Accelerating Classifier Training using AdaBoost within Cascades of Boosted Ensembles

A thesis presented in partial
fulfillment of the requirements

for the degree of

Master of Science in Computer Sciences

at Massey University, Auckland, New Zealand

Teo Susnjak

2009

Contents

List of Figures

List of Tables

List of Notations

Abstract

This thesis seeks to address current problems encountered when training classifiers within the framework of cascades of boosted ensembles (CoBE). At present, a significant challenge facing this framework are inordinate classifier training runtimes. In some cases, it can take days or weeks (Viola and Jones, 2004; Verschae et al., 2008) to train a classifier. The protracted training runtimes are an obstacle to the wider use of this framework (Brubaker et al., 2006). They also hinder the process of producing effective object detection applications and make the testing of new theories and algorithms, as well as verifications of others research, a considerable challenge (McCane and Novins, 2003).

An additional shortcoming of the CoBE framework is its limited ability to train classifiers incrementally. Presently, the most reliable method of integrating new dataset information into an existing classifier, is to re-train a classifier from beginning using the combined new and old datasets. This process is inefficient. It lacks scalability and discards valuable information learned in previous training.

To deal with these challenges, this thesis extends on the research by Barczak et al. (2008), and presents alternative CoBE frameworks for training classifiers. The alternative frameworks reduce training runtimes by an order of magnitude over common CoBE frameworks and introduce additional tractability to the process. They achieve this, while preserving the generalization ability of their classifiers.

This research also introduces a new framework for incrementally training CoBE classifiers and shows how this can be done without re-training classifiers from beginning. However, the incremental framework for CoBEs has some limitations. Although it is able to improve the positive detection rates of existing classifiers, currently it is unable to lower their false detection rates.

Acknowledgements

I would like to thank my supervisor Dr. Andre Barczak for his invaluable input, encouragement and patience over the last year. Thanks also to my colleagues at the Massey University Computer Science Department who helped me with technical issues and provided a fun and stimulating environment to work in. Thanks to the Computer Science faculty members at Massey University for always being encouraging and generous with their feedback and assistance.

Thanks to my friends for still being there despite not seeing much of me. My gratitude goes to Brian and Beverly for the genuine care and strength they have given me and also to my dear parents who have always been unwavering in their support. Thanks to Sarah for being so understanding and for providing me with much motivation during the last year.