

# Building Vietnam's License Plate Recognition System Based on OpenALPR

Hai Tran<sup>1</sup>, Giang Ma<sup>1</sup>, Thanh Nguyen<sup>2</sup>, Thanh Cao<sup>3\*</sup>

<sup>1</sup>Department of Information Technology, Ho Chi Minh University of Education (HCMUE), HCM, Vietnam-700000

<sup>2</sup>Faculty of Applied Sciences, Ho Chi Minh University of Technology and Education (HCMUTE), HCM, Vietnam-700000

<sup>3</sup>Faculty of Information Technology, Sai Gon University (SGU), HCM, Vietnam-700000

Email address: {haits,giangmn}@hcmue.edu.vn, thanhnlv@hcmute.edu.vn

**Abstract**— In this research, we proposed a solution for Vietnam's license plate recognition include many stages based on OpenALPR with OCR incremental training. This paper describes a practical approach for constructing an OpenALPR-based license plate recognition system. This paper outlines the procedures that must be taken in order to construct the system, including data gathering, detector training, OCR training, and system deployment. The test findings demonstrate that the system can recognize license plates with a high degree of accuracy and can be installed on a variety of platforms, thus satisfying users' practical requirements.

**Keywords**— Vietnam's license plate recognition, OpenALPR.

## I. INTRODUCTION

In Vietnam in general and Ho Chi Minh City in particular, the use of surveillance cameras is growing. Surveillance camera systems in Vietnam [1], however, only have recording capabilities and let users watch recorded footage in the event of an incident or a complaint. In order to increase the effectiveness of camera use, intelligent computer vision features must be integrated into cameras.

Vietnamese license plates [2] come in a variety of styles, including long and short plates for automobiles and old (8 characters) and modern (9 characters) plates for motorcycles. The long plates have a ratio of 4:1, while the short plates have a ratio of 4:3. Despite this, they have several things in common.

The license plate [3] has a white background with black letters and numbers. It contains one of the 20 letters A, B, C, D, E, F, G, H, K, L, M, N, P, S, T, U, V, X, Y, Z (motorcycles over 50cc do not use the letter A) combined with one of the following numbers: X, Y, Z (motorcycles over 50cc use letters instead). They are given out to personal automobiles, business project management boards, social groups, professional associations, non-public career units, and driving testing and training facilities. The second set of numbers is the registration order for the car, which has 05 natural digits and ranges from 000.01 to 999.99.

Specific dimensions are as follows:

- Width of the license plate: 190 mm
- Height of the license plate: 140 mm
- Width of letters and numbers: 22 mm
- Height of letters and numbers: 55 mm
- Space between the characters and the top of the plate: 6 mm
- Space between the characters and the bottom of the plate: 6 mm

- Thickness of the letters and numbers: 7 mm
- Space between characters and numbers: 5 mm
- Space between registration numbers: 12 mm.

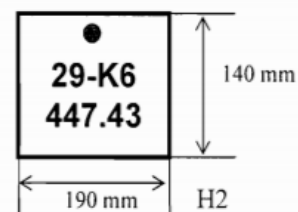


Fig. 1. Specifications of a motorbike license plate

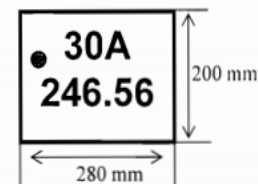


Fig. 2. a car's short license plate

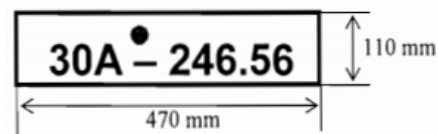


Fig. 3. a car's long license plate

Specific dimensions are as follows:

- Height of letters and numbers: 80 mm
- Width of letters and numbers: 40 mm
- Thickness of the letters and numbers: 10 mm
- Space between characters and numbers: 10 mm.
- Short license plate:
  - The width of the license plate: 280 mm
  - The height of the license plate: 200 mm
- Long license plate:
  - The width of the license plate: 470 mm
  - The height of the license plate: 110 mm

Numerous camera systems, including those made by ASTEC in Vietnam, Hikvision in China, and KBVISION in the US, are integrated into solutions that can run continuously, around-the-clock, in any weather. The cameras' shutter speeds can be automatically changed to accommodate various lighting situations, such as direct sunlight, light rain, or low light at night. To make sure the photos are of a high enough standard to be processed, the cameras are additionally attached

to high-power infrared transmitters. However, there aren't many open-source camera systems available right now that can be modified for the Vietnamese market. The US, Europe, and the Arab world are the only regions for which the OpenALPR [4,5] system can currently be customized for license plate recognition and vehicle type identification. The research topic focuses on adapting OpenALPR for Vietnamese automobiles and incorporating it into building systems like the campuses of Ho Chi Minh City University of Education's surveillance camera systems.

