# Super-Resolution Towards License Plate Recognition

**Valfride Nascimento**[1], **Rayson Laroca**[1], **David Menotti**[1]

[1]Department of Informatics, Federal University of Paraná (UFPR), Curitiba, Brazil

{vwnascimento, rblsantos, menotti}@inf.ufpr.br

***Abstract.*** *Recent years have seen significant developments in license plate recognition through the integration of deep learning techniques and the increasing availability of training data. Nevertheless, reconstructing license plates from low-resolution surveillance footage remains a challenge. To address this issue, we propose an attention-based super-resolution approach that incorporates sub-pixel convolution layers and an Optical Character Recognition* (OCR)*-based loss function. We trained the proposed architecture using synthetic images created by applying heavy Gaussian noise followed by bicubic downsampling to high-resolution license plate images. Our results show that the proposed approach for reconstructing these low-resolution images substantially outperforms existing methods in both quantitative and qualitative measures. Our source code is publicly available at* `https://github.com/valfride/lpr-rsr-ext/`.

## 1. Introduction

Super-resolution (SR) is widely used to enhance image or video quality by increasing resolution, particularly in domains such as medical imaging and surveillance [Liu et al. 2023]. Recent advancements in interpolation-based, example-based, and deep learning-based methods have significantly improved the capability to enhance low-resolution (LR) images and videos [Wang et al. 2021, Zhang et al. 2021b, Santos et al. 2022].

Despite these advances, SR remains challenging due to its ill-posed nature, where multiple solutions can exist in the high-resolution (HR) space [Wang et al. 2021]. Additionally, as the upscale factor increases, the computational complexity grows, and LR images may lack the necessary information for reconstructing desired details [Liu et al. 2023]. This study focuses on single-image super-resolution applied to license plate recognition, addressing the common occurrence of low-resolution and poor-quality images in real-world surveillance systems. While challenging conditions are typical in forensic applications, recent research in license plate recognition has primarily addressed scenarios with easily legible plates [Gong et al. 2022, Silva and Jung 2022, Laroca et al. 2023b].

Many researchers have proposed convolutional neural network-based approaches to tackle the SR problem [Lucas et al. 2019, Mehri et al. 2021, Liu et al. 2023]. Although these methods have demonstrated exceptional results, they often rely on computationally expensive deep architectures and prioritize increasing the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) metrics without considering the specific application. In the context of license plate recognition, such approaches do not effectively address confusion between closely resembling characters like ‘Q’ and ‘O’, ‘T’ and ‘7’, ‘Z’ and ‘2’, among others.

---

This work is a summary of a master’s dissertation [Nascimento 2023].

The second author served as an informal co-advisor for this work.

In this study, we propose a novel approach to enhance license plate super-resolution by integrating sub-pixel convolution layers and a Pixel Level Three-Fold Attention Module (PLTFAM). Our method extends [Mehri et al. 2021] by taking into account not only the image's pixel intensity values but also structural and textural information. To boost performance, we integrate an auto-encoder that extracts shallow features by squeezing and expanding the network with *PixelShuffle* (PS) and *PixelUnshuffle* (PU) layers. Additionally, we employ an Optical Character Recognition (OCR) model [Gonçalves et al. 2018] to extract features from license plate images during training, leading to enhanced SR performance and recognition rates.

In summary, the main contributions of this work are: **(i)** a SR approach that incorporates subpixel-convolution layers in combination with a PLTFAM; **(ii)** a novel perceptual loss that combines features extracted by an OCR model with L1 loss to reconstruct characters with the most relevant characteristics. This loss function allows the use of any OCR model for license plate recognition; and **(iii)** the datasets we built for this work, as well as the source code, are publicly available to the research community.

The rest of this work is structured as follows. Section 2 details our proposed network architecture and the new perceptual loss function. In Section 3, we present the experiments conducted and the corresponding results. Finally, Section 4 summarizes the findings, concluding this study.

## 2. Proposed Approach

This section presents our SR approach for enhancing feature extraction from low-resolution license plates. We incorporate ideas from [Zhang et al. 2023] to better capture structural and textural information from the license plate images. We also introduce a novel perceptual loss function that leverages an OCR model.

