SIBGRAPI 2024 | 37th Conference on Graphics, Patterns and Images

Enhancing License Plate Super-Resolution: A Layout-Aware and Character-Driven Approach

Valfride Nascimento[*], Rayson Laroca[†,*], Rafael O. Ribeiro[‡], William Robson Schwartz[§], David Menotti[*]

[*] Federal University of Paraná, Curitiba, Brazil
[†] Pontifical Catholic University of Paraná, Curitiba, Brazil
[‡] Brazilian Federal Police, Brasília, Brazil
[§] Federal University of Minas Gerais, Belo Horizonte, Brazil

[*] {vwnascimento,menotti}@inf.ufpr [†] rayson@ppgia.pucpr.br [‡] rafael.ror@pf.gov.br [§] william@dcc.ufmg.br

UFPR UNIVERSIDADE FEDERAL DO PARANÁ

Vision, Robotics and Imaging Laboratory

CNPq

CAPES

NVIDIA

Summary

1. Problem and Motivation
2. Contributions
3. Layout and Character Oriented Focal Loss
4. Architecture
5. Datasets
6. Quantitative Results
7. Qualitative Results
8. Ablation Study
9. Final Remarks

Problem and Motivation

- Low-resolution license plates are **common** in **real-world surveillance**.

Problem and Motivation

- Blurred and **low-quality** images **reduce** license plate recognition (LPR) **performance**

Problem and Motivation

- Need for an approach that i**mproves character reconstruction in low-resolution scenarios.**

Contributions

- **Novel Loss Function:** Layout and Character Oriented Focal Loss (LCOFL).

Contributions

- **Novel Loss Function:** Layout and Character Oriented Focal Loss (LCOFL).

- **Architecture improvements:** Incorporation of deformable convolutions and shared weights in attention modules.

Contributions

- **Novel Loss Function:** Layout and Character Oriented Focal Loss (LCOFL).
- **Architecture improvements:** Incorporation of deformable convolutions and shared weights in attention modules.
- **Training Strategy:** A GAN-based training approach with OCR-guided discriminator for generation of more realistic and recognizable LP images.

Contributions

- **Novel Loss Function:** Layout and Character Oriented Focal Loss (LCOFL).
- **Architecture improvements:** Incorporation of deformable convolutions and shared weights in attention modules.
- **Training Strategy:** A GAN-based training approach with OCR-guided discriminator for generation of more realistic and recognizable LP images.
- **Real-World Images:** Preliminary experiments with real data.

Layout and Character Oriented Focal Loss

Layout and Character Oriented Focal Loss

$$L_C = -\frac{1}{K}\sum_{k=1}^{K} w_k \log p_t(y_k^{GT}|x_k)$$

- Weighted cross-entropy to classify characters.

Layout and Character Oriented Focal Loss

$$L_C = -\frac{1}{K}\sum_{k=1}^{K} w_k \log p_t(y_k^{GT}|x_k)$$

- Weighted cross-entropy to classify characters.
- Penalizes misclassification from structural similarities.

Layout and Character Oriented Focal Loss

$$L_C = -\frac{1}{K}\sum_{k=1}^{K} w_k \log p_t(y_k^{GT}|x_k)$$

- Weighted cross-entropy to classify characters.
- Penalizes misclassification from structural similarities.
- Weights are updated based on confusion matrix after each epoch.

Layout and Character Oriented Focal Loss

$$L_P = \sum_{i=1}^{K} \left[D(x_k) \cdot A(y_k^{GT}) + \cdot A(x_k) \cdot D(y_k^{GT}) \right]$$

$$\mathrm{D}(c) = \begin{cases} \beta & \text{if } c \text{ is a digit} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{A}(c) = \begin{cases} \beta & \text{if } c \text{ is a letter} \\ 0 & \text{otherwise} \end{cases}$$

- Enforces the correct positional arrangement of letters and digits in the LP layout.

Layout and Character Oriented Focal Loss

$$L_P = \sum_{i=1}^{K} \left[D(x_k) \cdot A(y_k^{GT}) + \cdot A(x_k) \cdot D(y_k^{GT}) \right]$$

$$\mathrm{D}(c) = \begin{cases} \beta & \text{if } c \text{ is a digit} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{A}(c) = \begin{cases} \beta & \text{if } c \text{ is a letter} \\ 0 & \text{otherwise} \end{cases}$$

- Enforces the correct positional arrangement of letters and digits in the LP layout.
- Penalizes errors where a letter is reconstructed as a digit, or vice versa.

Layout and Character Oriented Focal Loss

$$L_S = \frac{1 - SSIM(S_i, H_i)}{2}$$

- **Guides** the network **to maintain** LP **layout and structura**l details.

