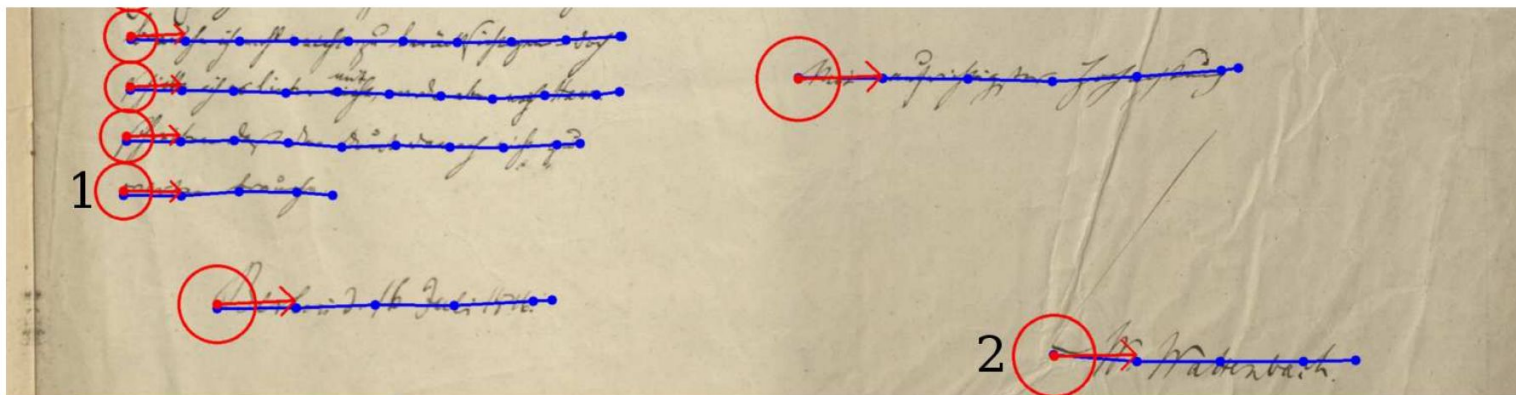# Transfer Learning For Handwritten Document Processing

Eric Burdett
MS Student - BYU

Start-Follow-Read

- End-to-End Full-Page Handwriting Recognizer [3]
  - Start of Line
  - Line Follower
  - Recognition
- Won 2017 ICDAR Competition on Handwritten Text Recognition

warten brauche.

W Wattenbach.

Start-Follow-Read - Does it Generalize?

1: Thursday, May 9, 1889
2: Went to Salt Lake to attend a
3: party given a Eldridge's. There was
4: present kate a Celia Sharp Katie
5: B Young Mel Sharp, Lottie and
6: Georgie Webber, Mose Thatcher and girl
7: Walt. Jennings, mr Teasdale. and others
8: It fell to my lot to take the
9: Webber's home.
10: Stayed at Eldriges that eve.

1: TUPSAAY,
2: Went to salt lake to attend a
3: parlyy gione a Adridgio. There was
4: purent Nate Celia Pharf lalie
5: B. Youngmel Charf, Loe
6: Beorgie Welhr, Mon Thatcher md giel
7: Walt, Zinmngs, Mr. Seardeli and others
8: lo fell to my lot to tatre the
9: Weblrrs home.
10: Stayed as Elaridges that en.
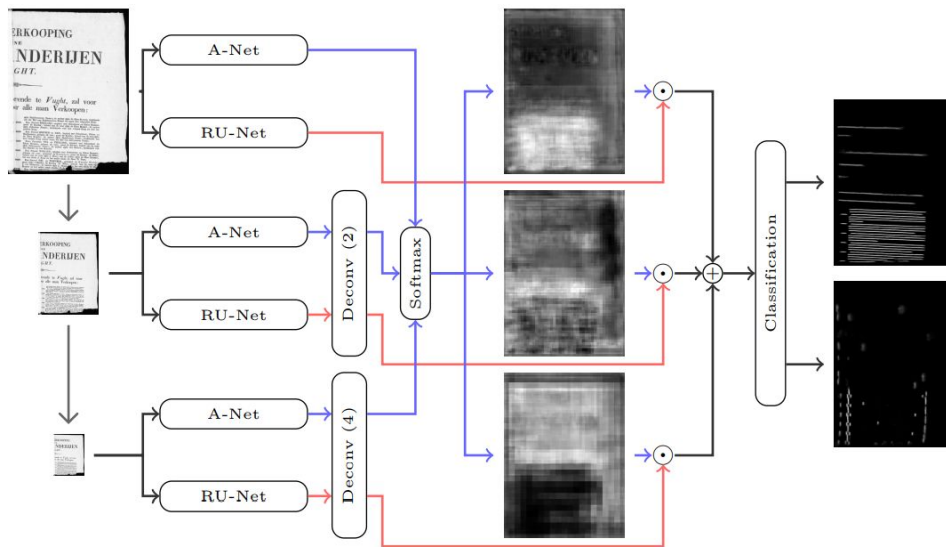
Start-Follow-Read - Does it Generalize?