### 2.1. Network Architecture Modifications

As illustrated in Figure 1, the proposed approach for license plate SR builds upon the network architecture of [Mehri et al. 2021] and [Zhang et al. 2023]. It includes a Shallow Feature Extractor (SFE), Residual Dense Blocks (RDBs), a Feature Module (FM) module, and a Reconstruction Module (RM). In a nutshell, the RM combines the output of the FM module with two long-skip connections, one from the end of the SFE module and the other from the input image, to produce the high-resolution output. Our specific modifications are discussed in the following paragraphs.

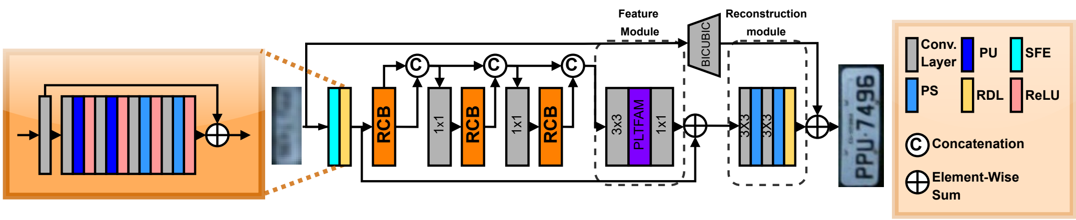



**Figure 1. The proposed architecture, which incorporates an autoencoder consisting of PS and PU layers for feature compression and expansion, respectively. This design aims to eliminate less significant features. The Two-fold Attention Module (TFAM) modules were replaced with PLTFAM modules throughout the network. PS = PixelShuffle; PU = PixelUnshuffle; RDL = Residual Dense Layers; SFE = Shallow Feature Extractor.**

The SFE block includes a $5 \times 5$ kernel convolution layer followed by an autoencoder that employs Depthwise-separable convolutional layers (DConvs), PU, and PS operations instead of conventional convolution layers, pooling, and upscale operations. The output of the layers is then combined with a skip connection from the initial convolution layers and processed by the RDBs.

In Figure 2, we present our modifications to the Multi-Path Residual Network (MPRNet)'s TFAM [Mehri et al. 2021] to create the PLTFAM. The design of this module is based on the following insights: **(i)** images are composed of the relationship between channels, where each channel contributes unique characteristics to form the final image, therefore, the extraction of these features is crucial for proper image restoration; **(ii)** the positional information of these essential features from the channels composing the images is required; **(iii)** traditional downscale and upscale operations rely on translational invariance and interpolation techniques, which are not able to learn a custom process for different tasks; and **(iv)** the module captures salient structure from the character fonts of the license plate, highlighting both structure and textural features in the image.

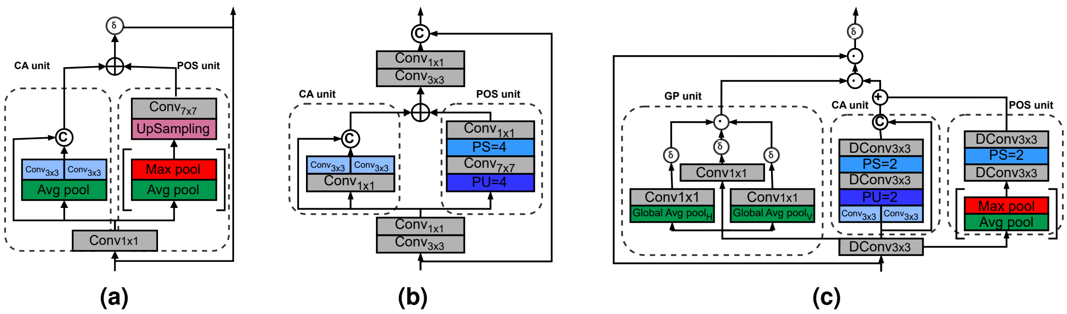



**Figure 2. Comparative illustration of the (a) TFAM in MPRNet [Mehri et al. 2021], (b) PixelShuffle Two-Fold Attention Module in our previous work [Nascimento et al. 2022], and (c) PixelShuffle Three-Fold Attention Module (ours).**

The Channel Unit (CA) module identifies and preserves relevant inter-channel relationship features by utilizing two parallel convolution layers. Their outputs are then concatenated and processed through convolution, PU, PS and DConv layers, summarizing features and enhancing image restoration.

