Deep Domain Adaption for Convolutional Neural Network (CNN) based Iris Segmentation: Solutions and Pitfalls

Ehsaneddin Jalilian and Adnreas Uhl

Department of Computer Science
University of Salzburg

BIOSIG – 2019

Outlines

1. Introduction and Problem Statement
2. Domain Adaptation Types
3. Discrepancy-Based (Homogeneous DDA)
4. Adversarial-Based (Homogeneous DDA)
5. Adversarial-Reconstruction-Based (Homogeneous DDA)
6. Heterogeneous DDA
7. Experimental Framework
8. Evaluations
9. Findings and Analysis
10. References

Introduction and Problem Statement

Introduction

- Segmentation of iris texture in eye images is a vital step in iris recognition.
- Showing superior performance, convolutional neural networks are increasing being used for iris segmentation.

Problem Statement

- Normally a domain shift exists between the source data (on which the network is trained) and the target data (*i.e.* collected from different sensors), causing the CNNs to fail in performing the segmentation in the target data well.
- Data labeling is extremely expensive and time-consuming process, and some times even the labels are not available.
- Using domain adaption techniques (*e.g.* DDA) we are able to cop with the domain shift between source data (for which the labels are available) and target data (for which the labels are not available).

Domain Adaptation Types

- Generally, domain adaption (DA) falls into two main categories base on the homogeneity of the source and target data:
  - Homogeneous DA: Feature spaces between the source and target domains are identical (the distribution of the features differ (e.g., changes in illumination, pose, and image quality)).
  - Heterogeneous DA: Feature spaces between the source and target domains are nonequivalent (RGB vs. depth, sketches vs. photos).

Figure. Smaple images from LFW[1], BCS[2], and CUFS[3] databases respectibely

[1] http://vis-www.cs.umass.edu/lfw/

[2] http://openaccess.thecvf.com/content_ICCV_2017_workshops/papers/w23/Xia_Detecting_Smiles_of_ICCV_2017_paper.pdf

[3] http://mmlab.ie.cuhk.edu.hk/archive/facesketch.html

Discrepancy-Based Approaches

- Discrepancy-Based Approaches: Fine-tun the deep network with labeled or unlabeled target data to diminish the domain shift (mostly used for classification and detection).
  - Class Criterion: Use the class label information (both in target and source) as the guide for transferring knowledge between different domains.
  - Statistic Criterion: Align the statistical distribution shift between source and target domains using available algorithms like: Maximum Mean Discrepancy (MMD), Gram matrix, ...
  - Architecture Criterion: Optimize the architecture of the network to minimize the distribution discrepancy.

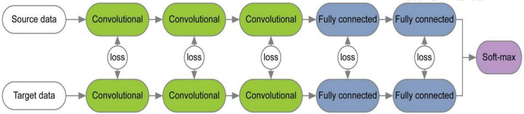

Figure. The two-stream architecture with related weights [1]

Adversarial-Based Approaches

- Adversarial networks consist of two models: A generative model that extracts the data distribution and a discriminative model that distinguishes whether a sample is from the generator or training dataset.
- Generative Models: Estimate generative models via an adversarial process (generative adversarial networks (GANs)).

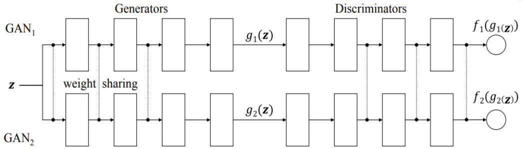

Figure. A generative adversarial network architecture [2]

Adversarial-Based (Homogeneous DDA)

- Nongenerative Models: The feature extractor learns a discriminative representation using labels in the source domain and maps target data to the same space through a domain-confusion loss.

