Adaptive Logo Recognition System

Dr.M. Rajababu[1], M.D.V.S. Lakshmi[2], P.Sai Sudha [3], K.H.V.Ravi Chandra[4], V.Mohan Sai Venkat[5]

[1,2,3,4,5]Department of IT,Aditya Engineering College(A), Surampalem, India
rajababu.makineedi@aec.edu.in

***Abstract*—** Logo Detection plays a vital role in different applications such as brand checking, copyright assurance, and visual look. In this paper, we explore the viability of convolutional Neural Networks (CNNs) and combination procedures for Logo Detection. We used three pre-trained CNN models VGG-16, DenseNet-201, and ResNet-50. Each model was applied independently on two datasets accomplishing the most extreme accuracies of 92.5% with DenseNet-201, 92.07% with VGG-16, and 90.45% with ResNet-50. Besides, we investigated the fusion model-VDeRe-24 to combine forecasts from these models, coming about in a combined result with an accuracy 94%. Our experimental findings highlight the potential fusion CNN model in making strides in Logo Detection accuracy, comparing with the state-of-the-art techniques for brand checking and copyright, assurance applications.

Keywords- Convolutional Neural Network, Feature Extraction, Feature Fusion, Logo detection.

1. Introduction

In later a long time, with the multiplication of computerized media and online stages, Logo Detection has ended up progressively vital for different applications such as brand observing, copyright assurance, and visual look. The capacity to precisely identify logos in pictures is fundamental for businesses to defend their brand personality and mental property rights, as well as for analysts and designers to construct compelling visual look frameworks.

Convolutional Neural Systems (CNNs) have developed as effective apparatuses for picture acknowledgment errands, and counting Logo Detection. These profound learning models have illustrated exceptional execution in recognizing complex designs and highlighting inside pictures, making them well-suited for Logo location errands. Among the available pre trained CNN models such as VGG-16, ResNet-50 and DenseNet-201 have appeared especially promising comes about in image classification assignments. Then we implemented fusion model VDeRe-24 with the combination of all three models.
In this consideration, we point to explore the viability of VGG-16, DenseNet-201, ResNet-50 and VDeRe-24 structures for Logo Detection. We prepare these models on a datasets Logo Detection dataset-LDDST and Small Logo Detection Dataset-SLDS of Logos, leveraging their pre-trained weights on large-scale picture datasets.

In the next section, we discuss about the literature review traversing the historical landscape of logo detection methodologies, before transitioning to our proposed framework. Here we unveil about our innovative approach, leveraging three pre trained CNN Models. Amidst this exploration, we introduce a ground-breaking approach-fusion model to elevate accuracy, epitomized by the pioneering VDeRe-24 fusion model. Our research further unfolds in the experiments and results section, where rigorous testing on the real-world datasets validates the efficacy of the methodology. Finally, we synthesize

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_59

our findings in the conclusion, shedding light on the broader implications of our research for logo detection.

2. LITERATURE REVIEW

The identification of logos, which can comprise text, shapes, images, or combinations [3] thereof, serves numerous purposes across diverse fields such as intelligent transportation, social media monitoring, and infringement detection. The central aim of logo detection is to ascertain the precise location of a specific logo within images or videos and accurately classify it. Previous logo detection methods are mostly handcrafted methods [5][7] proposed extracting texture and color features from original and augmented [2] images to detect logos in images. Early approaches to Logo Detection frequently depended on handcrafted highlights and conventional machine-learning calculations [5].

The main problem that has been faced while classifying depending upon whether they contain logo or not, and the main goal is to determine the origin of document with accuracy of the logo detection [1]. These strategies included extricating key visual highlights from pictures, such as color histograms, surface descriptors, and shape characteristics, and after that utilizing classifiers to recognize logos from foundation clutter [9]. Indexing the detected logo candidate region to a cloud for classifying or for concluding that the region is not of interest.Detection by empowering end-to-end learning of highlights specifically from picture information [4]. CNN-based approaches [5] have appeared prevalent execution in symbol discovery assignments, outperforming conventional strategies [5] in exactness and vigor.

