Domain Adaptation for CNN Based Iris Segmentation

Ehsaneddin Jalilian[1], Andreas Uhl[1] and Roland Kwitt[1]

[1]Department of Computer Sciences
University of Salzburg
Salzburg, Austria

BIOSIG 2017, Darmstadt

Outlines

Convolutional neural networks (CNNs) and iris segmentation

- CNNs demonstrated considerable success in solving key artificial vision challenges such as object detection, recognition, and segmentation
- Segmentation of iris texture in eye images is a key problem in iris recognition, which plays vital role in accuracy of the system
- In recent years, application of CNNs for iris segmentation has received some research attention[1] [2]

Problem statement

- Training CNNs requires adequate amount of labeled data
- Data labeling is extremely expensive and time-consuming process
- To confront this issue, we considered to adapt the domains of available labeled data to those of the targets, and train CNNs with the adapted data, and segment the target data, eliminating the need for the target data labels

Domain adaptation

- Given a source database $(X_s, Y_s, P(X_s))$ and a target database $(X_t, Y_t, P(X_t))$
- Under the domain difference scenario, we assume the conditional distributions of Y_s and Y_t are the same, but the marginal distributions of X_s and X_t differ in the two domains
- The distinction between two distributions is referred to as sample bias ϕ so as:

$$P_t = P_s(\phi(X_s), Y_s). \tag{1}$$

CNN based domain adaptation

- Using empirical risk minimization framework for supervised learning, we want to select an optimal parameter $\psi^{'}$, to minimize the following objective function

$$\psi_t^{'} = \arg\min_{\psi \in \Psi} \sum_{(x,y) \in X \times Y} \widetilde{P}_s(\phi(X_s), Y_s) g(x, y, \psi) = \arg\min_{\psi \in \Psi} \sum_{i=1}^{N} g(\phi(x_s), y_s, \psi). \tag{2}$$

- Weighting the images' intensities of source data by ϕ provides the solution to the minimization function

Linear Domain Adaptation

Linear intensity transfer

- Straight forward solution to weight the intensities of source data is using a linear normalization model:

$$b = (max(B) - min(B))\frac{a - min(A)}{max(A) - min(A)} + (min(B)). \qquad (3)$$

- Extract the intensity ranges of iris, non-iris, pupil regions in the target database
- Using this model, adapt the intensities in source data to those of the target
- Train the network with the adapted data and test it on the targets

Experimental framework

- Databases: Casia4i database[1], Iitd database[2], Casia5a database[3]
- Metrics: Segmentation error scores: nice1, nice2[4], and F1 score
- Network: Fully Convolutional Encoder-Decoder Network (FCN) [3]

Method	Adapted-target			Baseline(Source-target)		
Scores	nice1	nice2	f1	nice1	nice2	f1
Casia5a-casia4i	0.186	0.220	0.610	0.292	0.640	0.003
Casia5a-iitd	0.148	0.172	0.781	0.229	0.221	0.473
Casia4i-casia5a	0.066	0.194	0.730	0.274	0.406	0.341
Casia4i-iitd	0.121	0.141	0.808	0.218	0.219	0.724
Iitd-casia5a	0.062	0.185	0.739	0.049	0.117	0.830
Iitd-casia4i	0.299	0.319	0.569	0.315	0.584	0.045

Table: Segmentation scores for the linear-based domain adaptation method against the baseline (source-target) results

[1] http://biometrics.idealtest.org
[2] http://www4.comp.polyu.edu.hk/ csajaykr/database.php
[3] http://www.biometrics.idealtest.org
[4] http://nice1.di.ubi.pt/dates.htm

Linear Adaptation Experiment

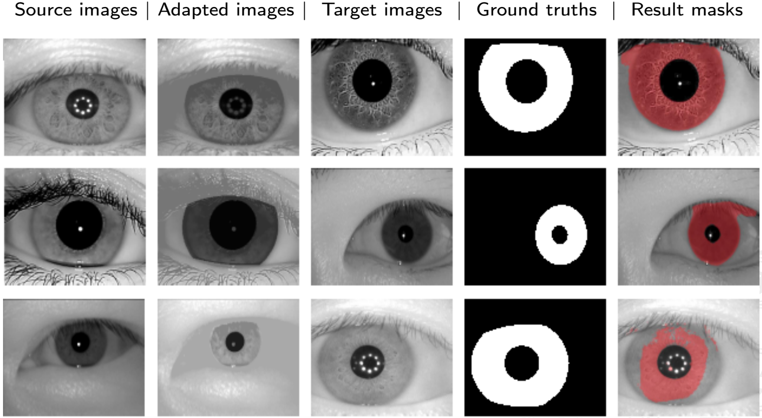

Figure: Sample adapted images and their corresponding segmentation results for Casia4i-Iitd (first row), Iitd-Casia5a (second row), and Casia5a-Casia4i (third row) database pairs (source-target) using the linear domain adaptation method

