A Text Recognition Augmented Deep Learning Approach for Logo Identification

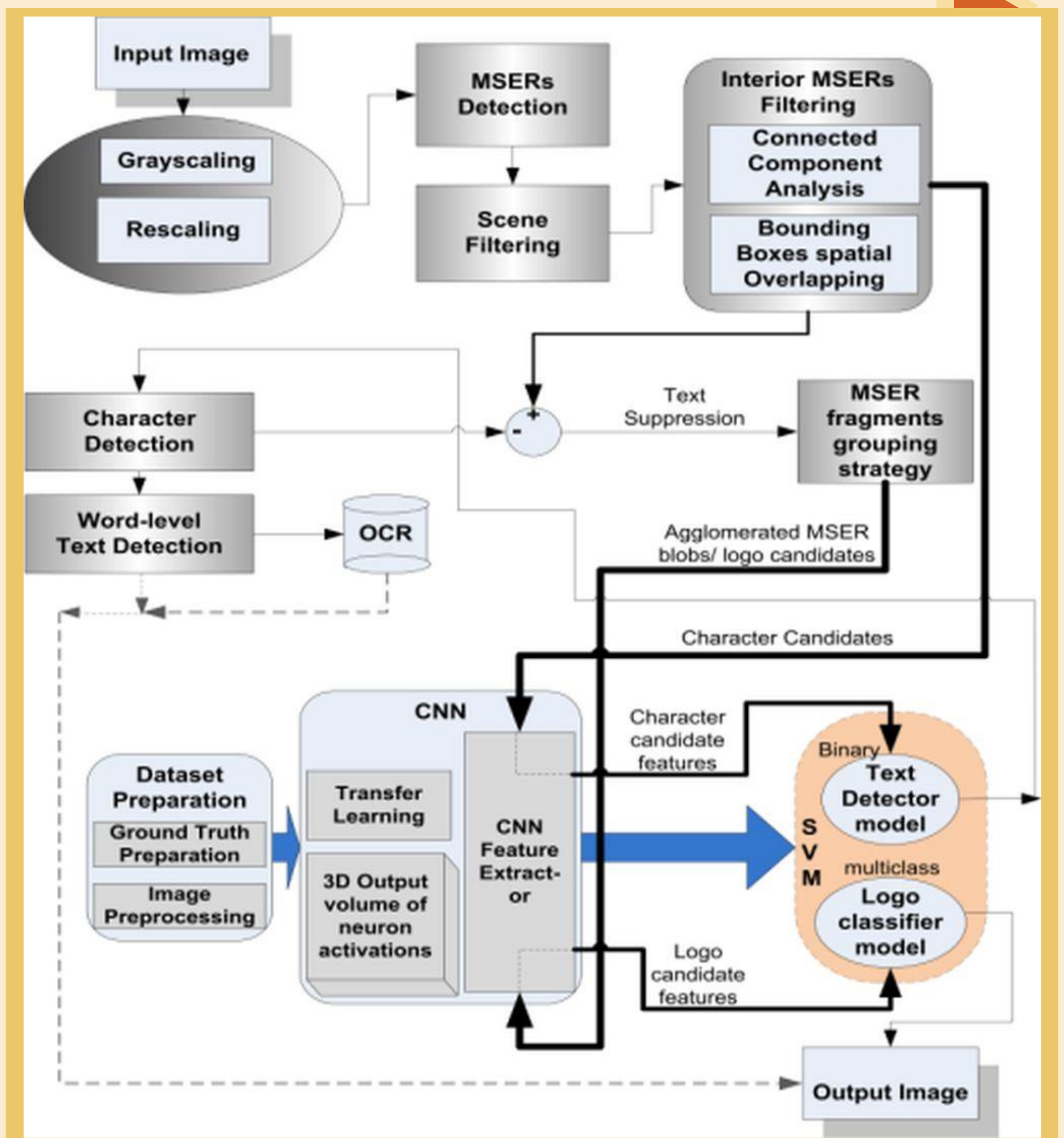
Moushumi Medhi[1], Shubham Sinha[2], Rajiv R. Sahay[3]
[1]Computational Vision Laboratory, Dept. of Computer Science and Engineering, Indian Institute of Technology Kharagpur, India
[2]Dept. of Computer Science and Technology, Indian Institute of Engineering Science and Technology, Shibpur, India
[3]Dept. of Electrical Engineering, Indian Institute of Technology Kharagpur, India
{medhi.moushumi, bitan1994, sahayiitm}@gmail.com

