Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

A comprehensive study on robustness of HOG and LBP towards image distortions

To cite this article: Hanwen Ren 2019 *J. Phys.: Conf. Ser.* **1325** 012012

A comprehensive study on robustness of HOG and LBP towards image distortions

Hanwen Ren

Brown University, Providence, RI, USA

renhanwen2015@163.com

Abstract. Autonomous driving has been attracting increasing attention in modern society. Lots of work have been conducted in this field focused on vehicle detection. Existing vehicle detection algorithms are developed based on the assumption that data come with good quality. However, the physical world is never free of distortions. In this paper, a comprehensive review study of vehicle detection feature extractors is conducted to test and compare their robustness towards image distortions in practical applications. The two tested mainstream feature extraction methods are Histogram of oriented gradient (HOG) and Local Binary Pattern (LBP). The specific tested distortion type includes Gaussian noise, Salt and Pepper noise, Gaussian blur, change of color saturation, change of contrast and change of brightness. The review test result shows that: 1). Gaussian noise has negative effects on both HOG and LBP; 2). A certain level of Gaussian Blur and decreasing contrast can do help to increase detection accuracy for LBP, HOG is not affected by them; 3). Color Saturation change, increasing contrast and brightness change will not affect the performance of both feature extraction methods too much when the intensity of these distortions stays in reasonable range; 4). The effect of Salt and Pepper noise is still uncovered based on our experiments.

1. Introduction

Autonomous driving is nowadays becoming an exciting field in both academia and industry. With the development in autonomous driving, drivers will have a larger chance to get themselves free from distracted, drunk or irrational human drivers. Fully developed autonomous driving will not only save people’s time and energy that consumes on road, but also reduce the possibility of encountering traffic accidents that caused by bad driving habits. Currently, some autonomous vehicles have been tested and drove on road like models from Tesla. However, there are still problems being revealed from time to time showing these cars are not fully autonomous yet. Many researches and companies in car industry are putting their hard effort into this field in order to make autonomous driving reach a higher level. One key factor in autonomous driving is vehicle detection. Failing to detect running vehicles on the road can result in serious consequences both physically and financially. Hence, many vehicle detection algorithms have been proposed.

There are many existing vehicle detection methods using different features of images. Some of them use low level features such as color, shadow, symmetry and edge [1] whereas other methods utilize more advanced methods like local feature descriptors such as HOG and haar-like features. Recently, there are even new proposed methods using convolutional neural network and get deep features from the image. Liujuan Cao et al. has reviewed that deep convolutional neural network with deep features can boost detection performance compared with handcraft designed features [2].

To provide fair comparison between existing deliberately designed vehicle detection methods and machine learning algorithms used in industry applications, several review works have been proposed. These review papers can be divided into two types: feature-based review and classification-based review. In feature-based review, Raad et al. take a look at background subtraction methods, feature based methods and motion based methods in vehicle detection [3]. They compared these methods by going through the basic idea behind them and provide readers with brief introductions. In classification-based review, a paper written by Bougharriou et al. reveals the great performance using Support Vector Machine with Histogram of oriented gradient feature. Their experiments show that support vector machine is a solid choice in vehicle detection scenario [4].

Even though many reviews have been proposed, there are only small portion of them focused on the perspective of feature robustness in physical world. Physical world is not perfect. There are lots of different types of distortion that may affect the quality of image [5] and the performance of these vehicle detection algorithms, such as noise, blurring and color modification, which makes the robustness of the detecting algorithms to be extremely important.

In this paper, we review the robustness of HOG and LBP on several existing distortions in physical word, including Gaussian noise, Salt and Pepper noise, Gaussian Blur, change of color saturation, change of color contrast and change of brightness. For these distorted images with different distortion levels that controlled by a parameter, we extract features using HOG and LBP in YCrCb color space and train a fixed linear SVM. For each distortion, we analyze the influence based on comparing the testing accuracies of original images and that of images containing different levels of noise. Based on these tests, several important conclusions will be drawn regarding to the robustness of feature extraction methods in practical application towards noises. Also, we introduce cross-dataset validation to verify these conclusions.