0: Airi
1: D
2: Chaloge &
3: B

ARU-Net

- State-of-the-Art Baseline Detection [4]
  - Deep U-Net (with residual units)
  - Spatial Attention Mechanism
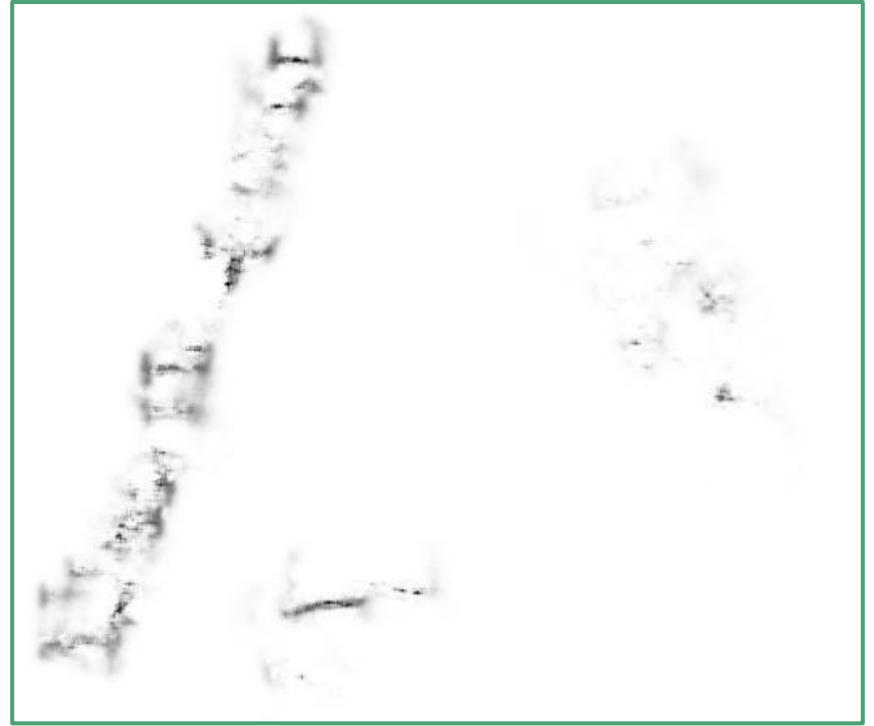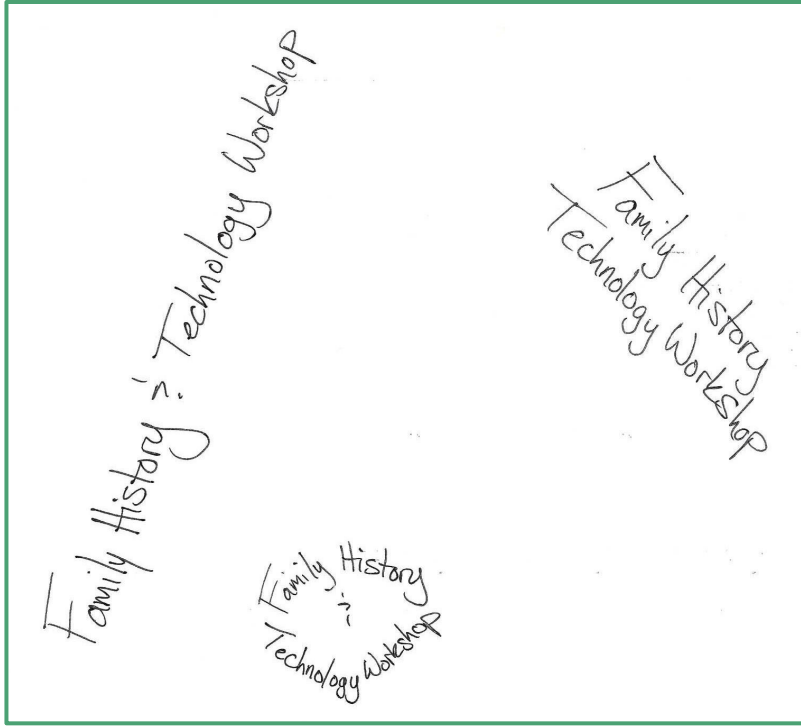- Winner of the 2019 ICDAR Competition on Baseline Detection

ARU-Net - Does it Generalize?

