Advances in Computer Vision and Pattern Recognition

Founding editor

Sameer Singh, Rail Vision, Castle Donington, UK

Series editor

Sing Bing Kang, Microsoft Research, Redmond, WA, USA

Advisory Board

Horst Bischof, Graz University of Technology, Austria
Richard Bowden, University of Surrey, Guildford, UK
Sven Dickinson, University of Toronto, ON, Canada
Jiaya Jia, The Chinese University of Hong Kong, Hong Kong
Kyoung Mu Lee, Seoul National University, South Korea
Yoichi Sato, The University of Tokyo, Japan
Bernt Schiele, Max Planck Institute for Computer Science, Saarbrücken, Germany
Stan Sclaroff, Boston University, MA, USA

UNCORRECTED PROOF

More information about this series at http://www.springer.com/series/4205

UNCORRECTED PROOF

Layout: **T1 Standard** Book ID: **420546_1_En** Book ISBN: **978-3-319-58346-4**
Chapter No.: **FM 1** Date: **24-5-2017** Time: **9:30 am** Page: **3/10**

Gabriela Csurka
Editor

Domain Adaptation in Computer Vision Applications

UNCORRECTED PROOF

Springer

Editor
Gabriela Csurka
Xerox Research Centre Europe
Meylan
France

ISSN 2191-6586 ISSN 2191-6594 (electronic)
Advances in Computer Vision and Pattern Recognition
ISBN 978-3-319-58346-4 ISBN 978-3-319-58347-1 (eBook)
DOI 10.1007/978-3-319-58347-1

Library of Congress Control Number: 2017939548

© Springer International Publishing AG 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

UNCORRECTED PROOF

To Gabriel, Elisabeth and Mikhaë

UNCORRECTED PROOF

Preface

While the proliferation of sensors being deployed in cell phones, vehicles, buildings, roadways, and computers allows for larger and more diverse information to be collected, the cost of acquiring labels for all these data remains extremely high. To overcome the burden of annotation, alternative solutions have been proposed in the literature to learn decision making models by exploiting unlabeled data from the same domain (data acquired in similar conditions as the targeted data) or also data from related but different domains (different datasets due to different conditions or provided by different customers). In many real-world machine learning scenarios, using only the data from the same domain might be insufficient and data or models borrowed from similar domains can significantly improve the learning process. Such a process, referred to as *domain adaptation*, aims to leverage labeled data in one or more related domains (sources), in order to build models for a target domain.

Domain adaptation is particularly critical for service companies, where all machine learning components deployed in a given service solution should be customized for a new customer either by annotating new data or, preferably, by calibrating the models in order to achieve a contractual performance in the new environment. While adaptation across domains is a challenging task for many applications, in this book, we focus on solutions for *visual applications*.

The aim of the book is to give a relatively broad view of the field by selecting a diverse set of methods which made different advances in the field. The book begins with a comprehensive survey of domain adaptation and transfer learning, including historical shallow methods, more recent methods using deep architectures, and methods addressing computer vision tasks beyond image categorization, such as detection, segmentation or visual attributes. Then, Chap. 2 gives a deeper look at dataset bias in existing datasets when different representations including features extracted from deep architectures are used. The rest of the book is divided into four main parts, following the same structure as the survey presented in Chap. 1.

Part I is dedicated to shallow domain adaptation methods, beginning with the widely used geodesic flow kernel (Chap. 3) and subspace alignment (Chap. 4). Both chapters propose solutions for selecting landmark samples in the source dataset. Chapter 5 presents domain-invariant embedding methods and Chap. 6 describes

UNCORRECTED PROOF

transductive transfer machines, a method that combines local feature space transformation with classifier selection and parameter adaptation. The first part ends with Chap. 7 that addresses domain adaptation cases where the access to the source data is constrained.

Part II is dedicated to deep adversarial discriminative domain adaptation methods. The first two methods presented use a confusion loss as an adversarial objective to adapt the source network towards the target data. The deep CORAL (Chap. 8) learns a nonlinear transformation that aligns correlations of activation layers of the deep model. The deep domain confusion network (Chap. 9) uses a maximum mean discrepancy based domain confusion loss to induce domain invariant representations. In contrast, Chap. 10 presents the domain-adversarial neural network that integrates a gradient reversal layer to promote the emergence of features discriminative for the main learning task and non-discriminate with respect to the domain shift.

Part III is a collection of contributions addressing domain adaptation problems different from classical image categorization. As such, Chap. 11 focuses on Fisher vector based patch encoding adaptation in the context of vehicle re-identification. Chapter 12 explores the adaptation of semantic segmentation models trained on synthetic images to correctly operate in real scenarios. Chapter 13 addresses the challenge of pedestrian detection by adapting a deformable part-based model trained on virtual-world data to real world data using structure-aware adaptive structural SVMs. Finally, Chap. 14 proposes a method to generalize semantic part detectors across domains.

Part IV concludes the book with unifying perspectives. On the one hand, Chap. 15 proposes to use multi-source domain generalization techniques for the purpose of learning cross-category generalizable attribute detectors. On the other hand, Chap. 16 proposes a common framework that unifies multi-domain and multi-task learning which can be flexibly applied also to zero-shot learning and zero-shot domain adaptation.

Overall, this comprehensive volume, designed to form and inform professionals, young researchers, and graduate students, is the first collection dedicated to domain adaptation for visual applications. In this book I wanted not only to address historically shallow and recent deep domain adaptation methods, but also contributions focused on object or object part detection, re-identification, image segmentation, attribute detection as well as present generic frameworks that unify domain adaptation with multi-domain, multi-task and zero-shot learning.

To give such a broad view, I brought together leading experts in the field to showcase their techniques. I would like to thank them specially for accepting my invitation and for their dedicated effort to share in this book their valuable experiences in the various chapters. Finally, I would also like to thank our Springer editors, Wayne Wheeler and Simon Rees, for their advice and their help in guiding me through the book production process.

Meylan, France
February 2017

Gabriela Csurka

UNCORRECTED PROOF

Contents

UNCORRECTED PROOF

Layout: T1 Standard | Book ID: 420546_1_En | Book ISBN: 978-3-319-58346-4
Chapter No.: FM 1 | Date: 24-5-2017 Time: 9:30 am | Page: 10/10

UNCORRECTED PROOF

Chapter 6
Adaptive Transductive Transfer Machines: A Pipeline for Unsupervised Domain Adaptation

Nazli Farajidavar, Teofilo de Campos and Josef Kittler

Abstract This chapter addresses the problem of transfer learning by unsupervised domain adaptation. We introduce a pipeline which is designed for the case where the joint distribution of samples and labels $P(\mathbf{X}^{src}, \mathbf{Y}^{src})$ in the source domain is assumed to be different, but related to that of a target domain $P(\mathbf{X}^{trg}, \mathbf{Y}^{trg})$, but labels $\mathbf{Y}^{trg}$ are not available for the target set. This is a problem of Transductive Transfer Learning. In contrast to other methodologies in this book, our method combines steps that adapt both the marginal and the conditional distributions of the data.

6.1 Introduction

The transfer learning (TL) taxonomy presented by Pan and Yang [355] and described also in Chap. 1 classifies TL approaches into three main categories: Inductive TL, when labeled samples are available in both source and target domains; Transductive TL, when labels are only available in the source set, and Unsupervised TL, when labeled data is not present. They also categorized the methods based on *instance re-weighting* (e.g., [91, 111]), *feature space transformation* (e.g., [45, 312]) and *learning parameters transformation* (e.g., [21, 55]).

The work presented in this chapter has its focus on Transductive TL, also known as Domain Adaptation (DA) problems. While different methods can potentially be combined to achieve a successful transductive transfer, in this work we have mainly restricted our attention to methods which focus on *feature space transformation*, *learning parameters transformation* and their combination.

N. Farajidavar (✉) · J. Kittler (✉)
University of Surrey, Guildford, UK
e-mail: N.Davar@surrey.ac.uk; Nazli.farajidavar@eng.ox.ac.uk

J. Kittler
e-mail: j.kittler@surrey.ac.uk

T. de Campos (✉)
Universidade de Brasília, Gama-DF, Brazil
e-mail: t.decampos@st-annes.oxon.org

© Springer International Publishing AG 2017

G. Csurka (ed.), *Domain Adaptation in Computer Vision Applications*,
Advances in Computer Vision and Pattern Recognition,
DOI 10.1007/978-3-319-58347-1_6

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

UNCORRECTED PROOF

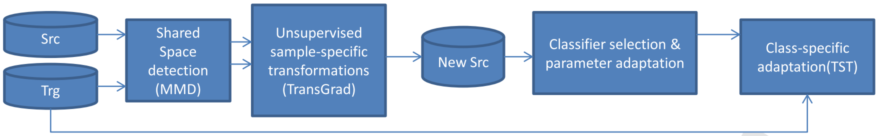


Fig. 6.1 The Adaptive Transductive Transfer Machine (ATTM)

Among the researchers following a similar line of work, Long et al. [312] proposed to do Joint Distribution Adaptation (JDA) by iteratively adapting both the marginal and conditional distributions modified Maximum Mean Discrepancy (MMD) algorithm [45]. JDA uses the pseudo target labels to define a shared subspace between the two domains. At each iteration, this method requires the construction and eigen decomposition of an $n \times n$ matrix whose complexity can be up to $O(n^3)$ where $n = n_{src} + n_{trg}$ is the total number of samples. Similarly, Gong et al. in [200] proposed a kernel-based domain adaptation method that exploits intrinsic low-dimensional structures in the datasets.

In this chapter[1] we propose a Transductive Transfer Machine (TTM) algorithm which combines methods that adapt the marginal and the conditional distribution of the samples, so that the source and target datasets become more similar. After adaptation, the transformed source domain data can be used to design a classifier for the target domain's samples. The TTM approaches this problem by combining four types of adaptation: (a) solving the task in a lower dimensional space that is shared between the two domains, (b) a set of local transformations to further increase the domain similarity, and (c) a set of class-conditional transformations aiming to increase the similarity between the posterior probability of samples in the source and target sets, (d) and finally we introduce the Adaptive TTM (ATTM), which uses two unsupervised dissimilarity measures before step (c) to perform classifier selection and automatic kernel parameter tuning.

Section 6.2 presents the core TTM components of our method and discusses the relation with previous works. This is followed in Sect. 6.3 by a description of our ATTM framework. In Sect. 6.4, the proposed pipeline is compared against other state-of-the-art methods and showing performance boost in cross-domain image classification, using various datasets. Section 6.5 concludes the paper.

6.2 Marginal and Conditional Distribution Adaptation

In order to introduce the ATTMdepicted in Fig. 6.1, we will first present its core component, *feature space transformation*, referred to as Transductive Transfer Machines (TTM), summarized in the steps below:

[1] This chapter is an amalgamation of the works published in [152–154] with additional analysis taking into account the works of other authors which were developed concurrently to our work.

420546_1_En_6_Chapter ✓ TYPESET DISK LE ✓ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

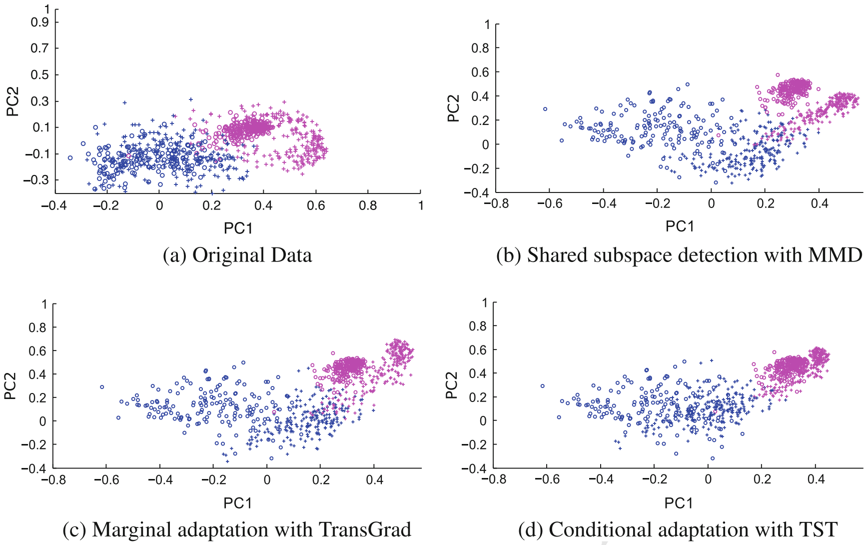


Fig. 6.2 The effect of different steps of the pipeline on digits 1 and 2 of the MNIST → USPS datasets, visualized in 2D through PCA. Source samples (MNIST) are indicated by stars, target ones (USPS) by circles, *red* indicates samples of digit 1 and *blue* indicates digit 2

1. A global linear transformation G' is applied to $\mathbf{X}^{src}$ and $\mathbf{X}^{trg}$ such that the marginal distribution of the source samples, $P(G'(\mathbf{X}^{src}))$ becomes more similar to that of target's, $P(G'(\mathbf{X}^{trg}))$. This is done by minimizing the MMD between the sets as described in Sect. 6.2.1.
2. Aiming to minimize the difference between the marginal distributions, a local transformation is applied to each transformed source domain sample $G''_i(G'(\mathbf{x}_i^{src}))$. This transformation, dubbed TransGrad, uses the gradient of the target data log-likelihood at each source sample. Details are in Sect. 6.2.2.
3. Finally, aiming to reduce the difference between the conditional distributions in source and target spaces, a class-based transformation is applied to each class of the transformed source samples $G'''_{y_i}(G''_i(G'(\mathbf{x}_i^{src})))$. While the first two steps are unsupervised transfer learning methods, this step is transductive, as it uses source labels. The transformation applies translation and scale transformations (TST) to each training set, as described in Sect. 6.2.3.

Fig. 6.2 illustrates these steps using a dataset composed of digits 1 and 2 from MNIST and USPS datasets. The first two principal components of the source data are used to project the data into a two dimensional space for a better visualization.

UNCORRECTED

6.2.1 Shared Space Detection with MMD

In the first step of TTM pipeline, we look for a shared space projection that reduces dimensionality of the data whilst minimizing the reconstruction error. As explained in [312], one possibility for that is to search for an orthogonal transformation matrix $\mathbf{W} \in \mathbb{R}^{f \times f'}$ such that the embedded data variance is maximized,

$$\max_{\mathbf{W}^\top \mathbf{W}=I} Tr(\mathbf{W}^\top \mathbf{X}\mathbf{H}\mathbf{X}^\top \mathbf{W}) , \tag{6.1}$$

where $\mathbf{X} = [\mathbf{X}^{src}; \mathbf{X}^{trg}] \in \mathbb{R}^{f \times n_{src}+n_{trg}}$ is the input data matrix that combines source and target samples, $Tr(\cdot)$ is the trace of a matrix, $\mathbf{H} = I - \frac{1}{n_{src}+n_{trg}} \mathbb{1}$ is a centering matrix where I is the identity matrix, $\mathbb{1}$ is a $(n_{src} + n_{trg}) \times (n_{src} + n_{trg})$ matrix of ones and f' is the dimensionality after the projection where $f' \leq f$.

The optimization problem can be efficiently solved by eigen decomposition. However, the above PCA-based representation may not reduce the difference between source and target domains, hence the need for a more appropriate transformation remains.

Following [213, 312, 354, 471] the empirical MMD measure, proposed in [354], is used as the distance measure to compare different distributions. This algorithm searches for a projection matrix, $\mathbf{W} \in \mathbb{R}^{f \times f'}$ which minimizes the distance between the means of the two distributions:

$$\left\| \frac{1}{n_{src}} \sum_{i=1}^{n_{src}} \mathbf{W}^\top \mathbf{x}_i - \frac{1}{n_{trg}} \sum_{j=n_{src}+1}^{n_{src}+n_{trg}} \mathbf{W}^\top \mathbf{x}_j \right\|^2 = Tr(\mathbf{W}^\top \mathbf{X}\mathbf{M}\mathbf{X}^\top \mathbf{W}) \tag{6.2}$$

where $\mathbf{M}$ is the MMD matrix and is computed as follows:

$$\mathbf{M}_{ij} = \begin{cases} \frac{1}{n_{src} n_{src}}, & \mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}^{src} \\ \frac{1}{n_{trg} n_{trg}}, & \mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}^{trg} \\ \frac{-1}{n_{src} n_{trg}}, & \text{otherwise.} \end{cases} \tag{6.3}$$

The constraint optimization problem then is to minimize Eq. (6.2) such that Eq. (6.1) is maximized, i.e., solve the following eigen-decomposition problem: $(\mathbf{X}\mathbf{M}\mathbf{X}^\top + \varepsilon \mathbb{I})\mathbf{W} = \mathbf{X}\mathbf{H}\mathbf{X}^\top \mathbf{W}\mathbf{D}$, giving the eigenvectors $\mathbf{W}$ and the associated eigenvalues in the form of the diagonal matrix $\mathbf{D}$. The effect is to obtain a lower dimensional shared space between the two domains. Consequently under the new representation $G'(\mathbf{x}) = \mathbf{W}^\top \mathbf{X}$, the marginal distributions of the two domains are drawn closer to each other, as the distance between their means is minimized. The effect of this transformation is shown[2] in Fig. 6.2b.

[2] Note however that in Fig. 6.2b a 2D view of feature space was generated using PCA and only two out of ten classes of digits in MNIST/USPS dataset are shown, while the MMD computation was

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

6.2.2 Sample-Based Adaptation with TransGrad

In the next step of the pipeline, we propose a sample-based transformation that shifts the source probability density function toward target clusters. Via the TransGrad step a set of local translations is applied to the source samples, making their distribution more similar to that of the target samples.

In general, target data may, but does not have to, lie in the same observation space. However, for the sake of simplicity, we shall assume that the transformation of the source to the target domain is locally linear, i.e., a sample's feature vector $\mathbf{x}$ from the source domain is shifted to the target space by

$$G''(\mathbf{x}) = \mathbf{x} + \gamma \mathbf{b}_{\mathbf{x}} \,, \tag{6.4}$$

where the f dimensional vector $\mathbf{b}_{\mathbf{x}}$ represents a local offset in the target domain and γ is a translation regulator. In order to impose as few assumptions as possible, we shall model the unlabeled target data, $\mathbf{X}^{trg}$ by a mixture of Gaussian probability density functions, $p(\mathbf{x}|\lambda) = \sum_{k=1}^{K} w_k p(\mathbf{x}|\lambda_k)$, whose parameters are denoted by $\lambda = \{w_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, k = 1, \cdots, K\}$ where w_k, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ denote the weight, mean vector and covariance matrix of Gaussian component k, respectively, and K denotes the number of components $p(\mathbf{x}|\lambda_k) = \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$.