2. Experiment setup

2.1 Tested Feature

Histogram of oriented gradient (HOG) is a widely used feature descriptor for object detection invented by Robert K. McConnell in 1986. Previous work from Navneet et al. shows HOG feature descriptor can outperform other existing feature sets for human detection [6]. Basically, this method will divide the whole image into small cells and calculate the gradient magnitude as well as direction for each pixel inside cell. Results from pixels will be compiled into bins where each bin contains the sum of gradient magnitude within a certain range of gradient direction. In addition, results from cell calculation can be normalized by all cells within a block where a block is a combination of several cells in order to get rid of the noise like illumination and shadow. This method can help us catch edges of vehicles that distinguish them from background.

Local Binary Pattern (LBP) is another popular visual descriptor used in computer vision field. It has already shown excellent performance in human face detection [7]. It is a method known for its low computational cost and high performance in texture classification. This method calculates a local binary representation of a pixel by comparing it with the chosen surrounding neighbors. Then, all binary representations are converted into values and a histogram is created based on these values. As there are more details and edges inside images that contain vehicles, LBP can help us capture these sharp changes in vehicles.

2.2 Classifier

For classification method, we use Support Vector Machine (SVM) in this experiment. SVM is set of supervised learning methods proposed by Vladimir Vapnik [8] and his colleagues, which belongs to a family of generalized linear classification [9]. SVM has been used in many pattern recognition scenarios and has already shown great performance in classification and forecasting problems [10, 11]. During training, SVM will try to get a boundary that can separate two given labels with the largest margin. When the training dataset is not linearly separable, a kernel trick can be performed to map the current feature space into higher dimensional space and make it linearly separable. In our experiment, vehicle images will have more sharp edges compared to background images. Thus, SVM can be trained to get the hyperplane that maximize the margin and make accurate predictions.

Figure 1. vehicle image samples from GTI dataset

Figure 2. road image samples from GTI dataset

2.3 Image dataset

The first image dataset A is from GTI vehicle image database [12]. This image dataset is used for researches on the vision-based vehicle detection tasks, which has 4000 vehicle rear images as well as 4000 road sequence images. The second dataset B comes from a combination of KITTI vision benchmark suite [13] and Udacity vehicle detection project. 4000 vehicle rear images are chosen from KITTI dataset and 4000 road sequence images are extracted from Udacity project. We primarily use dataset A for training and testing as all images come from one source. Dataset B is used for validation.

Figure 3. vehicle image samples from KITTI dataset

Figure 4. road image samples from Udacity dataset

3. Target Distortions

We choose six common Distortions from the list in Image database TID2013 [14].

3.1 Gaussian noise

Gaussian noise is one kind of noise whose probability density function follows that of Gaussian distribution. In digital images, this noise can be caused by thermal vibration of atoms and radiation of warm objects [15, 16]. Methods like median filter or Gaussian smooth are commonly used to reduce Gaussian noise. The following equation shows how we modify image pixels to add artificial Gaussian noise in all channels of RGB color space.

$$p(Z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

$$D(i,j) = D(i,j) + Z \tag{1}$$

Where $D(i,j)$ means the value in position (i, j) of image and $p(Z)$ means the probability of noise value equal to Z follows Gaussian distribution with mean μ and standard deviation σ.

3.2 Salt and Pepper noise

Salt and Pepper noise is also known as impulse noise. It is commonly caused by sudden disturbances in the image signal. It can also happen during the process of image transmission [14,17]. Median filter is an efficient way to reduce this kind of noise. The following equation shows how we modify image pixels to add artificial Salt and Pepper noise in all channels of RGB color space.

$$p(D(i,j) = 0 \ or \ 255) = P$$

$$p(D(i,j) = D(i,j)) = 1 - 2 \times P \tag{2}$$

Where P means the probability of encountering an high or low impulse.

3.3 Gaussian blur

Gaussian blur is an important type of distortions that often met in practical applications and frequently included in studies dealing with visual quality metrics. Gaussian blur can be used to reduce noise and details in image. The following equation shows how we modify image pixels to add Gaussian Blur by convoluting it as well as its surrounding pixels with Gaussian kernel in all channels of RGB color space.

$$D(i,j) = \sum_{x=-M/2}^{M/2} \left(\sum_{y=-M/2}^{M/2} \frac{D(i-x, j-y)}{2\pi\sigma^2} e^{-\frac{(x-\frac{M-1}{2})^2 + (y-\frac{M-1}{2})^2}{2\sigma^2}} \right) \tag{3}$$

Where M is the size of Gaussian Kernel.