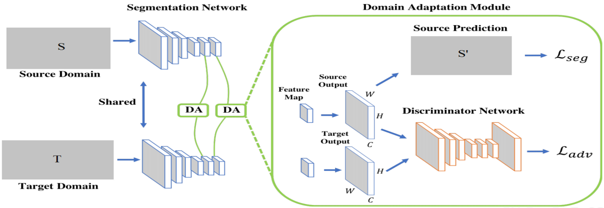

Figure. Non-generative Adversarial Learning Model (NALM) [3]

$$
\begin{aligned}
L_{seg}(I_s) &= -\sum_{h,w}\sum_{c\in C} Y_s^{(h,w,c)} \log(P_s^{(h,w,c)}) \\
L_{adv}(I_t) &= -\sum_{h,w} \log(D(P_t)^{(h,w,\mathbf{1})}) \\
L(I_s, I_t) &= \sum_i \lambda^i{}_{seg} L^i{}_{seg}(I_s) + \sum_i \lambda^i{}_{adv} L^i{}_{adv}(I_t)
\end{aligned}
\tag{1}
$$

Adversarial-Reconstruction Approaches

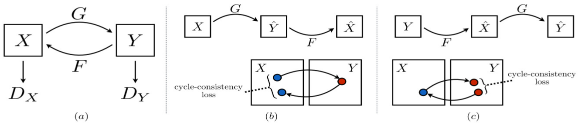

Figure. An Adversarial Reconstruction Model (CycleGAN) [4]

- Adversarial Reconstruction Models: Learn to translate an image from a source domain X to a target domain Y in the absence of paired examples. The reconstruction error is measured as the difference between the reconstructed and original images within each image domain by a cyclic mapping obtained via a GAN discriminator.

  Generator G: X ⇒ Y : translates images from X to Y
  Generator F: Y ⇒ X : translates images from Y to X
  Discriminator DX : scores how real an image of X looks (does this image look like X)
  Discriminator DY : scores how real an image of Y looks (does this image look like Y)

- CycleGAN loss function combines GAN loss and cycle-consistency loss:

$$\begin{aligned} L_{gan}(G, D_Y, X, Y) &= E_{y\sim Pdata(y)}[logD_Y(y)] + E_{x\sim Pdata(x)}[log(1 - D_Y(G(x)))] \\ L_{cyc}(G, F) &= E_{x\sim Pdata(x)}[||F(G(x)) - x||_1] + E_{y\sim Pdata(y)}[||G(F(y)) - y||_1] \\ L(G, F, D_x, D_y) &= L_{gan}(G, D_x, X, Y) + L_{gan}(F, D_x, Y, X) + \lambda L_{cyc}(G, F) \end{aligned} \tag{2}$$

- There exist many variants of adversarial reconstruction models like: DiscoGAN, SpGAN, DITDR [5],...

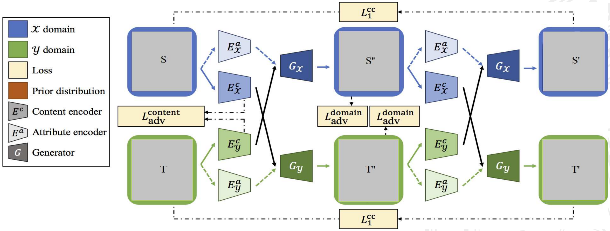

Figure. Diverse Image-to-image Translation via Disentangled Representations (DITDR)

Heterogeneous Domain Adaptation

- Discrepancy-Based Approaches: Neural style transfer (NST) model [6], extracts the contents of target data in higher layers, and the textures of the source data in lower layers, combining them using Gram Matrix.

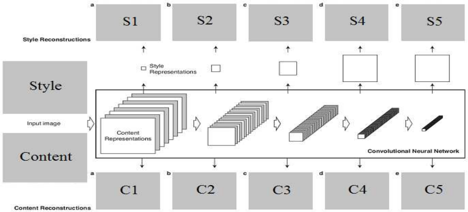

Figure. Heterogeneous Discrepancy-Based model [6]

- Adversarial-Based Approaches: E.g. weakly shared networks are used to synthesize images from text descriptions [7].
- Reconstruction-Based Approaches: E.g. CycleGan network is used to adapt between two heterogeneous domains.

Experimental Framework

- We selected some DDA, considering certain criteria:
  - Use a fully unsupervised scheme (for the target data)
  - Provide unpaired image-to-image adaptation
  - Deliver pixel level semantic affinity between the adapted domains.
- Methodes: We considered NALM, CycleGAN, DITDR, and NST.
- Databases: Casia4i (2640 images of 249 subjects), IITD (2240 images of 224 subjects), and Casia5a (1880 images of 47 subjects) databases.
- Evaluation: We developed six sets of unique database pairs, using the three available databases, terming them as "TargetToSource" (adapted from the "Target" to "Source").
- To evaluate the performance of DDA methods (not generating segmentation masks), we considered a fully convolutional neural network (FCN) [8].
- Iris segmentation accuracies were evaluated using *nice*1 and *nice*2 segmentation scores based on the NICE-I protocol[4].

