International Journal of Applied Sciences and Smart Technologies

Volume 5, Issue 2, pages 111–126

p-ISSN 2655-8564, e-ISSN 2685-9432

A Deep Learning Model for Identical National Flag Recognition in Selected African Countries

Halleluyah Oluwatobi Aworinde[1]*, Oladosu Oladimeji[2], Segun Adebayo[3], Akinwale Akinwunmi[1], Aderonke B. Sakpere[4], Olayanju Oladimeji[1]

[1]College of Computing and Communication Studies, Bowen University, Iwo, Nigeria
[2]Centre for Mathematical Modelling and Intelligent Systems for Health and Environment, Atlantic Technological University, Sligo, Ireland
[3]College of Engineering, Agriculture and Sciences, Bowen University, Iwo, Nigeria
[4]Department of Computer Science, University of Ibadan, Nigeria

[*]Corresponding Author: aworinde.halleluyah@bowen.edu.ng

(Received: 15-05-2023; Revised: 26-05-2023; Accepted: 22-06-2023)

Abstract

The national flags are among the symbolic representations of a country. They make us understand the country of interest in a particular issue. Therefore, they are commonly used in both private and government organizations. It has been discovered in recent times that the younger generation mostly and idly spend its time online; hence, knowing little about national flags. Additionally, some national flags (particularly in West Africa) are identical in nature. The likeness is in terms of layout, colours, shapes and objects on the national flags. Hence, there is a need to have a model for flag recognition. In this paper, national flag images of some West African countries were gathered to form a dataset. After this, the images were preprocessed by cropping out the irrelevant parts of the images. VGG-16 was used to extract necessary features and to develop the deep learning model. This contrasted with the existing handcrafted feature extraction and traditional machine learning techniques used on this subject matter. It was observed from this study that the proposed approach performed excellently well in predicting national flags; with an Accuracy of 98.20%, and an F1 score of 98.16%. In the future, it would be interesting to incorporate the national flag recognition into Human-Computer Interaction System. For instance, it could be used as flag recognition in some mobile and web applications for individuals with colour blindness.

This research work presents a robust model because of nature of the dataset used in this work compared to previous works.

Keywords: national flag, deep learning, multi-class, VGG, West Africa

1 Introduction

A National flag is one of the symbolic representations of a country. National flag together with a nation's coat of arms are used to emphasize national identity; hence, can be used in various gatherings including conferences, sports, and workshops [1]. Due to the importance of national flag to a nation, it can be seen virtually everywhere including social media websites [2], military operations, various (local and international) competitions and occasions, television or print media, soft or hard copy documents, the streets among other places [3]. National flag makes us understand the country of interest in a particular issue. Therefore, they are commonly used in both private and government organizations. Hence, flag recognition is very important to understand a country or organization of the subject of concern.

Additionally, despite the importance of the national flag, some national flags of some nations (particularly in west Africa) are identical in nature. The similarity is in terms of layout, colours, shapes and objects on the national flags as seen in Figures 1 and 2.

Figure 1 Two identical national flags based on layout and colour [4]

Volume 5, Issue 2, pages 111–126

p-ISSN 2655-8564, e-ISSN 2685-9432

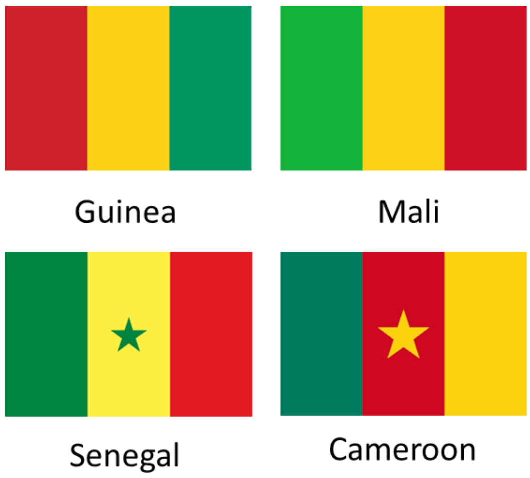

Figure 2 Four identical National flags based on object, shapes, colour and layout [5]

Furthermore, various colour blindness diseases including deuteranopia, protanopia and tritanopia incapacitate people with these diseases to identify and differentiate one national flag from another which affects their productivity in workplaces and schools.

