ELHAM ETEMAD
BSc (Computer Science), Isfahan University of Technology, 2010
MSc (Computer Engineering), Isfahan University of Technology, 2012

FACULTY OF COMPUTER SCIENCE

TITLE OF THESIS:	PERCEPTUAL SHAPE FEATURE BASED IMAGE CODING FOR VISUAL CONTENT CLASSIFICATION AND OBJECT RECOGNITION
TIME/DATE:	10:00 am, Wednesday, November 14, 2018
PLACE:	Room 3107, The Mona Campbell Building, 1459 LeMarchant Street

EXAMINING COMMITTEE:

Dr. Minglun Gong, Department of Computer Science, Memorial University of Newfoundland (External Examiner)

Dr. Evangelos E. Milios, Faculty of Computer Science, Dalhousie University (Reader)

Dr. Derek Reilly, Faculty of Computer Science, Dalhousie University (Reader)

Dr. Vlado Keselj, Faculty of Computer Science, Dalhousie University (Reader)

Dr. Qigang Gao, Faculty of Computer Science, Dalhousie University (Supervisor)

DEPARTMENTAL REPRESENTATIVE:	Dr. Michael McAllister, Faculty of Computer Science, Dalhousie University
CHAIR:	Dr. Sarah Wells, PhD Defence Panel, Faculty of Graduate Studies

ABSTRACT

The most essential technique in creating agents with ability to process and understand the content of visual data is object recognition, which includes image content classification, and object localization. Despite deep convolutional neural networks' (CNNs') performance gain in computer vision, there still are application scenarios with limited training data and computing power for which using deep CNNs based methods is not feasible. On the other hand, the human engineered image representations require less training data and computing power and can be enhanced by importing the domain specific knowledge. These representations may also benefit from the human vision characteristics in reducing the gap between computed image representations and human vision perception. In this thesis we have proposed four methods to improve image classification and object localization. All these methods utilize the perceptual shape features of image since it is proved that the human vision perception on objects mostly relies on shape features of the objects, while color and texture are utilized as extra sources to complete this perception. In the first method, we have created a static dictionary of perceptual shape features based on N-gram model and used that in combination with spatial pyramid matching to represent images. In the second method, a dynamic dictionary from image edge segments is formed where these segments are obtained from an octave of image in different scales. Third method considers the curve partitioning points as descriptive area of the image and created a dynamic dictionary from descriptors of these points. The proposed object localization method utilizes the perceptual shape features of the image to improve the predetermined location of objects in an image. The initial location may be obtained by any object recognition method, then the proposed method iteratively merges the edge segments with the detected object using a best first search strategy. These proposed methods have been evaluated on different benchmark image datasets. Judging on the overall performance of the proposed method, it is expected that the proposed methods would bring some useful alternatives to support efficient tool development for applications lacking training data or no training data at all.