Vehicle Detection using Upper Local Ternary Features with SVM Classification

Linn Linn Thike
University of Computer Studies, Yangon
lynnlynnthike@ucsy.edu.mm

Thin Lai Lai Thein
University of Computer Studies, Yangon
tllthein@ucsy.edu.mm

***Abstract*— Local Ternary Pattern (LTP) is an extension of Local Binary Pattern (LBP), which is known as a standard textural descriptor for various recognition systems. There are a lot of usages of LBP and LTP in the variety of pattern recognition in different fields. It is also efficient features for vehicle detection, recognition and monitoring. This paper proposed an algorithm that applied the operator to upper LTP (upper-LTP) matrix as feature vector to calculate instead of pattern stream generation. This paper analyzes the accuracy rate between the Complemented- Uniform Local Binary Pattern (Complemented-ULBP) which is our previous work, Uniform Local Binary Pattern (ULBP), simple Local Binary Pattern (LBP) and the proposed system, Upper-LTP, with SVM classification.**

Keywords—LBP, ULBP, Complemented-ULBP, LTP, Upper-LTP, SVM

I. Introduction

Parking Monitoring systems are currently one of the most important and essential system for today's developing world. Therefore, efficient vehicle detection methods are needed to get better Monitoring system. There are a lot of detection methods that are using different types of methods: counter, sensor, and image-based [1]. Counter-based counted the number of vehicles a car parking area by sensors apparatus at exit and entrance of the parking. But it can face a problem to install the apparatuses that are used to count based on the land condition because those devices are installing at the entrance of the parking. In sensor-based systems, there exist different types of sensors using in the parking monitoring system. Some common and developing sensors that are used today are: Infrared (IR) sensor, Ultrasonic Sensor, Microwave/Millimeter wave radar, Passive Acoustic Sensor, Photoelectric, Acceleration Sensor and Magnetic sensor, etc.,

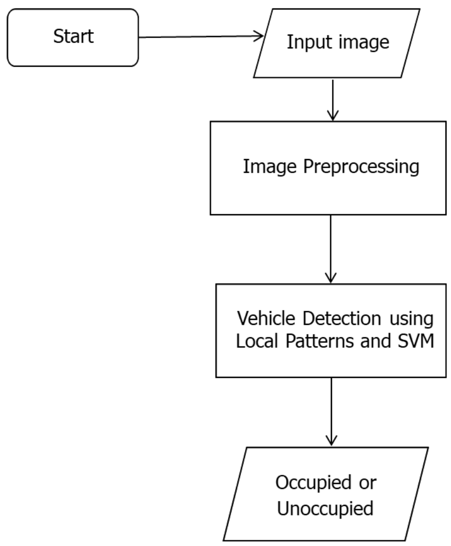
Although sensors can easily and ready to use, the cost is expensive and takes time to apply based on the parking space position and condition of the parking land. Therefore, image processing techniques are the most cost effective ways to use in the v by the use of existing surveillance cameras or webcams without extra expenses.

Baroffio and collages [2] developed the camera-based vehicle detection system by using hue histogram as features with SVM classification. They also compared two Visual sensor transmission: Compress-than-analysis (CTA) and Analysis-than-Compress techniques to transmit the image data or histogram features to the main server and to determine whether the space is occupied or not. PKlots dataset was used to classify which is offered online [4].Therefore, the transmitted videos or images are needed to compress and also have the enough quality to extract the essential information for the detection process. Also the blurring could occur because not only movements of the camera but also each sub space are extracted from the images of the whole parking lot. Therefore, it is necessary to develop that could handle the above problem better.

Rendondi and Baroffio compared two transmission methods compress-then-analysis (CTA) and analysis-then-compress (ATC) for image analysis in VSN [1] and evaluate the performance of object recognition. The authors considered to extract BRISK visual features that is the best optimization in faster compute and useful in both lower power and complexity hardware. In the compress-then analyze (CTA) case, firstly, images are compressed with JPEG at different rates with the JPEG quality factor QF from 1 to 100. Although the ATC perform analysis at low bitrates, the CAT also the best in results when enabling the bandwidth of high-quality images.

Foreground extraction is used by Sheng-Fuu Lin, Yung Yao Chen, and Sung-Chieh Liu [3] tried to develop vision based parking lot management system. Before foreground extraction, the color consistency is managed to adjust to the input image sequences from camera. Then foreground is extracted with the value of color intensity and shadow detection and removal is performed on the system. They divided the gray level values of 24-bits color images under three different weather; (1) from 100 to 130 gray level value is defined as common weather, (2) from 70 to 120 gray level value is a cloudy day, and (3) between 130 to 170 gray level is a sunny day.

