IMAGE STYLE TRANSFER FOR BRAIN VESSEL SEGMENTATION USING MULTI-MODAL MRI

SEMESTER PROJECT – SPRING 2022

Daniele Falcetta
Daniele.falcetta@eurecom.fr

EURECOM
Sophia Antipolis

SUPERVISORS: PROFESSOR M. ZULUAGA – F. GALATI

EURECOM
Sophia Antipolis

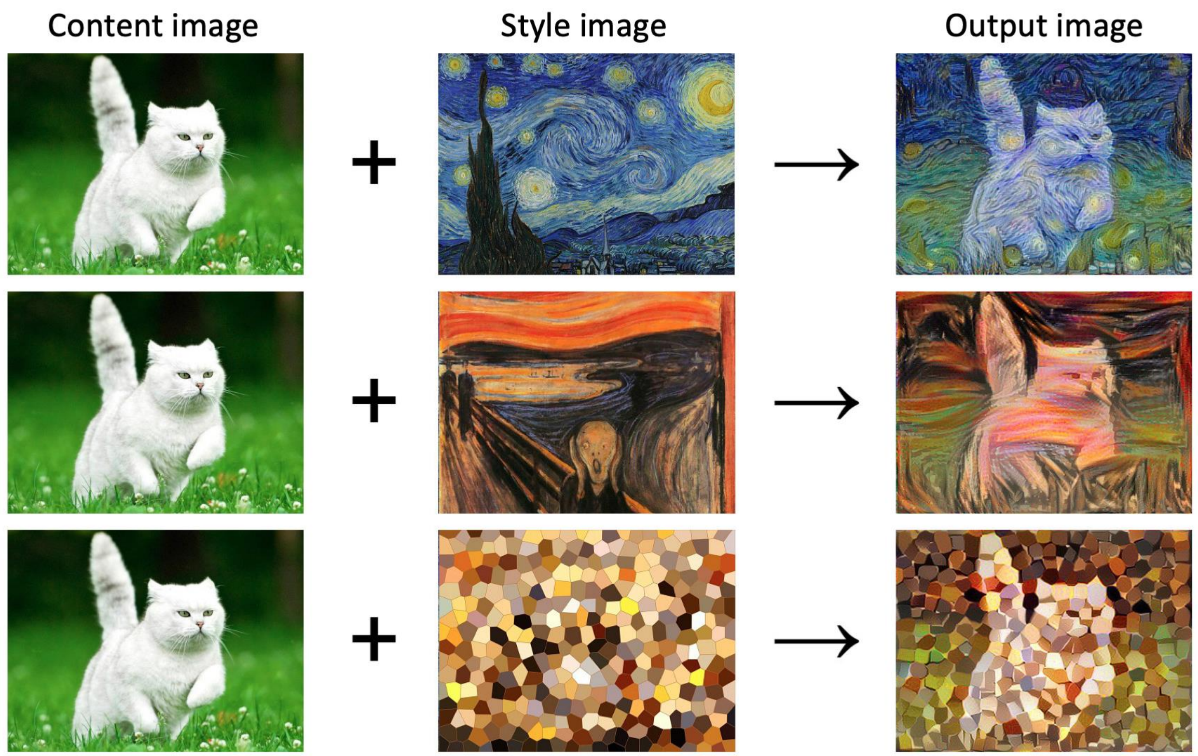
TABLE OF CONTENTS

IMAGE STYLE TRANSFER

INTRODUCTION

Task of **changing the style** of an image in one domain to the style of an image in another domain.

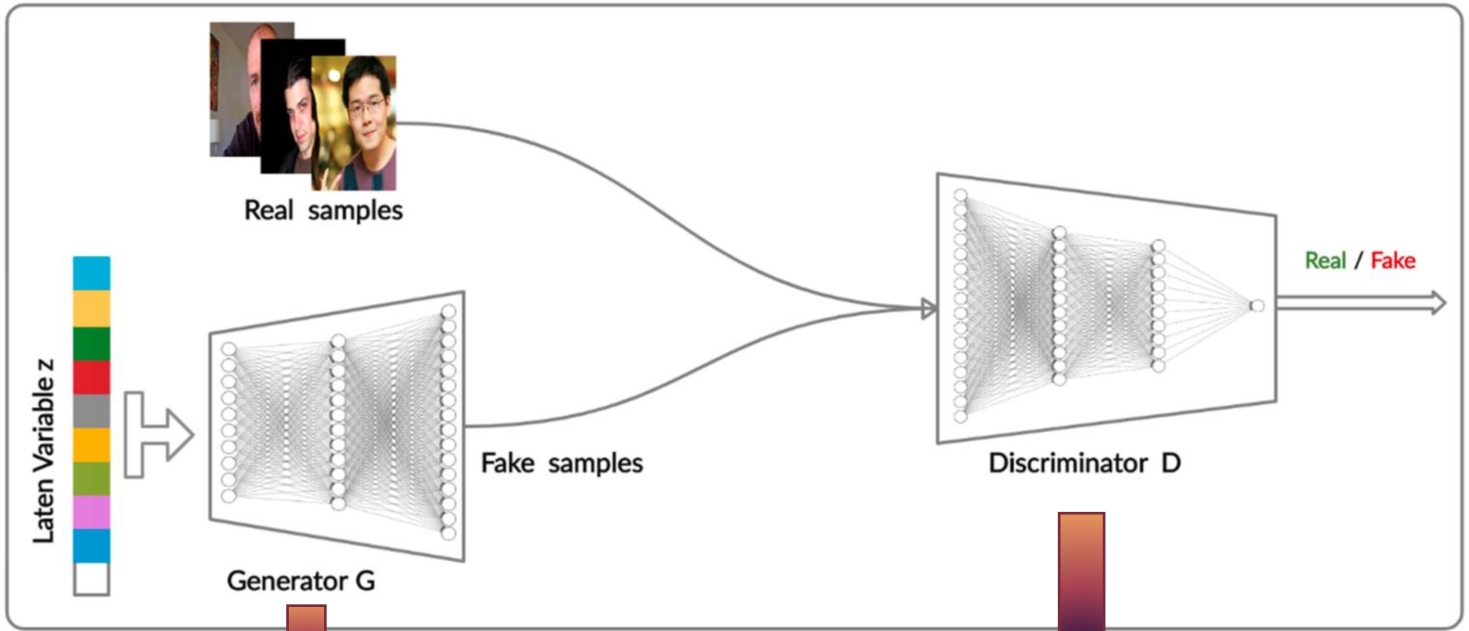
It takes **two input** images and **combines** them together in a way that the output image maintain the **structure** of the first image but appears to be ***painted*** in the style of the second one.



GENERATIVE ADVERSARIAL NETWORK

INTRODUCTION

- **Implicitly** learns to model the true distribution and generate sample.
- **Generator's Objective:** Fooling the Discriminator (making fake data look real)
- **Discriminator's Objective:** Distinguishing between real and fake images
- **Alternate Training/Optimization** between the two models *D* and *G* untill convergence
- **MinMax** Game: convergence at the ***Nash Equilibrium***



$$\mathcal{L}_{GAN}(G) = \mathbb{E}_{x \sim P_z(Z)}[log\,(1 - D(G(z)))]$$

$$\mathcal{L}_{GAN}(D) = \mathbb{E}_{x \sim P_{data}(x)}[\log D(x)] + \mathbb{E}_{x \sim P_z(Z)}[\log(1 - D(G(z)))]$$

CONDITIONAL GAN

INTRODUCTION

$$min\ max V(D,G) = \mathbb{E}_{x \sim Pdata(x)}[log\ D(x|y)] + \mathbb{E}_{x \sim Pz(x)}[log\ (1 - D(G(z|y)))]$$

