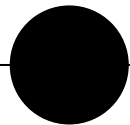


ReCycleGAN: Revisitando a CycleGAN



Redes Generativas (IA376N) – 2024S2

Gabriel Freitas (289.996)

Tiago Amorim (100.675)

Sumário

Tradução imagem-imagem

CycleGAN

Problema a resolver

Proposta

- Ferramentas

- Base de dados

- Cronograma

● Tradução Imagem-Imagem

Mapeamento entre domínios que compartilham relação (*image-to-image translation*):

Transferência de estilo (*style transfer*)

Aumento de qualidade (image enhancement)

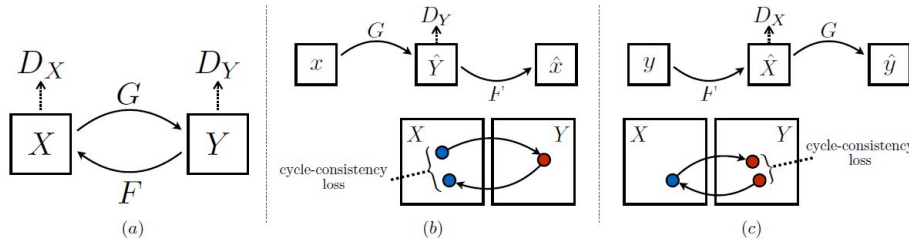
Remoção de ruído (image restoration)



Fonte: Jeong, Somi, et al. "Memory-guided unsupervised image-to-image translation." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021.

● CycleGAN

Zhu et al., em **Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks** (2017), mostrou uma forma de treinar uma rede para copiar o estilo de um conjunto de imagens em uma imagem de referência com bases de dados não pareados.



$$\begin{aligned} \mathcal{L}(G, F, D_X, D_Y) = & \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) \\ & + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) \\ & + \lambda \mathcal{L}_{\text{cyc}}(G, F), \end{aligned}$$

● Problema Atual

Grande avanço na qualidade das imagens geradas.

Treinamento e inferência nas redes mais novas depende de significativo recurso computacional (e.g.: modelos de difusão).

Real

Diferentes Modelos



Fonte: Jeong, Somi, et al. "Memory-guided unsupervised image-to-image translation." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021.

● Proposta de Projeto

Aplicar elementos de estudos mais recentes à CycleGAN e avaliar possíveis ganhos.

Novos elementos na arquitetura da rede.

Funções de perda adicionais.

Novas métricas de avaliação.

Comparativo com outras redes.

● Proposta de Projeto

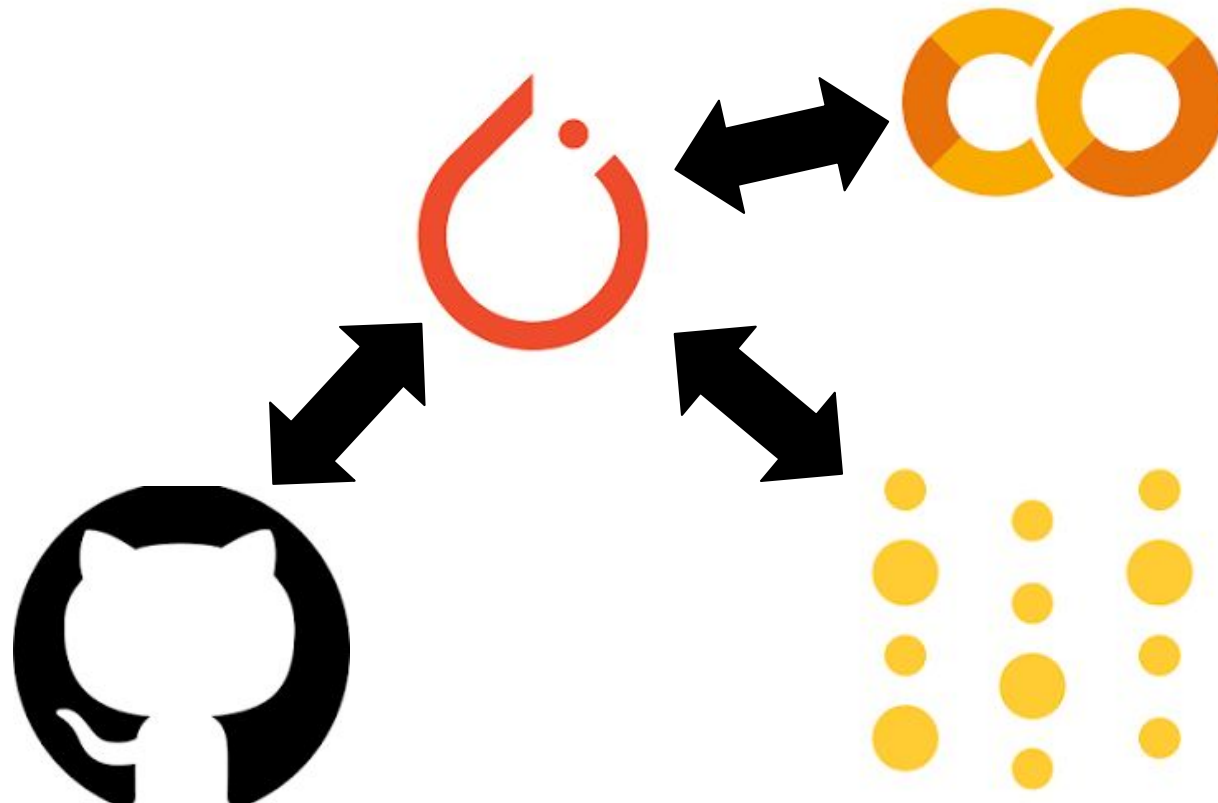
Skip connections
Camadas de atenção
Adaptadores LoRA