However, the recent improvement in Information Technology [6] which result to rise of Artificial Intelligence with fields such as Computer Vision and Machine Learning [7] which has made various processes to be automated including facial recognition, object detection among other things to solve problems in various domains [8]. Hence, the way forward is to develop a model which will make use of this Artificial Intelligence techniques to recognize and identify national flags to alleviate the challenges which are often encountered.

Currently, works on flag recognition are not much [9]; Hart et al. [10] proposed an interactive flag recognition system based on manual cropping of the flag from pictures. The system uses colour-based nearest neighbor classifier for the generation of list for user to make decision. In another approach, Jetley et al. [11] created dataset of flag images in which the authors used Micro-Structure Descriptor (MSD) by Liu et al. [12] for the extraction of features thereafter, SVM algorithm was used for the classification of the features and 99.2% accuracy was achieved. Similarly, Hao et al. [13] proposed a Color Threshold Determination (CTD) method to identify color flags in which Histogram Oriented Gradient (HOG) was used for feature extraction. Recognition accuracy of 97.1% was achieved under complex scene, sensitivity of 90.70% and specificity of 99.33%. However, with the emergence of Deep Neural Network (DNN) [14] which performs better than the traditional machine learning approach, Gu et al. [9], took a binary classification approach to flag recognition to differentiate between flag image and other images. In[15] the authors gathered 20,000 images of national flags after which color-based descriptor with Convolutional Neural Network (CNN) approach was proposed for National flag recognition, mean Average Precision of 89.5% was achieved. While Wu et al. [14] took another approach using binary Mask Region-Based Convolutional Neural Network (R-CNN) segmentation weights for flag recognition in the wild.

Therefore, the existing approaches cannot be applied on real world scenarios with a complex background and the existing of more than one object. Hence, this work aims at using multi-class deep learning approach to national flag recognition. For this paper, the main contributions include: (i)Proposing a deep learning model for countries' flag recognition, (ii) combining the proposed deep learning with a feature extractor (VGG-16) to enhance the performance of the proposed approach and (iii) developing a national flags dataset to train the proposed deep learning model.

2 Methodology

This work proposes to use deep learning approach for flag recognition. In this section, the description of the methodology used for automatic national flag recognition in West Africa countries is described.

2.1 Data Description and Preprocessing

The National flag image dataset was created manually by taking pictures of the national flag and from the internet using Google image search, Bing image search, Flickr, Facebook, and Instagram. The dataset entails Benin, Cameroon, Ghana, Guinea, Guinea-Bissau, Mali, and Senegal national flag images. The details of the national flags are given in Table 1.

Table 1: Dataset Distribution

National Flags	Number of Instances present
Benin	708
Cameroon	715
Ghana	714
Guinea	712
Guinea-Bissau	712
Mali	713
Senegal	713

The choice of these country is due to the challenges these flags present including intra-class and inter-class variations. Additionally, some of the flags of these countries have similar colours with different objects which make the identification of these flags to be painstaking as seen in Figures 1 and 2 hence, the model would ensure high performance.

2.2 Labeling

The ground truth labeling of the images was done manually by carefully examining the national flag images and assigning the corresponding label for each national flag image.

2.3 Image Resizing/Normalization

Since the necessary portion of the flag image has been gotten and unnecessary portion cropped out, each of the image maybe in different size. Therefore, there is need to resize the images to ensure that the images are configured (in height and weight) for the input layer of the training model.

2.4 Data augmentation

It is an effective technique to prevent classifier model from overfitting by providing randomly distorted training images to the model and thus, allowing the model to learn general features [15]. In this research, the training images were augmented by randomly zooming the image through 20%, by rotating the image with range value of

Volume 5, Issue 2, pages 111–126
p-ISSN 2655-8564, e-ISSN 2685-9432

40, width shift range of 20% and shear range of 20%. This will allow the model to learn to identify the flags irrespective of their distance with the camera (zoom).

