A machine learning model for predicting colour trends in the textile fashion industry in south-west Nigeria

Ajibolu Oyinkansola Gladys  
E-mail: oyinkansolaajibolu@gmail.com  
Akinola Solomon Olalekan  
E-mail: akinola.olalekan@dlc.ui.edu.ng  
Department of Computer Science,  
University of Ibadan, Ibadan, Nigeria

Abstract

Fashion is primarily based on adoption of trends by customers in the textile industry. Fashion trend forecasting is a complex process that aims at identifying future preferences of customers. The textile fashion industry is volatile, trends change very quickly. Fashion trend must be closely followed to increase sales amount. Colour forecasting is considered as one of the significant driving force in the textile fashion industry. If the supply of an item surpasses its demand, it would remain unsold thereby generating loss for the industry. How do we assist the textile manufacturing industries in solving the problem of under/over stocking? Tackling this question from a data driven vision perspective, we developed a model to forecast visual colour trends. In this study, a model for colour demand forecasting in the textile industry developed. This study involves the real life application of the model using real demand values for textile clothing by customers in the south west zone of Nigeria. We forecast future purchases based on historical demand data. Two approaches were combined for forecasting colour trends in the textile industry. The developed model first used the Convolutional Neural Network (CNN) for extracting hidden information/features/patterns from the image dataset. The extracted features were then applied to K- means algorithm for extracting the colours. The study proved that the two approaches used performed excellently well and that accurate colour forecasting can significantly enhance productivity and generate more sales for the textile industry.

Keywords: Convolutional Neural Network, K-means Clustering algorithm, Demand Forecasting, Colour Forecasting, Fashion Industry

1. INTRODUCTION

Forecasting is constantly encountered in most phases of our lives. Nearly everyone in almost every walk of life forecast to some extent. A forecast is a probabilistic estimate of a future value (Celia et al., 2003). It is the science of predicting future outcomes and it is used to determine the future targets of a business product. The role of demand forecasting has become increasingly important within businesses that have the maximization of customer service and the optimization of capital investment operating costs as their main objectives (Rexhausena et al., 2012). Thus, quick and effective supply chain management with flexible production schedules and appropriate inventory levels becomes critical for each stock keeping unit. For this reason, the demand forecasting process needs to be timely and accurate (Sanuwar, 2013). It is extremely important for a business to do proper forecasting before developing new products. Fashion trends are the popular styles of clothing and accessories at a particular time (www.masterclass.com). Forecasting plays a very important and crucial role in fashion industry. Poor forecasting results in stock outs or high inventory due to which a company can have huge losses. In particular, consumer-oriented markets such as fashion face uncertain demands, short life cycles which contributes more to the challenges of producing accurate forecast (Samaneh et al., 2015).
The textile Industries face several challenges regarding accurate forecasts. For instance, they have to place their production plans before exact knowledge about future demand is available. The identification of the number of stocks and the replenishment strategy are significant activities for the textile manufacturing industries. Therefore, features which affect production can be examined to increase sales. Maximizing the profits of fashion companies while minimizing the forecasting error and reducing the cost that result from excess capacity of production or from loss of potential revenues due to low demand is a major concern for the textile Industry.

The textile and retail businesses worldwide use color to segregate and market materials, products, and product lines. Color variety and new color or shade development are major driving forces in the production and marketing of textiles as well as numerous other products. The importance of color to the textile and retail businesses of today cannot be overstated. This is why we are focusing on detecting and classifying the colours of an image; so this will give textile Industries a heads up on what textile to produce either weekly or monthly or yearly. In this study, a fashion color trend predictor that can classify live images from various events pictures using Convolutional Neural Network (CNN) and also detect the dominant colors from these pictures using K Means Algorithm was developed. The objectives of this research include:

(i) Preprocess and classify the image datasets using OpenCV library and Convolutional Neural Network algorithm respectively.


(iii) Evaluate the accuracy of the developed model

The rest of this paper is organized as follows. Section two discusses some of the related works to this study while in Section 3, the methodology employed to achieve our goal is discussed. In Section 4, results obtained from this study are presented and discussed. Conclusion is made in Section 5.

