Image Augmentation and ResNet50: A Novel Approach for Car Logo Detection and Classification

by Kishan Rathod

Submission date: 01-May-2024 04:23PM (UTC+0530)
Submission ID: 2367723690
File name: G10.pdf (219.24K)
Word count: 2282
Character count: 12804

Image Augmentation and ResNet50: A Novel Approach for Car Logo Detection and Classification

Kishan Rathod[1], Vraj Patel[1], Abhi Bhimani[1], Ritika Chokshi[1], Trusha Patel[1]

[1]U & P. U. Patel Department of Computer Engineering, Chandubhai S Patel Institute of Technology (CSPIT), Faculty of Technology and Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Changa, India

Abstract:

This research focuses on developing a robust logo detection and recognition system for five car companies: Honda, Hyundai, Kia, Renault, and Suzuki. The dataset was manually generated for the five companies. To further diversify the dataset, three different image augmentation techniques were used: edge detection, color space transformation, and histogram equalization. After employing these data augmentation techniques, we were able to generate a more diverse dataset with different types of features, which helps to learn features in a more diverse manner. Furthermore, the dataset was used to train a pre-trained deep learning model known as ResNet50 for logo detection and classification. The results showcased the effectiveness and robustness of the proposed approach, achieving an impressive accuracy of 96% on the testing dataset. The model also showed a balanced performance across all classes, with precision, recall, and F1-scores ranging from 0.93 to 0.99. The consistent behaviour of the model in both training and validation, along with the confusion matrix analysis, highlights the robustness and generalization capability of the model. This research highlights the potential of combining image augmentation techniques with deep learning models for accurate logo detection and recognition tasks, which can be used in various application like industries including automotive, brand monitoring, and intellectual property protection.

1. Introduction

Logo detection and recognition have become an important part of various applications, including automotive industry, brand monitoring, and intellectual property protection. In the automobile sector, accurate logo detection and recognition helps in various type vehicle identification, asset management, and also targeted advertising strategies. However, developing a robust logo detection system poses several challenges because of variations in lighting conditions, viewpoints, and distortions and many other factors. Traditional logo detection approaches have been relied on the hand-crafted features and classical machine learning algorithms, such as matching of template, edge detection, and colour histograms [1][2]. Although these methods have shown promising results but they were used in a controlled environment and often struggled to generalize in real-world scenarios which have complex backgrounds, overlapping objects, and various types of illumination conditions. In recent years, deep learning techniques, particularly convolutional neural networks (CNNs), have demonstrated remarkable performance in various computer vision tasks, including object detection and recognition [3][4]. CNNs have the ability learn hierarchical representations from raw images, enabling them to capture visual patterns and features that are difficult to manually define using those traditional methods. This problem has led to a surge of interest in applying deep learning approaches to logo detection and recognition problems. Several studies has explored the use of CNNs for logo detection and recognition. Gomez et al. [5] proposed a two-stage approach in which a CNN is used for logo detection, followed by a second CNN for logo recognition. Their method achieved promising results on the FlickrLogos-32 dataset. Tuzun and Degerli [6] developed a CNN-based framework for real-time logo detection and recognition in video streams, which demonstrated its applicability in the broadcasting of sports matches. Despite these advancements, logo detection and recognition systems often face challenges when dealing with limited training data or imbalanced datasets, where certain logo classes have significantly fewer samples than others. So to overcome this problem, data augmentation techniques have been widely employed by artificially increasing the size and diversity of the training dataset [7][8]. In this research , we employed three different image augmentation techniques, i.e. edge detection, color space transformation, and histogram equalization. These techniques were applied to the original dataset, generating a larger and more diverse set of training samples. Subsequently, we leverage the power of deep learning by using a pre-trained deep learning model known as ResNet50 [9] for logo detection and classification.

