A Learning-Based Single-Image Super-Resolution Method for Very Low Quality License Plate Images

Alexandre Natã Vicente
Forensic Science Department
Civil Police of the Brazilian Federal District
Brasília, DF, Brazil, 70610-907
E-mail: ale.n.vicente@gmail.com

Helio Pedrini
Institute of Computing
University of Campinas
Campinas, SP, Brazil, 13083-852
E-mail: helio@ic.unicamp.br

***Abstract*—Spatial resolution enhancement of license plate images in real scenarios plays an important role in the fields of criminal investigation and forensic science. This paper presents a learning-based single-image super-resolution method that uses a priori knowledge of the input as the plate images captured at poor quality and very low resolution. The proposed method employs a decision tree to classify the input image and the classification results are used to weight the image patches in the reconstruction step. Additionally, a histogram equalization is performed to improve the performance of the classifier. Experiments conducted on synthetic and real-world images demonstrate that the proposed method is capable of producing satisfactory results.**

I. Introduction

Image enhancement methods for vehicle license plates play an important role in the fields of police investigation and forensic science, since the images are typically captured at poor quality and very low resolution, for instance, those from closed-circuit television (CCTV) with varying lighting and presence of noise. Conventional super-resolution techniques have serious limitations in such uncontrolled environments.

This work presents a super-resolution method for the aforementioned scenario, considering that the images meet the Brazilian license plate standard, which is a three-letter combination followed by a four-digit number. The proposed single image-based learning approach includes a histogram equalization pre-processing and employs a decision-tree learning algorithm.

Experimental results are presented and discussed, corroborating the good performance of the proposed method. Two data sets, one composed of synthetic characters and another composed of real license plate images acquired at very low resolution and with uncontrolled illumination conditions, are made available for future comparisons with other methods.

The remainder of the paper is organized as follows. Section II briefly reviews some image super-resolution methods. Section III describes the main stages of the proposed methodology. Section IV presents and analyzes the experimental results obtained by applying the method on synthetic and real images. Finally, Section V concludes the paper with final remarks and directions for future work.

II. Background

Super-Resolution (SR) [1], [2], [3], [4], [5], [6] is a family of methods used to enhance the spatial resolution of images. The main purpose of a super-resolution method is to increase the perceived detail and sharpness of images. Surveys on super-resolution methods and related concepts are provided by Nasrollahi and Moeslund [4] and van Ouwerkerk [5].

A major division in the family of SR methods refers to the amount of images used as input to the algorithm: single image or multiple images. The principles of the single-image and multiple-image super-resolution methods differ substantially.

In general, multiple-image methods aim at matching details present in the images and using the redundant information, as well as small differences among multiple images, to generate one or more high-resolution (HR) images.

The basic premise in multiple-image methods is that there is a low-resolution (LR) image sequence for a same scene, which is slightly different from others. Then, the LR images are modeled as degraded versions of the scene in HR [7]. The displacement between the images in low resolution should have subpixel precision, once the images would contain different information with integer pixel displacements.

On the other hand, single-image methods [2], [3] search for other sources of information to incorporate into the HR image. The single-image SR problem is inherently ill-posed, since many HR images can generate the observed LR image. Traditional interpolation methods (e.g. bilinear and bicubic interpolation) estimate the pixels of an HR image as a combination of LR pixels, operating as a smoothing process. Such smoothing can cause blurring artifacts in the super-resolved image when the natural image shows discontinuities such as edges and corners. Therefore, the purpose of single-image SR methods is to obtain better results than those using a priori statistical knowledge. Single-image SR methods can be further divided into two main categories: reconstruction-based methods and learning-based methods.

Reconstruction-based methods consist in defining a constraint for the HR image such that the reconstruction quality is increased. Despite the variety of methods existing in this category, the constraint is usually an explicitly and predetermined form of distribution or power function. Some examples include

978-1-5090-1897-0/16/$31.00 ©2016 IEEE

iterative back projection (IBP), projection onto convex sets (POCS), maximum likelihood (ML), maximum a posteriori (MAP), gradient profile prior (GPP) and fields of experts. Some reconstruction-based methods focus on edges, trying to rebuild them more clearly [8]. Other methods address the fact that the problem is ill-posed using regularization approaches, where a regularization parameter with a priori knowledge of the HR image reduces the universe of possible solutions to the ill-posed problem.