$Real\ Data\ x$

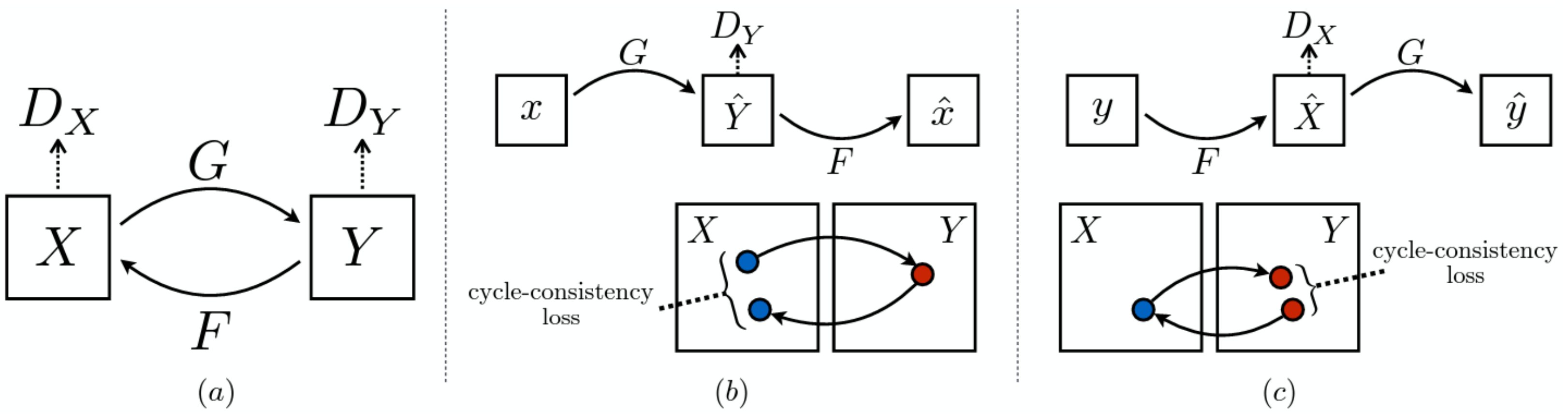
$Conditioning\ Input\ y$
$(e.g.\ \text{Image})$

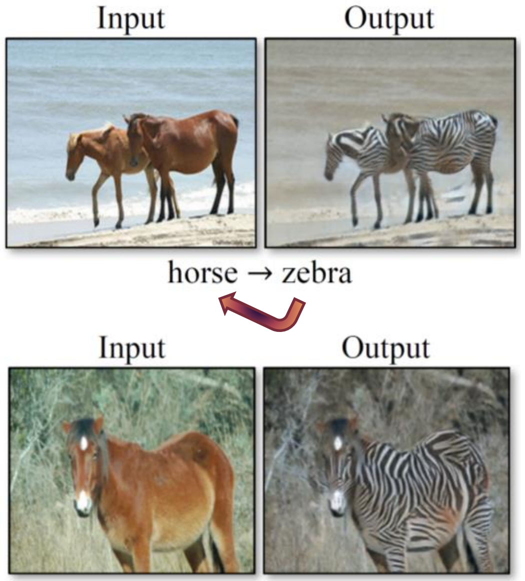
$Random\ Noise\ z$

Control the data generation process in a *supervised* manner.
Combines z and y into a joint hidden representation of real x, along with conditional variable y

CYCLE GAN – UNPAIRED TRAINING DATA

INTRODUCTION





Cycle Consistency Loss: Preserves original semantic content in the transformed image

$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1]$$

IMAGE SEGMENTATION

INTRODUCTION

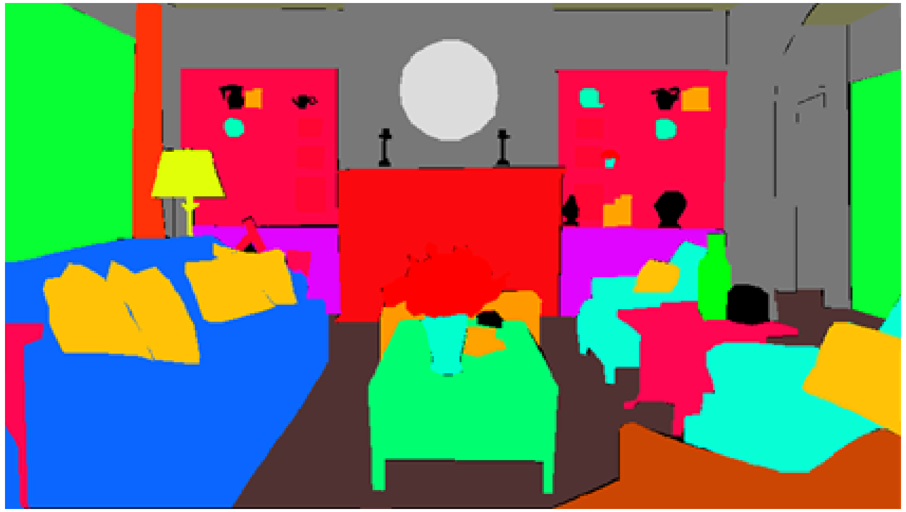
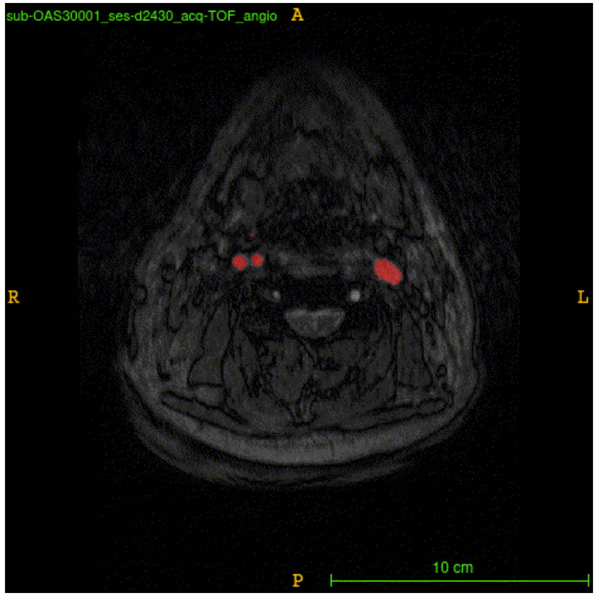


Image Segmentation is the Task of *clustering* together parts of an image that belong to the **same Object Class**.

The segmentation process is also called **Pixel-wise Classification**.
In other words, it involves partitioning images into multiple segments or objects.

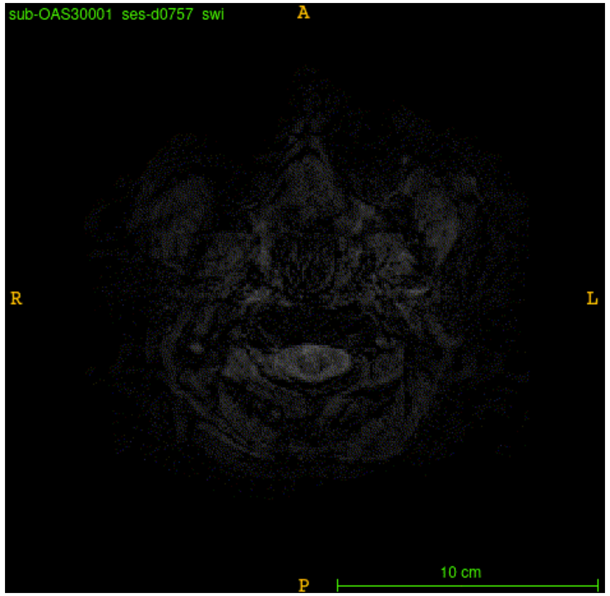
PROJECT GOAL

Investigate image style transfer techniques for brain vessel segmentation in Magnetic Resonance Images (MRI).



Time-Of-Flight Images (TOF)

- No Contrast needed
- Long acquisition Time ➔ Possible presence of Artifacts
- Possible signal loss from the vessels