The Positional Unit (POS) complements the CA module by determining the location of important features in the image. It extracts first-order statistics through pooling operations, combines the results, and processes them through DConvs and PS layers to restore the original feature map dimensions, highlighting relevant inter-channel relationship features and improving image restoration.

To extract critical structural, textural, and geometric features from the license plate, we incorporated the Geometrical Perception Unit (GP) branch. Inspired by [Zhang et al. 2023], it employs global average pooling in both vertical and horizontal directions. The output undergoes point-wise convolution and sigmoid function operations, followed by element-wise multiplication to obtain the final output.

The CA, POS, and GP units' outputs are combined through element-wise sum and multiplication operations, forming the attention mask. This mask enhances the input to the PLTFAM, effectively emphasizing key image features like inter-channel relationships, positional information, and structural details. This leads to improved image restoration.

The Residual Concatenation Blocks (RCBs) were improved by adding the PLT-FAM and dilated convolution layers to the bottleneck path of the adaptive residual blocks. This modification incorporates a wider context with an increased receptive field and preserves fine details in license plate images. Additionally, as shown in Figure 1, a reconstruction module was added to better aggregate fine details, consisting of two PS with a scale factor of 2, followed by DConv layers and consecutive RDBs.

### 2.2. Perceptual Loss

The proposed approach integrates a perceptual loss function to improve the accuracy of the SR method for license plate recognition. This loss function is tailored to consider the features anticipated by an OCR model, enhancing the system's accuracy.

$$PL = \frac{1}{n}\Big(\sum_{i=1}^{n}(H_i - S_i)^2 + \sum_{i=1}^{n}|f_{OCR}(H_i) - f_{OCR}(S_i)|\Big) \tag{1}$$

The perceptual loss function, as defined in Equation (1), consists of two terms: the Mean Squared Error (MSE) term and the feature extraction term. The MSE term measures the pixel value difference between the HR ($H_i$) and SR ($S_i$) license plate images, while the feature extraction term compares their representations from an OCR model ($f_{OCR}(\cdot)$). This loss function is adaptable to various OCR models, providing flexibility to incorporate novel models as they become available. We utilize the multi-task model proposed by [Gonçalves et al. 2018] due to its effectiveness in previous studies.

The MSE term penalizes significant errors between the expected and generated images, effectively improving overall image quality and preserving important structural information. Conversely, the feature extraction term, measured using the L1 loss, promotes robustness to noise and outliers while preserving sharp edges in the generated images. This dual approach enables a balanced evaluation, maintaining structural integrity while minimizing errors.

## 3. Experiments

Here, we describe the experiments carried out to validate the proposed method. We first discuss our experimental setup in Section 3.1. Subsequently, we provide a comprehensive analysis of the results obtained in Section 3.2. Finally, in Section 3.3, an ablation study is conducted to assess the contribution of each module integrated into the architecture.

### 3.1. Setup

We made use of license plate images obtained from two popular datasets: PKU [Yuan et al. 2017] and RodoSol-ALPR [Laroca et al. 2022a]. To the best of our knowledge, there is currently no public dataset that provides paired LR and HR images from real-world settings. Hence, we opted for these two datasets since they provide a wide range of scenarios under which the images were acquired.

The RodoSol-ALPR dataset consists of $20{,}000$ images, including vehicles with Brazilian license plates and Mercosur license plates[1]. It offers a diverse range of scenarios, including variations in license plate colors, lighting conditions, and character fonts

[1]Following [Laroca et al. 2022b, Nascimento et al. 2022, Silva and Jung 2022], we use the term "Brazilian" to refer to the layout used in Brazil prior to the adoption of the Mercosur layout.

(see Figure 3a). We followed the standard protocol defined in [Laroca et al. 2022a], allocating $40\%$ of the images for training, $20\%$ for validation, and $40\%$ for testing purposes.


(a) (b)

**Figure 3. a) Some license plate images from the RodoSol-ALPR dataset. The first two rows show Brazilian license plates, while the last two rows show Mercosur license plates. For scope reasons, we conduct experiments on license plates that have all characters arranged in a single row (i.e., 10K images); b) Examples from the PKU dataset, which also includes license plates with seven characters, albeit featuring varying layouts.**

The PKU dataset includes images grouped into G1-G5 representing various scenarios in mainland China, such as highways during the day (G1) and crosswalk intersections during the day or night (G5). Our experiments focused on G1-G3, totaling $2{,}253$ images with annotated license plate text [Zhang et al. 2021a]. The license plate images in the PKU dataset demonstrate high quality and legibility, as shown in Figure 3b. We followed the approach of [Zhang et al. 2021a, Laroca et al. 2022b], splitting $60\%$ of the images for training and validation, and the remaining $40\%$ for testing.