The emergence of deep learning, particularly CNNs [6] has significantly improved logo detection accuracy. Among CNN architectures, VGG-16, DenseNet-201 [8] and ResNet have emerged as popular choices. VGG-16 offers a solid foundation for image classification tasks with its deep design and relatively simple architecture. DenseNet[8], known for its dense network design and feature reuse, demonstrates efficient parameter utilization and has shown advancements in highlighting feature propagation[7]. ResNet, by introducing residual connections that mitigate the vanishing gradient problem, has achieved impressive performance in learning deep representations. On the LDDST[14][16] and SLDS[15] datasets, our suggested approach delivers state-of-the-art accuracy, possibly surpassing the capacities of these well-known CNN architectures for logo detection tasks.

A few ponders within the writing have investigated the utilization of pre-trained CNN models [6] for Logo Detection [3], fine-tuning and fast template matching [11] these models on Logo datasets to adjust them to particular acknowledgment errands. Exchange learning strategies, where pre-trained models[9] are utilized as highlight extractors and fine-tuned[10] with task-specific information, have demonstrated compelling in symbol discovery assignments with constrained preparing information.[8].In general, the writing survey highlights the advancement of Logo Detection strategies from handcrafted highlights to profound learning-based approaches[10]. It underscores the significance of CNN structures like VGG-16, DenseNet-201, and ResNet-50 in progressing the state-of-the-art in symbol location[6], as well as the potential of outfit strategies for advanced moving forward discovery execution[8].

3. PROPOSED METHOD

The Logo Detection system we propose utilizes three efficient Convolutional Neural Network (CNN) architectures: VGG-16, DenseNet-201, and ResNet-50, leveraging their respective attributes. Each of these models has illustrated fabulousness in picture acknowledgment assignments, making them appropriate candidates for Logo Detection.

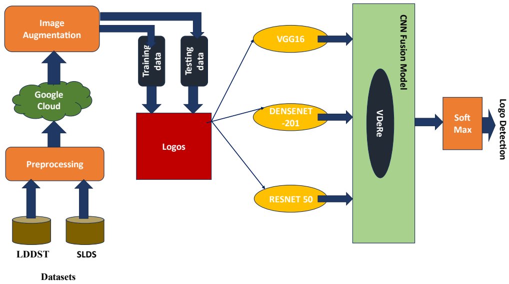

Fig.1 .Block Diagram for fusion model of ALORS

The system starts by pre-training the VGG-16, DenseNet-201 and ResNet-50 models on large-scale picture datasets to memorize nonexclusive visual representations. These pre-trained models are at that point fine-tuned on logo-specific datasets to adjust them to the errand of Logo Detection. Fine-tuning includes altering the demonstrated parameters to capture the particular highlights of logos, such as colors, shapes, and surfaces.

Once the personal models are fine-tuned, they are coordinated into a combination system to combine their forecasts successfully. A combination of expectations from different models permits the accumulation of differing data and upgrades the general discovery exactness and vigor.

In our proposed system, we utilize a late combination methodology, where the forecasts from VGG-16, ResNet-50, and DenseNet-201 are combined into a fusion model VDeRe-24 at the yield layer employing a weighted averaging approach. The weights allocated to each model's forecast are learned amid preparing and optimizing them to maximize discovery execution on the approval dataset. Figure-1 shows the bock diagram for fusion model of ALORS.

The coordinates demonstrate, comprising of the combined forecasts from ResNet-50, VGG-16 and DenseNet-201 shapes the spine of our Logo Detection framework. This bound-together show is competent in leveraging the complementary qualities of each person's CNN design, coming about in moving forward the discovery of exactness and flexibility to varieties in symbol appearance.

Our proposed system offers a flexible and viable arrangement for Logo Detection errands, profiting from the collective insights of different state-of-the-art CNN models. By coordinating assorted show models and utilizing combination methods, our system accomplishes prevalent execution in logo detection over real-world scenarios. Figure-2 shows the architecture of convolutional neural networks.