2. RELATED WORKS

Emmanuel et. al., (2019) in their study considered forecasting models like ARIMA (Autoregressive Integrated Moving Average), ETS (Exponential Smoothing), NNAR (Neural Network Autoregression) and an improved model DNNAR (Denoised Neural Network Autoregression) which was used for forecasting consumer fashion trends. Results showed that the DNNAR was the best performer among all the implemented models with the highest forecasting accuracy.

Celia et. al., (2003) work considered Artificial Neural Network (ANN), Seasonal Single Exponential Smoothing (SSES) and Winters three parameter model in forecasting sales in the fashion industry. The performance of the model was tested by comparing their coefficient of determination. The ANN gave the highest $R^2$ of 0.92 making it the best performing model among the other models.

Asli et. al., (2012) study presented the Adaptive Network based Fuzzy Inference System (ANFIS) which was used for forecasting demand in the clothing industry. The study was based on a real life application and the system was being tested by real demand values. Results showed that the demand forecasting system could help clothing manufacturers forecast demand accurately.

Andrea et. al., (2013) worked on the Fast Fourier Transform (FFT) algorithm, moving average and exponential smoothing which was individually carried out on a set of historical sales data to see which performed best. The Fourier Analysis returned a much more precise forecast in terms of MAPE (Mean Absolute Percentage Error) and Mean Absolute Deviation (MAD). It also gave more accurate results.

Ye Wang et al., (2012) considered the Support Vector Regression Method (SVR). It was used in forecasting demand and compared to the Radial Basis Function (RBF) Neural Network. Results showed that SVR is superior to RBF in prediction performance and effective for demand forecasting and was concluded that more scalable ways should be found to run these models in real world scenarios.

Elad et al., (2019) study explained the importance of the application of well based and effective forecasting models and suggests that accurate fashion trend forecasting can significantly enhance trend forecasting in textiles and also maximize profits and reduces loss that results from excess production.

Li-Xia Chang et al., (2009) presented the Statistical analysis, the Grey Theory Method and Rough Set Theory (RST) method, which was used for forecasting fashion colour. The combination of these methods greatly improved the accuracy of the prediction results and it was therefore concluded that intelligent prediction theories and algorithms should be attempted into the building of prediction models.

Pawan et al., (2019) considered the Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM) in forecasting demand for new clothings based on historical sales data. Performance was evaluated using the Weighted Mean Absolute Percentage Error (Wmape). The LSTM gave the best performance. It was recommended that more datasets should be provided for rigorous evaluation of these models and finding scalable ways to run the models in real world scenarios.

Kripesh Adhikari et al., (2017) presented an efficient method of activity recognition for fall detection using RGB and depth images. Experimental results showed that the Convolutional Neural Network is a very promising approach for pose recognition.

Teresa et al., (2018) applied the Moving average, weighted moving average, exponential smoothing, Holts method and Winters method in forecasting demand. Mean Absolute Deviation (MAD), Mean Squared Error (MSE) and R² (Determination Co efficient) was used for model performance. Results proved Simple Exponential Smoothing to be the best performing technique. It was concluded that the reliability of the models would be increased in future research by studying different types of artifical neural networks using back propagation and radial base function to predict demand.

Cagatay et al., (2019) considered Time Series Analysis and Regression methods like Linear regression, Bayesian regression, Neural Network regression, Decision Forest Regression and Boosted Decision Tree Regression in forecasting sales in this Study. Mean Absolute Error (MAE), Co efficient of Determination (R²) and Root Mean Squared Error (RMSE) was used in evaluating performance. After evaluation, Boosted Decision Tree Regression proved to be the best performing model.

The review of literature discuses the famous tools used in the fashion prediction works. However, to the best of our knowledge, none of the previous works addresses the problem of detecting dominant color of fashion trends to prevent wastage of materials in industries. Therefore, there is need to develop this model for the textile industries. The research leverages on the past knowledge of previously done works to create a better solution for this pending problem.

3.0 METHODOLOGY

Figure 1 shows the proposed model for this study, which contains two major phases: Classification of images with Convolutional Neural Network (CNN) and Classification of images into their dominant colours using K-means Clustering.
The flow decision diagram in Figure 1 shows the basic level of how our model is developed. From the data input, the images were classified using Convolutional Neural Network (CNN). The CNN classifier algorithm then classified the image data into the two classes we have. Once the class of “owanbe event” is chosen, then the k-means clustering algorithm detects and classifies the colours in frequency of the most dominant to the least dominant.