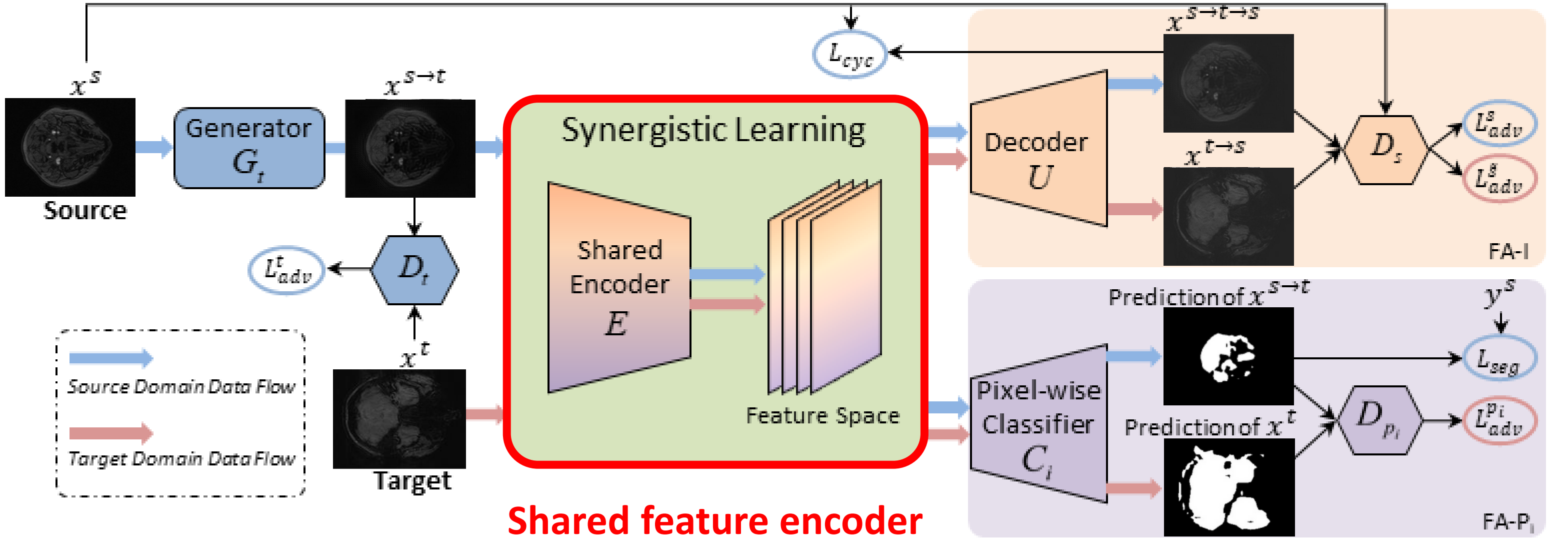
Susceptibility-Weighted Images (SWI)

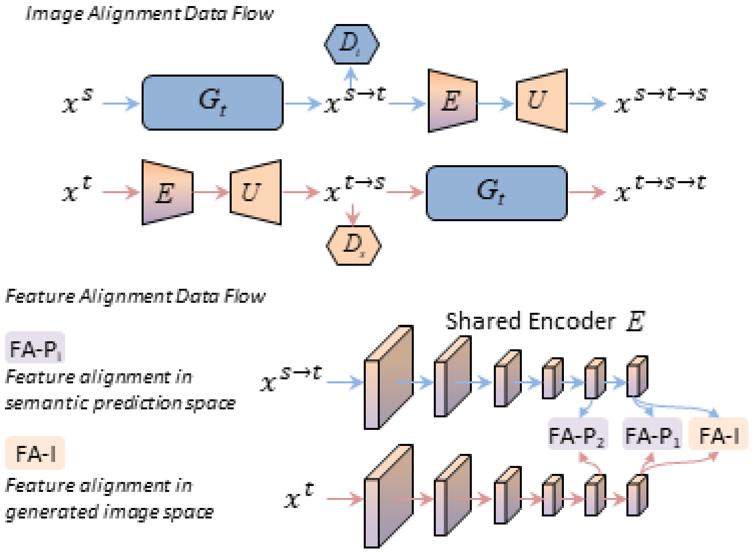
- Based on tissue magnetic susceptibility differences
- Require the injection of contrast
- Require post-processing

NETWORK ARCHITECTURE

SIFA

Synergic fusion of Adaptation from both Image and Feature perspective





Simultaneously, in one unified Framework:

1. **Transform** the appearance of images across domains
2. Enhance **Domain-Invariance** of the extracted features

1. Image Adaptation
2. Feature Adaptation

SIFA LOSSES

SIFA

- Proposed to overcome the **Vanishing Gradient problem** caused by the minmax loss in the original GAN
- **L2 Loss Function** instead of the CE one
- Allows the learning process to be more **stable**

$$\mathcal{L}_{\text{LSGAN}}(G) = \mathbb{E}_{x \sim P_z(Z)}\left[(D(G(z)) - c)^2\right]$$

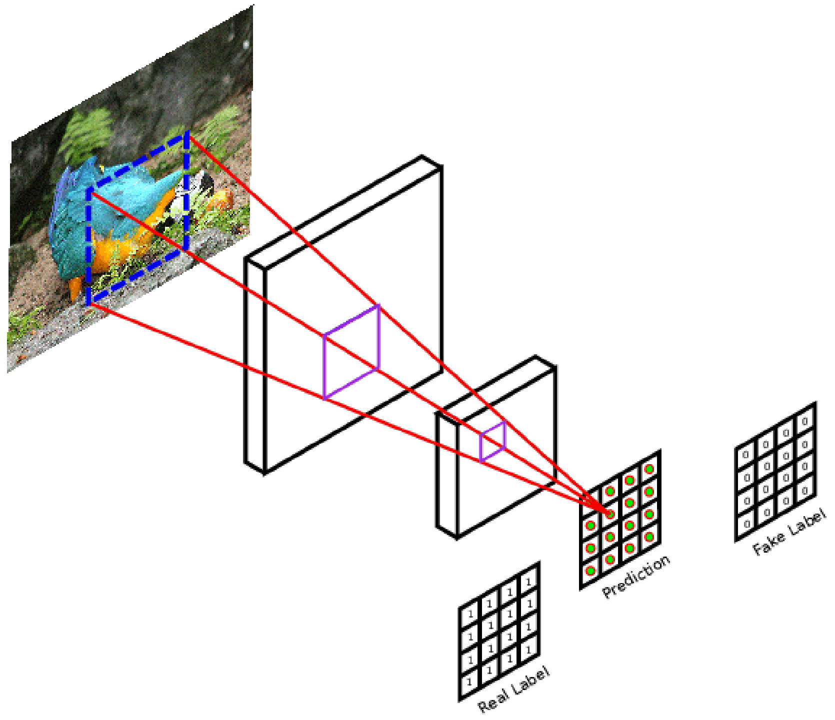
$$\mathcal{L}_{\text{LSGAN}}(D) = \mathbb{E}_{x \sim P_{data}(x)}\left[(D(x) - b)^2\right] + \mathbb{E}_{x \sim P_z(Z)}\left[(D(G(z)) - a)^2\right]$$

- **Hybrid** Loss
- Linear Combination of **Dice** and **Cross-Entropy** Loss
- $E \circ C$ serves as **Segmentation** Network for the Target
- It is trained using the sample pairs of $\{x^{s \to t}, y^s\}$

$$\mathcal{L}_{seg}(E, C) = H(y^s, \hat{y}^{s \to t}) + \alpha \cdot Dice(y^s, \hat{y}^{s \to t})$$

PATCH DISCRIMINATOR

SIFA



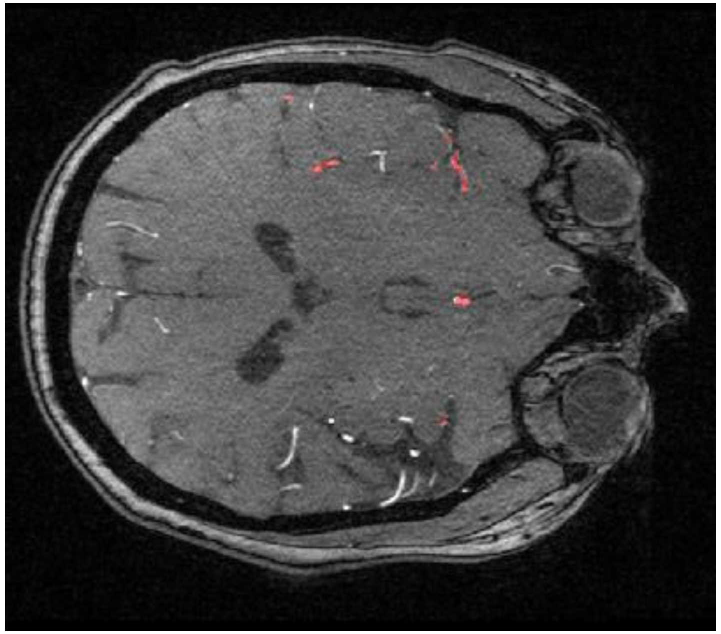
PatchGAN Discriminator

It does not work for the whole image but tries to classify each patch in an image as **Real** or **Fake**

This discriminator is run **convolutionally** across the image, **averaging** all responses to provide the ultimate output probability

