

V2IP: Videomonitoramento para Identificação de Pessoas e Veículos

Coordenador (a): David Menotti (UFPR)

Linha: (i) Ciência de dados, inteligência artificial, análise integrada e **aprendizado de máquina** aplicados a soluções tecnológicas de segurança pública, com foco em **videomonitoramento**

20/09/2022 - [11:30,11:50]

Coordenador Geral: **DAVID MENOTTI**

- Professor Associado IV - DInf / UFPR, desde 2015/2008 (UFOP)
- Bolsista de Produtividade CNPq - Pq-2
- Doutor em Ciência da Computação (UFMG)
 - *Informatique* (Université Paris-Est, França)
 - Bolsa CAPES - doutorado sanduíche/cotutela
 - especialidade: [Visão Computacional](#)
 - Aprendizado de Máquinas
- Experiência em Biometria (íris, olho) e Vigilância (veículos)
- InterForensics 2017, 2019 & 2021

Coordenador Associado: **WILLIAM ROBSON SCHWARTZ**

- Professor Associado II - DCC / UFMG, desde 2011
- Bolsista de Produtividade do CNPq - Pq-1D
- Doutor em Ciência da Computação (University of Maryland, EUA)
 - especialidade: Visão Computacional
- Experiência em Biometria (faces, gait) & Vigilância
- Coordenador Associado de projeto CAPES/PROFORENSES (2015-2019)
- Coordenador do PPGCC / UFMG
- InterForensics 2017, 2019 & 2021
- Projetos em Cooperação com David Menotti (>2012 - +20/+1k)

Coordenador Estrangeiro: **HUGO PROENÇA**

- Professor Associado - DInf / U. Beira Interior, Portugal, desde 2007
- Doutor em Ciência da Computação (UBI)
 - especialidade Reconhecimento de Padrões
- Experiência em Biometria Não Cooperativa
 - Ambientes de Videovigilância
- Destaque na comunidade mundial na área de Biometria
- Coorientação de doutorando do PPGInf/UFPR em 2019/2020

Colaborador Estrangeiro: **JOÃO NEVES**

- Professor Auxiliar - Universidade da Beira Interior
 - Autor de mais de 40 publicações científicas em Visão Computacional.
 - 7 projetos de pesquisa relacionados à Visão Computacional.
- Doutor em Ciências da Computação (UBI)
 - Especialidade: Visão Computacional
- Experiência em Videovigilância
 - Reconhecimento facial em imagens de baixa resolução
- Parceria com a DITEC/PF

Representante DITEC/PF: **RAFAEL OLIVEIRA RIBEIRO**

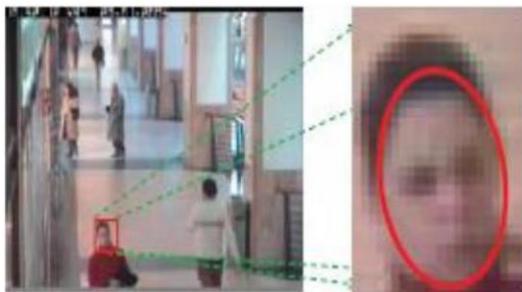
- Perito Criminal Federal (PCF) desde 2008
 - Serviço de Perícias em Audiovisual e Eletrônicos (SEPAEL) do
 - Instituto Nacional de Criminalística (INC)
- Chefe do SEPAEL de 04/2018 - 07/2020
 - Lançamento do software **PERITUS**
 - Sistema Integrado de Análise Forense de Imagens
- Representa INC desde 2019 no FISWG - agências GOV/PB/PV
 - *Facial Identification Scientific Working Group (FISWG)*
- Contato da área de **identificação humana** e **antropologia** da DITEC/PF

Representante SR-Bahia/PF: **JORGE DE ALBUQUERQUE LAMBERT**

- Perito Criminal Federal (PCF) desde 2008
 - Doutorando em Sistemas Robóticos na UnB desde 2016
 - Atuou na DITEC em Brasília de 2008 até 2019
 - Atuou no PROFORENSES ABCF/CAPES (2015-2020)
 - Atualmente PCF na Superintendência Regional da PF na BAHIA
- Pesquisador na área de *deep learning*
 - Reconhecimento de placas de automóveis em imagens de muito baixa qualidade