The HR images used in our experiments were created as follows. For each image, we first cropped the license plate region using the annotations provided. Afterward, we used the same annotations to rectify each license plate image so that it becomes more horizontal, tightly bounded, and easier to recognize. The rectified image is the HR image.

To generate LR versions of each HR image, we simulated lower-resolution effects based on [Zhang et al. 2021b]. We applied iterative random Gaussian noise to each HR image until reaching the desired degradation level for an LR image (i.e., SSIM $< 0.1$). To maintain the aspect ratio of the LR and HR images, we apply padding before resizing them to $20 \times 40$ pixels, resulting in an output shape of $80 \times 160$ pixels for an upscale factor of $4$. Examples of the generated license plate images for the RodoSol-ALPR and PKU datasets are shown in Figure 4a and Figure 4b, respectively.


(a) (b)

**Figure 4. a) Some HR-LR image pairs created from the RodoSol-ALPR dataset. b) Examples of HR-LR image pairs created from the PKU dataset.**

Our experiments were conducted using the PyTorch and Keras frameworks on a high-performance computer featuring an AMD Ryzen $9$ 5950X CPU, $128$ GB of RAM, and an NVIDIA Quadro RTX $8000$ GPU with $48$ GB of memory. We used the Adam optimizer with a learning rate of $10^{-4}$. The learning rate decreased by a factor of $0.3$

(down to $10^{-7}$) when no improvement in the loss function was observed. The training process was terminated after 20 epochs without any decrease in the loss function.

### 3.2. Experimental Results

In the literature, models are typically evaluated by the ratio of correctly recognized license plates to the total number of license plates in the test set [Wang et al. 2022, Silva and Jung 2022, Laroca et al. 2023a]. A license plate is considered correctly recognized if all characters are identified accurately. Considering our focus on low-resolution license plates, which are very common in forensic applications, we also report the recognition results considering partial matches (when at least 5 or 6 of the 7 characters are correctly recognized) as they may be useful in narrowing down the list of candidate license plates by incorporating additional information such as the vehicle's make and model.

The results of the license plate recognition experiment are shown in Table 1. The table demonstrates the recognition accuracy of HR and LR license plate images degraded by bicubic downsampling and recursive Gaussian noise. The difficulty of the task can be seen from the SSIM score, which ranges from 0 to 0.10, as illustrated in Figure 4a, where the license plate characters are barely distinguishable.

**Table 1. Results of our experiments. "All" refers to license plates where all characters were recognized correctly; $\geq 6$ and $\geq 5$ refer to license plates where at least 6 or 5 characters were recognized correctly, respectively.**

| | RodoSol-ALPR | | | PKU | | |
|---|---|---|---|---|---|---|
| | All | $\geq 6$ | $\geq 5$ | All | $\geq 6$ | $\geq 5$ |
| OCR (%) [Gonçalves et al. 2018] — no SR | | | | | | |
| HR | 96.6 | 98.6 | 99.0 | 99.4 | 99.9 | 99.9 |
| LR | 0.8 | 4.6 | 12.7 | 0.0 | 0.0 | 0.0 |
| OCR (%) [Gonçalves et al. 2018] — with SR | | | | | | |
| **Proposed** | **39.0** | **59.9** | **74.2** | **72.0** | **90.3** | **97.3** |
| [Nascimento et al. 2022] | 10.5 | 25.4 | 42.2 | 35.5 | 65.3 | 82.5 |
| [Mehri et al. 2021] | 1.45 | 7.0 | 17.4 | 22.5 | 49.2 | 70.6 |
| Average PSNR (dB) and SSIM | | | | | | |
| | | PSNR | SSIM | | PSNR | SSIM |
| **Proposed** | | 21.2 | **0.59** | | **18.3** | **0.61** |
| [Nascimento et al. 2022] | | **21.3** | 0.52 | | 18.1 | 0.54 |
| [Mehri et al. 2021] | | 16.8 | 0.38 | | 16.4 | 0.41 |

