

# MỘT MÔ HÌNH ĐỀ XUẤT CHO BÀI TOÁN NHẬN DẠNG KÝ TỰ TRÊN CONTAINER VẬN TẢI ĐƯỜNG THỦY

Lê Hoàng Thanh<sup>a\*</sup>

<sup>a</sup>Khoa Công nghệ Thông tin, Trường Đại học Nha Trang, Khánh Hoà, Việt Nam

## Lịch sử bài báo

Nhận ngày 08 tháng 01 năm 2017 | Chính sửa ngày 11 tháng 04 năm 2017

Chấp nhận đăng ngày 11 tháng 05 năm 2017

## Tóm tắt

Nhận dạng ký tự là một trong những lĩnh vực đang rất được quan tâm hiện nay nhằm xây dựng các chương trình có khả năng chuyển các hình ảnh thành các văn bản tài liệu. Kỹ thuật này được ứng dụng trong công tác quét và lưu trữ, đẩy nhanh việc nhập dữ liệu vào máy với ít lỗi hơn. Hiện tại trong lĩnh vực vận tải, công việc ghi nhận mã số của container được thực hiện chủ yếu bằng tay, vì vậy cần thiết phải có một giải pháp để nhận dạng tự động để tiết kiệm thời gian và công sức. Bài báo bước đầu đề xuất một mô hình giải quyết bài toán nhận dạng ký tự trên container dựa trên việc phân tách các ký tự và rút trích đặc trưng HOG. Kết quả bước đầu từ thực nghiệm chứng minh rằng mô hình đề xuất cho kết quả tương đối khả quan.

**Từ khóa:** HOG; Nhận dạng ký tự; Số container; SVM.

## 1. GIỚI THIỆU

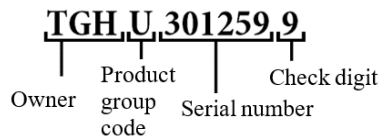
Nhận dạng ký tự là cần thiết và được ứng dụng rộng rãi trong nhiều lĩnh vực như nhận dạng các chữ số trên chi phiếu ngân hàng, mã số trên bì thư của dịch vụ bưu chính, hay các chữ số trên các biểu mẫu nói chung. Vấn đề nhận dạng ký tự nói chung là một thách thức lớn đối với các nhà nghiên cứu (Trần, 2013). Bài toán lớn luôn đặt ra phía trước vì sự phức tạp của việc nhận dạng ký tự phụ thuộc nhiều vào các yếu tố khác nhau như đặc thù của ứng dụng, hoặc chất lượng ảnh đầu vào. Có thể hình dung vấn đề này từ ứng dụng nhận dạng chữ viết tay so với nhận dạng bảng số xe, tuy cùng là bài toán nhận dạng ký tự nhưng cách tiếp cận và giải pháp là hoàn toàn khác biệt. Hệ thống nhận dạng thường bao gồm hai bước: (1) Rút trích đặc trưng từ ảnh; và (2) Học tự động từ các đặc trưng để có thể nhận dạng ký tự. Hiệu quả của hệ thống nhận dạng phụ thuộc vào các

phương pháp sử dụng ở hai giai đoạn trên. Hầu hết các hệ thống hiện nay đều sử dụng các đặc trưng cơ bản từ ảnh ký tự như đường biên, cạnh, độ dày, giá trị mức xám (Trần, 2013), Haar-like (Said, Tan, & Baker, 2000), với các xử lý đặc thù như lấy mẫu, dao động các điểm ảnh, biến đổi ảnh, thêm dữ liệu ảo (Freund & Schapire, 1996). Sau đó hệ thống nhận dạng huấn luyện các mô hình học tự động như  $k$ -NN, mạng neural, SVM, boosting (Du & Swamy, 2006). Các kỹ thuật áp dụng phụ thuộc rất nhiều vào từng bài toán cụ thể. Bài báo này đề xuất một mô hình giải quyết riêng cho bài toán nhận dạng ký tự trên container, bao gồm các bước chính là phân tách các ký tự trên container, rút trích đặc trưng ảnh và nhận dạng dựa trên máy phân lớp SVM.

## 2. QUY ƯỚC SỐ HIỆU CONTAINER

### 2.1. Hệ thống nhận biết

Hệ thống nhận biết của container bao gồm 4 thành phần sau: Mã chủ sở hữu (*owner code*), ký hiệu loại thiết bị (*product group code*), số serial (*serial number*) và chữ số kiểm tra (*check digit*), như được mô tả trong Hình 1.