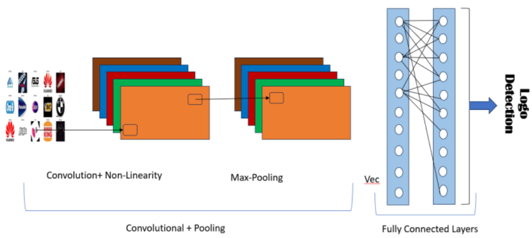

Fig. 2. CNN Architecture

A. Datasets

This study uses publicly accessible dataset LDDST[11] and SLDS[12]. The LDDST(a) included a total of 4334 logos from two classes and 555 logos from two classes. In SLDS(b) included a total of 754 logos from two classes and 270 classes from two classes.

(a)sample images from LDDST dataset

(b)sample images from SLDS dataset

B. Data Augmentation

In this we have a great images but might not be enough to train our image-recognizing program(CNN model)[7] effectively. To help for that learn better and we created more pictures from originals. Resize them, slightly rotate and even shift or stretch them a bit. There also flip them and add some fuzz or blur. By making more images in these ways, we gave wider variety to learn boosting accuracy for real-world images.

C. Training and Testing of Proposed datasets

When the Image augmentation is completed, the LDDST dataset get the total logos of 4889, SLDS dataset get the total logos of 1024. For these finished augmented photos, there is a 70:30 ratio between the training and testing datasets. The best results in deep learning are obtained with this splitting ratio. These training and test logos were then used to put CNN through the next level of training and testing. Table 1 displays the separation of two datasets into train and test.

S. No	DataSet(proposed)	Training set	Testing set	Total Logo Images
1.	LDDST	4334	555	4889
2.	SLDS	754	270	1024

Table 1: Training and Testing of our proposed datasets

4. EXPERIMENTAL RESULTS

A. Performance of Individual Models:

VGG-16: The VGG-16 model achieved an accuracy of 92% on the test dataset, demonstrating its effectiveness in logo detection. Despite its simplicity compared to more modern architectures, VGG-16's deep convolutional layers enable it to capture complex patterns and features exhibited in logos.

DenseNet-201: DenseNet-201 surpassed VGG-16's performance, achieving a precision of 92.5% on the test dataset. The densely connected layers in DenseNet-201 facilitate feature reuse and enable the model to learn more discriminative representations, resulting in higher detection accuracy.

ResNet-50: While slightly trailing behind DenseNet-201, ResNet-50 still performed admirably with an accuracy of 90.45% on the test dataset. Its residual connections mitigate the vanishing gradient problem, allowing more effective training of deeper networks.

S. No	CNN model	Accuracy in %	Precision	F1 score
1.	VGG-16	92.07	92.33	91.2
2.	DenseNet-201	92.5	93.66	91.9
3.	ResNet -50	90.45	89.88	90.1
4.	VDeRe-24	94.03	93.74	92.8

Table 2. comparison of Accuracy, Precision and F1 score on LDDST dataset.

B. Performance of Fusion Model:

By integrating the predictions of VGG-16, DenseNet-201 and ResNet-50 using a late fusion approach, we achieved a significant improvement in detection accuracy. The fused model achieved an accuracy of 94.03%, surpassing the individual model's performance. This highlights the complementary nature of the three architectures and the effectiveness of combining their predictions. Figure-3 shows the individual CNN performance for LDDST dataset on Four Models.

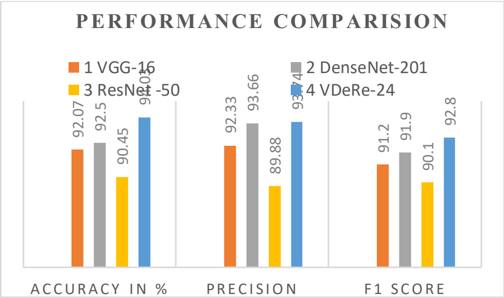

Fig. 3. Individual CNN performance for LDDST dataset on Four models.

Below are the Accuracy, Precision and F1 score comparison for SLDS dataset on four models.

