SIBGRAPI 2022 | 35th Conference on Graphics, Patterns and Images

UNIVERSIDADE FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Combining Attention Module and Pixel Shuffle for License Plate Super-Resolution

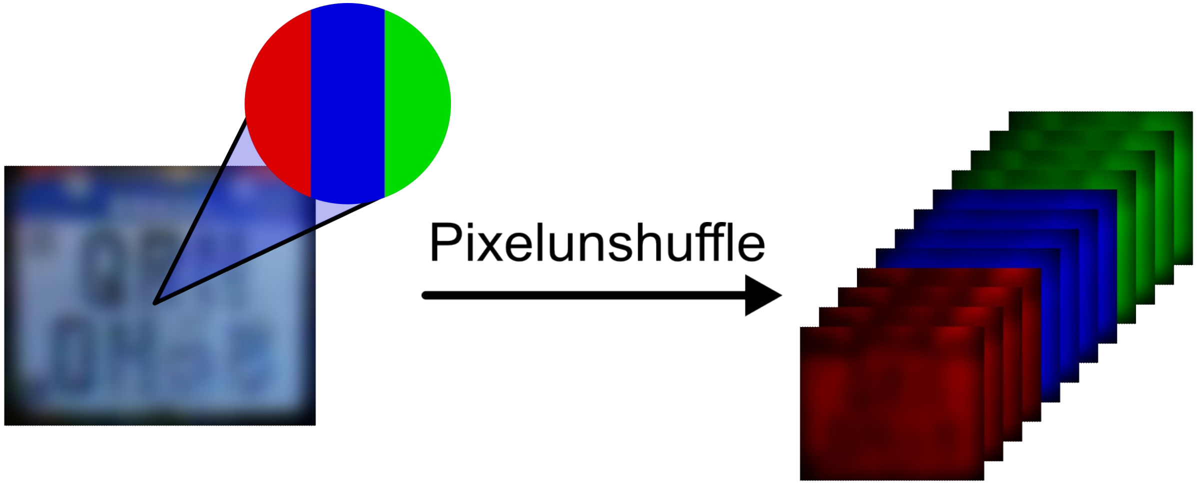
Valfride Nascimento[*], Rayson Laroca[*], Jorge de A. Lambert[†], William Robson Schwartz[‡], and David Menotti[*]

[*]Department of Informatics, Federal University of Paraná, Curitiba, Brazil
[†]Regional Superintendence at Bahia, Brazilian Federal Police, Salvador, Brazil
[‡]Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte, Brazil
[*]{vwnascimento,rblsantos,menotti}@inf.ufpr.br [†]lambert.jal@pf.gov.br [‡]william@dcc.ufmg.br

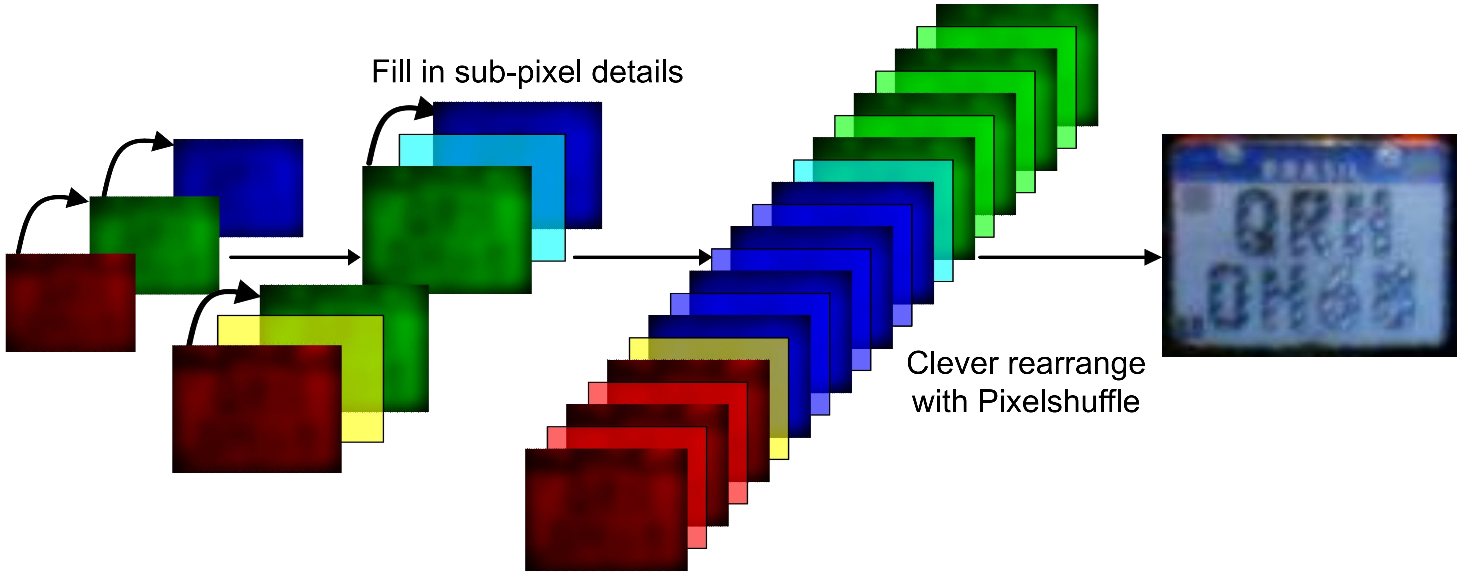
Sponsors: CAPES, CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico, UFRN PPG, SEBRAE RN
Promotion: SBC Sociedade Brasileira de Computação, CS-RA Joint Chapter IEEE Brazil Northeast
Realization: UFRN Universidade Federal do Rio Grande do Norte

