INTERNATIONAL JOURNAL OF COMPUTING, INTELLIGENCE, AND SECURITY RESEARCH (IJCISR)

Available Online @
ijcsir.fmsisndajournal.org.ng
IJCISR 2(1) 2023
Original Research

SELECTED NIGERIA FEDERAL UNIVERSITY CLASSIFICATION BY LOGO WITH MAXIMALLY STABLE EXTREMAL REGION (MSER) AND CONVOLUTIONAL NEURAL NETWORK (CNN)

[1*]Martins E. Irhebhude, [2]Adeola O. Kolawole and [3]Ali K. Hassan
[123]Department of Computer Science, Nigerian Defence Academy, Kaduna
*Corresponding author: mirhebhude@nda.edu.ng

Abstract

Logo recognition is important in various applications, including branding analysis, advertisement monitoring, and image retrieval. This paper presents an approach for the classification of selected logos of Federal Universities in Nigeria, leveraging the combination of Maximally Stable Extremal Regions (MSER) and Convolutional Neural Networks (CNN). The proposed methodology consists of two main stages: image region extraction using MSER and logo classification using CNN. In the first stage, the MSER algorithm was employed to detect stable and distinctive regions in the input images. This approach effectively captured the salient features of the logos. These MSER regions were bounded by polygon and preprocessed to generate a dataset for training a CNN model. In the second stage, a deep learning architecture based on CNN was designed and trained to classify the logos into the specified university classes. The model's performance was evaluated using accuracy, precision, recall, and F1-score. Experimental results demonstrated that the combination of MSER and CNN recorded 0.9626, 0.9561, 0.9622, and 0.9620 for accuracy, precision, recall, and F1 score respectively. The proposed logo recognition system was compared with two state-of-the-art models and recorded excellent performances. The model provides an efficient and effective solution for identifying and classifying logos in real-world scenarios. It offers potential applications in logo detection, logo-based image retrieval, and brand monitoring, facilitating various domains such as marketing analysis, copyright enforcement, and educational research. This approach's performance indicates its potential for further development and integration into practical systems.

Keywords: MSER, CNN, Logo recognition, Nigeria Federal University Logo

INTRODUCTION

Computer vision (CV) is a field of computer science that works on enabling computers to see, identify, and process images in the same way that human vision does, and then provide appropriate output. It is like imparting human intelligence and instincts to a computer. The art of viewing and recognizing different types of objects like humans is not an easy task to implement in computers, but the field of computer vision is improving (Irhebhude, 2015). CV as a field of study seeks to develop techniques that help computers to understand the content of digital images and videos (Brownlee, 2019).

Computer vision is concerned with the automatic extraction, analysis, and understanding of useful information from a single image or a sequence of images. It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding (Irhebhude, 2015).

A logo is a unique visual, word, or combination of words and graphic symbol(s) that identifies a company, public organization, an institution, or an individual's products and services. It is also seen as the means of identification for a corporation, organization, or institution all over the world (Alaei & Delalandre, 2014). A logo is also defined as a recognizable and distinct graphic design usually consisting of shapes, images, texts, or a combination of all (Hou et al., 2023). Logo detection has been widely applied in areas such as intelligent vehicle transportation systems (Yu et al., 2019), video advertising systems for product recommendations (Cheng et al., 2017), and television advertisements (Carvalho et al., 2021).

Given an image of a logo, detecting and separating a logo becomes a highly important task during intelligent logo-based image retrieval and classification. Thus, with logo detection, indexing of images is easily done thereby increasing access to documents stored by users. Therefore detection of a logo and separating it from the document has drawn the attention of researchers in the field of document image processing and document image analysis (Dixit & Shirdhonkar, 2016). Logo can be applied to the classification of objects which can be in the form of the Nigeria university logo. One of the key challenges in Nigeria's university logo recognition is the limited availability of annotated training data. While deep learning methods have shown great promise in logo recognition tasks, they heavily rely on large-scale annotated datasets for effective training (Bianco et al., 2017). However, for Nigeria university logos, such datasets may be scarce or non-existent. The scarcity of labeled data poses a significant hurdle in developing accurate and robust deep learning models for Nigeria university logo recognition. Addressing this challenge requires the creation of a comprehensive and representative dataset of Nigeria university logos, encompassing various logo designs, variations, and image conditions, to enable the development of effective recognition algorithms tailored to the unique characteristics of Nigeria university logos. Most previous methods are either time-consuming to deploy, require special hardware, or do not reach 100% detection accuracy (Bianco et al., 2017). This study uses the CNN-based classifier with MSER to recognize the selected Nigeria university logos. By leveraging the capabilities of the logo recognition system, Nigeria universities can enhance their brand management, protect their identity, and streamline various operational processes related to logo usage and recognition.

The remainder of the paper is as follows: section II describes the MSER model. Related studies are reported in section III. The methodology is discussed in section IV. section V discusses the experimental results. The summary of findings, conclusion, and recommendations, are described in section VI.

MAXIMALLY STABLE EXTREMAL REGION (MSER)

In computer vision tasks, the detection of interest points and features is important, several applications rely on these interest points as they possess distinguishing and highly invariant properties. MSER is a fast detection algorithm for affinely-invariant stable subset of extremal regions (Matas et al., 2004), MSER algorithm was designed to solve the problem of establishing correspondence between sets of images taken at different viewpoints. The term MSER depicts the collection of distinct regions found in grayscale images each of which describes an extremal property of the intensity function in the region and its outside boundaries. MSERs have properties that give them superior performance as stable local detector (Donoser & Bischof, 2006).

Matas et al. (2004) illustrated the concept of MSER as follows;

Image $\boldsymbol{I}$ is a mapping $I: D \subset \mathbb{Z}^2 \rightarrow S$.

Extremal regions are defined if:

- S is ordered
- An adjacency (neighbourhood) relation $A \subset D \times D$ is defined

RELATED LITERATURE

Several authors have used logos as an object to classify items for marketing analysis, copyright enforcement, educational research, etc. related works that used logos for classification are reported here.