Learning-based methods, also known as hallucination methods, use a machine learning step to estimate the details of the super-resolved image. These methods employ a pixel-based procedure, where the pixel values of the super-resolved image are estimated through statistical learning or patch-based procedures (also known as example-based). The super-resolved image is estimated using a dictionary of patches in low and high resolution. In these patch-based methods, the LR image is analyzed and the most likely HR patches to have originated the LR patches are considered.

III. Proposed Method

This section describes the proposed super-resolution method and is organized as follows: III-A presents the digital image formation model considered in our method; III-B reports the main specifications for the Brazilian vehicle license plates and illustrates the segmentation process of their characters; III-C provides an overview of the proposed method; III-D explores some relevant aspects of the classifier used in the method.

A. Image Formation Model

An image acquired by a conventional digital camera is captured from the scene light, which passes through a set of camera lenses and mirrors, and reaches a sensor matrix that converts the received light intensity into electrical signal. This signal forms the digital image. This process [9] can be modeled as

$$\mathrm{LR} = (\mathrm{Blur} * \mathrm{HR})\downarrow + \eta \quad (1)$$

where LR is the digital image obtained at low resolution, Blur is a smoothing filter for lens distortion and focus, HR is the image at high resolution, $\downarrow$ denotes the subsampling operator, and η is an additive noise.

B. Brazilian License Plates and Character Segmentation

Motor-vehicle license plates are usually standardized, however, such standardization varies from country to country and even state to state within a country. In Brazil, license plates form a national vehicle registration database and the standardization occurs across the country and is regulated by the National Traffic Council (CONTRAN). Its main regulations include Resolution 231, 241, 309, 372 and 510.

Resolution 231 and its amendments are the basis of Brazilian plate specifications, which define design and size for the plates and characters present in them. Since this work focuses on the characters present in the Brazilian automobile license plates, it is important to mention that this regulation determines a unique typeface known as *Mandatory*, whose height of the characters is 63 mm.

Resolution 510 completely redrafted the design and specifications of Brazilian plates to match them with a common standard of Mercosur, modifying the character font type to *FE Engschrift* and character height to 65 mm and 53 mm for automobiles and motorcycles, respectively. Since this resolution will only enter into force in 2017, this work does not consider the new design once there are no available license plates for testing.

Figure 1 illustrates the Mandatory's characters used in the license plates, whose standardization provides a unified and favorable scenario for the application of a super-resolution algorithm to explore features for a specific domain.

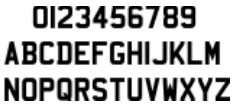

Fig. 1. Mandatory's characters used in the Brazilian license plates.

The input to the method proposed in this work is a single vehicle plate character. Therefore, it is necessary to segment the characters prior to their submission to the super-resolution algorithm. This segmentation can be done manually or through an automatic process. In this work, the segmentation was performed manually, according to the following rules: (i) each input does not contain pixels from neighboring characters and (ii) there is a margin of one or two pixels on every side of the bounding box that encompasses each character. Figure 2 illustrates the segmentation performed in the images.

Fig. 2. Illustration of character segmentation for a license plate at low resolution. Each colored bounding box contains a single image resulting from the four segmented characters.

C. Overview of the Proposed Method

Figure 3 illustrates the main stages of the proposed super-resolution method. Once the individual characters of the license plate have been segmented, each low-resolution (LR) image containing the character is provided as input. A padding operation with background color is applied to the input image. Then, an initial high resolution (HR) image is estimated through the bicubic interpolation technique. The HR reference images for digits 0, 1..., 9 are resized to the same dimensions as the HR image and the same padding operation is applied to this resized HR image.

The next step consists in building a dictionary of patches. The LR patches of the dictionary are obtained by applying both the blurring filter and the downsampling operator (Equation 1) to the reference images. The dictionary of patches is obtained by the association among HR and LR patches.

