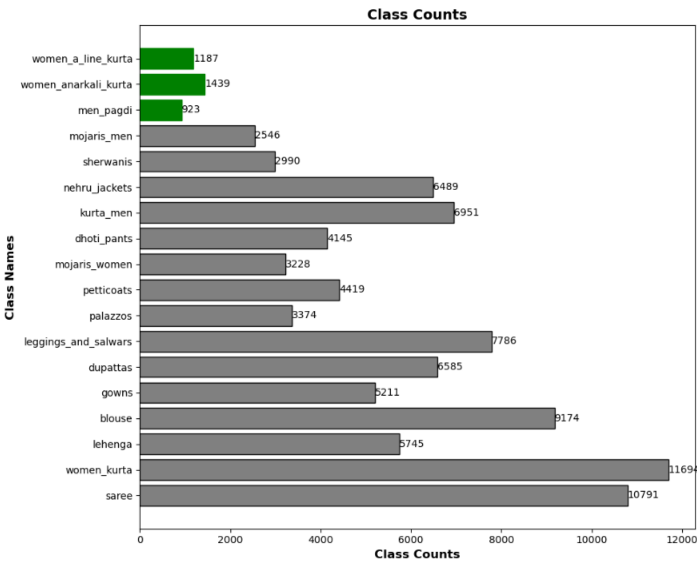
Gender	Categories
Women	Saree, Women Kurta, Leggings & Salwar, Palazzo, Lehenga, Dupatta, Blouse, Gown, Dhoti Pants, Petticoats, Women Mojari, Women Anarkali Kurta, Women A-Line Kurta
Men	Men Kurta, Nehru Jacket, Sherwani, Men Mojari, Men Pagdi

Table 1. The table contains the class of both men and women Ethnic wear in the dataset with an additional 3 classes[1].

To aid researchers in replicating and conducting further evaluations, we provide a detailed dataset split, as illustrated in Table 2. This transparently demonstrates how the dataset was partitioned into training, validation, and test sets to facilitate fair and robust evaluations of our proposed models and the existing CNN models, ResNet18, ResNet50, and ResNet101, as mentioned in our earlier sections.

Split	No. of Images
Train	94677
Validation	7809
Test	7808

Table 2. Statistics of the augmented IndoFashion dataset [1]



4. Experiments

We conducted an assessment of various state-of-the-art Image Classification models to fulfil our objective of classifying Indian ethnic apparel, following the approach adopted in the IndoFashion research. For our models, we utilised a ResNet backbone (ResNet-18, ResNet-50, and ResNet-101) that was pre-trained on the ILSVRC 2012 [12] dataset. The reported results are based on evaluations performed on the test subset of our dataset. To gauge the performance of our models, we employed Precision, Recall, and F1-Score as the evaluation metrics.

4.1 Preprocessing and Augmentation - In this section, we describe the preprocessing and augmentations applied to the image samples before feeding them into the CNN models. To ensure consistency in input size and reduce computational complexity, all the image samples were normalised to a fixed size of 128x256 pixels. This resizing ensures that all images have the same dimensions, facilitating efficient training of the CNN models on a larger dataset without imposing excessive computational demands. By maintaining a standardised input size, the training process becomes more streamlined, enabling the models to leverage ample computational resources effectively.

In the IndoFashion paper, various types of image augmentations were applied during the training process to address overfitting and enhance model performance. These augmentations were not applied during inference to ensure unbiased evaluations. The augmentations used in the training phase are outlined below:

1. Geometric Transformation: To introduce diversity in the training data, the images underwent geometric transformations such as horizontal flipping and random rotations by a small angle (10 degrees) with a probability of 50%. This approach allows the model to learn from variations in the image orientation and position, making it more robust.

2. Color Transformation: Color jitter was applied to the images by modifying the hue and saturation using a factor of 0.2. This color transformation enhances the model's ability to recognize objects under different lighting conditions and color variations, thereby improving its generalization capabilities.

By employing these augmentations during training, the study aimed to assess the impact of the additional classes on the accuracy of the ResNet models and ascertain how well the models adapt to the enriched dataset with newly introduced Indian ethnic apparel categories.

4.2 Hyperparameters - Experimental Setup For our experiments, we conducted the trials on Google Colab Pro + using a Hardware Accelerator GPU featuring the NVIDIA A100 Tensor Core GPU. During training, we employed the Adam optimizer with a learning rate of 1e-3. To prevent overfitting, we implemented a learning rate decay strategy when the validation loss showed no improvement for 5 consecutive epochs. Additionally, we applied early stopping with a patience of 25 successive epochs based on the validation loss to halt training if no significant improvement was observed.

5. Results and Evaluation

In the study, we utilised pre-trained ResNet18, ResNet50, and ResNet101 models, originally trained on the ILSVRC 2012 dataset. These models were trained on our customised dataset containing 18 classes. Interestingly, we observed that the less complex ResNet18 model outperformed the models trained on the IndoFashion Dataset, which had 15 classes.