Hendrick et al. (2018) proposed a deep learning pre-trained model for the classification of halal logos which was implemented in a mobile phone application with, an accuracy of 81.7 %. Similarly, Zulkeefli and Hashim (2022) proposed a deep learning technique to identify halal logos from different countries, the logos were identified using three existing techniques and algorithms namely;

YOLOv5, back propagation neural network, and MobileNetV2-SSD. An experiment was conducted by applying these techniques; the back propagation neural network had the best performance with F1-score, precision, and recall scores of 0.949, 0.960, and 0.940 respectively. To improve consumer product authentication, the technique was integrated into a smartphone application for identifying and tracking halal logos.

Yousaf et al. (2021) proposed patch-CNN for logo recognition. The technique works by using small patches of logos for training to solve the problem of misclassification, The classification was accomplished by dividing the logo images into small patches and a threshold was applied to drop no logo area according to ground truth. The architectures of AlexNet and ResNet were used for logo detection. This study proposed a segmentation-free architecture for logo detection and recognition. In literature, the concept of region proposal generation was used to solve the problem of logo detection, however, this technique suffered when the logo size was small. The CNN method proposed was specifically designed for extracting the detailed features from logo patches and attained an accuracy value of 0.9901.

Alsheikhy et al. (2020) proposed the use of a transfer learning technique for logo recognition with a small computational overhead. The proposed method was based on the Densely Connected Convolutional Networks (DenseNet) model and was trained and tested on publicly available FlickrLogos-32 dataset and achieved an accuracy of 92.8%.

Fehérvári and Appalaraju (2018) formulated logo detection and recognition as a few-shot problem using a two-step solution; consisting of a universal logo detector and a logo recognizer. An experiment was performed on the logo dataset built by the authors; Product Logo Dataset (PL2K) and FlickrLogos-32 dataset. Three state-of-the-art logo detector architectures were used namely; Faster R-CNN (Ren et al., 2015), Single Shot Multibox Detector (SSD) (Liu et al., 2015), and YOLOv3. The PL2k dataset had a recall rate of 97.73%. 93.52% and 92.10% on SSD, Faster R-CNN, and YOLOv3 respectively while on FlickrLogo-32 proposed technique recorded 60,04%, 79.87%, and 44.69% respectively.

Chen et al. (2018) proposed a vehicle logo recognition based on a capsule network using contained two convolutional layers and a fully connected network. Capsules are learned by a routing process, which is more effective than the pooling process in CNNs as it eliminates the use of too many parameters. The performance of the proposed model and CNN was evaluated on publicly available datasets. The CNN gave an accuracy of 99.35% while the capsule network recorded 100% accuracy which showed the robustness of the capsule model compared to CNN.

Pimkote and Kangkachit (2018) images of alcohol brand logos were classified using CNN. Initially, an all-in-one classifier was built to classify both alcohol and non-alcohol. The authors went further to build a classifier that could identify between each brand of alcohol by the logo. The proposed CNN architecture was evaluated on self-collected datasets of images from various websites. The classification performance was measured using precision and recall. The overall experimental result showed that the CNN recorded 89.17% and 89.16% precision and recall rates respectively.

Sharma et al. (2018) discussed the potential of deep learning-based object detectors. The Faster R-CNN and You Only Look Once version 2 (YOLOv2) were examined for automatic detection of signatures and logos from scanned administrative documents. Four different network models namely the Zhou Fan model (ZF model), Visual Geometry Group 16 Layers (VGG16), Visual Gait Generative Mode (VGGM), and YOLOv2 were considered for analysis to identify their potential in document image retrieval. The proposed approach detected signatures and logos simultaneously and achieved an accuracy of 80.2%.

Yufeng and Bo (2018) proposed a bicycle character identification method based on the Haar classifier; with the specific problem of changing ambient lighting and various bicycle types in the actual scene. Image filtering was adopted to eliminate noise interference in the environment. Secondly, image graying and normalization were performed to obtain a grayscale image of uniform sizes. Then, Haar-like features in the image were extracted, and the integral image was used to speed up feature extraction efficiently. Finally, the classifier was generated by using the AdaBoost algorithm achieving the highest accuracy record of 96.03%.

A CNN-based logo recognition scheme was proposed byWang et al. (2018) focusing on the shape context as a measurement for characterizing similarity between the target logo and others. The experimental results showed a 10% increase in performance on real images when compared to hand-crafted feature-based methods such as scale-invariant feature transform (SIFT).

A system for detection localization of multiple instances of trademark logos in sports videos was proposed by Manger et al. (2015), SIFT features and the local geometry of neighbouring features were considered to differentiate between different logos with ambiguous local features such as text-based logos. In contrast to other approaches, the technique did not rely on a training phase and no labeled data with annotated or absent logos was needed. The authors' approach focused on images of sports videos affected by compression artifacts, motion blur, small object sizes, occlusion, and several other artifacts. the system outputs 130 true positives and 436 false positives for the 7,323 frames.

Alaei and Delalandre (2014) proposed a system that first detects a few regions of interest (logo-patches), which likely contained the logos, in a document image. The detection method was based on the piecewise painting algorithm (PPA) and some probability features along with a decision tree. For logo recognition, a template-based recognition approach was proposed to recognize the logo which may be present in every detected logo-patch. The proposed logo recognition strategy used a search space reduction technique to decrease the number of template logo-models needed for the recognition of a logo in a detected logo-patch. The features used for search space reduction were based on the geometric properties of a detected logo-patch.

Llorca et al. (2013) proposed a vehicle logo recognition approach that is based on Histograms of Oriented Gradients (HOG) and Support Vector Machines (SVM) classifier. The system was specifically devised to work with images supplied by traffic cameras where the logos appeared with low resolution. A sliding window technique combined with a majority voting scheme was used to provide the estimated car manufacturer. The final global performance obtained was 92.59%.

