

Bridging Theory and Algorithms for Open-Set and Heterogeneous Domain Adaptations

by *Zhen Fang*

Thesis submitted in fulfilment of the requirements for
the degree of

Doctor of Philosophy in Computer Science

under the supervision of *Professor Jie Lu*
and

Professor Guangquan Zhang

University of Technology Sydney
Faculty of Engineering and Information Technology
Australian Artificial Intelligence Institute

June, 2021

Faculty of Engineering and Information Technology
Australian Artificial Intelligence Institute
University of Technology Sydney

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Zhen Fang, declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney. This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

Signature: Zhen Fang

Production Note:
Signature removed prior to publication.

Date: 18 June. 2021

*To my loving wife Yueting Peng,
my parents and my son Junlin
Fang.*

ACKNOWLEDGMENTS

This thesis concludes a wonderful journey at University of Technology Sydney for pursuing my Ph.D. degree in the past three and half years. I would like to take the chance to express my sincere gratitude to the people who inspired and helped me in many ways.

First of all, I would like to express the most sincere gratitude to my principal supervisor Professor Jie Lu and co-supervisor Professor Guangquan Zhang, for their wisdom, vision, expertise, guidance, enthusiastic involvement and persistent encouragement during the planning and development of this research work. Their wide knowledge and logical way of thinking have been of great value for me. Their understanding, encouraging and intensive guidance make this dissertation become possible. I am very grateful for their consistent and generous support throughout my whole doctoral study.

Second, I am obliged to my parents for their moral support, love, encouragement and blessings to complete this task. I am especially thankful to my wife Yueting Peng for her patience, love and encouragement during this journey.

I wish to express my appreciation to Dr Feng Liu, Dr Anjin Liu, Zhaoqing Liu, Jiahua Dong, Dr Junyu Xuan, Dr Hua Zuo, Dr Feng Gu, Dr Yiliao Song, Dr Yi Zhang, Dr Hang Yu, Dr Qian Zhang, Bin Wang, Mengjia Wu and Tianyu Liu and grateful thanks to research fellows at department for their help and motivation throughout my research work. I would also like to extend my special thanks to Lily Qian, Camila Cremonese and other staff members of Australian Artificial Intelligence Institute (AII) at University of Technology Sydney, for their timely help.

Zhen Fang

ABSTRACT

The availability of massive labeled datasets results in the successes of supervised learning in many applications such as object detection, speech recognition and natural language processing. However, the curse of domain mismatch arises if the test samples (target samples) and training samples (source samples) are from different domains. To overcome the mismatch between domains, researchers have proposed transfer learning, which aims to leverage knowledge from source samples with abundant labels to help train a classifier for target samples with insufficient labels.

Although many algorithms have been developed to solve transfer learning problem, most algorithms depend on additional assumptions to ensure their effectiveness. Many assumptions are difficult to be satisfied in real-world application. For example, the label sets of source and target samples are assumed to be same, however, in unsupervised setting, it is impossible to check this assumption. Those unrealistic assumptions limit the application of transfer learning.

This thesis mainly studies two important problems faced by many existing transfer learning algorithms: 1) how to transfer knowledge in the unsupervised setting, when the target label set is larger than the source label set; and 2) how to transfer knowledge, when the source and target samples are from different feature spaces.

Before solving problems 1) and 2), we first consider transfer learning which assumes that the feature spaces and label sets are both same. The setting is also called homogeneous domain adaptation. By studying the homogeneous domain adaptation, we aim to understand how the source knowledge can be used to help improve the target's performance under the simplest setting (i.e., homogeneous setting).

To solve problem 1), this thesis considers a challenging task: *unsupervised open-set domain adaptation* (UOSDA). First, to understand why source samples can be used to help target samples achieve good

performance in the open-set setting, we develop Probably approximately correct (PAC) learning theory for OSDA. Then, based on our theory, a kernel-based algorithm is proposed to solve problem 1). As an application of our OSDA theory, we also establish a theoretical foundation for *open-set learning* (OSL), an important sub-field in machine learning.

To solve problem 2), this thesis studies *semi-supervised heterogeneous domain adaptation* (SsHeDA) problem. Motivated by the compatibility condition in semi-supervised PAC theory, we explain the SsHeDA problem by proving its generalization error – why labeled source samples and unlabeled target samples help to reduce the target error. Guided by our theory, we devise a kernel-based algorithm.