ARU-Net - Does it Generalize?

The Point

- Incredible performance with enough labeled data
- Performance decreases as target domain differs from source domain
- Labeling data is costly
- Where do we go from here?

Transfer Learning

- The process of utilizing knowledge gained from one task and applying it to another related problem.
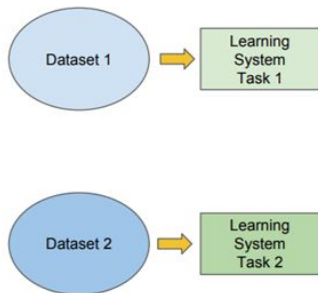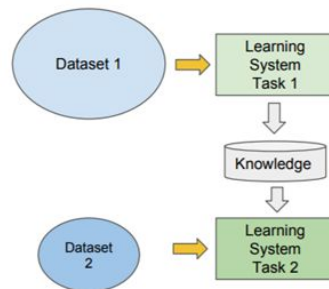
[12]

Types of Transfer Learning

[7]
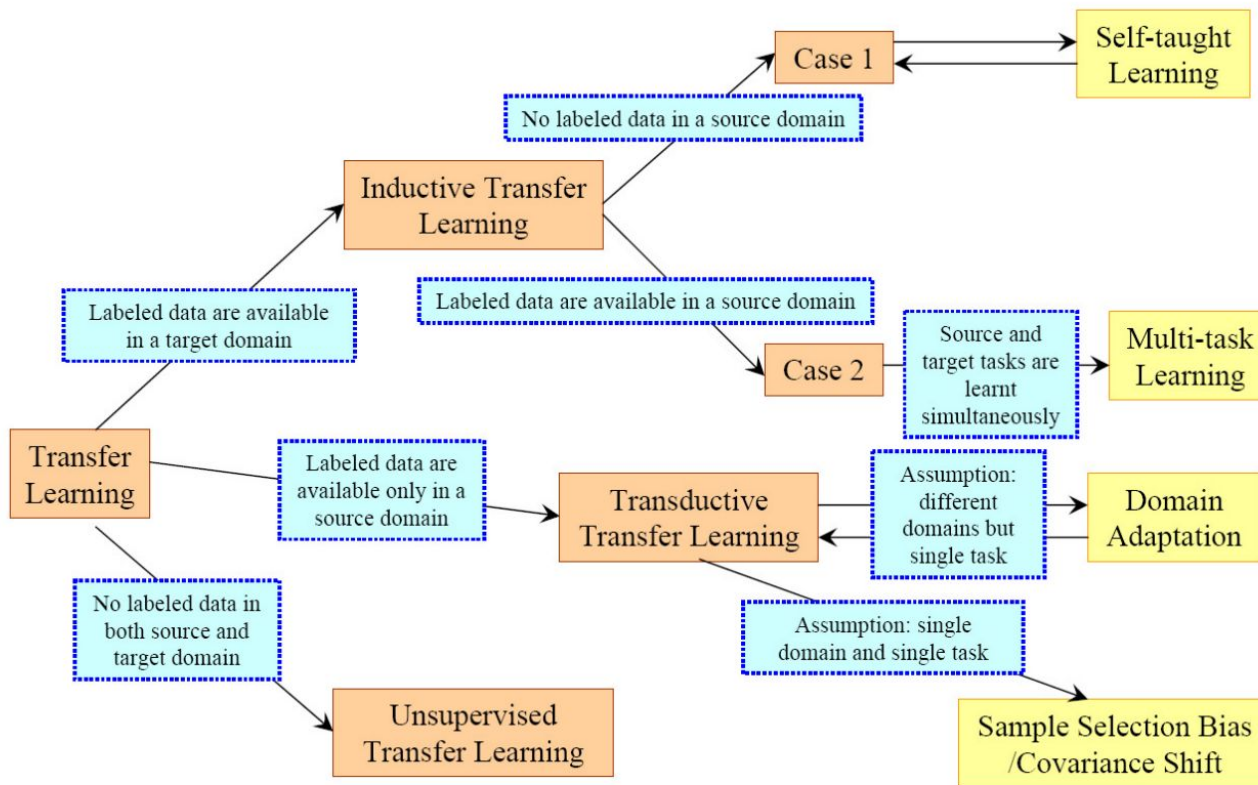
Inductive Transfer Learning

- Labeled data in source and target domains.
- Fine-tune on pretrained model
- Potential Benefits
  - Better Accuracy
  - Faster Training
  - Fewer Labeled Data in Target Domain