The problem of finding an optimal translation parameter $\mathbf{b}_{\mathbf{x}}$ can then be formulated as one of moving the source point $\mathbf{x}$ to a new location $G''(\mathbf{x}) = \mathbf{x} + \gamma \mathbf{b}_{\mathbf{x}}$ to increase its likelihood as measured using $p(G''(\mathbf{x})|\lambda^{trg})$. Using the Taylor expansion, in the vicinity of $\mathbf{x}$, the likelihood of $p(\mathbf{x} + \gamma \mathbf{b}_{\mathbf{x}})$ can be expressed as:

$$p(\mathbf{x} + \gamma \mathbf{b}_{\mathbf{x}}|\lambda) = p(\mathbf{x}|\lambda) + \gamma (\nabla_{\mathbf{x}} p(\mathbf{x}|\lambda))^{\top} . \mathbf{b}_{\mathbf{x}} \,. \tag{6.5}$$

We wish to maximize the $p(\mathbf{x} + \gamma \mathbf{b}_{\mathbf{x}}|\lambda)$ with respect to the unknown parameter, $\mathbf{b}_{\mathbf{x}}$. The learning problem then can be formulated as:

$$\max_{\mathbf{b}_{\mathbf{x}}} \left(p(\mathbf{x}|\lambda) + \gamma (\nabla_{\mathbf{x}} p(\mathbf{x}|\lambda))^{\top} . \mathbf{b}_{\mathbf{x}} \right) \quad \text{s.t. } \mathbf{b}_{\mathbf{x}}^{\top} . \mathbf{b}_{\mathbf{x}} = 1 \,. \tag{6.6}$$

The Lagrangian of Eq. (6.6) is $p(\mathbf{x}|\lambda) + \gamma (\nabla_{\mathbf{x}} p(\mathbf{x}|\lambda))^{\top} . \mathbf{b}_{\mathbf{x}} - \gamma' (\mathbf{b}_{\mathbf{x}}^{\top} . \mathbf{b}_{\mathbf{x}} - 1)$. Setting its gradient with respect to $\mathbf{b}_{\mathbf{x}}$ to zero

$$\nabla_{\mathbf{x}} p(\mathbf{x}|\lambda) - \gamma'' \mathbf{b}_{\mathbf{x}} = 0 \,, \tag{6.7}$$

where γ'' is considered as TransGrad's step size parameter and is equal to $\frac{2\gamma'}{\gamma}$, we find that the source data-point $\mathbf{x}$ should be moved in the direction of maximum gradient of the function $p(\mathbf{x}|\lambda)$. Accordingly, $\mathbf{b}_{\mathbf{x}}$ is defined as

---

(Footnote 2 continued)
done in a higher dimensional space with samples from all ten classes. For these reasons it may not be easy to see that the means of the source and target samples became closer after MMD.

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

$$\mathbf{b_x} = \nabla_{\mathbf{x}} p(\mathbf{x}|\lambda) = \sum_{k=1}^{K} w_k p(\mathbf{x}^{src}|\lambda_k).\boldsymbol{\Sigma}_k^{-1}(\mathbf{x} - \boldsymbol{\mu}_k) \; . \tag{6.8}$$

In practice, Eq. (6.4) translates $\mathbf{x}^{src}$ using the combination of the translations between $\mathbf{x}^{src}$ and $\boldsymbol{\mu}_k$, weighted by the likelihood of $G''(\mathbf{x}^{src})$ given λ_k. Up to our knowledge, this is the first time a sample-based transformation is proposed for transfer learning. The effect of this transformation can be seen in Fig. 6.2c.

6.2.3 Conditional Distribution Adaptation with TST

In order to reduce the class-conditional distribution mismatch between the corresponding clusters of the two domains, we used a set of linear class-specific transformations which we refer to as translation and scaling transformation, or TST. To achieve this, we assume that a Gaussian Mixture Model (GMM) fitted to the source classes can be adapted in a way that it matches to target classes. We follow Reynolds et al. [385] and use only diagonal covariance matrices in the GMM, making the complexity of the estimation system linear in f. In our experiments, we further simplify the model for this step of the pipeline by using only one Gaussian distribution per class which is not unrealistic considering the fact that what we are eventually interested in are compact and dense classes.

In order to adapt the class-conditional distributions, one can start with an attempt to match the joint distribution of the features and labels between the corresponding clusters of the two domains. However, in Transductive Transfer application scenarios, labeled samples are not available in the target domain. We thus use posterior probability of the target instances to build class-based models in the target domain. This relates to JDA [312], which uses pseudo-labels to iteratively update a supervised version of MMD. In our case, class-based adaptations are simplified to translation and scaling transformations, making the computational cost very attractive.

The proposed transformation adjusts the mean and standard deviation of the corresponding clusters from the source domain, i.e., each feature j of each sample $\mathbf{x}^i$ is adapted as follows:

$$G_{y^i}(x_j^i) = \frac{x_j^i - E^{src}[x_j, y^i]}{\sigma_{j,y_i}^{src}} \sigma_{j,y^i}^{trg} + E_{\Lambda_{src}}^{trg}[x_j, y^i] \; , \forall i = 1 \colon n_{src} \; , \tag{6.9}$$

where σ_{j,y^i}^{src} is the standard deviation of feature x_j of the source samples labeled as y^i and $E^{src}[x_j, y^i]$ is the joint expectation of the feature x_j and labels y^i defined by

$$E^{src}[x_j, y^i] = \frac{\sum_{i=1}^{n_{src}} x_j^i \mathbb{1}_{[y]}(y^i)}{\sum_{i=1}^{n_{src}} \mathbb{1}_{[y]}(y^i)} \; . \tag{6.10}$$

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

In Eq. (6.10) $\mathbb{1}_{[y]}(y^i)$ is an indicator function.[3]

An estimation of the target joint expectation is thus formulated as

$$E^{trg}[x_j, y] \approx E^{trg}_{\Lambda_{src}}[x_j, y] = \frac{\sum_{i=1}^{n_{trg}} x^i_j P_{\Lambda_{src}}(y|\mathbf{x}_i)}{\sum_{i=1}^{n_{trg}} P_{\Lambda_{src}}(y|\mathbf{x}_i)} \tag{6.11}$$

We propose to estimate the standard deviation per feature and per class using

$$\sigma^{trg}_{j,y^i} = \sqrt{\frac{\sum_{n=1}^{n_{trg}} (x^n_j - E^{trg}_{\Lambda_{src}}[x_j, y^i])^2 P_{\Lambda_{src}}(y^i|\mathbf{x}_n)}{\sum_{n=1}^{n_{trg}} P_{\Lambda_{src}}(y^i|\mathbf{x}_n)}} . \tag{6.12}$$

In summary, in a common DA problem, the joint expectation of the features and labels over source distribution, $E^{src}[x_j, y^i]$, is not necessarily equal to $E^{trg}[x_j, y^i]$. Therefore, one can argue that if the expectations in the source and target domains are induced to be similar, then the model Λ learned on the source data will generalize well to the target data. Consequently, the less these distributions differ, the better the trained model will perform.

Since the target expectation $E^{trg}_{\Lambda_{src}}[x_j, y^i]$ is only an approximation based on the target's posterior probabilities, rather than the ground-truth labels (which are not available in the target set), there is a danger that samples that would be miss-classified could lead to negative transfer, i.e., negative impact. To alleviate this, we follow Arnold et al.'s [18] suggestion and smooth out the transformation by applying the following mapping

$$G'''_{y^i}(x^i_j) = (1-\theta)x^i_j + \theta G_{y^i}(x^i_j) , \tag{6.13}$$

where $\theta \in [0, 1]$ is the transfer rate parameter. As it can be inferred from the MMD and TST equations, the effect of the first transformation is that it tries to find a shared subspace between the two domains to reduce the distributional mismatch at a global level second one is actually a class-specific transformation aiming to reduce the class-conditional mismatch among the clusters from one domain to another.

Iterative refinement of the conditional distribution. Matching the marginal distributions does not guarantee that the conditional distribution of the target can be approximated to that of the source. To our knowledge, most of the recent works related to this issue [55, 76, 378, 587] are Inductive Transfer Learning methods and they have access to some labeled data in the target domain which in practice makes the posteriors' estimations easier.

Instead, our class-specific transformation method (TST), reduces the difference between the likelihoods $P(G'''_y(\mathbf{x}^{src})|y = c)$ and $P(\mathbf{x}|y = c)$ by using the target posteriors estimated from a model trained on gradually modified source domain Eq. (6.13). Hence, these likelihood approximations will not be reliable unless we

[3]Our method uses insights from Arnold et al. [18], but Eqs. (6.10) and (6.11) rectify those from [18], as discussed in [154].

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

iterate over the whole distribution adaptation process and retrain the classifier model using $G_y'''(\mathbf{x}^{src})$.

Global dissimilarity as stopping criterion. In order to automatically control the number of the iterations in our pipeline, we introduce a domain dissimilarity measure inspired by sample selection bias correction techniques [99, 434]. Many of those techniques are based on weighting samples $\mathbf{x}_i^{src}$ using the ratio $w(\mathbf{x}) = \frac{P(\mathbf{x}|trg)}{P(\mathbf{x}|src)}$. This ratio can be estimated using a classifier that is trained to distinguish between source and target domains, i.e., samples are labeled as either belonging to class src or trg. Based on this idea, we use this classification performance as a measure of dissimilarity between two domains, i.e., if it is easy to distinguish between source and target samples, it means they are dissimilar. We coin this measure as Global Dissimilarity, $D^{\text{global}}(\mathbf{X}^{src}, \mathbf{X}^{trg})$ which is defined by the accuracy of a nearest neighbor domain classifier using a random split of training and test samples, each containing 50% of the data. If the domain dissimilarity is high, then more iterations are needed to achieve a better match between the domains.

Note that other methods could be used as stopping criteria. For instance by checking the incremental change in the transformation between two consecutive iterations we could stop the iterations in case that this measure is below a specific threshold, e.g., using the Frobenius norm between the covariances of the transformed source matrices of two consecutive iterative steps. However, we use $D^{\text{global}}(\mathbf{X}^{src}, \mathbf{X}^{trg})$ because this same measure is also engaged for selecting classifiers.

6.3 ATTM via Classifier Selection and Parameter Adaptation

We do not assume that source and target domain samples follow the same distribution, so the best performing learner for the source set may not be the best for the target set. We propose to use dissimilarity measures between source and target sets in order to select the classifier and adjust its kernel parameters. The empirical results showed that the optimization of SVM using grid search in the parameter space with cross-validation on the source led to over-fitting. We therefore prefer to use Kernel LDA (KDA) [57] and PCA+NN classifiers as the main learners.

To select between these classifiers and to adapt the KDA kernel length-scale parameter, we propose to use two measures. The first is the **Global Dissimilarity** between the source and target distributions, described in Sect. 6.2.3. The second measure, coined **Clusters Dissimilarity** ($D^{\text{clusters}}(\mathbf{X}^{src}, \mathbf{X}^{trg})$), is proportional to the average dissimilarity between the source and target clusters, computed using the average of the distances between the source class centers and their nearest target cluster center. The target clusters centers are obtained using K-means on the target data, initialized using source class centers. We therefore assume that there is no shuffling in the placement of the clusters from one domain to another.

The proposed **Clusters Dissimilarity** is similar to the cross-domain sparse-shot similarity of Xu et al. [549] which is used for multi-source motion tracking. Xu

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

et al. proposed to use object motion tracking data in each domain and compared tracks across domains using the Kullback-Leibler Divergence between GMMs that represent them.[4]

When both dissimilarity measures indicate that the cross-domain datasets are very different, the choice of a nonparametric classifier such as Nearest Neighbor (NN) is preferred, requiring no optimization during training. When the two domains are similar at the global level, the choice of a parametric classifier such as KDA is more sensible, however, with care taken, to avoid over-fitting on the training set. So if the local dissimilarity is high, the kernel parameters must be adapted.

Following the common practice in the vision community (e.g., [521]), we initially set σ parameter of the Radial Basis Function (RBF) kernel in KDA to

$$\sigma = \frac{1}{n_{src}^2} \sum_{i,j}^{n_{src}} |\mathbf{x}_i - \mathbf{x}_j|_1, \forall \mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}^{src} \tag{6.14}$$

where ℓ^1 norm is used in the kernel function. This is then adapted using a linear function of the cluster dissimilarity measure

$$\sigma' = \sigma \gamma''' D^{\text{clusters}}(\mathbf{X}^{src}, \mathbf{X}^{trg}) \;, \tag{6.15}$$

where γ''' is a constant value which is empirically set to be the average cluster dissimilarity obtained in a set of cross-domain comparisons. This was devised based on the fact that the credibility of a classifier is inversely proportional to the dissimilarity between training and test samples. In the case of KDA, the best way to tune its generalization ability is via the kernel length-scale.

Note that the cluster dissimilarity measure can only be computed if enough samples are available in both source and target sets or if they are not too unbalanced. When these conditions are not satisfied, our algorithm avoids kernel-based method and selects the NN classifier. The parameter selection and model adaptation mechanism is summarized in Table 6.1, where the arrows pointing up ($\Uparrow$) indicate high dissimilarity and arrows pointing down ($\Downarrow$) indicate low dissimilarity.[5]

In conclusion, our ATTM pipeline will use a PCA+NN classifier as its main learner model if the global dissimilarity between the two domains is high or there are not enough source samples and consequently not enough cluster-wise samples to train a highly reliable learner model or to further adjust the classifier parameters. In any other circumstances, the model will use the KDA classifier and adjusts the kernel length-scale if required.

Computational complexity. The proposed TTM method for feature space adaptation has computational cost as follows:

1. MMD: $O(n^2)$ for constructing the MMD matrix, $O(nf^2)$ for covariance computation and $O(f^3)$ for eigen decomposition.

[4] Table 6.3 shows these two measures computed on all datasets, discussed later.

[5] The measures were judged as high or low based on a subset of values observed in Table 6.3.

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

Table 6.1 Classifier selection and length-scale adaptation

D^{global}	$D^{clusters}$	Classifier	How to set σ'
$\Uparrow$	$\Uparrow$	NN	—
$\Downarrow$	$\Downarrow$	KDA	$\sigma' = \sigma^{src}$ Eq. (6.14)
$\Downarrow$	$\Uparrow$	KDA	$\sigma' = \sigma^{src}\gamma''' D^{\text{clusters}}(\mathbf{X}^{src}, \mathbf{X}^{trg})$

2. TransGrad: $O(nK)$ for Expectation step of GMM computation, $O(nKf)$ for the computation of covariance matrices and $O(K)$ for the Maximization step of the GMM computation. Once the GMM is built, the TransGrad transformation itself is $O(nKf)$.
3. TST: $O(Cnf)$ for class-specific TST transformations where C is the number of classes.
4. NN classifier: zero for training and $O(n^2 f)$ for reapplying the classifier.

For each of the T iterations, the classifier is re-applied and TST is computed. Therefore, the overall complexity of our training algorithm is dominated by the cost of training a GMM (which is low by using diagonal covariances) and iteratively applying a classifier. The core transformations proposed in this pipeline, TransGrad and TST are $O(nKf)$ and $O(Cnf)$, respectively,i.e., much cheaper than most methods in the literature. MMD is the only component whose complexity is greater than linear on n, but it is executed only once and its main cost comes from eigen decomposition, for which there is a wide range of optimized libraries available.

By adding the classifier selection step and kernel adaptation to TTM, we obtain ATTM, shown in algorithm 4. The classifier selection step uses the computation of the $D^{\text{clusters}}(\mathbf{X}^{src}, \mathbf{X}^{trg})$, which costs $O(n^2)$, as it uses K-means clustering, but this is executed only once. TST, which has linear cost, is the main part of the algorithm. As it uses source labels, it is iterated. The most expensive part of the loop is the re-training and application of classifiers.

Algorithm 4: ATTM: Adaptive Transductive Transfer Machine

Input: $\mathbf{X}^{src}, \mathbf{Y}^{src}, \mathbf{X}^{trg}$

1. Search for the shared subspace between the two domains (MMD, Sect. 6.2.1)
2. TransGrad: apply local adjustments to the source marginal distribution (Sect. 6.2.2)
3. Select the appropriate classifier (Sect. 6.3), if it is kernel-based, tune σ using Eq. (6.15)

while $T < 10$ *and* $|D^{global}(G^t(\mathbf{X}^{src}), \mathbf{X}^{trg})| > threshold$ **do**

4. Find the feature-wise TST transformation Eqs. (6.9), (6.11), 6.12)
5. Transform the source domain clusters Eq. (6.13)
6. Retrain the classifier using the transformed source

Output: $\mathbf{Y}^{trg}$

UNCORRECTED PROOF

6.4 Experimental Evaluation

In the experiments of this chapter, we used three public benchmark datasets: the USPS/MNIST [110], COIL20 [341] and Caltech+office (OC10) [407]. These are widely used to evaluate computer vision and transfer learning algorithms, enabling us to compare our results with other state-of-the-art methods. In most of our experiments, we used their standard features, available from their website: raw images for USPS, MNIST and COIL20; and SURFBOV for OC10. In Sect. 6.4.1, we show results using DeCAF [128] features.

Preliminary evaluations. In a preliminary evaluation of the characteristics of the domains and classifiers, we evaluated a set of widely used classifiers on all the datasets using a fivefold cross-validation, reporting mean accuracy measure in Table 6.2. In the case of the NN classifier, we further projected our full space into its principal components (PCA), retaining 90% of the energy. As one can note in most of the experiments KDA is the winning classifier. SVM is also a strong learner but it requires optimization of parameters C and σ, which can make it optimal for the source domain, but not necessarily for the target. It is worth noting that PCA+NN's performance is remarkably close to that of KDA on the first two datasets and it is significantly superior on the DSLR dataset.

The two cross-domain dissimilarity measures are shown in Table 6.3. These results justify the design options shown in Table 6.1 so NN is used for the digits datasets (MNIST↔USPS) and where the number of source samples was not adequate for an accurate parameter adaptation (D↔W), and KDA is used for the remaining transfer tasks, with kernel parameters set based on $D^{\text{clusters}}(\mathbf{X}^{src}, \mathbf{X}^{trg})$.

Probing and benchmark results. We performed probing experiments to evaluate the relevance of each component of the proposed system. The simplest design, labeled TTM0 refers to an iterative version of TST [154]; TTM1 is the combination of the MMD and TST; and finally TTM2 adds to TTM1 the samplewise marginal adaptation (TransGrad) applied before TST (see Fig. 6.1). We have also carried out experiments to show that our proposed classifier selection and model adaptation techniques (ATTM) improve the performance of both TTM and JDA algorithms significantly. We compared our methods with four state-of-the-art approaches [200, 312, 354, 438] using the same public datasets and the same settings as those of [200, 312]. The results are in Table 6.4. Further comparisons with other DA methods such as Sampling Geodesic Flow (SGF) using the Grassmann manifolds [206] are reported in [200].