After data augmentation, for evaluation purposes, the dataset was partition into two, 70% was used for training and 30% was used for testing.

2.5 Model Design

In recent times, it has been proved that Deep Convolutional Neural Network (DCNN) is efficient for object recognition. Nevertheless, a lot of training images are required for this. Therefore, to make the best use of DCNN for national flag recognition, we need to use the pre-trained DCNN instead of training a CNN from scratch and this comes with several benefits such as less computational power needed since the CNN processes the images just once to extract the features. Also, this process requires less data as the CNN architecture has activations that could be used for feature extraction to obtain a high accuracy [18].

Therefore, for the feature extraction pre-trained VGG 16 [19] was used for the extraction of high-level and the rich features. These extracted features include shapes, patterns, colors, and textures, among others. The earlier layers of the network detect low-level features like edges and textures, while the deeper layers learn more complex features that represent objects in the flags. The extracted CNN features were trained by adding flatten layer then two custom dense layers and an output layer which is suitable for multi-class problems. These dense layers make use of similarity measurements by analyzing the activation values of specific features that have been extracted in the initial stage. The similarity measurement helps in comparing different objects among the flags. Features with an activation value of zero indicate that they are not activated. The number of features (feature index) will continue to decrease from the flatten to the output layer which entails the number of classes (0-6) where the final prediction is made based on the activated features. A sample of the activation visualization of the features is shown in Figures 3-6.