Although it could reduce the accuracy the gray level could change based on the vehicle color.



Figure(1). Vehicle Detection System Overview

Jian Li and Hanyi Du [8] proposed Extended Gradient Local Ternary Pattern (EGLTP) which contained properties of the original LTP, Semantic Local Binary Patterns (S LBP) and HOG. EGLTP is implemented with SVM classifier to train and to detect the different types and sizes of vehicles. EGLTP is compared with HOG, uniform LBP, GLBP, and HOG-uniform LBP features by using the horizontal axis (FPR) and the vertical axis (FNR) which are computed based on true/false positive/negative samples. It is needed the spatial linear because the smaller number FPR, the lower detection rate for EGLTP.

In 2016, Pedro G.F [7] proposed a NR-IQA method which uses LTP descriptors with machine learning. It provides a better performance improvement than LBP. Moreover, LTP is more robust to noise in uniform areas, but not to gray-level transformation because of noise insensitive. LTP descriptors are not able to detect milder image degradation. Therefore, a multiple LTP is introduced for texture extraction. The histograms of the texture are training as features. Their method predicted image quality without references, i.e., the method is no-reference (NR).

In our previous work [5], background is subtracted by calculating the MSE value between the background image and input images. Before MSE value is calculated, the system applied Complemented-ULBP which finds the negative value of the image before Uniform Local Binary Pattern is calculated. The system is compared with the Uniform Local Binary Pattern (ULBP) and it got the slightly better result by 10% accuracy rate and it reduced the MSE value by 1:2 ratios.

In this paper, a detection system that is using only upper binary matrix or lower binary matrix of LTP instead of using both matrixes which are the resulted matrix of LTP is proposed. Also it is compared the proposed Upper-LTP with Complemented-ULBP, ULBP and Simple LBP patterns features for performance measurements by using SVM classification.

The paper organized with the following sessions. Firstly, the session is described the system overview and techniques. Next, it presents our proposed method, Upper-LTP and Complemented-ULBP with SVM classification. The section 4 addresses experimental results. The last section concludes the paper.

II. Vehicle Detection System

There are two steps for vehicle detection:

- *Feature Extraction:* Textural Features or Pattern Features like Local Binary Patterns, Local Ternary are popular to use in Vehicle Detection systems. Also they are widely used in facial features and other areas.
- *Classification:* SVM is widely used to classify the two classes like 0 and 1. Other classification and machine learning system are also used for wider areas with multiple dimensions data.

A. Uniform Local Binary Pattern (ULBP)

Ojala et al. [6] introduced Local Binary Patterns (LBP) in 2009 . It divided the input images pixels equally into sub dimension, eg.9x9, each sub-dimension defined with a rotation R, center pixels C and neighbors P. Then a histogram h is calculated by dividing the intensity value of C with neighbors P, that is the calculation of the gray scale invariant. If the result is positive, its value is defined as 1 and if it is not then 0.

The example variant LBP calculation and the primitives ' value are described in Figure 2 and Figure 3. The LBP is defined in equation (1).

$$LBP_{R,P}(t_c) = \sum_{p=0}^{p-1} s\left(t_p - t_c\right)2^p, \qquad (1)$$

Where t_p is the neighbor pixel, t_c is the grayscale of the center pixel *C*, the radius of the neighborhood, *R*, the total number of neighbors, *P*, and the LBP calculation is shown in equation (2)

$$S(t) = \begin{cases} 1 & t \geq 0 \\ 0 & otherwise \end{cases}, \qquad (2)$$

Uniform-LBP (ULBP) is a LBP which it is measured uniformly pattern that is mostly at 2, i.e. transitions is defined in the range 0 and 1 ≤ 2.

The local binary patterns are mostly uniform in natural images. The equation (3) calculated uniform pattern values and 58 uniformed patterns are calculated [7].