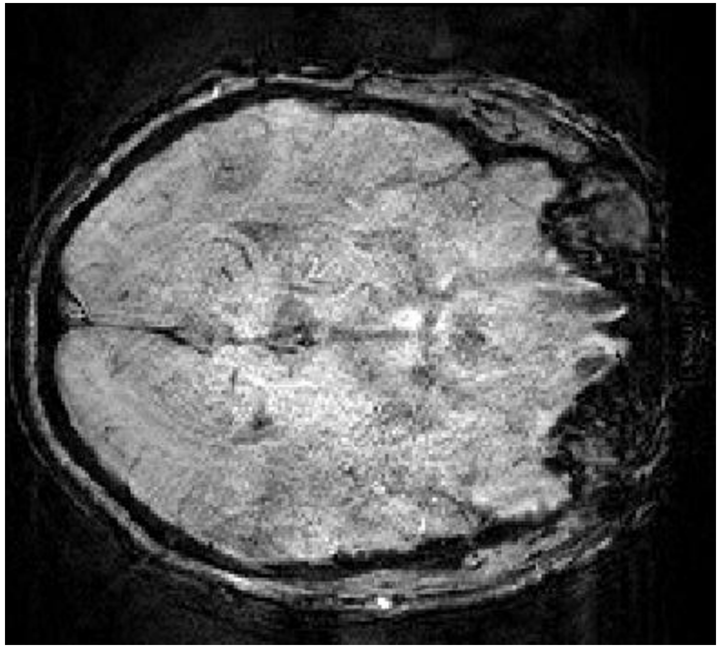
OUR DATASET

DATASET AND PREPROCESSING



- **TOF** Images as **Source** Domain
- **Size** (232, 576, 768)
- Average **Spacing** [0.59 0.29 0.29]
- **Both** Vessel and Brain **Mask**

≈20 **3D Images** for each Domain



- **SWI** Images as **Target** Domain
- **Size** (80, 192, 256)
- Average **Spacing** [2.00 0.89 0.89]
- **No** Segmentation **Mask**

PREPROCESSING

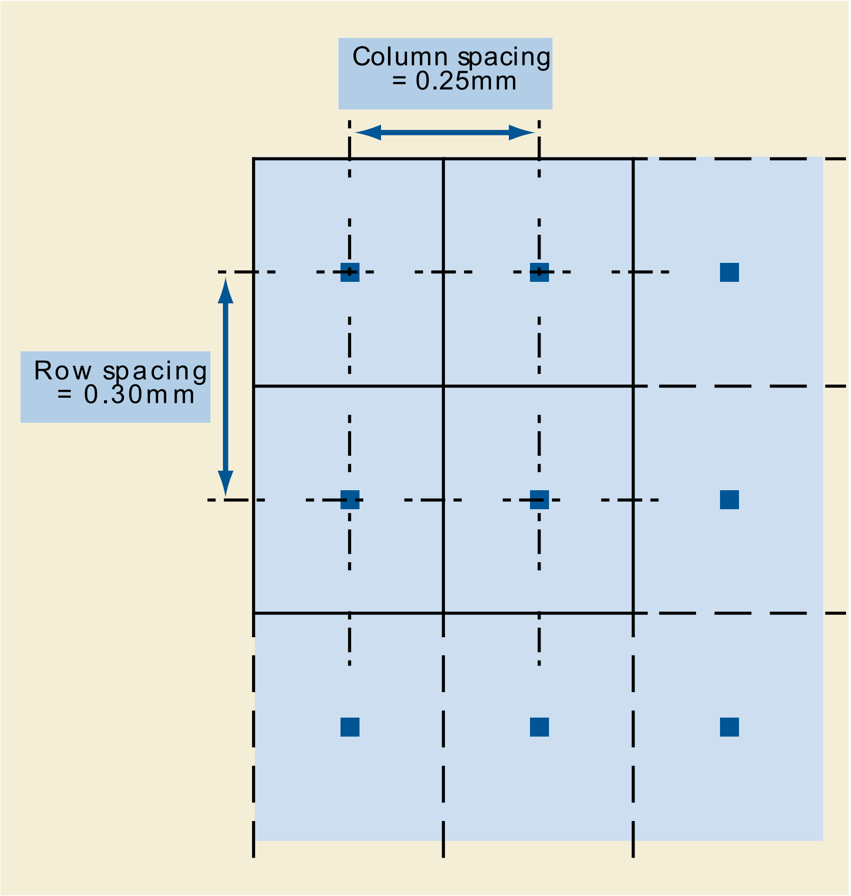
DATASET AND PREPROCESSING

- **Spacing Extraction:** Voxel spacing values indicate the real-world size or scale of each voxel in the 3D slice stack

  X and Y Spacings define the **distance** from the **center** of one pixel to the center of an adjacent one, in **millimeters**

- **Slicing** and **Reshaping** according to the SWI spacing

  Same X and Y spacings for any image



EURECOM
Sophia Antipolis

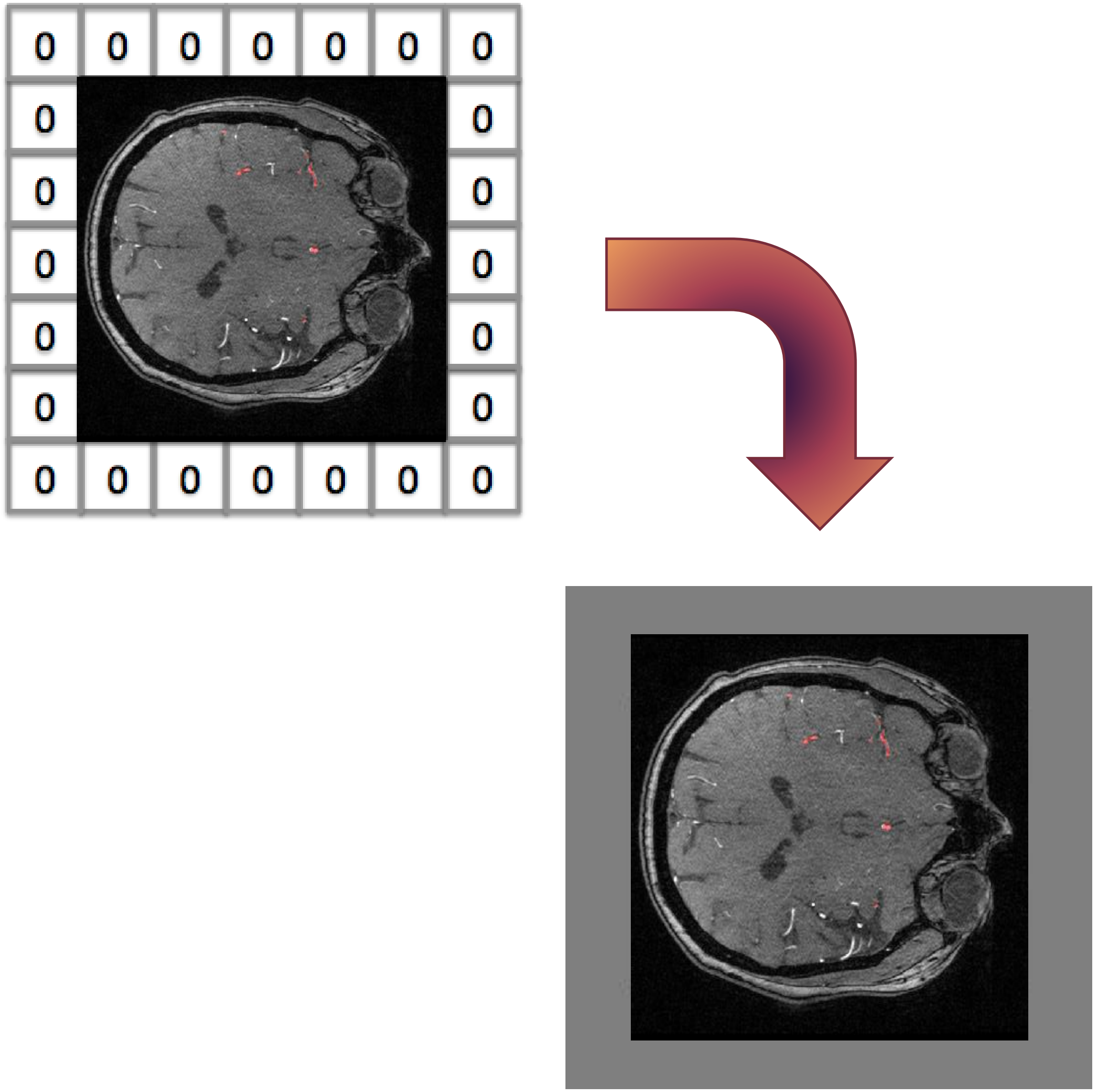
PREPROCESSING

DATASET AND PREPROCESSING

- **Standardization** using mean μ and std σ computed over the **Volume** of the slice
- **Normalization** using **Min** and **Max** computed over the whole Dataset
- Center **Crop**, **Padding**, **One-Hot Encoding**