Videomonitoramento

- Grande quantidade de dados visuais vem sendo gerada por câmeras
- Análise não é uma tarefa simples (custosa / tediosa)
 - principalmente quando se trata de dados de **baixa qualidade** ou
 - os agentes de interesse da cena se localizam **distantes** da câmera ou
 - onde a visualização da face seja **parcial** pela **pose** ou por uso de **máscaras**



Objetivo geral:

Aplicar técnicas de visão computacional e aprendizado de máquina, usando o estado da arte (***deep learning***) para:

- a criação de **soluções tecnológicas** de **segurança pública e forense** com foco em videomonitoramento para resolver problemas:
 - identificação de **pessoas** e
 - reconhecimento de **placas veiculares**,
- especialmente para dados desafiadores, tais como
 - vídeos de **baixa resolução** e vídeos contendo pessoas e veículos **distantes** da câmera ou onde a visualização da face seja parcial pela **pose** ou por uso de **máscaras**.

Objetivo Principal/Específicos:

- (P1) reconhecimento de placas de **veículo** a partir de vídeos digitais em cenário de baixa qualidade / resolução ou baixa visibilidade;
- (P2) identificação de **pessoas** onde a imagem da face seja **parcial** por uso de máscaras ou pela pose;
- (P3) **reconhecimento facial em vídeos**;
- (P4) reconhecimento de pessoas pela forma de **andar** (*gait*).

Objetivos específicos:

1. Cotar e adquirir equipamentos necessários à pesquisa / ao projeto.
1. Coletar e analisar a literatura e trabalhos correlatos já apresentados à comunidade científica relacionadas aos problemas-alvo.

Objetivos específicos:

3. Coleta e compilação de bases de dados referentes aos quatro problemas-alvo.
 - a. Criar/coletar bases de referência em **placas de veículos** em cenários desafiadores (P1).
 - b. Criar/coletar bases de referência em identificação de **pessoas** usando **máscaras** (P2).
 - c. Criar/coletar bases de referência em identificação de pessoas em vídeo (P3).
 - d. Criar/coletar bases de referência para reconhecimento de pessoas a partir do **gait** (P4).

Objetivos específicos:

4. Explorar conceitos de aprendizado em profundidade (*Deep Learning*) em todos os problemas-alvo para uma melhor descrição dos dados (P1, P2, P3, P4).

Objetivos específicos:

5. Investigação e proposição de novas soluções para os quatro problemas-alvo.
 - a. Desenvolver algoritmo de **super-resolução** que seja capaz de reconstruir a região da placa de licenciamento para então realizar a identificação do veículo (P1).
 - b. Aprimorar soluções para o reconhecimento de indivíduos a partir da região **ocular** com foco em imagens da face usando **máscara** ou com oclusão na **face** (P2).
 - c. Desenvolver algoritmo capaz de reconhecer **faces** em **vídeos** de baixa qualidade por meio de fusão de informações capturadas a partir de *frames* relevantes do vídeo (P3).
 - d. Criação e validação de algoritmo robusto à direção da caminhada para identificação de pessoas pelo *gait* a partir da estimação do **esqueleto** 2D e 3D (P4).

Objetivos específicos:

6. Fortalecer a formação de recursos humanos nacional e internacionalmente a nível de pós-graduação *stricto sensu* para a área temática de Segurança Pública.
6. Integrar os algoritmos ao sistema PERITUS, desenvolvido pelo INC/DITEC/PF, para uso da polícia e institutos de perícia, validadas pela parceria com o SEPAEL/DPER/INC/DITEC/PF e SETEC/SR/PF/BA.

Objetivos específicos:

- 8. Publicar e divulgar os resultados obtidos.
- 8. Realizar encontros técnicos de pesquisa (trimestrais) entre os membros participantes.
- 8. Realizar workshops técnicos de pesquisa (anuais) entre os membros participantes.
- 8. Submeter relatório do projeto à agência financiadora.