The aim of this research is to use OpenALPR to create a high-precision license plate recognition system employing security cameras:

- Investigate and apply OpenALPR for Vietnam's license plate recognition.
- Integrate into surveillance camera systems.
- Build the feature of recognizing Vietnamese license plates.

This research aims to improve the efficacy of surveillance camera systems and partially address the problems found in the application problem based on my contributions and suggestions.

## II. BACKGROUND AND RELATED WORK

A sequential mechanism governs how OpenALPR works. The output is the text from the license plate in the input image after multiple processing stages and input image processing.

### *i. License plate detection*

For each input image, the detecting stage is performed once. To identify probable license plate regions (x, y, width, and height), it employs the LBP algorithm, which is frequently used for face recognition. The subsequent stages receive each of these regions for additional processing. GPU acceleration can be used to increase performance because the detection stage typically requires the most work [6,7].

### *ii. Binarization*

There will be several occurrences of this step (and every stage that follows), one for each potential region for a license plate. For each area of the license plate, the binarization process will generate several black and white images. The best probability of locating every character is provided by using many black and white photos. For instance, if a single black and white image is too dark or too bright, characters may be missed. Different parameters are used with the Sauvola and Wolf-Jolien methods during binarization. The following steps involve processing each black and white image.

### *iii. Character Analysis*

Character-sized matching regions will be sought after in the license plate area during the character analysis stage. This will be accomplished by locating all linked color blobs in the vicinity of the license plate. The next step is to search for color blobs with peaks and valleys that are lined up with those of other blobs of a similar size and that are around the same width and height as characters from a license plate. The license plate area will see several iterations of this study. First, smaller character cells will be recognized, then gradually larger characters. A prospective license plate area will be

eliminated immediately if nothing is discovered there. If any viable characters are discovered, the character region will be preserved, and processing will move on to the next phase.

### *iv. Plate Edges*

In this stage, the edges of the license plate will be searched for. The Detection stage will be responsible for identifying a region that may contain a license plate. Usually, a region slightly larger or smaller than the actual license plate will be detected, but the exact top/bottom/left/right edges of the license plate will not be searched for. [8]

The first step is to find all the lines for the license plate region. In the `platelines.cpp` class, the license plate image will be processed, and a list of horizontal and vertical lines will be computed. The `platecorners.cpp` class will use this list, as well as the height of the characters (calculated in `Character Analysis`) to find the most suitable edges of the license plate. It will use some configurable weights to determine which edge is the most reasonable. It will try to use a default edge (based on the ideal width/height of the license plate) to see if it fits.

### *v. Deskew*

The deskew step [4] will remap the area of the license plate to a standard size and orientation using the results of the plate edges. This will result in a perfectly straight license plate image that is not rotated or tilted.

### *vi. Character Segmentation*

The character segmentation stage [9] aims to separate every character from the image of the license plate. To identify spaces between the characters on the license plate, it will employ a vertical histogram. Additionally, during this stage, the character cells will be cleaned up by having minor blemishes, broken pieces, and irregular character areas removed.

### *vii. Optical Character Recognition (OCR)*

Each character box will be independently analyzed by the OCR step. It will determine all potential characters and their confidence level for each character image. [10]

### *viii. Post Processing*

The character segmentation stage will attempt to isolate all characters in the license plate image. It will use a vertical histogram to find gaps between the characters of the license plate. This stage will also clean up the character cells by removing small spots, broken fragments, and non-standard character regions.

The OCR stage will analyze each character cell independently. For each character cell image, it will calculate all possible characters and their confidence level.

Given a list of all characters and their confidence levels, this step will determine the best possible letter combinations. Post-processing will remove all characters below a specific confidence threshold. It will also have a soft threshold defined by the programmer, where characters within this threshold will still be added to the list of possibilities. However, a blank character may be added because a low-confidence character may not actually be part of the license plate.

Post-processing will also handle license plate region verification if requested. For example, if you tell OpenALPR that a certain plate is a Vietnam plate, it will try to match the result with the Vietnam format template (for example: [digit] [digit] [letter] [letter] - [digit] [digit] [digit] [digit]). For example, if the top three results are:

- 63B2 - 18041
- G382 - 18041
- 63B2 - 18041

The third result will match the template, while the other two will not. Therefore, this step will mark the third result as the best match.