Non-linear intensity transfer

- In the linear adaptation, all the source intensity ranges get normalized to "a single average intensity range of that region in the target database"
- The target intensity ranges follow a non-linear distribution
- To address this, calculate the mean of corresponding minimum values and apply kernel smoothing regression on the data to get a polynomial function $f(x)$:

$$f(x) = p_1 x^n + p_2 x^{n-1} + ... + p_n x + p_{n+1}. \tag{4}$$

- Select a min for each sample, and estimated the max using the polynomial

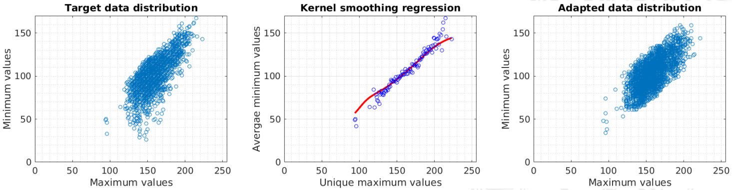

Figure: Sample non-linear data adaptation steps

Non-linear Damion Adaptation

Method	Adapted-target (NB)			Baseline(Source-target)		
Scores	nice1	nice2	f1	nice1	nice2	f1
Casia5a-casia4i	0.274	0.353	0.098	0.292	0.640	0.003
Casia5a-iitd	0.266	0.305	0.498	0.229	0.221	0.473
Casia4i-casia5a	0.027	0.074	0.859	0.274	0.406	0.341
Casia4i-iitd	0.102	0.095	0.812	0.218	0.219	0.724
Iitd-casia5a	0.034	0.088	0.813	0.049	0.117	0.830
Iitd-casia4i	0.208	0.174	0.374	0.315	0.584	0.045

Table: Segmentation scores for the non-linear-based domain adaptation method against the baseline (source-target) results

Non-linear Damion Adaptation

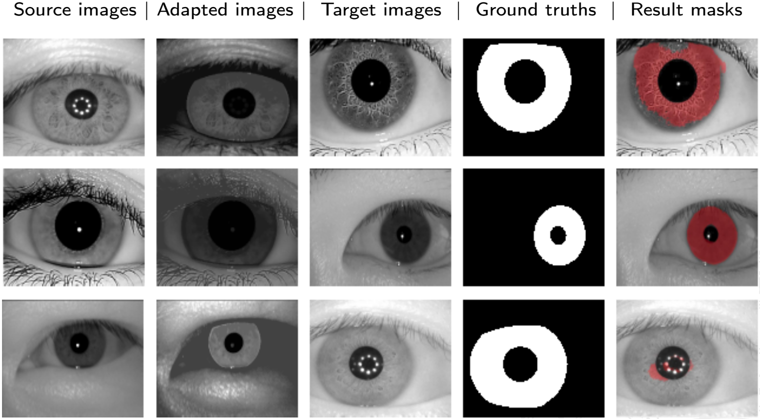

Figure: Sample adapted images and their corresponding segmentation results for Casia4i-Iitd (first row), Iitd-Casia5a (second row), and Casia5a-Casia4i (third row) database pairs (source-target) using the non-linear domain adaptation method

Alternative Approach

Training with minimum labeled data

- With the aim of minimizing the number of labeled data required to train the FCN, and maintaining optimal segmentation scores
- Decreased the number of labeled samples required to train the FCN stepwise
- Tested the FCN on the corresponding databases

Database	Casia5a			Casia4i			Iiitd		
Score	nice1	nice2	f1	nice1	nice2	f1	nice1	nice2	f1
15 pcs	0.075	0.082	0.875	0.205	0.263	0.502	0.089	0.097	0.856
25 pcs	0.064	0.077	0.896	0.099	0.115	0.814	0.077	0.083	0.879
50 pcs	0.050	0.070	0.909	0.078	0.068	0.841	0.063	0.070	0.889
100 pcs	0.021	0.040	0.921	0.038	0.039	0.926	0.035	0.037	0.941

Method	target -target		
Scores	nice1	nice2	f1
Casia5a-casia5a	0.019	0.038	0.925
Casia4i-casia4i	0.033	0.038	0.937
Iiitd-iitd	0.027	0.032	0.951

- Optimal segmentation scores can be achieved using (ap) 100 training samples
- Slightly lower, but very close scores can be achieved with 50 to 25 samples

Conclusion and Future Work

Conclusion

- We proposed two domain adaptation methods for CNN based iris segmentation
- Feature representations affecting the weights during training process are not limited to tonal distributions, and further features such as geometric properties of iris, non-iris, and pupil are definitely affecting this process
- Tonal distribution (intensity ranges of iris, non-iris, and pupil) plays a key role in generalization of FCNs on new iris data that differs from the training data

- **Future work**
  - We will investigate the relations between the two proposed methods and the reasons for the different results
  - We also explore more feature representations which encourage further distinctions between two domains, hoping to be able to develop a more comprehensive domain adaptation method

Thank you, Remarks?

References

Ehsaneddin Jalilian and Andreas Uhl.
Iris segmentation using fully convolutional encoder–decoder networks.
In *Deep Learning for Biometrics*, pages 133–155. Springer, 2017.

Nianfeng Liu, Haiqing Li, Man Zhang, Jing Liu, Zhenan Sun, and Tieniu Tan.
Accurate iris segmentation in non-cooperative environments using fully convolutional networks.
In *Biometrics (ICB), 2016 International Conference on*, pages 1–8. IEEE, 2016.

Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla.
Segnet: A deep convolutional encoder-decoder architecture for image segmentation.
arXiv preprint arXiv:1511.00561, 2015.