Trabalhos aceitos no **SIBGRAPI 2022 (Qualis-A3 h5-21)**

* Conference on Graphics, Patterns and Images

* Natal de 24 à 27 de outubro de 2022

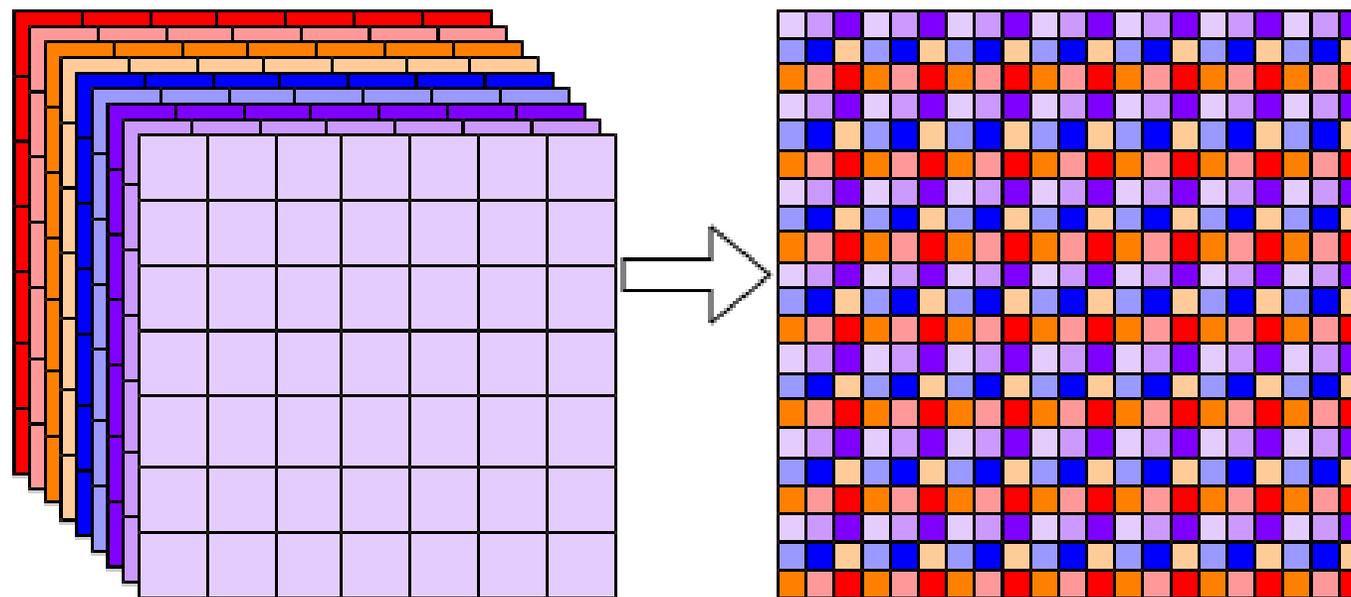
- **Combining Attention Module and Pixel Shuffle for License Plate Super-resolution**