Table 6.2 Evaluation of four classifiers using fivefold cross-validation on individual domains

Classifier	MNIST	USPS	COIL1	COIL2	Caltech	Amazon	Webcam	DSLR
PCA+NN	91.97	93.64	99.02	98.91	38.80	60.59	79.58	**76.95**
LR	86.15	89.22	92.36	92.22	56.27	72.46	80.01	67.49
KDA	**94.05**	94.84	**100.00**	**99.71**	**58.16**	**78.73**	**89.54**	63.94
SVM	91.80	**95.28**	99.72	99.44	57.17	74.86	86.44	75.80

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

Table 6.3 Cross-domain dissimilarities between domains ($src \rightarrow trg$), with datasets abbreviated as M: MNIST, U: USPS, CO1: COIL1, CO2: COIL2, C: Caltech, A: Amazon, W: Webcam, and D: DSLR

src	M	U	CO1	CO2	C	C	C	A	A	A	W	W	W	D	D	D
trg	U	M	CO2	CO1	A	W	D	C	W	D	C	A	D	C	A	W
D^{clusters} (%)	3.4	3.2	2.6	2.5	3.2	3.3	3.1	3.1	3.5	3.6	3.7	3.5	3.7	3.5	3.4	3.3
D^{global} (%)	9.8	9.8	6.3	5.6	5.5	7.8	7.9	6.1	7.4	0.8	7.5	7.2	5.1	7.8	7.9	4.7

UNCORRECTED PROOF

Table 6.4 Recognition accuracies on DA tasks with datasets abbreviated as in Table 6.3. Comparisons in column two start with the baseline accuracy obtained using NN and PCA followed by the results of the algorithms discussed. The last two columns show the effect of the classifier selection and model adaptation techniques (6.3) on JDA and TTM algorithms

Transfer task	NN base-line	PCA base-line	TCA [354]	TSL [438]	GFK (PLS, PCA) [200]	JDA (1NN) [312]	TTM0 (TST, 1NN)	TTM1 (MMD + TTM0)	TTM2 (TransGrad + TTM1)	AJDA (Adapt.JDA)	ATTM (Adapt.TTM2)
M → U	65.94	66.22	56.28	66.06	67.22	67.28	75.94	76.61	77.94	67.28	**77.94**
U → M	44.70	44.95	51.05	53.75	46.45	59.65	59.79	59.41	61.15	59.65	**61.15**
CO1 → 2	83.61	84.72	88.47	88.06	72.50	89.31	88.89	88.75	93.19	**94.31**	92.64
CO2 → 1	82.78	84.03	85.83	87.92	74.17	88.47	88.89	88.61	88.75	**92.36**	91.11
C → A	23.70	36.95	38.20	44.47	41.4	44.78	39.87	44.25	46.76	58.56	**60.85**
C → W	25.76	32.54	38.64	34.24	40.68	41.69	41.02	39.66	41.02	48.81	**62.03**
C → D	25.48	38.22	41.40	43.31	41.1	45.22	50.31	44.58	47.13	45.86	**50.32**
A → C	26.00	34.73	27.76	37.58	37.9	39.36	36.24	35.53	39.62	40.43	**42.92**
A → W	29.83	35.59	37.63	33.90	35.7	37.97	37.63	42.37	39.32	49.83	**50.51**
A → D	25.48	27.39	33.12	26.11	36.31	39.49	33.75	29.30	29.94	38.21	**39.49**
W → C	19.86	26.36	29.30	29.83	29.3	31.17	26.99	29.83	30.36	**35.80**	34.02
W → A	22.96	31.00	30.06	30.27	35.5	32.78	29.12	30.69	31.11	38.94	**39.67**
W → D	59.24	77.07	87.26	87.26	80.89	89.17	85.98	89.17	89.81	89.17	**89.81**
D → C	26.27	29.65	31.70	28.50	30.28	31.52	29.65	31.25	32.06	28.31	**32.41**
D → A	28.50	32.05	32.15	27.56	36.1	33.09	31.21	29.75	30.27	37.47	**38.73**
D → W	63.39	75.93	86.10	85.42	79.1	89.49	85.08	**90.84**	88.81	89.49	88.81
Average	43.06	49.23	50.35	52.34	50.00	54.88	54.12	55.10	56.20	59.17	**60.72**

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

As one can note, all the DA methods improve the accuracy over the baseline. Furthermore, our ATTM method generally outperforms all the other methods. The main reason for that is that our method combines three different feature adaptation techniques with a further classifier parameter adaptation step.

In most of the tasks, both TTM1, 2 algorithms show comparative performance with respect to the JDA [312]. The average performance accuracy of the TTM1 and TTM2 on 16 transfer tasks is **55.10** and **56.20%**, respectively, where the performance improved by **0.22** and **1.32%** compared to the best performing baseline method JDA [200]. Moreover in almost all datasets, TTM2 wins over TTM1 due to its initial domain dissimilarity adjustments using the TransGrad. On average, our methods (TTM1, TTM2 and ATTM) give better results than JDA [312] (and AJDA) because the MMD-based transformation of JDA is coarser than ours. Furthermore, in JDA [312] the number of iterations is a predefined constant, in our algorithm we based this number on a sensible measure of domain dissimilarity described earlier. Moreover, the proposed TTM guarantees an acceptable level of performance about five times faster than the best performing state-of-the-art approach. GFK performs well on some of the OC10 experiments but poorly on the others. The reason is that the subspace dimension should be small enough to ensure that different sub-spaces transit smoothly along the geodesic flow, which may not be an accurate representation of the input data. JDA and TTM perform much better by learning a more accurate shared space.

We also evaluated the proposed classifier selection and model adaptation techniques on JDA [312] and TTM [153]. The results are indicated by AJDA and ATTM in Table 6.4. Their performance shows that the model adaptation significantly enhances the final classifier. One should note that in the cases where our model adaptation technique selects the NN classifier as the main learner of the algorithm, the results remain steady. The performance gains of **4.59** and **4.29%** in ATTM and AJDA, respectively, validate the proposed dissimilarity measures for model selection and adaptation. The proposed model adaptation step of the pipeline selected the NN classifier for MNIST↔USPS and for DSLR→Webcam. For all other transfer problems, KDA was chosen and σ adaptation was used.

Shared subspace projection methods. After developing our MMD-based algorithm, we came across alternative subspace projection methods [25, 164]. In [25] the author proposes the Domain Invariant Projection (DIP) where a Grassmann manifold latent subspace is used to project the data and the MMD measure is subsequently applied for evaluating the source and target domains dissimilarity. The aim is to find a representation of the data that is invariant across different domains. Alternatively, they propose a second projection, DIP-CC, that not only minimizes the distance between the distribution of the projected source and target, but also yields better classification performance. The algorithm searches for a projection that encourages samples with the same labels to form a more compact cluster which is achieved by minimizing the distance between the projected samples of each class and their mean.

In contrast to the manifold alignment methods that use local statistical structure of the data [519, 520, 577], the authors of [164] exploit the global covariance sta-

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

Table 6.5 Recognition accuracies obtained with 1NN classifiers on target domains using different shared subspace projection methods, compared to MMD, i.e., the first step of our TTM

DA experiment	DIP [25]	DIP-CC [25]	SA [164]	MMD
C → A	50.0	**51.8**	39.0	46.1
C → W	47.6	**47.7**	36.8	38.0
C → D	49.0	**51.4**	39.6	45.9
A → C	**43.3**	43.2	35.3	40.6
A → W	46.7	**47.8**	38.6	40.0
A → D	42.8	**43.3**	37.6	31.9
W → C	37.0	**37.1**	32.3	31.3
W → A	**42.5**	41.1	37.4	31.9
W → D	86.4	85.3	80.3	**89.2**
D → C	**39.0**	35.8	32.4	33.4
D → A	40.5	**41.0**	38.0	31.2
D → W	86.7	84.0	83.6	**87.5**
average	**51.0**	50.8	44.2	45.6

tistical structure of the two domains during the adaptation process. The source data is projected onto the source subspace and the target data onto the target subspace in contrast to most domain adaptation methods in the literature. This method, called Subspace Alignment (SA), is totally unsupervised and does not require any target labels. SA makes use of the correlated features in both domains where some of these features can be specific to one domain yet correlated to some other features in the other one allowing the method to use both shared and domain specific features.

In Table 6.5 we compare these state-of-the-art latent subspace detection methods (DIP, DIP-CC, and SA) with the MMD-PCA-based method which we used in our TTM framework. As one can note, some of these methods outperform MMD-based subspace projection at the cost of a higher computational complexity. All these subspace detection methods could replace the first step of our pipeline and potentially improve the final classification performance. However, given that MMD is the step with the highest asymptotic cost of our pipeline (see Sect. 6.3), we advocate that it is important to use the simplest unsupervised subspace transformation method and focus on the transductive part of the algorithm to improve performance.

6.4.1 Using Stronger Features (DeCAF)

Following the same experimental setting, we present further results for OC10 dataset. The previous sections show the results obtained using the original standard feature extraction method for these datasets (bags of SURF features). Owing to the success

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

of deep CNN methods, a newer feature extraction method has become standard, known as Deep Convolutional Activation Features (DeCAF) [127]. State-of-the-art method [275] Following [275], we used the output from the sixth layer as the visual features, leading to 4,096-dim DeCAF6 features. In this set of experiments we compare our TTM and ATTM methods with the methods that aim to solve the DA task by adapting the classifiers hyperplanes or by means of auxiliary classifiers, namely; the Adaptive SVM (A-SVM) [555], Domain Adaptation Machine (DAM) [135] and DA-M2S [74].

In [135], the author proposed a multiple source domain adaptation method referred to as DAM by leveraging a set of pre-learned classifiers independently trained with the labeled patterns from multiple source domains. More specifically, DAM was introduced as a data dependent regulator constrained by Least-Squares SVM (LS-SVM), which forces the target classifier to share similar decision values with auxiliary classifiers from the relevant source domains on the unlabeled patterns of the target domain. For a single source domain scenario, the experiments were repeated 10 times by using randomly generated subsets of source and target domain samples and the mean performance is reported in Table 6.6.

The DA-M2s method of [74] is an extension of the DAM method where from each RGB image data two nonlinear features are extracted, one describing the depth information and the other containing visual information. Using the Kernel Canonical Correlation Analysis (KCCA), the correlation between these two types of features is maximized. For the OC10 dataset (which have no depth maps), the method DA-M2s w/o depth represents source and target domains as two views of the same object classes. DA-M2s and LS-SVM are built on top of adaptive SVM (SVM-A) [555], which is a general framework to adapt one or more existing classifiers of any type to a new target dataset.

Note that in Table 6.6 the baseline without any transformation using the DeCAF features and NN classifier is significantly better than the results of Table 6.4, simply because the DeCAF features are better than SURF. As one can see our TTM and ATTM methods both outperform the other state-of-the-art approaches in most of the cases gaining 2.9 and 5.96% average performance enhancements over the best performing state-of-the-art method of DA-M2S (w/o depth), respectively. One should note that in both state-of-the-art approaches, DAM [135] and DA-M2S [74], the model has access to a small number of labeled samples from the target domain while our model does not benefit from that.

Sensitivity of TransGrad parameters. To evaluate sensitivity of TransGrad parameters, we ran TTM varying values of the regulator γ'' of the TransGrad step Eq. (6.7), and the results are in Fig. 6.3a. For all datasets, the performance improved as γ'' grows but it plateau for $\gamma'' \geq 5$. For this reason we used $\gamma'' = 5$ in all experiments of this chapter.

UNCORRECTED PROOF

420546_1_En_6_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

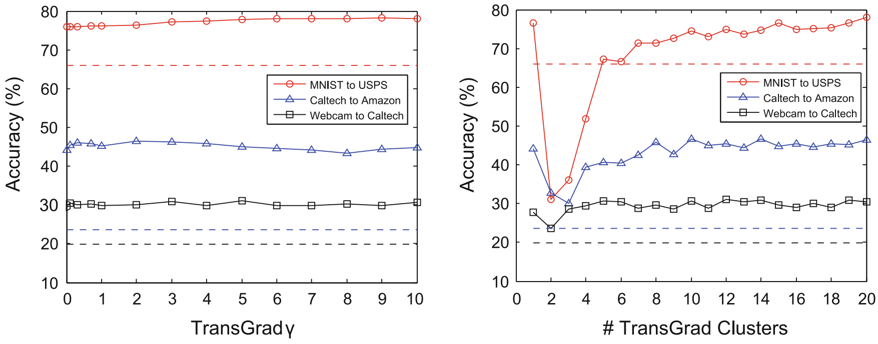
Table 6.6 Results obtained on the OC10 dataset using DeCAF features. The Baseline, JDA and TTM columns show the results achieved using the 1-NN classifier

Transfer task	Baseline DeCAF	SVM-A [555]	DAM [135]	DA-M2S [74]	JDA [312]	TTM (1NN)	ATTM
C → A	85.70	83.54	84.73	84.27	89.77	89.98	**92.17**
C → W	66.10	81.72	82.48	82.87	83.73	86.78	**90.84**
C → D	74.52	74.58	78.14	75.83	86.62	89.17	**92.99**
A → C	70.35	74.36	76.60	78.11	82.28	83.70	**86.55**
A → W	64.97	70.58	74.32	71.04	78.64	**89.81**	89.15
A → D	57.29	96.56	93.82	**96.62**	80.25	81.36	90.45
W → C	60.37	85.37	**87.88**	86.38	83.53	80.41	83.44
W → A	62.53	96.71	96.31	**97.12**	90.19	88.52	92.27
W → D	98.73	78.14	81.27	77.60	100	100	**100**
D → C	52.09	91.00	**91.75**	91.37	85.13	82.90	82.28
D → A	62.73	76.61	79.39	78.14	91.44	90.81	**91.65**
D → W	89.15	83.89	84.59	83.31	98.98	98.98	**98.98**
Avg	70.33	83.95	84.06	84.97	87.55	87.87	**90.90**

We also ran TTM with varying number Gaussian components K in the TransGrad step for the target GMM. Theoretically as the number of GMM components increases the translations get more accurate and the performance becomes more stable. We plot the classification accuracy w.r.t. K in Fig. 6.3b. Note that for $K = 1$, TransGrad contributes to an improvement over the baseline, as it induces a global shift toward the target set. But in general, for values of K smaller than the number of classes, we do not actually expect TransGrad to help, as it will shift samples from different classes toward the same clusters. This explains why the performance increases with K for $K > 2$. Based on this result, we adopted $K = 20$ in all other experiments of this chapter.

Timing comparison. We have compared the execution time of our TTM algorithm against JDA [312] in the transfer task from the MNIST digits dataset to the USPS digits dataset. Both algorithms were implemented in Matlab and were evaluated on a Intel Core2 64bit, 3 GHz machine running Linux. We averaged the time measured over five experiments. The JDA algorithm took 21.38 ± 0.26 s and our full TTM framework took 4.42 ± 0.12 s, broken down as: 0.40 ± 0.01 seconds to find the appropriate shared space using the MMD, 1.90 ± 0.06 to perform the sample-wise marginal distribution adaptations using TransGrad and finally 2.42 ± 0.12 s to apply the iterative conditional distribution adaptations (TST). The time complexity obviously will grow for both AJDA and ATTM due to kernel computation of the KDA classifier.

UNCORRECTED PROOF



(a) TransGrad γ'' (b) TransGrad GMM components' size

Fig. 6.3 The effect of different γ'' values and number of GMM clusters in the TransGrad step of our framework on the final performance of the pipeline for three cross-domain experiments. Constant lines show the baseline accuracy for each experiment

6.5 Conclusion and Discussion

In this chapter, we introduced transductive transfer machines (TTM), which aim to adapt both the marginal and conditional distributions of the source samples so that they become more similar to those of target samples, leading to an improvement in the classification results in DA scenarios. The proposed TTM pipeline consists of the following steps: first, a global linear transformation is applied to both source and target domain samples, so that their expected values match. In the next step, a novel method applies a sample-based transformation to source samples. This leads to a finer adaptation of their marginal distribution, taking into account the likelihood of each source sample given the target PDF. Finally, we proposed to iteratively adapt the class-based posterior distribution of source samples using an efficient linear transformation whose complexity only depends on the number of features. In addition, we proposed the use of two unsupervised comparison measures, Global and Clusters Dissimilarities. The former is used both to automatically determine the number of iterations needed and also to select the pipeline's main learner model. The latter measure, Clusters Dissimilarity, is used for adjusting the classifier's parameters for the new target domain. Our approach was shown to outperform state-of-the-art methods on various datasets, with a lower computational cost.

Our work [153] was one of the first to show that although DeCAF features lead to a step change in both discrimination power and generalization of image descriptors, they actually do not "undo the domain bias," as argued in [127]. DeCAF features can in fact be improved by applying feature space transformation using DA methods, and our method (ATTM) delivers improvement in performance, outperforming all the methods published prior to [152].

UNCORRECTED PROOF

420546_1_En_6_Chapter ✓ TYPESET DISK LE ✓ CP Disp.:25/5/2017 Pages: 140 Layout: T1-Standard

Acknowledgements This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014307/2 and the MOD University Defence Research Collaboration in Signal Processing.

UNCORRECTED PROOF

References

1. Martín Abadi, Ashish Agarwal, Paul Barham, Eugeneo Brevd, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from www.tensorflow.org.
2. P-A Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008.
3. Ayan Acharya, Eduardo R. Hruschka, Joydeep Ghosh, and Sreangsu Acharyya. Transfer learning with cluster ensembles. In *ICML Workshop on Unsupervised and Transfer Learning (WUTL)*, 2012.
4. Ankur Agarwal and Bill Triggs. A local basis representation for estimating human pose from cluttered images. In *Asian Conference on Computer Vision (ACCV)*, 2006.
5. Pulkit Agrawal, Ross Girshick, and Jitendra Malik. Analyzing the performance of multilayer neural networks for object recognition. In *European Conference on Computer Vision (ECCV)*, 2014.
6. Julien Ah-Pine, Marco Bressan, Stéphane Clinchant, Gabriela Csurka, Yves Hoppenot, and Jean-Michel Renders. Crossing textual and visual content in different application scenarios. *Multimedia Tools and Applications*, 42(1):31–56, 2009.
7. Zeynep Akata, Florent Perronnin, Zaid Harchaoui, and Cordelia Schmid. Label-embedding for attribute-based classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
8. Zeynep Akata, Scott Reed, Daniel Walter, Honglak Lee, and Bernt Schiele. Evaluation of output embeddings for fine-grained image classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
9. Samir Al-Stouhi and Chandan K. Reddy. Adaptive boosting for transfer learning using dynamic updates. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, 2011.
10. Rahaf Aljundi, R'emi Emonet, Damien Muselet, and Marc Sebban. Landmarks-based kernelized subspace alignment for unsupervised domain adaptation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
11. Rahaf Aljundi and Tinne Tuytelaars. Lightweight unsupervised domain adaptation by convolutional filter reconstruction. In *ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2016.