Fig. 4. A sample of license plate

### III. VIETNAM'S LICENSE PLATE RECOGNITION FRAMEWORK

If you are using *Word*, use either the Microsoft Equation Editor or the *MathType* add-on (<http://www.mathtype.com>) for equations in your paper (Insert | Object | Create New | Microsoft Equation or MathType Equation). "Float over text" should not be selected.

#### A. Technical stack

Use the links below to get the scripts required for training Vietnamese license plate recognition as well as the OpenALPR open-source code:

- [github.com/openalpr/train-ocr](https://github.com/openalpr/train-ocr)
- [github.com/openalpr/train-detector](https://github.com/openalpr/train-detector)
- [github.com/openalpr/openalpr/releases/tag/v2.2.0](https://github.com/openalpr/openalpr/releases/tag/v2.2.0)

The scripts must also be run on Python 3.x/2.x and OpenCV 3.0.0 for Detection training.

#### B. Data processing

Gathering the Vietnam's license plate as follow:

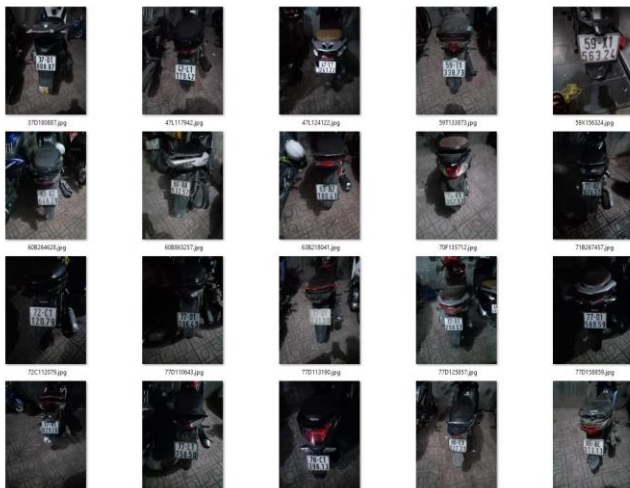


Fig. 5. Gathering license plate photo data

Run the `openalpr-utils-binarizefontsheet.exe` program to

create the boxes from each image.