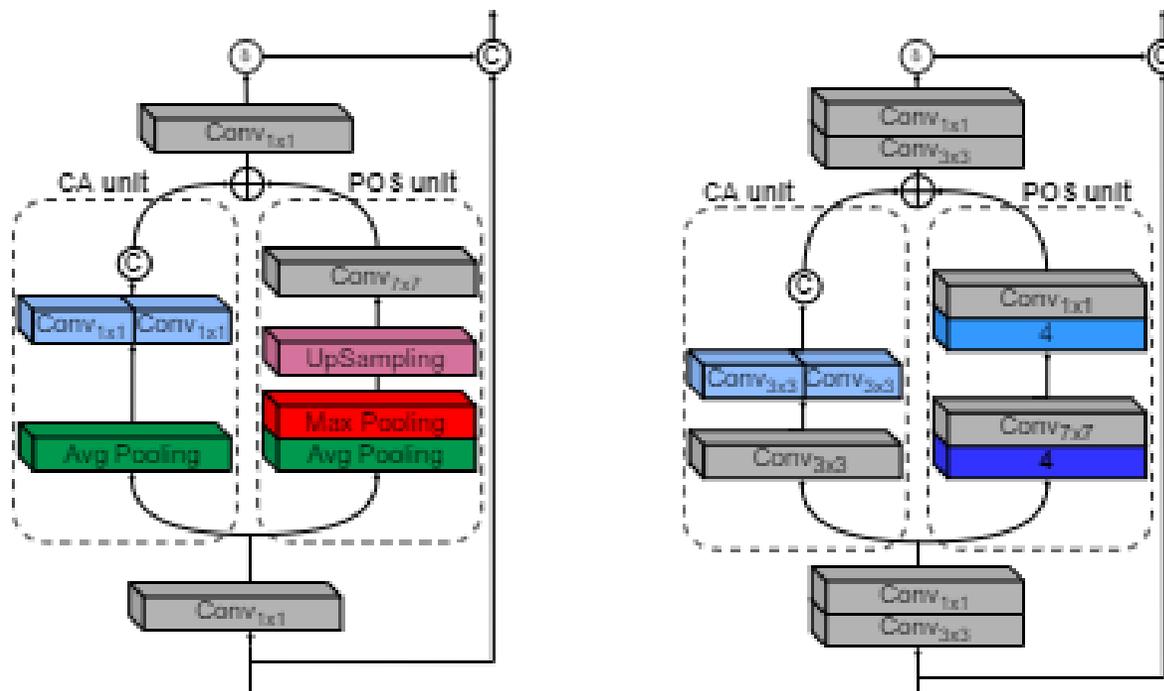
[Valfride Nascimento](#), Rayson Laroca, Jorge Lambert, William Schwartz, David Menotti

- **Face Super-resolution Using Stochastic Differential Equations**

[Marcelo dos Santos](#), Rayson Laroca, Rafael Ribeiro, João Neves, Hugo Proença, David Menotti

(P1): Reconhecimento de placas de **veículo** a partir de vídeos digitais em cenário de baixa qualidade / resolução ou baixa visibilidade;





- Ilustração comparativa dos módulos de atenção: à esquerda *Two-Fold Attention Module* (Mehri et al., 2021) e à direita *PixelShuffle Two-fold Attention Module* (Proposta)

- Os experimentos foram conduzidos em placas veiculares extraídas do dataset RodoSol-ALPR (Laroca et al., 2022). Entre suas 20.000 imagens temos 5.000 referentes aos layouts:



- As imagens em Alta-resolução e seus pares em Baixa-resolução foram geradas como descrito a seguir:
 - Degradação da imagem levando em consideração o índice de similaridade (SSIM)





Resultados Qualitativos para imagens com SSIM no intervalo]0, 0.75]

Tabela: Taxa de reconhecimento (%) dos experimentos até o presente momento.

SSIM	Cars			Motorcycles			Cars & Motor.			
	All	≤ 6	≤ 5	All	≤ 6	≤ 5	All	≤ 6	≤ 5	
No super-resolution										
HR	90.9	97.5	98.9	95.2	99.5	99.9	92.8	98.4	99.4	
Proposed model & baselines trained and tested with]0, 0.75] SSIM images										
Proposed	69.8	82.6	88.9	66.1	78.3	85.1	68.1	80.7	87.2	
LR-LPR (no SR) [3]	61.4	78.0	86.5	47.0	68.8	80.4	54.9	73.9	83.7	
MPRNet [1]	48.2	66.1	75.7	50.0	65.0	74.6	49.0	65.6	75.2	
Average PSNR (dB) and SSIM for tests with]0, 0.75] SSIM images										
		PSNR			SSIM					
Proposed		26.4			0.89					
MPRNet [1]		19.7			0.79					

- (P2): Identificação de **pessoas** onde a imagem da face seja **parcial** por uso de máscaras ou pela pose;
(P3): Reconhecimento facial em vídeos;