**Hình 1. Một ví dụ về mã quy ước nhận biết container**

#### 2.1.1. Mã chủ sở hữu

Mã chủ sở hữu bao gồm 3 chữ cái viết hoa được thống nhất và đăng ký với cơ quan đăng kiểm quốc tế thông qua cơ quan đăng kiểm quốc gia hoặc đăng ký trực tiếp với Cục container Quốc tế (BIC). Sau khi đăng ký, việc sở hữu mã mới được chính thức công nhận trên toàn thế giới.

Một hãng có thể sở hữu một hoặc nhiều mã khác nhau, mặc dù BIC hạn chế điều này, và đưa ra những điều kiện nhất định cho việc đăng ký nhiều mã. Ở Việt Nam, đến đầu năm 2010, có 6 công ty đăng ký mã tiếp đầu ngữ với BIC, chi tiết như Bảng 1.

**Bảng 1. Mã chủ sở hữu của các công ty Việt Nam**

| Mã BIC | Tên công ty         |
|--------|---------------------|
| GMD    | Gemadep             |
| GMT    | Gematrans           |
| NSH    | Nam Trieu shipping  |
| VCL    | Vinashin-TGC        |
| VNL    | Vinalines container |
| VNT    | Vinashin-TGC        |

### 2.1.2. Ký hiệu loại thiết bị

Là một trong ba chữ cái dưới đây viết hoa, tương ứng với từng loại thiết bị:

- U: Container chở hàng (*freight container*);
- J: Thiết bị có thể tháo rời của container chở hàng (*detachable*);
- Z: Đầu kéo (*trailer*) hoặc moóc (*chassis*).

Việc sử dụng bất kỳ chữ cái nào không thuộc ba chữ cái trên (U, J, Z) làm ký hiệu loại thiết bị được coi là không tuân theo tiêu chuẩn ISO 6346 (ISO, 1995).

### 2.1.3. Serial number

Đây chính là số hiệu của container, gồm 6 chữ số. Nếu số serial không đủ 6 chữ số, thì các chữ số 0 sẽ được thêm vào phía trước để thành đủ 6 chữ số. Chẳng hạn, nếu số serial là “1234”, thì số serial đầy đủ sẽ là “001234”. Số serial do chủ sở hữu container đặt ra, nhưng đảm bảo nguyên tắc mỗi số chỉ sử dụng duy nhất cho một container.

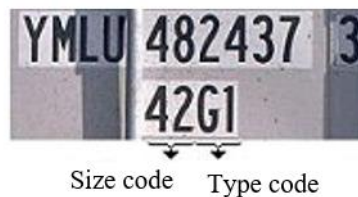
### 2.1.4. Chữ số kiểm tra (*Check digit*)

Là một chữ số đứng sau số serial, dùng để kiểm tra tính chính xác của chuỗi ký tự đứng trước đó. Với mỗi chuỗi ký tự gồm tiếp đầu ngữ và số serial, áp dụng cách tính chữ số kiểm tra container, sẽ tính được chữ số kiểm tra cần thiết. Việc sử dụng số kiểm tra là để giảm thiểu rủi ro sai sót trong quá trình nhập số container. Thực tế là số container được nhiều đối tượng sử dụng nhiều lần, trên nhiều chứng từ do đó khả năng nhập sai số là rất lớn. Mỗi số container sẽ tương ứng với một chữ số kiểm tra. Do đó, việc nhập sai số phần lớn sẽ bị phát hiện do chữ số kiểm tra khác với thực tế. Tuy vậy, điều này không phải

tuyệt đối, bởi nếu sai 2 ký tự trở lên thì có thể số kiểm tra vẫn đúng, và sai sót không bị phát hiện ra (ISO, 1995).

## 2.2. Mã kích thước và mã kiểu

Mã kích thước bao gồm 2 ký tự, chữ cái hoặc chữ số. Ký tự thứ nhất biểu thị chiều dài container, chữ số 4 trong ví dụ tại Hình 2 thể hiện chiều dài container này là 40 feet (~12.192m). Ký tự thứ hai biểu thị chiều rộng và chiều cao container, trong ví dụ này chiều cao là 8 feet 6 inches (~2.591m). Mã kiểu cũng gồm 2 ký tự, ký tự thứ nhất cho biết kiểu container, ký tự “G” thể hiện container hàng bách hóa. Ký tự thứ hai biểu thị đặc tính chính liên quan đến container, ký tự “1” nghĩa là container có cửa thông gió phía trên. Tóm lại, “42G1” trong Hình 2 thể hiện container bách hóa dài 40 feet, cao 8 feet 6 inches, thông gió phía trên. Tiêu chuẩn ISO 6346 quy định chi tiết ý nghĩa các mã kích thước và mã kiểu.