UNCORRECTED PROOF

© Springer International Publishing AG 2017

G. Csurka (ed.), *Domain Adaptation in Computer Vision Applications*,
Advances in Computer Vision and Pattern Recognition,
DOI 10.1007/978-3-319-58347-1

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

12. Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191. American Mathematical Soc., 2007.
13. Rie Kubota Ando and Tong Zhang. A framework for learning predictive structures from multiple tasks and unlabeled data. *Journal of Machine Learning Research*, 6:1817–1853, 2005.
14. Christophe Andrieu, Nando de Freitas, Arnaud Doucet, and Michael I. Jordan. An introduction to mcmc for machine learning. *Machine Learning*, 50(1):5–43, 2003.
15. Relja Arandjelovic and Andrew Zisserman. All about VLAD. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
16. Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 33(5):898–916, 2011.
17. Andreas Argyriou, Theodoros Evgeniou, and Massimiliano Pontil. Convex multi-task feature learning. *Machine Learning*, 73(3):243–272, 2008.
18. Andrew Arnold, Ramesh Nallapati, and William Cohen. A comparative study of methods for transductive transfer learning. In *ICDM Workshop (ICDMW)*, 2007.
19. Mathieu Aubry, Daniel Maturana, Alexei Efros, Bryan Russell, and Josef Sivic. Seeing 3d chairs: exemplar part-based 2d-3d alignment using a large dataset of CAD models. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
20. Mathieu Aubry and Bryan C. Russell. Understanding deep features with computer-generated imagery. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
21. Yusuf Aytar and Andrew Zisserman. Tabula rasa: Model transfer for object category detection. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
22. Jimmy Ba and Rich Caruana. Do deep nets really need to be deep? In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
23. Artem Babenko, Anton Slesarev, Alexandr Chigorin, and Victor S. Lempitsky. Neural codes for image retrieval. In *European Conference on Computer Vision (ECCV)*, 2014.
24. Vijay Badrinarayanan, Alex Handa, and Roberto Cipolla. SegNet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. *CoRR*, arXiv:1505.07293, 2015.
25. Mahsa Baktashmotlagh, Mehrtash Harandi, Brian Lovell, and Mathieu Salzmann. Unsupervised domain adaptation by domain invariant projection. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
26. Mahsa Baktashmotlagh, Mehrtash Harandi, Brian Lovell, and Mathieu Salzmann. Domain adaptation on the statistical manifold. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
27. Mahsa Baktashmotlagh, Mehrtash Harandi, and Mathieu Salzmann. Distribution-matching embedding for visual domain adaptation. *Journal of Machine Learning Research*, 2:1–30, 2016.
28. Evgeniy Bart and Shimon Ullman. Cross-generalization: Learning novel classes from a single example by feature replacement. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.
29. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *European Conference on Computer Vision (ECCV)*, 2006.
30. Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine Learning*, 20(3):151–175, 2010.
31. Shai Ben-David, John Blitzer, Koby Crammer, and Fernando Pereira. Analysis of representations for domain adaptation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2006.
32. Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 35(8):1798–1828, 2013.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

33. Tamara L. Berg, Alexander C. Berg, and Jonathan Shih. Automatic attribute discovery and characterization from noisy web data. In *European Conference on Computer Vision (ECCV)*, 2010.
34. Tamara L. Berg, Alexander Sorokin, Gang Wang, David A. Forsyth, Derek Hoiem, Ian Endres, and Ali Farhadi. It's all about the data. *Proceedings of the IEEE*, 98(8):1434–1452, 2010.
35. Thomas Berg and Peter Belhumeur. Poof: Part-based one-vs.-one features for fine-grained categorization, face verification, and attribute estimation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
36. Alessandro Bergamo and Lorenzo Torresani. Exploiting weakly-labeled web images to improve object classification: a domain adaptation approach. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2010.
37. Stanely Bileschi. CBCL StreetScenes challenge framework, 2007.
38. Jeff A. Bilmes. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. Technical report, International Computer Science Institute, 1998.
39. Arijit Biswas and Devi Parikh. Simultaneous active learning of classifiers & attributes via relative feedback. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
40. John Blitzer, Mark Dredze, and Fernando Pereira. Biographies, bolly-wood, boomboxes and blenders: Domain adaptation for sentiment classification. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2007.
41. John Blitzer, Dean P. Foster, and Sham M. Kakade. Zero-shot domain adaptation: A multi-view approach. Technical report, University of California, Berkeley, 2009.
42. John Blitzer, Sham Kakade, and Dean P. Foster. Domain adaptation with coupled subspaces. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2011.
43. John Blitzer, Ryan McDonald, and Fernando Pereira. Domain adaptation with structural correspondence learning. In *International Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2006.
44. Erik Bochinski, Volker Eiselein, and Thomas Sikora. Training a convolutional neural network for multi-class object detection using solely virtualworld data. In *IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS)*, 2016.
45. Karsten M. Borgwardt, Arthur Gretton, Malte J. Rasch, Hans-Peter Kriegel, Bernhard Schölkopf, and Alex J. Smola. Integrating structured biological data by kernel maximum mean discrepancy. *Bioinformatics*, 22:49–57, 2006.
46. Léon Bottou. *Online Algorithms and Stochastic Approximations*. Cambridge University Press, 1998.
47. Lubomir Bourdev, Subhransu Maji, and Jitendra Malik. Describing people: A poselet-based approach to attribute classification. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
48. Konstantinos Bousmalis, Nathan Silberman, David Dohan, Dumitru Erhan, and Dilip Krishnan. Unsupervised pixel-level domain adaptation with generative adversarial networks. *CoRR*, arXiv:1612.05424, 2016.
49. Konstantinos Bousmalis, George Trigeorgis, Nathan Silberman, Dumitru Erhan, and Dilip Krishnan. Domain separation networks. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2016.
50. Steve Branson, Catherine Wah, Florian Schroff, Boris Babenko, Peter Welinder, Pietro Perona, and Serge Belongie. Visual recognition with humans in the loop. In *European Conference on Computer Vision (ECCV)*, 2010.
51. Michael D. Breitenstein, Fabian Reichlin, Esther Koller-Meier, Bastien Leibe, and Luc Van Gool. Online multi-person tracking-by-detection from a single, uncalibrated camera. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 31(9):1820–1833, 2011.
52. Jane Bromley, James W Bentz, Léon Bottou, Isabelle Guyon, Yann LeCun, Cliff Moore, Eduard Säckinger, and Roopak Shah. Signature verification using a "siamese" time delay neural network. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(04):669–688, 1993.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

53. Gabriel J. Brostow, Julie Fauqueur, and Roberto Cipolla. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters*, 30(2):88–89, 2009.
54. Gabriel J. Brostow, Jamie Shotton, and Roberto Cipolla. Segmentation and recognition using structure from motion point clouds. In *European Conference on Computer Vision (ECCV)*, 2008.
55. Lorenzo Bruzzone and Mattia Marconcini. Domain adaptation problems: A dasvm classification technique and a circular validation strategy. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 32:770–787, 2010.
56. Daniel J. Butler, Jonas Wulff, Garrett B. Stanley, and Michael J. Black. A naturalistic open source movie for optical flow evaluation. In *European Conference on Computer Vision (ECCV)*, 2012.
57. Deng Cai, Xiaofei He, and Jiawei Han. Efficient kernel discriminant analysis via spectral regression. In *IEEE International Conference on Data Mining (ICDM)*, 2007.
58. Jian-Feng Cai, Emmanuel J. Candés, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *Journal on Optimization*, 20(4):1956–1982, 2010.
59. Guanqun Cao, Alexandros Iosifidis, Ke Chen, and Moncef Gabbouj. Generalized multi-view embedding for visual recognition and cross-modal retrieval. *CoRR*, arXiv:1605.09696, 2016.
60. Kevin M. Carter. *Dimensionality reduction on statistical manifolds*. ProQuest, 2009.
61. Rich Caruana. Multitask learning: A knowledge-based source of inductive bias. *Machine Learning*, 28:41–75, 1997.
62. Rui Caseiro, Joao F. Henriques, Pedro Martins, and Jorge Batista. Beyond the shortest path: Unsupervised domain adaptation by sampling subspaces along the spline flow. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
63. Lluís Castrejón, Yusuf Aytar, Carl Vondrick, Hamed Pirsiavash, and Antonio Torralba. Learning aligned cross-modal representations from weakly aligned data. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
64. Yee Seng Chan and Hwee Tou Ng. Domain adaptation with active learning for word sense disambiguation. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2007.
65. Olivier Chapelle, Bernhard Schlkopf, and Alexander Zien. *Semi-supervised learning*. MIT Press, 2006.
66. Ken Chatfield, Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. Return of the devil in the details: Delving deep into convolutional nets. In *BMVA British Machine Vision Conference (BMVC)*, 2014.
67. Ratthachat Chatpatanasiri, Teesid Korsrilabutr, Pasakorn Tangchanachaianan, and Boonserm Kijsirikul. A new kernelization framework for mahalanobis distance learning algorithms. *Neurocomputing*, 73(10):1570–1579, 2010.
68. Rita Chattopadhyay, Jieping Ye, Sethuraman Panchanathan, Wei Fan, and Ian Davidson. Multi-source domain adaptation and its application to early detection of fatigue. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2011.
69. Kamalika Chaudhuri, Sham M. Kakade, Karen Livescu, and Karthik Sridharan Sridharan. Multi-view clustering via canonical correlation analysis. In *International Conference on Machine Learning (ICML)*, 2009.
70. Chao-Yeh Chen and Kristen Grauman. Inferring analogous attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
71. Chenyi Chen, Ari Seff, Alain L. Kornhauser, and Jianxiong Xiao. DeepDriving: Learning affordance for direct perception in autonomous driving. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
72. Huizhong Chen, Andrew Gallagher, and Bernd Girod. Describing clothing by semantic attributes. In *European Conference on Computer Vision (ECCV)*, 2012.
73. Liang-Chieh Chen, Sanja Fidler, and Raquel Yuille, Alan L. Urtasun. Beat the MTurkers: Automatic image labeling from weak 3D supervision. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

UNCORRECTED PROOF

74. Lin Chen, Wen Li, and Dong Xu. Recognizing rgb images by learning from rgb-d data. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
75. Lin Chen, Qiang Zhang, and Baoxin Li. Predicting multiple attributes via relative multi-task learning. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
76. Minmin Chen, Kilian Q. Weinberger, and John Blitzer. Co-training for domain adaptation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2011.
77. Minmin Chen, Zhixiang Xu, Kilian Q. Weinberger, and Fei Sha. Marginalized denoising autoencoders for domain adaptation. In *International Conference on Machine Learning (ICML)*, 2012.
78. Ning Chen, Jun Zhu, Jianfei Chen, and Bo Zhang. Dropout training for support vector machines. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2014.
79. Qiang Chen, Junshi Huang, Rogerio Feris, Lisa M Brown, Jian Dong, and Shuicheng Yan. Deep domain adaptation for describing people based on fine-grained clothing attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
80. Wei-Yu Chen, Tzu-Ming Harry Hsu, and Yao-Hung Hubert Tsai. Transfer neural trees for heterogeneous domain adaptation. In *European Conference on Computer Vision (ECCV)*, 2016.
81. Xianjie Chen, Roozbeh Mottaghi, Xiaobai Liu, Sanja Fidler, Raquel Urtasun, and Alan Yuille. Detect what you can: Detecting and representing objects using holistic models and body parts. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
82. Xinlei Chen and Abhinav Gupta. Webly supervised learning of convolutional networks. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
83. Yisong Chen, Guoping Wang, and Shihai Dong. Learning with progressive transductive support vector machine. *Pattern Recognition Letters*, 24(12):845–855, 2003.
84. Haibin Cheng, Pang-Ning Tan, and Rong Jin. Localized support vector machine and its efficient algorithm. In *SIAM International Conference on Data Mining (SDM)*, 2007.
85. Boris Chidlovskii, Stéphane Clinchant, and Gabriela Csurka. Domain adaptation in the absence of source domain data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, 2016.
86. Boris Chidlovskii, Gabriela Csurka, and Shalini Gangwar. Assembling heterogeneous domain adaptation methods for image classification. In *CLEF online Working Notes*, 2014.
87. Sun-Wook Choi, Chong Ho Lee, and In Kyu Park. Scene classification via hypergraph-based semantic attributes subnetworks identification. In *European Conference on Computer Vision (ECCV)*, 2014.
88. Sumit Chopra, Suhrid Balakrishnan, and Raghuraman Gopalan. DLID: Deep learning for domain adaptation by interpolating between domains. In *ICML Workshop on Challenges in Representation Learning (WREPL)*, 2013.
89. Sumit Chopra, Raia Hadsell, and Yann LeCun. Learning a similarity metric discriminatively, with application to face verification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.
90. Brian Chu, Vashisht Madhavan, Oscar Beijbom, Judy Hoffman, and Trevor Darrell. Best practices for fine-tuning visual classifiers to new domains. In *ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2016.
91. Wen-Sheng Chu, Fernando De la Torre, and Jeffery F. Cohn. Selective transfer machine for personalized facial action unit detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
92. Mircea Cimpoi, Subhransu Maji, and Andrea Vedaldi. Deep filter banks for texture recognition and segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
93. Dan Cireşan, Ueli Meier, Jonathan Masci, and Jürgen Schmidhuber. Multi-column deep neural network for traffic sign classification. *Neural Networks*, 32:333–338, 2012.
94. Stéphane Clinchant, Gabriela Csurka, and Boris Chidlovskii. Transductive adaptation of black box predictions. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2016.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

95. David Cohn, Les Atlas, and Richard Ladner. Improving generalization with active learning. *Machine Learning*, 15(2):201–221, 1994.
96. Brendan Collins, Jia Deng, Kai Li, and Li Fei-Fei. Towards scalable dataset construction: An active learning approach. In *European Conference on Computer Vision (ECCV)*, 2008.
97. Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The Cityscapes dataset for semantic urban scene understanding. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
98. Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Scharwächter, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset. In *CVPR Workshop on The Future of Datasets in Vision (FCV)*, 2015.
99. C. Cortes, M. Mohri, M. Riley, and A. Rostamizadeh. Sample selection bias correction theory. In *Proceedings of the 19th international conference on Algorithmic Learning Theory*, ALT'08, pages 38–53, Berlin, Heidelberg, 2008. Springer-Verlag.
100. Corinna Cortes, Mehryar Mohri, and Afshin Rostamizadeh. Algorithms for learning kernels based on centered alignment. *Journal of Machine Learning Research*, 13(1):795–828, 2012.
101. Nicolas Courty, Rémi Flamary, Devis Tuia, and Alain Rakotomamonjy. Optimal transport for domain adaptation. *CoRR*, arXiv:1507.00504, 2015.
102. Elliot J. Crowley and Andrew Zisserman. In search of art. In *ECCV Workshop on Computer Vision for Art Analysis*, 2014.
103. Elliot J. Crowley and Andrew Zisserman. The state of the art: Object retrieval in paintings using discriminative regions. In *BMVA British Machine Vision Conference (BMVC)*, 2014.
104. Elliot J. Crowley and Andrew Zisserman. The art of detection. In *ECCV Workshop on Computer Vision for Art Analysis, (CVAA)*, 2016.
105. Gabriela Csurka, Boris Chidlovskii, and Stéphane Clinchant. Adapted domain specific class means. In *ICCV workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2015.
106. Gabriela Csurka, Boris Chidlovskii, Stéphane Clinchant, and Sophia Michel. Unsupervised domain adaptation with regularized domain instance denoising. In *ECCV workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2016.
107. Gabriela Csurka, Boris Chidlovskii, and Florent Perronnin. Domain adaptation with a domain specific class means classifier. In *ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2014.
108. Gabriela Csurka, Christopher Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray. Visual categorization with bags of keypoints. In *ECCV Workshop on Statistical learning in computer vision (SLCV)*, 2004.
109. Gabriela Csurka, Diane Larlus, Albert Gordo, and Jon Almazan. What is the right way to represent document images? *CoRR*, arXiv:1603.01076, 2016.
110. Yan Le Cun, B. Boser, J. S. Denker, R. E. Howard, W. Habbard, L. D. Jackel, and D. Henderson. Handwritten digit recognition with a back-propagation network. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 1990.
111. Wenyuan Dai, Yuqiang Chen, Gui-rong Xue, Qiang Yang, and Yong Yu. Translated learning: Transfer learning across different feature spaces. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2008.
112. Wenyuan Dai, Qiang Yang, Giu-Rong Xue, and Yong Yu. Boosting for transfer learning. In *International Conference on Machine Learning (ICML)*, 2007.
113. Wenyuan Dai, Qiang Yang, Gui-Rong Xue, and Yong Yu. Self-taught clustering. In *International Conference on Machine Learning (ICML)*, 2008.
114. Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.
115. Hal Daumé III. Frustratingly easy domain adaptation. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2007.
116. Hal Daumé III. Frustratingly easy domain adaptation. *CoRR*, arXiv:0907.1815, 2009.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

117. Hal Daumé III, Abhishek Kumar, and Avishek Saha. Co-regularization based semi-supervised domain adaptation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2010.
118. Hal Daumé III and Daniel Marcu. Domain adaptation for statistical classifiers. *Journal of Artificial Intelligence Research*, 26(1):101–126, 2006.
119. Jason V. Davis, Brian Kulis, Prateek Jain, Suvrit Sra, and Inderjit S. Dhillon. Information-theoretic metric learning. In *International Conference on Machine Learning (ICML)*, 2007.
120. Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
121. Thomas Deselaers, Bogdan Alexe, and Vittorio Ferrari. Localizing objects while learning their appearance. In *European Conference on Computer Vision (ECCV)*, 2010.
122. Chris Ding, Tao Li, Wei Peng, and Haesun Park. Orthogonal nonnegative matrix tri-factorizations for clustering. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2005.
123. Santosh Divvala, Ali Farhadi, and Carlos Guestrin. Learning everything about anything: Webly-supervised visual concept learning. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
124. Piotr Dollár, Christian Wojek, Bernt Schiele, and Pietro Perona. Pedestrian detection: a benchmark. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
125. Piotr Dollár, Christian Wojek, Bernt Schiele, and Pietro Perona. Pedestrian detection: an evaluation of the state of the art. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 34(4):743–761, 2012.
126. Jeff Donahue, Judy Hoffman, Erik Rodner, Kate Saenko, and Trevor Darrell. Semi-supervised domain adaptation with instance constraints. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
127. Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. *CoRR*, arXiv:1310.1531, 2013.
128. Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. In *International Conference on Machine Learning (ICML)*, 2014.
129. David L. Donoho. Compressed sensing. *Transactions on Information Theory*, 52:1289–1306, 2006.
130. Mark Dredze and Koby Crammer. Online methods for multi-domain learning and adaptation. In *International Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2008.
131. Mark Dredze, Alex Kulesza, and Koby Crammer. Multi-domain learning by confidence-weighted parameter combination. *Machine Learning*, 79(1):123–149, 2010.
132. Alain Droniou and Olivier Sigaud. Gated autoencoders with tied input weights. In *International Conference on Machine Learning (ICML)*, 2013.
133. Kun Duan, Devi Parikh, David Crandall, and Kristen Grauman. Discovering localized attributes for fine-grained recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
134. Lixin Duan, Ivor W. Tsang, and Dong Xu. Domain transfer multiple kernel learning. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 34(3):465–479, 2012.
135. Lixin Duan, Ivor W. Tsang, Dong Xu, and Tat-Seng Chua. Domain adaptation from multiple sources via auxiliary classifiers. In *International Conference on Machine Learning (ICML)*, 2009.
136. Lixin Duan, Ivor W. Tsang, Dong Xu, and Steve J. Maybank. Domain transfer SVM for video concept detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
137. Lixin Duan, Dong Xu, and Shih-Fu Chang. Exploiting web images for event recognition in consumer videos: A multiple source domain adaptation approach. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