International Journal of Applied Sciences and Smart Technologies

Volume 5, Issue 2, pages 111–126

p-ISSN 2655-8564, e-ISSN 2685-9432

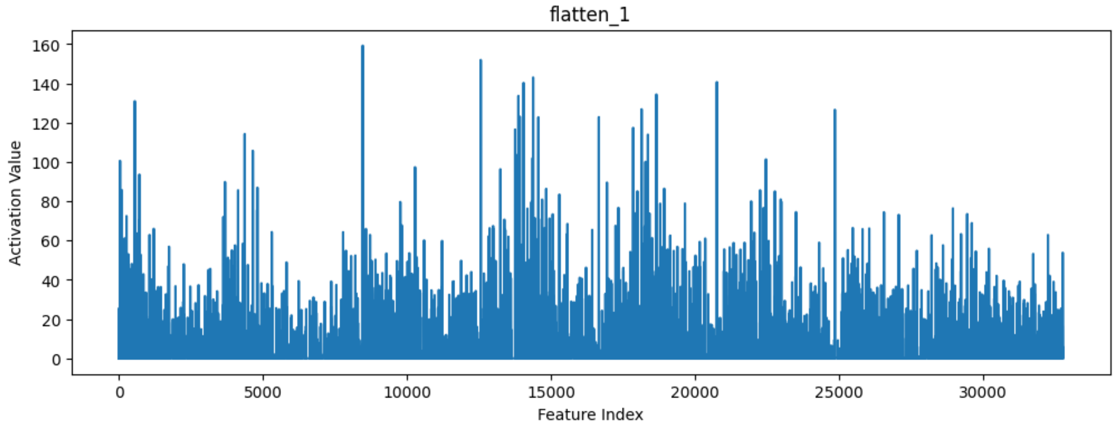

Figure 3: Activation visualization of the features for the flatten layer

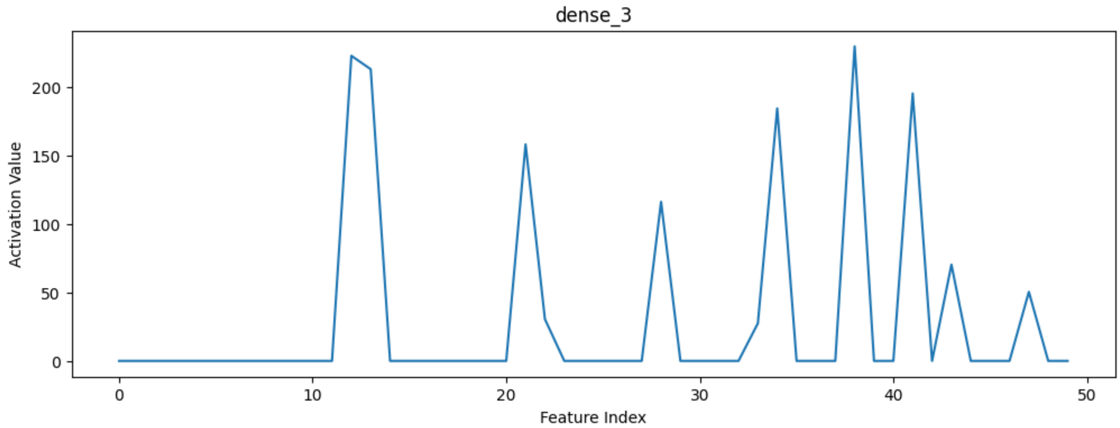

Figure 4: Activation visualization of the features for the first dense layer

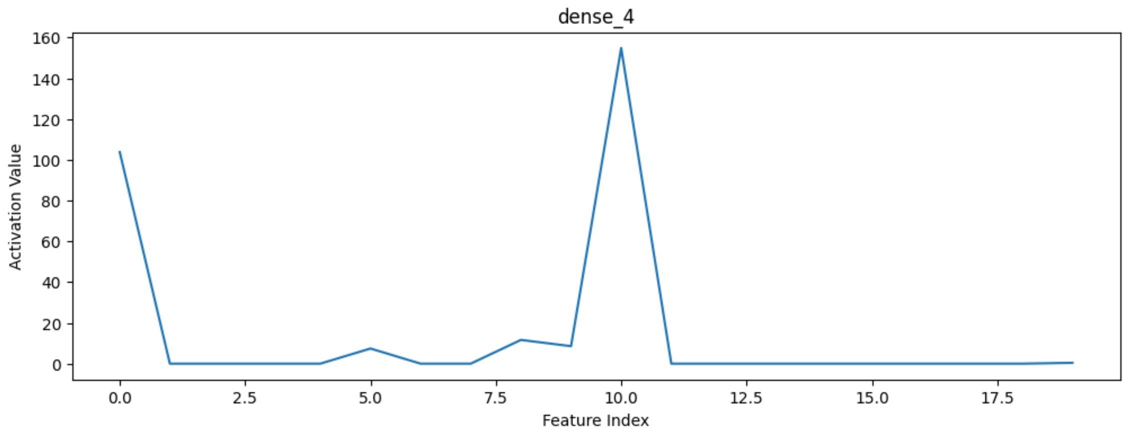

Figure 5: Activation visualization of the features for the second dense layer

International Journal of Applied Sciences and Smart Technologies
Volume 5, Issue 2, pages 111–126
p-ISSN 2655-8564, e-ISSN 2685-9432

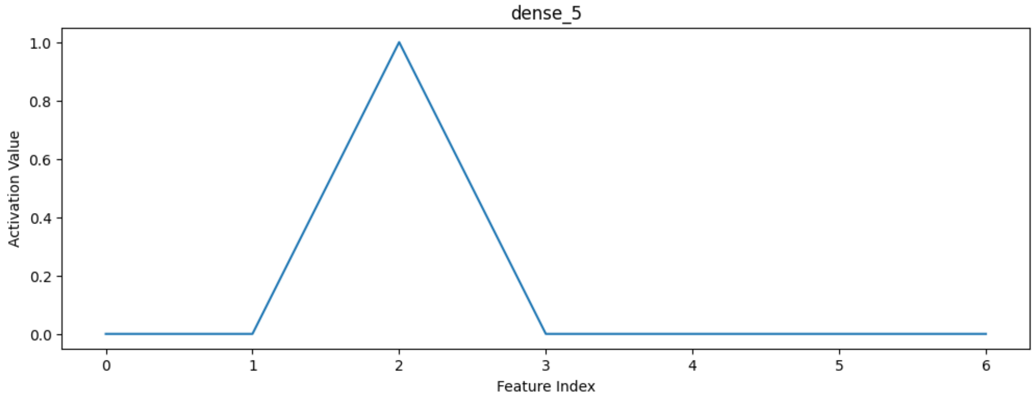

Figure 6: Activation visualization of the features for the output layer

In VGG 16, as shown in Figure 7 the input image is transmitted through a group of convolutional layers with a 3x3 receptive field. It is assumed that the convolution stride is 1 pixel. For spatial pooling, five max-pooling layers with a stride of two are used (down sampling). A 2x2 pixel window is used for max-pooling layers, which succeed some of the convolutional layers. After the set of convolutional layers, there are three fully connected layers with channel sizes of 4096, 4096, and 1000, respectively.