Table 3. comparison of Accuracy, Precision and F1 score on SLDS dataset

S.No.	CNN model	Accuracy in %	Precision	F1 score
1	VGG-16	79.65	79.2	79.06
2	DenseNet-201	78.63	77.49	77.46

3	ResNet -50	77.94	77.03	76.89
4	VDeRe-24	85.04	84.67	84.07

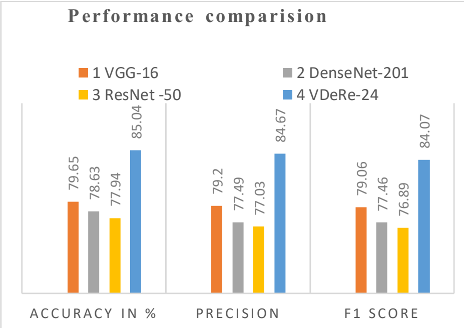

Fig.4. Performance comparison on SLDS dataset

With a value of 92.5 and the best precision value of 93.66, the LDDST dataset with CNN DenseNet-201 yielded the best accuracy results. Using CNN DenseNet-201 on the SLDS dataset, the best precision of 83.47 and the best training accuracy of 84.56 were achieved.

Table 4. Accuracy, Precision and F1 score comparison for two datasets

S.No.	**Dataset**	**Fusion Model**	**Accuracy in %**	**Precision**	**F1 score**
1	LDDST	VDeRe-24	94.03	93.74	92.8
2	SLDS	VDeRe-24	90.67	89.45	90.07

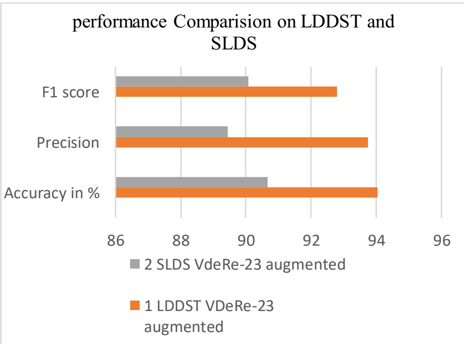

Fig.5. performance Comparison of fusion model on LDDST and SLDS Datasets

with Fusion model (VDeRe-24).we got the highest testing accuracy of 94.03% on LDDST dataset when compared to the accuracy of 90.67% on SLDS dataset

C. Discussions

The superior performance of DenseNet-201 and the fusion model underscores the importance of dense connections and ensemble learning in Logo Detection tasks. The slight disparity in performance between the model and the fused model suggests that each model captures different aspects of logo features. Combining their predictions allows for a more comprehensive understanding of logo characteristics, leading to improved detection accuracy. The Adaptive Logo Recognition System(ALRS) consists of three separate models and one fusion model. Logo recognition was accomplished by the CNN, DenseNet-201 and fusion CNN VDeRe-24. We tested the performance of the proposed model with alternative methods using the LDDST and SLDS datasets. The results are shown in Table 5.

Table 5. The Comparison of results of the proposed method with previously methods

S. No	**Dataset**	**Author**	**Method**	**Accuracy %**
1.	Tobacco-800	H.Wang[1]	CAN (2010)	84.2
2.	Tobacco -800		CAN (2010)	76.6
3.	Tobacco -800	B.Moitte[12]	Nearest neibour	87.08
4.	Tobacco -800	R. K. Jain[13]	Logo Net	82.2
5.	Kvasir	N.Ghatwary [7]	DenseNet-201	71.6
6.	MICCAI 15	N.Ghatwary [8]	AlexNet	78
7.	SLDS	ours	Proposed method	90.67
8.	LDDST	ours	Proposed method	94.03

By doing the aforementioned comparisons, we can demonstrate that our approach outperforms other methods by achieving higher recognition accuracy 94.03 on LDDST and 90.67 on SLDS datasets.