Hình 2. Ví dụ về mã quy ước kích thước và kiểu của container

## 2.3. Các dấu hiệu khai thác

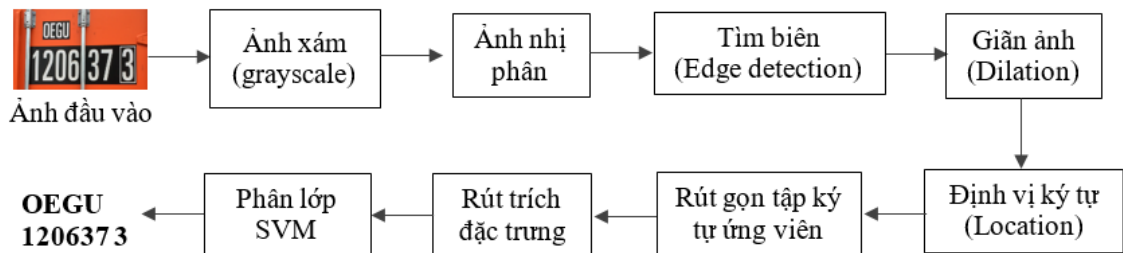
Các dấu hiệu trong khai thác gồm loại bắt buộc và loại không bắt buộc. Dấu hiệu bắt buộc bao gồm tải trọng container, cảnh báo nguy hiểm điện và container cao. Đối với trọng lượng tối đa sẽ được ghi trên cửa container, số liệu tương tự như trong biển chứng nhận an toàn CSC. Một số container cũng thể hiện trọng lượng vỏ, trọng tải hữu ích hay lượng hàng xếp cho phép. Dấu hiệu không bắt buộc bao gồm khối lượng hữu ích lớn nhất, mã quốc gia. Thông tin về khối lượng hữu ích lớn nhất sẽ được dán trên cửa container, phía dưới dấu hiệu trọng lượng container tối đa. Mã quốc gia gồm 2 chữ cái viết tắt thể hiện tên quốc gia sở hữu container. Trong ví dụ ở Hình 3, “KR” quy ước container đến từ Hàn Quốc (ISO, 1995).



Hình 3. Ví dụ về mã quy ước tải trọng và mã quốc gia của container

### 3. MÔ HÌNH ĐỀ XUẤT

*Tìm biên ảnh:* Một biên trong một hình ảnh nói chung có thể được định nghĩa như là đường bao của ảnh hoặc đường viền ngăn cách khu vực ảnh liền kề có đặc điểm tương đối khác biệt theo một số tính năng quan tâm, một trong số đó là độ đo về sự thay đổi đột ngột về cấp xám. Biên có thông tin quan trọng góp phần hướng tới phân tích và giải thích thông tin hình ảnh nhận dạng (Luong, 2003). Kết quả thực nghiệm bước đầu cho thấy thuật toán tìm biên Canny tương đối thích hợp trong việc phân tách ký tự trong trường hợp bài toán cụ thể này.



Hình 4. Các bước của mô hình đề xuất

*Giãn ảnh nhị phân:* Tập hợp B thường thì được coi như là một phần tử cấu trúc trong giãn nhị phân, cũng như trong các phép toán hình thái khác, tập hợp A là tập hợp các phần tử của hình ảnh gốc (Luong, 2003). Với A và B là các tập hợp trong  $Z^2$ , thì phép giãn nhị phân của A theo B ( $A \oplus B$ ) được định nghĩa qua công thức (1) (Luong, 2003).

$$A \oplus B = \{z \mid (B)_z \cap A \neq \emptyset\} \tag{1}$$

Phép giãn ảnh được dùng trong trường hợp bài toán này với mục đích làm rõ các đối tượng đã tìm biên ở bước trước, tránh tình trạng bị mất nét.