List of Publications

Referred journals:

- [1] **Zhen Fang**, Jie Lu, Feng Liu, Junyu Xuan, and Guangquan Zhang, Open Set Domain Adaptation Theoretical Bound and Algorithm, *IEEE Transactions on Neural Networks and Learning Systems*, 2020. Published. (ERA A*)

- [2] Yiyang Zhang, Feng Liu, **Zhen Fang**, Bo Yuan, Guangquan Zhang and Jie Lu, Learning from a Complementary-label Source Domain: Theory and Algorithms, *IEEE Transactions on Neural Networks and Learning Systems*, 2021. Accepted for publication. (ERA A*)

International conferences:

- [1] **Zhen Fang** and Jie Lu, Anjin Liu, Feng Liu, and Guangquan Zhang, Learning Bounds for Open-Set Learning, *International Conference on Machine Learning (ICML)*, 2021. Accepted for publication. (CORA A*)

- [2] **Zhen Fang** and Jie Lu, Feng Liu, and Guangquan Zhang, Unsupervised Domain Adaptation with Sphere Retracting Transformation, *International Joint Conference on Neural Networks (IJCNN)*, 2019. Published. (CORA A)

- [3] Li Zhong, **Zhen Fang**, Feng Liu, Jie Lu, Bo Yuan, and Guangquan Zhang, How does the Combined Risk Affect the Performance of Unsupervised Domain Adaptation Approaches ?, *Association for the Advancement of Artificial Intelligence (AAAI)*, 2020. Published. (CORA A*)

- [4] Yiyang Zhang, Feng Liu, **Zhen Fang**, Bo Yuan, Guangquan Zhang and Jie Lu, Clarinet: A One-step Approach Towards Budget-friendly Unsupervised Domain Adaptation, *International Joint Conference on Artificial Intelligence (IJCAI)*, 2020. Published. (CORA A*)
-

Papers submitted in referred journals:

- [1] **Zhen Fang**, Jie Lu, Feng Liu and Guangquan Zhang, Semi-supervised Heterogeneous Domain Adaptation: Theory and Algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021 (Revise and Resubmit). Submitted for publication. (ERA A*)
- [2] Feng Gu, Jie Lu, **Zhen Fang** and Guangquan Zhang, Neighbor-Searching Discrepancy-based Real Concept Drift, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020 (Revise and Resubmit). Submitted for publication. (ERA A*)
- [3] Li Zhong, **Zhen Fang**, Feng Liu, Bo Yuan, Guangquan Zhang and Jie Lu, Bridging the Theoretical Bound and Deep Algorithms for Open Set Domain Adaptation, *IEEE Transactions on Neural Networks and Learning Systems*, 2021 (Minor Revision). Submitted for publication. (ERA A*)

Contents

Certificate	ii
Dedication	iv
Acknowledgments	iv
Abstract	vi
List of Publications	viii
List of Figures	xvi
List of Tables	xviii
List of Symbols	1
1 Introduction	1
1.1 Background	1
1.2 Problem Statement	2
1.3 Research Innovation and Contributions	6
1.3.1 Research Innovation	6
1.3.2 Research Contributions	7
1.4 Organization of the Thesis	10
2 Literature Review	12
2.1 Transfer Learning	12
2.1.1 Categorization of Transfer Learning	13
2.1.2 Basic Strategies for Transfer Learning	14
2.1.3 Deep Learning in Transfer Learning	17
2.2 Domain Adaptation	19
2.2.1 Unsupervised Homogeneous Domain Adaptation .	20

2.2.2	Unsupervised Open-Set Domain Adaptation and Open-Set Learning	26
2.2.3	Semi-supervised Heterogeneous Domain Adaptation	31
3	A Manifold-based Algorithm for Unsupervised Homogeneous Domain Adaptation	34
3.1	Introduction	34
3.2	Proposed Algorithm	37
3.2.1	Fuzzy c-means Clustering	37
3.2.2	Manifold Regularization	38
3.2.3	Manifold Fuzzy Clustering	39
3.2.4	Influence of Parameter λ	42
3.3	Experimental Results	43
3.3.1	Baseline Algorithms	44
3.3.2	Synthetic Data	44
3.3.3	Real World Datasets	46
3.3.4	Hyper-parameter Settings	47
3.3.5	Experimental Results	48
3.3.6	Parameter Analysis	49
3.3.7	Insufficient Source Samples	53
3.4	Summary	55
4	Source Domain Reconstruction for Unsupervised Homogeneous Domain Adaptation	58
4.1	Introduction	58
4.2	Problem Setting and Theoretical Analysis	61
4.2.1	Problem Setting	61
4.2.2	Proposed Theoretical Problems	63
4.2.3	Theoretical Analysis for SDR	64
4.2.4	Selected Algorithm Spaces	69
4.3	Proposed Algorithm	70
4.3.1	MixUp in Domain Adaptation	70
4.3.2	Intermediate Domains	71