Furthermore, we conducted experiments with the ResNet50 model, both with and without image augmentation, and found that the F1 scores were comparable between the two approaches. Additionally, we explored the more complex ResNet101 model, trained on both the modified IndoFashion dataset with 18 classes and the original IndoFashion dataset. Surprisingly, these two ResNet101 models demonstrated similar F1 scores despite the complexity difference.

Models Trained on 18 Classes of IndoFashion Dataset			
Model	Precision	Recall	F1-Score
R-18(NA)	87.58	86.83	86.92
R-18(F)	88.99	87.81	87.81
R-18(J)	87.22	86.14	86.17
R-18(J+F)	88.49	87.45	87.39
R-50(NA)	87.41	86.48	86.37
R-50(F)	88.78	88.08	88.03
R-50(J)	87.17	86.35	86.40
R-50(J+F)	89.11	88.32	88.33
R-101(NA)	88.06	87.03	87.01
R-101(F)	88.18	87.22	87.27
R-101(J)	87.21	86.33	86.29
R-101(J+F)	88.82	88.22	88.23

Table illustrates the impact of employing various augmentations during training. To maintain clarity, we use the following notations: R for ResNet, (J) for Jitter augmentation, (F) for Flipping, and (J+F) for Jitter and Flipping combined, and (NA) for No Augmentation.

The table provided showcases the detailed performance metrics for each model under various augmentation settings. Specifically, it presents the precision, recall, and F1-score for each ResNet model with different augmentation combinations, as well as the case with no augmentation. Notably, the augmentation techniques of Jitter and Flipping, either individually or combined, generally improved the model's performance across all ResNet architectures.

By adding the three classes (Men Pagdi, Women Anarkali Kurta, and Women A-Line Kurta) to the existing IndoFashion dataset, several effects and implications can be inferred:

Test Statistics for 3 specific classes (added)			
Model	Recall	F1-score	Error Rate
Men Pagdi			
R-50(J+F)	85.19	92	14.81
Women Anarkali Kurta (sub-class)			
R-50(J+F)	96.06	97.99	3.94
Women A-Line Kurta (sub-class)			
R-50(J+F)	75.96	86.34	24.04
Women Kurta			
R-50(J+F)	88.4	93.84	11.6

Improved Classification Performance for Unique Class "Men Pagdi" - The model achieved a relatively high F1-score of 92% for the unique class "Men Pagdi," indicating that it could accurately classify this new class. The higher F1-score suggests that the model was able to effectively distinguish "Men Pagdi" from other classes, demonstrating the dataset's ability to generalize to new and distinct categories.

High F1-Score for Subclass "Women Anarkali Kurta" - The model achieved an impressive F1-score of 97.99% for the subclass "Women Anarkali

Kurta." This high F1-score indicates that the model accurately classified the subclass, which is a subset of the existing class "Women Kurta." The strong performance suggests that the additional subclass contributed valuable information to the model, enhancing its ability to recognize fine-grained variations within the "Women Kurta" category.

Challenges in Classifying Subclass "Women A-Line Kurta": The model obtained a relatively lower F1-score of 86.34% for the subclass "Women A-Line Kurta." This lower score may be attributed to the limited number of samples available for this subclass compared to other classes in the dataset. The smaller sample size could lead to challenges in effectively learning the distinctive features of this subclass, resulting in a slightly lower classification performance.

Impact on Existing Class "Women Kurta" - The F1-score for the existing class "Women Kurta" remained relatively high at 93.84%. The addition of new subclasses did not negatively impact the model's ability to classify the original "Women Kurta" category. Instead, it suggests that the model was able to adapt to the inclusion of more granular subcategories without compromising its performance on the broader class.

Model Generalization and Data Imbalance: The model exhibited strong generalization capabilities for Men Pagdi and Women Anarkali Kurta, as indicated by their high F1-scores and recall values, even with the introduction of new subclasses. However, the F1-score for Women A-Line Kurta was relatively lower, suggesting the need for more data samples to enhance the model's performance and generalization for this specific subclass. The dataset's inherent class imbalance, evident from the class counts, could present challenges for classes with limited samples, such as Women A-Line Kurta. To address this issue and improve overall model performance, it is essential to balance the dataset by ensuring sufficient samples for each class, thus providing the model with a more representative and diverse training set.

In our study, we encountered an imbalanced dataset, which presented a challenge in effectively training our models. To address this issue, we adopted two techniques to handle the class imbalance problem. Firstly, we computed class weights for each class based on the number of samples they contained. Classes with fewer samples were assigned higher weights, while classes with more samples received lower weights. This approach ensured that the loss function gave more importance to underrepresented classes during training.

Secondly, we incorporated the computed class weights into the Cross Entropy Loss function (nn.CrossEntropyLoss) during model training. By doing so, we effectively assigned higher importance to the samples from underrepresented classes, helping to mitigate the impact of the dataset's imbalance on the learning process.

Upon evaluating the performance of various model architectures, we observed that the ResNet-50 model exhibited the best results among all the models we tested. Consequently, we selected this model as our base architecture to explore the impact of additional data augmentation techniques on the modified IndoFashion Dataset containing 18 classes.

