

Title: Development of a Feature-Selective Neuroevolution Method and its Relevance in Medical CAD Applications

Name: Maxine Yen Ling Tan

Promoters: Prof. Rudi Deklerck, Prof. Bart Jansen

Abstract

Computer-Aided Diagnosis (CAD) systems have been receiving increasing attention in recent years. CAD has been defined as a diagnosis made by radiologists with the benefit of information generated by computerized image analysis. The significant focus on developing CAD systems for lung cancer has primarily been because it is by far the leading cause of cancer death among both men and women in the United States. The primary reason for significant optimism in CAD for lung cancer detection is the observation that those patients who are diagnosed with early stage lung cancer, and who undergo curative resection have a much better prognosis, with five-year survival rates rising to 40 to 70%. Many feature selection and classification methods are proposed for lung cancer detection and diagnosis. A feature selection method that automatically *selects* features at the same time as it evolves neural networks called Feature Selective NeuroEvolution of Augmenting Topologies (FS-NEAT) was proposed by Whiteson et al. In this work, we propose a novel feature selection method called Feature Deselective NeuroEvolution of Augmenting Topologies (FD-NEAT), which begins with fully-connected inputs and automatically *deselects* irrelevant or redundant inputs. FD-NEAT, FS-NEAT and traditional NEAT are compared in some mathematical problems, and in a challenging race car simulator domain (RARS). The results show that FD-NEAT outperforms FS-NEAT in terms of network performance and feature selection, and evolves networks that offer the best compromise between network size and performance. The performance of FD-NEAT is also compared with that of two other established classifiers, namely support vector machines (SVMs) and fixed-topology neural networks in a novel computer-aided lung nodule detection system for computed tomography (CT) images. A set of 235 randomly-selected cases from the publicly-available Lung Image Database Consortium (LIDC) database was used for training, and 125 independent cases for testing. The results show that the CAD system equipped with any of the three classifiers performs well with respect to other methods described in literature.