

Image To Image Translation for Domain Adaptation

[Zak Murez et al. Image to Image Translation for Domain Adaptation. CoRR , 2017.](#)

Prash Goel

Anirudh Ajith

Arihant Samar

Motivation



- CNNs work well when given large amounts of annotated data.
- For some domains, annotated data is expensive to create/obtain.
- Image Translation for Domain Adaptation attempts to ameliorate this problem by using labelled training data from one domain to generate realistic training data corresponding to another domain.

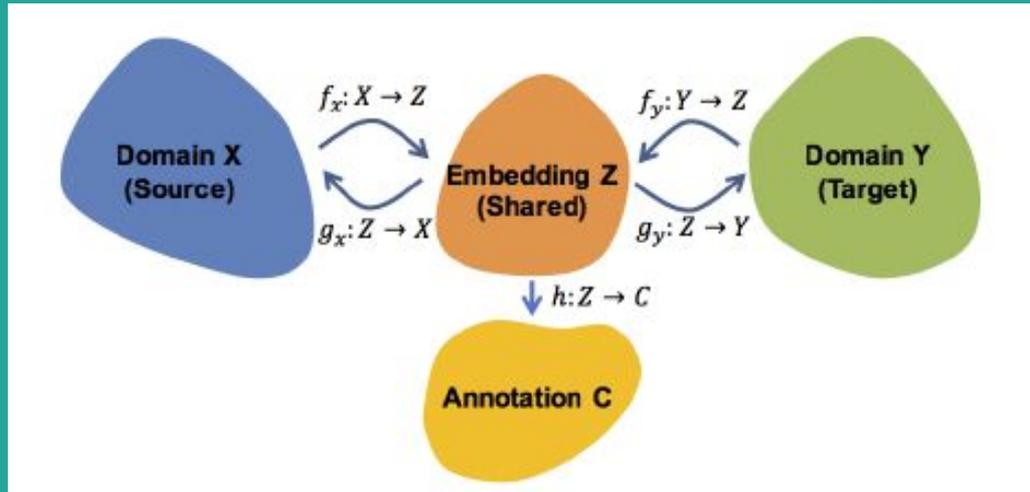


Method



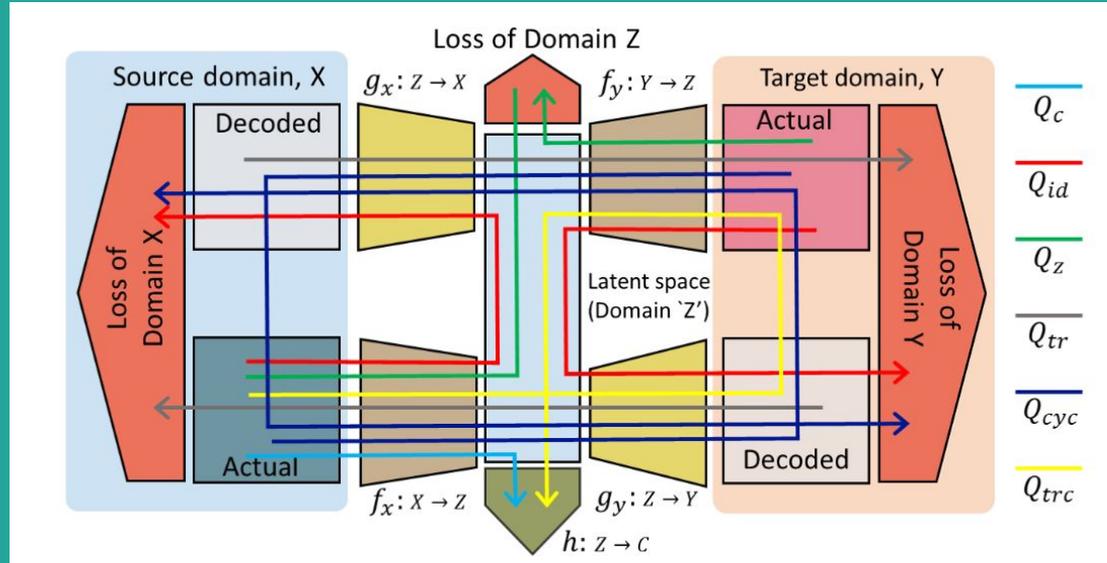
The authors of the paper propose to use an intermediate shared domain which facilitates the conversion between the source and the target domains.

They propose to use the shared domain representation as the input data for the required task.



The shared domain should be such that:

- The shared domain feature vectors should be domain agnostic
- It should be possible to reconstruct the image in the original domain from the shared domain.
- Cycle consistency should be maintained.