NIST Launches Studies into Masks' Effect on Face Recognition Software, July 27, 2020, <https://www.nist.gov/news-events/news/2020/07/nist-launches-studies-masks-effect-face-recognition-software>

Homem mascarado adentra em supermercado de Palmas de Monte Alto e rouba 9 mil reais, 27 de Novembro de 2018, <https://alertabahia.com.br/homem-mascarado-adentra-supermercado-palmas-monte-alto-rouba-9-mil-reais/>

Super-resolução e reconhecimento facial

Baixa Resolução



Super-Resolução



Referência



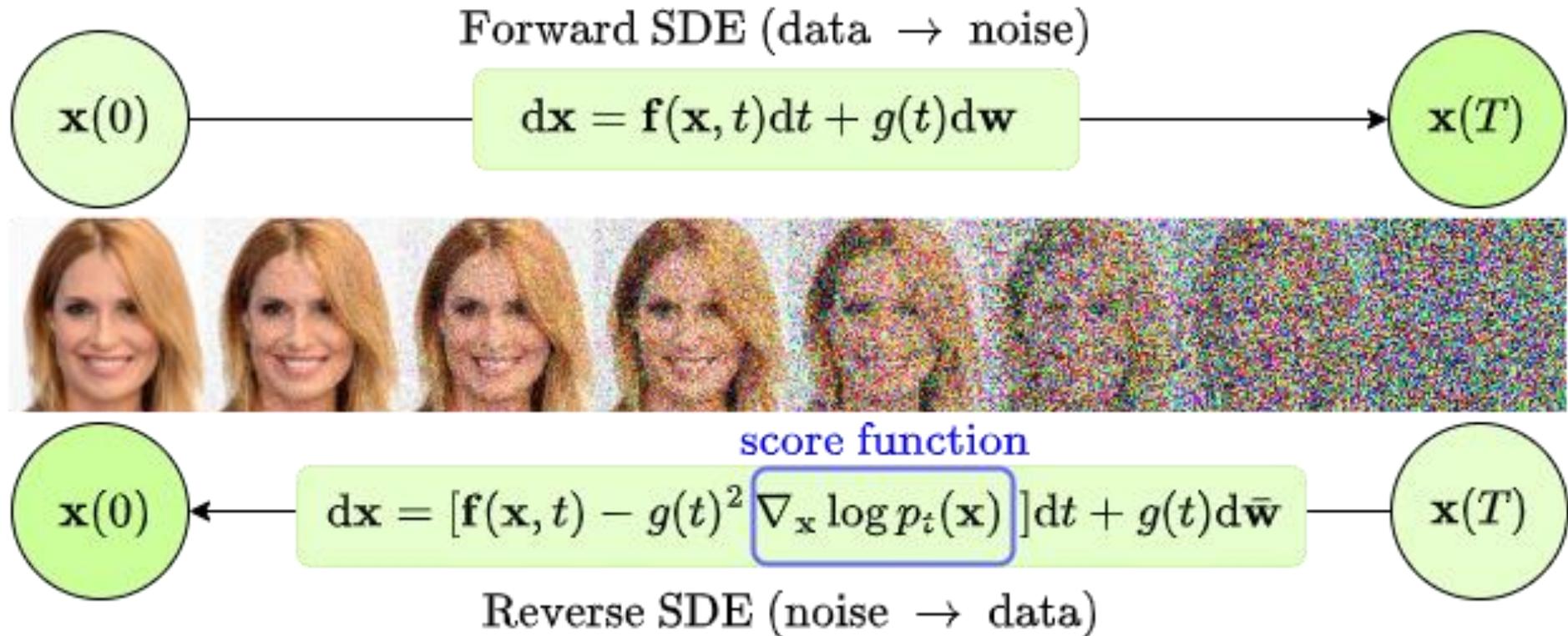
Modelos de difusão

Forward Diffusion (data \rightarrow noise)



Reverse Diffusion (noise \rightarrow data)