  256x256x1 Images
  256x256x2 Masks

- Transformation into **Tfrecords**



SIFA OUTPUTS (IMAGES)

EXPERIMENTS AND RESULTS

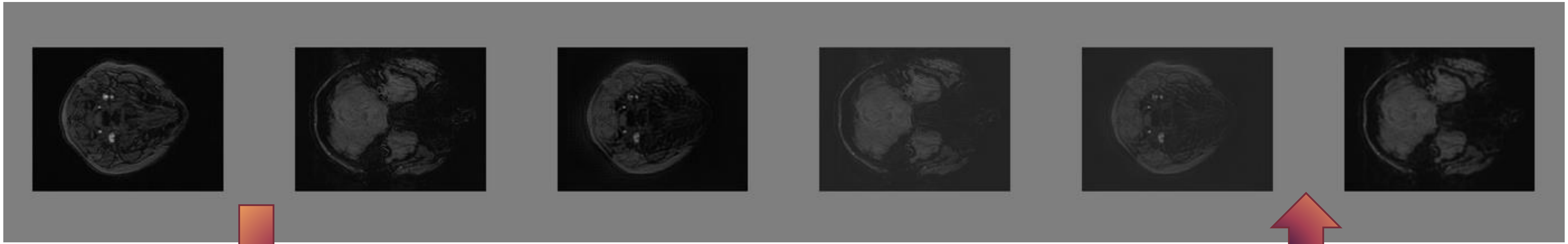
X_s : *Input from the* ***Source*** *Domain*

X_t : *Input from the* ***Target*** *Domain*

$X_{s\to t\to s}$: ***Reconstructed Source*** *Image*

$X_{t\to s\to t}$: ***Recontruscted Target*** *Image*

X_s X_t $X_{s\to t}$ $X_{t\to s}$ $X_{s\to t\to s}$ $X_{t\to s\to t}$



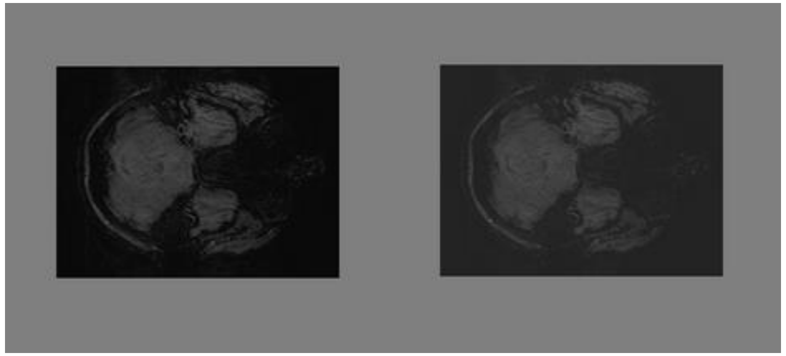
$X_{s\to t}$: *Translated Image from* ***Source*** *to* ***Target***

$X_{t\to s}$: *Translated Image from* ***Target*** *to* ***Source***

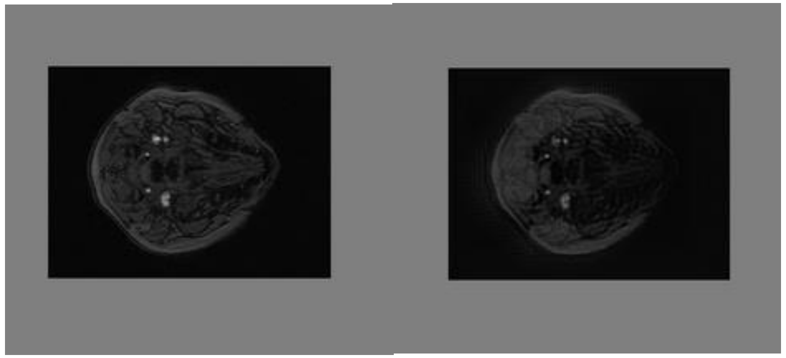
SIFA OUTPUTS (SEGMENTATION MASKS)

EXPERIMENTS AND RESULTS

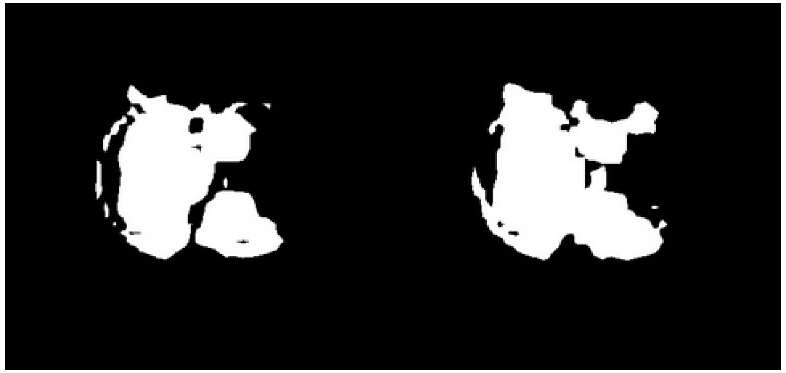
X_t $X_{t \to s}$



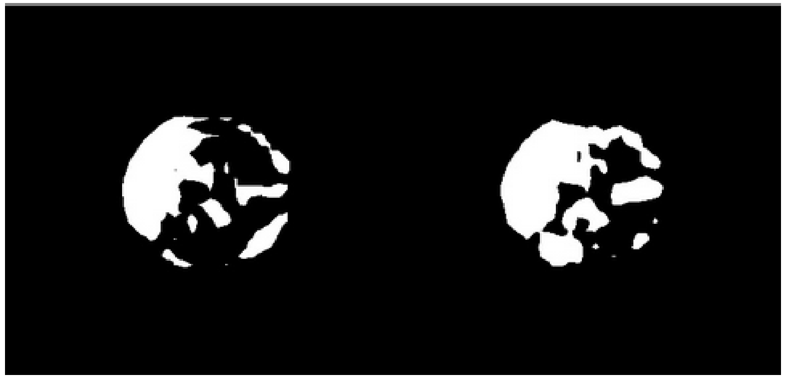
X_s $X_{s \to t}$



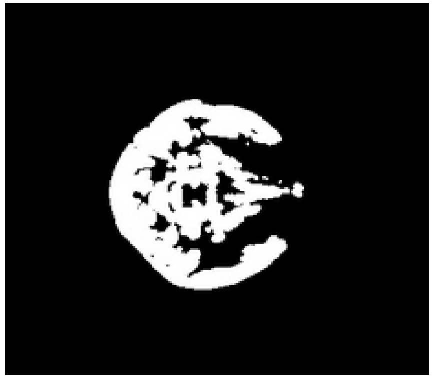
$Predicted\ Y_t$ $Predicted\ Y_{t \to s}$



$Predicted\ Y_s$ $Predicted\ Y_{s \to t}$



$Ground\ Truth\ Y_s$



EXPERIMENTS

EXPERIMENTS AND RESULTS

	Step #	Learning Rate	Segmentation parameter	Used Mask
Experiment 1#	40k	2e-3	1	Vessel
Experiment 2#	40k	2e-3	1	Brain
Experiment 3#	10k	1e-4	1	Brain

- **Batch size:** 4 Images for each domain (max allowed by the GPU)
- **Step:** Complete network optimization using a single batch as input
- **Segmentation parameter:** If 0, all the Seg Losses = 0 during the whole training (Cycle-GAN)
- **Used Masks:** either vessels or brain as segmentation masks