2. Methodology

2.1 Data Generation

For this research, the data were manually generated for the five car companies, i.e Hondo, Hyundai, Kia, Renault, and Suzuki. For each of the car companies around 40-70 images were captured. Classifying car logos with this amount of data may result in underfitting and lower performance of our machine learning model. To increase the size of the dataset, we further applied the data augmentation technique.

2.2 Data Augmentation

To further extend our dataset, we used data augmentation techniques. In this study, three different image augmentation techniques were applied to the original dataset containing logo images from five car companies. These techniques included edge detection, color space transformation, and histogram equalization. The augmented images were generated using the Keras library's ImageDataGenerator class, which provided a convenient and useful way to apply various transformations to the input images very quickly and easily.

2.2.1 Edge Detection

Edge detection is a fundamental image processing technique that identifies boundaries or edges within an image. In this study, the Canny edge detection algorithm was applied to the grayscale version of the original images. The resulting edges detected in the images capture the essential structural information of the logos, which would be beneficial for our machine learning model to learn features based on the shapes of the logos.

2.2.2 Color Space Transformation

Color space transformation involves converting the representation of color information in an image from one color mode to another color mode. In this study, the original RGB (Red-Green-Blue) images were converted to the hue (Hue, Saturation, Value) color space. This transformation will help the model learn the features based on the colors because the HSV color space separates the color information (hue and saturation) from the intensity information (value).

2.2.3 Histogram Equalization

Histogram equalization is one of the technique used to enhance the contrast of an image by redistributing the pixel intensities across the entire range of values. This technique was applied to the grayscale version of the original images, resulting in images with improved visibility of details and a better contrast between the logo and the background.

2.2.4 ImageDataGenerator Class

For each of the three augmentation techniques, 20 new images were generated from the original images using various transformations provided by the ImageDataGenerator class in the Keras library. These transformations included rotation, shifting, brightness, shearing, flipping, and centering. After data augmentation, the total number of images in each class is shown in Table 1.

Table 1. Number of images in each class after data augmentation

Class	Number of images after augmentation
Honda	1688
Hyundai	1343
Kia	1295
Renault	987
Suzuki	1119

2.2.5 Data Splitting

The augmented dataset was then split in 70-30 ratio where 70% of the dataset was used for training and 30% of the dataset was used for testing the pre- trained model. Furthermore, we divided the dataset into a 50:50 ratio for the validation set, where 50% of the dataset was used for training validation and 50% for testing validation. The split datasets were then used to train the pre-trained deep learning model called ReseNet50 for logo detection using the keras library.

2.2.6 ResNet 50

ResNet 50 is renowned for its exceptional depth, comprising 50 layers. The architecture is structured into several building blocks, each containing convolutional layers, batch normalization, ReLU activations, and residual connections. A total of 50 epochs were used to train the model. By exposing the model to a wider range of variations during training, the augmentation process improved the model's ability to generalize and accurately detect logos in new, unseen images. Overall, the image augmentation process played a crucial role in this research by increasing the size and diversity of the training dataset, which improved the performance and robustness of the logo detection model. The results of are shown in the next section.

3. Results

To evaluate the performance of the proposed logo detection model, we conducted experiments on an augmented dataset containing images of logos from five car companies. Honda, Hyundai, Kia, Renault, and Suzuki. The model was trained for 50 epochs, and to determine the training and validation accuracy and their

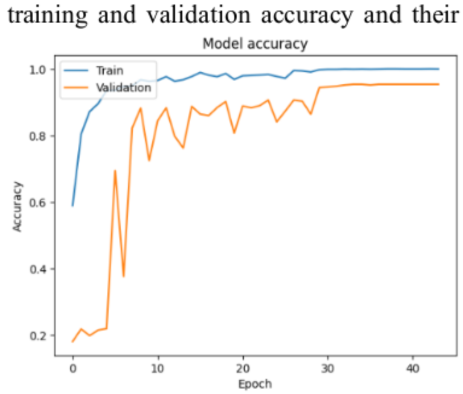

Figure 1. Model accuracy

performances, we plotted a graph for the 50 epochs which is shown in figure 1. From the figure, we can clearly see that after 30 initially for training data, the accuracy was around 0.6, but for the validation set it was nearly 0.2. However, as the model continued to be trained, both validation and training accuracy increased. After 30 epochs, both accuracy values remained stable, where the training accuracy reached 0.99 and the validation accuracy was 0.95. Furthermore, to observe the loss of the model, we also plotted another graph, as shown in Figure 2.

