The effect of image enhancement algorithms on convolutional neural networks

José A. Rodríguez-Rodríguez, Miguel A. Molina-Cabello, Rafaela Benítez-Rochel, Ezequiel López-Rubio
Department of Computer Languages and Computer Science. University of Málaga
Bulevar Louis Pasteur, 35, 29071, Málaga, Spain
Biomedic Research Institute of Málaga (IBIMA)
C/ Doctor Miguel Díaz Recio, 28, 29010, Málaga, Spain

***Abstract*—Convolutional Neural Networks (CNNs) are widely used due to their high performance in many tasks related to computer vision. In particular, image classification is one of the fields where CNNs are employed with success. However, images can be heavily affected by several inconveniences such as noise or illumination. Therefore, image enhancement algorithms have been developed to improve the quality of the images. In this work, the impact that brightness and image contrast enhancement techniques have on the performance achieved by CNNs in classification tasks is analyzed. More specifically, several well known CNNs architectures such as Alexnet or Googlenet, and image contrast enhancement techniques such as Gamma Correction or Logarithm Transformation are studied. Different experiments have been carried out, and the obtained qualitative and quantitative results are reported.**

***Index Terms*—convolutional neural networks, image classification, brightness, image enhancement techniques**

I. Introduction

Convolutional Neural Networks (CNNs) are deep neural networks whose convolutional layers alternate with subsampling layers, reminiscent of simple and complex cells in the primary visual cortex [1]. Due to their recognition accuracy, CNNs are highly useful in several disciplines as speech recognition [2] and natural language processing [3] but are dominant in computer vision: image classification [4], image segmentation [5], objects in image detection, [6], image style transfer [7], etc. In many cases, these methods even exceed the performance that a human can provide.

The ImageNet Large Scale Visual Recognition Challenge (ILSRVC) [8] has encouraged the development of new and better CNN architectures over the years. Architectures like AlexNet [9], in 2012, VGGNet [10], GoogleNet [11] in 2014, and ResNet [4] in 2015 have brought a high number of improvements and innovations to the field of object recognition. It is well known the importance of network architecture in achieving better performances by making changes in different layers of the network, using several types of activation functions or learning algorithms. However, the impact of image quality on the classification results of such networks is often overlooked. It is fundamental to notice at this point that preprocessing techniques for image classification are critical to training an efficient neural network [12]. In most studies, networks are trained and tested on high-quality image data sets, but in real-world applications, image quality may be low due to external factors so final recognition can be degraded.

In this work, it has been studied whether image enhancement algorithms can be used in image pre-processing to improve the performance of CNNs. More specifically, our focus is on the effect of contrast enhancement techniques. Image enhancement is the procedure of improving the quality and information content of original data before processing [13]. Its purpose is to amplify certain image features for analysis, diagnosis, and display. Typically, image enhancement includes intensity and contrast manipulation, noise reduction, edges sharpening and filtering, etc. Contrast enhancement plays a crucial role in image processing applications, such as digital photography, medical image analysis, remote sensing, LCD processing, and scientific visualization. The problem of image contrast enhancement is subjective and application dependent, and this justifies the presence of numerous algorithms in the literature. We have analyzed the impact on CNN's of four different techniques: Gamma Correction (GC) [13], Histogram Equalization (HE) [14], Contrast Limited Adaptive Histogram Equalization (CLAHE) [15] and Logarithm Transformation (LT) [13]. All of them are well known spatial domain techniques, they operate on the pixels of an image, thus it enhances the overall contrast of an image.

The next section describes the methodology which guides this work. Then, Section III exhibits the considered methods, the selected dataset, and the obtained results. Finally, conclusions are provided in Section IV.

II. Methodology

In this section, our proposed methodology to improve the object detection performance of Convolutional Neural Networks is detailed. Let us assume that ϕ is the pixel value (in digital numbers) of a digital image under ideal illumination conditions. Let H be the number of bits employed to represent ϕ, so that $\phi \in \left\{0, ..., 2^H - 1\right\}$. Low illumination conditions can be modeled as a division of ϕ by a positive brightness scale factor b to yield the low illumination pixel value (in digital numbers):

$$\tilde{\phi} = \frac{1}{b}\phi \quad (1)$$

The low illumination pixel value is quantized as follows:

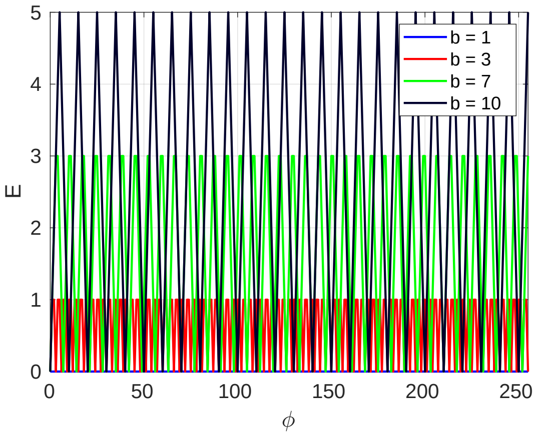

Fig. 1: Quantization Error (E) for values of the bright scale (b) in the set $\{1,3,7,10\}$.

$$\hat{\phi} = \text{floor}\left(\tilde{\phi} + \frac{1}{2}\right) = \text{floor}\left(\frac{1}{b}\phi + \frac{1}{2}\right) \tag{2}$$

where floor rounds towards $-\infty$. Therefore, low illumination incurs in an absolute quantization error which is given by:

$$E = \left| b \cdot \hat{\phi} - \phi \right| = \left| b \cdot \text{floor}\left(\frac{1}{b}\phi + \frac{1}{2}\right) - \phi \right| \tag{3}$$

The higher the brightness scale factor b, the higher the absolute error E. This means that, the lower the illumination, the more information which is lost (see Figure 1). In order to overcome this inconvenience, we propose to apply image enhancement procedures to the quantized low illumination values $\hat{\phi}$ to yield restored pixel values $\bar{\phi}$ which are as close as possible to the ideal values ϕ. This way the information loss is alleviated, which facilitates the object detection task of the Convolutional Neural Network, which is provided $\bar{\phi}$ as input.

III. Experimental Results

In this paper, several experiments have been performed in order to analyze the impact of four contrast enhancement techniques on the performance of three well-known convolutional neural network architectures. This section is organized as follows. First, the methods used are explained (in Subsection III-A). Next, it is detailed the dataset employed in the experiments (in Subsection III-B). After that, a qualitative analysis is carried out to get a good perception of how the cited techniques improve the contrast of two images (in Subsection III-C. Finally, quantitative results are presented (in Subsection III-D).

A. Methods

It has been considered three well-known architecture for the analysis: AlexNet [9] (8 layers), GoogleNet [11] (22 layers) and ResNet [4] architecture (50 layers). These CNNs have been implemented through Torchvision package from PyTorch[16] library and these models had been pre-trained using a sub-set of images from ImageNet dataset published for ILSRVC2012 challenge [17].

