University of Adelaide

Deep learning for multi-label scene classification

by

Junjie Zhang

A thesis submitted in fulfillment for the
degree of Master

Under Supervised by
Chunhua Shen and Javen Shi
School of Computer Science

August 2016

Declaration of Authorship

I, Junjie Zhang, declare that this thesis titled and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at the University of Adelaide.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at the University of Adelaide or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ____________________________________

Date: ____________________________________

Abstract

Scene classification is an important topic in computer vision. For similar weather conditions, there are some obstacles for extracting features from outdoor images. In this thesis, I present a novel approach to classify cloudy and sunny weather images. Inspired by recent study of a deep convolutional neural network and the spatial pyramid matching, I generate a model based on the ImageNet dataset. Starting with parameters learned from other classification tasks, I fine-tune the model using outdoor images. Experiments demonstrate that our classifier can achieve state-of-the-art accuracy.

Multi-label learning is a variant of supervised learning where the task is to predict a set of examples, which can belong to multiple classes. This is a variant of popular multi-class classification problems in which each sample has one class label only. It can apply to a wide range of applications, which include text categorisation, semantic image labelling etc.. A lot of research work has been done on multi-label learning with different approaches. In this thesis, I train a neural network from scratch based on the generated artificial images. The model is learned by minimising an error function based on the Hamming distance, through the backpropagation optimisation. The model has high capability of generalisation.

Acknowledgements

I am grateful to my main supervisor, Prof. Chunhua Shen, and co-supervisor, Dr. Qinfeng Shi, whose expertise, understanding, and support made it possible for me to work on the neural network that is of great interest to me. It is a pleasure working with them. Prof. Shen's dedication and keen interest to help his students had been solely and mainly responsible for completing my work. His timely advice, meticulous scrutiny, and scholarly advice have helped me to accomplish the tasks.

I would like to thank for my co-supervisor, Qinfeng Shi, for his time and effort on helping me understand research work and knowledge of machine learning.

I thank PhD candidate Teng Li who cooperated to work on the Weather Classification project. We set up experiment environment, analysed test results and discussed approaches. His prompt inspirations, timely suggestions with kindness, and enthusiasm work attitude have helped to achieve the perfect classification accuracy.

I am extremely thankful to research fellow, Guoshen Lin, for his kind help and discussion about the weather classification and the multi-label classification.

I would also like to thank staffs and visitors in the Australian Center for Visual Technologies (ACVT). Attending the reading group has enriched my knowledge in Computer Vision and learned a lot from them.

Finally, I would like to warmly thank the staffs in the School of Computer Science at the University of Adelaide. Julie Mayo, Sharyn Liersh and Jo Rogers have done excellent jobs on administration which helps students to be focus on their research work.

Contents

List of Figures

List of Tables

Abbreviations

SVM	**S**upport **V**ector **M**achine
ANNs	**A**rtificial **N**eural **N**etworks
ROI	**R**ogin **o**f **I**nterest
SIFT	**S**cale **I**nvariant **F**eature **T**ransform
SGD	**S**tochastic **G**radient **D**escent
BP	**B**ack**P**ropagation
MLE	**M**aximum **L**ikelihood **E**stimation
MAP	**M**aximum **A** **P**osteriori
CNN	**C**onvolutional **N**eural **N**etwork
ReLU	**Re**ctified **L**inear **U**nit
SPM	**S**patial **P**yramid **M**atching
SPP	**S**patial **P**yramid **P**ooling
GPU	**G**raphics **P**rocessing **U**nit
BR	**B**inary **R**elevance Classifier
CC	**C**lassifier **C**hains Classifier
CML	**C**ollective **M**ulti-**L**abel Classifier