Table of Contents

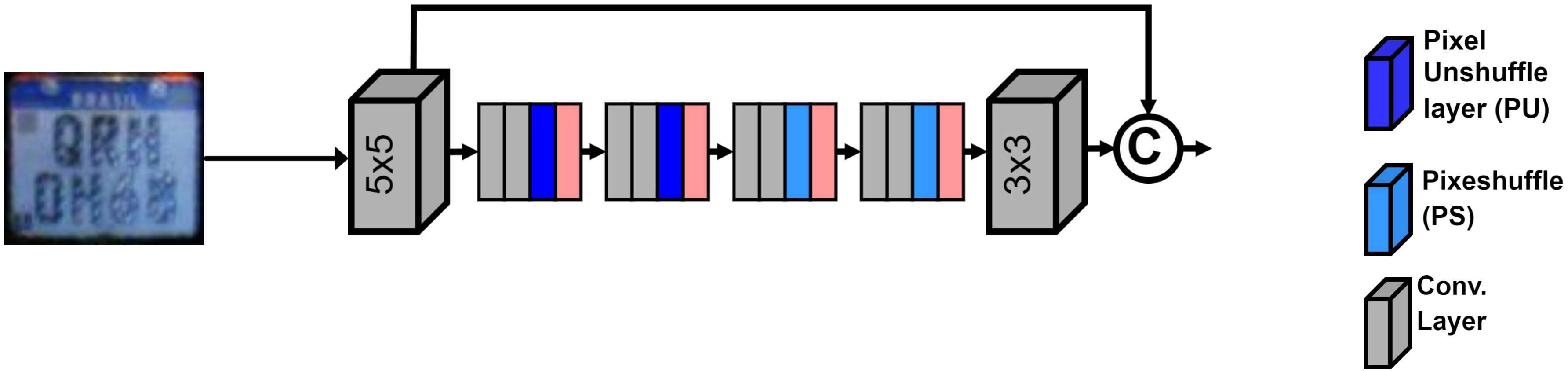
Sub-pixel Convolution Layer



Sub-pixel Convolution Layer

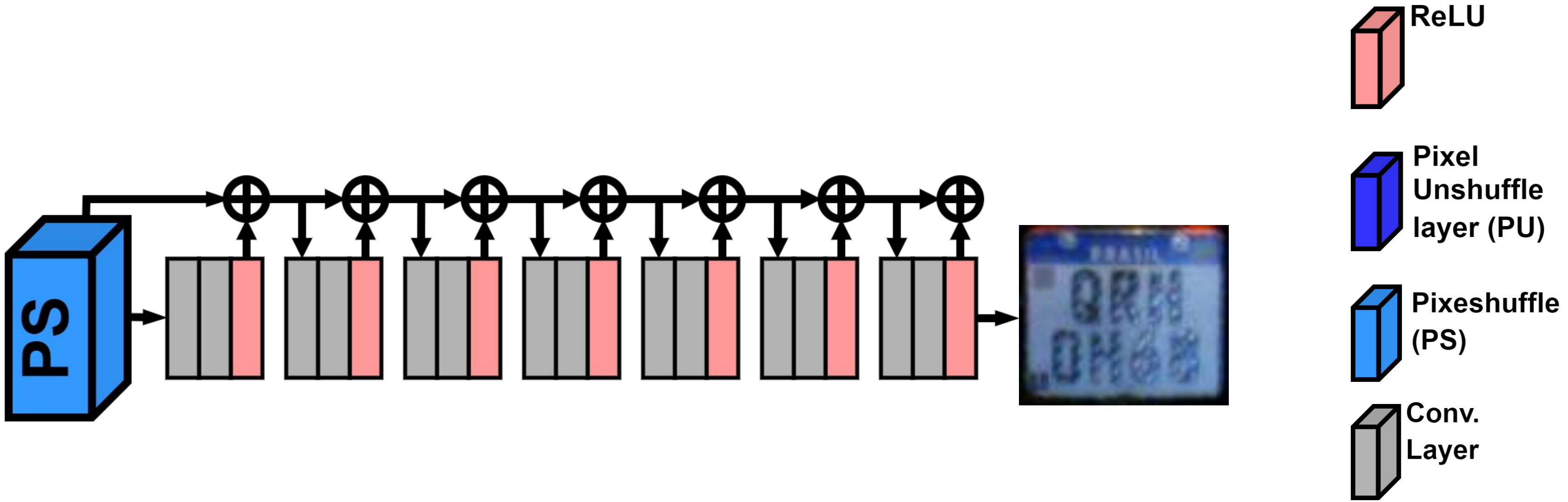


Shallow Feature Extraction (SFE)



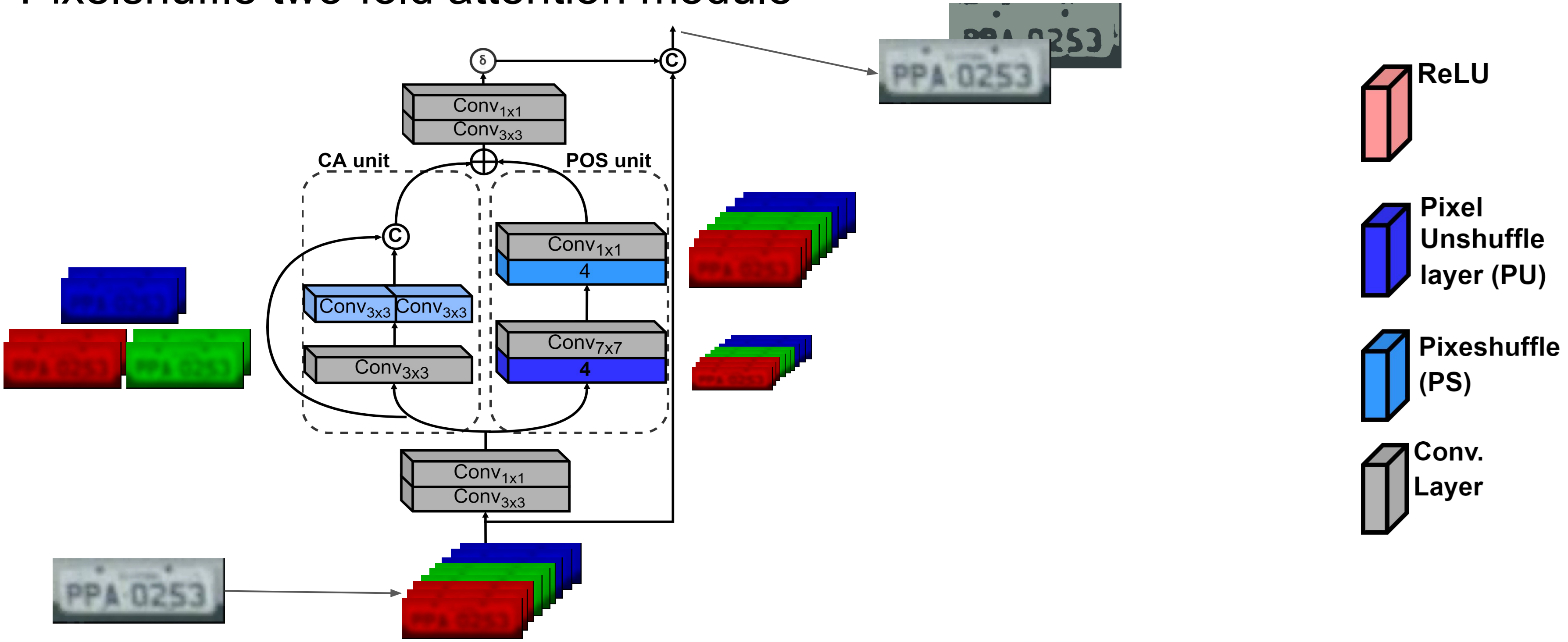
[1] A. Mehri, P. B. Ardakani, and A. D. Sappa, "**MPRNet: Multi-path residual network for lightweight image super resolution**," in ***IEEE Winter Conference on Applications of Computer Vision***, 2021, pp. 2703–2712.