5.CONCLUSION

Our study presents a comprehensive investigation of logo detection using CNN models, including VGG-16, DenseNet-201, and ResNet-50, as well as a fused model combining their predictions. Through experimentation and evaluation on a test dataset, we have demonstrated the effectiveness of these models in accurately identifying logos within images. The three models, VGG-16, DenseNet-201, and ResNet-50, exhibited strong performance, with accuracy ranging from 90% to 93%. Each model leverages unique structural features to capture distinct aspects of logo characteristics, contributing to their success in logo detection tasks. Moreover, the fusion of predictions from VGG-16, DenseNet-201, and ResNet-50 yielded a notable improvement in accuracy, achieving a performance of 94%. This fusion approach effectively combines the strengths of multiple CNN architectures, resulting in enhanced detection accuracy compared to robustness and accuracy of logo detection systems. Overall, our study contributes to the growing body of literature on computer vision applications and provides valuable insights for advancing logo detection algorithms in various real-world scenarios. Future work may involve exploring other fusion techniques, such as early or attention-based fusion, to further improve detection accuracy and robustness. Additionally, evaluating the transferability of the fused model to different.

6. REFERENCES

[1] H. Wang, “Document logo detection and recognition using bayesian model,” In Proceeding of ICPR, pp. 1961-1964,2010.

[2] G. Zhu and D. Doermann, “Automatic document logo detection,” In proceedings of ICDAR, pp. 864-868, 2007.

[3] S. Hou, Jiacheng Li, W. Min “Deep Learning for Logo Detection: A Survey” ,2022.

[4] H. Wang and Y. Chen, “Logo detection in document images based on boundary extension of features rectangles,” In Proceedings of ICDAR, pp. 1335-1339, 2009.

[5] R. Jain and D. Doermann, “Logo retrieval in document images,” In Proceedings of DAS, pp. 135-139, 2012.

[6] O. Fialka and M. Cadik, “Fft and convolution performance in image filtering on gpu,” In Proceeding of International Conference on Information Visualisation,pp. 609-614, 2006.

[7] T. Mahalakshmi, R. Muthaiah, and P.Swaminathan, “An overview of template matching technique in image processing,” Rsearch Journal of Applied Sciences, Engineering and Technology, 4(24), pp. 5469-5473, 2012.

[8] N. Ghatwary, X. Ye, and M. Zolgharni, :Esophageal abnormality detection using DenseNet bassed faster R-CNN with Gabor features,” IEEE Access, vol. 7, pp. 84374-84385,2019.

[9] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in 2016 IEEE Conference on Computer Vision and Pattern recognition(CVPR), pp. 770-778, Las Vegas, NV, USA, December 2015.

[10] S. Lowtehr, V. Chandran, and S. Sridharan, “Recognition of logo images using invariants defined from hihger-order spectra,” In Proceedings of Asian Conference on Computer Vision, 2001.

[11] J. Lewis, “Fast template matching,” in Vision Interface, pp. 120-123, 1995.

[12] Raju, S. Viswanadha, A. Vinaya Babu, G. V. S. Raju, and K. R. Madhavi. "W-Period Technique for Parallel String Matching." IJCSNS 7, no. 9 (2007): 162.

[13] B. Miotti, “A general framework for graphical item retrieval and identification in printed and handwritten documents,” Universit`a Degli Studi di Firenze, Dipartimento di Sistemi e Informatica, PhD Thesis, 2010.

[14] R. K. Jain, T. Watasue, T. Nakagawa, S. Takahiro, Y. Iwamoto, R. Xiang, and Y. W. Chen, “LogoNet: Layer-aggregated attension centerNet for logo detection,” in IEEE international conference on Consumer Electronics, 2021, pp. 1-6

[15] Madhavi, K. Reddy, A. Vinaya Babu, A. Anand Rao, and S. V. N. Raju. "Identification of optimal cluster centroid of multi-variable functions for clustering concept-drift categorical data." In Proceedings of the International Conference on Advances in Computing, Communications and Informatics, pp. 124-128. 2012.

[16] https://www.kaggle.com/datasets/gauravduttakiit/fakereal-logo-detection

[17] https://www.kaggle.com/datasets/mediatechlab/gneutralspeech?rvi=1

[18] https://www.kaggle.com/datasets/prosperchuks/fakereal-logo-detection-dataset