Bianco et al. (2017)proposed a deep learning-based method for logo recognition. The approach involved a recognition pipeline comprising a logo region proposal and a specially trained CNN for logo classification, even with imprecise localization. The method was evaluated using the FlickrLogos-32 database, exploring the impact of synthetic versus real data augmentation and image pre-processing on recognition performance. Additionally, the paper systematically investigated the benefits of various training choices, including class-balancing, sample-weighting, and explicit modeling of the background class. Experimental results confirmed the feasibility of the proposed method, demonstrating its superiority over state-of-the-art approaches in logo recognition. The highest accuracy of 0.958 was obtained.

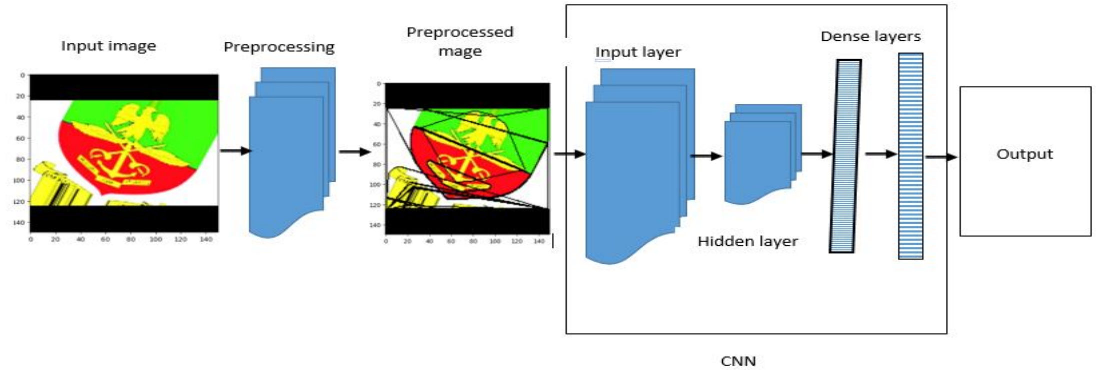
Hou et al. (2023) surveyed deep-learning solutions for logo detection. It is worthy noting that existing works on logo detection are divided across different scenarios of small-size logo detection, multi-scale logo detection, logo detection in complex scenes, long-tail logo detection, open-set logo detection, few-shot logo detection, etc. The authors identified applications of logo detection to include; brand monitoring, intelligent transporting, and document categorization. Logo classification was divided into traditional machine learning and deep learning-based approaches depending on the feature extraction style. It was concluded that logo detection has been efficient with great achievement however, further improvement is still required.

Pan et al. (2016) proposed a solution to the difficulty in identifying TV logos as a result of factors like illumination, occlusion, noise, etc. the authors employed the use of efficient CNN and MSER algorithms for extraction of candidate frames, the geometric constraint was also designed to remove non-logo objects from the candidate frames. Results obtained from the study show that the proposed technique outperformed previous state-of-art methods. This study is based on the technique by Pan et al. (2016) but however skipped the geometric constraint recommended by the authors. This was done to achieve a more simplified technique with comparable performance.

Kosala et al. (2023) in their study presented a method to detect logos in vehicles using a combined technique of MSER, Vertical Sobel, and Histogram of Oriented Gradient (HOG) with SVM. The proposed technique was able to eliminate the limitation of recognizing objects of different sizes. Results showed that the proposed approached significantly improved accuracy and computation time. The technique used handcrafted means for classifying extracted features for recognition. This approach fails usually when the volume of data goes into millions; hence the need for a more robust recognition algorithm.

PROPOSED METHODOLOGY

The pipeline that integrates MSER, CNN, and Synthetic Minority Over-sampling Technique (SMOTE) oversampling involves a series of steps to enhance the performance of a logo recognition system, especially in scenarios with imbalanced datasets. This study presents an MSER efficient and robust detection method widely used in text detection and is categorized as a connected component (Karanje, 2018). MSER was selected as the detection method due to its fairly light run-time complexity of; $O(n\,log(log(n)))$ where n is the number of image pixels (Pan et al., 2016). The robustness of to blur scale also makes the method more advantageous in processing captured images (Pan et al., 2016). This step produced the connected component and put a polygon around the component detected which helps the target of the proposed CNN-based classifier for feature extraction and logo classification. The architecture of the proposed model for university logo detection is illustrated in figure 1, the CNN and MSER are used in the logo detection process, it consists of several key components and layers which are briefly described.



Input image

The structure of this input depends on the specific architecture of the CNN and the requirements of the task it is designed for. Generally, a colour image is represented as a three-dimensional tensor, where the dimensions correspond to width, height, and color channels. For this research, the input image has been resized to (50, 150, 3). The distribution of our dataset is explained subsequently.

Dataset

Presently Nigeria has 51 Federal Universities (National Universities Commission, 2021), 62 State Universities, and 47 Private Universities (NUC, 2020); making it a total of 160 universities in Nigeria. Self-gathered dataset was used for this study. The self-gathered data is based on logos from selected Nigeria universities. It is difficult, time, and resource-consuming to obtain logos from the universities directly, hence the use of the internet to obtain logo images from selected university's websites. The dataset consists of 10,918 images of 24 Nigeria university logos. The selected universities were included because of the relatively adequate number of logos obtained for the experiment. Table 1 and Figure 2 give a summary of the description and visual of the logo dataset collected, the number of images gathered, and the names of the universities. All images were gathered from the google search images taking into consideration creative common licenses of the image.