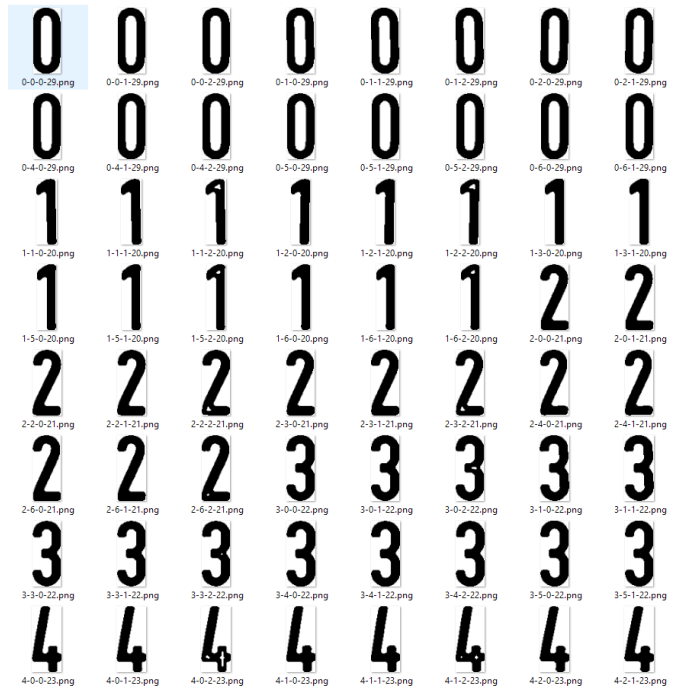


Fig. 6. Valid character output

#### C. OCR training

Perform character recognition training for license plates using OCR and OpenALPR.

```
009V0PR1HS7916PVXAYNBQBAMTCBR2VZJCJL6RX9X9LKKYSF0VGNR3V
MVAGJX31KAM7AM0SKLB4B95LQIENVZVL63RUVXPQ3019R1STZTDLPOS
GFGGR8F5MT6A8BF1YE5FYZDBPCVH900RHJXHVAPW65ZCP09HSQDLIT1W
XQ9MYR1AD8HEHUZNLQ4UG976AW770APWNEB871S1K1DSK2TJ5M104UG
6JY8TR131TQ0SJV0GE98PR39P982RW9MSB50HGNFV05VAZ2TKBC160B
F3FNZWFHQ7Y0RE94UJ1GBXPK9GGHUG3DR1DUW780VPBF0HEW7X1D22U4
1WB0Q26L T60DUBGXQ15A0QZ2XA100U0MKV0XR0GCR6AE6GC7GSAS1Z
N9X24RZZ4RDL9786CMQNT1U1JY5U1YCN6QZD7PHVS9Y4QE2A63BCS50V
HS4K8Y0018TLZJ5ECCFR0T3M1K9B2H37Z45WJTMGW3DBU4J1V0PCE57
9LL8P7L2AAP4PLAJ1B8CZ3SWB118543KMS05926NTBFT79NXQ424FDJ6
OT86TAD0G59JXAGTU9617350V8EVCY614311TLX55K114FELAV11W050P
Y11HY8MEXBKCVLJ4K1E5P9M66QE90LG8964M9B96KGWAL62R8LX34S
JDRUB6000ESTRWVIFVHLLH5B10MR1U8QJLYL20GD401SYG06EE3SEJ
Y6C7E80BRVJ1HXZXM9X0G801756VXDRF016M7J1Y1FSJ5P2HXRCBQMZ
VNCU19421F6GFZVDRSECB6LV006MKV0E87QTLUT57SBV0KDDIP6XS2LY
GKU93NT7F2QA972Y0VKJU0BUJ1P2Y0070M2GGZU4JH0NVLWDDRD6B07
SPK2DSANLU50C2ZKD4S0MU127Y0U1VMV4QWRXC92JLXTA96DWDD9QLP0
R7ZD7MK050HRD7AZEG8X37LHDOHDZA1LW7J68G5D9M4FEN40GD6YN1N
F9VTQYU00NP6XRBF3PT0AQ06GSKXUREAJ8Q55BRX7TAE6QU21B8J
SXX1820EFDV69V7Z8FRER7L1DQQTVHPQMBE703CVGE1MKFX3BDBUY92
GTDHLU8A3CQYM39RSN39GCMZINFZ76Q1UQK308JMO99UF9ZQRW3D4HD
VADFMKWU443JW4TL3KARF86695UC6ENXGSF4BPGP2XQPGG1G91T85019
6N62FHHYHTH009XB0U94PYADCCORSJ6C0S2C8Q7QUNQLRUVCSXF3U
MG7AM176NLPJYRT
```

Fig. 7. A sample of OCR training file

#### D. License plate detection training

The common location of the license plate in the image will be found by the detector. The Local Binary Pattern (LBP) technique is used to determine the license plate region. Both negative and positive pictures are necessary to train the detector.

The following stage is to train the detector after gathering

hundreds to thousands of photos of license plates. The training command must first be set up to use the exact size.

Change the variables WIDTH, HEIGHT, and COUNTRY in the prep.txt command to reflect Vietnam (WIDTH: 28, HEIGHT: 20, and COUNTRY: "vn"). The total number of pixels should be close to 650px, and the width and height must be appropriate to the size of the license plate.

**E. Vietnam's license plate recognition system**

Number equations consecutively with equation numbers in parentheses flush with the right margin, as in (1). First use the equation editor to create the equation.

The Windows operating system or a webserver can be used to deploy the Vietnamese license plate recognition system.