Reconstruction Module (RM)



[1] A. Mehri, P. B. Ardakani, and A. D. Sappa, “**MPRNet: Multi-path residual network for lightweight image super resolution**,” in ***IEEE Winter Conference on Applications of Computer Vision***, 2021, pp. 2703–2712.

Pixelshuffle two-fold attention module



Perceptual Loss

$$PL = \frac{\sum_{i=0}^{n} \left(H_i - S_i \right) \left(1 + \alpha \cdot D\left(H_i, S_i \right) \right)}{n}$$

$$D\left(H_i,\ S_i \right) = \frac{Levenshtein\left(H_i,\ S_i \right)^2}{7} + \left(1 - SSIM\left(H_i,\ S_i \right) \right)$$

RodoSol-ALPR

The RodoSol-ALPR [3] comprises 20,000 images taken by static cameras at pay rolls located in the Brazilian state of Espírito Santo. There are 5,000 images of each of the following combinations of vehicle type and License Plate layout:

[3] R. Laroca et al., “**On the cross-dataset generalization in license plate recognition,**” in International Conference on Computer Vision Theory and Applications (VISAPP), Feb 2022, pp. 166–178. The instructions for downloading the RodoSol-ALPR dataset are available at https://github.com/raysonlaroca/rodosol-alpr-dataset

Building the Low-Resolution/Quality Dataset

- Cropped the LP region using the annotations provided by the authors;
- LPs are rectified and tagged as the HR image;
- Degradation is added until the desired SSIM score between the HR and LR images is reached.

SSIM Score Comparison

Qualitative Results

[1] A. Mehri, P. B. Ardakani, and A. D. Sappa, "**MPRNet: Multi-path residual network for lightweight image super resolution**," in ***IEEE Winter Conference on Applications of Computer Vision***, 2021, pp. 2703–2712.

Quantitative Results

Proposed model & baselines trained and tested with]0,0.75] SSIM images									
	Cars			**Motorcycles**			**Cars & Motor.**		
	All	≤ 6	≤ 5	All	≤ 6	≤ 5	All	≤ 6	≤ 5
Proposed	**69.8**	**82.6**	**88.9**	**66.1**	**78.3**	**85.1**	**68.1**	**80.7**	**87.2**
LR-LPR (no SR) [2]	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

[1] A. Mehri, P. B. Ardakani, and A. D. Sappa, "**MPRNet: Multi-path residual network for lightweight image super resolution**," in ***IEEE Winter Conference on Applications of Computer Vision***, 2021, pp. 2703–2712.

[2] G. R. Gonçalves et al., "**Multi-task learning for low-resolution license plate recognition**," in ***Iberoamerican Congress on Pattern Recognition* (CIARP)**, Oct 2019, pp. 251–261.

Quantitative Results

Average PSNR (dB) and SSIM for tests with]0 , 0 . 75] SSIM images		
	PSNR (db)	**SSIM**
Proposed	26.4	0.89
MPRNet [1]	19.7	0.79

[1] A. Mehri, P. B. Ardakani, and A. D. Sappa, “**MPRNet: Multi-path residual network for lightweight image super resolution**,” in ***IEEE Winter Conference on Applications of Computer Vision***, 2021, pp. 2703–2712.

Final Remarks and Future Work

- Final remarks
  - The main intuition of our approach is to exploit channel reorganization and learning capabilities from Pixelshuffle;
  - We added an autoencoder with PS and PU layers for shallow feature extraction as the network input and a Reconstruction module in its end for better image quality;
  - In *PLTFAM* we exploited PS and PU instead of the original *MaxPool*, *AvgPool* and *upscaling* ones from the original architecture.

Final Remarks and Future Work

- Future Work:
  - Collect videos where the LP is perfectly legible on one frame but illegible on another;
  - Develop new methods for real-world scenarios;
  - Carry out experiments in cross-dataset setups.

Thank you!

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

Visão
Robótica
Imagem

RODOSOL

CAPES

CNPq

NVIDIA