*Định vị đối tượng:* Để khoanh vùng các đối tượng có khả năng là ký tự, bài báo sử dụng thuật toán định vị dựa trên láng giềng 8. Mỗi đối tượng ứng viên trong ảnh sẽ được khoanh vùng bởi một hình chữ nhật có kích thước  $n \times m$

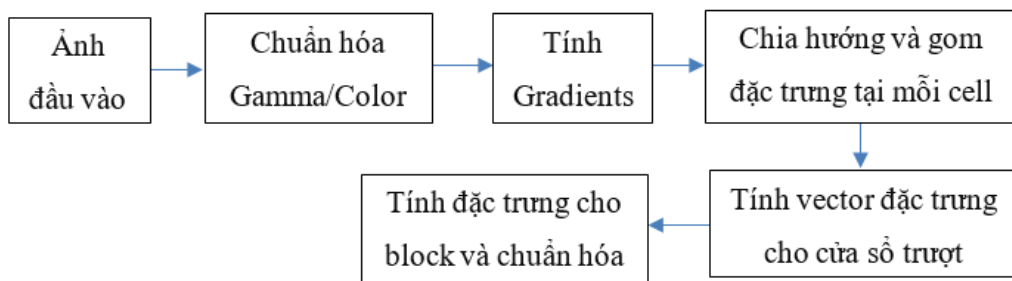
*Rút gọn tập ứng viên:* Sau bước định vị các đối tượng trong ảnh ta sẽ được một tập các ký tự ứng viên để tiến hành nhận dạng. Tuy nhiên để rút ngắn thời gian thực hiện cho bước phân loại tiếp theo, bài báo đề xuất việc sơ loại các đối tượng dựa trên tỷ lệ kích thước chúng. Từ thực nghiệm cho thấy một đối tượng chỉ có thể là ký tự trên container nếu thỏa mãn các điều kiện như công thức (2).

$$\frac{0.486}{1.083} - \alpha \leq \frac{m}{n} \leq \frac{0.486}{1.083} + \alpha \quad 40px - \beta \leq n < m \leq 200px + \beta \quad (2)$$



**Hình 5.** Ví dụ về việc khoanh vùng các ký tự ứng viên đã được khoanh vùng

*Rút trích đặc trưng:* Hiện tại có nhiều phương pháp để rút trích đặc trưng của một ảnh. Trong bài toán nhận ký tự quang học nói chung, phương pháp chia lưới ký tự hiện vẫn đang được sử dụng phổ biến. Tuy nhiên, bài báo này đề xuất sử dụng các phương pháp rút trích đặc trưng dựa trên gradient ảnh là HOG (Histogram of Oriented Gradients) như mô tả trong Hình 6. Các kỹ thuật này đã được chứng minh tính hiệu quả trong việc rút trích đặc trưng phục vụ đầu vào cho các bài toán nhận dạng (Dalal & Triggs, 2005).



**Hình 6.** Các bước rút trích đặc trưng HOG

Kỹ thuật HOG sử dụng cửa sổ trượt để rút trích đặc trưng ảnh đầu vào. Ảnh sẽ được chia thành các ô có kích thước  $8 \times 8$  pixel,  $2 \times 2$  ô tạo thành 1 block. Gradient vector được tính tại mỗi pixel, như vậy sẽ có 64 gradient vector trong 1 cell. Các gradient vector này sẽ được đặt vào histogram gồm  $n$  bin. Giả sử  $n$  được chọn là 9 thì dãy các bin của histogram sẽ nằm trong khoảng từ 0 -  $180^0$ , mỗi bin ứng với phạm vi  $20^0$  (Dalal & Triggs, 2005). Giá trị bố trí vào các bin của histogram chính là tỉ trọng độ lớn của gradient vector gần với bin đó nhất. Hay nói cách khác, ta chia tỷ trọng độ lớn của gradient vector vào 2 bin gần nhất.

Như đã nói ở trên, giá trị của 9 bin trong histogram chính là tích lũy độ lớn của 64 gradient vector trong 1 cell ( $8 \times 8$  pixel). Giả sử mức xám mỗi pixel trong 1 cell được nhân 1.5 lần (tăng độ tương phản) thì độ lớn của các gradient vector cũng được nhân lên 1.5 lần. Như vậy giá trị của mỗi bin trong histogram cũng sẽ được nhân 1.5 lần. Do đó, để bất biến với mọi thay đổi về độ sáng của ảnh ta cần phải chuẩn hóa histogram. Tuy nhiên trên thực tế không cần phải chuẩn hóa từng histogram, thay vào đó sẽ thực hiện việc chuẩn hóa theo block (Dalal & Triggs, 2005). Trong đó, mỗi block sẽ được biểu diễn bằng cách ghép các histogram của 4 ô thành một vector gồm 36 thành phần (4 histogram  $\times$  9 bin/histogram). Trong bài toán dò ký tự container, cửa sổ trượt được thiết lập có kích thước là  $64 \times 128$  pixel sẽ được chia thành 105 block (7 block ngang  $\times$  15 block dọc). Mỗi block được biểu diễn bởi 36 giá trị, như vậy tổng cộng sẽ có 3780 giá trị cho 1 cửa sổ dò tìm. Hay nói cách khác mỗi cửa sổ sẽ được biểu diễn bởi một vector đặc trưng bao gồm 3780 chiều.