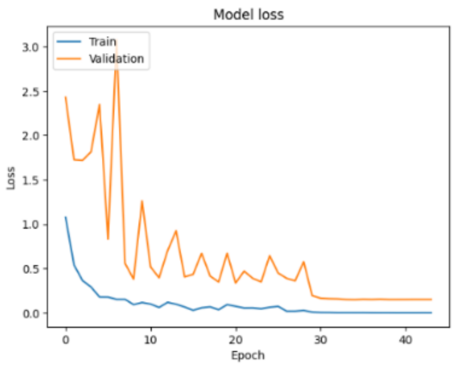

Figure 2. Model Loss

We can clearly see that the initial loss was high for both the training and validation sets. However, gradually decreased as the epochs ran and the model learned the features more accurately. Furthermore, we tested the model with 30% testing data and evaluated the performance of the testing data using evaluation metrics like precision, recall, F1 score, and accuracy, as shown in table 2.

Table 2. Evaluation Metrices

Class	Precision	Recall	F1-Score	Support
Honda	0.97	0.94	0.96	284
Hyundai	0.93	0.96	0.94	205
Kia	0.97	0.99	0.98	174
Renault	0.96	0.95	0.96	86
Suzuki	0.96	0.98	0.97	156
Accuracy			0.96	905
Macro average	0.96	0.96	0.96	905
Weighted average	0.96	0.96	0.96	905

From Table 1, we can see that the model achieved an impressive accuracy of 0.96, indicating that it correctly classified 96% of the logos across all classes in the testing set. The precision, recall, and F1-score of all classes ranged between 0.93 and 0.99, demonstrating the robustness of the model. The macro- and weighted average metrics also showcase a balanced view of the performance across classes because all classes yielded the same score of 0.96, demonstrating consistent performance. Furthermore, we plotted the confusion matrix for the testing dataset, as shown in Figure 2.

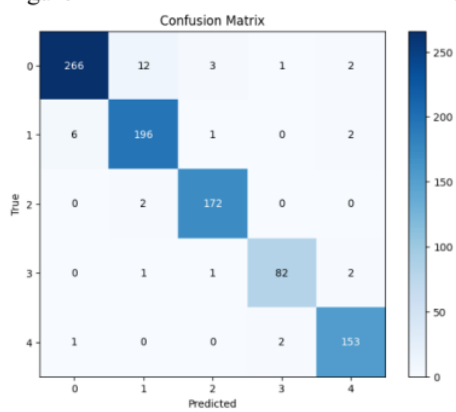

Figure 2. Confusion matrix of testing data

From the figure 2, we can see that all the classes have very few misclassifications and demonstrated a good ratio of correct classifications. Overall, the results demonstrate the effectiveness of the proposed approach, which combines image augmentation techniques and the ResNet50 architecture, in accurately detecting and classifying logos from various car companies. The high accuracy, balanced performance across classes, and consistent training/validation behaviour highlight the model's robustness and generalization capability.