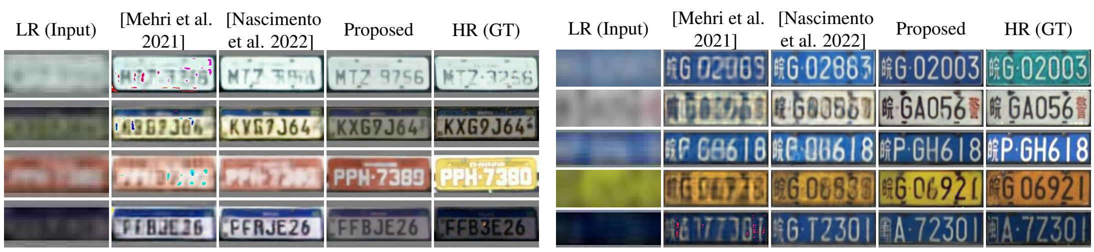
The proposed SR network outperformed the baseline models [Mehri et al. 2021, Nascimento et al. 2022] (see the second section of Table 1). The multi-task OCR model [Gonçalves et al. 2018] demonstrated remarkable improvement when applied to images reconstructed by our SR approach in both datasets, particularly in the PKU dataset, with a 14.8% higher recognition rate compared to the method proposed in our preliminary method [Nascimento et al. 2022] and a 26.7% higher accuracy compared to MPRNet [Mehri et al. 2021] for license plates with more than five correct characters.

For completeness, we detail in Table 1 the PSNR and SSIM obtained by each approach. Similar to what was observed in [Zhang et al. 2018, Lin et al. 2021], the PSNR metric seems inappropriate for this particular application, as our approach and the one proposed in [Nascimento et al. 2022] reached comparable values, despite ours leading to significantly better results achieved by the OCR model. The SSIM metric, on the other hand, seems to better represent the quality of reconstruction of license plate images, as the proposed method achieved considerably better SSIM values in both datasets.

The improved OCR network performance can be attributed to the effective extraction of textural and structural information enabled by the proposed GP unit, along with the optimized channel scaling and reorganization facilitated by the CA and POS units, which utilize pyramid and PS layers.

The variation in accuracy between the two datasets can be attributed to the diversity present in the RodoSol-ALPR dataset, which includes a range of layouts, lighting conditions, and character fonts, while PKU largely comprises license plates with a uniform layout, with less variation in the conditions under which the images were collected.

The visual comparison of the generated SR images using our technique and the baselines [Mehri et al. 2021, Nascimento et al. 2022] supports the results of the recognition experiments. Figure 5a and Figure 5b present four pairs of LR and corresponding SR images, along with the original HR image for reference. These images clearly demonstrate that our proposed approach outperforms its preliminary version [Nascimento et al. 2022] and MPRNet [Mehri et al. 2021] in terms of perceptual quality.




(a) (b)

**Figure 5. a) Typical examples of the images generated by the proposed approach and baselines in the RodoSol-ALPR dataset. b) Representative samples of the images generated by the proposed approach and baselines in the PKU dataset. GT = Ground Truth.**

Common issues observed in images produced by MPRNet include blurriness, where character edges blend into the license plate background, resulting in artifacts. This blurriness can also cause multiple characters' edges to blend together, leading to visual distortions. In contrast, the architecture proposed in [Nascimento et al. 2022] successfully reconstructs characters but introduces strong undulations, making them appear as part of the license plate background in some cases (as seen in the first row of Figure 5a). Our proposed model, however, consistently generates clear character edges, accurately reconstructs the original font, and avoids missing characters or incomplete lines.

When uncertain about character reconstruction, our model tends to generate characters most congruent with the LR input, as seen in the last row of Figure 5a and Figure 5b (e.g., "3" reconstructed as "J" and "Z" reconstructed as "2"). Incorporating a lexicon or vocabulary could address this issue by guiding the network to recognize characters based on specific license plate layouts. Furthermore, the network tends to generate similar background colors for different images, as observed in the second row of Figure 5a and the first row of Figure 5b; however, our analysis indicted that this has minimal impact on the achieved recognition results.