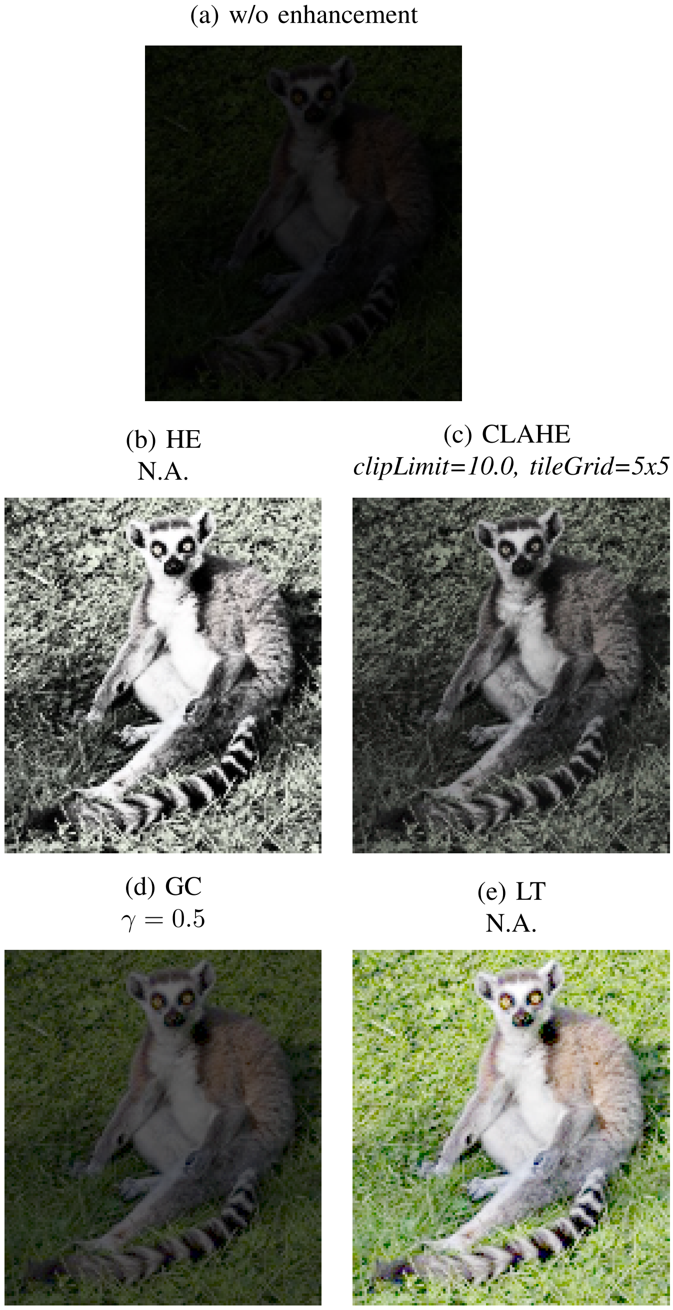

Fig. 2: Comparison between the contrast enhancement techniques applied over the Image 5944 from the ImageNet dataset (ILSRVC2012) where bright scale was set to six: (a) without contrast enhancement, (b) HE, (c) CLAHE, (d) GC and (e) LT.

In addition, among the different contrast enhancement algorithms which can be found in the literature, this work has considered Gamma Correction (GC), Histogram Equalization

(a) w/o enhancement

(b) HE
N.A.

(c) CLAHE
clipLimit=10.0, tileGrid=5x5

(d) GC
$\gamma = 0.5$

(e) LT
N.A.

Fig. 3: Comparison between the contrast enhancement techniques applied over the Image 328 from the ImageNet dataset (ILSRVC2012) where bright scale was set to one: (a) without contrast enhancement, (b) HE, (c) CLAHE, (d) GC and (e) LT.

(HE), Contrast Limited Adaptive Histogram Equalization (CLAHE) and Logarithm Transformation (LT). LT and HE do not have any parameter to tune (noted as *N.A.*), GC has one parameter (γ) and CLAHE has two parameters (*tileGrid* and *clipLimit*). GC, CLAHE and HE have been implemented through OpenCV library [18], while LT through logarithm function from NumPy library [19].

The reported experiments have been carried out on a 64-bit Personal Computer with a six-core Intel Core i7-9750 4.50 GHz CPU, 16 GB RAM, and an NVIDIA GeForce RTX 2080 GPU.

B. Dataset

A sub-set of one thousand (1,000) images has been selected randomly, following to a uniform distribution, from the ILSRVC2012 validation set that is formed by 50,000 images in total. The images are encoded in RGB format where each color is encoded with 8 bits. Also, in this dataset, the images have been hand labeled with the presence or absence of 1,000 object categories. The ILSRVC2012 dataset can be downloaded from its website[1].

C. Qualitative Results

In this subsection, a comparison between the proposed contrast enhancement algorithms is presented in a visual way. The idea is to get a better intuition about the effect that these algorithms produce on RGB images.

Figure 2 shows a Madagascar cat, corresponding with the Image 5944, and their enhanced versions. The original image has been obtained using a bright scale of five and it is placed at the top of the figure in Figure 2 (a). The pixel values range of this image goes from 0 to 50, approximately. In addition, it is observed that HE and CLAHE are good transformation because they increase the contrast and brightness at the same time. As known, CLAHE limits the noise amplification thus it could be more suitable than HE to improve the CNN performance. Also, it can be deduced GC improves the contrast and brightness compared with the original image and it keeps a better color balance compared with the image when HE or CLAHE are applied. Although, the best distinguished image is generated by LT, where it can be observed an improvement, in contrast, brightness and balance of colors.

As commented before, the original image in Figure 2 was synthesized to be dark, therefore it could be interesting to apply the enhancement algorithms to another brighter image. For this purpose, it has been generated Figure 3 that shows the same comparison for Image 328 with a bright scale of one. This image is a bird called a kite. In this case, LT is not the best technique. Indeed, it seems to have a worse performance compared with GC, HE and CLAHE that generates images with a similar appearance. The most different image is when LT is applied because the image is brighter.