Table 1 Summary of Dataset

S/No	UNIVERSITY	No Of IMAGES
1	Ahmadu Bello University Zaria (ABU)	640
2	Abubakar Tafawa Balewa University of Technology (ATBU)	487
3	Bayero University Kano (BUK)	489
4	Federal University Dutsin-Ma (FUDMA)	480
5	Federal University, Lokoja (FULOKOJA)	461
6	Federal University of Agriculture, Abeokuta (FUNAAB)	489
7	Federal University of Technology Akure (FUTA)	480
8	Federal University of Technology Minna (FUTMINNA)	484

9	Federal University of Technology Owerri (FUTO)	486
10	Modibbo Adamawa Federal University of Technology, Yola (MAUTECH)	475
11	Michael Okpara University of Agriculture (MOAU)	483
12	Nigerian Defence Academy (NDA)	461
13	National Open University of Nigeria (NOUN)	474
14	University Of Nigeria Nsukka (NSUKKA)	482
15	Obafemi Awolowo University (OAUIFE)	469
16	Federal University of Agriculture, Makurdi (UAM)	455
17	Usman Dan Fodiyo University Sokoto (UDUSOK)	449
18	University of Abuja (UNIABUJA)	460
19	University of Benin (UNIBEN)	396
20	University of Calabar, Calabar (UNICAL)	112
21	University of Ibadan (UNIIBADAN)	273
22	University of Ilorin (UNIILORIN)	575
23	University of Maiduguri (UNIMAID)	371
24	University Of Port Harcourt (UNIPORT)	487
	TOTAL	10918

Figure 2: Visual Samples of Selected Federal University Logo of Nigeria

Data augmentation enables practitioners to significantly increase the diversity of data available for training models, without collecting new data. For data augmentation in large neural networks, techniques such as cropping, padding, and horizontal flipping are commonly used for training (Ho et al.,

2019). Image augmentation is useful in artificially creating variations in images to expand any existing dataset, this creates new and different images from the existing image dataset that represent a comprehensive set of possible images. In training a model for logo detection, large data sets are required but log datasets suffer from scarcity hence the need for data augmentation to reduce the problem of scarcity (Hou et al., 2023).
In the dataset used for the experiment, the images of the university logos appeared zoomed in, zoomed out, and cropped (Figure 3). It is not practical to see logos that are flipped; if a logo with a direction other than facing up is used to test a network, the network will fail as it can only recognize logos facing up. In this study, rotation augmentation is used to help the network detect logos in different orientations, to avoid this problem. Although, sufficient sample images were not available; the study adopted a manual augmentation strategic and generated more samples using rotation techniques.

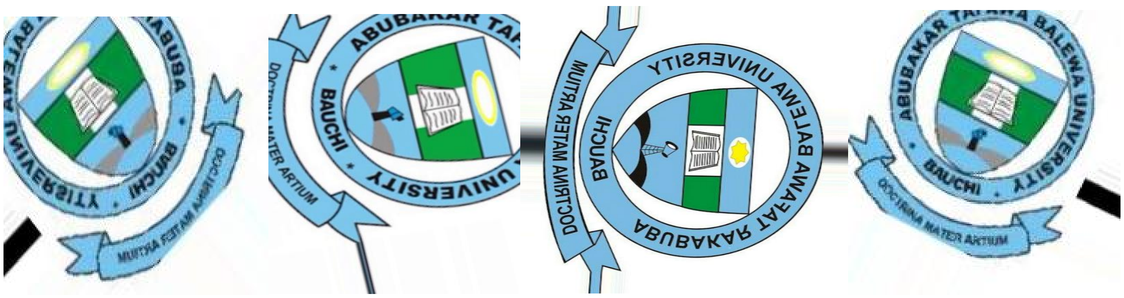


Figure 3: Sample of Image of a Selected University Logo Augmented

Data Preprocessing

The MSER is introduced for detecting connected components from the input logo image (Pan et al., 2016) after which the CNN model detects and classifies logos to the corresponding group. This study, however, is varied from the work by Pan et al. (2016) by not introducing the geometric constraint. The algorithm is applied to the input image as shown in figure 1, this extracts candidate boxes on the image, and detects, and collects distinct regions found in the image. This step produces the connected component and the output of the process is an image with a polygon like shape drawn around the detected components. The drawn boundary is useful for feature extraction and classification of logo in the CNN model.
Oversampling is a technique used in the field of machine learning to address class imbalance in a dataset. Class imbalance occurs when one class of data (the minority class) is significantly underrepresented compared to another class (the majority class). In such cases, machine learning models tend to perform poorly because they may be biased toward the majority class and have difficulty making accurate predictions for the minority class (Mohammed et al., 2020).
Oversampling is a method to mitigate this issue by increasing the number of instances in the minority class. This is typically done by creating synthetic or duplicate instances of the existing minority class data or by generating new data points that are similar to the existing minority class data, (Perez-Ortiz et al., 2016) Synthetic Minority Over-sampling Technique (SMOTE) was applied in this study to address the imbalanced dataset. It is applied after the dataset was split into three categories of training, validation and testing. Each category is run through the oversampler to correct the imbalance of the dataset since the spliting is stratified.

CNN-Based Classifier

The CNN-based classifier network architecture as applied in Bianco et al. (2017) for classification of logos from the Flickr-32 and logos-32plus dataset was used in this paper. The network architecture was first used by Krizhevsky et al. (2017), and it has proven to be of high performance on small RGB images. The network architecture was developed starting with three convolutional layers interleaved by ReLU (Rectified Linear Unit) non-linearities and pooling layers. After every pooling block, the data dimensions are cut in half by the pooling layers. The last layer of the network is made up of two fully-connected layers and a softMax classifier. The parameters were tweaked and layers were added until the desired outcome was reached. Table 2 presents the entire network structure

Table 2: Network Architecture

	Layer type	Layer name	Number of neurons	Activation Function	Hyperparameters
1	convolutional layer	Conv2D	16	Relu	448
2	convolutional layer	Conv2D	16	Relu	2320
4	Pooling layer	poolmax	-	-	0
5	Dropout	dropout	-	-	0
6	convolutional layer	Conv2D	32	Relu	4640
7	convolutional layer	Conv2D	32	Relu	9248
9	Pooling layer	maxpool	-	-	0
10	Dropout	dropout	-	-	0
11	convolutional layer	Conv2D	32	Relu	9248
12	convolutional layer	Conv2D	32	Relu	9248
14	Pooling layer	maxpool	-	-	0
15	Dropout	dropout	-	-	0
16	convolutional layer	Conv2D	64	Relu	18496
17	convolutional layer	Conv2D	64	Relu	36928
18	Dropout	dropout	-	-	0
19	Fully Connected of size	dense	128	Relu	524416
20	Dropout	dropout	-	-	0
21	Fully Connected of size 24	dense	24	softmax	3096

RESULTS AND DISCUSSION

This implementation and analysis of the result which includes the dataset distribution, training output that helps in visualizing the performance of the model during training and testing, and the accuracy of the model on the unseen data are discussed in this section. Figure 4 shows how a sample image is transformed after different regions of connected components were detected using the MSER algorithm. The logo image can be seen before passing through MSER and after detection by the algorithm, a bounding box with a red line is drawn around the detected regions this is to help emphasize the target of the proposed CNN-Based Classifier for feature extraction and logo classification.

