AFRIFASHION1600: A Contemporary African Fashion Dataset for Computer Vision

Wuraola Fisayo Oyewusi Olubayo Adekanmbi Sharon Ibejih

Opeyemi Osakuade Ifeoma Okoh

Mary Salami

Data Science Nigeria, Lagos,Nigeria

{wuraola,olubayo,sharon,osakuade,ifeoma,mary}@datasciencenigeria.ai

Abstract

This work presents AFRIFASHION1600, an openly accessible contemporary African fashion image dataset containing 1600 samples labelled into 8 classes representing some African fashion styles. Each sample is coloured and has an image size of 128 x 128. This is a niche dataset that aims to improve visibility, inclusion, and familiarity of African fashion in computer vision tasks.AFRIFASHION1600 dataset is available here.

1. Introduction

Fashion is everywhere. It is one of the main ways that humans present themselves to others. It signals what individuals want to communicate about their sexuality, wealth, professionalism, subcultural and political allegiances [4]. African fashion is as diverse and dynamic as the continent and the people who live there. Contemporary African Fashion puts Africa at the intersection of world cultures and globalized identities, displaying the powerful creative force and impact of newly emerging styles [6]. The rise of digital technologies and techniques has improved the way that fashion is curated, analyzed and preserved in context. As the revolution of computer vision with artificial intelligence (AI) is underway, AI is enabling a wide range of application innovations from electronic retailing, personalized stylist, to the fashion design process. Some use cases for computer vision in fashion include specific tasks but are not limited to image synthesis, detection, analysis, and recommendation [1].

1.1. Fashion Datasets for Computer Vision

The development of fashion datasets has fueled advances in clothing recognition. In 2017, Han Xiao, et al. [14] of Zalando (Europe's largest online fashion platform) introduced the Fashion-MNIST dataset, a popular built-in image dataset for deep learning. The goal was to present a dataset that was more challenging to classify compared to the original MNIST which was introduced by LeChan et al in 1988 [8]. Fashion MNIST has 70000 images categorized into 10 classes which later showed more complexity than the original MNIST. Yamaguchi et al.[15] in 2012, provided a novel dataset with over 138000 images obtained from Chictopia, which were used to improve pose identification and demonstrate a prototype application for pose-independent visual garment retrieval.

In 2015, Kiapour et al.[7] introduced a street-shop dataset containing 404,683 shop photos collected from 25 different online retailers and 20,357 street photos, street photos. It also has a total of 39,479 clothing item matches between street and shop photos. Liu et al in 2016 [9] introduced the DeepFashion dataset which is one of the largest clothing datasets that contains 800000 images with massive attributes, clothing landmarks, as well as cross-pose/cross-domain correspondences of clothing pairs. The fashion landmark dataset was presented by Liu et al.[10] in 2016 with over 120000 images used for pose and scale variations. The Modanet dataset was constructed in 2018 by Zheng et al.[16] which has more than 55,000 fully-annotated images with pixel-level segments, polygons, and bounding boxes covering 13 categories.

1.2. Africa Fashion Dataset for Computer Vision

Previous works that involve fashion datasets are big, have a wide range of popular fashion image classes but none that we know of at the time of this project have representation of what contemporary African fashion look like. Our emphasis is on contemporary African fashion because Africa is home to some of the world's most dramatic dress practices, including textiles, jewellery, coiffures, and there is infinite combinations of all of these elements [13] . This lack of representation could be due to different factors, which include but not limited to availability of ready to use images online, individualization of African fashion items styles where consumers visit their tailors for custom fits depending on an occasion and thus not suited to the method of production of popular fashion wears.

In this work we present AFRIFASHION1600, a contemporary African fashion dataset curated to improve visibility, inclusion and familiarity of African fashion in computer vision tasks. While this is a small dataset of only 1600 images in 8 distinct fashion item classes, a good dataset needs to represent a sufficiently challenging problem to make it both useful and to ensure its longevity [3]. We hope it can help demystify computer vision use cases for upcoming communities of Africans who are learning about Artificial Intelligence and as a foundation that other interesting projects can be built one. AFRIFASHION was curated from openly available images online and will be open-sourced but all images belong to the original copyright owners.Table 1 shows basic information about the presentation of the dataset and Table 2 shows detailed explanations about the image class labels,class names and description.