4.4	Experiments	72
4.4.1	Datasets	72
4.4.2	Experimental Setup	77
4.4.3	Experimental Results	79
4.4.4	Parameter Sensitivity	81
4.5	Summary	82
5	Unsupervised Open-Set Domain Adaptation: Theory and Algorithm	85
5.1	Introduction	85
5.2	Theoretical Analysis	88
5.2.1	Problem Setting and Necessary Concepts	89
5.2.2	Concepts and Notations	89
5.3	Proposed Algorithm	94
5.3.1	Distribution Alignment	96
5.3.2	Manifold Regularization	97
5.3.3	Open-Set Loss Function	98
5.3.4	Overall Reformulation	99
5.3.5	Training	99
5.4	Experiments and Evaluations	102
5.4.1	Real World Datasets	102
5.4.2	Baseline Algorithms	103
5.4.3	Experimental Setup	104
5.4.4	Experimental Results	107
5.4.5	Open-Set Parameters Analysis	108
5.4.6	Parameter Sensitivity, Ablation Study, Convergence Analysis.	109
5.5	Summary	112
6	Learning in Open-World	116
6.1	Introduction	116
6.2	Open-Set Learning Theory	118
6.2.1	Problem Setting and Concepts	118

6.2.2	Transfer Between Domains	120
6.2.3	Construction of Ideal Auxiliary Domain	126
6.2.4	Empirical Estimation for IAD Risk	128
6.2.5	Main Theoretical Results	130
6.3	A Principle Guided OSL Algorithm	132
6.3.1	Datasets	134
6.3.2	Open-Set Learning Demonstration	135
6.3.3	Experimental Setup	136
6.3.4	Experimental Evaluation and Result Analysis	136
6.4	Discussion	137
6.5	Summary	138
7	Bridging Theory and Algorithm for Semi-supervised Heterogeneous Domain Adaptation	141
7.1	Introduction	141
7.2	Problem Setting and Concepts	143
7.2.1	Problem Setting	144
7.2.2	Concepts and Notations	144
7.3	Theoretical foundation for SsHeDA	145
7.3.1	Uniform Sample Complexity	146
7.3.2	Compatibility and Transfer Error Rate	147
7.3.3	Obstacle to Extend SsHoDA Theory	148
7.3.4	Theoretical Analysis	154
7.4	Bring SsHeDA Theory into the Reality	160
7.4.1	Loss Function Motivated by Theorem 18	160
7.4.2	Target Distribution Alignment	161
7.4.3	Overall Loss Function	162
7.5	Kernel-based Algorithm for SsHeDA	162
7.5.1	Loss Function in KHDA	163
7.5.2	Reformulation of the KHDA Loss Function	164
7.5.3	Analytical Solution	168
7.5.4	KHDA Algorithm	172
7.6	Experiments and Evaluations	173

7.6.1	Datasets and Baseline Algorithms	173
7.6.2	Experimental Setup	176
7.6.3	Experimental Results	176
7.6.4	Ablation Study	178
7.7	Summary	179
8	Conclusions and Future Directions	186
8.1	Conclusions	186
8.2	Future Study	188
A	Necessary Lemmas	191
B	Proofs for Chapter 5	195
B.1	Proof of Theorem 5	196
B.2	Proof of Corollary 5.1	199
C	Proofs for Chapter 6	202
C.1	Proof of Theorem 9	209
C.2	Proof of Theorem 10	211
C.3	Proof of Theorem 11	212
C.4	Proof for Theorem 12	214
	Bibliography	223