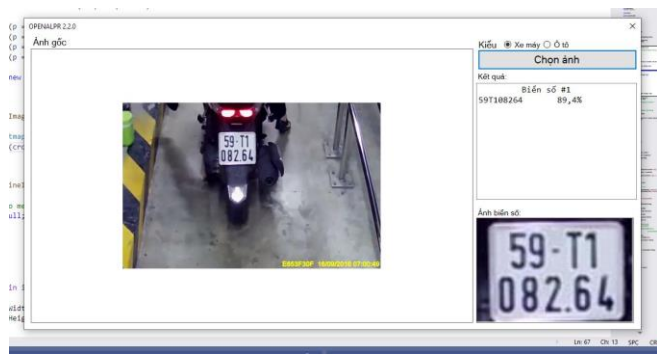


Fig. 8. Vietnamese license plate recognition system for Windows OS

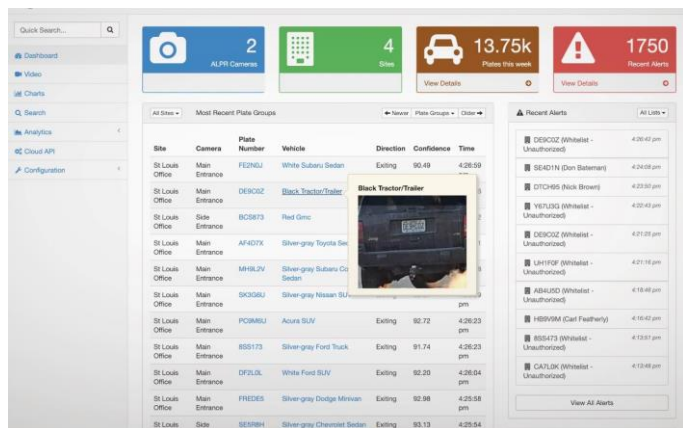


Fig. 9. Vietnamese license plate recognition system for webserver

**IV. EXPERIMENTAL RESULTS AND DISCUSSION**

**A. Experimental Environment**

The program is written using the Visual Studio 2019 Express application in C# language and tested on the Asus Vivobook S15 S510UQ laptop with Core i5 8250 3.4GHz, 12GB RAM, NVIDIA Geforce GTX 940M.

The OpenALPR source code used for research is the open-source version OpenALPR 2.2.0 and was customized on Visual Studio Code, compiled on Python 2.7.

**B. License plate recognition workflow**

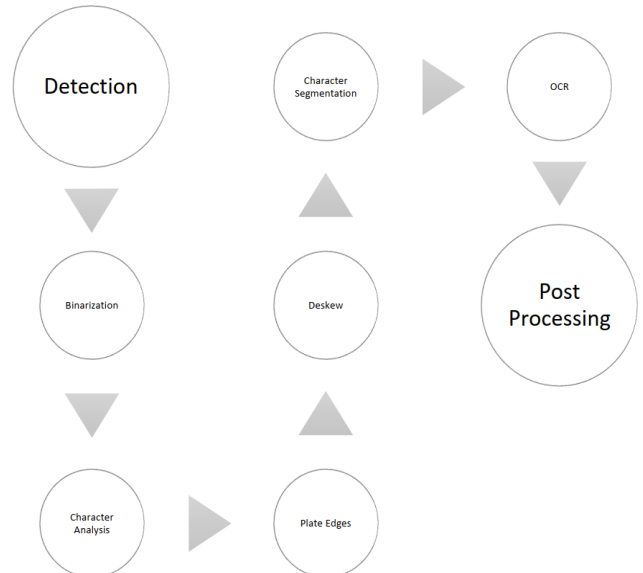


Fig. 10. License plate recognition workflow

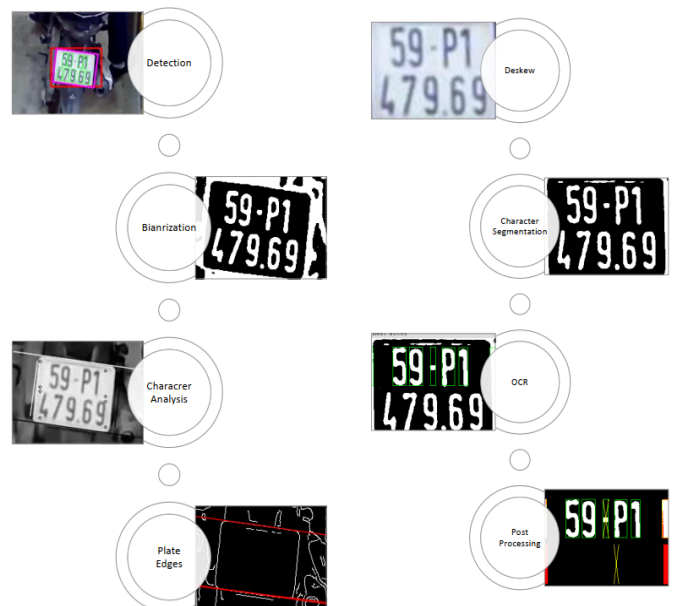


Fig. 11. An example of license plate recognition workflow

**B. Experimental results**

120 images of vehicles going through the parking lot scanner are included in the test dataset.

**V. CONCLUSION**

In this research, we proposed a solution for Vietnam's license plate recognition include many stages based on OpenALPR with OCR incremental training. The experimental result in the real application shows the feasibility of the proposed solution in Vietnam application environment.

This paper emphasizes the need for programming abilities, familiarity with software development tools, and open-source libraries in order to design a license plate recognition system. However, creating a license plate recognition system is now simpler and can be implemented on a variety of platforms