3.4 Color saturation

Color saturation determines how colorful the image looks like. During image acquisition and processing, changes in color saturation may occur. In addition, image compression can also cause changes in color saturation [13]. The following equation shows how we modify image pixels to change color saturation in Cr and Cb channels of YCrCb color space.

$$D(i,j) = 128 + (D(i,j) - 128) \times K \tag{4}$$

Where K is the control parameter of color saturation.

3.5 Change of contrast

Contrast defines the difference in luminance that makes object distinguishable. Increasing contrast will make objects easier to be captured by human visual system. The following equation shows how we modify image pixels to change contrast in all channels of RGB color space.

$$D(i,j) = D(i,j) \times C \tag{5}$$

Where C is the multiplication factor.

3.6 Change of Brightness

Brightness of the image means the perception revoked by the luminance of a visual target. The following equation shows how to modify image pixels to change brightness in all channels of RGB color space.

$$D(i,j) = D(i,j) + B \quad (6)$$

Where B is a constant add to each pixel.

4. Experiment result

The baseline of testing accuracy using HOG on dataset A is 97.31% and that on Dataset B is 100%. For LBP, the baseline of testing accuracy in dataset A 88.25% is and that on Dataset B is 100%. The baseline accuracy means the accuracy we get using the original images. As dataset B is a dataset where vehicle images and road sequence images come from different sources, the high testing accuracy on both features can be caused by this fact. Thus, dataset B is only used for validation and we focused more on dataset A in discussion. The A in the legend of plots means the experiment is performed on dataset A(asterisk marker) while B means dataset B(circle marker).

4.1 Gaussian Noise

Our setting in adding Gaussian Noise is to fix μ to be 0 and change standard deviation σ. The value set of σ is (5, 10, 15, 20, 25, 30, 35, 40, 45, 50).

Figure 5. different levels of Gaussian Noise

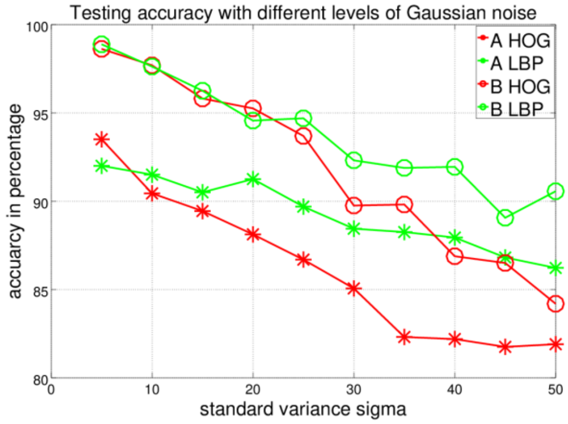

Figure 6. Gaussian Noise testing result

For Gaussian noise, testing accuracies of dataset A using both features drop with the increasing variance σ of the gaussian model. Accuracy with HOG feature drops more than that with LBP feature. After adding the Gaussian noise, the edge information of vehicles are affected and when the variance is high enough, the background image can also have sharp gradient change like a fake edge. This will make SVM harder to get the hyperplane to separate the training set. Thus, HOG and LBP feature extraction will both be affected by the Gaussian noise. Validation set B also show same effects with increasing variance σ. Thus, both HOG and LBP feature are not robust to this noise.

4.2 Salt and Pepper Noise
Our setting in adding Salt and Pepper Noise is to increase the probability of encountering a high or low impulse P. The value set of P is (0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1).