Each neuron in the fully connected layer receives information from the activations of the neuron in the layer below. The number 1000 represents the total number of categories in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). However, for this research 7 was used since the number of the countries considered is 7. The last layer is called the soft-max layer. All the hidden layers are outfitted by the rectification (ReLU) non-linearity layer [17] given in equation (1). VGG-16 architecture's main benefit is that it generalizes well to other datasets [19].

$$f(x) = \max(0, x) \tag{1}$$

Furthermore, SoftMax [21] given in equation (2) was applied to produce probability values for each class. The class with the highest probability value is the identified flag. Multiclass cross entropy was used as the loss function.

$$\sigma(\vec{z})_I = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{2}$$

International Journal of Applied Sciences and Smart Technologies
Volume 5, Issue 2, pages 111–126
p-ISSN 2655-8564, e-ISSN 2685-9432

Where: σ is the softmax, $\vec{z}$ is the input vector, e^{z_i} is the standard exponential function for input vector, K is the number of classes and e^{z_j} is the standard exponential function for output vector.

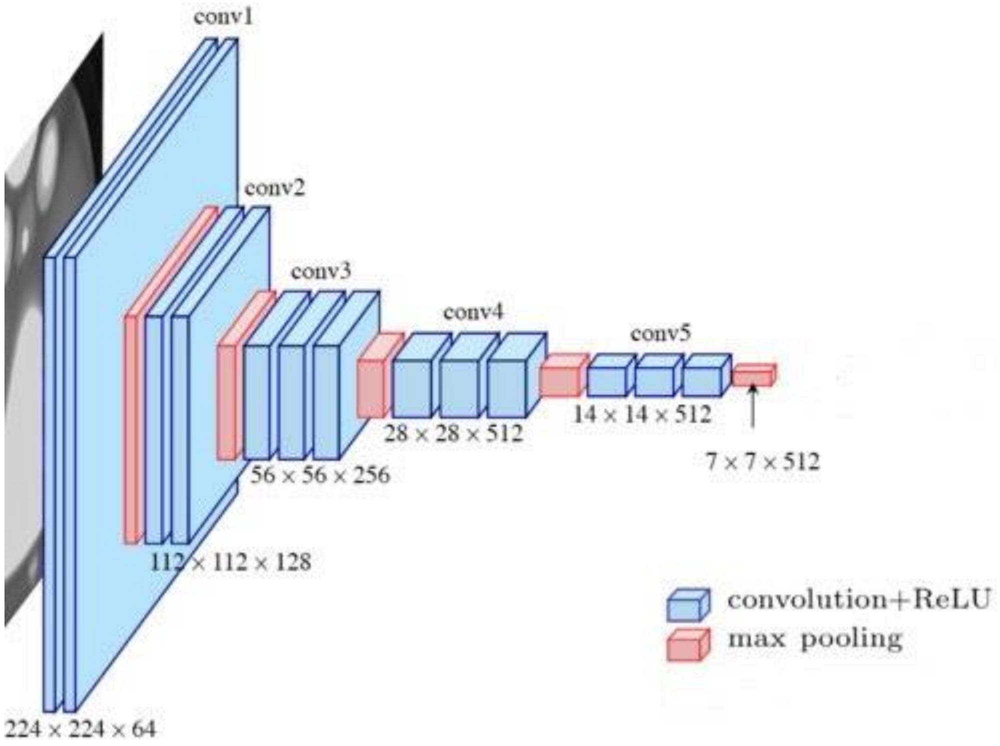

Figure 7: The architecture of VGG-16 [20]

In this research, Paper space Gradient (available at: https://www.paperspace.com/), an open source web application to create and edit live code was used in order to maximize the use of software and hardware resources made available for researchers by Paper space 30GB of RAM and 5GB Disk allocated by paper space which was used for developing the model. All the algorithms were implemented with the use of TensorFlow framework with other libraries including OpenCV for manipulating and image preprocessing.