Equações diferenciais estocásticas



Modelos de difusão discretos:

- Score matching with Langevin dynamics (SMLD)
- Denoising Diffusion Probabilistic Models (DDPM)

Modelos contínuos:

- SDE-VE (SMLD)
- SDE-VEcs (etapa de correção)
- SDE-VP (DDPM)
- SDE-subVP

Algoritmo

Algorithm 1 Predictor-Corrector (PC) sampling

N: Number of discretization steps for the reverse-time SDE

M: Number of correction steps

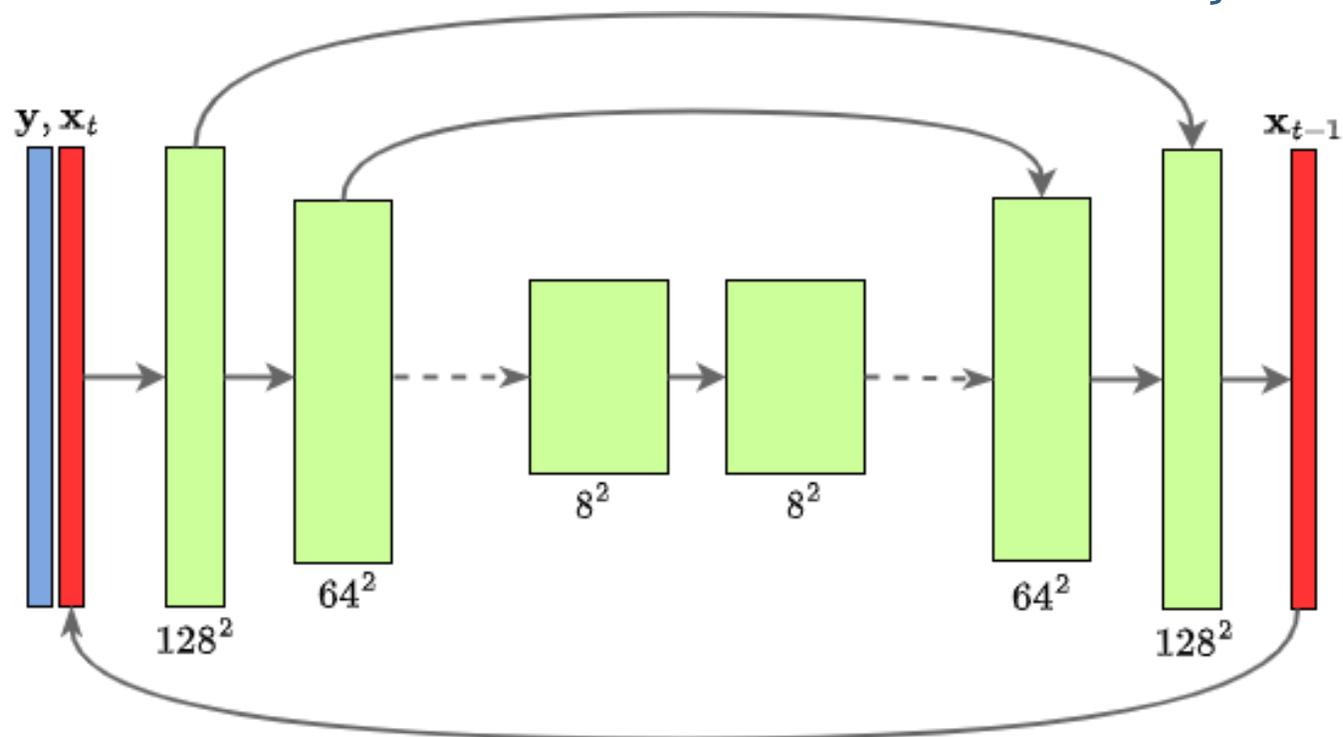
```
1: Initialize  $\mathbf{x}_T \sim p_T(\mathbf{x})$ 
2: for  $i = N - 1$  to 0 do
3:    $\mathbf{x}_i \leftarrow \text{Predictor}(\mathbf{x}_{i+1})$ 
4:   for  $j = 1$  to  $M$  do
5:      $\mathbf{x}_i \leftarrow \text{Corrector}(\mathbf{x}_i)$ 
6:   end for
7: end for
8: return  $\mathbf{x}_0$ 
```