INTRODUCTION

- Logo/brand name detection and recognition in unstructured and highly unpredictable natural images has always been a challenging problem because of the extensively varying types, classes and patterns of logos.
- To address this challenging problem, we propose to combine maximally stable extremal region (MSER) based text detection and recognition with logo/brand image recognition using convolutional neural networks (CNN).

Why text detection/recognition?

- Reduces computational overhead in subsequent logo detection stages by progressively removing regions less likely to contain logo.
- Facilitates the additional task of optical character recognition (OCR) which recognizes the text associated with the logo image. This is useful in cases where logo detection mechanism fails and thereby augments the accuracy of the overall logo recognition system.

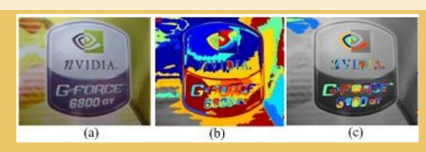


Block diagram of the proposed algorithm

01 METHODOLOGY

Detection of Msers

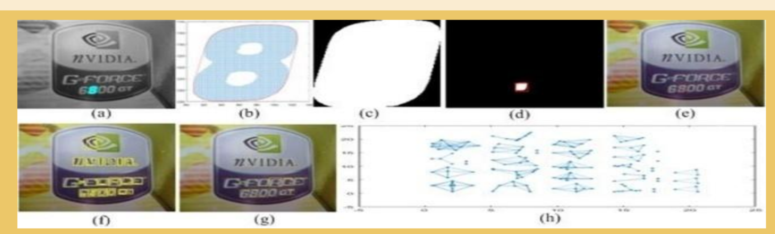
- Image Pre-processing
- Detection of stable covariant regions
- Background subtraction



(a) Original image. (b) Detected MSERs. Both fine and large structures are detected in the initial step to ensure detection of regions-of-interest (ROIs) with no missing foreground objects. (c) MSER filtered image using shape parameters.

Text detection

- Interior MSER filtering
- Minimum bounding rectangles
- Character detection and Word level detection
- Application of OCR

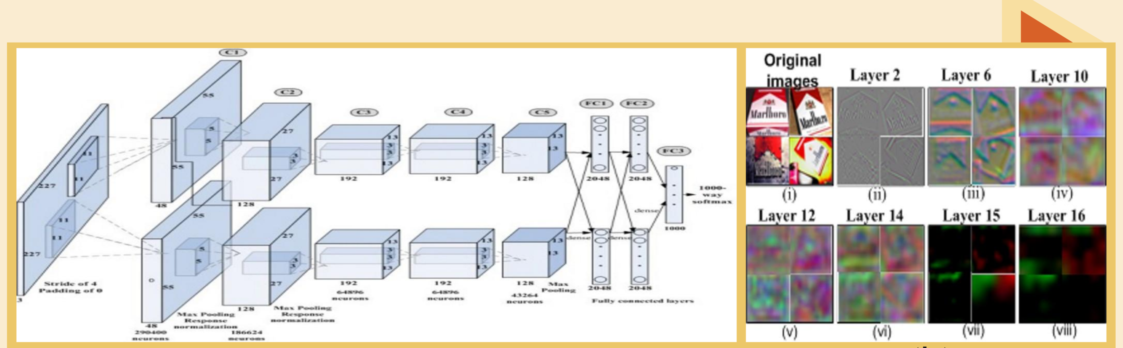


(a) Detected MSER. (b) Pixels of the detected MSER enclosed by its convex envelope. (c) Binary convex hull image with pixels within the hull set to 1. (d) Bounding box of the binary convex hull image. (e) Bounding box for the MSER in the original image. (f) MBRs of all the filtered MSERs. (g) Elimination of inner MBRs and merging of partially overlapping MBRs. (h) Graph of the connected components of MSER regions in (f).