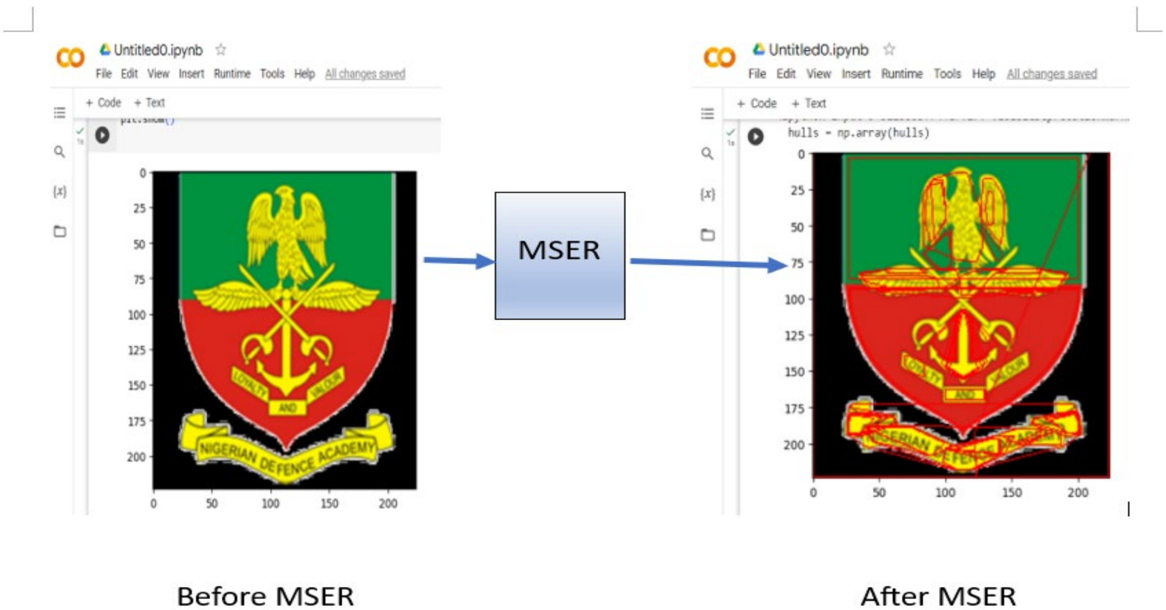


Figure 4: Sample Logo Image before and after passing through MSER

Dataset split

The dataset consists of 10918 images from 24 classes, this is split into three (3) categories; training set, validation set, and testing set. The dataset contained unbalanced classes which was addressed by applying the SMOTE sampling algorithm to avoid the bias of classifying the class with the highest representation in the dataset.

Training Set

The training set consists of 50% of the dataset totalling 5459 images. This class was then oversampled to achieve a uniform number of images per class and 7680 images were obtained. It is a critical component of the machine learning workflow and serves as the foundation for the model's learning process. Figure 5 shows the data distribution of the training set before and after over-sampling.