*Nhận dạng ký tự:* Để thực hiện việc phân lớp ký tự, bài báo sử dụng phân lớp SVM với chiến lược “one-versus-all” (Alpaydin, 2014). Như vậy với tổng cộng 36 ký tự sẽ tương ứng với việc xây dựng 36 bộ phân lớp SVM nhị phân.

#### 4. KẾT QUẢ THỰC NGHIỆM

Để huấn luyện mô hình, bài báo sử dụng tập dữ liệu chuẩn *Chars74K* được xây dựng từ các ảnh chụp của Google Street View (De Campos, Babu, & Varma, 2009). Các ký tự được chọn bao gồm số và chữ cái in hoa, hầu hết được trích từ bộ *GoodImg* trong

*Chars74K*. Đối với các ký tự số, vì số lượng bộ *GoodImg* khá ít nên toàn bộ các ký tự trong *BadImg* cũng được bổ sung đưa vào quá trình huấn luyện.



**Hình 7. Minh họa tập dữ liệu Chars74K dùng để huấn luyện**

Như vậy tổng cộng số lượng các ký tự đưa vào để huấn luyện là 6087, với chi tiết được liệt kê trong Bảng 2. Sau quá trình huấn luyện, một tập bao gồm 104 ảnh chụp container lấy từ Internet được sử dụng để đánh giá hiệu năng của mô hình. Bên cạnh việc đo độ chính xác dựa trên số lượng ảnh nhận dạng đúng, cần thiết có sự đánh giá tổng quát khả năng phát hiện và định vị đúng ký tự của mô hình thông qua độ đo *recall* và *precision*. Trong đó *recall* càng cao phản ánh khả năng phát hiện ra được ký tự càng cao, còn độ đo *precision* phản ánh tính tin cậy trong kết luận là ký tự container. Kết quả thực nghiệm thể hiện trong Bảng 3.

**Bảng 2. Chi tiết số lượng ký tự sử dụng để huấn luyện mô hình**

| Ký tự | Số lượng | Ký tự | Số lượng | Ký tự | Số lượng | Ký tự | Số lượng |
|-------|----------|-------|----------|-------|----------|-------|----------|
| 0     | 155      | 9     | 71       | I     | 302      | R     | 389      |
| 1     | 128      | A     | 558      | J     | 77       | S     | 342      |
| 2     | 118      | B     | 115      | K     | 92       | T     | 312      |
| 3     | 72       | C     | 215      | L     | 215      | U     | 92       |
| 4     | 73       | D     | 191      | M     | 149      | V     | 84       |
| 5     | 81       | E     | 446      | N     | 363      | X     | 80       |
| 6     | 84       | F     | 79       | O     | 382      | Y     | 67       |
| 7     | 61       | G     | 143      | P     | 159      | Z     | 55       |
| 8     | 42       | H     | 193      | Q     | 35       | W     | 67       |

**Bảng 3. Kết quả thực nghiệm đánh giá độ chính xác của mô hình và khả năng phát hiện và định vị ký tự**

| Tổng số ảnh                     | Số lượng ảnh sai<br>(có ký tự sai, hoặc bị sót) | Tổng số ký<br>tự thực tế | Số lượng khoanh<br>vùng ký tự       | Số lượng ký tự phát<br>hiện và định vị đúng |
|---------------------------------|---|--------------------------|-------------------------------------|---|
| 104                             | 29  | 1067                     | 1207                                | 927   |
| <i>Accuracy</i> = 75/104 ~72.1% |   |                          | <i>Recall</i> = 927/1067 ~ 86.8%    |   |
|                                 |   |                          | <i>Precision</i> = 927/1207 ~ 76.8% |   |



## 5. KẾT LUẬN

Có thể thấy rằng bài toán nhận dạng ký tự quang học vẫn là một trong những thách thức lớn cần được giải quyết, các bước thiết kế thuật toán cần phải phù hợp với đặc thù của ảnh trong tập dữ liệu. Bài báo này nhằm nỗ lực đưa ra một hướng giải quyết mới cho việc nhận dạng các ký tự trên container vận tải đường thủy. Một trong những đề xuất quan trọng của bài báo đó là quá trình phân tách ký tự và rút trích đặc trưng HOG trước khi đưa vào bộ phân lớp SVM.

