

## A Study of Script Identification Techniques

Rumaan Bashir, Kaiser J. Giri

Islamic University of Science and Technology, Awantipora, J&K  
[rumaan.bashir@islamicuniversity.edu.in](mailto:rumaan.bashir@islamicuniversity.edu.in), [kaiser.giri@islamicuniversity.edu.in](mailto:kaiser.giri@islamicuniversity.edu.in)

### Abstract

In the recent decades script identification is performed in order to recognize the script(s) of the text present in an image of a document. Since numerous scripts exist in today's world to write the languages therefore a variety of techniques to address this have been developed over time. Some techniques have been applied to the whole image document while as some only target limited areas as a block or even a word. This paper provides a view of most of the techniques which are applied to detect scripts at word-level, line-level, block-level or page-level. Here, a gist of the methods used for script identification is given.

### Keywords

Scripts, Script Identification Techniques, Script Identification.

### 1. INTRODUCTION

Document Image Understanding & Analysis, the area of Computer Science dealing with processing of document images, demands script identification specifically when the possibility of multiple scripts is high and identification is required to proceed to subsequent stages of document processing. Script Identification becomes significantly important when the indexing, retrieval, recognition and understanding are directly going to affect the permissible operations like searching a particular image, sorting images, classifying or categorizing images, selecting suitable script from the given text area or even retrieving document images having text written in a specific script. The fundamental job in script identification is to formulate a method to determine the necessary description of the features present in a script document which almost is sufficient to recognize the document's script for different applications. The methods used for the Identification of Scripts are generally classified into two classes, *local methods* & *global methods* (Bashir & Quadri, 2014).



Figure 1: Common Categories of Script Identification Techniques.