List of Figures

1.1	Thesis Structure. \mathcal{X}_s and \mathcal{X}_t are the feature spaces for source and target domains, respectively. \mathcal{Y}_s and \mathcal{Y}_t are the label sets for source and target domains, respectively.	11
3.1	(a) Source domain and target domain are given. (b) The domain-invariant feature algorithms learn a map to seek a new feature space where domains are similar and assume the learned small error source classifier can be generalized to the target domain. (c) The domain-invariant feature algorithms match the marginal distributions but discrepancy between conditional distributions are enlarged. (d) Our algorithm (MFC) learns the geometric structure among target samples and source class centers while utilizing the fuzzy c-mean clustering to help classification rather than seeking an unknown feature space.	35
3.2	Two types of distribution discrepancy between source domain and target domain.	40
3.3	Comparisons of baselines and the proposed MFC algorithm on the synthetic samples	45
3.4	Parameter λ study of MFC	50
3.5	Parameter p study of MFC	51
3.6	Analysis for the level of fuzziness.	52
3.7	Accuracy and standard deviation on different data proportions.	53

4.1	To connect source domain and target domain, we construct a source domain sequence, where domains varies continuously. Then, we select an intermediate source domain as the reconstructed source domain.	59
4.2	Parameter analysis for proposed algorithm DMU on ρ_0 . .	83
4.3	The parameter analysis for proposed algorithm DMU on T .	84
5.1	Aim of UOSDA. (a) The original source and target samples are given. (b) UCSDA algorithm matches the source and target samples, leading to negative transfer. Because the unknown target samples interfere with distribution matching. (c) UOSDA algorithm classifies known target samples into the correct known classes and recognizes the unknown target samples as unknown.	87
5.2	The horizontal axis is the difference in the open-set parameters $\delta = \alpha - \gamma$. In the figures, the difference δ is not larger than α , since the parameter γ is required to be larger than or equal to 0. If $\delta > 0$, α is larger than γ . If $\delta < 0$, γ is larger.	114
5.3	Parameter sensitivity study, ablation study and convergence analysis of the proposed DAOD algorithm.	115
6.1	The left figure shows the auxiliary distribution U and marginal distribution P_X (red curve: distribution P_X ; blue lines: distribution U introduced in Definition 21; grey regions: the regions for unknown classes). The right figure shows the marginal distribution $Q_U^{0,\beta}$ of ideal auxiliary domain generated by U , $P_{X Y \in \mathcal{Y}_k}$ and $L_{0,\beta}$ (blue lines and curve: $Q_U^{0,\beta}$ defined in Definition 21).	122

6.2 (a) is the training samples for double-moon dataset. (b) is the decision regions under closed-set learning setting for double-moon dataset. (c) is the decision regions under open-set learning setting for double-moon dataset. (d) is the relationship between error and sample size. (e) is parameter analysis for β . (f) is parameter analysis for t . 134

List of Tables

3.1	Accuracy (%) on Synthetic datasets.	46
3.2	Introduction of datasets.	46
3.3	Accuracy (%) on Office+Caltech10 datasets using DeCAF6 features.	56
3.4	Accuracy (%) on Office-31 datasets using ResNet50 features.	56
3.5	Accuracy (%) on Office-Home datasets using ResNet50 features.	56
3.6	Accuracy (%) on ImageCLEF-DA using ResNet50 features.	57
3.7	Accuracy (%) on Amazon Review datasets.	57
3.8	Accuracy (%) on Office+Caltech10 and Amazon Review datasets with $\lambda = 0, 10, +\infty$	57
4.1	Introduction of datasets.	72
4.2	Accuracy (%) on Amazon Review datasets.	73
4.3	Accuracy (%) on Office-Home datasets using ResNet50 features.	74
4.4	Accuracy (%) on Office+Caltech256 datasets using DeCAF6 features.	74
4.5	Accuracy (%) on Office+Caltech256 datasets using SURF features.	75
4.6	Accuracy (%) on ImageCLEF-DA datasets using ResNet50 features.	75
4.7	Accuracy (%) on Office31 datasets using ResNet50 features.	76
4.8	Accuracy (%) on Office31 datasets using DeCAF6 features.	76