D. Quantitative Results

The main objective of this work is to study four well-known contrast enhancement algorithms to improve the performance of a CNN when illumination is low or, in other words, for dark input images. Regarding the algorithms, LT and HE cannot be optimized because they do not have any parameters to tune. But CLAHE and GC have some parameters that can be chosen to get the best performance. This way, in this Subsection it is presented the next parametric analysis: GC and CLAHE optimization study, and a comparison between the other considered algorithms and the optimized ones.

In order to compare the performance achieved by each proposal, the accuracy-1 (*acc-1*) has been selected as the

[1] http://www.image-net.org/challenges/LSVRC/2012/

performance metric. This measure provides real values from 0 to 1, where higher is better. It is defined as follows:

$$acc\text{-}1 = \frac{SP\text{-}1}{TP} \quad (4)$$

where SP-1 is the number of successful predictions considering the greatest probability at the model output; and TP is the total number of predictions carried out.

First of all, continuing with the analysis of Subsection above, in Table I is shown the five best predictions and their corresponding probabilities for Image 5944 and 328. In all the studied cases the prediction label is Madagascar cat and kite, respectively, as expected by the ground table values. Regarding the probability results, GC and LT produce a considerable probability boost for Image 5944. But these are not good algorithms for Image 328 where the photograph is well illuminated, indeed probability is reduced, especially for LT algorithm.

From this analysis, it can be intuited that the proposed contrast enhancement algorithms studied in this paper have different behavior depending on the image brightness.

Then, the effect of gamma correction on the performance is shown in Figure 4, where four configurations for parameter γ have been tuned. The selected configurations are $\gamma = \{0.1, 0.3, 0.5, 0.7\}$. As it can be observed, the improvement is higher for darker darker images. All CNNs present similar acc-1 drop around 10% for the range of bright scale used in this study. Furthermore, the best performance is reached by ResNet-50, as expected, because this CNN has more number of layers than the other ones. From these plots, it could be stated that the optimal value of γ is around 0.5 which it will be used in the final comparison.

Respecting the CLAHE optimization analysis, it is shown in Figure 5. As commented in Section II, CLAHE has two degrees of freedom in the present implementation: *tileGrid* and *clipLimit*. The parameter *tileGrid* is related with the window size employed to apply the histogram equalization algorithm, while *clipLimit* which is related with the gain limit used by the algorithm. The former was fixed to 5x5, so that, the considered parameter to be optimized in the present study is *clipLimit*, where the considered tuned configurations are $clipLimit = \{5, 10, 30, 50\}$. Figure 5 shows the CLAHE optimization plots where it can be deduced there is a similar improvement than the GC plots shown in Figure 4. In this case, the optimal value of *clipLimit* seems to be more clear for GoogleNet and ResNet-50 to be around 10. However, it is less clear for AlexNet because the curves are crossed in different points. A fair choice of *clipLimit* could be 10 following the criteria of maximizing accuracy-1 for higher values of bright scale.

At last, Figure 6 shows a comparison between the optimized and the other enhancement techniques. The best performance for higher values of bright scale is reached by LT, specially for darker images. Indeed, only this algorithm has a positive slope relative to the bright scale. However, LT has not a good performance for bright scale values below four. Furthermore,