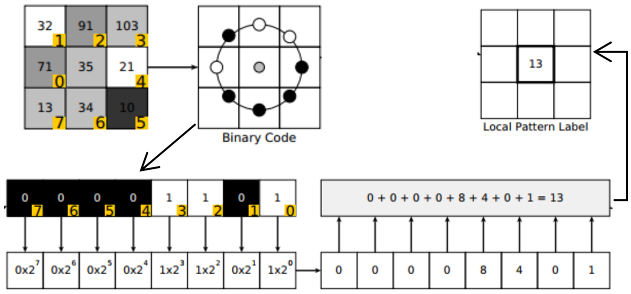


Figure 2. Example calculation of the gray scale variant LBP

$$U(LBP_{P,R}(t_c)) = \left| \begin{matrix} s(g_{P-1} - g_c) - \\ s(g_0 - g_c) \end{matrix} \right| + \sum_{P=1}^{P-1} \left| \begin{matrix} s(g_P - g_c) - \\ s(g_{p-1} - g_c) \end{matrix} \right| \quad (3)$$

where, $g_P$is the gradient value of the neighbor, and g_c is the gradient value of the center pixel.

B. *Local Ternary Pattern (LTP)*

The Local Ternary Pattern (LTP) is an extension of the LBP [10] with patch wise feature extraction. The LBP operator is formally defined as:

$$\hat{S}(t) = \begin{cases} 1, & t \geq \gamma \\ 0, & -\tau < t < \gamma \\ -1, & t < -\gamma, \end{cases} \quad (4)$$

where γ is a threshold which determines how intensity changes can be considered as an edge. After using the above equation, the ternary codes are computed. Then two codes are split for each of the patterns: a positive (upper pattern) and a negative (lower pattern). Those codes are treated as two separate channels of LBP.

The basic feature extraction of a pixel using LTP is described in figure (3) with 3x3 neighbor, the radius (R=1), neighbors (P=8) and the threshold τ is set to five. By using equation (4), three values are generate (1,0,-1).

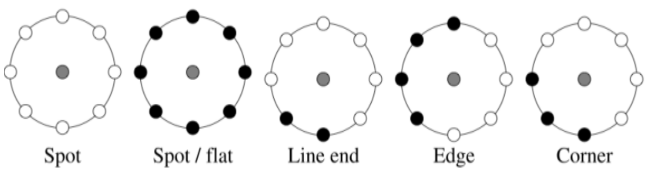


Figure 3. The texture pattern detected by the LBP

Then LTP codes are divided by two LBP patterns using just values with positives. The upper pattern is created by converting the negative replace with zero and the replacement of zeros into positive values, the lower pattern is defined. The calculation steps are shown in figure (4).

When figure (4) and (5) are comparing that the LTP is the extended pattern of the LBP and, two pattern maps is resulted. It is defined the maps as two separate channels of LBP descriptors. Then results are combined at to generate texture features. Next session proposed an algorithm that used just the histogram of the upper patterns of LTP as feature vector for classification compared with our previous work.

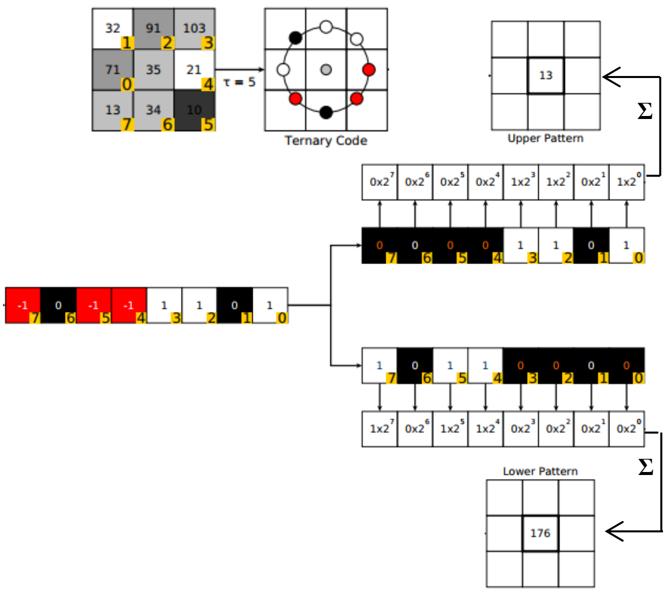


Figure 4. Example calculation of LTP

C. *Support Vector Machine (SVM)*

Support Vector Machine (SVM) is a popular and desirable classification algorithm especially for high dimensional data. Most researchers considered that it give the better accuracy than other classification algorithm. It is a learning system that is the classifiers for two classes. For example, if the set of training examples *D* is

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}, \quad (2)$$

where $x_i = (x_{i1}, x_{i2}, \dots, x_{ir})$ is a *r*-dimensional input vector, y_i is the classes (output value) $y_i \in \{1, -1\}$. SVM finds a linear function of the form is shown in equation (3).