(a) Character proposals. (b) Detected characters. (c) Chain of overlapping characters. (d) Word level detection. (e) OCR text recognition on full image. (f) OCR text recognition on our detected text regions.

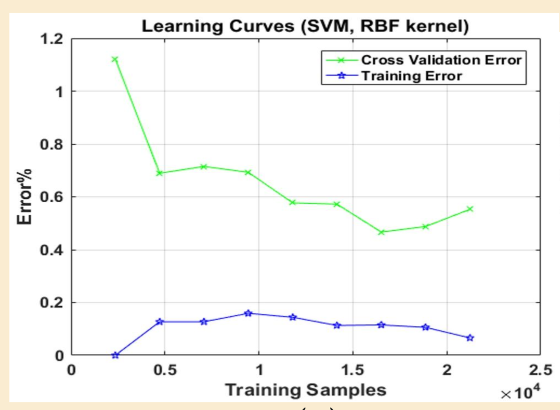
Logo Detection and Classification

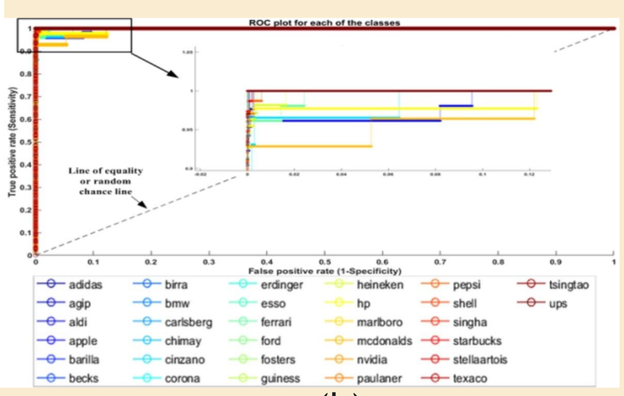
- Text suppression
- Component analysis to agglomerate non-text MSERs and form logo proposals
- Transfer Learning and feature extraction using a pre-trained Alexnet CNN [1] model
- Multiclass SVM classifier for logo classification
- 32 logo classes:Flickr-32 [2]+MICC logo datasets [3]



(a) (b)
(a) Architecture of the pre-trained Alexnet [1]. (b) Features extracted by several hidden layers of our pre-trained CNN

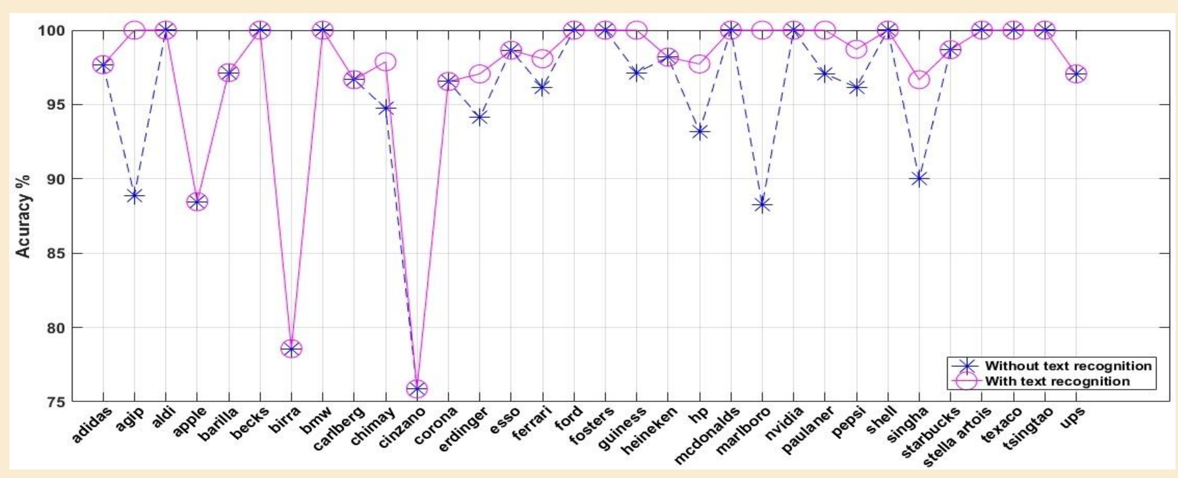
02 RESULTS





(a) (b)

(a) Learning Curves for SVM classifier used for logo recognition. Radial Basis Function (RBF) kernel is used to train the SVM for logo classification. Cross validation error increases and the classifier suffers from overfitting as training sample size exceeds 16.5k images which amounts to 70% of the total data. Hence the training size of the classifier was limited to 60% of the total data available. (b) ROC curves. ROC curves nearly follow the vertical axis for most part.