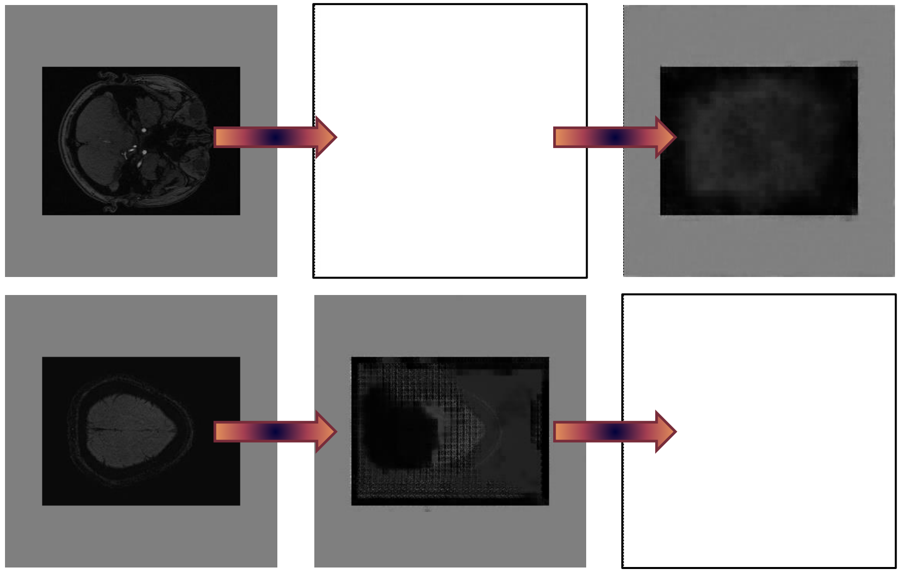
EXPERIMENT #1: VESSEL MASK (40K)

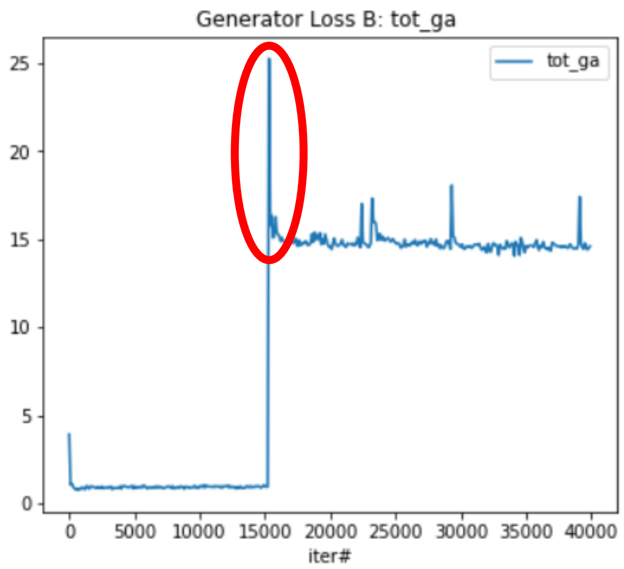
EXPERIMENTS AND RESULTS

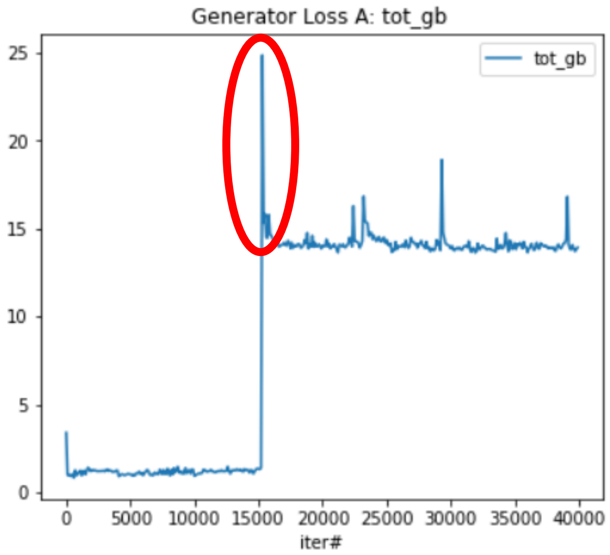
Iter >14999

S

T







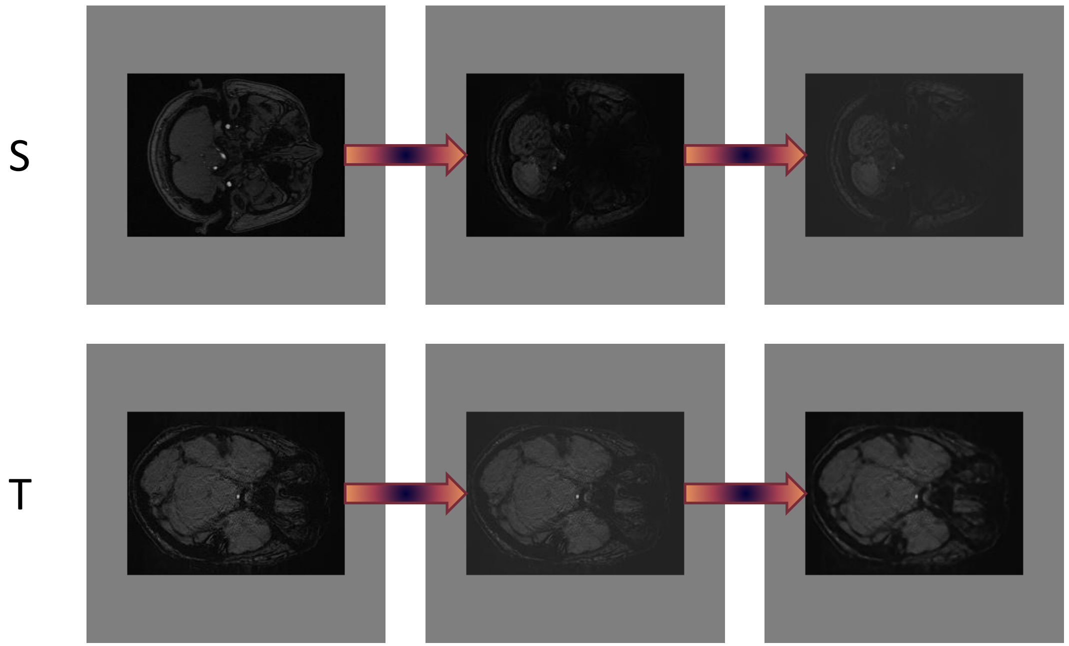
Predicted Y_s | *Predicted* $Y_{s \to t}$ | *Ground Truth* Y_s

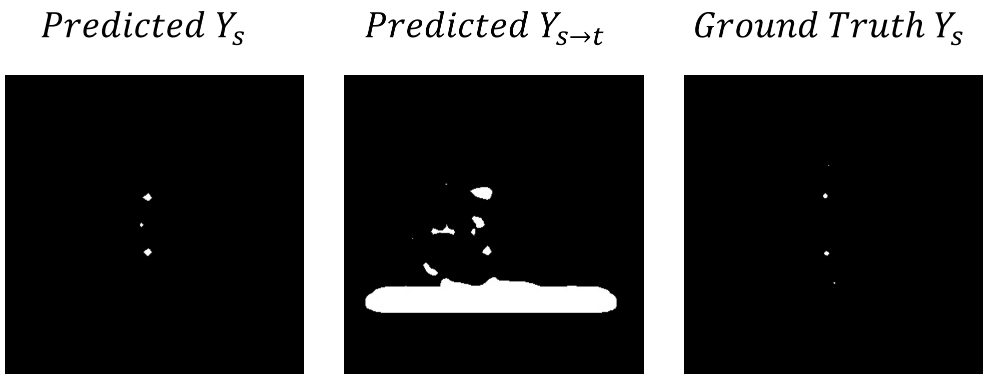
After **15k Iters** the Losses of the Generator rapdly grow up: **Discriminators win** the *game* and the Generators are not able to create new transformed images anymore nor segmentaion masks

EXPERIMENT #1: VESSEL MASK (40K)