$$f(x) = < w.x > +b \quad (3)$$

So that an input vector x_i is assigned to the positive class if $f(x_i) \geq 0$ and to the negative class otherwise, ie., it is defined by equation (4). In our system, if it is positive, it will be defined as occupied and otherwise it is considered as empty.

$$y = \begin{cases} 1 \; if < w.x_i > +b \geq 0 \\ -1 \; if < w.x_i > +b < 0 \end{cases} \quad (4)$$

D. Precision and Recall

The performance of the system is calculated by not only precision and recall but also by processing time cost. The precision and recall is calculated by the equation (5), (6) and the processing time is calculated including the SVM training time.

$$P = \frac{T_P}{T_p + F_P} \quad (5)$$

$$R = \frac{T_p}{T_p + F_N} \quad (6)$$

where, precision rate, P, recall rate, R, true positive, $\mathrm{T_P}$, False positive, $\mathrm{F_P}$, and False negative, $\mathrm{F_N}$. Processing time is evaluated on Time per patch and also including the SVM training time.

III. The Proposed System

An algorithm that used the upper LTP pattern as features which is resulted from the LTP operator is proposed in this paper. It is shown in table (1).

In this algorithm, the vehicles that are occupied and not occupied images are preprocessed by converting RGB to grayscale. Then the image adjustment is applied to have better intensity of the image and then the LTP operator is applied and two vector matrixes, upper and lower patterns, are resulted. Instead of using both vectors, we used just only upper LTP pattern as feature to classify with the SVM classification by using the equation (4) and calculation of the LTP are shown in figure(4).

Figure 5. Parking database for ShweDagon Pagoda

In this paper, we compared the performance of the proposed system, previous works Complemented-ULBP and the Simple LBP, ULBP as features vector by classifying with SVM classification. The experimental results are described in the next session.

A. Database

For training SVM we used our own datasets which is taken at the ShweDagon Pogada's south car parking, shown n figure 5, in Feb, 2019 and also PKLot dataset [9] The ShweDagon video database is taken at Shwedagon Pagoda's north parking in 14.11.2019 when the weather is Sunny during 8:00 AM to 12:00 AM and the number of total pictures is shown in table (5). The video file's resolution is 1080x1920 pixels with the frame rate 30.02 frames per second. Also it is used PKlots dataset to train and test SVM classifier in the system.

TABLE I. THE UPPER LOCAL TENARY PATTERN ALGORITHM

1. Preprocessed the images by converting the color RGB into gray scale.
2. Then intensity values are adjusted.
3. The LTP operator is applied to all images.
4. Only the histogram of upper LTP patterns are extracted as features patterns.
5. The SVM is trained with both the vehicle and spaces images without vehicles' upper LTP patterns.
6. The histogram of input image's upper LTP patterns is used to classify that the space has vehicle or not.
7. The result will be display.

In the PKLot, the images are captured in different weather and light condition; sunny, cloudy, and overcast. The whole parking images for each parking lot and segmented sub-images for the spaces are offer. They were captured at the parking lots of the Federal University of Parana (UFPR) and the Pontical Catholic University of Parana (PUCPR), located in Curitiba, Brazil.

They captured the images 5-minute time-lapse interval, more than 30 days under overcast, sunny, and rainy periods with JPEG color format (quality 100%) resolution (1280x 720) pixels. Those are stored in three different folder names: UFPR04, UFPR05 and PUCPR. The First two captured form 4th and 5th floors of the UFPR building and third is captured 10th floor of the PUCPR building. In each of above folder, there contained three sub-folders based on the weather; overcast, sunny, and rainy. All space's information which is occupied or not are classified and they are given with .xml file. The total parking space and their images numbers are shown in Table 2 and parking spaces images are shown in figure 6.

Figure 6: PKlot dataset example image. The parking spaces that considered as not fully visible have not been extracted.

IV. EXPERIENMENTAL RESULTS

The SVM classifier is already training with PKlot dataset and ShweDagon Dataset by using the proposed Upper-LTP feature, Complemented-ULBP, ULBP and Simple LBP feature extraction stage of the proposed system. Then they are trained with SVM classifier respectively. The input images are also implemented with the upper features extractions and then prediction with SVM classifier in order to know the vehicles are existed or not in the spaces.