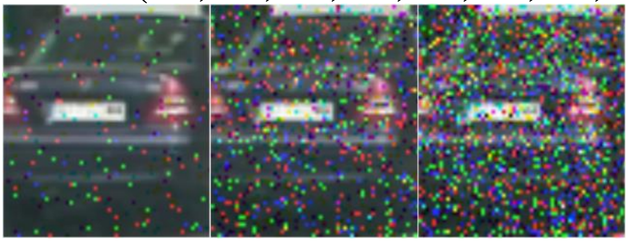

Figure 7. different levels of Salt and Pepper Noise

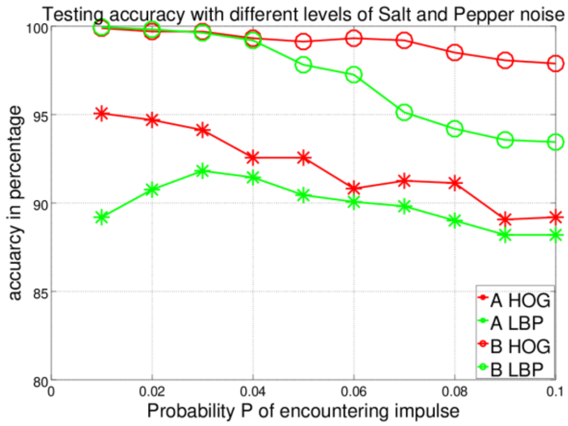

Figure 8. Salt & Pepper Noise testing result

For Salt and Pepper Noise, we expect to see the accuracies of both testing results drop with the increasing probability of encountering high or low impulses. The reason is that when high or low impulses appear in the image, all gradient related to them will be relatively larger in magnitude and the direction will point towards these noises. For HOG, both vehicle images and non-vehicle images will have big gradient magnitudes added to the histogram vectors. As a result, vehicle images and no-vehicle images will have more characteristics in common. However, from the results we observe testing accuracy of dataset A with HOG drops with the increasing probability of encountering impulses while that of LBP feature remains the same. To our surprise, validation set B shows the opposite pattern, accuracy with HOG are stable and that with LBP drops. As two datasets give us opposite results on how accuracies change with increasing probability of impulse, we cannot draw any conclusion for now. More studies are needed to check the effect of Salt and Pepper noise.

4.3 Gaussian Blur
Our setting in adding Gaussian Blur is to increase the size of Gaussian kernel M. The value set of M is (5, 7, 9, 11, 13, 15, 17, 19, 21, 23).

Figure 9. different levels of Gaussian Blur

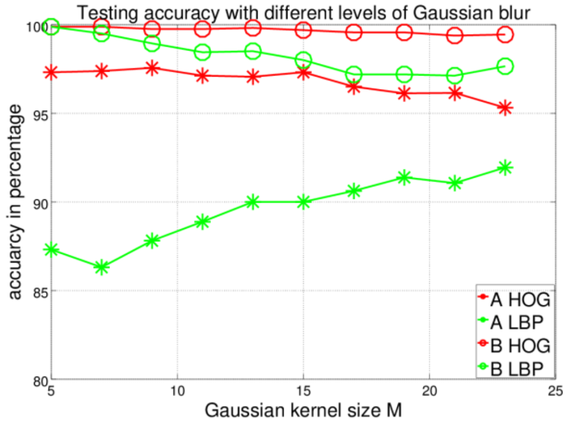

Figure 10. Gaussian Blur testing result

For Gaussian Blur, testing accuracy of dataset A with HOG remains stable and that with LBP increases a little with Gaussian kernel size. Validation dataset B shows stable accuracies for both HOG and LBP features. Gaussian Blur serves to smooth the image which will not change the relationships of value comparison relationship between certain pixel and its surrounding pixels too much. Also, Gaussian blur may remove some small details in road sequences which may be previously be captured as vehicle edges, which will help to increase the learning accuracy. For HOG, the feature tracks large gradient as well as the place where they occur. Thus, detailed structure in road sequence hard to affect the performance as they are not exactly like vehicle edge by being sparse or discontinuous. However, for LBP, patterns are merged into bins and the location information is lost as the feature vector is combined sums, which makes it highly depends on the gradient. The difference in designs of two feature extractors may explain why Gaussian blur can help to increase the accuracy with LBP but not HOG. Overall, both HOG and LBP feature are robust to this noise.

4.4 Color Saturation Change

Our setting in increasing color saturation is to increase the factor K used in the equation in Cr and Cb channel of YCrCb color space. The value set of K is (1.09, 1.18, 1.27, 1.36, 1.45, 1.54, 1.63, 1.72, 1.81, 1.9). For decreasing color saturation, the value set of K is (0.91, 0.82, 0.73, 0.64, 0.55, 0.46, 0.37, 0.28, 0.19, 0.1).

