EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Mathematisch-Naturwissenschaftliche Fakultät

Lernbasierte Computer Vision

Bachelorarbeit

A Comparison of Synthetic-to-Real Domain Adaptation Techniques

Eberhard Karls Universität Tübingen
Mathematisch-Naturwissenschaftliche Fakultät
Wilhelm-Schickard-Institut für Informatik
Lernbasierte Computer Vision
Peter Trost, `peter.trost@student.uni-tuebingen.de`, 2019

Bearbeitungszeitraum: 24.05.2019-23.09.2019

Betreuer/Gutachter: Prof. Dr. Andreas Geiger, Universität Tübingen

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe und dass alle Stellen, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen sind, durch Angaben von Quellen als Entlehnung kenntlich gemacht worden sind. Diese Bachelorarbeit wurde in gleicher oder ähnlicher Form in keinem anderen Studiengang als Prüfungsleistung vorgelegt.

Peter Trost (Matrikelnummer 4039682), June 13, 2019

Abstract

Template

Acknowledgments

If you have someone to Acknowledge ;)

Contents

1. Introduction

What is this all about?

Cite like this: [GPAM+14]

1.1. Problem Statement

TODO: what you have to do here :)

2. Related Work

2.1. Adapting Visual Category Models to New Domains

see [SKFD10]

Notes while reading:

2.1.1. Abstract

- one of the first studies of domain shift in context of object recognition
- method that adapts object models acquired in a particular visual domain to new imaging conditions by learning a transformation that minimizes the effect of domain-induced changes in the feature distribution **TODO: is copy paste, rephrase this!**
- supervised learning
- no labeled examples in the new domain needed
- could also be applied to non-image data
- authors also contribute a freely available multi-domain object database

2.1.2. Introduction

- kernel-based, nearest-neighbor classifiers (**TODO: look these up**) often fail on other visual domains than the one trained on
- often want to label *target* visual domain that doesn't have labels yet while having access to *source* domain that has labeled examples
- insufficient using object classifiers trained on source domain **TODO: include Figure 1, look up SVM[-bow] and NBNN**
- domain shift can affect feature distribution and cause the classifier to fail its prediction
- causes of visual domain shift include changes in camera, image resolution, lighting, background, viewpoint, post-processing **TODO: rephrase!**

- introduce domain adaptation technique based on cross-domain transformations
- key idea: regularized non-linear transformation that maps points in the source domain (green) closer to those in the target domain (blue) using supervised data from both domains. The input consists of labeled pairs of inter-domain examples that are known to be either similar (black lines) or dissimilar (red lines). The output is the learned transformation, which can be applied to previously unseen test data points. **TODO: include Figure 2, rephrase!**
- **TODO: look up '[theoretic] metric learning'**

2.1.3. Domain Adaptation Using Regularized Cross-Domain Transforms

general domain adaptation model in linear setting:
let source domain be $\mathcal{A}$ and target domain $\mathcal{B}$. Vectors $\mathbf{x} \in \mathcal{A}$, $\mathbf{y} \in \mathcal{B}$. Learn transformation W from $\mathcal{B}$ to $\mathcal{A}$ (and W^T from $\mathcal{A}$ to $\mathcal{B}$). Let dimensionality of $\mathbf{x}$ be d_A and of $\mathbf{y}$ be d_B then the transformation matrix W is $d_A \times d_B$. Resulting inner product similarity function between $\mathbf{x}$ and the transformed $\mathbf{y}$ as

$$\mathrm{sim}_W(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T W \mathbf{y}$$

to avoid overfitting use regularization function for W denoted as $r(W)$.

TODO: lookup Mahalanobis metric learning method, information theoretic metric learning (ITML)

($\|W\|$: square root of sum of squares of elements)

2.2. Adversarial Discriminative Domain Adaptation

see [THSD17]

TODO: Look up cross entropy loss

2.3. Adversarial Dropout Regularization

see [SUHS17]

experiments on p4d dataset and VisDA-classification
TODO: read this again for the benchmarks and comparison of domain adaptation techniques

2.4. Adversarial Feature Learning

see [DKD16]

TODO: look up jensen-shannon divergence

2.5. Adversarially Learned Inference

see [DBP+16]

2.6. CyCADA: Cycle Consistent Adversarial Domain Adaptation

see [HTP+17]

- synthetic datasets cheaper and more accurate in classification than real ones
- per-pixel label accuracy drops from 93%(real) to 54%(synthetic)
- while translating from synth to real semantic information might be lost (e.g translating line-drawing of a cat to a picture of a dog)
- CyCADA uses cycle consistency and semantic losses
-

3. Datasets

4. Domain Adaptation Techniques

4.1. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

see [ZPIE17]

- image-to-image translation: extracting characteristics of an image and translating it to another style while preserving the characteristics (rgb to greyscale, painting to photo,..)
- special in this approach: no paired images necessary (datasets with paired images are far more expensive)
- create mapping $G : X \rightarrow Y$ from source domain X to target domain Y
- The Generator has to trick the discriminator into believing $G(x), x \in X$ is actually a real sample y from the target domain Y (matches distribution $p_{\text{data}}(y)$)
- problem of mode collapse: any inputimage will be translated to the same output image
- add cycle-consistency constraint: create mapping $F : Y \rightarrow X$ and add contraint $F(G(x)) \overset{!}{\approx} x$
- objective for mapping/generator G and discriminator D_Y:
  $\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{y\ p_{\text{data}}(y)}[\log D_Y(y)] + \mathbb{E}_{x\ p_{\text{data}}(x)}[1 - \log D_Y(G(x))]$
- analogous for mapping/generator F and discriminator D_X
- generators try to minimize the objective, discriminators try to maximize it
- cycle consistency loss:
  $\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x\ p_{\text{data}}(x)}[||F(G(x)) - x||_1] + \mathbb{E}_{y\ p_{\text{data}}(y)}[||G(F(y)) - y||_1]$
- full objective:
  $\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) + \lambda \mathcal{L}_{\text{cyc}}(G, F)$
- solve: $G^*, F^* = \arg\min_{G,F} \max_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y)$

5. Conclusion

To conclude...

A. Blub

Bibliography

[DBP+16] Vincent Dumoulin, Ishmael Belghazi, Ben Poole, Olivier Mastropietro, Alex Lamb, Martin Arjovsky, and Aaron Courville. Adversarially learned inference. *arXiv preprint arXiv:1606.00704*, 2016.

[DKD16] Jeff Donahue, Philipp Krähenbühl, and Trevor Darrell. Adversarial feature learning. *CoRR*, abs/1605.09782, 2016.

[GPAM+14] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 2672–2680. Curran Associates, Inc., 2014.

[HTP+17] Judy Hoffman, Eric Tzeng, Taesung Park, Jun-Yan Zhu, Phillip Isola, Kate Saenko, Alexei A. Efros, and Trevor Darrell. Cycada: Cycle-consistent adversarial domain adaptation. *CoRR*, abs/1711.03213, 2017.

[SKFD10] Kate Saenko, Brian Kulis, Mario Fritz, and Trevor Darrell. Adapting visual category models to new domains. In Kostas Daniilidis, Petros Maragos, and Nikos Paragios, editors, *Computer Vision – ECCV 2010*, pages 213–226, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[SUHS17] Kuniaki Saito, Yoshitaka Ushiku, Tatsuya Harada, and Kate Saenko. Adversarial dropout regularization. *CoRR*, abs/1711.01575, 2017.

[THSD17] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. *CoRR*, abs/1702.05464, 2017.

[ZPIE17] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *CoRR*, abs/1703.10593, 2017.