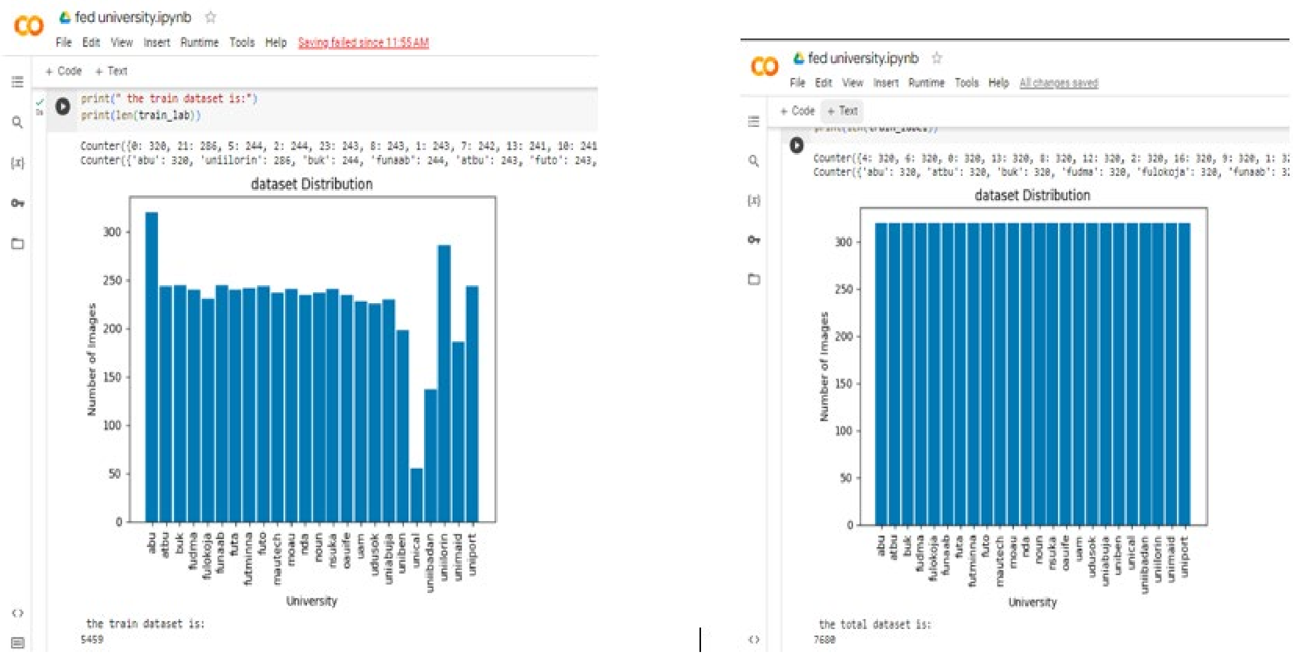


Figure 5: Training set Distribution Before and After Oversampling

Validation Set

The validation set consists of 40% of the dataset which is equal to 4368 images. and 6144 images after over-sampling. This is a subset of the data that is used during the training process to assess the performance of a machine-learning model and make decisions regarding its hyperparameters and configuration. It serves as an intermediary step between the training set and the testing set. Figure 6 shows the distribution of the validation set before and after over-sampling.

Figure 6: Validation set Distribution Before and After oversampling

Testing Set

The testing set consists of 10% of the dataset, amounting to 1092 images and 1536 images after over-sampling. Testing set is crucial for evaluating how well the model generalizes its learned patterns to new, unseen data. This data was used to assess the performance of the model after it had been trained on a separate dataset. Figure 7 shows the distribution of the testing set before and after over-sampling.

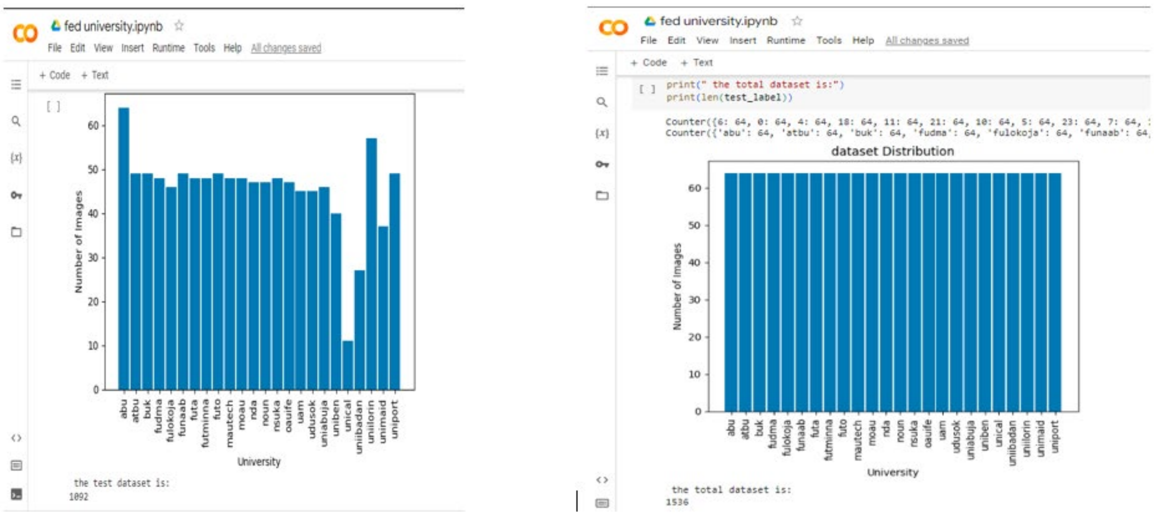


Figure 7: Testing Set Distribution Before and After oversampling

Model Evaluation

This section presents the performance of the model on the unseen data (that is, testing data), a plot of training accuracy and validation accuracy, and the plot of training and validation loss for evaluation.

Testing the model

Figure 8 shows the code snippet for evaluating the performance of the model and the output of the testing. The test score of 0.1391 and test accuracy of 0.9648 which is equivalent to 96.48% was achieved.

```
score = model.evaluate(train_datagen.flow(new_test, test_label, shuffle = False), verbose=2)
print('Test score:', score[0])
print('Test accuracy:', score[1])

48/48 - 19s - loss: 0.1391 - accuracy: 0.9648 - 19s/epoch - 395ms/step
Test score: 0.13910046219825745
Test accuracy: 0.96484375
```

Figure 8: Output for model Testing

The plot of training and validation accuracy for the first part of the training session at 50 epochs was obtained, this is illustrated in Figure 9. The blue curve represents the training accuracy while the brown curve indicates the validation accuracy, the two curves rose with an increase in epoch and got the highest accuracy at epoch 50. The figure shows that the model performed well without overfitting.

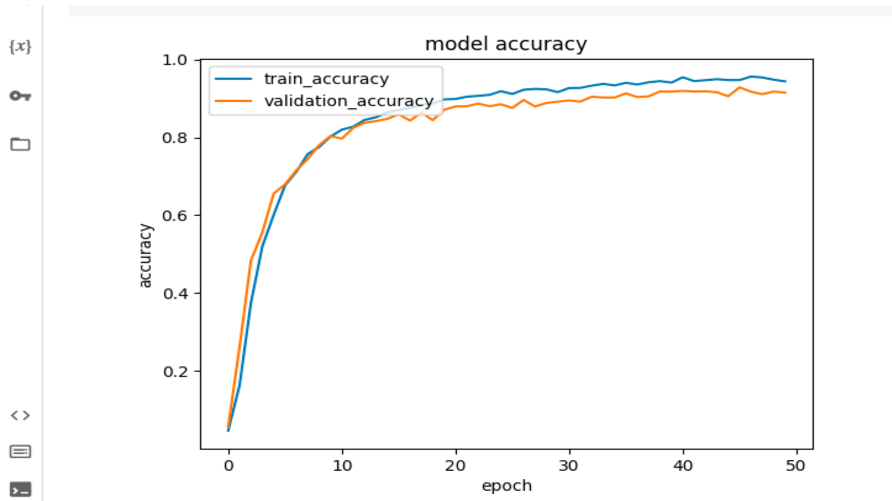


Figure 9: Training and Validation Plot

The performance of the model after running the experiment is visualized in the confusion matrix as displayed in Figure 10. The matrix shows the true class against the predicted class representing all 24 categories of selected Nigeria Federal University logos. This shows how the model correctly predicts a positive class instance as positive, a true positive occurs when the actual condition is positive, and the model's prediction is also positive. From the figure each class has equal numbers of predicted logos (64) this is a result of balancing the dataset.