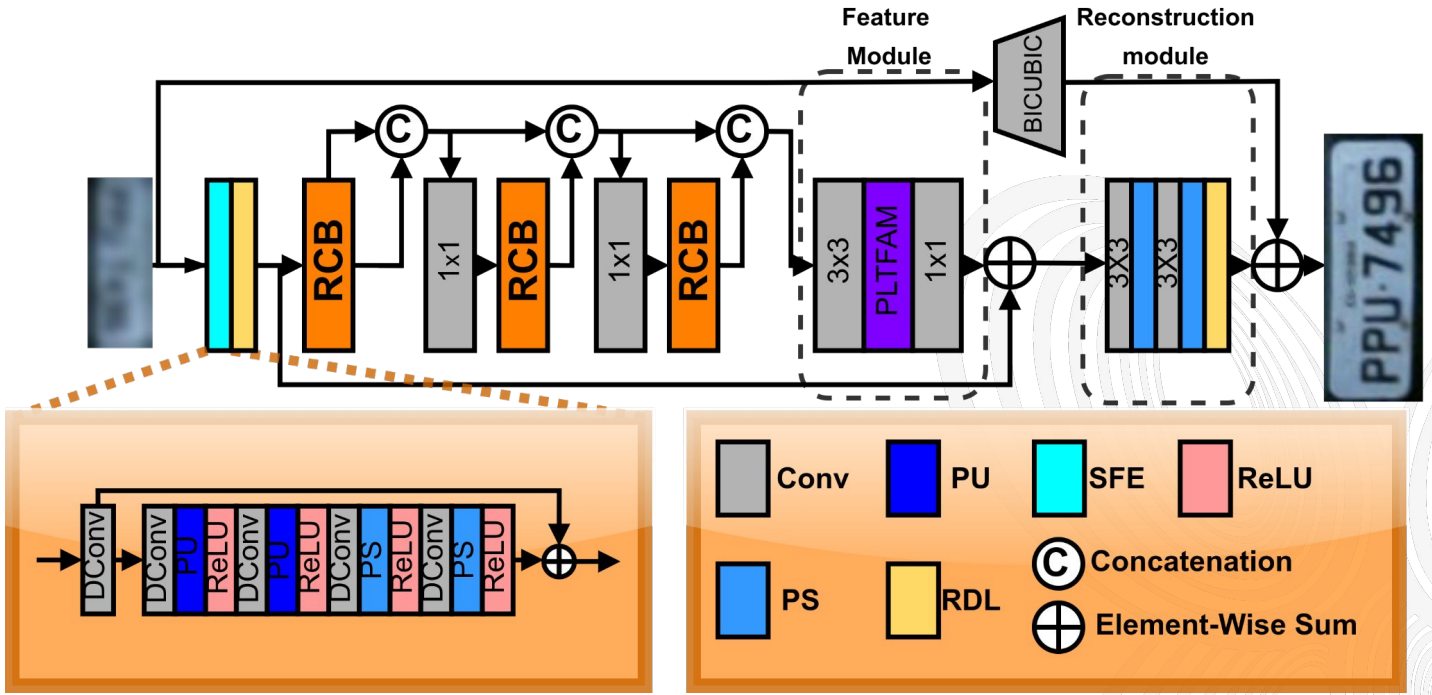
Layout and Character Oriented Focal Loss

$$L_S = \frac{1 - SSIM(S_i, H_i)}{2}$$

- **Guides** the network **to maintain** LP **layout and structura**l details.
- **Penalizes deviations** in **texture**, **structure**, and **contrast**.

Network Architecture

Architecture



Datasets

RodoSol-ALPR[1]

- **Input Resolution (LR):** 16×48 *px*;
- **Output Resolution (HR):** 32×96 *px*.

Preliminary Real-World Images

- **Input Resolution (LR):** 16×48 *px*;
- **Output Resolution (HR):** 32×96 *px*.

[1] R. Laroca, E. V. Cardoso, D. R. Lucio, V. Estevam, and D. Menotti, "**On the Cross-dataset Generalization in License Plate Recognition**" in *International Conference on Computer Vision Theory and Applications (VISAPP)*, Feb 2022, pp. 166-178.

Quantitative Results

Quantitative Results

Test Images	# Correct Characters					
	RodoSol-ALPR			*Preliminary Real-World Images*		
	All	≥ 6	≥ 5	All	≥ 6	≥ 5
HR (No SR)	98.5%	99.9%	99.9%	90.6%	98.7%	100%
LR (No SR)	**0.8%**	**4.1%**	**11.7%**	**9.9%**	**28.0%**	**56.2%**
Proposed model & baselines tested						
LR + SR (SR3 [6])	*43.1%*	*82.2%*	*82.2%*	31.7%	63.7%	80.1%
LR + SR (PLNET [5])	39.0%	59.9%	74.2%	*36.3%*	*67.2%*	*82.5%*
LR + SR (Proposed)	**49.8%**	**71.2%**	**83.3%**	**39.5%**	**70.2%**	**83.1%**

[5] Nascimento, V., Laroca, R., Lambert, J.D.A., Schwartz, W.R. and Menotti, D., "**Super-resolution of license plate images using attention modules and sub-pixel convolution layers.**" In **Computers & Graphics**, 2023, 113, pp.69-76.