138. Lixin Duan, Dong Xu, and Ivor W. Tsang. Domain adaptation from multiple sources: A domain-dependent regularization approach. *Transactions on Neural Networks and Learning Systems*, 23(3):504–518, 2012.
139. Lixin Duan, Dong Xu, and Ivor W Tsang. Learning with augmented features for heterogeneous domain adaptation. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 36(6):1134–1148, 2012.
140. Lixin Duan, Dong Xu, Ivor W Tsang, and Jiebo Luo. Visual event recognition in videos by learning from web data. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 34(9):1667–1680, 2012.
141. John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. Technical report, EECS Department, University of California, Berkeley, 2010.
142. Miroslav Dudík, Robert E. Schapire, and Steven J. Phillips. Correcting sample selection bias in maximum entropy density estimation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2005.
143. Alan Edelman, Tomás A. Arias, and Steven T. Smith. The geometry of algorithms with orthogonality constraints. *Journal of Matrix Analysis and Applications*, 20(2):303–353, 1998.
144. David Eigen, Christian Puhrsch, and Rob Fergus. Depth map prediction from a single image using a multi-scale deep network. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
145. Ian Endres, Vivek Srikumar, Ming-Wei Chang, and Derek Hoiem. Learning shared body plans. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
146. Markus Enzweiler and Dariu M. Gavrila. Monocular pedestrian detection: Survey and experiments. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 31(12):2179–2195, 2009.
147. Victor Escorcia, Juan Carlos Niebles, and Bernard Ghanem. On the relationship between visual attributes and convolutional networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
148. Marc Everingham, Luc Van Gool, Chris Williams, John Winn, and Andrew. Zisserman. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88(2):303–338, 2010.
149. Theodoros Evgeniou and Massimiliano Pontil. Regularized multi-task learning. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2004.
150. Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9, 2008.
151. Chen Fang, Ye Xu, and Daniel N. Rockmore. Unbiased metric learning: On the utilization of multiple datasets and web images for softening bias. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
152. Nazli FarajiDavar, Teofilo deCampos, and Josef Kittler. Adaptive transductive transfer machines. In *BMVA British Machine Vision Conference (BMVC)*, 2014.
153. Nazli FarajiDavar, Teofilo deCampos, and Josef Kittler. Transductive transfer machines. In *Asian Conference on Computer Vision (ACCV)*, 2014.
154. Nazli FarajiDavar, Teofilo deCampos, Josef Kittler, and Fei Yang. Transductive transfer learning for action recognition in tennis games. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
155. Nazli FarajiDavar, Teofilo deCampos, David Windridge, Josef Kittler, and William Christmas. Domain adaptation in the context of sport video action recognition. In *BMVA British Machine Vision Conference (BMVC)*, 2012.
156. Ali Farhadi, Ian Endres, and Derek Hoiem. Attribute-centric recognition for cross-category generalization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
157. Ali Farhadi, Ian Endres, Derek Hoiem, and David Forsyth. Describing objects by their attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

UNCORRECTED PROOF

158. Manaal Faruqui and Chris Dyer. Improving vector space word representations using multilingual correlation. In *Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2014.
159. Li Fei-Fei, Rob Fergus, and Pietro Perona. One-shot learning of object categories. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 28(4):594–611, 2006.
160. Li Fei-Fei, Rob Fergus, and Pietro Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. *Computer Vision and Image Understanding*, 106(1):57–70, 2007.
161. Christiane Fellbaum. *WordNet: An Electronic Lexical Database*. Bradford Books, 1998.
162. Pedro F Felzenszwalb, Ross B Girshick, David McAllester, and Deva Ramanan. Object detection with discriminatively trained part-based models. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 32(9):1627–1645, 2010.
163. Robert Fergus, Li Fei-Fei, Pietro Perona, and Andrew Zisserman. Learning object categories from google's image search. In *IEEE International Conference on Computer Vision (ICCV)*, 2005.
164. Basura Fernando, Amaury Habrard, Marc Sebban, and Tinne Tuytelaars. Unsupervised visual domain adaptation using subspace alignment. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
165. Basura Fernando, Amaury Habrard, Marc Sebban, and Tinne Tuytelaars. Subspace alignment for domain adaptation. *CoRR*, arXiv:1409.5241, 2014.
166. Vittorio Ferrari and Andrew Zisserman. Learning visual attributes. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2007.
167. Michael Fink. Object classification from a single example utilizing class relevance pseudo-metrics. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2004.
168. Yoav Freund and Robert Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.
169. Andrea Frome, Greg S. Corrado, Jon Shlens, Samy Bengio, Jeff Dean, Marc'Aurelio Ranzato, and Tomas Mikolov. Devise: A deep visual-semantic embedding model. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2013.
170. Yanwei Fu, Timothy M. Hospedales, Tao Xiang, Zhengyong Fu, and Shaogang Gong. Transductive multi-view embedding for zero-shot recognition and annotation. In *European Conference on Computer Vision (ECCV)*, 2014.
171. Yanwei Fu, Timothy M. Hospedales, Tao Xiang, and Shaogang Gong. Learning multi-modal latent attributes. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 36(2):303–316, 2014.
172. Yanwei Fu, Timothy M. Hospedales, Tao Xiang, and Shaogang Gong. Learning multi-modal latent attributes. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 36(2):303–316, 2014.
173. Zhenyong Fu, Tao Xiang, Elyor Kodirov, and Shaogang Gong. Zero-shot object recognition by semantic manifold distance. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
174. Adrien Gaidon and Eleonora Vig. Online domain adaptation for multi-object tracking. In *BMVA British Machine Vision Conference (BMVC)*, 2015.
175. Adrien Gaidon, Qiao Wang, Yohann Cabon, and Eleonora Vig. Virtual worlds as proxy for multi-object tracking analysis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
176. Adrien Gaidon, Gloria Zen, and José A. Rodriguez-Serrano. Self-learning camera: Autonomous adaptation of object detectors to unlabeled video streams. *CoRR*, arXiv:1406.4296, 2014.
177. Chuang Gan, Ming Lin, Yi Yang, Yueting Zhuang, and Alexander G. Hauptmann. Exploring semantic inter-class relationships (SIR) for zero-shot action recognition. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2015.
178. Chuang Gan, Chen Sun, Lixin Duan, and Boqing Gong. Webly-supervised video recognition by mutually voting for relevant web images and web video frames. In *European Conference on Computer Vision (ECCV)*, 2016.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

179. Chuang Gan, Tianbao Yang, and Boqing Gong. Learning attributes equals multi-source domain generalization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
180. Chuang Gan, Yi Yang, Linchao Zhu, Deli Zhao, and Yueting Zhuang. Recognizing an action using its name: A knowledge-based approach. *International Journal of Computer Vision*, pages 1–17, 2016.
181. Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. *CoRR*, arXiv:1409.7495, 2014.
182. Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *International Conference on Machine Learning (ICML)*, 2015.
183. Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor S. Lempitsky. Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 2016.
184. Jean-Luc Gauvain and Chin-Hui Lee. Maximum a posteriori estimation for multivariate gaussian mixture observations of markov chain. *Transactions on Speech and Audio Processing*, 2(2):291–298, 1994.
185. Liang Ge, Jing Gao, Hung Ngo, Kang Li, and Aidong Zhang. On handling negative transfer and imbalanced distributions in multiple source transfer learning. In *SIAM International Conference on Data Mining (SDM)*, 2013.
186. Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The KITTI dataset. *International Journal of Robotics Research*, 32:1231–1237, 2013.
187. Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
188. Pascal Germain, Amaury Habrard, François Laviolette, and Emilie Morvant. A PAC-Bayesian approach for domain adaptation with specialization to linear classifiers. In *International Conference on Machine Learning (ICML)*, 2013.
189. Muhammad Ghifary, W. Bastiaan Kleijn, and Mengjie Zhang. Domain adaptive neural networks for object recognition. *CoRR*, arXiv:1409.6041, 2014.
190. Muhammad Ghifary, W. Bastiaan Kleijn, Mengjie Zhang, and David Balduzzi. Domain generalization for object recognition with multi-task autoencoders. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
191. Muhammad Ghifary, W Bastiaan Kleijn, Mengjie Zhang, and David Balduzzi. Deep reconstruction-classification networks for unsupervised domain adaptation. In *European Conference on Computer Vision (ECCV)*, 2016.
192. Ross Girshick, Jeff Donahue, Trevor Darrell, and Jagannath Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
193. Ross Girshick, Forrest Iandola, Trevor Darrell, and Jitendra Malik. Deformable part models are convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
194. Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Domain adaptation for large-scale sentiment classification: A deep learning approach. In *International Conference on Machine Learning (ICML)*, 2011.
195. Daniel Goehring, Judy Hoffman, Erik Rodner, Kate Saenko, and Trevor Darrell. Interactive adaptation of real-time object detectors. In *International Conference on Robotics and Automation (ICRA)*, 2014.
196. Boqing Gong, Kristen Grauman, and Fei Sha. Connecting the dots with landmarks: Discriminatively learning domain invariant features for unsupervised domain adaptation. In *International Conference on Machine Learning (ICML)*, 2013.
197. Boqing Gong, Kristen Grauman, and Fei Sha. Reshaping visual datasets for domain adaptation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2013.
198. Boqing Gong, Kristen Grauman, and Fei Sha. Learning kernels for unsupervised domain adaptation with applications to visual object recognition. *International Journal of Computer Vision*, 109(1):3–27, 2014.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

199. Boqing Gong, Jianzhuang Liu, Xiaogang Wang, and Xiaoou Tang. Learning semantic signatures for 3d object retrieval. *Transactions on Multimedia*, 15(2):369–377, 2013.
200. Boqing Gong, Yuan Shi, Fei Sha, and Kristen Grauman. Geodesic flow kernel for unsupervised domain adaptation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
201. Shaogang Gong, Marco Cristani, Shuicheng Yan, and Chen Change Loy. *Person re-identification*. Springer, 2014.
202. Yunchao Gong, Qifa Ke, Michael Isard, and Svetlana Lazebnik. A multi-view embedding space for modeling internet images, tags, and their semantics. *International Journal of Computer Vision*, 106(2):210–233, 2014.
203. Yunchao Gong, Liwei Wang, Ruiqi Guo, and Svetlana Lazebnik. Multi-scale orderless pooling of deep convolutional activation features. In *European Conference on Computer Vision (ECCV)*, 2014.
204. Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
205. Raghuraman Gopalan. Learning cross-domain information transfer for location recognition and clustering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
206. Raghuraman Gopalan, Ruonan Li, and Rama Chellappa. Domain adaptation for object recognition: An unsupervised approach. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
207. Raghuraman Gopalan, Ruonan Li, and Rama Chellappa. Unsupervised adaptation across domain shifts by generating intermediate data representations. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 36(11), 2014.
208. Albert Gordo, Jon Almazán, Jerome Revaud, and Diane Larlus. Deep image retrieval: Learning global representations for image search. In *European Conference on Computer Vision (ECCV)*, 2016.
209. Philippe-Henri Gosselin, Naila Murray, Hervé Jégou, and Florent Perronnin. Revisiting the Fisher vector for fine-grained classification. *Pattern Recognition Letters*, 49(11):92–98, 2014.
210. Kristen Grauman, Gregory Shakhnarovich, and Trevor Darrell. Inferring 3D structure with a statistical image-based shape model. In *IEEE International Conference on Computer Vision (ICCV)*, 2003.
211. Doug Gray, Shane Brennan, and Hai Tao. Evaluating appearance models for recognition, reacquisition, and tracking. In *International Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, 2007.
212. Arthur Gretton, Karsten M Borgwardt, Malte J Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(1):723–773, 2012.
213. Arthur Gretton, Karsten M. Borgwardt, Malte J Rasch, Bernhard Schlkopf, and Alex J. Smola. A kernel method for the two sample problem. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2007.
214. Arthur Gretton, Alex Smola, Jiayuan Huang, Marcel Schmittfull, Karsten Borgwardt, and Bernhard Schölkopf. Covariate shift by kernel mean matching. In Joaquin Quiñonero Candela, Masashi Sugiyama, Anton Schwaighofer, and Neil D. Lawrence, editors, *Dataset Shift in Machine Learning*. The MIT Press, 2009.
215. Gregory Griffin, Alex Holub, and Pietro Perona. Caltech-256 object category dataset. Technical report, Californian Institute of Technologie, 2007.
216. Matthieu Guillaumin, Daniel Küttel, and Vittorio Ferrari. Imagenet auto-annotation with segmentation propagation. *International Journal of Computer Vision*, 110(3):328–348, 2014.
217. Ralf Haeusler and Daniel Kondermann. Synthesizing real world stereo challenges. In *German Conference on Pattern Recognition (GCPR)*, 2013.
218. Haltakov Haltakov, Christian Unger, and Slobodan Ilic. Framework for generation of synthetic ground truth data for driver assistance applications. In *German Conference on Pattern Recognition (GCPR)*, 2013.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

219. Jihun Ham, Daniel D Lee, Sebastian Mika, and Bernhard Schölkopf. A kernel view of the dimensionality reduction of manifolds. In *International Conference on Machine Learning (ICML)*, 2004.
220. David J. Hand. Classifier technology and the illusion of progress. *Statistical Science*, 21:1–15, 2006.
221. Ankur Handa, Viorica Patraucean, Vijay Badrinarayanan, Simon Stent, and Roberto Cipolla. Synthcam3d: Semantic understanding with synthetic indoor scenes. *CoRR*, arXiv:1505.00171, 2015.
222. Ankur Handa, Viorica Patraucean, Vijay Badrinarayanan, Simon Stent, and Roberto Cipolla. Understanding real world indoor scenes with synthetic data. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
223. David R. Hardoon, Sandor Szedmak, and John Shawe-Taylor. Canonical correlation analysis: An overview with application to learning methods. *Neurocomputing*, 16(12):2639–2664, 2004.
224. Maayan Harel and Shie Mannor. Learning from multiple outlooks. In *International Conference on Machine Learning (ICML)*, 2011.
225. Bharath Hariharan, Jitendra Malik, and Deva Ramanan. Discriminative decorrelation for clustering and classification. In *European Conference on Computer Vision (ECCV)*, 2012.
226. Adam W. Harley, Alex Ufkes, and Konstantinos G. Derpanis. Evaluation of deep convolutional nets for document image classification and retrieval. In *International Conference on Document Analysis and Recognition (ICDAR)*, 2015.
227. Richard Hartley, Jochen Trumpf, Yuchao Dai, and Hongdong Li. Rotation averaging. *International Journal of Computer Vision*, 103(3):267–305, 2013.
228. Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.
229. Hironori Hattori, Vishnu Naresh Boddeti, Kris M. Kitani, and Takeo Kanade. Learning scene-specific pedestrian detectors without real data. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
230. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, arXiv:1512.03385, 2015.
231. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
232. Geoffrey E. Hinton, Oriol Vinyals, and Jeff Dean. Distilling the knowledge in a neural network. In *NIPS Workshop on Deep Learning and Representation Learning*, 2014.
233. Martin Hirzer, Csaba Beleznai, Peter M. Roth, and Horst Bischof. Person re-identification by descriptive and discriminative classification. In *Scandinavian Conference (SCIA)*, 2011.
234. Frank Lauren Hitchcock. The expression of a tensor or a polyadic as a sum of products. *Journal of Mathematics and Physics*, 6(1):164–189, 1927.
235. Judy Hoffman, Trevor Darrell, and Kate Saenko. Continuous manifold based adaptation for evolving visual domains. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
236. Judy Hoffman, Sergio Guadarrama, Eric S. Tzeng, Ronghang Hu, Jeff Donahue, Ross Girshick, Trevor Darrell, and Kate Saenko. LSDA: Large scale detection through adaptation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
237. Judy Hoffman, Saurabh Gupta, and Trevor Darrell. Learning with side information through modality hallucination. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
238. Judy Hoffman, Brian Kulis, Trevor Darrell, and Kate Saenko. Discovering latent domains for multisource domain adaptation. In *European Conference on Computer Vision (ECCV)*, 2012.
239. Judy Hoffman, Erik Rodner, Jeff Donahue, Trevor Darrell, and Kate Saenko. Efficient learning of domain-invariant image representations. In *International Conference on Learning representations (ICLR)*, 2013.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

240. Judy Hoffman, Eric Tzeng, Jeff Donahue, Yangqing Jia, Kate Saenko, and Trevor Darrell. One-shot adaptation of supervised deep convolutional models. *CoRR*, arXiv:1312.6204, 2013.
241. Judy Hoffman, Eric Tzeng, Jeff Donahue, Yangqing Jia, Kate Saenko, and Trevor Darrell. One-shot adaptation of supervised deep convolutional models. In *International Conference on Learning representations (ICLR)*, 2014.
242. Alex Holub, Pietro Perona, and Michael C. Burl. Entropy-based active learning for object recognition. In *CVPR Workshop on Online Learning for Classification (OLC)*, 2008.
243. Jiayuan Huang, Alex Smola, Arthur Gretton, Karsten Borgwardt, and Bernhard Schölkopf. Correcting sample selection bias by unlabeled data. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2007.
244. Sheng Huang, Mohamed Elhoseiny, Ahmed Elgammal, and Dan Yang. Learning hypergraph-regularized attribute predictors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
245. Sung Ju Hwang, Fei Sha, and Kristen Grauman. Sharing features between objects and their attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
246. Sung Ju Hwang and Leonid Sigal. A unified semantic embedding: Relating taxonomies and attributes. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
247. Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning (ICML)*, 2015.
248. Vidit Jain and Eric Learned-Miller. Online domain adaptation of a pre-trained cascade of classifiers. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
249. Omar Javed, Saad Ali, and Mubarak Shah. Online detection and classification of moving objects using progressively improving detectors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.
250. Dinesh Jayaraman and Kristen Grauman. Zero-shot recognition with unreliable attributes. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
251. Dinesh Jayaraman, Fei Sha, and Kristen Grauman. Decorrelating semantic visual attributes by resisting the urge to share. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
252. I-Hong Jhuo, Dong Liu, D.T. Lee, and Shih.-Fu. Chang. Robust visual domain adaptation with low-rank reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
253. Shuiwang Ji and Jieping Ye. An accelerated gradient method for trace norm minimization. In *International Conference on Machine Learning (ICML)*, 2009.
254. Yangqing Jia, Mathieu Salzmann, and Trevor Darrell. Learning cross-modality similarity for multinomial data. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
255. Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. *CoRR*, arXiv:1408.5093, 2014.
256. Wei Jiang, Eric Zavesky, Shih-Fu Chang, and Alex Loui. Cross-domain learning methods for high-level visual concept classification. In *International Conference on Image Processing (ICIP)*, 2008.
257. Thorsten Joachims. Transductive inference for text classification using support vector machines. In *International Conference on Machine Learning (ICML)*, 1999.
258. Jungseock Joo, Shuo Wang, and Song-Chun Zhu. Human attribute recognition by rich appearance dictionary. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
259. Ajay J Joshi, Fatih Porikli, and Nikolaos Papanikolopoulos. Multi-class active learning for image classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
260. Toshihiro Kamishima, Masahiro Hamasaki, and Shotaro Akaho. Trbagg: A simple transfer learning method and its application to personalization in collaborative tagging. In *IEEE International Conference on Data Mining (ICDM)*, 2009.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