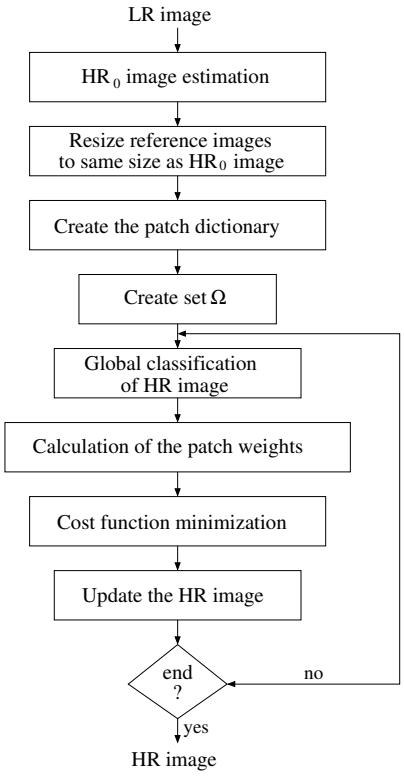

Fig. 3. Diagram of the proposed super-resolution method.

From the dictionary of patches, the set Ω is built, which stores the possible HR patches corresponding to each patch in the LR image. To determine which HR patches can match the patch in LR image, a similarity measure is employed among the patch in the LR image to be super-resolved and all LR patches that belong to the pre-built dictionary. Correlation was the similarity measure employed in this step, since it is very fast to compute it and due to the fact that this measure needs to be calculated multiples times.

A classifier is then applied to the HR image currently estimated. A decision tree was chosen as classifier. As output, we obtain the classification result and the probability information computed at training time.

From the results obtained with the classifier, we calculate the weights for the patches that must be used through Equation 2, which aims at minimizing a cost function. Gradient descent algorithm was employed in this minimization task.

After estimating the HR image that minimizes Equation 2, this estimate is updated and the algorithm stopping criterion (number maximum of iterations) is verified. If the criterion is satisfied, the algorithm finishes and the super-resolved image is considered the estimated HR image. Otherwise, the classification process continues.

$$\widehat{\text{HR}} = \arg\min \Bigg[\|\text{LR} - (\text{Blur} * \text{HR})\downarrow\|_2^2 + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w(i,j,k) \|R(i,j)\text{HR} - x(i,j,k)\|_2^2 \Bigg] \quad (2)$$

where $\widehat{\text{HR}}$ is the estimated HR image, LR is the low-resolution image (input to the algorithm), Blur is the blurring filter, $\downarrow$ is the downsampling operator, λ is the regularization parameter that controls the relative contribution between the fidelity factor to input $\|\text{LR} - (\text{Blur} * \text{HR})\downarrow\|_2^2$ and the learning factor $\left[\sum_{i,j} \sum_{k \in \Omega(i,j)} w(i,j,k) \|R(i,j)\text{HR} - x(i,j,k)\|_2^2\right]$, (i,j) is the pixel position at low resolution corresponding to the upper left corner of the LR patch, Ω is a set that stores the HR patches corresponding to each LR patch, $w(i,j,k)$ are the weights computed for each patch of low resolution (index k) belonging to $\Omega(i,j)$, $R(i,j)$ is the operator that selects the HR patch (size $n \times n$) of HR image in the position corresponding to (i,j) of low resolution (it is necessary to convert the position (i,j) for its corresponding in the HR image), $x(i,j,k)$ is the k-th HR patch that belongs to $\Omega(i,j)$.

The cost function, expressed in Equation 2, was chosen because it provides a trade-off, controlled by λ, between fidelity and learning. It also allows the weights of the patches used in the learning to be computed and smoothed taking both the similarity among the patches and global classification/similarity into account. Furthermore, the L^2-norm makes the equation differentiable, which is desirable for several minimization methods, as the gradient descent used in this work.

D. Algorithm Details

Certain steps of the proposed algorithm require a more detailed explanation: (i) the construction of the patch dictionary and set Ω; (ii) the classifier used in the experiments; (iii) the impact of the histogram manipulation for the classifier performance and, consequently, for the proposed super-resolution method; (iv) the weight calculation. This section provides additional information on such aspects.

1) Construction of patch dictionary and set Ω: Given an HR image, HR patches with dimension $m \times m$ and overlapping are created by simply copying the pixels that belong to the patch with dimensions $m \times m$ (for instance, patch 1 is composed of pixels (1:m-1, 1:m-1), patch 2 is composed of pixels (2:m, 1:m-1) and so on). To create the LR patches, initially the blurring filter (Equation 1) is applied. Then, the patches with dimension $m \times m$ are created in a similar way for the image resulting from the blurring filter application. The downsampling operator is applied to these patches, generating the LR patches. Finally, the dictionary is created by associating the LR patches and its corresponding HR patch. Figure 4 illustrates the main stages for creating the patches.