Accuracy plot of the proposed logo detection algorithm.
Average accuracy: 95.74% (without text identification)
: 97.17% (with text identification)

Method	Year	Precision	Recall	F-Measure
Romberg et al. [2]	2011	0.98	0.61	0.752
Revaud et al. [4]	2012	≥ 0.98	0.73	0.836
Romberg et al. [5]	2013	0.99	0.83	0.903
Farajzadeh [6]	2015	0.93	0.86	0.894
Iandola et al. [7]-AlexNet	2015	0.73	Not reported	Not reported
Liu et al [8]	2016	0.96	0.86	0.907
Oliveira et al. [9]-caffenet	2016	0.93	0.89	0.905
Proposed method	-	**0.98**	**0.96**	**0.97**

The performances of a number of significant works in logo detection using FlickrLogos-32 dataset are listed in the Table.

Assessment of the proposed method using three different types of logo images. (a) Word formation and subsequent OCR application. (b) Logo candidate generation. (c) Final Output Images.

(a) Character proposals. (b) Detected characters. (c) Detected text region. (d) Logo region proposals. (e) Output image. Text lines (noise), placed too close to the logo, form part of the logo and our logo recognition mechanism fails. The problem can be addressed by incorporating more training data with increased noise levels.

Other failure cases: blurred images, purely text logo images.

03 CONCLUSION

- A computationally feasible and accurate, size and orientation invariant algorithm is developed that detects and classifies logos in complex, natural backgrounds.
- The presence of any text in the image is detected followed by its recognition to ameliorate the problem of incorrect logo image identification.
- The detected text related to the logo serves as a backup in cases of failure of identification of images of logos.
- The proposed methodology has significantly reduced the computational expense by eliminating the sliding window approach on full original image but extracts only the most probable ROIs using MSER and CNN classifier which improves upon the state-of-the-art logo detection and recognition algorithms.

References

[1] A. Krizhevsky, I. Sutskever, and J. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2012, 1097-1105.
[2] S. Romberg, L. G. Pueyo, R. Lienhart, and R. V. Zwol. Scalable logo recognition in real-world images. *Proc. 1st ACM Intl. Conf. on Multimedia Retrieval*, 2011, 965–968.
[3] H. Sahbi, L. Ballan, G. Serra, and A. Bimbo. Context-Dependent logo matching and recognition. *IEEE Transactions on Image Processing, 22* (3), IEEE, 2013, 1018–1031
[4] J. Revaud, M. Douze, and C. Schmid. Correlation-based burstiness for logo retrieval. *Proc. 20th ACM Intl. Conf. on Multimedia*, 2012, 965-968.
[5] S. Romberg and R. Lienhart. Bundle min-hashing for logo recognition. *Proc. 3rd ACM Conf. on Intl. Conf. on Multimedia Retrieval*, 2013, 113–120.
[6] N. Farajzadeh. Exemplar-based logo and trademark recognition. *Machine Vision and Applications, 26* (6), 2015, 791-805.
[7] F. N. Iandola, A. Shen, P. Gao, and K. Keutzer. DeepLogo: Hitting Logo Recognition with the Deep Neural Network Hammer. *arXiv*: 1510.02131, 2015.
[8] Y. Liu, J. Wang, Z. Li, and H. Li. Efficient logo recognition by local feature groups. *Multimedia Systems*, 2016, 1-9.
[9] G. Oliveira, X. Frazão, A. Pimentel, and B. Ribeiro. Automatic Graphic Logo Detection via Fast Region-based Convolutional Networks. *arXiv*:1604.06083, 2016.