3.1 Phases for image classification
This can be broken down into four stages:

- **Phase One (Data Gathering Phase)**: In this phase, data gathering is quite important, not just any images of the two classes are gotten, specific images of the two classes are gotten and if they satisfy the features required, the images are loaded.

- **Phase Two (Data Preprocessing Phase)**: the image dataset was loaded into the system, resized to the best option and also reduced according to some noisy images that may disturb the result.

- **Phase Three (Data Representation Phase)**: In this phase, the features were extracted by the convolutional neural network itself. Data Augmentation also took place. Data Augmentation is the duplication and re-scaling of duplicated images to create more data sets for better results. Basically, it’s an automatic image generator.

- **Phase Four (Data Model Building and Evaluation Phase)** In this phase, the complete model was built, tested and evaluated. The expected result is an accurately classified image with its dominant colours classified with K-means clustering.

This study was implemented on Google colab. Colaboratory notebook is a free cloud service that uses Python programming language. It provides a wide range support for deep learning applications using popular libraries. Python 3 was used along with many other libraries like Tensorflow, pydrive for connection to the data in the google drive, keras libraries for keras model, Matplotlib, Sklearn and OpenCv.

3.2 Source and Nature of Data Used
The dataset consists of two classes: Owanbe images and Official images. Each class consists of 500 images, making a total of 1000 images in all. The owanbe images consist of images of people in a wedding party while mostly standing and colourful. The official images on the other hand consists of official events for example conferences, seminars and office meeting, the features in these images is that there are not much people in the images, chairs and tables are very vividly shown and most people were sitting down. The datasets (images) was personally collected from events/occasions across the south west region of Nigeria and also from events photographers who donated images for this project.

3.3 Feature Extraction
In this phase, the network will perform a series of convolutions and pooling operations during which the features are detected. The phase consists of alternating trainable fully convolutional layers which learn high level features. Features are nothing but the unique signatures of the given image or unique properties that defines an image. Features are extracted in order to differentiate between the images. Features extraction are used in almost all machine vision algorithms.

We can train few algorithms using the features extracted from the image. This is widely used in machine learning.

Image Classification method based on human posture:
The primary focus of our work is object detection of key features. The network is trained to detect and classify different body positions of the detected people in the images. These features are based on:
- Standing Posture: the model detects the standing position
- Sitting Posture: the model detects the sitting position

3.4 Data Preprocessing
Feature encoding and data augmentation are the two data preprocessing techniques that were used to preprocess the dataset.

3.4.1 Feature Encoding: It is the process of transforming a categorical variable into a continuous variable and using them in the model (Prashanth, 2018). Some of the features to encode are the categorical features. Most of the Machine learning algorithms cannot handle categorical variables unless we convert them to numerical values. The encoding techniques used was one hot encoding. One hot encoding creates new (binary) columns, indicating the presence of each possible value from the original data.

3.4.2 Data Augmentation: Data augmentation is usually used in CNN when a better Performance is required. It is a common technique to improve results and introduce variability in the dataset. It is the process of increasing the amount and diversity of data. We do not collect new data, rather we transformed the already present dataset (collected images). Data Augmentation can multiply a single image up to 50 times its quantity. Virtually, we used 50,000 datasets but in reality we just used 1000 datasets. It was also used in resizing and rescaling our images before being fed into the algorithm.

3.5 Image Classification Tool Employed
The classifier used to classify the images into event type for this study was Convolutional Neural Network (CNN).

3.5.1 Convolutional Neural Network (CNN)
Convolutional neural network (CNN) is a special architecture of artificial neural networks, proposed by Yann LeCun in 1988. A Convolutional Neural Network is a Deep Learning algorithm which can take in an input image, assign importance to various aspects(objects) in the image and be able to
differentiate one from the other (www.towardsdatascience.com).

All CNN models follow a similar architecture, as shown in the figure 2.

The blocks in a CNN model consists of the following:

3.5.1.1 Input Layer
The input layer is usually a multidimensional array of data where data are fed into the Network (Ferreira et al., 2017). It is the initial data for the neural network. Input data can be image pixels or its transformation, patterns, image signals, etc.