3 Results and Discussions

To evaluating this model, the dataset was portioned into two. 70% of the dataset was used for training while the remaining was used for testing. Evaluation metrics such as Accuracy, F1-Score and recall were used to evaluate the model.

International Journal of Applied Sciences and Smart Technologies
Volume 5, Issue 2, pages 111–126
p-ISSN 2655-8564, e-ISSN 2685-9432

Table 2: Performance Evaluation

Metric	**Score**
Accuracy	98.2%
Recall	0.8902620087336245
Precision	0.9280771384659074
F1 Score	0.9816383477392395

Additionally, Streamlit [23] was used to deploy the model to web application after which independent test set was supplied to the web app for prediction as seen in Figures 8 and 9.

Figure 8. Result of Independent test on Ghana

Figure 9 Result of Independent test on Guinea

Failure cases

Due to high similarity among these flags, there were failure cases in the experiment as shown in figures 10 and 11.

Figure 10.Result of Independent test on Cameroon wrongly predicted as Mali

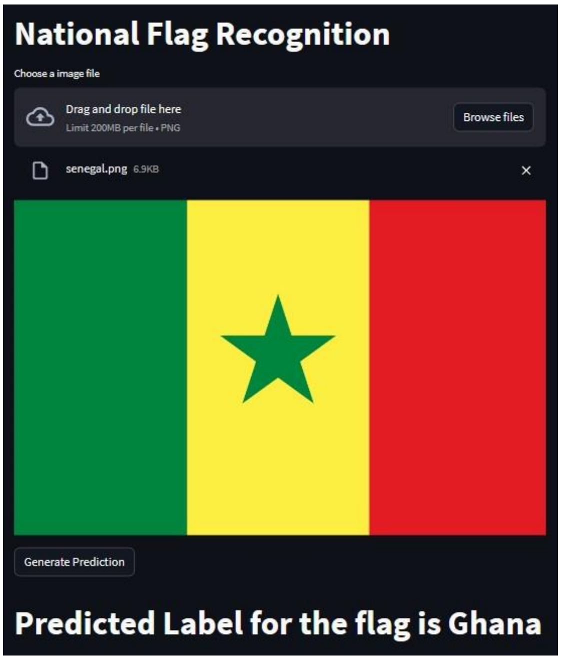

Figure 11. Result of Independent test on Senegal wrongly predicted as Ghana

Based on our research, this proposed deep learning approach to national flag recognition, entails convolution layers, SoftMax, dense and other special layers including dataset of 4987 National flag images. Most of the existing works used handcrafted feature extraction techniques. However, deep learning approach was used in this work to extract various level features after which the extracted features were passed to additional layer that enhanced the accuracy of our model by focusing on the necessary features. This work, compared to previous works on this subject matter, such as [7] and [13]has better performance in terms of accuracy and precision.

This approach has lots of advantages such as the use of VGG-16, a pretrained model to extract necessary features; then the extracted features are passed through additionally convolutional and dense layers thereby neglecting irrelevant features and enhancing the accuracy of the model.

The dataset has a huge part to play in the performance of the deep learning model, through national flag images gotten from various sources with different backgrounds made the model to be more robust.

4 Conclusions

The aim of this work is to develop a model that would detect and identify national flags in an image. For this reason, a model was developed using multi-class learning approach. The model was trained on flag image dataset which was collected both locally and from the internet summing up 4987 images. This is the first West African flag dataset on this subject matter as far as we are concerned.

The result of this work shows that high accuracy can be obtained with the use of dimensionality reduction (feature extraction and selection). It was observed from this study that the proposed approach performed excellently well in predicting National flags; with Accuracy of 98.20% and F1 score of 98.16%. As a result of the tedious task of gathering dataset, this work was able to gather over 4987 images. Thus, it would also be recommended that more data (flag images) be added to the dataset. It would be interesting to propose more approaches to national flag recognition which will also serve as comparison with this work in future. Furthermore, the application of association rule mining techniques for improved prediction, since some flags have common objects, shapes, and layout. Hence, association rule mining will help unveil these combination patterns in flags for enhanced prediction.