Name	Description	Data Samples	File Size
African Fashion Dataset	Folder containing subfolders of images in respective classes	1600	22 MB
AfricanFashionDataset.csv	CSV format of the image_ids and their labels in a DataFrame	1600	51 KB

Table 1. Files contained in the AFRIFASHION1600 dataset

2. Methodology

Figure 1 shows the process flow for the curation of AFRIFASHION1600.

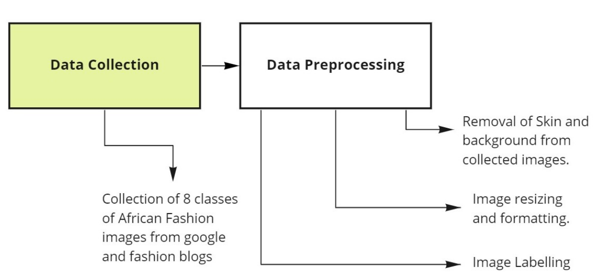

Figure 1. Stages of AFRIFASHION1600 Curation

Class Label	Class Name	Description
0	African_Blouse	This refers to different styles of blouses worn by women in different textile types such as Ankara, Lace, Linen.etc African_Blouse is worn across the continent.
1	African_Shirts	These are different styles of tops/shirts worn by men in different textile types such as Ankara, Lace, Linen.etc
2	Agbada	This attire consists of a Top/Shirt with matching trousers and a flowing gown is worn over these. This type is also called a grand boubou
3	Buba and Trouser	This is also called a dashiki trouser set. It consists of African tops/shirts and trouser
4	Gele	This is referred to as women's cloth head scarf/head tie that is commonly worn in many parts of Africa. It is called duku in Malawi and Ghana
5	Gown	This refers to different styles of a gown worn by women in different textile types such as Ankara, Lace, Linen. etc. Gowns are usually a single piece of clothing without an additional piece
6	Skirt and Blouse	This attire refers to a combination of a female African blouse worn as a top and a skirt(a separate outer garment) worn as the lower part of a dress that covers a person from the waist downwards
7	Wrapper and Blouse	This attire refers to a combination of a female African blouse worn as a top and a garment(wrapper) tied across the waist as the lower part of a dress

Table 2. Class Labels, Class Names, Number of Images and Description of Images in the AFRIFASHION1600 dataset

2.1. Data Collection

Images in the AFRIFASHION1600 dataset were curated from openly available web sources like Google images and blogs focused on African fashion. The differing sources of images,low volume of images without humans, complex poses of humans wearing the attires,fewer representation of certain classes of fashion images impacted the availability of usable images.

2.2. Data Processing

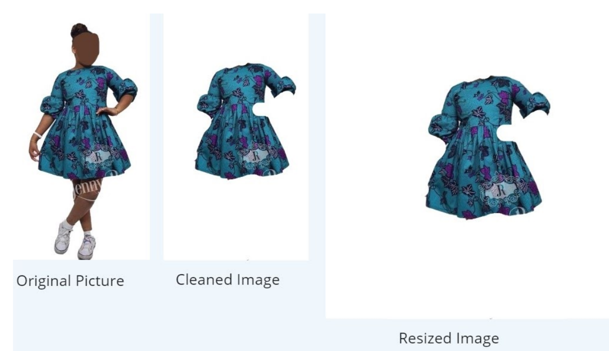

Figure 2.2 displays the image processing progression while curating the AFRIFASHION1600 dataset and the preprocessing steps are listed below :

1. Data Cleaning to remove background and skin in images, a simple in house deep learning model. AFRIRAZER [12] was used for this process

2. Cleaned coloured images were resized of 128 x 128, therefore each image has a shape of (128, 128, 3).

3. All cleaned images were converted to PNG image format for uniformity and then labelled by in house team according to the proper classes.