3.5.1.2 Convolution Layer
The main building block of CNN is the convolutional layer (Krig, 2016). Convolution is a mathematical operation to merge two sets of information. It reforms the original function into a new representation of the similarity of two functions in a certain window (Yu Han Liu, 2018). In our case the convolution is applied on the input data using a convolution filter to produce a feature map. Convolution is an efficient way of feature extraction, skilled in reducing data dimension and producing a less redundant dataset which is the feature map (Yu Han Liu, 2018). The first convolutional layer extracts low level meaningful features like edges, corners, lines etc. Next convolutional layer extracts higher level features but the highest level features are extracted in the last convolutional layer (Samer et al., 2015).
3.5.1.3 Pooling Layer

After a convolution operation we usually perform pooling to reduce the dimensionality. The pooling layer is responsible for reducing the spatial size of the convolved feature (www.towardsdatascience.com). This is to decrease the computational power required to process the data through dimensionality reduction (Sumit, 2018). Pooling brings much benefits to CNN. It aims at preventing overfitting by reducing data dimensionality. Pooling layers downsample each feature map independently, reducing the height and width, keeping the depth intact.

3.5.1.4 Fully Connected Layer

The fully connected layer will serve as a classifier on top of the extracted features. They will assign a probability for the object on the image being what the algorithm predicts it is (Rokas, 2019). Fully connected layers performs classification based on the feature extracted by the previous layers. It takes the end result of the convolution and pooling process and reaches a classification decision giving the final probabilities for each label (missinglink.ai).

3.5.1.5 Output Layer

After multiple layers of convolution, we would need the output in the form of a class. The convolution and pooling layers would only be able to extract features and reduce the number of parameters from the original images. However, to generate the final output, we need a fully connected layer to generate an output equal to the number of classes we need (Gupta, 2017).

From the above explanation, convolution and pooling are the two main extractors in the CNN algorithm. The CNN algorithm being implemented acts as both the feature extractor and the Classifier.

3.6 Dominant Colour Detection Tool Employed

This is a stage where the classified images were loaded for detection of colours in the images using k-means algorithm.

3.6.1 K-means Clustering Algorithm

K-means algorithm, was applied to the domain of dominant colour detection to give the frequency of colours in an image. The k-means algorithm is an algorithm to cluster n objects based on attributes into k partitions, where k < n (Hartigan and Wong., 1979; Zha et. al, Dec 2001). The way k-means works is, it takes a cluster of colours that look similar and group them together in ranking. We used k-means for colour detection because it is the best for separating the colours in the image with clusters while grouping the same images together. Simply speaking k-means clustering is an algorithm to classify or to group the objects based on attributes/features into K number of group. K is positive integer number (Hartigan & Wong, 1979; Zha et al, 2001).

\[
J = \sum_{j=1}^{K} \sum_{i=1}^{n} ||x_i^{(j)} - c_j||^2
\]

(1)

Where:

\( J \) = Objective function
\( K \) = Number of clusters
\( N \) = Number of cases
\( x_i \) = Case i and \( c_j \) = Centroid for cluster j

The grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid.

The algorithm steps taken to achieve this purpose are presented below in Fig 4.
In Figure 4 flow chart, at random we select ‘k’ points not necessarily from the dataset and assign each data point to the closest cluster. We then compute and place the new centroid of each cluster and reassign the data points to the new closest clusters. If any reassignment takes place, we return to step 3. Else the model is ready. This model enabled the detection and classification of colours.

3.7 Performance Metrics

The evaluation metrics for the model built are: Accuracy and Log Loss

- **Accuracy**: This gives the ratio between the correctly predicted outcome and the total sum of all predictions

\[
\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}
\]  \hspace{1cm} (2)

- **Binary Cross Entropy Loss**: Measures how far away from the true value. It also describes the loss between two probability distributions.

\[
H(p, q) = - \sum_x p(x) \log q(x).
\]  \hspace{1cm} (3)

Where:

\( p(x) \) is the true probability distribution
\( q(x) \) the predicted probability distribution.

The loss tells us how wrong our model predictions are. The lower the value of log loss, the higher the accuracy.