- **Predictor:** faz a previsão da próxima imagem
- **Corrector:** efetua eventuais correções. Internamente possui parâmetro r que controla a suavização das imagens

Arquitetura da rede de super-resolução para estimar a *score function*

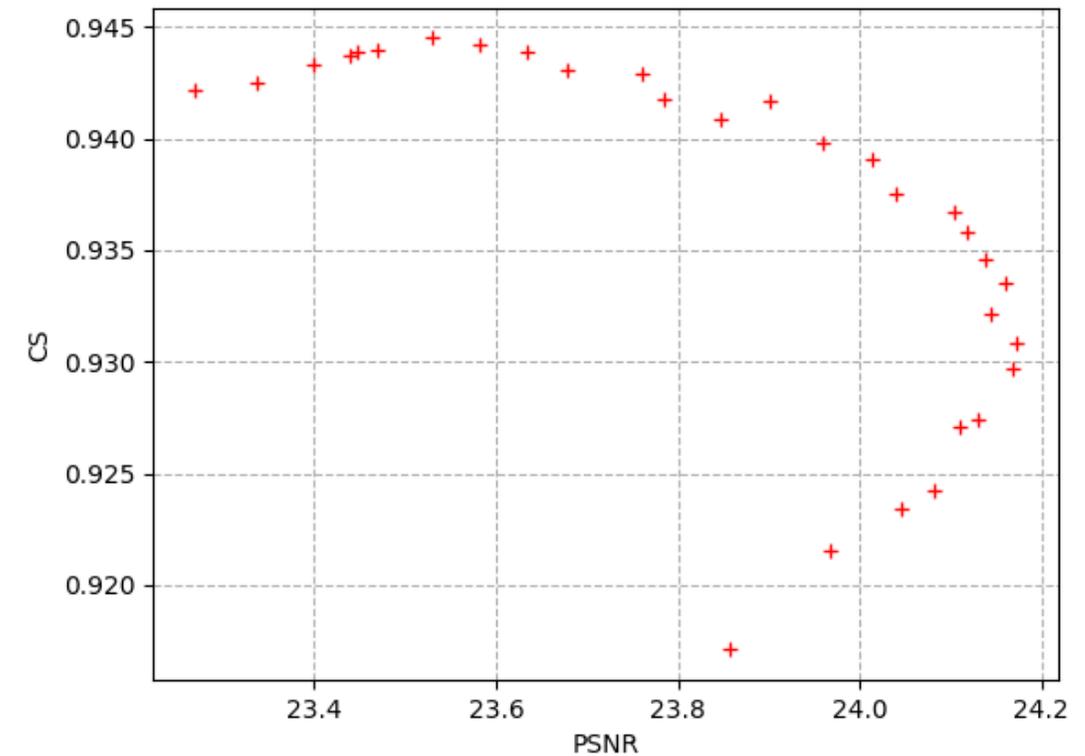
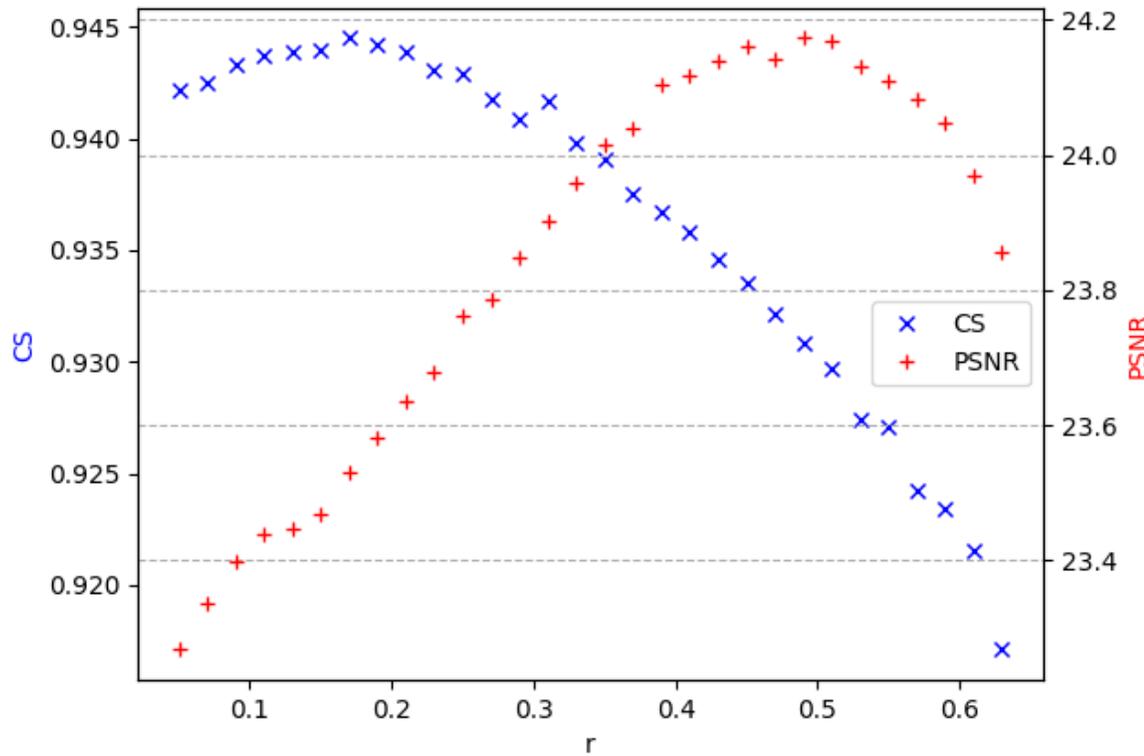
U-net com blocos ResNet e camadas de atenção



