Preface

My research work aims to develop a platform for the analysis and recognition of the Assamese scripts/documents through the man-machine dialogues and interaction. It is becoming increasingly important to have information available in digital format for effective access, reliable storage and long term preservation. The advancement of information technology (IT) and the need for automatic data processing catalyze this process. This thesis provides a design of an Optical Character Recognition (OCR) system for handwritten Assamese text/documents.

This thesis is organized in seven chapters as follows:

The first chapter provides an overview of the general background of character recognition (CR) and the problem setting. Motivation behind the present work and the major contributions are also briefly described.

In chapter two a review of literature have been made. It also includes the development of CR systems as per its historical evolution. The OCR Systems developed for Oriental Languages and Research in Indian Scripts is also reviewed in this chapter.

Chapter three contains various methodologies available in the literature regarding the development of a CR system. It also emphasizes the methodology adopted in the present study for the analysis and recognition of handwritten Assamese text/documents.
Chapter four is devoted to the overview of the Assamese script. It also includes the detailed analysis of Assamese script in the linguistic point of view as well as the structural features associated with it.

Chapter five deals with the overview of the neural network classifier, which is the heart of our recognition system. It includes the details of the ‘multi-layer feedforward backpropagation neural network’ used as a specific classifier for the recognition module in the present study.

Chapter six details the analysis and the experimental results found in the present study. It includes the results of the isolated character recognition experiments done on various neural network architectures as well as the results of the fundamental image preprocessing operations and systems applicability on real-life document images of handwritten Assamese script.

Chapter seven is devoted to the results & overall conclusions of the present study and scope for future work for further extending the work.

Guwahati

August, 2009

Masih Saikia.

Guwahati

August, 2009