4. Conclusion

In conclusion, this research aimed to develop a better system for spotting and recognizing five different car logos. We focused on five major companies: Honda, Hyundai, Kia, Renault, and Suzuki. We used image augmentation methods to make our dataset bigger and more diverse as it was manually made. This helped the model learn better and become more robust and efficient. We trained a deep learning model called ResNet50 on the data with a 70-30 split on the dataset. Our results showed that our approach worked well. The model could spot logos with an accuracy of 96% on the training dataset. This means that our method could be useful for identifying logos in real-world situations. Overall, our research shows that combining image augmentation with deep learning is effective for logo detection. Here, we tackle a common problem in computer vision: identifying logos. It is important because it can help in many areas like advertising and identifying cars. Traditionally, this was done using methods like template matching and edge detection, but they did not work well in real-world situations. Therefore, the use of deep learning approaches showed promise in this area of computer vision. Our method used a pre-trained model called ResNet50, which is known for its 50-layer convolution and reputed architecture. Using image augmentation, we made our dataset bigger and more varied. This helped the model learn better and perform well on different types of logos. Our results showed that the model could accurately detect logos from different car companies, even under challenging conditions like varied lighting and backgrounds. This suggests that our approach could be useful in practical applications where accurate logo detection is important. Overall, our research demonstrates the potential of combining deep learning with image augmentation for logo detection tasks.

5. References

[1] Romberg, S., Pueyo, L. G., Lienhart, R., & Van Zwol, R. (2011). Scalable logo recognition in real-world images. Proceedings of the 1st ACM International Conference on Multimedia Retrieval, 1-8.

[2] Sahbi, H., Puissant, A., & Sergent, G. (2005). Logo recognition by a fusion of color and shape descriptors. Pattern Recognition and Image Analysis, 3687, 639-646.

[3] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097-1105.

[4] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 580-587.

[5] Gomez, L., Malik, J., & Romera, M. (2020). Logo recognition and brand detection in images with deep learning. arXiv preprint arXiv:2006.03721.

[6] Tuzun, A., & Degerli, Y. (2021). Real-time logo detection and recognition in video streams using deep learning. IEEE Access, 9, 78541-78557.

[7] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.

[8] Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image

classification using deep learning. arXiv preprint arXiv:1712.04621.

[9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770-778.

Image Augmentation and ResNet50: A Novel Approach for Car Logo Detection and Classification

ORIGINALITY REPORT

10%	6%	9%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	"Computer Vision – ECCV 2016 Workshops", Springer Science and Business Media LLC, 2016 Publication	1%
2	Mayuri Popat, Sanskruti Patel, Yagnik Poshiya, Achyut Krishna Sai. "Brain Tumor Image Segmentation using Box-Unet Architecture", 2023 First International Conference on Advances in Electrical, Electronics and Computational Intelligence (ICAEECI), 2023 Publication	1%
3	koreascience.or.kr Internet Source	1%
4	Md. Rahad Islam Bhuiyan, Sami Azam, Sidratul Montaha, Risul Islam Jim et al. "Deep learning-based analysis of COVID-19 X-ray images: Incorporating clinical significance and assessing misinterpretation", DIGITAL HEALTH, 2023 Publication	1%

5 doaj.org 1%
Internet Source

6 Faisal Binzagr. "Explainable AI-driven model for gastrointestinal cancer classification", Frontiers in Medicine, 2024 1%
Publication

7 ijsrst.com 1%
Internet Source

8 Dagmara Stasiowska. "Rocket launch induced G-force impact on Apis mellifera queens' fertility - results of a pilot study", Research Square Platform LLC, 2022 1%
Publication

9 Tan, Khang Siang, Wei Hong Lim, and Nor Ashidi Mat Isa. "Novel initialization scheme for Fuzzy C-Means algorithm on color image segmentation", Applied Soft Computing, 2013. 1%
Publication

10 Pavlo Tertychnyi, Cagri Ozcinar, Gholamreza Anbarjafari. "Low-quality fingerprint classification using deep neural network", IET Biometrics, 2018 1%
Publication

11 Ye Meng, Sujuan Hou, Jing Wang, Weikuan Jia, Yuanjie Zheng, Awudu Karim. "An Adaptive 1%

Representation Algorithm for Multi-scale Logo Detection", Displays, 2021
Publication

12	repositories.lib.utexas.edu Internet Source	1 %

Exclude quotes	Off	Exclude matches	< 1%
Exclude bibliography	On		