Contents

Acknowledgements v

Abstract(English/Français) vii

List of figures xv

List of tables xviii

I Introduction 1

1 General introduction 3
- 1.1 Motivations . . . 3
- 1.2 What is novelty detection about ? . . . 4
- 1.3 Thesis objectives . . . 7
- 1.4 Thesis contributions . . . 8

2 Introduction to remote sensing 11
- 2.1 A brief history of remote sensing . . . 11
- 2.2 Remote sensing of the Earth by satellites/airborne platforms . . . 12
  - 2.2.1 Optical remote sensing . . . 13
  - 2.2.2 Sensors characteristics . . . 15
- 2.3 Remote sensing classification and change detection . . . 18
  - 2.3.1 Change detection in remote sensing . . . 19
  - 2.3.2 Classification in remote sensing . . . 21
- 2.4 Summary . . . 23

3 Introduction to machine learning 25
- 3.1 Learning from data . . . 25
- 3.2 Generalization . . . 28
- 3.3 Learning paradigms . . . 33
  - 3.3.1 Supervised learning . . . 33
  - 3.3.2 Unsupervised learning . . . 34
  - 3.3.3 Semi-Supervised learning . . . 37

Morsier, Frank de
Semi-supervised and unsupervised kernel-based novelty detection with ...
2014

digitalisiert durch
IDS Basel/Bern, im Auftrag der
Schweizerischen Nationalbibliothek

3.4 Learning with kernels . . . 38
3.4.1 The kernel trick . . . 39
3.4.2 Mercer kernels . . . 40
3.4.3 Reproducing kernel Hilbert spaces (RKHS) . . . 42
3.4.4 Support Vector Machines (SVM) . . . 43
3.4.5 Cost-Sensitive SVM (CS-SVM) . . . 46
3.5 Summary . . . 47

4 Novelty detection in machine learning **49**
4.1 The novelty, the outlier and the anomaly . . . 49
4.2 Novelty detection in applications . . . 51
4.3 General formulation . . . 52
4.3.1 Supervised novelty detection . . . 54
4.3.2 Unsupervised novelty detection . . . 54
4.3.3 Semi-supervised novelty detection . . . 54
4.4 State-of-the-art in novelty detection . . . 55
4.4.1 Density estimation approaches . . . 55
4.4.2 Reconstruction-based approaches . . . 58
4.4.3 Support Vector Machines and boundary approaches . . . 59
4.4.4 Neural network approaches . . . 63
4.4.5 Information theoretic approaches . . . 64
4.5 Summary . . . 64

II Novelty detection in clusters **65**

5 Background **67**
5.1 Data made of overlapping clusters . . . 67
5.2 Support Vector Clustering . . . 68
5.2.1 Describing cluster boundaries with support vectors . . . 68
5.2.2 Cluster assignment . . . 69
5.3 Entire solution path algorithms . . . 70
5.3.1 OC-SVM entire solution path . . . 70
5.3.2 CS-SVM entire solution path . . . 72
5.4 Nested decision boundaries . . . 73
5.4.1 Nested OC-SVM solution path . . . 73
5.4.2 Nested CS-SVM solution path . . . 74
5.5 Summary . . . 76

6 Unsupervised change detection via hierarchical clustering **77**
6.1 Clustering and outlier hierarchy . . . 77
6.2 Hierarchical clustering algorithms . . . 79
6.2.1 Hierarchical support vector clustering . . . 79

6.2.2 Hierarchical Gaussian process clustering 80
6.2.3 Cluster assignment from hierarchical adjacency matrices 81
6.3 Validity measure for hierarchical clustering 82
6.4 Cluster merging criteria 85
6.4.1 Outlier dispersion as merging mechanism 85
6.4.2 Penalizing different core sizes 87
6.4.3 Setting the merging level 88
6.4.4 Computational considerations 88
6.5 Experiments 89
6.5.1 Experimental setup 89
6.5.2 Results 90
6.6 Conclusions 92

7 Semi-supervised novelty detection via SVM entire solution path 95
7.1 Detecting unknown changes under limited supervision 95
7.2 Semi-supervised novelty detection as a two-class problem 97
7.3 Unsupervised parameter selection based on the cluster assumption 98
7.4 Computational complexity 100
7.5 Experiments 101
7.5.1 Experimental setup 102
7.5.2 Numerical results 106
7.5.3 Free parameters sensitivity 108
7.5.4 Algorithms runtime and convergence analysis 110
7.6 Conclusions 110

III Novelty detection in manifolds 113

8 Background 115
8.1 High dimensional data on manifolds 115
8.2 Linear subspaces 118
8.2.1 Sparse Subspace Clustering 119
8.2.2 Low-Rank Subspace Clustering 120
8.2.3 Sparse and low-rank Subspace Clustering 122
8.3 Non-linear subspaces 123
8.4 Summary 124

9 Unsupervised classification via non-linear low-rank and sparse subspace representation 125
9.1 Non-linear subspace representation 125
9.1.1 Kernel Low-Rank and Sparse Subspace Clustering (KLRSSC) . . . 126
9.1.2 Kernel Low-Rank Subspace Clustering (KLRSC): $\lambda \to 0$ 129
9.1.3 Kernel Sparse Subspace Clustering (KSSC): $\lambda \to \infty$ 130

9.1.4 Handling noise . . . 131
9.1.5 Handling outliers . . . 132
9.1.6 Handling missing entries . . . 132
9.2 Scalable clustering . . . 132
9.2.1 Collaborative representation . . . 132
9.2.2 About scalability . . . 133
9.3 Parameters and model selection . . . 134
9.4 Subspace independence measures . . . 135
9.4.1 Linear subspaces . . . 135
9.4.2 Non-linear subspaces . . . 136
9.5 Experiments . . . 137
9.5.1 Experiment 1: Face and object clustering . . . 137
9.5.2 Experiment 2: Unsupervised hyperspectral image classification . . 146
9.6 Conclusions . . . 152

10 Semi-supervised one-class classification via graphs **153**
10.1 Introduction . . . 153
10.2 One-class label propagation . . . 154
10.3 One-class constrained spectral clustering . . . 155
10.3.1 Spectral orthogonal subspace projection . . . 156
10.3.2 Spectral learning . . . 157
10.3.3 Constrained spectral clustering . . . 157
10.4 One-class classifier with graph-based regularization . . . 159
10.5 Experiments . . . 159
10.5.1 Experimental setup . . . 159
10.5.2 Results . . . 160
10.6 Conclusions . . . 163

IV Conclusions **167**

Conclusions **169**
Towards more realistic detection approaches . . . 169
Contributions . . . 170
Further perspectives . . . 171

A Appendix A **173**
A.1 Accuracy evaluation metrics . . . 173

B Appendix B **177**
B.1 Landsat . . . 177
B.1.1 Bastrop fires . . . 177
B.1.2 Gloucester floods . . . 179

B.2 ROSIS 180
B.2.1 Pavia DFC 180
B.2.2 Pavia center 181
B.3 Faces and objects databases 182

C Appendix C **185**
C.1 Additional face & object clustering results 185

Bibliography **190**

Curriculum Vitae **215**