The economic feasibility of maintaining large number of documents in digital image formats has created a tremendous demand for robust ways to access and manipulate the information these images contain. This requires research in the area of document image understanding, specifically in the area of document image recognition as well as document image retrieval. There have been many excellent attempts in building robust document analysis systems in industry, academia and research labs. One way to provide traditional database indexing and retrieval capabilities is to fully convert the document to an electronic representation which can be indexed automatically. Unfortunately, there are many factors which prohibit complete conversion including high cost, low document quality, and non-availability of OCRs for non-European languages.

Word spotting has been adopted and used by various researchers as a complementary technique to Optical Character Recognition for document analysis and retrieval. The various applications of word spotting include document indexing, image retrieval and information filtering. The important factors in word spotting techniques are pre-processing, selection and extraction of proper features and image matching algorithms. The Euclidean based algorithm is considered to be a faster matching algorithm. In the word spotting literature the Euclidean based algorithm has been used successfully to compare the features extracted from word images. However, the problem with this approach is that interpolation of images to get same width leads to loss of very useful informations. Dynamic Time Warping based algorithm is more preferable than Euclidean based algorithm.

In this thesis, a new approach based on Weighted Euclidean distance based algorithm has been used innovatively to compare two word images. The various features, i.e., projection profiles, word profiles and transitional features are extracted from the word images which are compared via Weighted Euclidean based algorithm with greater speed and higher accuracy. The experiments have been conducted on a large printed document databases of English language. The average precision rates achieved for this language were 95.48%. The time taken for matching every two images was 0.03 milli-seconds.

A second line of research performed in this thesis considers keyword spotting for Hindi documents, the task of retrieving all instances of a given word or phrase from a collection of documents. During the course of this thesis, a novel learning based keyword spotting system using recurrent neural networks was developed. To the knowledge of the author, this is the first time that recurrent neural network based method is explored to Indian languages documents. In a set of experiments, its superior performance over state-of-the-art reference systems is shown.

Finally, the above system is applied for exhaustive recognition. The main findings of this thesis are that supervised learning in the form of training can increase the retrieval accuracy. The key to success lies in a well-balanced trade-off between data quality and data quantity when choosing the elements to be added to the training set. Performance evaluation using datasets from different languages shows the effectiveness of our approaches. Extension works are recommended that need further consideration in the future to further the state-of-the-art in document image retrieval.