EXPERIMENTS AND RESULTS

Iter 14999



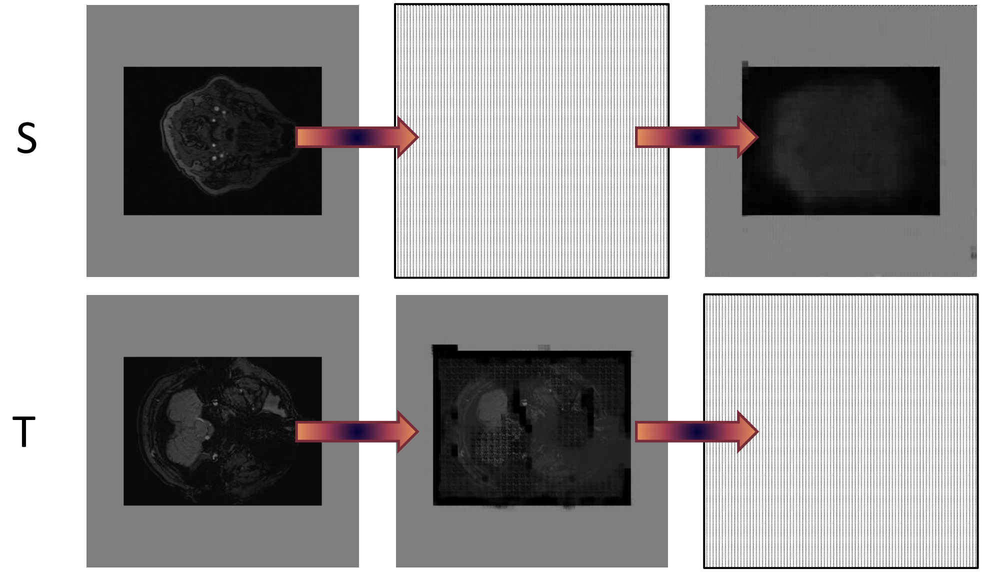


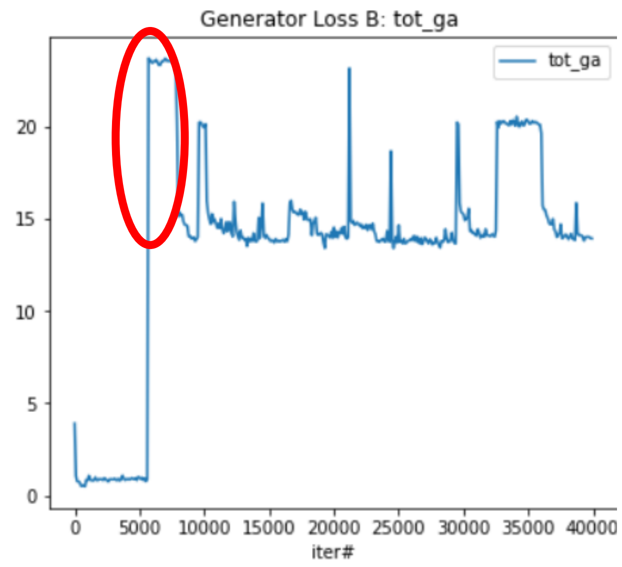
- Vessel are too tiny: the network is not able to segment and fails.
- Great imbalance between background and vessel

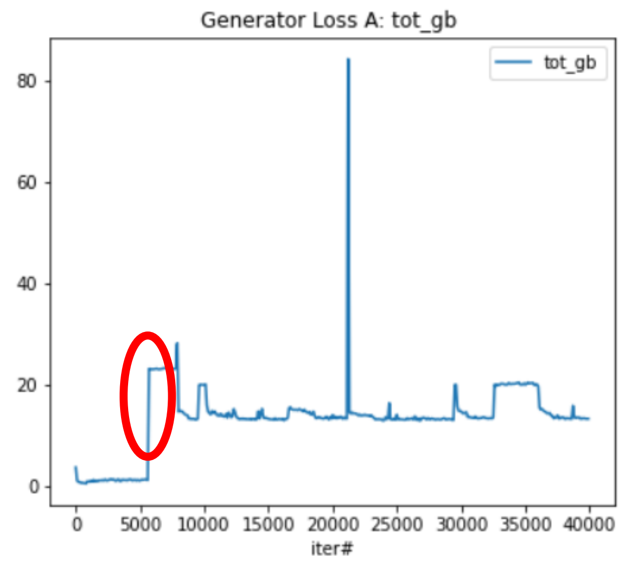
EXPERIMENT #2: BRAIN MASK (40K)

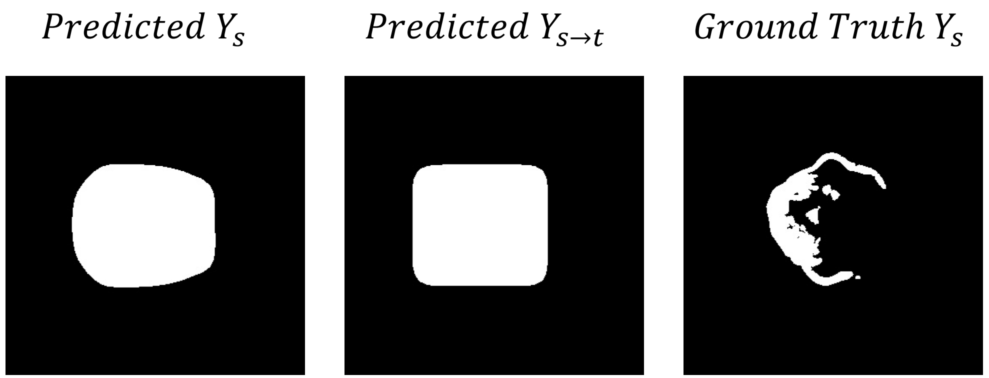
EXPERIMENTS AND RESULTS

Iter >5399









After just **5k Iters** the Generators' Losses increase fastly: Same phenomena as before

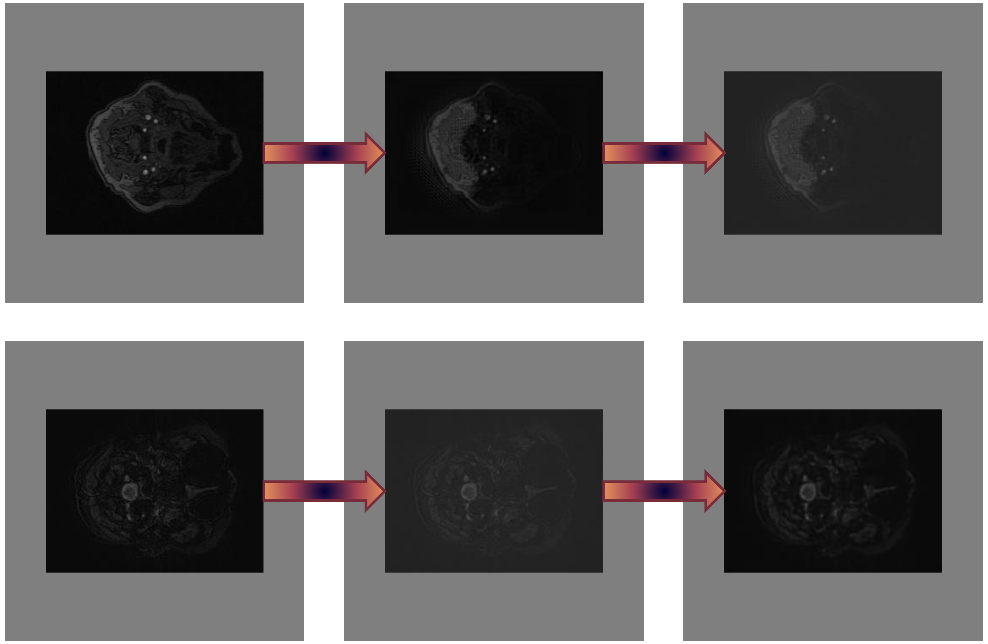
EXPERIMENT #2: BRAIN MASK (40K)

EXPERIMENTS AND RESULTS

Iter 5399

S

T

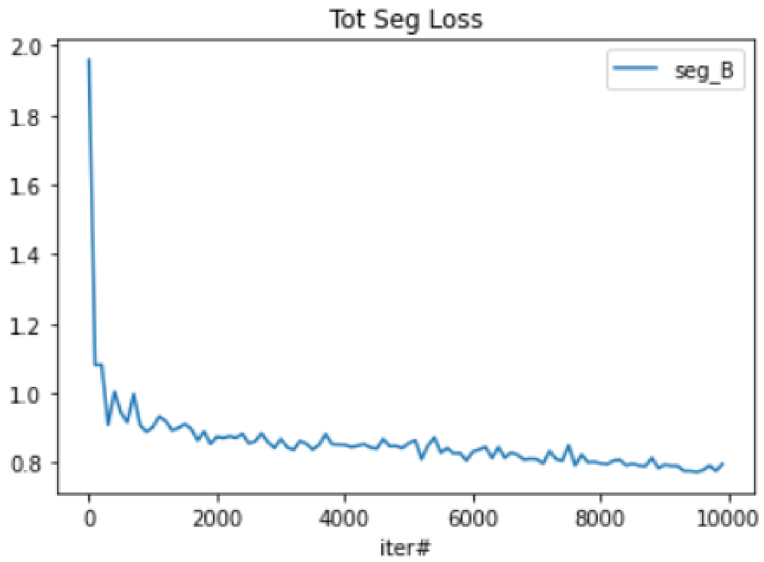


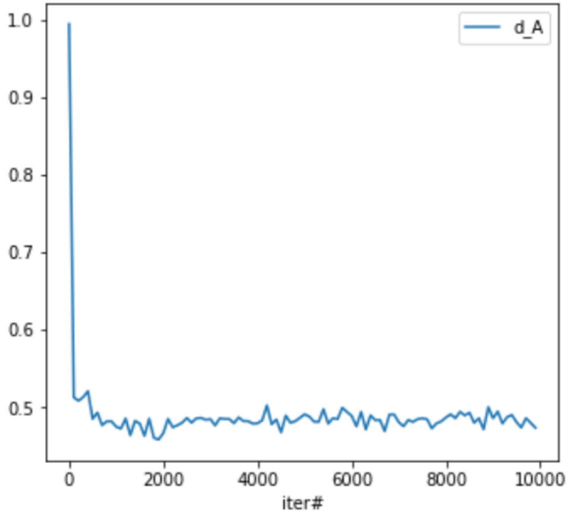
Predicted Y_s | *Predicted* $Y_{s \to t}$ | *Ground Truth* Y_s