[4] http://nice1.di.ubi.pt/evaluation.htm

Evaluations

- Applying the generative based methods (CycleGan, DITDR): They are not able to preserve the geometric properties of the image contents (i.e. iris circle shape) as well as the iris texture (which holds the unique information utilized for recognition) in the reconstructed images.

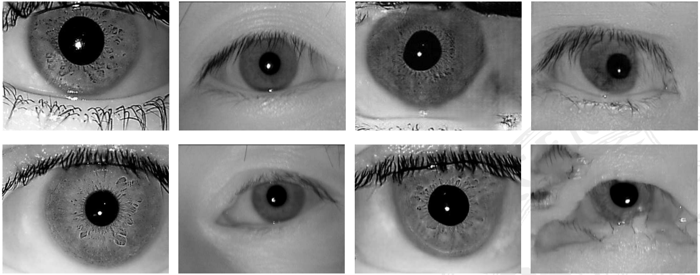

Figure. Sample images from IITD and Casia5a databases (first and second columns), and their corresponding adaptions using CycleGAN and DITDR models (third and fourth columns) respectively

Evaluations

- Applying NALM, NST models and comparing the results to those of the shallow (non-linear) DA algorithm proposed in [9] (as the only existing DA methode for CNN based iris segmentation).
- It should be noted that [9] uses the target data labels to map the distributions between the two domains.

Method	NST		NALM		[9]		Baseline	
Scores	nice1	nice2	nice1	nice2	nice1	nice2	nice1	nice2
Casia5aToCasia4i	0.14	0.30	0.21	0.36	0.02	0.07	0.27	0.40
Casia5aToIITD	0.05	0.16	0.07	0.22	0.03	0.08	0.04	0.11
Casia4iToCasia5a	0.30	0.62	0.29	0.61	0.27	0.35	0.29	0.64
Casia4iToIITD	0.10	0.09	0.17	0.19	0.20	0.17	0.31	0.58
IITDToCasia5a	0.24	0.22	0.15	0.21	0.26	0.30	0.22	0.22
IITDToCasia4i	0.08	0.10	0.07	0.10	0.10	0.09	0.21	0.21

Table: **Segmentation scores for NST, NALM, and [9] (Nonlinear) methods against the baseline**

Evaluations

- Examplary results obtained by applying the non-generative adversarial learning model (NALM) to our experimental data.

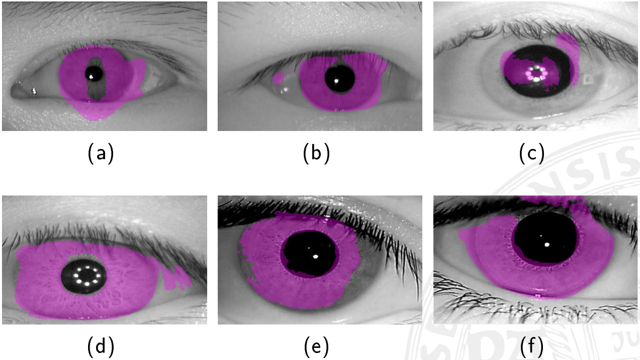

Figure. Sample adapted images and the segmentations for: Casia5aToCasia4i (a), Casia5aToIITD (b), Casia4iToCasia5a (c), Casia4iToIITD (d), IITDToCasia5a (e), IITDToCasia4i (f) using NALM

Evaluations

- Examplary results obtained by applying the neural style transfer model (NST) to our experimental data.
- While there exist some artifacts, the model is able to preserve the geometric properties of the image contents (*e.g.* iris circle shape) and the iris texture in the reconstructed images.