4.0 IMPLEMENTATION

4.1 Image Data Preprocessing

Training Detail

With reference to Figure 3, an image of size 200 x 200 is fed as input into the CNN consisting of four CNN layers and three fully connected layers with a ReLU and Sigmoid as last layer. Each convolutional layer uses a filter of 2 x 2 and a stride of 2. Each convolutional layer is followed by a max-pooling layer of filter size 2 x 2 and a stride of 2. The CNN is trained with 1000 images.
4.2 Image Classification

Convolutional Neural Network was used to classify the images after the dataset had been preprocessed. Figures 4 and 5 show the validation loss, as well as the training and validation accuracy. Table 1 shows the summary of the validation accuracy and loss values.

![Figure 5: Training and Validation Accuracy over 20 epochs](image)

![Figure 6: Training and Validation loss over 20 epochs](image)

Table 1: Training/Validation Accuracy and Loss Values

<table>
<thead>
<tr>
<th>Validation Accuracy</th>
<th>Validation Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.89</td>
<td>0.28</td>
</tr>
</tbody>
</table>

4.3 Testing Phase

Here in the Testing phase, the model retrieves the testing datasets from the folder in the google drive. We have manually set the test dataset to 18 images, 9 from official event, 9 from Owanbe Party Events.

4.3.1 Image Classification Result

The Test datasets in a folder was referenced from the google drive being used. The classification result brought out two classifications. The official event and the Owanbe event as seen in the images below.
4.3.2 Colour Classification Result
4.3.2.1 Image Loading
Once the images were classified and stored, we loaded the images from the stored place on the computer or google drive. The image was read using OpenCv library. The opencv Library allows smooth image reading.
4.3.2.2 Plotting the Image to Graph

Figure 9: Image loading from the computer

Figure 10: Image conversion to array

Figure 11: 3D Graph of the Image
From Figure 4.11, one can easily see that the data points are forming groups. Some places in a graph are denser, which we can think as different colour dominance on the image. We tried to achieve these clusters through k-means clustering.

**4.3.2.3 Plotting the Dominant Colour**

| 154 75 104 |
| 230 202 225 |
| 52 18 53 |
| 140 173 200 |
| 209 109 176 |

Figure 12: colours displayed in the order of dominance \((k = 5)\)

Figure 12 shows the final result. The colours were being arranged in an order that portrayed how dominant they were in the image.

**4.4 Results and Discussion**

In a binary classification algorithm such as logistics regression, the goal is to minimize the cross-entropy function. The whole purpose of the loss function is to return high values for bad predictions and low values for good predictions. Therefore the lower the values of log loss, the higher the value of accuracy. The log loss helps in the prevention of error by streamlining down the probability of the classification.

The Accuracy is the value used to describe how accurate the model will be in prediction. The Training Accuracy for our model was 89.9% and the validation Accuracy was 87.5%. This is a good result as the model predicted our images fairly. The Log Loss for our training was 28.5% and validation was 31.3%. This is a good result because the lower the log loss, the more accurate the model will be. The above results suggest what a promising tool CNNs can be, especially to solve visual recognition tasks.

Some existing works in the literature prove that our model did excellently well:

In Mohamed El Amine Elforaici *et. al.*, (2018) work, the feature based models, which combined Support Vector Machine (SVM) and Convolutional Neural Network (CNN) gave an accuracy of 78.6% and 88.1% for the standing and walking positions respectively. Also, in Kripesh Adhikari (2017) work, the Convolutional Neural Network (CNN) was used in the activity recognition for fall detection which gave an accuracy of 74%.
5.0 CONCLUSION

We have presented a study for the demand forecast of coloured textiles in the south western region of Nigeria. In this study, we investigated the effect of K means clustering algorithm and the Convolutional Neural Network in the demand forecasting problem in the textile manufacturing company. The model first used the Convolutional Neural Network to extract hidden/underlying information (features) from the image dataset. The extracted features were then applied to K means algorithm for clustering the images into several disjoint clusters of colour in their order of dominance. Image datasets consist of images gotten from all forms of events across the south western region of Nigeria. This images were used for training, testing and evaluating the performance of the developed model.

Experimental results in terms of accuracy and log loss showed that the developed demand forecasting model produced excellent forecasting result and according to this experiment, it can be concluded that the model can be a very effective and useful tool for the textile manufacturing industries. The vision based forecasting model can be very useful for the industry for accurate production planning and the marketing departments can easily make analysis considering future demands for specific colours by using this developed model.

References

ISSN 2277-5528


DOI: [10.1088/1742-6596/1087/6/062032](https://doi.org/10.1088/1742-6596/1087/6/062032)