261. Takafumi Kanamori, Shohei Hido, and Masashi Sugiyama. Efficient direct density ratio estimation for non-stationarity adaptation and outlier detection. *Journal of Machine Learning Research*, 10:1391–1445, 2009.
262. Biliana Kaneva, Antonio Torralba, and William T. Freeman. Evaluating image features using a photorealistic virtual world. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2282–2289, 2011.
263. Pichai Kankuekul, Aram Kawewong, Sirinart Tangruamsub, and Osamu Hasegawa. Online incremental attribute-based zero-shot learning. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
264. Ashish Kapoor, Kristen Grauman, Raquel Urtasun, and Trevor Darrell. Active learning with gaussian processes for object categorization. In *IEEE International Conference on Computer Vision (ICCV)*, 2007.
265. Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. Large-scale video classification with convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
266. Robert E Kass and Paul W Vos. *Geometrical foundations of asymptotic inference*. Wiley. com, 2011.
267. Koray Kavukcuoglu, Marc'Aurelio Ranzato, and Yann LeCun. Fast inference in sparse coding algorithms with applications to object recognition. *CoRR*, arXiv:1010.3467, 2010.
268. Aditya Khosla, Tinghui Zhou, Tomasz Malisiewicz, Alexei A. Efros, and Antonio Torralba. Undoing the damage of dataset bias. In *European Conference on Computer Vision (ECCV)*, 2012.
269. Daniel Kifer, Shai Ben-David, and Johannes Gehrke. Detecting change in data streams. In *International Conference on Very large Data Bases (VLDB)*, 2004.
270. Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning representations (ICLR)*, 2015.
271. Brendan F. Klare, Serhat S. Bucak, Anil K. Jain, and Tayfun Akgul. Towards automated caricature recognition. In *International Conference on Biometrics (ICB)*, 2012.
272. Adriana Kovashka, Devi Parikh, and Kristen Grauman. Whittlesearch: Image search with relative attribute feedback. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
273. Adriana Kovashka, Sudheendra Vijayanarasimhan, and Kristen Grauman. Actively selecting annotations among objects and attributes. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
274. Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A Shamma, Michael Bernstein, and Li Fei-Fei. Visual genome: Connecting language and vision using crowdsourced dense image annotations. *CoRR*, arXiv:1602:07332, 2016.
275. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep Convolutional Neural Networks. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2012.
276. Anders Krogh and Jesper Vedelsby. Neural network ensembles, cross validation, and active learning. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 1995.
277. Roland Kuhn, Jean-Claude Junqua, Patrick Nguyenand, and Nancy Niedzielski. Rapid speaker adaptation in eigenvoice space. *Transactions on Speech and Audio Processing*, 8(6):695–707, 2000.
278. Brian Kulis, Kate Saenko, and Trevor Darrell. What you saw is not what you get: Domain adaptation using asymmetric kernel transforms. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
279. Abhishek Kumar and Hal Daumé III. Learning task grouping and overlap in multi-task learning. In *International Conference on Machine Learning (ICML)*, 2012.
280. Neeraj Kumar, Alexander C. Berg, Peter N. Belhumeur, and Shree K. Nayar. Attribute and simile classifiers for face verification. In *IEEE International Conference on Computer Vision (ICCV)*, 2009.

UNCORRECTED PROOF

281. Abhijit Kundu, Yin F. Li, Daellert, Fuxin Li, and James M. Rehg. Joint semantic segmentation and 3D reconstruction from monocular video. In *European Conference on Computer Vision (ECCV)*, 2014.
282. Daniel Küttel and Vittorio Ferrari. Figure-ground segmentation by transferring window masks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
283. Shrenik Lad and Devi Parikh. Interactively guiding semi-supervised clustering via attribute-based explanations. In *European Conference on Computer Vision (ECCV)*, 2014.
284. Kevin Lai, Liefeng Bo, Xiaofeng Ren, and Dieter Fox. A large-scale hierarchical multi-view rgb-d object dataset. In *International Conference on Robotics and Automation (ICRA)*, 2011.
285. Kevin Lai and Dieter Fox. 3D laser scan classification using web data and domain adaptation. In *Robotics: Science and Systems Conference (RSS)*, 2009.
286. Kevin Lai and Dieter Fox. Object recognition in 3D point clouds using web data and domain adaptation. *International Journal of Robotics Research*, 29(8):1019–1037, 2010.
287. Christoph H. Lampert, Hannes Nickisch, and Stefan Harmeling. Learning to detect unseen object classes by between-class attribute transfer. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
288. Gert Lanckriet, Nello Cristianini, Peter Bartlett, Laurent El Ghaoui, and Michael I. Jordan. Learning the kernel matrix with semidefinite programming. *Journal of Machine Learning Research*, 5:27–72, 2004.
289. Hugo Larochelle, Dumitru Erhan, and Yoshua Bengio. Zero-data learning of new tasks. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2008.
290. Ryan Layne, Timothy M. Hospedales, and Shaogang Gong. Re-id: Hunting attributes in the wild. In *BMVA British Machine Vision Conference (BMVC)*, 2014.
291. Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
292. Christopher J. Leggetter and Philip C. Woodland. Maximum likelihood linear regression for speaker adaptation of continuous density hidden markov models. *Computer Speech and Language*, 9(2):171–185, 1995.
293. Bastian Leibe and Bernt Schiele. Analyzing appearance and contour based methods for object categorization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003.
294. Anat Levin, Paul Viola, and Yoav Freund. Unsupervised improvement of visual detectors using co-training. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
295. Elizaveta Levina and Peter J. Bickel. Maximum likelihood estimation of intrinsic dimension. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2004.
296. Li-Jia Li, Hao Su, Li Fei-Fei, and Eric P Xing. Object bank: A high-level image representation for scene classification & semantic feature sparsification. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2010.
297. Ruonan Li and Todd Zickler. Discriminative virtual views for cross-view action recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
298. Wei Li and Xiaogang Wang. Locally aligned feature transforms across views. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
299. Wen Li, Lixin Duan, Dong Xu, and Iwor W. Tsang. Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 36(6):1134–1148, 2014.
300. Wenbin Li and Mario Fritz. Recognizing materials from virtual examples. In *European Conference on Computer Vision (ECCV)*, 2012.
301. Liang Liang and Kristen Grauman. Beyond comparing image pairs: Setwise active learning for relative attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
302. Xuejun Liao, Ya Xue, and Lawrence Carin. Logistic regression with an auxiliary data source. In *International Conference on Machine Learning (ICML)*, 2005.
303. Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C.Lawrence Zitnick. Microsoft COCO: Common objects in context. In *European Conference on Computer Vision (ECCV)*, 2014.

UNCORRECTED PROOF

304. Ming-Yu Liu and Oncel Tuzel. Coupled generative adversarial networks. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2016.
305. Xiuwen Liu, Anuj Srivastava, and Kyle Gallivan. Optimal linear representations of images for object recognition. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 26:662–666, 2004.
306. Joan M. Llargues, Juan Peralta, Raul Arrabales, Manuel González, Paulo Cortez, and Antonio M. López. Artificial intelligence approaches for the generation and assessment of believable human-like behaviour in virtual characters. *Expert Systems With Applications*, 41(16):7281–7290, 2014.
307. Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
308. Jonathan L. Long, Ning Zhang, and Trevor Darrell. Do convnets learn correspondence? In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
309. Mingsheng Long, Yue Cao, Jianmin Wang, and Michael I. Jordan. Learning transferable features with deep adaptation networks. In *International Conference on Machine Learning (ICML)*, 2015.
310. Mingsheng Long, Guiguang Ding, Jianmin Wang, Jiaguang Sun, Yuchen Guo, and Philip S. Yu. Transfer sparse coding for robust image representation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
311. Mingsheng Long, Jianmin Wang, Guiguang Ding, Sinno Jialin Pan, and Philip S. Yu. Adaptation regularization: a general framework for transfer learning. *Transactions on Knowledge and Data Engineering*, 5(26):1076–1089, 2014.
312. Mingsheng Long, Jianmin Wang, Guiguang Ding, Jiaguang Sun, and Philip S. Yu. Transfer feature learning with joint distribution adaptation. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
313. Mingsheng Long, Jianmin Wang, Guiguang Ding, Jiaguang Sun, and Philip S. Yu. Transfer joint matching for unsupervised domain adaptation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
314. Mingsheng Long, Jianmin Wang, and Michael I. Jordan. Deep transfer learning with joint adaptation networks. *CoRR*, arXiv:1605.06636, 2016.
315. David G Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
316. Ping Luo, Xiaogang Wang, and Xiaoou Tang. A deep sum-product architecture for robust facial attributes analysis. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
317. Andy Jinhua Ma, Jiawei Li, Pong C. Yuen, and Ping Li. Cross-domain person reidentification using domain adaptation ranking svms. *Transactions on Image Processing*, 24(5):1599–1613, 2015.
318. Bingpeng Ma, Yu Su, and Frédéric Jurie. Local descriptors encoded by Fisher vectors for person re-identification. In *ECCV Workshop on Re-Identification (Re-Id)*, 2012.
319. Laurens van der Maaten, Minmin Chen, Stephen Tyree, and Kilian Weinberger. Learning with marginalized corrupted features. In *International Conference on Machine Learning (ICML)*, 2013.
320. Dhruv Mahajan, Sundararajan Sellamanickam, and Vinod Nair. A joint learning framework for attribute models and object descriptions. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
321. Tomasz Malisiewicz, Abhinav Gupta, and Alexei A Efros. Ensemble of exemplar-svms for object detection and beyond. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
322. Yishay Mansour, Mehryar Mohri, and Afshin Rostamizadeh. Domain adaptation: Learning bounds and algorithms. In *Annual Conference on Learning Theory (COLT)*, 2009.
323. Yishay Mansour, Mehryar Mohri, and Afshin Rostamizadeh. Domain adaptation with multiple sources. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2009.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

324. Yishay Mansour, Mehryar Mohri, and Afshin Rostamizadeh. Multiple source adaptation and the Rényi divergence. In *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2009.
325. Javier Marín, David Vázquez, David Gerónimo, and Antonio M. López, López. Learning appearance in virtual scenarios for pedestrian detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
326. Francisco Massa, Bryan C. Russell, and Mathieu Aubry. Deep exemplar 2D-3D detection by adapting from real to rendered views. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
327. Giona Matasci, Michele Volpi, Mikhail Kanevski, Lorenzo Bruzzone, and Devis Tuia. Semi-supervised transfer component analysis for domain adaptation in remote sensing image classification. *Transactions on Geoscience and Remote Sensing*, 53(7):3550–3564, 2015.
328. Nikolaus Mayer, Eddy Ilg, Philip Hausser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
329. Stephan Meister and Daniel Kondermann. Real versus realistically rendered scenes for optical flow evaluation. In *ITG Conference on Electronic Media Technology (CEMT)*, 2011.
330. Roland Memisevic and Geoffrey E. Hinton. Unsupervised learning of image transformations. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.
331. Microsoft. Microsoft Research Cambridge Object Recognition Image Database. http://research.microsoft.com/en-us/downloads/b94de342-60dc-45d0-830b-9f6eff91b301/default.aspx, 2005.
332. Stephen Milborrow, John Morkel, and Fred Nicolls. The MUCT Landmarked Face Database. In *Annual Symposium of the Pattern Recognition Association of South Africa*, 2010. http://www.milbo.org/muct.
333. Erik G. Miller, Nicholas E. Matsakis, and Paul A. Viola. Learning from one example through shared densities on transforms. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
334. Fatemeh Mirrashed, Vlad I. Morariu, Behjat Siddiquie, Rogerio S. Feris, and Larry S. Davis. Domain adaptive object detection. In *Workshops on Application of Computer Vision (WACV)*, 2013.
335. Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing Atari with deep reinforcement learning. In *NIPS Workshop on Deep Learning*, 2013.
336. Roozbeh Mottaghi, Xianjie Chen, Xiaobai Liu, Nam-Gyu Cho, Seong-Whan Lee, Sanja Fidler, Raquel Urtasun, and Alan Yuille. The role of context for object detection and semantic segmentation in the wild. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
337. Yair Movshovitz-Attias, Takeo Kanade, and Yaser Sheikh. How useful is photo-realistic rendering for visual learning? *CoRR*, arXiv:1603.08152, 2016.
338. Damian Mrowca, Marcus Rohrbach, Judy Hoffman, Ronghang Hu, Kate Saenko, and Trevor Darrell. Spatial semantic regularisation for large scale object detection. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
339. Krikamol Muandet, David Balduzzi, and Bernhard Schölkopf. Domain generalization via invariant feature representation. In *International Conference on Machine Learning (ICML)*, 2013.
340. Kevin Murphy, Antonio Torralba, and William T. Freeman. Using the forest to see the trees: a graphical model relating features, objects, and scenes. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2003.
341. S. A. Nene, S. K. Nayar, and H. Murase. Columbia Object Image Library (COIL-20). Technical report, CUCS-005-96, February 1996.
342. Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y. Ng. Reading digits in natural images with unsupervised feature learning. In *NIPS Workshop on Deep Learning and Unsupervised Feature Learning (DLUFL)*, 2011.

UNCORRECTED PROOF

343. Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, and Andrew Y Ng. Multimodal deep learning. In *International Conference on Machine Learning (ICML)*, 2011.
344. Jie Ni, Qiang Qiu, and Rama Chellappa. Subspace interpolation via dictionary learning for unsupervised domain sadaptation. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
345. Li Niu, Wen Li, and Dong Xu. Visual recognition by learning from web data: A weakly supervised domain generalization approach. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
346. Hyeonwoo Noh, Seunghoon Hong, and Bohyung Han. Learning deconvolution network for semantic segmentation. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
347. David Novotny, Diane Larlus, and Andrea Vedaldi. I have seen enough: Transferring parts across categories. In *BMVA British Machine Vision Conference (BMVC)*, 2016.
348. Naveen Onkarappa and Angel D. Sappa. Synthetic sequences and ground-truth flow field generation for algorithm validation. *Multimedia Tools and Applications*, 74(9):3121–3135, 2015.
349. Maxime Oquab, Léon Bottou, Ivan Laptev, and Josef Sivic. Learning and transferring mid-level image representations using convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
350. Vicente Ordonez, Jia Deng, Yejin Choi, Alexander C. Berg, and Tamara L. Berg. From large scale image categorization to entry-level categories. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
351. Ivan V. Oseledets. Tensor-train decomposition. *Journal on Scientific Computing*, 33(5):2295–2317, 2011.
352. Sakrapee Paisitkriangkrai, Chunhua Shen, and Anton van den Hengel. Learning to rank in person re-identification with metric ensembles. *CoRR*, arXiv:1503.01543, 2015.
353. Mark Palatucci, Dean Pomerleau, Geoffrey E. Hinton, and Tom M. Mitchell. Zero-shot learning with semantic output codes. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2009.
354. Sinno J. Pan, James T. Tsang, Ivor W.and Kwok, and Qiang Yang. Domain adaptation via transfer component analysis. *Transactions on Neural Networks*, 22(2):199–210, 2011.
355. Sinno J. Pan and Qiang Yang. A survey on transfer learning. *Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010.
356. Sinno Jialin Pan, Xiaochuan Ni, Jian-Tao Sun, Qiang Yang, and Zheng Chen. Cross-domain sentiment classification via spectral feature alignment. In *International Conference on World Wide Web (WWW)*, 2010.
357. Pau Panareda-Busto, Joerg Liebelt, and Juergen Gall. Adaptation of synthetic data for coarse-to-fine viewpoint refinement. In *BMVA British Machine Vision Conference (BMVC)*, 2015.
358. Jeremie Papon and Markus Schoeler. Semantic pose using deep networks trained on synthetic RGB-D. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
359. Devi Parikh and Kristen Grauman. Interactively building a discriminative vocabulary of nameable attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
360. Devi Parikh and Kristen Grauman. Relative attributes. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
361. Amar Parkash and Devi Parikh. Attributes for classifier feedback. In *European Conference on Computer Vision (ECCV)*, 2012.
362. Novi Patricia and Barbara Caputo. Learning to learn, from transfer learning to domain adaptation: A unifying perspective. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
363. Genevieve Patterson and James Hays. SUN attribute database: Discovering, annotating, and recognizing scene attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
364. Xingchao Peng, Baochen Sun, Karim Ali, and Kate Saenko. Learning deep object detectors from 3D models. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.