Image	Algorithm	Predictions	Probability
5944	W/O	**1.- Madagascar cat**	**0.7288**
		2.- indri	0.0393
		3.- Egyptian cat	0.0354
		4.- meerkat	0.0202
		5.- whippet	0.0126
	HE	1.- **1.- Madagascar cat**	**0.8068**
		2.- indri	0.1730
		3.- meerkat	0.0029
		4.- Egyptian cat	0.0021
		5.- mongoose	0.0016
	CLAHE	**1.- Madagascar cat**	**0.7912**
		2.- indri	0.1607
		3.- badger	0.0082
		4.- meerkat	0.0049
		5.- mongoose	0.0040
	GC	1.- **1.- Madagascar cat**	**0.9677**
		2.- indri	0.0204
		3.- Egyptian cat	0.0020
		4.- meerkat	0.0014
		5.- whippet	0.0008
	LT	1.- **1.- Madagascar cat**	**0.9584**
		2.- indri	0.0321
		3.- Egyptian cat	0.0017
		4.- meerkat	0.0008
		5.- banana	0.0006
328	W/O	**1.- kite**	**0.9442**
		2.- bald eagle	0.0195
		3.- coucal	0.0056
		4.- limpkin	0.0020
		5.- house finch	0.0012
	HE	**1.- kite**	**0.9673**
		2.- bald eagle	0.0110
		3.- coucal	0.0056
		4.- limpkin	0.0015
		5.- hen	0.0006
	CLAHE	**1.- kite**	**0.9454**
		2.- bald eagle	0.0087
		3.- coucal	0.0084
		4.- prairie chicken	0.0018
		5.- house finch	0.0014
	GC	1.- **kite**	**0.9191**
		2.- bald eagle	0.0368
		3.- coucal	0.0068
		4.- house finch	0.0029
		5.- bittern	0.0018
	LT	1.- **kite**	**0.7836**
		2.- bald eagle	0.1677
		3.- coucal	0.0042
		4.- bittern	0.0030
		5.- house finch	0.0024

TABLE I: Predictions for Images 5944 (Figure 2) and 328 (Figure 3) provided by GoogleNet. Each column exhibits the image, the algorithm, the predicted class and its probability, respectively. Bright scale was set to 6 for the first image and 1 for the second one. Best five predictions ordered by their probability are shown for each tuned configuration. Ground truth class is highlighted in **bold**.

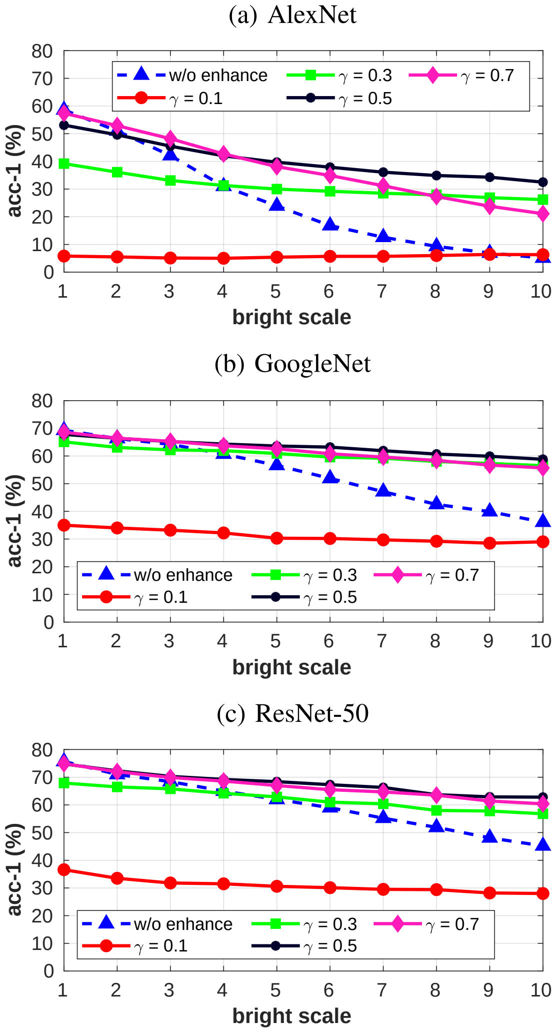

Fig. 4: Gamma Correction Analysis to find the optimal value of γ for the CNNs under study: (a) AlexNet, (b) GoogleNet, (c) ResNet-50.