For the construction of the patch dictionary, each of one the reference images is used as HR image to the process described. Patches generated for all these reference images constitute the final dictionary of patches.

For the set Ω, the determination of which HR patches can match a patch of the LR image is performed through a similarity measure (correlation) among the patch of the LR image to be super-resolved and all LR patches that belong to the dictionary previously built. If the similarity measure is greater than a specified threshold T, then the HR patch corresponding to the

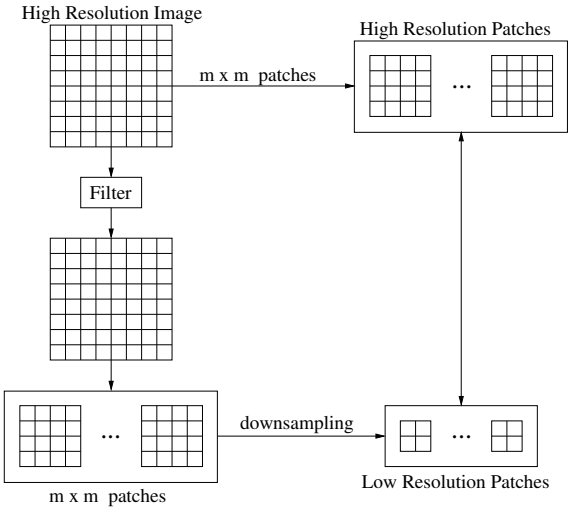

Fig. 4. Construction scheme for the patch dictionary used in the proposed method.

LR patch belonging to the dictionary is included into the set Ω as possible corresponding HR patch. For each position (i, j) in the LR image, only the *nPosition* patches with the highest measures of similarity (correlation) are kept in the set. Thus, set Ω can be represented as

$$\Omega(i,j) = \left\{ \begin{array}{l} \left\{ X^1_{m\times m}, X^2_{m\times m}, ..., X^{k_0}_{m\times m} \right\} \text{ reference } 0 \\ \left\{ X^1_{m\times m}, X^2_{m\times m}, ..., X^{k_1}_{m\times m} \right\} \text{ reference } 1 \\ \quad ... \\ \left\{ X^1_{m\times m}, X^2_{m\times m}, ..., X^{k_9}_{m\times m} \right\} \text{ reference } 9 \end{array} \right\} \quad (3)$$

where (i, j) is the position of the pixel at low resolution corresponding to the upper left corner of the LR patch, $X^1_{m\times m}, X^2_{m\times m}, ..., X^{k_j}_{m\times m}$ are the HR patches belonging to the dictionary whose its corresponding of low resolution satisfied the adopted criterion of similarity. To facilitate the computation of Equation 5, it is useful to separate the HR patches according to the reference that originated them (0, 1,... 9). Finally, the number of elements in the set Ω must also be limited to a pre-defined number (maximum number of patches of set Ω) and, if the number of elements is larger than this pre-defined number, the patches with lower similarity value must be removed from Ω.

2) Classifier and effect of histogram manipulation: A decision-tree classifier was employed in our method for the image classification task. This classifier was chosen because its outcome is easy to interpret, which is a desirable characteristic and also useful for criminology, since the results can be made available for users in forensic science. Furthermore, the decision-tree classifier provides the probability of each class obtained in the training set for each leaf. This information is used as weights ($w(i, j, k)$ in Equation 2) for patches in the set Ω.

The descending order of a given similarity metric between the image to be super-resolved and each reference was used as descriptors. Algorithm 1 presents a pseudocode for obtaining the descriptors.

Algorithm 1: Extraction of descriptors for the classifier.

input : RI$\{i\}$: reference image for pattern i
output: HR_image: image for which the descriptors are being calculated

```
for i = 1 to number of patterns do
    similarity(i) ← similarity_metric(HR_image, RIresize{i})
end
[orderedSimilarity, sortIndex] ← sort(similarity, 'descend');
descriptors ← sortIndex;
```

To illustrate the computing of descriptors, suppose an image (*HR_image*) whose similarity vector for a given similarity metric has resulted in the following vector: 0.12, 0.85, 0.34, 0.22, 0.15, 0.36, 0.72, 0.11, 0.28, 0.30 for $i = 1, 2, ..., 10$, respectively. Then, the descending order array (*orderedSimilarity*) would be 0.85, 0.72, 0.36, 0.34, 0.30, 0.28, 0.22, 0.15, 0.12, 0.11, whereas *sortIndex* (consequently, the descriptors) would be 2, 7, 6, 3, 10, 9, 4, 5, 1, 8.