TABLE II. PKLOT DATASET

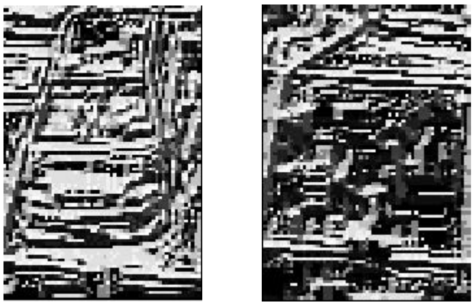
Parking lot	Weather condition	# of days	# of images	# of parking spaces occupied	empty	total
UFPR04	Sunny	20	2,098	32,166 (54.98%)	26,334 (45.02%)	58,400
(28 parking spaces)	Overcast	15	1,408	11,608 (29.47%)	27,779 (70.53%)	39,387
	Rainy	14	285	2,351 (29.54%)	5,607 (70.46%)	7,958
	Subtotal		3,791	46,125 (43.58%)	59,720 (56.42%)	105,845
UFPR05	Sunny	25	2,500	57,584 (57.65%)	42,306 (42.35%)	99,890
(45 parking spaces)	Overcast	19	1,426	33,764 (59,27%)	23,202 (40.73%)	56.966
	Rainy	8	226	6,078 (68.07%)	2,851 (31.93%)	8,929
	Subtotal		4,152	97,426 (58.77%)	68.359 (41.23%)	165,785
PUCPR	Sunny	24	2,315	96,762 (46.42%)	111,672 (53.58%)	208,433
(100 parking spaces)	Overcast	11	1,328	42,363 (31.90%)	90,417 (68.10%)	132,780
	Rainy	8	831	55,104 (66,35%)	27,951 (33.65%)	83,056
	Subtotal		4,474	194,229 (45.78%)	230.040 (51.46%)	424,269
TOTAL			12,417	337,780 (48.54%)	358.119 (51.46%)	695.899

(a)

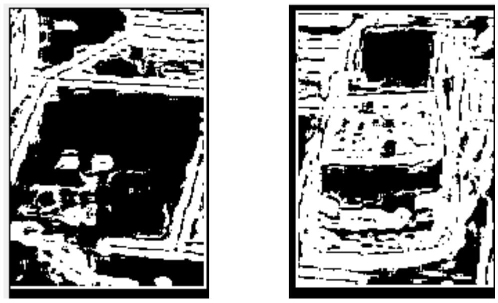
(b)

Figure 7. (a) RGB to grayscale conversion (b) Preprocessing for Complemented-ULBP

The proposed Upper-LTP and three types of LBP: simple LBP, ULBP and Complemented-ULBP are used to extract features, shown in figure 8, SVM classifier training and prediction is also performed on those three features. The training process is performed on the PKlots dataset's 8800 images and the prediction is processed with 70 images. The processing time cost and accuracy rate is defined by precision and recall values for those three features are shown in table 3 and table 4 respectively. The preprocessing stages of the images are shown in figure (7).



(a)



(b)

Figure 8 (a) Complemented-ULBP pattern images
(b) Upper-LTP pattern images

TABLE III. PROCESSING TIME OF THE EXPERIMENT

SVM	LBP	ULBP	Complemented-ULBP	Upper-LTP
Training	399.14 secs	388.48 secs	341.73 secs	529.68 secs
Prediction	48.28 secs	36.35 secs	30.13 secs	26.325 secs

Although it is better for prediction processing time, the accuracy rate is less than the ULBP and Complemented-ULBP in false negative rate, because of the color intensity value of the vehicles. By the experimental results, the Complemented-ULBP feature is the best of four in saving processing time rather than simple LBP, ULBP and proposed Upper-LTP. Because it training time is less than the other three features. The accuracy rate is 69% for ULBP, 79% for Complemented-ULBP and 81% for proposed Upper-LTP.

TABLE IV. ACCURACY RATE OF EXPERIMENT

Accuracy	LBP	ULBP	Complemented-ULBP	Upper-LTP
False Positive	11	5	10	4
False Negative	13	11	-	17

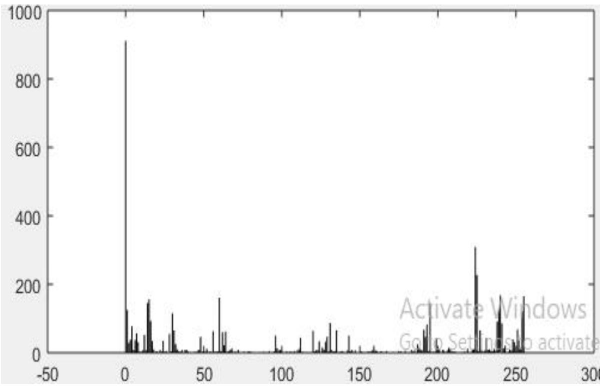
But the proposed Upper-LTP saves better time in the prediction. It is occurred when the vehicle color is black or mixing with the background. But it can reduce the entire false positive rate than the other three features. The comparison of the accuracy rate is shown in table 4.