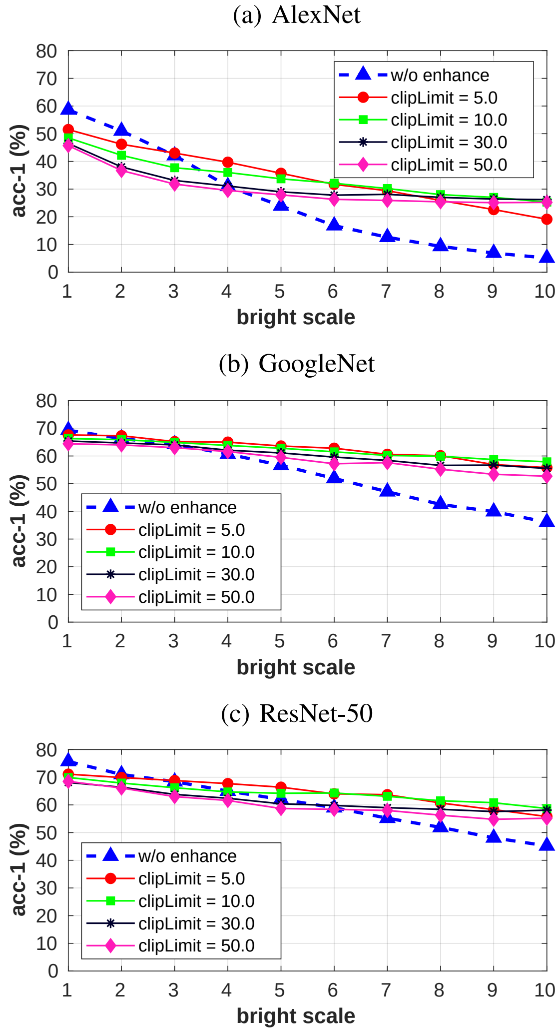

Fig. 5: CLAHE algorithm Analysis to find the optimal value of clipLimit for the CNNs under study: (a) AlexNet, (b) GoogleNet, (c) ResNet-50.

it is also observed that all the proposed techniques improves acc-1 for higher values of bright scale. CLAHE and HE have similar behavior as expected, because CLAHE is based on HE. Respecting to GC, it has better performance than the plots based on HE.

Regarding considered CNNs, the greatest improvement happens when LT is applied on AlexNet CNN where it can be seen an acc-1 boost around 40% compared with no enhanced images.

IV. CONCLUSIONS

In this paper four well-known contrast enhancement algorithms are proposed to improve the accuracy (acc-1) of CNNs when illumination is low: HE, CLAHE, GC, and LT. With this objective, these enhancement techniques have been applied on three pre-trained CNNs: AlexNet, GoogleNet, and ResNet-50. Also, a qualitative analysis shows how the proposed algorithms exhibit different behavior depending on the brightness level of the input images. Next, it has been demonstrated these four algorithms improve the acc-1 metric for dark images. However, it presents some issues when the images become brighter, especially the LT algorithm. In this demonstration, first GC and CLAHE optimization studies have been performed in order to get the optimal parameter values for each one of them. In this sense, regarding the tuned configurations, this optimization process concludes the optimal parameter for GC and CLAHE are $\gamma = 0.5$ and $clipLimit = 10.0$ (with $tileGrid = 5x5$), respectively. Additionally, experimental results confirm that LT gets the best performance in terms of acc-1 when applied on the three CNNs for dark images. An increase of 40% is achieved for a bright scale of 10. Although, this algorithm worsens the performance of the CNNs for low values of bright scale (from 1 to 3, approximately). GC, HE and CLAHE are more robust than LT in terms of brightness because they have a better behavior for brighter images and achieve a moderated