In order to investigate the effect of the similarity metric and histogram manipulation on the performance of the proposed super-resolution method, the decision tree behavior was evaluated for a magnification factor of 4 using cross-validation, more specifically using leave-one-out cross-validation (LOOCV), and the set of images described in the Section IV-B. The similarity metrics evaluated in this work were correlation and structural similarity index (SSIM) [10]. The following histogram manipulations were compared:

1) no histogram manipulation.
2) histogram equalization applied only to the training and test images of the decision tree.
3) histogram equalization applied to the training and test images of the decision-tree, as well as the reference images.
4) contrast stretching applied to the training and test images of the decision-tree, as well as the reference images.
5) image normalization (pixel intensities are divided by L^1-norm of the image) applied to the training and test images of the decision-tree, as well as the reference images.

Each one of these scenarios was evaluated by using the reference images with white background and gray background, illustrated in Figure 5 and the descending order of the similarity metric as descriptors, and SSIM or correlation as similarity measure. The average error rates achieved with LOOCV are reported in Tables I and II.

Fig. 5. Reference images with white and gray background.

It is possible to observe that the best performance for the classifier occurs when the histogram equalization technique is

TABLE I
AVERAGE ERROR VALUES USING CORRELATION AS SIMILARITY METRIC.

Scenario	White Background	Gray Background
a	46.0%	45.0%
b	35.5%	38.5%
c	38.5%	30.5%
d	46.0%	44.5%
e	46.0%	45.0%

TABLE II
AVERAGE ERROR VALUES USING SSIM AS SIMILARITY METRIC.

Scenario	White Background	Gray Background
a	48.5%	43.5%
b	19.5%	21.0%
c	30.5%	33.5%
d	37.0%	35.0%
e	45.5%	52.5%

applied only to the training and test images with decision tree (that is, not applying the equalization to the reference images). Reference images with white background and SSIM generated the best combination.

Since SSIM [10] demonstrated to be more promising than the correlation measure for the classification performance, we used SSIM as similarity measure in the classification step. Furthermore, SSIM is consistent with the human visual system and is based on structural information, defined as

$$\text{SSIM}(\mathbf{f},\mathbf{g}) = \frac{(2\mu_f\mu_g + C_1)(2\sigma_{fg} + C_2)}{(\mu_f^2 + \mu_g^2 + C_1)(\sigma_f^2 + \sigma_g^2 + C_2)} \quad (4)$$

where μ_f is the mean of $\mathbf{f}$, μ_g is the mean of $\mathbf{g}$, σ_f^2 is the variance of $\mathbf{f}$, σ_g^2 is the variance of $\mathbf{g}$, and σ_{xy} is the covariance of $\mathbf{f}$ and $\mathbf{g}$. The SSIM measure is in the range of $[0..1]$, such that closer it is to 1, the more similar the images $\mathbf{f}$ and $\mathbf{g}$ are.

3) Weight calculation: Let $P(c)$, $c = 0, 1, ..., 9$ the probabilities for each class achieved by the classifier in the training set corresponding to the leaf for the classification of the estimated HR image, r the pattern that originated the element $k \in \Omega(i,j)$, and $S(k)$ the similarity measure calculated for patch k in the construction of set Ω. Then, the value w is calculated as

$$w(i,j,k) = P(r)\,\frac{S(k)}{\sum_{n\in\Omega(i,j)} S(n)} \quad (5)$$

IV. EXPERIMENTAL RESULTS

This section describes and evaluates the results obtained with the proposed method. Synthetic and real low-resolution images containing license plate characters were used in our experiments to assess the effectiveness of the method. The data sets are publicly available at [11].