The models used are

- **Encoders** are CNNs which convert the images from source and target domains to a feature vector in the shared domain.
- **Decoders** reconstruct the source and target domain images from the feature vectors in the shared domain using deconvolutional layers.
- **Classifier** which maps the feature vectors in the latent space to the annotations.
- **Discriminator** which classifies whether a feature vector was generated from the source or the target domain.
- **Discriminators** which try to identify if an image is real or fake.

Multiple losses were used.

- A **cross-entropy loss** for classifier h to classify source images to the corresponding annotations.
- A pixel wise **image reconstruction loss** to ensure that reconstructed images are close to original images.
- **Discriminator cross-entropy losses** to correctly classify an image in one of the two classes as described above.
- **Cycle consistency loss** which are again pixel wise image reconstruction loss.
- A **classification loss** for classifier which classifies the feature vector generated from target domain (where the image in target domain is constructed from source domain images).

All these losses are added together after multiplying with suitable weightage hyperparameters as mentioned in the paper.

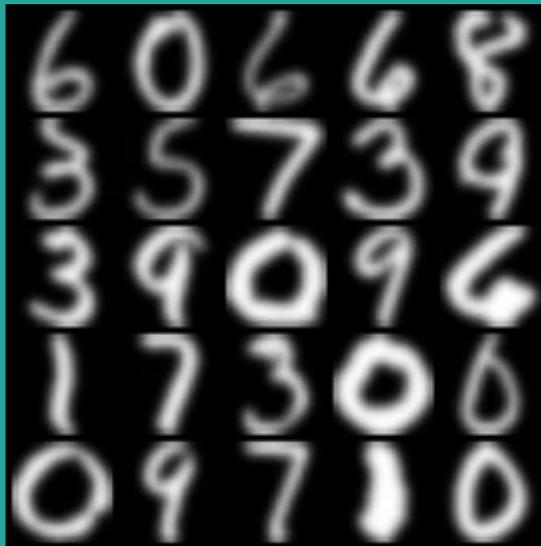
Datasets



Digits



MNIST



USPS



SVHN

Office



Amazon

Webcam

DSLR

Results





MNIST->USPS



USPS->MNIST



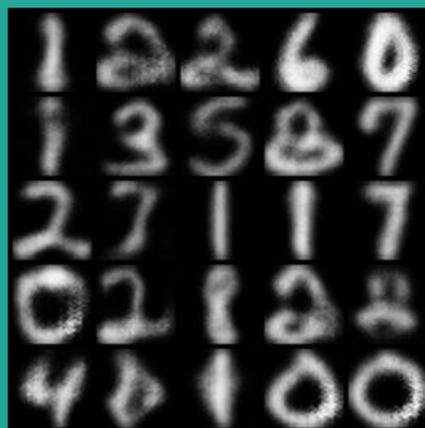
MNIST->SVHN



SVHN->MNIST



USPS->SVHN



SVHN->USPS



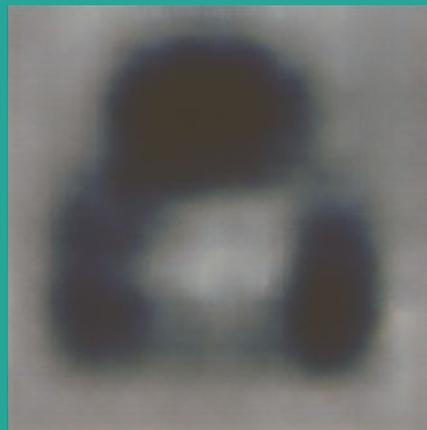
Original DSLR Domain image



DSLR -> Amazon



Original Amazon Domain image



Amazon -> Webcam

Method	Our implementation	Results from paper
Amazon \rightarrow Webcam	76.35	75.3
Webcam \rightarrow Amazon	52.21	52.1
Webcam \rightarrow DSLR	96.18	99.0
Amazon \rightarrow DSLR	69.9	71.1
DSLR \rightarrow Webcam	94.3	96
DSLR \rightarrow Amazon	47.2	50.0

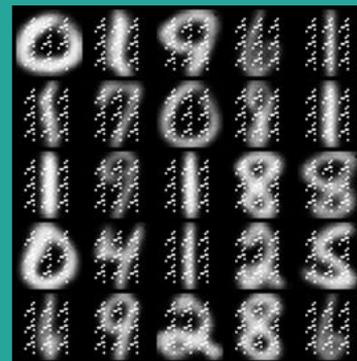
Limitations



- Heavy dependence on weight initialization
- Inability to generate proper images for Office dataset
- Bad performance when generating RGB images



USPS->MNIST



MNIST->USPS



Original DSLR Domain image



DSLR -> Amazon



Original Amazon Domain image



Amazon -> Webcam

Conclusion



- ✓ Implemented the framework from scratch
- ✓ Replicated their results on two datasets
- ✓ Discovered undisclosed limitations of the framework upon further experimentation

Thank you