Figure 11. different levels of increasing Color Saturation

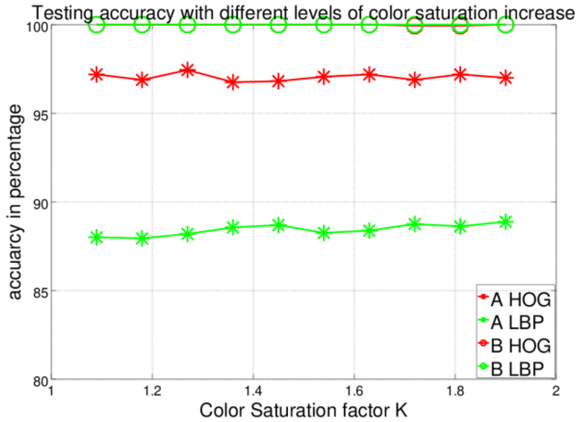

Figure 12. Increasing Color Saturation testing result

Figure 13. different levels of decreasing Color Saturation

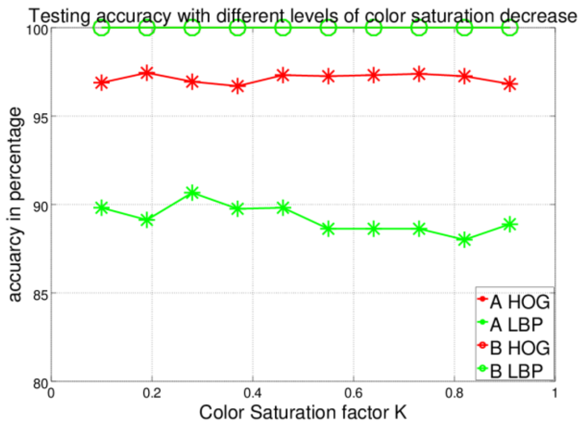

Figure 14. Decreasing Color Saturation testing result

For color saturation change, both HOG and LBP seems to be stable. When color saturation changes, the first channel, which is a grayscale of the image stays the same. As we extract features from YCrCb color space, as long as the grayscale image stays the same, we do not lose any edge information that can help to distinguish between vehicles and road sequences.

4.5 Contrast Change

Our setting in increasing contrast is to increase the factor C which all RGB color space pixel will be multiplied with. The value set of C is (1.09, 1.18, 1.27, 1.36, 1.45, 1.54, 1.63, 1.72, 1.81, 1.9). For decreasing contrast, the value set of C is (0.91, 0.82, 0.73, 0.64, 0.55, 0.46, 0.37, 0.28, 0.19, 0.1). The parameter choices here make sure when using the biggest contrast change parameter, the vehicle is still noticeable by human eyes instead of a complete white image, which will happen when contrast factor C is extremely high.