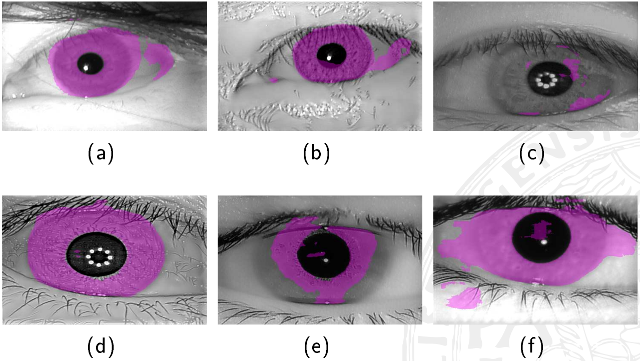

Figure. Sample adapted images and the segmentations for: Casia5aToCasia4i (a), Casia5aToIITD (b), Casia4iToCasia5a (c), Casia4iToIITD (d), IITDToCasia5a (e), IITDToCasia4i (f) using NST

Evaluations

- To assess the actual performance of the methods within an iris segmentation framework, we compared the obtained results to those of some traditional iris segmentation algorithms: Osiris [5], Caht [6], and Wahet [7], as well as our FCN network (when trained and tested on an identical database, assuming that the training labels are available).

Method	Osiris		Caht		Wahet		FCN	
Scores	nice1	nice2	nice1	nice2	nice1	nice2	nice1	nice2
Casia5a	0.018	0.033	0.036	0.151	0.024	0.083	0.008	0.002
Casia4i	0.056	0.067	0116	0.147	0.060	0.084	0.044	0.043
IITD	0.055	0.075	0.113	0.156	0.137	0.176	0.053	0.059

Table: Segmentation scores obtained by applying Osiris, Wahet, Cahet, and FCN network on the experimental databases

[5] Viterbi algorithm on the gradient map of anisotropic smoothed iris

[6] Contrast-adjusted hough transform

[7] Weighted adaptive Hough and ellipsopolar transform

Findings and Analysis

- Models which are able to split and treat the content and the texture information (i.e. NST), perform comparably better on iris domains possessing diverged (heterogeneous) feature spaces.
- NALM model delivers promising results mostly on databases that possess higher content affinities (e.g. IITDToCasia4i).
- Nonetheless, both approaches obtained promising results adapting between Casia4i and IITD databases, performing superior to the [9], and the baseline.
- As a key strength, in these approaches no target label data is used in the adaptation process.
- The worst results for both approaches (expectedly) are obtained on Casia4iToCasia5a adaptation, which seems to be due to the high divergence of the feature spaces between these two databases, and the subsequent difficulties networks have to learn the mapping specially from Casia4i to Casia5a.

Findings and Analysis

- While experimental results proved expediency of existing unsupervised DDA models, yet in most cases, the adaptations and their subsequent segmentations were far from the optimal (to be used for recognition).
- We can conclude that many issues still remain to be addressed when it comes to CNN based iris segmentation using deep domain adaptation techniques.
- Analyzing the performance, we can also predict that unsupervised-heterogeneous deep DA methods would attract more attention for CNN based iris segmentation applications in future.

End

Thank you, Remarks?

References

Artem Rozantsev, Mathieu Salzmann, and Pascal Fua.
Beyond sharing weights for deep domain adaptation.
IEEE transactions on pattern analysis and machine intelligence, 41(4):801–814, 2019.

Ming-Yu Liu and Oncel Tuzel.
Coupled generative adversarial networks.
In *Advances in neural information processing systems*, pages 469–477, 2016.

Tsai Yi-Hsuan, Hung Wei-Chih, Schulter Samuel, Sohn Kihyuk, Yang Ming-Hsuan, and Chandraker Manmohan-Krishna.
Learning to adapt structured output space for semantic segmentation.
2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 7472–7481, 2018.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei Efros.
Unpaired image-to-image translation using cycle-consistent adversarial networks.
In *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.

Lee Hsin-Ying, Tseng Hung-Yu, Huang Jia-Bin, Kumar Singh Maneesh, and Yang Ming-Hsuan.
Diverse image-to-image translation via disentangled representations.
CoRR, abs/1808.00948, 2018.

Leon A.Gatys, Alexander S.Ecker, and Matthias Bethge.
A neural algorithm of artistic style.
CoRR, abs/1508.06576, 2015.

Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaolei Huang, Xiaogang Wang, and Dimitris N. Metaxas.
Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks.
In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5907–5915, 2017.

Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla.
Segnet: A deep convolutional encoder-decoder architecture for image segmentation.
IEEE transactions on pattern analysis and machine intelligence, 39(12):2481–2495, 2017.

Ehsaneddin Jalilian, Andreas Uhl, and Roland Kwitt.
Domain adaptation for cnn based iris segmentation.
In *International Conference of the Biometrics Special Interest Group (BIOSIG'17)*, pages 51–60, Darmstadt, 2017.