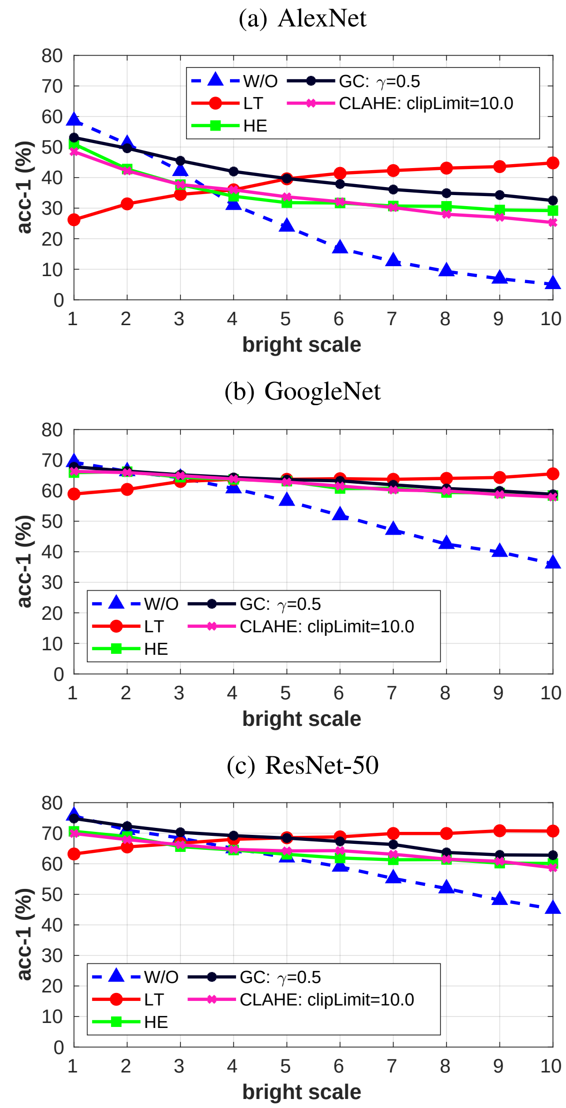

Fig. 6: Comparison between the contrast enhancement techniques used in this paper: (a) without contrast enhancement, (b) HE, (c) CLAHE, (d) GC and (e) LT.

performance improvement for dark images. As a conclusion, it can be stated that LT is a good option to be used as an unsupervised pre-processing algorithm in CNNs to improve acc-1 when images are darker, while the other considered enhancement algorithms are good choices for a wider range of brightness.

ACKNOWLEDGMENT

This work is partially supported by the Ministry of Economy and Competitiveness of Spain under grants TIN2016-75097-P and PPIT.UMA.B1.2017. It is also partially supported by the Ministry of Science, Innovation and Universities of Spain under grant RTI2018-094645-B-I00, project name Automated detection with low cost hardware of unusual activities in video sequences. It is also partially supported by the Autonomous Government of Andalusia (Spain) under grant UMA18-FEDERJA-084, project name Detection of anomalous behavior agents by deep learning in low cost video surveillance intelligent systems. All of them include funds from the European Regional Development Fund (ERDF). The authors thankfully acknowledge the computer resources, technical expertise and assistance provided by the SCBI (Supercomputing and Bioinformatics) center of the University of Málaga. They have also been supported by the Biomedic Research Institute of Málaga (IBIMA).They also gratefully acknowledge the support of NVIDIA Corporation with the donation of two Titan X GPUs used for this research. The authors also thankfully acknowledge the grant of the Universidad de Málaga.

REFERENCES

[1] T. N. Hubel, D. H.; Wiesel, "Receptive fields of single neurones in the cat's striate cortex," *The Journal of Physiology*, vol. 148, no. 3, pp. 574–591, 1959.

[2] Y. Zhang, M. Pezeshki, P. Brakel, S. Zhang, C. L. Y. Bengio, and A. Courville, "Towards end-to-end speech recognition with deep convolutional neural networks," 2017.

[3] W. Wang and J. Gang, "Application of convolutional neural network in natural language processing," 07 2018, pp. 64–70.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.

[6] M. A. Molina-Cabello, R. M. Luque-Baena, E. López-Rubio, and K. Thurnhofer-Hemsi, "Vehicle type detection by ensembles of convolutional neural networks operating on super resolved images," *Integrated Computer-Aided Engineering*, vol. 25, no. 4, pp. 321–333, 2018.

[7] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 2414–2423.

[8] Stanford Vision Lab, "Imagenet, large scale visual recognition challenge 2012 (ilsvrc2012)," http://www.image-net.org/challenges/LSVRC/2012/, 2012, [Online; accessed 31-October-2019].

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014.

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.

[12] K.-M. Koo and E.-Y. Cha, "Image recognition performance enhancements using image normalization," *Human-centric Computing and Information Sciences*, vol. 7, 12 2017.

[13] R. Gonzalez and R. Woods, *Digital Image Processing*. Upper Saddle River, New Jersey: Prentice Hall, 2002.

[14] S. Patel and M. Goswami, "Comparative analysis of histogram equalization techniques," in *2014 International Conference on Contemporary Computing and Informatics (IC3I)*, 2014, pp. 167–168.

[15] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics gems IV*. Academic Press Professional, Inc., 1994, pp. 474–485.

[16] PyTorch Contributors, "PyTorch Documentation. Version 1.2.0," https://pytorch.org/docs/1.2.0/, 2019, [Online; accessed 22-March-2020].

[17] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, 2015.

[18] OpenCV Contributors, "OpenCV Documentation. Version 4.1.1," https://docs.opencv.org/4.1.1/, 2020, [Online; accessed 22-March-2020].

[19] NumPy Contributors, "NumPpy Documentation. Version 1.17.2," https://numpy.org/devdocs/, 2020, [Online; accessed 22-March-2020].