[7]
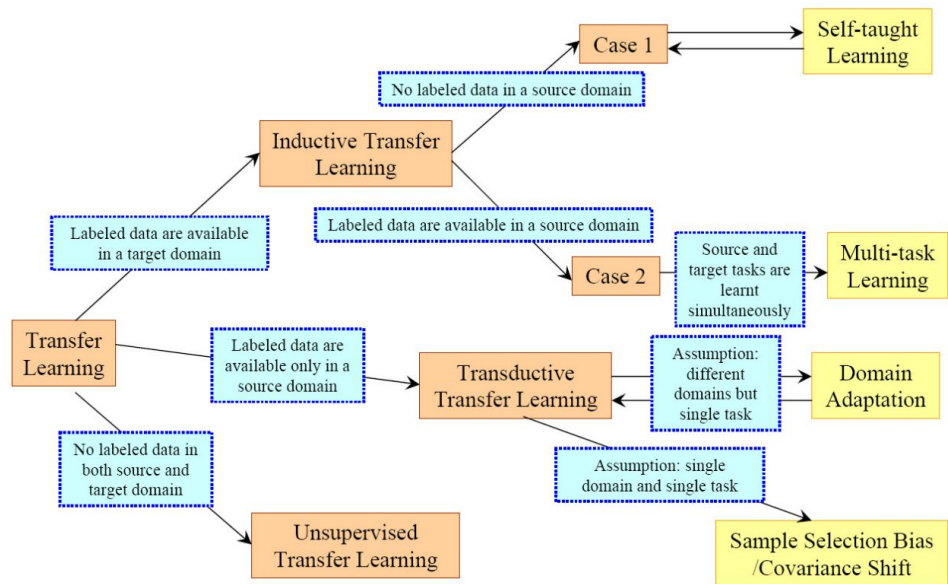
Transductive Transfer Learning

- Labeled data in source, Unlabeled data in target
- Access to unlabeled target data during training
- Potential Benefits
  - Better accuracy
  - Less/No labeled data needed in target domain
  - Align the feature representations in the source and target domains

[7]
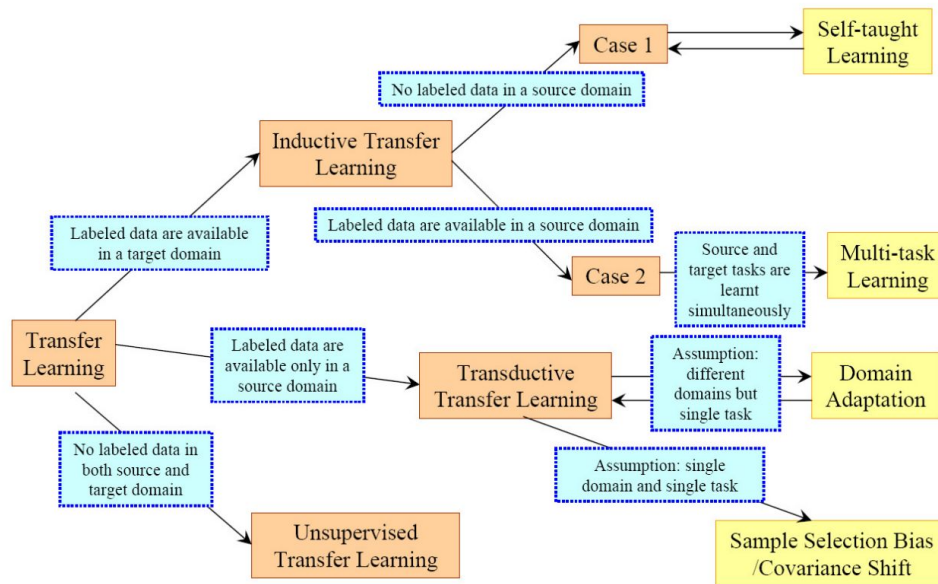
Feature Representation Transfer

- Identify good feature points that apply to both the source and target domain

[10]

Feature Representation Transfer

Labeled Data   Unlabeled Data

[10]