Finally, it is noteworthy that our model exhibits superior adaptability and ease of training when compared to approaches relying on generative adversarial networks, which often present instability and fall into mode collapse [Li et al. 2021, Saharia et al. 2023]. The attention-based design allows for more straightforward training, making it accessible to a broader range of practitioners and researchers in the field. Moreover, although we did not conduct specific experiments related to execution time, we anticipate that our approach delivers enhanced efficiency when contrasted with architectures based on diffusion models, which are known for being computationally expensive.

### 3.3. Ablation Study

We conducted an ablation study to assess the individual contributions of each unit in our integrated architecture. Four baselines were established: i) replacing the autoencoder with a $5 \times 5$ DConv layer for shallow feature extraction [Mehri et al. 2021]; ii) omitting the TFAM module and adjusting the output shape of the preceding layer; iii) substituting PS and PU layers with transposed and strided convolution layers [Shi et al. 2016]; and iv) replacing the perceptual loss with MSE, commonly used in SR [Wang et al. 2021, Liu et al. 2023]. The results are summarized in Table 2.

**Table 2. Recognition rates (%) achieved in the ablation study.**

| Approach | RodoSol-ALPR | | | PKU | | |
|---|---|---|---|---|---|---|
| | All | $\geq 6$ | $\geq 5$ | All | $\geq 6$ | $\geq 5$ |
| Proposed (w/o autoencoder) | 32.7 | 55.0 | 70.1 | **73.8** | 90.2 | 96.6 |
| Proposed (w/o TFAM) | 33.3 | 55.0 | 69.6 | 73.1 | 90.1 | 96.6 |
| Proposed (w/o PS and PU layers) | 34.3 | 54.8 | 68.5 | 70.4 | 89.9 | 96.7 |
| Proposed (w/o perceptual loss) | 35.6 | 57.3 | 71.9 | 72.4 | **91.4** | 97.1 |
| Proposed | **39.0** | **59.9** | **74.2** | 72.0 | 90.3 | **97.3** |

The experiments on the RodoSol-ALPR dataset showed that each component of the proposed system significantly contributed to its performance. The complete system achieved a recognition rate of 39.0%, while the best version without one of the components achieved 35.6%. Removing the autoencoder unit had the most detrimental effect, resulting in a recognition rate of 32.7%, as it plays a crucial role in extracting shallow features and guiding the network's reconstruction process.

On the other hand, the recognition rates on the PKU dataset were primarily improved by incorporating the PS and PU layers. We hypothesized that the other units are not necessary for this dataset, as it contains less complex images compared to RodoSol-ALPR. This might explain why some authors focused their ablation studies exclusively on the largest and most diverse dataset used in their experiments [Zhang et al. 2021a, Qin and Liu 2022, Wang et al. 2022].

## 4. Conclusions

This work proposes a novel super-resolution approach to enhance the recognition of low-resolution license plates. Our method combines subpixel-convolution layers (PS and PU) with PLTFAM and introduces a novel perceptual loss function that integrates OCR features with L1 loss and MSE. Through this approach, it reconstructs characters with distinctive attributes while simultaneously enhancing the overall image quality.

Our approach, utilizing PS and PU layers for custom scale operations, outperforms conventional methods by leveraging both structural and textural features. An autoencoder with PS and PU layers extracts shallow features and generates an attention mask, emphasizing relevant information during the super-resolution process.

We achieve superior recognition rates compared to notable baselines on two publicly available datasets from Brazil and mainland China. Notably, on the RodoSol-ALPR dataset, our method enables a recognition rate of $39.0\%$ to be achieved by the OCR model, outperforming baseline methods by a significant margin. Similarly, on the PKU dataset, our approach surpasses the baselines with an OCR recognition rate of $72.0\%$.

All datasets and source code used in our experiments are made available to encourage further research and development in license plate recognition super-resolution.

Last but not least, this work has generated the following contributions:

- A preliminary version of the proposed method was published at the *2022 Conference on Graphics, Patterns and Images (SIBGRAPI)* **[Nascimento et al. 2022]**;
- The proposed super-resolution method was published in the *Computers & Graphics* journal **[Nascimento et al. 2023]**;
- This study was recently honored at the *Workshop of Theses and Dissertations*, hosted by the committee of the *2023 Conference on Graphics, Patterns, and Images (SIBGRAPI)*, as **the best master's thesis completed in 2023**;
- This study was recently awarded the *Prêmio UFPR de Excelência Acadêmica*, being recognized by the Postgraduate Program in Informatics at the Federal University of Paraná (UFPR) as **the best master's thesis completed in 2023**.