Metodologia

- Treinamento: Base Flickr - Faces - HQ (FFHQ)
- Imagens de baixa resolução geradas com *downsampling* de 8 x (128 -> 16)
- Testes: Base CelebA-HQ
- Extração de características com VGG-Face
- Métricas: SSIM, PSNR e consistência.
 - Similaridade de cosseno (CS) entre as características
- Análise da influência da suavização da imagem nas métricas PSNR e CS

Resultados



Resultados qualitativos



Resultados quantitativos

Model	PSNR \uparrow	SSIM \uparrow	CONSISTENCY \downarrow	CS \uparrow
GFP-GAN [3]	21.5326 \pm 1.5273	0.6006 \pm 0.0709	37.2256 \pm 12.4622	0.8689 \pm 0.0581
SPARNet [4]	24.3686 \pm 1.7844	0.7223 \pm 0.0679	13.6512 \pm 4.9063	0.9307 \pm 0.0301
SR3 [5]	22.9581 \pm 1.8370	0.6605 \pm 0.0758	1.3715 \pm 0.7904	0.9370 \pm 0.0244
SDE-VP	22.7171 \pm 1.8107	0.6448 \pm 0.0787	0.1074 \pm 0.0592	0.9330 \pm 0.0262
SDE-subVP	22.6455 \pm 1.8047	0.6428 \pm 0.0797	0.1433 \pm 0.1212	0.9300 \pm 0.0261
SDE-VE	23.5101 \pm 1.9492	0.6879 \pm 0.0797	0.0454 \pm 0.0357	0.9443 \pm 0.0222

Resultados preliminares - Cenário real - SCface



Aumento no CS de **64 %** para **70%**



Coleta de dados reais

- Placas de Veículos (bitrate / resolução)
 - Doação de Câmeras
 - Instalação de câmeras no campus da UFPR
 - Cooperação com a PF - Claudio Viana
- Rostos/Faces
 - Cooperação com a PF para pesquisa na base de passaportes (gerenciado pela própria PF)
 - Cooperação com a PM do Paraná para imagens em ambientes realísticos



Obrigado(a).