Domain Adversarial Training

$$\frac{\partial L_y}{\partial \theta_f}$$

$$-\lambda \frac{\partial L_d}{\partial \theta_f}$$

feature extractor $G_f(\cdot; \theta_f)$

gradient reversal layer

features $\mathbf{f}$

input $\mathbf{x}$

$$\frac{\partial L_y}{\partial \theta_y}$$

loss L_y

class label y

label predictor $G_y(\cdot; \theta_y)$

domain classifier $G_d(\cdot; \theta_d)$

domain label d

$$\lambda \frac{\partial L_d}{\partial \theta_d}$$

loss L_d

forwardprop

$$\frac{\partial L_y}{\partial \theta_f}$$

backprop (and produced derivatives)

SYN Numbers ➜ SVHN

Blue ➜ Source Activations

Red ➜ Target Activations
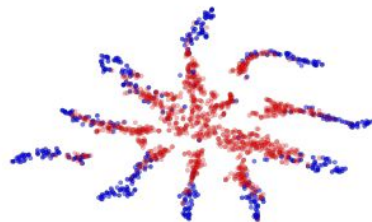
(a) Non-adapted

(b) Adapted

[9]

Domain Adversarial Training

[1]

CycleGAN

[5]
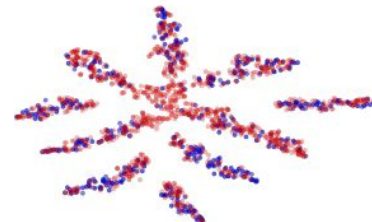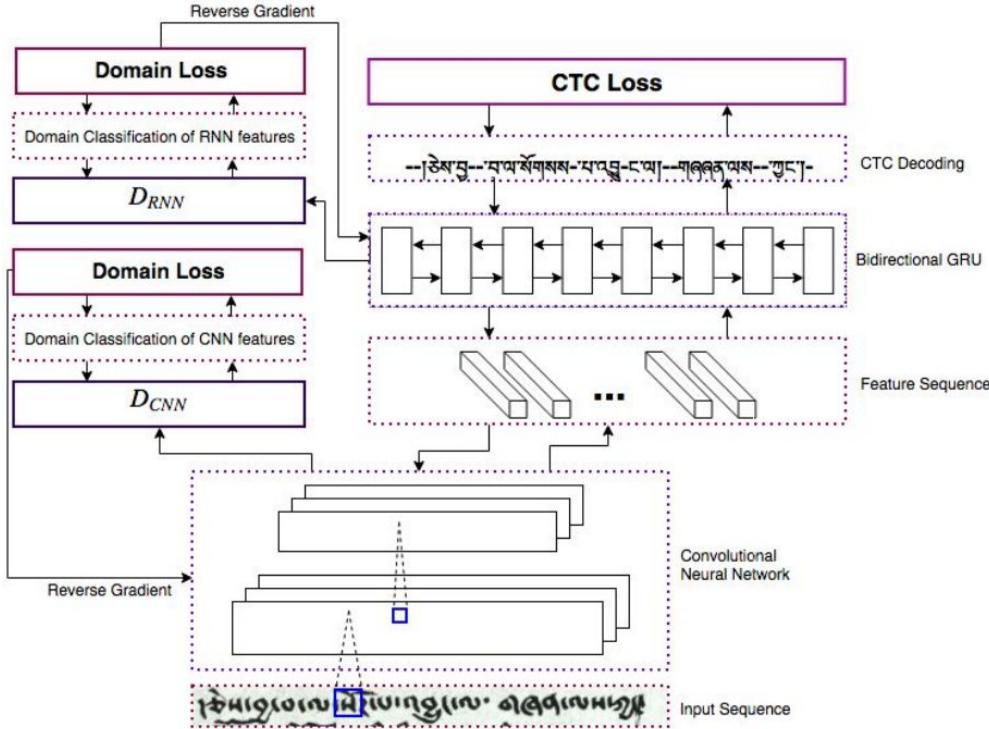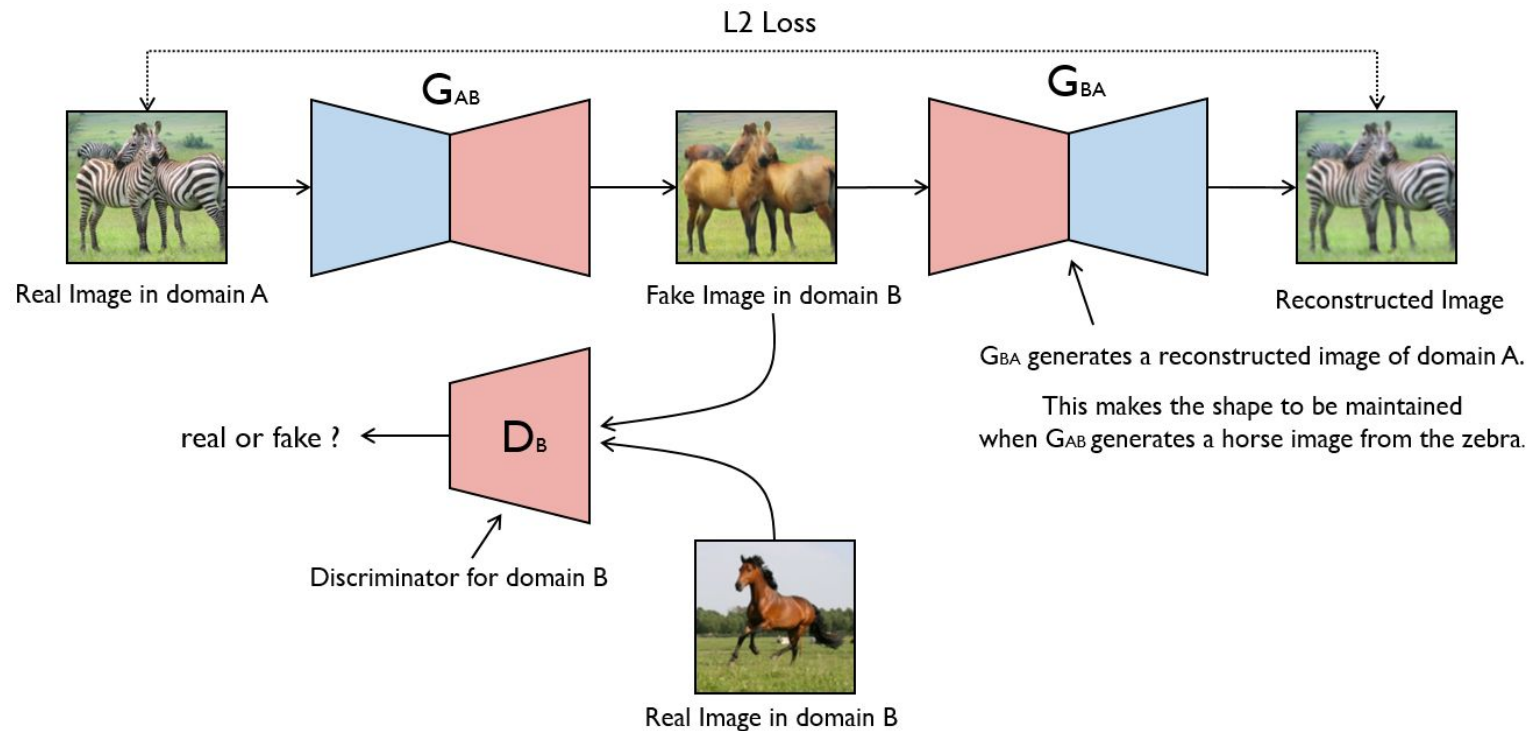
CycleGAN