Figure 15. different levels of increasing Contrast

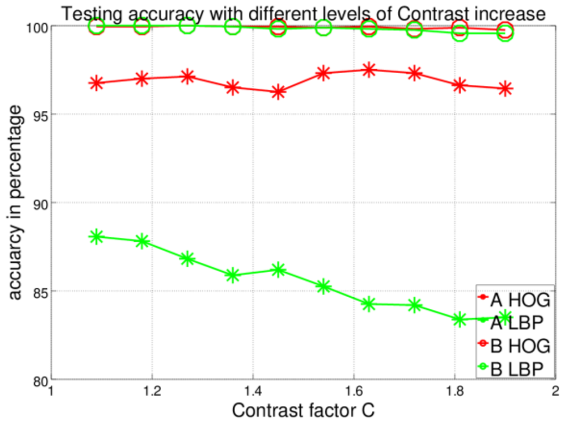

Figure 16. Increasing Contrast testing result

Figure 17. different levels of decreasing Contrast

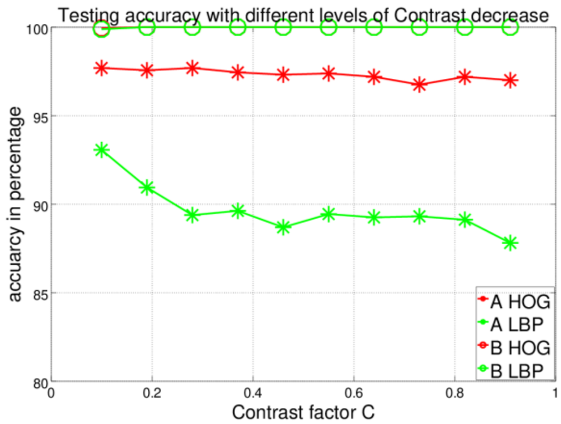

Figure 18. Decreasing Contrast testing result

For Contrast change, the testing accuracy of both features are stable when the contrast factor C do not change too much. However, when contrast increases, LBP accuracy goes down with increasing C. When C increases, some pixel value will become 255 and saturate. Thus, more and more pixels will saturate with increasing C factor. In this way, the binary patterns in LBP feature can be changed. Also, the length of feature vector in LBP is much shorter compared to that in HOG. With fewer dimensions, it will be harder for SVM to get a hyperplane to separate two groups. On the other hand, when we decrease contrast, the difference between pixels will be smaller and the relationship between pixels can also change as the values of channel are always integers. With the change of local binary pattern and gradient, when C is low enough, only strong edges will survive. Thus, when we decrease contrast, it helps us to remove some weak edges in road sequences and focus more on the strong edges of vehicles. It may explain why we see accuracy increases a little bit when C is 0.1 for LBP. What we can take away from this experiment is that maybe we can decrease contrast as a preprocess stage in further vehicle detection procedure because it helps to remove small noisy edge and increase accuracy. Overall, both HOG and LBP are robust to contrast change.

4.6 Brightness change

Our setting in increasing brightness is to increase the value B which is added to each RGB color space pixel. The value set of B is (10, 20, 30, 40, 50, 60, 70, 80, 90, 100). For decreasing brightness, the value set of C is (-10, -20, -30, -40, -50, -60, -70, -80, -90, -100). The parameter choices here make sure when using the biggest brightness increase parameter, the vehicle is still noticeable by human eyes instead of a complete white image, which will happen when increase parameter B is extremely high. Also, for the smallest brightness decrease parameter, we make sure that the vehicle is noticeable by human eyes instead of a complete black image, which will happen when decrease parameter B is extremely low.

Figure 19. different levels of increasing Brightness

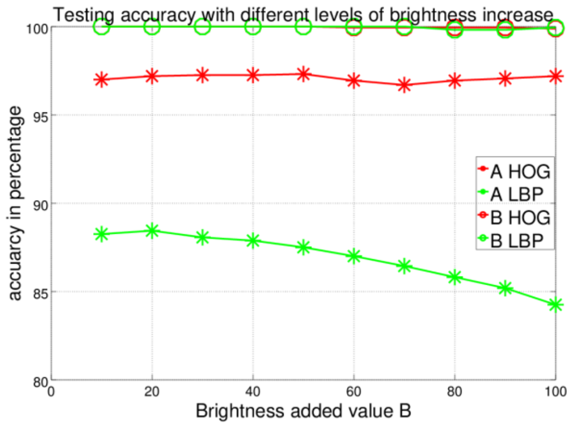

Figure 20. Increasing Brightness testing result

Figure 21. different levels of decreasing Brightness

 1325 (2019) 012012 doi:10.1088/1742-6596/1325/1/012012

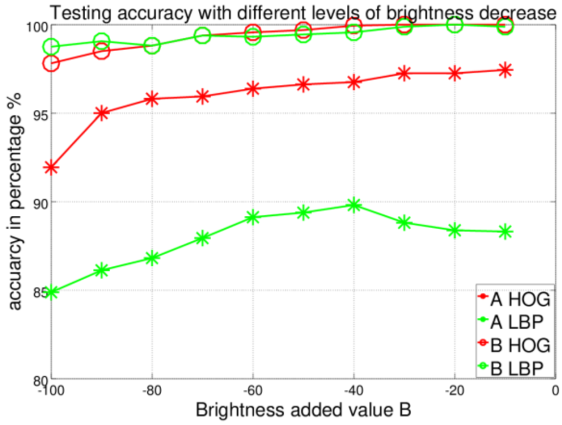

Figure 22. Decreasing Brightness testing result

For both increasing and decreasing of brightness, testing accuracies of two datasets with HOG and LBP are relatively stable. When we add or subtract the values to modify brightness, the values can saturate to 0 or 255 and change the local gradient as well as local binary pattern related to them. This may explain the slight changes in the testing accuracies between different levels of brightness change. Also, with more and more pixels saturate, we lose information from images and test accuracies are expected to drop as a result.

