Chapter 7
Conclusions and Future Work

In our research we were investigating the use of informative vector machines for face recognition and its performance with other kernel classifiers such as SVM. Throughout our research we have studied the IVM algorithm and found problems related to it and proposed several improvements. Experiments carried out for faces made us achieve several goals those were set out at the beginning.

7.1 Research Achievements

One of the major goals we were perusing in our research was improvement of classification accuracy of faces. Having an accuracy of over 97% with all types of kernels makes IVM a good classification algorithm for faces. But it is clear that none of IVM type classifiers can perform better than SVM RBF kernel classifier. But IVM* ARD type classifiers (which implements algorithm 3 in chapter 4) indicated a very close accuracy compared to SVM which makes it an alternative solution that can be implemented in applications that currently use SVM solutions.

Another important aspect in machine learning research that we were considering in our research was the sparsity of solutions for face recognition with our algorithms. In our research we were able to reprove the high sparsity of Gaussian process classifiers compared to SVM. But we were only able to achieve sparsity with ARD kernels. Despite slightly low accuracy compared to SVM, IVM* with ARD kernels indicated a special ability to provide highly sparse solutions. Interestingly non-ARD kernels could not provide any significant level of sparseness compared to SVM.

To our knowledge the dimension reductions with ARD kernels for frontal faces is a novel solution that has not been observed before with another kernel classifier algorithm. In our experiments we had to restrict diagonal ARD metrics when creating ARD kernels due to computational limitations. Experimental results clearly indicated both dimension reductions and sparse solutions make improvements in both storage capacities and classification speeds compared to SVM. This is one of the major achievements in our research.
7.2 Future research

As outcomes from our research we can propose several ideas for future research. They can be categorized mainly as improvements to IVM algorithm and it’s applications to face recognition.

All of our experiments in this research were restricted to holistic frontal faces with limited variations and complex variations such as pose variations, expressions, aging and illuminations were not considered. Results that were achieved in this research with informative vector machines algorithms such as sparse solutions and dimension reductions may not be applicable for faces with different variations. We believe that future research with kernel classifiers described in the thesis with various types of face databases would help to come up with more concrete conclusions.

The most significant result from our research was the distribution of ARD values on face images. As we explained in the last chapter most of the important facial components had lower ARD values which can be ignored without any loss of accuracy. Since we don’t have a theoretical explanation we would like to propose more research on this observation. Future research in this area can be useful for face recognition as well as understanding the nature of ARD kernels.

There are several weaknesses in IVM algorithm that we found during our research. One major weakness with IVM is the randomness and inability of finding optimal solutions. Our attempts to find a solution with common points and Expectation Propogation failed to provide a solution to this problem. But we were able to specify a strategy to reduce randomness, which was successful with our face recognition. We don’t have a theoretical proof as to whether it would be appropriate under all circumstances. Also the use of ARD kernels allowed us to propose a stopping condition, which again cannot be fully justified to be applicable to all classification problems. Research focused to improve the

We had to empirically find the number of data vector for training due lack of method in IVM. This was another weakness we found with IVM that need to be further research. Also optimization of ARD kernels took very long time though the non-ARD kernels were considerably faster. Use of faster optimization methods or approximation methods with ARD kernels would make IVM a much suitable algorithm in domains with large data sets.

Many features in IVM such as sparsity, dimension reduction in feature space, faster classification times and low storage requirements makes it a suitable algorithm to be implemented in real world applications. Especially it can be a very good solution for applications where both low storage capacities and recognition speeds are critical such as embedded systems, surveillance
systems and robotics. We hope our research will motivate further industrial oriented research for face recognition using informative vector machines.