A. Synthetic Images

This experiment aims at evaluating the proposed method under a controlled scenario, where LR images and their corresponding HR images are available to facilitate the comprehension of the main aspects of the method. Since RMSE and SSIM are metrics sensitive to image translations, a rectangle was cropped from the reference images precisely encompassing the character pixels. Then, both metrics were calculated by traversing this cropped reference image along the estimated HR image, such that the lowest RMSE and the highest SSIM values are used as the final metric outcomes.

From pictures of characters 0 to 9 in high resolution (height of 40 pixels and width of 28 pixels), LR images were generated (height of 10 pixels and width of 7 pixels) with different values of signal-to-noise ratio by adding a white Gaussian noise to the LR images. The signal-to-noise ratio values used were 5 dB, 10 dB and 15 dB. For each value, 20 LR images were generated for each character, resulting in 600 LR images. They were divided into groups, training and test sets, with equal distribution with respect to the number of images and noise level.

The classifier was developed on the training set and then applied to each sample in the test set. For comparison purpose, the following methods available in the literature were also applied to the test sets: iterative back projection (IBP) [12], kernel ridge regression (KRR) [2], gradient profile prior (GPP) [8] and bicubic interpolation.

The configuration used in this experiment was: images with white background, scaling factor of 4, patch size of 16×16, 3×3 mean filter as blurring filter, threshold $T = 0.8$ for similarity measure, *nPosition* = 3, maximum number of patches $n = 1000$ for set Ω, $\lambda = 0.2$, maximum number of iterations $N = 20$. Table III and Figure 6 show the results for the synthetic data set.

TABLE III
AVERAGE VALUES (IN PERCENTAGE) OF RSME AND SSIM FOR SYNTHETIC IMAGES SIMULATED WITH SNR EQUAL TO 5, 10 AND 15.

Method	SNR = 5 dB		SNR = 10 dB		SNR = 15 dB	
	RMSE	SSIM	RMSE	SSIM	RMSE	SSIM
Ours	27.24	55.71	17.32	77.28	14.15	85.70
IBP	40.18	33.39	33.53	51.02	29.53	51.17
KRR	31.46	41.60	26.03	52.93	23.42	58.77
GPP	39.53	37.08	31.12	52.96	25.85	63.07
Bicubic	31.46	41.61	26.03	52.93	23.42	58.77

From the achieved results, it is possible to observe that our method produces quantitative superior results when compared to the other evaluated approaches. Furthermore, even in severe image conditions due to low resolution and presence of noise, our method was capable of providing satisfactory qualitative results both in terms of edge definition and contrast response. For instance, the IBP method did not generate proper HR images, whereas KRR had similar results to the bicubic interpolation.

B. Real Images

In this scenario, the experiment uses characters extracted from Brazilian license plate images at low resolution under

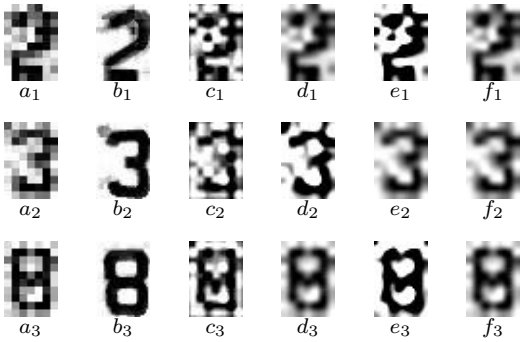

Fig. 6. Examples of super-resolved synthetic images obtained with the proposed method. Images "a", "b", "c", "d", "e" and "f" were scaled by a factor of 4 through nearest neighbor, proposed method, IBP, KRR, GPP and bicubic interpolation, respectively. Images with index 1 have SNR = 5, whereas others have SNR = 10.

uncontrolled lighting conditions. The proposed method was applied to a set of real images composed of 200 images (20 images for each digit 0, 1,..., 9).

The classifier was trained using 90% of the samples and applied to the 10% remaining. The configuration used in this experiment was: images with gray background, scaling factor of 4, patch size of 16×16, 3×3 mean filter as blurring filter, threshold $T = 0.8$ for similarity measure, *nPosition* = 3, maximum number of patches $n = 1000$ for set Ω, $\lambda = 0.2$, maximum number of iterations $N = 50$. Figure 7 shows the qualitative results for the real images obtained with the evaluated approaches. A quantitative comparison is not possible, since there is no ground-truth for the real license plate images.