CycleGAN
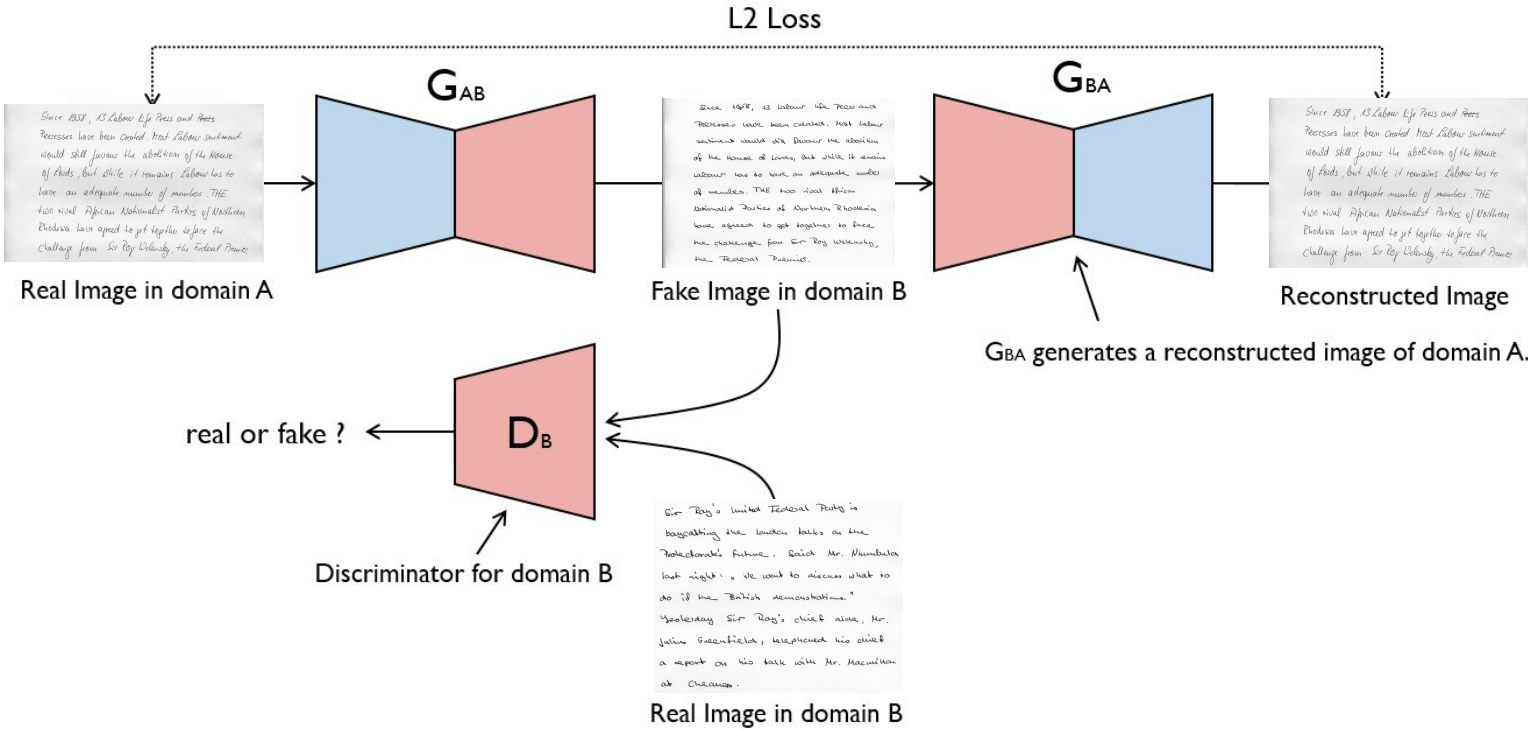
CycleGAN - Chinese Characters

SIMHEIM Font

Generated Characters

SIMHEIM Font

Generated Characters

[2]
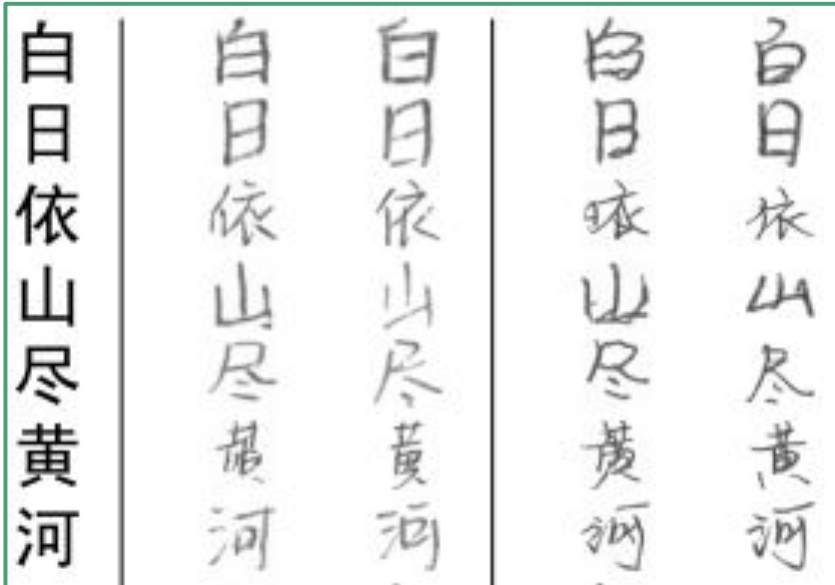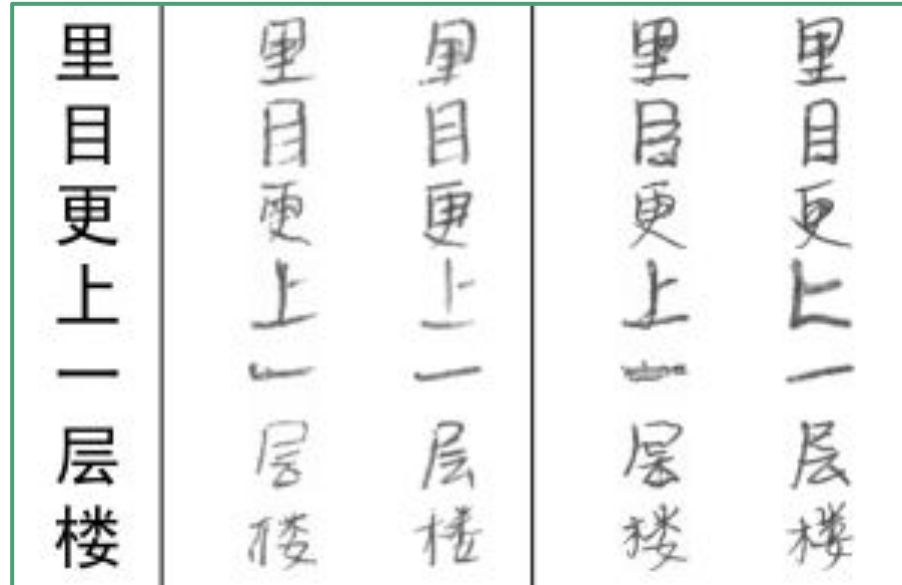
Other Transductive Transfer Learning Ideas