TABLE V. SHWE DAGON DATASET

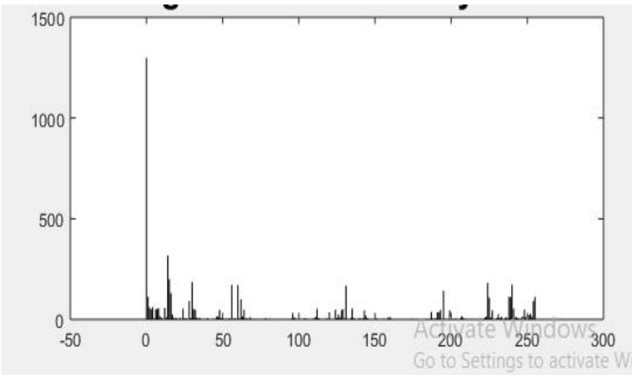
Parking Lot	Weather Condition	No. of images	No of Spaces		
			Empty	Occup-ied	Total
South	Sunny	30	159 (22%)	561 (78%)	720

V. CONCLUSION

As described in experimental results, the proposed system takes the processing time than Complemented-ULBP which is our previous proposed feature. It is better results in accuracy rate than the rest except the false negative is greater because of the color intensity value of the vehicle. Therefore, it is still needed to reduce the false negative rate by reducing color intensity in future.



(a)



(b)

Figure 7 (a) the histogram feature of vehicle,
(b) the histogram feature of no-vehicle space

REFERENCES

[1] P.Almeida, Luiz S. Oliveira, E.Silva Jr., A.Britto Jr.,A.Koerich, “Parking Space Detection using Textural Descriptors”, 2013 IEEE International Conference on Systems, Man, and Cybernetics.

[2] L.Baroffio, L.Bondi, M.Cesana, A.Redondi, M.Tagliasacchi, “A Visual Sensor Network for Parking Lot Occupancy Detection in Smart Cities”, International Journal of Computer Trends and Technology (IJCTT) 2015.

[3] Sheng-Fuu Lin, Yung-Yao Chen, and Sung-Chieh Liu, 2006 IEEE International Conference on Systems, Man, and Cybernetics October 8-11, 2006, Taipei,Taiwan.

[4] P.Ammeida, S.Oliveira, S.Britto Jr, J.Silva Jr, Koerich,“PKLot-A Robust Dataset for Parking Lot Classification”,UFPR, PUCPR (Curitiba), UEPG,P Ponta Grossa, PR, Brazil.

[5] Linn Linn Thike, Thin Lai Lai Thein, “Parking Space Detection using Complemented-ULBP Background Subtraction”, 2019 IEEE Global Conference on Consumer Electronics (IEEE GCCE 2019), Osaka, Japan.

[6] T.Ojala, M.Piektikainen, and I.Maenpa, “Multiresolution gary-scale and rotation invariant texture classification with local binary patterns”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no.7, 2002.

[7] M. Opitz, M. Diem, S. Fiel, F. Kleber, and R. Sablatnig, “Endto-end text recognition using local ternary patterns, mser and deep convolutional nets,” in *Document Analysis Systems (DAS), 2014 11th IAPR International Workshop on*. IEEE, 2014, pp. 186–190.

[8] J. Li, H. Du, Y. Liu, K. Zhang and H. Zhou, "Extended Gradient Local Ternary Pattern for Vehicle Detection," 2014 IEEE 17th International Conference on Computational Science and Engineering, Chengdu, 2014, pp. 1882-1885, doi: 10.1109/CSE.2014.345.

[9] Almeida, P., Oliveira, L. S., Silva Jr, E., Britto Jr, A., Koerich, A., PKLot – A robust dataset for parking lot classification, Expert Systems with Applications, 42(11):4937-4949, 2015.

[10] Garcia Freitas, P.; Da Eira, L.P.; Santos, S.S.; Farias, M.C.Q. On the Application LBP Texture Descriptors and Its Variants for No-Reference Image Quality Assessment. J. Imaging 2018, 4, 114.

[11] Suruliandi, A. & Meena, Kartik & Rose, R.. (2012). Local binary pattern and its derivatives for face recognition. Computer Vision, IET. 6. 480-488. 10.1049/iet-cvi.2011.0228.