Tuy kết quả chưa thực sự cao, nhưng bài báo đã có sự khảo sát tương đối kỹ càng đặc thù của việc nhận dạng ký tự container để từ đó đề xuất một cách giải quyết với kết quả ban đầu tương đối khả quan. Sở dĩ kết quả chưa cao có thể một phần do mô hình được huấn luyện trên tập dữ liệu *Chars74K* trong khi thử nghiệm trên các ảnh chụp thực tế. Một trong những hướng phát triển để nâng cao độ chính xác của mô hình đó là tạo một bộ dữ liệu riêng, trong đó các ký tự được cắt từ ảnh chụp container.

## TÀI LIỆU THAM KHẢO

- Alpaydin, E. (2014). *Introduction to machine learning*. Massachusetts, USA: MIT Press.
- Cha, S. H. (2001). *Use of distance measures in handwriting analysis*. (Doctoral Thesis), State University, New York. Retrieved from [https://www.researchgate.net/profile/Sung-Hyuk\\_Cha/publication/33787096\\_Use\\_of\\_distance\\_measures\\_in\\_handwriting\\_analysis\\_electronic\\_resource/links/00b495177e4a98d3b7000000.pdf](https://www.researchgate.net/profile/Sung-Hyuk_Cha/publication/33787096_Use_of_distance_measures_in_handwriting_analysis_electronic_resource/links/00b495177e4a98d3b7000000.pdf)
- Dalal, N., & Triggs, B. (2005). *Histograms of oriented gradients for human detection*. Paper presented at The IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), USA.
- De Campos, T. E., Babu, B. R., & Varma, M. (2009). *Character recognition in natural images*. Paper presented at VISAPP 2009, Portugal.
- Du, K. L., & Swamy, M. N. (2006). *Neural networks in a soft computing framework*. Berlin, Germany: Springer Science & Business Media.
- Freund, Y., & Schapire, R. E. (1996). *Experiments with a new boosting algorithm*. Paper presented at Machine Learning: The Thirteenth International Conference, Italy.
- International Organization for Standardization – ISO. (1995). *ISO 6346*. Retrieved from <https://www.iso.org/standard/20453.html>
- Luong, M. B. (2003). *Nhập môn xử lý ảnh số*. Hà Nội, Việt Nam: NXB Khoa học và Kỹ thuật.

- Said, H. E., Tan, T. N., & Baker, K. D. (2000). Personal identification based on handwriting. *Pattern Recognition*, 33(1), 149-160.
- Trần, H. C. (2013). *Nhận dạng ký tự quang học*. Hà Nội, Việt Nam: Trường Đại học Công Nghiệp Hà Nội.
- Trier, Ø. D., Jain, A. K., & Taxt, T. (1996). Feature extraction methods for character recognition-a survey. *Pattern Recognition*, 29(4), 641-662.

## A NOVEL APPROACH FOR SHIPPING CONTAINER CODE RECOGNITION

Le Hoang Thanh<sup>a\*</sup>

<sup>a</sup>The Faculty of Information Technology, Nhatrang University, Khanhhoa, Vietnam

\*Corresponding author: Email: thanhhlh@ntu.edu.vn

### Article history

Received: January 08<sup>th</sup>, 2017 | Received in revised form: April 11<sup>th</sup>, 2017

Accepted: May 11<sup>th</sup>, 2017

---

### Abstract

*Optical character recognition is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded texts, whether from a scanned document, a photo of a document, a scene-photo or from subtitle text superimposed on an image. It is widely used as a form of information entry from printed data records including, passport documents, invoices, bank statements, computerized receipts, business cards, mail, printouts of static-data, or any suitable documentation. Currently, in logistic, container code recognition is mainly done manually, so it is necessary to have a solution for automatic identification to save time and effort. This paper proposes a novel model for code recognition which can be applied for shipping containers widely used in logistics. The obtained experimental results have proved that the proposed model produces satisfactory confidence on a benchmark dataset.*

**Keywords:** Character recognition; Container number; HOG; SVM.

---