References

[1] Y. Moroi, The Flag and the Anthem in Japan, Peace Review, 32(2) (2020) 212–218.

[2] O. Orti, R. Tous, M. Gomez, J. Poveda, L. Cruz, and O. Wust, Real-Time Logo Detection in Brand-Related Social Media Images, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), (2019) 125–136.

[3] D. M. Montserrat, Q. Lin, J. Allebach, and E. J. Delp, Logo detection and recognition with synthetic images, in IS&T International Symposium on Electronic Imaging, Imaging and Multimedia Analytics in a Web and Mobile World 2018, (2018) 3371–3377.

[4] W. Smith, flag of Benin, Encyclopedia Britannica, (2013).

[5] W. Smith, flag of Guinea-Bissau, Encyclopedia Britannica, Jul. (2011).

[6] J. L. Hennessy and D. A. Patterson, Computer Architecture: A Quantitative Approach," in Morgan Kaufmann, San Francisco, CA, sixth edition., (2019).

[7] N. C. Thompson, K. Greenewald, K. Lee, and G. F. Manso, The Computational Limits of Deep Learning, 4 (2020) 2–5.

[8] V. Wiley and T. Lucas, Computer Vision and Image Processing: A Paper Review, International Journal of Artificial Intelligence Research, 2(1) (2018) 22.

[9] M. Gu, K. Hao, and Z. Qu, Flag detection with convolutional network, CSAI '18: Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence, ACM, (2018) 258–262.

[10] E. Hart, S.-H. Cha, C. C. Tappert, E. ; Hart, and S.-H. Cha, Interactive Flag Identification, (2004).

[11] S. Jetley, A. Vaze, and S. Belhe, Automatic flag recognition using texture based color analysis and gradient features, 2013 IEEE 2nd International Conference on Image Information Processing, IEEE ICIIP 2013, (2013) 464–469.

[12] G. H. Liu, Z. Y. Li, L. Zhang, and Y. Xu, Image retrieval based on micro-structure descriptor, Pattern Recognition, 44(9) (2011) 2123–2133,

[13] K. Hao, Z. Qu, and Q. Gong, Color flag recognition based on HOG and color features in complex scene, Ninth International Conference on Digital Image Processing (ICDIP 2017), (2017) 104200A.

[14] Y. Lecun, Y. Bengio, and G. Hinton, Deep learning, Nature, 521 (7553) (2015) 436–444.

[15] Y. Said and M. Barr, Countries flags detection based on local context network and color features, Multimedia Tools and Applications, (2021) 14753–14765.

[16] S. Wu, FlagDetSeg : Multi-Nation Flag Detection and Segmentation in the Wild, no. 5.

[17] L. Perez and J. Wang, The Effectiveness of Data Augmentation in Image Classification using Deep Learning, Convolutional Neural Networks Vis. Recognit, 11 (2017) 1–8.

[18] I. Gogul and V. S. Kumar, Flower species recognition system using convolution neural networks and transfer learning, 2017 4th International Conference on Signal Processing, Communication and Networking, ICSCN 2017, (2017).

[19] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, (2015) 1–14.

[20] M. Ferguson, R. Ak, Y. T. T. Lee, and K. H. Law, Automatic localization of casting defects with convolutional neural networks, Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017, (2017) 1726–1735.

[21] T. Kaur and T. K. Gandhi, Automated brain image classification based on VGG-16 and transfer learning, Proceedings - 2019 International Conference on Information Technology, ICIT 2019, (2019) 94–98.

[22] S. Sharma, S. Sharma, and A. Anidhya, Understanding Activation Functions in Neural Networks, International Journal of Engineering Applied Sciences and Technology, 4(12) (2020) 310–316.

[23] S. Shukla, A. Maheshwari, and P. Johri, Comparative Analysis of Ml Algorithms & Stream Lit Web Application, Proceedings - 2021 3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2021, (2021) 175–180.

This page intentionally left blank