ABSTRACT

This report consists of the basic description of Hindi OCR, a software system. In this project various image pre-processing, features extraction and classification algorithms have been explored and implemented, to design high performance OCR software for Indian Language Hindi based on Devanagari script.

During the research phase, accessible papers on existing OCR systems like Bangla, Nepali relevant to our project have been studied and some useful portions have been reviewed and summarized in this report.

In this project emphasis has been laid on recognizing the individual base consonants and vowels which can be later extended to recognize complex derived letters & words.

A working fragmented along with its Graphical User Interface was being achieved during this five months internship period.