UNCORRECTED PROOF

365. Bojan Pepik, Michael Stark, Peter Gehler, and Bernt Schiele. Teaching 3D geometry to deformable part models. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
366. Florent Perronnin, Christopher Dance, Gabriela Csurka, and Marco Bressan. Adapted vocabularies for generic visual categorization. In *European Conference on Computer Vision (ECCV)*, 2006.
367. Florent Perronnin, Yan Liu, Jorge Sánchez, and Hervé Poirier. Large-scale image retrieval with compressed Fisher vectors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
368. Florent Perronnin, Jorge Sánchez, and Yan Liu. Large-scale image categorization with explicit data embedding. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
369. Florent Perronnin, Jorge Sánchez, and Thomas Mensink. Improving the fisher kernel for large-scale image classification. In *European Conference on Computer Vision (ECCV)*, 2010.
370. Leonid Pishchulin, Arjun Jain, Mykhaylo Andriluka, Thorsten Thormahlen, and Bernt Schiele. Articulated people detection and pose estimation: reshaping the future. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
371. Leonid Pishchulin, Arjun Jain, Christian Wojek, Mykhaylo Andriluka, Thorsten Thormählen, and Bernt Schiele. Learning people detection models from few training samples. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
372. Peter Prettenhofer and Benno Stein. Cross-language text classification using structural correspondence learning. In *Annual Meeting of the Association for Computational Linguistics(ACL)*, 2010.
373. Amazon Mechanical Turk. http://www.mturk.com.
374. Guo-Jun Qi, Charu Aggarwal, and Thomas Huang. Towards semantic knowledge propagation from text corpus to web images. In *International Conference on World Wide Web (WWW)*, 2011.
375. Guo-Jun Qi, Charu Aggarwal, Yong Rui, Qi Tian, Shiyu Chang, and Thomas Huang. Towards cross-category knowledge propagation for learning visual concepts. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
376. Guo-Jun Qi, Xian-Sheng Hua, Yong Rui, Jinhui Tang, and Hong-Jiang Zhang. Two-dimensional active learning for image classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
377. Qiang Qiu, Vishal M. Patel, Pavan Turaga, and Rama Chellappa. Domain adaptive dictionary learning. In *European Conference on Computer Vision (ECCV)*, 2012.
378. Brian Quanz, Jun Huan, and Meenakshi Mishra. Knowledge transfer with low-quality data: A feature extraction issue. *Transactions on Knowledge and Data Engineering*, 24(10):1789–1802, 2012.
379. Piyush Rai, Avishek Saha, Hal Daumé III, and Suresh Venkatasubramanian. Domain adaptation meets active learning. In *ACL Workshop on Active Learning for Natural Language Processing (ALNLP)*, 2010.
380. Rajat Raina, Alexis Battle, Honglak Lee, Benjamin Packer, and Andrew Y. Ng. Self-taught learning: transfer learning from unlabeled data. In *International Conference on Machine Learning (ICML)*, 2007.
381. Anant Raj, Vinay P. Namboodiri Namboodiri, and Tinne Tuytelaars. Subspace alignment based domain adaptation for rcnn detector. In *BMVA British Machine Vision Conference (BMVC)*, 2015.
382. Nikhil Rasiwasia, Jose Costa Pereira, Emanuele Coviello, Gabriel Doyle, Gert R. G. Lanckriet, Roger Levy, and Nuno Vasconcelos. A new approach to cross-modal multimedia retrieval. In *ACM Multimedia*, 2010.
383. Mohammad Rastegari, Abdou Diba, Devi Parikh, and Alireza Farhadi. Multi-attribute queries: To merge or not to merge? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

384. Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. CNN Features off-the-shelf: an Astounding Baseline for Recognition. *CoRR*, arXiv:1403.6382, 2014.
385. Douglas A. Reynolds, Thomas F. Quatieri, and Robert B. Dunn. Speaker verification using adapted Gaussian Mixture Models. *Digital Signal Processing*, 10(1):19–41, 2000.
386. Stephan R. Richter, Vibhav Vineet, Stefan Roth, and Vladlen Koltun. Playing for data: Ground truth from computer games. In *European Conference on Computer Vision (ECCV)*, 2016.
387. Stephan R. Richter, Vibhav Vineet, Stefan Roth, and Koltun Vladlen. Playing for data: Ground truth from computer games. In *European Conference on Computer Vision (ECCV)*, 2016.
388. Erik Rodner and Joachim Denzler. Learning with few examples by transferring feature relevance. In *BMVA British Machine Vision Conference (BMVC)*, 2009.
389. Erik Rodner, Judy Hoffman, Jeff Donahue, Trevor Darrell, and Kate Saenko. Towards adapting imagenet to reality: Scalable domain adaptation with implicit low-rank transformations. *CoRR*, arXiv:1308.4200, 2013.
390. José A. Rodríguez-Serrano, Harsimrat Sandhawalia, Raja Bala, Florent Perronnin, and Craig Saunders. Data-driven vehicle identification by image matching. In *ECCV Workshop on Computer Vision in Vehicle Technologies: From Earth to Mars (CVVT)*, 2012.
391. José A. Rodríguez-Serrano, Florent Perronnin, Gemma Sánchez, and Josep Lladós. Unsupervised writer adaptation of whole-word HMMs with application to word-spotting. *Pattern Recognition Letters*, 31(8):742–749, 2010.
392. Marcus Rohrbach, Sandra Ebert, and Bernt Schiele. Transfer learning in a transductive setting. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2013.
393. Marcus Rohrbach, Michael Stark, György Szarvas, Iryna Gurevych, and Bernt Schiele. What helps where – and why? semantic relatedness for knowledge transfer. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
394. Bernardino Romera-Paredes, Hane Aung, Nadia Bianchi-Berthouze, and Massimiliano Pontil. Multilinear multitask learning. In *International Conference on Machine Learning (ICML)*, 2013.
395. German Ros, Sebastian Ramos, Manuel Granados, Amir H. Bakhtiary, dAVID Vázquez, and Antonio M. López. Vision-based offline-online perception paradigm for autonomous driving. In *Winter Conference on Applications of Computer Vision (WACV)*, 2015.
396. German Ros, Laura Sellart, Joanna Materzyńska, David Vázquez, and Antonio M. López. The SYNTHIA dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
397. German Ros, Simon Stent, Pablo F. Alcantarilla, and Tomoki Watanabe. Training constrained deconvolutional networks for road scene semantic segmentation. *CoRR*, arXiv:1604.01545, 2016.
398. Chuck Rosenberg, Martial Hebert, and Henry Schneiderman. Semisupervised self-training of object detection models. In *Workshops on Application of Computer Vision (WACV/MOTION)*, 2005.
399. Peter M. Roth, Sabine Sternig, Helmut Grabner, and Horst Bischof. Classifier grids for robust adaptive object detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
400. Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.
401. Artem Rozantsev, Vincent Lepetit, and Pascal Fua. On rendering synthetic images for training an object detector. *Computer Vision and Image Understanding*, 137:24–37, 2015.
402. Artem Rozantsev, Mathieu Salzmann, and Pascal Fua. Beyond sharing weights for deep domain adaptation. *CoRR*, arXiv:1603.06432, 2016.
403. Evgenia Rubinshtein and Anuj Srivastava. Optimal linear projections for enhancing desired data statistics. *Statistics Computing*, 20(3):267–282, 2010.
404. Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.

UNCORRECTED PROOF

405. Olga Russakovsky, Li-Jia Li, and Li Fei-Fei. Best of both worlds: human-machine collaboration for object annotation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
406. Bryan C. Russell, Antonio Torralba, Kevin P. Murphy, and William T. Freeman. LabelMe: a database and web-based tool for image annotation. *International Journal of Computer Vision*, 77:157–173, 2008.
407. Kate Saenko, Brian Kulis, Mario Fritz, and Trevor Darrell. Adapting visual category models to new domains. In *European Conference on Computer Vision (ECCV)*, 2010.
408. Avishek Saha, Piyush Rai, Hal Daumé III, Suresh Venkatasubramanian, and Scott L. DuVall. Active supervised domain adaptation. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, 2011.
409. Jorge Sánchez, Florent Perronnin, Thomas Mensink, and Jakob Verbeek. Image classification with the Fisher Vector: Theory and practice. *International Journal of Computer Vision*, 105(3):222–245, 2013.
410. Ramachandruni N. Sandeep, Yashaswi Verma, and C. V. Jawahar. Relative parts: Distinctive parts for learning relative attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
411. Scott Satkin, Jason Lin, and Martial Hebert. Data-driven scene understanding from 3D models. In *BMVA British Machine Vision Conference (BMVC)*, 2012.
412. Shreyas Saxena and Jakob Verbeek. Heterogeneous face recognition with cnns. In *ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2016.
413. Timo Scharwächter, Markus Enzweiler, Uwe Franke, and Stefan Roth. Efficient multi-cue scene segmentation. In *German Conference on Pattern Recognition (GCPR)*, 2013.
414. Walter J Scheirer, Neeraj Kumar, Peter N Belhumeur, and Terrance E Boult. Multi-attribute spaces: Calibration for attribute fusion and similarity search. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
415. Johannes Schels, Jörg Liebelt, Klaus Schertler, and Rainer Lienhart. Synthetically trained multi-view object class and viewpoint detection for advanced image retrieval. In *International Conference on Multimedia Retrieval (ICMR)*, 2011.
416. Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 1997.
417. Florian Schroff, Antonio Criminisi, and Andrew Zisserman. Harvesting image databases from the web. In *IEEE International Conference on Computer Vision (ICCV)*, 2007.
418. Pierre Sermanet, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. *CoRR*, arXiv:1312.6229, 2013.
419. Burr Settles. active learning literature survey. Technical Report Computer Sciences Technical Report 1648, University of Wisconsin-Madison, 2010.
420. H. Sebastian Seung, Manfred Opper, and Haim Sompolinsky. Query by committee. In *Annual ACM workshop on Computational Learning Theory (CLT)*, 1992.
421. Alireza Shafaei, James J. Little, and Mark Schmidt. Play and learn: Using video games to train computer vision models. In *BMVA British Machine Vision Conference (BMVC)*, 2016.
422. Shai Shalev-Shwartz. *Online Learning: Theory, Algorithms, and Applications.* PhD thesis, Hebrew University, 7 2007.
423. Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2011.
424. Abhishek Sharma and David W. Jacobs. Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
425. Abhishek Sharma, Abhishek Kumar, Hal Daumé III, and David W. Jacobs. Generalized multiview analysis: A discriminative latent space. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

426. Pramod Sharma, Chang Huang, and Ram Nevatia. Unsupervised incremental learning for improved object detection in a video. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
427. Pramod Sharma and Ram Nevatia. Efficient detector adaptation for object detection in a video. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
428. Viktoriia Sharmanska, Novi Quadrianto, and Christoph H. Lampert. Augmented attribute representations. In *European Conference on Computer Vision (ECCV)*, 2012.
429. Sumit Shekhar, Vishal M. Patel, Hien V. Nguyen, and Rama Chellappa. Generalized domain-adaptive dictionaries. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
430. Haoquan Shen, Shoou-I Yu, Yi Yang, Deyu Meng, and Alex Hauptmann. Unsupervised video adaptation for parsing human motion. In *European Conference on Computer Vision (ECCV)*, 2014.
431. Xiaoxiao Shi, Wei Fan, and Jiangtao Ren. Actively transfer domain knowledge. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, 2008.
432. Xiaoxiao Shi, Qi Liu, Wei Fan, Philip S. Yu, and Ruixin Zhu. Transfer learning on heterogeneous feature spaces via spectral transformation. In *IEEE International Conference on Data Mining (ICDM)*, 2010.
433. Zhiyuan Shi, Yongxin Yang, Timothy M Hospedales, and Tao Xiang. Weakly supervised learning of objects, attributes and their associations. In *European Conference on Computer Vision (ECCV)*, 2014.
434. Hidetoshi Shimodaira. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of Statistical Planning and Inference*, 90(2):227–244, 2000.
435. Jamie Shotton, Andrew Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, and Andrew Blake. Real-time human pose recognition in parts from single depth images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
436. Abhinav Shrivastava, Saurabh Singh, and Abhinav Gupta. Constrained semi-supervised learning using attributes and comparative attributes. In *European Conference on Computer Vision (ECCV)*, 2012.
437. Xiangbo Shu, Guo-Jun Qi, Jinhui Tang, and Wang Jingdong. Weakly-shared deep transfer networks for heterogeneous-domain knowledge propagation. In *ACM Multimedia*, 2015.
438. Si Si, Dacheng Tao, and Bo B. Geng. Bregman divergence-based regularization for transfer subspace learning. *Transactions on Knowledge and Data Engineering*, 22(7):929 –942, 2010.
439. Behjat Siddiquie, Rogerio S Feris, and Larry S Davis. Image ranking and retrieval based on multi-attribute queries. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
440. Olivier Sigaud, Clment Masson, David Filliat, and Freek Stulp. Gated networks: an inventory. *CoRR*, arXiv:1512.03201, 2015.
441. Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *European Conference on Computer Vision (ECCV)*, 2012.
442. Karen Simonyan and Andrew Zisserman. Two-stream convolutional networks for action recognition in videos. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
443. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, arXiv:1409.1556, 2014.
444. Ajit Singh, Paul Singh, and Geoffrey J. Gordon. Relational learning via collective matrix factorization. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2008.
445. Josef Sivic, and Andrew Zisserman. Video google: A text retrieval approach to object matching in videos. In *IEEE International Conference on Computer Vision (ICCV)*, 2003.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

446. Brandon Smith and Li Zhang. Collaborative facial landmark localization for transferring annotations across datasets. In *European Conference on Computer Vision (ECCV)*, 2014.
447. Alex Smola, Arthur Gretton, Le Song, and Bernhard Schölkopf. A hilbert space embedding for distributions. In *Algorithmic Learning Theory*, 2007.
448. Yainuvis Socarras, Sebastian Ramos, David Vázquez, Antonio M. López, and Theo Gevers. Adapting pedestrian detection from synthetic to far infrared images. In *ICCV Workshop on Visual Domain Adaptation and Dataset Bias (VisDA)*, 2013.
449. Richard Socher, Milind Ganjoo, Christopher D. Manning, and Andrew Y. Ng. Zero-shot learning through cross-modal transfer. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2013.
450. Richard Socher and Fei-Fei Li. Connecting modalities: Semi-supervised segmentation and annotation of images using unaligned text corpora. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
451. Khurram Soomro, Amir Roshan Zamir, and Mubarak Shah. UCF101: A dataset of 101 human actions classes from videos in the wild. *CoRR*, arXiv:1212.0402, 2012.
452. César De Souza, Adrien Gaidon, Yohann Cabon, and Antonio M. López. Procedural generation of videos to train deep action recognition networks. *CoRR*, arXiv:1612.00881, 2016.
453. Bharath K Sriperumbudur, Arthur Gretton, Kenji Fukumizu, Bernhard Schölkopf, and Gert RG Lanckriet. Hilbert space embeddings and metrics on probability measures. *Journal of Machine Learning Research*, 11:1517–1561, 2010.
454. Anuj Srivastava and Xiuwen Liu. Tools for application-driven linear dimension reduction. *Neurocomputing*, 67:136–160, 2005.
455. Nitish Srivastava, Geoffrey E. Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 2014.
456. Severin Stalder, Helmut Grabner, and Luc Van Gool. Exploring context to learn scene specific object detectors. In *International Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, 2009.
457. Michael Stark, Michael Goesele, and Bernt Schiele. A shape-based object class model for knowledge transfer. In *IEEE International Conference on Computer Vision (ICCV)*, 2009.
458. Michael Stark, Michael Goesele, and Bernt Schiele. Back to the future: Learning shape models from 3D CAD data. In *BMVA British Machine Vision Conference (BMVC)*, 2010.
459. Ingo Steinwart. On the influence of the kernel on the consistency of support vector machines. *Journal of Machine Learning Research*, 2:67–93, 2002.
460. Hao Su, Charles R. Qi, Yangyan Yi, and Leonidas Guibas. Render for CNN: viewpoint estimation in images using CNNs trained with rendered 3D model views. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
461. Hao Su, Fan Wang, Yangyan Yi, and Leonidas Guibas. 3D-assisted feature synthesis for novel views of an object. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
462. Yu Su, Moray Allan, and Frédéric Jurie. Improving object classification using semantic attributes. In *BMVA British Machine Vision Conference (BMVC)*, 2010.
463. Masashi Sugiyama, Shinichi Nakajima, Hisashi Kashima, Paul von. Buenau, and Motoaki Kawanabe. Direct importance estimation with model selection and its application to covariate shift adaptation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2008.
464. Baochen Sun. *Correlation Alignment for Domain Adaptation*. PhD thesis, University of Massachusetts Lowell, 8 2016.
465. Baochen Sun, Jiashi Feng, and Kate Saenko. Return of frustratingly easy domain adaptation. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2016.
466. Baochen Sun, Xingchao Peng, and Kate Saenko. Generating large scale image datasets from 3d cad models. In *CVPR Workshop on The Future of Datasets in Vision (FDV)*, 2015.
467. Baochen Sun and Kate Saenko. From virtual to reality: Fast adaptation of virtual object detectors to real domains. In *BMVA British Machine Vision Conference (BMVC)*, 2014.