- Self-Supervised Learning [6]
    - Fine-Tune model on images from the target set that classified with high confidence
- Style-Transfer [11]
    - Apply handwriting style from target set to source set as pre-processing step

Looking Forward

- Expand on transductive transfer learning for handwriting recognition
- Apply these techniques using a source domain other than a system font
  - Tibetan Characters [1]
  - Chinese Characters [2]
- The Goal: Produce a system that utilizes the power of transfer learning to achieve good performance on unlabeled datasets

References

[1] S. Keret, L. Wolf, N. Dershowitz, E. Werner, O. Almogi and D. Wangchuk, "Transductive Learning for Reading Handwritten Tibetan Manuscripts," in *15th International Conference on Document Analysis and Recognition*, Sydney, Australia, 2019.

[2] B. Chang, Q. Zhang, S. Pan and L. Meng, "Generating Handwritten Chinese Characters using CycleGAN," in *Winter Conference on Applications of Computer Vision (WACV)*, Lake Tahoe, NV/CA, 2018.

[3] C. Wigington, C. Tensmeyer, B. Davis, W. Barrett, B. Price and S. Cohen, "Start, Follow, Read: End-to-End Full-Page Handwriting Recognition," in *European Conference on Computer Vision*, Munich, Germany, 2018.

[4] T. Gruning, G. Leifert, T. Straub, J. Michael and R. Labahn, "A Two Stage Method for Text Line Detection in Historical Documents," *International Journal on Document Analysis and Recognition (IJDAR),* vol. 22, no. 3, pp. 285-302, 2019.

[5] J.-Y. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," in *International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017.

[6] V. Frinken and H. Bunke, "Evaluating Retraining Rules for Semi-Supervised Learning in Neural Network Based Cursive Word Recognition," in *10th International Conference on Document Analysis and Recognition*, Barcelona, Spain, 2009.

References

[7] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering,* vol. 22, no. 10, pp. 1345-1359, 2010.

[8] J. Yosinski, J. Clune, Y. Bengio and H. Lipson, "How transferable are features in deep neural networks?," in *Advances in Neural Information Processing Systems (NIPS)*, Montreal, Canada, 2014.

[9] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand and V. Lempitsky, "Domain-Adversarial Training of Neural Networks," *Journal of Machine Learning Research,* vol. 17, no. 1, pp. 1-35, 2016.

[10] U. V. Marti and H. Bunke, "A full English sentence database for off-line handwriting recognition," in *Proceedings of the 5th International Conference on Document Analysis and Recognition*, Bangalore, India, 1999.

[11] R. Gomez, A. F. Biten, L. Gomez, J. Gibert, M. Rusinol and D. Karatzas, "Selective Style Transfer for Text," in *Proceedings of the 15th International Conference on Document Analysis and Recognition*, Sydney, Australia, 2019.

[12] D. Sarkar, "A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning," Towards Data Science, 14 November 2018. [Online]. Available: https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a. [Accessed 20 February 2020].

References

[13] R. Vijay, "Image-to-Image Translation using CycleGAN Model," Towards Data Science, 14 November 2019. [Online]. Available: https://towardsdatascience.com/image-to-image-translation-using-cyclegan-model-d58cfff04755. [Accessed 22 February 2020].