Perdas baseadas em redes
pré-treinadas
Suavização do Gerador

FID	PSNR
SSIM	LPIPS

Percepção de usuários
Redes de difusão

Ferramentas



● Problemas e Bases de Dados



Transferência de estilo: **Nexet 2017**

50.000 imagens de câmeras automotivas.

Anotações de condição de luz e local.

Foco em 9.380 imagens de dia e noite em Nova York.



● Problemas e Bases de Dados

Remoção de ruídos: **O-Haze, I-Haze, D-Hazy**

Imagens embaçadas e de referência.

1400 imagens com mapas de profundidade para sintetizar cenas embaçadas.

Será utilizada como se fosse uma base não-pareada.



Ground truth image



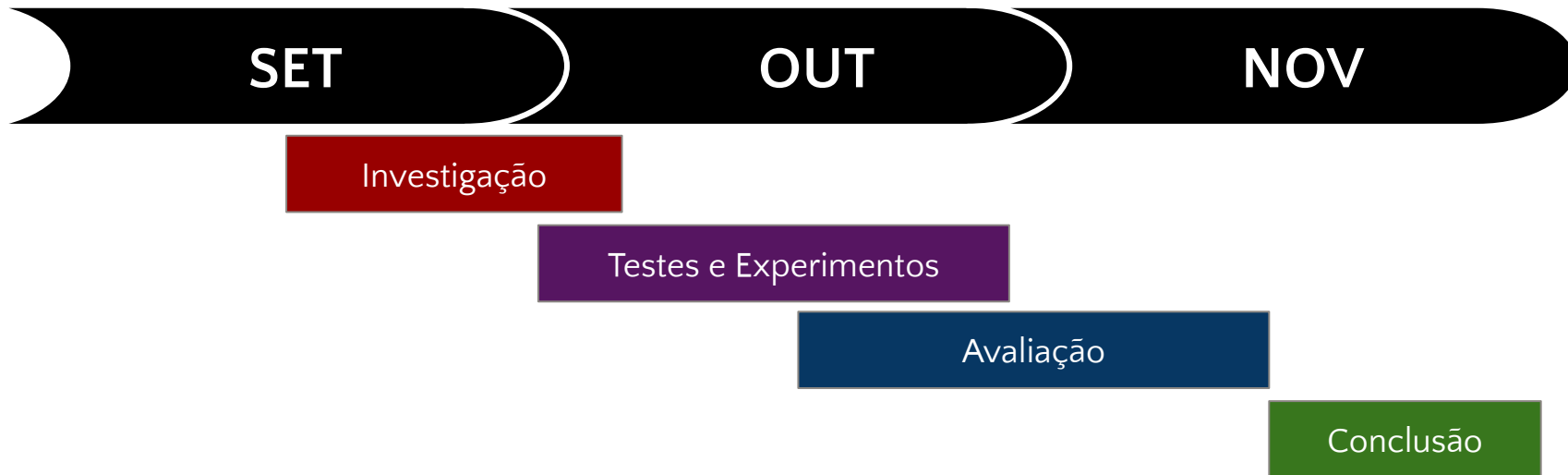
Depth map



Hazy image



● Cronograma



● Perguntas?



ReCycleGAN