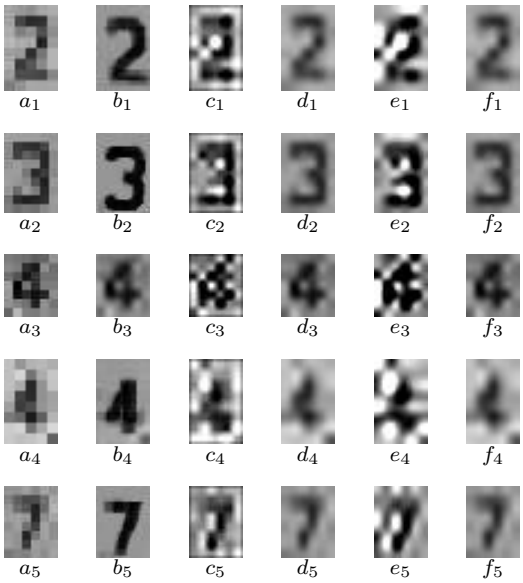

Fig. 7. Examples of super-resolved real images obtained with the evaluated approaches. Images "a", "b", "c", "d", "e" and "f" were scaled by a factor of 4 through nearest neighbor, proposed method, IBP, KRR, GPP and bicubic interpolation, respectively.

It is possible to observe that the proposed method generates images with better results in terms of edge definition and contrast. Even for some images where the results are degraded, the resulting images obtained with our method are similar to bicubic interpolation (b_3 in Figure 7). Furthermore, it is worth mentioning that the other evaluated methods failed under such severe conditions of noise and varying lighting.

V. Conclusions and Future Work

This work addressed the super-resolution problem in images that follow the Brazilian license plate standard. A single image-based learning method for very low quality license plate images is proposed and discussed, which explores the prior knowledge that the images contain only digits with known format.

The influence of histogram manipulation in the performance of the super-resolution method was also discussed, which is little analyzed in other approaches available in the literature. Although the proposed method has been applied to real images at very low resolution, presence of noise, and uncontrolled lighting conditions, the qualitative results demonstrated its effectiveness, outperforming other super-resolution techniques on the investigated specific domain.

As directions for future work, we intend to expand the data sets used in our experiments, as well as evaluate the use of other classifiers during the learning process.

Acknowledgments

The authors are thankful to FAPESP (grant #2011/22749-8), CNPq (grants #307113/2012-4 and #487186/2013-3), Federal District Funding Agency (FAP/DF), Ilaraine Acácio Arce Expert Forensic Foundation (FPCIAA), University of Brasília (UnB) and Civil Police of the Brazilian Federal District (PCDF) for their financial support.

References

[1] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," in *European Conference on Computer Vision*. Springer, 2014, pp. 184–199.

[2] K. I. Kim and Y. Kwon, *Example-based Learning for Single-Image Super-Resolution*. Berlin, Heidelberg: Springer, 2008, pp. 456–465.

[3] ——, "Single-Image Super-Resolution using Sparse Regression and Natural Image Prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1127–1133, 2010.

[4] K. Nasrollahi and T. B. Moeslund, "Super-Resolution: A Comprehensive Survey," *Machine Vision and Applications*, vol. 25, no. 6, pp. 1423–1468, 2014.

[5] J. van Ouwerkerk, "Image Super-Resolution Survey," *Image and Vision Computing*, vol. 24, no. 10, pp. 1039–1052, 2006.

[6] G. Polatkan, M. Zhou, L. Carin, D. Blei, and I. Daubechies, "A Bayesian Nonparametric Approach to Image Super-Resolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 2, pp. 346–358, 2015.

[7] R. Gerchberg, "Super-Resolution through Error Energy Reduction," *Journal of Modern Optics*, vol. 21, no. 9, pp. 709–720, 1974.

[8] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, "Image Super-Resolution using Gradient Profile Prior," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.

[9] Y. Tian, K.-H. Yap, and Y. He, "High Resolution Vehicle License Plate Reconstruction using Soft Recognition Learning," in *8th International Conference on Information, Communications and Signal Processing*. IEEE, 2011, pp. 1–5.

[10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[11] SR-LicensePlate, "License Plate Images," 2016, https://github.com/alevicente/SR-LicensePlate/tree/master/Datasets/.

[12] M. Irani and S. Peleg, "Improving Resolution by Image Registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, no. 3, pp. 231–239, 1991.