[6] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi,"**Image super-resolution via iterative refinement**," **IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)**, vol. 45, no. 4, pp. 4713–4726, 2023. Google Research.

Quantitative Results

Test Images	# Correct Characters					
	RodoSol-ALPR			*Preliminary Real-World Images*		
	All	≥ 6	≥ 5	All	≥ 6	≥ 5
HR (No SR)	98.5%	99.9%	99.9%	90.6%	98.7%	100%
LR (No SR)	**0.8%**	**4.1%**	**11.7%**	**9.9%**	**28.0%**	**56.2%**
Proposed model & baselines tested						
LR + SR (SR3 [6])	*43.1%*	*82.2%*	*82.2%*	31.7%	63.7%	80.1%
LR + SR (PLNET [5])	39.0%	59.9%	74.2%	*36.3%*	*67.2%*	*82.5%*
LR + SR (Proposed)	**49.8%**	**71.2%**	**83.3%**	**39.5%**	**70.2%**	**83.1%**

[5] Nascimento, V., Laroca, R., Lambert, J.D.A., Schwartz, W.R. and Menotti, D., "**Super-resolution of license plate images using attention modules and sub-pixel convolution layers.**" In **Computers & Graphics**, 2023, 113, pp.69-76.

[6] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi,"**Image super-resolution via iterative refinement**," **IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)**, vol. 45, no. 4, pp. 4713–4726, 2023. Google Research.

Quantitative Results

Test Images	# Correct Characters					
	RodoSol-ALPR			*Preliminary Real-World Images*		
	All	≥ 6	≥ 5	All	≥ 6	≥ 5
HR (No SR)	98.5%	99.9%	99.9%	90.6%	98.7%	100%
LR (No SR)	**0.8%**	**4.1%**	**11.7%**	**9.9%**	**28.0%**	**56.2%**
Proposed model & baselines tested						
LR + SR (SR3 [6])	*43.1%*	*82.2%*	*82.2%*	31.7%	63.7%	80.1%
LR + SR (PLNET [5])	39.0%	59.9%	74.2%	*36.3%*	*67.2%*	*82.5%*
LR + SR (Proposed)	**49.8%**	**71.2%**	**83.3%**	**39.5%**	**70.2%**	**83.1%**

[5] Nascimento, V., Laroca, R., Lambert, J.D.A., Schwartz, W.R. and Menotti, D., "**Super-resolution of license plate images using attention modules and sub-pixel convolution layers.**" In **Computers & Graphics**, 2023, 113, pp.69-76.

[6] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi,"**Image super-resolution via iterative refinement**," **IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)**, vol. 45, no. 4, pp. 4713–4726, 2023. Google Research.

Qualitative Results

[5] Nascimento, V., Laroca, R., Lambert, J.D.A., Schwartz, W.R. and Menotti, D., "**Super-resolution of license plate images using attention modules and sub-pixel convolution layers.**" In **Computers & Graphics**, 2023, 113, pp.69-76.

[6] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi,"**Image super-resolution via iterative refinement**," **IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)**, vol. 45, no. 4, pp. 4713–4726, 2023. Google Research.

Qualitative Results

SR3 [6]

PLNET [5]

PROPOSED

HR

SR3 [6]

PLNET [5]

PROPOSED

HR

[5] Nascimento, V., Laroca, R., Lambert, J.D.A., Schwartz, W.R. and Menotti, D., "**Super-resolution of license plate images using attention modules and sub-pixel convolution layers.**" In **Computers & Graphics**, 2023, 113, pp.69-76.

[6] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi,"**Image super-resolution via iterative refinement**," **IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)**, vol. 45, no. 4, pp. 4713–4726, 2023. Google Research.

Ablation Study

Recognition Rates (RR) Achieved With Different Components Integrated Into The Proposed Approach

Approach	**RR**
Proposed (w/o ArchMod, GAN-style, and LCOFL)	39.0%
Proposed (w/o LCOFL)	45.9%
Proposed (w/o ArchMod and LCOFL)	47.6%
Proposed (w/o GAN-style and LCOFL)	47.7%
Proposed (w/o ArchMod and GAN-style)	48.2%
Proposed (w/o ArchMod)	49.2%
Proposed (w/o GAN-style)	49.4%
Proposed	49.8%

Final Remarks

- Our approach involved the implementation of LCOFL for character reconstruction according with the LP layout.

- LCOFL effectively mitigates confusions between structurally similar characters.

- Modifications to the archicture and training procedure led to state of the art results.

Final Remarks - Future Works

- Conduct experiments on various LP layouts.
- Complete the paired real-world LR/HR dataset and make it publicly accessible for researchers.
- Ensure the dataset supports multi-image super-resolution.

THANK YOU!

Vision, Robotics and Imaging Laboratory