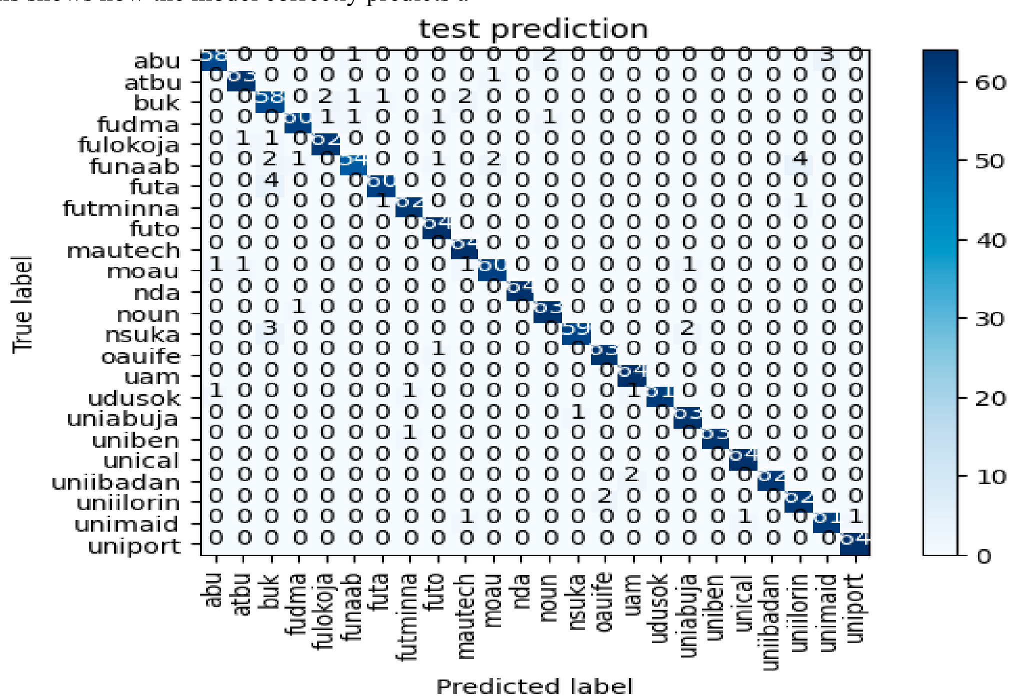


Figure 10: Confusion Matrix for the proposed model

The dark blue diagonal lines show the correctly classified number of logos which is also referred to as true positive values. Six (FUTO, MAUTECH, NDA, UAM, UNICAL, and UNIPORT) of the twenty-four classes of university logos reported a true positive value of 64 which indicates the images were all classified to belong to the right class of university. while the other eighteen classes reported varying numbers of incorrectly identified logos that were misclassified into other categories. This demonstrates the proposed model's capability and effectiveness in logo identification.

Figure 11 is the model's performance showing the true positive, false positive, and false negative rates of the experiment.

class	True Positive
abu	58
atbu	63
buk	58
fudma	60
fulokoja	62
funaab	54
futa	60
futminna	62
futo	64
mautech	64
moau	60
nda	64
noun	63
nsuka	59
oauife	63
uam	64
udusok	61
uniabuja	63
uniben	63
unical	64
uniibadan	62
uniilorin	62
unimaid	61
uniport	64

(a) True Positive

class	False Positive
abu	2
atbu	2
buk	10
fudma	2
fulokoja	3
funaab	3
futa	2
futminna	2
futo	3
mautech	4
moau	3
nda	0
noun	3
nsuka	1
oauife	2
uam	3
udusok	0
uniabuja	3
uniben	0
unical	1
uniibadan	0
uniilorin	5
unimaid	3
uniport	1

(b) False Positive

class	False Negetive
abu	6
atbu	1
buk	6
fudma	4
fulokoja	2
funaab	10
futa	4
futminna	2
futo	0
mautech	0
moau	4
nda	0
noun	1
nsuka	5
oauife	1
uam	0
udusok	3
uniabuja	1
uniben	1
unical	0
uniibadan	2
uniilorin	2
unimaid	3
uniport	0

(c) False Negative

Figure 11: TP, FP, and FN Rates of Model's Performance

The average mean of precision, recall, and F1 Score obtained from testing the proposed model across all twenty-four logo classes is displayed in Figure 12. The average values for precision stood at 0.963 while recall and F1 score reported 0.962 each.

```
[40] import statistics
     mAP = statistics.mean(precision)
     print (f'mean avearage precision is ', mAP)

     mean avearage precision is  0.9628198399004961

[41] mAR = statistics.mean(recall)
     print (f'mean avearage recall is ', mAR)

     mean avearage recall is  0.9622395833333334

     mAF1 = statistics.mean(F1)
     print (f'mean avearage recall is ', mAF1)

     mean avearage recall is  0.9620985446873259
```

Figure 12: Mean Average Values for F1, Recall, and Precision

The proposed model was compared with selected state-of-the-art models using accuracy, precision, recall, and F1-measures as basis. Table 3 shows the performances obtained.

Table 3: Comparative Analysis with State-of-the-art Models

Architecture	Accuracy	Precision	Recall	F1-Measure
MSER+ geometric constraint + CNN (Pan et al., 2016)	0.93	-	-	-
MSER + vertical sobel + CNN (Kosala et al., 2023) (as reported)	-	93.97	82.26	87.72
MSER + CNN (ours)	0.9626	0.9626	0.9622	0.9620

CONCLUSION

In conclusion, this paper presented an approach for logo recognition of selected Federal Universities in Nigeria, utilizing a combination of Maximally Stable Extremal Regions (MSER) and Convolutional Neural Networks (CNN). The results obtained from the experiments demonstrated the effectiveness of the proposed approach in accurately recognizing and classifying the logos of the selected Federal Universities in Nigeria. The combination of MSER and CNN effectively captured the salient features of the logos, including colour, shape, and texture, leading to high-performance recognition. The proposed system offers potential applications in various domains, including branding analysis, advertisement monitoring, and image retrieval. It can be used for marketing analysis, copyright enforcement, and educational research, among others. By accurately identifying and classifying logos, the system provides valuable insights into brand visibility and recognition.