4. Images of the same class names were stored in the same image self-named folders in this format "foldername_id.png".

   Figure 2 shows examples of cleaned images

Figure 2. Class names and sample images for the Africa Fashion dataset

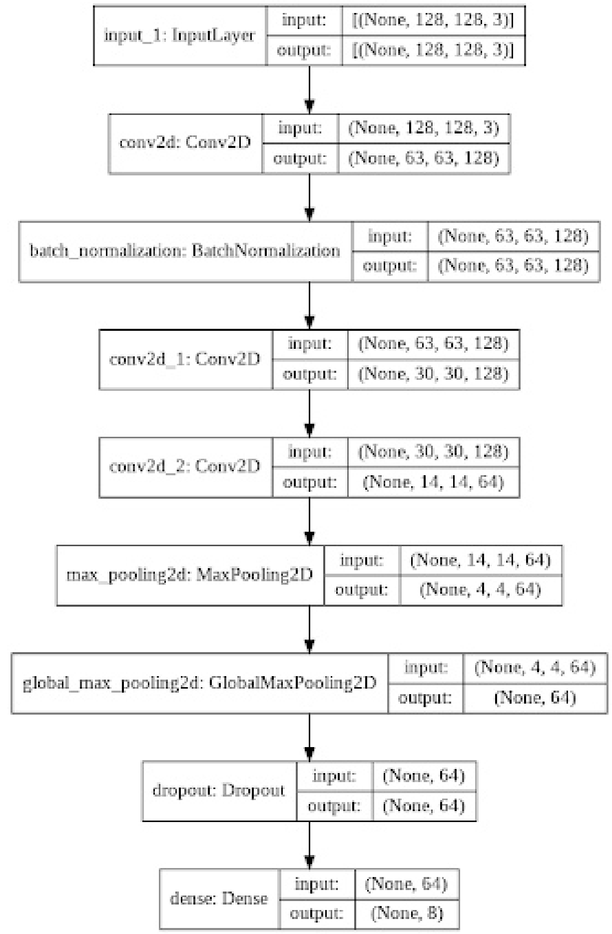

Figure 3. Model architecture using the Convolutional Neural Network

3. Classification Experiment and Results

AFRIFASHION1600 is a collection of images and it is suitable for a range of computer vision tasks such as image classification, image synthesis etc. For the purpose of this work, we experimented with image classification. The dataset having a total of 1,600 images in 8 classes is shuffled and split equally among classes into train, test, and validation with a proportion of 80:10:10 respectively, for the classification experiment.

3.1. Classification of AFRIFASHION1600 using Convolutional Neural Network (CNN)

A Convolutional neural network (CNN) is a type of artificial neural network commonly used in analyzing images. Compared to a regular neural network, CNN layers are organized in 3 dimensions: width, height, and depth [11]. The model architecture used, as shown in Figure 3 takes an input shape of (128 X 128 X 3) representing the width, height, and the number of channels for each input image, and then convolves up to the Dense layer, which is fully connected to 8 output neurons for our multi-class classification problem. Each of the hidden layers uses the ReLU activation function except for the output layer with a softmax function. After the training process, the performance of the model was evaluated with fbeta_score. When it comes to multi-class cases, F1-Score should involve all the classes. To do so, we require a multi-class measure of Precision and Recall to be inserted into the harmonic mean. Such metrics may have two different specifications, giving rise to two different metrics: Micro F1-Score and Macro F1-Score. Macro F1-Score is the harmonic mean of Macro-Precision and Macro-Recall while that of Micro-Average F1-Score is just equal to Accuracy [5].

After evaluation, the model had a test fbeta_score of 40.04. This poor performance is as expected from a model that was trained on small data of about 1,300 images. Although no form of data transformation or augmentation was applied to get this result, we assume that in practice there could be a slight improvement if data preprocessing is involved and hyperparameters tuned. However, we decided to explore transfer learning from popular pretrained deep learning .

3.2. Classification of AFRIFASHION1600 using Transfer Learning

Data gathered from multiple sources proved that pretrained ConvNets along with fine-tuning policies is better or, at least, equal as well as deep networks trained from scratch. Also, fine-tuning leads to faster convergence than training from the scratch. Authors have examined transfer learning in a variety of ways [6].The pretrained models leveraged in this work were accessed from the tensorflow.keras.application API [2]

For this experiment no data augmentation was done and the we aimed to get the base score for each model without tuning. Table**??** shows the test fbeta_score from each pretrained model.The pretrained models were trained on trained on ImageNet, a dataset of over 14 million images with 1000 classes [16].The model weights were loaded into ConvNet. However, the fully-connected layer at the top of each pretrained model was excluded from our network layers since the input size from our dataset differed from their respective default input sizes.