5. Conclusion

In this paper, we tested the robustness of HOG and LBP towards six image distortions (Gaussian noise, Salt and Pepper noise, Gaussian blur, color saturation change, contrast change and brightness change) on vehicle detection scenario. We fixed our learning method to be SVM and tested on two datasets. The conclusion we can make here is that both HOG and LBP features are very robust to Gaussian blur, color saturation change and decreasing of contrast. Surprisingly, we even observed Gaussian blur and decreasing of contrast can help us improve learning performance for LBP. For increasing contrast and brightness change, if the parameters of these changes are within reasonable range, these two features are relatively stable. In addition, both HOG and LBP features will be affected by Gaussian Noise. However, Salt and Pepper Noise can have ambiguous effects on these features because results on different datasets show opposite patterns. Further searches can be done to better understand this issue. Finally, HOG feature seems to be overall more stable than LBP feature, but it also costs more computational power. Thus, we can choose HOG for performance and LBP for speed in real practice.

References

[1] Yang, Z. & Pun-Cheng, L. S. C. “Vehicle detection in intelligent transportation systems and its applications under varying environments: A review”, Image and Vision Computing, Vol. 69, pp. 143–154 ,2018.

[2] Cao, L., Jiang, Q., Cheng, M. & Wang, C. “Robust vehicle detection by combining deep features with exemplar classification”, Neurocomputing, Vol. 215, pp. 225–231, 2016.

[3] Hadi, R. A., Sulong, G. & George, L. E. “Vehicle Detection and Tracking Techniques : A Concise Review”, Signal & Image Processing : An International Journal, Vol. 5, pp. 1–12, 2014.

[4] Bougharriou, S., Hamdaoui, F. & Mtibaa, “A. Linear SVM classifier based HOG car detection”, 2017 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), 2017.

[5] Karen Panetta, Long Bao, Sos Agaian, “A Human Visual “No-Reference” Image Quality Measure,” IEEE Instrumentation & Measurement Magazine, vol. 19, no. 3, pp. 34-38, June 2016.

[6] Davis, M. & Sahin, F. "HOG feature human detection system", 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016.
[7] Satpathy, A., Jiang, X. & Eng, H.-L. "LBP-based edge-texture features for object recognition", IEEE Trans. Image Process. 23, 1953–1964, 2014.
[8] Nayak, J., Naik, B. & Behera, H. S. "A Comprehensive Survey on Support Vector Machine in Data Mining Tasks: Applications & Challenges", International Journal of Database Theory and Application, Vol. 7, pp. 169–186, 2015.
[9] Qi, X., Silvestrov, S. & Nazir, T. "Data classification with support vector machine and generalized support vector machine", 2017.
[10] Wang, Z. & Xue, X. "Multi-Class Support Vector Machine", Support Vector Machines Applications, pp. 23-48, 2014.
[11] Pu, J.-Y., Jun-yi, P. U. & Xiu-ren, L. E. I. "Two-stage fast training method based on core vector machine and support vector machine", Journal of Computer Applications, Vol. 32, pp. 419–424, 2013.
[12] J. Arróspide, L. Salgado, M. Nieto, "Video analysis based vehicle detection and tracking using an MCMC sampling framework", EURASIP Journal on Advances in Signal Processing, vol. 2012, Article ID 2012:2, Jan. 2012.
[13] Geiger2013IJRR, Andreas Geiger and Philip Lenz and Christoph Stiller and Raquel Urtasun, Vision meets Robotics: The KITTI Dataset, International Journal of Robotics Research (IJRR), 2013.
[14] Ponomarenko, N. et al. "Image database TID2013: Peculiarities, results and perspectives", Signal Processing: Image Communication, Vol. 30, pp. 57–77, 2015.
[15] Boyat, A. K. & Joshi, B. K. "A Review Paper : Noise Models in Digital Image Processing", Signal & Image Processing : An International Journal, Vol. 6, pp. 63–75, 2015.
[16] Karen Panetta, Long Bao, Sos Agaian, "Sequence-to-Sequence Similarity-Based Filter for Image Denoising," IEEE Sensors journal, vol. 16, no. 11, 2016.
[17] Karen Panetta, Long Bao, Sos Agaian, "A New Unified Impulse Noise Removal Algorithm Usig a new Referenced Sequence-to-Sequence Similarity Detector," IEEE Access, vol. 6, pp. 37225-37236, 2018.