To augment the dataset, we implemented several data augmentation techniques not present in the original models. These techniques included Random Shear, Random Brightness, Shear and Brightness combination, Translate and Rotate, as well as Jitter and Flip. For each augmentation, we employed the previously computed class weights and the Weighted Loss Function.
Upon evaluating the five models with different augmentation techniques, we obtained the following results:

Models trained on 18 classes of IndoFashion Dataset

Model	Precision	Recall	F1-Score	Error Rate
R-50(S)	87.05	86.58	86.57	13.42
R-50(B)	87.42	86.76	86.79	13.24
R-50(B+S)	88.00	87.73	87.72	12.27
R-50(T+R)	88.32	88.26	88.25	11.74
R-50(J+F)	87.90	87.62	87.61	12.38

Notably, the R-50(T+R) model, which involved the Translate and Rotate augmentation, yielded the best performance compared to other models and demonstrated a performance level similar to the best-performing model of the original ResNet50 model trained on the IndoFashion Dataset.

These results highlight the effectiveness of our approach in handling the dataset's class imbalance problem and the significant impact of data augmentation techniques in enhancing the models' generalization ability. By utilizing these findings, researchers and practitioners can design more robust and accurate models for image classification tasks, even with challenging and imbalanced datasets.

6. Conclusions

In this research, we explored the Indian Ethnic Apparel Classification model using the IndoFashion dataset. By augmenting the dataset with three new classes, including one unique class and two subclasses, we evaluated the impact on model performance. Our findings revealed that the addition of these classes improved the model's ability to recognize new and diverse categories, resulting in impressive F1-scores for the unique class "Men Pagdi" and the subclass "Women Anarkali Kurta." However, challenges were observed in classifying the subclass "Women A-Line Kurta," emphasizing the need for a larger sample size for more accurate representation. By applying data augmentation techniques, we effectively handled the dataset's class imbalance and enhanced the generalization capabilities of our models. The ResNet-50 model, augmented with translation and rotation, yielded the best performance, indicating the potential for robust image classification models for Indian ethnic apparel.

7. Future work:

In future research, we plan to address the challenges presented by the subclass "Women A-Line Kurta" by acquiring more data samples to improve model performance and generalization. Additionally, we will explore more sophisticated data augmentation techniques and investigate the effectiveness of transfer learning from pre-trained models on larger datasets. Furthermore, we intend to extend the IndoFashion dataset to incorporate even more diverse classes, including regional and traditional Indian ethnic apparel, to further enhance the model's ability to recognize a broader range of clothing categories. Lastly, we aim to explore advanced techniques such as domain adaptation and ensemble learning to improve model robustness across different sources of Indian ethnic apparel images. Overall, our research lays the foundation for the development of accurate and efficient image classification models for the rich and diverse world of Indian ethnic apparel.

8. References

[1] P. S. Rajput and S. Aneja, “IndoFashion : Apparel Classification for Indian Ethnic Clothes.” arXiv, Apr. 06, 2021. Accessed: Oct. 22, 2022. [Online]. Available: http://arxiv.org/abs/2104.02830

[2] V. Umaashankar, G. S. S, and A. Prakash, “Atlas: A Dataset and Benchmark for E-commerce Clothing Product Categorization.” arXiv, Aug. 12, 2019. Accessed: Oct. 22, 2022. [Online]. Available: http://arxiv.org/abs/1908.08984

[3] H. Jain *et al.*, “Generative Fashion for Indian Clothing,” in *8th ACM IKDD CODS and 26th COMAD*, Bangalore India: ACM, Jan. 2021, pp. 415–415. doi: 10.1145/3430984.3431057.

[4] Y. Ge, R. Zhang, L. Wu, X. Wang, X. Tang, and P. Luo, “DeepFashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images.” arXiv, Jan. 23, 2019. Accessed: Feb. 21, 2023. [Online]. Available: http://arxiv.org/abs/1901.07973

[5] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, “DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 1096–1104. doi: 10.1109/CVPR.2016.124.

[6] “Online Shopping India Mobile, Cameras, Lifestyle & more Online @ Flipkart.com.” https://www.flipkart.com (accessed Jul. 31, 2023).

[7] “Online Shopping site in India: Shop Online for Mobiles, Books, Watches, Shoes and More - Amazon.in.” https://www.amazon.in/ (accessed Jul. 31, 2023).

[8] “Online Shopping India - Shop Online for Branded Shoes, Clothing & Accessories in India | Myntra.com,” *Myntra*. https://www.myntra.com/ (accessed Jul. 31, 2023).

[9] “Wedding bridal sarees, Designer lehengas sarees, Salwar kameez, Embroidered sarees salwar kameez, Bridal wedding lehengas, Traditional sarees lehngas from India,” *Utsav Fashion*. https://www.utsavfashion.com (accessed Jul. 31, 2023).

[10] “Buy Online Sherwani, Wedding Kurta Pajama, Kurta for Men, Nehru Jacket, Indo Western, Suits & Blazers.” https://www.bodylinestore.com/ (accessed Jul. 31, 2023).

[11] “Google Images.” https://images.google.com/ (accessed Jul. 31, 2023).

[12] “ImageNet.” https://www.image-net.org/challenges/LSVRC/2012/ (accessed Jul. 31, 2023).