



# International Journal on Recent Researches In Science, Engineering & Technology

(Division of Computer Science and Engineering)

A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. It is an absolutely free (No processing charges, No publishing charges etc) Journal Indexed in JIR, DIIF and SJIF.

Research Paper

Available online at: [www.jrrset.com](http://www.jrrset.com)

Chief Editor : Dr. M.Narayana Rao, Ph.D., Rtd. Professor, NIT, Trichy.

ISSN (Print) : 2347-6729

ISSN (Online) : 2348-3105

Volume 3, Issue 4,  
April 2015.

JIR IF : 2.54

DIIF IF : 1.46

SJIF IF : 1.329

---

## Detection of Scene Text Based On Machine Learning Classifiers

I.Vasudevan, S.Gunasekaran ,T.Dineshkumar

<sup>1</sup>Assistant Professor, Department of Computer science and Engineering, V.S.B Engineering college

<sup>2</sup>Assistant Professor, Department of Computer science and Engineering, V.S.B Engineering college

<sup>3</sup>Student, Computer science and Engineering, V.S.B Engineering college

Karur, Tamilnadu, India

[ivasu78@gmail.com](mailto:ivasu78@gmail.com)

**Abstract** - In computer vision society, the detection of text and its recognition in camera captured images have been considered as very important issue. In this paper, we present algorithm for scene text detection based on machine learning methodology where the system will be trained based on data sets. Two machine learning classifiers used, among which one helps in the generation of candidate word regions and the other filters out nontext ones. To be clear, Maximally Stable Extremal Region (MSER) algorithm has been used to extract Connected Components (CC) in images. Then clustering of these extracted CCs takes place so that the candidate regions will be generated. Next we concern about normalization of candidate word regions and determine whether each region contains text or not. AdaBoost classifier has been trained for clustering process based on adjacency relationship and based on multilayer perceptrons system has been trained for the filtering of nontext ones. Our paper is not based on heuristics rules instead based on training and testing of the system which increases accuracy and ignore complexity issues.