## Acknowledgments

This work was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES) (*Programa de Cooperação Acadêmica em Segurança Pública e Ciências Forenses # 88881.516265/2020-01*), and also by the National Council for Scientific and Technological Development (CNPq) (# 309953/2019-7 and # 308879/2020-1).

## References

Gonçalves, G. R. et al. (2018). Real-time automatic license plate recognition through deep multi-task networks. In *Conf. on Graphics, Patterns and Images*, pages 110–117.

Gong, Y. et al. (2022). Unified Chinese license plate detection and recognition with high efficiency. *Journal of Visual Communication and Image Representation*, 86:103541.

Laroca, R., Cardoso, E. V., Lucio, D. R., Estevam, V., and Menotti, D. (2022a). On the cross-dataset generalization in license plate recognition. In *International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 166–178.

Laroca, R., Estevam, V., Britto Jr., A. S., Minetto, R., and Menotti, D. (2023a). Do we train on test data? The impact of near-duplicates on license plate recognition. In *International Joint Conference on Neural Networks (IJCNN)*, pages 1–8.

Laroca, R. et al. (2022b). A first look at dataset bias in license plate recognition. In *Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 234–239.

Laroca, R. et al. (2023b). Leveraging model fusion for improved license plate recognition. In *Iberoamerican Congress on Pattern Recognition (CIARP)*, pages 60–75.

Li, W., Fan, L., Wang, Z., Ma, C., and Cui, X. (2021). Tackling mode collapse in multi-generator GANs with orthogonal vectors. *Pattern Recognition*, 110:107646.

Lin, M., Liu, L., Wang, F., Li, J., and Pan, J. (2021). License plate image reconstruction based on generative adversarial networks. *Remote Sensing*, 13(15):3018.

Liu, A. et al. (2023). Blind image super-resolution: A survey and beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5):5461–5480.

Lucas, A. et al. (2019). Generative adversarial networks and perceptual losses for video super-resolution. *IEEE Transactions on Image Processing*, 28(7):3312–3327.

Mehri, A., Ardakani, P. B., and Sappa, A. D. (2021). MPRNet: Multi-path residual network for lightweight image super resolution. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 2703–2712.

Nascimento, V. (2023). Super-resolution towards license plate recognition. Master's thesis, Federal University of Paraná (UFPR).

Nascimento, V. et al. (2023). Super-resolution of license plate images using attention modules and sub-pixel convolution layers. *Computers & Graphics*, 113:69–76.

Nascimento, V., Laroca, R., Lambert, J. A., Schwartz, W. R., and Menotti, D. (2022). Combining attention module and pixel shuffle for license plate super-resolution. In *Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 228–233.

Qin, S. and Liu, S. (2022). Towards end-to-end car license plate location and recognition in unconstrained scenarios. *Neural Computing and Applications*, 34:21551–21566.

Saharia, C. et al. (2023). Image super-resolution via iterative refinement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):4713–4726.

Santos, M. et al. (2022). Face super-resolution using stochastic differential equations. In *Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 216–221.

Shi, W. et al. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1874–1883.

Silva, S. M. and Jung, C. R. (2022). A flexible approach for automatic license plate recognition in unconstrained scenarios. *IEEE Transactions on Intelligent Transportation Systems*, 23(6):5693–5703.

Wang, Y., Bian, Z.-P., Zhou, Y., and Chau, L.-P. (2022). Rethinking and designing a high-performing automatic license plate recognition approach. *IEEE Transactions on Intelligent Transportation Systems*, 23(7):8868–8880.

Wang, Z. et al. (2021). Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10):3365–3387.

Yuan, Y. et al. (2017). A robust and efficient approach to license plate detection. *IEEE Transactions on Image Processing*, 26(3):1102–1114.

Zhang, L. et al. (2021a). A robust attentional framework for license plate recognition in the wild. *IEEE Transactions on Intelligent Transportation Systems*, 22(11):6967–6976.

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 586–595.

Zhang, Y. et al. (2021b). Residual dense network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(7):2480–2495.

Zhang, Y., Huang, Y., Wang, K., Qi, G., and Zhu, J. (2023). Single image super-resolution reconstruction with preservation of structure and texture details. *Mathematics*, 11:216.