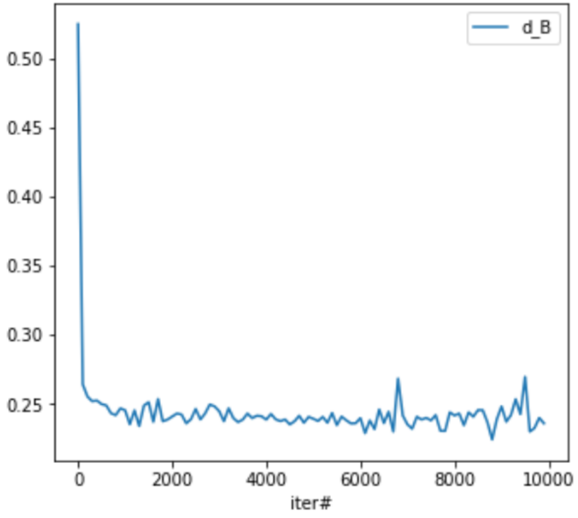
- Brain mask is bigger than the previous vessels mask
- Reduced imbalance between **Background** and the actual **Brain**

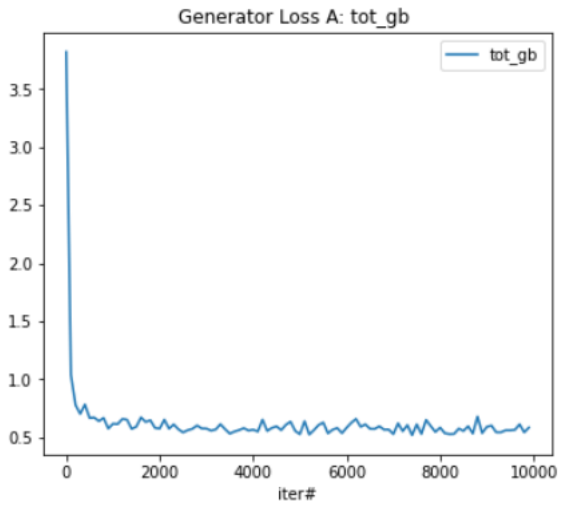
EXPERIMENT #3: BRAIN MASK (10K), LOWER LR

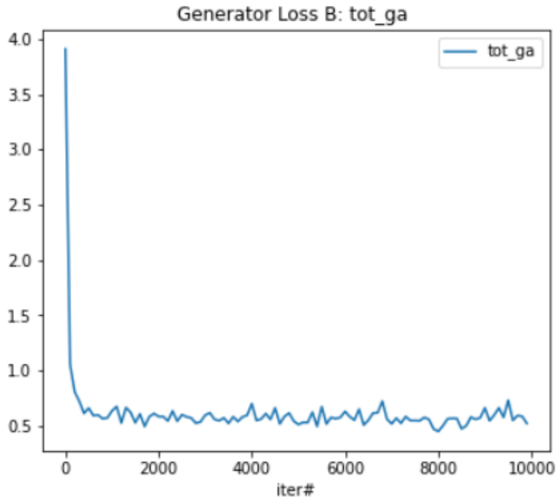
EXPERIMENTS AND RESULTS







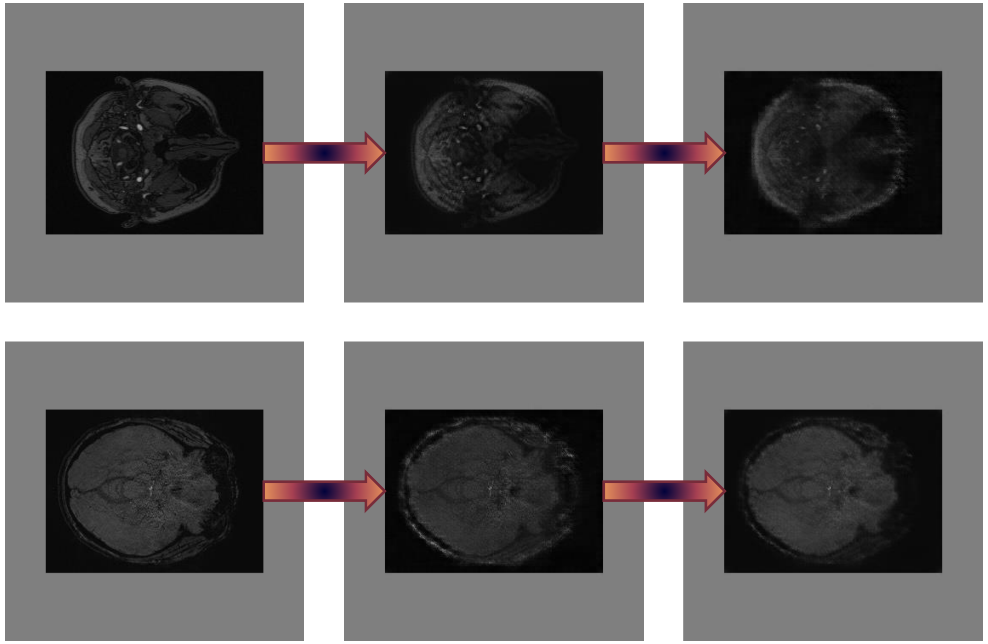




EURECOM
Sophia Antipolis

EXPERIMENT #3: BRAIN MASK (10K), LOWER LR

EXPERIMENTS AND RESULTS



- Transformed Images are more blurry than before
- The pixel-wise classifier C is not able to learn how to segment small pixels in the image

SUMMARY AND CONCLUSIONS

CONCLUSION

SUMMARY

- Good transformed images quality but possible fading effect.
- **Instability:** the network is complex to fine-tune and **heavy** to train, convergence is not always reached.
- **Vessel Masks** are too unbalanced towards the background to be predicted, better results with **Brain Masks.**

POSSIBLE FUTURE EXPERIMENTS

- **Convergence Issues:** Fine-tune loss hyper-parameters so that all the objective function terms have the same importance.
- **Fine-tuning:** Run different experiments with lower lr, different number of steps, decay, etc.
- **Unstable Losses**: Try a different Discriminator Loss in order to avoid any possible divergence and instability issues.
- Use **Target Masks** to have a quantitative segmentation metric and to compute a validation score (Dice Coefficient)

THANK YOU FOR YOUR ATTENTION

PROFESSOR:
MARIA ZULUAGA

STUDENT:
DANIELE FALCETTA

TEACHING ASSISTANT:
FRANCESCO GALATI

SEMESTER PROJECT – SPRING 2022

EURECOM
Sophia Antipolis

BIBLIOGRAPHY

- Synergistic Image and Feature Adaptation: Towards Cross-Modality Domain Adaptation for Medical Image Segmentation. Cheng Chen, Qi Dou, Hao Chen, Jing Qin, Pheng-Ann Heng
- Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros
- Image-to-Image Translation with Conditional Adversarial Networks. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A.Efros, Berkeley AI Research (BAIR) Laboratory, UCBerkeley.
- Image Segmentation Using Deep Learning: A Survey. Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, Demetri Terzopoulos
- Susceptibility-Weighted Imaging: Technical Aspects and Clinical Applications, Part 1. E.M. Haacke, S. Mittal, Z. Wu, J. Neelavalli and Y.-C.N. Cheng. American Journal of Neuroradiology January 2009, 30 (1) 19-30; DOI: https://doi.org/10.3174/ajnr.A1400 (http://www.ajnr.org/content/30/1/19)

- https://radiopaedia.org/articles/time-of-flight-angiography-1
- https://radiopaedia.org/articles/susceptibility-weighted-imaging-1
- https://dicom.innolitics.com/ciods/rt-dose/image-plane/00280030