5.1	Introduction of datasets.	102
5.2	Acc(OS*) and Acc(OS) (%) on Office-31, Office-Home and PIE Datasets.	108
6.1	The performance on dataset CIFAR-10 is evaluated by macro-averaged F1-scores in 11 classes (10 known classes and 1 unknown class). We report the experimental results reproduced by [176]. A larger score is better.	139
6.2	The performance on dataset MNIST is evaluated by macro- averaged F1-scores in 11 classes.	140
6.3	The performance on MNIST, SVHN, CIFAR-10, CIFAR+10 and CIFAR+50 are evaluated by macro-averaged F1 scores. We report the experimental results reported by [180] . . .	140
6.4	Ablation study on dataset MNIST, Omnilot, MNIST-Noise and Noise.	140
7.1	Parameters for different tasks on KHDA.	176
7.2	Accuracy (%) with standard error on Office+Caltech256 dataset from SURF \rightarrow DeCAF6.	181
7.3	Accuracy (%) with standard error on Office+Caltech256 dataset from DeCAF6 \rightarrow SURF.	181
7.4	Accuracy (%) with standard error on Reuters-21578 dataset.	182
7.5	Accuracy (%) with standard error on Wikipedia dataset.	182
7.6	Accuracy (%) with standard error on Multilingual Reuters Collection (MRC) dataset.	183
7.7	Accuracy (%) with standard error on ImageCLEF dataset from ResNet-50 \rightarrow VGG-16.	184
7.8	Accuracy (%) with standard error on ImageCLEF dataset from VGG-16 \rightarrow ResNet-50.	185
7.9	Accuracy (%) of ablation study of KHDA on Office+Caltech256, Wikipedia and Reuters-21578 datasets.	185

List of Symbols

\mathcal{X}	Feature space
\mathbf{x}	Sample from \mathcal{X}
\mathcal{Y}	Label set
\mathbf{y}	Label (one-hot vector) from \mathcal{Y}
\mathcal{X}_s	Source feature space
\mathbf{x}_s	Source sample from \mathcal{X}_s
\mathcal{S}	Labeled source samples
\mathcal{X}_t	Target feature space
\mathbf{x}_l	Labeled target sample from \mathcal{X}_t
\mathcal{T}_l	Labeled target samples
\mathbf{x}_u	Unlabeled target sample from \mathcal{X}_t
\mathcal{T}_u	Unlabeled target samples
\mathcal{Y}_s	Source label set
\mathbf{y}_s	Label (one-hot vector) from \mathcal{Y}_s
\mathcal{Y}_t	Target label set
\mathbf{y}_t	Label (one-hot vector) from \mathcal{Y}_t
P_{XY}	Joint distribution
$P_{X_s Y_s}$	Source joint distribution
$P_{X_t Y_t}$	Target joint distribution
P_X	Marginal distribution
P_{X_s}	Source marginal distribution
P_{X_t}	Target marginal distribution
$P_{Y X}$	Conditional distribution
$P_{Y_s X_s}$	Source conditional distribution
$P_{Y_t X_t}$	Target conditional distribution

ℓ	Loss function
\mathbf{h}	Hypothesis function
\mathcal{H}	Hypothesis space
$k(\cdot, \cdot), k_s(\cdot, \cdot), k_t(\cdot, \cdot)$	Reproducing kernel (RK), source RK, target RK
$\mathcal{H}_k, \mathcal{H}_{k_s}, \mathcal{H}_{k_t}$	RKHS with kernels $k(\cdot, \cdot), k_s(\cdot, \cdot), k_t(\cdot, \cdot)$
$d_{\mathcal{H}}^{\ell}, d_{\mathbf{h}, \mathcal{H}}^{\ell}$	$\mathcal{H}\Delta\mathcal{H}$ -divergence, disparity discrepancy
$D_{\mathbf{h}}, D_{\mathcal{F}}$	Projected MMD distance, MMD distance
χ	Compatibility
$\text{err}, \widehat{\text{err}}$	Transfer error rate, empirical transfer error rate
Λ	Combined risk
$\mathbf{T}_s, \mathbf{T}_t$	Source and target transformations
$\mathcal{F}_s, \mathcal{F}_t$	Source and target transformation spaces
$R_P^{\alpha}, R_Q^{\alpha}$	α -risks corresponding to P_{XY}, Q_{XY}
$R_{P,k}, R_{Q,k}$	Partial risks for known classes for P_{XY}, Q_{XY}
$R_{P,u}, R_{Q,u}$	Partial risks for unknown classes for P_{XY}, Q_{XY}
R_s, R_t	Source risk, target risk risk
$\widehat{R}_s, \widehat{R}_t$	Empirical source risk, empirical target risk
R_t^*	Partial risk on known target classes
R_t^{K+1}	Partial risk on unknown target classes