UNCORRECTED PROOF

468. Baochen Sun and Kate Saenko. Deep coral: Correlation alignment for deep domain adaptation. In *ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2016.
469. Chen Sun, Chuang Gan, and Ram Nevatia. Automatic concept discovery from parallel text and visual corpora. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
470. Chen Sun, Sanketh Shetty, Rahul Sukthankar, and Ram Nevatia. Temporal localization of fine-grained actions in videos by domain transfer from web images. In *ACM Multimedia*, 2015.
471. Qian Sun, Rita Chattopadhyay, Sethuraman Panchanathan, and Jieping Ye. A two-stage weighting framework for multi-source domain adaptation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2011.
472. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
473. ben Tan, Yangqiu Song, Erheng Zhong, and Qiang Yang. Transitive transfer learning. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2015.
474. Ben Tan, Erheng Zhong, Michael Ng, and K. Qiang Yang. Mixed-transfer: Transfer learning over mixed graphs. In *SIAM International Conference on Data Mining (SDM)*, 2014.
475. Kevin Tang, Vignesh Ramanathan, Li Fei-Fei, and Daphne Koller. Shifting weights: Adapting object detectors from image to video. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2012.
476. Ran Tao, Arnold WM Smeulders, and Shih-Fu Chang. Attributes and categories for generic instance search from one example. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
477. Geoffrey R. Taylor, Andrew J. Chosak, and Paul C. Brewer. OVVV: Using virtual worlds to design and evaluate surveillance systems. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.
478. Unity Technologies. Unity Development Platform.
479. George R. Terrell. The maximal smoothing principle in density estimation. *Journal of the American Statistical Association*, 85(410):470–477, 1990.
480. Giorgos Tolias, Ronan Sicre, and Hervé Jégou. Particular object retrieval with integral max-pooling of CNN activations. In *International Conference on Machine Learning (ICML)*, 2016.
481. Tatiana Tommasi and Barbara Caputo. The more you know, the less you learn: from knowledge transfer to one-shot learning of object categories. In *BMVA British Machine Vision Conference (BMVC)*, 2009.
482. Tatiana Tommasi and Barbara Caputo. Safety in numbers: learning categories from few examples with multi model knowledge transfer. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
483. Tatiana Tommasi and Barbara Caputo. Frustratingly easy NBNN domain adaptation. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
484. Tatiana Tommasi, Novi Patricia, Barbara Caputo, and Tinne Tuytelaars. A deeper look at dataset bias. In *German Conference on Pattern Recognition (GCPR)*, 2015.
485. Tatiana Tommasi, Novi Quadrianto, Barbara Caputo, and Christoph H. Lampert. Beyond dataset bias: Multi-task unaligned shared knowledge transfer. In *Asian Conference on Computer Vision (ACCV)*, 2012.
486. Tatiana Tommasi and Tinne Tuytelaars. A testbed for cross-dataset analysis. In *ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2014.
487. Jonathan J Tompson, Arjun Jain, Yann LeCun, and Christoph Bregler. Joint training of a convolutional network and a graphical model for human pose estimation. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
488. Simon Tong and Daphne Koller. Support vector machine active learning with applications to text classification. *Journal of Machine Learning Research*, 2:45–66, 2002.
489. Antonio Torralba and Alexei A. Efros. Unbiased look at dataset bias. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

490. Antonio Torralba, Kevin P. Murphy, and William T. Freeman. Sharing visual features for multiclass and multiview object detection. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 29(5):854–869, 2007.
491. Lorenzo Torresani, Martin Szummer, and Andrew Fitzgibbon. Efficient object category recognition using classemes. In *European Conference on Computer Vision (ECCV)*, 2010.
492. Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, and Manohar Paluri. C3d: Generic features for video analysis. *IEEE International Conference on Computer Vision (ICCV)*, 2015.
493. Ledyard R. Tucker. Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31(3):279–311, 1966.
494. Shubham Tulsiani, João Carreira, and Jitendra Malik. Pose induction for novel object categories. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
495. Shubham Tulsiani and Jitendra Malik. Viewpoints and keypoints. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
496. Eric Tzeng, Coline Devin, Judy Hoffman, Chelsea Finn, Xingchao Peng, Sergey Levine, Kate Saenko, and Trevor Darrell. Towards adapting deep visuomotor representations from simulated to real environments. *CoRR*, arXiv:1511.07111, 2015.
497. Eric Tzeng, Judy Hoffman, Trevor Darrell, and Kate Saenko. Simultaneous deep transfer across domains and tasks. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
498. Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *NIPS Workshop on Adversarial Training, (WAT)*, 2016.
499. Eric Tzeng, Judy Hoffman, Ning Zhang, Kate Saenko, and Trevor Darrell. Deep domain confusion: Maximizing for domain invariance. *CoRR*, arXiv:1412.3474, 2014.
500. Jasper R.R. Uijlings, Koen E.A. van de Sande, Theo Gevers, and Arnold W.M. Smeulders. Selective search for object recognition. *International Journal of Computer Vision*, 104(2):154–171, 2013.
501. Laurens van der Maaten. Barnes-Hut-SNE. *CoRR*, arXiv:1301.3342, 2013.
502. Manik Varma and Andrew Zisserman. A statistical approach to material classification using image patch exemplars. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 31(11):2032–2047, 2009.
503. David Vazquez, Antonio M. López, Javier Marín, Daniel Ponsa, and David Gerónimo. Virtual and real world adaptation for pedestrian detection. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 36(4):797–809, 2014.
504. David Vázquez, Antonio M. López, Daniel Ponsa, and David Gerónimo. Interactive training of human detectors. In Angel D. Sappa and Jordi Vitrià, editors, *Multimodal Interaction in Image and Video Applications Intelligent Systems*, pages 169–184. Springer, 2013.
505. David Vázquez, Antonio M. López, Daniel Ponsa, and Javier Marín. Cool world: domain adaptation of virtual and real worlds for human detection using active learning. In *NIPS Workshop on Domain Adaptation: Theory and Applications (DATA)*, 2011.
506. David Vázquez, Antonio M. López, Daniel Ponsa, and Javier Marín. Virtual worlds and active learning for human detection. In *International Conference on Multimodal Interaction (ICMI)*, 2011.
507. Andrea Vedaldi, Varun Gulshan, Manik Varma, and Andrew Zisserman. Multiple kernels for object detection. In *IEEE International Conference on Computer Vision (ICCV)*, 2009.
508. Andrea Vedaldi, Siddarth Mahendran, Stavros Tsogkas, Subhrajyoti Maji, Ross Girshick, Juho Kannala, Esa Rahtu, Iasonas Kokkinos, Matthew B Blaschko, Daniel Weiss, Ben Taskar, Karen Simonyan, Naomi Saphra, and Sammy Mohamed. Understanding objects in detail with fine-grained attributes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
509. Ramakrishna Vedantam, Xiao Lin, Tanmay Batra, C. Lawrence Zitnick, and Devi Parikh. Learning common sense through visual abstraction. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

510. V.S.R. Veeravasarapu, Rudra Narayan Hota, Constantin Rothkopf, and Ramesh Visvanathan. Model validation for vision systems via graphics simulation. *CoRR*, arXiv:1512.01401, 2015.
511. V.S.R. Veeravasarapu, Rudra Narayan Hota, Constantin Rothkopf, and Ramesh Visvanathan. Simulations for validation of vision systems. *CoRR*, arXiv:1512.01030, 2015.
512. V.S.R. Veeravasarapu, Constantin Rothkopf, and Ramesh Visvanathan. Model-driven simulations for deep convolutional neural networks. *CoRR*, arXiv:1605.09582, 2016.
513. Sudheendra Vijayanarasimhan and Kristen Grauman. What's it going to cost you?: Predicting effort vs. informativeness for multi-label image annotations. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
514. Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In *International Conference on Machine Learning (ICML)*, 2008.
515. Alexei Vinokourov, Nello Cristianini, and John Shawe-Taylor. Inferring a semantic representation of text via cross-language correlation analysis. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2003.
516. Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.
517. Catherine Wah, Steve Branson, Pietro Perona, and Serge Belongie. Multiclass recognition and part localization with humans in the loop. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
518. Catherine Wah, Steve Branson, Peter Welinder, Pietro Perona, and Serge Belongie. The Caltech-UCSD birds-200-2011 dataset. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011.
519. Chang Wang and Sridhar Mahadevan. Manifold alignment without correspondence. In *AAAI International Joint Conference on Artificial Intelligence (IJCAI)*, 2009.
520. Chang Wang and Sridhar Mahadevan. Heterogeneous domain adaptation using manifold alignment. In *AAAI International Joint Conference on Artificial Intelligence (IJCAI)*, 2011.
521. Heng Wang, Muhammad Muneeb Ullah, Alexander Kläser, Ivan Laptev, and Cordelia Schmid. Evaluation of local spatio-temporal features for action recognition. In *BMVA British Machine Vision Conference (BMVC)*, 2009.
522. LiMin Wang, Yu Qiao, and Xiaoou Tang. Motionlets: Mid-level 3d parts for human motion recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
523. Liwei Wang, Yin Li, and Svetlana Lazebnik. Learning deep structure-preserving image-text embeddings. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
524. Meng Wang and Xiaogang Wang. Automatic adaptation of a generic pedestrian detector to a specific traffic scene. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
525. Weiran Wang, Raman Arora, Karen Livescu, and Jeff Bilmes. On deep multi-view representation learning. In *International Conference on Machine Learning (ICML)*, 2015.
526. Xiaoyang Wang and Qiang Ji. A unified probabilistic approach modeling relationships between attributes and objects. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
527. Xiaoyu Wang, Gang Hua, and Tony X. han. Detection by detections: Non-parametric detector adaptation for a video. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
528. Xiaoyu Wang, Ming Yang, Shenghuo Zhu, and Yuanqing Lin. Regionlets for generic object detection. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
529. Xin-Jing Wang, Lei Zhang, Xirong Li, and Wei-Ying Ma. Annotating images by mining image search results. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 30:1919–1932, 2008.
530. Xuezhi Wang, Tzu-Kuo Huang, and Jeff Schneider. Active transfer learning under model shift. In *International Conference on Machine Learning (ICML)*, 2014.
531. Xuezhi Wang and Jeff Schneider. Flexible transfer learning under support and model shift. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

532. Yang Wang and Greg Mori. A discriminative latent model of object classes and attributes. In *European Conference on Computer Vision (ECCV)*, 2010.
533. Kilian Q Weinberger and Lawrence K Saul. Unsupervised learning of image manifolds by semidefinite programming. *International Journal of Computer Vision*, 70(1):77–90, 2006.
534. Kilian Q. Weinberger and Lawrence K. Saul. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 10:207–244, 2009.
535. Karl Weiss, Taghi M. Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big Data*, 9(3), 2016.
536. Jason Weston, Samy Bengio, and Nicolas Usunier. Large scale image annotation: learning to rank with joint word-image embeddings. *Machine Learning*, 81(1):21–35, 2010.
537. Zheng Whang, Yangqiu Song, and Changshui Zhang. Transferred dimensionality reduction. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, 2008.
538. Lieve Van Woensel, Geoff Archer Archer, and Darja Panades-Estruch, Laura abd Vrscaj. Ten technologies which could change our lives. Technical report, EPRS - European Parlimentary Research Service, January 2015.
539. Phil C. Woodland. Speaker adaptation: techniques and challenges. In *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, 1999.
540. Bo Wu and Ram Nevatia. Improving part based object detection by unsupervised, online boosting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.
541. Jianxiong Xiao, James Hays, Krista A. Ehinger, Aude Oliva, and Antonio Torralba. Sun database: Large-scale scene recognition from abbey to zoo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
542. Min Xiao and Yuhong Guo. Semi-supervised subspace co-projection for multi-class heterogeneous domain adaptation. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, 2015.
543. Chang Xu, Dacheng Tao, and Chao Xu. A survey on multi-view learning. *CoRR*, arXiv:1304.5634, 2013.
544. Jiaolong Xu, Sebastian Ramos, David Vázquez, and Antonio M. López. Domain adaptation of deformable part-based models. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 36(12):2367–2380, 2014.
545. Jiaolong Xu, Sebastian Ramos, David Vazquez, and Antonio M. López. Hierarchical adaptive structural svm for domain adaptation. *CoRR*, arXiv:1408.5400, 2014.
546. Jiaolong Xu, Sebastian Ramos, David Vázquez, and Antonio M. López. Hierarchical adaptive structural SVM for domain adaptation. *International Journal of Computer Vision*, 119(2):159–178, 2016.
547. Jiaolong Xu, David Vázquez, Antonio M. López, Javier Marín, and Daniel Ponsa. Learning a part-based pedestrian detector in a virtual world. *Transactions on Intelligent Transportation Systems*, 15(5):2121–2131, 2014.
548. Jiaolong Xu, David Vázquez, Krystian Mikolajczyk, and Antonio M. López. Hierarchical online domain adaptation of deformable part-based models. In *International Conference on Robotics and Automation (ICRA)*, 2016.
549. Xiang Xu, Shaogang Gong, and Timothy M. Hospedales. Cross-domain traffic scene understanding by motion model transfer. In *International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Stream (ARTEMIS)*, 2013.
550. Zheng Xu, Wen Li, Li Niu, and Dong Xu. Exploiting low-rank structure from latent domains for domain generalization. In *European Conference on Computer Vision (ECCV)*, 2014.
551. Zhijie Xu and Shiliang Sun. Multi-source transfer learning with multi-view adaboost. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2010.
552. Makoto Yamada, Leonid Sigal, and Michalis Raptis. No bias left behind: Covariate shift adaptation for discriminative 3d pose estimation. In *European Conference on Computer Vision (ECCV)*, 2012.
553. Fei Yan and Krystian Mikolajczyk. Deep correlation for matching images and text. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

554. Yuguang Yan, Qingyao Wu, Mingkui Tan, and Huaqing Min. Online heterogeneous transfer learning by weighted offline and online classifiers. In *ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2016.
555. Jun Yang, Rong Yan, and Alexander G. Hauptmann. Cross-domain video concept detection using adaptive SVMs. In *ACM Multimedia*, 2007.
556. Jun Yang, Rong Yan, and Alexander G. Hauptmann. Cross-domain video concept detection using adaptive svms. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.
557. Liu Yang, Liping Jing, Jian Yu, and Michael K. Ng. Learning transferred weights from co-occurrence data for heterogeneous transfer learning. *Transactions on Neural Networks and Learning Systems*, 27(11):2187–2200, 2015.
558. Meng Yang, Lei Zhang, Xiangchu Feng, and David Zhang. Fisher discrimination dictionary learning for sparse representation. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
559. Qiang Yang, Yuqiang Chen, Gui-Rong Xue, Wenyuan Dai, and Yu. Yong. Heterogeneous transfer learning for image clustering via the socialweb. In *Annual Meeting of the Association for Computational Linguistics(ACL)*, 2009.
560. Yi Yang and Deva Ramanan. Articulated human detection with flexible mixtures of parts. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 35(12):2878–2890, 2013.
561. Yongxin Yang and Timothy M. Hospedales. A unified perspective on multi-domain and multi-task learning. In *International Conference on Learning representations (ICLR)*, 2015.
562. Yongxin Yang and Timothy M. Hospedales. Deep multi-task representation learning: A tensor factorisation approach. In *ICLR*, 2016.
563. Yongxin Yang and Timothy M. Hospedales. Multivariate regression on the grassmannian for predicting novel domains. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
564. Yongxin Yang and Timothy M. Hospedales. Trace norm regularised deep multi-task learning. *CoRR*, arXiv:1606.04038, 2016.
565. Ting Yao, Yingwei Pan, Chong-Wah Ngo, Houqiang Li, and Tao Mei. Semi-supervised domain adaptation with subspace learning for visual recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
566. Yi Yao and Gianfranco Doretto. Boosting for transfer learning with multiple sources. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
567. Dong Yi, Zhen Lei, and Stan Z. Li. Deep metric learning for practical person re-identification. *CoRR*, arXiv:1407.4979, 2014.
568. Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
569. Felix X. Yu, Rongrong Ji, Ming-Hen Tsai, Guangnan Ye, and Shih-Fu Chang. Weak attributes for large-scale image retrieval. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
570. Xiaodong Yu and Yiannis Aloimonos. Attribute-based transfer learning for object categorization with zero/one training example. In *European Conference on Computer Vision (ECCV)*, 2010.
571. Bianca Zadrozny. Learning and evaluating classifiers under sample selection bias. In *International Conference on Machine Learning (ICML)*, 2004.
572. Matthew D. Zeiler. ADADELTA: an adaptive learning rate method. *CoRR*, arXiv:1212.5701, 2012.
573. Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. *CoRR*, arXiv:1311.2901, 2013.
574. Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *European Conference on Computer Vision (ECCV)*, 2014.
575. Matthew D. Zeiler, Dilip Krishnan, Graham W. Taylor, and Rob Fergus. Deconvolutional networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/5/2017 Pages: 354 Layout: T1-Standard

576. Zheng-Jun Zha, tao Mei, Meng Wang, Zengfu Wang, and Xian-Sheng Hua. Robust distance metric learning with auxiliary knowledge. In *AAAI International Joint Conference on Artificial Intelligence (IJCAI)*, 2009.
577. Deming Zhai, Bo Li, Hong Chang, Shiguang Shan, Xilin Chen, and Wen Gao. Manifold alignment via corresponding projections. In *BMVA British Machine Vision Conference (BMVC)*, 2010.
578. Cha Zhang, Raffay Hammid, and Zhengyou Zhang. Taylor expansion based classifier adaptation: Application to person detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
579. Kun Zhang, Bernhard Schölkopf, Krikamol Muandet, and Zhikun Wang. Domain adaptation under target and conditional shift. *Journal of Machine Learning Research*, 28(3):819–827, 2013.
580. Ning Zhang, Jeff Donahue, Ross Girshick, and Trevor Darrell. Part-based R-CNNs for fine-grained category detection. In *European Conference on Computer Vision (ECCV)*, 2014.
581. Ning Zhang, Manohar Paluri, Marc'Aurelio Ranzato, Trevor Darrell, and Lubomir Bourdev. Panda: Pose aligned networks for deep attribute modeling. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
582. Yu Zhang and Dit-Yan Yeung. Transfer metric learning by learning task relationships. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2010.
583. Zhanpeng Zhang, Ping Luo, Chen Change Loy, and Xiaoou Tang. Facial landmark detection by deep multi-task learning. In *European Conference on Computer Vision (ECCV)*, 2014.
584. Ziming Zhang and Venkatesh Saligrama. Person re-identification via structured prediction. *CoRR*, arXiv:1406.4444, 2014.
585. Rui Zhao, Wanli Ouyang, and Xiaogang Wang. Person re-identificationby saliency learning. *CoRR*, arXiv:1412.1908, 2014.
586. Shuai Zheng, Junge Zhang, Kaiqi Huang, Ran He, and Tieniu Tan. Robust view transformation model for gait recognition. In *International Conference on Image Processing (ICIP)*, 2011.
587. Erheng Zhong, Wei Fan, Jing Peng, Kun Zhang, Jiangtao Ren, Deepak Turaga, and Olivier Verscheure. Cross domain distribution adaptation via kernel mapping. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2009.
588. Joey Tianyi Zhou, Sinno Jialin Pan, Ivor W. Tsang, and Yan Yan. Hybrid heterogeneous transfer learning through deep learning. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2014.
589. Joey Tianyi Zhou, Ivor W. Tsang, Sinno Jialin Pan, and Mingkui Tan. Heterogeneous domain adaptation for multiple classes. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2014.
590. Mianwei Zhou and Kevin C. Chang. Unifying learning to rank and domain adaptation: Enabling cross-task document scoring. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2014.
591. Fan Zhu and Ling Shao. Enhancing action recognition by cross-domain dictionary learning. In *BMVA British Machine Vision Conference (BMVC)*, 2013.
592. Xiangxin Zhu, Carl Vondrick, Charles C. Fowlkes, and Deva Ramanan. Do we need more training data? *International Journal of Computer Vision*, 119(1):76–92, 2016.
593. Xiaojin Zhu, Andrew B. Goldberg, Ronald Brachman, and Thomas Dietterich. *Introduction to semi-supervised learning*. Morgan & Claypool Publishers, 2009.
594. Yin Zhu, Yuqiang Chen, Zhongqi Lu, Sinno J. Pan, Gui-Rong Xue, Yong Yu, and Qiang Yang. Heterogeneous transfer learning for image classification. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2011.
595. Yuke Zhu, Roozbeh Mottaghi, Eric Kolve, Joseph J. Lim, and Abhinav Gupta. Target-driven visual navigation in indoor scenes using deep reinforcement learning. *CoRR*, arXiv:1609.05143, 2016.

UNCORRECTED PROOF

596. C. Lawrence Zitnick and Piotr Dollár. Edge boxes: Locating object proposals from edges. In *European Conference on Computer Vision (ECCV)*, 2014.
597. C. Lawrence Zitnick, Ramakrishna Vedantam, and Devi Parikh. Adopting abstract images for semantic scene understanding. *Transactions of Pattern Recognition and Machine Analyses (PAMI)*, 38(4):627–638, 2016.
598. Laurent Zwald and Gilles Blanchard. On the convergence of eigenspaces in kernel principal component analysis. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2005.

UNCORRECTED PROOF

420546_1_En_BOOKBACKMATTER ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:**25/5/2017** Pages: **354** Layout: **T1-Standard**