Model	Top-1 Accuracy on ImageNet	Model Accuracy on AFRIFASHION1600	Parameters for AFRIFASHION1600
VGG16	71.3	83.41	14,780,244
VGG19	71.3	84.30	20,089,940
Xception	79.0	69.57	21,123,644
ResNet50V2	76.0	66.46	23,826,964
DenseNet121	75.0	83.66	7,168,156
DenseNet201	77.3	81.62	18,567,764
InceptionResNetV2	80.3	70.46	54,385,908

Table 3. The pre-trained models, their test accuracy on our dataset compared to that of ImageNet, and the total params used. The hyperparameters were constant across the models, with a learning rate of 0.001, a training epoch of 100, and a batch size of 64.

Figure 3 shows Seven pretrained models that were explored in this experiment, VGG19 with 20,089,940 parameters had the highest accuracy on AFRIFASHION1600 with a score of 84.30% while ResNet50V2 with 23,826,964 parameters had the least accuracy score of 66.46%. DenseNet121 achieved an accuracy score of 83.66 with only 7,168,156 parameters

4. Conclusion

We presented AFRIFASHION1600, a contemporary African fashion dataset curated to improve visibility, inclusion, and familiarity of African fashion in computer vision tasks. While this is a small dataset of only 1600 images in 8 distinct fashion item classes, it provides a good enough challenge for machine learning models as practice and a foundation to build other interesting computer vision tasks.The authors hope the work encourages young Africans who are exploring applications of Artificial Intelligence, to build with, to curate and include similar datasets in other spheres they may have been under explored.

References

[1] Wen-Huang Cheng, Sijie Song, Chieh-Yun Chen, Shintami Chusnul Hidayati, and Jiaying Liu. Fashion meets computer vision: A survey. *arXiv preprint arXiv:2003.13988*, 2020. 4321

[2] Francois Chollet et al. Keras, 2015. 4324

[3] Gregory Cohen, Saeed Afshar, Jonathan Tapson, and Andr e van Schaik. Emnist: an extension of mnist to handwritten. *Proceedings of the IEEE*, 2017. 4322

[4] Jennifer Craik. *Fashion: the key concepts*. Berg Publishers, 2009. 4321

[5] Margherita Grandini, Enrico Bagli, and Giorgio Visani. Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756*, 2020. 4323

[6] Mehdi Habibzadeh, Mahboobeh Jannesari, Zahra Rezaei, Hossein Baharvand, and Mehdi Totonchi. Automatic white blood cell classification using pre-trained deep learning models: Resnet and inception. In *Tenth international conference on machine vision (ICMV 2017)*, volume 10696, page 1069612. International Society for Optics and Photonics, 2018. 4324

[7] M Hadi Kiapour, Xufeng Han, Svetlana Lazebnik, Alexander C Berg, and Tamara L Berg. Where to buy it: Matching street clothing photos in online shops. In *Proceedings of the IEEE international conference on computer vision*, pages 3343–3351, 2015. 4321

[8] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. 4321

[9] Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, and Xiaoou T ang. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 4321

[10] Ziwei Liu, Sijie Yan, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Fashion landmark detection in the wild. In *European Conference on Computer Vision*, pages 229–245. Springer, 2016. 4321

[11] John Paul A Madulid and Paula E Mayol. Clothing classification using the convolutional neural network inception model. In *Proceedings of the 2019 2nd International Conference on Information Science and Systems*, pages 3–7, 2019. 4323

[12] Wuraola Fisayo Oyewusi, Gbemileke Onilude, Olubayo Adekanmbi, and Olalekan Akinsande. Afrirazer: A deep learning model to remove background and skin from traditional african fashion images. 4323

[13] Victoria L Rovine. African fashion design: Global networks and local styles. *Africultures*, 69:104–109, 2006. 4322

[14] Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017. 4321

[15] Kota Yamaguchi, M Hadi Kiapour, Luis E Ortiz, and Tamara L Berg. Parsing clothing in fashion photographs. In *2012 IEEE Conference on Computer vision and pattern recognition*, pages 3570–3577. IEEE, 2012. 4321

[16] Shuai Zheng, Fan Yang, M Hadi Kiapour, and Robinson Piramuthu. Modanet: A large-scale street fashion dataset with polygon annotations. In *Proceedings of the 26th ACM international conference on Multimedia*, pages 1670–1678, 2018. 4321, 4324