REFERENCES

Alaei, A., & Delalandre, M. (2014). A complete logo detection/recognition system for document images. 2014 11th IAPR International Workshop on Document Analysis Systems, Tours, France.

Alsheikhy, A., Said, Y., & Barr, M. (2020). Logo Recognition with the Use of Deep Convolutionla Neural Network. *Engineering, Technology & Applied Science Research*, *10*(5), 6191-6194. https://doi.org/10.48084/etasr.3734

Bianco, S., Buzzelli, M., Mazzini, D., & Schettini, R. (2017). Deep learning for logo recognition. *Neurocomputing*, 23-30. https://doi.org/10.1016/j.neucom.2017.03.051

Brownlee, J. (2019). *A Gentle Introduction to Computer Vision*. Retrieved 22 November 2023 from https://machinelearningmastery.com/what-is-computer-vision/

Carvalho, P., Pereira, A., & Viana, P. (2021). Automatic tv logo identification for advertisement detection without prior data. *Applied Sciences*, *11*(16), 7494. https://doi.org/10.3390/app11167494

Chen, R., Jalal, M. A., Mihaylova, L. S., & Moore, R. K. (2018). Learning Capsules for Vehicle Logo Recognition. *2018 21st International Conference on Information Fusion (FUSION)*, 565-572. https://doi.org/10.23919/ICIF.2018.8455227

Cheng, Z.-Q., Wu, X., Liu, Y., & Hua, X.-S. (2017). Video ecommerce++: Toward large scale online video advertising. *IEEE transactions on multimedia*, *19*(6), 1170-1183. https://doi.org/10.1109/TMM.2016.2647386

Dixit, U. D., & Shirdhonkar, M. (2016). Automatic logo detection and extraction using singular value decomposition. 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India.

Donoser, M., & Bischof, H. (2006). Efficient maximally stable extremal region (MSER) tracking. 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06), New York, NY, USA.

Fehérvári, I., & Appalaraju, S. (2018). Scalable Logo Recognition Using Proxies. *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 715-725.

Hendrick, Wang, C.-M., Aripriharta, Jhe, C.-G., Tsu, P.-C., & Jong, G.-J. (2018). The halal logo classification by using NVIDIA DIGITS. 2018 International Conference on Applied Information Technology and Innovation (ICAITI), Padang, Indonesia.

Ho, D., Liang, E., & Liaw, R. (2019). *1000x Faster Data Augmentation*. https://bair.berkeley.edu/blog/2019/06/07/data_aug/

Hou, S., Li, J., Min, W., Hou, Q., Zhao, Y., Zheng, Y., & Jiang, S. (2023). Deep learning for logo detection: A survey. *ACM Transactions on Multimedia Computing, Communications and Applications*, *20*(3), 1-23. https://doi.org/10.1145/3611309

Irhebhude, M. E. (2015). *Object detection, recognition and re-identification in video footage* Loughborough University].

Kosala, G., Harjoko, A., & Hartati, S. (2023). MSER-Vertical Sobel for Vehicle Logo Detection. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *7*(5), 1239-1245.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *COMMUNICATIONS OF THE ACM*, *60*(6).

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S. E., Fu, C.-Y., & Berg, A. C. (2015). SSD: Single Shot MultiBox Detector. European Conference on Computer Vision,

Llorca, D. F., Arroyo, R., & Sotelo, M. A. (2013). Vehicle logo recognition in traffic images using HOG features and SVM. 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), The Hague, The Netherlands.

Manger, D., Müller, M., & Kubietziel, M. (2015). Filtering local features for logo detection and localization in sports videos. 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia.

Matas, J., Chum, O., Urban, M., & Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, *22*(10), 761-767.

Pan, D., Shi, P., Qiu, Z., Sha, Y., Zhongdi, X., & Zhoushao, J. (2016). TV Logo Classification Based on Convolutional neural network. International Conference on Information and Automation Ningbo, China.

Pimkote, P., & Kangkachit, T. (2018). Classification of alcohol brand logos using convolutional neural networks. 2018 International Conference on Digital Arts, Media and Technology (ICDAMT), Phayao, Thailand.

Ren, S., He, K., Girshick, R. B., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*, 1137-1149.

Sharma, N., Mandal, R., Sharma, R., Pal, U., & Blumenstein, M. (2018). Signature and Logo Detection using Deep CNN for Document Image Retrieval. *2018 16th International*

Conference on Frontiers in Handwriting Recognition (ICFHR), 416-422.

Wang, Y., Yang, W., & Zhang, H. (2018). Deep learning single logo recognition with data enhancement by shape context. 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil.

Yousaf, W., Umar, A., Shirazi, S. H., Khan, Z., Razzak, I., & Zaka, M. (2021). Patch-CNN: Deep learning for logo. *Journal of Intelligent & Fuzzy Systems*, 3849–3862.

Yu, Y., Guan, H., Li, D., & Yu, C. (2019). A cascaded deep convolutional network for vehicle logo recognition from frontal and rear images of vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *22*(2), 758-771. https://doi.org/10.1109/TITS.2019.2956082

Yufeng, D., & Bo, Z. (2018). Intelligent identification method of bicycle logo based on Haar classifier. 2018 5th International Conference on Systems and Informatics (ICSAI),

Zulkeefli, S. N. I., & Hashim, N. (2022). Comparison of CNN-based Algorithms for Halal Logo Recognition. *Journal of System and Management Sciences*, *12*(5), 155-168.